AI-Driven Assistive Technologies for Inclusive Learning in Special Education Environments

Dr. T. Sureshkumar[1], Dhenesh Raj[2], Dr.R.Vijayalakshmi[3], Dr.J.Karthikeyan[4], Dr.R.Vasanthan[5], Ogabek Solayev[6], Urazbaeva Yulduz[7]

[1]Librarian, National College (Autonomous), Affiliated to Bharathidasan University, Tiruchirappalli, Tamil Nadu, India -6200001
sureshkumart@nct.ac.in

[2]Department of Environmental Science, Graphic Era Hill University, Dehradun, Uttarakhand, India
draj@gehu.ac.in

[3](Corresponding Author), College Librarian, A.D.M.College for Women (Autonomous)Affiliated to Bharathidasan University, Nagapattinam - 611 001.
library@adjadmc.ac.in"

[4]Assistant Professor, Head, Career Guidance and Placement Cell, National College(Autonomous), Tiruchirappalli Tamilnadu. India. jkarthikeyan@nct.ac.in"

[5]Associate Professor, Department of English, Nagaland University, Kohima Campus, Meriema Kohima, Nagaland, India
vasanthan@nagalanduniversity.ac.in
library@adjadmc.ac.in"

[6]Department of Pedagogy and psychology, Urgench state university, Urgench Uzbekistan,
ogabek.s@urdu.uz
0009-0009-4203-5467

[7]Department of psychology, Mamun university, Khiva. Uzbekistan,
urazbayeva_yulduz@mamunedu.uz

HOW TO CITE:
T. Sureshkumar, Dhenesh Raj, R. Vijayalakshmi, J. Karthikeyan, R. Vasanthan, Ogabek Solayev, Urazbaeva Yulduz (2026). AI-Driven Assistive Technologies for Inclusive Learning in Special Education Environments. International Journal of Special Education, 41(1s), 20-35.

ABSTRACT:
Inclusive education of learners with disabilities is still limited by a lack of personalization, poor assistive infrastructure, and cognitive and sensory variability. Old methods of teaching are not usually dynamic and thus learners are not as engaged and do not learn effectively. This paper will solve these problems by suggesting an AI-based assistive learning model that can be used in a special education setting. The overall goal is to make intelligent technologies more accessible, personalized and adaptive to promote learning. The suggested research methodology is a combination of machine learning, natural language processing, and computer vision in order to create multimodal assistance systems, such as speech-to-text systems, emotion-aware tutoring agents, or adaptive content recommendation engines. The system capitalizes on real-time data about learners to customize the instructional plans according to cognitive profiles and behavioral reactions. The

experimental assessment was done in a variety of learners groups and between the traditional assistive devices and the suggested AI-based system. Findings show a tremendous growth: learning engagement was enhanced by 42.6, task completion rates were enhanced by 37.8, accuracy to understand the content was improved by 33.4, and response time was reduced by 28.1% as compared to traditional techniques. Also, the level of accessibility to visually and hearing-impaired learners increased by 45.2, which means a strong inclusivity. The application of this study to scalable implementation into smart classrooms, integration with learning platforms, and real-time adaptive interventions to various disabilities are all applications of this study. The framework assists educators to have actionable information and encourages learner autonomy. To sum up, AI-based assistive technologies offer a revolutionary method of inclusive learning to allow personalized, efficient, and accessible learning. The suggested system shows a significant potential of bridging the educational gap and providing the students with special needs with equal educational opportunities.

Keywords: AI in Education, Assistive Technologies, Inclusive Learning, Special Education, Adaptive Learning Systems, Multimodal Interaction.

COPYRIGHT STATEMENT:
Copyright: © 2026 Authors.
Open access publication underthe terms and conditions of the Creative CommonsAttribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Nomenclature Used:
SEN = Special Educational Needs
NDD = Neurodevelopmental Disorders
CV = Computer Vision
NLP = Natural Language Processing
ML = Machine Learning

I. Introduction

A. Background on Inclusive Education and Special Learning Needs

Inclusive education is a core value of making sure that every learner, irrespective of his or her cognitive, sensory and physical disabilities, receives equal access to quality educational experiences. The last two decades have seen the world education policies focusing more on the importance of implementing students with disabilities into the mainstream learning systems. Although there have been legislative advancements, there still exist vast differences between the theoretical implementation of inclusive education and the actual practice of inclusive education, especially concerning the learners with neurodevelopmental disorders, sensory impairments, and learning disabilities. The conventional classroom layouts, teaching strategies, and curricular frameworks were mainly created to suit neuro typical learners and students with special needs are underserved and often disinterested.

Special education is one that accommodates a wide range of students such as those with autism spectrum disorder, dyslexia, attention deficit hyperactivity disorder (ADHD), visual, hearing, and physical impairments. The groups are unique in terms of their cognitive profiles and require different instruction methods. Barua et al. (2022) observe that AI-based customized assistive technologies are slowly gaining acceptance in meeting the diversity of the learning needs of children with neurodevelopmental disorders.

Traditional education systems frequently do not have the infrastructure, the resources, and trained staff to provide truly personalized support on large scale.

B. Challenges in Traditional Assistive Learning Systems

Although traditional assistive technologies are perennial, they have some very serious limitations that limit their transformative capabilities. Fixed tools, like printed Braille materials, pre-programmed augmentative communication devices, and fixed captioning systems, are not able to dynamically respond to changing learner needs. Such systems are not usually interoperable with larger educational ecosystems, and are designed in isolation. In addition, most of the assistive technology is prohibitively costly, and needs significant technical skills to install and maintain, posing further challenges especially in environments with limited resources. In their extensive study on assistive technology in special education, Saabi et al. (2025) found that the various gaps in systemic gaps such as inconsistent implementation, inadequate evidence-based analysis, and standardized frameworks on technology integration were identified.

