

ANIME RECOMMENDATION CHATBOT USING HYBRID FILTERING & TRANSFORMERS WITH IMPACT OF GENRE DIVERSITY

¹Vipul Jadhav and ²Sunita Jena

¹Student and ²Assistant 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

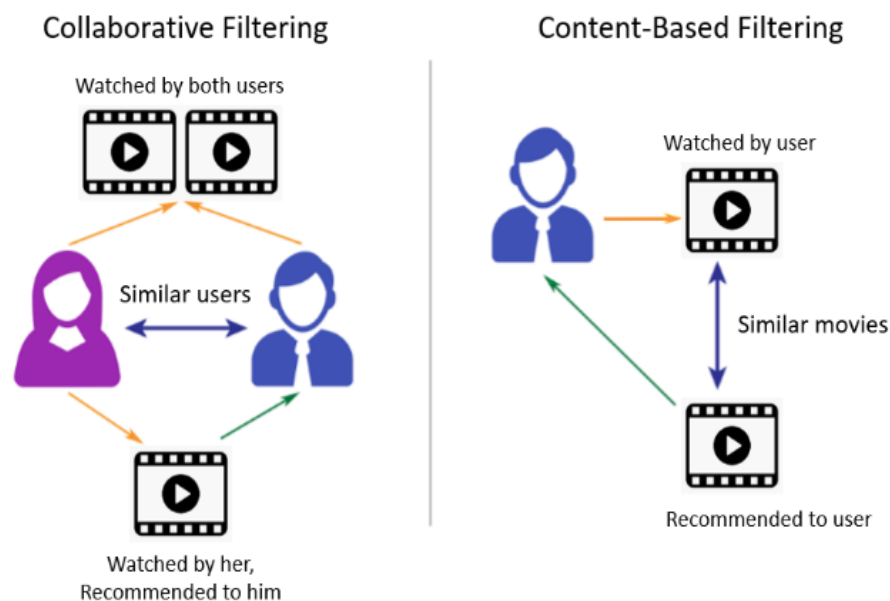


Fig 1. Collaborative filtering vs Content-Based Filtering Approach.

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

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

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.

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.

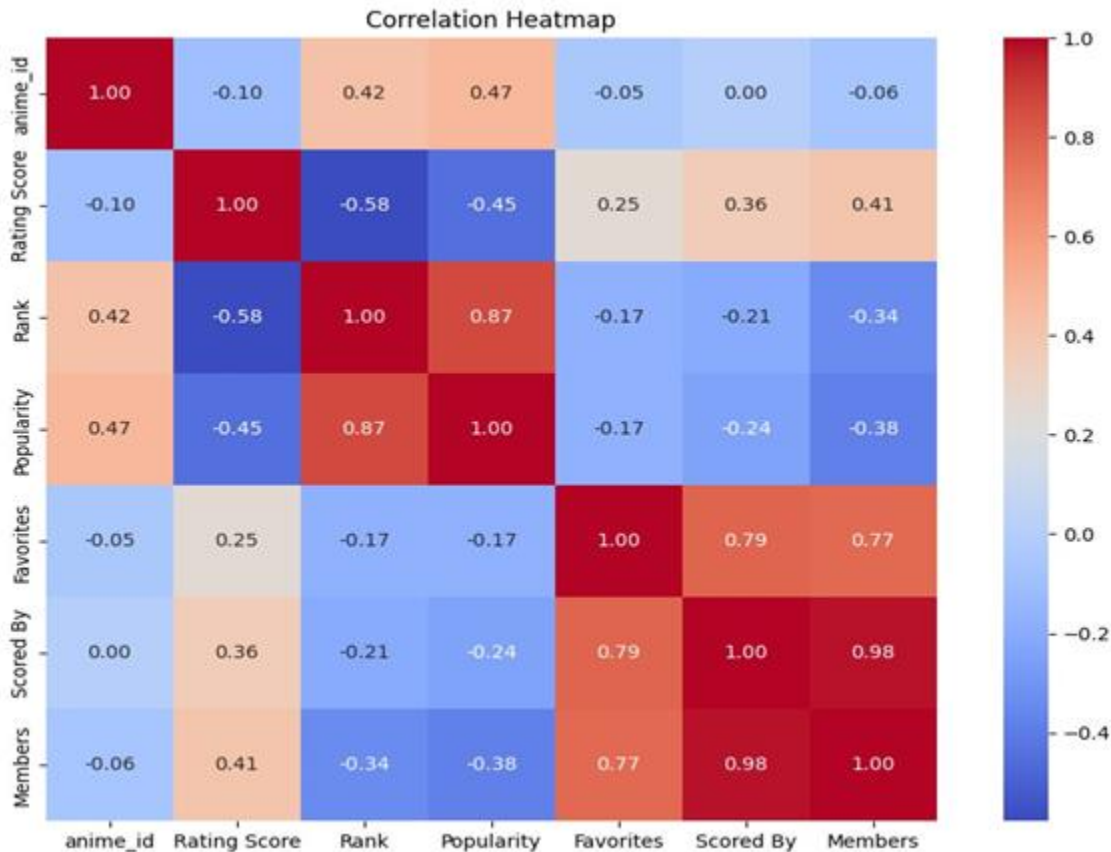


Fig 2: Correlation Heatmap of the Variables.

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.

```

# 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..

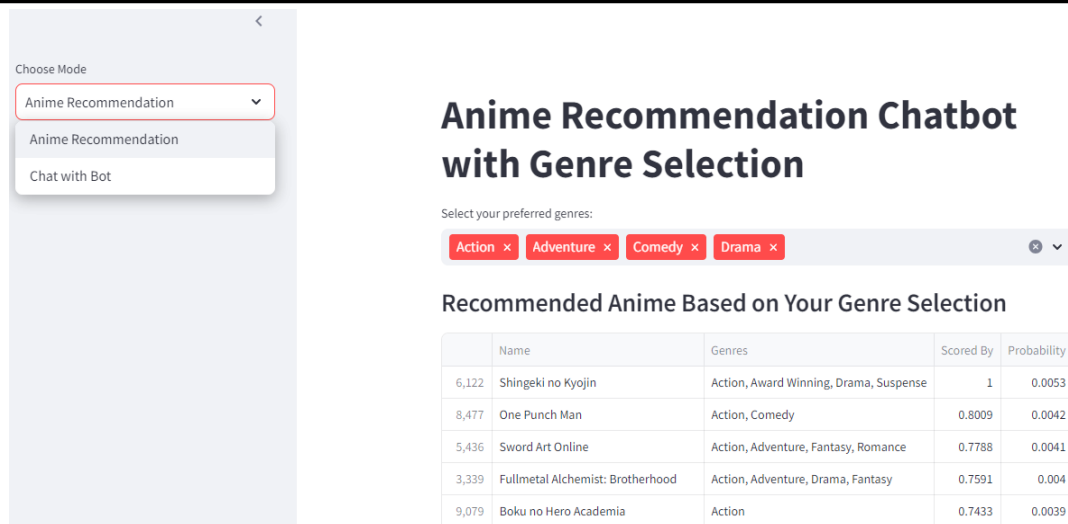


Fig 5: Snapshot of Chatbot using Streamlit & ngrok.

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 rpyr 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

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.

ACKNOWLEDGEMENT

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.

-
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.