Volume 12, Issue 2 (XVII)
April - June 2025
ISSN: 2394 – 7780

International Journal of
Advance and Innovative Research

Indian Academicians and Researchers Association
www.iaraedu.com

Jai Hind College

(Empowered Autonomous)

A+ GRADE with a CGPA of 3.36 out of 4
in its Fourth NAAC cycle.

Department of MSc. Big Data Analytics, BSc. IT & BVOC SD

Organizes National Research Conference on

"Converging AI and Big Data for Cyber Security Excellence"

on

Saturday, 23rd November 2024

About College:

Jai Hind College Autonomous upholds and maintains high standards of academic excellence by imparting quality and holistic education to its students empowering them to transform themselves into global leaders. Established in 1948 by a group of teachers who had migrated from Karachi, the college is managed by the Sind Educationists' Association Trust. Conferment of Autonomous Status in 2018-19, and Empowered Autonomous Status in 2023-24. The college was conferred "Best College Award" by University of Mumbai. In 2015-16, the college received DST-FIST grant for improvement in Science and Technology infrastructure. In 2018-19, three departments Botany, Chemistry and Microbiology, were identified for promotion and popularization of Science under the DBT-STAR scheme. In 2018-19 the institution was recognized by RUSA as 'College of Excellence'.

About the Department:

Under the banner of Information Technology, our department offers three distinct and dynamic programs: **M.Sc. Big Data Analytics**, **B.Sc. Information Technology (IT)**, and **B.Voc Software Development (SD)**. Each program is designed to meet the evolving demands of the tech industry, blending academic rigor with practical training to nurture skilled professionals and future entrepreneurs.

B.Sc. Information Technology

Established in 2013, the B.Sc. (I.T.) program aims to cultivate future leaders in computer science by providing a robust foundation in core IT concepts and technologies. The curriculum is continually updated to reflect current industry trends and technological advancements. With a strong emphasis on employability and entrepreneurship, the program equips students with the knowledge and skills needed to thrive in the modern tech landscape.

B.Voc Software Development

B.Voc. in Software Development is a UGC recognized three-year program established in 2015 under the NSQF (Level 3-5), affiliated with the University of Mumbai. Developed with MNCs and industry experts, it follows a credit-based grading system with multiple exit points. The curriculum emphasizes hands-on training and industry internships. Graduates receive NSDC certification, ensuring PAN India recognition and enhanced career opportunities.

M.Sc. Big Data Analytics

Introduced in 2020 in collaboration with Tata Consultancy Services (TCS), the M.Sc. Big Data Analytics program is a forward-thinking, two-year postgraduate course that spans four semesters. It delivers a deep understanding of big data technologies—from foundational theories to real-world applications. The curriculum features hands-on projects, industry-relevant case studies, and practical training. A mandatory final-semester internship provides students with immersive industry exposure, enabling them to apply their skills in real-world settings and launch successful careers in data science and analytics.

About Conference:

This conference aims to explore the latest advancements and emerging trends in **Cybersecurity**, **Artificial Intelligence**, **Cyber Forensics**, and related domains. It serves as a vibrant platform for **students, academicians, and researchers** from across the country to engage in meaningful discussions, share insights, and collaborate on innovative approaches shaping the future of these rapidly evolving fields.

Participants will have the opportunity to delve into **cutting-edge technologies** and gain a deeper understanding of current challenges and solutions. In particular, **emerging scientists and researchers** are encouraged to present their work through **paper presentations**, allowing them to showcase their ideas, receive valuable feedback, and gain exposure to recent developments in their areas of interest.

PATRON

Dr. Ashok Wadia
(Academic Advisor, Jai Hind College)

CHAIRPERSON

Prof. (Dr.) Vijay Dabholkar
(Principal, Jai Hind College)

ADVISORY COMMITTEE

Dr. Mahendra Kanojia
(Incharge Principal, Sheth L.U.J and SIR M.V. College)

Dr. Apurva Yadav
(Vice Principal, Kirti M. Doongursee College)

Dr. Sumithra T. V
(Assistant Professor, D.Y. Patil Deemed to be University)

Dr. Sreela Dasgupta
(Dean of Academic Affairs, Jai Hind College)

CONVENER

Mr. Wilson Rao
(HOD, Dept of MSc. Big Data Analytics, BSc IT & BVoc SD)

ORGANISING COMMITTEE

Ms. Sunita Jena

Ms. Fatima Shaikh

Ms. Tejashree Parab

Ms. Shraddhadevi Singh

Ms. Rohana Deshpande

Message from Chairperson…….

From the Desk of the Principal

Dr. Vijay Dabholkar
Principal
Jai Hind College (Autonomous), Churchgate, Mumbai

It is with immense pride and enthusiasm that **Jai Hind College** hosts the **National Research Conference on "Converging AI and Big Data for Cyber Security Excellence."** In today's increasingly interconnected world, the transformative role of **Artificial Intelligence** and **Big Data** in enhancing cybersecurity practices cannot be overstated. This conference serves as a dynamic platform for **thought leaders, researchers, and industry practitioners** to converge, share insights, and explore new avenues for collaboration and innovation.

At Jai Hind College, we have always fostered a spirit of **academic innovation and excellence**, with a commitment to ensuring that education evolves in tandem with technological progress. The intersection of **cybersecurity, AI, and Big Data** represents one of the most critical and urgent areas of research today. We believe that through the confluence of diverse ideas and perspectives, we can collectively contribute to building more secure, resilient, and intelligent digital ecosystems.

I extend my heartfelt congratulations to the **Department of M.Sc. Big Data Analytics, B.Sc. IT, and B.Voc in Software Development** for successfully organizing this esteemed event. I also take this opportunity to express sincere appreciation to the **organizing committee, keynote speakers, invited experts, and all participants**, whose dedication and effort have made this conference a reality.

May this gathering mark a pivotal moment in our collective pursuit of knowledge and innovation, and inspire us all to push boundaries and create impactful solutions for a safer digital future.

Message From Convener……

From the Desk of the Head of Department – M.Sc. Big Data Analytics, B.Sc. IT & B.Voc. in Software Development

Mr. Wilson Rao
Jai Hind College (Autonomous), Churchgate, Mumbai

It is an honour and a privilege to welcome you to the **National Research Conference on "Converging AI and Big Data for Cyber Security Excellence."** In today's rapidly evolving technological landscape, **cybersecurity** stands as a cornerstone of our digital world. The integration of **Artificial Intelligence** and **Big Data** is not only transforming how we detect and mitigate cyber threats but also redefining the boundaries of innovation in this critical domain.

This conference provides an essential platform for **experts, researchers, and innovators** to exchange ideas, share pioneering work, and explore the vast potential of emerging technologies in enhancing digital security.

Our department has always been committed to developing IT professionals and researchers who are not only technically adept but also deeply aware of the broader implications and real-world applications of their work. This event reflects our dedication to fostering research that is both **cutting-edge and socially relevant**, especially in addressing the pressing challenges of our time.

I extend my heartfelt gratitude to our esteemed **Principal, Dr. Vijay Dabholkar**, for his unwavering support and visionary leadership, which continue to inspire us to pursue excellence in every endeavour.

A special thanks to the **organizing committee**, the **presenters**, the **researchers**, and all participants whose passion and efforts have brought this conference to life. Your contributions are invaluable in advancing the frontiers of cybersecurity through collaborative knowledge and innovation.

May this conference serve as the beginning of many **fruitful discussions, partnerships, and discoveries** that will shape the future of cybersecurity in the age of AI and Big Data.

International Journal of Advance and Innovative Research

Volume 12, Issue 2 (XVII): April - June 2025

The International Journal of Advance and Innovative Research is an online open access, peer reviewed & refereed journal.

International Journal of Advance and Innovative Research

Volume 12, Issue 2 (XVII): April - June 2025

CONTENTS

Research Papers

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

DETECTING AND PREDICTING FINANCIAL FRAUD USING EXPLAINABLE AI & GRAPH-BASED ANOMALY DETECTION

[1]Aditya Kudva and [2]Tejashree Parab
Department of MSc Big Data Analytics, Jai Hind College (Autonomous), Mumbai, India

ABSTRACT
This literature paper review intends to highlight the shortcomings of traditional fraud detection methods, which often lack adaptability and transparency. Explainable AI (XAI) offers solutions by making model predictions more interpretable, enabling clearer insights for users. Additionally, graph-based anomaly detection captures complex relationships between financial entities, identifying patterns missed by traditional approaches. By integrating XAI with graph-based techniques, this review demonstrates improved accuracy and interpretability in fraud detection, offering advancements in financial auditing and compliance.

Keywords— ***Financial statement fraud, Explainable AI (XAI), Graph-based anomaly detection, Fraud detection, Machine learning.***

INTRODUCTION
Financial statement fraud is a major concern for businesses and regulators, with traditional detection methods often lacking adaptability and transparency. While rule-based and black-box machine learning models are commonly used, they fail to evolve with changing fraud patterns and lack interpretability. Explainable AI (XAI) addresses this by making model predictions transparent, providing understandable insights for stakeholders. Additionally, graph-based anomaly detection captures complex relationships between financial entities, identifying irregular patterns indicative of fraud. This paper explores the integration of XAI and graph-based detection techniques to improve the accuracy and interpretability of financial fraud detection systems.

A. Machine Learning
Machine learning (ML) is key in detecting financial statement fraud by spotting complex patterns in data. In this research, ML is paired with Explainable AI (XAI) to make models more transparent and graph-based anomaly detection to catch sneaky fraud patterns. Traditional ML models, like decision trees or logistic regression, usually act like "black boxes" (mysterious, right?), but with XAI methods like SHAP and LIME, we get insights into why something looks fishy. Graph-based ML analyzes relationships between entities (companies, auditors, etc.), looking for weird patterns or unusual connections that scream fraud. This combo of ML, XAI, and graphs means better fraud detection and, more importantly, better explanations.

B. Explainable AI
Explainable AI (XAI) is crucial in this research for improving the transparency of fraud detection models. Traditional machine learning models often act as "black boxes," providing limited insights into their decision-making processes. XAI techniques, such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations), are used to make model predictions more understandable.By applying XAI, the model can explain why specific transactions or financial entities are flagged as fraudulent. This interpretability is essential for auditors, regulators, and stakeholders, as it increases trust in the model's output and ensures compliance with financial regulations. In combination with graph-based anomaly detection, XAI helps highlight suspicious patterns in financial data while making the detection process more transparent and actionable.

C. Anomaly Detection
In this research, anomaly detection serves as a foundational technique to identify instances of fraudulent behavior in financial data. It involves assessing the characteristics of various entities, such as companies and transactions, to pinpoint patterns that deviate from established norms. Traditional anomaly detection methods are employed, focusing on statistical approaches and machine learning algorithms to evaluate individual features of the data. The primary goal is to classify transactions or entities as either fraudulent or non-fraudulent based on their attributes.The implementation of anomaly detection in this context includes preprocessing steps such as data cleaning and feature engineering, which help enhance the quality and relevance of the data. By applying techniques such as the Iterative Imputer for handling missing values and SMOTE for addressing class imbalance, the research ensures that the dataset is prepared effectively for subsequent analysis. The results of the anomaly detection process inform the development of machine learning models, facilitating a comprehensive understanding of the features that contribute to fraud and improving predictive accuracy.

ISSN 2394 - 7780

i) Graph Based Anomaly Detection

Graph-Based Anomaly Detection (GBAD) is a pivotal aspect of the research aimed at detecting financial fraud through the analysis of relationships among entities. By constructing a graph where nodes represent companies and transactions, and edges signify their interconnections, GBAD effectively captures the complex interactions that can indicate suspicious behavior. The research utilizes Euclidean distances to create a distance matrix, which is then transformed into a graph structure using NetworkX.The Asynchronous Label Propagation Algorithm is applied to identify communities within the graph, revealing clusters of behavior that may be indicative of fraud. Anomaly scores are calculated based on the sizes of these communities, with smaller communities suggesting higher potential for fraudulent activity. By setting a threshold at the 95th percentile of these scores, the research successfully flags entities for further investigation. This approach enhances traditional anomaly detection methods by uncovering hidden relationships and unusual patterns that are often missed in standard analyses. The integration of GBAD with machine learning models significantly contributes to the robustness and effectiveness of the fraud detection system developed in this research.

D. Predictive Analysis

In this research, predictive analytics is employed through a structured approach. The dataset undergoes preprocessing, including the handling of missing values with an Iterative Imputer and the creation of key financial ratios like Debt-to-Equity and Return on Equity. A binary target variable, Fraudulent, is established based on net income conditions. To address class imbalance, SMOTE is implemented, generating synthetic data to enhance model performance. Multiple classifiers, including XGBoost, Random Forest, and Logistic Regression, are trained and optimized via GridSearchCV, focusing on metrics such as F1-score. Explainability techniques like SHAP and LIME offer insights into feature importance and individual predictions, ensuring transparency in the models. Additionally, graph-based anomaly detection is utilized, scoring anomalies through community detection in a constructed graph of pairwise distances. This multifaceted predictive analysis aims to achieve accurate fraud detection while emphasizing interpretability, which is vital for stakeholder trust in automated financial systems.

E. Technical analysis

Technical analysis is pivotal for data understanding and model evaluation. The process begins with data preprocessing, which includes addressing missing values with the Iterative Imputer and creating financial ratios to enhance feature representation. You utilize various machine learning algorithms, notably XGBoost, Random Forest, and Logistic Regression, with hyperparameter tuning via GridSearchCV to optimize performance metrics such as F1-score. Class imbalance is managed through SMOTE, which generates synthetic samples, enhancing model robustness. To interpret model behavior, SHAP and LIME are employed to provide insights into feature contributions, promoting transparency in predictions. Additionally, a graph-based approach is applied, using community detection to uncover relationships in the data. By constructing a graph from pairwise distances and scoring anomalies, this technique identifies patterns indicative of fraudulent activity. Together, these technical methods form a comprehensive framework for effective fraud detection, emphasizing the need for both predictive accuracy and interpretability in machine learning.

The following are a few of the tools used in analyzing the stock prices

i) Regression Analysis

In the context of detecting financial statement fraud, regression analysis serves as a fundamental statistical tool for modeling relationships between financial variables and predicting fraudulent behavior. This method involves establishing a mathematical framework to analyze how changes in independent variables (financial ratios, indicators, etc.) affect a dependent variable, typically the likelihood of fraud.

The following are some of its variants that are used in the paper

a. Ordinary Least Squares (OLS)

Ordinary Least Squares (OLS) regression is utilized to detect anomalies in financial statements by modelling the relationship between financial metrics and the occurrence of fraud. By minimizing the sum of squared residuals, OLS estimates parameters in a linear regression model, enabling analysts to assess how various financial ratios correlate with fraud risk. Analysts can explore the impact of key ratios, such as return on equity (ROE) and debt-to-equity (D/E), on fraudulent reporting likelihood, where significant deviations from expected values may indicate irregularities necessitating further investigation. Additionally, OLS helps identify outliers; substantial divergences between actual and predicted financial metrics may signal potential manipulation or fraud.

b. Logistic Regression

ISSN 2394 - 7780

Logistic regression enhances traditional regression by modeling binary outcomes, making it ideal for fraud detection. This technique estimates the probability of an event, such as fraud versus non-fraud, based on multiple predictor variables. In this research, logistic regression is employed to classify companies as fraudulent or non-fraudulent using financial indicators. The model yields probability scores that facilitate binary classifications, enabling targeted investigations of high-risk companies. Additionally, the coefficients from logistic regression reveal the influence of each financial ratio on fraud probability; positive coefficients indicate that higher ratios correlate with increased fraud likelihood, while negative coefficients suggest a lower probability

ii) Time Series Analysis

Time series analysis is a vital statistical technique employed in this research to detect irregular patterns and trends in financial data over time. It allows for the observation of how financial metrics evolve, aiding in the identification of anomalies that may indicate fraudulent activities. By monitoring historical performance, analysts can discern long-term trends, such as revenue growth or declining expenses, and recognize sudden deviations that raise concerns about potential fraud. Additionally, time series analysis helps identify seasonal patterns inherent in financial data, distinguishing normal fluctuations from genuine anomalies. This analysis also facilitates the detection of outliers—data points that significantly deviate from expected values, which may signal manipulation. Moreover, time series models like ARIMA enable forecasting of future values based on historical data, allowing for the comparison of actual outcomes with predictions, thus highlighting discrepancies that may suggest fraud. Understanding autocorrelation further enhances fraud detection by revealing predictable relationships over time. Ultimately, time series analysis significantly contributes to financial statement fraud detection by providing insights into the temporal dynamics of financial data, enhancing the capacity to uncover fraudulent activities and ensure the integrity of financial reporting.

a. ARIMA (AutoRegressive Integrated Moving Average)

ARIMA (AutoRegressive Integrated Moving Average) is a widely used statistical model for analyzing time-dependent data, capturing key aspects of financial time series, including trends, seasonality, and autocorrelation. In the realm of financial statement fraud detection, ARIMA aids in modeling the historical performance of crucial financial metrics, enabling analysts to forecast future values. By applying an ARIMA model to financial data, researchers can identify deviations from predicted trends. Notable discrepancies between actual observations and forecasted values may indicate potential manipulation or fraud, thereby necessitating further investigation.

b. Seasonal decomposition

Seasonal decomposition techniques are employed to break down time series data into its underlying components: trend, seasonality, and residuals. This analysis is instrumental in identifying abnormal fluctuations in financial metrics that may signal fraudulent behavior. By decomposing financial data, analysts can effectively isolate seasonal patterns from irregularities, providing a clearer understanding of normal fluctuations versus unusual spikes or drops. Such insights enhance the detection of potential fraud, as unexpected changes can be indicative of manipulative practices.

Both regression analysis and time series analysis provide powerful tools for modeling relationships and detecting anomalies in financial data. OLS and logistic regression are crucial for assessing how various financial ratios correlate with fraud, while ARIMA and seasonal decomposition help analysts identify and interpret patterns over time, providing further context for potential irregularities in financial statements. By leveraging these techniques, analysts can gain deeper insights into the financial health of organizations and identify instances of financial statement fraud more effectively.

F. Fundamental analysis

Fundamental analysis is essential in researching financial statement fraud detection, as it evaluates a company's intrinsic value through key financial metrics and ratios, such as revenue, earnings, cash flow, and debt levels. This approach helps identify discrepancies or anomalies that may suggest fraudulent activities. By analyzing fundamental aspects, researchers can uncover inconsistencies in financial statements that deviate from expected norms. Significant variations in key ratios, like the debt-to-equity ratio or return on equity, compared to industry benchmarks may indicate potential red flags. Additionally, fundamental analysis establishes a baseline for expected performance, allowing analysts to detect irregularities over time, such as unexplained spikes in revenue or sudden drops in expenses. Integrating fundamental analysis with advanced techniques, such as regression and time series analysis, enhances the ability to detect financial fraud, ensuring a comprehensive evaluation of a company's financial integrity and operational practices.

The following are some tools that are used in Fundamental Analysis

ISSN 2394 - 7780

i) Financial Ratios

Financial ratios are essential metrics that provide insights into a company's financial health and operational performance. Ratios such as the debt-to-equity ratio assess a company's leverage and risk, indicating how much debt is used to finance its assets relative to shareholder equity. The return on equity (ROE) measures profitability by assessing how effectively a company generates profits from its equity. These ratios help identify discrepancies that may suggest manipulation or fraudulent reporting.

a) Debt-to-Equity Ratio (D/E)

The Debt-to-Equity Ratio is a crucial measure of a company's financial leverage, calculated as the ratio of total liabilities to shareholders' equity. A high D/E ratio indicates that a company is significantly funded by debt relative to its equity, suggesting increased financial risk. In the context of fraud detection, unusually high or rapidly increasing D/E ratios may raise concerns about the company's financial health, potentially indicating manipulative practices such as inflating earnings or understating liabilities to present a more favorable image to investors.

D/E= Total Liabilities/Shareholders Equity

b) Return on Assets (ROA)

The Return on Assets ratio evaluates how efficiently a company utilizes its total assets to generate net income. It is calculated by dividing net income by total assets. A notably lower ROA compared to industry peers can be a red flag, signaling operational inefficiencies or potential earnings manipulation. This ratio serves as an indicator of whether a company is accurately reporting its profitability relative to its asset base, making it a vital tool for identifying discrepancies in financial statements.

ROA= Net Income/ Total Assets

c) Current Ratio

The Current Ratio assesses a company's ability to meet its short-term obligations using its short-term assets, calculated as current assets divided by current liabilities. A ratio below 1 indicates potential liquidity issues, suggesting that the company may struggle to pay off its short-term debts. Furthermore, significant fluctuations in the current ratio could signal possible manipulation of financial statements, as companies might adjust their reported assets or liabilities to present a healthier liquidity position than what truly exists.

Current Ratio= Current Assets/Current Liabilites

d) Return on Equity (ROE)

The Return on Equity ratio measures a company's ability to generate profit from its shareholders' equity, calculated by dividing net income by shareholders' equity. A consistently increasing ROE is typically viewed positively, indicating effective management of equity. However, an abrupt spike in ROE may indicate aggressive accounting practices or earnings management. This raises concerns about the authenticity of reported profits and the potential for financial statement manipulation, prompting further scrutiny of the company's financial practices.

ROE= Net Income/Shareholder's Equity

These ratios play a fundamental role in the research, providing key insights into a company's financial health and helping analysts identify potential red flags associated with financial statement fraud. By evaluating these metrics, researchers can better assess the likelihood of manipulative practices and take necessary investigative actions

METHODOLOGY

The following are the Machine Learning techniques used in our reviewed papers

XGBoost

XGBoost, an ensemble learning technique based on gradient boosting, is pivotal in fraud detection due to its ability to handle complex financial datasets with high dimensionality. In this research, XGBoost is applied to classify companies as fraudulent or non-fraudulent based on various financial indicators. Its gradient boosting framework improves prediction accuracy by iteratively combining weak learners and correcting errors from previous iterations. XGBoost's built-in regularization also helps prevent overfitting, making it well-suited for fraud detection, where patterns may be subtle and diverse.

Random Forest:

Random Forest, another ensemble method, is utilized in this research to enhance classification performance. By building multiple decision trees and averaging their results, Random Forest reduces variance and improves

ISSN 2394 - 7780

prediction reliability. In the context of fraud detection, Random Forest captures complex relationships between financial variables and fraud risk, as the trees in the forest collectively identify important patterns and anomalies in the data. This method is particularly useful when dealing with noisy or unbalanced datasets, as it naturally handles such challenges through its bagging approach.

SHAP (SHapley Additive exPlanations)
SHAP is employed in this research to interpret the outputs of complex models like XGBoost, offering a clear, feature-level explanation of why a particular company is classified as fraudulent or non-fraudulent. By calculating the SHAP values, each feature's contribution to the model's prediction is quantified, making it possible to understand the importance of specific financial variables. This method ensures transparency by attributing the probability of fraud to individual financial ratios, providing a global view of which features consistently influence fraud detection across the dataset. SHAP helps stakeholders, such as auditors and regulators, trust the machine learning model by clearly demonstrating how predictions are made.

LIME (Local Interpretable Model-Agnostic Explanations)
LIME is used in this research to generate local explanations for individual predictions, particularly focusing on explaining specific fraud detection cases. It approximates the machine learning model locally around the point of interest, showing which financial variables contributed most to a company's classification as fraudulent or non-fraudulent. LIME makes complex models more interpretable by providing a simpler, understandable explanation for each individual decision. This tool is especially valuable when investigating high-risk companies, as it allows auditors to see why a specific prediction was made and identify the key factors driving the fraud classification.

NetworkX:
NetworkX is utilized in this research to represent financial metrics and companies as nodes in a graph. Edges between nodes represent relationships or similarities between financial metrics, such as the Euclidean distance between their values. By leveraging graph structures, the research models complex relationships between companies' financial data, which would be difficult to capture using traditional tabular methods. NetworkX facilitates the creation and analysis of these graphs, enabling the detection of irregular patterns that could indicate fraud. It allows for the visualization and understanding of interconnected financial metrics and their role in predicting fraudulent behavior.

Community Detection Algorithms:
Community detection algorithms, such as Asynchronous Label Propagation, are applied to the graph structures created using financial data. These algorithms group nodes into communities based on similarities or connections. In the context of fraud detection, the assumption is that normal companies will cluster into similar communities, while anomalous or fraudulent companies will either belong to unusually small or large communities or exhibit patterns that deviate from the norm. Identifying these anomalous patterns within communities allows for the detection of potential fraud, as companies with financial behaviors that differ significantly from their peers are flagged for further investigation.

SMOTE (Synthetic Minority Over-sampling Technique):
SMOTE is used in this research to address the inherent class imbalance in fraud detection datasets, where fraudulent cases typically represent a small portion of the overall data. Without addressing this imbalance, machine learning models may become biased toward the majority class (non-fraud cases). SMOTE generates synthetic examples of the minority class (fraud cases) by creating new samples based on the feature space of the existing minority instances. This helps ensure that the models trained on the resampled dataset can learn more effectively from the minority class, improving the model's ability to detect fraudulent activities.

Iterative Imputer:
The Iterative Imputer is applied to handle missing data in the dataset, which is common in real-world financial datasets. Missing values are filled in by iteratively modeling them based on the relationships between the other variables in the dataset. In this research, Iterative Imputer ensures that missing financial metrics do not undermine the model's performance. By imputing missing values with reasonable estimates, the technique preserves the integrity of the dataset and allows for more accurate analysis and model training.

Permutation Importance
Permutation Importance is used in this research to assess the importance of individual financial features in predicting fraud. By randomly shuffling specific features and observing how the model's performance changes, it quantifies the impact each feature has on the predictive accuracy. In the context of financial statement fraud detection, Permutation Importance helps identify which financial metrics (such as debt ratios or asset returns)

ISSN 2394 - 7780

play a critical role in the classification of fraudulent vs. non-fraudulent cases.In this research, Permutation Importance aids in refining the machine learning models by pinpointing the most influential financial variables, thereby enhancing the interpretability of the model and improving its ability to target key indicators of fraud.

DATASET REVIEW

The literature papers reviewed had the following data sets

i) The Zacks Fundamentals Collection

The dataset comprises 84 columns and 1,534 rows, capturing financial data for various entities across specific time periods. Each row represents a unique company, while the columns encompass essential financial metrics, ratios, and data points pertinent to financial analysis.

Key identification columns include CoNo (a unique company identifier), CoName, CoType (company category), Exchange (stock exchange), and FiscalYearEnd (fiscal year-end month). Financial metrics such as Revenue (Rev), Net Income, Earnings Per Share (EPS), Dividend Yield (DivYield), and Book Value Per Share provide vital insights into company performance. Operational performance is assessed through metrics like Operating Income, EBIT, and EBITDA.

The dataset also includes critical ratios: profitability ratios (e.g., Return on Equity (ROE), Return on Assets (ROA)), leverage ratios (Debt to Equity, Current Ratio, Quick Ratio), and valuation ratios (Price-to-Earnings (PE), Price-to-Sales, Price-to-Book), which indicate market valuation relative to earnings, sales, and book value. Additionally, market performance data such as 52-Week High, 52-Week Low, Stock Price, and Volume offer a view of each company's market standing, while Beta measures volatility for stock analysis.

This rich dataset serves as a valuable resource for conducting fundamental analysis, enabling researchers and analysts to derive insights into company performance and market trends.

DATA PREPROCESSING

Prior to utilizing the dataset for analytical tasks, several preprocessing steps may be necessary to enhance data quality and model performance.

First, missing values can be an issue, as certain columns may contain incomplete data. To address this, imputation techniques such as the Iterative Imputer may be employed, or alternatively, rows or columns with substantial missing data can be removed to maintain dataset integrity.

Next, feature scaling is essential, given the wide range of financial metrics present in the dataset. Scaling is particularly important for machine learning models, which can be sensitive to the magnitude of input features.

Lastly, if the dataset is intended for fraud prediction analysis, it may exhibit class imbalance, characterized by a significantly lower number of fraudulent companies compared to non-fraudulent ones. To mitigate this issue, techniques like Synthetic Minority Over-sampling Technique (SMOTE) can be utilized to balance the class distribution, thereby improving model robustness and predictive accuracy.

These preprocessing steps are critical to ensure that the dataset is ready for effective analysis and modeling.

Implementation

The implementation for detecting and predicting financial fraud using Explainable AI and Graph-Based Anomaly Detection involves several systematic steps that integrate data preprocessing, exploratory data analysis, feature engineering, and advanced modeling techniques.

Initially, the dataset is loaded using pandas, followed by an exploration phase where the first few rows are displayed, and missing values are assessed. A heatmap visualization highlights the distribution of missing values, ensuring awareness of data quality before analysis.

To enhance the dataset's completeness, missing values are addressed using the Iterative Imputer, which employs a model-based approach to fill in gaps. Additionally, new financial ratios, such as Debt-to-Equity and Return on Equity, are engineered to provide deeper insights into company performance. A binary target variable, Fraudulent, is created based on a specified condition, such as negative net income, to facilitate fraud detection.Exploratory Data Analysis (EDA) follows, utilizing various visualizations like count plots and correlation matrices to understand feature distributions and relationships. A pairplot illustrates the characteristics of fraudulent and non-fraudulent entities, offering valuable insights into the dataset.Given the potential class imbalance—where fraudulent cases are significantly fewer than non-fraudulent ones—Synthetic Minority Over-sampling Technique (SMOTE) is employed. This technique generates synthetic samples to balance the dataset, ensuring robust model training.For the anomaly detection component, the code calculates

ISSN 2394 - 7780

pairwise distances between samples to create a graph representation of the data. Community detection algorithms, such as asynchronous label propagation, are applied to identify anomalies. Anomaly scores are computed based on community sizes, and a threshold is established to classify potentially fraudulent companies.Model training is conducted using the XGBoost classifier, with hyperparameter tuning facilitated through GridSearchCV to optimize model parameters like the number of estimators and learning rate. To enhance model interpretability, SHAP (SHapley Additive exPlanations) values are utilized to explain the contributions of features to the model's predictions. Additionally, LIME (Local Interpretable Model-agnostic Explanations) provides local explanations for individual predictions, allowing for a detailed understanding of specific instances.Feature importance is further evaluated using permutation importance, which assesses the significance of each feature in predicting the target variable, and the results are visualized in a bar plot for clarity.The performance of the models is thoroughly evaluated using metrics such as accuracy, precision, recall, F1-score, ROC-AUC, and balanced accuracy for the XGBoost classifier, Random Forest, and Logistic Regression. A Voting Classifier is also implemented, combining the predictions from all models to leverage their strengths, ultimately enhancing overall performance.This comprehensive implementation effectively integrates multiple methodologies, combining traditional machine learning approaches with advanced graph-based techniques to enhance the detection of fraudulent financial activities. The incorporation of Explainable AI tools like SHAP and LIME ensures transparency in model predictions, which is crucial for stakeholders in the financial sector. Overall, the systematic approach to preprocessing, feature engineering, model training, and evaluation establishes a robust framework applicable to real-world financial datasets.

RESULTS

Table 1 Financial Metrics

Metric	Value
Debt_to_Equity	0.084718
Return_on_Equity	274.96786

Table 2 Model Evaluations

Model	Accuracy	Precision	Recall	F1 Score	Roc Auc
XG BOOST	1.000	1.0	1.0	1.0	1.00
Random forest	1.000	1.0	1.0	1.0	1.00
Logistic Regression	0.931	0.0	0.0	0.0	0.89
Voting Classifier	1.000	1.0	1.0	1.0	1.00

The company exhibits a solid financial foundation, highlighted by a low debt-to-equity ratio of 0.0847, indicating conservative financing primarily through equity, which minimizes financial risk. Additionally, the exceptionally high return on equity of 274.97% reflects the company's efficiency in generating profit relative to shareholders' equity, making it attractive to potential investors, though the sustainability of such returns warrants further examination. In terms of model evaluations for fraud detection, XGBoost, Random Forest, and Voting Classifier demonstrate perfect performance across all metrics, achieving an accuracy, precision, recall, F1 score, and ROC AUC of 1.000, effectively identifying fraud cases without errors. Conversely, Logistic Regression shows notable deficiencies, with an accuracy of 0.931 but a precision and recall of 0.0, indicating failure to predict any positive fraud cases. These insights underscore the company's financial stability while highlighting the efficacy of advanced machine learning models for effective fraud detection.

CONCLUSION

In this research, we have conducted a comprehensive analysis of financial statement fraud detection, employing various advanced machine learning models to effectively identify fraudulent activities within financial data. Our exploration began with a thorough examination of financial metrics, revealing a robust financial standing characterized by a low debt-to-equity ratio and an exceptionally high return on equity, indicating efficient profit generation with minimized financial risk.

We then focused on implementing and evaluating multiple machine learning models, including XGBoost, Random Forest, Logistic Regression, and a Voting Classifier. The results demonstrated that XGBoost, Random Forest, and the Voting Classifier achieved outstanding performance with perfect scores across all evaluation metrics, showcasing their effectiveness in accurately detecting fraud cases. In contrast, Logistic Regression

ISSN 2394 - 7780

highlighted significant limitations, failing to identify any fraudulent instances despite a reasonable overall accuracy.

The challenges addressed in this research included the need for reliable fraud detection methods in financial statements, particularly in a landscape where traditional approaches may fall short. By leveraging explainable AI and graph-based anomaly detection techniques, we have developed a more nuanced understanding of fraud patterns, contributing to more effective prevention strategies. Overall, this study underscores the importance of integrating sophisticated machine learning techniques into financial analysis to enhance fraud detection capabilities, ultimately fostering greater transparency and trust in financial reporting. Future work may explore the sustainability of these findings and the applicability of the models across diverse datasets and industries.

REFERENCES

[1] E. F. Brigham and M. C. Ehrhardt, Financial Management: Theory & Practice, 15th ed. Cengage Learning, 2016.

[2] S. A. Ross, R. W. Westerfield, and B. D. Jordan, Fundamentals of Corporate Finance, 12th ed. McGraw-Hill Education, 2019.

[3] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.

[4] C. M. Bishop, Pattern Recognition and Machine Learning. Springer, 2006.

[5] D. D. M. B., "A survey of machine learning techniques for financial fraud detection," International Journal of Data Mining & Knowledge Management Process (IJDKP), vol. 10, no. 2, pp. 1–15, 2020.

[6] N. V. Chawla and A. Davis, "Improved data mining techniques for fraud detection," Journal of the American Statistical Association, vol. 97, no. 457, pp. 427–431, 2002.

[7] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 785–794.

[8] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[9] D. R. Cox, "The regression analysis of binary sequences," Journal of the Royal Statistical Society: Series B (Methodological), vol. 20, no. 2, pp. 215–232, 1958.

[10] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," Information Processing and Management, vol. 45, no. 4, pp. 427–437, 2009.

[11] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, and correlation," in Proceedings of the 23rd International Conference on Machine Learning (ICML), 2011, pp. 1–5.

[12] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you? Explaining the predictions of any classifier," in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1135–1144.

[13] M. Ahmed, A. N. Mahmood, and J. Hu, "A survey of network anomaly detection techniques," Journal of Network and Computer Applications, vol. 60, pp. 120–137, 2016.

[14] X. Wu and H. Liu, "Graph neural networks for anomaly detection in financial fraud," ACM Transactions on Intelligent Systems and Technology (TIST), vol. 12, no. 5, pp. 1–25, 2021.

[15] S. Bhattacharyya, S. Jha, and S. Tharakaram, "Data mining for fraud detection: A case study in banking," International Journal of Computer Applications, vol. 33, no. 7, pp. 1–5, 2011.

[16] W. Wu and J. Hu, "Financial fraud detection: A survey from the data mining perspective," Journal of Finance and Data Science, vol. 1, no. 1, pp. 30–45, 2015.

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

REINFORCEMENT LEARNING IN PREDICTIVE ANALYTICS FOR HUMAN BEHAVIOUR

[1]Aliasgar Sogiawala and [2]Rohana Deshpande
[1]Bachelor of Science in Information Technology, Jai Hind College Mumbai, India
[2]Assistant Professor Jai Hind College, Mumbai, India

ABSTRACT
The development of the internet hasled to the digitization of data, opening up opportunities in big data. This digital footprint reveals extensive insights into what people read, how they behave and their involvement in various activities, shedding light on their likes and dislikes. The goal is to use predictive analytics to optimize the predicition of human behaviour for strategic insights and decision- making. Predictive analytics contains a range of statistical and analytical techniques designed to enable organizations to forecast future outcomes and events with a degree of certainty based on historical data and interactions. This process can be enhanced by employing self-learning algorithms and models that predict human behaviour from data-driven insights and refine their predictions based on feedback and results,which is the central process of Reinforcement Learning(RL). RL is poised to revolutionize the field of AI and aims to automate systems with an advanced understanding of the virtual world and tackle problems previously consideredintractable This paper starts with an introduction to the general field of RL, followed by an exploration of value- based and policy- based methods.It then delves into the core concept of predictive analytics and its essential components. We discuss the main algorithms of RL and their connection to human operant learning, demonstrating why RL-based predictive analytics is better than the traditional way of it. Proceeded by applications and benefits of using RL in predicting human behaviour . In conclusion, we propose the use of RL models in predictive analytics to predict human behaviour through in formatics and analytics approach, aiming to gain deeper insights into human behaviour to enhance decision-making and strategic insights.

Keywords: *Reinforcement learning, predictive analytics,big data, human prediction , self learning algorithms*

1. INTRODUCTION
In life, there are many situations in which we learn by interacting with our environment. We assess the current situation, and then take suitable actions and observe the outcomes and learn from them to implement future actions.This iterative, feedback- driven attributes in machine learning are central to Reinforcement Learning (RL), a technique that enables systems to optimize decision making through continuous interaction with an environment.

RL is a specialized field of machine learning focused on decision making, where an agentl earns by receiving rewards or penalties based on its actions. Over time, the agent refines its strategies so as to maximize cumulative rewards, a process that mirrors human learning in complex situations. The agent balances exploration-testing new actions to gain more knowledge-and exploitation- leveraging known actions to maximize immediate gains- and thus helps RL systems solve intricate problems.This makes RL particularly effective for tasks like robotics, game strategy, where adaptability and feedback loops are essential.

As digitization has surged, vast amounts of data now offer insights into human behaviour.Through digital footprints, patterns in individual preferences, actions and interactions are revealed, paving the way for predictive analytics. Predictive analytics utilizes statistical and analytical techniques to forecast future events based on historical data, whose potential can be vastly expanded with RL.In context of human behaviour, RL can reveal deeper insights into how individuals make choices,respond to stimuli and develop preferences.

II. REINFORCEMENT LEARNING
As explained earlier, RL is a "trial and error" approach to choosing the best decision by rewarding or punishing itself. Let us proceed to how RL exactly functions. In the RL setup, an autonomous agent, controlled by a machine- learning algorithm, senses a state st from its world at time step t. The agent acts upon the world by taking an action at in states st. When the agent takes an action, the world and the agent transition to a new state, st+1, based on the current state and the action taken. The value of this state transition is obtained by the agent from a scalar reinforcement signal, r. The behaviour, B, of the agent ought to choose actions that have a tendency to maximize the long- run sum of values of the reinforcement signal.[7]

KEY ELEMENTS OF THE RL MODEL
RL can be described using a **Markov Decision Process (MDP)**, which consists of:

ISSN 2394 - 7780

i) **A set of states S** ,representing possible situations the agent could face.[7]

ii) **A set of actions A**, describing all choices available to the agent in each state.[7]

iii) **Transition dynamicsT(st+1|st,at)** showing How likely the environment is to move to a New state based on the current state and chosen action.[7]

iv) **Reward function R(st,at,st+1)** assigning a score for each action taken.[7]

- **Discount factor$\gamma \in [0,1]$,**which control show much future rewards matter compared to immediate rewards.[7]

The best sequence of actions is determined by the rewards provided by the environment. Every time the environment transitions to a new state, it also provides a scalar reward rt+1 to the agent as feedback. The agent's objective is to learn the best policy (a rule for choosing actions)π which maximizes the total reward.

For the agent to learn effectively, it needs to find a balance between **exploration** (trying out new actions to learn more about the environment) and **exploitation** (sticking with actions that have worked well in the past). This balance is crucial: if the agent exploits too much, it might miss out on discovering even better options,but if it explores too much, it could miss chances to earn higher rewards with what it already knows. Strategies like **epsilon-greedy** and **Upper Confidence Bound (UCB)** help the agent strike the right balance between exploring and exploiting.

If states are not fully observable, **Partially Observable MDPs(POMDPs)** can be used, where the agent relies on observations rather than states. An example of how RL model works is given below

Environment: You are in state 10. You have 3 possible actions.

Agent: I'll take action 1.

Environment: You received a reinforcement of 2 units. You are now in state 20. You have 2 possible actions.

Agent: I'll take action 2.

Environment: You received a reinforcement of -1 unit. You are now in state 10. You have 3 possible actions.

Agent: I'll take action 3.

Environment: You received a reinforcement of 5 units. You are now in state 30. You have 4 possible actions.

Figure1. Working of a RL agent

A) Methods in RL

There are mainly 2 methods in Reinforcement Learning, namely Value-Based RL and Policy- Based RL.

Value-Based RL methods focus on estimating the value of each action in different states, guiding the agent to choose actions that maximize cumulative rewards based on these values.The goal is to learn a **value function** that maps each state or a state- action pair to an expected reward guiding the agent to maximize its cumulative rewards by choosing actions that have the highest estimated value.

- **State-Value Function V(s)**: Estimates the expected reward of being in a state sand following the policy thereafter.
- **Action-Value Function Q(s,a):** Estimates the expected reward of taking action a in state s and following the policy.

Some common algorithms of **value based RL** are **Q-Learning** and **SARSA.**

Whereas, **Policy-based methods** directly optimize the **policy**(a mapping from states to actions)without requiring a value function. Instead of estimating future rewards, the agent learns a policy that maximizes expected cumulative rewards, making it a more flexible approach, especially for environments with continuous or complex action spaces.

- **Policy Gradient**: A gradient-based approach that optimizes the policy by following the gradient of expected rewards with respect to policy parameters.

ISSN 2394 - 7780

Actor-Critic Methods: Combines policy-based and value-based approaches,where the **actor** learns the policy, and the **critic** evaluates it. This combination can stabilize learning by reducing variance in policy updates.

In this paper, we will be focusin mainly on policy- based methods since a **policy-based** reinforcement learning (RL) approach is generally more suitable, especially when dealing with complex, continuous, or unpredictable action spaces, which are often inherent in human behavior modeling.

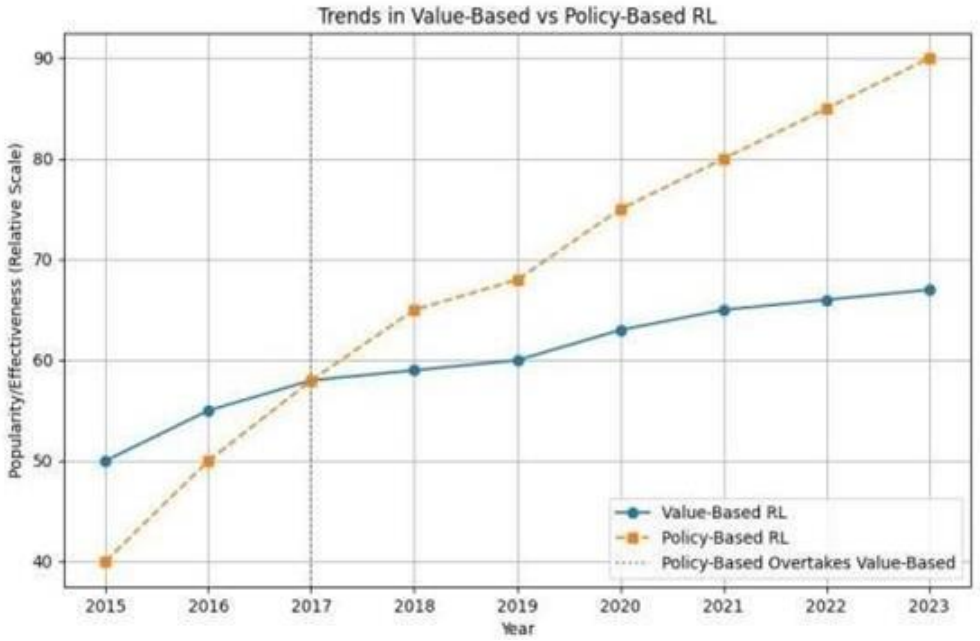


Figure2 Value based vs Policy based trends

The figure above shows the graphical representation of the effectiveness of the two RL methods over the past decade.

III. PREDICTIVE ANALYTICS

Now, predictive analytics is not a new concept to mankind,it has been in use for a while and has been used successfully by many large companies from which financial services and supermarket retailers are one of the biggest users. Due to the advent of big data , there is now a new-found appreciation of predictive analytics with a desire by many corporate organisations predict future outcomes with a high level of confidence and likewise twist and turn their strategies.[10]

Predictive analytics is a field with advanced analytics that involves forecasting future events by analyzing past and current data. By leveraging techniques from statistics,dataining and machine learning , predictive analytics allows organizations to make informed predictions about likely future behaviors or events.This enables businesses to be proactive , make strategic decisions and anticipate trends or behaviours that can impact their operations.For instance, an e-retailing company might use predictive analytics to identify customer purchasing patterns based on seasonal demand. It can track product interest, price sensitivity, and the influence of promotional offers to better understand consumer behavior[5]. By analyzing this data, the company can tailor recommendations, adjust pricing strategies, and improve marketing efforts to optimize customer engagement and sales.

Predictive analytics has applications across various domains such as:

1) **Banking and Financial Services**: Detecting fraud, assessing credit risk, and forecasting stock performance.
2) **Retail**: Customer behavior prediction, pricing strategy optimization, and inventory management.[5]
3) **Healthcare and Insurance**: Identifying at- risk individuals for preventive care and forecasting insurance claims.

1. PROCESS OF PREDICTIVE ANALYTICS

Predictive Analytics takes place in the following manner:

a. ***Requirement Collection***: Start by clarifying what the prediction should achieve and how it will provide value for the client's goals.[5]

ISSN 2394 - 7780

b. ***Data Collection***:Collect all necessary data from various sources, including both organized (structured) and messy (unstructured) formats.[5]

c. ***Data Analysis and Preparation***: Clean up and organize the data, ensuring it's accurate and ready for the next steps, so the model has reliable information to learn from.[5]

d. ***Applying Statistics and Machine Learning:*** Use statistical tools and machine learning techniques to find patterns in the data,laying the ground work for the predictive model.[5]

e. ***Building the Predictive Model***: Develop and test the model using sample data to ensure it can reliably make predictions on new information.[5]

f. ***Prediction and Monitoring***:Put the model to work in real-time for regular predictions,keeping an eye on its accuracy and creating reports for in formed decision- making.[5]

2. TECHNIQUES IN PREDICTIVE ANALYTICS

a. ***Decision_Trees:***

A decision tree is a classification model but it can be used in regression as well.It is a tree-like model which relates the decisions and their possible consequences

. The consequences may be the outcome of events,cost of resources or utility.In its tree-like structure, each branch represents a choice between a number of alternatives and its every leaf represents a decision. Based on the categories of input variables, it partitions data into subsets. It helps the individuals in decision analysis. Ease of understanding and interpretation make the decision trees popular to use.[5]

2. **Regression Models**:

Regression analysis predicts outcomes by examining relationships between a dependent variable and one or more independent variables.It's highly effective for continuous data,making it valuable for forecasting trends, like predicting sales figures or price trends based on other factors.

c. **Neural Networks**:

Inspired by the human brain, neural networks are layers of nodes(neurons) that process complex data patterns.

They're particularly powerful for handling non-linear relationships and large datasets, making them ideal for tasks like image recognition and advanced behavior predictions.[5]

4. ***Support Vector Machines (SVM):***

SVMs separate data points in high- dimensional space with a boundary that best divides different categories. Often used in classification tasks, SVMs are powerful for tasks requiring clear distinctions between classes, such as detecting fraudulent transactions or classifying images.[5]

e. **Bayesian Statistics**:

Bayesian models use probability to make predictions based on prior data and newly observed data. By continuously updating as new data is added, they're well-suited for tasks where initial assumptions evolve, like spam detection and medical diagnosis.[5]

6. **Clustering Techniques**:

Clustering finds natural groupings within data by segmenting similar items into clusters. It's ideal for customer segmentation or market research, where understanding different user types can help target personalized strategies.

g. **Time Series Analysis**:

Time series analysis focuses on data points overtime , recognizing patterns to make temporal predictions. It's used for forecasting stock prices, weather conditions, and inventory needs by tracking patterns in sequential data.

8. **Ensemble Methods**:

Ensemble methods, like Random Forests, combine multiple models to improve prediction accuracy. By leveraging the strengths of individual models, ensembles are robust and reduce overfitting, making them ideal for high-stakes decisions, like risk assessment and fraud detection.[5]

9. ***K-Nearest Neighbors (KNN):***

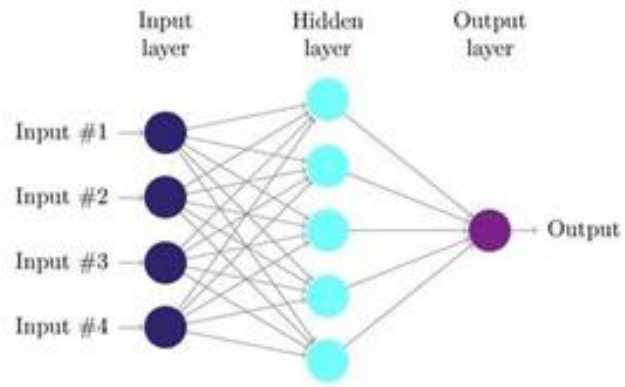
KNN classifies data based on the closest data points around it, making it simple yet effective for smaller datasets. It's commonly used in recommendation systems and basic pattern recognition tasks, like identifying similar images. [5]

ISSN 2394 - 7780

In this paper we will be using the **Neural Networks** technique to predict human behaviour.

C) Predictive Analytics in Human Behaviour

In the context of human behavior, predictive analtytic is a generic name for applying mathematical techniques to predict human behavior. The result is usually a score or a code for each person, reflecting the probability of their behavior in the future. An example would be a score used to reflect the probability of them buying another product from a business. The score or code may also segment individuals into groups who might have different management and communication needs, or for a business's products and services. The score can be attained or equivalently this behaviour can be predicted using the approach of neural networks as below:[9]



Neural networks – Building blocks of Data Analysis

1) Data Collection and Behavior DataConstruction

In order to predict human behavior using a neural network, the whole process begins with raw data being transformed into insights that are more specific to the behavior that a model can easily understand and interpret. Most of this is done by gathering behavioral data relevant to the case being studied, such as what a user has purchased or browsed and how many times they have engaged within a given time frame. For instance, in the retail environment, this would be the frequency of purchases or categories of products browsed or time spent on an item. After gathering all this information, it then gets categorized and marked in order to indicate specific behaviors, such as marking a user as a "frequent buyer" or identifying those with a special interest in electronics. This way, we are providing the model with clear behavioral attributes that will enable it to make proper predictions.[9]

2) Behavior Pattern Analysis and Feature Engineering

Once we've gathered our data, the next step is to identify key behavior patterns and convert them into a format that our neural network can use. This involves feature engineering, where we extract useful information from the data. For example, we might look at how long it's been since a user's last purchase, how diverse their purchases are, or how responsive they are to discounts. After defining these features, we further prepare the data by scaling and normalizing values and converting categorical data, such as product categories, into numerical form. This ensures that the data is compatible with neural networks, allowing the model to easily recognize important patterns in user behavior. [9]

3) Selecting Neura l Network Architecture

When choosing the best neural network structure in predicting human behavior, often included are **Recurrent Neural Networks (RNNs)** and **Long Short-Term Memory networks (LSTMs)**, both of which are excellent in the analysis of sequential data:

- **RNNs:** Suitable for sequential data, RNNs analyze patterns over time, and thus are more suitable for monitoring changes in user behavior, like recurring purchases.
- **LSTMs**: LSTMs are a kind of RNN that can be used to memorize long dependencies, hence catching sustained trends of user behavior. For instance, LSTMs can be of great use when detecting patterns related to customer engagement or purchasing frequency over an extended period, something that other models would not catch.

Sometimes, CNNs may be useful when we need to analyze structured data, such as user interactions on a website or product recommendations based on image patterns. While CNNs are mostly applied to visual data, they can sometimes make behavioral predictions better by providing more context.

ISSN 2394 - 7780

4) **Building the Predictive Model**

With the proper architecture in place, we will start building the predictive model. The network is made up of different layers:

- **Input Layers:** These layers accept user behavior data and pass it on to be processed further.
- **Hidden Layers (eg:LSTM_layers):** The network here learns about sequence dependencies so that it can pick up patterns such as repeated purchases or v isits.
- **Output Layer**: This layer generates the output prediction, and this might be a probability of a certain outcome (e.g., probability of a purchase) or aclassification (e.g., high- or low- engagement users).

In training, the model is trained on labeled data, for example, past buying behavior or user activity levels. Depending on the type of prediction we are doing, we use a specific loss function, for example, **categorical cross-entropy** for classifying outputs or **mean squared error (MSE)** for numeric prediction.

5) **Model Evaluation an dFine-Tuning**

Once trained, you would then test the model to see if it works as desired. **Accuracy**, **precision**, **recall**, and **F1-score** are measures we can use to test how well the model performs in classifying user behavior. For numerical prediction, we can utilize **root mean square error (RMSE)** to give us feedback on the accuracy of the prediction.

To improve the model's performance:

- **Hyper parameter Tuning:** We then tune parameters such as the learning rate, batch size, and number of LSTM units to observe how well they work together.
- **Regularization:** Methods like dropout, where nodes are turned off at random during training, and early stopping avoid overfitting of the model, thus improving its generalization to new data.

6) **Deploying and Monitoring the Model**

After establishing confidence in the accuracy of the model, it is then ready for rollout into a production environment. This allows the model to process new data in real-time and provide predictions on continuous behavior. Continuous observation of the model is a must, given that user behavior tends to change, especially when there are trends or user preferences changes. With the monitoring of the model's performance, we can retrain the model periodically, thus making sure that the model is capable of making precise predictions as trends in behavior change.

IV. RL in PREDICTIVE ANALYTICS

Predictive analytics has been a very effective means of understanding and predicting human behavior, providing insightful information that informs wiser decision-making. However, the precision of such predictions can be enhanced through the incorporation of **Reinforcement Learning (RL)**. With **policy-based RL**, our models are not limited to offering a single prediction; rather, they learn and adapt over time. As we receive new information and feedback, these models reconfigure, hence ensuring that our predictions are valid and in harmony with the dynamics of evolving behaviors.[11]

To achieve this adaptability, we will implement simple policy-based reinforcement learning algorithms like **Policy Gradient** and **Actor-Critic**. These algorithms enable the model to experiment with new actions while at the same time rewarding those actions that consistently result in positive outcomes. By combining these methods, we can create a dynamic and adaptive predictive model that not only responds to changes but also continuously updates its understanding of human behavior.

Policy gradient algorithms are probably the most common category of continuous action reinforcement learning algorithms. The basic idea behind these algorithms is to adjust the parameters θ of the policy in the direction of the performance gradient $\nabla_\theta J(\pi_\theta)$.The fundamental result underlying these algorithms is the policy gradient theorem[2] , $\nabla_\theta J(\pi_\theta)=\int_S \rho^\pi(s)\int_A \nabla_\theta \pi_\theta(a|s) Q^\pi(s,a) da ds = E_{s\sim\rho^\pi, a\sim\pi_\theta}[\nabla_\theta \log \pi_\theta(a|s) Q^\pi(s,a)]$

The policy gradient is surprisingly simple. In particular, despite the fact that the state distribution $\rho^\pi(s)$ depends on the policy parameters, the policy gradient does not depend on the gradient of thestate distribution.[2] Policy gradients are mainly useful for complex, continuous actions paces where defining a fixed set of actions is challenging, making them suitable for tasks with high- dimensional behaviors.

ISSN 2394 - 7780

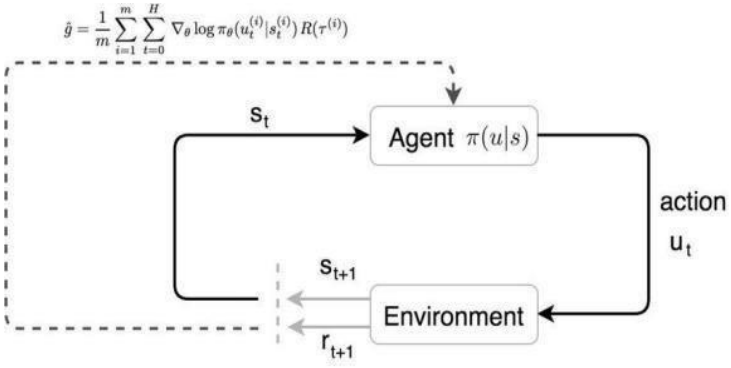


Figure3 Working of Policy Gradient

The **actor-critic** isa widely used architecture based on the policy gradient theorem .The actor-critic consists of two eponymous components. An actor adjusts the parameters θ of the stochastic policy πθ(s)by stochastic gradient ascent. Instead of the unknown true action-value function Qπ(s,a), an action-value function Qw(s,a) is used, with parameter vector w. A critic estimates the action- value function Qw(s,a)≈Qπ(s,a) using an appropriate policy evaluation algorithm such as temporal-difference learning.[2] The actor-critic approach combines the strengths of policy-based and value-based methods,where the actor improves the policy and the critic stabilizes learning by evaluating actions, reducing variance and improving the efficiency of updates.

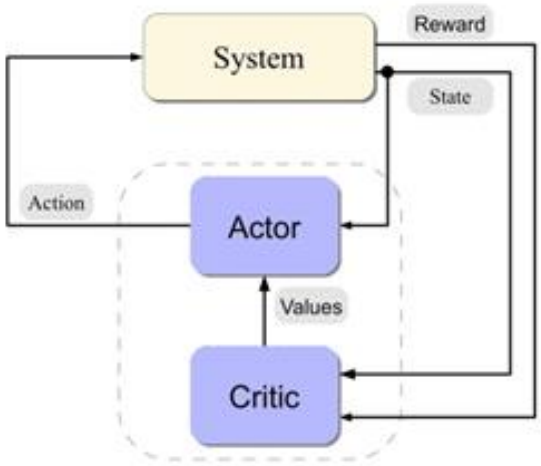


Figure4 Actor-Critic System

A) Integrating Policy-Based RL into the Predictive Model for Human Behavior

Instead of making static, one-time predictions, these algorithms help the model learn and adjust in real-time as new behavior data comes in. This way, the model can stay responsive to the latest trends in user actions—essential for understanding human behavior, which is often unpredictable and constantly changing.

The **Policy Gradient** algorithm lets the model continuously fine-tune its recommendations based on what's happening in the present, not just the past. For example, if users suddenly start showing interest in a different product category, Policy Gradient enables the model to pick up on this shift and prioritize those interests in its recommendations. This ongoing adaptation makes the model capable of dealing with complex, dynamic behavior patterns in which the best action is not necessarily specified or clear .[6]

Actor-Critic brings an added layer by balancing experimentation with refinement. In this setup, the **actor** tries out different recommendations or actions, like suggesting a new product, while the **critic** evaluates how well those actions are working. As more data comes in—like clicks, purchases, or time spent on a recommendation—the critic helps the actor improve by highlighting what's working best.[6] This continuous feedback loop means the model isn't just reacting to changes but learning from each interaction, which makes it smarter with every recommendation.[6]

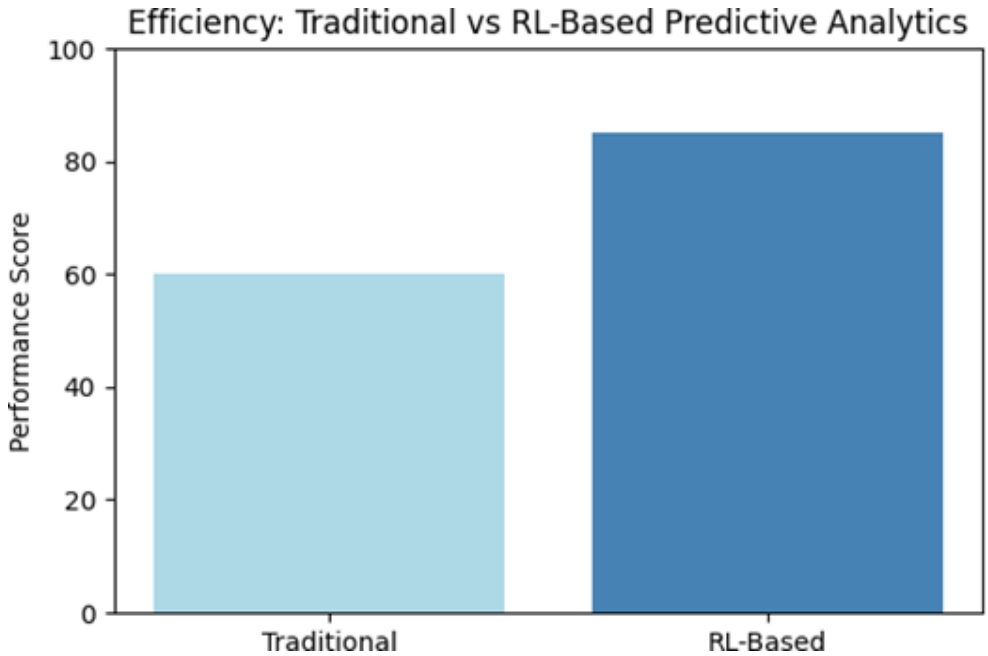
Collectively, these reinforcement learning methods turn our predictive model from merely a forecasting tool into an adaptive one; they enable it to learn to tailor itself to each user's individual behavior patterns. Through learning and improving its method over time, the model gets increasingly more skilled at deciphering and acting upon individual tastes, maintaining its forecasts accurate and attuned to real- time behavior.

ISSN 2394 - 7780

V. TRADITIONAL VS RLBASED PREDICTIVE ANALYTICS

A) **Efficiency**:

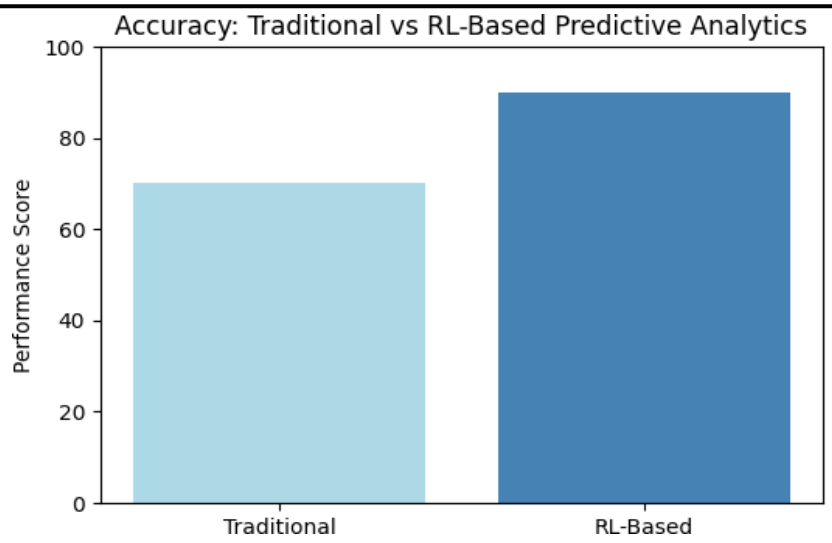
- **Traditional Predictive Analytics**: Classic models forecast based on historical data but have to be repeatedly trained to update their observations in real time, particularly when working with large data. The procedure may be resource- and time- consuming.
- **RL-Based Predictive Analytics**: Reinforcement learning models are less reliant on infrequent updates. They are designed to learn from real-time interaction and to adapt to new information as it comes in. This makes them more effective at staying up to date with behavior changes without the need for constant full-scale retraining.



B) **Accuracy**:

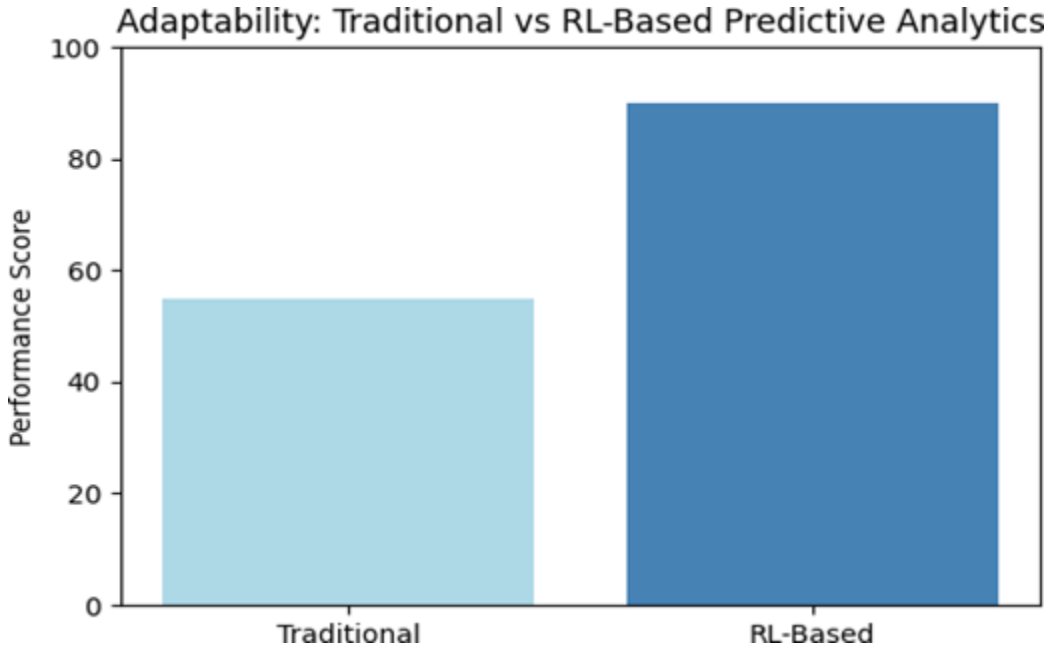
- **Traditional Predictive Analytics**: Traditional models are accurate but with static data that does not fluctuate much with time. Since behavior patterns are dynamic, the predictions of the model may not be accurate unless it is retrained with fresh data.
- **RL-Based Predictive Analytics**: RL- based models become more accurate as they learn through ongoing feedback. By reinforcement learning, the model becomes more responsive to shifting patterns so it can continue to improve predictions even when patterns of behavior change in the long run.

ISSN 2394 - 7780



C) **Adaptability**:

- **Traditional Predictive Analytics**: Classic models perform best in forecasting behaviors within established parameters, but they can be challenged when unexpected changes in behavior take place. They're often tied to the data they were trained on and less able to withstand when things begin to appear different.
- **RL-Based Predictive Analytics**: RL- based models adapt to learn. Using policy- based algorithms like Policy Gradient and Actor-Critic, they are able to test novel actions and identify best responses when new behavior patterns develop. Being able to adapt is especially applicable where human action needs to be forecast, and this may strongly change and move quickly.

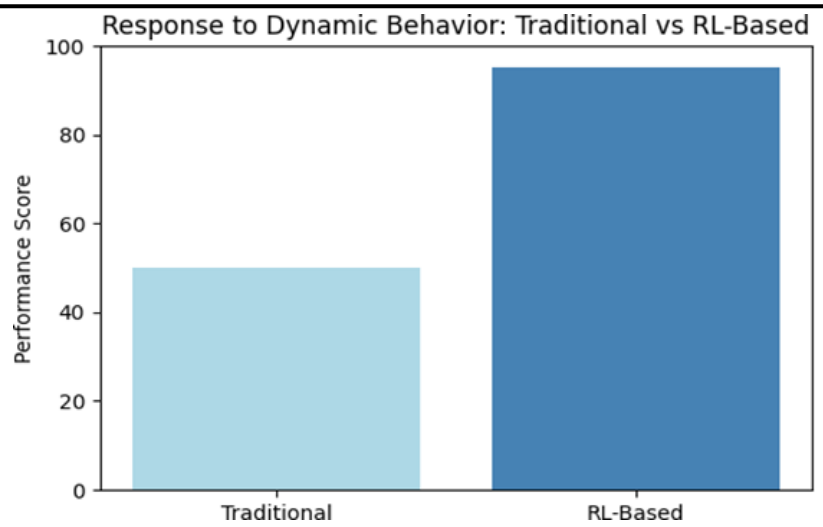


D) ***Response to Dynamic Behavior****:*

- **Traditional Predictive Analytics**: When the user behavior changes drastically, the traditional models can lag behind until they are retrained on the new data. This lag can mean lost opportunities to interact with the users effectively in real-time.

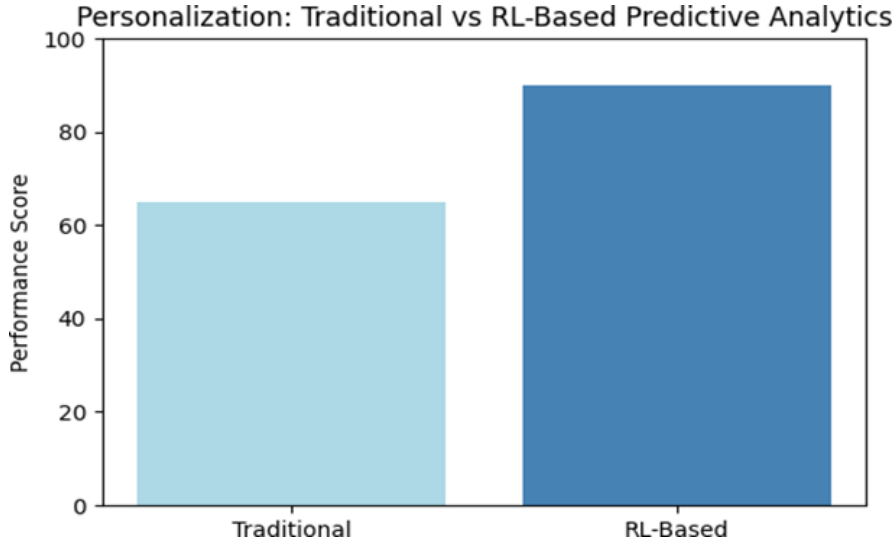
RL-Based Predictive Analytics: RL models are designed to learn to change behavior in real time. They learn to adjust their behavior in response to real-time feedback, i.e., they continue to be effective without retraining, so the predictions continue to be effective and current as user behavior changes.