Moreover, the conventional systems lack proper mechanisms of capturing and addressing real-time learner behavioural and emotional conditions. A student with frustration, anxiety or cognitive fatigue needs instant pedagogical change which can only be offered through the dynamic tools. Failure to close this gap is a contributing factor to high dropout rates, low academic achievement and low confidence in learners with disabilities.

C. Role of Artificial Intelligence in Transforming Special Education

In a paradigm shift, artificial intelligence can be used to overcome the limitations of traditional assistive systems. AI can make educational settings responsive and personal adaptive environments by facilitating real-time data processing and pattern recognition, as well as adaptive decision-making. Machine learning algorithms can be used to understand the interaction patterns of the learners and produce individualized learning paths, whereas natural language processing can help to provide the learners with speech or language challenges with an easier way to communicate. The computer vision technologies allow interacting via gestures and recognizing emotions, which can give teachers a better understanding of the engagement and understanding of the learners. Kooli and Chakraoui (2025) highlight the fact that the beneficial effects of AI-driven assistive technologies in inclusive education can be measured in its accessibility and pedagogical efficacy, yet they also bring up significant ethical and policy issues that need to be discussed on the proactive level.

The introduction of AI in special education is not just a technological update but a philosophical rebranding one that has the learner at the center of educational process. Harkins-Brown et al. (2025) express that AI implementations in special education span between intelligent tutoring systems and automated assessment systems, both of which can help create a more equitable and responsive learning experience. Fitas (2025) also emphasizes the importance of AI in breaking language barriers and assisting learners with varying linguistic and cognitive features.

D. Research Gap and Motivation

Although there is an increasing interest in AI-based learning tools, the applicability of multimodal AI systems in special education is unexploited. Most of the literature to date has either studied individual parts, like speech recognition or emotion detection, but not a combination of both into a unified adaptive system in real-time. A systematic review of AI in inclusive education provided by Li et al. (2025) revealed that there is still a gap in the available empirical evidence of the holistic, multimodal approach. Moreover, the implementation of such systems in the various categories of disabilities in a single architecture has not been sufficiently proven. The rationale behind this research is the necessity to design, deploy, and test a comprehensive AI-based system that will cater to the entire range of the needs of the learners in special education.

E. Contributions and Paper Organization

The paper contributes to the research of AI-assisted inclusive education in a number of ways. To begin

with, it suggests a new multilayer AI-assisted design, which unites machine learning, NLP, and computer vision together into a single adaptive learning platform. Second, it presents an emotionally sensitive tutoring agent with the ability to model the cognitive state in real-time and adapt instruction. Third, it proposes a thorough experimental analysis of several categories of disabilities, which showed that it had a significant positive impact on the engagement, comprehension, accommodation, and task completion. Fourth, it offers a scalable architecture that can be implemented in smart classrooms and be integrated with existing educational systems. Alsaleh (2026) emphasizes that AI can impact individuals with disabilities in various aspects, such as communication, mobility, and cognitive support, which justifies the extensive applicability of the suggested framework. Ahmed et al. (2025) also highlight the transformative possibility of AI in the development of personalized and inclusive education by putting it in the context of a global discussion on educational equity and technological innovation.

The rest of this paper will be structured as follows: Section II will be a review of literature and the research gaps. In Section III the proposed system architecture is presented. The methodology and model development is described in Section IV. Section V details the experimental design and strategy of evaluation. The results are discussed in section VI. Section VII has conclusive remarks and suggestions on the future research.

II. Literature Review and Related Work

A. Overview of Existing Assistive Technologies in Special Education

The assistive technologies landscape in the special education has been changing significantly in the last decade, shifting towards the stagnant and hardware-based tools to dynamic and software-driven ones. Initial assistive technologies were mainly compensatory (screen readers on the visually impaired and hearing aids on the auditory processing disorder), with a smaller number of technologies that offered both compensatory and assistive features to their users. Memisevic and Hadzic (2025) give an in-depth overview of assistive technology in inclusive education and categorize tools in terms of sensory, motor, cognitive, and communicative. In their analysis, they find that they have seen an increase in adoption rates but implementation fidelity within institutions is not uniform.

Navas-Bonilla et al. (2025) performed a systematic review of the available information on the types, tools, and nature of technology-enabled inclusive education, finding more than forty types of digital interventions. Their results suggest that one of the most popular tools are tablet-based applications, text-to-speech software, and augmented communication devices. Nevertheless, the review also notes that there are considerable differences in the quality of evidence as many tools are not rigorously empirically validated. Chavan et al., (2026) also explored the use of AI and assistive technologies in healthcare, and they indicated that the experience in a clinical setting could be generalized to an educational setting to enhance monitoring and adaptive response capacities.

B. AI Applications in Adaptive and Personalized Learning

The capabilities of artificial intelligence have proved to be rather promising in facilitating adaptive and customized learning experiences. Machine-learning-driven intelligent tutoring systems can simulate the competency of individual learners, forecast the performance curve, and produce instructional resources. A scoping review of the AI integration in inclusive education by Pagliara et al. (2024) revealed that adaptive learning platforms can considerably enhance the engagement of learners and their academic performance when supported by evidence-based pedagogical principles. In the same manner, Paglialunga and Melogno (2025) surveyed AI-based interventions in students with learning disabilities and found that NLP-based reading helper tools and automated feedback systems are beneficial to significant improvements in literacy and comprehension.

The use of emotion recognition and affective computing is one of the most promising areas. Face analysis and physiological responses could identify frustration, confusion, or engagement, allowing

instructors to preemptively intervene with systems able to detect these factors. A survey of the views of the stakeholders conducted by Tsvetkova et al. (2025) found that emotion-aware tutoring is one of the trends of AI-based assistive technologies according to the Delphi survey, and experts believe that it will be widely used in the coming five years. Alnajdi et al. (2025) studied acceptance and use of AI-based assistive technology among students with visual disability in universities and concluded that perceived usefulness and ease of use are the main determinants of the acceptance, where physical self-esteem is a significant moderator.

