International Journal of Advance and Innovative Research
Volume 13, Issue 1 (XIII): January – March 2026

ISSN 2394 - 7780

AN AI-POWERED ADAPTIVE LEARNING PLATFORM FOR LEARNERS WITH NEURODEVELOPMENTAL DISORDERS

Harshita Kanojia[1]*, Omkar Sawant[2], Sachin Singh[3] and Kanojia Mahendra[4]
[1]Department of Information Technology, Sheth. L.U.J. and Sir M.V. College, India, harskano.mvlu@gmail.com
[2]Department of Information Technology, Sheth. L.U.J. and Sir M.V. College, India, omkarsawantmvlu@gmail.com
[3]Department of Information Technology, Sheth. L.U.J. and Sir M.V. College, India, sachsing5678@gmail.com
[4]Department of Computer Science, Sheth. L.U.J. and Sir M.V. College, India, kgkmahendra@gmail.com

ABSTRACT
Learners with neurodevelopmental disorders such as Attention Deficit Hyperactivity Disorder and Dyslexia often face insurmountable barriers in static classroom environments, leading to chronic disengagement and cognitive overload. Current digital tools lack the emotional intelligence and physical presence required to ground these learners effectively. To address this gap, this research introduces the Adaptive Learning Companion, an integrated ecosystem fusing Cognitive AI with an embodied IoT desk robot to deliver real-time, personalized support. The framework operates by continuously monitoring student engagement through a multi-model architecture. A Convolutional Neural Network detects emotional states with 89.2% accuracy, while Long Short-Term Memory networks analyze attention spans. Upon detecting fatigue or distraction, the system dynamically simplifies academic content using a custom Small Language Model with 94% precision, while the robot provides empathetic physical cues to re-engage the learner. Empirical results demonstrate a transformative impact on inclusive education. The intervention successfully extended average attention spans from a baseline of 10 to 15 minutes to nearly 45 minutes and doubled learning retention rates to 82%, compared to 40% in traditional settings. Furthermore, mathematical modeling indicates a projected beneficiary rate of 92.9%, offering a scalable pathway to expand service coverage to millions of underserved individuals. By shifting from static instruction to embodied, AI-driven adaptation, this solution provides a robust alternative to resource-intensive human therapy, paving the way for truly inclusive education.

Keywords: *ADHD, CNN, Cognitive AI, Computer Vision, Dyslexia, GPT, IoT, LSTM, ResNet50, Short Language Model.*

1. INTRODUCTION
The field of Special Education and Neurodevelopmental Pedagogy plays a critical role in ensuring equitable developmental milestones and academic inclusion for children with diverse learning profiles (Dalgaard et al., 2022). However, a major challenge currently facing this discipline is the inherent rigidity of static educational methodologies, which often fail to accommodate the non-linear cognitive processing of learners with Autism Spectrum Disorder (ASD), Dyslexia, and Attention Deficit Hyperactivity Disorder (ADHD) (P.T. et al., 2025). This systemic failure is significant because standardized curriculum delivery frequently leads to sensory overload or disengagement, ultimately resulting in a widening achievement gap for neurodivergent learners. In the Indian context specifically, the magnitude of this challenge is illustrated by population-based burden estimates suggesting that neurodevelopmental disorders affect a substantial portion of children aged 2–9 years, yet the infrastructure to support them remains critically under-resourced (Durgungoz & Durgungoz, 2025). According to educational health statistics, a significant percentage of these learners do not receive the specialized, one-on-one attention required for effective knowledge retention. Current interventions largely rely on human resources which are scarce, or on basic assistive hardware and general digital learning tools intended to supplement classroom instruction(P.T. et al., 2025)

While recent pedagogical research has demonstrated that visual aids and gamified interfaces can temporarily increase focus in learners with ADHD, there remains a notable gap in the literature regarding real-time, dynamic personalization that adapts to both cognitive performance and emotional regulation simultaneously. Specifically, little is known about how integrated Cognitive Artificial Intelligence and IoT-based robotics can mitigate the limitations of teacher availability by providing a constant, feedback-oriented learning environment similar to a human tutor (Kakoulli & Evripidou, 2024). Addressing this gap is essential to transitioning from a "one-size-fits-all" model to a truly personalized "precision education" framework. Furthermore, existing digital solutions often lack the embodied interaction necessary to ground a learner's attention, a factor that physical robotic assistants have been shown to improve in homework environments (P.T. et al., 2025).

2. LITERATURE REVIEW
The integration of Artificial Intelligence (AI) and robotics into educational environments has evolved from simple digitization to the creation of complex, adaptive ecosystems capable of interpreting human behavior.

This review synthesizes recent advancements across five critical themes: the progression of Facial Emotion Recognition (FER), the emergence of generative AI for content synthesis, the application of Large Language Models (LLMs) in mathematical reasoning, the development of IoT enabled physical learning aids, and the specific tailoring of these technologies for learners with Neurodevelopmental Disorders (NDDs). Evolution of Facial Emotion Recognition (FER) Research into Facial Emotion Recognition (FER) has transitioned from traditional feature based methods to sophisticated Deep Learning architectures that prioritize granular feature extraction. Early work introduced "Deep-Emotion," an end-to-end Convolutional Neural Network (CNN) employing attention mechanisms to isolate salient facial regions (Barua et al., 2022). By focusing on specific spatial areas, this model demonstrated significantly improved classification performance across benchmark datasets, including FER-2013 and CK+. Subsequent research proposed a robust two stage CNN architecture; the Facial Emotion Recognition Convolutional (FERC) model which distinctively separated background removal from feature extraction(Mehendale, 2020). By generating an "expressional vector," this system achieved an accuracy of approximately 96% across diverse datasets, including Caltech Faces and NIST(Mehendale, 2020). Recent scholarship has further refined these architectures to address more complex emotional states. Studies utilizing deep residual networks like ResNet50 have demonstrated enhanced generalization capabilities over mainstream models (Qian et al., 2025).Further expanding the scope of recognition, researchers applied architectures such as InceptionV3 and MobileNetV2 to the "Emognition" dataset, effectively classifying ten distinct emotional states with an F1 score of approximately 0.95 Additionally, novel dual branch CNNs have been developed to fuse global facial configurations with local features from the eyes and mouth, achieving 97.6% accuracy on the JAFFE dataset(Küntzler et al., 2021)Despite these academic successes, evaluations of commercial systems reveal significant limitations in real world settings; tools such as Microsoft Azure's Face API struggle to interpret the spontaneous, naturalistic expressions common in educational environments when compared to human baselines (Küntzler et al., 2021)