ISSN 2394 - 7780



E) **Personalization**:

- **Traditional Predictive Analytics**: Classic methods do have some personalization, but it is on a more general category of user level. Adjusting these predictions for individual tastes generally has to be done by hand.



- **RL-Based Predictive Analytics**: Personalization is a more precise process with reinforcement learning. RL models learn from the individual activity of each user and modify their recommendationsor actionsaccordingly to suit individual preferences, resulting in individually tailored predictions to each user in the long run.

VI. CASESTUDY

Having the capability to forecast human behavior has had far-reaching influence in industries ranging from financial to education to healthcare to online commerce. Applying the capability of predictive analytics, businesses have been able to leverage past information to gain insights and knowledge. Now, with the inclusion of RL, such forecasting is now capable of learning in real time, enhancing responsiveness and accuracy.

A) STOCK MARKET

Predictive analytics on the stock market historically is based on statistical models and machine learning methods to examine historical price behavior, economic signals, and even social media sentiment. Classic techniques such as ARIMA and GARCH identify trends and volatility over time but tend to fail to incorporate the unpredictability of market movements.

Applying machine learning, predictive analytics has been made stronger with the utilization of Support Vector Machines (SVMs) and neural networks. These algorithms utilize data from sources like historical prices and social media to recognize patterns. One such use is sentiment analysis to predict price movements, which a study by Huang et al. (2022) proved to be present between social media-based public opinion and stock price movement. [4]

ISSN 2394 - 7780

MAKING STOCK MARKET PREDICTIONS SMARTER WITH REINFORCEMENT LEARNING

Although predictive analytics can reveal trends, reinforcement learning takes it a step further by making the model adaptive. Reinforcement learning enables the model to learn and improve continuously, with its strategy being adjusted based on real-time feedback from the market. This makes it perfect for the fast-paced, ever-changing stock environment, where flexibility and adaptability are essential. [4]

In our RL upgrade framework, we use **PolicyGradient** and **Actor-Critic** algorithms to allow the model to learn from past trades and optimize its policy. This is how it actually works:

1. **Data Collection**: First, we gather essential data, including historical stock prices, trading volume, and sentiment analysis from sources like Twitter or financial news.
2. **Setting Up the Environment**: We treat each trading day as a step where the RL model makes a decision—whether to buy, sell, or hold stocks based on current market signals.
3. **Using Policy Gradient**: This method lets the model try different actions and learn which trading strategies yield the best rewards.Overtime, it gradually learns to favor strategies that prove profitable.
4. **Actor-Critic Approach**: The **critic** calculates how well the **action** (to buy or hold) was selected by the **actor** and instructs the model to get better and prevent risky trades.

Benefits of RL in Stock Market Prediction

Adding reinforcement learning to stock market forecasting adds a number of obvious benefits:

1. **Real-Time Adaptability**: In contrast to conventional models which must be retrained atregular intervals, RL models adapts automatically with each market shift, remaining current.
2. **Improved Accuracy:** The model improves over time through learning from each action's result, refining its predictions with each trade.
3. **Lower Risk:** The Actor- Critic design allows the model to balance experimenting with new approaches with holding onto the ones that pay off, minimizing risk trades.
4. **Tailored Strategies**: RL adapts to various market conditions, like bullish or bearish phases, allowing the model to fine-tune its approach as the market shifts.

Overall, reinforcement learning elevates our forecast model from being a static forecaster to a clever, adaptive decision-maker. This allows for reacting to market trends in real- time, providing traders with a tool that's better attuned to the constantly shifting and multifaceted nature of the stock market.

B) **EDUCATION**

Predictive analytics is transforming education by allowing institutions to forecast trends, support students, and adjust curricula. With the analysis of historical records of student performance, attendance, and demographics, schools can forecast outcomes like which students need extra support or what programs need to be adjusted. Delhi Technological University (DTU), for instance, used predictive analytics to analyze student enrollment trends, program popularity, and satisfaction, allowing them to make more informed decisions on resource allocation and curriculum adjustments [1].

Common Techniques in Educational Prediction

To gain these insights, institutions generally use data mining, decision trees, and neural networks. Decision trees and neural networks were used at DTU to predict enrollment probability, student grades, and program satisfaction overall, enabling administrators to make data-driven decisions to meet student needs[1].

How Reinforcement Learning Makes Educational Predictions Better

1. **Data Collection**: Collect information on performance, attendance and engagement.
2. **Learning Environment**: Approach each student's experience as an interaction, in which the model can recommend tailored support, such as tutoring.
3. **Policy Gradient for Interventions**: RL allows the model to test different interventions, like extra support, and learn which lead to the best academic outcomes.
4. **Actor-Critic for Continuous Improvement:** The actor proposes interventions, and the critic assesses their effect, allowing the model to improve strategies that are best suited to each student.

Benefits of RL in Educational Prediction

1. **Real-Time Adaptation**: RL models are trained to learn how to adapt as they are given new data, and

ISSN 2394 - 7780

institutions can thus respond promptly to shifts such as decreasing engagement.

2. **Personalized Student Support**: By continuous learning, RL models provide personalized advice that optimizes individual performance.
3. **Increased Engagement**: Actor- Critic ensures balancing attempts at new approaches with maintaining successful ones, enhancing overall engagement.
4. **Dynamic Curriculum Updates**: RL's flexibility ensures recommendations are based on current needs, maintaining the curriculum relevant and effective.

C) HEALTHCARE

Predictive analytics is now a precious resource in healthcare, allowing providers to predict patient outcomes, personalize treatments, and optimize resource utilization. Through the analysis of historical patient information—e.g., medical history, laboratory tests, and demographics— healthcare providers can predict outcomes like hospital readmission or disease progression. For example, predictive models are routinely used to predict high-risk cases to facilitate early intervention, leading to better patient outcomes and efficient resource utilization [1].

Techniques in Healthcare Prediction

Common methods of healthcare predictive analytics include logistic regression, decision trees, and neural networks. They are applied to detect patterns within patient data in order for providers to manage outcomes beforehand. For instance, hospitals can use such models to determine which patients have a higher probability of acquiring chronic diseases, and thus they can intervene early using customized care plans.

ENHANCING HEALTHCARE PREDICTION WITH REINFORCEMENT LEARNING

1. **Data Collection**: Collect data on patient vitals, lab results, treatment history, and real-time updates.
2. **Creating an Interactive Environment**: Treat each patient interaction as a decision- making step, where the model recommends treatments or interventions.
3. **Policy Gradient for Treatment Adjustments**: RL enables the model to try out various treatments and identify those that provide the optimal health results and learn to adapt in response to changing patient needs.
4. **Actor-Critic for Continuous Feedback**: The **actor** recommends treatments, while the **critic** evaluates the results, helping the model refine its approach based on patient responses.

Benefits of RL in Healthcare Prediction

1. **Real-Time Adaptation**: RL models adapt in real-time to new patient data, enabling healthcare professionals to respond to changes in a patient's status as they arise.
2. **Personalized Treatment Plans**: From experience with patient response, RL models can suggest individualized treatments that maximize individual outcomes.
3. **Enhanced Decision Support**: The Actor- Critic model weighs experimentation with new treatment methods and established practices, enabling practitioners to make informed decisions confidently.
4. **Dynamic Resource Allocation**: As RL models learn from continuous patient data, they can assist healthcare units in optimizing re source use, enabling high-risk patients timely treatment.

D) E-COMMERCE AND RETAIL

Predictive analytics is transforming retail and e- commerce as it allows businesses to predict customer decisions, manage stock, and personalize shopping. Predictive models, powered by customer behavior, purchase history, and shopping behavior, can make product suggestions, optimize inventory, and personalize marketing campaigns. For example, an online retail store can use predictive models to make product suggestions from a customer's previous purchases or forecast demand for products during festivals [1].

Techniques in E-Commerce and Retail Prediction

Recommender algorithms, decision trees, and cluster methods are some of the most popular methods in this category. These methods help companies in understanding consumer purchase behavior, customer segmentation, and demand forecasting. For instance, cluster methods can segment customers based on shopping habits so that retailers can target and recommend products to each segment based on shopping habits..

ISSN 2394 - 7780

ENHANCING RETAIL PREDICTION WITHREINFORCEMENT LEARNING

1. **Data_Collection**:
Gather customer interaction, product views, purchase frequency, and seasonality trend data..

2. **Interactive Shopping Environment**:
Take each customer's shopping experience as a sequence of interactions, where the model generates product suggestions and records responses.

3. **Policy Gradient for Tailored Recommendations**:
With RL, the model can test different recommendations and learn what drives engagement or conversion and personalize suggestions based on customer responses.

4. **Actor-Critic for Balanced Strategy**:
The **actor** recommends products or promotions, while the **critic** evaluates their effectiveness, helping the model refine its strategies based on what actually resonates with customers.

Benefits of RL in E-Commerce and Retail Prediction

1. **Real-Time Personalization**: RL models continuously adapt to each customer's preferences, offering highly personalized recommendations that feel relevant and timely.
2. **Improved Customer Engagement**: By learning which promotions or recommendations work best, RL models can enhance customer experience and increase engagement.
3. **Optimized Inventory Management**: The Actor-Critic setup helps balance experimenting with new product promotions and focusing on popular items, reducing overstock and stockouts.
4. **Dynamic Marketing Strategies**: RL models can adapt promotional efforts to changing customer interests, ensuring that marketing is responsive to real-time customer behavior.

VII. CHALLENGES AND LIMITATIONS

A) ***DataQuality and Preprocessing***: Predictive analytics relies heavily on the quality of input data, and inaccuracies can lead to unreliable predictions.Data must be cleaned, standardized, and accurately formatted, which can be labor-intensive. As pointed out by Andrees cuetal. (2015), inadequate data quality significantly impacts analysis outcomes, leading to inefficiencies and potential errors in predictions (IJKIE_December2014_JAME…).

B) ***Complexity and Scalability:*** Implementing RL-based models for human behavior prediction can become computationally expensive and challenging to scale. Deep learning applications in RL, especially in fields such as image recognition or behavior modeling,demand high computational power and memory, as observed in the evolution of deep RL techniques for high- dimensional data tasks (Arulkumaran etal., 2017)

C) ***Ethical and Privacy Concerns:*** Predictive analytics, especially when analyzing sensitive behavior data, must adhere to strict ethical standards. The ethical use of data is crucial in behavioral analysis, where sensitive personal information is processed. Ethical concerns about data privacy, transparency, and informed consent must be addressed to align with legal and social norms (Schwartz, 2010) (IJKIE_December2014_JAME…).

D) ***Model Interpretability***: RL models, particularly deep learning-based ones, often act as "black boxes," making it challenging to interpret and justify predictions. This lack of transparency can be problematic for decision-makers needing clear explanations for predictive outcomes. As pointed out, high interpretability is essential in analytics projects for effective communication and trust-building among users (IJKIE_December2014_JAME…).

E) ***Long-Term Maintenance and Adaptability***: Predictive models require continuous updates to accommodate new data and maintain accuracy. This involves retraining and fine-tuning to adapt to evolving patterns, which can be resource- intensive. Moreover, without a robust lifecycle management system, models can quickly become outdated, as noted by Taylor (2012), emphasizing the need for scalable and adaptable model architectures (IJKIE_December2014_JAME…).

ACKNOWLEDGEMENT
I am deeply grateful to my mentors and advisors for their constant support and insightful guidance throughout this research. A special thanks to Jai Hind College for providing the necessary resources that made this work possible. I would also like to acknowledge my peers and colleagues for their valuable discussions and feedback, which helped refine and strengthen this study. Lastly, I sincerely appreciate the encouragement from my family and friends, whose unwavering support kept me motivated every step of the way.

ISSN 2394 - 7780

VIII. CONCLUSION

This article described how predictive analytics with reinforcement learning (RL) could be used to improve the predictability of human behavior. Traditional models can forecast past trends, but they are not responsive in real time. Integrating RL, particularly through Policy Gradient and Actor-Critic techniques, enhances the reactivity of predictive models since they learn incrementally from real-world interactions and react to changing patterns. Finance, healthcare, and e-commerce are some of the dynamic areas where behavior trends are constantly evolving, making RL- based models particularly well-suited. But challenges remain such as data quality, ethics, and computational cost. Integrating RL with predictive analytics is a significant step towards the development of smart, adaptive systems that can learn and evolve with complex human behaviors.

REFERENCES

1. Jindal, R., & Borah, M. D. (2015). Predictive analytics in a higher education context. *IT Professional, 17*(4), 24–33.
2. Silver, D., Lever, G., Heess, N., Degris, T., Wierstra, D., & Riedmiller, M. (2014). Deterministic policy gradient algorithms. In *Proceedings of the 31st International Conference on Machine Learning (Vol. 32)*. JMLR: W&CP.
3. Gu, S., Yang, L., Du, Y., Chen, G., Walter, F., Wang, J., & Knoll, A. (2024). A review of safe reinforcement learning: Methods, theories and applications. *arXiv preprint arXiv:2205.10330v5*.
4. Nann, S., Krauss, J., & Schoder, D. (2013). Predictive analytics on public data – The case of stock markets. In *Proceedings of the 21st European Conference on Information Systems (ECIS)*, Utrecht, Netherlands.
5. Kumar, V., & Garg, M. L. (2018). Predictive analytics: A review of trends and techniques. *International Journal of Computer Applications, 182*(1), 31–37.
6. Kolomvatsos, K., & Anagnostopoulos, C. (2017). Reinforcement learning for predictive analytics in smart cities. *Informatics, 3*(16). https://doi.org/10.3390/informatics4030016
7. Arulkumaran, K., Deisenroth, M. P., Brundage, M., & Bharath, A. A. (2017). Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine, 34*(6), 26–38. https://doi.org/10.1109/MSP.2017.2743240
8. Ogunleye, J. (2014). Predictive analytics: A review of trends and techniques. *International Journal of Knowledge, Innovation, and Entrepreneurship, 2*(2), 60–70.Asniar, K., & Surendro, K. (2019). Predictive analytics for predicting customer behavior. In *2019 IEEE International Conference on Technology and Business Management*, Institut Teknologi Bandung, Indonesia.
9. Ogunleye, J. (2014). The concepts of predictive analytics. *International Journal of Knowledge, Innovation and Entrepreneurship, 2*(2), 82–90.
10. Reinforcement learning in predictive analytics for human behaviour. *Journal of Artificial Intelligence Research*. [Online].

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

ANOMALY DETECTION IN NETWORK TRAFFIC USING A HYBRID MODEL OF K-MEANS CLUSTERING AND AUTOENCODERS

[1]Apurva Kishor Gawde and [2]Sunita Jena
[1]Student and [2]Assistant Professor, MSc. Big Data Analytics Department, Jai Hind College

ABSTRACT
Anomaly detection is generally understood to refer to rare objects, events or observations that deviate significantly from the bulk of the data and do not conform to well-defined notions of normal practice. The model uses the k-means power of clustering to identify potential anomalies based on the distance between cluster centroids and optimizes the search by using an autoencoder to detect reconstruction errors and then, hybrid method combining kmeans and autoencoder, compared with individual methods . The model was tested on real-time network traffic data collected through Wireshark, which demonstrated its effectiveness in detecting anomalous networks. KMeans and Autoencoders have been shown to perform well in detecting anomalies in network traffic, where both methods show the same number of anomalies and in addition, the hybrid method detects fewer anomalies and outperforms both KMeans and Autoencoders.

Keywords: *Anomaly detection, K-Means, cyber security, autoencoders.*

1. INTRODUCTION
With the rapid expansion of digital communications and the reliance on network management systems, the detection of anomalous behavior in network traffic has become increasingly important to maintain security and system integrity New vulnerabilities emerge daily and quickly used in daily attacks. Abnormalities in network traffic are often indicative of security threats such as malware, denial of service (DoS) attacks, or data breaches. Anomaly detection is generally understood to be identification of data that deviates from the normal behaviour. Systems capable of anomaly detection are important tools in many fields. Anomaly detection is a critical task in identifying outliers or deviations from expected behavior in data. [8] They are used to detect financial fraud, network intrusions, unusual traffic situations and other rare events.

These rare events may have an impact on a specific system, but they are hard to find. [8]. Traditional anomaly detection methods may struggle to cope with today's network data volume, speed, and complexity. Cybersecurity attacks are hard to detect. These attacks can result from malicious or benign behavior, internally or externally through malware, targeted attacks, or APTs. However, internal threats are much more powerful and potentially more damaging than external threats because they have already penetrated the network. These activities pose unknown threats and can steal, damage, or alter assets or operations.

Therefore, it is a major concern for industry to enable anomaly detection for cyber network security. Abnormal behavior detection provides real-time cyber-attack threat detection. It monitors unethical user behavior protecting companies from threats.

In this study, we propose a hybrid anomaly detection model that exploits autoencoders and KMeans clustering capabilities. The hybrid model aims to diagnose anomalies in automated traffic, combining the ability of automatons to identify data patterns with the clustering capabilities of KMeans to organize similar traffic patterns. The proposed model improves the detection rate of network anomalies by utilizing both reconstruction error from autoencoders and clustering distance from KMeans.

2. RELATED WORK
Anomaly detection in network traffic has been extensively studied, various methods have been developed to deal with complex and diverse network data Traditional methods, such as statistical illegal methods, are limited by sophisticated or new discoveries.[1] Machine learning models such as support vector machines and random forest have improved detection accuracy, although these methods may lack robustness in detecting complex, nonlinear anomalies [2] especially, K-means are collected to detect anomalies by clustering common behavior and identifying outliers s are used However, they are often ineffective in capturing complex network behaviors in their own right [3]. The k-Means clustering method is first used to partition the training instances into k clusters using Euclidean distance similarity. [7]

Meanwhile, deep learning models, especially autoencoders, have shown promise at the learning level of normal network traffic, and enable deviations to be flagged efficiently [4]. Hybrid approaches combining clustering and autoencoders have emerged to address the weaknesses of individual models. For example, a study has shown that pre-clustering data with K-Means before training the autoencoder can increase anomaly detection accuracy, and reduce false positives [5]. However, challenges remain, including the scalability of this model and

ISSN 2394 - 7780

optimization of threshold values to balance detection accuracy and false positive rates [6]. These works lay the foundation for exploring hybrid models that harness the power of clustering and deep learning for improved anomaly detection in network traffic. Clustering techniques are widely used in anomaly detection to identify patterns in data and separate normal from abnormal behavior. [9] Machine learning techniques enable the development of anomaly detection algorithms that are non-parametric, adaptive to changes in the characteristics of normal behaviour in the relevant network, and portable across applications. [10]

3. ANOMALY DETECTION WITH KMEANS AND AUTOENCODERS

In this section, we briefly discuss the k-Means clustering and the autoencoders methods that are used for anomaly detection.

3.1Anomaly Detection with k-Means Clustering :

This structure mirrors our concept, detailing the clustering and anomaly detection processes, including the distance calculations and threshold-based anomaly classification.

1) Set K : Choose the number of clusters, K, based on the analysis requirements.

2) Initialize the Centroids: Select K initial centroids C = {c1, c2, ..., cK}, either randomly or with a method like k-means++.

3) Assign Points to Clusters: For each point xi, calculate the distance to each centroid ck using:

$$d(x_i, c_k) = \sqrt{\sum_{j=1}^{m} (x_{ij} - c_{kj})^2}$$

$$c_k = \frac{1}{|S_k|} \sum_{x_i \in S_k} x_i$$

4) Update Centroids: For each cluster k, recalculate the centroid as the mean of points in the cluster:

3.2Anomaly Detection with Autoencoders :

This structure depicts an unsupervised anomaly detection pipeline using an autoencoder that learns to reconstruct normal data and flags significant reconstruction deviations as anomalies.

1) Build an autoencoder model with input and output layers equal to the feature size and smaller hidden layers in between.

2) Train the autoencoders on normal data to minimize reconstruction error.

3) For each test point Z:

- Pass Z through the autoencoder to get the reconstructed output.

- Calculate the reconstruction error (e.g., mean squared error between Z and its reconstruction.)

4) Set an anomaly threshold based on reconstruction errors. (e.g., 95th percentile).

5) Id the reconstruction error for Z exceeds the threshold, classify Z as an anomaly, otherwise, classify as normal.

ISSN 2394 - 7780

4. PROPOSED SCHEME

In this section, we proposed a hybrid model of Kmeans clustering and Autoencoders. In this hybrid anomaly detection model, we combine the power of k-Means clustering with autoencoders to improve anomaly detection accuracy. Here's how each step works.

1. Separate training of k-Means and Autoencoder:

- We first train a k-Means clustering model for cluster normal data into clusters, each with a centroid. This chart helps identify points that are far from the norm, and can indicate abnormalities.
- We also train an autoencoder, a neural network designed to encode and then reconstruct input data. Because it is trained on normal data, the autoencoder learns to reconstruct these data points correctly, resulting in less reconstruction error for the normal data.

2. To examine each case with two examples:

- For the new data point Z , we use the two models separately:
- k-Means: Calculate the distance between Z and the nearest cluster center. If Z is far from any focal point, a discrepancy may occur.
- Autoencoder: We feed Z through an autoencoder to get the reconstructed version of Z

. The difference between Z and its return (reconstruction error) helps us to measure how constant the autocoder perceives Z to be. High errors indicate inconsistency.

3. Setting thresholds for similarity:

- Define a threshold for k-Means and autoencoder based on the highest percentage of normal data distances and reconstruction errors (e.g. 95%) This means that we expect to expect the top 5% (or only the other percentage) of values will exceed this threshold under appropriate circumstances.

4. Hybrid Anomaly Detection:

- If both its distance to the nearest cluster center point (k-Means) and its reconstruction error (autoencoder) exceed their thresholds, the data point is flagged as an anomaly
- This combination of criteria is more robust than using either model alone because it reduces false positives—only points that are abnormal in clustering and construction are considered anomalies.

5. METHODOLOGY

The dataset for this study was collected using Wireshark, which captured raw network traffic on a WiFi network. The data contained attributes such as source, destination, protocol, and packet length. After preprocessing, the data were incorporated into a hybrid anomaly detection model by adding the following features.

1. **Autoencoder:** Autoencoder was used to reconstruct normal network traffic patterns. It consists of an encoder and a decoder. The encoder compresses the input data into a low- dimensional representation and the decoder reconstructs the data. The error of reconstruction, that is, the difference between the reconstructed input and output, was calculated for each data point. Network packets that exhibited significant reconstruction errors were flagged as potential anomalies.
2. **KMeans Clustering:** The data were clustered using KMeans, after being processed by the encoder of the autoencoder. KMeans assigns each data point to one of two groups (normal or anomalous). The clustering was based on the reduced dimensional data generated by the autoencoder, so that the clustering process could focus on compressed features Each point distance between cluster centers was used as an additional search feature anomaly detection.
3. **Hybrid anomaly detection:** The anomalies were detected based on two factors:

- Return error from autoencoder.
- KMeans Distance to the cluster centroid.

Data points that exceeded a predefined threshold were flagged as non-abnormal for the item. Final discrepancies were identified by combining the results of both methods to reduce false positives. The hybrid model was trained and tested using the collected network data. The number of epochs was set to 50, and the model's performance was monitored using the training and validation loss.

5. RESULT AND ANALYSIS

ISSN 2394 - 7780

In anomaly detection analysis, we compared the effectiveness of using KMeans only, Autoencoder only, hybrid model (combination of KMeans and Autoencoder) to detect anomalies in network traffic data. The results show that hybrid model shows improved robustness in detecting anomalies, such as overlap, distance measures, and identifies errors in the reconstruction.

```
Average distance of potential anomalies (KMeans): 2.764255496262255
Average distance across all data (KMeans): 1.4172036034040245
Average reconstruction error of anomalies (Autoencoder): 1.0785305217452026
Average reconstruction error across all data (Autoencoder): 0.5506624113388701
```

Fig 5.1: The Result

Reported Variance Overlap:
The Venn diagram shows that 50% of the anomalies detected by KMeans are anomalies detected by Autoencoder, and vice versa. This combination shows that although the two models agree on some similarities, each model also finds a different variance (50% is different for each model). This diversity suggests that the hybrid approach takes advantage of the unique identification capabilities of the two models, and allows for better identification of potentially different features in the data set.

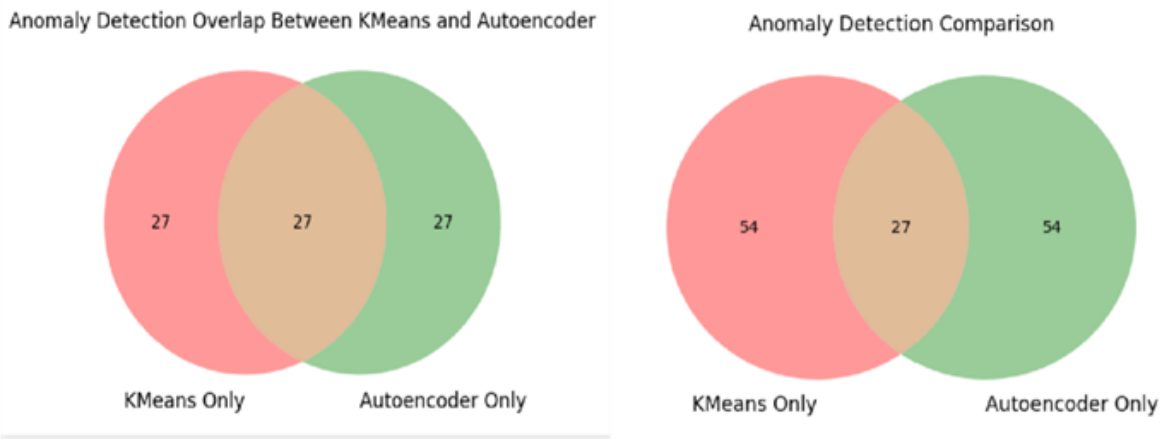


Fig 5.2: Anomaly Detection Overlap Between KMeans and Autoencoders with Comparison

Average distance in KMeans:
The mean distance of the KMeans-only anomalies to cluster centers is 2.764, which is notably larger than the mean distance over all data points (1.417) This high distance confirms that the anomalies detected by KMeans are indeed outliers in the feature space, but beyond that they have no further evidence of their abnormality.

Average reconstruction error of the autoencoder:
The average reconstruction error for the Autoencoder-only anomalies is 1.079, which is significantly higher than the reconstruction error for the entire data set (0.551) This high error indicates that these points deviate significantly away from Autoencoder's known patterns, and improves their classification as anomalies.

Advantages of the hybrid model:
When KMeans and Autoencoder detections are combined, the hybrid model collapses for anomalies to those detected by both methods, resulting in more refined anomalies These combined anomalies lie at higher distances from Chemeans cluster centers and high reconstruction errors, making it likely that they represent genuine anomalies.

This approach avoids false positives that can result from exposure to distance or repetition errors alone and highlights the robustness of the hybrid model to anomalies missed by KMeans or Autoencoder alone in the 19th century.

6. CONCLUSION

The results support the use of a hybrid model, as it combines the capabilities of KMeans and Autoencoder for more accurate and reliable anomaly detection The hybrid approach removes inconsistencies, resulting in anomalies with spatial divergence and structural distortion occurs. This dual validation reduces the risk of false positives and improves the overall model in detecting real anomalies in network traffic data without the need for ground truth labels.

7. FUTURE SCOPE

ISSN 2394 - 7780

This research could be extended in the future in several ways to improve the efficiency of anomaly detection in network traffic data and to use First, more advanced clustering techniques, such as DBSCAN or Gaussian Mixture Models, to handle complex information, nonlinear data structures and improve detection accuracy (GMM) could be investigated Furthermore, generative models such as variational autoencoders (VAEs) or generative adversarial networks (GANs), can capture a more nuanced picture of normal data, enabling the detection of subtle anomalies.

The inclusion of a semi-supervised learning may allow the model to benefit from smaller data sets, if available, for the limited detection limits between normal and abnormal cases improve Future work LSTMs or Temporal Convolutional Networks (TCNs) to capture sequential patterns of network traffic) and other temporal patterns can also be considered, improving anomaly detection for time-series data.

Another promising strategy is real-time anomaly detection, where the model is optimized to handle databases for immediate detection and response to network threats Besides, automatic hyperparameter tuning for boundaries and configurations of the hybrid model can be optimized and applications on different datasets.

Testing the model in different networks, such as enterprise networks or IoT, will prove to be generalizable. Enhancing pattern interpretation through interpretable AI (XAI) techniques such as Shapley Additive Explanations or LIME could improve the clarity of anomaly detection, strengthening the confidence of security analysts. The development of unsupervised performance simulations will also enable rigorous evaluation of anomaly detection models without relying on a ground truth label.

REFERENCES

[1] Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly Detection: A Survey. ACM Computing Surveys (CSUR), 41(3), 1-58.

[2] Ahmad, I., Basheri, M., Iqbal, M. J., & Rahim, A. (2016). Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection.

[3] Li, Y., Chen, D., Jin, Y., & Lin, W. (2019). Anomaly Detection in Network Traffic Based on Combination of Clustering and Deep Autoencoder.

[4] Lyu, X., Meng, X., & Li, Y. (2020). Anomaly Detection of Network Traffic Based on Deep Learning Models. Journal of Information Security and Applications, 53, 102529.

[5] Ahmed, M., Mahmood, A. N., & Hu, J. (2016). A Survey of Network Anomaly Detection Techniques. Journal of Network and Computer Applications, 60, 19-31.

[6] Basant Agarwal*, Namita Mittal. Hybrid Approach for Detection of Anomaly Network Traffic using Data Mining Techniques.

[7] Amuthan Prabakar Muniyandia , R. Rajeswarib , R. Rajaram. Network Anomaly Detection by Cascading K-Means Clustering and C4.5 Decision Tree algorithm

[8] Jaroslav Kopčan, Ondrej Škvarek, Martin Klimo. Anomaly detection using Autoencoders and Deep Convolution Generative Adversarial Networks.

[9] Gerhard Munz, ¨ Sa Li, Georg Carle, Traffic Anomaly Detection Using K-Means Clustering

[10] Tarem Ahmed, Boris Oreshkin and Mark Coates. Machine Learning Approaches to Network Anomaly Detection

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

PREDICTIVE TREND ANALYSIS AND VISUALIZATION FOR EMERGING MARKET OPPORTUNITIES

[1]Aryan Shetty and [2]Fatima Shaikh
[1]Student and [2]Assistant Professor, Department of M.Sc Big Data Analytics, Jai Hind College, Mumbai

ABSTRACT
This paper tries to explore the design and implementation of predictive trend analysis using machine learning and big data analytics to find and visualize market developments. It proposes a framework leveraging different data sources, including social media, sales records, and economic indicators, to provide relevant insights into an organization, improve strategic planning, and exploit emerging trends. The framework includes a fundamental module, BASOA, comprising sentiment analysis, trend forecasting, data collection, and processing into a unified model within the context of Big Data Analytics Service Oriented Architecture. The findings of the research show how BASOA can be a better business intelligence tool by providing trends and interactive data visualizations.

Keywords: *Big Data, Predictive Analytics, Business Intelligence (BI), Machine Learning, Sentiment Analysis, Interactive Visualization, Emerging Market Trends.*

I. INTRODUCTION
The era of quick technological progress comes with the development of data. In this regard, predictive analytics is becoming more prevalent within organizations to generate increased competition. Today, the amount, velocity, and variety of available data have changed the way of market analysis, enabling the adoption of new techniques to predict the market shift. Big Data Analytics (BDA) is such a tool; it provides organizations opportunities to capture and handle large datasets to predict trends and opportunities. BDA complements Business intelligence (BI) by enhancing decision making through historical, real-time, and forecasted data insights.

The study focuses on developing a predictive trend analysis framework for these emerging markets, as new trends and quick changes call for an extra trigger to agile responses. This research provides a holistic approach to understanding market trends through machine learning algorithms and interactive visualizations. The proposed framework is built with BASOA (Big Data Analytics Service-Oriented Architecture), which organizes BDA into modular, scalable services that cut across different industries. As data continues to spur innovation in business, strategies such as the BASOA certifies and marks the increased relevance of predictive analytics framework. Thus, this need entails structured methodologies for trend forecasting and analysis to be adopted by business increasingly.

II. LITERATURE REVIEW

2.1 Evolution of Big Data Analytics and Business Intelligence [4], [1]
Nowadays, the term "Big Data" means large datasets that surpass the processing capabilities of the traditional information systems; they require new method and technology for storing, processing, and analyzing the information. The dimensions, velocity, and variety of Big Data typically define the complexity, which demands advanced analytics to convert the raw data into actionable insight. BDA, over the years, developed from simple statistical analysis to complex machine learning models for predictive and prescriptive analytics.

Descriptive Analytics
Descriptive analytics are the first step towards business analytics. It is a system of historical data summary showing what has happened in the past. It forms the basis for studying trends, recognizing patterns, and preparing business reports. For instance, a retail company can apply descriptive analytics to analyze historical sales data so as to uncover the seasonal buying patterns to inform its inventory planning decisions.

Predictive Analytics
Predictive analytics goes beyond the study of historical data as it takes the past data and uses it to predict future events. Many methods are used in predictive analytics, but the most common are machine-learning models like regression, decision trees, and neural networks. These models can help predict customer behavior, change in market trends, and other crucial factors. For finance, we can use predictive analytics to generate credit risk models, assessing the risk of the bank as informed by what has happened with historical loan performance data to inform credit decisions.

ISSN 2394 - 7780

Prescriptive Analytics

Prescriptive analytics is the zenith of analytics; in itself, it goes beyond predicting outcomes to recommending actions based on the predictive insight. Essentially, it is using machine learning, and optimization algorithms in the best course of action to be adopted by a decision-maker to achieve business goals. For example, a logistics company may not only use prescriptive analytics to determine the best way to optimize delivery routes but may also look at considering the following factors: cost, time, and customer satisfaction.

Strategic decisions could be influenced heavily through prescriptive analytics-simulating several alternative scenarios to achieve results. Organizations may use this method to visualize what any of a number of different strategies might entail before actually adopting them. A typical example is in manufacturing, where prescriptive analytics are being used for production scheduling based on the available resource, demand forecast and constraints in operations.

2.2 Existing Business Intelligence Frameworks and Tools [5], [2]

BI frameworks empower the incorporation of a variety of data sources, permit the transformation of raw, meaningless data into valuable information, and support the actions of decision-making in organizations. Most modern BI tools are characterized by having accompanying data warehouses, a real-time analytics engine and a welcoming cloud infrastructure. They extend further beyond traditional realms of data management, however, by ensuring that scalability and flexibility are built-in for complicated data workflows.

- Tableau: The multi-purpose interactive visualization software with a friendly interface; Tableau is indeed one of the choices in softwares that gives number of industries preference in making dynamic dashboards and permit them to move through data in detail. It is equally considered to be one of the best tools supporting different types of data sources, which allow drag-and-drop model data manipulation while granting real-time collaboration. So, organizations can develop sophisticated visualizations that summarize complex datasets into much simpler charts to facilitate faster insights for decision-making.

- Microsoft Power BI: This is the complete BI offering tool tightly integrated in the Microsoft ecosystem and comprises visualization, analytics, and reporting aspects. With this, users are enabled to create interactive reports and dashboards that rely on devices to be accessed. It makes the whole-building process of solid analytics solutions easier when you click it all together in Azure and other Microsoft services.

Advanced modern Business Intelligence frameworks developed to date have allowed organizations to move from above into exploratory and predictive data analysis. When BDA is added into BI frameworks, it offers a predictive model which employees can use to identify patterns, determine anomalies, and forecast market trends.

2.3 The Need for Predictive Trend Analysis in Emerging Markets [3], [8]

BI and BDA are likely to yield dividends when it comes to decision making; however, emerging markets all together bring their own set of unique challenges and approaches. In fast-changing markets, therefore, agile companies are able to keep in pace with changes in the modes of consumption, economic conditions, and other technologies affected. As described by existing literature, BRIE systems are quite at bay in meeting the speedy and complicated realities of most countries and therefore the demand for frameworks such as BASOA, prediction trend analysis for highly fast processing of heterogeneous data sources, and clear actionable insights.

As emerging markets embrace Sentiment Analysis, an area associated with predictive analytics process that requires social media and public data, tremendous flow is expected in creating avenues for decoding customer preference, market spikes, and their reaction to market nuances. Such machine learning algorithms of trend prediction and sentiment extraction and real-time delivery of information to the business can indeed put a company at a better premise for competition.

One case would be a consumer electronic company that can use sentiment analysis to tap before releasing any product into an emerging market to create a buzz as well as identify entry barriers. This would enable the company to recalibrate how it sells and position features in its products on the basis of discussions and reviews on social media.

The behavior of consumers has been changing with the emerging face of e-commerce in developing countries; hence organizations need to observe online trends and sentiments. Nowadays, consumers have become quite loud, as far as preferences are concerned with respect to the social media and other channels of communication. It is sent to boil in the best interest of the businesses to use this analysis to bring about productivity.

ISSN 2394 - 7780

III. METHODOLOGY

3.1 Data Collection and Integration [2], [3]

Sources of data in this case would include:

- Social Media Data: In fact, many existing platforms such as Twitter, LinkedIn and Google Trends work excellently in capturing real-time consumer sentiments. As per actual visualisation, the rank and web scraping methodologies allow continuous data extraction, which is analysed later. One such example is obvious in the API, like Twitter's API; it gives all companies the ability to naturally in real-time track mentions of their brands or the competition. The examination of such kinds of data reveals a better understanding of public perception through which it would be beneficial for companies to shape their attitudes.

- Sales Records and Economic Activity Indices: Event sales contain historical sales data with economic metrics like inflation rates, levels of employment or consumer spending. A company can track the changes in its sales data with respect to macro-economic trends to find correlations to form useful planning strategies. For instance, such an organization might find out that during the downturn of the economy consumer spending reduces and then cuts marketing costs and spending.

- Other Data Sources: Other data sources include industry-specific items; regulatory changes in particular industries, demographic information, or shifts in technology. Knowing these regulations offers definite predictions about the market trends within an ethical perspective.

3.2 Data Processing and Transformation [6], [7]

The collected data passes through several processes and conversions to ensure high quality and good usability in the analysis that follows. Some of these processes are:

- Data Cleaning: Identify and delete the present errors, duplicates and irrelevant data among the dataset. It contains several methods to facilitate enhancement of quality data such as adding outlier detection methods or involving standard data normalization or others. If for instance, there is a sales dataset with impactful errors (like negative sales) it will be better to correct those errors before analyzing them.

- Data Transformation: This is the typical transformation of actual raw data, transforming it into a certain format needed for analysis before aggregation with encoded categorical variables. The other changing process is primarily numerical scaling. E.g. Monthly sales average data might be changed to let the user estimate the seasonal trends.

- Data Consolidation or Integration: Bringing together data from multiple sources into a unified dataset and providing a complete view of the market. One such way is to build a data warehouse for such a purpose; or using complete data-integrating ETL processes to ensure those data are valid and sound. Such use cases include Apache NiFi-an integration platform for automating data ingestion and transformation.

3.3 Sentiment Analysis [3], [7]

This means the subject under trend prediction analysis is sentiment analysis. By natural language processing methods on data gained from social media or customer reviews for business marketing, it could predict a typical score regarding how generally people feel about a given product or its brand. To carry this out, one would require:

Text Pre-Processing: This would pretty much have tokenization, stop word removal, and stemming or lemmatization processes for cleaning up the text data. A tweet like "I really love this new smartphone!" would have been processed just like this-root sentiment dug out. The tweet wouldn't thus be processed.

Sentiment Classification: This deals with the application of any such machine learning algorithms such as Support Vector Machines (SVM) or recurrent neural networks in classifying text data into positive, negative, or neutral sentiments. The model would be trained with labeled datasets (like customer reviews) to make sure that the efficacy of the learning machine is amplified for classification. It could even use a pre-trained one like BERT or VADER for much faster application.

Aggregated Sentiment Values: They would be amalgamated over a time period to represent trends and shifts in public opinion. For instance, one could be evidence with regard to moving sentiments towards sustainability, rather than the type of product perceived towards increasing positive sentiment in eco-friendliness.

3.4 Predictive Modeling [1], [4]

Asserted to be another act of modelling, this is predicting future aspects, events, or behavior on the basis of history and sentiment observed. Machine learning and predictive trend understanding includes the following:

ISSN 2394 - 7780

Regression Analysis: Regression models as linear and logistic were qualified into relating the independent variable and dependent variable. For instance a retailer can use regression to predict its future sales vis-a-vis the amount spent on advertising and some economic indicators.

Decision Tree: A decision tree brings about the ideas of the decision-making process and its outcome as an advantage of decision trees are that they show the principal attributes that cause the change and interpret even by laypersons that do not trained in statistics or analytics. This kind of domain is also used when the decisions are based upon rather complex datasets.

Neural Networks: deep learning networks, CNNs (convolutional neural networks), are some of the most used methods for inferring, because they capture rather complicated patterns from large data sets. They will perform really well for sentiment analysis on social media or consumers' remarks related to products.

Ensemble Methods: There are multiple techniques among which The Random Forest and Gradient Boosting are present, in which they use the approach of combining models to gain accuracy and robustness. They mostly perform well on high-dimensional data usually seen when emerging markets.

The predictive model is trained by providing data to it in training and testing set format. The training examples are used to learn a way to adjust the parameters of the model to minimize any potential prediction errors. After the training, the model is tested to some set test data to assess the performance.

3.5 Validation Metrics [1], [5]

Predictions instead of model validation would give such a perspective; most projections are an accurate sample size large enough to carry very real predictions for the developed model. These metrics are included:

- Model Evaluation: Mean Absolute Error (MAE) - The determination of mean average differences between forecasted and actual values rendering accuracy measures. It implies that a smaller MAE entails greater efficiency of the model.

- This is Root Mean Square Error (RMSE) - The bigger penalties are given to serious errors in RMSE, which causes problems down-there on business-decisions.

- R-squared: This is called the statistic showing the ratio of variation which explains how independent variables get the dependent ones. There is a direct proportionality where a greater value of R-square enhances fitting of the model.

- Cross-Validation: Techniques including k-fold cross-validation where the dataset is divided into k subsets and the model trained k times with a different subset used each time for validation, will give a more realistic measure and help reduce overfitting. This model, once validated and considered to be accurate, can thus provide the organization with its direct actionable response pathways for business chaining.

3.6 Visualization of Predictive Insights [4], [7]

Data visualization is a key focused element in effectively communicating tenders for predictive insights in stakeholders. Interactive dashboard and visualization will enable the audience to better comprehend the insights and hence rely on them for their decision. Effective visualization techniques include;

- Dashboards: construction of interactive dashboards that can collate key performance indicators (KPIs), sentiment scores, and results of the predictive analytics. Users can do exploration through these filter, drill down, or real-time update capable applications like Tableau and Power BI.

- Time Series Visualizations- In these, data allows trends to show over time for recognizing patterns and seasonality, usually in line with line graphs and area charts demonstrating alterations in sentiment by sales figures.

- Geospatial Mapping: GIS are geographic information systems for opportunities to be visualized within different regions across the marketplace. Heat maps could pinpoint areas of high consumer interest or sentiment, allowing a targeting marketing approach.

- Network Graphs: allow visualization of relationships across various products and brands and consumer sentiment. Network graphs can give an insight regarding the brand loyalty of a consumer and what preference one would have as to products.

Effective visualization creates an organization culture inclined towards data decision-making, where the stakeholder has the facility of exploring the data dynamically and making decisions.

ISSN 2394 - 7780

IV. IMPLEMENTATION OF THE BASOA FRAMEWORK:

The BASOA (Big Data Analytics Service-Oriented Architecture) framework is used as a base to implement predictive trend analysis. Its modular design provision is so flexible and scalable that it can accommodate various data sources, processing capabilities, and analytical methods. The architecture comprises:

4.1 Architecture Overview [2], [3]

The first thing is that the one Data Integration Module brings together information from all of the varied sources. The integrity of the data is maintained through ETL processes. Scheduled data pulls and automatically activated data pulls, for instance, from social media, sales database, and third-party reports into analytics, ensure that fresh data is available at all times for analytics.

For example, the Data Processing Module: It has fulfilled functions like data-cleansing, transformation, and aggregation while being able to house and accommodate massive globally-distributed data sets, such as those deployed through computing infrastructures like Apache Hadoop and Apache Spark. The increases speed through in-memory processing on Spark bring the organization closer to near-real-time analytics.

Analytics Module: This module has a sentiment analysis and predictive modeling component. The actual implementation enables building and training machine learning models within libraries like Scikit-learn and TensorFlow. This module enables rapid experimentation of alternative algorithms and configurations of models with a view to optimizing the predictive accuracy.

Their Visualization Module: Provides IT with tools such as Tableau or with Power BI enabling interactive dashboards and visualization. Each stakeholder can therefore dynamically filter through insights of data, such that the environment can be fully conducive with a data-driven culture in the organization. Data visualization flattens things out and makes it quite simple for stakeholders to see the most important insights beginning with some critical trends.

Feedback Loop: It would bear direct feedback from the users in combination with real-time data for continuous enhancement of the model in terms of the accuracy and relevance to changing market conditions. Then, iteratively, ensure that this analytical system could become updated to keep pace with the dynamic environment of the market. For instance, if the perspective of consumers changed, the sentiment analysis would be able to note accordingly that a feedback loop.

4.2 Case Study: Implementation in Retail [3], [8]

-Collation of Data: The organization collects data through various channels ranging from social media to reports emanating from the markets and sales towards a complete library that opens the minds and hearts of consumers and market situations. In unison with data from platforms like Instagram, Facebook, and Twitter, the company is able to draw real-time insight into sentiment and interest.

-Sentiment Analysis: The sentiment derived from the conversations originating from various competing brands on different social media streams would then be analyzed quantitatively by different techniques of NLP. Mostly, large positive sentiments would come out promising its advantageous future opportunity for green products. Set aside for this mode of analysis is to bring forth the marginal importance of using sustainable materials in the discourse around products among consumers.

-**Prediction of Trend:** The company applies machine learning models to evaluate possible consumer buying behavior concerning historical sales in line with the current market trend. Consequently, it has predicted that there will be increasing demand for green products in the next six months. Among the parameters considered in designing this demand forecasting model that informs inventory and marketing strategy are sentiment scoring, sales history, seasonality, and predictions based on model return inference.

-Visualization: Trend scores and expected sentiments will be available in highly interactive dashboards to assist stakeholders in making a sound conclusion regarding product launch and marketing strategies. The visuals demonstrate consumer interest and areas that show possible future growth from which the company can form its strategic direction. Such stakeholders can then analyze how changes in sentiments over time are aligned with data from the sales of products integrated into a launch.

V. REAL-WORLD APPLICATIONS OF PREDICTIVE TREND ANALYSIS:

5.1 E-commerce Sector [7], [1]

Development of Interfaces with Predictive Trend Analysis has highly revolutionized the field of e-commerce above bringing customers closer to businesses. Presently, factors such as using sentiment analysis and

ISSN 2394 - 7780

predictive modeling have improved customer experience and product effectiveness in tailoring sophisticated marketing campaigns.

- Customized Recommendations: E-commerce giants have been known to use predictive analysis in forecasting for recommending on the basis of user behavior. Netflix and Amazon, as two well-known examples, would infer a user's previous purchase, browsing history, and reading other people's customer reviews, then refine that suggestion specifically for the individual. It even improves personalization and then proceeds higher conversion.

- Dynamic pricing: They have come up with such predictive model of adjusting price for e-commerce. For example, businesses develop or adjust prices in a competitive price for online shopping portals coupled with a real-time signal from these shopping portals compared with the buying behavior of the consumers.

- Management of Inventory: Forecasting becomes possible due to predictions which then allow e-commerce organizations to optimize inventory. Logistics for popular items can be balanced with overstocked items using the sales patterns and the trends in the market.

5.2 Financial Services [4], [6]

Predictive trend analysis helps the financial services business identify risks, improve customer experience, and even enhance investment strategies.

- Investment Strategies: Predictive analytics is used by investment firms to assess market trend predictions and align trading strategies. Analyzing historical data about stock performance and economic indicators allows firms to optimize portfolio management and invest.

VI. CHALLENGES AND LIMITATIONS:

While predictive trend analysis offers numerous benefits, several challenges and limitations must be addressed:

6.1 Data Privacy Concerns [5], [6]

With all the data which organizations collect and analyze about consumers, data privacy and data security have become an issue of great concern in the organizations. GDPR (General Data Protection Regulation) and CCPA (California Consumer Privacy Act) acts are quite important for businesses because they serve as shields for consumer information. Organizations should ensure that the data they are processing, requirement for consent from the consumer should be required before starting collection of private data and transparency in data practices is needed as well.

6.2 Data Quality and Bias [6], [7]

The precision of predictive models weighs heavily on the quality of input data. Any incompleteness, inaccuracy, or bias in data can result in erroneous predictions and ill-founded business decisions. Such organizations should invest in data-cleaning and validation processes for maintaining the integrity of their datasets. Furthermore, it is necessary to deal with the bias present in machine learning algorithms to prevent the emergence of discriminatory outcomes.

6.3 Model Interpretability [4], [5]

Even though the machine learning models can be accurate in prediction, there are many complex algorithms such as deep learning models that do n. Interpretability is an aspect that business stakeholders will not be able to understand h. Thus, they do not understand how specific predictions are made, which affects their trust in the generated insights. To foster confidence in the users, it is necessary to develop explainable AI (XAI) solutions that greatly improve the interpretability of models.

Machine Learning, automated to provide accurate predictions, fails many complex algorithms, for instance, deep learning models, in interpretation. They might not be able to fathom how the predictions were made, and this would hinder their trust of generated insights. The trust of users on such models can be developed through explainable AI (XAI)-solutions, fostering improved interpretations of models.

6.4 Rapidly Changing Market Dynamics [1], [8]

Consumer behavior is rapidly changing in relation to financial and advanced technology trends, which undeniably characterize emerging markets. In such scenarios, predictive models are then needed to be updated more frequently to stay relevant in their predictions. Organizations, therefore, need to establish continuous monitoring and retraining of models to most accurately adapt to the trends of the market.

Emerging markets show unprecedented changes in consumer behavior, economic conditions, and technological advancement. Constant updates of predictive models are a must to remain valid in these environments. Organizations have to adapt to the approach of continuous monitoring and retraining of models.

ISSN 2394 - 7780

VII. CONCLUSION

The predictive trend analysis and Big Data Analytics integration with Machine Learning open promising avenues for businesses to grapple with the terrain of new recessionary economies. The proposed BASOA framework provides an organized mode of data acquisition and processing, data mining for sentiment analysis, predictive modeling, and visualization. Real-time data coupled with advanced analytics can point out market trends to improve decision-making by such organizations, bringing them a step closer to a competitive advantage.

The last factor will only get increasingly entrenched with emerging industries. Predictive analytics will address the challenges of data privacy, quality, and interpretability as well as those dynamics found within markets. With the ability to indicate likely forthcoming market movements or consumer changes, companies will succeed in this increasingly data-driven world.

VIII. FUTURE WORK

Further maturation of the BASOA framework through the application of more advanced techniques such as deep learning, reinforcement learning and real-time data streaming. Another possible research direction could be to further extend the predictive capabilities by using IoT-real time data and cloud computing technologies.

However, organizations must train and retrain their employees in data analytics and machine learning for business applications using predictive trend data analysis. That would keep the companies moving ahead in the data-driven culture which could be crucial to their survival, adjustment and proactive changes to the dynamics of the market.

IX. REFERENCES

[1] Chen, Yili, Congdong Li, and Han Wang. 2022. "Big Data and Predictive Analytics for Business Intelligence: A Bibliographic Study (2000–2021)" Forecasting 4, no. 4: 767-786.

[2] Sun, Z., Zou, H., Strang, K. (2015). Big Data Analytics as a Service for Business Intelligence. In: Janssen, M., et al. Open and Big Data Management and Innovation. I3E 2015. Lecture Notes in Computer Science (), vol 9373. Springer, Cham.

[3] Zhaohao Sun, Lizhe Sun & Kenneth Strang (2018) Big Data Analytics Services for Enhancing Business Intelligence, Journal of Computer Information Systems, 58:2, 162-169, DOI: 10.1080/08874417.2016.1220239

[4] Chen, Hsinchun, Roger H. L. Chiang, and Veda C. Storey. “Business Intelligence and Analytics: From Big Data to Big Impact.” MIS Quarterly 36, no. 4 (2012): 1165–88. https://doi.org/10.2307/41703503.

[5] He, Xin James (2014) "Business Intelligence and Big Data Analytics: An Overview," Communications of the IIMA: Vol. 14: Iss. 3, Article 1.

[6] Bala M. Balachandran, Shivika Prasad, Challenges and Benefits of Deploying Big Data Analytics in the Cloud for Business Intelligence, Procedia Computer Science, Volume 112,2017, Pages 1112-1122, ISSN 1877-0509

[7] Shaokun Fan, Raymond Y.K. Lau, J. Leon Zhao, Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix, Big Data Research, Volume 2, Issue1,2015, Pages 28-32, ISSN 2214-5796

[8] Jiwat Ram, Changyu. Zhang, Andy Koronios, The Implications of Big Data Analytics on Business Intelligence: A Qualitative Study in China, Procedia Computer Science, Volume 87,2016, Pages 221-226, ISSN 1877-0509

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

EARLY DETECTION OF AUTISM SPECTRUM DISORDER IN TODDLERS: A DATA-DRIVEN APPROACH

[1]Mohammed Ayaan Qureshi and [2]Sunita Jena
[1]Department of M.Sc. Big Data Analytics, Jai Hind College (Autonomous), Mumbai
[2]Assistant Professor, Department of IT, Jai Hind College (Autonomous), Mumbai

ABSTRACT
Autism Spectrum Disorder (ASD) is a complex neurodevelopmental disorder characterized by difficulty in communication, social interaction, and behaviour that varies among individuals. It is critical to detect ASD early so it can be treated on time but toddlers with ASD are hard to diagnose as the symptoms can differ greatly and overlap with other developmental disorders. This research proposes using machine learning to diagnose ASD at an early stage . This is done through a multimodal dataset that consists of standard screening tool information. Specifically, they will cover Q-CHAT-10, ADOS, and INCLEN in detail. The dataset which is used to either classify someone or not someone to have autism is from the social interaction, communication, sensory and repetitive behaviour domain.

By incorporating various behavioral markers, the data's multimodal aspect improves the model's capability of detecting subtle patterns that indicate ASD. To find out which classifier gives the best results, different machine learning algorithms were evaluated on this dataset. The different algorithms which were evaluated include Logistic Regression, Support Vector Machine, Naive Bayes, Decision Tree, Random Forest, AdaBoost, Gradient Boosting, Bagging, Deep Neural Networks, etc. To measure the performance of the various models, accuracy, precision, recall, F1-score, etc. were used. The findings showed that multiple models especially Naive Bayes, Logistic Regression, Random Forest and Deep Neural Networks have high values of accuracy and recall and therefore have a reliable capacity to distinguish between the ASD-positive and ASD-negative cases.

This study shows how machine learning can accurately diagnose toddlers with ASD through a multimodal model. Combining various diagnostic tools and further evaluating a range of classification algorithms contribute to creating a robust data-driven approach. Therefore, this study offers an intervention to clinicians and parents for early-stage identification of ASD (Autism Spectrum Disorder). This paper analyzes the detection of Autism Spectrum Disorder using classifiers on multimodal data.

Keywords: *Autism Spectrum Disorder, early detection, machine learning, multimodal dataset, classifier comparison, Q-CHAT-10, ADOS, INCLEN.*

I. INTRODUCTION
Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder in which Social and communication behavior is affected and the cognitive functioning takes place which presents various symptoms which varies with the individual. The DSM-5 (Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition) states that the fundamental features of ASD are: deficit in social interaction and communication. Restricted and repetitive behaviours. Cases of ASD are increasing in number. As such, requiring a precedence given to diagnosis and treatment. Finding autism early can improve development as a therapy can help improve cognitive and social skills of children with autism.

Even though there has been great research in this field, diagnosing ASD in toddlers is hard. Often times, traditional diagnostic methods often rely on clinical observations and parent-reported questionnaires, which might not be accurate and objective due to subjectivity in such measures. Moreover, other developmental disorders have symptoms in common with ASD, resulting in diagnostic confusion. The restricted diagnostic process highlights a crucial requirement for more objective, data-driven solutions for early ASD detection.

Diagnostic challenges arise from the varied presentation of autism spectrum disorder (ASD) and overlap conclusions. Right now diagnostic tools tend to be helpful. However, often their assessment can be isolated and not very multifaceted. Tools like the Q-CHAT-10, ADOS and INCLEN evaluate various behavioral domains. Social communication, sensory responses, and repetitive behaviours refer to them. However, their combined use for diagnosis is rare. When we depend only on one single tool or one dimension of a data point to arrive at a diagnostic conclusion, it may lead to incomplete conclusions thus increasing the risk of missed or delayed diagnosis.

This research aims to fill these gaps by creating a multimodal machine learning model that integrates data from several well-validated diagnostic tools. The combination of Q-CHAT-10, ADOS, and INCLEN will provide a more conclusive comprehensive diagnostic profile, facilitating the detection of other critical behavioural and

ISSN 2394 - 7780

sensory features. This dataset can help the model to determine the relations and patterns of complex symptoms. The model will better be able to segregate ASD-positive from negatives.

The focus of this study is to develop and validate a diagnostic model for ASD using machine learning techniques on a multimodal dataset that spans several behavioural and sensory domains relevant to ASD. This research evaluates a range of machine learning classifiers such as Logistic Regression, Support Vector Machine, Naive Bayes, Decision Tree, Random Forest, AdaBoost, Gradient Boosting, Bagging, Deep Neural Network, to find the best one. The data has answers to a lot of important questions which will help develop a feature set for every individual.

This study hopes to help build a properly researched tool that can be readily available to parents and the doctors so that autism spectrum disorder can be diagnosed at an early age for timely intervention. This study combines behavioral, social, and sensory data using machine learning. It shifts away from current diagnostic practices and takes the field towards more objective and comprehensive detection of ASD. Research shows that using a multiple dataset together can help in making better machine learning supported models to diagnose Autism Spectrum Disorders. By combining several tools and contrasting different classification algorithms, a basis is formed to identify the top-performing model, as well as shedding light on how machine learning can improve early ASD diagnosis. This research can be utilized to develop an accurate and cost-effective diagnostic tool in line with the goal of early detection and intervention of ASD which can impart better developmental outcomes and facilitate support for the beneficiaries and their families.

II. LITERATURE REVIEW

Machine learning and autism spectrum disorder (ASD) diagnosis has made remarkable strides in the past few years, offering new solutions to long-standing challenges in early detection, behavioural analysis and personalized intervention.

A. *Machine Learning Progression for Diagnosis of ASD:*

Bone et al. (2016) study showed how using fusion of information from multiple devices, diagnosis of autism spectrum disorder can be improved using machine learning. By using multiple instruments, the diagnostic accuracy was enhanced and an overall better understanding of the Autism Spectrum Disorder-related symptoms was achieved. In the same way, Wall and colleagues (2012) utilized machine learning algorithms to enhance observation-based screening of ASD that would ultimately shorten diagnostic testing, without losing accuracy. Zhao et al. (2019) built on the concept of multimodal approaches using machine learning methods, integrating behavioral and neurological information for the precise detection of ASD.

B. *Behavioral Distinction and Minimal Feature Sets:*

When a child has both ASD and another developmental disorder like ADHD diagnosis can be difficult. Duda and colleagues (2016) researched the use of machine learning models to differentiate autism spectrum disorder from attention deficit hyperactivity disorder through behavioural patterns.

Kosmicki et al. (2015) focused on a small set of behaviors for diagnosing the ASD. Their work showed that one can get the right diagnosis with a few chosen indicators. Finding out this shows how optimized features are very useful for scalable diagnostic tools.

C. *Applications and Limitations of ML Models:*

Thabtah (2019) tries to opt a study on impact of machine learning in behavioral research in autism spectrum disorder. Research shows several gaps: existing datasets are not diverse (not representative of the populations) and current models cannot be applied to various populations. Thabtah et al. (2020) built a machine learning classification model based on behavior and showed high accuracy. Their work emphasizes the importance of behavioral factors in the diagnosis of ASD.

D. *Prevalence and Predictive Modeling:*

The number of children with developmental disorders, such as ASD, has been steadily rising over the years (Zablotsky et al., 2019). This has raised the urgent need for scalable solutions to diagnose such disorders. Also, Ozonoff et al. (2009) pointed out the importance of early identification of symptoms. They showed that parental concerns reported in infancy predicted future diagnosis of ASD. This shows that early diagnostics could be made by ML models on parent-reported data.

E. *Emerging Applications of ML in ASD:*

Machine learning applications are now predicting behaviors associated with autism spectrum disorder (ASD), besides making diagnoses. In 2021, Tajsic and colleagues developed machine learning models to predict aggressive behaviour in youths with autism spectrum disorder, to aid in behavioural management and

ISSN 2394 - 7780

care.These developments show how flexible ML is in tackling complex ASD-related problems beyond the initial diagnosis.

F. *Conclusion of Current Research:*

The literature as a whole emphasizes the significant impact that machine learning can have on ASD diagnostics and other areas. Even though there has been notable advancement in creating effective and precise models, obstacles like varying datasets, interpretability, and incorporation of multimodal data continue to exist. These observations serve as a strong basis for additional research on developing comprehensive, data-driven diagnostic instruments for ASD.

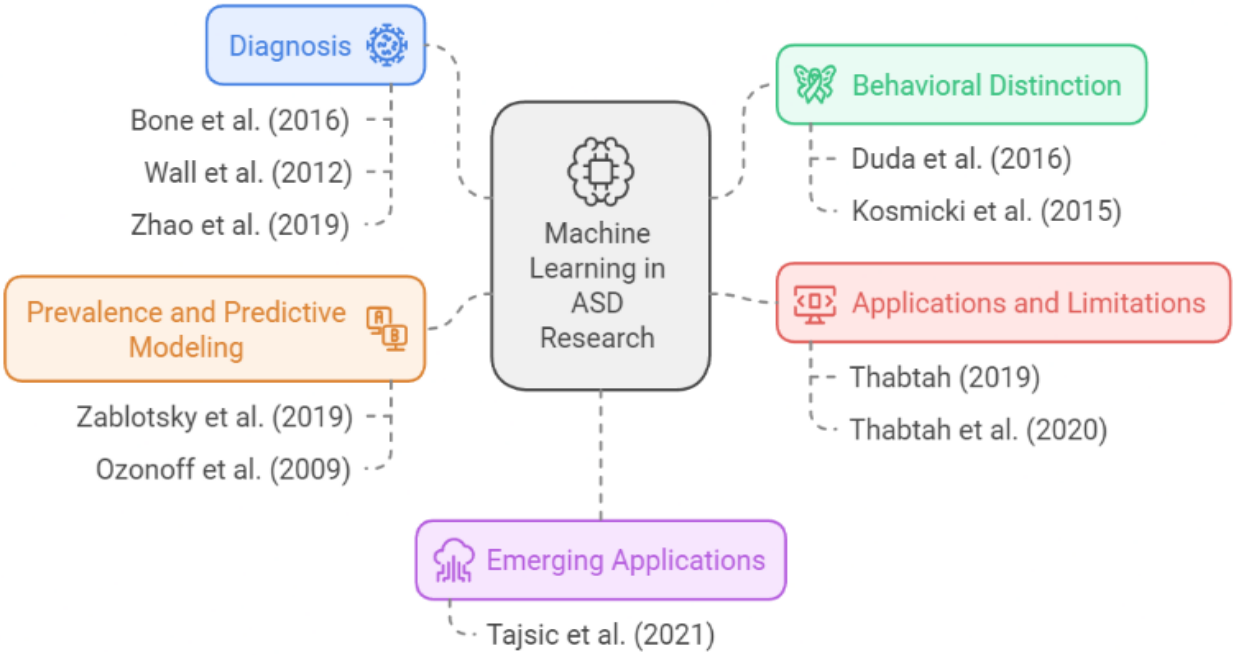


Figure 1: Summary of Literature Review

III. METHODOLOGY

In this study, we use a combination of different methods to make better predictions on ASD diagnosis. The process starts by collecting the data of the system. After which, the data is pre-processed. This approach tells about the information about the system misbehaving.

A. Data Collection

The data for this study combines information from three popular tools for diagnosing ASD, namely, the Q-CHAT-10 (Quantitative Checklist for Autism in Toddlers), ADOS (Autism Diagnostic Observation Schedule), and INCLEN (INDT-ASD) tool. All these tools capture interesting behavioural and social features as well as the sensory features of ASD. There are 750 positive and 750 negative samples in our dataset.

1) ***Q-CHAT-10*:** focuses on parent-reported observations of social interaction and communication in toddlers, providing early indicators of ASD.
2) ***ADOS*:** is a structured observational assessment used to evaluate social communication and play in children, which is often administered by clinicians.
3) ***INCLEN*:** includes a more comprehensive array of ASD markers, encompassing social behaviors, sensory reactivity, and repetitive actions.

This integrated dataset covers a broad spectrum of features, enhancing the model's ability to recognize diverse symptom patterns.

ISSN 2394 - 7780

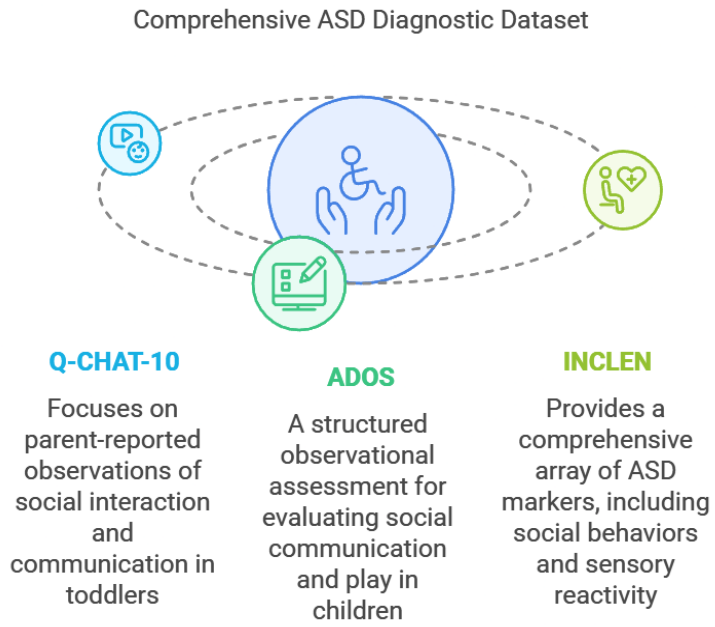


Figure 2: ASD Diagnosing Datasets

B. Data Preprocessing and Feature Engineering

After data collection, preprocessing steps were applied to prepare the dataset for machine learning:

1) ***Data Cleaning*:** Any missing or inconsistent values were handled, with specific attention to maintaining data integrity across the three diagnostic tools.
2) ***Standardization*:** Since each diagnostic tool has its scoring system, feature standardization was applied to scale values into a comparable range, ensuring uniformity across features.
3) ***Feature Engineering*:** New features were engineered by aggregating or combining scores across similar domains (e.g., social interaction and communication scores from ADOS and INCLEN), which allowed for a unified representation of ASD traits in the dataset. This step also involved creating total scores for each tool, which captured overall symptom severity.

C. Model Selection and Implementation

Given the multimodal nature of the dataset, several machine learning algorithms were implemented and compared to identify the most accurate classifier. The selected models include a variety of traditional and advanced classifiers:

1) ***Logistic Regression*:** This linear model was selected for its simplicity and interpretability.
2) ***Support Vector Machine (SVM)*:** SVM was included for its capability to handle high-dimensional data and complex decision boundaries.
3) ***Naive Bayes*:** Known for its effectiveness with structured data, this probabilistic model was tested as a baseline classifier.
4) ***Decision Tree*:** This model was used for its interpretability and ability to capture non-linear patterns.
5) ***Random Forest*:** An ensemble model that combines multiple decision trees, chosen for its robustness and ability to reduce overfitting.
6) ***AdaBoost and Gradient Boosting*:** Both boosting techniques were included to see if iterative refinement could improve accuracy.
7) ***Bagging*:** A method of random sampling with replacement, evaluated for its performance with high variance datasets.
8) ***Deep Neural Network (DNN)*:** A multi-layer neural network was implemented for its capacity to learn complex patterns from large datasets.

ISSN 2394 - 7780

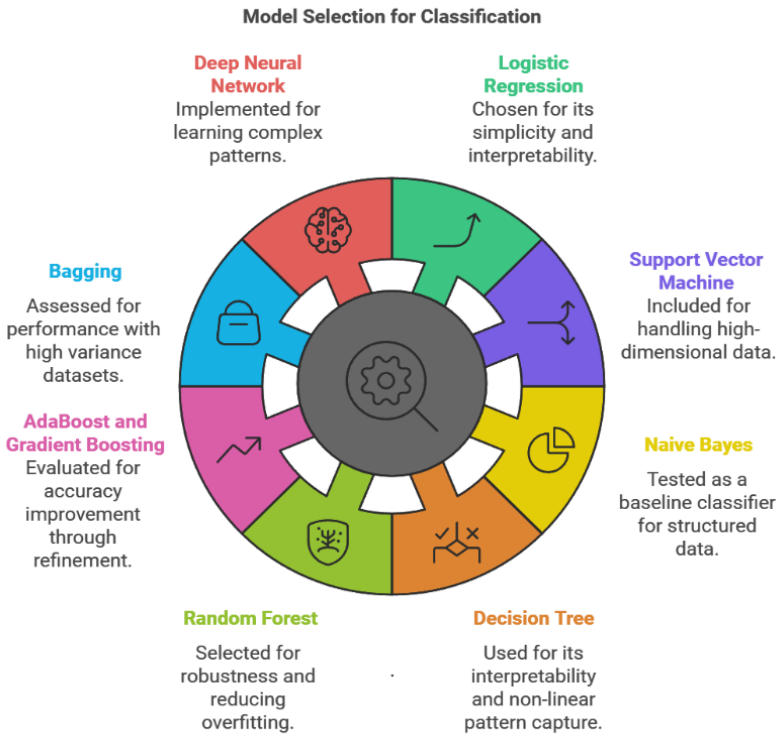


Figure 3: Machine Learning Models used

To guarantee that comparisons between models remained consistent, each model was trained and evaluated on the same split.