C. Limitations of Current Approaches

The existing AI-supporting educational software, in spite of its potential, has a number of structural shortcomings. The majority of the systems are specific to a single type of disability, and thus cannot be used in heterogeneous classrooms. H N et al. (2026) surveyed assistive educational technologies with students with neurodevelopmental disorders and discovered that most of the current systems do not consider comorbidities or the evolving needs of learners over time. Also, the privacy of data, discrimination of algorithms, and the lack of interpretability of AI decision-making will be critical ethical issues that are not sufficiently covered in the literature.

Numerous intriguing prototypes created in the research environment do not make it to the classroom because they require relatively expensive hardware, are not economical, or require specialized training. According to Fitas (2025), ethical guidelines that guide the use of AI in education should consider the cultural diversity, language differences, and power imbalance between the creators of technology and end-users. Unless AI systems are consciously designed to focus on these aspects, they may contribute to but not alleviate educational inequities.

D. Comparative Analysis of Existing Frameworks

A comparative study of the current AI-based models of inclusive education shows that there are a variety of methodological practices and different levels of system integration. The vast majority of frameworks focus on either personalization or accessibility, and seldom do they fulfill both at the same time as a single architecture. The presented framework of the given study is unique as it incorporates the multimodal inputs, such as text, speech, gesture, and biometric signals, into a coherent adaptive system. Ahmed et al. (2025) highlight that the development of inclusive education with the help of AI should be based on the interdisciplinary collaboration and systematic analysis, which will serve as the design principles of the developed framework. Kooli and Chakraoui (2025) also suggest policy-level solutions to ethical and sustainable implementation of AI in education, which will offer a governance structure to the technical input of this work.

Table I. Comparative Summary of Related Work in AI-Driven Inclusive Education

Reference	AI Technology Used	Target Disability	System Integration	Evaluation Metrics	Scalability	Key Limitation
Barua et al. (2022)	ML, NLP, CV	Neurodevelopmental disorders	Low	Engagement, accuracy	Limited	No real-time adaptation
Pagliara et al. (2024)	ML, Adaptive algorithms	Mixed disabilities	Moderate	Learning outcomes	Moderate	No emotion awareness
Alnajdi et al. (2025)	AI-assistive tools	Visual impairment	Low	Acceptance, usability	Limited	Single disability focus
Paglialunga & Melogno (2025)	NLP, AI feedback	Learning disabilities	Moderate	Literacy, comprehension	Moderate	No multimodal input

Kooli & Chakraoui (2025)	Policy-AI framework	All disabilities	Conceptual	Policy impact	High (proposed)	No empirical evaluation
Fitas (2025)	AI-NLP, Ethics AI	Language barriers, SEN	Low	Accessibility, equity	Moderate	Ethical framework only
Navas-Bonilla et al. (2025)	Digital tools, AI-apps	Mixed	Low	Tool classification	Limited	Review only, no system
Tsvetkova et al. (2025)	Emotion AI, Delphi	General SEN	Conceptual	Expert consensus	High (future)	No current deployment
H N et al. (2026)	AI-assistive review	NDD and other	Moderate	Comprehensive review	Moderate	Lacks unified framework
Li et al. (2025)	AI empirical review	Inclusive education	Moderate	Engagement, outcomes	Moderate	No real-time component
Ahmed et al. (2025)	AI personalization	Students with disability	Moderate	Personalization index	High	Limited empirical proof
Proposed Framework	ML + NLP + CV (Multimodal)	Multi-disability	High (Unified)	6 key metrics	High	Computational cost

As observed in the comparative analysis above, the current frameworks are rather unidimensional in nature, with little inter-disability, inter-AI-modality and inter-real-time adaptive capabilities integration. The multimodal fusion, singular architecture and empirically verified evaluation of the proposed framework is unique to fill these gaps.

E. Identified Research Gaps

In the context of the overall review, three main areas of research gap can be defined: (1) the lack of integrated multimodal AI systems, which could effectively support multitask learners with various types of disabilities; (2) the scarcity of real-time adaptive systems, which could be dynamically responsive to both cognitive and emotional processes; (3) the lack of empirical validation of AI systems in the context of heterogeneous populations of The proposed architecture and experimental evaluation is a direct way to fill all three gaps in this study.

III. System Architecture of Ai-Driven Assistive Framework

The suggested assistive framework with AI is a multilayer intelligent architecture that is aimed to assist various needs of learners in real-time. This system incorporates the data acquisition, intelligent processing and the interactive application delivery in tightly coupled modules. The framework, in its simplest form, allows sustained observation of the behavior of learners, their mental activities, and sensory reactions, and converting raw signals of the input into practical instructions. The system integrates machine learning, natural language processing, and computer vision in a single pipeline, going beyond the constraints of the disjointed assistive tools. The architecture is highly accessible, responsive and scalable and deployment is viable in a variety of educational contexts including individual therapy sessions or inclusive smart classrooms with a variety of learner profiles.

A. Multilayer Architecture (Data, Processing, Application Layers)

The system architecture consists of three main layers which are the Data Acquisition Layer, the AI Processing Layer and the Application Delivery Layer. The Data Acquisition Layer engages the sensors, devices and channels of user interaction to receive multimodal input streams of video, audio, touch gestures, eye-tracking information and biometrics information. The main intelligent modules, i.e., the feature extraction, model inference, emotion recognition and adaptive content selection are placed at the AI Processing Layer. It is a layer that uses the principles of microservice based design to allow scalability of single AI components to be independent and allows them to be updated in a modular fashion. Application Delivery Layer offers input processed results in the form of customized learning materials, instant feedback, and teacher dashboard via user-friendly interfaces. Inter-layered communication is enhanced by a safe API middleware framework that complies with low-latency, high-throughput exchange of data and adheres to educational data privacy regulations such as FERPA and GDPR. The layered structure is fault tolerant because in case of failure of a single module, the failure does not propagate to affect the whole system. Load balancing directs computational load between cloud and edge computing nodes, and load balancing can optimize response times especially important to adaptive interventions that are real-time.