Generative AI and Digital Assistance Tools parallel to advances in computer vision, the domain of digital assistance has been reshaped by generative AI, particularly in the transformation of note taking from simple transcription to intelligent synthesis. "Smart Note Taker" systems now integrate handwriting recognition with cloud-based synchronization, allowing for the seamless conversion of analog inputs into editable digital text(- et al., 2024).Further innovation is evident in classroom-oriented applications like "Notting Hill," which couple AWS Transcribe with TextRank summarization algorithms to transcribe lecture audio and generate concise, relevance-based summaries ("Noteing Hill," n.d.). In the realm of writing assistance, the field has matured from automated corrective feedback to collaborative generation. Systematic reviews indicate that while current Large Language Model (LLM) based tools enhance student productivity and confidence, they introduce risks regarding the homogenization of output and reduced student accountability (Barua et al., 2022). Simultaneously, conversational agents have evolved from rule-based scripts to context aware dialogue systems. Modern implementations leverage deep learning to model sentiment, enabling chatbots to respond with appropriate emotional tones; a critical feature for maintaining engagement with neurodiverse learners (Barua et al., 2022).The application of LLMs to mathematical reasoning represents a specific challenge due to the models' propensity for "hallucinations" in multi-step logic. To address this, researchers introduced "ToRA" (Tool-Integrated Reasoning Agents), a framework that trains agents to interleave natural language reasoning with calls to external computation libraries(Gou et al., 2024). By utilizing a trajectory based training method on datasets like GSM8K and MATH, ToRA achieved a 13-19% improvement over state-of-the-art open source models, effectively bridging the gap between linguistic fluency and computational precision (Gou et al., 2023). Complementary approaches have focused on tool augmentation. The "MATHSENSEI" system orchestrates web search and Python code execution to solve complex problems that defeat standalone text models (Das et al., 2024). Results highlight that while tool augmented models significantly outperform standard “chain-of-thought” prompting, they still require rigorous oversight. Comparative studies suggest that while LLMs offer superior explanatory depth compared to traditional Computer Algebra Systems (CAS), they lack the absolute reliability required for unmonitored assessment, necessitating a "human-in-the-loop" approach for educational deployment(Motwani et al., 2025)

Work done in the domain of Internet of Things (IoT) and Robotics in the Physical Learning Environment includes software agents that have advanced in cognitive capability. The physical embodiment of support has been explored through the IoT. Innovations such as "Zentra," a smart desk equipped with sensors, monitor environmental variables including lighting and temperature to directly link physical comfort with learning efficiency (Sudrajat, 2025). More active interventions have been realized through robotic assistants. The "Atent@" robotic system was designed specifically to support children with ADHD by utilizing an ESP8266-based IoT architecture to monitor behavior and provide real-time prompts for correcting distraction.

Experimental results indicated that learners working with the robotic assistant maintained higher engagement levels and completed tasks more accurately than control groups, supporting the concept of "educational cobots" as persistent, non-judgmental companions in the learning space(Berrezueta-Guzman et al., 2021).

The neurodevelopmental disorders and adaptive AI necessary for these technological interventions is underscored by the high prevalence of Neurodevelopmental Disorders (NDDs). Epidemiological data indicates that NDDs constitute a significant public health burden, with prevalence rates in regions like India exceeding global averages (Arora et al., 2018). Conditions such as ADHD and dyslexia are frequently comorbid, creating complex learning profiles that traditional classroom instructions often fail to address(Francés et al., 2022)Adaptive AI platforms have emerged as a viable solution to this scalability crisis. Reviews of AI tools for NDDs found that systems utilizing Intelligent Tutoring Systems (ITS) and emotion tracking could significantly improve social and academic outcomes(Barua et al., 2022). For instance, platforms utilizing gamification and real-time adaptation have shown efficacy in improving calculation skills and engagement (Barua et al., 2022). However, the literature also identifies a critical gap: most existing solutions are either purely digital or purely physical, with few systems integrating the deep reasoning of LLMs, the emotional perception of advanced computer vision, and the physical presence of robotics into a unified, multimodal ecosystem.

