Archives available at journals.mriindia.com

International Journal on Advanced Computer Theory and Engineering

ISSN: 2319 - 2526

Volume 15 Issue 01s, 2026

AI-Driven Inclusive Assessment Tool for Education

[1]Devanand Patil, [2]Kartik Ugale, [3]Payal Naphade, [4]Divya Jadhao

[1 2 3 4] *Department of Computer Engineering, SSBT's College of Engineering and Technology, Jalgaon, Maharashtra 425001, India*

Email: [1] *pdevas531@email.com*, [2]*ugalekartik2004@email.com*, [3] *payalnaphade2004@email.com*, [4] *jadhaodivya27@email.com*

Peer Review Information

Submission: 08 Dec 2025

Revision: 25 Dec 2025

Acceptance: 10 Jan 2026

Keywords

AI-Driven Assessment, Inclusive Education, Educational Technology, Personalized Learning, Learning Analytics

Abstract

Traditional exam systems usually depend on fixed question sets and manual grading. Such methods fail to address differences in students' skills or accessibility needs. In this work, an AI-based assessment platform is presented that focuses on flexibility and inclusivity. The system uses Natural Language Processing (NLP) techniques to automate the evaluation of descriptive answers. It applies adaptive logic to pick multiple-choice questions (MCQs) that match a learner's level of understanding. Handwritten responses are checked with Optical Character Recognition (OCR) to keep the marking process accurate and fair. For better accessibility, Speech-to-Text (STT) and Text-to-Speech (TTS) tools are added, helping students who face visual or motor challenges take the test comfortably. The overall system follows a modular architecture, connecting React.js for the user interface, Node.js for backend operations, and Python-based AI modules for question generation and scoring. This paper focuses on the system design, methodology, and prototype-level feasibility analysis of an inclusive AI-driven assessment framework.

Introduction

Assessment forms the foundation of the educational process, influencing both teaching practices and learning outcomes. Traditional evaluation methods-such as handwritten answer sheets, static sets of multiple-choice questions (MCQs), and manually graded responses-often face challenges of inefficiency, subjectivity, and limited accessibility. These conventional systems slow down feedback cycles and fail to accommodate the diverse needs of learners. Students with visual or motor impairments, and those who prefer handwritten responses, frequently experience barriers in existing digital examination environments.

Recent advancements in Artificial Intelligence (AI) present opportunities to overcome these challenges through automation, adaptivity, and inclusivity in assessment. Transformer-based models enable semantic understanding for descriptive answers, while adaptive algorithms dynamically adjust question difficulty based on learner performance, ensuring fairness and personalization. Optical Character Recognition (OCR) supports digitization and evaluation of handwritten content, and Speech-to-Text (STT) along with Text-to-Speech (TTS) technologies enhance accessibility for differently-abled learners.

The proposed system is being developed as a modular and scalable framework integrating these AI-driven components into a unified assessment platform. It consists of four primary modules:

1. **Transformer-based semantic grading** for automated descriptive answer evaluation.
2. **Adaptive MCQ engine** that adjusts difficulty according to learner performance.

3. **OCR-enabled handwriting processing** for evaluating pen-and-paper submissions.
4. **STT/TTS interfaces** for accessibility support and inclusive interaction.

The work focuses on practical implementation aspects such as algorithmic design, adjustment, technology stack configuration, and RESTful API integration. A teacher-in-the-loop mechanism is also included to ensure fairness and maintain alignment between automated grading and human judgment. Future evaluation studies will focus on validating the system's reliability, inclusivity, and usability across diverse learner profiles.

Objectives

1. Develop a modular assessment platform supporting typed, handwritten, and spoken inputs.
2. Provide accessibility through STT/TTS and WCAG-compliant user interfaces.
3. Automate descriptive grading using transformer-based semantic models.
4. Implement an adaptive MCQ engine for proficiency-based question delivery.
5. Establish evaluation metrics for analyzing system performance, inclusivity, and usability.