B. Data Acquisition: Learner Profiles, Behavioral and Sensory Inputs

The input stage of the framework is made up of data acquisition, which involves the holistic multimodal expression of the cognitive, behavioral and sensory state of every learner. The profiles of learners are built based on first assessment which identifies disability type, cognitive base, communication preferences and the learning history of the learner. As the system gathers interaction information, these profiles will be dynamically updated. The behavioral inputs consist of touch response patterns, task completion patterns, the route of navigation and error rates, which are all proxies of levels of engagement and comprehension. Integrated hardware such as RGB-D cameras to capture images, microphones to analyze speech, eye-tracking sensors to test the gaze, physiological sensors to measure heart rate variability and galvanic skin response are used to gather sensory inputs. These data streams are coherentized and time stamped to facilitate coherent multimodal fusion at the processing layer. The privacy-sensitive data gathering techniques such as preprocessing and anonymization algorithms on-device are given special consideration so that the sensitive learner data is not compromised anywhere in the pipeline. The system has consent management modules allowing the guardians and educators to set the parameters of data collection according to the institutional policies and ethical standards.

C. AI Modules: Machine Learning, NLP, Computer Vision Integration

The AI Processing Layer consists of three closely-knit functional modules. The Machine Learning Module uses ensemble classifiers and deep neural networks to model the competence trajectories of learners, forecast the gaps in knowledge, and make recommendations on adaptive content. Gradient boosting and recurrent neural network designs are employed to learn the temporal dynamics of learning and long-range dependency of the interaction sequences of the learner. The Natural Language Processing Module allows speech-to-text transcription, semantic analysis of student input, automatic grammatical feedback and dialogue management to conversational tutoring agents. Transformer-trained models that are fine-tuned on learning corpora can be used to nuance understand the language of learners at different levels of verbal proficiency. Computer Vision Module is used to analyse facial action units in real-time, estimate gaze direction and recognise gestures to infer cognitive and emotional states based on non-verbal behavioural information. A multimodal fusion engine integrates the results of each of the three modules with a weighted late-fusion approach that uses dynamically set confidence weights of signal quality and contextual relevance to yield a combined state of the learner which is used to make instructional adaptation decisions.

D. Communication and Interface Design for Accessibility

The Application Delivery Layer shows the outputs of the system in a collection of available interfaces created in accordance with the Web Content Accessibility Guidelines (WCAG 2.1) and Universal Design of Learning (UDL). Interfaces modalities consist of visually accessible dashboards using screen readers, voice navigation with menus, high contrast display, customizable fonts and sets of symbols, and haptic feedstrots as a channel of tactile communication. Symbol-based interaction panels and predictive communication assistant's augmentative and alternative communication (AAC) interfaces are used to help learners with severe communication impairments to interact. Teacher dashboards show real-time analytics of learner progress, engagement metrics, and estimates of cognitive load, which can be used to intervene in pedagogy in a timely manner. All interfaces are device-agnostic, and can be deployed on tablets, desktop computers, smart whiteboards and wearable assistive devices. The communication layer uses low-bandwidth optimization methods to provide functionality in connectivity-constrained environments, and expands the applicability of the framework to underserved educational environments around the world.

E. Feedback and Adaptation Mechanisms

The most critical element of operation in the proposed framework is feedback and adaptation that completes the instructional loop between the detection of the learner state and the pedagogical response. Feedback mechanism exists at two timescales, i.e. immediate feedback, which is given during milliseconds of the learner input, offers corrective hints, encouraging reinforcement or alternative form of explanation; and longitudinal feedback, which is accumulated across the sessions, informs sequencing of the curriculum, learning difficulty progression, and priority of learning objectives. The reinforcement learning agent that drives adaptive content selection is conditioned by historical data on learner outcome, and maximizes cumulative learning gain by continually improving its instructional policy. As sustained frustration/disengagement states are recognized in the emotion recognition module, the system dynamically adjusts task difficulty, content modality, or provides a motivational talk via the conversational tutoring agent. Educator alert notifications are activated when the signals to learner distress are over the set thresholds, allowing human intervention when AI-mediated interventions are not sufficient. Explicit learner agency is also facilitated by the adaptation engine, which supports students in making preferences and suppressing automated modifications, which are key to self-regulated learning behavior as well as metacognitive awareness and long-term educational success.

IV. Methodology and Model Development

This part discusses the methodological basis of the assistive learning framework proposed to use AI. It covers the data preprocessing strategies, adaptive algorithm generation, the cognitive state modeling, the personalization mechanisms as well as the overall system workflow. It is done by giving mathematical formulations to provide reproducibility and rigor.

A. Data Preprocessing and Feature Extraction Techniques

Raw multimodal data obtained as a result of interaction with learners is preprocessed in a structured pipeline before it is inputted into the model. It can be assumed that the raw input data is D = {x i, y i (i=1 to N) where x i is the multimodal feature vector and y i is the label of the learning outcome. Time-series behavioral data have missing values that are filled in by the following forward-fill interpolation:

$$\hat{x}_t = x_{t-1} + \left[\frac{x_{t+k} - x_{t-1}}{k}\right] \quad (1)$$

With t being time index and k gap length. Min-max scaling is used to perform feature normalization to keep all feature values in the range [0,1]:

$$x_i' = \frac{x_i - x_{min}}{x_{\max} - x_{min}} \quad (2)$$

Principal Component Analysis (PCA) is used to reduce dimensionality of high-dimensional visual feature vectors, and only keep components that cumulatively capture 95% of the variance. Audio inputs are converted to spectral features with Mel-frequency cepstral coefficients (MFCCs), and provide a 40-dimensional feature representation per audio frame. The behavioral sequences are transformed into temporal features by sliding window segmentation, W = 30 samples, S = 10 samples.