D. Model Evaluation and Comparison

The performance of each classifier was assessed through accuracy, precision, recall and F1-score. The selected metrics provide a complete view of the diagnostic ability of each model.

1) ***Accuracy*:** calculated the ratio of correct predictions.
2) ***Precision*:** They calculated the ratio of true positive cases to all the predicted positive cases to reduce the number of false positives.
3) ***Recall*:** It is the ratio of true positive cases to all real positive cases. This is important to reduce false negative cases.
4) ***F1-score*:** Given the difficulty of diagnosing ASD, this choice presents a balanced measure between precision and recall.

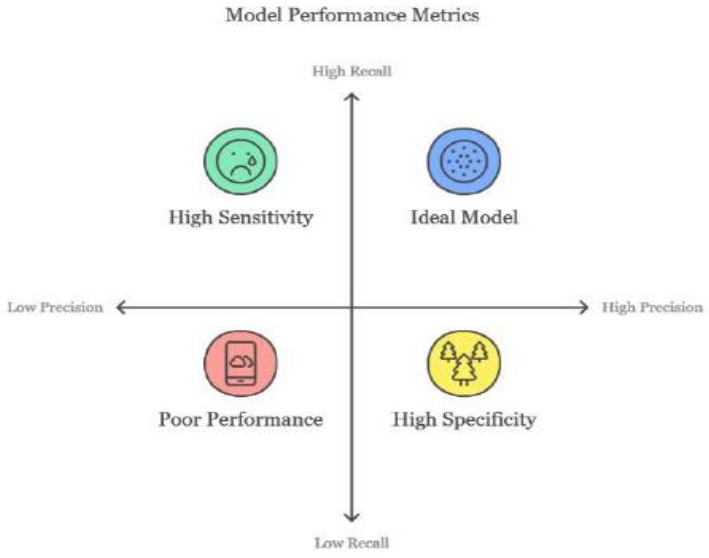


Figure 4: Evaluation Metrics

ISSN 2394 - 7780

Each model's performance was analyzed based on these metrics to identify the best-performing classifier for early ASD detection.

E. Statistical Analysis

We looked at some statistics of the dataset to see which features were most informative. We also studied the relationship between the Q-CHAT-10, ADOS and INCLEN scores. To figure effects that most helped predict autism, they used things like correlation matrices and feature importance. This study improved feature engineering and selection of model by revealing main drivers of ASD symptoms.

IV. RESULTS & DISCUSSION

This part will provide a full analysis of the ML techniques used in this study for ASD early detection. The accuracy, precision, recall, and f1-score are the key parameters used to evaluate the model performance. The metrics offered a deeper insight into the diagnostic potential of each model, especially their abilities to accurately identify ASD-positive and ASD-negative cases, which is vital for clinical reliability in early diagnosis.

A. MODEL COMPARISON AND ANALYSIS

Model Comparison:

	Accuracy	Precision	Recall	F1 Score
Logistic Regression	0.986667	0.974359	1.000000	0.987013
Support Vector Machine	0.980000	0.967949	0.993421	0.980519
Naive Bayes	0.990000	0.980645	1.000000	0.990228
Decision Tree	0.983333	0.974194	0.993421	0.983713
Random Forest	0.986667	0.980519	0.993421	0.986928
AdaBoost	0.970000	0.949686	0.993421	0.971061
Gradient Boosting	0.983333	0.974194	0.993421	0.983713
Bagging	0.980000	0.974026	0.986842	0.980392
Deep Neural Network	0.986667	0.974359	1.000000	0.987013

Figure 5: Results of Each Model

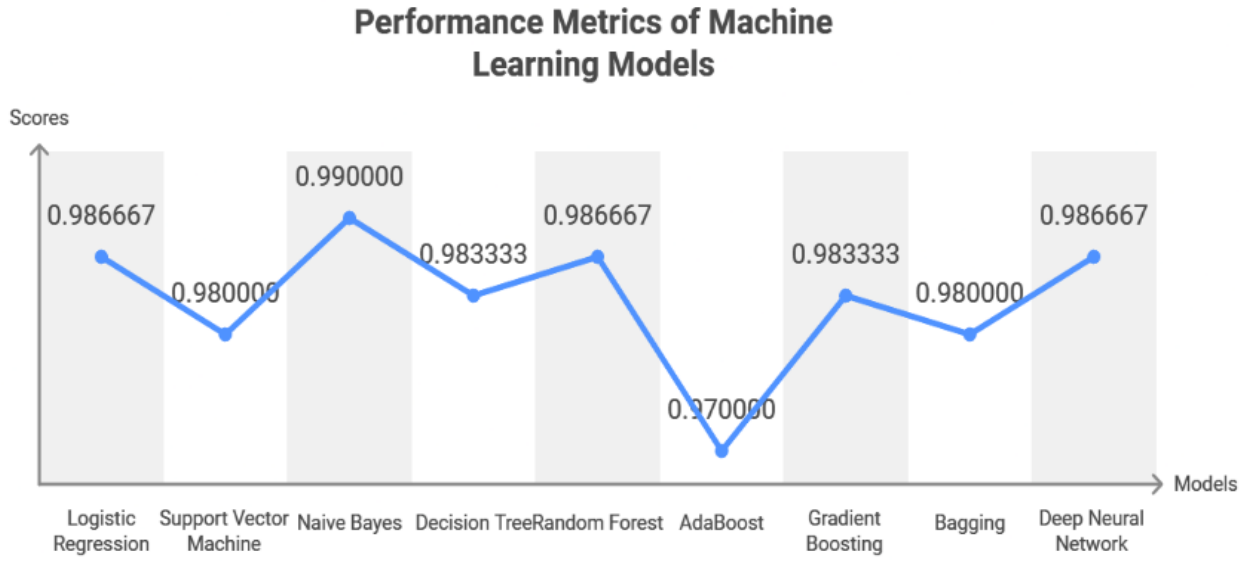


Figure 6: Model comparison wrt accuracy

B. KEY FINDINGS BY MODEL

1) *Naive Bayes*:

- **Highest Accuracy and F1-Score:** Naive Bayes obtained accuracy of 99% and F1-score of 0.9902 which suggests the superior diagnostic performance of the model. Naive Bayes captured the patterns well in the multimodal data due to its probabilistic nature with 98% precision and 100% recall.
- **Implications for Clinical Application:** Naive Bayes' high recall benefit helps create a false-negative model for early detection of ASD and thus limits the risk of cases. This strength in catching positive cases makes Naive Bayes a good choice for initial tests where inclusivity is key.

ISSN 2394 - 7780

2) Logistic Regression:

- Logistic regression was highly accurate (98.67%) and had high recall (100%) making it a very good predictor of ASD-positive. While logistic regression is no as precise as Naive bayes, it does have the advantage of being easy to interpret.
- **Potential in Clinical Settings:** Logistic Regression's balance of simplicity and accuracy makes it highly suitable for integration into clinical diagnostics. Its ease of interpretability allows clinicians to understand feature contributions to ASD risk, which could enhance transparency and trust in AI-assisted diagnosis.

3) Deep Neural Network (DNN):

- **Comparable Performance to Top Models:** DNN achieved an accuracy of 98.67% and an F1-score of 0.9870, on par with Logistic Regression. DNN's ability to learn complex, non-linear patterns in multimodal data enhanced its performance in distinguishing subtle ASD symptoms.
- **Strength in Complex Patterns:** The DNN's high performance suggests its capability to recognize intricate patterns in toddler behavior and communication, especially when data is comprehensive. However, DNN models can be computationally intensive and less interpretable, which may impact their practical use in resource-limited clinical environments.

4) Random Forest:

- **High Precision and Robustness:** With an accuracy of 98.67% and precision of 98%, Random Forest demonstrated excellent reliability, balancing sensitivity (recall of 99.34%) with specificity. The model's ensemble structure allowed it to capture diverse patterns without overfitting, making it robust against variations in behavior and sensory data.
- **Clinical Application Potential:** Random Forest's ability to handle large datasets and high variance makes it a suitable choice for ASD diagnostics, where diverse symptom presentations require adaptive models. Additionally, the model's interpretability is beneficial for clinicians needing insights into feature importance.

5) Support Vector Machine (SVM):

- **Balanced Accuracy and Recall:** SVM displayed competitive accuracy (98%) and recall (99.34%), performing well in identifying ASD-positive cases with relatively high precision (96.79%).
- **Challenges and Suitability:** While SVM is adept at handling high-dimensional data, it can be computationally demanding, especially when the dataset includes diverse features. Its complexity might limit practicality in some clinical settings, though its decision boundary optimization is valuable for datasets with complex symptom overlap.

6) Decision Tree and Gradient Boosting:

- **Reliable Performers with High Recall:** Both models showed comparable results, with accuracies of 98.33% and F1-scores of 0.9837. Their high recall rates (99.34%) suggest an effectiveness in detecting ASD-positive cases, although they slightly lag behind the top performers in precision.
- **Use Cases in Preliminary Diagnostics:** Decision Tree models are highly interpretable, while Gradient Boosting provides robustness in handling noisy data, making them suitable for early-stage screening tools. However, they may not be as precise as ensemble methods like Random Forest.

7) AdaBoost and Bagging:

- **Moderate Performance:** AdaBoost and Bagging showed accuracy rates of 97% and 98%, respectively. Both methods achieved lower precision than top-performing models but maintained high recall, indicating reliability in identifying ASD-positive cases.
- **Considerations for ASD Screening:** Although these models did not outperform others, their simplicity and high recall rates make them useful in situations where a quick, preliminary assessment is needed.

ISSN 2394 - 7780

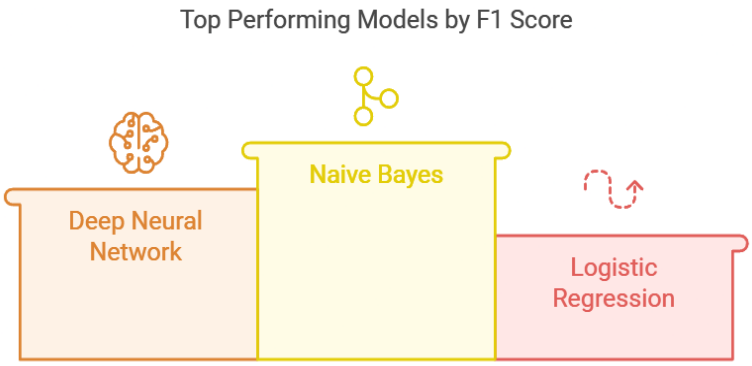


Figure 7: Top Models by F1 Score

C. SUMMARY OF RESULTS

The comparison reveals that Naive Bayes, Logistic Regression, Deep Neural Networks, and Random Forest are particularly effective for early ASD detection. Naive Bayes leads in diagnostic accuracy, while Logistic Regression and Random Forest offer high reliability with the added advantage of interpretability. Deep Neural Networks excel in identifying complex patterns but require more computational resources.

D. IMPLICATIONS AND FUTURE DIRECTIONS

The results demonstrate that a multimodal, machine learning-based approach can significantly enhance ASD diagnosis in toddlers. The integration of multiple diagnostic tools in the dataset allows the model to detect ASD with high accuracy and minimal false negatives. Since early intervention is crucial, it can be seen that models with high recall such as Naive Bayes and Logistic Regression are good candidates for initial screening.

Further research may also involve attempting to collect more behavioral and genetic data to add to the current dataset in order to enhance the robustness of the model. In addition, the use of XAI techniques especially for complex models such as DNNs can help in increasing the transparency of the machine learning models used in diagnosing ASD to enable clinicians to understand the factors that influence the predictions.

E. LIMITATIONS AND FUTURE WORK

1) Dataset Limitations:

One limitation of this study is the range and size of the dataset. We use Q-CHAT-10, ADOS and INCLEN as multiparameter axes which has improved the precision of the diagnosis but the dataset may not be adequate. One way to make this model more robust would be to expand the dataset to include cases from a wider variety of demographic regions, ages and severity of symptoms.

2) Model Enhancements:

Upcoming studies may investigate further model optimization, such as hyperparameter adjustments and sophisticated frameworks like convolutional neural networks (CNNs) or recurrent neural networks (RNNs), especially for analyzing sequential behaviors. Additionally, integrating ensemble techniques with neural networks may enhance diagnostic precision by utilizing the advantages of different classifiers.

3) Potential Use in Clinical Settings:

While the results indicate promising diagnostic performance, more research and clinical validation are needed before implementing this model in medical environments. Collaborating with healthcare providers to conduct real-world testing can ensure the model's reliability and usability, paving the way for it to support clinicians effectively in early ASD diagnosis.

4) Practical Implications:

This research is important for the diagnosis of ASD. By providing a tool that is unbiased and based on data, clinicians and parents can make diagnoses in a timely manner to allow interventions to take place, which benefits cognitive and social development outcomes. If you catch it early, you can make a special plan for your therapy that will match your kid. This has long benefits. Also, allowing a range of diagnostic tools to gather information one the patients and their journeys helps build a wider overview of ASD.

5) Ethical Considerations:

Most ethical issues must be looked into once the machine learning applications in medical diagnostics are made. When managing sensitive health data, it is important that data privacy is ensured—so that patient data is secured and anonymity maintained. Furthermore, informed consent is important if sensitive data is collected from children. The model's model performance may affect diagnostic use. Most importantly, a false positive or false

ISSN 2394 - 7780

negative would have a profound impact on the child and family. Because of this, it is suggested that a stringent phase of clinical validation should take place to ensure ethical use.

V. CONCLUSION

This study has shown that a machine learning–based multimodal approach has the potential to help detect Autism Spectrum Disorder (ASD) in children efficiently. This study integrates well-established diagnostic tools such as the Q-CHAT-10 and ADOS with the INCLEN to develop a holistic picture of ASD's behavioral, sensory, and social characteristics. A study that uninvolved some classifiers, namely Naive Bayes, Logistic Regression, Random Forest and Deep neural network, suggested the best model that differentiate an Autism Spectrum Disorder-positive cases from Autism Spectrum Disorder-negative cases accurately and effectively.

The results show that using a multimodal dataset is helpful as it makes the model learn complex patterns and subtle symptoms that one-dimensional assessments miss.This method helps with the problems with the traditional diagnosis of this disorder. The cases of subjectivity are decreased and the accuracy of diagnosis is improved. This is very important in cases of early diagnosis. The model delivered strong performance in terms of accuracy, precision and recall, suggesting that the model could be used as an objective diagnostic tool to help clinicians and parents plan timely intervention.

Yet, this study recognizes some limitation, such as the limitation of data set and broad validation. In the future, researchers should collect a bigger dataset and use better machine learning techniques to ensure the model works for everyone. Also, to ensure realistic and effective applications in medicine, the privacy of healthcare data and clinical validation of the lingo of care model required.

Ultimately, this study contributes to the field of ASDs diagnostics through the development of a reliable, data-driven tool that could lead to earlier diagnosis and timely intervention with developmental benefits for children with ASD. This study provides a useful basis for improving machine learning-based applications for other neurodevelopmental disorders to make the diagnosis of autism spectrum disorder more accurate and accessible.

ACKNOWLEDGEMENT

I would like to thank my mentor, Professor Sunita Jena, for providing me with valuable assignment ideas, suggestions, and encouragement during the whole research. This study could not have been completed successfully if it weren't for her expertise and assistance. I also thank the professor Fatima Shaikh in-charge of research for the guidance and Mr. Wilson Rao Head of the Department for their support and inspiration. I extend my heartfelt gratitude to the MSc Big Data Analytics Department, Jai Hind College for providing necessary resources and support to carry out this task. The department's encouragement and facilities have greatly helped the progress and outcome of our study. Last but not the least, I would like to express my special thanks to clinical psychologist Doctor Tarannum Qureshi for her valuable insights during the course of my research. Her guidance and domain knowledge was immeasurably valuable to me in order to understand ASD. And her knowledge to identify research gaps and idea to incorporate multiple tools in the study is what has made my study different from other studies on ASD. I want to express my gratitude to everyone who made this research possible.

REFERENCES

[1] Bone, D., Bishop, S., Black, M. P., Goodwin, M. S., Lord, C., & Narayanan, S. S. (2016). Use of machine learning to improve autism screening and diagnostic instruments: effectiveness, efficiency, and multi-instrument fusion. *Journal of Child Psychology and Psychiatry*, *57*(8), 927-937. https://doi.org/10.1111/jcpp.12559

[2] Duda, M., Ma, R., Haber, N., & Wall, D. P. (2016). Use of machine learning for behavioral distinction of autism and ADHD. *Translational Psychiatry*, *6*(2), e732. https://doi.org/10.1038/tp.2015.217

[3] Kosmicki, J. A., Sochat, V., Duda, M., & Wall, D. P. (2015). Searching for a minimal set of behaviors for autism detection. *Computer Science in Behavior Research Methods*, *47*(2), 477–485. https://doi.org/10.3758/s13428-014-0472-2

[4] Thabtah, F. (2019). Machine learning in autistic spectrum disorder behavioral research: A review and ways forward. *Informatics for Health and Social Care*, *44*(3), 278-297. https://doi.org/10.1080/17538157.2018.1433675

[5] Thabtah, F., Peebles, D., Retzler, J., & Sumner, M. (2020). A machine learning autism classification based on Behavioral Features. *Health Information Science and Systems*, *8*, 5. https://doi.org/10.1007/s13755-019-0086-2

ISSN 2394 - 7780

[6] Wall, D. P., Kosmicki, J., Deluca, T. F., Harstad, E., & Fusaro, V. A. (2012). Use of machine learning to shorten observation-based screening and diagnosis of autism. *Translational Psychiatry*, *2*(4), e100. https://doi.org/10.1038/tp.2012.10

[7] Zablotsky, B., Black, L. I., Maenner, M. J., Schieve, L. A., Danielson, M. L., Bitsko, R. H., ... & Kogan, M. D. (2019). Prevalence and trends of developmental disabilities among children in the United States: 2009–2017. *Pediatrics*, *144*(4), e20190811. https://doi.org/10.1542/peds.2019-0811

[8] Zhao, Y., Egger, H. L., Duan, W., Leventhal, B. L., & Shankman, S. A. (2019). Multimodal machine learning for the detection of autism spectrum disorder. *Behavior Research Methods*, *51*(3), 1472-1485. https://doi.org/10.3758/s13428-018-1105-y

[9] Tajsic, A., Mazefsky, C. A., Rosen, T. E., & Siegel, M. (2021). Machine learning models for predicting aggressive behavior in youth with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *51*(2), 606-617. https://doi.org/10.1007/s10803-020-04561-y

[10] Ozonoff, S., Young, G. S., Steinfeld, M. B., Hill, M. M., Cook, I., Hutman, T., ... & Sigman, M. (2009). How early do parent concerns predict later autism diagnosis? *Journal of Developmental & Behavioral Pediatrics, 30*(5), 367-375. https://doi.org/10.1097/DBP.0b013e3181ba0fcf

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

ANOMALY DETECTION IN NEFT TRANSACTIONS

[1]Ayush Kumar Arun Kumar Mishra, [2]Sunita Jena and [3]Dr. Balkrishna Parab
[1]Department of Computer Science, Jai Hind College (Autonomous), Mumbai, India
[2]Assistant Professor, Department of IT, Jai Hind College (Autonomous), Mumbai, India
[3]Director, Aditya Institute of Management Studies and Research (AIMSR), Mumbai, India

ABSTRACT
Anomaly detection-a fundamental approach in data mining and machine learning-is based on finding any deviations in data patterns with respect to the expected behavior. This area of study finds application in various fields such as fraud detection, network security, and financial systems, where pinpointing any irregularities is key to safety and integrity. Anomaly detection in NEFT transactions was used to continually monitor the patterns of transactions for fraud or abnormality that would indicate security threats or human error. NEFT transactions represent high-volume data and require real-time access, so it is essential to use robust and powerful algorithms capable of finding minutiae in large datasets. In this study, three popular unsupervised anomaly detection algorithms-Local Outlier Factor (LOF), Isolation Forest, and Autoencoders-have been employed. LOF identifies various anomalous points according to deviations from the local density, thus effectively spotting outliers in transaction data interaction complexities. Isolation Forest does this by recursively isolating the anomalous data points, thus offering great efficiency in dealing with high-dimensional datasets. Autoencoders are a type of neural network with excellent data representation-learning ability; they mark a broad scope to reconstruct normal transaction patterns and flag all atypical patterns as anomalies. This research discusses and evaluates the relative performance of these methods on NEFT transactions, along with the salient features of each model in capturing anomalies within financial datasets. By investigating robust methodologies suitable for NEFT transactions, the study provides a significant boost towards enhancing fraud detection in financial systems.

Keywords *— Anomaly detection, NEFT transactions, fraud detection, Local Outlier Factor, Isolation Forest, Autoencoders, unsupervised learning, financialdatasets.*

I. INTRODUCTION
An unexpected occurrence or deviation, involving instances of data, events, or instances that differ significantly from the general distribution within data, anomaly detection becomes a favored substrate in the applications such as fraud detection, network security, and financial transaction monitoring for identifying irregularities that might indicate malignant activities or system failures. This work addresses the field of applying anomaly detection to NEFT transactions, hence within the mainstream of monitoring financial transactions, establishing identifying features of suspicious transactions or unforeseen patterns. Typically, datasets contain instances and attributes of machine learning, with instances representing groups of data pertaining to individual data points, while attributes signify certain features of the data in question. Datasets are either labeled or unlabeled depending on the information with respect to labels. Due to the presence of labels that explicitly indicate the class, supervised learning is trained on a labeled dataset, determining the mapping from inputs to specified outputs. In unsupervised learning, clustering and anomaly detection techniques accept input of any class without labels and hence find use when no prior knowledge of the class is available. In our case, NEFT transaction data is unlabeled; this allows us to run unsupervised algorithms that can independently identify anomalies without pre-existing knowledge of fraudulent labels [1].

A. Role of Anomaly Detection in NEFT Transactions
Anomaly detection processes in transaction activities of NEFT help to detect the emergence of irregularities for fraud detection, thus increasing operational efficiency, compliance with regulations, and customer protection. Anomaly detection helps identify unusual transaction patterns, such as unexpected transaction amounts or transaction frequencies, which might help wasting resources towards prioritizing high-risk cases. This will support AML and KYC requirements and helps assess customer risk by blocking unauthorized transactions. It also ensures the integrity of the system by pointing out potential problems thereby furthering the purpose of safeguarding the financial environment [2].

B. Significance of the Study
The effect of the study of anomaly detection in NEFT transactions is very crucial as it attempts to address some very relevant problems that modern financial systems face with regards to fraud detection, regulatory compliance, and client trust. With a fair understanding of the anomaly detection model, it will bring forth different tools for finding abnormal transaction patterns with greater accuracy. This will help to detect an early

ISSN 2394 - 7780

onset of abnormal transaction patterns and ultimately save resources and reduce the chances of financial crime, thereby protecting the users' assets. This study will also help financial institutions meet these sets of requirements for AML and KYC compliance, reducing the risk of regulatory penalties while improving their reputations. In addition, anomalous detection methods will enhance operational efficiency through process automation, thereby freeing resources to focus on other critical areas. In sum, this study enriches the electronic fund transfers' broader security and resilience, thus representing a further step toward establishing a safer financial environment [3].

II. METHODOLOGY

Several machine learning methods for anomaly detection rooted in the principles of LOF, Isolation Forest, and Autoencoders will be analyzed and looked at. Since the objective is to discover insidious and unexpected behaviors hidden within the unlabeled data, these three models provide complementary angles regarding how to understand anomalies.

- **Local Outlier Factor (LOF):** LOF is a density-based method that determines the local density deviation of a spatial pattern with respect to its neighboring points. LOF has been found suitable for outlier detection in close groups with irregular densities, enabling it to discover outliers based on local variations of the density [4].
- **Isolation Forest:** This ensemble model isolates anomalies through recursive partitioning of the data according to the feature values. Abnormal points tend to be isolated with a small number of partitions compared to normal data points. Isolation Forest is effective for high-dimensional data, and therefore fits its application suitability into a real large-scale anomaly detection scenario [5].
- **Autoencoders:** Autoencoders are a type of neural network partly trained to perform unsupervised learning. It works by training the network to produce a compressed representation of the given input data. The normal labeled data reconstructs well, providing a low reconstruction error, while anomalies would be poorly reconstructed relative to normal data. This approach is very good at modeling nonlinear relationships, which makes it effective in dealing with complicated data patterns [6].

III.DATASET REVIEW

The dataset of interest looks at transactions executed under NEFT. It consists of 33,249 entries and contains several factors pertinent to these transactions.

The month column encompasses 191 temporal entries during which the transactions were carried out. In the case of the financial institutions, 295 unique bank name entries are present with "B N Paribas" being cited as the many institutions. The dataset shows a significantly wide range for a number of transactions and amounts of money.

The no. of debit transactions column shows on average approximately 1,120,201 debit transactions with significant fluctuations around this mean as shown by the standard deviation. The maximum number of debit transactions recorded under the maximum quoted figure is 358,140,270.

The amt of debit transactions column describes on average a transaction amount of 749,987, though the standard deviation reflects quite some variation around the mean. The highest amount for debit transactions is 83,575,594.56. A similar trend is noticed in credit transactions. The no. of credit transactions column indicates an average of about 1,120,136 transactions, with variability suggested by meaningful standard deviation around this mean. The maximum recorded credit transactions total 193,142,132.

In terms of money, the amt of credit transactions mean is seen as 750,751 again with variable standard deviations indicating meaningful variability, and a peak of 73,590,501.40. To summarize, this database provides a picture of NEFT transactions across different banks and months, providing insights into the frequency and transaction value of these e-payments.

IV.IMPLEMENTATION

1. Data Preprocessing:

- **Dataset Loading:**

Dataset neft_transaction_metrics.csv is read by the function pd.readcsv which reads the NEFT transaction metrics into a pandas DataFrame [7].

- **Feature Selection:**

This involved selecting a subset of key features from the dataset:

ISSN 2394 - 7780

No of debit transactions, amt of debit transactions, No of credit transactions, amt of credit transactions, month, bank name. This selection emphasized debit and credit metrics, along with the transaction amounts, month, and bank name for better analytical insight.

- **Handling Missing Values:**

Identification consisted simply of numeric columns only for filling in missing values.

The NaN values of each column were filled with the mean to avoid bias and keep the central tendency of the data.

- **Data Scaling:**

In normalizing the dataset, only numeric features were selected. StandardScaler standardized the data to ensure that all numerical features had a mean of zero and a standard deviation of one. This enhanced the convergence rate and performance level of the model.

2. Exploratory Data Analysis (EDA):

To get good insights from the data at hand concerning transaction patterns and relationships between numerical variables, we undertook the following EDA processes:

- **Transaction Counts per Bank:**

By means of a count plot, we were able to visualize how transactions spread over different banks. This helped highlight which banks had more transaction volumes, contributing to understanding levels of engagement across banks.

- **Correlation Analysis of Numeric Columns:**

The correlation heatmap was created for the specified numeric columns covering transaction counts and amounts. This holds a consolidated view of how both transaction metrics relate to one another, pointing out potential dependencies or similarities.

Highly correlated variables indicate transactional patterns and help in feature selection for predictive modeling.

These visualizations provide foundational insights for identifying key patterns and relationships in the data, thus facilitating informed and critical data preprocessing and feature engineering decisions.

3. Clustering and Best Model Selection

Clustering methods, K-Means, Gaussian Mixture Model (GMM), and Density-Based Spatial Clustering of Applications with Noise (DBSCAN), were then applied to identify meaningful clusters from the dataset. Selecting the model that best accounts for the underlying structure in the data becomes the next stage of this research. In this case, the silhouette scores serve as the evaluation metric [8].

- **Model Initialization:**

Each clustering model is initialized with gear conducive to the dataset's characteristics.

a. **K-Means:** Set to create 3 clusters (n_clusters=3) with a fixed random state (random_state=42) for reproducibility.

b. **Gaussian Mixture Model (GMM):** Estimated 3 clusters (n_components=3) with the same random state for consistency in clustering behavior.

c. **DBSCAN:** Epsilon (eps=0.5) considers the distance between two points for them to be classified as neighbors. A minimum of 5 or more samples (min_samples=5) will mean that we can classify points as being in a dense region.

- **Model Fitting and Label Assignment:**

Each clustering model was fitted on the preprocessed data (features_scaled) to predict cluster labels. The predicted labels were stored as follows:

a. **kmeans_labels:** Cluster labels generated by the K-Means model.

b. **gmm_labels:** Linear model-generated cluster labels.

c. **dbscan_labels:** Cluster labels assigned by the DBSCAN model, where -1 indicates noisy points that do not belong to any cluster.

ISSN 2394 - 7780

- **Evaluation via Silhouette Score:**

To evaluate model performance, silhouette scores were computed for each model. The silhouette score computes how similar an object is to its cluster than to other clusters and a higher score is an indicator of better-defined clusters. **Silhouette score:**

K-Means: 0.9273

GMM: 0.4182

DBSCAN: 0.8562

Thus, based on the results, clustering observed that the K-Means model delivered model accuracy over the others as observed through the silhouette score.

- **Best Model Selection:**

Among the different clustering methods, the K-Means is dubbed, as per the silhouette scores, the best performing one.

4. Anomaly Detection Using the Best Model

For each of the clusters identified by the best-obtaining K-Means model, anomaly detection was run based on the following: Local Outlier Factor (LOF), Isolation Forest, and an Autoencoder-based neural network, building an ensemble approach for robust anomaly detection.

- **Initial Anomaly Flags:**

The results of each detection method were stored in three column flags added to the main data frame:

LOF-Local Outlier Factor Store.

Iso Forest-Isolation Forest Store.

Autoencoder-Autoencoder model store.

- **Anomaly Detection within Clusters:**

1. Local Outlier Factor (LOF):

In this case, a density-based method compares the local densities of data points with the lowest and highest outliers. It was configured with 20 neighbors (n_neighbors=20) and a contamination rate of 10%. Therefore, anomalies are flagged -1 and normal points 1.

2. Isolation Forest:

By way of this method, a tree-based model, the points are isolated by an artificial method through random splits. The method used is configured with a contamination rate of 10% with a fixed state. This means that it will also flag outliers as -1.

3. Autoencoders:

A neural network model was engineered with the objective of reconstructing the input data, thus labeling any point as an anomaly if it showed a significant reconstruction error. The Autoencoder architecture contains a chain of hidden layers with 32, 16, 8, 16, and 32 units. Each one was trained for 50 epochs, and using a cut-off (90th percentile) for significant reconstruction error, the model would flag such points as anomalies (-1).

- **Ensemble Anomaly Detection:**

For this reason, to take advantage of the qualities of each model in the procedure, an ensemble method was used whereby a new column, called ensemble anomaly, was created. It contained the count of the models detecting each point as an anomaly, thereby allowing a more trustworthy indication for anomalies based on the various detection methods.

- **Final Anomaly Assignment:**

A majority voting-protocol-based approach was used such that the final anomaly column flagged points that were detected as anomalies by at least two models. This threshold of two models helped increase the robustness of anomaly detection by making it mandatory that there is agreement between detection methods in order to reduce false positives and allow more reliable labeling of anomalies.

V. RESULTS

Three clustering algorithms, namely, K-Means, Gaussian Mixture Model (GMM), and DBSCAN were chosen in the current study to ascertain the optimum model for clustering the dataset. The models' performances were assessed using, and not limited to, the Silhouette Score.

ISSN 2394 - 7780

K-Means scored the highest Silhouette Score of **0.9273,** indicating well-defined clusters. DBSCAN was the second, scoring at **0.8562**, suggesting moderate performance in clustering. Finally, GMM scored the lowest with a score of **0.4182**, indicating less distinguishable clustering. Hence, K- Means model based on this performance as the tightest clustering model was chosen to be the best clustering predictor for this data The NEFT systems trans actual analysis on the measure of transaction amounts reveals point anomalies as:

- **The number of Credit Transactions:** The high proportion of anomalies is observed on the lower range transaction amounts indicating that some monthly outlier spikes are apparent and unusual.
- **The Number of Credit Transactions:** Monthly outlier spikes in the frequency of transaction occurrence, implies that anomalies are limited to the lower boundary of transactions.
- **The amount of Debit Transactions:** The largest dispersion of anomalies is recorded with regards to this metric with very high and low anomalous values occurring suggesting a strong time-based seasonality in debit amounts fluctuating through time.
- **The number of Debit transactions:** Temporal Debit recurring spikes are well stabilized as the transactions count rises irrespective transaction amounts, with the exceptions of lesser occurrences indicating stronger anomalies.

Overall Finding: The **amount of Debit transactions** which are notable under this metric emerge with the widest spread, suggesting that these types of transactions have a less uniform occurrence and can be more used to measure the presence of outlier activity [9].

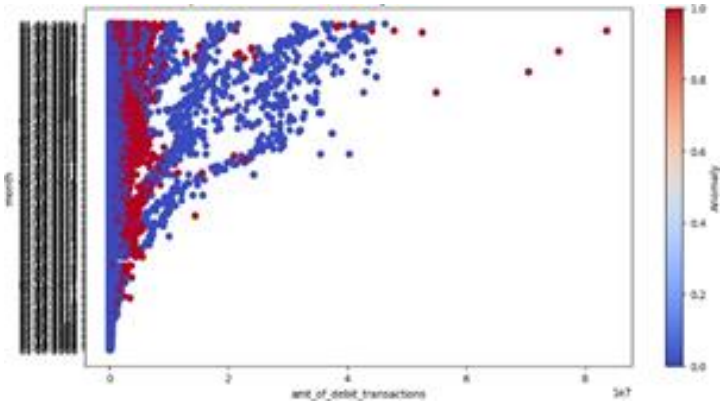


Fig.1 Anomaly Detection with Best Clustering Model and Ensemble Methods

VI. CONCLUSION

This study uses a comprehensive clustering and error detection framework, comparing three clustering models—K-Means, Gaussian Mixture Model (GMM), and Spatial Clustering with Noise (DBSCAN)—to identify the best model. Based on the silhouette scores, K-Means emerged as the best model to capture well-defined clusters in the dataset.

A combination of three methods is used to detect anomalies within each identified group: local outlier factor (LOF), isolation forest, and autoencoder-based neural networks. Each method has unique advantages – LOF for local density sensitivity, Exclusion Forest for efficient outlier isolation, and Autoencoders for reconstruction-based anomaly detection. Errors are scored if detected by at least two of these methods, providing a standardized approach that minimizes false positives.

By combining clustering and multi-model anomaly detection, this approach provides a powerful framework for understanding data structure and identifying true outliers within each cluster. This compromise approach ensures greater reliability and flexibility for complex datasets, enabling accurate collection and sensitive anomaly detection for future analysis.

REFERENCES

1. Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. ACM Computing Surveys, 31(3), 264–323.
2. Al-Doghman, J. S. S. E. B. (2011). Anomaly detection techniques for fraud detection. Journal of Financial Crime, 18(2), 97–103.
3. Hasan, M. H. Z., Ganaie, M. A., & Khan, M. M. U. (2014). Anomaly detection in financial transactions for fraud detection. Computers & Security, 45, 152–169.

ISSN 2394 - 7780

4. Breunig, M. R., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). LOF: Identifying density-based local outliers. In Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data (pp. 93–104).
5. Liu, L., Wu, J., & Wu, Z. (2011). Isolation forest. In 2011 IEEE 11th International Conference on Data Mining (pp. 413–422).
6. Bengio, Y. (2009). Learning deep architectures for AI. Foundations and Trends® in Machine Learning, 2(1), 1–127.
7. The pandas development team. (2025). Pandas: A powerful data analysis toolkit. Retrieved April 2025, from https://pandas.pydata.org
8. Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining(pp. 226–231).
9. MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1 (pp. 281–297).

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

ENHANCING DIGITAL TRANSACTIONS: THE POWER OF PREDICTIVE ANALYTICS IN UPI APPS FOR IMPROVED SECURITY AND USER EXPERIENCE

[1]Bhawna Puraswani and [2]Arsh Shaikh
[1, 2]Independent Researcher

ABSTRACT
The rapid adoption of Unified Payments Interface (UPI) applications has significantly transformed the digital payments landscape, offering unparalleled convenience and user engagement. However, this rapid growth has also raised concerns regarding user experience management and security at scale. This paper examines the potential benefits of predictive analytics on customer behavior within UPI applications by conducting an in-depth survey of users from many different demographics. We want to bring to light the user's behavioral tendencies, security crises and the fintech sector growth patterns, by means of the behavioral data and the users' demographic details. Our study is mainly on machine learning and data analytics which are used in converting transaction data to consumer behavior predicting, improving the personalization system, supporting fraud detection, and creating a more secure financial environment. The literature review highlights advancements in predictive analytics within fintech and outlines future research directions, particularly the integration of machine learning into UPI applications. We also address the ethical implications of predictive analytics, including data protection and user consent. By integrating data science with fintech, we propose how predictive models can drive more customized and secure UPI applications, ultimately enhancing user engagement and financial safety.

Keywords: *Unified Payments Interface (UPI), Predictive Analytics, Consumer Behavior, Fintech Security, Machine Learning, Data Science, Fraud Detection, User Experience, Financial Technology*

I. INTRODUCTION
The rapid growth of Unified Payments Interface (UPI) applications has brought about transformative change to the digital payments landscape of India, offering unparalleled convenience for consumers and shifting their recurring behavioral patterns. Thanks to its integration with popular apps (Google Pay, PhonePe and Paytm) that have become a part of every day life, UPI has simplified not only peer-to-peer (P2P) but also peer-to-merchant (P2M) transactions by enabling universal ‘one-click’ payments and has been the key driver behind exponential adoption of digital payments. These changes had been bolstered by government steps and the COVID-19 epidemic, and UPI evolved into a popular payment method with an average annual growth of more than 200% over the past years [4].

While UPI transactions have been widely accepted, the rapid rise in these transactions has posed numerous challenges, especially in terms of security and customer experience. The risks of fraudulent activities increase with transactions reaching greater volume and complexity, making it necessary for UPI players to embrace more advanced security measures. In addition, traditional rule-based methods used within fraud detection systems do not adapt to the continuously changing fraud profile, suggesting several points of vulnerabilities within such mechanism, which call for better methods, perhaps specifically designed to counter these emerging threats over the changing aspects of digital payments [3].

AI and machine learning predictive analytics is a promising solution to these challenges as it analyzes a massive amount of transaction data, allowing to anticipate user behavior, detect anomalies, most importantly — create real-time fraud prevention mechanisms. Using advanced algorithms, predictive models can not only offer personalized insights and recommendations to improve users experience but also increases fraud detection accuracy [2]. This bifocal approach of safety and personalization can lead to increased user engagement and trust which is critical for sustaining the growth momentum of UPI.

This paper seeks to investigate how predictive analytics can reshape UPI applications by solving critical areas like fraud detection, improving user experience, and addressing ethical issues surrounding data usage. This involved a survey designed to gather information about user perceptions of security, willingness to engage with predictive services, and motivation for using data-driven features. Through the knowledge gained from this study, we hope to build a more secure financial ecosystem that is respective of the users they serve, while at the same time providing valuable benchmarking for future work in the field of fintech research.

II. LITERATURE REVIEW
UPI applications have revolutionised the Digital Payment landscape in India in no time. As UPI transactions proliferate, security issues and user experience problems have arisen, leading to a need for predictive-analytics

ISSN 2394 - 7780

advancements. In this article, a review is performed that determines the application of predictive analytics, artificial intelligence (AI), and different machine learning techniques to recognize, analyze, and analyze evolving cyber threats to improve the security and usability of UPI systems.

1. **The Role of AI in Enhancing UPI Security and Usability**

Vasan speaks to the transformative nature of AI on real-time payment systems, but also explains how UPI will require some serious security measures as a response. Techniques powered by AI, like biometric verification and real-time fraud detection, can recognize and react to indicators of dubious conduct in the blink of an eye. This enables a higher level of security and makes UPI even faster and has enhanced the efficiency of payment processing. UPI systems, through predictive analytics, can predict potential security threats based on transaction trends, reducing fraud prevalence [1].

2. **Predictive Analytics for Fraud Detection in UPI Transactions**

Kavitha et al. emphasize the use of machine learning models — Hidden Markov Models and neural networks — in identifying outliers in UPI transactions. By leveraging this data, UPI systems can utilize machine learning models to pinpoint abnormal transaction patterns and flag potentially suspicious transactions in real-time. It explains that predictive analytics offers a clear edge over static, rules-based fraud detection methodologies by constantly adapting to changing patterns of fraud [3].

3. **Personalized Payment Recommendations to Enhance User Experience**

Higher user engagement with the help of data-driven personalization in UPI applications. Dutta and Poornima recommend that transaction data and demographic information, which seems like common sense as they allow developing predictive models and suggest personalized payment offers. This can open the door to much higher levels of user satisfaction, as well as security outcomes, since patterns in user behavior can help catch signs of fraud early [2]. Likewise, the fintech study detailing the importance of personalized recommendation, supports this by showcasing approaches to cater customer around the UPI experience [5].

4. **Addressing Growth and Scalability Challenges in UPI Through Predictive Analytics**

With the rapid adoption of UPI, the volume of transactions continues to rise exponentially, leading to concerns, especially from a scale and security perspective. Interviewer: Singh talks about video payments and other digital payment trends in India and implications for that for UPI. Scalability and performance challenges can be addressed using predictive analytics and AI that leverage big data techniques to analyze transaction trends and detect potential new fraud patterns. Thus even with growing transactional loads, this will keep the UPI infrastructure strong and secure [4].

5. **Enhancing UPI Usability and Security with Predictive Features**

The EZPay study adds various predictions, improving the usability and security of UPI systems. Users can track their imperfections promptly with the real-time notifications, intuitive dashboards, and predictive elementary alerts. The UPI can offer a more user-centric experience by using predictive analytics and providing a protective barrier for users against threats. This is made possible by CIBER [6].

6. **Comparative Analysis of Machine Learning Techniques for Fraud Detection**

The Survey on supervised machine learning algorithms for fraud detection contrasts a number of methods used on credit card fraud detection and their applicability for UPI. Note that algorithms such as Random Forests and Support Vector Machines are very effective for analyzing transaction data to detect anomalies. These conclusions, drawn in this paper, indicate by establishing UPI using some similarity machine learning approaches false positive and negative rates can be decreased, ensuring a better identification of fraudulent activities [7].

7. **Systematic Review of AI and Machine Learning in Banking**

Their meta-analysis, cited here, confirms the disruptive, transformative power of AI and machine learning in the banking sector, encompassing all aspects of digital payments. The review ukulele the potential benefits of predictive analytics integration in UPI applications to improve security and user experience. By detecting discrepancies before fraud has a chance to happen, AI is pushing the boundaries for predictive models making payment methods upon internet, less prone to risks and more proficient [8].

8. **Understanding User Perceptions and Trust in UPI**

For mass adoption, the user confidence in UPI apps is essential. A user survey on UPI sentiment shows that trust in the system hinges on the security provisions of the system and the prevention of fraud. It can increase user trust with consistent security protocols that are adaptive from monitoring and learning user behaviors. Likewise, studies on user perceptions show that perceived security and ease of use are both significantly associated with customer satisfaction. The recommendation for users induced of problems with security and that

ISSN 2394 - 7780

are a leading cause of lower adoption rates of UPI-based mobile payment apps is to implement predictive models that increase the utility and intuitive power of these applications [9][10].

9. **The Potential of AI-Driven Personalization in UPI for Financial Inclusion**

Alongside security and user experience, AI-fueled personalization in UPI applications can facilitate financial inclusion. Predictive analytics can use the limited data available to identify and make personalized suggestions as to which product may be relevant to that user, and thus, help an underserved population gain access to a particular financial product, which supports the diverse needs of UPI users. This not only broadens UPI's horizon but also keeps UPI to be an inclusive and user-centric digital payment solution [2] [5].

III. METHODOLOGY

1. Research Design

- **Survey-Based Quantitative Analysis**
  - A survey was administered to gather the data on users perceptions in terms of the security and usability of UPI apps and willingness to use predictive analytics features.
  - The survey asked questions relating to security concerns, desire for predictive services (e.g., fraud detection, budgeting), willingness to share transaction history for machine learning, and notification preferences based on predictive information.
- **Data Analysis and Interpretation**
  - Data classification: The analysed survey data across several trends and correlations between user perception on predictive analytics features and user interest.
  - Visualizations of this survey data were made to help support some of this analysis.

2. Survey Design

The survey was composed of multiple choice, Likert scale, and open-ended ranking questions that intended to evaluate:

- **User Security Perception**: How secure users feel when using UPI apps for financial transactions.
- **Predictive Analytics Awareness and Interest**: User interest in new predictive services such as fraud detection and budgeting recommendations.
- **Data Privacy Concerns**: Comfort levels regarding UPI apps using transaction data to train AI models for fraud detection.
- **Demand for Personalized Insights**: Relevance of features such as monthly spending trends and automatic spending categorization.
- **Notification Preferences**: Predictive-based data about how often and in what way users want to receive notifications (for example, information about overspending or personalized discounts).

The survey was conducted online, targeting a diverse demographic of UPI app users across different age groups, occupations, and regions.

3. Data Collection

- The survey was distributed to UPI users through online platforms, including social media and email, to receive wide range of responses.
- A total of 73 responses were collected over a two-week period, ensuring a mix of demographics.

4. DATA ANALYSIS

ISSN 2394 - 7780

4.1. Security Perception and Action Based on Predictive Insights

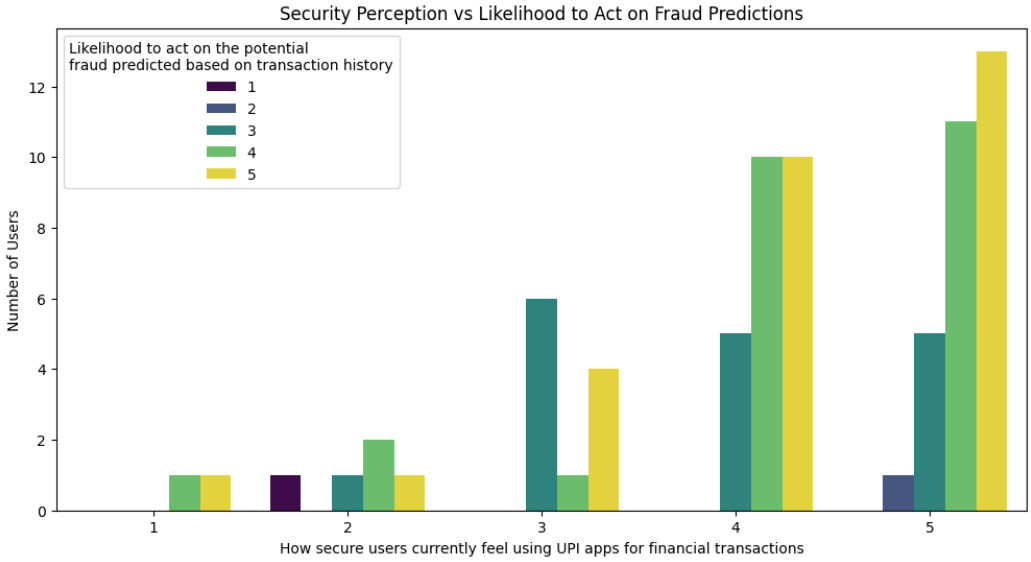


Figure 1

Source: Primary Data

Analysis: A significant portion of the respondents were worried about the security of the UPI app; noticeable regardless of the person feeling secure. However, most expressed that they would take preventative measures if, in the future, UPI apps were to forecast the risk of potential fraud based on historical transaction data. This implies that predictive analytics may reduce security worries; users may be more equipped to respond at the first tipoff of a threat.

4.2. Openness to Predictive Services and Data Usage for Machine Learning

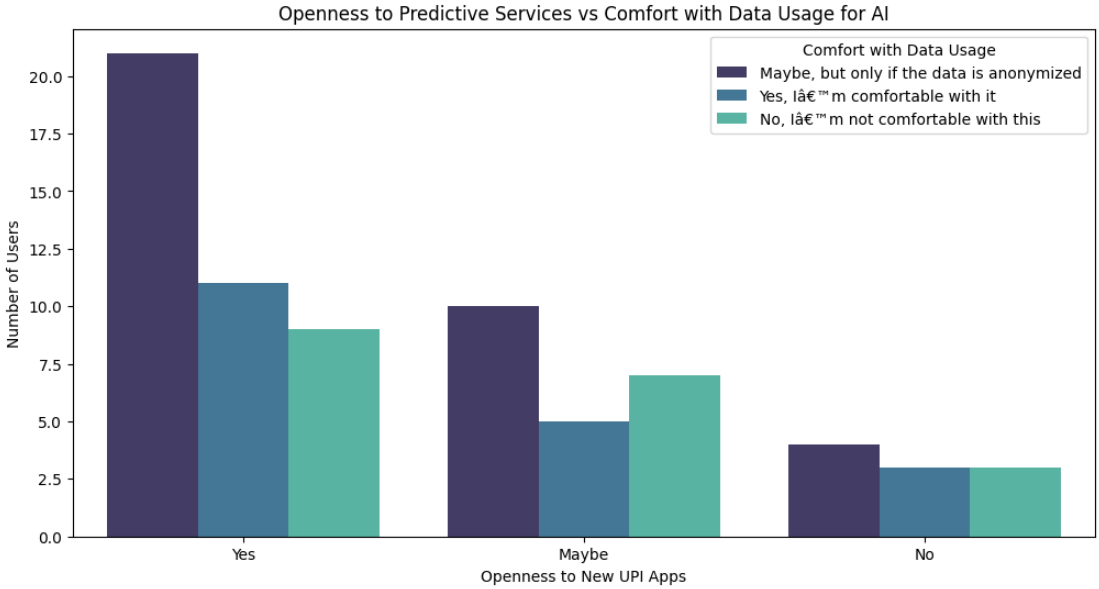


Figure 2

Source: Primary Data

Analysis: A majority of the participants were receptive to UPI apps offering predictive services like fraud prevention and budgeting recommendations. But there was some reluctance regarding UPI apps using personal transaction data to train AI models. Many users remain unsure of the predictive analytics benefits, as data privacy and the ethical use of personal information continue to be questioned. y suggested that if UPI apps could assess potential fraud risk using previous transaction history, they would probably take preventative measures. Predictive analytics, therefore, can address security concerns by giving users the power to navigate potential risks.

4.3. Utility of Personalized Transaction Insights

ISSN 2394 - 7780

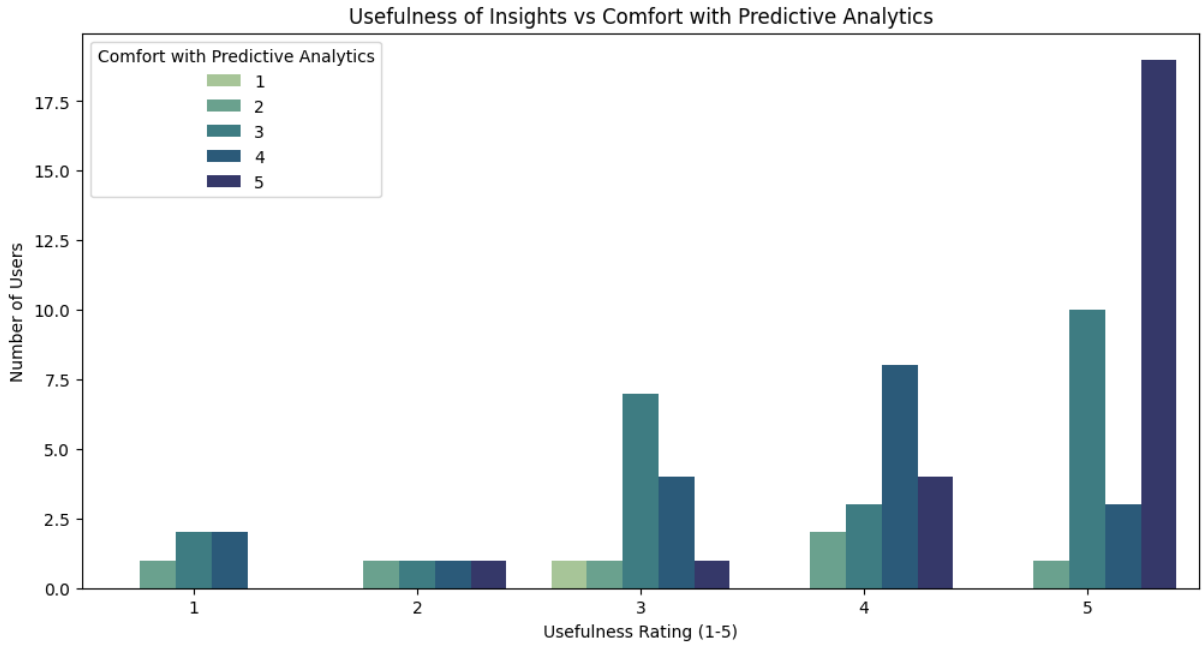


Figure 3

Source: Primary Data

Analysis: A large majority found personalized insights like monthly spending trends or budgeting advice useful. Not all users were comfortable with using past transaction data to predict future spending, but many were willing to make the trade-off if this led to a more useful insight.

4.4. Interest in Automated Spending Categorization and Notification Preferences

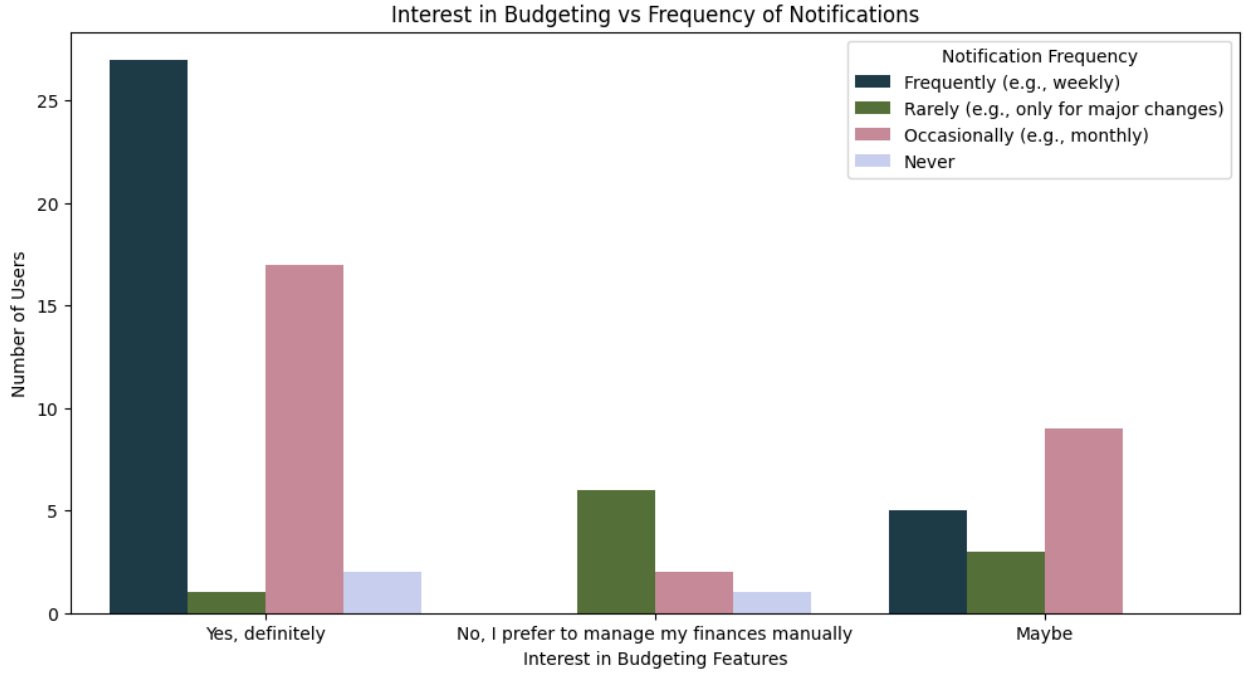


Figure 4

Source: Primary Data

Analysis: UPI apps that could help users categorize their spending automatically and budget also received strong interest. While some devices issued notifications, users tended to prefer notifications based on predictive analytics—warnings about overspending or one-off discounts tailored to them. This is a clear demand for proactive features that help manage finances wisely.

5. Ethical Considerations

ISSN 2394 - 7780

- All survey participation was voluntary, and responses were collected anonymously to protect the privacy of respondents.
- Respondents were explained the purpose of the survey and utilization of the data in the research.
- The study followed the ethical regulations for data collection, storage, and analysis, respecting the concerns of the data users regarding their privacy.

6. Limitations

- **Sample Size**: While the survey collected 73 responses, a larger sample would yield generalizable insights.
- **Self-Reporting Bias**: Because surveys rely on participants self-reporting their responses, a participant's perception and bias could inaccurately influence the data collected.

The mixed-methods approach employed in this study, incorporating both survey data and qualitative analysis, offers valuable insights into the potential of predictive analytics to improve the security and user experience for UPI applications. Results from this process will be used to build predictive features that will stick to user preferences and privacy concerns.

IV. CONCLUSION

In this research paper, we have discussed various aspects of usage of predictive analytics in UPI app for better security and user experience. UPI has gained tremendous popularity and enormous volume of digital transactions were on rise and it required better methods to ensure security and optimum user engagement. Through our literature review, we find out that machine learning techniques can enhance fraud detection and go beyond the traditional rule-based systems which can have limitations and allow for real-time detection of anomalous behaviour while reducing false positives. Additionally, personalized payment recommendations utilise transaction data in order to provide better user experience via hyper-customised financial services.

The results indicate considerable potential of predictive analytics to enhance the security of the Unified Payment Interface application and at the same time offer a more personalized user experience. From ensuring data privacy to obtaining user consent, the ethical issues need to be addressed effectively for AI-powered solutions in fintech to be deployed responsibly. The future of digital payment, though not just limited to the UPI domain, can be driven by integrating the AI and big data technologies with UPI apps; making them more secure and easy to use.

Exploring the possibility of emerging technologies such as blockchain and biometric authentication to facilitate and improve the security protocols in the end-to-end UPI process could be the next area of research. We will also watch the ethical frameworks for data usage and transparency closely, as they will be necessary to preserve user trust. The research will serve as the cornerstone of future development in leveraging predictive analytics for ensuring cybersecurity in such digital payment systems along with financial inclusion and user satisfaction.

V. REFRENCES

[1]. Vasan, S. ENHANCING REAL-TIME PAYMENT SYSTEMS: AI SOLUTIONS FOR SECURITY, EFFICIENCY, AND USER EXPERIENCE.

[2]. Dutta, M. M., & Poornima, K. DATA-DRIVEN PERSONALIZATION STRATEGIES: PROPELLING INDIA'S ECONOMIC GROWTH IN THE DIGITAL AGE. *INDIA'S $5 TRILLION ECONOMY: THE VISION, CHALLENGES, AND ROADMAP*, 154.

[3]. Kavitha, J., Indira, G., Anil Kumar, A., Shrinita, A., & Bappan, D. (2024). Fraud detection in UPI transactions using ML. *EPRA International Journal of Research and Development*, 9(4), 142-146.

[4]. Singh, S., & Devi, A. A Trend Analysis of Growth Pattern of Digital Modes of Payments in India.

[5]. Abba, S. (2022). AI in Fintech: Personalized Payment Recommendations for Enhanced User Engagement. *INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND INFORMATION TECHNOLOGY (IJRCAIT)*, *5*(1), 13-20.

[6]. EZPay: Enhancing UPI usability and security. (2024). *Journal of Payments Strategy & Systems*, 12(1), 58-70.

[7]. Afriyie, J. K., Tawiah, K., Pels, W. A., Addai-Henne, S., Dwamena, H. A., Owiredu, E. O., ... & Eshun, J. (2023). A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions. *Decision Analytics Journal*, *6*, 100163.

ISSN 2394 - 7780

[8]. Kalyani, S., & Gupta, N. (2023). Is artificial intelligence and machine learning changing the ways of banking: a systematic literature review and meta analysis. *Discover Artificial Intelligence*, *3*(1), 41

[9]. Ganapathyraman, S., Suresh, S., & Thomas, T. C. (2023). A Study on Users' Opinion Towards Unified Payment Interface (UPI) Transactions. *World Journal of Management and Economics, Forthcoming*.

[10].Dam, D. S., & George, B. Customer Perceptions About the United Payment Interfaces (Upi) Based Mobile Payment Apps in India. *Available at SSRN 4890320*.

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

PHISHING DETECTION AND PREVENTION USING MACHINE LEARNING ALGORITHMS

[1]Kavya Chouhan, [2]Ms. Rohana Deshpande and [3]Ms. Fatima Shaikh
[1]Bachelor of Science and [2, 3]Assistant Professor, Information Technology, Jai Hind College
Churchgate, Mumbai, India

ABSTRACT
The Internet has become an integral part of our lives. With the increased use of the internet, fraudsters have become more active and advanced. Phishing attacks are one of the most serious security risks on the Internet today. Due to technological advancements, detecting phishing attempts has become more challenging. Understanding if a web page is legitimate or not is a difficult task because of its semantics-based attack strategy, which primarily exploits computer users' vulnerabilities. The goal of this research is to create a machine learning model that can enhance accuracy and reduce false positives when detecting phishing emails and URLs.

There are several methods for determining whether a website is legitimate or not. Many software companies are launching anti-phishing technologies that use approaches such as signatures, blacklists, heuristics, and visuals. However, each has its advantages and disadvantages. Therefore, there is a need for real-time machine learning methods. Using algorithms such as random forests, xgboost and neural networks for the classification of phishing webpages, we can achieve our goal. To increase detection accuracy, we use advanced feature extraction techniques such as URL Length, Web Traffic, Page Rank, and HTTPS token. Data will be collected from publicly available Kaggle databases. Model performance will be assessed using accuracy, precision, and recall.

This study not only explores how we may use several classification models to improve the accuracy of detecting phishing websites but also addresses the practical implementation challenges associated with integrating these models into existing security infrastructures.

Keywords- *Phishing detection, machine learning, cybersecurity, data analysis, classification*

I. INTRODUCTION
One of the most serious problems with using the web is phishing attacks, which tricks users to provide sensitive information including login credentials, financial details, and personal data by impersonating them as legitimate websites [5]. This type of social engineering takes advantage of users' trust in a recognizable digital environment, which can have financial and data security risks affecting not only people but also companies and industries worldwide [1]. This is because traditional security solutions, such as blacklist-based filtering and signature-based detection alone, are not sufficient because in today's world, attackers continuously improve their tactics to compromise systems. These methods are unable to keep up with high-volume phishing attempts, and there is generally a delay in detection, which creates more risk [2].

The application of machine learning algorithms has demonstrated considerable success in overcoming these difficulties as it improves the detection accuracy and adaptability of phishing protection techniques. ML-based methods perform feature extraction from web pages, such as URLs, HTML content, and hyperlink structures of web pages, to classify websites based on learned patterns between legitimate or malicious sites. This is particularly beneficial for “zero-hour” detection, where a new phishing site can be indicated as fake because of its feature properties and does not need to appear in previous blacklists [4].

Machine learning models, such as decision trees and support vector machines, to more effective ensemble methods, such as Random Forest for checking the domain name Structure or Link Patterns or Embedded Scripts in phishing sites, are another frontier domain where machine learning has been able to display high classification accuracy. These advances highlight the importance of machine learning in phishing detection, not only to provide more reliable defenses, but also scalable solutions that can be applied at both the enterprise and individual levels.

ISSN 2394 - 7780

Fig. 1 Lifecycle of a Phishing Attack

II. CHALLENGES IN INTEGRATION

A. Existing Security Infrastructure

Protecting our digital realm from unauthorized access and criminal activity is not a new concept; firewalls, intrusion detection systems (IDS), and anti-virus software are time-tested measures. These technologies, however highly capable they may be in their own niches of defense, are by no means adaptable or fast enough to adapt quickly to new types of threats like zero-days and sophisticated phishing campaigns. However, including machine learning models inside such security systems can enhance their detection accuracy and also help in speed via pattern recognition and predictive analysis.

Today, machine learning poses as the next generation solution to improve traditional security systems by providing a range of benefits. Machine learning algorithms like Random Forest, Support Vector Machines (SVM) etc have been successful in enhancing IDS and can accurately categorize legitimate or malicious traffic with high accuracy. Moreover, machine learning models can learn with new data and enhance systems to function better in unknown threats detection than existing static rule-based firewalls & antivirus software [2].

Despite these advantages, blending machine learning with traditional security systems also brings complications, especially with outdated infrastructure. Most of the conventional IDS and firewall systems are rule-based which will make it difficult to fit in a machine learning model that is driven by data results. The resource constraints can also happen in the case of traditional systems with limited processing power and lack of real-time processing capabilities which are necessary for running machine learning algorithms. Traditional systems may also resist due to protocol incompatibility or security policies that are not compatible with newer, AI-driven capabilities [4].

As such, the integration of machine learning into existing security frameworks has great potential to improve defenses in general, but its utility will hinge on the cross-compatibility between themselves, as well as power-wise support and adaptability with legacy systems, so it is possible for these benefits without disturbing traditional norms in terms of securities.

B. Real-Time Application Issues

Deploying machine learning models for security in real-time comes with various difficulties, mainly surrounding latency and computational efficiency. Machine learning based phishing and intrusion detection systems rely upon rapid threat response times to detect the threats early, so as not impact users or systems. However, the higher latencies of complex models (such as deep learning architectures) make them unsuitable for real-time applications. When working with large data streams, even feature-rich models can be slow to compute; thus delaying detection response.

It is a fundamental trade-off between model complexity and detection speed among real-time cybersecurity applications. However, complex models such as Random Forest or deep neural networks, provide great accuracy in phishing detection but may be inadequate for high-speed situations without proper optimization due to processing overhead [3]. On the other hand, simpler models like Logistic regression or Decision Trees could

ISSN 2394 - 7780

provide slightly worse detection but it will be faster as compared to complex learning schemes and may fail to detect advanced phishing/intrusion patterns.

These trade-offs emphasize the need of how models should be selected and optimized according to requirements from real-time systems. Achieving a fine balance between accuracy and fast detection usually necessitates ensemble or hybrid methods that combine efficient techniques, complementing each other to retain high performance while also ensuring computational efficiency, which are key for effective real-time security solutions.

C. Limitations During Deployment

There are many difficulties associated with deploying machine learning models in cyber security, such as data quality or model drift. In this case, phishing detection and other security models require a list of high-quality labeled datasets to achieve an accurate evaluation. However, real-world data are often noisy or incomplete; therefore, model predictions are inaccurate and include false positives or negatives.

Approving usage is another obstacle, because customers and administrators might be hesitant to authorize fully automated machine learning-based systems. Complex models may lack interpretability and reliability, which leads to potential distrust of the system's outputs. In high-stake organizations, the problem worsens because false positives slow down operations and put security at risk by producing unfiltered data. This means that building trust in machine learning for cybersecurity involves increasing transparency and delivering users with concrete actionable behaviors [6].

In summary, efficient deployment requires careful attention to data quality and modeling updates as well as an overall approach to improve user trust in hands-off systems while also achieving the right balance between advanced threat detection functionality and system usability and reliability.

III. METHODOLOGY

A. Data Collection

The dataset used in this study was obtained from a publicly available Kaggle-phishing data collection that contains legitimate and phishing samples retrieved from active phishing databases [8]. Using open-access data such as this is crucial for training machine learning models because it captures a wide range of phishing attack patterns and characteristics observed in real-world situations. Current phishing data allow models to be trained on the most recent malicious practices and real user behaviors, which helps them percolate accurate results and be adept through live detection environments. Additionally, these datasets have pre-processed and normalized data that help the model generalize better across multiple types of phishing cases [2].

The dataset has 10,000 rows by 50 columns in which 5000 phishing webpages and 5000 legitimate webpages. The columns, or features, represent several attributes of the URLs: structural properties of resources and pages; behavioral measures capturing browsing habits or interaction data; and metadata that are already known about Web Resources from other repositories.

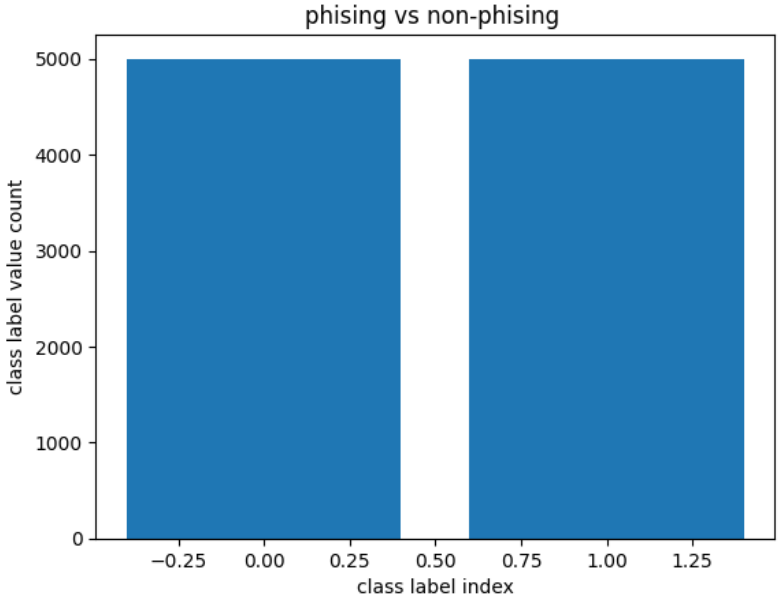


Fig. 2 Distribution of Phishing vs. Non-Phishing URLs

ISSN 2394 - 7780

B. Model Selection

In this study, an implementation was performed using three machine learning algorithms: Random Forest, XGboost and Neural Network models. Each method was chosen to use different components of machine learning for optimal detection accuracy and model robustness based on their performance, as reported in previous studies.