Literature Review

Artificial Intelligence (AI) has significantly influenced modern educational assessment systems, providing new methods to automate grading, personalize learning, and enhance inclusivity. However, despite the growing adoption of AI in education, several studies indicate persistent gaps in adaptability, multimodal support, and accessibility for differently-abled learners. This section reviews prior works that collectively form the foundation of the proposed inclusive assessment framework.

Owan et al. [1] examined the potential of AI tools in educational measurement and assessment, emphasizing automation, reliability, and consistency in grading. Their research demonstrated that AI can reduce human bias and ensure faster evaluation, yet the study primarily focused on digital environments. It lacked mechanisms for handling handwritten submissions or accessibility features for students with disabilities. The proposed system builds upon this foundation by expanding automation to multimodal inputs such as handwriting, speech, and text, thereby ensuring fairness and inclusivity for all learners.

Mitra [2] developed an AI-powered adaptive learning model aimed at improving accessibility for disabled learners. The model dynamically adjusted content difficulty based on learner performance, offering speech narration and adaptive interfaces. The study reported improved participation among visually and motor-impaired students but did not integrate assessment-based feedback or semantic evaluation. In contrast, the proposed system incorporates adaptive testing alongside descriptive evaluation to measure understanding through both objective and descriptive formats, thus achieving a comprehensive learning assessment approach.

Çela et al. [3] explored the role of AI in vocational and technical education, focusing on skill-based assessment and performance monitoring. Their findings highlighted how AI can automate practical evaluations and align assessment criteria with industry standards. However, the study largely concentrated on quantitative skill measurement, overlooking qualitative evaluation such as descriptive or conceptual understanding. The present system addresses this gap by introducing transformer-based semantic grading that evaluates the meaning and context of written responses rather than depending solely on keywords or numerical data.

Mehrabi and Morphew [4] proposed an AI-based assessment tool for identifying partial and full mastery levels among students in large engineering classrooms. Their model successfully analyzed performance analytics to classify learner proficiency levels, ensuring scalability and objectivity. Nevertheless, it lacked qualitative feedback mechanisms and multimodal evaluation. The proposed framework extends this approach by combining adaptive MCQ selection, NLP-based descriptive evaluation, and feedback generation using large language models, enabling continuous and personalized assessment cycles.

Singh and Kumar [5] presented an OCR-based deep learning approach for recognizing and processing handwritten answers. Their research achieved over 90% accuracy in text extraction from high-quality scans, proving the feasibility of pen-and-paper digitization. However, their study focused solely on text recognition and did not address semantic grading or integration with adaptive systems. The current work builds upon this by integrating OCR with NLP-based evaluation, where recognized handwritten responses are semantically analyzed using BERT-based embeddings to produce context-aware grading and feedback. This combination enables the inclusion of traditional handwritten exams in digital AI-powered assessment systems.

In summary, existing literature shows strong progress in AI-driven automation and adaptive education; however, it often fails to integrate adaptability, accessibility, and multimodal assessment within a single framework. The proposed system addresses these gaps by combining adaptive multiple-choice evaluation, transformer-based descriptive grading, OCR-supported handwritten assessment, and STT/TTS-enabled accessibility. This unified approach aims to provide a fair, efficient, and inclusive evaluation environment that supports diverse learner needs, including persons with disabilities (PWD).