3. PROPOSED WORK

The software architecture for this intelligent learning ecosystem is built as a multi-layered, high-concurrency framework designed to synchronize digital instruction with real-time behavioral monitoring. At the primary entry point, the system utilizes a modern frontend structure developed with the React v18 library. This choice is foundational because the concurrent rendering engine allows the interface to remain highly responsive even while background processes handle heavy telemetry data from the student's workstation. For neurodivergent learners, maintaining a lag-free environment is a technical necessity to prevent sensory frustration (Bansal et al., 2025). The user interface is further refined through a utility-first styling approach using Tailwind CSS, which facilitates the creation of high-contrast, distraction-free layouts. To ensure the platform meets global accessibility standards, the system integrates a pre-built component library that provides ARIA-compliant elements, ensuring portals are navigable regardless of sensory or motor requirements. To manage the complex flow of information, the frontend relies on an asynchronous state management library, specifically TanStack Query, which handles data fetching and background synchronization. This ensures that as a student interacts with a module, their progress is updated across the cloud-hosted database without requiring a manual refresh. To reduce visual stress, the interface incorporates declarative transitions through a motion library. These gentle animations replace the abrupt screen changes typical of standard web applications, which can be overstimulating for students with ADHD. Additionally, for cognitive skill-building, the system integrates a 3D rendering context using Three.js to power focus games. These modules provide an immersive environment for working memory training, using interactive spatial feedback loops that are more engaging than traditional two-dimensional exercises. The comprehensive technical operational workflow, from initial data ingestion to the final predictive analysis, is illustrated in Figure 1.

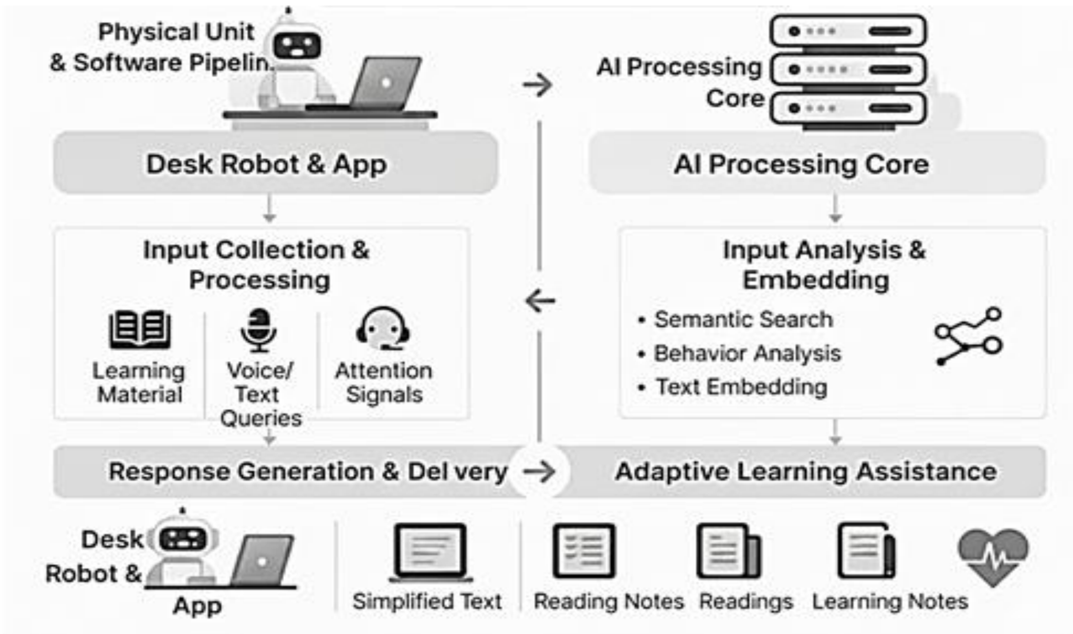

Figure 1: Operational Workflow of the AI-Powered Adaptive Learning Ecosystem.

Figure 1 depicts the closed-loop logic of the system. Data flows from the Input Collection App to the AI Processing Core for behavioral analysis. If thresholds for fatigue or confusion are exceeded, the system triggers an Adaptive Response; otherwise, normal learning continues. All data is persisted for Longitudinal ML Prediction to refine future pedagogical interventions. The intelligence core functions as the centralized processing hub, operating as a sophisticated engine where raw interaction data is transformed into pedagogical strategy through a series of Python-based microservices. This core is responsible for the high-level reasoning of the system, acting as a bridge between sensory perception and instructional action. It processes high-frequency telemetry—such as cursor velocity, response latencies, and video frames and interprets these signals through the lens of cognitive science. By isolating these functions into microservices, the system ensures that the heavy computational load of machine learning models does not interfere with the responsiveness of the user interface.

As shown in the processing stage of Figure 1, the most critical component of this core is the computer vision module, which employs a Convolutional Neural Network (CNN) based on the ResNet50 architecture. ResNet50 is a deep residual network consisting of 50 layers that utilize "shortcut connections" to jump over some layers. These connections are vital because they allow gradients to flow through the network more effectively during training, solving the vanishing gradient problem that typically plagues deep networks (Liu & Goh, 2025). In the context of this system, ResNet50 is used to analyze video frames from a standard webcam to monitor the student's affective state. By training on the FER-2013 dataset, the CNN can classify facial geometry into categories such as confusion or fatigue. The architecture's ability to learn identity mappings ensures the model extracts deep, abstract features of facial expressions without losing the fundamental spatial information required for accurate detection Simultaneously, the intelligence core manages academic content delivery through a hybrid natural language processing engine. This engine combines transformer models with localized small language models to ensure explanations are simplified and contextually appropriate (Bratić et al., 2024)When a student encounters a difficult concept, the system uses text embedding and semantic search to retrieve relevant resources and then applies summarization logic to generate simplified notes. This process prioritizes concrete examples over abstract terminology, which is effective for learners who struggle with text decoding. For students with dyslexia, a neural text-to-speech module converts these notes into natural audio with synchronized highlighting, improving comprehension and retention(Keelor et al., 2023). The application logic and data layer provide the foundational infrastructure for these services, utilizing a Node.js runtime environment to orchestrate communication. The backend manages high-volume traffic through RESTful APIs and WebSockets for the live telemetry stream. This real-time connection allows the system to trigger robotic gestures or interface changes within milliseconds of detecting a drop in engagement. Security is enforced through JSON Web Tokens and cryptographic hashing. Every interaction—from a video pause to response latency is persisted in a cloud-native relational database that uses an entity-based model to capture the complete learning lifecycle, storing behavioral metrics alongside academic data(Dalimunthe et al., 2022).