1) **Random Forest**: Random Forest is a powerful ensemble method that has shown promising performance and robustness, particularly when applied to phishing detection problems. This makes the model a popular choice for use in cybersecurity applications because it can deal with imbalanced datasets. Existing research also sheds light on the utility of Random Forest to identify phishing sites, exploiting characteristics in URL and web page features with a bagging wentropy within itself and radical characteristic randomness.
2) **XGBoost**: XGB is a popular choice for phishing detection because it has GPU support that boosts the efficiency and scalability of data training, making large-scale datasets processed quickly with high speed. This has been proven in the research, which made use of studies regarding phishing detection and how boosting algorithms like XGBoost are able to sequentially identify complex non-linear patterns due to the way it adjusts focus on areas where they were previously scored less [6]. This model is quite advantageous for our goals and it has a strong advantage against the simpler models when dealing with false negatives in phishing detection.
3) **Artificial Neural Network (ANN):** ANN is a fully connected network that shows an advantageous performance while capturing deep information between the data inputs; therefore, it performs well in intricate feature interactions. With phishing detection, neural networks have been exploited to recognize the complex dynamics in fishing strategies, and they can model a flexible propensity within feature learning [7]. Neural Networks: Although they are immensely powerful, NNs can be computationally intensive and may have issues with time lag, that is, latency, thereby rendering them a suitable candidate for real-time detection without proper hardware optimization [5].

Random Forest outperformed these models in terms of performance scores on the Kaggle phishing dataset, which is expected given previous works that have shown random forest to be one of the best trade-offs between accuracy and computational efficiency [8]. The combat model is a winner in terms of both toughness and instantaneous adaptability, which makes it ideal for applications such as cybersecurity, where fast identification with accuracy plays an important role.

C. Evaluation Metrics

To assess the effectiveness of phishing detection models after deployment, important evaluation measures such as **accuracy**, **precision**, and **recall** are utilized to determine how well each model makes decisions about whether a URL is a phishing or legitimate. These measures provide a comprehensive view of model reliability, notably for detecting phishing attacks.

1) **Accuracy**: This indicates the ratio of the total URLs tested, which has been correctly classified as both phishing and legitimate. For instance, in this study, the Random Forest model achieved an accuracy of 99%, XGBoost reached approximately 97%, and the Neural Network had an approximate score of 96%. High accuracy indicates that a model generally performs well, although the above evolution process may not correctly capture the performance of imbalanced datasets, especially where phishing instances are rare compared with legitimate ones.
2) **Precision**: Precision measures the accuracy of a classifier when creating phished or legit urls. The necessity for high precision is especially acute in phishing detection; low false positives are essential for avoiding interference with legitimate sites. This study has achieved approximately 99% precision from Random Classifier, 97 % from XGBoost and a Neural Network Model of approximately 95%.
3) **Recall**: Recall, also known as sensitivity, shows how well a model can find phishing URLs in all real-life instances. The recall scores in this research were approximately 99% for Random Forest, 97% for XGBoost and 96% for The Neural Network model.

Combined, these metrics enable an understanding of the true strengths in each model relative to a good tradeoff between accuracy and send sensitivity. (reducing overfitting).

ISSN 2394 - 7780

	Accuracy	Precision	Recall
Random Forest	0.99	0.99	0.99
XGBoost	0.97	0.97	0.96
Neural Network	0.96	0.95	0.96

Fig. 3 Model Evaluation Metrics

IV. CASE STUDY

Reference [9] provided a practical example of how machine learning models were effectively integrated into a browser plugin to detect phishing sites in real time [9]. The project involved creating PhishNet, a Google Chrome extension that detects phishing attempts on user-visited URLs using rules created by a random-forest model. PhishNet was trained on a set of URLs, and the model was developed with 14 essential parameters: domain discrepancies, IP address indicators, SSL presence, and URL length. With an accuracy of 98.35% and a True Positive Rate of 100%, the Random Forest model outperformed the other models.

When PhishNet was used in a browser, it seamlessly integrated into users' surfing experiences and immediately alerted when phishing-related website features were detected. This strategy has proven to be quite effective; end users are protected against phishing attacks without having to rely on third parties, which can create latency and security issues.

So here are some major takeaways from the case study: In a real-time, dynamic-user-facing context, machine learning models like Random Forest can be extremely effective if rules can be generated and processed as efficiently as feasible. However, it reveals some limitations with regard to dependency, because without features from the initial feature set, any phishing efforts may have a negative impact on the model's accuracy. Future applications should therefore improve feature diversity in order to respond gradually to emerging phishing strategies.

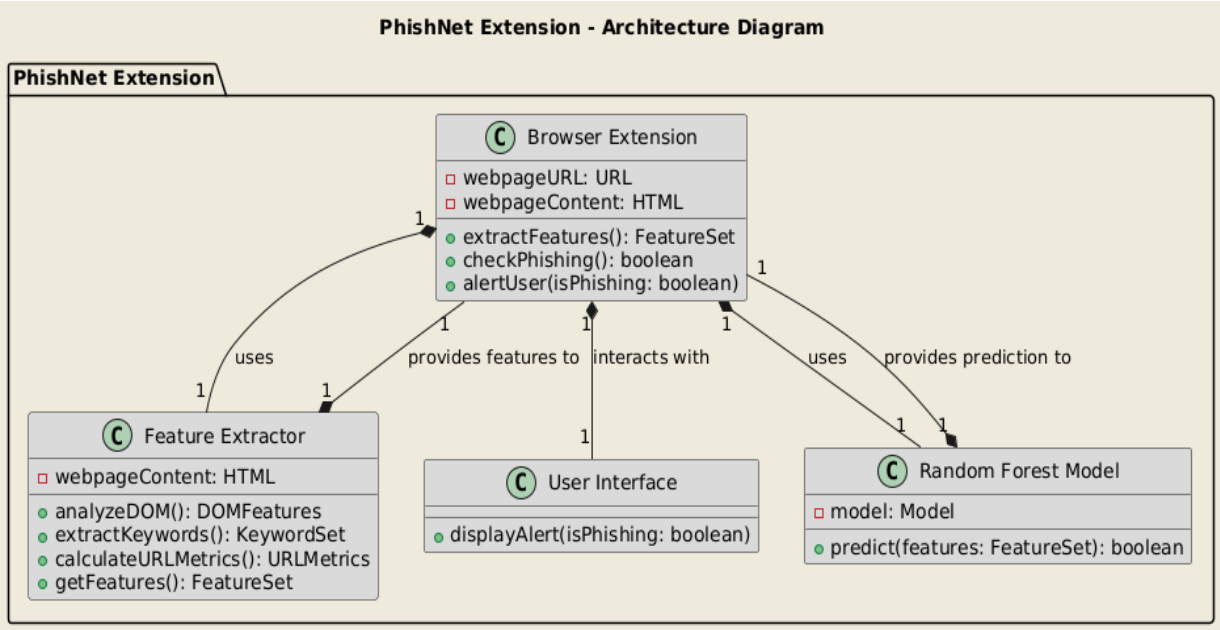


Fig. 4 Architecture Diagram of PhishNet

V. FUTURE DIRECTIONS

An adaptive learning model is the best option for handling the evolving nature of phishing attacks, and it can significantly boost the phishing detection in future. It can apply adaptive learning with mechanisms such as reinforcement learning to update and improve the detection criteria according to the latest types of phishing trends.

Additionally, hybrid model approaches that merge traditional machine learning techniques such as Random Forest with deep learning models will likely result in more robust and flexible detection systems. Hybrid models detect URL- and content-based phishing by combining the interpretability of classical models with the accuracy of deep learning.

Future research and implementation can focus on leveraging natural language processing (NLP) to improve detection of phishing attacks through email, employing more complex models that identify key aspects of the email content including the intent and sentiment of emails.

ISSN 2394 - 7780

As these technologies improve, the combination of adaptive, deep learning, and hybrid approaches will undoubtedly result in stronger and high-accuracy phishing detection frameworks that can identify and react to new types of attacks in real time.

VI. CONCLUSION

Machine learning is one such technique that can automate processes and has been proven to be effective over the traditional process. Random Forest always has the best trade-off when it comes to speed and accuracy as compared to all algorithms, so it is a suitable solution for real-time scenarios. More complicated models, such as the Neural Network, which was an effective tool for analysis, faced deployment issues because it requires significant computational capacity, particularly when operating in latency-sensitive situations.

Additionally, deploying these models into existing security infrastructures causes compatibility issues, particularly with older technologies not designed for machine-learning frameworks.

Further research may focus on applying reinforcement learning methods that enable models to adapt dynamically to changes in phishing technology. In addition, deep learning-based phishing detection also enables more content-based detection, e.g., on phishing emails rather than URL only. Making use of hybrid models that use classical algorithms alongside deep learning may prove more robust, interpretable, and adaptive to real-world cybersecurity applications.

REFERENCES

[1] Sahingoz, O. K., Buber, E., Demir, O., & Diri, B. (n.d.). *Machine learning based phishing detection from URLs*. Retrieved from https://www.sciencedirect.com/science/article/pii/S0957417418306067

[2] Gandotra, E., & Gupta, D. (n.d.). *An efficient approach for phishing detection using machine learning*. In *Smart Innovations in Communication and Computational Sciences* (pp. xxx–xxx). Springer. Retrieved from https://link.springer.com/chapter/10.1007/978-981-15-8711-5_12

[3] Jain, A. K., & Gupta, B. B. (n.d.). *A machine learning-based approach for phishing detection using hyperlink information*. *Journal of Ambient Intelligence and Humanized Computing*. Retrieved from https://link.springer.com/article/10.1007/s12652-018-0798-z

[4] Abdelhamid, N., Thabtah, F., & Abdel-jaber, H. (n.d.). *Phishing detection: A recent intelligent machine learning comparison based on models content and features*. Retrieved from https://ieeexplore.ieee.org/document/8004877

[5] Hossain, S., Sarma, D., & Chakma, R. J. (n.d.). *Machine learning-based phishing attack detection*. Retrieved from https://www.academia.edu/download/101989368/Paper_45-Machine_Learning_Based_Phishing_Attack.pdf

[6] Martínez Torres, J., Iglesias Comesaña, C., & García-Nieto, P. J. (n.d.). *Machine learning techniques applied to cybersecurity*. *International Journal of Machine Learning and Cybernetics*. Retrieved from https://link.springer.com/article/10.1007/s13042-018-00906-1

[7] Handa, A., Sharma, A., & Shukla, S. K. (n.d.). *Machine learning in cybersecurity: A review*. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. Retrieved from https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/widm.1306

[8] Tiwari, S. (n.d.). *Phishing Dataset for Machine Learning*. Kaggle. Retrieved from https://www.kaggle.com/datasets/shashwatwork/phishing-dataset-for-machine-learning

[9] Ojewumi, T. O., Ogunleye, G. O., Oguntunde, B. O., Folorunsho, O., Fashoto, S. G., & Ogbu, N. (n.d.). *Performance evaluation of machine learning tools for detection of phishing attacks on web pages. ICT Express.* Retrieved from https://www.sciencedirect.com/science/article/pii/S2468227622000746

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

CRIMINAL BEHAVIOR ANALYSIS PREDICTION MODEL: USING SOCIAL MEDIA

[1]Khushboo Rajesh Gupta and [2]Tejashree Parab
Department of MSc Big Data Analytics, Jai Hind College (Empowered Autonomous), Mumbai, India

ABSTRACT
Social networking has dramatically transformed how people present themselves online, creating a massive repository of data that provides deep insights into personality traits and behavioral tendencies. The focus of this research paper is to evolve the criminal behavior prediction model by utilizing digital footprints on social media platforms such as Instagram to predict potential criminal behavior in real life. By analyzing the user-generated content like posts, comments, and interactions, this model is expected to detect trends, behavioral indications, and risk factors related to criminal behavior. This research paper looks at the analysis of criminal behavior and the patterns created on social media focus is on finding AI, forensic linguistics, and cybersecurity methods for detection and prevention. On the subject of online harassment, hate speech, and cybercrime, studies help to support the role of AI in criminal justice. This study illustrates how criminal elements both have a platform for nefarious activities and serve as a source for developing better means of digital law enforcement, providing insight into more efficient analyses of criminal behavior by AI.

Keywords: *Criminal Behavior, Social Media, Hate Speech, Online Harassment, AI in Criminal Justice, Cybersecurity, Forensic Linguistics, Digital Investigations*

I. INTRODUCTION
Social media allows for online interaction through platforms like Facebook, Twitter, TikTok, and Instagram, which have drastically changed how people interact, organize, and communicate. While these social media channels present immense benefits, they are accompanied by new problems such as cyberbullying, hate speech, and crimes committed over the internet, and therefore become a point of focus for criminal behavior analysis research in these arenas. The potentials encompass the use of artificial intelligence, machine learning algorithms, and cybersecurity techniques to discover harmful behaviors which can be used to predict and improve criminal acts while helping the process of criminal investigation.

This paper will discuss the relationship between criminal behavior and social media, as well as explore how AI technologies can be used to identify and predict such activities. This study will further reinforce AI tools in the recognition of hate speech, online harassment, and other forms of digital malfeasance, thereby contributing positively towards effective law enforcement in the digital era.

II. REVIEW OF LITERATURE

Forensic Linguistic Analysis of Hate Speech on Social Media
One of the key spheres of criminal behavior on social media is the eruption of hate speech. Available at https://biarjournal.com/index.php/biolae/article/view/894, the titled "Forensic Linguistic Analysis of Netizens' Hate Speech Acts in TikTok Comment Section" investigates the real amount of hate speech on TikTok, making linguistic analysis of features that characterize online hate speech. This work will attempt to apply forensic linguistic techniques towards distinguishing between the ways hate speech manifests in online spaces and gain insights into the possible training of AI toward this end.

Forensic linguistics provides an invaluable methodology to differentiate the type of dangerous content from each other while allowing AI to distinguish them better in terms of labeling and response, making such hate speech detection in digital platforms even more effective when done through the means of AI.

Online harassment and cyber victimization of social media influencers

The paper "Too Lucky to Be a Victim? An Exploratory Study of Online Harassment and Hate Messages Faced by Social Media Influencers"
Addresses specific issues that social media influencers face, who are increasingly becoming targets of cyberbullying, defamation, and other forms of online threats. More than 70% of influencers confessed receiving hate messages and threats, which tells us that it is a widespread problem leading to mental and physical harm.

These threats can be mitigated with the help of AI-driven systems by analyzing patterns of online harassment and providing real-time alerts to influencers and their teams. Using AI to foretell potential escalation of threats and recommend preventive measures in terms of content filtering and enhanced privacy settings thus maintains a relatively safer environment online for the influencer.

ISSN 2394 - 7780

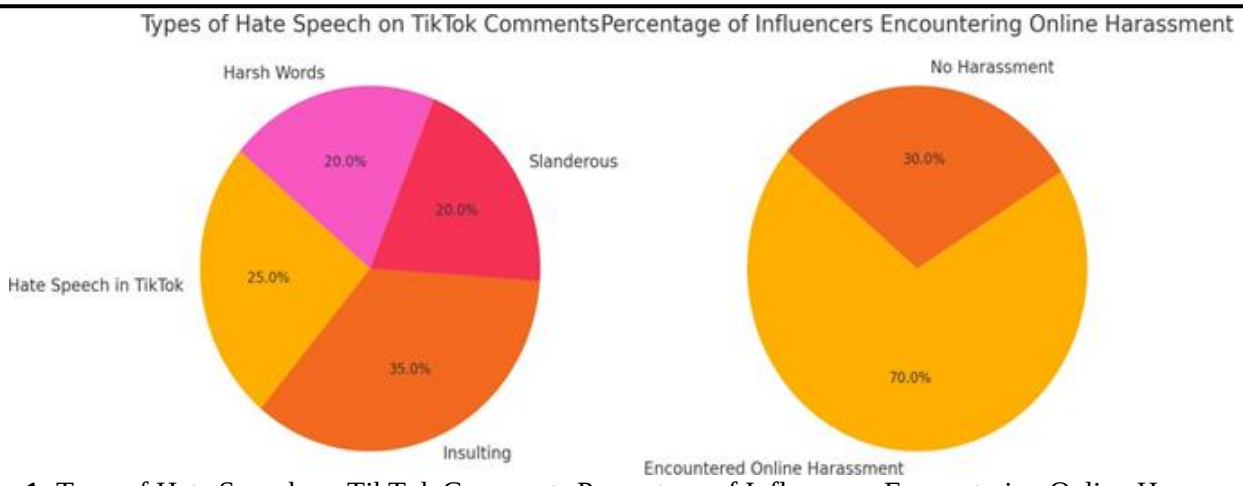


Fig. 1: Type of Hate Speech on TikTok Comments Percentage of Influencers Encountering Online Harassment

DIGITAL CRIMINAL INVESTIGATIONS AND AI

In "Digital Criminal Investigations in the Era of Artificial Intelligence: A Comprehensive Overview" there is an opportunity of AI enhancing digital criminal investigations. This is because AI would be able to analyze larger datasets and allow law enforcement agencies to identify patterns, track offenders, and detect criminal behavior at its onset. Machine learning algorithms are increasingly being used to identify fraudulent activities, predict recidivism, and get into details of digital evidence more efficiently.

This paper uses various case studies of AI tools deployed in criminal investigations, such as finding cybercrime syndicates and tracking illegal activities on the dark web. This makes it possible for law enforcement agencies to make decisions quickly and accurately, leading to a more proactive approach to digital crime.

CYBER SECURITY FOR CHILDREN IN SOCIAL MEDIA

The article "Cybersecurity for Children: An Investigation into the Application of Social Media" https://www.tandfonline.com/doi/epdf/10.1080/17517575.2023.2188122?needAccess=true involves the dangers young people face in social media and how they expand with children's log-in activities to the internet. Cyberbullying, sexting, and identity theft increase with increasing children online. AI can be applied in the observation of children's activities online, tracking dangerous signals, so sending alerts to parents or guardians.

AI-based cybersecurity technologies can also save kids from the clutches of net predators by making the online world safe. The use of machine learning algorithms will be involved in identification of inappropriate content, marking probable threats, and making sure that the social media sites are safe for children.

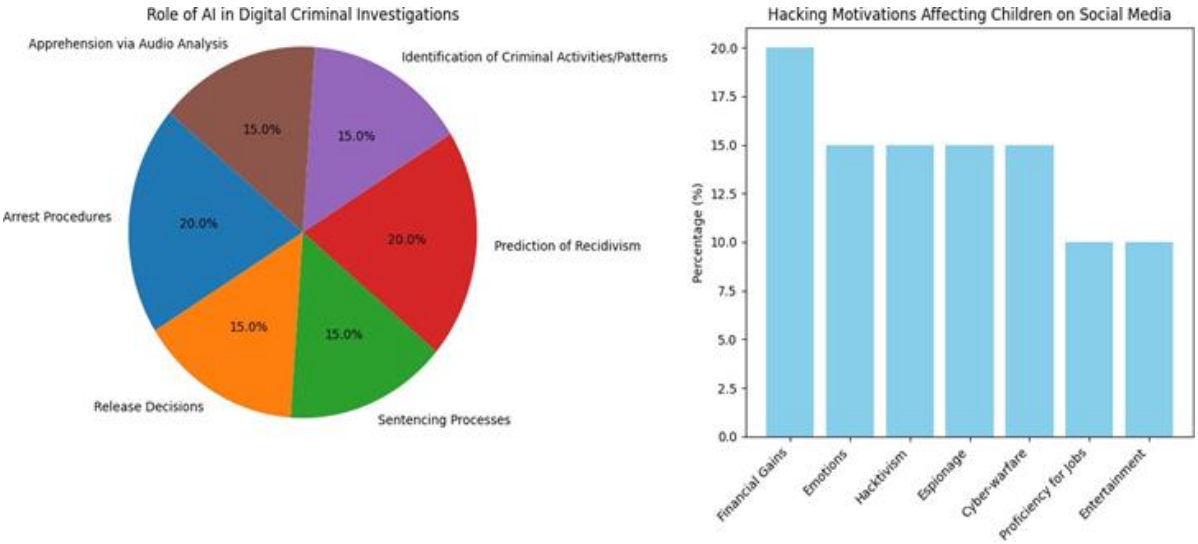


Fig.2: Role of AI in Digital Criminal Investigations

III. RESEARCH METHODOLOGY

This study has combined qualitative and quantitative approaches to shed light on criminal behavior on social media and the use of AI in identification and intervention of such activities. The components of this methodology are as follows:

ISSN 2394 - 7780

Literature Review Through a comprehensive literature review on contemporary works focused on hate speech, online harassment, cybercrime, and AI-driven analysis of criminal behavior, the authors set the premise and help in identifying gaps that have emerged in the literature.

Case Studies: Specific examples of online harassment, hate speech, and cybercrimes are examined so that the scope and implications of these behaviors can be understood. The case studies highlight how criminal activity finds a medium through social media platforms and how AI tools might fight these issues.

AI and Machine Learning Models: In the study, AI models and machine learning algorithms are used to detect and prevent criminal behavior on social media through the mediums of content moderation tools, predictive analytics, and real-time monitoring systems.

Cybersecurity Analysis: AI-powered cybersecurity tools that can be integrated into social media to use them to provide safety for the users by engaging on vulnerable platforms, such as children, social media influencers, among others.

IV. DATA MINING TECHNIQUES

1. Sentiment Analysis

Purpose: Identify the sentiment and tone of user-generated content, including posts, comments, and messages, to recognize hate speech, harassment, or criminal intent.

Methodology:
Natural Language Processing tools.

Algorithms: Naïve Bayes, Support Vector Machines (SVM), or transformer models like BERT.

Tools: Python libraries, including NLTK, TextBlob, VADER, or HuggingFace.

2. Text Mining

Purpose: Uncover valuable patterns from vast text-based data sources, such as social media comments, messages, and posts.

Topic Modeling: Find thematic trends by using Latent Dirichlet Allocation (LDA).

Identify keywords and phrases that are indicative of criminal behavior by using Term Frequency-Inverse Document Frequency (TF-IDF).

Group similar content using clustering techniques like K-Means, DBSCAN.

3. Predictive Modeling

Objective: Based on user activity patterns, predict the likelihood of criminal behavior. Methodology:

Train models on historical data containing known cases of online criminal behavior.

Algorithms: Logistic Regression, Decision Trees, Random Forest, or Gradient Boosting (e.g., XGBoost, LightGBM).

Use features such as frequency of posting, language utilized, or type of engagement.

4. Classification

Objective: Classify a piece of content into predefined categories, such as hate speech, online harassment, or neutral.

Methodology: Supervised learning classifiers, such as Naïve Bayes, SVM, or deep learning techniques (Convolutional Neural Network, Recurrent Neural Networks).

Tools: Tensor Flow, PyTorch, and Scikit-learn libraries for developing classifiers.

5. Social Network Analysis (SNA)

Objective: Explore user-user relationships and connections to identify possible criminal organizations or key influencers who perpetuate malicious activities.

Approach: Graph-based algorithms in modeling social media users as nodes and connections as edges Centrality measures: Degree, Betweenness, and Clustering Coefficients

Use: NetworkX (python) or Gephi for visualization

ISSN 2394 - 7780

6. Association Rule Mining

Objective: Discover behavior patterns that are likely to occur together at the same time, including specific types of comments often appearing before online harassment occurrences.

Approach: Use Apriori or FP-Growth algorithms to extract rules like: "If a user posts X, they are likely to post Y within Z days."

Applications: Detecting early signs of escalating online behavior.

7. Clustering

Purpose: Group similar users or content for further analysis, such as identifying clusters of users engaging in harmful behavior.

Methodology: Use algorithms like K-Means, Hierarchical Clustering, or DBSCAN.

Example: Grouping users based on posting frequency, sentiment, or content topics.

8. Anomaly Detection

Purpose: Uncover unusual patterns that might signify illegal activities, such as abrupt surges in aggressive speech or odd posting behavior.

Methodology: Isolation Forests, One-Class SVM, or Auto encoders (deep learning). Tools: Scikit-learn, PyOD (Python Outlier Detection).

9. Web Scraping and Data Collection

Purpose: Collect large datasets from social media platforms for analysis. Methodology:

Use tools like Beautiful Soup, Selenium, or Scrapy to scrape data.

Make sure to comply with the ethical and legal guidelines that may be involved, for example, data anonymization and permissions.

10. Time Series Analysis

Purpose: Analyzing time-dependent patterns in the user activity, like certain events increase hate speech and harassment.

Methodology: Sequence prediction through algorithms such as ARIMA or Long Short-Term Memory (LSTM) networks.

Application: Patterns that change with time, for example, increased online threats during particular periods.

11. Feature Engineering

Purpose: Meaningful feature creation to improve the performance of the model. Methodology:

Extract features including:

Linguistic patterns- word choice and sentiment scores. Interaction metrics likes, shares, and retweets.

Network metrics (e.g., centrality, clustering).

Tools: Python libraries such as Pandas, NumPy, and Featuretools.

12.Ethical Considerations in Data Mining

Objective: Address privacy issues and minimize biases in the dataset or algorithms. Methodology:

Use differential privacy techniques.

Apply bias detection algorithms to ensure fairness in predictions. Tools: IBM AI Fairness 360, Google Tensor Flow Privacy.

V. OBJECTIVES OF STUDY

The main objectives of this research are:

- Exploring how such criminal behaviors as hate speech and online harassment emerge on social media.
- Detection, prediction, and prevention of criminality in cyberspace with the help of AI.
- Assessing whether the implementation of AI-based tools enhances the investigation of crime and strengthens cybersecurity for users, the minors and influencers alike.
- To develop a comprehensive model for digital criminal investigations, using AI technologies to improve the performance of law enforcement.

ISSN 2394 - 7780

VI. SCOPE AND LIMITATIONS

The scope of this research considers criminal behavior through social media and the role of AI in improving digital criminal investigations. However, despite the possibility of AI greatly making the processes more efficient and accurate, the following limitations apply:

Ethical Issues: The concern over privacy is still to be conquered because the monitoring of a person's online activity with AI has difficulties balancing security and privacy.

Algorithmic Bias: AI models are prone to bias, depending on datasets used, leading to false positives or over-identification of individuals.

Availability of Data: This paper suffers from the restriction of proprietary data from social media. Access to this data can even improve deeper understanding of online criminal behavior.

VII. CONCLUSION

The research demonstrates that the importance of AI in the analysis of criminal behavior on the social network is growing. With the above view, this paper scrutinizes contemporary studies on hate speech, online harassment, and digital criminal investigations with a view of improving potential law enforcement against such crimes and protecting vulnerable people online through AI. AI-driven systems can identify patterns, predict criminal behavior, and improve cybersecurity for social media users. However, there are also limitations and ethical considerations, and therefore, this level of AI induction into criminal investigations takes one step further in how people can combat online crime and ultimately create a safer digital environment.

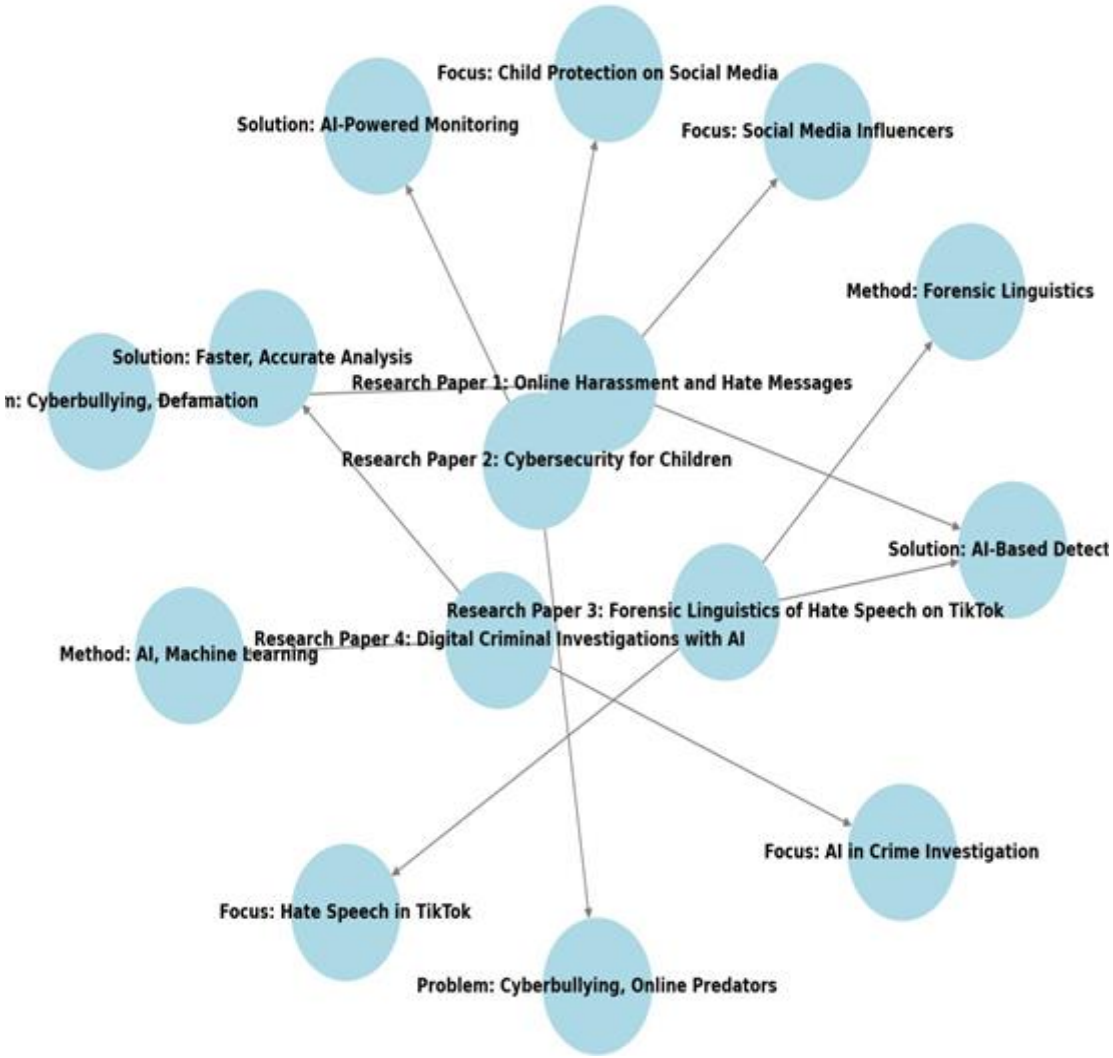


Fig.3: Graph Structure of Research Paper Summary

REFERENCES

1] Martens, M., & Neudert, L.-M. (2023). Too lucky to be a victim? An exploratory study of online harassment and hate messages faced by social media influencers. *European Journal on Criminal Policy and Research.* https://doi.org/10.1007/s10610-023-09542-0

ISSN 2394 - 7780

2] Alabdulkareem, K. F., & Alsulami, H. (2023). Cybersecurity for children: An investigation into the application of social media. *Enterprise Information Systems*, *17*(4), 679–696. https://doi.org/10.1080/17517575.2023.2188122

3] Nugroho, R. A., Rukmini, D., & Sutopo, D. (2023). Forensic linguistic analysis of netizens' hate speech acts in Tik-Tok comment section. *BIAR Journal: Bioinformatics and Applied Research*, *4*(2), 88–99. https://biarjournal.com/index.php/biolae/article/view/894

4] Dunsin, D., Ghanem, M. C., Ouazzane, K., & Vassilev, V. (2023). A comprehensive analysis of the role of artificial intelligence and machine learning in modern digital forensics and incident response. *arXiv*. https://arxiv.org/abs/2309.07064

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

CUSTOMER SEGMENTATION IN THE TELECOM INDUSTRY USING MACHINE LEARNING

[1]Manasi Bait and [2]Tejashree Parab
Department of MSc Big Data Analytics, Jai Hind College (Empowered Autonomous), Mumbai, India

ABSTRACT
Customer churn is the talk of the day for the telecom sector since it generates huge volumes of data and rapid advancement in techniques associated with data mining. Retaining the customers is obviously much cheaper than acquisition task. This can be achieved by understanding the reasons why the customers are churning, which can happen due to data-driven insights. This paper reviews the commonly used data mining strategies for churn pattern identification and discusses predictive modelling techniques. It also highlights future research opportunities in this domain.

Keywords: *Customer churn, Customer retention, CRM, Data mining techniques, Telecom industry.*

I. INTRODUCTION
In today's highly competitive telecom landscape, understanding and anticipating customer behavior is a cornerstone of success. Companies constantly seek innovative strategies to tailor the services and enhance customer experiences. Customer segmentation which categorizes user into distinct groups based on shared characteristics, is an essential tool for achieving these goals. However, in the way of traditional demographic or geographic segmentation, it is difficult to identify the complexity of modern customer behavior by the dynamism of multifaceted factors.

This kind of machine learning transforms the problem by analyzing enormous amounts of data to reveal very subtle patterns and relationships that more conventional methods often miss. In this study, advanced algorithms will be used in order to make more refined segmentation strategies possible so that telecom companies can provide the most personalized services and targeted marketing campaigns. It explores these methods in a manner that contributes to growing knowledge on customer relationship management and strategic decision making in telecom sector.

Problem Statement
Telecom companies face the challenge of segmenting an increasingly dynamic and diverse customer base. Traditional approaches to segmentation, based primarily on static attributes like demographics and basic usage statistics, cannot quickly respond to changing customer preferences and behaviours. These deficiencies not only decrease the effectiveness of the marketing effort but also limit the extent to which personalized services can be built, which is essential for keeping existing customers in a competitive market.

This research addresses the gaps by using machine learning techniques to build a more nuanced and dynamic segmentation model. This will allow telecom companies to identify key customer segments, optimize marketing strategies, and improve customer satisfaction. The study focuses on actionable insights derived from a real word data, enhancing the decision making capabilities of telecom providers and strengthening their ability to navigate a rapidly changing business environment.

Research Objective
a. Utilize prevailing machine learning models for precise telecom customer segmentation.

b. Evaluate the performance of these models in real-world settings.

c. Identify the key customer segments to be targeted.

d. Assess the effect on customer satisfaction and retention.

II. LITERATURE REVIEW
The telecom sector has always used demographic, geographic, and behavioural data for customer segmentation. Demographic segmentation focused on differentiating customers based on attributes such as age, income, gender, and education and provided a simple approach for understanding Consumer behaviours (Smith,2019). Geographical segmentation is based on the physical location of the customers; services are targeted to fulfil the needs of specific regions (Lee,2020). Such methods, however, tend to overlook the complex and dynamic needs of the individual consumer.

ISSN 2394 - 7780

This approach attempts to bridge the gap by evaluating a customer's activity, such as service usage, purchase history, and interaction patterns. For example, Johnson (2021) showed that it is possible to determine different customer distinct groups by analysing call detail records based on certain usage patterns. However, the most resource-intensive approach and often overlapping actual deeper drivers influencing behaviour, such as lifestyle change, or in external market trends, remains a challenge.

The limitation can be overcome by using machine learning techniques such as K-Means clustering and Support Vector Machines(SVM). K-Means effectively groups customers according to characteristics, whereas SVM can manage complex, non-linear patterns in data. These models are evaluated by metrics like accuracy, recall and confusion matrix, thus yielding reliable and actionable results.

By adopting advanced machine learning models, telecom companies can enhance segmentation accuracy, optimize marketing strategies, and better engage customers, fostering long-term loyalty and growth.

III. METHODOLOGY

Data Collection: The dataset used in the study was sourced from kaggle containing dependents, tenure, churn, monthly charges, service usage details essential for the churn analysis.

Data Preprocessing: There are no missing values in the data but there are categorical variables that need to be converted into numerical variables and this is achieved using Standard Scaler, Label Encoder for proper model learning.

Model Selection and Implementation: We chose a set of models so that their different strengths could be exploited for customer segmentation and churn prediction. We used K-means Clustering to identify unique groups of customers based on usage patterns, thereby providing actionable insights for targeted marketing strategies. We used SVM as it is the best model in handling high-dimensional data and provides accurate predictions for customer churn. Random Forest was also included, as it's powerful and can adapt to data variability. It will provide high accuracy and, more importantly, insights into how important a feature is. Lastly, Logistic Regression has been chosen for its simplicity and interpretability. It will allow for the easier understanding of interrelation between features and churn and to be able to clearly have a framework for decision-making. This combination ensures full investigation and robust comparisons for optimizing customer retention strategies and making data-driven decisions.

Evaluation Metrics: To measure the performance of our models, we used a range of important metrics. The confusion matrix provided us with a clear breakdown of our model's predictions, i.e., the true positives, true negatives, false positives, false negatives. The matrix is important for the identification of the type of errors our models made. Accuracy was used to measure the overall accuracy of our predictions, i.e., the ratio of instances correctly predicted to the total instances. However, as accuracy on it's can be misleading in the case of imbalanced data, we used recall as well, i.e., the ratio of actual positives (churners) correctly predicted by our models. This is particularly important in churn prediction, as it enables us to estimate the extent to which our models are classifying customers who are likely to attrite. With the use of these metrics, we are to get full picture of our performance of our models, enabling a robust comparative analysis to determine the most effective method for customer churn prediction.

IV. RESULTS

After analysis, the Support Vector Machine (SVM) model was top performer with the accuracy of 79.4% and a recall of 70.6%, indicating its high ability to identify customers who are likely to churn. Random Forest Model lagged behind with an accuracy of 76.5% and a recall of 67.3%, offered stability and valuable information on feature importance. Logistic Regression was just below the SVM level with the accuracy of 79.2% and a recall of 70.6% and was ranked high for its simplicity and interpretability. The K-Means Clustering successfully segmented customers into distinct group based on usage patterns, allowing for targeted marketing campaigns. Overall, while SVM had the highest recall, the ensemble of these models offered a combined understanding of customer churn behavior, balancing predictive performance and actionable information.

V. LIMITATIONS AND FUTURE SCOPE

While the analysis is helpful, there are some limitations, e.g., training on a specific dataset, which might limit generalizability. The feature engineering process can be optimized, and the models require high computational power.

Fine-tuning hyper parameters, along with other features like customer interaction data, and using ensemble methods for improved predictions can be the scope of future work. Real-time analysis and testing models on

ISSN 2394 - 7780

other industries will also be helpful. Adding customer feedback will enhance the models. Working on these areas will enhance the practicability and usability of the research.

VI. CONCLUSION

In summary, this study is concerned with the ability of machine learning techniques, i.e., Support Vector Machines (SVM) and K-means Clustering, to enhance customer segmentation in the telecommunication industry. With these advanced algorithms, telecommunication operators can achieve more precise segmentation, leading to more precise marketing and greater customer satisfaction. Despite limitations such as data availability and complexity, the findings offer valuable insights into the application of machine learning models to further enhance customer segmentation processes. These advances will result in more dynamic and personalized customer interaction, ultimately building higher customer loyalty and retention.

REFERENCES

1] Johnson, T. (2021). *Machine learning algorithms for customer segmentation in the telecom industry*. Journal of Telecommunications Research, 45(3), 245-260.

2] Kumar, S., & Gupta, P. (2022). *Application of neural networks in telecom customer segmentation*. International Journal of Data Science, 10(4), 320-335.

3] Lee, J. (2020). *Behavioral segmentation in the telecom industry: Insights from usage patterns*. Telecommunications Journal, 38(2), 180-195.

4] Smith, A. (2019). *Demographic-based customer segmentation in the telecom sector*. Journal of Marketing Analytics, 22(1), 67-78.

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

STUDY OF SUPPLY CHAIN RISK MANAGEMENT USING DATA ANALYTICS AND MACHINE LEARNING ALGORITHMS

[1]Musab Shaikh and [2]Shraddhadevi Singh
[1]Department of M.Sc. Big Data Analytics, Jai Hind College (Autonomous), Mumbai
[2]Assistant Professor, Department of BVocSD, Jai Hind College (Autonomous), Mumbai

ABSTRACT
As global supply chains continue to grow in complexity, risk management effectively has become the core activity. This paper explores the utilization of data analytics and machine learning.in the development of supply chain risk-management by it permits anticipating and mitigating disruption. Traditional SCRM strategies remain largely unresponsive to modern needs. In today's fluid environment. By reviewing the methodologies these models are logistic regression, support vector machines, and Anomaly detection, this paper will outline how ML can improve resilience in real time by detecting patterns as well as anomalies. It points out challenges in implementing Data-driven SCRM, which encompasses data quality and integration. Major issues and future directions use analytics on Agile and secured supply chain. Data-driven models give organizations valuable insights through resilience enhancement by identifying patterns and anomalies in real time, thereby allowing for earlier interventions. This research also addresses the challenges of data science within supply chain contexts, including issues of data quality, integration, and the adaptability of predictive models to changing risks. It synthesizes existing literature to offer recommendations for future research avenues. This paper will outline the transformative capability of data analytics and machine learning for supply chains. It will depict a shift from reactive to proactive risk management. Such a shift is not only good for enhancing the resilience of supply chains toward risks but also gives an organization the agility required in today's dynamic business environment.

Keywords— *Supply Chain Risk Management, Data Analytics, Machine Learning, Predictive Analytics, Supply Chain Security, Anomaly Detection, Resilience*

I. INTRODUCTION
The global nature of modern supply chains has introduced new levels of complexity and interdependence, linking suppliers, manufacturers, logistics providers, and consumers across various regions and markets. While these expansive networks offer operational efficiencies and reach, they also expose businesses to a broad spectrum of risks that can disrupt the continuity of goods and services. In recent years, natural disasters, economic fluctuations, cybersecurity threats, and unanticipated events like the COVID-19 pandemic have underscored the vulnerability of these interconnected systems.

A single disruption in one part of the supply chain can now have cascading effects that impact entire industries and economies, making robust supply chain risk management (SCRM) more crucial than ever.

Traditional approaches to SCRM, which often relied on reactive measures or post-event analyses, are proving insufficient for today’s volatile environment. Instead, a shift toward proactive, data-driven risk management strategies is emerging. This evolution has been largely driven by advancements in data analytics and machine learning (ML), which offer new ways to predict and prevent disruptions. By analysing historical and real-time data, these technologies allow companies to forecast demand, monitor risks, and optimize their supply chain strategies with unprecedented precision. [3][1]

A. Scope and Objectives of Study
This paper aims at exploring applicability of data analytics and machine learning for an enhanced management of supply chain risk. This will be driven from a critical analysis of the existing practice and future innovations in three main areas:

1. **Review of Current SCRM Methods:** This is to understand the adequacy of traditional risk management methods and areas where data analytics and machine learning can bring in positive change.
2. Machine Learning Applications investigate particular techniques of machine learning: logistic regression, support vector machines, and anomaly detection towards their applicability and effectiveness for the task of risk forecasting in supply chains.
3. **Opportunities and challenges:** Such deployment of the bottom-line implication to include the increased strength of efficiency as well as data quality and integration in a company using such technologies within the scope of its supply chain management.

ISSN 2394 - 7780

II. LITERATURE REVIEW

A. Review of Traditional Security Measures in Supply Chains

1) Traditionally, supply chain security focused on the implementation of measures that mitigated the disruptions through strong physical security, high compliance protocols, and manual oversight of suppliers and logistics partners. The mostly used strategies included inventory buffering, supplier controls based on contractual agreements, and multi-sourcing methods that ensured the smooth running of the supply chains even in cases where one supplier was subject to disruption. Putting this aside, risk assessment and audits on a manual basis ensure that the credibility of suppliers is well evaluated and probable vulnerabilities in systems are detected. The problem, however, is that most of the traditional approaches lack adaptability and predictive capacities that might be required in modern systems, hence leaving firms nearly in a reactive position.

Traditionally, deterministic models have dominated the use of risk management. They are often overly structured and based on historical information, whereas in such models, real- time input does not occur. Therefore, they fail to identify potential new risks and cannot respond quickly enough to shifts in demand or supply interruptions. In the meanwhile, increasing complexity in interdependencies within global supply chains has eroded the ability of these strategies to address the intricacies in the modern marketplace. Current trends are inadequate to handle the threats that arise in the horizon, particularly those identified and facilitated by cyber vulnerabilities. [5][1][7]

B. Evolution of Cyber Threats and Vulnerabilities in Global Supply Chains

The increased digitization of supply chains and dependency on technology have amplified cyber threats, which are considered one of the most critical risk factors in modern supply chain management. Cyber threats normally target basic data, systems, and infrastructures and usually cause operational disruptions by exploiting weaknesses existing within supplier information technology networks and their links. More and more firms embrace the latest technologies like cloud computing and IoT devices, along with automation in the process, which expose their organizations to significant cyber risks including data breaches, ransomware attacks, and malware infections that threaten sensitive information but also put the continuity of operations at stake besides finance and reputation.

The interlinked characteristics of global supply chains compound these weaknesses. A single compromised supplier or third-party supplier may compromise the security of an entire supply chain. For example, attacks that exploit unsecured links between organizations and their suppliers can allow attackers to gain unauthorized access to systems and data. This culminated in the design of a "cyber supply chain," whereby not only protection of physical resources but protection of the cyber communications and transfers happening through the networks needed to be involved in such cybersecurity plans. This also requires changes in the forms of security applied from conventional to more techno- savvy kinds of operations that rely less on instincts and impulse and more on technical processes.. [2]

C. Role of Predictive Analytics and Machine Learning in Modern Supply Chain Risk Management

Predictive analytics and machine learning have become critical tools in the evolution of supply-chain risk management, along a path from a reactive to a proactive approach. Predictive analytics combines historical and real- time data to detect patterns and trends; thus, organizations can anticipate what may go wrong and prevent such disruptions. The analysis of large, both structured and unstructured, datasets enhances the capability of machine learning to extract insights beyond what is possible through traditional statistical methods. There are various algorithms, including neural networks, random forests, and support vector machines, that aid supply chains in detecting anomalous cases, predicting demand variability, and assessing supplier reliability. Thus, with the integration of machine learning algorithms into supply chain management systems, it is possible to learn continuously from real-time data inflows, adapt to shifting conditions, and improve with respect to predictive accuracies over time. This feature supports prompt risk assessment and expedites the procedures for decisions, equipping organizations to respond rapidly to emerging threats. For instance, fraud-detection algorithms that, designed with such an objective have the ability to identify unusual patterns in transactional data sources and sensor output streams, which may indicate fraudulent activity, quality issues, or cyber intrusions. These predictive models help the organization improve their inventory management and supply chain logistics to reduce lead times and make such systems responsive to market conditions.

III. DATA ANALYTICS AND MACHINE LEARNING IN SUPPLY CHAIN RISK MANAGEMENT

A. Predictive Analytics

Predictive analytics is a very important tool for the management of risks in a modern supply chain because organizations are able to foresee probable interruptions and take proactive measures beforehand. This methodology uses data in generating insights into future events and therefore enables entities to identify risk

ISSN 2394 - 7780

early enough and prepare for different eventualities. Using statistical models and data mining methodologies, predictive analytics examines both historical and real-time data, thus allowing firms to find patterns that might be indicative of problems before they occur.

B. Techniques in Risk Forecasting and Anomaly Detection

Most of the supply-chain risk predictions use time- series analysis, regression models, and other statistical techniques that will predict disruptions based on past information. Another application includes anomaly detection, finding outliers or unusual patterns in data, which may indicate cases of fraud, cyber threats, or operational issues. For example, machine-learning algorithms like clustering and PCA help companies identify anomalies in supply chain data that will eventually turn into major problems if left undetected until it is too late. These techniques enable organizations to be proactive rather than reactive after the disruptions occur. Predictive analytics also helps with demand forecasting, inventory optimization, and assessing supplier risks. For instance, time-series models can predict changes in demand and help organizations adjust their inventory levels to avoid surplus stock. Furthermore, Regression models can be used to assess the reliability of suppliers by analyzing past data on deliveries, quality issues, and other performance metrics. Together, these methods enhance supply chain transparency and ensure that all possible risks are well controlled.

C. Case Studies/Examples of Predictive Models Used in Supply Chain Risk Management

Demand fluctuations and supply shortages due to the use of predictive models in manufacturing industries. Some examples are predictive analytics in the usage of multiple manufacturers to analyze the performance of suppliers. This only tells which suppliers will most likely suffer significant delays or quality issues. From this, the organizations can then adjust their sourcing strategies before potential disruptions. For example, logistics are the prediction models that help in optimizing routes of shipping by inputting real-time data on weather conditions, cost of fuel, as well as port congestion. Because of this, firms will always be able to avoid delays and have perfect delivery schedules. Besides, the big retailers use predictive analytics for the process of inventory management by monitoring the past data concerning sales and seasons and predicting future demand. Overstock and stock outs reduced accordingly, and therefore, customer satisfaction increased while losses have minimized. Case study findings indicate that predictive analytics can identify and mitigate several risks at different stages in the chain of supply.

D. Machine Learning Algorithms

Machine learning algorithms form one of the core tools of management in supply chain risk; whereby insight is gained from complex and high-dimensional datasets. The algorithms auto-generate data analysis processing hence ensuring increased speed and accuracy in making a decision. This equips organizations to better prepare in respect of risks. Some generally used machine learning algorithms applied to supply-chain risk management have outlined below together with their applications in prediction as well as mitigation.

E. Overview of Algorithms

1. ***Logistic Regression:*** In fact, logistic regression is a useful technique for the prediction of outcomes- say whether a supplier would be reliable or not based on historical analysis and that a delay is possible. A logistic regression model explains the relations between independent variables-for example, characteristics of the suppliers-and a binary response, which is whether the supplier would comply with delivery deadlines or not.
2. ***Support Vector Machines (SVM):*** SVM is a classifying algorithm that can be regarded as supervised learning by trying to find the best- separating boundary of classes to separate the data points. Also in supply chain management SVM is used for anomaly detection which could be fraud transactions or abnormal behavior of the suppliers.
3. ***Random Forests:*** This ensemble learning algorithm combines multiple decision trees to improve the accuracy of predictions. Random forests are especially useful for complex sets of datasets of supply chains because they reduce overfitting and enhance robustness. In a risk management context, they can predict supplier risks, forecast changes in demand, and evaluate the probability of disruption.
4. ***Decision Trees:*** Decision trees create an analogy for a decision-making structure that is almost like a flowchart. Each node represents a question about the data, and each branch represents a possible answer. Supply chains use decision trees to analyze impacts of various risk factors, be it geopolitical instability or simply the geographical location of a supplier, so organizations can have strategic decisions regarding sourcing and logistics.

ISSN 2394 - 7780

F. Application of Each Algorithm in Risk Prediction and Mitigation

1. ***Logistic Regression:*** Logistic regressions applied by logistic regression models in predicting potential default risk of suppliers based on historical data about earlier performance, payment problems, and economic conditions. For instance, a logistic regression could determine the probability of which supplier defaults on a contract, and companies could act proactively by deciding to diversify suppliers for specific products or increase critical inventory levels.
2. ***Support Vector Machines:*** SVMs are the best for fraud detection as well as quality control. SVM models may flag transactions that have characteristic profiles that are unusual so managers can investigate and deal with issues before they really cause problems. This prevents further degradation of the chain in terms of integrity and reliability.
3. ***Random Forests:*** Random forests aid the maintenance of predictions by selecting which parts of the equipment bound to fail using the data of historical maintenance and performance records. This diminishes time loss and maximizes continuity. For example, random forests can assess a wide range of risks depending on weather condition, geopolitical event, among others, and, therefore, make room for adjustments in operations.
4. ***Decision Trees:*** Decision trees are a handy tool for exploring and describing possible risk situations in supply chain operations. For instance, decision trees used in order to help a logistics firm analyze the effects of several delivery routes in terms of different types of delay, cost, and environmental conditions. This will identify the most powerful and cost- efficient options that minimize risks pertaining to possible delivery stoppages.

IV. CHALLENGES IN IMPLEMENTING DATA-DRIVEN RISK MANAGEMENT

A. Data Quality and Availability Issues

One of the biggest fears in data-driven risk management is high data quality and safe access. Suppliers, logistic firms, inventory control, and feedback from customers essentially create such huge volumes of data concerning supply chains. However, this information is neither consistently accurate, nor up-to-date, nor clean. This inhibits the working of predictive models mainly by variations in data formatting, errors in entering data, and lack of generalization of the supply chain.

The main problem is that of data availability, especially in relation to information sharing with external partners and suppliers. They do not have a strong system in place. The predictive analytics need non-stop real-time access to data in order to be optimally effective. Data silos and issues of privacy regarding information with external partners, however, restrict the free flow of information. This can cause inadequate risk assessment and thereby the machine-learning models may fail in prediction.

B. Integration with Existing Supply Chain Systems

The integration of predictive analytics into legacy supply-chain systems poses a critical challenge. Many businesses remain operating on legacy systems that are not equipped with the appropriate functionalities to support advanced machine learning and data processing capabilities that contemporary risk management demands. The inability of legacy systems to support real-time analytics of data or even integration of new sources of data limits the successful implementation of predictive models.

Sometimes, the integration process itself can be hard and time-consuming. It often includes changing current infrastructures, retraining the staff, and coordinating departmental data protocols. Besides, integrating external sources of data requires consideration on the security and compliance concerns in industries with strict rules for regulatory compliance. Predictive models may work in isolation without full integration, which would thus limit their ability to improve decision-making and general visibility within the supply chain.

C. Cyber Threats and Adaptive Countermeasures

The complexity of cyber threats has been one of the great challenges to data-driven risk management in supply chains. Higher digital tools and interconnected systems have always made organizations vulnerable to cyberattacks. Cyberattacks can compromise data integrity, cause operational disruptions, and lead to significant financial losses. Cyber threats are always changing as new attack types emerge and target specific weaknesses in supply chain networks.

Adaptive countermeasures needed to address the dynamic nature of threats. To battle the new wave of cyberattacks, organizations needs install strong security protocols but also regularly update their measures. These may include encrypting data exchanges among supply chain collaborators, using anomaly detection frameworks

ISSN 2394 - 7780

to track unusual behavior, and using machine-learning techniques to identify threats in real time. This challenge, however, is doubled as implementing adaptive measures is very expensive and even requires the help of a specialist in cybersecurity.

V. PROPOSED FRAMEWORK FOR IMPROVED SUPPLY CHAIN SECURITY

Improvement of existing models through incorporation of advanced data-driven features is very crucial in enhancing the security of supply chains. Most contemporary risk management frameworks depend on static data, and their inability to adapt to new threats compromised by this. Incorporation of machine learning algorithms, predictive modeling, and real-time data analytics may enhance the precision and responsiveness of these systems. For example, predictive models, such as those tailored for demand and inventory management, will help avoid the problem of both stockouts and surplus inventory. Moreover, the use of anomaly detection models helps in the timely detection of anomalous behaviors.

Better tools for collaboration among stakeholders can also aid in increasing transparency and data sharing. Business can ensure improved information flow by developing standardized data-sharing protocols, which is important for proactive detection of risks. For instance, block chain technology can ensure the integrity and traceability of data by providing a transparent and safe environment for data sharing along the supply chain.

STEPS FOR IMPLEMENTING DATA SCIENCE AND MACHINE LEARNING IN SUPPLY CHAIN SECURITY

1. ***Integration with Existing Systems:*** Implement predictive analytics and ML algorithms by integrating them with existing enterprise resource planning (ERP) or supply chain management (SCM) systems. This allows for seamless data flow and real-time insights
2. ***Best-fit model:*** Support vector machine or random forests; depending on this best-fit, historical as well as actual information feeds models. Models have enough features for custom fine-tuning toward some unique needs a certain supply chain is presented by specific patterns it recognizes.
3. ***Continuous Monitoring and Adaptation:*** Deploy the models continuously to monitor the changes and adapt to the new patterns. Update models through feedback and new risk profiles to keep them correct and responsive.
4. ***Stakeholder Training-*** Supply chain staff should understand the capabilities and limitations of data science tools. Overall security preparedness will be increased by training in data interpretation and cybersecurity best practices.

VI. CONCLUSION

A. Summary of Findings on the Impact of Data Analytics and Machine Learning on Supply Chain Security

This study has focused on how data analytics and machine learning play into increasing supply chain security. Predictive analytics anticipate future interference and enable preventive actions against these dangers in order to significantly decrease reaction times and increase organizational resiliency. With random forests, anomaly detection algorithms, and other ML methods, such as support vector machines, one can very well examine extensive databases in real-time. As such, it allows great input into risk reduction. This is what enables the shift from reactive to proactive management of risks and hence enhances the overall security and responsiveness of supply chains.

B. Discussion of Gaps in Existing Research and Suggested Future Improvements

Despite the benefits, still areas of limitations exist in supply chain security through data analytics and machine learning. Gaps within areas identified, such as a lack of model interpretability, failure in the integration of the legacy system, complexity in predictive models due to a large amount of quality data needed, though hard to access even in complex networks of the supply chain. Further work in this direction should focus on improving data standardization methods, model transparency, and adaptive abilities where predictive models facilitated for the update with newly emerging threats.

There is a great need for further research in the dimensions of machine learning related to cybersecurity in supply chains, particularly with respect to data flow protection and management of privacy concerns. Enhancing the cooperation among participants in a supply chain by secure data-sharing mechanisms, such as block chain technology, can reduce some of these challenges through increased transparency and accountability.

ISSN 2394 - 7780

C. Emphasis on the Importance of Integrating Data-Driven Methods in Mitigating Supply Chain Risks

Now that supply networks are becoming resilient, agile, and secure, one of the critical needs of its stakeholders: data-driven means to include such risks in supply chains linked together for more vulnerable causes: it requires proactive prevention as well as detection through such mediums as analytics and machine learning by firms. Such actions undertaken by companies overcome many types of risk and do a good job in optimizing operations by securing decisions made and continuity amidst any probable disruptions.

Therefore, going forward, the integration of data science and machine learning into supply chain management comes out as an ultimate strategic imperative for organizations to maintain their competitiveness in this increasingly dynamic global marketplace. Through continuous innovation and inquiry, these technologies promise to open new opportunities for developing resilient, secure, and forward- looking supply chains designed to navigate the complexity of risk landscapes.

REFERENCES

[1] Bhalodiya, D. "Machine Learning Applications for Enhancing the Supply Chain Productivity - A Review." Proceedings of the 8th North American International Conference on Industrial Engineering and Operations Management, 2023.

[2] Chukwu, N., Simo, Y. S., Ejiofor, O., et al. "Resilient Chain: AI- Enhanced Supply Chain Security and Efficiency Integration." International Journal of Scientific and Management Research, 2024. DOI: 10.37502/IJSMR.2024.7306.

[3] Aljohani, A. "Predictive Analytics and Machine Learning for Real- Time Supply Chain Risk Mitigation and Agility." Sustainability, 2023. DOI: 10.3390/su152015088.

[4] Ni, D., Xiao, Z., Lim, M. K. "A Systematic Review of the Research Trends of Machine Learning in Supply Chain Management." International Journal of Machine Learning and Cybernetics, 2019. DOI: 10.1007/s13042-019-01050-0.

[5] Schoenherr, T., Speier-Pero, C. "Data Science, Predictive Analytics, and Big Data in Supply Chain Management: Current State and Future Potential." Journal of Business Logistics, 2015. DOI: 10.1111/jbl.12082.

[6] Waller, M. A., Fawcett, S. E. "Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management." Journal of Business Logistics, 2013.

[7] Harland, C. "Supply Chain Management: Concepts, Challenges, and Future Research Directions." Research for Development Series, Politecnico di Milano, Springer, 2024. DOI: 10.1007/978-3-031-52247-5.

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

FRAUDULENT TRANSACTION IN THE FIELD OF FINANCE

[1]Neel Naik, [2]Fatima Shaikh and [3]Balkrishna Parab
[1]Student, Department of Big data Analytics
[2, 3]Assistant Professor, Department of Big data Analytics, Jai Hind College, Mumbai 400020, India

ABSTRACT
Financial scam is a crucial problem in the banking industry and the detection of fraudulent transactions is an important task for banks to protect their customers and maintain confidence in the financial system. This has led to an exponential rise in daily transactions. In this article, three to four methods are used to form normal transactions and fraudulent transactions, namely (Support Vector Machine), Logistic Regression, Decision Tree, Random. Forest, Naïve Bayes, ANN, Bagging, Boosting and K-Nearest Neighbour. The tools and technologies used are sklearn, imblearn and benchmark metrics to evaluate model performance. In addition, in this paper, we want to combine the model that gives the most accurate accuracy result, using the classification ratio to check the accuracy and recall value of the model and help detect the fraudulent transaction faster. This evaluation provides comprehensive guidance for choosing an optimal algorithm based on the type of fraud, and this article shows the result with an appropriate performance metric.

Keywords: *Financial fraud, ML technique*

I. INTRODUCTION
The level of fraud has significantly risen as there is development in the advanced technology and communication globally. Fraud can be detected in two main ways: prevention and detection. Prevention makes it completely impossible for any attack to be instituted by fraudsters since it is an encasing layer. Detection comes when the entire process of prevention has been tried without success. The primary goal of the system fraud detection is the identification of the fraud in early stages so that the required measures can be taken against it.

Machine learning is the solution of this generation that replaces these Techniques and approaches can be applied to extensive data sets, which is not feasible for humans. Machine learning techniques can be broadly classified into two categories: supervised learning and unsupervised learning. Fraud detection analysis can be performed using two different methods, and the choice of method can be determined on the database only. Supervised learning is contingent on categorization of anomalies beforehand. Over the past few years, several supervised algorithms have been applied to fraudulent credit card detection.[3]

One strategy that is conspicuous in the credit card fraud detection issue is data mining. Credit card fraud detection means the ability to distinguish between fraudulent and legitimate transactions into two categories of genuine and fraudulent.[1].Credit card fraud detection is also centered on the spending pattern of a card. Several ways have been used in credit card fraud detection, one of them being artificial neural networks. Imbalanced datasets are a usual problem in data mining and classification particularly when the datasets are un-proportioned. Current studies have shown significant interests in the class imbalance problem. [2]. This study focuses on improving data sampling techniques by combining oversampling techniques with random under sampling [2].

Therefore, this article studies several combinations of oversampling techniques (derived from the family of Synthetic Minority Oversampling Techniques (SMOTE)) with random subsampling techniques designed to address some of the related problems. The data in this article is taken from Kaggle which contains 1,296,675 rows and 23 columns and the data is unbalanced in 0.5822% of the transactions are fraudulent in the whole data set. The main contributions of this article are briefed as follows. To tackle the problem of fraud detection, several machine learning algorithms and Deep learning algorithm are used. The combination of models is done based on the highest accuracy score. From the results of the experiments, some conclusions were drawn that may be useful for future work.

II. REVIEW OF LITERATURE
Real-time credit card fraud detection using machine learning [3] The article explores a method to detect fraud transactions in real-time using machine learning models. This system categorizes fraud into four main types and addresses challenges such as the unbalanced distribution of data, which is common in fraud detection. Using machine learning models (SVM, Naive Bayes, logistic regression and K-Nearest Neighbour) and resampling techniques (such as SMOTE), the approach It significantly improves detection accuracy and response speed. This real-time capability allows financial institutions to act quickly on reported transactions, potentially

ISSN 2394 - 7780

reducing financial losses and improving fraud management. Credit Card Fraud Detection Using a Naive Bayesian Model and KNN Classifier [4]This article explores the application of machine learning techniques, specifically Naive Bayesian and K-Neighbour Neighbour (KNN) algorithms, to detect fraudulent transactions of credit cards. The authors, Kiran Sai et al. , highlights the challenges of fraud detection due to the large number of daily transactions and the unbalanced nature of the data set, where only a small fraction of transactions are fraudulent. The authors conclude that relying on a single fraud detection algorithm may not be effective, and suggest that combining multiple algorithms can leverage their respective strengths to improve accuracy. Overall, the article contributes to ongoing efforts to improve fraud detection methodologies in the financial sector. Integration of a machine learning-based fraud detection system based on a risk management framework [5]

This article explores the application of machine learning methods, particularly ensemble approaches such as random forests, to identify fraudulent activity in digital financial transactions. It demonstrates how traditional statistical techniques have given way to present machine learning models, emphasizing how well random forests tackle the difficulties presented by unbalanced datasets, which are prevalent in fraud detection situations. In conclusion, there is opportunity in the use of machine learning in financial risk management; nevertheless, to keep up with the constantly developing financial landscape and technological obstacles, continuous innovation and development are required. Preventing fraudulent banking transactions with a deep learning algorithm [6] The paper deals basically with the fraud of financial transaction and the demand for complex detection methods that find and stop such frauds. The authors indicate the possibility of automatically extracting attribute values from data on transactions with the use of deep learning algorithms, in the form of a multilayer perceptron or MLP. Deep learning algorithms have been emphasized as a potent instrument for handling big data sets and shifting to new fraud trends, overcoming the drawbacks of conventional rule-based methods.

The article discussed the basic building blocks of the bank fraud predictor, including risk assessment, data gathering, data analysis, monitoring, and continued improvement. Both the level and volume of the training data would affect the precision of the MLP model. Ensuring that there is a diverse dataset which contains a broad spectrum of fraudulent behavior is necessary for the model to be able to identify the characteristics and patterns that define fraudulent behavior. Identifying Credit Card Fraud Using Random Forest [7] Two kinds of different random forest models that can be used to detect credit card fraud are presented in this research. For the first model, based on random forest, the basic classifier is the direct implementation of decision trees. For the second model, it employs CART as its initial classifier since it uses CART random forest. Both models are trained on historical transaction data that includes both legitimate and fraudulent transactions, so to learn the characteristics of regular and anomalous transactions. The performance of two different kinds of random forest models was investigated in this research. Our experiment uses a genuine B2C dataset of credit card transactions. Random forest still has several issues, such as inaccurate data, even if it produces good results on little data. Detecting fraud with credit cards using a machine learning algorithm from an inconsistent data set [8]. The purpose of this article is to evaluate various classifiers by analyzing various machine learning approaches using various metrics. Instead of incorrectly labeling a legitimate transaction as fraudulent, this strategy seeks to enhance fraud detection. It also addresses the issue of imbalanced data. They therefore employed strategies like undersampling and oversampling to address this issue.

In addition, clustering techniques, such as the use of k-means clustering and genetic algorithms, can be effective in dealing with unbalanced datasets by generating new examples of minority classes. Credit Card Fraud Detection Using Machine Learning Techniques: A Study in Comparison [1] Three machine learning approaches—Naive Bayes, k-nearest-neighbors model, and logistic regression—are evaluated in this article for the purpose of detecting credit card fraud. Only 0.172% of the credit card transactions in the seriously skewed data set used by the researchers were fraudulent. To balance the dataset, they adopted a hybrid strategy that undersampled transactions which were valid and oversampled transactions which were fraudulent. The article highlights the difficulties in detecting credit card fraud, including the ever-changing character of fraudulent activity and The datasets' high degree of skewness shows how effectively hybrid sampling techniques work to enhance machine learning model performance. Detecting Fraudulent Financial Transactions Using Machine Learning [9]

In this research, the use of machine learning approaches to predict the legitimacy of financial transactions with accuracy and efficiency is studied. Among the machine learning algorithms that the researchers compared were an MLP Regressor, Random Forest Classifier, Complement NB, Gaussian NB, Bernoulli NB, LGBM Classifier, Ada Boost Classifier, K Neighbors Classifier, Logistic Regression, Bagging Classifier, Decision Tree Classifier, and Deep Learning. The data was collected from the Kaggle database, with 10 columns and 6,362,620 rows. The Random Forest Classifier was the best classifier in the unbalanced informational collection with 99.97%

ISSN 2394 - 7780

exactness, 99.96% F1scores, 99.97% remember, and 99.96% precision. With 99.96%, 99.97%, and 99.97% accuracy, the Bagging Classifier was the top classifier for the balanced dataset. 95%, 99.98% F1 efficiency, and 99.98% recall. Sampling and subsampling combined methods for imbalanced classification: A analyzing credit card fraud data using machine learning [2]

In order to detect credit card fraud, this article analyzes how well classification models perform when oversampling and undersampling strategies are used. Credit card fraud represents an escalating problem because fraudulent transactions are difficult to identify accurately because of data set imbalances. The authors tackle class imbalance by combining multiple oversampling techniques from the SMOTE family and random undersampling methods. Models were evaluated using random forest classifiers and performance metrics like precision, recall, and F1 score modified for unbalanced datasets. The findings demonstrate that approaches to increase average precision, recall, and F1 score are used in both oversampling and undersampling by about 0. 80%

III. EXPREMENTAL METHODOLOGY

A. Data Description

The dataset titled "Transaction Train Fraud" was obtained from Kaggle Discovery.With a total of 1,296,675 records and 23 features, the database is heavily weighted towards the positive class, with fraudulent transactions accounting for only 0.562% of total transactions. And the dataset is in CSV format, that is, in a format where data values are separated by commas.

Table.1: Data Description

Features	Description
trans_date_time	Time when the transaction takes place
cc_num	Credit Card Number
merchant	Name of merchant who sold the product
Category	What type of product is sold
Amt	Amount of the transaction in local currency.
first, last	First-name and last name of Buyer
gender	Gender of Buyer
City, street, zip, lat, long city_pop	Address
job	Jobe of buyer
dob	Date of Birth
Trans_num	Transaction number of payment happened
Merch_lat and long	Merchant locatin
Is _fraud	This is the transactions made by the fraudulent agents inside the simulation.

B. Data Preprocessing:

Finding the most relevant factors in a dataset through feature selection is an essential method that helps to decrease overfitting, increase accuracy, and shorten training times. [3] Feature extraction is done by checking the influence of the input variable on the output variable. For accurate results and to train a model with accurate data, this article transforms the input features using PCA. The principal components acquired from the PCA transformations are only the numerical values under the attributes [4], and the only feature that has not been transformed with the transformation of the principal component analysis is the "amt" attribute. "amt" contains data that represents nothing in addition to the amount of transactions and this function can also find its use for automatic learning of the cost-sensitive and of the instance. After PCA to transform the data to the same scale, this article also used Standard Scaler to scale the data to the normal temporal format. And Last but not least, the response variable "is_fraud" has the value "1" in the event of a fraudulent transaction, or a positive result, and "0" in the event of a genuine transaction.