SYSTEM ARCHITECTURE

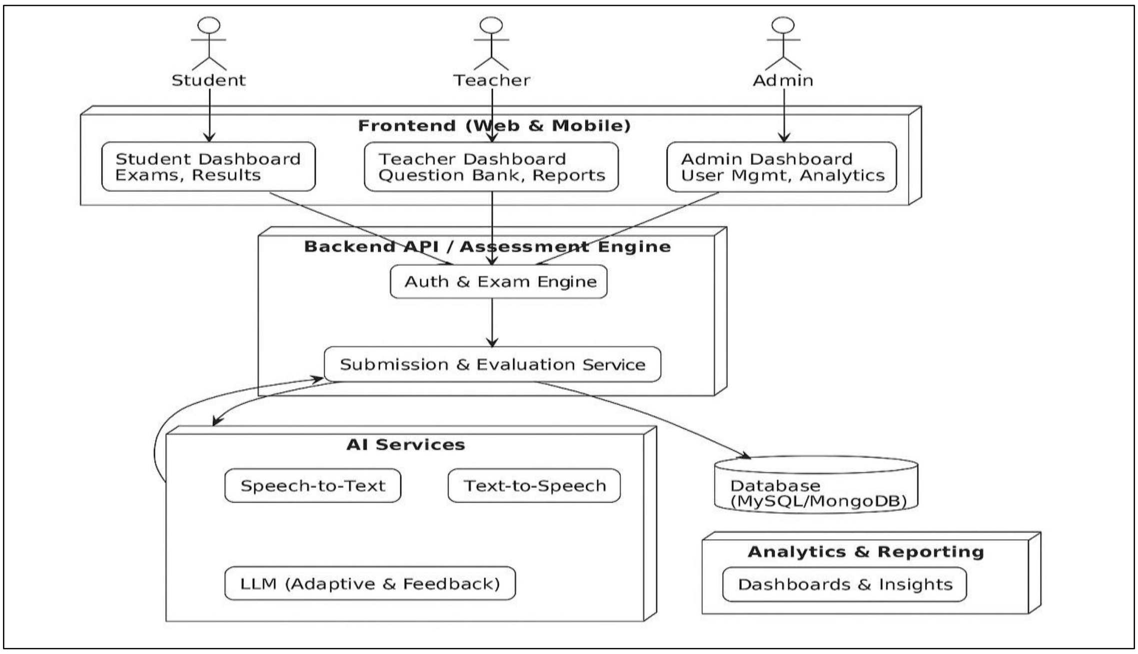

Figure 1: High-level system architecture: Frontend Interface, Backend Application Layer, AI microservices Layer, and Database Layer

The system follows a service-oriented architecture to ensure modularity, scalability, and ease of maintenance. Each functional unit is designed as an independent component, allowing efficient communication and seamless integration through RESTful APIs. The major components of the system are described below:

- **Frontend(React.js):** Provides role-based interfaces for students, teachers, and administrators. Accessibility features such as keyboard navigation, large-font display, and contrast toggling are included to support users with visual or motor impairments.
- **Backend(Node.js and Express.js):** Acts as the core application layer responsible for authentication using JSON Web Tokens (JWT), request validation, API routing to AI microservices, logging, and system orchestration.
- **AI Microservices (Python):** Comprise independent services for (a) semantic evaluation of descriptive answers, (b) OCR-based handwriting processing, (c) adaptive MCQ selection, and (d) Speech-to-Text and Text-to-Speech handling. Each microservice exposes RESTful endpoints for backend communication.
- **Database Layer:** Uses MySQL to store transactional data such as user profiles, submissions, and evaluation records. An optional Elasticsearch module can be integrated to enable fast text search and analytical querying over large datasets.

Methodology

The proposed system follows a modular, AI-driven architecture designed to adapt to learner performance and support multiple assessment types, including objective, descriptive, and handwritten evaluation. Each component will integrate through a unified backend API, ensuring interoperability, scalability, and real-time adaptability. The overall workflow employs Artificial Intelligence (AI), Natural Language Processing (NLP), and deep learning models to make assessments both personalized and inclusive.Descriptive Answer Evaluation (AI-Based Semantic Scoring)

For evaluating descriptive or long-answer questions, the system will use semantic similarity analysis powered by deep learning models. Traditional keyword-matching systems fail to understand conceptual meaning; therefore, a fine-tuned BERT (Bidirectional

Encoder Representations from Transformers) model will be employed to measure semantic equivalence between the students' response and the instructor's reference answer. The workflow will proceed as follows:

1. **Input Processing:** The instructor will upload model answers and optional rubrics via the web interface. The student's response will be collected and preprocessed (tokenization, stop-word removal, punctuation normalization) using the NLTK and Transformers libraries in Python.
2. **Embedding Generation:** Both model and student answers will be encoded into 768-dimensional contextual embeddings using a fine-tuned BERT model from HuggingFace. These embeddings will capture deep semantic meaning rather than surface-level word similarity.
3. **Similarity Computation:** The cosine similarity between the reference embedding (R) and the student embedding (S) will be computed as:

$$Similarity(R, S) = \frac{(R \cdot S)}{(||R|| \times ||S||)} \quad (1)$$

4. **AI-Based Scoring:** The computed similarity score will be passed to a Python-based scoring engine that maps it to a mark range based on predefined thresholds. Scores above an empirically chosen threshold (e.g., 0.85) indicate high conceptual understanding, while lower scores trigger AI-assisted feedback.
5. **Feedback Generation:** Using a Langflow pipeline connected to a large language model (LLM), the system will generate personalized feedback, such as "Your explanation lacks mention of the core concept of Newton's Law of Motion."

The complete semantic evaluation process will run as a dedicated Python microservice, deployed alongside the backend. It will communicate via REST APIs developed in Node.js and Express.js, while the React.js frontend will visualize both the similarity score and AI-generated feedback for the student. This approach aims to achieve fully automated, AI-assisted descriptive answer evaluation that closely replicates human grading.

Adaptive MCQ Selection (AI-Driven Difficulty Adjustment)

For multiple-choice questions (MCQs), the system will implement an adaptive AI algorithm to dynamically adjust question difficulty in real time based on student performance. Instead of presenting static quizzes, the platform will continuously monitor student responses and modify subsequent questions accordingly.

Each question in the dataset carries a difficulty value $Dq \in [0, 1]$. The learner's proficiency score Pt evolves after each question using the adaptive upate formula:

$$Pt + 1 = Pt + \eta(rt - Pt) \quad (2)$$

where rt represents the correctness of the student's current answer (1 for correct, 0 for incorrect), and η is the learning rate controlling adaptability. The next question's difficulty will be matched to the updated proficiency level Pt+1, ensuring personalized progression.

In implementation:

1. The adaptive logic will be created using a visual AI workflow in **Langflow**, where nodes represent decision chains-answer correctness, proficiency computation, and next-question selection.
2. A **TensorFlow**-based prediction model within Langflow will enhance adaptability by predicting the optimal next question based on historical data (accuracy trends, time spent, and response confidence).
3. The **Node.js backend** will trigger this Langflow pipeline through REST API calls after each question submission.
4. The resulting JSON response, containing the next question ID and its difficulty level, will be rendered in real time by the **React.js frontend**.

This hybrid combination of **Langflow + TensorFlow** will allow adaptive questioning without complex backend logic. The AI model will continuously refine its understanding of the learner's skill level, ensuring a personalized and engaging assessment experience.

Pen-and-Paper Processing (OCR + NLP Integration)

To handle handwritten submissions, the system will incorporate Optical Character Recognition (OCR) and NLP pipelines:

1. Scanned answer sheets will be uploaded through the frontend.
2. Images will be preprocessed using **OpenCV** for noise removal, binarization, and skew correction.
3. Cleaned images will be processed by **Tesseract OCR (Python)** to extract text.
4. Extracted text will undergo grammar correction and tokenization before being semantically evaluated through the same BERT-based scoring mechanism used for descriptive answers.
5. If either OCR confidence or semantic similarity falls below the predefined threshold, the system will flag the response for manual verification to maintain reliability and fairness.