B. Design of Adaptive Learning Algorithms

The adaptive learning engine uses a Q-learning reinforcement model whereby the system is learning an optimal policy of instruction pi that maximizes the expected cumulative reward R in a learner interaction episode. S is the state space that represents the current level of cognition, engagement index and recent error rate, whereas A is the space of available instructional interventions, e.g., content difficulty adjustment, switching modality, and providing hints.

Algorithm 1: Adaptive Learning Policy Optimization

INPUT: Learner state s ∈ S, Action space A, Reward function R

OUTPUT: Optimal policy π* for instructional adaptation

INITIALIZE: Q(s,a) = 0 ∀ s ∈ S, a ∈ A; ε = 1.0; γ = 0.9; α = 0.01

BEGIN

FOR each learning episode e = 1 to E DO

Observe initial learner state s_0

FOR each time step t = 0 to T DO

IF rand() < ε THEN select random action a_t (exploration)

ELSE select a_t = argmax_a Q(s_t, a) (exploitation)

Execute action a_t; observe reward r_t and next state s_(t+1)

Update: Q(s_t,a_t) ← Q(s_t,a_t) + α[r_t + γ·max_a Q(s_(t+1),a) - Q(s_t,a_t)]

Decay: ε ← max(ε_min, ε · ε_decay)

s_t ← s_(t+1)

END FOR

END FOR

RETURN π*(s) = argmax_a Q(s,a)

END

C. Emotion Recognition and Cognitive State Modeling

Emotion recognition is a key element of the adaptive capacity of the framework that can be used to infer affective states of learners in real-time based on facial, physiological, and behavioral cues. The emotion recognition sub-module uses a convolutional neural network (CNN) model that has been trained on sequences of facial action units (AU) obtained on the AffectNet and FER+ datasets. Heart rate variability (HRV) and galvanic skin response (GSR) are physiological measures, which are integrated with visual images by a multimodal attention system. The cognitive state model is an internal state of the learner, defined as a Markov process with transitions of cognitive states, engaged, confused, fatigued and frustrated, and transition probabilities that are learned by interaction sequences.

The emotion and cognitive state estimation pipeline is controlled by the following equations:

$$E(t) = softmax\left(W_e \cdot \left[f_{v(t)};\ f_{p(t)}\right] + b_e\right) \quad (3)$$

The probability of emotions at time t can be expressed as equation (3) where f v(t) and f p (t) are the vectors of the visual and physiological features, respectively, W e and b e are the learned weight matrix and bias. The softmax guarantees valid output in terms of probabilities of emotion classes.

$$P(s_(t+1) \mid s_t, a_t) = A_(s_t, s_(t+1)) \cdot B_(s_(t+1), o_t) \quad (4)$$

The cognitive state transition probability in the Hidden Markov Model is defined in equation (4), where A is the state transition matrix, B is the emission probability matrix, s t is the current state, a t is the instructional action and o t is the current observation.

$$L_e = -\Sigma\, y_k \cdot \log(p_k) + \lambda \cdot ||W||^2 \quad (5)$$

The emotion classification loss function with cross-entropy loss, and L2 regularization term is shown in equation (5). In this case, y k is the actual emotion label, p k is the expected probability and lambda is used to regulate the regularization to avoid overfitting.

$$C(t) = \alpha \cdot E_{conf(t)} + \beta \cdot E_{frust(t)} + \gamma \cdot \left(1 - E_{eng(t)}\right) \quad (6)$$

The Cognitive Load Index C(t) is calculated using equation (6) as a weighted average of confusion probability Econf, frustration probability Efrust and the complement of the engagement probability Eeng with alpha, beta, gamma being empirically tuned weighting coefficients.

$$\Delta d = sign(C(t) - \theta_d) \cdot step_{size} \quad (7)$$

Difficulty adjustment delta d is controlled by equation (7) that depends on the difference between the cognitive load C(t) and the threshold theta d. The step-size is the size of the content difficulty increment or decrease used by the adaptive engine.

D. Personalization Strategies and Recommendation Systems

The personalization engine uses a collaborative-content filtering strategy that combines both collaborative and content filtering techniques to produce personalized learning content recommendations. A feature profile vector v i, which is a collection of estimates of cognitive level, indicator of learning style, previous performance scores and accommodation requirements related to disability, represents each learner. The content items are also represented in feature vectors cj that represent the level of pedagogy, mode, and skill area, and accessibility qualities. R(i,j), the recommendation score of learner i and content item j is a weighted average of the collaborative filtering prediction of matrix factorization and content based similarity scores based on the cosine distance in the shared embedding space.

The collaborative filtering uses alternating least squares (ALS) matrix factorization to factor the learner-content interaction matrix M into matrices of latent factors P and Q, and minimize reconstruction error on observed interactions. The content-based component calculates the semantic similarity between the knowledge state demonstrated by the learner and prerequisite structure of the candidate content with the help of the curriculum graph embeddings based on the graph convolutional networks. The list of final recommendations is obtained through re-ranking of the candidate items based on diversity-conscious post-processing step that would expose one to diverse content modalities to avoid excessive specialization and encourage a balanced set of skills. Temporal recency weighting is used to make sure that recommendations are based on the most recent condition of the learner, and not past averages, which is especially important in the case of learners who may vary dramatically within and between sessions in terms of cognitive and emotional conditioning. There is also explicit preference feedback in the system, using lightweight rating interfaces modified to be accessible and learners directly affect their instructional trajectory.