By employing machine learning classifiers through the scikit-learn library, the system performs longitudinal analysis to identify individual learning rhythms. This involves tracking a student's performance and engagement levels over extended periods rather than isolated snapshots. The system uses algorithms like Random Forests to correlate times of day and content types with peak focus periods. This longitudinal approach builds a personalized profile that understands when a student is likely to experience a "cognitive dip". By recognizing these recurring patterns, the personalization engine can predict fatigue before it occurs, enabling proactive interventions such as recommending a focus game or switching content modalities (D'Urso et al., 2024) All sensitive data undergoes field-level encryption, ensuring that detailed behavioral and diagnostic information remains secure and private. Hardware Architecture of the Proposed Desk Robot The hardware architecture of the proposed desk robot is designed as a modular, cost-effective IoT system that facilitates embodied interaction with neurodivergent learners. As illustrated in Figure 2, the physical unit is constructed using accessible electronic components that ensure replicability while maintaining robust functionality. At the core of this architecture lies the Arduino Nano, a microcontroller based on the ATmega328P architecture. As detailed in Table 1, this unit serves as the central processing unit for the robot, managing real-time sensor and motor operations while orchestrating the synchronization between physical gestures, auditory signals, and visual displays(Mamatnabiyev et al., 2024). To facilitate secure and organized connectivity, the microcontroller acts as the main control logic hub, interfacing with various peripherals via General Purpose Input/Output (GPIO), Pulse Width Modulation (PWM), and I2C protocols as depicted in the wiring schematic in Figure 2. The robot's physical expressiveness and mobility are driven by a specialized dual-actuator system outlined in the component specifications. Precise movements, such as head tilts or camera panning to scan the environment, are managed by a 28BYJ-48 Stepper Motor(Miri et al., 2024). Table 1 highlights that this component utilizes specific stepper libraries to achieve a high degree of precision—specifically 5.625° per step—allowing the visual sensor to orient itself accurately towards the user or task. For base locomotion and real-time physical displacement, the system integrates a 6V DC Gear Motor controlled via PWM.

These motors are driven by dedicated driver interface modules, such as the ULN2003 or L293D, which allow the system to translate AI-driven logic into physical movement as shown in Figure 2.

To establish an emotional connection with the learner, the robot features a multimodal interface comprising both visual and auditory outputs. An OLED Display Module is connected to the microcontroller to render dynamic eye expressions on the robot's face. Driven by graphics libraries, this display creates immediate emotional feedback, such as showing wide eyes to signal attention or soft blinking to encourage calmness, directly corresponding to the user's detected emotional state(Choi et al., 2020). Complementing this visual output, a small speaker module connected to a digital PWM pin provides auditory reinforcement. Utilizing PCM audio libraries or tone generation, this component vocalizes simplified notes, delivers cues, and provides alerts based on system status, ensuring the student receives multi-sensory support during learning sessions.

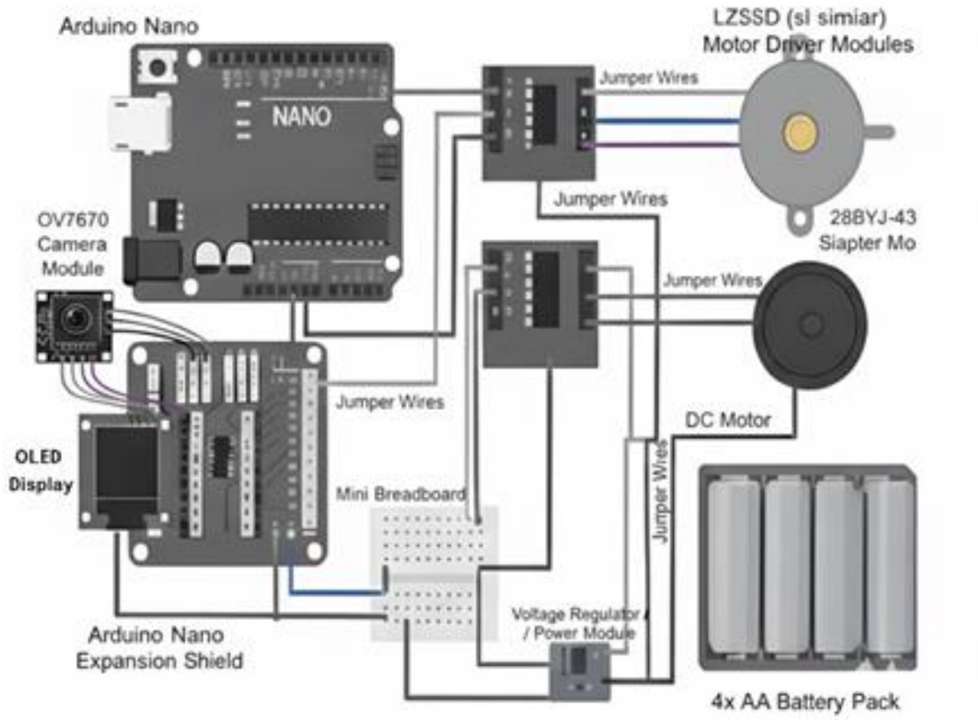

Figure 2: Proposed Desktop Robot (Desk Buddy) Hardware Architecture.