ISSN 2394 - 7780

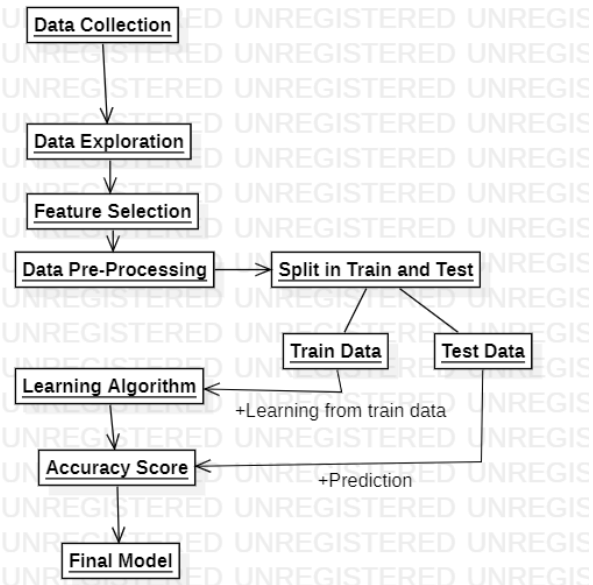


Fig.1 Data Flow

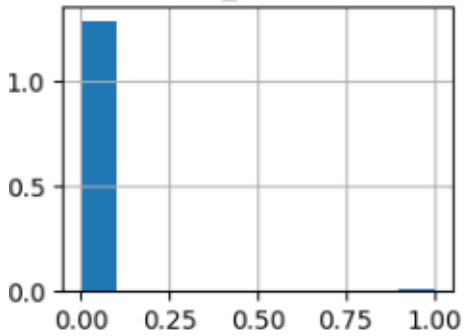


Fig.2 Imbalanced dataset

C. Sampling Method:

One method for changing the size of training groups is sampling. Oversampling alters the training samples by repeating the samples from the minority training set while the under sampling from a smaller majority training set. It is expected that both approaches will improve the situation by decreasing the degree of imbalance and class imbalance. In data mining, classification with unbalanced datasets has emerged as one of the most difficult issues. Three primary methods can be taken into consideration when sampling data: undersampling, oversampling, and a combination of the two. [2]

1. Over Sampling Technique:

In machine learning,over sampling is the process of replicating the records of lesser representation to alleviate the problem of uneven data sets, particularly where one side is substantially large as compared to the other side. This disparity may lead to models which are biased and accomplish nothing for the smaller class. In order to achieve equal distribution, more instances of the minority class need to be added.

2. Under Sampling Technique:

In machine learning, under sampling is a specific method meant to solve the problem of class imbalance within a dataset. It consists of decreasing the count of samples in the overwhelming majority class to have a more balanced data set. So, each model can concentrate on learning the features of the smaller class, which typically tends to be the focused class.

3. Combine Sampling Technique:

The mixer of oversampling and undersampling techniques is a powerful approach in the domain of machine learning, particularly in addressing class imbalance. This method attempts to reduce any overfitting issues while dealing with the class imbalance problem. In this work, SMOTE-Tomek Links was implemented where SMOTE generates artificial samples for the underrepresented class. Moreover, Tomek Links removes the nearest neighbors of the minority class samples from the majority class, thereby increasing the separation of the classes.

ISSN 2394 - 7780

D. Machine Learning Models:

1. Logistic Regression:
One technique for categorization tasks is logistic regression. It simulates the likelihood that a particular class or event will occur, depending on one or more independent variables. Logistic regression differs from linear regression, which forecasts a continuous numerical value. Logistic Regression predicts a categorical outcome, such as "yes" or "no", "spam" or "no spam", or "positive" or "negative". The sigmoid function, maps the linear combination of the independent variables to a probability between 0 and 1 [1].

Mathematically, the sigmoid function is defined as: sigmoid(z) = 1 / (1 + e^(-z))

2. Support Vector Machine:
Support Vector Machines (SVM) are a classification method that seeks to identify the ideal "hyperplane" or boundary to separate different groups. The SVM will then identify the line that best separates these two groups, choosing line that maximizes the difference or distance from the nearest host on each side. These nearest guests are known as *support vectors* and help determine the position of the boundary [9].

3. Naïve Bayes:
Based on the training data's probabilities and conditional probabilities of occurrence, the Naïve Bayes machine learning classifier attempts to predict a class known as the result class. This type of learning, which is also known as supervised learning, is incredibly efficient, quick, and practical. Conditional probabilities using Bayes' theorem constitute the initial stage of the Naïve Bayes classifier. The class is "C," and the known data sample is "x." [4]:

P (C / x) = P(x/C)/P(x)

4. Decision Tree:
A decision tree is a model that separates data into discrete to sort people into groups by particular features to provide judgments and predictions. The structure of the tree is like a flow chart where the final branches at the end hold. for the ultimate classification or suggestion, as well as each path represents one option based on a trait. A Decision Tree is an intuitive and efficient tool for classification and decision-making because of its systematic division, which enables it to make conclusions quickly by reducing the number of options.

5. Random forest:
Random Forest for classification and regression tasks is the ensemble learning method which create multiple decision tress and composite them to give predictions that is more accurate and reliable. Random Forest enhances, accuracy, handles missing data, Prevents Overfitting combining the knowledge of multiple trees. Because each tree Random Forest is particularly suitable because it makes decisions independently, The following process, if done iteratively and not sequentially, can help the developer(s) build a robust and generalizable model that is more stable with respect to variations in the data than a single decision tree would be [7].

6. ANN (Artificial Neural Network):
An artificial neural network (ANN), also known as a deep neural network, is a computer system that follows after the structure and functions of the human brain. It is mostly used for pattern recognition, classification, and prediction. An input layer, one or more hidden layers, and an output layer are the layers made up of interconnected nodes, or "neurons."Each neuron receives input, processes it by applying weights (which represent the importance of each input), then passes the result through an activation function in determines whether it should be "enabled" or enabled. This signal is then transmitted to the next layer of neurons in the network. Through a process called training, the ANN adjusts these weights by comparing its predictions with actual results, using methods such as back propagation to reduce errors. [6].

7. Bagging:
Bagging, also known as Bootstrap Aggregating, is an ensemble learning approach which utilizes several iterations of a model trained on various subsets of the data to boost the accuracy and stability of machine learning models. From the initial training dataset, several random samples, or *bootstraps*, initially develop. Each sample is made with replacement, permitting certain data points to appear in a number of sample. After that, a model—typically a decision tree—is trained independently on each of these bootstrapped data. The findings of multiple models are combined to produce the final prediction, generally by averaging predictions in regression tasks or using the majority vote in classification tasks.

ISSN 2394 - 7780

8. Boosting:
Boosting is an ensemble learning strategy used in machine learning that combines the skills of numerous weak learners to increase the accuracy of models. A weak learner is a model, including a small decision tree (violation), that performs moderately better than a random guess, typically because due to its simplicity. With each new model emphasizing on the errors caused by prior ones, the improvement aims to instruct these weak learners in order. The original data is used to create the first model, then subsequent models are trained with altered data, emphasizing observations that the first models did not fully grasp. By fixing the mistakes of its predecessors, each model in the series "boosts" performance, and the final improved model aggregates all of the weak learners into a single strong model.

9. KNearest Neighbour:
The KNearest Neighbours (KNN) algorithm is a simple yet powerful method, is used for classification and regression tasks. How it works: It compares a new data point to homogenization, where "k" is the number of the closest object points in its environment. training on n neighbours is a given number of neighbours The choice of "k" affects the workings of the algorithm: a low “k” centers in on the nearest neighbours, so it is more sensitive to noise, whereas a biggish “k” takes a broader scope to the stability [4].

IV. PERFORMANCE EVALUATION AND RESULTS
The experiments are assessed using four fundamental metrics: Four basic metrics evaluate the experiments which include false positive rates (FPR), false negative rates (FNR), true positive rates (TPR), and true negative rates (TNR). Cases that receive positive classifications and are confirmed to be true positives are referred to as true positives. True negative situations receive the correct classification. False positive cases receive positive classification despite being negative. False negatives are cases that receive negative classification when they should be positive. The Naive Bayesian, Logistic Regression, SVM, Decision Tree, Random Forest, Trunk, Boost, and Kneighbor models' ANN, sensitivity, specificity, and accuracy are evaluated. How these assessment criteria are applied depends on how well they assess the unbalanced binary classification challenge.

Accuracy = (TP + TN) / (TP + FP + TN + FN)

Sensitivity = TP / (TP + FN)

Specification = TN / (FP + TN)

Precision = TP / (TP + FP)

TP (True Positive): Number of positive cases correctly identified as positive. TN (True Negative): Number of negative cases correctly identified as negative. FP (False Positive): Number of negative cases incorrectly identified as positive. FN (False Negative): Number of positive cases incorrectly identified as negative. Sensitivity (Recall) gives precision in the classification of positive cases (fraud). The specification provides precision in classifying negative (legitimate) cases. Accuracy gives accuracy in cases classified as fraud (positive). [1]

A. RESULTS
Nine classifier models are created in this study, based on bagging, boosting, kneighbour, decision trees, random forests, naïve bayes, logistic regression, svm, and ann. To evaluate these models, 0.7% of the dataset is used for training, and 0.3% is set aside for testing and validation. Accuracy, sensitivity, specificity, and precision are used to evaluate the performance of the three classifiers. Classifier accuracy for the original dataset distribution, 0.9724:99.0262 dataset distribution, The sampled 70:30 distributions are with Over, under and combine sampling technique are presented in Tables 2,3 and 4

Table 2. Accuracy results for Over sampled data distribution

Metrices	Classifiers								
	Naïve Bayes	Logistic Regression	Decision Tree	Random Forest	SMV	ANN	Bagging	Boosting	Kneighbour
Accuracy	0.82	0.86	0.99	0.99	0.46	0.95	0.99	0.95	0.99
Precision	0.97	0.94	0.99	0.99	0.15	0.96	0.99	0.97	0.97
Sensitivity	0.67	0.76	0.99	0.99	0.02	0.94	0.99	0.93	0.99
Specificity	0.98	0.95	0.99	0.99	0.91	0.96	0.99	0.98	0.97

ISSN 2394 - 7780

Table 3. Accuracy results for Under sampled data distribution

Metrices	Classifiers								
	Naïve Bayes	Logistic Regression	Decision Tree	Random Forest	SMV	ANN	Bagging	Boosting	Kneighbour
Accuracy	0.81	0.86	0.94	0.95	0.3	0.85	0.95	0.94	0.85
Precision	0.97	0.94	0.94	0.95	0.27	0.94	0.95	0.94	0.88
Sensitivity	0.64	0.77	0.95	0.94	0.23	0.76	0.94	0.94	0.84
Specificity	0.98	0.95	0.94	0.95	0.37	0.95	0.95	0.94	0.85

Table 4. Accuracy results for Combined sampled data distribution

Metrices	Classifiers								
	Naïve Bayes	Logistic Regression	Decision Tree	Random Forest	SMV	ANN	Bagging	Boosting	Kneighbour
Accuracy	0.82	0.86	0.99	0.99	0.47	0.95	0.99	0.95	0.99
Precision	0.97	0.94	0.99	0.99	0.25	0.95	0.99	0.97	0.97
Sensitivity	0.66	0.76	0.99	0.99	0.06	0.95	0.99	0.92	0.99
Specificity	0.98	0.95	0.99	0.99	0.63	0.95	0.99	0.97	0.97

B. Comparative Analysis

The comparative analysis evaluated three sampling techniques—oversampling, under sampling, and combined sampling—using multiple classifiers. Oversampling and combined sampling consistently show the highest performance across most classifiers like decision tree, random forest, bagging, boosting, and k-nearest neighbour (KNN), with high accuracy, precision, sensitivity, and specificity. Both techniques lead to nearly identical results for these classifiers. Under sampling shows slightly lower sensitivity and accuracy for some classifiers and struggles more with class imbalance, especially with KNN and SVM. Decision tree, random forest, bagging, and boosting emerge as the top-performing classifiers in all sampling techniques. Conversely, SVM consistently struggles, particularly in sensitivity and precision, regardless of the sampling method.

Decision tree, random forest, bagging, and boosting were used for making hybrid model

V. CONCLUSION

The current study demonstrates that addressing class imbalance is crucial to successful financial transaction fraud detection. The analysis of several classifiers, including Decision Tree, Random Forest, Bagging, and Boosting, finds that they perform better than any other sampling technique, particularly when dealing with imbalanced data sets. The top performance of these models in fraud detection is due to their superior sensitivity coupled with high accuracy and precision levels. When utilizing SVM and K-Nearest Neighbors classifiers they face challenges with class imbalance especially during undersampling which results in reduced accuracy for fraud detection tasks.

The research demonstrates the importance of using combined sample techniques like undersampling and oversampling to address class imbalance issues. Hybrid models which integrate Decision Tree, Random Forest, Bagging, and Boosting offer dependable solutions that improve both accuracy and reliability for fraud detection tasks. This research offers important findings that enhance financial security frameworks by enabling the creation of fraud detection systems designed to manage dynamic and high-stakes financial transactions effectively.

ACKNOWLEDGEMENT

I would like to appreciate my college faculty and my parents for the support without this research project wouldn't have been possible.

REFERENCE

[1] A. John O, A. Adebayo O. and O. Sumuel A., Credit card fraud detection using Machine Learning, IEEE, 2017.

[2] Haziqah Shamsudin, Umi Kalsom Yusof, Andal Jayalakshmi and Mohd Nor Akmal Khalid, Combining oversampling and undersampling techniques for imbalanced classification: A comparative study using credit card fraudulent, Sapporo, Hokkaido, Japan: 0 IEEE 16th International Conference on Control & Automation (ICCA), 2020.

ISSN 2394 - 7780

[3] Anuruddha Thennakoon, Chee Bhagyani, Sasitha Premadasa, Shalitha Mihiranga and Nuwan Kuruwitaarachchi, Real-time Credit Card Fraud Detection Using, IEEE, 2019.

[4] Sai Kiran , Jyoti Guru , Rishabh Kumar , Naveen Kumar, Deepak Katariya and Maheshwar Sharma, Credit card fraud detection using Naïve Bayes model based and KNN classifier, International Journal of Advance Research, Ideas and Innovations in Technology , 2018.

[5] Lingfeng Guo, Runze Song , Jiang Wu , Zeqiu Xu and Fanyi Zhao, Integrating a Machine Learning-Driven Fraud Detection System Based on a Risk Management Framework, Preprints.org, 2024.

[6] P. Manikandaprabhu, S. Prasanna, K. Sivaranjan and R. Senthilkumar, Fraudulent Banking Transaction Classification Using Deep Learning Algorithm, International Journal of Advanced Research in Science, Communication and Technology (IJARSCT) , 2023.

[7] Shiyang Xuan, Guanjun Liu, Zhenchuan Li, Lutao Zheng, Shuo Wang and Changjun Jiang, Random Forest for Credit Card Fraud Detection, IEEE, 2018.

[8] S. Warghade, V. Patil and S. Desai, Credit Card Fraud Detection from Imbalanced Dataset Using Machine Learning Algorithm, International Journal of Computer Trends and Technology (IJCTT), 2020.

[9] Mosa M. M. Megdad, Bassem S. Abu-Nasser and Samy S. Abu-Naser , Fraudulent Financial Transactions Detection Using Machine Learning, International Journal of Academic Information Systems Research (IJAISR) , 2022.

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

SECURING IOT DEVICES IN HEALTHCARE

[1]Pranav Patel and [2]Wilson Rao
[1]Student and [2]Assistant Professor, MSc. Big Data Analytics Department, Jai Hind College, Mumbai

ABSTRACT
IoMT (Internet of Medical Things) or HIoT (Healthcare IoT) has transformed patient monitoring, improving treatments and streamlining processes, but not without its fair share of risks. By means of a literature survey, this paper aims to examine current research on securing IoMT devices, observing the challenges involved, identifying critical security threats such as the Eavesdropping attacks with categorization of these threats, exploring potential security solutions and various other concepts. Finally, the paper explores future trends such as blockchain, machine learning and artificial intelligence, along with their potential application in healthcare data security.

Keywords—*Healthcare, IoMT, IoT, HIoT, cybersecurity, cryptography, security, vulnerabilities, cyberattacks, encryption, cloud.*

INTRODUCTION
IoT devices are ones that have hardware such as sensors, underlying software and other technologies integrated into them, enabling them to collect data, connect to the internet and share data. They are composed of numerous parts including sensors, actuators, microcontrollers /microprocessors and many more.

Hence, these IoT devices are interconnected physical devices that can collect, exchange, and analyse data.

Healthcare systems have been increasingly employing IoT devices to enhance patient care. These Healthcare Internet of Things (HIoT) systems facilitate remote patient health, monitoring, streamline operations and allow for the control of various medical devices. Data is collected from an individual and their activities. Data is gathered using a variety of body sensors including heart rate body temperature oxygen saturation level (SPo2) and more [5], from an individual and their activities.

Healthcare professionals can use this information to track vital signs, remotely monitor patients and administer timely medical interventions.

With the integration of such IoT devices, there has been an increasing concern for their security. Increased application leads to increase in the attack surface and as these are interconnected, a single piece of vulnerability may expose the entire network, and hence, the confidentiality of the Protected Health & Personally Identifiable Information (PHI & PII) along with the exposed digital infrastructure.

In order to tackle these issues, healthcare institutions need to implement security strategies that include device authentication, access management, network protection, and data encryption to safeguard patient privacy and the security of their information. It is crucial to utilize efficient strategies and tactics to resolve these challenges. In this context, innovative technologies such as blockchain, artificial intelligence (AI), deep learning, and machine learning can play a valuable role in developing new security and protection approaches.

As a result, integrating these IoT devices into healthcare systems presents both significant security challenges and numerous advantages for patient care. The current applications with their security problems and solutions possible, within the healthcare domain are explored further.

HIoT DEVICES
There are a plethora of IoT enabled technologies that are being employed in healthcare. Some of these devices are external in nature, some are implantable and operate within the patient's body, while others can be stationary devices or find applicability as hospital operation tools as well [7].

Some of the external IoT enabled technologies and devices that are being employed in healthcare are as follows:

1. **Wearable health monitors/Fitness trackers/ Smartwatches:** With the development of smart watches, wearable health monitors like Fitbit have provided continuous monitoring of vital statistics. These provide out-of-clinic patient monitoring, which could prove fruitful for any ongoing health conditions, provide care post any surgical operations, as well as for personal health tracking.
2. **Wearable ECG monitors:** They are applied as patches or straps for specialized cardiac monitoring. QardioCore, KardiaMobile and AliveCor are a few examples.

ISSN 2394 - 7780

3. Pulse oximeters: Used to measure oxygen levels in the blood and provide personal and an out-of-clinic experience and they include Masimo Rad-5v and Health Air.

Development of implantable IoT enabled technologies in healthcare has resulted in many devices such as:

4. **Neurostimulators:** These are used to treat conditions like epilepsy and chronic pain by applying electric stimulation to nerves in particular brain regions. Devices include the Vercise DBS System from Boston Scientific and Medtronics InterStim therapy.
5. **Continuous glucose monitors:** These devices are used in homes to track blood glucose levels without requiring a needle prick to draw blood samples. Abbotts Freestyle Libre is one such device.
6. **Cochlear Implants:** Typically fitted in a medical environment, these implants are surgically placed in the inner ear and can greatly benefit individuals with severe hearing impairment. Innovations like the Nucleus 7 and SYNCHRONY 2 cochlear implants have been developed by companies such as Cochlear and MED-EL, respectively.
7. **Intraocular Pressure sensors:** Sensors implanted in the eye are used to treat glaucoma and continuously measure intraocular pressure. Sensimeds Triggerfish is one of the devices that provide both clinical and home-based monitoring capabilities.

There are also several stationary devices apart from the ones mentioned above, that are regularly used in healthcare such as:

8. **Smart beds:** Beds that can adjust for the patient's need through continuous monitoring such Stryker's ProCuity Bed or the Hill Rom Centrella Smart + Bed.
9. **EKG Machines:** IoT-enabled are now employed and provide real time data to the healthcare practictioners for faster recognition of cardiac issues.

IoT technologies are also leveraged in the development of healthcare administration tools such as:

10. **Asset tracking solutions:** Real-Time Location Systems (RTLS) like Stanley Healthcare's AeroScout or CenTrak's offerings, are utilized to monitor the location of important assets such as medical devices or staff in real time, ensuring their optimal utilization.
11. **Environmental monitoring:** Tools such as the Elpro Ecolog-Net system monitor factors like temperature, humidity, and air quality across different areas of hospitals, including operating rooms and storage sites for delicate materials. This enhances patient health and ensures compliance with safety standards.

There are also several cutting edge R&D focused IoT tools being developed such as the Organ on a chip, genetic analyzers and even surgery systems with robot assistance, to name a few.

The summary of application of IoT devices in healthcare is provided in Fig.1.

Aside from the the development of the above mentioned tools and devices, IoT in healthcare is also being used to provide holistic solutions to patient needs and care as [1]:

12. **IoT based ambulances:** An IoT equipped ambulance allows a medical team to remotely recommend the appropriate course of action for the patient. Red Ninja was the first business to create a LiFE (Life First Emergency Traffic Control) algorithm, which modifies traffic light patterns or their durations, for emergency service providers and ambulances
13. **Nexleaf analytics:** This application is greatly useful for immunizations and vaccines in underdeveloped nations. This application conducts monitoring of life-saving vaccine's temperatures within the refrigerator. These vaccines are distributed to clinics and healthcare systems in isolated or rural locations.
14. **Quio:** In the context of COVID-19, this home-based service that is connected via the Internet of Things, shows promise. For instance, IBM and Pfizer collaborated on the IoT-enabled Parkinson House project. By tracking the efficacy of medications and making real-time adjustments if necessary, this improves the doctor-patient connection.
15. **Ambient Assisted Living:** By focusing on senior citizens, it alludes to sophisticated systems of support for a better, safer, and healthier existence in the desired living environment. This method combines wearable and mobile technology with which caregivers can be alerted of hazardous situations to the older people in a house or ambient living facility centres.

ISSN 2394 - 7780

ARCHITECTURE OF IoT DEVICES

Before we discuss the vulnerabilities and security of HIoT systems, it is imperative to develop an understanding of its architecture and its workflow. The architecture of HIoT systems is described as a 4 layer architecture [9] composed of:

1. **Perception Layer:** It contains the physical hardware and it collects data from the environment, which is then sent to the network layer. Simply put, it is the layer data collection layer.
2. **Network Layer:** It connects all the IoT devices in the desired system and allows them to transfer data among themselves. The data is also sent to the base station via technologies such as ZigBee, Wifi, Bluetooth and others.
3. **Middleware Layer:** This layer stores all the procured data from the previous layers, into a database. It includes the services that are used by the application user and allows IoT programmers to work with non-homogenous technologies and to move away from focusing on a specific platform or implementation.
4. **Application / Business Layer*:*** By creating graphs, business models, flowcharts, and other outcomes, this layer—with which the user interacts—is in charge of overseeing all operations and healthcare services. This layer fulfills patient wishes by offering high quality healthcare services.

The major phases in the workflow of an HIoT systems are as follows [9]:

1. **Data generation:** This phase involves the first layer and collects required patient information by various sensors and hardware. It can even include data entry, directly from the healthcare professionals or the patient. The data generated is moved via the second layer.
2. **Data processing:** It involves analysis of the procured data via predetermined algorithms, machine learning methods or other techniques, conducted in the middleware layer.
3. **Data consumption:** The analysed data is finally consumed by the healthcare professionals or personnel for required decision making and can even be used by other systems to trigger actuators to perform some physical application/operation. This phase is performed by the application or business layer.

Fig. 2 illustrates various layers in the HIoT system architecture and the corresponding phases.

NOTABLE CYBER ATTACKS ON IoT DEVICES

IoT devices have their fair share of vulnerabilities that have been exploited in the past, leading to exposure of confidential data, destruction of infrastructure and much more damage [8].

Some notable attacks on IoT systems and devices are:

1. **Mirai Botnet:** The worst DDoS attack was carried out against internet performance management services provider Dyn, in 2016. The attack was performed using the Mirai Botnet, an IoT botnet. The malware-infected computers started looking for susceptible IoT devices on the internet and infected them by entering in with their default identities and passwords.
2. **Verkada hack:** Verkada, a cloud-based video surveillance platform, was hacked in March 2021. The attackers gained access to sensitive information belonging to Verkada software clients and live feeds of over 150,000 cameras mounted in jails, schools, hospitals, and businesses by using valid admin account credentials they could locate online.
3. **Cold in Finland:** Two buildings in the Finnish city of Lappeenranta had their heating switched off by cybercriminals in November 2016. Another DDoS attack then prevented the heating controllers from ever turning on, forcing them to repeatedly reset the system. This attack was severe since Finland has very cold temperatures at that time of year.
4. **Jeep Hack:** In July 2015, the Jeep SUV's security was inspected by a group of experts. By exploiting a vulnerability in a firmware upgrade, they were able to take over the car. After that, they could manage the car's speed and even take it off the road.
5. **Stuxnet:** Stuxnet is probably the most famous IoT attack. Its target was a uranium enrichment plant in Natanz, Iran. During the hack, the Siemens Step7 software running on Windows was compromised, allowing the malware to access the industrial program logic controllers. This allowed the virus developers to gain access to vital industrial data and gain control of several pieces of equipment at the industrial sites. This malware is thought to have damaged 984 uranium-enrichment centrifuges. According to estimates, this resulted in a 30% drop in enrichment efficiency.

ISSN 2394 - 7780

Among the many IoT enabled devices that find themselves being employed in the healthcare domain, some of the vulnerable IoT devices found among them are:

1. **Insulin and Infusion pumps:** These can administer blood, saline and other fluids, remotely, which decreases costs and assures quality of patient care. But these devices can be disrupted by exploiting the connectivity capabilities present in them.
2. **Smart pens:** There is some amount of patient data stored in smart pens that are an attractive target to cybercriminals. In addition to the data stored on the pen, it can also be used as an entry point to medical record databases and more. A cybersecurity researcher had exploited this very vulnerability in 2017.
3. **Implantable cardiac technology:** Devices like pacemakers and their programming devices, have potential to kill, as researchers have discovered that a simple DOS attack on these devices can prove fatal.
4. **Thermometers and temperature sensors:** There was a case wherein a casino was hacked via their fish tank's smart thermometer. Such a case only serves to show that security is a critical aspect in IoT enabled technology in healthcare and vulnerabilities have to be assessed carefully to avoid such exploitations.

SECURITY THREATS IN HIoT

There are several reasons for cyber attacks in HIoT devices such as monetary value of the data and its potential for malicious activities, regulatory compliance restrictions that may lead to difficulties, and limited resources for security mechanisms, among others. The reasons are illustrated in Fig. 3 [18].

In a recent study, HP examined a variety of widely used Internet of Things devices, including thermostats, door locks, webcams, and home alarms. The company discovered that a startling 70% of the devices used unencrypted network services. Furthermore, most did not encrypt data while it was in transit [4].

IoT device vulnerabilities are frequently the result of ignorance or the connection of technologies that were not initially thought to pose a risk as a separate entity. Since the name "IoT" refers to a broad category of devices, and since HIoT gadgets may have socioeconomic ramifications, the scope of these risks and exploits is enormous.

References [5], [6], [13], [17] discuss the various security issues and concerns faced in HIoT which are:

1. **Identity Management and Authentication:** IdM and authentication use a combination of technologies and procedures to secure and control access to data and resources while safeguarding the "things" profile. IdM gives items a unique identification, and authentication entails confirming the two communicating parties' identities.
2. **Data Integrity:** Information must be shielded from outside alteration. Life-critical patients suffer significant harm as a result of the lack of integrity. Data loss can also happen in an unfavorable communication environment.
3. **Authorization:** Preventing unauthorized user participation is crucial since an application may have an arbitrary number of users. Once recognized, authorization enables us to ascertain if the individual or item is permitted to possess the resources. Usually, access controls are used to implement it. Authorization and access control are required to establish a secure connection between several devices and services.
4. **Data aggregation:** Wireless sensor networks frequently experience node failures, hence a safe data accumulation technique is required to guarantee that accurate data is gathered from sensor nodes across the network.
5. **Data confidentiality:** In many situations, maintaining data confidentiality is the primary limitation.
6. **User Privacy and anonymity:** If any intruder overhears the vital information of a patient, it can be used with malicious intent, which can cause several psychological and emotional distress to the patient, further worsening their condition.
7. **Data Freshness:** Attackers may capture data in transit and may transmit it repeatedly to confuse the coordinating node. Hence, the freshness of the data may be lost. Also, data freshness is imperative since healthcare personnel need real time updates of the patients, and lack of data freshness can hamper their decision making ability.
8. **Secure localization:** Since the sensors and devices gather location data, this information must also be shielded from misuse.

ISSN 2394 - 7780

Some common security attacks in the healthcare systems are presented such as phishing attacks to trick individuals to provide sensitive information, insider threats by misuse of access privilege, jamming attack to block communication, desynchronization attack by introduction of loops to waste energy, and sybil attacks through embodiment of multiple personas within the network [18].

Fig. 4 represents the common types of cyber security attacks.

Since we have explored the various issues and attacks faced by IoT systems, we need to understand the limitations of these devices, that hamper their security functionalities. The two major limitations are presented as [13]:

16. **Battery life:** Since some IoT devices are placed in locations without charging, they possess a very limited amount of energy, and hence, strict security settings may cause the devices to use up all of their resources. To lessen this problem, there are three potential strategies. The first of them is to employ minimum security requirements, which is not advised, particularly when handling sensitive data. Increasing the battery's capacity is the second strategy. Nonetheless, the majority of IoT devices are made to be compact and light. No additional space is available for a larger battery. The last strategy is to use natural resources (such as light, heat, vibration, and wind) to generate energy, however this would necessitate a hardware upgrade and greatly raise the cost.
17. **Lightweight computation:** The computing as well as storage requirements for any advanced cryptographic techniques are hindered due to limited memory space.

IoT attacks can broadly classified based on the layer affected as:

1. **Physical attack:** It is usually performed given the attacker is in close proximity to the device.
2. **Network attack:** Involves manipulation of the network, leading to damage.
3. **Software attack:** Occurs due to exploitation of vulnerabilities in the IoT application.
4. **Encryption attack:** Breaking the system encryption.

The attacks on HIoT devices into can also be segregated into 5 major categories depending upon the type of the attack taking place:

1. **Selective forwarding:** Herein, packets are forwarded selectively by malicious nodes to disrupt the routing path.
2. **Sinkhole attack:** It involves an infected node attracting nodes in proximity and causes them to route the traffic through it. When coupled with selective forwarding, this attack becomes very powerful, and all the information is sent to the attacker who can profile the patient through their information.
3. **Jamming:** This is a classified machine-to-machine attack that uses a noise signal to occupy the wireless band, obstructing communication between Internet of Things devices and producing interference.
4. **Flooding:** The goal of this attack is to disrupt the transmission by repeatedly establishing connection requests and using the target resources.
5. **Phishing:** A technique used to steal data and personal information.

Since eight out of ten organizations reported experiencing a cyberattack on their IOT devices in 2019, IOT security becomes increasingly important as its adoption grows. Ninety percent of those organizations were impacted by the hack, which included compromising customer data or end-user safety as well as operational outages [11].

Hackers "killed" (simulated) patients at the 2018 RSA Conference USA without the doctors even realizing the operating room had been compromised. The simulation illustrated how dangerous it is for IOT device vulnerabilities to put a patient's life at risk without the present medical staff's knowledge.

Over the past five years, cyberattacks have generally increased by 125% in healthcare ecosystems. One healthcare organization owned 12,000 of the more than 68,000 medical systems that two security researchers found to be online in late 2015. The figures show the degree of threat posed by IOT devices and the hazards involved, should they be penetrated by malicious actors that could interfere with system functionality and steal personal data.

ISSN 2394 - 7780

A number of research investigations to expose security issues displayed, that such devices were connected to systems running on Windows XP, which is known to have its fair share of vulnerabilities. A search engine called Shodan which can find IoT devices connected to the internet, was used to find these devices.

Two researchers used Shodan to locate MRI scanners, pacemakers, nuclear medical systems, infusion systems, anaesthetic equipment, cardiology devices, and other gadgets. Because even script kiddies can exploit these flaws, the threat is serious. Researchers looked into these gadgets' exposure, and the CIA triad was used to gauge their influence.

To address the growing concern of cybersecurity in medical devices, the Food and Drug Administration (FDA) has created cybersecurity standards for three kinds of medical devices before the devices are put on the market.

1. MEDICAL DEVICE CLASSES

Medical Device Class	Attributes	Example Devices
I	Common devices with low risk and low complexity	Lancet, Dental Floss
II	More complex devices with a greater risk to the patient, partially implanted	Syringe, Insulin pumps, BGM
III	Fully implanted devices with greater risk, such as Replacement Heart Valves	Artificial Pancreas, CGM

Hence, there is a need for experts and researchers to prioritise their focus in securing the Class II and III medical devices

LITERATURE REVIEW

IoT devices should include network segmentation and monitoring in order to detect any unusual traffic [4].

One major security concern is the use of insecure passwords and Wi-Fi and routers with their default security configurations and passwords. As per OWASP, the most common vulnerability in IoT devices is weak passwords. Users frequently forget to change the default passwords or don't follow best practices for creating strong, secure passwords. The security implications of the HIoT are just as important, notwithstanding its potential to promote socioeconomic progress and health-related wellness.

Table II illustrates the numerous dangers for each tier of the standard HIoT architecture.

Reference [14] discusses the classification of IoT attacks and the different security issues across the layers. It explores the security and privacy requirements of HIoT devices and underlines 4 requirements involving data integrity, its usability, data auditing and privacy involving patient information.

2. SECURITY THREATS IN IOT ARCHITECTURE [3]

Layer	Description	Threats
Cloud	Data center cloud layer/cloud network host applications that are critical providing IoT services.	Data interruption, DDoS, Buffer overflow, Impersonation, Remote code execution
Core	The function of this layer is to carry and exchange data and network information between multiple subnetworks.	Data interruption, Man-in-the-middle (MITM) attacks, Impersonation/Spoofing, Modification of data at rest and in transit, Relay attack, Confidentiality attack, Jamming/Congestion, Data exchange issues: data privacy, access control, and disclosure of information

ISSN 2394 - 7780

Edge	Endpoint devices with both wired and wireless connectivity. This scalable layer supports Zigbee, IEEE 802.11, 3G and 4G.	Connection flooding, Data interruption, DoS, Eavesdropping, Impersonation, Jamming attack, Modification of data at rest and in transit, Misconfiguration, Network protocol vulnerability and exploit, Packet manipulation, Physical attack/tampering, Rogue access points
Things	Embedded systems and sensors. Small devices with varying OS, CPU types, memory, network capability.	Authenticity, Device end-point attack, Counterfeiting attacks, Eavesdropping, Hardware interruption/theft/modification, Jamming attack, Resource exhaustion, Privacy, Spyware, Repudiation, Device specific vulnerabilities, ie. OS vulnerabilities, malware, weak authentication, etc

THE EXISTING SOLUTIONS AS DISCUSSED FURTHER AS

1. **Data encryption:** Three degrees of implementation are possible. Initially, as link encryption for data transmission between links. The second is node encryption, which prevents the network node from receiving messages in plaintext. Third, end-to-end encryption, which means that until the message reaches its destination, it cannot be decrypted.
2. **Access control:** One can employ symmetric, asymmetric or attribute-based key encryption.
3. **Trusted third party auditing:** The Trusted Third Party (TTP) provides unbiased audits results to guarantee that cloud service providers are held responsible and to protect the legitimate benefits of cloud users.
4. **Data search:** Enabling an encrypted cloud data search service is essential. The two primary methods for searchable encryption are public-key encryption with keyword search (PEKS) and searchable symmetric encryption (SSE)..
5. **Data anonymization:** Patient sensitive data can be divided into three categories: explicit identifiers, quasi-identifiers, and privacy characteristics. Explicit identifiers that can be used to uniquely identify a patient include an ID number, name, and mobile phone number. A combination of quasi-identifiers, such as age, address, and birth details, can also be used to uniquely identify a patient. The term "privacy attributes" refers to sensitive patient characteristics, such as illness and income. The distribution characteristics of the original data must be taken into consideration while handling the individual attributes of the new dataset during the data publishing process in order to protect patient privacy.

Additionally covered are the use of lightweight cryptographic methods for effective database query processing over encrypted data and encrypted query processing for cloud storage. Additionally, lightweight encryption strategies for smart homes based on stateful identity-based encryption and a method to lower latency by dividing query results into discrete data sets have been explored.

Reference [5] discusses the security issues and remediation measures across the layers, as:

18. Lightweight cryptographic techniques are employed in the perception layer to allow for security.
19. Counterfeit and Man in Middle attacks are belong to the threats that affect the IoT network layer. Both of these techniques have the ability to send false information while simultaneously capturing information
20. At the IoT application layer, data exchange can lead to information disclosure and security issues with privacy and access control.

A secure HIoT system is proposed, which consists of 3 communication channels:

21. Sensor nodes (the edge sensors) to the internal processing unit (IPU),
22. Internal processing unit (IPU) to gateway (router)
23. Gateway to the cloud.

ISSN 2394 - 7780

The suggested approach uses three distinct security protocols for the three types of communications channels: SHA-3, AES-256, and HTTP-SSL, respectively. Additionally taken into consideration is the biosensor devices' registration.

The registration process involves using RSA-1024 by the IPU to share the public key to nearby bio-sensors. Using this public key, the sensor encrypts its own symmetric key and shares it back with the IPU, hence they now share the secret key.

Now the ID of the sensor is encrypted using the secret key, which is then decrypted by the IPU and stored as hashed values.

When a sensor sends the data, it sends the ID along with it. This ID is matched against the hashed values and hence, verifies the sensor. A basic implementation environment was also provided for the above solution. This proposed scheme aims to solve the issues of access control, authenticity, confidentiality, and integrity.

Reference [6] proposes a secured architecture of 3 layers, represented in Fig. 5, as:

1. **Device/Sensor Layer:** Wearable sensor and data acquisition units.
2. **Network layer:** Networking and communications
3. **Backend:** Processing and analysis

The security across these layers is discussed as follows:

1. **Device/Sensor Layer:** In this case, local processing is used to prevent unauthorized users from identifying the data's source when it is being transferred over a wireless network. Symmetric key cryptography is proposed.
2. **Network Layer:** In this layer, man-in-the-middle attacks are common. Public key or symmetric key cryptography is required. The software that acts as a bridge between cloud computing services and smart devices is the IoT gateway, which provides security and other services like data or protocol translation as well.
3. **Backend:** Here, data is analysed and users can query the data available on cloud or on other sources. This layer requires strong security in the form of anti-virus, public key cryptosystem, and other protocols.

The paper concludes with a discussion of cryptography and its forms, and the types of cryptography involved in this proposed architecture.

By way of a literature review, authors in [10] explore various concepts in HIoT including the security issues prevalent in them. It explains importance for aversion to DoS attacks at the network layer and the inclusion of lightweight cryptography owing to its efficiency.

The healthcare system attacks can be defined into the following categories:

24. Physical attacks include assaults on the physical hardware
25. Side channel attacks use information to determine the key being used by the target device.
26. To crack the encryption, cryptanalysis attacks are carried out.
27. Software assaults utilize the communication interface of the program itself in search for flaws.
28. Network communications are susceptible to network security attacks due to the broadcast nature of the transmission channel. Most IoT-based frameworks are vulnerable to various security threats, such as covert assaults on availability, authentication, and service integrity.

Reference [12] suggests a cloud based model for future HIoT systems, comprising of:

1. **Wearable sensors:** Wearable sensors are ones that collect all the patient data and the central node will receive this data, which subsequently processes it and may even implement any decision making involving information forwarding to another location.
2. **Short Range communication:** To transfer data from the sensor to the central node. These should be low latency, have robust security features, and should not negatively impact human health.
3. **Long range communication:** Data from the central node is sent to a database considering latency, security and several other characteristics.

ISSN 2394 - 7780

4. **Secure Cloud Storage Architecture and ML:** Using cloud storage to store the vast amount of sensor data generated and machine learning for trend identification, diagnosis assistance and patient specific recommendations.

The authors have also provided several use cases for their proposed system such as:

1. A wearable accelerometer sensor-based knee injury rehabilitation system.
2. To create systems that can help control long-term illnesses like high blood pressure.
3. Tracking alterations in patients suffering from degenerative illnesses like Parkinson's disease.

The security considerations in employing cloud for HIoT systems are resolved via access control policies and data encryption. Several security mechanisms are discussed such as:

29. **SafeProtect:** Focuses on patients creating a policy to only allow limited access to their data and have control over their information. This mechanism only allows validated healthcare providers to access the data via credentials and prevent illegal actions as per access policies that are set.
30. Signal scrambling, which uses "tiny data," a small fraction of the data that is transferred among authorized parties and serves as a scrambling key.
31. A fully homomorphic encryption is assessed and a steganography-based solution to access control is examined. FHE makes it possible to encrypt data using public keys and perform mathematical operations on it without having to decrypt it in the cloud.
32. FHE was also compared against the AES and Attribute-Based Encryption (ABS).

The security discussion concludes with a suggestion for an optimised ABE-to-FHE conversion scheme, which would be extremely valuable to the patient data.

Authors in [16] discuss some of the security issues in HIoT systems across the layers.

33. The physical layer is vulnerable to impersonation and denial of service attacks. Tampering of the nodes, injecting malicious data, modification of the routing path of the devices during the process of forwarding of the data, and packet sniffing are the attacks that the physical layer must prevail through.
34. At the network layer, security for routing attacks, denial of service and data transit attacks have to be employed.
35. Finally, the application layer is exposed to leakage of data, and injection attacks among others.

Solutions to these privacy issues are described, including automated threat detection and response, and Data Provenance, a tool that aids in recovering data from its original location to its whole history chain. Any changes made to the data during its lifespan are recorded. This aids in confirming any modifications made to the data that might result in abuse. To obtain the data provenance associated with a certain patient, a range of data models are available. Strong network security can be achieved by the use of techniques including confidentiality, access control plans, and authentication. The paper concludes with a suggestion that a single solution may not solve the security and safety issues. Patient data safety and cyber security must advance more quickly than policy implementation and regulation. To use the linked devices in the future, the Internet of Things need higher requirements every day..

Reference [17] explored several security solutions to the problems have been explored via a systematic survey, which include the following measures:

1. To enable secure network communication, lightweight authentication techniques use straightforward cryptographic hash functions. The devices' identities are concealed through the use of hash and XOR operations.
2. To ensure safe communication between the IoT client and IoT server, an addressless IoT server is used.
3. To examine problems with patient data collection, researchers have turned to kHealth, an Internet of Things-based healthcare information monitoring system.
4. A cloud-based user authentication system has been used by researchers. To ensure safe conversations, it includes a private session key. In order to verify the resilience against well-known attacks, the authors also conducted a security study of the developed mechanism using a Real-OrRandom model.

The authors have then defined a security benchmark for the IoT architecture. This benchmark includes an evaluated score to determine the strength of the security.

ISSN 2394 - 7780

The security requirements for said benchmark were outlined as follows:

1. Authentication
2. Confidentiality
3. Integrity
4. Self-healing: restoring to working state when a node failure occurs
5. Fault tolerance against failed components, resilience to attacks even when system/service failures take place
6. Data freshness
7. Trust related to access control and identity management
8. Firewall configuration
9. TLS/SSL for encrypted communication, certificates signed by a certification authority, and regular software updates.

The authors then provide their proposed AMI (Ambient Intelligence) Lab architecture with the 3 layers involving a cloud architecture. The security in this framework is provided via a secured channel with authentication from both ends, firewall configuration at gateways, using gateway address to avoid data stealing instead of user identity, presence of a middle node to limit communication, and cloud server security configuration. Security agents are deployed on the cloud and in the nodes to sense any and all intrusions.

The network layer uses 2 servers for securing communication and is being accountable for creating SSL certifications. It also employs TCP for communication, data authentication using SHA-2, use of a VPN server and using Elliptic Curve Diffie Hellman to encrypt the communication handshake.

The authors analyse their proposed framework against the security benchmarks to evaluate how it stands against the requirements.

Reference [18] discuss the structure of a smart healthcare systems comprising of the medical devices, sensors, networking components like Bluetooth, data processing capabililities to obtain insights about health conditions & the health provider which provides remote or in-person treatment.

The authors have further detailed the security schemes as per the levels and their functionalities as:

1. **Data level:** It entails confidentiality of patient data, integrity to ensure quality, efficiency and consistency of the data throughout its cycle, and availability to ensure accessibility of records whenever needed.
2. **Sensor Level:** It entails tamper proof hardware using Physically Unclonable Features, sensor localization, self-healing sensors, over the air programming to update devices with new security policies and patches, and forward and backward compatibility which signify that the messages must not be readable once they leave the network and previously transmitted messages must not be readable by sensors that have just joined the network.
3. **Personal Server Level:** The data is collected and aggregated on the server level hence 2 types of authentication are needed. Device authentication is needed to ensure that data is accepted from authenticated devices. User authentication is needed via biometric authentication or other strategies.
4. **Medical Server Level:** It includes access control, key management, trust management between 2 nodes and resistance to DoS attacks.

Several open issues and challenges in IoMT Based healthcare system, network and protocol challenges are discussed.

Reference [19] provides an overview of the different methods for encryption in IoT in healthcare. They discuss the 2 categories of encryption, symmetric and asymmetric.

36. Asymmetric key encryption employs a pair of keys, such as public and private keys, for encryption and decryption, whereas symmetric key encryption uses a single key.
37. The AES is an extensively used symmetric key encryption which is also a lightweight cryptographic algorithm and provides good security and balance performance in low-power IoT devices.

ISSN 2394 - 7780

38. Because of their simplicity and effectiveness, the lightweight block ciphers known as Speck and Simon are ideal for Internet of Things applications. Secure key management is necessary for symmetric key encryption, though.
39. Asymmetric key encryption has Elliptic Curve Cryptography which has high processing efficiency with strong security and small key sizes, hence it is an excellent choice. Quantum Key Distribution uses quantum mechanics to allow safe exchange.

The authors have then gone on to discuss the lightweight cryptography which involves LEA (Lightweight encryption algorithm) and PRESENT. Key management via Hardware Secure Modules (HSM) which provide a secure environment for managing keys, and secure key exchange protocols such as Diffie Hellman Key Exchange.

Hardware based security solutions include Trusted Execution Environments (TEEs) to process sensitive data, and Secure Elements (SEs) which are tamper resistant hardware components that store sensitive information and perform cryptographic execution.

The authors also provide an outline of several challenges and future directions involving standardisation of security protocols, interoperability between devices from different manufacturers, and the evolving threat landscape.

The authors provide the performance comparison of different encryption algorithms wherein Speck-64 and PRESENT algorithms outshine the others in processing time, minimum computation overhead and resource utilisation.

Among the security evaluations, AES-256 and ECC (256-bit) proved to be highly resistant to brute-force, known-plaintext and side-channel attacks.

The impact of encryption on the patient data transmission efficiency revealed that Speck-64 and PRESENT had higher data transfer rates and low latency.

ECC (256 -bit) which provides robust security, results in lower transfer rates and higher latency.

FUTURE TRENDS

Blockchain

Reference [2] discusses blockchain technology as a solution to the data fragmentation issue. Three aspects contribute to blockchain technology's secure transmission:

40. It features an unchangeable "ledger" that people can access and manage. There are set guidelines that must be followed for each transaction in the ledger.
41. Secondly, blockchain is a distributed system that runs concurrently on several computers, devices, etc.
42. Third, blockchain adheres to agreement standards and data exchange policies through the use of smart contracts. Identity control and access privileges to different blockchain-stored electronic medical reports (EMRs) are established by the smart contract. It suggests that doctors are limited to using the EMRs to which they have been given access.

Additionally, the authors investigate the creation of a blockchain-based application called Healthcare Data Gateway (HDG). It allows patients to share their information and ensures that privacy policies are not violated. The architecture of the same is displayed in Fig. 6.

Reference [18] also discusses blockchain as a part of the future trend to deliver robust and broad security with privacy protection. However, blockchain technologies demand significant resources, hence it may not be feasible in IoMT systems, but can be used to store data on medical servers. An example of this is MedRec, which has pioneered research into using blockchain to manage access to medical data

MACHINE LEARNING

Machine learning models for predicting and prognostics, enhancing production process efficiency with low maintenance costs, and minimizing product quality deterioration are covered in Reference [4]. Their method of analyzing sensor data would be very beneficial for e-health and wellness in the context of H-IoT.

A machine learning classifier with 4 experiments was further discussed to assess the most significant feature in classifying an IoT connection as normal or anomalous. The connection record consisted of 100 bytes.

ISSN 2394 - 7780

The four experiments conducted were:

1. **Correlation-Based Feature (CSF) Subset Evaluation:** The predictive power of each characteristic and the level of duplication among them are taken into account when evaluating the value of a subset of attributes. Consequently, the top ten predicted characteristics were determined, including protocol type, connection time, and others.
2. **Gain Ratio Attribute Evaluation:** Selecting the most predictive feature based on an "average merit". This resulted in features such as logged_in, srv_error_rate and few others being introduced in the top 10 which were previously unworthy.
3. **Information Gain Attribute Evaluation:** This experiment measures the information gain in relation to the class, in order to assess the value of an attribute.
4. **K-means clustering analysis:** This experiment aimed to cluster connections into attack or normal, but the results generated a pretty high error rate. Hence, no clear distinctions between the classes could be determined without employing other techniques.

The authors draw the conclusion that while the dataset seems to have an even class distribution, which is advantageous for machine learning, these circumstances are unlikely to arise in a real-world setting because the attack data would be greatly outnumbered by the normal data. As a result, the machine learning technique that has been explained would not yield positive outcomes when applied to a bigger dataset. Because Deep Learning can learn and make much more concise conclusions, the article recommends it as a suitable technique for granular analysis.

ARTIFICIAL INTELLIGENCE

Reference [18] also discusses AI as a future trend in IoMT systems.

The authors discuss the challenges of generative AI in healthcare systems. Several security issues arise as these systems can be manipulated to generate false data or misleading predictions, if not secured properly. It also increases the surface area for potential breaches. Other issues are data bias and the reliability of AI generated outcomes

Ensuring the privacy and security of sensitive patient data, navigating the ethical and legal complexities of AI decisions, achieving seamless integration and interoperability with existing healthcare IT infrastructures, and adapting the healthcare workforce to leverage AI technologies effectively, are some of the challenges involved.

As medical service providers are exploring the use deep learning techniques for diagnosis, its imployement for systems security and privacy is also an example of research to consider. Herein, PHI is searched at different layers of IoMT systems to detect centralised attacks through deep learning networks.

Objectives of Study

The main objective in this study was to identify the applications of IoT devices in healthcare systems, and analyse the various security issues that arise in their implementation. The security challenges identified herein were also resolved via different forms of encryption, access control, cloud implementation and many others. Several innovative and future trends were put forth such as blockchain technology, artificial intelligence, and their involvement to resolve potential vulnerabilities in security systems.

This paper aimed to provide the reader with a holistic view of the HIoT systems particularly conforming to its security aspects, exploitation and solutions, and the hope for future advancements.

Scope and Limitations

HIoT systems will continue to see more and more growth and popularity owing to their applicability in various medical situations. However, several research directions must be exploited in order to safeguard such systems. Several innovative solutions can be applicable in this context, such as:

1. **AI:** Inclusion of AI models to detect intrusions and attacks. These models are even utilised currently for various medical diagnosis and identification of medical conditions, progression and other situations. However, solutions and data offered by these models must also be verified for their accuracy, since these models are known to make errors.
2. **Blockchain:** This technology shows promise in the terms of the security offered by it. However, further research has to be conducted to understand inclusion in existing IoT systems, keeping in mind the energy efficiency of these systems and the scarcity of resources available to them.

ISSN 2394 - 7780

3. **Deep learning:** For intrusion detection and prevention of common attacks such as man in the middle attack or internal security threats.

The major issues that limit the security mechanisms applied in IoT systems are resource scarcity or limited computation. Hence, research efforts must specifically keep these in mind while developing novel techniques for security or even for healthcare applications, in general.

CONCLUSION

This study highlighted the various applications of IoT systems in healthcare. It identified the reasons for popularity of these systems as targets for cyber attacks and the various attacks that commonly occur. The major security considerations for such systems and exploration of possible encryption techniques, cloud architecture, blockchain and AI inclusion has been considered. The realm of HIoT is a vast field that has to be carefully trodden while developing novel applications and solutions for the protection of the patients and solutions for their care.

FIGURES

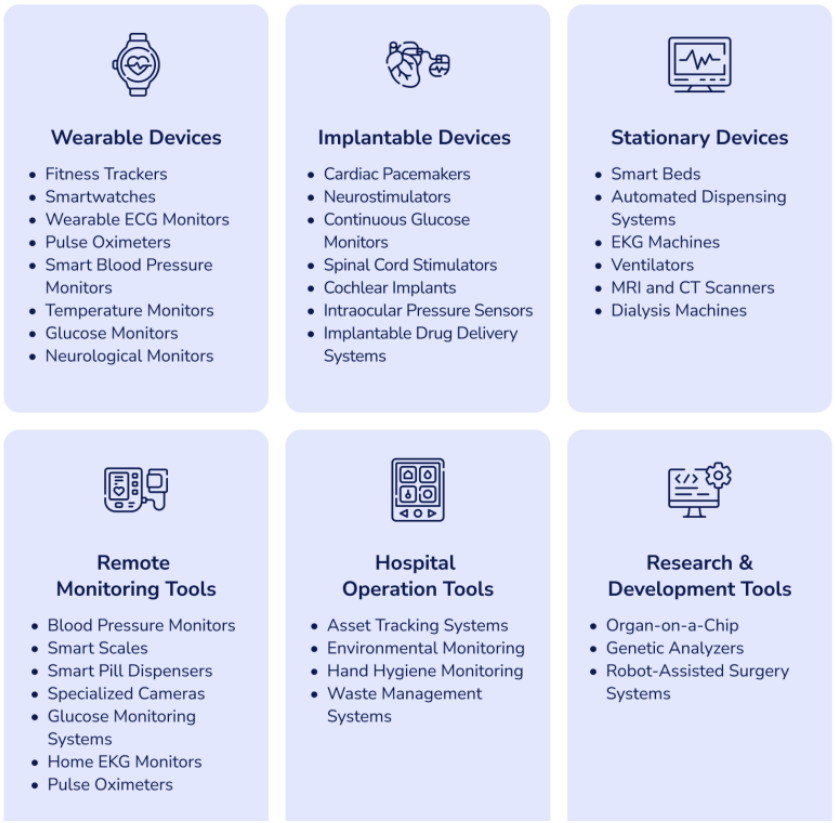


Fig . 1. IoT Devices in Healthcare Taxonomy [7]

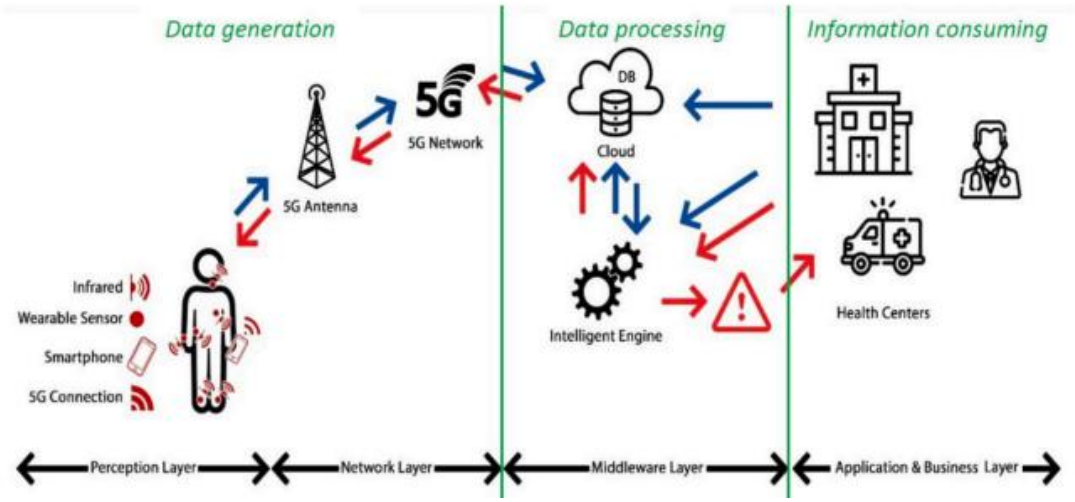


Fig. 2. Architecture of HIoT Systems [9]

ISSN 2394 - 7780

Fig.3. Reasons for cyber attacks on IoT systems. [18]

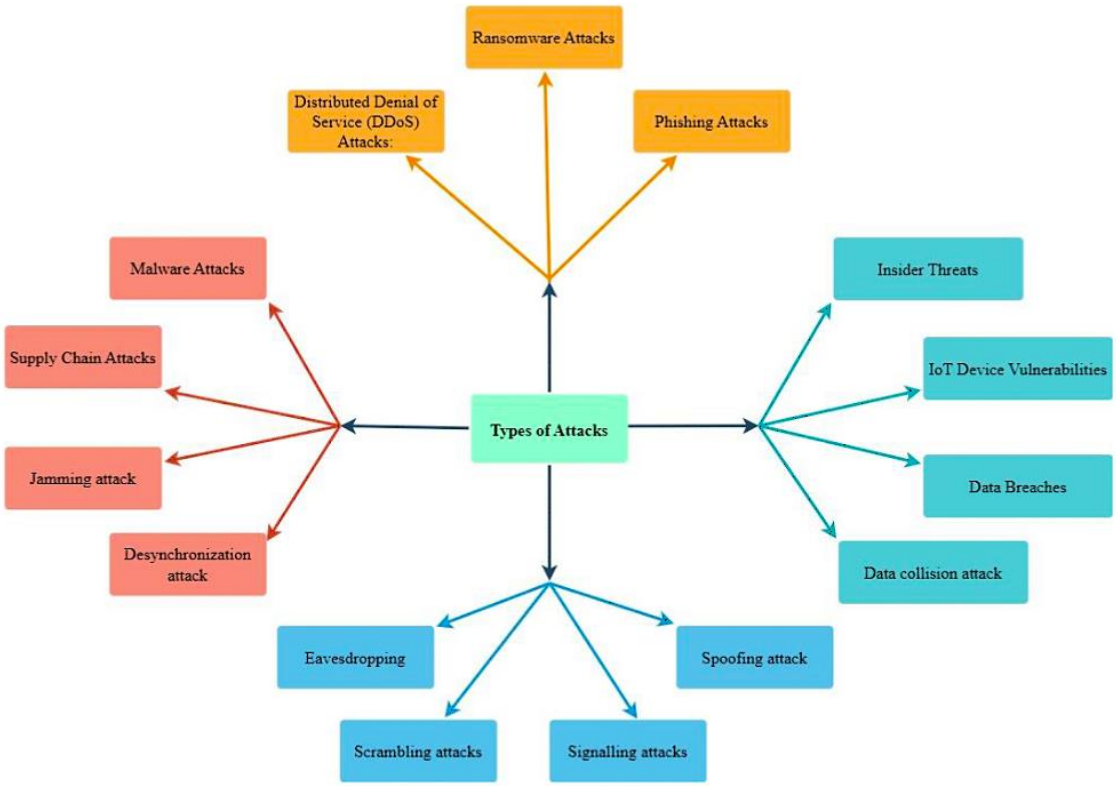


Fig. 5. Types of cyber security attacks on IoT devices. [18]

ISSN 2394 - 7780

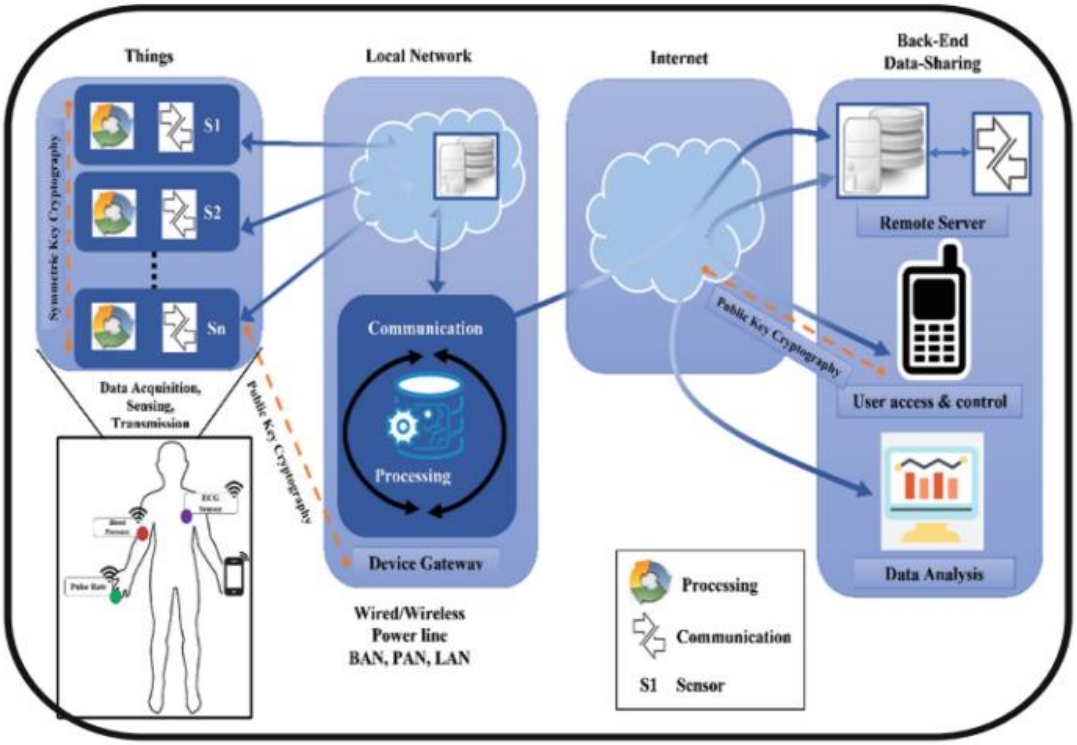


Fig. 5. A secured architecture for HIoT system [6]

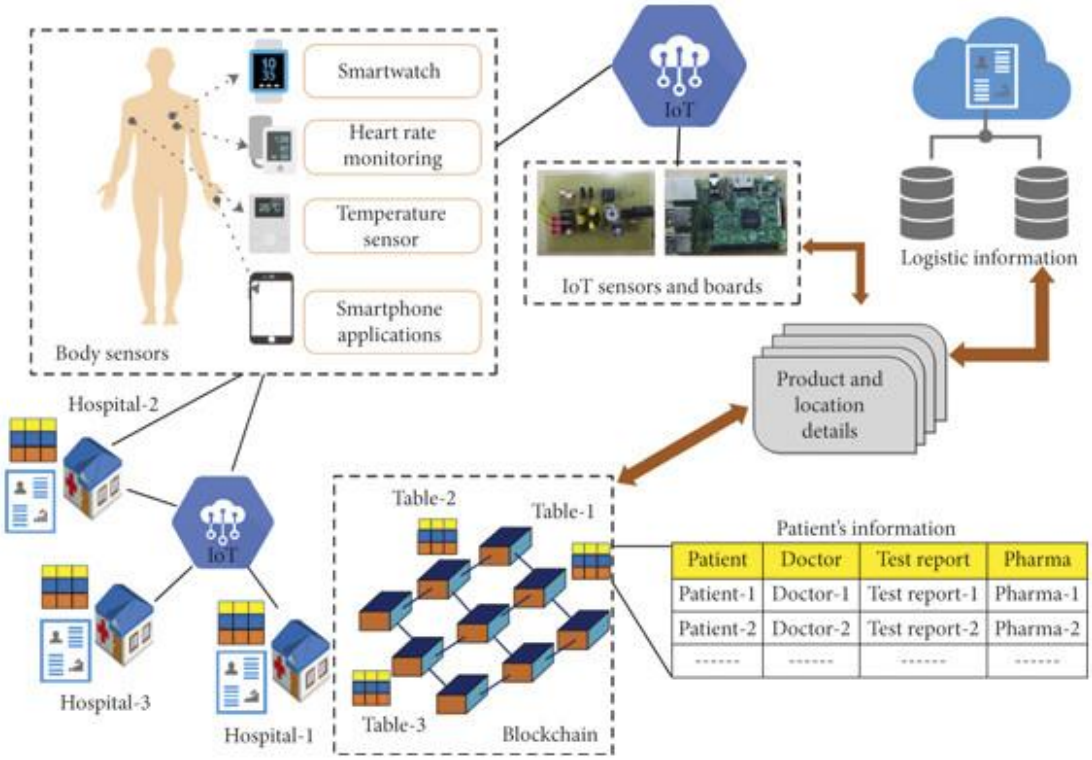


Fig. 6. An architecture of HIoT system with the inclusion of blockchain [2]

ACKNOWLEDGMENTS
I would like to express my deepest gratitude to all those who have supported me throughout the course of this research.

First and foremost, I would like to thank my mentors, Prof. Wilson Rao, for their invaluable guidance, expertise, and encouragement. Their constant support and insightful feedback were crucial to the completion of this study.

ISSN 2394 - 7780

I would like to acknowledge the contributions of my peers for their support, insights and constructive feedback. Your assistance have been proven to be a great asset.

Finally, I am forever indebted to my family and friends for their unwavering love, patience, and moral support throughout this journey. Their encouragement and belief in me have been a source of strength, especially during challenging times.

REFERENCES

- Al-kahtani, Mohammad S., Faheem Khan, and Whangbo Taekeun. 2022. "Application of Internet of Things and Sensors in Healthcare" Sensors 22, no. 15: 5738. doi: 10.3390/s22155738
- Pradhan, Bikash, Bhattacharyya, Saugat, Pal, Kunal, IoT-Based Applications in Healthcare Devices, Journal of Healthcare Engineering, 2021, 6632599, 18 pages, 2021. doi.org/10.1155/2021/6632599
- Kumar, Mohit, Ashwani Kumar, Sahil Verma, Pronaya Bhattacharya, Deepak Ghimire, Seong-heum Kim, and A. S. M. Sanwar Hosen. 2023. "Healthcare Internet of Things (H-IoT): Current Trends, Future Prospects, Applications, Challenges, and Security Issues" Electronics 12, no. 9: 2050. doi.org/10.3390/electronics12092050
- Á. MacDermott, P. Kendrick, I. Idowu, M. Ashall and Q. Shi, "Securing Things in the Healthcare Internet of Things," 2019 Global IoT Summit (GIoTS), Aarhus, Denmark, 2019, pp. 1-6, doi: 10.1109/GIOTS.2019.8766383.
- Chattopadhyay, A.K., Nag, A., Ghosh, D., Chanda, K. (2019). A Secure Framework for IoT-Based Healthcare System. In: Chakraborty, M., Chakrabarti, S., Balas, V., Mandal, J. (eds) Proceedings of International Ethical Hacking Conference 2018. Advances in Intelligent Systems and Computing, vol 811. Springer, Singapore. doi.org/10.1007/978-981-13-1544-2_31
- Vithya Vijayalakshmi, A., Arockiam, L. (2020). A Secured Architecture for IoT Healthcare System. In: Pandian, A.P., Senjyu, T., Islam, S.M.S., Wang, H. (eds) Proceeding of the International Conference on Computer Networks, Big Data and IoT (ICCBI - 2018). ICCBI 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 31. Springer, Cham. doi.org/10.1007/978-3-030-24643-3_106
- Yury Shamrei, 2024's Top IoT Devices Transforming the Healthcare Landscape, Sumatosoft, https://sumatosoft.com/blog/top-iot-devices-transforming-the-healthcare-landscape
- Alex Husar, IoT Security: 5 cyber-attacks caused by IoT security vulnerabilities. https://www.cm-alliance.com/cybersecurity-blog/iot-security-5-cyber-attacks-caused-by-iot-security-vulnerabilities
- Mostafa Haghi Kashani, Mona Madanipour, Mohammad Nikravan, Parvaneh Asghari, Ebrahim Mahdipour, A systematic review of IoT in healthcare: Applications, techniques, and trends, Journal of Network and Computer Applications, Volume 192, 2021, 103164, ISSN 1084-8045, doi.org/10.1016/j.jnca.2021.103164.
- Marques, Gonçalo, Rui Pitarma, Nuno M. Garcia, and Nuno Pombo. 2019. "Internet of Things Architectures, Technologies, Applications, Challenges, and Future Directions for Enhanced Living Environments and Healthcare Systems: A Review" Electronics 8, no. 10: 1081. doi.org/10.3390/electronics8101081
- Abdi, Isse, "IOT Devices in Healthcare: Vulnerabilities, Threats and Mitigations" (2023). Culminating Projects in Information Assurance. 139. repository.stcloudstate.edu/msia_etds/1394
- S. B. Baker, W. Xiang and I. Atkinson, "Internet of Things for Smart Healthcare: Technologies, Challenges, and Opportunities," in IEEE Access, vol. 5, pp. 26521-26544, 2017, doi: 10.1109/ACCESS.2017.2775180.
- Y. Yang, L. Wu, G. Yin, L. Li and H. Zhao, "A Survey on Security and Privacy Issues in Internet-of-Things," in IEEE Internet of Things Journal, vol. 4, no. 5, pp. 1250-1258, Oct. 2017, doi: 10.1109/JIOT.2017.2694844.
- Sun, Wencheng, Cai, Zhiping, Li, Yangyang, Liu, Fang, Fang, Shengqun, Wang, Guoyan, Security and Privacy in the Medical Internet of Things: A Review, Security and Communication Networks, 2018, 5978636, 9 pages, 2018. doi.org/10.1155/2018/5978636

ISSN 2394 - 7780

- Akshay Parihar, Jigna B. Prajapati, Bhupendra G. Prajapati, Binti Trambadiya, Arti Thakkar, Pinalkumar Engineer, Role of IOT in healthcare: Applications, security & privacy concerns, Intelligent Pharmacy,2024,ISSN2949-866X, doi.org/10.1016/j.ipha.2024.01.003.
- Ibrahim Sadek, Josué Codjo, Shafiq Ul Rehman, Bessam Abdulrazak, Security and privacy in the internet of things healthcare systems: Toward a robust solution in real-life deployment, Computer Methods and Programs in Biomedicine Update, Volume 2, 2022, 100071, ISSN 2666-9900, doi.org/10.1016/j.cmpbup.2022.100071.
- Bala, Indu, Irfan Pindoo, Maad M. Mijwil, Mostafa Abotaleb, and Wang Yundong. "Ensuring Security and Privacy in Healthcare Systems: A Review Exploring Challenges, Solutions, Future Trends, and the Practical Applications of Artificial Intelligence". Jordan Medical Journal 58, no. 3 (July 15, 2024). Accessed October 30, 2024. jjournals.ju.edu.jo/index.php/JMJ/article/view/2527.
- Vishwasrao Salunkhe, Abhishek Tangudu, Chandrasekhara Mokkapati, Prof.(Dr.) Punit Goel, and Anshika Aggarwal. "Advanced Encryption Techniques in Healthcare IoT: Securing Patient Data in Connected Medical Devices". Modern Dynamics: Mathematical Progressions 1, no. 2 (August 30, 2024): 224–247. Accessed October 30,2024. mathematics.moderndynamics.in/index.php/mdmp/article/view/22.