E. Implementation Details and System Workflow

The suggested framework is realized with the help of a Python-based stack of the backend that includes PyTorch as a deep learning model trainer and inference engine, FastAPI as a RESTful service delivery tool, and Apache Kafka as a real-time data streaming engine. The frontend interfaces are coded with React.js and accessible extensions and ARIA. The data about learner profile is stored in MongoDB NoSQL database, time-series interaction logs are processed with the help of InfluxDB to

make time queries effectively. The Model training utilizes the acceleration of computation on the NVIDIA A100 instances. The entire system process begins with the loading of learner profile at the session beginning, then proceeds to the continuous multimodal data recognition, real-time model deduction, adaptive content selection, interface display and generation of learning analytics at the session conclusion. During live sessions, educator dashboards are updated every 5 seconds and can give near real-time access to the state of learners and interventions by the system.

V. Experimental Setup And Performance Evaluation

In this part, the experimental design will be described, along with the characteristics of the dataset, evaluation measures, baseline comparisons, and validation methods that will be used to rigorously evaluate the performance of the proposed framework in different groups of learners and types of disabilities.

A. Dataset Description and Learner Group Categorization

The experimental assessment was based on a composite dataset based on three sources, such as the engagement and boredom detection data collected by the DAiSEE dataset, and the emotion recognition benchmarking data collected by the EmotiW dataset, as well as a proprietary institutional dataset collected by a special education school over 12 weeks with informed parental consent. The institutional dataset included logs of interactions among 186 students (aged 6-18) of five disability groups: autism spectrum disorder (n=42), dyslexia (n=38), hearing impairment (n=35), visual impairment (n=33), and ADHD (n=38). Data were collected in terms of records of the sessions, logs of response latency, sequence of tasks completed, physiological data, and teacher notes. Stratified sampling was used to divide the dataset into training (70%), validation (15%), and test (15%) subsets to ensure that the proportion of disability categories was preserved in partitions.

B. Evaluation Metrics

The six main measures used in performance evaluation were (1) Learning Engagement Score (LES) measured in terms of interaction frequency and duration of sustained attention; (2) Task Completion Rate (TCR) was measured as the percentage of assigned learning tasks that were completed successfully per session; (3) Comprehension Accuracy (CA) measured in terms of post-task assessment scores; (4) Response Time (RT) which was based on the average lat The difference of the performance between the two was statistically evaluated by means of paired t-tests with the Bonferonni correction in case of multiple comparisons with the significance level of $p < 0.05$.

C. Baseline Models and Comparative Methods

Three comparison systems were used to compare them: (1) Static Assistive Tool (SAT), which is a traditional non-adaptive system that includes screen readers and built-in AAC devices; (2) Rule-Based Adaptive System (RBAS), a deterministic adaptive learning platform that adjusts difficulty according to pre-set performance levels without machine learning; and (3) Single-Mode AI System (SMAS), a un The Multimodal AI-Driven Assistive Framework (MADAF) was tested in comparison to all three baselines in all six metrics of evaluation. All the systems were run on the same hardware set-ups, to allow fair comparison and educator training was equated across conditions, to reduce confounding instructional factors.

D. Experimental Scenarios and Validation Techniques

he evaluation was done in four scenarios: (1) Individual learning sessions in which learners with dyslexia developed literacy skills; (2) Mathematics concepts instruction in which learners with ADHD developed mathematics concepts; (3) Communication skills building in which learners with autism spectrum disorder developed communication skills; and (4) Science content instruction with multimodal accessibility support in which learners with visual and hearing impairments developed science concepts. The scenarios were administered in each case in a period of more than eight weeks with a twice-weekly session. To evaluate generalization of the model, cross-validation was done through 5-fold stratified cross-validation and ablation studies were done to determine the

contribution of each AI module (ML, NLP, CV) to the overall system performance. The inter-rater reliability of educator annotations was evaluated with the help of the Cohen Kappa, with the mean agreement being $\varkappa = 0.84$.

E. Statistical Analysis and Result Interpretation

Experimental data were statistically analyzed using a full set of parametric and non-parametric tests that were suitable to the characteristics of the data distribution of each metric. Shapiro-Wilk tests of normality have confirmed the approximate normality of the LES, TCR, and CA measures, which allowed the use of paired t-tests to compare groups. Response time data were skewed to the right, and required Wilcoxon signed-rank tests when comparing baselines. Cohen d was used to calculate effect sizes when making parametric comparisons and rank-biserial correlation when making non-parametric tests because it gives a standardized effect size, which does not depend on the size of the sample. All the reported percentages of improvement are the mean differences calculated using the entire test cohort, with 95% confidence intervals given to describe the uncertainty in estimating. To determine the patterns of different performances and possible interaction effects between system components and learner's characters, subgroup analyses in terms of disability were provided, to reveal information about the contextual factors that moderate AI-driven adaptive interventions.

VI. Results and Discussion

This section will provide quantitative experimental data and interpret the system performance in six key metrics, four disability specific scenarios and three baseline comparisons. There is a result table and a discussion of analytical results attached to each subsection.

A. Performance Comparison with Conventional Assistive Systems

Table II. Overall Performance Comparison Across Systems

Metric	SAT (Baseline)	RBAS	SMAS	MADAF (Proposed)
Learning Engagement Score (%)	54.3	61.8	67.2	77.4
Task Completion Rate (%)	58.6	65.4	70.1	80.7
Comprehension Accuracy (%)	61.2	67.9	72.5	81.6
Response Time (seconds)	8.4	7.1	6.5	6.0
Accessibility Index (%)	48.7	57.3	63.8	70.7
System Usability Scale (0–100)	62.1	69.4	74.8	84.3

Table II shows the overall performance of all the four systems using six evaluation metrics. The offered MADAF system performs better than all the baseline strategies in all the metrics. The Static Assistive Tool (SAT) demonstrates the worst performance, which is a manifestation of the inherent weakness of the conventional tools, which is not adaptive and is used in dynamic learning contexts. Rule-Based Adaptive System (RBAS) exhibits a slight improvement to SAT, which indicates that even deterministic adaptation can be meaningful. The Single-Modal AI System (SMAS) continues to enhance the performance with an AI-based speech processing, justifying the usefulness of machine intelligence in education. But the overall multimodal integration of MADAF has the highest scores in each dimension, with a learning engagement score of 77.4% versus 54.3 of SAT, a task completion rate of 80.7% versus 58.6 of SAT, and a system usability score of 84.3 versus 62.1 of

SAT. The time delay decreased to 6.0 seconds as opposed to the 8.4 seconds indicates the real-time adaptive effectiveness of the suggested architecture. All these findings support the hypothesis that multimodal AI integration can support education on a significantly higher level than traditional and partial AI methods, and prove the principles underlying the design of the MADAF framework.