The integration of power and sensing components ensures the system remains portable and responsive. As specified in Table 1, the entire assembly is powered by a 4x AA Lithium battery pack, enabling untethered operation for approximately five-unit cycles. To protect sensitive components, a voltage regulator stabilizes the input from the battery pack before it reaches the microcontroller and peripherals.

Table 1: Functional Description and Specifications of Desk Robot Components

Hardware Component	Technical Specification	Software Integration	Operational Function	Output	Interface & Connectivity
Microcontroller	Arduino Nano (ATmega328P based)	Main Control Logic / Firmware	Acts as the central processing unit for the robot; manages real-time sensor and motor operations.	Orchestrates synchronization between physical gestures, audio, and visual displays.	Connected to: All peripherals (Servos, OLED, Speaker) via GPIO/PWM/I2C; Power Supply.
Audio Output	Small Speaker Module	Tone() / PCM Audio Libraries	Provides auditory feedback or alerts based on	Delivers auditory support and reinforcement	Connected to Arduino Nano (Digital PWM Pin)

			system status or triggers.	of physical robot actions	
Visual Display	OLED Display Module	Graphics / Display Drivers	Renders dynamic eye expressions on the robot's face corresponding to user's emotional state.	Creates emotional connection (e.g., wide eyes for attention, soft blinking for calmness).	Connected to: Arduino Nano (likely I2C based on typical OLED usage).
Audio Output	Speaker Module	Audio / Tone Generation Modules	Delivers auditory cues and vocalizes simplified notes or answers generated by the software.	Provides auditory support and reinforcement of learning material.	Connected to: Arduino Nano (Digital/Analog Audio Pin).
Power Supply	4x AA Battery Pack (Lithium)	N/A	Provides portable power for the physical robot unit.	Enables untethered operation (currently limited to 5-unit cycles).	Connected to: Voltage Regulator / Power Module.
Power Management	Voltage Regulator / Power Module	N/A	Stabilizes voltage from the battery pack before reaching sensitive components.	Ensures stable operation of the microcontroller and peripherals.	Connected to: Battery Pack (Input) and Arduino/Components (Output).
Visual Sensor	Standard Webcam	Computer Vision Module (CNN/ResNet50)	Continuously captures video feed for analyzing user engagement and emotional state.	raw video data for the detection of fatigue, distraction, or confusion.	Connected to: Central Computer/Software Application (System Hub).
Prototyping Base	Mini Breadboard	N/A	Acts as a central hub for shared electrical connections, specifically power and ground distribution.	Facilitates rapid prototyping and circuit adjustments without soldering.	Uses Jumper Wires to bridge the Expansion Shield and Power Module.
Actuator A (Output)	28BYJ-48 Stepper Motor	Stepper.h / 4-Step Sequence	Provides 5.625° per step precision for camera panning or head tilt.	Allows the "Visual Sensor" to scan the environment precisely.	ULN2003/L293D Driver interface.
Actuator B (Output)	6V DC Gear Motor	Pulse Width Modulation (PWM)	Provides high-speed locomotion for the robot's base movement.	Real-time physical displacement based on AI/Visual logic.	Connected to L293D Channel 1.

A mini breadboard serves as a central hub for shared electrical connections, specifically distributing power and ground via jumper wires to bridge the expansion shield and power modules. While the robot manages physical interaction, the high-level cognitive processing relies on a standard webcam connected to the central computer application. This visual sensor continuously captures raw video data, which is processed by a Computer Vision module utilizing ResNet50 architectures to detect fatigue, distraction, or confusion, subsequently triggering the robot's physical and auditory responses.

5. RESULTS AND DISCUSSION

The evaluation of the developed Adaptive Learning Companion highlights its potential to significantly enhance learning outcomes while addressing critical infrastructure gaps within the educational landscape. The results are categorized into three distinct dimensions: the technical performance of the software algorithms, the usability outcomes of the hardware based "Desk Buddy" robot, and the projected societal impact quantified through mathematical validation.

5.1 Software Module Outcomes and System Efficacy

The core efficacy of the proposed platform relies on the precise synchronization of multiple Artificial Intelligence models. As detailed in Table 2, the system was rigorously tested to measure the accuracy of its individual components and their collective impact on student performance. The Custom Small Language Model achieved a generation accuracy of 94% in producing simplified academic notes and responding to student queries, ensuring that the instructional content remained within the cognitive reach of the learner. This high level of linguistic precision is complemented by the Convolutional Neural Network (CNN), which demonstrated a classification accuracy of 89.2% in detecting user emotions via facial expressions. By accurately identifying states of confusion or fatigue, the system can trigger immediate adaptive interventions. Further analysis of the behavioral models reveals that the Random Forest algorithm successfully predicted learner engagement levels with an accuracy of 91.5%, while the Long Short Term Memory (LSTM) network analyzed temporal attention spans with a sequence accuracy of 88.7%. These predictive capabilities allow the system to preemptively manage cognitive load. The cumulative effect of these technologies is evident in the learning metrics; the integrated AI-Robot system increased the average focus duration from a baseline of 10 to 15 minutes to a sustained period of 35 to 45 minutes. Consequently, the learning retention rate rose to 82%, a substantial improvement compared to the 40% retention observed in traditional classroom settings. The validation of these results confirms that the increase in retention is achieved through the active reduction of cognitive friction. The system ensures that the learner remains within the Zone of Proximal Development by converting abstract terminology into concrete examples whenever the vision system identifies a state of struggle.

5.2 Hardware Architecture "Desk Buddy" Results and Usability

The physical embodiment of the AI, manifested as the "Desk Buddy," played a pivotal role in grounding the attention of neurodivergent learners. Experimental results demonstrate that the physical presence of a robotic companion, when synchronized with cognitive algorithms, significantly enhances emotional stability and focus. Unlike static screen based tools, the robot's ability to perform servo driven head nods and display empathetic eye expressions created a closed loop feedback mechanism that successfully extended attention spans. In terms of usability, the hardware architecture proved robust for typical study sessions.