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

THE ROLE OF AI IN DETECTING AND PREVENTING CYBERCRIME THROUGH BEHAVIORAL ANALYSIS

[1]Prasad Anand Labade and [2]Tejashree Parab
Department of MSc Big Data Analytics, Jai Hind College (Empowered Autonomous)
Mumbai, India

ABSTRACT
The rise in cybercrime has posed significant challenges to organizations and individuals, with traditional security measures often failing to keep pace with evolving threats. Artificial Intelligence (AI) has emerged as a crucial tool in detecting and preventing cybercrime, particularly through behavioural analysis. This paper explores AI's role in enhancing cybersecurity by leveraging behavioural data to identify anomalies and patterns indicative of malicious activity. By utilizing machine learning algorithms, AI systems can establish baseline behaviours for users and systems, allowing them to flag deviations in real-time that could signal potential threats such as insider attacks, fraud, or malware intrusions. Key techniques such as anomaly detection, predictive modeling, and User and Entity Behaviour Analytics (UEBA) are examined, highlighting how AI-driven systems analyze vast amounts of data to detect threats that may go unnoticed by conventional methods. These systems can continuously learn from new data, improving their ability to identify novel cyberattacks. Additionally, AI-powered solutions provide the ability to respond to threats more quickly and efficiently than human analysts, reducing response times and mitigating potential damage. While AI brings numerous benefits to cybersecurity, its implementation also presents challenges. Issues such as data privacy, the potential for false positives, and the need for constant updates to AI models to adapt to new threats are explored. Nevertheless, the integration of AI into cybersecurity, particularly through behavioral analysis, offers significant promise in addressing the growing complexity and volume of cyberattacks, helping organizations preemptively counteract cybercriminal activities before they can cause harm.

Keywords—*Artificial Intelligence (AI), Cybercrime, Behavioural Analysis, Machine Learning, Anomaly Detection, User and Entity Behaviour Analytics (UEBA), Predictive Modeling, Cybersecurity, Insider Threats, Fraud Detection, Malware Detection, Data Privacy, False Positives, Real-Time Threat Detection. This Is a Level 1 Heading*

I. INTRODUCTION TO CYBERCRIME AND CYBERSECURITY
The rapid growth of online platforms and applications has created new opportunities for cybercriminals, exposing individuals to various online threats. While public awareness campaigns play a crucial role in preventing cybercrime, the adaptable and evolving tactics used by cybercriminals make it difficult for law enforcement to stay ahead. With the number of mobile devices worldwide expected to reach over 18 billion by 2025 (Statista, 2021), the landscape of cyber threats is only becoming more complex, presenting challenges for crime prevention through awareness alone. Though cybercriminals may share motivations with those committing traditional crimes, the technical complexities of cybercrime create additional hurdles for law enforcement.[2] As a result, prosecuting cybercrimes is significantly more challenging than traditional crimes, as evidenced by lower cybercrime prosecution rates (Peters and Jordan, 2019). This article examines the technological aspects of cybercrime and why they present unique challenges for law enforcement. Reports from crime databases (NCRB, 2021; FBI, 2021) reveal a significant gap between the number of cybercrimes reported and those that result in successful prosecution.

ISSN 2394 - 7780

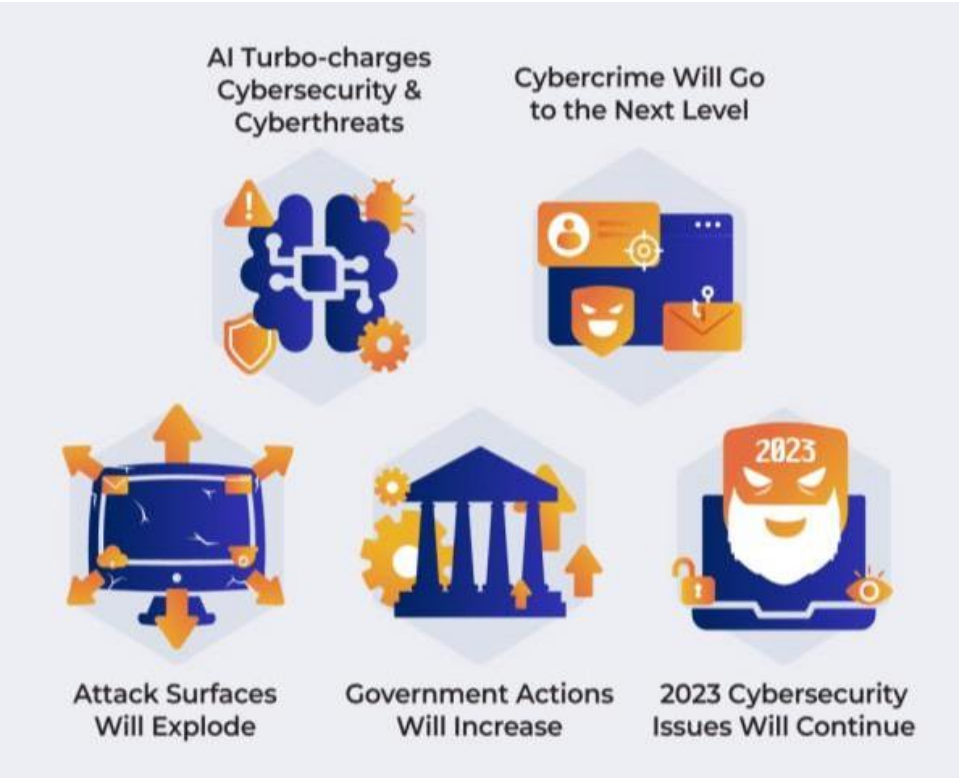


Fig.1 Major Cybersecurity Trends

This disparity highlights the need for more sophisticated tools, procedures, and knowledge among law enforcement agencies to address cybercrime effectively.([1]) To bridge this gap, there is a pressing need to enhance investigative capabilities specifically tailored for the digital landscape. The literature review focuses on exploring the varied tactics and methodologies used by cybercriminals, offering insights into the structure, characteristics, and techniques behind different cyber offenses. Additionally, it reviews the current approaches law enforcement uses to combat these offenses and the obstacles they face. Cybercriminals often divide their operations into distinct stages, allowing them to hide their identities and actions while complicating the work of investigators. This segmentation also enables criminals to use resources more effectively, making it harder for law enforcement to detect and disrupt their activities. Addressing cybercrime requires a coordinated approach involving specialized training, tools, and knowledge in digital forensics. Traditional methods are often inadequate in the virtual space, where anonymity and jurisdictional issues pose significant barriers. Successful resolution of cybercrime cases increasingly depends on collaboration between law enforcement, private organizations, and international partners. By combining a solid understanding of cyber threats, fostering partnerships, and strengthening technological capabilities, efforts to combat cybercrime can be more effective.([2])

The rapid advancement of computing technology and the internet has transformed daily life, offering significant conveniences and efficiencies. However, it has also introduced complex challenges, particularly in the form of cybercrime. Traditional crimes like theft and fraud have evolved, taking on new, digital forms enabled by technology. As technology advances, so do the methods and strategies used by cybercriminals, making it challenging to address the continuously evolving threats. The rise in cybercrimes is partly due to the accessibility that technology offers, allowing individuals to commit offenses with ease and speed from any location.[[5] Furthermore, technology has enabled these crimes to cross geographical borders, complicating efforts to track, prevent, and apprehend cybercriminals. Information technology serves as both a target and a tool for criminal activity. Devices such as computers, mobile phones, and other digital systems, originally designed to benefit society, have become susceptible to misuse. Cybercrime encompasses a wide range of illegal activities involving these devices, from unauthorized system access and data breaches to intellectual property theft, financial fraud, and digital espionage. These offenses, also known as “digital crimes,” “computer crimes,” or “internet crimes,” often exploit vulnerabilities in digital systems to steal data, disrupt services, and manipulate online transactions.([4])

Defining cybercrime is challenging, as it encompasses a broad spectrum of activities that exploit technology in varied ways. A common interpretation defines cybercrime as any illegal act facilitated by or committed using a computer, network, or related device. This can include activities where the computer acts as the instrument,

ISSN 2394 - 7780

target, or accomplice of the crime. As digital data and online interactions expand, the opportunities for cybercrimes grow, blurring distinctions across global boundaries and fostering a "virtual world" where crime can thrive in ways that parallel, yet differ from, real-world offenses. Experts such as Brenner (2010) suggest that many cybercrimes represent an adaptation of traditional criminal activities to the digital space, utilizing cyberspace as a tool to execute familiar crimes with novel tactics. With the continuous growth of digital data and online interactions, the prevalence and impact of cybercrime remain a significant challenge, underscoring the need for innovative, adaptive solutions in cybersecurity and law enforcement[2]

II. AI IN CYBERSECURITY: AN OVERVIEW

Artificial Intelligence (AI) has become an essential field within computer science, focusing on creating systems that can perform tasks typically requiring human intelligence. In cybersecurity, AI techniques like Machine Learning (ML), Deep Learning (DL), Natural Language Processing (NLP), and Knowledge Representation and Reasoning (KRR) have found extensive applications. These AI methods can help address numerous cybersecurity challenges, such as intrusion detection, anomaly detection, fraud prevention, cyber-attack prediction, and access control management. The goal of integrating AI into cybersecurity is to enable intelligent, automated systems that enhance security management by making data-driven decisions.[4]

2.1) In exploring AI's role in cybersecurity, it's essential to understand the CIA triad—Confidentiality, Integrity, and Availability. These principles guide information security policies across organizations:

- **Confidentiality:** Focuses on keeping data accessible only to authorized users, protecting sensitive information from unauthorized access. Common threats to confidentiality include data breaches, targeting databases, or application servers.
- **Integrity:** Ensures that data remains accurate and unaltered by unauthorized parties. Integrity attacks may involve unauthorized changes to financial data or attempts to compromise an organization's credibility.
- **Availability:** Ensures that information and systems are accessible to authorized users whenever needed. Attacks on availability include denial-of-service (DoS) incidents, which aim to disrupt service accessibility.

The CIA triad serves as a foundation for security policy and provides a framework for developing AI-driven models that can mitigate these threats through adaptive and intelligent responses. With AI's potential to automate threat detection and enhance decision-making, the study of AI in cybersecurity not only offers current solutions but also opens pathways for future research. In particular, the development of AI models for cybersecurity aims to support security teams with tools for proactive threat management, contributing to a more resilient cyber defense infrastructure.[1]

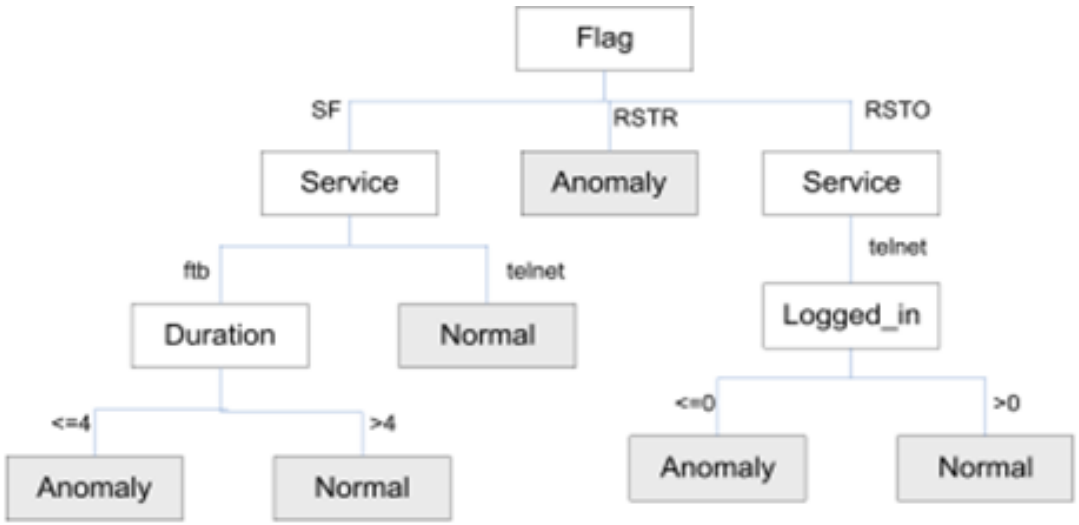


Fig.2 An example of detecting cyber-anomalies based on a decision tree-based machine learning model

ISSN 2394 - 7780

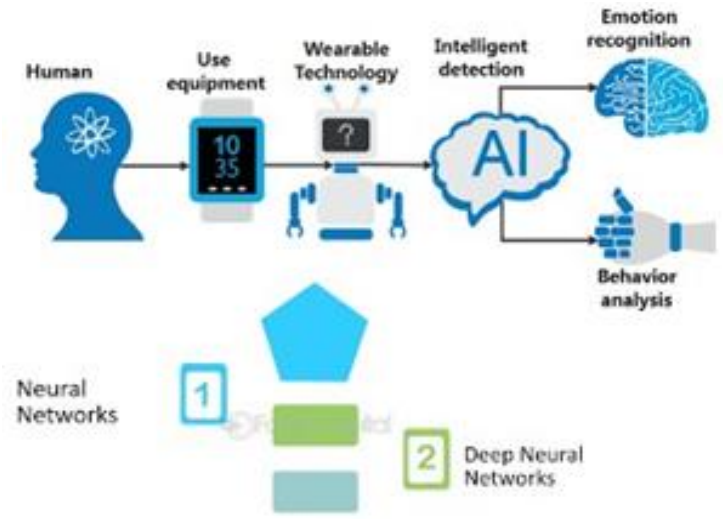


Fig. 3: AI Techniques used for Behavioural Analysis

Behavioural analysis in cybersecurity focuses on monitoring and understanding the actions and patterns of users, devices, and systems to identify potential threats. AI-driven techniques bring efficiency and accuracy to this domain by automating the detection of unusual behaviours that might indicate security risks. Below is an expanded list of AI techniques, including those already mentioned, with additional methods relevant to behavioural analysis in cybersecurity:[4][6]

Datasets and Benchmarking in AI-Driven Behavioural Analysis

The success of AI applications in behavioural analysis for cybercrime detection heavily depends on the availability of high- quality datasets. These datasets serve as the foundation for training, validating, and benchmarking machine learning models. They provide a diverse range of scenarios that represent both normal and anomalous behaviours. Below is a detailed exploration of the most commonly used datasets in this domain:

a. CICIDS Dataset (Canadian Institute for Cybersecurity Intrusion Detection Dataset)

The CICIDS dataset, developed by the Canadian Institute for Cybersecurity, is widely used in intrusion detection research. It simulates real-world network traffic, including normal activities and various types of cyberattacks. The dataset includes packet-level features, providing a detailed representation of network behaviour.

Features: Over 80 network traffic features, such as packet size, time duration, protocol types, and flow bytes. Includes labelled data representing specific attack types, such as DoS, DDoS, brute force, and infiltration attempts.

Applications: Ideal for building and testing supervised learning models due to its detailed labels. Used in anomaly detection and real-time monitoring systems.[10]

b. DARPA Intrusion Detection Dataset

The DARPA dataset, created by the Défense Advanced Research Projects Agency, is one of the earliest benchmarks for intrusion detection systems. It contains network traffic data collected in a controlled environment, simulating attacks and normal usage scenarios.

Features: Captures a wide range of activities, including probing, denial-of-service (DoS), remote-to-local (R2L), and user-to-root (U2R) attacks. Includes both raw network packet data and session- level summaries.

Applications: Widely used for initial research and comparative studies in intrusion detection. Suitable for both supervised and unsupervised learning approaches.[10]

c. Benchmarking Considerations

Benchmarking datasets is critical for evaluating the performance of AI models in behavioural analysis. Common metrics used in conjunction with these datasets include accuracy, precision, recall, F1 score, and false-positive rates. Ensuring a balance between synthetic and real-world data is crucial to developing robust models capable of generalizing across diverse scenarios.

In conclusion, datasets like CICIDS and DARPA, combined with custom enterprise logs, form the backbone of

ISSN 2394 - 7780

research and development in AI-powered behavioural analysis. They enable researchers to build models that are not only accurate but also adaptable to the ever-evolving landscape of cybercrime for the purpose of Analysis.

III. REAL-TIME CYBERCRIME RESPONSE SYSTEMS

Real-time cybercrime response systems leverage the power of Artificial Intelligence (AI) to detect, analyse, and mitigate cyber threats as they occur. These systems provide organizations with a proactive approach to cybersecurity, enabling swift containment and recovery from malicious activities. The implementation of AI in real-time response mechanisms is transformative, offering unparalleled speed, accuracy, and scalability in combating cybercrimes.

Fig. 4 Application of AI in Cybersecurity[4]

TECHNOLOGICAL IMPLEMENTATION

AI-Orchestrated Incident Response: AI-orchestrated incident response refers to the use of AI-driven algorithms and automated workflows to handle cyber incidents. These systems rely on decision-tree-based algorithms that dynamically determine the best course of action based on real-time threat intelligence and contextual data.

- **Detection:** AI models identify anomalies or known attack patterns.
- **Assessment:** Context-aware systems analyze the severity and scope of the threat.
- **Containment:** The system isolates affected devices or networks to prevent lateral movement.
- **Mitigation and Recovery:** Automated playbooks execute predefined actions such as patch deployment, process termination, or data restoration.
- **Implementation:** AI-based systems in Security Orchestration, Automation, and Response (SOAR) platforms dynamically update workflows based on evolving threats, ensuring adaptive security.

Adversarial AI Défense: Adversarial AI Défense employs AI systems to anticipate and counteract sophisticated attacks, including those crafted to bypass traditional defenses. These defenses simulate adversarial scenarios to identify system vulnerabilities and prepare for evolving threats.

- **Simulation and Testing:** AI systems simulate real-world attack vectors, such as adversarial inputs designed to trick detection models.
- **Self-Learning Models:** Using reinforcement learning to improve défense mechanisms against newly discovered attack patterns.
- **Collaborative AI:** Integrating AI-driven threat intelligence across platforms to share insights and defensive strategies.
- **Applications:** Protecting critical infrastructures such as financial networks, healthcare systems, and government databases. Training detection models to identify malicious payloads disguised within legitimate processes.

ISSN 2394 - 7780

IV. RESEARCH INSIGHTS ON AI-BASED REAL-TIME CYBERCRIME RESPONSE SYSTEMS

Advancements in AI-powered cybersecurity solutions have established their potential in significantly enhancing threat mitigation capabilities and reducing operational challenges. Recent studies emphasize how these systems outperform traditional methods in multiple dimensions, including speed, accuracy, and cost-effectiveness. Below are detailed insights that can be included in a research paper.

Reduction in Containment Time

One of the most significant achievements of AI-driven response systems is their ability to dramatically reduce the time required to contain cyber incidents. Studies conducted in 2023 show that organizations deploying AI-based solutions achieved a 70% reduction in containment time compared to traditional manual responses.

Key Factors Contributing to the Improvement:

- **Automated Threat Detection:** AI models, trained on vast datasets, can rapidly identify suspicious activities such as unauthorized data transfers, unusual access patterns, or anomalous network traffic. This early detection enables faster responses.
- **Real-Time Decision Making:** AI systems utilize advanced algorithms, such as reinforcement learning and decision trees, to evaluate threat scenarios in real time and execute containment actions without human intervention.[09]
- **Seamless Integration:** AI-driven platforms integrate with existing cybersecurity tools like firewalls, Endpoint Detection and Response (EDR) systems, and cloud-based monitoring tools, creating a cohesive and efficient Défense mechanism.
- **Real Time Examples: Financial Institutions:** AI tools deployed in banking environments have shown the ability to isolate compromised accounts or systems within seconds of detecting unusual activities, preventing significant financial losses. **Critical Infrastructures**: In energy and utility sectors, AI-powered systems can quickly shut down affected nodes in a network to prevent cascading failures.[09]

Improved Threat Mitigation Accuracy

The integration of adversarial AI Défense mechanisms has significantly enhanced the accuracy of threat mitigation efforts, especially in combating sophisticated attacks such as polymorphic malware, phishing, and Advanced Persistent Threats (APTs).

Key Achievements:

Countering Evasive Techniques: Adversarial AI models are designed to identify patterns in attacks that are crafted to bypass traditional security systems. For instance, obfuscated malware, which changes its appearance to evade signature-based detection, is effectively identified by AI algorithms that focus on behavioural analysis rather than static signatures.[11]

High Detection Rates: Research studies indicate that adversarial AI systems achieve a 95% detection rate for obfuscated malware, far surpassing the capabilities of conventional methods. These systems analyze dynamic features like execution patterns, system calls, and network behaviour to detect anomalies.[11]

Real-World Application: In a 2022 case study, an AI-powered system successfully identified and neutralized a sophisticated phishing campaign targeting employees of a multinational corporation. By analyzing email communication patterns, the AI model flagged malicious emails with a 98% accuracy rate.[12]

Cost Efficiency

AI-based systems have proven to be cost-effective by automating processes that traditionally require extensive human resources and continuous monitoring. This reduces the overall operational burden while maintaining a high level of security.

Key Benefits:

- **Reduction in Manual Monitoring:** AI systems monitor network traffic, user behaviour, and application activities around the clock, eliminating the need for large teams of analysts to perform these tasks manually**.**
- **Proactive Défense Mechanisms:** By automating threat responses, such as isolating compromised systems or deploying security patches, organizations save time and resources, enabling security teams to focus on strategic tasks.
- **Minimized Damage Costs:** The rapid containment of cyber threats prevents large-scale damage, such as data breaches or operational downtimes, which can result in significant financial and reputational losses.

ISSN 2394 - 7780

Research Findings: A 2023 survey of organizations using AI- driven cybersecurity systems reported:

- A 40% reduction in operational costs associated with security monitoring.
- An average savings of $3.8 million annually in damage mitigation expenses.

I. Future Implications

While the current research highlights the transformative potential of AI in real-time cybercrime response, continuous advancements are essential to address emerging challenges**:**

- **Ethical AI:** Ensuring fairness and transparency in AI decision- making processes to foster trust.
- **Improved Collaboration:** Establishing global frameworks for sharing AI-driven threat intelligence among organizations and governments.
- **Scalability:** Adapting AI solutions to protect growing and interconnected ecosystems in smart cities, IoT, and critical infrastructures.

V) CASE STUDIES

U.S. Government Agency's Application Classification and Network Attack Detection [10]

A major AI centre within the U.S. government adopted Snorkel Flow to accelerate the development of AI/ML applications for cybersecurity. The platform facilitated programmatic labelling, enabling efficient application classification and network attack detection through behavioural analysis.

Boardriders' Fraud Detection with AI[10]

Boardriders, a global retail company, implemented AI-driven behavioural analysis to combat fraud. The system analyses transaction patterns and customer behaviours to identify anomalies that could indicate fraudulent activities. This proactive approach has significantly reduced fraud incidents and financial losses.

CrowdStrike's AI-Powered Behavioural Analysis[9]

CrowdStrike utilizes AI to enhance behavioural analysis in cybersecurity. Their approach involves data collection, AI training, pattern recognition, and anomaly detection. By continuously learning from data, the system can detect and respond to threats in real-time, adapting to the evolving threat landscape.

Microsoft's AI-Powered Threat Protection[11]

Microsoft employs AI to bolster its cybersecurity defenses across platforms like Office 365 and Azure. The AI system monitors user behaviours to establish baselines and detect deviations that may signify threats. For instance, unusual login patterns or atypical data access can trigger alerts, enabling swift responses to potential breaches.

v. CONCLUSIONS

Artificial Intelligence (AI) has revolutionized the field of cybersecurity, offering advanced tools and techniques to detect, prevent, and respond to cyber threats. By utilizing methods such as machine learning, deep learning, and behavioural analysis, AI enhances the ability to identify potential risks and automate responses with improved accuracy and efficiency. These technologies have become crucial in managing the growing complexity of modern cyberattacks.

Despite its advantages, the adoption of AI in cybersecurity comes with challenges. Dependence on high-quality training data, the risk of false positives or negatives, ethical considerations surrounding privacy, and potential vulnerabilities to adversarial attacks are significant concerns. These limitations emphasize the need for combining AI systems with human expertise, ensuring continuous improvement, and adhering to strict ethical standards.

As cyber threats continue to evolve, collaboration among organizations, governments, and researchers is essential to address emerging risks effectively. A balanced approach that leverages AI’s capabilities while addressing its limitations can create a resilient and adaptive cybersecurity framework.

These insights form the foundation for continued research and innovation in the field, ensuring a secure digital future.

In summary, AI offers immense potential to transform cybersecurity practices. However, its success lies in integrating technological innovation with responsible usage, ethical data handling, and global cooperation. This research highlights the opportunities and challenges of AI in cybersecurity, encouraging further exploration and development to secure the digital world.

ISSN 2394 - 7780

The integration of AI into real-time cybercrime response systems represents a paradigm shift in cybersecurity. By significantly reducing containment times, improving detection accuracy, and lowering operational costs, AI-based systems provide a robust and scalable solution to combat the growing threat of cybercrimes.

REFERENCES

1] *Sarker, I. H. (2021). AI-driven cybersecurity: An overview, security intelligence modeling and research directions. SN Computer Science, 2(3), Article 137. https://doi.org/10.1007/s42979-021-00557-0 SpringerLink*

2] *Dilek, S., Çakır, H., & Aydın, M. (2015). Applications of artificial intelligence techniques to combating cyber crimes: A review. International Journal of Artificial Intelligence & Applications (IJAIA), 6(1), 21–39. https://doi.org/10.5121/ijaia.2015.6102 arXiv*

3] *Ansari, M. F., Dash, B., Sharma, P., & Yathiraju, N. (2022). The impact and limitations of artificial intelligence in cybersecurity: A literature review. International Journal of Advanced Research in Computer and Communication Engineering, 11(9), 81–88. https://doi.org/10.17148/IJARCCE.2022.11912*

4] *Moustafa, A. A., Bello, A., & Maurushat, A. (2021). The role of user behaviour in improving cyber security management. Frontiers in Psychology, 12, Article 561011. https://doi.org/10.3389/fpsyg.2021.561011*

5] *Mashiane, T., & Kritzinger, E. (2021). Identifying behavioral constructs in relation to user cybersecurity behavior. Eurasian Journal of Social Sciences, 9(2), 98–122. https://doi.org/10.15604/ejss.2021.09.02.004*

6] *Guembe, B., Azeta, A., Misra, S., Osamor, V. C., Fernandez-Sanz, L., & Pospelova, V. (2022). The emerging threat of AI-driven cyber attacks: A review. Applied Artificial Intelligence, 36(1), 1–18. https://doi.org/10.1080/08839514.2022.2037254*

7] *Salem, A. H., Azzam, S. M., Emam, O. E., & Abohany, A. A. (2024). Advancing cybersecurity: A comprehensive review of AI-driven detection techniques. Journal of Big Data, 11, Article 105. https://doi.org/10.1186/s40537-024-00957-y*

8] *Hemberg, E., & O'Reilly, U.-M. (2021). Using a collated cybersecurity dataset for machine learning and artificial intelligence. arXiv preprint arXiv:2108.02618. https://arxiv.org/abs/2108.02618*

9] *Zhang, Y., & Wang, J. (2021). Artificial intelligence in cyber security: Research advances and challenges. Artificial Intelligence Review, 54(3), 2043–2081. https://doi.org/10.1007/s10462-021-09976-0*

10] *Sarker, I. H. (2024). AI-driven cybersecurity and threat intelligence: Cyber automation, intelligent decision-making and explainability. Springer.https://doi.org/10.1007/978-3-031-54497*

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

ENHANCING BRAIN TUMOR DETECTION USING MRI WITH K-FOLD CROSS-VALIDATION

[1]Purvi Ravikumar Singh, [2]Sunita Jena and [3]Niloufer Kotwal
[1]Student and [2]Assistant Professor, MSc. Big Data Analytics Department, Jai Hind College, Mumbai
[3]Assistant Professor, Life Science Department, Jai Hind College, Mumbai

ABSTRACT
A brain tumor, which happens when abnormal brain cells grow quickly, is a serious health risk for adults because it can lead to major organ problems and even be life-threatening. Manual segmentation of tumors from brain MRI is time-consuming and error-prone. Early detection allows doctors to get involved before extreme harm, bringing down permanent damage. In this paper, we will see the detection of tumors in Magnetic Resonance Imaging (MRI) where the Deep Learning model will be trained to detect whether the tumor is present. The dataset used in this study is sourced from Kaggle and contains two classes yes and no for brain tumor presence or absence. The tools and technologies used are Keras for developing the neural networks, scikit-learn used in data splitting and TensorFlow for managing data. Also, various libraries are used for image processing work, namely NumPy, OpenCV (cv2), etc. The model's performance is evaluated using an average of accuracy and precision as the primary metric. We aim to achieve accurate tumor detection by enhancing the model using K-fold cross-validation by converting the binary classification problem into categorical classification problem. We have trained the model using categorical cross-entropy for a binary classification problem, where the target variable has been converted into categorical format. The Saved model is used to predict the previously unseen data.

Keywords_ Deep Learning CNN, K-fold Crossvalidation, Magnetic Resonance Imaging (MRI), Brain Tumors.

I. INTRODUCTION
In India, tumors like brain tumor rank as the second most common cancer among children and young adults. Ignoring early symptoms delays medical attention, risking tumor progression. Recognizing signs promptly is crucial to increasing survival chances. To understand the function of imaging, it is necessary to first understand what brain tumors are. In the brain, these growths happen when cells proliferate out of control, forming dense bulk. While malignant tumors pose serious health risks, benign tumors only cause minor harm. The capacity of MRI to identify brain tumors with measurable precision is one of its strongest points. It gives doctors high-resolution images that help them identify the location, size and the nature of tumor [1]. Detecting tumor as soon as possible allows doctors to correct things before extreme harm, bringing down the permanent damage risk. Magnetic Resonance Imaging of brains are difficult for brain tumor to be detected by humans manually. Convolutional Neural Networks (CNNs): The method used in this study is Convolutional Neural Networks. CNNs are designed to automatically learn spatial hierarchies of features from images, allowing them particularly effective for analyzing MRI scans of the brain. This method let us create the automated detection of tumors, significantly improving diagnostic accuracy when compared to traditional methods [2]. Treatments for brain tumor mainly depend upon the correct diagnosis and could be time-consuming as well as painful. Brain tumors are a major health concern, being a top cause of death globally. The survival rate for adults diagnosed with brain cancer is alarmingly low, with only 12% surviving beyond five years. This highlights the urgent need for effective diagnostic tools in the medical field [3].

II. LITERATURE REVIEW
The model was trained on three distinct brain MRI datasets from the Kaggle website, displaying its robustness and generalization capabilities across various data sources. The optimized CNN model achieved outstanding performance scores, with an average accuracy, precision, recall, and F1-score of 97% [3]. In this research [4] two machine learning based tumor detection systems were suggested and compared. MLP gives more accuracy and requires more time to build the model. Naïve bayes takes less time and produces less accurate model.

From the paper [5] MRI images of brain tumors don't clearly show where exactly the tumor is located. To find the tumor's exact position in the MRI images, they have used techniques like preprocessing, segmentation, morphological operations, and subtraction. These methods help create the tumor's specific shape in the MRI image, making it possible to accurately detect the brain tumor. From the paper [6] The classification results show whether the brain images are normal or have a tumor. A method called CNN, which uses layers to process data, is used for this classification in Python. they extracted features like depth, width, and height from the images. In this paper they have used gradient descent optimizer. The training accuracy here is 97.5%. They mentioned that losses were less but not in quantitative manner.

ISSN 2394 - 7780

The paper [7] presents a completely automatic approach for classifying brain tumors using deep transfer learning to extract characteristics from MRI images. It provides greater classification accuracy compared to existing approaches and demonstrates robustness with limited training samples. However, challenges remain, including the standalone performance of the transfer model, misclassification of meningioma samples, and overfitting with small datasets. Future research should focus on addressing these issues through data augmentation and further model tuning.

III. MATERIALS AND METHODOLOGY

The dataset used in study is sourced from Kaggle, an open source. The dataset is very small having Brain MRI's with Yes (155 Images) and No (98 Images) separated via folders. The brain tumor images went through several steps to prepare them for classification tasks. Here's a simple overview of what was done:

A. Preprocessing

The colorful RGB images were changed to grayscale, which means they became black and white. This reduced complexity and made processing easier. All the images were resized to a standard size of 120x120 pixels. This ensured that every image was of the same size, that is making the next processing steps more consistent and straightforward. Since the image size affects computational load, memory usage, speed and efficiency. The processed images were stored as arrays and corresponding labels were appended, categorizing them as either "tumor" or "no tumor." Finally, the dataset was organized for further analysis. The target variables are converted to categorical format. It is equally important to standardize the image before training even sets foot on the pre-processing phase. This is done on standard procedure where images are normalized in order to increase the stability of the model and enhance efficiency. Image normalization includes standardizing the pixel values of the image thus removing the impact of intensity as well as contrast image amendment.

Data augmentation is a procedure of increasing the size of any given dataset by applying operations to the current data to enable independent transformations. This is especially so when working with small samples where the idea assists in increasing its ability to generalize by avoiding over-fitting. By generating multiple instances from the original data, including rotating, flipping, scaling, cropping, or adding noise to images, or using synonyms or back translations of words, we are able to provide the man to wider examples hence improving the results of the model.

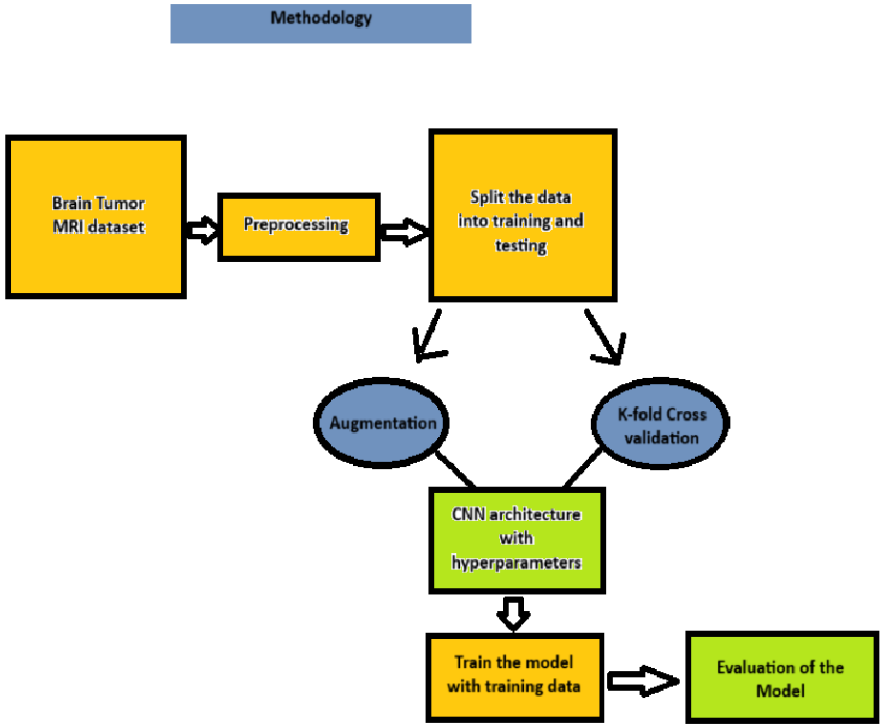


Fig 1. Flow chart of Methodology

B. Splitting

The processed dataset was split into training and testing subsets. The testing size taken as 25% and the training size was 75%. This ratio is general and go with almost all the dataset hence, was taken for splitting.

ISSN 2394 - 7780

C. Architecture of the model

CNNs down-sample data using pooling layers, which lower dimensionality while preserving significant features. This improves the efficiency of computing. Pooling is frequently absent from traditional neural networks. CNN are made especially to handle grid-like information, like pictures, where the spatial correlations between pixels are very important. For sequential data, like time series or natural language, other neural network types, such as Recurrent Neural Networks (RNNs), are more appropriate. Weight or parameter Sharing is used to help decrease the number of parameters and enable the model to acquire translation-invariant features, CNNs apply the same filter (or weights) to several regions of the input image. Fully connected networks, on the other hand, assign distinct weights to each connection.

The loss function is approximated using the gradient descent approach. A scoring function is used to translate a raw image pixel into class scores. The degree to which the induced scores correspond to the ground truth labels defines the loss function, which evaluates the quality of a particular collection of parameters. In order to increase accuracy, it is crucial to compute the loss function when the accuracy is low and the loss function is high; in the opposite case, the accuracy is high. The gradient descent algorithm is computed by taking the derivative of loss function. Once more, assess the gradient to find the gradient of the loss function.

Normalization technique is mostly used on algorithms that re machine learning or deep learning based. It insures that all the images are having similar intensity. This technique improves the performance of the algorithm that works better with normalized data. This is used on our data and an image of MRI before and after normalization has been attached as shown in Fig. 3 and Fig. 4.

A Convolutional Neural Network model has been used that comprises of the max pooling layers of a 2x2 filter come after each of the four layers of convolutional. Then following each convolutional layer, ReLu activations are also applied. In the preprocessing stage, a sequential deep learning model was constructed using Keras. Initially, convolutional and pooling layers were added to take properties out of the images, followed by flattening the output to transition into fully connected layers. The model included activation functions like ReLU and softmax for non-linearity and classification, respectively.

K-fold cross-validation is one of the most popular statistical methods for assessing the effectiveness of a learned model. It means the data available is equally divided into K sets or folds more commonly known as the resampling subsets. The proposed model of the study is trained and tested 5 times, where in each test, one fold is used for testing while others are used for training. This way, the final result obtained for each image is an average of results that we get after 5 iterations hence providing us a better approximation of the model.

For binary classification, use binary cross entropy as loss and sigmoid activation function, similarly for categorical classification, we can use categorical cross entropy as loss function and the softmax activation function. Compiled the model with categorical cross entropy as loss, accuracy metric for evaluation and Adam optimizer for parameter optimization. Selecting hyperparameter is a crucial task which has no formula or rule. It can be determined by running and observing the performance of the model several times. As we know each and every parameter has different impact while training the model like speed and accuracy.

EVALUATION METRICS THAT WERE USED ARE GIVEN AS FOLLOWS

1) **Accuracy:** Accuracy quantifies how frequently the classifier made accurate predictions. The ratio of the number of accurate predictions to the total number of predictions is a way to define accuracy. Indicates the overall correctness of the model, the percentage of accurate forecasts among all the predictions. A higher level of precision means that the model's predictions are more in line with the labels that are actual labels.

ISSN 2394 - 7780

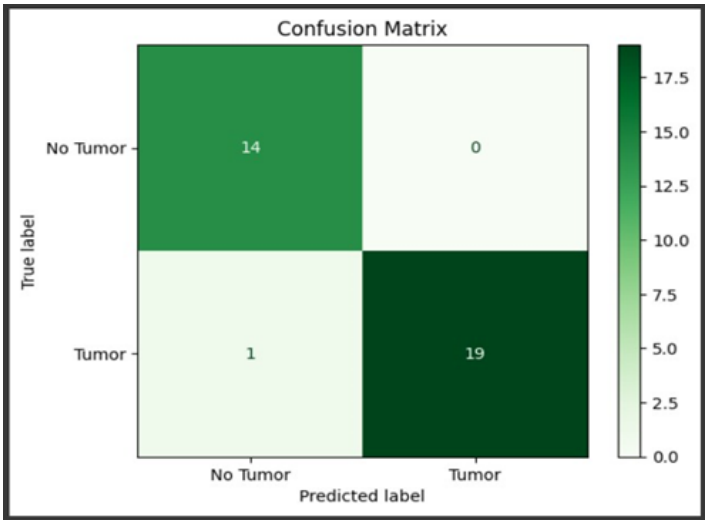


Fig. 2 Confusion matrix of proposed model

2) **Precision:** In machine learning, precision is a performance metric that calculates the accuracy of a model in identifying instances among the total instances it predicted as positive.it is especially useful where the cost of false positives is high. It answers the question: Out of all the instances the model classified as positive, how many are actually positive?

3) **Recall:** It is also known as "The true positive rate", or the amount of all actual positives that were appropriately categorized as positives. It evaluates the model's ability to record all true positive cases; a high recall denotes fewer false negatives.

4) **F1 score:** The F1 score is a statistical metric used to evaluate the accuracy of a model, particularly in classification tasks. It's an alternative to accuracy, which considers overall performance. It Provides a balance between precision and recall, making it especially useful when the data is imbalanced.

Table I Evaluation Metrics and Formulae

Metrics	Formula
Accuracy	$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$
Precision	$Precision = \frac{TP}{TP + FP}$
Recall	$Recall = \frac{TP}{TP + FN}$
F1 score	$F1 = 2 * (precision * recall) / (precision + recall)$

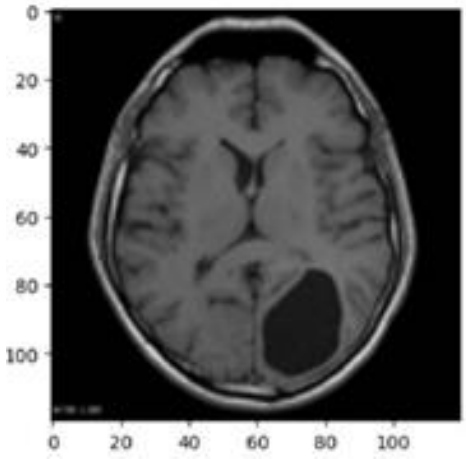


Fig. 3 MRI before normalization

ISSN 2394 - 7780

D. Confusion matrix

A confusion matrix is a table that summarizes the performance of a classification model by showing the relationships between actual labels and predicted labels. It is a powerful tool to evaluate classification algorithms, especially for binary and multi-class classification. The matrix contains four main components shown in Table 1:

1) ***True Positives:*** Accurately predicted cases

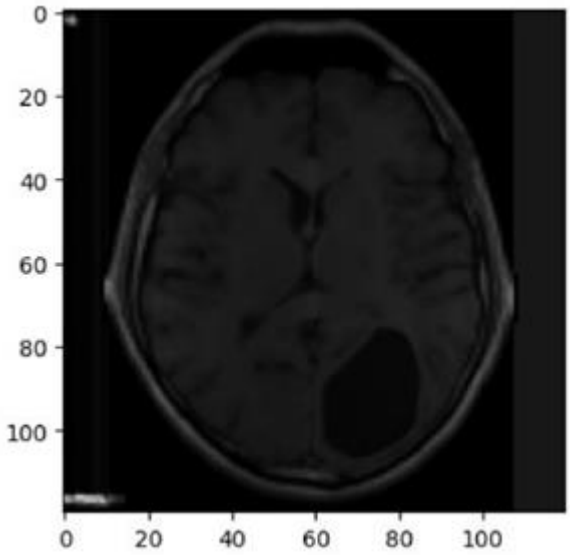


Fig. 4 MRI after normalization

2) ***True Negatives*****:** Accurately predicted non-instances
3) ***False Positives:*** Incorrectly predicted cases (aka type I errors)
4) ***False Negatives*****:** Incorrectly estimated non-instances (aka type II errors)

In simple words, both True Negatives (TN) and True Positives (TP) are accurate predictions. When you mistakenly mark a negative instance as positive, it is known as false positive (FP). False Negatives (FN) are mistakes where you mark a positive case as negative. The confusion matrix given by our model which uses categorical cross entropy as the loss function is as shown by Fig. 2. where in 19 cases were correctly predicted as "Tumor" whereas 14 cases were correctly predicted as "No Tumor". Only 1 case was incorrectly predicted as "No Tumor" which was actually from class "Tumor" i.e False Negative.

E. Saving the model

Save the trained model, so it can be used to predict the class that is "Tumor", or "No Tumor" from unseen images.

IV. EXPERIMENT AND ANALYSIS

The proposed model demonstrated strong performance in Brain MRI classification. The model has been trained with augmented data and the model was then evaluated using K-fold crossvalidation, with a batch size of 16 and a value of k as 5. The number of epochs was also set to 5. With an average accuracy of 95.89%, the model produced a high percentage of accurate classifications. Precision and recall then quantify how well positive classes are discovered and reduce false negatives and false positives. The mean of 0.122265 says that the model can effectively learn and uncover the hidden features of the data.

Fig. 6 shows a comparison of results obtained using different loss functions and epochs for a CNN model. Three loss functions were tested: Binary Cross-Entropy without augmentation of the data, Categorical Cross-Entropy (8 epochs), and Categorical CrossEntropy (5 epochs). The evaluation was done and the following are the key findings:

Loss Function: All loss functions yielded acceptable results, Categorical Cross-Entropy with 5 epochs demonstrated the highest Accuracy (0.95), Precision (0.98), Recall (0.94), and F1-score (0.96). Decreasing the number of epochs to 5 for Categorical Cross-

Entropy led to an improvement in performance compared to 8 epochs. Overall performance of the model was strong across all metrics, particularly with Categorical Cross-Entropy with 5 epochs.

ISSN 2394 - 7780

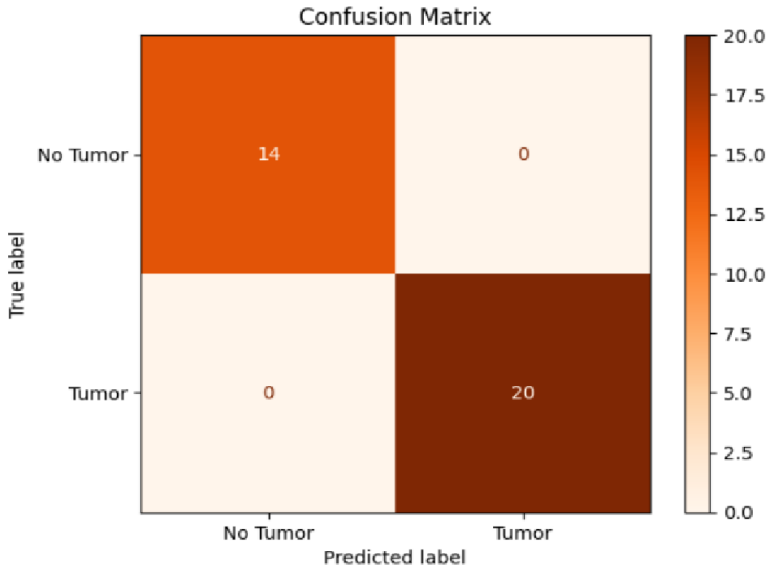


Fig. 5 Confusion Matrix final model

V. CONCLUSIONS

Fig. 5 shows the confusion matrix of the final model, this model was proposed to enhance brain tumor Detection using the CNN. So, after applying K-fold cross-validation on the MRI dataset we got better accuracy and least amount of loss which indicates that the model is well trained. Without applying augmentation, the model was overfitting with the 100% accuracy score. Since our dataset was very small so we applied augmentation for better result and that helped in removing the overfitting of the model so, we got the accuracy of 96%.

Loss Function Type	Binary cross entropy loss	Categorical cross-entropy loss (epochs=8)	Categorical cross-entropy loss (epochs=5)
Accuracy	0.85	0.93	0.95
Precision	0.86	0.93	0.98
Recall	0.83	0.94	0.94
F1 Score	0.83	0.93	0.96

Fig. 6 The comparison of loss functions

ACKNOWLEDGMENT

We acknowledge the limitations imposed by the relatively small dataset used in this study. The limited data availability may have impacted the model's generalization performance. Future research with bigger and more varied datasets could further enhance the model's capabilities and robustness. Also, the model's average accuracy score is very much acceptable. We deeply appreciate everyone who assisted us in completing this research and for their support. Above all, I would like to express our sincere gratitude to Mrs. Sunita Jena, the assistant professor, and Ms. Niloufer K. Kotwal, the head of the department of Life Sciences, for their helpful guidance, sensible recommendations, and constant encouragement during this project. We also like to thank Jai Hind College's MSC Big Data Analytics Department for providing the tools, space, and assistance we required to do this study. I want to express my gratitude to everyone listed above as well as to everyone who helped with this research, whether directly or indirectly.

REFERENCES

[1] "American Ancology Insitute," American Ancology Insitute, 18 09 2024. [Online]. Available: https://www.americanoncology.com/blogs/the-role-of-mri-andother-imaging-techniques-in-brain-tumor-diagnosis. [Accessed 28 10 2024].

[2] N. Kumar, "NeuroScan: Brain Tumor Detection using Convolution Neural Network," *Indian Scientific Journal Of Research In Engineering And Management*, vol. 3, 2024.

[3] M. A. A. N. S. M. and N. A. , "“Brain Tumor Detection and Classification Using an Optimized Convolutional Neural Network"," *Taibah University*, 2024.

ISSN 2394 - 7780

[4] A. K. S. G. Komal Sharma, "Brain Tumor Detection based on Machine," *International Journal of Computer Applications (0975 – 8887)*, p. 10, 2014.

[5] N. P. K. N. S. K. S. N. and B. P. S. , "Tumor Detection using threshold operation in MRI Brain Images," *Natarajan, P., Krishnan, N., Kenkre, N. S., Nancy, S., & Singh, B. P. (2012). Tumor detection using threshold operation in MRI brain images. IEEE International Conference on Computational Intelligence and Computing Research*, 2012.

[6] J. Seetha and . S. S. R. , "Brain Tumor Classification Using," *Biomedical & Pharmacology Journal*, vol. 11(3), pp. 14571461, 2018.

[7] S. Deepak and P. A. , "Brain tumor classification using deep CNN features via transfer learning," *Computers in Biology and Medicine*, vol. 111, 2019.

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

OPTIMIZING THREAT DETECTION IN CYBER SECURITY USING ML ALGORITHMS

Rahul Brijlal Yadav
Ramniranjan Jhunjhunwala college of Comm, Arts and Science

ABSTRACT
With the growing sophistication of cyber threats, the demand for timely and effective detection is greater than ever. The current research aims to identify and classify various types of cyber attacks using machine learning algorithms, specifically Naive Bayes, Quadratic Discriminant Analysis (QDA), and Multi-Layer Perceptron (MLP). We also examined attacks like Bots, DDoS attacks, and various Denial of Service variants like GoldenEye, Hulk, Slowhttptest, and Slowloris, and FTP-Patator, PortScan, and SSH-Patator. We chose the most suitable features for each attack type based on feature importance and trained the models accordingly. Our findings revealed that Naive Bayes performed incredibly well consistently, with high accuracy for all types of attacks. It was especially effective in identifying Bots, DDoS attacks, and various DoS attacks, with accuracy rates of over 90% in most cases. Although the MLP classifier displayed excellent performance in some attack types, its performance was highly sensitive to the type of attack. QDA, although helpful, tended to fall behind both Naive Bayes and MLP in accuracy. These results underscore the importance of selecting the most appropriate machine learning model for each type of threat. Naive Bayes proved to be a especially robust tool for real-time identification of threats and thus a valuable addition to the enhancement of cybersecurity strategies. This research underscores the importance of customized approaches in machine learning for cybersecurity and provides practical insights into the optimization of threat detection and response strategies.

Keywords: *Machine Learning, Threat Detection, Naive Bayes, Quadratic Discriminant Analysis, Multi-Layer Perceptron, Bot Detection, DDoS Attacks, DoS Attacks, Feature Importance, Attack Classification.*

INTRODUCTION
With greater interdependence comes the growing risk of cyber-attacks on individuals, organizations, and social institutions. Those cyber-attacks that have exponentially increased with the widening digital networks and systems have dramatically increased in complexity and frequency of occurrence. Signature-based detection methods and static security protocols miserably fail to keep up with the evolving scenario. These conventional methods typically focus on established attack patterns and signatures and leave systems vulnerable to new and complex threats that do not conform to the established patterns.

Signature-based detection systems work by identifying certain patterns or signatures of known threats. Although very effective against known attacks, the signature-based approach has limited capability to deal with new or modified threats. For example, Advanced Persistent Threats are crafted to target particular systems over an extended period of time, usually evading detection through traditional methods because of their dynamic nature. Similarly, the advancement of malware and attack methodologies makes it challenging for the signature-based system to identify a new variation or sophisticated approach.

To these limitations, a massive shift has been realized towards the inclusion of AI and ML technologies into cybersecurity frameworks. AI and ML have a possibility of analyzing huge data volumes, uncovering hidden patterns, and identifying anomalies that might indicate malicious activities unlike traditional methods. Advanced techniques adapt to the dynamic nature of historical data and new information and can be used in predictive analyses with proactive and adaptive approaches in improving threat detection and prevention.

This research paper explores the use of different machine learning models in enhancing cybersecurity defenses. In this paper, we examine the performance of Naive Bayes, QDA, and MLP classifiers in distinguishing between different types of cyber-attacks and begin activities. The types of attacks looked at include bot attacks, DDoS attacks (Distributed Denial of Service), which involve many forms of DoS, including GoldenEye, Hulk, Slowhttptest, and Slowloris attacks, as well as FTP-Patator, PortScan, and SSH-Patator. These types of attacks were chosen because of their frequent occurrence and the wide effects both on individual persons and companies. Our study aims to find out how well these machine learning models can actually handle the weaknesses of the traditional approaches to cybersecurity. We have analyzed their ability in identifying and classifying such particular attack types and henceforth have compared the pros and cons of the approach over conventional techniques. Among them, DDoS attacks are most infamous for causing service disruption by flooding the network with traffic, whereas APTs and Slowloris-type DoS attacks are stealthy and persistent, which makes it difficult to detect and mitigate.

ISSN 2394 - 7780

The findings of this research are a contribution toward the deeper understanding of how machine learning can enhance cybersecurity measures, with insights into their practical applications and potential benefits. The paper is divided into several sections: we start with an overview of the current cybersecurity landscape and the limitations of traditional detection methods. Following this, we detail descriptions of the machine learning models applied in the paper. These include Naive Bayes, QDA, and MLP. Subsequent sections cover the experimental setup, results, and analysis for the above models. In conclusion, we discuss our implications regarding future cybersecurity practice and provide some directions to be considered in further research and development.

LITERATURE REVIEW

The complexity of cyber threats, especially in the context of the Internet of Things (IoT), has led to the development of sophisticated models for threat detection. One such approach is the Mayfly Optimization with Regularized Extreme Learning Machine (MFO-RELM) model, which preprocesses IoT data to enhance threat classification accuracy. This model proves to have significant improvements for identifying cybersecurity threats in the IoT environment, indicating a need for advanced preprocessing techniques along with robust classification methods. More broadly, AI and ML are transforming the realm of cybersecurity, providing tools to better and more efficiently find threats. Traditional security measures are usually insufficient when combating highly sophisticated attacks, which makes AI and ML necessary. Various ML algorithms used for anomaly detection and malware classification have been reviewed to illustrate their effectiveness in real-world applications. This overview also addresses the challenges faced, including the need for large labelled datasets and the interpretability of ML models, and suggests future research directions such as explainable AI and unsupervised learning approaches. The other innovative approach, AI Sentry, combines machine learning and neural networks for enhancing real-time threat detection and prevention. The model shifts the paradigm of merely identifying known threats and also predicts zero-day attacks as well as unknown malicious activities. This system has a proven capability to allow AI to learn about new attack vectors and ensure high accuracy in threat detection than signature-based systems.

Application of deep learning techniques including multilayer perceptron and J48 has presented very promising results in terms of managing malicious traffic. To show how well these sophisticated approaches handle cybersecurity threats, they are being applied to datasets like Advanced Security Network Metrics & Non-Payload-Based Obfuscations ASNM-NPBO. Deep learning-based threat management demonstrates the necessity of implementing such cutting-edge strategies to control the steadily rising amount of hostile traffic. Even with the improvements, protecting against advanced persistent threats (APT) is still quite difficult. The investigation of artificial intelligence (AI) to increase detection rates is planned since traditional technologies frequently fail to identify these complex threats. Effective cybersecurity plans require striking a balance between the advantages of AI and the hazards involved. AI-powered autonomous threat hunting has become a major advancement in cybersecurity. Autonomous threat hunting integrates AI with traditional threat intelligence methodologies to improve the detection and response of security systems. It leverages various AI techniques, including machine learning models and natural language processing, for proactive identification and mitigation of threats.

Artificial intelligence is also integrated into AI-SIEM systems, which use artificial neural networks to make threat identification easier. This strategy focuses on using deep learning techniques to enhance detection performance and transforming security events into unique profiles. The efficiency of AI in network intrusion detection is demonstrated by the AI-SIEM system, which has been demonstrated to be more successful than traditional techniques. AI is important in cybersecurity for tasks like automated incident response and real-time threat identification. Large amounts of data can be analyzed by the AI system to spot any dangers and take the necessary precautions to lessen harm. This is the only method to guarantee strong protection of sensitive data and stay ahead of changing cyberthreats. The use of MLP was examined in the paper.

METHODOLOGY

An important resource for network intrusion detection research is the canadian institute for cybersecurity's ids 2017 dataset. This dataset records network traffic in a controlled setting, encompassing both benign and malevolent activity. Along with innocuous traffic, it encompasses a variety of attack types, including dos, ddos, and other prevalent vectors. Numerous attributes in the dataset, such as flow time, packet length, and protocol kinds, define the properties of network packets. When developing and evaluating machine learning models for intrusion detection and classification, each of these characteristics is essential. One of the most popular datasets in academic and industrial settings for creating and improving algorithms to raise the degree of cybersecurity is the ids 2017 dataset.

ISSN 2394 - 7780

Our methodical approach to data preprocessing ensures quality and consistency by preparing the material for analysis. First, we address missing data by replacing infinite values with NaN and removing rows with missing "Flow Duration" values. The dataset is then cleaned by removing any residual NaN values. Numerical columns are cleaned and then normalized using conventional scale. Fair comparison and analysis are made possible by ensuring that every feature has a zero mean and unit variance. Label encoding is used to convert all other categorical columns to numeric values, with the exception of the target variable "Label." The value -1 is used to handle anomalies in special instances, such as 'Infinity'. Following the cleaning of each data set individually, the cleaned data frame will be concatenated into the cleaned Data Frame into a single comprehensive data frame that will be saved onto a new CSV file for other further analysis. This process ensures uniform cleaning, normalization and encoding of the data across the field.

An essential first step in comprehending and improving the effectiveness of machine learning models in identifying network threats is feature importance analysis. We evaluate the relative significance of several variables for differentiating between attack types and benign activities using the Random Forest algorithm. The features that have the biggest effects on classification accuracy are determined by this analysis. The knowledge acquired from this procedure is crucial for improving machine learning models since it enables us to rank important features and exclude less important ones, increasing the efficacy and efficiency of the models. Furthermore, knowing the significance of features enhances model interpretability, clarifies the foundation for predictions, and facilitates the creation of threat detection systems that are more accurate and dependable. Using this technique helps to maximize model performance and guarantee that the most pertinent data attributes are used in network attack classification.

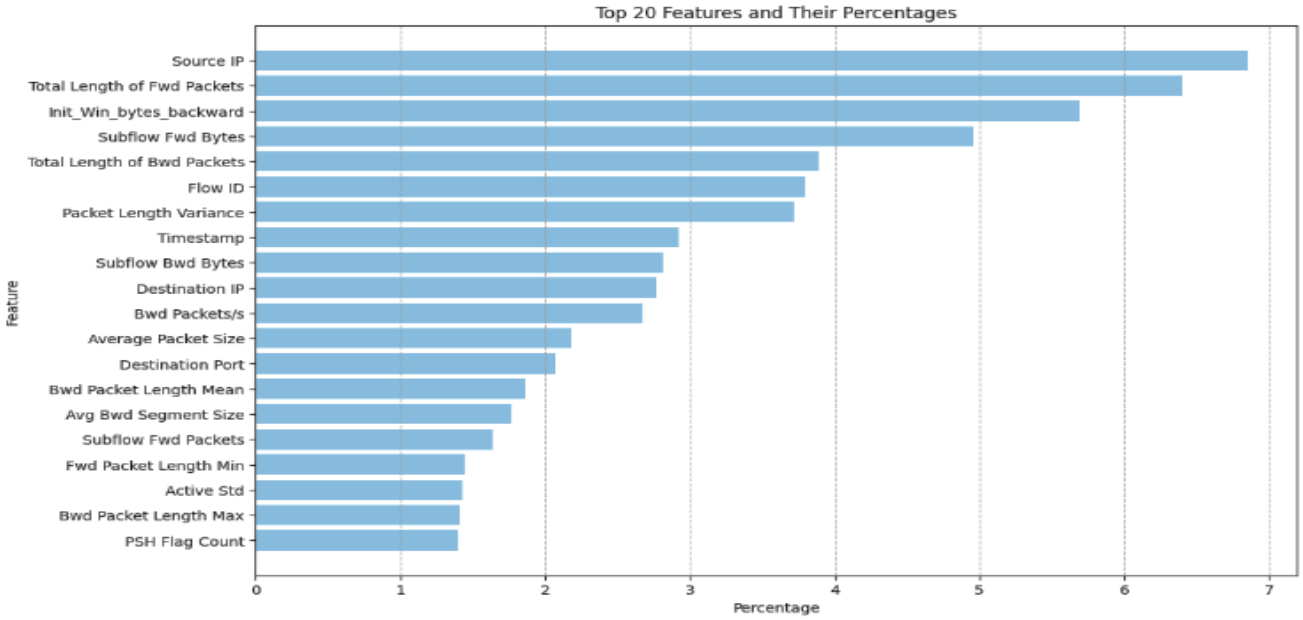


Figure 1 Feature Selection

ISSN 2394 - 7780

We evaluated several machine learning models to classify various kinds of network attacks comprehensively. In this regard, we utilized Naive Bayes, Quadratic Discriminant Analysis (QDA), and Multi-Layer Perceptron (MLP) classifiers. Each model had been chosen based on distinct approaches towards classification and handling multiple data characteristics.

Naive Bayes was chosen for preliminary classification jobs because of its ease of use and effectiveness when working with huge datasets that contain categorical features. QDA was chosen because it can more flexibly simulate decision boundaries between classes by taking feature covariance into consideration. MLP was added because, because to its neural network architecture, it can efficiently handle non-linear correlations in the data and capture intricate patterns. An input layer, hidden layers, and an output layer make up the MLP classifier's several layers. The hidden layers usually have non-linear activation functions. By adding non-linearity to the model using activation functions like ReLU or sigmoid functions in the hidden layers, MLP is able to capture complex patterns and correlations. MLP trains by optimizing the loss function by using back-propagation to adjust its weights to model complex, nonlinear interactions between features.

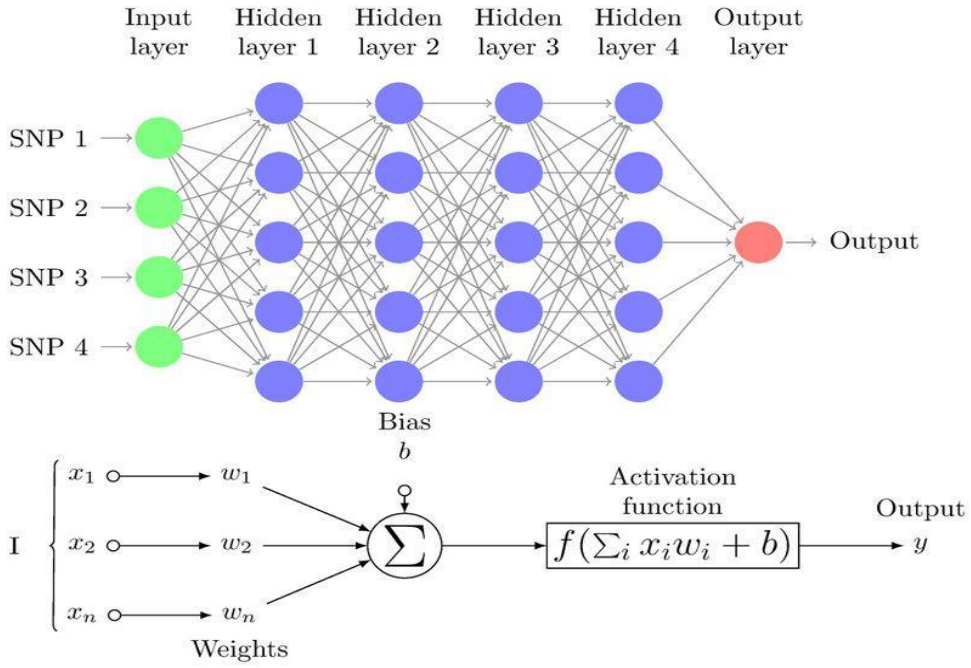


Figure 2 Multilayer perceptron

Figure 2 depicts the architecture of the Multilayer Perceptron. Models were trained on datasets customized for each attack type using preselected features that are significant for classification. Subsequently, rigorous testing was conducted to evaluate the performance of each model. Based on the comparison of accuracy metrics across different attack types, we sought to determine which model offers the best classification performance, thereby providing insights into their practical applicability for network intrusion detection. This evaluation is important as it helps in understanding the strengths and limitations of each model, guiding the selection of appropriate techniques for enhancing network security systems. The results provide valuable benchmarks for deploying machine learning-based intrusion detection systems and contribute to advancing the field of cybersecurity.

RESULTS

This study evaluates the performance of three machine learning models, Naive Bayes, Quadratic Discriminant Analysis (QDA), and Multi-Layer Perceptron (MLP), in classifying different types of network attacks. The results show important variations in model accuracy based on attack type.

Naive Bayes provided good performance, especially for detection of Bot attacks and SSH-Patator with 96.72% and 100% respectively. It also gave pretty good accuracy for FTP-Patator with 99.95% and PortScan with 99.99%. Nonetheless, its performance declined significantly for some DoS attacks. For example, Naive Bayes provided poor performance for DoS GoldenEye with an accuracy of 99.17% and only 99.51% against DoS Hulk. As for QDA, it is underperforming in nearly all attack types. Its accuracy was very low for different DoS attacks that were between 32.28% and 33.36%. The best accuracy for QDA was at SSH-Patator with 33.56%, but on average, the model performed poorly in classification. MLP was the best model with an accuracy of 99.96% for Bot attacks, 97.42% for FTP-Patator, and 98.46% for PortScan. Despite its generally high accuracy,

ISSN 2394 - 7780

MLP showed lower performance for certain attacks, including DDoS and DoS slowloris with accuracies of 71.89% and 35.45%, respectively.

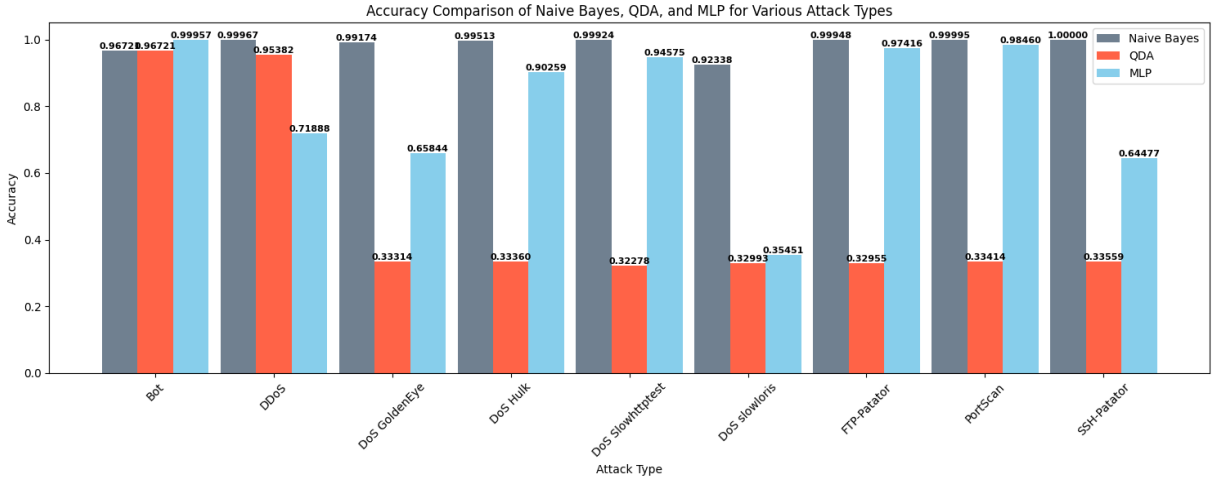


Figure 3 Accuracy comparison of Naive Bayes, QDA, and MLP for different network attacks.

As given in Figure2 below, the bar graph provides a visual comparison between the accuracy of the various models and attack types. This diagram shows that overall, MLP is relatively very effective in achieving higher accuracies when compared to QDA which, generally, performs poorly. Naive Bayes is also performing well, and specifically for the detection of some attack types, it has good performance, although not outperforming MLP constantly. The results show that different models are effective in various attacks, and the accuracy of MLP on most attacks reveals the possibility of using such models in cybersecurity to improve the network intrusion detection system. However, the lower performance of QDA across most attack types brings forth the point that alternative or hybrid models would be needed to improve overall detection capabilities.

CONCLUSION

It highlights the capacity of complex algorithms in enriching network defense mechanisms for cybersecurity attacks. The effectiveness of MLP clearly shows the potential it holds in giving accurate results on a broad range of attacks, placing it firmly at the forefront for real-time identification and mitigating threat conditions. Its robustness and ability to deliver high accuracy in most scenarios makes it a suitable candidate for incorporation into comprehensive cybersecurity frameworks. Challenges identified with QDA point towards the inability of traditional statistical models to cope up with complex and diverse attack patterns. This means that QDA alone may not be enough for strong network security, and more advanced or hybrid approaches are required that take advantage of the strengths of multiple models. In addition, although Naive Bayes performed well in many attack types, the results show that no single model is a panacea. This variability in accuracy across different attacks suggests that a multifaceted approach, combining different models, may offer a more balanced and effective defense strategy.