B. Improvement Analysis in Percentage Metrics

Table III. Percentage Improvement of MADAF over Baselines

Metric	MADAF vs. SAT (%)	MADAF vs. RBAS (%)	MADAF vs. SMAS (%)
Learning Engagement	+42.6	+25.2	+15.1
Task Completion Rate	+37.8	+23.4	+15.1
Comprehension Accuracy	+33.4	+20.2	+12.6
Response Time Reduction	+28.6	+15.5	+7.7
Accessibility Index	+45.2	+23.4	+10.8
System Usability Scale	+35.7	+21.5	+12.7

Table III estimates the extent of the improvement that MADAF would have made compared to each baseline system. The largest improvements in the Accessibility Index are also the most significant, with MADAF showing a 45.2% relative improvement over SAT, which highlights the transformative nature of multimodal accessibility design. The effectiveness of the emotion-aware adaptive engine in maintaining the learner motivation and attention is observed in the learning engagement improvements of 42.6% over SAT. The adaptive difficulty calibration and real-time management of cognitive load results in task completion increases of 37.8% that directly translate to quantifiable productivity gains. The 28.6 percent decrease in response time compared to SAT suggests that the mechanism of predictive prefetching and content pre-rendering of the system is effective to reduce the latency of instruction. The gains of comprehension of 33.4% verify that the personalized content sequencing promotes knowledge retention and conceptual understanding. The gradual decrease in the improvement margins between SAT, RBAS and SMAS justifies the role played by each layer of the architecture and proves that the entire multimodal integration is required to reach the optimal performance as opposed to any single component.

C. Usability and Accessibility Enhancements

Table IV shows a disaggregated usability and accessibility data by individual modalities of the interface, which gives us an understanding of which dimensions MADAF yields the largest benefits. The accuracy of speech-to-text increased significantly by 31.5% in comparison to the SAT (71.2) to 93.6% (MADAF), which is based on the domain-specific language model of the NLP module that was trained on special education vocabulary and learner speech patterns such as non-standard articulation profile typical of ASD and hearing impairment groups.

Table IV. Accessibility and Usability Results by Interface Modality

Interface Modality	SAT	RBAS	SMAS	MADAF
Speech-to-Text Accuracy (%)	71.2	74.8	82.4	93.6
Screen Reader Compat. (%)	68.4	72.1	75.8	91.3
Symbol/AAC Interface Use (%)	52.3	60.7	66.2	80.4
Gesture Recognition Acc. (%)	N/A	N/A	61.4	88.2

Educator Satisfaction (SUS)	63.8	70.2	76.5	86.7

Compared to SAT, semantic HTML structuring and extensive WCAG 2.1 compliance, screen reader compatibility was 91.3% in MADAF, which was 68.4% in SAT. There was a 53.7% improvement in symbol and AAC interface use between SAT and MADAF, and the predictive engine of the AAC interface made a great contribution to the usability of non-verbal learners. The high gesture recognition of 88.2% in MADAF and 61.4% in SMAS show that the multi-stream hand pose estimation pipeline in the computer vision module is the best. Teacher satisfaction assessed through the System Usability Scale increased to 86.7 compared to 63.8 and MADAF was rated as excellent (SUS > 85) regarding usability, and this result shows that MADAF is practically acceptable to use in the classroom.

D. Impact on Different Disability Categories

Table V will give a more detailed breakdown of the performance of MADAF within the five categories of disabilities being assessed, with significant results of varying effectiveness. The learners with dyslexia had the highest rate of task completion (82.4) and a high accuracy of text understanding (84.1) which could be explained by the adaptive phonological decoding support and the text-to-speech rendering with dyslexia-friendly font overlay provided by the NLP module. The learners with ADHD scored highly on engagement (78.6%), and task completion (80.9%), and served as positive beneficiaries of the gamification features and short-burst activity sequencing introduced by the adaptive content engine.

Table V. Performance Results by Disability Category (MADAF System**)**

Disability Category	LES (%)	TCR (%)	CA (%)	RT (sec)	AI (%)
ASD (n=42)	74.8	78.3	79.6	6.3	68.4
Dyslexia (n=38)	79.2	82.4	84.1	5.7	69.8
Hearing Impairment (n=35)	76.4	80.1	81.3	5.9	73.2
Visual Impairment (n=33)	77.8	81.6	82.7	6.1	74.6
ADHD (n=38)	78.6	80.9	80.2	5.8	68.9
Overall Mean	77.4	80.7	81.6	6.0	71.0

Visual impaired learners had the best accessibility index (74.6) and good comprehensibility (82.7) that indicates the effectiveness of multimodal content presentation with a combination of audio description, simulation of tactile feedback, and integration of a screen reader. The learners with hearing impairments demonstrated the same level of improvement and were associated with the accessibility index of 73.2% due to real-time captioning accuracy and sign language recognition integration. Although ASD learners had the lowest task completion rate (78.3), compared to other groups, they showed a 41.2% improvement compared to their SAT baseline performance which affirms the usefulness of predictable interface structure, visual schedule support, and lower-sensory overload design principles included in the MADAF interface. All these findings indicate the capacity of the framework to provide substantive, disability-responsive assistance to a heterogeneous group of learners.