Accessibility and Inclusive Interaction

Accessibility features will be directly integrated into the assessment process to support differently-abled learners:

1. **Text-to-Speech (TTS):** Questions and feedback will be converted into spoken form using the Google Speech Synthesis API.
2. **Speech-to-Text (STT):** Students will be able to answer verbally, and their responses will be transcribed into text using the Google Speech Recognition API.
3. Both voice-based interactions will be processed through the same AI evaluation pipeline to ensure uniform grading across all response modes.

System Integration and Flow

The complete architecture will integrate all modules seamlessly:

1. **Frontend (React.js):** Displays adaptive MCQs, descriptive answer editors, and accessibility features.
2. **Backend (Node.js + Express.js):** Manages routes, data flow, and communication with Langflow and Python services.
3. **AI Services (Langflow + TensorFlow + BERT):** Handle adaptive question selection, semantic evaluation, and feedback generation.

All modules will interact through RESTful APIs and share a central JSON dataset. This architecture ensures modularity, real-time adaptability, and inclusivity, making the system capable of addressing diverse educational needs effectively.

Prototype Evaluation And Feasibility Analysis

The proposed AI-Driven Inclusive Assessment Tool is currently under implementation and testing. Based on system design and theoretical evaluation, the following observations demonstrate the feasibility of the proposed architecture and will guide future large-scale empirical evaluation.

1. **Automated Descriptive Evaluation:** The NLP-based answer evaluation module is designed to produce consistent grading results comparable to human evaluators, with the goal of reducing subjectivity in descriptive answer assessment.

2. **Pen-Paper Evaluation:** The Optical Character Recognition (OCR) and NLP integration enables the system to automatically process handwritten responses and assess them semantically. The module is designed to handle standard-quality scanned answer sheets effectively.

3. **Adaptive Question Selection:** The adaptive MCQ system dynamically adjust the difficulty level of questions according to learner performance, providing a more accurate estimation of proficiency with fewer questions than static tests

4. **Accessibility for PWD Learners:** The inclusion of Text-to-Speech (TTS), Speech-to-Text (STT), and an accessible interface will help students with disabilities to comfortably take tests without manual typing or reading barriers.

These expected results will be verified during the final testing phase.

Conclusion and Future Work

The proposed system aims to create an inclusive AI-driven assessment platform combining adaptive MCQ selection, semantic descriptive evaluation, OCR-based handwritten answer processing, and accessibility through speech interfaces. The model is being developed to enhance fairness and automation in examinations while supporting PWD candidates. Future work includes multilingual support, privacy-focused model optimization, and the exploration of rubric-driven and logic-aware grading models to improve factual correctness beyond baseline semantic similarity.

References

[1] Owan, V.J., Abang, K.A., Idika, D.O., Etta, E.O., Bassey, B.A.: Exploring the Potential of Artificial Intelligence Tools in Educational Measurement and Assessment. Eurasia Journal of Mathematics, Science and Technology Education, 19(8), 1–15 (2023)

[2] Mitra, S.: AI-Powered Adaptive Education for Disabled Learners. SSRN Electronic Journal (2024)

[3] Çela, E., Eappen, P., Vedishchev, A.: Artificial Intelligence in Vocational Education and Training. IGI Global (2024)

[4] Mehrabi, A., Morphew, J.W.: AI Skills-Based Assessment Tool for Identifying Partial and Full-Mastery within Large Engineering Classrooms. In: ASEE Annual Conference Proceedings (2024)

[5] Singh, A., Kumar, R.: Automated Handwritten Text Recognition Using Deep Learning and OCR. IEEE Access, 11, 35421–35433 (2023)

[6] Korada, L., Sikha, V.K., Siramgari, D.: AI and Accessibility: A Conceptual Framework for Inclusive Technology. International Journal of Emerging Research, 12(4), 45–53 (2024)

[7] Nagajayant Nagamani. (2022). Explainability and Robustness Trade-offs: Ensuring Safety and Fairness in Large-Scale AI Deployments. International Journal of Intelligent Systems and Applications in Engineering, 10(3s), 484–494. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/8017