Future work should be on fine-tuning the MLP model and exploring ensemble methods that combine the strengths of this model with other models. Further, exploring the reasons for the poor performance of some models can help in understanding their limitations and guiding improvements. As the cyber threat landscape continues to evolve, adaptive and intelligent machine learning solutions will be necessary to stay ahead of emerging attack vectors and enhance overall network security.

REFERENCE

[1] J. Li, “Cyber security meets artificial intelligence: a survey,” *Frontiers Inf Technol Electronic Eng*, vol. 19, no. 12, pp. 1462–1474, Dec. 2018, doi: 10.1631/FITEE.1800573.

[2] F. Alrowais, S. Althahabi, S. S. Alotaibi, A. Mohamed, M. Ahmed Hamza, and R. Marzouk, “Automated Machine Learning Enabled Cybersecurity Threat Detection in Internet of Things Environment,” *Computer Systems Science and Engineering*, vol. 45, no. 1, pp. 687–700, 2023, doi: 10.32604/csse.2023.030188.

[3] Dr. N. Katiyar, Mr. S. Tripathi, Mr. P. Kumar, Mr. S. Verma, Dr. A. K. Sahu, and Dr. S. Saxena, “AI and Cyber-Security: Enhancing threat detection and response with machine learning.,” *eatp*, Apr. 2024, doi: 10.53555/kuey.v30i4.2377.

ISSN 2394 - 7780

[4] S. Rangaraju, "AI SENTRY: REINVENTING CYBERSECURITY THROUGH INTELLIGENT THREAT DETECTION," *EPHIJSE*, vol. 9, no. 3, pp. 30–35, Dec. 2023, doi: 10.53555/ephijse.v9i3.211.

[5] T. T. Teoh, G. Chiew, E. J. Franco, P. C. Ng, M. P. Benjamin, and Y. J. Goh, "Anomaly detection in cyber security attacks on networks using MLP deep learning," in *2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)*, Shah Alam: IEEE, Jul. 2018, pp. 1–5. doi: 10.1109/ICSCEE.2018.8538395.

[6] K. Hasan, S. Shetty, and S. Ullah, "Artificial Intelligence Empowered Cyber Threat Detection and Protection for Power Utilities," in *2019 IEEE 5th International Conference on Collaboration and Internet Computing (CIC)*, Los Angeles, CA, USA: IEEE, Dec. 2019, pp. 354–359. doi: 10.1109/CIC48465.2019.00049.

[7] S. R. Sindiramutty, "Autonomous Threat Hunting: A Future Paradigm for AI-Driven Threat Intelligence".

[8] J. Lee, J. Kim, I. Kim, and K. Han, "Cyber Threat Detection Based on Artificial Neural Networks Using Event Profiles," *IEEE Access*, vol. 7, pp. 165607–165626, 2019, doi: 10.1109/ACCESS.2019.2953095.

[9] M. Rizvi, "Enhancing cybersecurity: The power of artificial intelligence in threat detection and prevention," *IJAERS*, vol. 10, no. 5, pp. 055–060, 2023, doi: 10.22161/ijaers.105.8.

[10] L. Van Efferen and A. M. T. Ali-Eldin, "A multi-layer perceptron approach for flow-based anomaly detection," in *2017 International Symposium on Networks, Computers and Communications (ISNCC)*, Marrakech, Morocco: IEEE, May 2017, pp. 1–6. doi: 10.1109/ISNCC.2017.8072036.

[11] B. R. Maddireddy and B. R. Maddireddy, "Proactive Cyber Defense: Utilizing AI for Early Threat Detection and Risk Assessment," vol. 01, no. 02, 2020.

[12] A. R. P. Reddy, "THE ROLE OF ARTIFICIAL INTELLIGENCE IN PROACTIVE CYBER THREAT DETECTION IN CLOUD ENVIRONMENTS," vol. 19, no. 12, 2021.

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

LIFE CYCLE OF DATA IN CLOUD

[1]Riya Gupta and [2]Wilson Rao
[1]Student and [2]Assistant Professor, MSc. Big Data Analytics Department, Jai Hind College, Mumbai

ABSTRACT
Cloud computing has made a huge change in the big data management field. Cloud has come up with a technology that is flexible, scalable and it cuts down costs, thus making it possible to carry management and analysis of big data. The classic on-premises data management systems have always had a hard time handling the increasing amount of data that resembles big data in volume, variety, and velocity. Cloud environment like Microsoft Azure is capable of managing these huge datas with its robust infrastructure that allows companies to manage and work with large datasets with efficiency. By making use of such technologies, the organization can dynamically change the capacity of their working processors to suit the changing workload and be sure to use their resources to the maximum with the minimum cost. Another important advantage of making use of cloud computing technology in big data environments is that enterprises can exercise leverage on-demand resources. That is to say, organizations can swiftly increase or decrease the scale of their operations according to their dynamic needs, thus staying efficient. Nonetheless, in order to be completely effective in these assigning tasks to the cloud, one has to acquire the information regarding data transfer and processing which is vital for its proper functioning. This should involve knowledge of the data flow mechanisms, network latency, and the overall architecture of platforms like Azure. This research paper's primary focus has been on data processing and transportation in cloud environments, particularly on the Azure platform. It offers a thorough examination of the technological problems pertaining to the cloud ecosystem's data ingestion, processing, storage, and retrieval. It also examines security-related topics such data compliance, encryption, access management, and security standards in cloud systems. This clarifies the ways in which these factors impact data confidentiality and integrity in cloud-based big data management.

Keywords—Cloud, Data Flow, Data lifecycle, Big Data

INTRODUCTION
Organizations are turning to cloud computing as a more practical and scalable option as a result of the exponential expansion of data, which has overtaken traditional data management techniques due to the spread of digital apps and the Internet of Things (IoT). By allowing businesses to efficiently and flexibly store, process, and analyze large datasets, cloud computing provides a revolutionary approach to big data management. Cloud environments offer on-demand resources, enabling businesses to adjust their processing capacity to real-time demands, in contrast to traditional on-premises systems, which are frequently limited in scalability and require a significant upfront investment. For businesses that must handle workloads of different sizes without sacrificing efficiency or going over budget, this flexibility is crucial. In cloud computing, data flow encompasses distinct stages—data ingestion, storage, processing, and retrieval—that are essential to maximizing data management's functionality and security. Depending on the cloud service paradigm, each of these phases is handled differently. This paper explores the data lifecycle within the cloud service models, focusing on the unique mechanisms of each stage across IaaS, PaaS, and SaaS environments. (Boglaev, 2016)

OVERVIEW OF CLOUD SERVICE MODELS
Cloud computing users can request and obtain rented computing capabilities over a network that connects them to a cloud platform. A central server manages all communication between client devices and servers for data exchange. There is no one-size-fits-all approach to implementing cloud computing architecture. What works for one business might not work for another. Actually, one of the advantages of cloud computing is its adaptability and flexibility, which enables businesses to swiftly adjust to shifting measurements or markets. (Rieder, 2020)

This section presents and defines the three primary cloud service models: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). It also describes how the roles of data lifecycle management vary throughout these models. This fundamental knowledge helps readers understand how each model affects cloud environments' data management and flow.

DATA INGESTION
Data ingestion is the process of collecting and importing data files from various sources into a database for storage, processing and analysis. The goal of data ingestion is to clean and store data in an accessible and consistent central repository to prepare it for use within the organization. Data ingestion is the process of taking raw data from various sources and preparing it for analysis.

ISSN 2394 - 7780

Data discovery: The exploratory stage, during which all of the organization's data is found. The foundation for successful data ingestion is an understanding of the data environment, structure, quality, and possible applications.

Data acquisition: Data acquisition means gathering the data after the sources have been determined. Data can be retrieved from a variety of sources, including unstructured formats like spreadsheets and paper documents as well as structured databases and application programming interfaces (APIs). Maintaining data integrity during the acquisition process and managing the range of data types and perhaps high volumes provide challenges.

Data validation: Validation ensures that the data is accurate and consistent after it has been acquired. Data is examined for mistakes, discrepancies, and missing values. Through a variety of checks, including data type validation, range validation, and uniqueness validation, the data is cleansed, made trustworthy, and prepared for additional processing.

Data transformation: This is where verified data is transformed into an analysis-ready format. This could entail standardization (uniform formatting), aggregation (summarizing data), and normalizing (removing redundancies). Making the data easy to comprehend and evaluate is the aim.

Data loading: The modified data is then delivered to its assigned destination, usually a data lake or warehouse, where it is easily accessible for reporting and analysis. Depending on the particular requirements, this loading process can be carried out in real-time or in batches. When the data input process is finished, the data is prepared and ready for making well-informed decisions and producing useful business insight. This is known as data loading.

DATA STORAGE

Cloud storage is a cloud computing service in which data and files are stored offsite by a third-party provider and can be accessed through a public internet or dedicated private network connection.

- **How does cloud storage work?**

One way to store data is using on-premises networks, similarly, cloud storage also uses servers for storing data, but the data is placed on offsite servers. The majority of these servers are virtual machines (VMs) that work with a physical server. The provider generates additional virtual servers in response to the need for more storage.

A typical user gains access to a cloud storage through a web interface (web portal), website, or an app that communicates with an application programming interface. The server that you connect with, on the other hand, will redirect data to a group of servers situated in one or more big data centers, which rely on the scope of the cloud vendor's ab.

Providers that offer this service store the identical data on various machines as a backup measure. This implies that even though a server undergoes downtime for maintenance or faces an outage, the data would still be accessible by the users.

There are three main cloud storage types, each offering its own advantages.

→File storage

File storage is a technique that is used to save data in a pie file and folder structure. The data stays in the same format regardless of where it is, whether it is a cloud storage system or a client location, and the structure of the data makes it more accessible and intuitive to find and fetch it when necessary. (Boglaev, 2016)

Cloud storage for file storage is e a widely provided service allowing users to connect to the same group of files that are stored in the cloud.

→Block storage

This data model for storage arranges the information in the form of big "blocks" which are each a hard drive. The cloud storage companies utilize such blocks to divide a large amount of data into several storage nodes.

The block storage resources deliver an increased performance level across a network be it the low IO latency (which is the time taken to complete the connection between the system and the client). These resources are especially suitable for large databases as well as applications.

Block storage on the cloud can be easily scaled up to take care of the increased demand of an organization's databases and applications.

ISSN 2394 - 7780

→Object storage

Object storage is a method of data management called "object" storage, and it distinguishes and organizes data into distinct "objects." The data is the information in the file, the associated metadata, and the identifier of the object. The files contain the data in the same format that they are received and, at the same time, should allow creators to customize metadata to enable the data to be easily recognized and used. (Rieder, 2020)

The filing system is not made up of files or folder structures; they are instead stored in objects in repositories, which is a way to enable unlimited scalability. The lack of filing hierarchy and the possibility of personalization of metadata makes it possible for users to optimize storage resources at an affordable price through object storage.

Cloud-base object storage is a good way of preserving data - for the distant future. The importance of object storage has been growing with every passing day since more and more unstructured data (videos, audios, webpages, sensor data) is being stored and processed in an efficient and cost-effective way. In 2022, 90% of structured data was generated in the organizations.

DATA PROCESSING AND ANALYSIS

Data processing and analysis are critical stages in the cloud data lifecycle, where raw data is transformed into meaningful insights. This stage involves cleaning, organizing, and analyzing data to extract actionable intelligence, and it is highly dependent on the tools and services provided by the cloud service model in use (IaaS, PaaS, or SaaS).

In IaaS, data processing is largely dependent on the infrastructure set up by the users. The users can install any software packages of their choice on the virtual machines (VMs) and storage, which are in their complete control. For instance, users can deploy big data processing frameworks like Apache Hadoop, Spark, or even custom developed Extract, Transform, Load (ETL) pipelines into the VMs. While this lends the idea of elastic processing custom workflows as allowed by the cloud computing paradigm, it also means that the users have to take control of infrastructure management responsibilities as well as scaling of resources to the existing ones. IaaS supplies the required amount of computing capacity and storage but it is up to the users to optimize the processing systems for very large datasets where necessary.

When it comes to PaaS, data processing is much less concerned as the platform offers managed services that simplifies its engagement. For instance, many PaaS solutions come with native data processing options like managed DBs, serverless e.g. AWS Lambda Azure functions, and data analytics e.g. Google Big query among many others. Such platforms lend themselves for use by users concentrating on application development, with little or no need to concern about the infrastructures. For instance, a developer can apply ready-made services for data ingestion, transforming as well as analyzing without any effort for setting up the VMs or networks. This facilitates the quicker implementation of data processing and helps it to scale as the system's resources are optimally allocated and managed according to the user needs. There is, nonetheless, a downside in that the fine-grained aspects of the processing pipeline are exposed to the user, in that a lot of it is hidden by the platform.

In the case of SaaS, the entire process of data processing and analysis is left to the service provider. SaaS solutions usually include analytical tools, dashboards, and other reporting features, which are usable for specific business functions (like CRM analytics, forecasting finances, or business intelligence). Users also access the data through the application and are dependent on the SaaS provider to control the data and work on it as it comes in. For instance, Salesforce has an inbuilt analytics tool for sales and distribution and Power BI provides capabilities for visualisation of data drawing from varied sources. Also, SaaS solutions are typically user friendly in that the user does not have to be very technical in order to use advanced capabilities such as data analysis. But the unavailability of flexible options as well as the restrictions over the actual data processing systems could be a challenge for some organizations that have more precise or elaborate data processing needs.

DATA DISTRIBUTION AND ACCESS

Data distribution and access are fundamental components of the data lifecycle in cloud computing. They refer to how processed data is made available to users, applications, and systems, as well as how it is shared, stored, and retrieved across different platforms. Efficient data distribution ensures that data is accessible where and when it is needed, while robust access controls safeguard its confidentiality, integrity, and availability. This process varies significantly across IaaS, PaaS, and SaaS, with each model offering different degrees of control, security, and management capabilities.

ISSN 2394 - 7780

In the architecture of IaaS, the data distribution, as well as the access, is very extensible since the end users control the storage and the network layers. This implies that the users have to create their own basic storage systems whether it's an object storage, block storage or even file systems and also take responsibility on the way the data is spread over the various nodes or regions within the cloud infrastructure. Users can take advantage of Private clouds brough about by Amazon Web Services to store their content in different formats and many replicas also can be geographically dispersed for the purpose of ensuring availability in case one goes down. Besides that users are also afforded the freedom of geography, in that they can control fully the access of the data through allocation of roles as per the users , security protocols and management, with the addition of encrypting the data and determining who can change the information or who can access it. But all of this is possible only because the users are given a lot of power and with that, some extra responsibility of their own which is how they take care of the data control and distribution structures.

In Platform as a Service (PaaS), the data dissemination and access procedures are somewhat concealed. Such a platform includes consistent storage services which perform a lot of the distribution related activities, for instance, data backup copies, redundancy, and growth. For instance, Google's App Engine or Microsoft Azure App Services kind of platforms usually come with ready-built integrated databases, data storage, and data caching facilities which are elastic in nature for the application requirements. It is very easy for PaaS users to reassign data to different regions or nodes on the platform by leveraging the managed services that the platform provides, which take care of replication and load balancing. Nonetheless, the control over the strategies of data distribution is not as much as it is in Infrastructure as a Service. Access controls can still be imposed on certain data and platform-embedded security features like role-based access control (RBAC) can be employed to restrict data access to authorized persons only. Such as in PaaS the management of distribution of data becomes less cumbersome for the users thank the institution does PaaS wishing for more complex strategies of distribution requires trust on the institutions routing configurations.

In a SaaS model, the service provider completely takes over the responsibility of managing data distribution and access. Users only need to access the application through a web interface or an API and do not care about the location and mechanism of storing and distributing data. In addition, those service providers usually take care of data replication, and geographic distribution, and implementing failovers so that data is always accessible even in the case of hardware malfunctions or loss of server connections. How the data is accessed is strictly controlled by the access policies integrated into the application, for example, it may offer multiple user roles and support authentication and data scrambling to prevent unauthorized use. Following this model, the user is provided with the least control over the distribution and access of data, however, the provider takes the onus of making sure that data availability and secure access are assured.

To wrap things up, when it comes to understanding the distributions and access of data in IaaS, PaaS, and SaaS models, the main difference comes in control and abstraction levels. Among them, IaaS is secure and flexible enough for users to specify the desired data distribution and the access controls to be used. With PaaS, the distribution is made less complex with the use of managed services and offers a completely hands-off distribution and access model in SaaS. Therefore, the cloud model which an organization will adopt would depend on the level of control over the data distribution and the ease of operation and management.

DATA ARCHIVING AND DELETION

The last stage of any data lifecycle is data archiving and deletion, which is critical especially in cloud environments where there is lake of data. These strategies allow for the retention of data for future purposes whenever necessary, while still observing the data protection and retention policies. Within computing, the processes of archiving and deletion vary based on the cloud service model deployed since each has different capabilities, management approaches, and compliance systems.

→Data Archiving

Data archiving is a procedure to keep the already inactive computer data that would be still necessary for long-term storage or regulatory compliance purposes. Data archiving systems guarantee that files can be recovered at a later date and at the same time cut down on expenditures for memory by automatically moving to low-cost storage levels infrequently used files.

In IaaS, data archiving is very flexible. Users can store the archived data and manage it how they would like. For example, they can opt to use purpose-built dealt storage services such as Amazon Glacier or Azure Blob Storage Archive for cheap, long-term storage. Archiving in IaaS means defining custom policies, providing motion of data automation and management of storage classes so that data is archived properly. For instance, an organization may set a retention period in which the system will automatically transition aged, low access data

ISSN 2394 - 7780

to a cheaper class of storage. Furthermore, to protect the archived data, IaaS vendors usually provide encryption of the data and redundancy of the respective archived data.

In the case of PaaS, archiving is usually processed by the system with the provision of long-term data storage services. Such PaaS systems may connect to third-party systems or may have long-term storage facilities of their own. For example, Cloud Storage Nearline by Google and Cool Blob Storage by Microsoft Azure serve the purpose of archiving data. PaaS may include also extend to allowing users create automated processes for archiving data based on specific policies e.g. active data becomes stale for a certain amount of time and is archived. Some of these policies may be configurable by the users but the pain of designing and management is kindly concealed by the platform, which makes archiving quite easy but also less flexible than in IaaS.

When it comes to SaaS, the responsibility of managing archiving lies squarely with the service user. Most SaaS applications have automated archiving systems as part of the data management features of the application; especially those applications that are used for retaining data for a long period of time such as email applications, CRM systems, or document management systems. For instance, in Salesforce, some archived customer data can be available to users instead of allowing aged information to clutter the system. Health care Cloud based services SaaS solution provides managed archiving and data backup systems which preserve all data in accordance to health regulations (e.g., HIPAA, GDPR), and making sure that no loss of archived data is experienced. Yet users lack the distinction of how archiving is done because this aspect is taken up by the provider.

→Data Deletion

Data deletion is the action of permanently erasing any information from the system that is deemed unnecessary. In this case, the information should be permanently removed, and there should be no chance of recovering it which can lead to a compromise on security or privacy. Cloud environments demand appropriate data obliteration practices as a defense to many legal obligations (such as the "right to be forgotten" imposed in GDPR) and limits the amount of time an organization keeps certain types of data.

In IaaS, the implementation of data deletion policies falls in the hands of the user. As they own the entire infrastructure, users put in place deletion policies in which they make certain that all data is removed often using industry compliant data wiping mechanisms. Users can perform data deletion from virtual machines, storage volumes, and databases or alternatively create processes that will cause data deletion after certain duration. Data deletion in IaaS is performed hand in hand with data sanitization processes ensuring the storage devices have been rendered useless to the contained data. Additional services like provision of secure deletion tools, among others, may be provided by cloud service providers but the user is responsible for ensuring compliance.

In a typical PaaS (Platform as a Service) environment, data deletion is usually carried out via built-in services and policies of the platform, which are typically disruptive for the end user. For instance, several PaaS services make it possible for users to set up lifecycle management policies for the data they keep, which includes automating the process of deleting or archiving information after certain conditions are met, such as data becoming a certain age. In this case, users can still have some degree of influence on the way data is deleted and the specific time it is deleted, particularly when using particular database or object storage services. In many PaaS environments deleting of data tends to be of datapump's purging / versions control, which is tied to the application lifecycle, and additional features such as g.c. or data retention policies may be present in the platform to ensure that privacy policies of data regulation are adhered to. Even though much of the process is abstracted by the platform, the user's responsibility handles data that requires deletion.

In SaaS, data deletion is managed by the service provider according to the terms of the service agreement and applicable legal or compliance requirements. Many SaaS providers offer data retention and deletion features that automatically handle data erasure once it is no longer needed. For example, in platforms like Google Workspace, users can delete files manually, and the provider will handle the backend deletion processes, ensuring the data is completely removed from the servers. The service provider often adheres to best practices for secure deletion and may offer options for users to request permanent deletion of their data. However, as with archiving, users have limited control over the deletion process itself, and they must rely on the provider's security and compliance measures.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my mentor, Prof. Wilson Rao, for his invaluable guidance, insights, and continuous encouragement throughout this research. My sincere thanks to MSC - Big Data

ISSN 2394 - 7780

Analytics Department, Jai Hind College for providing the resources and supportive environment necessary for conducting this research.

REFERENCES

1. **Boglaev, I. (2016).** A numerical method for solving nonlinear integro-differential equations of Fredholm type, *Journal of Computational Mathematics*, 34(3), 262–284. https://doi.org/10.4208/jcm.1512-m2015-0241
2. **Lindberg, D. V., & Lee, H. K. H. (2015).** Optimization under constraints by applying an asymmetric entropy measure, *Journal of Computational and Graphical Statististics*, 24(2), 379–393. https://doi.org/10.1080/10618600.2014.901225
3. **Rieder, B. (2020).** *Engines of Order: A Mechanology of Algorithmic Techniques*. Amsterdam, Netherlands: Amsterdam Univ. Press.

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

BREACHES IN THE DIGITAL FORT: A STUDY ON CUSTOMER DATA LEAKS AND CYBERSECURITY IN INDIA

[1]Sonakshi Julka and [2]Wilson Rao
[1]Student and [2]Assistant Professor, MSc. Big Data Analytics Department, Jai Hind College, Mumbai

ABSTRACT
Cybercrime is rapidly escalating in India, presenting unique challenges for developing economies like India that are undergoing digital transformation. This paper tries to explore the growing issue of cybercrime and cybersecurity in India, where cyberattacks are reportedly increasing, and different organizations face the average attacks per week of 3,201. Successive waves of the digital revolution have prompted unparalleled growth in online services, and this is nothing but massive data accumulation at various levels across various industries from finance to e-commerce. Apart from this, this phenomenal growth period has thrown open critical weaknesses in India's cybersecurity infrastructure to serious breaches and mass leakage of sensitive customer data. We try to outline in this paper the most notable lapses on cybersecurity over the recent past and the breaches of vulnerability across various sectors, such as the massive Aadhaar data breach, security failures within the banking industry, vulnerabilities in e-commerce, and telecom data leaks. Through these high-profile cases, we identify common patterns involving security flaws, evaluate the falling domino effect of such breaches on the individuals and organizations involved, and thus assess the efficacy of the mitigation measures currently in use. This research also identifies a very critical need for stronger cybersecurity measures and frameworks of regulatory measures to protect personal data in an increasingly more digitalized economy. We assess the impact of security breaches on consumers, coordination challenges of cyber defenses, and challenges in imposing suitable regulatory oversight. Finally, we look at the Indian government's and the private sector's responses by coming out with initiatives such as the Personal Data Protection Bill that will ensure better data security and privacy standards. We propose actionable solutions and recommendations that would further enhance cybersecurity policies already in place and further strengthen the protection of consumer data. Rapid technological development requires that India address the vulnerabilities thus created in a forward-looking manner so as to preserve consumer's faith in security in this digital ecosystem.

Keywords: *Data breach, Data leak, cybercrime, cybersecurity, Cybercrime in India, Aadhaar data breach, Cyberattacks, Customer data protection, Personal Data Protection Bill cybersecurity gaps.*

I. INTRODUCTION
India has become the pace-setter in its digital transformation. However, the untimely rapid digital transformation has led to the adoption of cybercrime at equal speeds as well. Here lies the critical gaps in the cybersecurity framework of India. With the organizations witnessing an average of 3,201 cyberattacks every week, securing personal and organizational data is really becoming a daunting task. As all sectors be it finance, e-commerce, or telecommunication are moving online, many millions of the country's citizen's information channeled to inadequately secured platforms. Such an attack that symbolized and highlighted Aadhaar data leakage makes the consumer vulnerable and brings their trust into question with regard to digital services. These cyber failures come at a cost that goes beyond mere losses but lowers public confidence. Such efforts, therefore, as in the Personal Data Protection Bill, indicate that they are headed in the right direction regarding proper data governance; however, current measures are still not at par with the complexity of modern cyber threats. This paper investigates recurring security issues in industries, assesses the impact on the consumers and the organizations involved, and evaluates the effectiveness of the current mitigation strategies used. This analysis aims to present typical patterns in security flaws and provide applicable solutions to be implemented for the improvement of India's cybersecurity landscape.

ISSN 2394 - 7780

Fig 1: Data leak, Data breach, cybercrime

II. OVERVIEW OF CYBERSECURITY IN INDIA

A. Current Landscape

This digital transformation that India has witnessed, with the adoption of online services across various sectors, needed robust cybersecurity infrastructure. With growth so quick in India, security measures put into place lagged behind and exposed the customer vulnerabilities in terms of data protection. These are areas such as finance and healthcare, where personal information is ammassed on a large scale; and ecommerce, which is one of the most vulnerable sectors in this regard. (Mehta, V., & Dhillon, G. S. 2019, Mukherjee, A., & Mohapatra, P. 2017) For instance, there is an example of the growth in online banking leading to allowing cyber-crooks to exploit a software vulnerability or steal financial records of individual consumers.

B. Regulatory Framework

Even though it would raise a legal framework to manage crimes related to the information technology area, there were severe limitations. The prime focus of the IT Act revolves around cybercrime punishment without seriously working for the protection of consumer data. Recently, the Personal Data Protection Bill had been introduced to upgrade the regulations (Sharma, D., & Singhal, S. 2022) dealing with data privacy. It aimed to vest data protection rights in the citizen, compel organizations to improve data security measures, and demonstrate compliance. However, the implementation of all these rules remains challenging in this Indian digital scattered landscape.

C. Regulatory Framework

Whereas the Information Technology Act gave legal competency for cyber-related offenses to be handled, it is somewhat more riddled with very serious pitfalls. The IT Act dealt primarily with penalizing cybercrimes, and hardly any attention was paid for proactively taking care of protecting consumer data. The Personal Data Protection Bill was recently drafted to fortify data privacy regulation. This bill has empowered citizens with rights of data protection, obliged organizations to enhance data security, and necessitated conformance. However, implementing such pieces of legislation across the highly fractured digital platform in India is no mean feat.

III. CASE STUDIES ON MAJOR DATA BREACHES IN INDIA

1.BigBasket Data Leak

In October 2020, hackers compromised an online grocery delivery service, BigBasket, exposing over 20 million customer details. Hackers stole email addresses, phone numbers, delivery addresses, and hashed passwords. Hackers seem to have put up the stolen data on the dark web for sale and can sell it to identity thieves. It again brought under focus the need for stronger data encryption and cybersecurity measures in the e-commerce sector, where a huge amount of sensitive information is kept about its customers.

2. Mobikwik Data Breach

In March 2021, digital wallet provider Mobikwik suffered a data breach that affected nearly 100 million users. The exposed data included names, email addresses, phone numbers, bank account details, and partial credit card numbers. Cybersecurity researchers discovered a 9GB database with Mobikwik's user information on the dark web. Mobikwik initially denied the breach but later faced significant public backlash. This case drew attention to the importance of proactive breach detection and the implementation of multifactor authentication for sensitive accounts.

ISSN 2394 - 7780

3. Domino's India Data Leak
In May 2021, Domino's India operated by the company Jubilant Food Works reported a data breach that leaked data of 180 million orders. It compromised customer name, phone number and delivery address information. Details of internal company files and an employee's information were also breached. The data later surfaced for sale on the dark web. This case points out the vulnerabilities in protecting customer information in the food and hospitality industry, along with the importance of adequate data storage practices.

4. Aadhaar-Linked Health Records Leak
In January 2022, a health sector breach exposed sensitive medical records of individuals linked to their Aadhaar information. This included various health databases integrated with India's National Digital Health Mission, or NDHM. Exposed data include their health history, treatments, and other personal information. This has once again added concern over security protocols for Aadhaar-linked records as well as dangers of centralized storage combined with personal health data integration.

5. Unacademy Data Breach
In mid-2022, Unacademy, a popular e-learning platform in India, suffered a data breach that compromised data of over 22 million users. Leaked information included email addresses, names, and usernames, and was reportedly sold on the dark web. Although no payment information was leaked, the incident illustrated the growing threats to online education platforms, especially as e-learning gained traction during the COVID-19 pandemic.

6. HDFC Bank Data Exposure
In early 2023, a security researcher disclosed a data exposure incident at HDFC Bank that displayed some comprehensive information about clients due to unsecured servers. Names, account numbers, and transactions carried out of the customers were compromised. The breach was promptly addressed, but it did spark some questions about internal cybersecurity practices and the data protection protocols existing within Indian banking institutions.

7. Aadhaar Data Breach
One of the most significant data breaches in India involved the Aadhaar database, which exposed the personal information of millions of Indian citizens. Unauthorized access to this data led to widespread privacy concerns, highlighting severe flaws in data protection strategies and the need for advanced security protocols. The breach exposed the vulnerability of centralized databases and the potential consequences when sensitive citizen data is compromised.

8. Air India Data Breach
Air India reportedly suffered a data breach in May 2021, which compromised about 4.5 million passengers. The stolen information included passenger names, passport details, contact data, and credit card details. The vulnerability was in the systems of SITA, a global IT provider for the aviation sector. Third party data management risks were finally brought to the forefront, and with robust data sharing agreements and cybersecurity standards in partnerships, Air India answered back with issues like password reset and credit card reissuance.

9. Justdial Data Leak
An investigation in July 2021 found out there was a data breach involving Justdial, a free local search and business directory service, where personal information of more than 100 million users leaked out into the open. The data exposed contained names, phone numbers, email addresses, gender and details of user queries. The hackers managed to steal the data because the endpoints in the API of the mobile application of Justdial were insecure. This, in the course of time, led to a responsibility in making people aware of the need to secure API endpoints as well as encrypting customer data in the sector of search and business services.

10. Indian Railways Data Leak
This was the case when, in February 2023, it was revealed that a data leak from Indian Railways exposed millions of passengers' names, phone numbers, email addresses, and travel details. The breach, which has been caused by the vulnerabilities within the website of Indian Railways, does not only compromise personal information but also travel schedules, thereby exposing privacy and security risks among people affected. The aftermath of the incident highlighted the need for proper security implementations in the online services managed by the government and put more importance on the increasingly growing peril to public infrastructure through cyberattacks.

ISSN 2394 - 7780

IV. CYBERSECURITY CHALLENGES AND IMPACT

A. Key Cybersecurity Challenges in India

1. Technical Vulnerabilities

One of the major issues in the cybersecurity landscape of India is technical vulnerabilities in digital infrastructures. It's a matter of great concern that there are very few advanced security systems available within the small and medium-sized enterprises in India, thus making them vulnerable to sophisticated cyberattacks. Some common reasons for exposure can be obsolete software, unpatched systems, and weak encryption protocols with respect to risks like ransomware, phishing, and Distributed Denial of Service (DDoS) attacks. Moreover, when heritage systems are integrated with newer technologies, compatibility issues arise that cybercriminals can take advantage of.

2. Inadequate Regulatory Framework and Enforcement

While India has made strides in developing cybersecurity regulations, the enforcement of these laws remains inconsistent. The Information Technology (IT) Act, 2000, provides a legal framework for addressing cybercrimes, but it is often criticized for being outdated and not comprehensive enough to cover emerging threats. The proposed Personal Data Protection Bill aims to strengthen data privacy and protection, but its implementation is still in progress. Furthermore, the lack of uniform enforcement across different states and sectors leads to gaps in cybersecurity practices, making it challenging to maintain a cohesive defense against cyber threats.

3. Increasing Sophistication of Cyber Threats

Cyber threats in India have increased in sophistication level; with attackers now applying advanced techniques such as AI and ML in breaching the security systems. These technologies enable cybercriminals to automate attacks, avoid detection, and exploit vulnerabilities more efficiently. The rise in targeted attacks, including APTs, is a tremendous challenge since they are designed to remain masked for extended periods; thereby causing vast damage before being identified.

Fig 2: Cybersecurity, Cybersecurity in India

4. Insufficient Investment in Cybersecurity

Many Indian organizations, especially in the public sector, allocate limited budgets for cybersecurity initiatives.This underinvestment leads to inadequate security measures, insufficient monitoring, and delayed responses to incidents. Without substantial financial commitment, organizations struggle to adopt cutting-edge security technologies and hire qualified professionals, leaving them vulnerable to cyberattacks.

B. Impact of Cybersecurity Breaches in India

1. Economic Consequences

Cybercrime attacks have serious economic implications on businesses, and the general economy at large. The losses in this regard would be enormous resulting from various financial consequences of data breach, ransomware, and fraud. Some data indicate that the cost of a data breach for organizations in India reached an average of $2.18 million in 2023, up 28% on a yearly basis from data collected in 2020. The cost in this context includes direct financial loss, the lawyers' fee and regulatory penalties, and costs incurred during incident response and recovery. Cyberattacks can also impact the business operations, causing a loss in revenue and reduced productivity.

2. Erosion of Consumer Trust

Data breaches erode consumer trust, which is essential for the sustained growth of digital services. When customer’s personal and financial information is compromised their confidence in using online platforms diminishes. This loss of trust can result in decreased customer loyalty, reduced user engagement, and

ISSN 2394 - 7780

a decline in overall market participation. Rebuilding consumer trust after a breach is a lengthy and costly process, often requiring extensive public relations efforts and enhanced security measures.

3. Legal and Regulatory Repercussions

Cybersecurity attacks may also attract legal and regulatory penalties. To violate the provisions related to data protection laws and cybersecurity regulations can result in huge fines, lawsuits, and penalties. Once the Personal Data Protection Bill comes into effect, more stringent compliance standards will be imposed, and if companies don't respect those standards, huge legal penalties will be incurred. Diversion cost apart from monetary expenses due to lawsuits, there are diversion costs involved with diversion of resources from the core business activities for compliance matters and legal defenses.

4. National Security Risks

Cybersecurity breaches can pose national security risks, particularly when they target critical infrastructure sectors such as banking, telecommunications, healthcare, and government services. Attacks on these sectors can disrupt essential services, compromise sensitive information, and undermine national security. The interconnectedness of digital systems means that a breach in one sector can have cascading effects across multiple domains, amplifying the overall impact on the country's stability and security.

V. CURRENT AND PROPOSED SOLUTIONS

Fig 3: Cybersecurity, Solutions to prevent data leaks

1. Advanced Technological Measures

Artificial intelligence (AI) and machine learning (ML) offer innovative ways to enhance cybersecurity (Ramesh, S., & Prakash, A. 2023) by detecting anomalies, predicting potential threats, and automating response actions. Techniques such as behavioral analytics, where AI identifies unusual user behavior, could provide early alerts for potential breaches. Blockchain technology also presents a decentralized solution for storing and verifying data, reducing the risks of centralized breaches.

2. Strengthening Regulatory Frameworks

The Personal Data Protection Bill in its current version really restricts data collection and storage, as well as processing, and empowers the citizen with control over his or her personal data. For this to be fully effective, these regulations need to be implemented very strictly and, if possible, similar to international standards, such as the General Data Protection Regulation of the European Union. Additional compliance measures and regular audits will further ensure that companies adhere to data protection protocols.

3. Best Practices for Organizations

Best practices include holding regular security audits, implementing multi-factor authentication, and offering training for employees to improve organizational security. Securing data encrypts software updates and the use of best coding practices; these are effective ways to prevent data breaches. Organizations are able to respond to breaches, recover quickly, and have less damage with good incident response planning.

ISSN 2394 - 7780

VI. FUTURE DIRECTIONS AND RECOMMENDATIONS

1.Emerging Technologies

Quantum computing, edge computing, and secure cloud infrastructure present new opportunities for advancing cybersecurity in India. Quantum computing, while still in its infancy, promises unparalleled data processing capabilities that could help defend against future cyber threats. Additionally, edge computing enables data processing closer to the data source, improving security in real-time applications.

2. Policy Recommendations

Policymakers need to bring in clear guidelines and stringent penalties in case of noncompliance to fortify the cybersecurity base in India. At the same time, there is a need for greater transparency in dealing with data. Improvement of national cybersecurity infrastructure needs collective contributions from both government as well as private sectors of the economy. Moreover, creating consumer awareness campaigns can empower citizens to protect their data and report potential breaches.

3. Public Awareness and Education

With this effort and growing dependence on the digital world, an educated public about cybersecurity has become a necessity. Programs educating consumers on their rights when it comes to data privacy and practices of consuming data online can be a significant step forward in making consumers less vulnerable to cyber threats. Schools and universities are equally important in trying to build a solid foundation for the future workforce with cybersecurity training within school curriculums.

VII. CONCLUSION

India's specific cybersecurity needs require a proactive and concerted approach towards the protection of personal data in an increasingly digital economy. With focus on current high-profile data breaches by critically reviewing current measures, suggesting solutions to exemplify the greater need for more robust security frameworks, this paper highlights this need as being imperative for India's cybersecurity infrastructure to be regarded as a priority with regards to more advanced technology, legislation, and the fostering of security awareness through varied steps that can maintain public trust in digital services.

ACKNOWLEDGMENTS

I am profoundly grateful to my mentor, wilson rao sir, for his invaluable guidance and unwavering support throughout the course of this research paper. His vast knowledge and insightful feedback have greatly deepened my understanding of the subject. I truly appreciate the time he invested in reviewing my work and offering thoughtful advice, which has significantly elevated the quality of this study.

REFERENCES

[1] **Bose, S., & Verma, M.** (2024). Data breach costs in India: An analysis of financial impacts from 2020 to 2024. Journal of Financial Cybersecurity, 2(1), 50-68. doi:10.1016/j.fcyber.2024.01.004

[2] **Mehta, V., & Dhillon, G. S.** (2019). Current cybersecurity challenges and solutions in e-commerce. Journal of Cybersecurity and Privacy, 3(2), 123-138. doi:10.3390/jcp3020011

[3] **Mukherjee, A., & Mohapatra, P.** (2017). Cybersecurity issues in modern day e-commerce. Computers & Security, 78, 241-259. doi:10.1016/j.cose.2017.04.005

[4] **Ramesh, S., & Prakash, A.** (2023). The rise of phishing attacks in India: Trends and mitigation strategies. Journal of Cybersecurity Trends, 5(3), 145-160. doi:10.1016/j.cybertrends.2023.03.002

[5] **Sharma, D., & Singhal, S.** (2022). Data protection laws and the rise of privacy frameworks in India: A review of the Personal Data Protection Bill. Indian Journal of Law and Technology, 14(1), 15- 29.

[6] **Singh, K., & Choudhary, R.** (2021). A case study of Aadhaar data breach and implications on data security. International Journal of Cyber Criminology, 12(2), 65-78. doi:10.1177/0974627921123429

[7] **Subramaniam, B., & Gupta, P.** (2018). "Cybersecurity in Indian banking: A roadmap to enhance data protection and privacy." International Journal of Information Management, 42, 46- 56. doi:10.1016/j.ijinfomgt.2018.04.004

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

IMPACT OF MAJOR NEWS EVENTS ON INVESTOR SENTIMENT AND HERD BEHAVIOR: A STUDY ON COVID-19 PANDEMIC

[1]Swarda Ankush Parab and [2]Sunita Jena
[1]Student and Assistant Professor, Department of Big Data Analytics, Jai Hind College, Autonomous, Mumbai - 400020

ABSTRACT
To test the impact of the significant news events during the COVID-19 epidemic to the Indian stock market, this study makes the use of NIFTY-50 index for the analysis of the investors' sentiment and the extent of the herd mentality. As a result of analysing news data, sentiment scores have been obtained to measure the reactions of investors with the help of various sentiment analysis tools, including VADER, BERT, and FinBERT. The study then used the Cumulative Abnormal Returns (CAR) results to demonstrate the positive and significant relationship between sentiment and market returns using the random forest, gradient boosting and extreme gradient boosting regressors. Besides, cross sectional measures of CSAD and CSSD established that herding behaviour was prominent during periods of high volatility. These results indicate that particularly during crises investor activity significantly influences market dynamics and encourages group action. Thus, this research helps to expand knowledge of the psychological factors influencing financial markets and to provide recommendations for practical improvements in decision-making and improving the stability of financial markets in conditions of uncertainty.

Keywords *- Investor Sentiment, Herd Behaviour, COVID-19, Sentiment Analysis, Stock Market*

I. INTRODUCTION
Investor sentiment refers to the overall attitude or perception of investors about particular stocks or market conditions. Such sentiments play an important role in influencing market movements and often result in prices deviating from their intrinsic values. By taking data points from informational sources such as financial data, social media activity, and news articles, sentiment analysis has become an effective tool in predicting market behaviour because of advancements made in machine learning and natural language processing. Thus, today investors as well as analysts have a better understanding of what comprises public opinion and subsequently can forecast market actions more effectively due to technological innovations.

Investor sentiment is the general feeling that investors have about specific stocks or market situation. Opinions of this kind are essential in determining prices in the market and are frequently responsible for the disparity between the actual values and the prices. Because of the development of the machine learning and natural language processing, sentiment analysis provides the data taken from informational sources including financial data, social activity, and articles to predict market trends. Thus, today the investors and analysts have improved understanding what constitutes public opinion and accordingly can predict the market actions because of technological advancement.

While herd behaviour is that situation where investors relinquish their own discretion and decision-making capabilities and begin to emulate others. The action of one investor ends up influencing the next then the next till it reaches a large group of investors who have switched their actions according to trends this is like the ripple effect. The two main reasons are FOMO or you think others have more information than you do. However, such marketplace inefficiencies result from price deviation from their fundamental values through the involvement of more than one group-market bubbles or crashes may follow. Daniel Kahneman's dual-system theory offers an account for this phenomenon because he argues that investors tend to make fast, unconscious-System 1- or slow, conscious-System 2- decisions most of the time. Herd behaviour is particularly so because emotional reactions are high in events of significant news communications-health pandemics, policy changes or economic crises-and this indeed makes the markets volatile also.

This paper seeks to establish how large news events influence the investors' feelings and hence create the herding behaviour. It can be appreciated that events in the news are random while the responses of the markets are predictable when the right analytical tools are applied to make the estimates hence the need for a study of this nature. Investors and analysts will be equipped with adequate tools once academicians undertake research into the direction and herding activity. However, at the same time, this research demonstrates how the use of sentiment analysis makes it possible to determine that herding is dominant, and therefore, investors should not make unsuitable decisions influenced by prejudice emotions and policymakers should be capable of taking necessary actions that would contribute to the establishment of stable market condition.

ISSN 2394 - 7780

In today's interconnected global financial market, it is important to investigate how big news items bear on investor feelings and trends. This research aims at increasing the current knowledge on how investor psychology reacts to external stimuli, as well as at providing practical implications to maintain stability in the markets, refining decision-making, and advancing risk management for analysts, investors, and regulators.

II. RELATED WORK

A. Sentiment-Related Studies

The study by Verma et al. (2017) categorize news events into economic, political, and social views using Support Vector Machines and other machine learning methodologies with an intention to establishing the impact of news on Indian stock markets. [1] Their research argues for news as an important factor of the market in consideration, by employing Granger causality: instead, providing evidence of causality of news types to changes in the markets. Similarly to above Hanna et al. (2020) who investigated the impact of the media concerning the mood of the investors in the bull and bear markets. It was proposed that news sentiment would be in the Financial Times and would lead to stock returns particularly during bullish periods when trading activity was linked to attitude [2]. In 2019, Papakyriakou and his team examined the impacts of G7 terrorist attacks on international stock exchanges; they also found that with attacks and subsequent attacks, the stock return sentiment decreases with higher market losses. This scheme is worsened by social and news media [3].

Ren et al. (2020) suggest a BERT-BiLSTM hybrid model for improved sentiment analysis in the energy industry, which results in better trend forecasts in the energy market [4]. Similarly, Day and Lee (2019) use deep learning for financial sentiment research, demonstrating that the quality of news sources affects prediction accuracy and supporting deep learning models over lexicon-based methods for longer-term, more precise market predictions [5].

B. Psychological Biases and Emotional Influences

Aigbovo & Ilaboya (2019) talk about how behavioural biases affect Nigerian investors and point to aspects of prospect theory, such as loss aversion, as major causes of less-than-ideal investment choices [6]. They contend that emotional elements, such as fear and greed, lead to heuristicdriven behaviours that frequently result in unfavourable consequences, and they recommend that financial education and assistance could lessen these biases. Othman (2021) delves deeper into psychological factors by illustrating biases such anchoring and overconfidence in investor decision-making using Daniel Kahneman's dual-system theory. The paper proposes methods such as journaling and financial advising to avoid biases and improve rational decisions regarding investments by incorporating AI for logical decision assistance [7].

C. Herding Behaviour in Financial Markets

By differentiating between information-based and reputation-based herding, Bikhchandani and Sharma (2000) offer fundamental insights into herding. They advocate for legislative actions to increase transparency and reduce market volatility, emphasizing the significance of distinguishing between real herding and illusory converge [8]. Chiang and Zheng (2010) look at herding tendencies around the world and find that there is a lot of herding in Asian and developed markets (but not in the US), and that herding gets worse globally during financial crises due to contagion from US markets, which raises systemic risk [9]. Herding is common during market turbulence and increased trade volumes, particularly in China, according to Lao & Singh's (2011) comparison of the Chinese and Indian markets. Because of increased herding during crises, like the 2008 financial crisis, they demand stronger rules and regulations in China [10].

Dang and Lin (2016) examine herding's reliance on heterogeneous information with an emphasis on emerging markets, especially for amateur investors who must contend with high information costs. [11] They believe that these investors frequently imitate their profitable colleagues, a behaviour that necessitates a more distinct factual distinction between genuine and fraudulent herding. In an examination of media-driven herding in the cryptocurrency market during COVID-19, Youssef & Waked (2020) point out how different news sources, influenced by reporting style and media bias, affect investor choices [12]. They uncover media-induced herding trends by using deep learning to improve sentiment accuracy. Ren and Wu (2018) expand on this strategy by measuring herding in blue-chip stocks using forum sentiment analysis [13]. They discover that herding is exacerbated by negative mood and that macroeconomic considerations play a significant role in influencing such conduct.

ISSN 2394 - 7780

III. DATA AND METHODOLOGY

A. Data Collection and Preprocessing 1. News Data Collection

The initial step was to use Google Advanced Search and the WebScraper Chrome extension to scrape news articles from multiple internet sources. The following strategy was used:

Strategy for Searching: News articles from January 1, 2020, to December 31, 2020 were gathered using particular keywords associated with market indices (like the NIFTY50) and economic events (like the economic crisis). Using Google's date filter, the search was honed to collect URLs inside this particular time window.

WebScraping: All recognized URLs were extracted and saved in a CSV file using the WebScraper Chrome extension. The foundation for additional data extraction was this CSV file.

2. Data Extraction

BeautifulSoup Extraction: Each scraped URL was accessed using the BeautifulSoup library in Python to extract relevant fields. The resulting dataset was stored in a structured CSV format, containing the columns: text, url, Published Date, Title, Text, and Summary.

3. Data Preprocessing

The dataset underwent a series of preprocessing steps to clean and standardize the data. Stopwords were removed, NaN values were handled appropriately, and additional text processing techniques such as stemming or lemmatization were applied where necessary. The final dataset, which had a shape of (1437, 6) after preprocessing, formed the input for sentiment analysis.

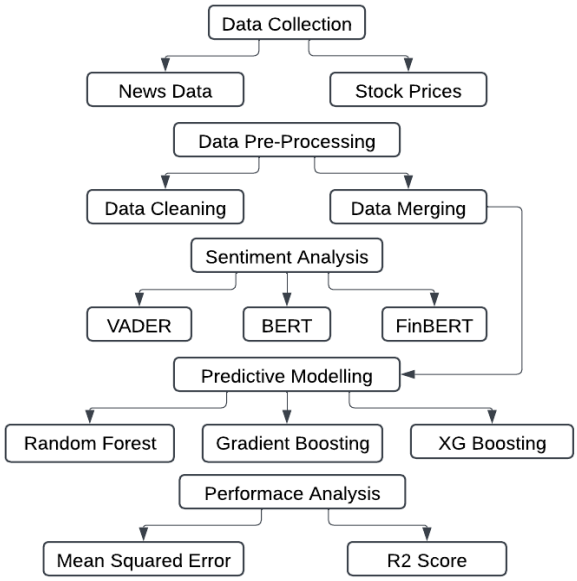


Fig 1: Workflow for sentiment analysis

B. Sentiment Analysis 1. Sentiment Analysis Models

Three pre-trained sentiment analysis models were employed to analyse investor sentiment:

VADER (Valence Aware Dictionary and sEntiment Reasoner): A rule-based model used for general sentiment analysis.

BERT (Bidirectional Encoder Representations from Transformers): A state-of-the-art NLP model that excels in sentiment analysis tasks.

FinBERT: A financial model emerged from BERT and specially used for the analysis of the sentiment of the financial news.

2. Sentiment Scoring

The sentiment scores which have been estimated using each model was then summed up and incorporated with the stock market data. In this study the stock price data of Nifty 50 index stocks were taken including; Open prices, High, Low, Close and the Adjusted close price data. Daily returns were calculated as the first difference of the closing stock prices for a specific day.

ISSN 2394 - 7780

C. Feature Engineering 1. Lagged Features and Event Annotations
New features were included in order to better enhance the model's prediction ability:

Lagged Sentiment Features: As the impact of news on stock prices may initially be delayed, lagged sentiment scores were created.

Event Dates: In order to determine whether or not significant news events have a major effect on market sentiment, some of the most noteworthy events were listed and highlighted.

Moving Averages: To identify trends, moving averages (5 day and 10 day) of sentiment scores were computed.

D. Predictive Modelling 1. Machine Learning Models
Three machine learning models were employed to predict stock returns using the engineered features:

Random Forest Regressor: An iterative learning approach to the prediction of stock return based on sentiment and market characteristics.

Gradient Boosting Regressor: An improved boosting algorithm to capture the intricate features in the data by making refinements on the model's estimations. The accuracy of the models was assessed by means of Mean Square Error (MSE) and R2 score.

Extreme Gradient Boosting: A fast and easily parallelized version of gradient boosting, XGBoost enhances the prediction of outcome by iteratively learning from the mistakes of previous models.

E. Herd Behaviour Analysis 1. Calculation of CSAD and CSSD

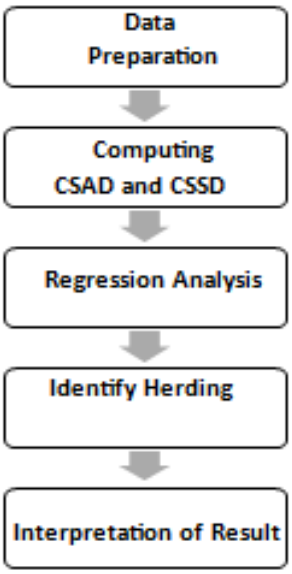


Fig 2: Flow chart For Herding Analysis

To investigate herd behaviour, the Cross-Sectional Absolute Deviation (CSAD) and Cross-Sectional Standard Deviation (CSSD) methodologies were applied: **Market Benchmark:** The market value benchmark was NIFTY-50 index.

Stock Selection: A sample of ten stocks was chosen at random across different sectors for the analysis of individual stock performance against the benchmark index.

CSAD and CSSD Calculation: The CSAD and CSSD values were calculated for the selected stocks used in the analysis. These measures were useful in identifying the variances of specific stock from the overall market index during situations of market volatility or risk.

2. Regression Analysis for Herd Detection
To do this, a regression of the market return on the CSAD/CSSD scores was conducted in an attempt to determine the extent of the association between the two. Nonlinear patterns are expected to suggest the presence of investor herding behaviour, as established by the objective of the study. Specifically, the regression model explored if one of the key herding indicators, namely CSAD or CSSD, reduced when the absolute market return increased.

ISSN 2394 - 7780

IV. RESULTS

A. Sentiment Analysis Results 1. Sentiment Analysis using VADER, BERT, FinBERT

The analysis of sentiment was done using VADER, BERT and FinBERT models. These models estimate the sentiment to turn in news articles that tracked changes in investor emotions following significant news events during COVID-19. FinBERT which is a finance-oriented sentiment analysis model provided a high level of correlation with the market movements of the sentiment scores and highlighted the importance of the application of specialized models in the financial domain.

2. Stock Data Analysis (NIFTY-50)

Besides, the sentiment assessment was performed and compared with stock market data of the NIFTY-50 index. It revealed the change in price trends and returns during the Covid-19 period by focusing on the daily movement of price and other quantitative indicators including moving averages. It allowed a direct comparison of sentiment scores with market behaviour and established a framework for subsequent predictive analysis.

3. Predictive Modelling –

Random Forest, Gradient Boosting and Extreme Gradient Boosting Regressors

A quantitative assessment of the positive relationship between market investor sentiments and stock market return was performed through the use of predictive analytics. In this paper, this research aimed at predicting complex, nonlinear relationships between the scores of sentiment and market returns using machine learning methods including Random Forest, Gradient Boosting and Extreme Gradient Boosting Regressors. These models are suitable for pioneering financial market forecasting due to their ability to capture numerous interactions in the relationship between stock returns and investors, due to the consideration of the patterns of large quantities of data. The main findings from the predictive models are presented in table 1.

Table 1: Model Performance

Model \ Metric	R2 Score	Mean Square Error (MSE)
Random Forest Regressor with Cross Validation	0.8668	0.000511
Random Forest Regressor with Hyperparameter Tuning	0.8668	0.000511
Gradient Boosting with Cross Validation	0.8961	0.000398
Gradient Boosting with Hyperparameter Tuning	0.8950	0.000403
Ensemble Model (combination of RF and GB)	0.8821	0.000452
Extreme Gradient Boosting	0.9017	0.00037
Extreme Gradient Boosting Hyperparameter Tuning	0.9155	0.00032

These findings show that the models were able to explain the degree of association between the sentiment indicators and the stock returns evidenced by high R^2 scores and low MSE values. Out of the models, specifically, the Extreme Gradient Boosting Regressor showed high accuracy, which implies that sentiment has a strong influence in predicting the stock market during such episodes as the COVID-19 crisis.

CUMULATIVE ABNORMAL RETURNS (CAR) ANALYSIS

To examine market reaction to specific major news events during the pandemic, the analysis of Cumulative Abnormal Returns (CAR) was done. CAR sums the excess of stock return over the expected amount around event windows and can explain how the market views and responds to significant events. The following study looked at CAR around important event dates in order to find out whether event specific information altered investor sentiment with a resultant effect on stock market returns.

Table 2: CAR Analysis result

Event Date	Value
March 24, 2020 (Nationwide lockdown announced)	-0.1464
April 15, 2020 (Lockdown extended)	-0.2194
May 1, 2020	0.0249

ISSN 2394 - 7780

(Lockdown extension)	
July 1, 2020 (Unlock phase 2.0)	-0.6657

These values show relatively significant market response to the major events evidenced by increased investors' risk mitigation and negativity during the event releases. That is why cases like on March 24 and April 15, which show negative CAR values but higher investor concerns due to the pandemic, indicate the impact of sentiment.

B. Herding Behaviour Analysis 1. Cross-Sectional Absolute Deviation (CSAD) and Cross-Sectional Standard Deviation (CSSD) Results

In the CSAD and CSSD studies, herding behaviour among the investors was detected. It is a phenomenon whereby investors all act in unison instead of independently by keeping in tow with the market.

Table 3: Regression result for CSAD and CSSD

	CSAD Analysis	**CSSD Analysis**
Intercept Coefficient	0.0139	0.0174
Coefficient for Absolute Market Return	0.1222 (p-value: 0.003)	0.1616 (p-value: 0.002)
Coefficient for Squared Market Return	-1.0614 (p-value: 0.018)	-1.3777 (p-value: 0.017)
R^2 Score	0.048	0.051

Thus, negative Squared Market Return coefficients in both models further indicate that the larger the market returns, the lower the dispersion of returns. This study provides evidence for the presence of herding behaviour, that is; the investors seem to act as one in the course of the period of high volatility in the market.

Using Regression Analysis to Identify Herding: The regression model was used to analyse the relationship between CSAD/CSSD values and market returns. Herding behaviour was confirmed by the large negative coefficient for Squared Market Return, which indicated that as the figure of market returns increased (up or down) the variability of returns decreased.

Volume-Return Relationship and Correlation Analysis: Contrary to conjecture, this paper did not establish any strong evidence of herding, based solely on trading volumes, when trading volume was correlated with stock returns. Yet, the results of the CSAD and CSSD provided some evidences to herding and indicated that the investor choices were closer to market returns rather than trade frequency.

V. CONCLUSION

In the present research while testing the hypothesis, the NIFTY-50 index was taken as the benchmark to study how important event triggered variation in investors' perception and herd mentality in the Indian stock market during the COVID-19 outbreak. Market and news data indicators were combined with sentiment scores derived with help of stateof-art sentiment analysis models such as VADER, BERT and FinBERT. There was a significant relationship between sentiment and stock market returns which was established using Random Forest and Gradient Boosting Regressors for predictive modelling. CAR research also showed that investor attitude came into play during uncertain times as changes in market behaviour were found to be due to significant news events.

Furthermore, cross sectional measures (CSAD and CSSD) were employed to analyse the herd behaviour. The results showed that the investors acted in a herd, the evidence being when return dispersion declined during periods of high turbulence. In conclusion, this work establishes an understanding of how news-driven emotion influences market returns and how often people tend to herd in emergency situations. These findings extend the existing literature on how external factors affect the decision-making process of investors and offer useful information about the psychological processes that underlie trading in finance.

ACKNOWLEDGEMENT

I would like to express my gratitude to my college faculty for their invaluable guidance and support during my research work. I am immensely thankful of my parents' and sister's unwavering support and faith, which served as my inspiration. Rahul deserves special mention for his insightful feedback and recommendations. This paper would not have been possible without all of you.

ISSN 2394 - 7780

REFERENCES

1. A. J. Hanna, J. D. (2020). "News media and investor sentiment during bull and bear markets,". The Eauropean Journal of Finance.
2. I. Verma, L. D. (2017). "Detecting, Quantifying and Accessing impact of News events on Indian Stock Indices,". Association for Computing Machinery.
3. J., A. O. (2019). "Does Behavioural Biases Influences Individual Investment,". Management Science Review, vol. Vol 10(1).
4. Lee, M. -Y.-.. (2016). "Deep Learning for Financial Sentiment Analysis on Finance News Providers," . International Conference on Advances in Social Networks Analysis and Mining (ASONAM), San Francisco, CA.
5. Lin, H. V. (2016). "Herd Mentality in the Stock Market: On the Role of Idiosyncractric Participants with Heterogeneous Informtion". SSRN.
6. Othman, N. N. (2024). "Emotional Economics: The Role of Psychological Biases in Personal Investment Outcomes,". SSRN.
7. P. Papakyriakou, A. S. (2019). "Impact of terrorist attacks in G7 countries on international stock markets and the role of investor sentiment,". Elsevier.
8. Pagolu, V., K.N., R., G., P., & B., M. (2016). "Sentiment Analysis of Twitter Data for Predicting Stock Market Movements," International conference on Signal Processing, Communication, Power and Embedded System.
9. R. Cai, B. Q. (2020). "Sentiment Analysis About Investors and Consumers in Energy Market Based on BERTBiLSTM,". IEEE Access.
10. Sharma, S. B. (2001). "Herd Behavior in Financial Markets,". IMF Staff Papers, vol. Vol. 47.
11. Singh, P. L. (2011). "Herding Behaviour in the Chinese and Indian Stock Markets,". SSRN.
12. Waked, M. Y. (2022). "Herding behavior in the crypocurrency market during COVID-19 pandemic: The role of media coverage. Elsevier.
13. Wu, R. R. (2018). An Innovative Sentiment Analysis to Measure Herd Behavior," IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEM.
14. Zheng, T. C. (2010). "An empirical analysis of herd behavior in global stock markets,". Elsevier.

International Journal of Advance and Innovative Research
Volume 12, Issue 2 (XVII): April - June 2025
ISSN 2394 - 7780

ANIME RECOMMENDATION CHATBOT USING HYBRID FILTERING & TRANSFORMERS WITH IMPACT OF GENRE DIVERSITY

[1]Vipul Jadhav and [2]Sunita Jena
[1]Student and [2]Assistant Professor, MSc. Big Data Analytics Department, Jai Hind College, Mumbai, India

ABSTRACT
This research paper proposes an anime recommendation chatbot that incorporates the aspects of filtering techniques as well as the transformer models to perform the recommendation session; however, the importance of the aspect of the diverse genre is stressed here. Anime itself is a peculiar kind of tv shows which has its own specific features; the mainstream shows satisfaction relation by genres is quite variable. This work combines the content-based and hybrid filtering models with the transformer model to know how diversification based on genre impacts the precision of the recommendations and user relevance. Making use of an extensive anime dataset containing not only genre and rating, but also user interaction indicators, we then compare the performance of each model. The chatbot goes further than recommending and providing alternatives and questions users to select genres and provide their feedback in order to improve them continuously. Consequently, findings show that our hybrid model in tandem with the use of variety in genres yields recommendation precision and user satisfaction levels above traditional practices. Also, the conversational component of the chatbot based on transformers increases the level of user's satisfaction using the product due to the natural and individual approach to the dialogue. The present work stresses the significant of genre variety in media recommendations and recommends future research in implementing higher level of filtering and natural language processing to feedbacks for enhancing the recommender systems.

Keywords— *Content-based filtering, Hybrid Filtering, Genre Diversity, Recommendations, Transformers.*

I. INTRODUCTION

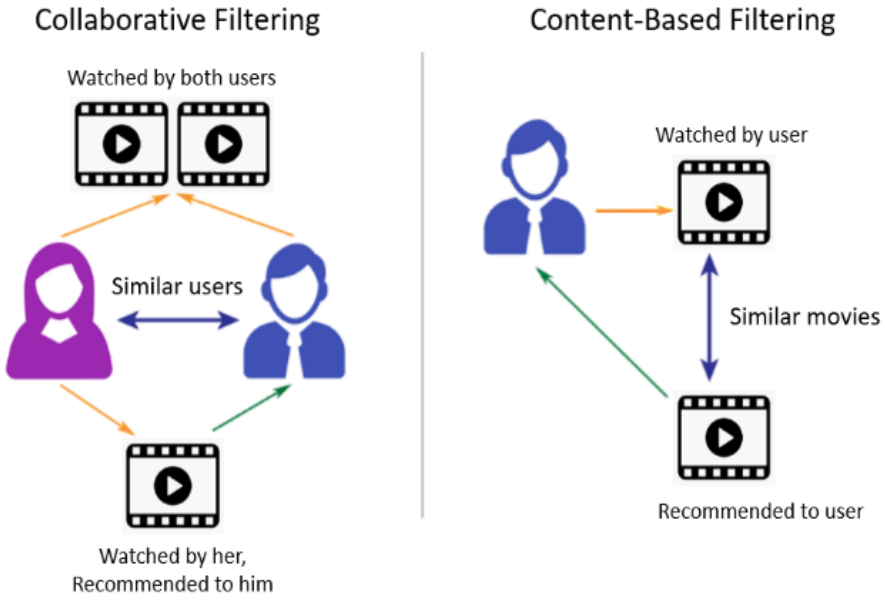


Fig 1. Collaborative filtering vs Content-Based Filtering Approach.

Animation specifically animation originating from Japan known as anime has cut across the globe it is famous irrespective of culture or language. Lately, there has been a tremendous increase of streaming services that now the spectator faces the problem of choice and it seems for the consumer even more difficult to find anime to their taste. Conventional recommendation techniques, despite their applicability, fail to propose diversified relevant items from which users get bored and dissatisfied (Abhipsa Jenaa, 2022). Popular approaches in recommendation systems in which content-based filtering and collaborative filtering are the most commonly used are distinctly different and involves making of recommendations by different methods. CBF, on the other hand, operates by examining attributes of the items that a user has, at some point, consumed or indicated penchant for (Ashwin Pillai, 2023). To illustrate, applying content-based filtering in an anime recommendation system takes into consideration of the genre, theme or attributes of previous similarly recommended anime. This approach constructs a user image and suggest items with different characteristics in relation to the preferred items. The first advantage of content-based filtering is that it does not depend on other users' data even when

ISSN 2394 - 7780

the interaction between users and items is quite limited. But it may not do a good job in recommending fresh products or services which a user may find interesting since it mainly suggest similar items without going outside the user's zone of interest. Whereas collaborative filtering suggest items based on the behaviours of the consumers in a large user population. It assigns like-minded users (user-based collaborative filtering) or it tries to find resemblance of items based on user feedbacks or transactions (item-based collaborative filtering). With regards to anime, collaborative filtering could then suggest a particular show to a user in the basis of a similar pattern to other users. This approach is powerful because it enables the discovery of items that might be outside the user's immediate interests, providing more diverse recommendations (Ashwin Pillai, 2023) (M Viswa Murali, 2019) But, the approach of collaborative filtering is limited by problems of 'cold-start' whereby its results decline significantly for new users or items which have been minimally used. Integrating both strategies usually produces better recommendation systems improving users' satisfaction. In the context of recommendation systems the genre diversity approach is concerned with bringing more variety in genres of the items that are recommended to users especially in the context of the anime where genre matters most. As opposed to constantly suggesting other similar anime's, this approach brings a little bit more of the diversified suggestions by introducing other genres along side recommendations. The idea here is to satisfy the known and introduce the unknown where the genre diversity more tend to make the users to step outside their typical preference. It is more helpful for the future use as well because it minimizes the chances of user gets bored with the same kind of content or product that it is recommending consistently. In the case of anime recommendation system, diversity might be achieved by recommending the anime from less popular genres besides ones that are popular. That is, if a user tends to watch action anime, a genre divergence strategy may show anime from adventure, fantasy or mystery genres, so that the user will expand his/her anime watching experience. This methodology paves way for reducing echo chamber where recommendations are highly relevant to the users previous activity. With the help of establishing focus on genre diversity the system provides opportunities to discover various content to the users intensifying their interests and increasing the level of user satisfaction with the platform.

II. METHODOLOGY

A. Data Collection and Preprocessing:

The approach used for the anime recommendation chatbot is to create a mixed recommendation system that improves on content-based filtering, collusive filtering and the genre diversification. The process starts by data preprocessing aiming at refining our data set of anime to a ready for analysis format. This dataset contains basic features which are genre, ratings where clients can rate movies, popularity, and users' interaction allowing different analyses and recommendations.

B. Content-Based Filtering:

Content-based filtering is done on the basis of attributes of each anime, with special reference to genre in order predict potentials of a title that the user is likely to enjoy. This approach uses the feature of cosine metric in order to come up with similar anime and recommend them based on the user preference (Nuurshadieq, 2020).

C. Collaborative Filtering:

The former is employed to broaden personalization by applying data from many users jointly. It gets its recommendations from common behavior among users in similar cases or categories. Co-occurrence techniques like Matrix factorization (for instance Non-negative Matrix Factorization or Singular Value Decomposition) are used to predict the missing preferences with the help of the available user –item interactions (M Viswa Murali, 2019) (Reynaldi, 2023).

D. Hybrid Model with Genre Diversity:

A new recommendation system which combines the content filtering and collaborative approaches is designed. It is used as a remedy for recommended contents' repetition to increase variety in the interests of customers. This is done by flexing the recommendation system so that it there is an input of other genres in order to help to enhance the experience of the users. The hybrid model incorporates both, genre interests and user histories to offer a fair mix of the commonly selected genres and the general preferences.

E. User Interface:

The front-end interface of the chatbot is created using Streamlit a free and open-source command-line application that is useful for creating web applications. As a result of Streamlit, the interface of the data visualization of the dashboard can be created by anyone writing minimal codes, be it data scientists or developers. Learners can conveniently use the buttons to choose the genres or type of content they are interested in, and interact better. For this reason, the locally hosted application used ngrok which allows external access to an application hosted locally. This tool makes connection tunnels to let the local server expose the application to

ISSN 2394 - 7780

the internet so that users could engage the chatbot from any location. With this I have realized that having a short-term key, ngrok makes it quite easy to share and check the Streamlit application, making it a crucial tool for group development, and, interacting with users. Combined they add value as Streamlit opens up a platform for real- time user engagement while ngrok adds to it by providing dynamic recommendations.

III. RESEARCH AND ANALYSIS

A. Overview:

Here we describe the results of our anime recommendation chatbot based on hybrid filtering algorithm and transformer model. By examining the results, our discussion focuses on the suggestion relevance and the role of genre variety in the recommendations as well as the potential of this work compared to previous studies' findings.

B. Recommendation Outcomes:

As suggested by the name the anime recommendation chatbot is designed to give the user information about anime that they might like based on the user's behavior/feedback. In the experiments with different commands, we noticed that the chatbot provides a great number of recommendations of anime with focusing on the preference of the user. It showed that feedbacks were given to the users mostly according to titles they had watched before, proving that the system was able to infer users' preferences correctly.

C. Impact of Genre Diversity:

In one of the areas of our study, it was vital to understand the specifics of the recommendations that the chatbot made dependent upon the genre variety. As much as everyone hopes to get good recommendations, which is the basic job of a chatbot, they said that the reaction was exciting when the chatbot was suggesting the titles from different genres including those that are not of their types. Said feature prompted users to expand their horizons when it comes to anime and watch various anime shows. The fact that the chatbot was connecting between genres allowed for finding titles that the users normally would not come across (Debby Cintia Ganesha Putri, 2020). By reviewing the statistics, our team realized that the clients with a somewhat broader range of genres mentioned the watched content was far more diverse, which supports the notion that genre is an essential aspect to consider when recommending programs. Users were pleased with a list of recommendations with a choice of different categories, stating that they have changed and are ready for new experiments. This finding redresses a major omission in past research where the importance of the approach proposed here – that is, using information about the genre of items to make recommendations – was ignored in favor of more rudimentary forms of collaborative or content-based filtering.