VII. Conclusion and Future Scope

This study has introduced, developed, and tested a holistic AI-based assistive learning system to support inclusive learning among students with various disabilities. The Multimodal AI-Driven Assistive Framework (MADAF) proposed

combines machine learning, natural language processing, and computer vision into a single multilayer framework that can deliver adaptive instructions in real-time. Experimental testing of 186 learners with five types of disability showed that the system had a significant performance benefit compared to conventional and partial AI baseline systems with 42.6, 37.8, 33.4, 28.1, and 45.2 improvements in the engagement of learners, completion of tasks, accuracy in understanding, response time, and index of accessibility, respectively. There are four major scientific contributions of this work. First, the multimodal fusion architecture is a new combination of different sensing modalities and AI inference engines into a single educational support platform. Second, the emotion-sensitive reinforcement learning agent facilitates truly responsive instructional adjustment that is fine-tuned to instant-by-instant cognitive and affective learner condition. Third, the disability-stratified experimental design offers finer-grained evidence on the effectiveness of systems difference, which can further develop the evidence base on AI in special education. Fourth, the scalable deployment structure exhibits practical feasibility to real world classroom implementation. The modular nature of the framework makes it a scaled solution that can be deployed in smart classrooms, remote education platforms, and in hybrid classrooms. Pathways to integration with full-fledged learning management systems such as Moodle, Canvas and Google Classroom are known, and adoption can proceed without wholesale institutional change in infrastructure. Future directions in research are the design of federated learning systems that allow improving models privately across distributed school networks, more sophisticated AI ethics systems that regulate the transparency of algorithmic decision-making and the sovereignty of learner data, and cross-platform integration with mobile and wearable assistive systems to support beyond classroom settings. Neuroimaging-based cognitive state modeling and integration of large language models in conversational tutoring are interesting areas to explore in the next generation.

References

Ahmed, S., Rahman, M. S., Kaiser, M. S., & Hosen, A. S. M. S. (2025). Advancing personalized and inclusive education for students with disability through artificial intelligence: Perspectives, challenges, and opportunities. Digital, 5(2), 11. https://doi.org/10.3390/digital5020011

Alnajdi, S. M., Salem, M. A., & Elshaer, I. A. (2025). Examining the acceptance and use of AI-based assistive technology among university students with visual disability: The moderating role of physical self-esteem. Bioengineering, 12(10), 1095. https://doi.org/10.3390/bioengineering12101095

Alsaleh, A. (2026). The influence of artificial intelligence on individuals with disabilities. Acta Psychologica, 262, 106010. https://doi.org/10.1016/j.actpsy.2025.106010

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

Fitas, R. (2025). Inclusive education with AI: Supporting special needs and tackling language barriers. AI Ethics, 5, 5729–5757. https://doi.org/10.1007/s43681-025-00824-3

Chavan, R., Shakadwipi, P. A. J., Singavi, D., Sonawane, R., & Burse, T. (2026). A Smart Barcode Identifier & Recovery System for Lost Individuals in Mass Gatherings. International Journal on Advanced Computer Theory and Engineering, 15(1S), 18–21. https://doi.org/10.65521/ijacte.v15i1S.1299

H N, N. K., Patil, C. M., K V, S., B, J., A P, A., & Mahadevaswamy. (2026). Inclusive learning through AI: A comprehensive review of assistive educational technologies for students with neurodevelopmental disorders and other disabilities. Journal of Engineering Education Transformations, 39(Special Issue 2), 332–340. https://doi.org/10.16920/jeet/2026/v39is2/26041

Harkins-Brown, A. R., Carling, L. Z., & Peloff, D. C. (2025). Artificial intelligence in special education. Encyclopedia, 5(1), 11. https://doi.org/10.3390/encyclopedia5010011

Kooli, C., & Chakraoui, R. (2025). AI-driven assistive technologies in inclusive education: Benefits, challenges, and policy recommendations. Sustainable Futures, 10, 101042. https://doi.org/10.1016/j.sftr.2025.101042

Li, J., Yan, Y., & Zeng, X. (2025). Exploring artificial intelligence in inclusive education: A systematic review of empirical studies. Applied Sciences, 15(23), 12624. https://doi.org/10.3390/app152312624

Memisevic, H., & Hadzic, S. (2025). Assistive technology in inclusive education. In G. Bennett & E. Goodall (Eds.), The Palgrave encyclopedia of disability. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-40858-8_508-1

Navas-Bonilla, C. R., Guerra-Arango, J. A., Oviedo-Guado, D. A., & Murillo-Noriega, D. E. (2025). Inclusive education through technology: A systematic review of types, tools and characteristics. Frontiers in Education, 10, 1527851. https://doi.org/10.3389/feduc.2025.1527851

Pagliara, S. M., Bonavolonta, G., Pia, M., Falchi, S., Zurru, A. L., Fenu, G., & Mura, A. (2024). The integration of artificial intelligence in inclusive education: A scoping review. Information, 15(12), 774. https://doi.org/10.3390/info15120774

Paglialunga, A., & Melogno, S. (2025). The effectiveness of artificial intelligence-based interventions for students with learning disabilities: A systematic review. Brain Sciences, 15(8), 806. https://doi.org/10.3390/brainsci15080806

Saabi, N., Chlioui, I., & Radgui, M. (2025). A review of assistive technology in special education. Engineering Proceedings, 112(1), 45. https://doi.org/10.3390/engproc2025112045

Tsvetkova, P., Lekova, A., Simov, A., & Mitevska, M. (2025). Identifying future trends in AI-driven assistive technologies: Insights from a national Delphi survey of stakeholder perspectives. Societies, 15(9), 246. https://doi.org/10.3390/soc15090246