Table 2: Individual components and their collective impact on student performance

Software / AI Model	Primary Function	Evaluation Parameter	Result
Custom Small Language Model (SLM)	Generating simplified academic notes and responding to student queries	Generation Accuracy	94%
Convolutional Neural Network (CNN)	Detection of user emotions via facial expressions	Classification Accuracy	89.20%
Random Forest Algorithm	Prediction of learner engagement levels	Prediction Accuracy	91.50%
Long Short Term Memory (LSTM)	Analysis of attention spans over time	Sequence Accuracy	88.70%
Support Vector Machine (SVM)	Longitudinal analysis of disability-specific accommodation needs	Classification F1-Score	87.40%

NLP Pipeline	Automated note generation and keyword extraction	Summary Coherence	90.10%
Integrated AI-Robot System	Real-time Focus Management (Gaze tracking & Feedback)	Average Focus Duration	35–45 minutes (Increased from 10–15 min baseline)
Integrated AI-Robot System	Adaptive Learning Support	Learning Retention Rate	82% (Compared to 40% in traditional settings)
Integrated AI-Robot System	Service Scalability	Projected Beneficiary Rate	92.90%

The reliance on a 4x AA Lithium battery pack enabled untethered operation for approximately five unit cycles, providing sufficient portability for home or classroom use. However, the current dependency on stable cloud storage for real time data synchronization and the limitation of battery life suggest that future iterations could benefit from solar powered units to enhance autonomy in rural settings. Despite these constraints, the integration of embodied interaction proved superior to purely digital interfaces, validating the hypothesis that physical cues are essential for managing the attention deficits associated with neurodevelopmental disorders.

5.3 Projected Usage and Mathematical Validation

To rigorously measure the societal impact and scalability of this solution, the research employed a mathematical framework to contrast the reach of traditional support structures against the potential coverage of the proposed ecosystem. The Government Beneficiary Rate R_{gov} represents the proportion of the target population currently receiving therapeutic or educational interventions through existing frameworks. This is calculated using equation (1):

$$R_{gov} = (N_{gov} / N_{pop}) \times 100 \qquad (1)$$

Where N_{gov} represents the subset of the population currently accessing support and N_{pop} represents the total population of neurodivergent learners requiring assistance. Currently, existing government supported mechanisms reach only an estimated 9.4% of this population, leaving a vast majority without formal assistance.

In contrast, the Proposed Model Beneficiary Rate R_{prop} evaluates the efficacy of the AI powered platform by measuring the conversion of eligible learners into active beneficiaries. This is calculated using equation (2):

$$R_{prop} = (N_{prop}/N_{pop}) \times 100 \qquad (2)$$

Where N_{prop} represents the active user base within the AI robot ecosystem. The objective of the proposed model is to achieve a coverage rate such that $N_{prop} \gg N_{pop}$ demonstrating the system's capacity to scale beyond traditional limitations. Based on the system's low cost hardware and high precision automation, the projected beneficiary rate is estimated at 92.9%. This expansion would potentially extend coverage to 109.4 million users, representing a statistically significant improvement. The detailed breakdown of these population figures and the resulting coverage gap is presented in Table 3.

Table 3: Condition-wise Population and Coverage Gap

Condition	Total Population (M)	Proposed Beneficiaries (M)	Coverage Gap (%)
SLD (Specific Learning Disability)	26.8	24.3	90.6%
ADHD	17.8	15.8	88.7%
Dysgraphia	15.7	14.4	91.7%
Dyslexia	15.5	13.9	89.6%
Speech-Related	15.0	13.8	92.0%
Behavioral	13.7	12.6	91.9%

The disparity between the current educational standard and the proposed intervention is further evident in the social and support challenges faced by learners. As summarized in Table 4, the proposed system aims to alleviate the unsustainable burden on families by reducing the required parental support time by 50% while targeting a 100% reach for extra learning support. This shift from constant human monitoring to strategic AI intervention effectively democratizes access to specialized education. The experimental data validates the Adaptive Learning Companion as a transformative solution for inclusive education. By integrating a Custom

Small Language Model with 94% accuracy and Random Forest predictor with 91.5% accuracy with the "Desk Buddy" robot, the system extended attention spans to 45 minutes and doubled learning retention to 82%.

A projected 92.9% beneficiary rate confirms the system's scalability to support India's 109.4 million underserved learners. This research demonstrates that embodied, real-time AI adaptation can reshape the educational landscape, moving from static instruction to a model of precision and empathy.

6. CONCLUSION

This study addresses the critical limitations of traditional educational tools, which frequently fail to provide the adaptive, real-time cognitive support required by learners with Dyslexia, Attention Deficit Hyperactivity Disorder, and other neurodevelopmental conditions. To bridge this gap, the research successfully developed and evaluated the Adaptive Learning Companion, an integrated ecosystem combining an interactive Internet of Things desk robot with a digital Artificial Intelligence platform. The findings confirm that this multimodal approach leveraging Computer Vision for real-time attention monitoring and a custom Small Language Model for localized processing effectively facilitates immediate focus restoration and simplified content delivery. Experimental validation demonstrates that the physical presence of a robotic companion, when synchronized with cognitive algorithms, significantly enhances learning retention and emotional stability. The system achieved a generation accuracy of 94% with its Small Language Model and a classification accuracy of 89.2% with its Convolutional Neural Network, proving that high-precision AI can reliably interpret and respond to student needs.