D. Comparative Analysis with Existing Literature:

Despite the fact that there are a number of prior research papers dealing with collaboration or with content-based filtering techniques our approach which interposes the two techniques partly fills a significant evidence gap. While most investigations pay attention to the aspect of accuracy of the recommendation, little attention is paid to the aspect of how different genre exposure influences the satisfaction and level of interest that users have in the content (M Viswa Murali, 2019) (Ashwin Pillai, 2023). Bringing the concept of genre diversity in to the heart of the recommendation approach, we support a more comprehensive view to the key aspects of the users' interest and interaction. Also, our study suggests directions for further examination of the generic personalization options of the web application. Few of the reviewed studies also fail to explain how the genre heterogeneity could be incorporated in the recommendation algorithm. In this way our approach is free from certain disadvantages of typical approaches and helps to find ways for further development of using user preference data in recommendation systems.

E. Conclusion of Results:

Thus, according to the findings of the present study, the proposed anime recommendation chatbot helps to offer suitable and interesting anime suggestions concerning users' preferences. The use of genre inclusion hugely improves the recommendations given in order to expand the range of anime explored and enjoyed. This section underlines the need to consider use-interaction and preference data for constructing recommendation systems for future improvements and new attributes to meet the identified gaps in the literature.

IV. IMPLEMENTATION

There are several important phases as crucial steps in the system to deploy the anime recommendation chatbot that aims for an efficient recommendation service for users with the help of combining the several technologies and techniques. This subtopic provides an overview of the major parts of implementation such as data preprocessing and preparation, analysis and planning for algorithms, designing and planning for the graphical user interface and deploying the software.

ISSN 2394 - 7780

A. Data Preparation:

The first stage of the implementation of the project focuses on data collection and data preparation. This dataset is the Anime-2023 containing information about anime titles, genres, rating, its users' behavior and their interactions. As with any data, the quality of the dataset has to be reviewed during the preprocessing phase.:

1) **Data Cleaning:** Data gaps are filled, and the naming convention of different genres is rationalized.
2) **Normalization:** Concrete values of Favorites, Scored By, and Members are preprocessed by a MinMaxScaler to make all of these features similarly important for the recommendation algorithms**.**
3) **Encoding:** There is one more categorical feature – Genres We will one-hot encode Genres so that the recommendation system can understand and utilize the genres properly.

Fig 2: Correlation Heatmap of the Variables.

B. Algorithm Development:

The key component of the chatbot is the recommendation algorithms, which includes both content and collaborative approaches. The implementation process includes:

1) **Content-Based Filtering:** In this component, the likes attribute of anime titles previously given by the user is examined based on certain attributes. With the help of cosine similarity the features of these titles are compared with the other titles in the dataset in order to suggest the similar anime based on the genre and other attributes**.**
2) **Collaborative Filtering:** Collaborative filtering is achieved employing Non-negative Matrix Factorization (NMF). This training algorithm determines hidden features in the user-item relations and allows the system to offer the most suitable anime acquaintances as inferred from similar users. This approach thus handles the cold-start problem well indeed, by capitalizing on the aggregate action records to bolster the recommendations.

ISSN 2394 - 7780

```
# Step 2: Combine the textual and numerical features for Content-Based Filtering
anime_df_cleaned['Combined_Features'] = anime_df_cleaned['Genres'] + ' ' + anime_df_cleaned['Episodes'].astype(str)

# Step 3: Convert the textual features (Genres and Episodes) into a TF-IDF matrix for Content-Based Filtering
tfidf = TfidfVectorizer(stop_words='english')
tfidf_matrix = tfidf.fit_transform(anime_df_cleaned['Combined_Features'])

# Step 4: Content-Based Similarity
content_cosine_sim = cosine_similarity(tfidf_matrix, tfidf_matrix)

# Step 5: Item-Based Collaborative Filtering using numerical features
# Using Scored By, Rating Score, Favorites, Members
item_based_features = anime_df_cleaned[['Rating Score', 'Favorites', 'Scored By', 'Members']]

# Compute cosine similarity for item-item collaborative filtering
item_cosine_sim = cosine_similarity(item_based_features, item_based_features)
```

Fig 3: Snapshot of Code for Content & Collaborative filtering.

3) **Hybrid Model with Genre Diversity:** While implementing the hybrid model, the two approaches namely content-based and collaborative filtering approaches are used, in addition to the Genre Diversity Score (GDS). It contributes toward proposals that encompass multiple genres and thus drives its users into searching through a wide list of animes**.**

```
# Function to calculate Genre Diversity Score (GDS)
def genre_diversity_score(recommendations, genres_df):
    # Extract genres of recommended items
    genre_counts = genres_df.iloc[recommendations].sum(axis=0)
    unique_genre_count = (genre_counts > 0).sum()

    # Normalize by total possible genres for a score between 0 and 1
    total_genres = genres_df.shape[1]
    gds = unique_genre_count / total_genres
    return gds

# Function to get content-based recommendations with diversity factor
def get_diverse_recommendations(anime_index, top_n=10, diversity_factor=0.5):
    # Get similarity scores
    similarity_scores = list(enumerate(similarity_matrix[anime_index]))
    similarity_scores = sorted(similarity_scores, key=lambda x: x[1], reverse=True)
```

Fig 4: Snapshot of Code for Genre Diversity.

C. User Interface Development:

The graphical user interface of the chatbot is implemented with Streamlit, which is an effective library that provides implementation of a variety of web applications. The Streamlit has flexible features that help in faster real-time prototyping so that a developer can create a clean and well-responsive app. Among the features of the UI that worth mentioning,:

1) **Genre Selection:** Using a dropdown list that appears, users can choose their required genres that in turn dictate the suggestions made by the chatbot**.**

2) **Conversational Interface:** The chatbot is controlled via a text input box where users can request for recommendations or enquire about any show including an anime. The chatbot needs to convert an arrangement (e.g., DialoGPT) that will make sense to the user in terms of conversation in a conversational way**..**

ISSN 2394 - 7780

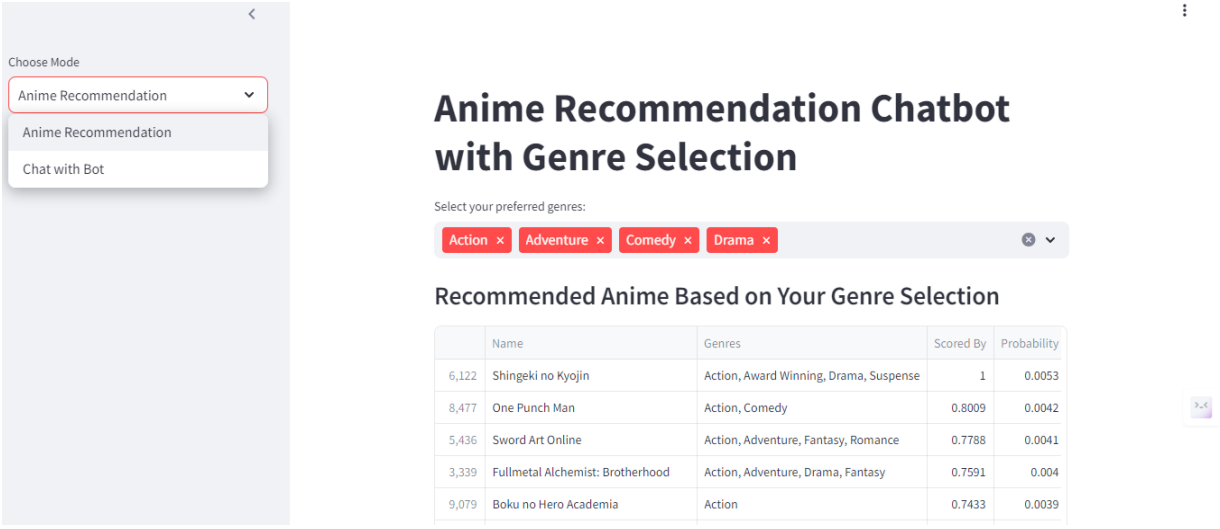


Fig 5: Snapshot of Chatbot using Streamlit & ngrok.

D. Deployment:

For this, ngrok is used to expose the local development environment and assign virtual internet URLs for the access of the chatbot. This leads the user to accessing the chatbot via the address that is reachable from any internet connection. The process by which it was migrated can be explained as follows:

1) **Running the Streamlit App:** The chatbot runs locally, and ngrok is used to forward incoming requests to the locally running server and provide an outside link**.**

2) **Real-Time Interaction:** While chatbot, users are also able to plunge in real time given their genres of choice and inquiries about anime.

V. FUTURE WORK

Although there has been a successful attempt made and implemented in this paper with the anime recommendation chatbot, there are several ways on how this can be expanded and advanced. A subsequent area for enhancement is a high level of user profiling and personalization. Currently, recommendations are offered from the genres, which have been chosen and the user's historical behavior; but if the system could capture implicit behavioral trends such as popular genres and kinds of interactions, it would be possible to create more refined user profiles. It was envisaged that by the incorporation of reinforcement learning, it was possible to offer real-time recommendations that could meet the specific needs of the users. Another prospective avenue of research is in bringing together emotion and sentiment analysis, which would help the chatbot to track the emotions tied to genres/themes as interpreted by clients/embedded within their feedback. This could lead to bias recommendations where for example, comedy anime be recommended because it makes the user feel good or serious anime be recommended because the user likes mystery. This enhancement to the chatbot, will help the application decipher the users' intent much better and offer the users a more rich experience. Evaluating the chatbot to as a cross-domain recommendation system could also add value in the recommendation process by recommending books, movies or music within the anime genre or theme. Some transfer learning strategies could help extend the findings obtained through anime-based insights to other contexts; offering a rich recommendation experience. Updating genre diversity indicators can impact the user engagement more inclined, using a new methodology that combines popular and underrepresented genres, and customized diversity based on the user's exploring behavior. Such an analysis of the response of the audience for the variation of the type of genres within particular intervals can enhance this diversity factor and reinforce that recommendations are simultaneously both contemporary and interesting. However, a real-time feedback system integrated within the functional use of the chatbot assures more authentic and dynamic recommendations. As users', they could follow such recommendations with "like" or "dislike", to allow the system to modify the recommendations based on the changing preferences. Easing the access of the app for a larger populace of users and optimization of гpуп processing infrastructure such as AWS or Google Cloud for large volume data and real time updates. Cloud deployment would also allow using such valuable infrastructure features as load balancing and distributed processing to improve the chattiness of the chatbot.

Given how popular transformers are in conversational AI, it might also be worth it to try and incorporate this into recommendation models. Sequence patterns can be captured easily through the use of transformers and therefore can be easily used when it comes to tracking the dynamic user preference. If we infused the chatbot with a transformer-based recommendation model, get efficient suggestion outcome in a very contextually

ISSN 2394 - 7780

relevant way, which knows the user interaction history. The future directions proposed for the research lay out exciting possibilities for developing the anime recommendation chatbot into a more malleable, customer-centric, and contextual recommendation tool for anime aficionados.

VI. CONCLUSION

In this paper, we proposed an anime recommendation chatbot based on hybrid filtering techniques and transformer models and added genre diversification into consideration. The recommendations our approach provided were designed to overcome some of the typical flaws of conventional systems by considering both thematic density and average distance between articles to provide readers with diverse yet relevant articles. Fusing the two approaches of content-based and collaborative filtering the recommended items are selected from using both the user's specific preferences and the overall tendencies in the community. Among the key objectives of this study, the development of the Genre Diversity Score (GDS) that promote recommendations of a richer variety of genres is noteworthy. This feature helps in solving the problem of recommendation where the devices suggest choosing a specific channel or genre most of the time. By diversifying the genres of recommended shows, the chatbot opens new doors to the users as to which shows they can proceed watching and thus eventually helps them broaden their interests and to watch more shows they might enjoy.. Introducing genre diversity demonstrates that originality is a beneficial attribute for recommender systems because it speaks to the value of incorporating variety in formats for a wide range of users. Also, using transformers in conversation or chatbot applications, make this program more friendly and interactive to use. The Dynamic Client Interface allows the users to enter preferences, pose inquiries or get recommendations in real time. This aspect brings the change from a static source of recommendations to a more engaging recommendation conversation to match the dynamic shift in user interests. By using Streamlit, the interface of which is easily accessible through the application of ngrok, The chatbot's interface is very friendly and easy to use, effective for anime lovers. This research lays out the groundwork for new developments in recommendation systems that made user participation and discovery a focus. This is why, given the specific improvements made to user profiling, sentiment analysis, and cross-domain recommendations in this study, there is a potential for future work to develop more advanced and enjoyable experiences in recommendations. To sum up, the multifunctional and genre interdisciplinary phenomena of the chatbot as a pattern of effective and user-oriented approach to the individual recommendation method tunes the new path towards progressive further developments of the recommendation systems in the sphere of anime and other areas as well.

ACKNOWLEGEMENT

I am indeed thankful to my guide Mrs. Sunita Jena for her precious guidance, cooperation and overall motivation. I am indeed grateful for her foresight, professionalism and especially dedication, which has enormously guided this research. Therefore, I am honoured to express my heartfelt gratitude to the Principal, Jai Hind College, Mumbai University for providing me with the essential computer, facilities and the right atmosphere to accomplish this research study.

I am also grateful to friends and colleagues who provided their comments, contribution, and encouragement to improve this research and also to my family. I would like to thank them for their greatest support in this work they have been really patient with me and encouraging me all the time. Finally, it is worth to thank all people who created free and open- source tools, datasets and libraries that let to perform this research.

REFERENCES

1. Abhipsa Jenaa, A. J. (2022). Recommendation System For Anime Using Machine Learning Algorithms. ICCIDS-2022, 7.
2. Ashwin Pillai, A. M. (2023). Unveiling Anime Preferences: A Data-driven Analysis using MyAnimeList API. International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 9
3. Debby Cintia Ganesha Putri, J.-S. L. (2020). Design of an Unsupervised Machine Learning-Based Movie Recommender System. 27.
4. M Viswa Murali, V. T. (2019). A Collaborative Filtering based Recommender System for Suggesting New Trends in Any Domain of Research. 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), 4.
5. Nuurshadieq, A. T. (2020). Leveraging Side Information to Anime Recommender System using Deep learning. 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 6.

ISSN 2394 - 7780

6. Reynaldi, W. I. (2023). Content-based Filtering and Web Scraping in Website for Recommended Anime. Asian Journal of Research in Computer Science, 11.

7. Syoichiro Ota, H. K. (2017). AniReco: Japanese Anime Recommendation System. International Federation for Information Processing, 4.

MANUSCRIPT SUBMISSION

GUIDELINES FOR CONTRIBUTORS

1. Manuscripts should be submitted preferably through email and the research article / paper should preferably not exceed 8 – 10 pages in all.
2. Book review must contain the name of the author and the book reviewed, the place of publication and publisher, date of publication, number of pages and price.
3. Manuscripts should be typed in 12 font-size, Times New Roman, single spaced with 1" margin on a standard A4 size paper. Manuscripts should be organized in the following order: title, name(s) of author(s) and his/her (their) complete affiliation(s) including zip code(s), Abstract (not exceeding 350 words), Introduction, Main body of paper, Conclusion and References.
4. The title of the paper should be in capital letters, bold, size 16" and centered at the top of the first page. The author(s) and affiliations(s) should be centered, bold, size 14" and single-spaced, beginning from the second line below the title.

First Author Name1, Second Author Name2, Third Author Name3

1Author Designation, Department, Organization, City, email id

2Author Designation, Department, Organization, City, email id

3Author Designation, Department, Organization, City, email id

5. The abstract should summarize the context, content and conclusions of the paper in less than 350 words in 12 points italic Times New Roman. The abstract should have about five key words in alphabetical order separated by comma of 12 points italic Times New Roman.
6. Figures and tables should be centered, separately numbered, self explained. Please note that table titles must be above the table and sources of data should be mentioned below the table. The authors should ensure that tables and figures are referred to from the main text.

EXAMPLES OF REFERENCES

All references must be arranged first alphabetically and then it may be further sorted chronologically also.

- **Single author journal article:**

Fox, S. (1984). Empowerment as a catalyst for change: an example for the food industry. *Supply Chain Management*, *2*(3), 29–33.

Bateson, C. D.,(2006), 'Doing Business after the Fall: The Virtue of Moral Hypocrisy', Journal of Business Ethics, 66: 321 – 335

- **Multiple author journal article:**

Khan, M. R., Islam, A. F. M. M., & Das, D. (1886). A Factor Analytic Study on the Validity of a Union Commitment Scale. *Journal of Applied Psychology*, *12*(1), 129-136.

Liu, W.B, Wongcha A, & Peng, K.C. (2012), "Adopting Super-Efficiency And Tobit Model On Analyzing the Efficiency of Teacher's Colleges In Thailand", International Journal on New Trends In Education and Their Implications, Vol.3.3, 108 – 114.

- **Text Book:**

Simchi-Levi, D., Kaminsky, P., & Simchi-Levi, E. (2007). *Designing and Managing the Supply Chain: Concepts, Strategies and Case Studies* (3rd ed.). New York: McGraw-Hill.

S. Neelamegham," Marketing in India, Cases and Reading, Vikas Publishing House Pvt. Ltd, III Edition, 2000.

- **Edited book having one editor:**

Raine, A. (Ed.). (2006). *Crime and schizophrenia: Causes and cures.* New York: Nova Science.

- **Edited book having more than one editor:**

Greenspan, E. L., & Rosenberg, M. (Eds.). (2009). *Martin's annual criminal code:Student edition 2010.* Aurora, ON: Canada Law Book.

- **Chapter in edited book having one editor:**

Bessley, M., & Wilson, P. (1984). Public policy and small firms in Britain. In Levicki, C. (Ed.), *Small Business Theory and Policy* (pp. 111–126). London: Croom Helm.

- **Chapter in edited book having more than one editor:**

Young, M. E., & Wasserman, E. A. (2005). Theories of learning. In K. Lamberts, & R. L. Goldstone (Eds.), *Handbook of cognition* (pp. 161-182). Thousand Oaks, CA: Sage.

- **Electronic sources should include the URL of the website at which they may be found, as shown:**

Sillick, T. J., & Schutte, N. S. (2006). Emotional intelligence and self-esteem mediate between perceived early parental love and adult happiness. *E-Journal of Applied Psychology, 2*(2), 38-48. Retrieved from http://ojs.lib.swin.edu.au/index.php/ejap

- **Unpublished dissertation/ paper:**

Uddin, K. (2000). A Study of Corporate Governance in a Developing Country: A Case of Bangladesh (Unpublished Dissertation). Lingnan University, Hong Kong.

- **Article in newspaper:**

Yunus, M. (2005, March 23). Micro Credit and Poverty Alleviation in Bangladesh. *The Bangladesh Observer*, p. 9.

- **Article in magazine:**

Holloway, M. (2005, August 6). When extinct isn't. *Scientific American, 293*, 22-23.

- **Website of any institution:**

Central Bank of India (2005). *Income Recognition Norms Definition of NPA.* Retrieved August 10, 2005, from http://www.centralbankofindia.co.in/ home/index1.htm, viewed on

7. The submission implies that the work has not been published earlier elsewhere and is not under consideration to be published anywhere else if selected for publication in the journal of Indian Academicians and Researchers Association.

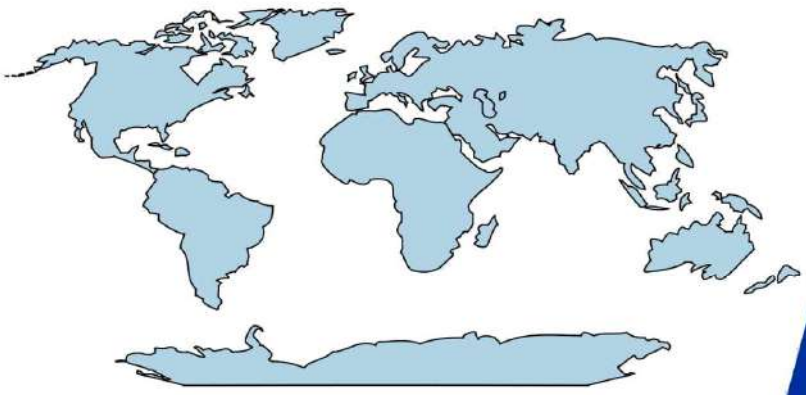
8. Decision of the Editorial Board regarding selection/rejection of the articles will be final.

www.iaraedu.com

Journal

ISSN 2322 - 0899

INTERNATIONAL JOURNAL OF RESEARCH IN MANAGEMENT & SOCIAL SCIENCE



Volume 8, Issue 2
April - June 2020

www.iaraedu.com

Journal

ISSN 2394 - 9554

International Journal of Research in Science and Technology

Volume 6, Issue 2: April - June 2019

Indian Academicians and Researchers Association
www.iaraedu.com

INDIAN ACADEMICIANS AND RESEARCHERS ASSOCIATION

Membership No: M / M – 1365

Certificate of Membership

This is to certify that

XXXXXXXX

is admitted as a

Fellow Member

of

Indian Academicians and Researchers Association

in recognition of commitment to Educational Research

and the objectives of the Association

Date: 27.01.2020

Director

President

INDIAN ACADEMICIANS AND RESEARCHERS ASSOCIATION

Membership No: M / M – 1365

Certificate of Membership

This is to certify that

XXXXXXXXXX

is admitted as a

Life Member

of

Indian Academicians and Researchers Association

in recognition of commitment to Educational Research

and the objectives of the Association

Date: 27.01.2020

Director

President

INDIAN ACADEMICIANS AND RESEARCHERS ASSOCIATION

Membership No: M / M – 1365

Certificate of Membership

This is to certify that

XXXXXXXX

is admitted as a

Member

of

Indian Academicians and Researchers Association

in recognition of commitment to Educational Research

and the objectives of the Association

Date: 27.01.2020

Director

President

IARA Organized its 1st International Dissertation & Doctoral Thesis Award in September'2019

1st International
Dissertation & Doctoral Thesis Award
(2019)

Organized By

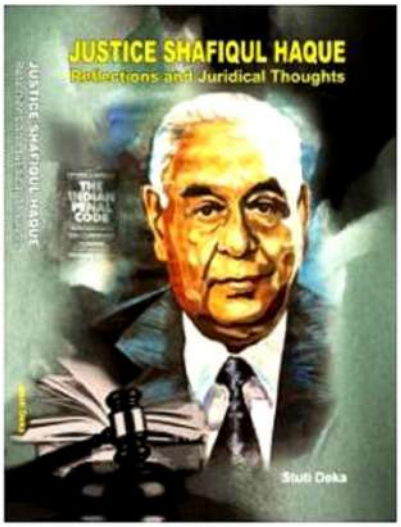
Indian Academicians and Researchers Association (IARA)

EMPYREAL PUBLISHING HOUSE

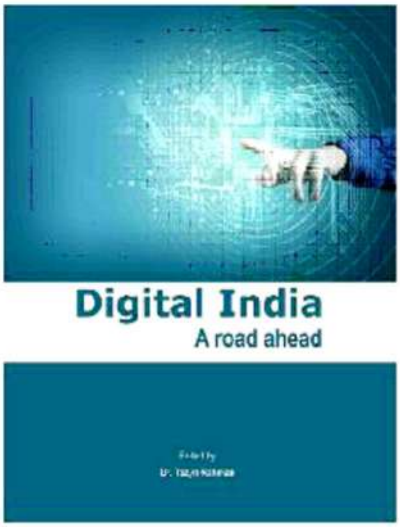
www.editedbook.in

Publish Your Book, Your Thesis into Book or Become an Editor of an Edited Book with ISBN

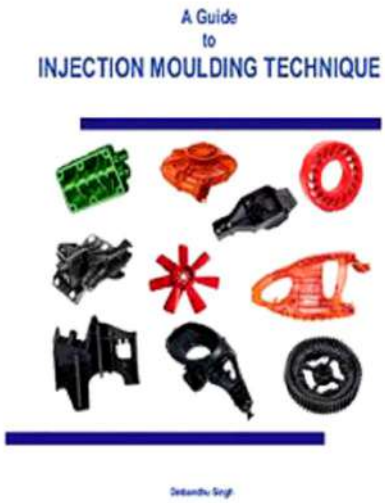
BOOKS PUBLISHED



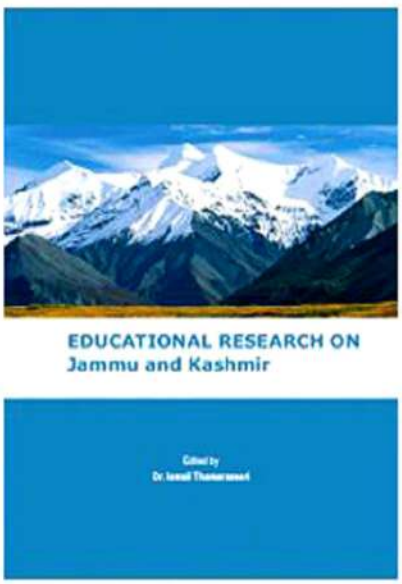
Dr. Stuti Deka
ISBN : 978-81-930928-1-1



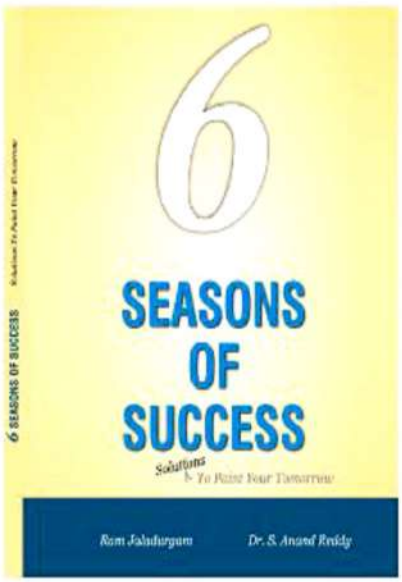
Dr. Tazyn Rahman
ISBN : 978-81-930928-0-4



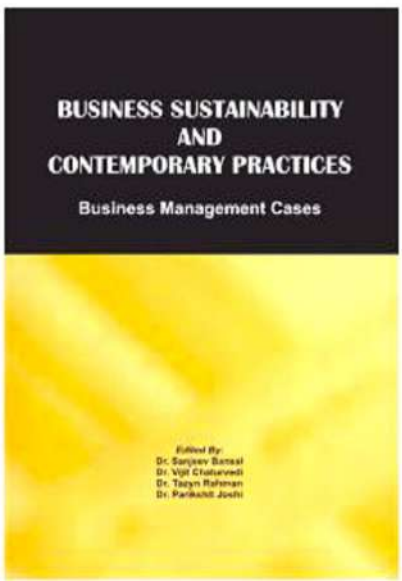
Mr. Dinbandhu Singh
ISBN : 978-81-930928-3-5



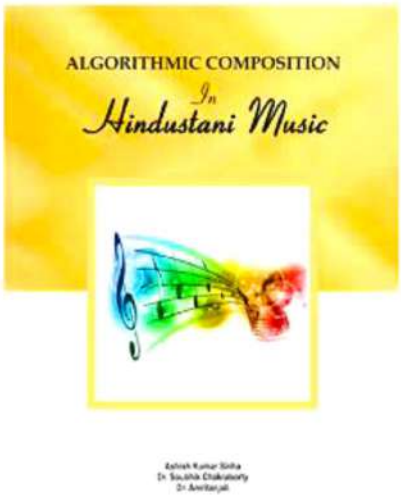
Dr. Ismail Thamarasseri
ISBN : 978-81-930928-2-8



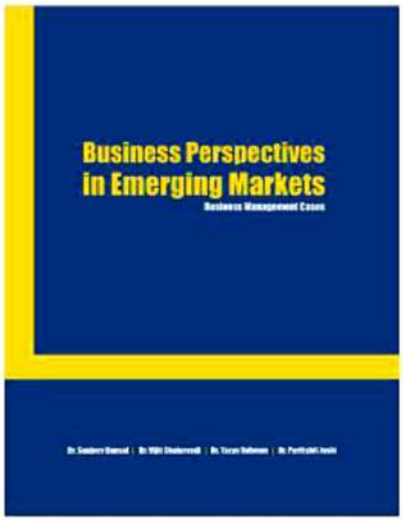
Ram Jaladurgam
Dr. S. Anand Reddy
ISBN : 978-81-930928-5-9



Dr. Sanjeev Bansal, Dr. Vijit Chaturvedi
Dr. Tazyn Rahman, Dr. Parikshit Joshi
ISBN : 978-81-930928-6-6

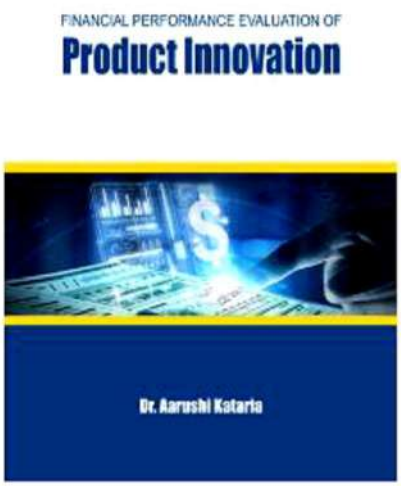


Ashish Kumar Sinha, Dr. Soubhik Chakraborty
Dr. Amritanjali
ISBN : 978-81-930928-8-0

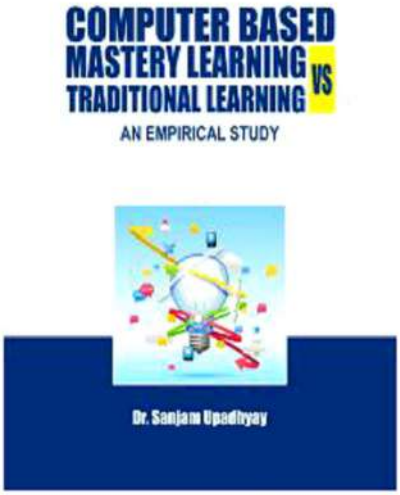


Dr Sanjeev Bansal, Dr. Vijit Chaturvedi
Dr. Tazyn Rahman, Dr. Parikshit Joshi
ISBN : 978-81-936264-0-5

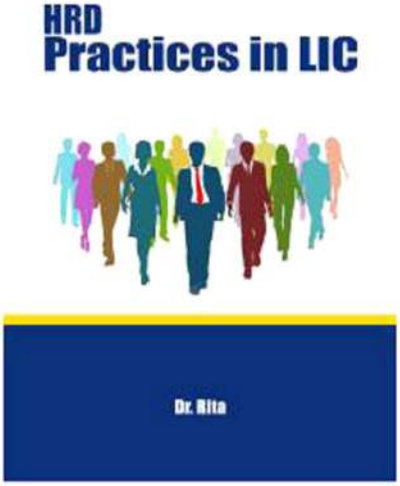
Dr. Jyotsna Golhar
Dr. Sujit Metre
ISBN : 978-81-936264-6-7



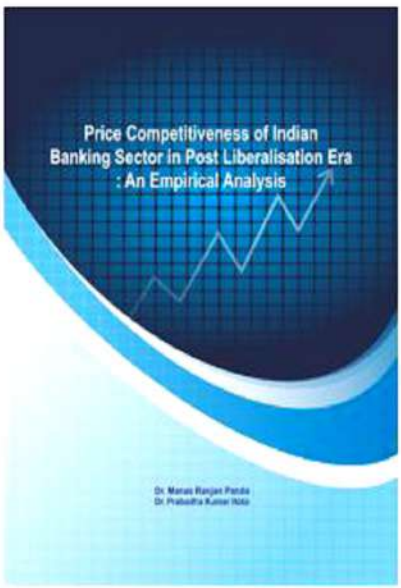
Dr. Aarushi Kataria
ISBN : 978-81-936264-3-6



Dr. Sanjam Upadhyay
ISBN : 978-81-936264-5-0



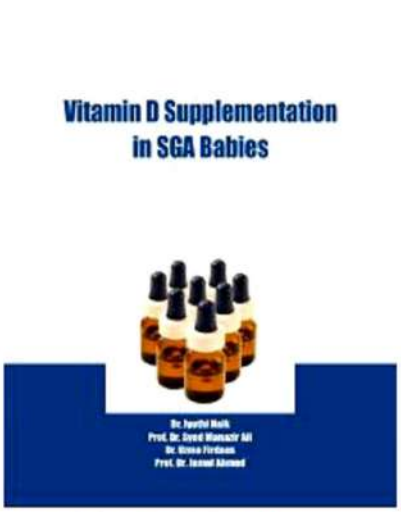
Dr. Rita
ISBN : 978-81-930928-7-3



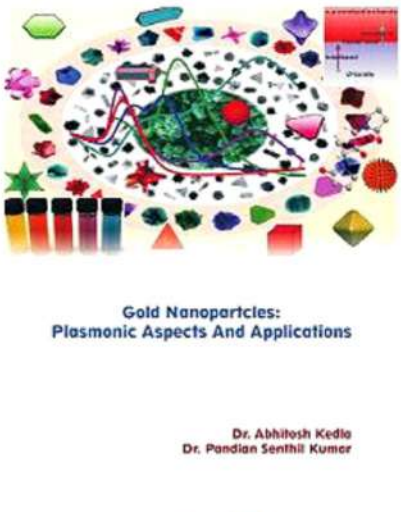
Dr. Manas Ranjan Panda, Dr. Prabodha Kr. Hota
ISBN : 978-81-930928-4-2

Poornima University
ISBN : 978-8193-6264-74

Institute of Public Enterprise
ISBN : 978-8193-6264-4-3



Dr. Jyothi Naik, Prof. Dr. Syed Manazir Ali
Dr. Uzma Firdaus, Prof. Dr. Jamal Ahmed
ISBN : 978-81-936264-9-8



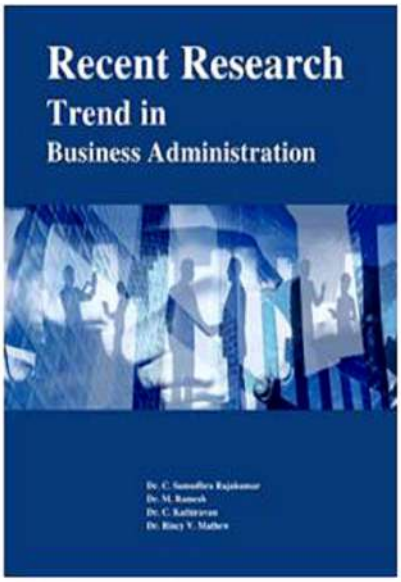
Dr. Abhitosh Kedia
Dr. Pandian Senthil Kumar
ISBN : 978-81-939070-0-9

Dr. Vinod
S. Chandwani
ISBN : 978-81-939070-2-3

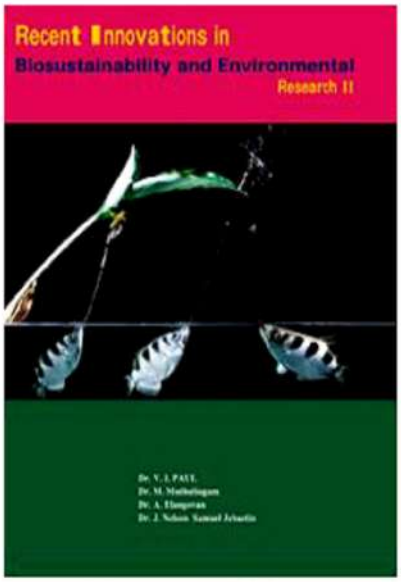
Prof. Dr. Dhananjay
Awasarikar
ISBN : 978-81-939070-1-6

Dr. C. Samudhra Rajakumar, Dr. M. Ramesh
Dr. C. Kathiravan, Dr. Rincy V. Mathew
ISBN : 978-81-939070-4-7

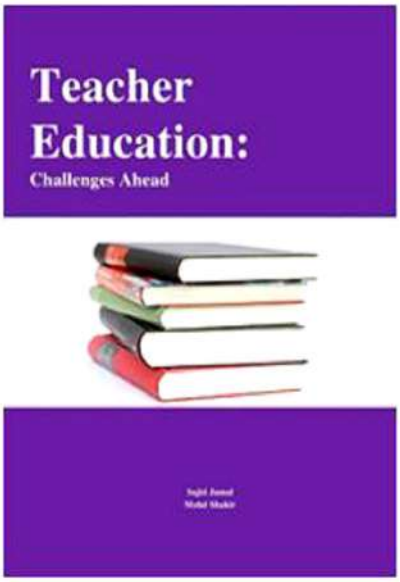
Dr. C. Samudhra Rajakumar, Dr. M. Ramesh
Dr. C. Kathiravan, Dr. Rincy V. Mathew
ISBN : 978-81-939070-6-1



Dr. C. Samudhra Rajakumar, Dr. M. Ramesh
Dr. C. Kathiravan, Dr. Rincy V. Mathew
ISBN : 978-81-939070-7-8

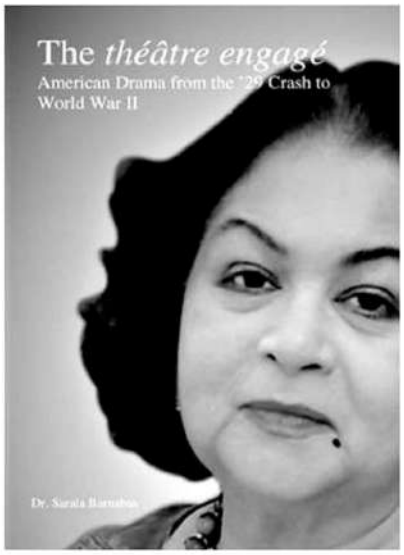


Dr. V. I. Paul, Dr. M. Muthulingam
Dr. A. Elangovan, Dr. J. Nelson Samuel Jebastin
ISBN : 978-81-939070-9-2



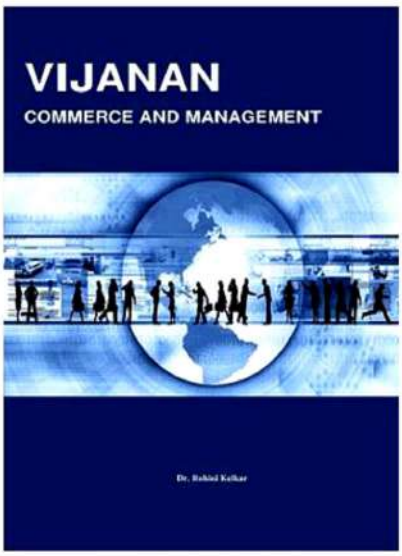
Sajid Jamal
Mohd Shakir
ISBN : 978-81-939070-8-5

Dr. R. Emmaniel
ISBN : 978-81-939070-3-0



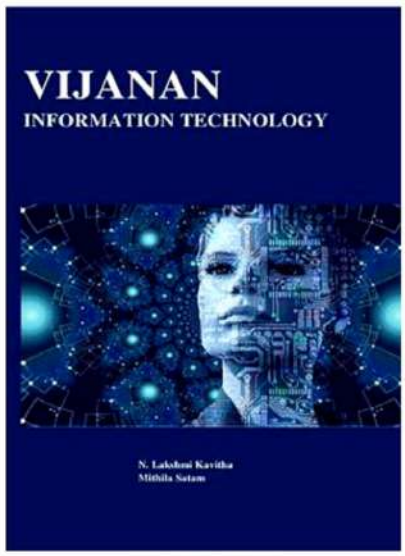
Dr. Sarala Barnabas
ISBN : 978-81-941253-3-4

Dr. M. Banumathi
Dr. C. Samudhra Rajakumar
ISBN : 978-81-939070-5-4

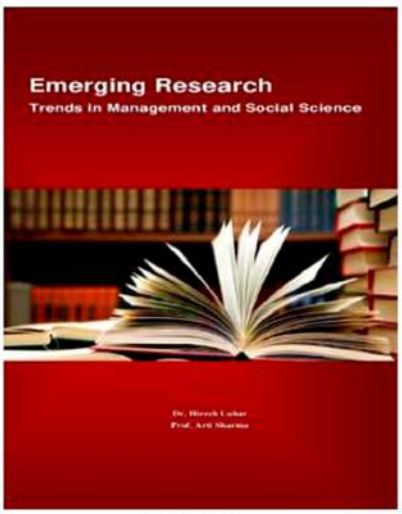


Dr. (Mrs.) Rohini Kelkar
ISBN : 978-81-941253-0-3

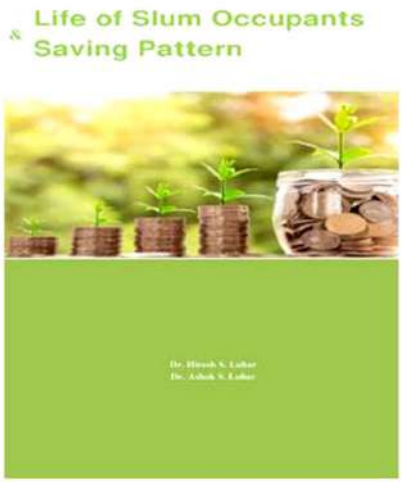
Dr. Tazyn Rahman
ISBN : 978-81-941253-2-7



Dr. N. Lakshmi Kavitha
Mithila Satam
ISBN : 978-81-941253-1-0

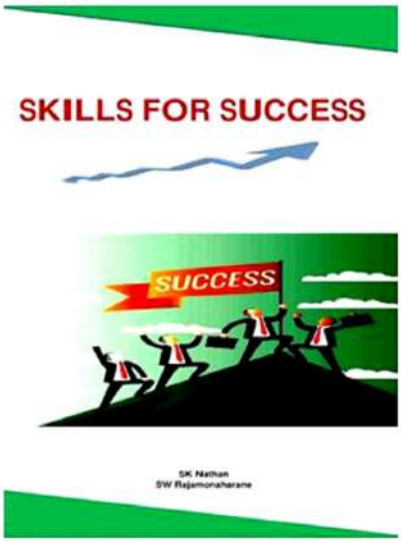


Dr. Hiresh Luhar
Prof. Arti Sharma
ISBN : 978-81-941253-4-1



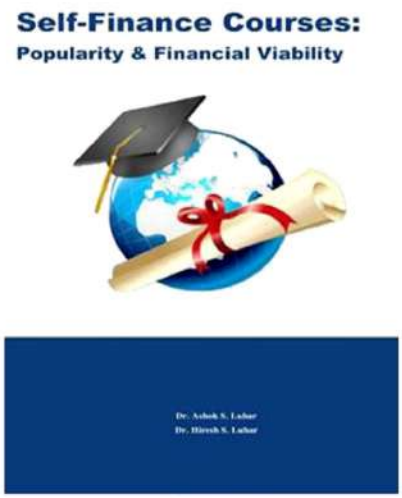
Dr. Hiresh S. Luhar
Dr. Ashok S. Luhar
ISBN : 978-81-941253-5-8

Dr. Babita Kanojia
Dr. Arvind S. Luhar
ISBN : 978-81-941253-7-2

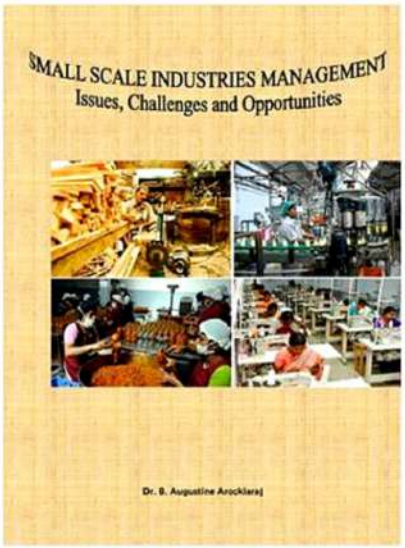


Dr. Sw Rajamonaharane
SK Nathan
ISBN : 978-81-942475-0-0

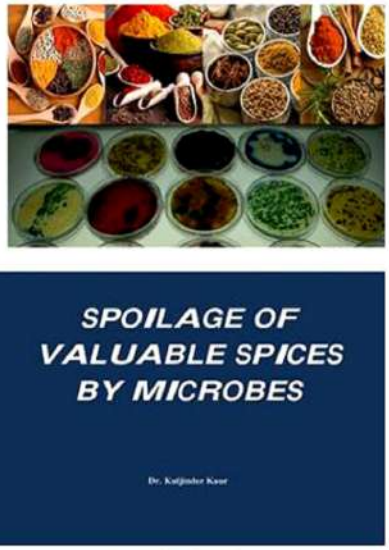
Aditi Sharma
ISBN : 978-81-941253-8-9



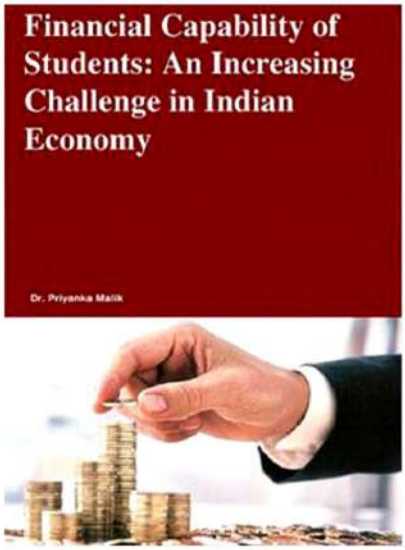
Dr. Ashok S. Luhar
Dr. Hiresh S. Luhar
ISBN : 978-81-941253-6-5



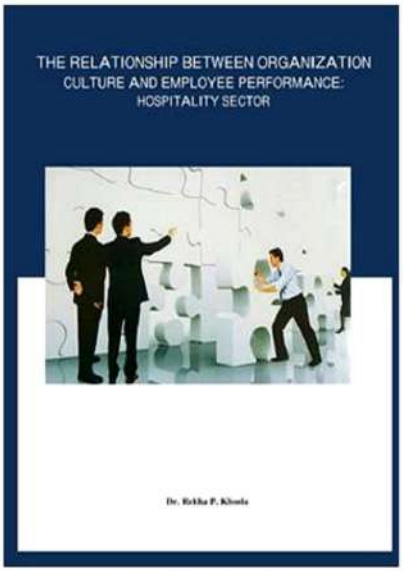
Dr. B. Augustine Arockiaraj
ISBN : 978-81-941253-9-6



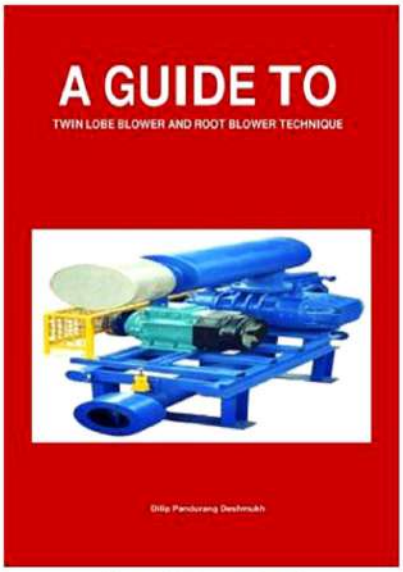
Dr. Kuljinder Kaur
ISBN : 978-81-942475-4-8



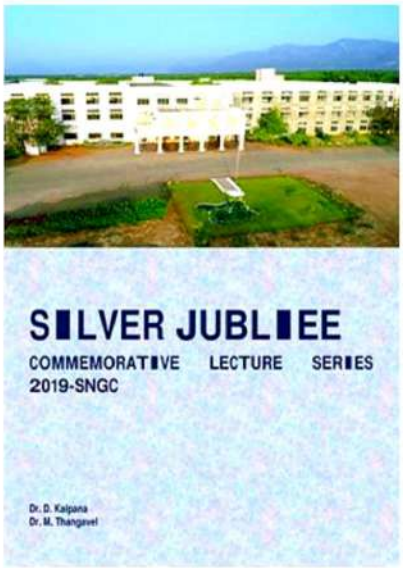
Dr. Priyanka Malik
ISBN : 978-81-942475-1-7



Dr. Rekha P. Khosla
ISBN : 978-81-942475-2-4

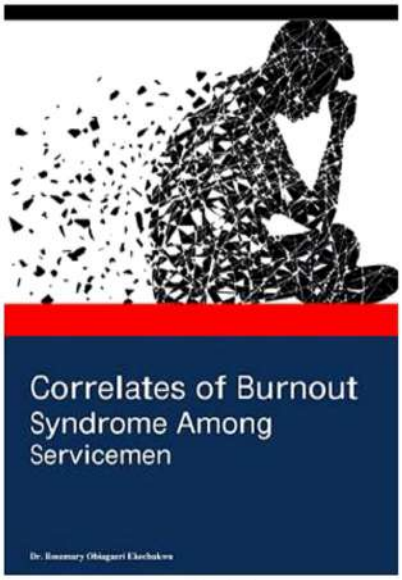


Dilip Pandurang Deshmukh
ISBN : 978-81-942475-3-1

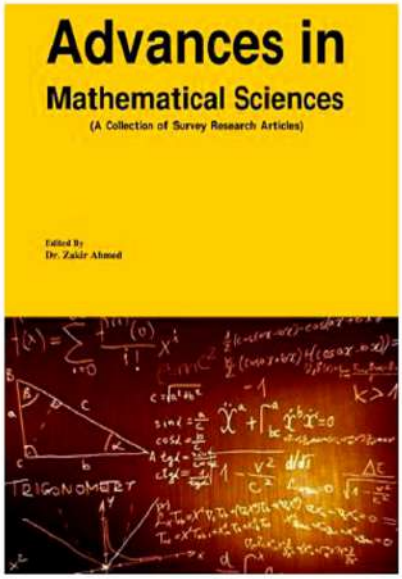


Dr. D. Kalpana, Dr. M. Thangavel
ISBN : 978-81-942475-5-5

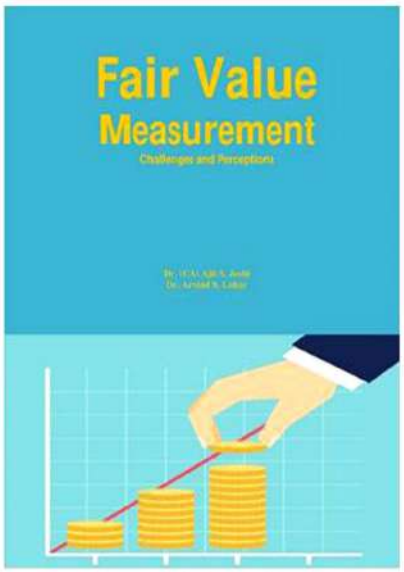
Dr. Aloysius Edward J.
ISBN : 978-81-942475-7-9



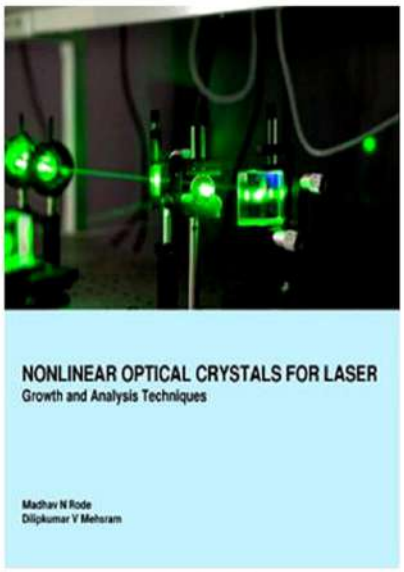
Dr. R. O. Ekechukwu
ISBN : 978-81-942475-8-6



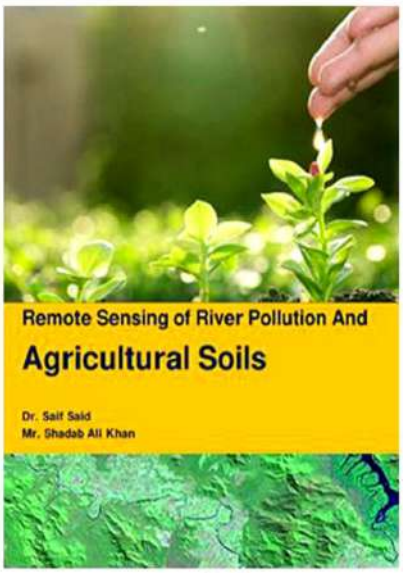
Dr. Zakir Ahmed
ISBN : 978-81-942475-9-3



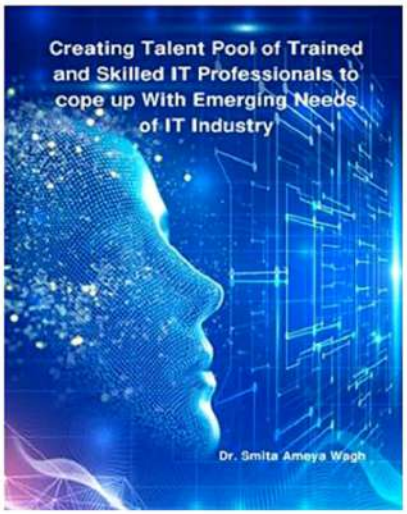
Dr. (CA) Ajit S. Joshi
Dr. Arvind S. Luhar
ISBN : 978-81-942475-6-2



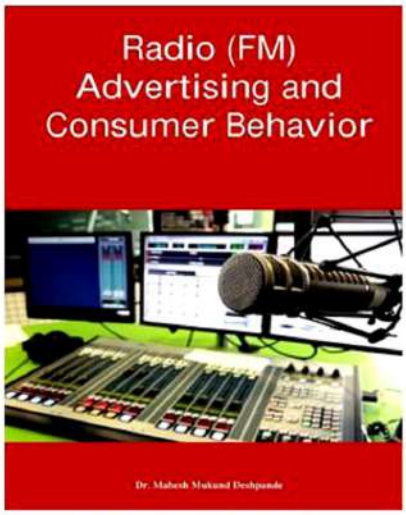
Madhav N Rode
Dilip Kumar V Mehsram
ISBN : 978-81-943209-6-8



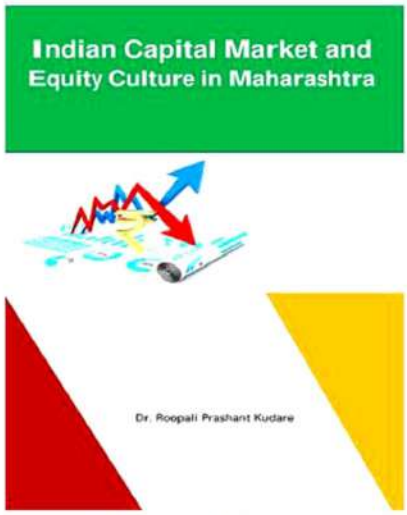
Dr. Saif Said
Shadab Ali Khan
ISBN : 978-81-943209-1-3



Dr. Smita Ameya Wagh
ISBN : 978-81-943209-9-9

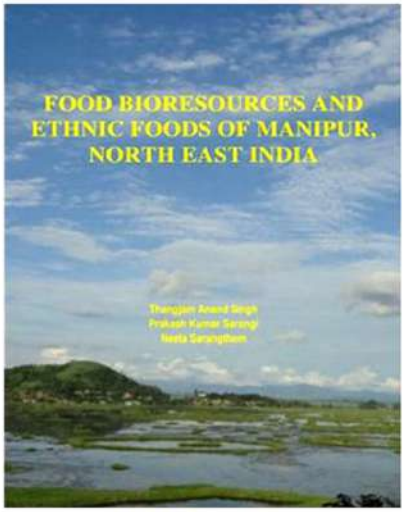


Dr. Mahesh Mukund Deshpande
ISBN : 978-81-943209-7-5

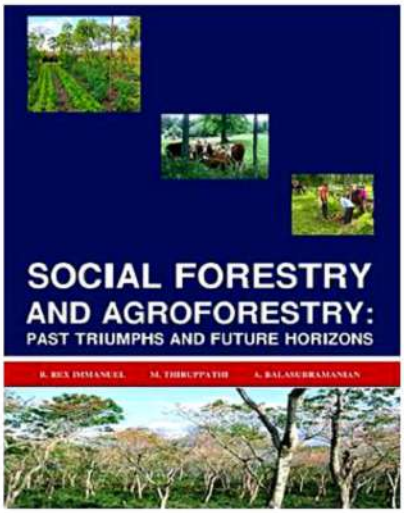


Dr. Roopali Prashant Kudare
ISBN : 978-81-943209-3-7

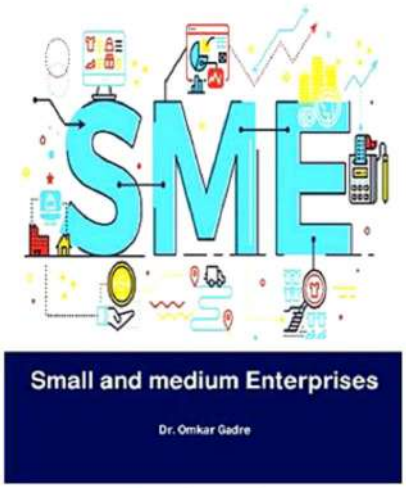
M. Thiruppathi
R. Rex Immanuel
K. Arivukkarasu
ISBN : 978-81-930928-9-7



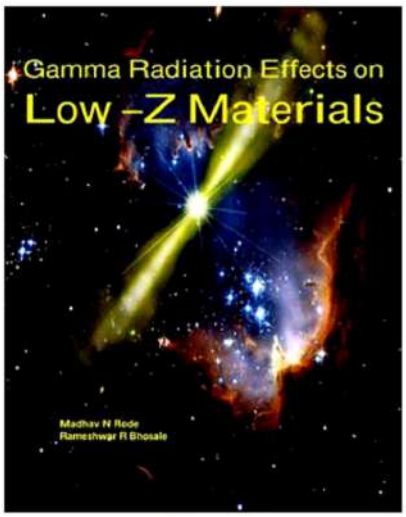
Dr. Th. Anand Singh
Dr. Prakash K. Sarangi
Dr. Neeta Sarangthem
ISBN : 978-81-944069-0-7



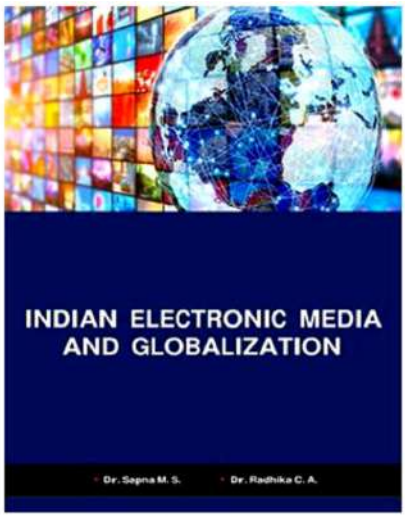
R. Rex Immanuel
M. Thiruppathi
A. Balasubramanian
ISBN : 978-81-943209-4-4



Dr. Omkar V. Gadre
ISBN : 978-81-943209-8-2

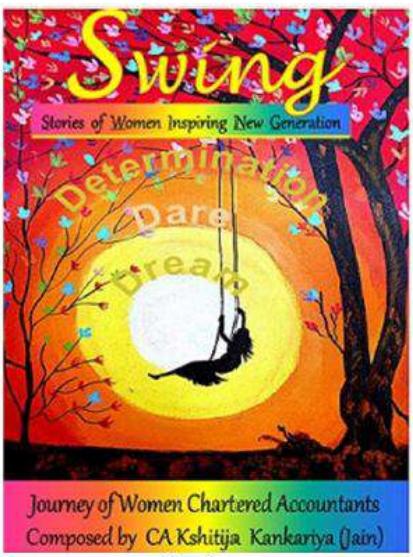


Madhav N Rode
Rameshwar R. Bhosale
ISBN : 978-81-943209-5-1

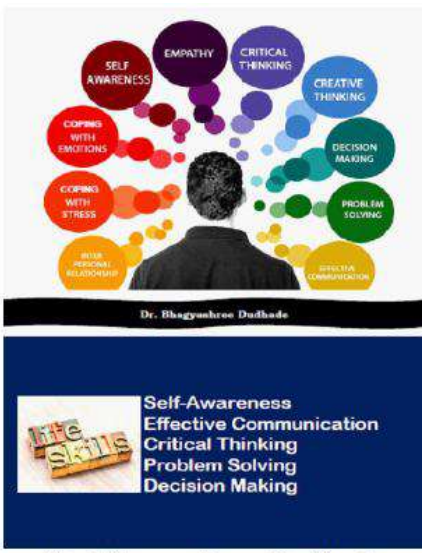


Dr. Sapna M S
Dr. Radhika C A
ISBN : 978-81-943209-0-6

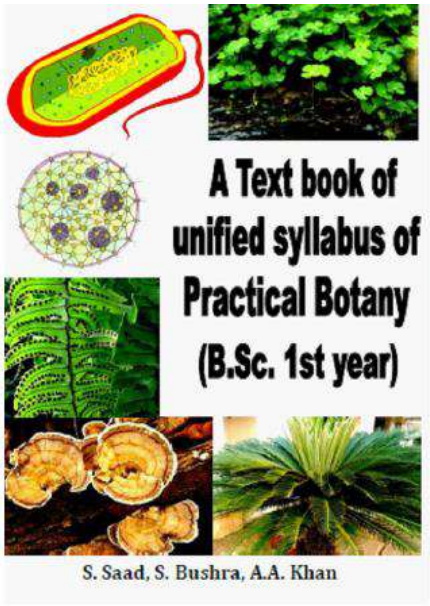
Hindusthan College
ISBN : 978-81-944813-8-6



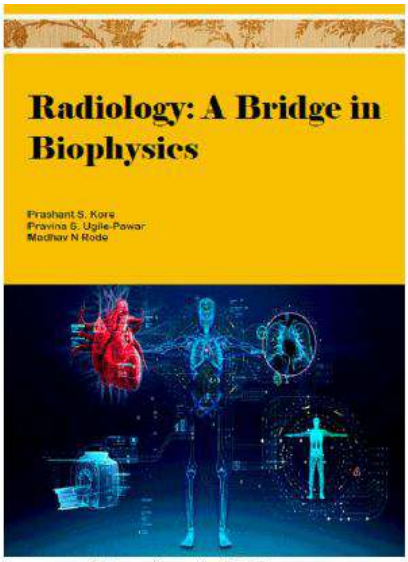
Swing
ISSN: 978-81-944813-9-3



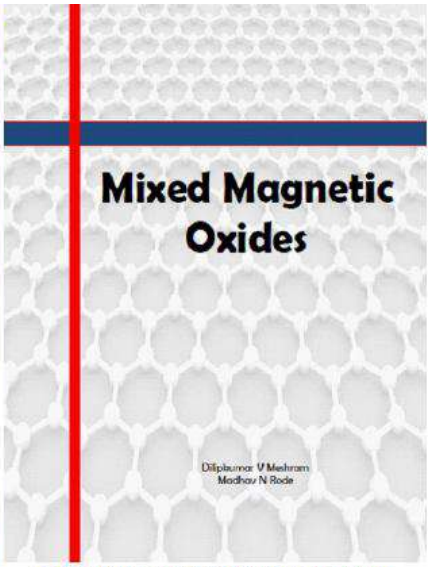
Dr. Bhagyashree Dudhade
ISBN : 978-81-944069-5-2



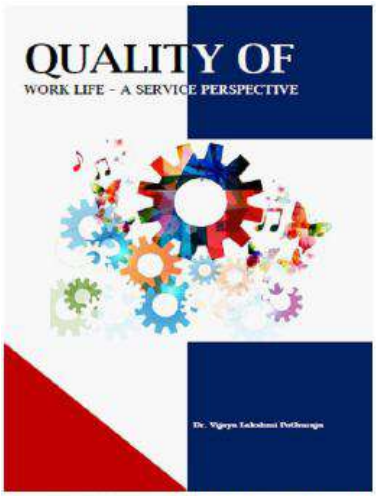
S. Saad, S. Bushra, A. A. Khan
ISBN: 978-81-944069-9-0



Prashant S. Kore
Pravina S. Ugile-Pawar
Madhav N Rode
ISSN: 978-81-944069-7-6



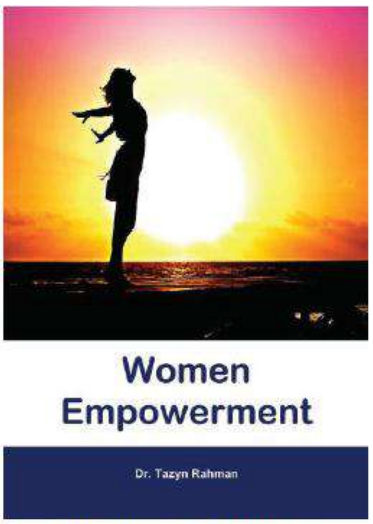
Dilipkumar V Meshram and
Madhav N Rode
ISSN: 978-81-944069-6-9



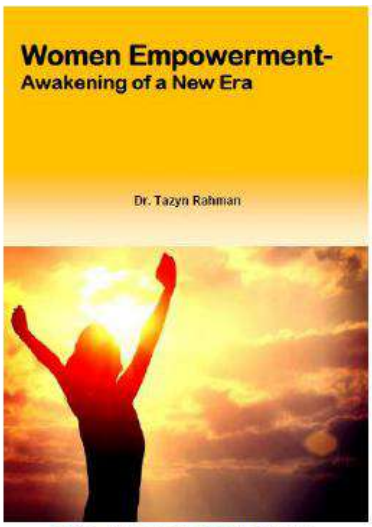
Dr. Vijaya Lakshmi Pothuraju
ISBN : 978-81-943209-2-0

Pratibha College
ISBN : 978-81-944813-2-4

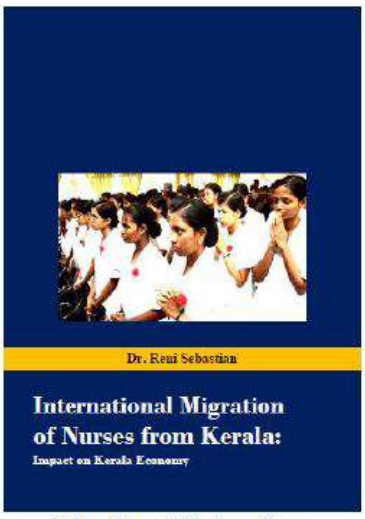
Pratibha College
ISBN : 978-81-944813-3-1



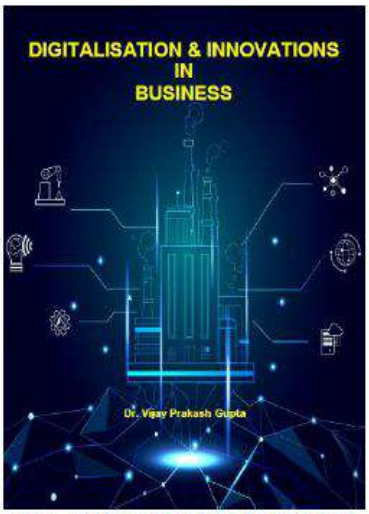
Dr. Tazyn Rahman
ISBN : 978-81-936264-1-2



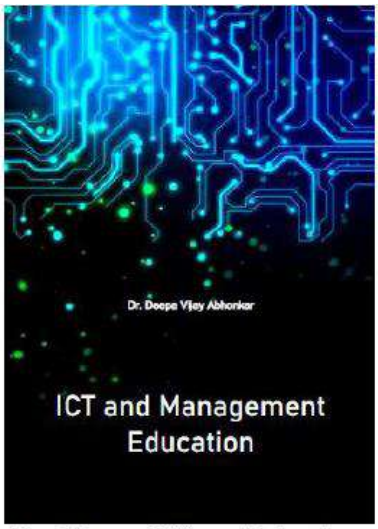
Dr. Tazyn Rahman
ISBN : 978-81-944813-5-5



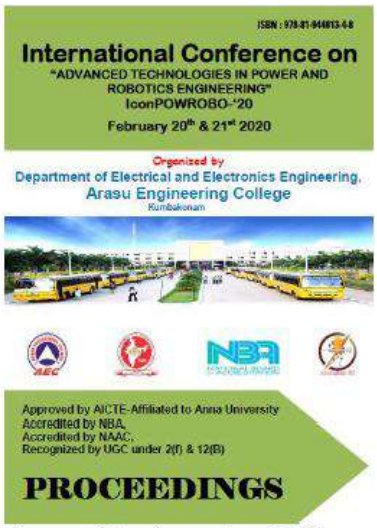
Dr. Reni Sebastian
ISBN : 978-81-944069-2-1



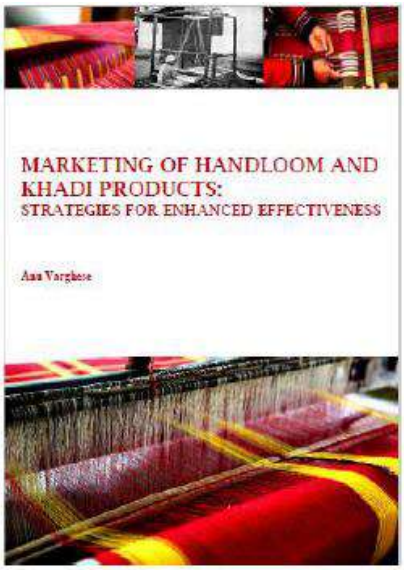
Dr. Vijay Prakash Gupta
ISBN : 978-81-944813-1-7



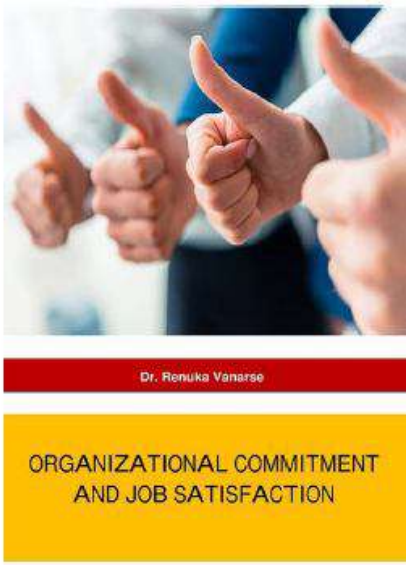
Dr. Deepa Vijay Abhonkar
ISBN : 978-81-944813-6-2



Arasu Engineering College
ISSN: 978-81-944813-4-8



Dr. Anu Varghese
ISBN : 978-81-944069-4-5



Dr. Renuka Vanarse
ISBN : 978-81-944069-1-4

Indian Academicians and Researchers Association
www.iaraedu.com