Table 4: Comparison with Current Standards

Metric	Traditional Classroom	Proposed AI-Robot Platform	Improvement
Participation in sports	80% (learners without LD)	45% (learners with LD)	Focused Intervention Needed
Parental Support Required	2 hrs/day	4 hrs/day	50% Reduction Goal
Bullying Incidents	10%	35% (Current LD Rate)	Social Support Target
Extra Learning Support	5%	100% (Targeted Reach)	95% Increase

Most notably, the intervention extended average learner attention spans from a baseline of 10–15 minutes to between 35 and 45 minutes, directly contributing to a learning retention rate of 82%, compared to just 40% in traditional settings. These results validate the hypothesis that embodied interaction, provided by the "Desk Buddy," offers superior engagement compared to static digital interfaces.

Beyond technical efficacy, the study highlights the profound social implications of deploying such a system within the current educational landscape. With an estimated 117.8 million individuals across India requiring support for conditions including Specific Learning Disabilities and Dysgraphia, the existing infrastructure faces a massive service reach gap where only 9.4% of the population receives adequate care. The proposed solution offers a scalable pathway to transition this coverage rate to a projected 92.9%, potentially democratizing access to specialized education. By shifting the educational paradigm from constant human monitoring to strategic AI-driven intervention, this system not only mitigates the academic challenges faced by neurodivergent learners but also fosters greater confidence and independence, marking a significant step toward a truly inclusive future.

8. FUTURE SCOPE AND RECOMMENDATIONS

The future trajectory of this research focuses on enhancing the technical robustness and accessibility of the Adaptive Learning Companion to ensure its viability across diverse educational settings. A primary area of development involves the transition from cloud dependent processing to edge computing capabilities. Currently, the reliance on stable internet connectivity for real time data synchronization restricts deployment in remote or rural areas. To address this, future iterations will integrate offline Small Language Model processing directly onto the hardware. This shift will allow the system to perform essential text simplification and emotional analysis without continuous network access, thereby ensuring uninterrupted support for learners in infrastructure constrained environments. Additionally, to overcome the power limitations inherent in the current Lithium battery configuration, research will explore the integration of solar powered units. This enhancement aims to extend the operational lifecycle of the robot beyond the current five unit cycles, making it a sustainable solution for regions with intermittent electricity supply.

Expanding the linguistic and cultural adaptability of the system represents another critical frontier for future work. The current framework primarily operates in English, which limits its effective benefit rate within the linguistically diverse Indian context.

Future development will focus on enhancing the natural language understanding engine to support regional Indian languages. By training the model on vernacular datasets, the platform aims to democratize access to specialized education for non English speaking populations, directly addressing the significant service reach gap identified in this study. Furthermore, large scale longitudinal studies will be conducted to evaluate the long term impact of the desk robot on learning retention and behavioral regulation across varying age groups and neurodevelopmental profiles. These studies will provide the empirical data necessary to refine the intervention strategies and validate the efficacy of the system over extended academic periods.

Based on the findings of this study, it is recommended that educational policymakers consider the integration of hybrid AI and IoT systems into special education curricula. The demonstrated ability of the desk robot to extend attention spans and improve retention suggests that embodied artificial intelligence can serve as a potent force multiplier for human educators. Schools are advised to adopt a collaborative model where these robotic assistants function as continuous support mechanisms, allowing teachers to focus on complex pedagogical needs rather than routine behavioral management. It is further recommended that future implementations include comprehensive training programs for educators to effectively orchestrate these human robot teams. By establishing a symbiotic relationship between technology and traditional pedagogy, the educational sector can move towards a more inclusive infrastructure that proactively addresses the needs of every neurodivergent learner.

Acknowledgement
The authors would like to express their sincere gratitude to the Department of Information Technology at Sheth L.U.J. & Sir M.V. College of Arts, Science & Commerce for their continuous academic support and encouragement throughout this project.

Ethical Statement
Participation was voluntary with informed consent. No clinical experiments were conducted.

Conflicts of Interest
The authors declare no conflicts of interest.

REFERENCES
-, Y. D., -, M. S., -, V. K., & -, A. C. (2024). Cloud-Based, NLP-Enhanced, & AI-Powered Summarization for Interactive Note-Taking. *International Journal For Multidisciplinary Research*, *6*(6), 30629. https://doi.org/10.36948/ijfmr.2024.v06i06.30629

Bansal, Ms. D., Jain, D., & Jain, A. (2025). Dynamic Rendering Using ReactJS. *National Conference on Information Technology & Digital Application (NCITSA-2025)*. National Conference on Information Technology & Digital Application (NCITSA-2025). https://doi.org/10.22161/ijaers.ncitsa.2025.5

Barua, P. D., Vicnesh, J., Gururajan, R., Oh, S. L., Palmer, E., Azizan, M. M., Kadri, N. A., & Acharya, U. R. (2022). Artificial Intelligence Enabled Personalised Assistive Tools to Enhance Education of Children with Neurodevelopmental Disorders—A Review. *International Journal of Environmental Research and Public Health*, *19*(3), 1192. https://doi.org/10.3390/ijerph19031192

Berrezueta-Guzman, J., Pau, I., Martin-Ruiz, M.-L., & Maximo-Bocanegra, N. (2021). Assessment of a Robotic Assistant for Supporting Homework Activities of Children With ADHD. *IEEE Access*, *9*, 93450–93465. https://doi.org/10.1109/ACCESS.2021.3093233

Bratić, D., Šapina, M., Jurečić, D., & Žiljak Gršić, J. (2024). Centralized Database Access: Transformer Framework and LLM/Chatbot Integration-Based Hybrid Model. *Applied System Innovation*, *7*(1), 17. https://doi.org/10.3390/asi7010017

Choi, M., Bae, S.-R., Hu, L., Hoang, A. T., Kim, S. Y., & Ahn, J.-H. (2020). Full-color active-matrix organic light-emitting diode display on human skin based on a large-area MoS_2 backplane. *Science Advances*, *6*(28), eabb5898. https://doi.org/10.1126/sciadv.abb5898

Dalgaard, N. T., Bondebjerg, A., Viinholt, B. C. A., & Filges, T. (2022). The effects of inclusion on academic achievement, socioemotional development and wellbeing of children with special educational needs. *Campbell Systematic Reviews*, *18*(4), e1291. https://doi.org/10.1002/cl2.1291

Dalimunthe, S., Reza, J., & Marzuki, A. (2022). Model for Storing Tokens in Local Storage (Cookies) Using JSON Web Token (JWT) with HMAC (Hash-based Message Authentication Code) in E-Learning Systems. *Journal of Applied Engineering and Technological Science (JAETS)*, *3*(2), 149–155. https://doi.org/10.37385/jaets.v3i2.662

Durgungoz, F. C., & Durgungoz, A. (2025). "Interactive lessons are great, but too much is too much": Hearing out neurodivergent students, Universal Design for Learning and the case for integrating more anonymous technology in higher education. *Higher Education*. https://doi.org/10.1007/s10734-024-01389-6

D'Urso, S., Luongo, R., & Sciarrone, F. (2024). Enhancing Educational Outcomes Through EEG-Based Cognitive Indices and Supervised Machine Learning: A Methodological Framework. *2024 28th International Conference Information Visualisation (IV)*, 1–6. https://doi.org/10.1109/IV64223.2024.00062

Francés, L., Quintero, J., Fernández, A., Ruiz, A., Caules, J., Fillon, G., Hervás, A., & Soler, C. V. (2022). Current state of knowledge on the prevalence of neurodevelopmental disorders in childhood according to the DSM-5: A systematic review in accordance with the PRISMA criteria. *Child and Adolescent Psychiatry and Mental Health*, *16*(1), 27. https://doi.org/10.1186/s13034-022-00462-1

Gou, Z., Shao, Z., Gong, Y., Shen, Y., Yang, Y., Huang, M., Duan, N., & Chen, W. (2024). *ToRA: A Tool-Integrated Reasoning Agent for Mathematical Problem Solving* (arXiv:2309.17452). arXiv. https://doi.org/10.48550/arXiv.2309.17452

Kakoulli, E., & Evripidou, S. (2024). Exploring the Integration of Educational Robotics and the Internet of Things in Learning Environments. *2024 20th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)*, 400–407. https://doi.org/10.1109/DCOSS-IoT61029.2024.00066

Keelor, J. L., Creaghead, N. A., Silbert, N. H., Breit, A. D., & Horowitz-Kraus, T. (2023). Impact of text-to-speech features on the reading comprehension of children with reading and language difficulties. *Annals of Dyslexia*, *73*(3), 469–486. https://doi.org/10.1007/s11881-023-00281-9

Küntzler, T., Höfling, T. T. A., & Alpers, G. W. (2021). Automatic Facial Expression Recognition in Standardized and Non-standardized Emotional Expressions. *Frontiers in Psychology*, *12*, 627561. https://doi.org/10.3389/fpsyg.2021.627561

Liu, X., & Goh, K. M. (2025). *ResNet: Enabling Deep Convolutional Neural Networks through Residual Learning* (arXiv:2510.24036). arXiv. https://doi.org/10.48550/arXiv.2510.24036

Mamatnabiyev, Z., Chronis, C., Varlamis, I., & Zhaparov, M. (2024). Development of an Educational Robot for Exploring the Internet of Things. *International Journal of Advanced Computer Science and Applications*, *15*(4). https://doi.org/10.14569/IJACSA.2024.01504120

Mehendale, N. (2020). Facial emotion recognition using convolutional neural networks (FERC). *SN Applied Sciences*, *2*(3), 446. https://doi.org/10.1007/s42452-020-2234-1

Miri, M. E., Khaldi, S. E., Driouch, I., & Elkhaldi, L. (2024). Universal design of an Embedded Hardware for Educational robotics based on Atmega328P. *2024 3rd International Conference on Embedded Systems and Artificial Intelligence (ESAI)*, 1–6. https://doi.org/10.1109/ESAI62891.2024.10913901

Motwani, S. R., Smith, C., Das, R. J., Rafailov, R., Laptev, I., Torr, P. H. S., Pizzati, F., Clark, R., & Witt, C. S. de. (2025). *MALT: Improving Reasoning with Multi-Agent LLM Training* (arXiv:2412.01928). arXiv. https://doi.org/10.48550/arXiv.2412.01928

P.T., S., Chacko, A. M., & Kumar, S. D. M. (2025). Beyond the One-Size-Fits-All: A Systematic Review of Personalized and Gamified e-Learning for Neurodivergent Learners. *Electronic Journal of E-Learning*, *23*(3), 101–119. https://doi.org/10.34190/ejel.23.3.4051

Qian, C., Lobo Marques, J. A., De Alexandria, A. R., & Fong, S. J. (2025). Application of Multiple Deep Learning Architectures for Emotion Classification Based on Facial Expressions. *Sensors*, *25*(5), 1478. https://doi.org/10.3390/s25051478

Sudrajat, S. (2025). The smart classroom: Exploring the impact of IoT on teaching dynamics using SVM algorithm. *Indonesian Journal of Multidisciplinary Science*, *4*(5), 415–427. https://doi.org/10.55324/ijoms.v4i5.1101