

Design and Development of an AI-Based Mock Interview Platform for Skill Assessment

Prachi Gupta¹, Aditya Saini², Aman Kumar Sharma³, Aman Kumar Pandey⁴, Ahad Ali⁵
^{1,2,3,4,5}Department of Computer Science and Engineering, Moradabad Institute of Technology, Moradabad, India

Abstract— Interview preparation remains a significant challenge for students due to limited access to personalized practice sessions and structured evaluation. Traditional mock interviews rely heavily on human experts, making them expensive, time-consuming and difficult to scale. This paper presents the design and implementation of an AI-driven Mock Interview Platform that automates interview simulation and performance assessment. The system utilizes generative language models for domain-specific question generation, speech-to-text technology for response transcription, and Natural Language Processing (NLP) techniques for semantic evaluation of answers. A weighted scoring mechanism evaluates response relevance, completeness and clarity to generate structured feedback reports. The platform was tested with 40 students across multiple technical domains, where 75% of users showed measurable improvement in answer structuring after repeated sessions. The proposed system demonstrates a scalable, cost-effective and intelligent solution for enhancing interview readiness and employability skills, bridging the gap between academic preparation and industry expectations.

Keywords— Artificial Intelligence, Generative AI, Natural Language Processing, Speech Recognition, Automated Interview Evaluation, Employability Skill Assessment

I. INTRODUCTION

Interviews constitute a critical stage in the recruitment process, serving as a primary mechanism for assessing a candidate's technical knowledge, communication skills and problem-solving ability. Despite possessing adequate academic knowledge, many students struggle to perform effectively in interviews due to limited practice opportunities, performance anxiety, and a lack of structured feedback. Traditional mock interview sessions depend heavily on human evaluators, making them costly, time-intensive and difficult to scale across large student populations.

With the rapid advancement of Artificial Intelligence (AI) [1], particularly in generative language models [2] and Natural Language Processing (NLP) [3], automated systems are now capable of simulating human-like conversational interactions. AI-driven platforms can generate domain-specific interview questions, analyze candidate responses using semantic understanding and provide structured performance feedback in real time. Such systems offer scalable, consistent and cost-effective alternatives to conventional interview preparation methods.

This paper proposes the design and implementation of an AI-driven Mock Interview Platform that automates interview simulation, response evaluation and structured feedback generation to enhance interview readiness and employability skills.

II. RELATED WORK

Several platforms and research studies have addressed interview preparation using digital tools and artificial intelligence techniques. Google Interview Warmup [4] provides AI-generated questions and textual keyword-based feedback; however, it lacks structured scoring mechanisms and detailed performance analytics.

Platforms such as Pramp rely on peer-to-peer mock interviews, enabling real-time interaction but depending heavily on participant availability and subjective evaluation.

Recent research in affective computing has explored speech emotion recognition and facial expression analysis using deep learning models to evaluate candidate behavior [5], [6]. While these approaches contribute to behavioral assessment, they often focus on single-modality analysis and do not integrate complete interview simulation with automated semantic evaluation.

Automated grading systems for short-answer evaluation have long utilized semantic similarity measures and dependency graph alignments [12]. Recent advancements using generative AI and large language models [7], [8] further enhance contextual evaluation capabilities. Additionally, transformer-based architectures [9] have significantly improved contextual representation learning, enabling more accurate response evaluation beyond simple keyword matching.

Despite these advancements, most existing systems provide either static question banks, peer-based practice, or limited evaluation mechanisms. There remains a gap in developing an integrated platform that combines dynamic AI-based question generation, speech-to-text processing, semantic response evaluation, structured scoring, and automated feedback within a unified framework.

A comparative analysis of existing systems and the proposed platform is presented in Table I.

TABLE I. COMPARATIVE ANALYSIS OF EXISTING INTERVIEW PREPARATION SYSTEMS

System Study	Question Generation	Response Evaluation	Structured Feedback	Limitation
Google Interview Warmup	AI-based	Keyword-based	Limited	No scoring Mechanism
Pramp	Peer-based	Human Evaluation	Subjective	Not Scalable
Emotion Recognition Systems	No	Behavioural Analysis	Partial	No content evaluation
Proposed System	AI-generated	Semantic NLP-based	Structured and Automated	Internet Dependent

The proposed system addresses these limitations by integrating automated question generation, semantic response analysis, and structured scoring within a scalable AI-driven framework.

III. PROPOSED SYSTEM

The proposed AI-Driven Mock Interview Platform is a web-based intelligent system designed to automate interview simulation, response evaluation and structured feedback generation within a unified framework. The system aims to replicate real-world interview conditions while ensuring objective and scalable assessment.

The platform operates in a structured workflow. Initially, users select interview parameters such as domain, job role and experience level. Based on these inputs, a generative AI model dynamically produces domain-specific interview questions aligned with industry standards. The questions are presented in textual or voice format to simulate a realistic interview environment.

User responses are recorded either as text or speech. In the case of spoken responses, a speech-to-text module transcribes the audio into textual form. The transcribed response is then processed by the AI analysis engine, which applies NLP and semantic similarity techniques to evaluate relevance, correctness, clarity and completeness.

A weighted scoring mechanism generates performance metrics, and the system produces a structured feedback report highlighting strengths and areas for improvement. All session data is securely stored to enable performance tracking across multiple attempts.

Unlike existing systems that rely solely on keyword matching or peer-based evaluation, the proposed platform integrates dynamic question generation, semantic analysis, automated scoring, and structured feedback within a single scalable architecture.

IV. SYSTEM ARCHITECTURE

The proposed AI-driven Mock Interview Platform follows a modular and scalable architecture, where each component performs a specific functional role. As illustrated in Fig. 1, the system integrates user interaction, AI-based question generation, response processing, semantic evaluation and structured feedback generation within a unified framework.

The architecture ensures efficient data flow, secure storage, and independent module scalability. The complete workflow of the system can be represented as:

User → Question Generator → Response Recorder → Speech-to-Text (STT) → NLP Analysis → Evaluation Engine → Feedback Module → Database

This structured pipeline enables automated interview simulation and objective performance assessment.

A. User Authentication and Dashboard

This module acts as the system entry point. It provides secure login and registration mechanisms, ensuring controlled access. Upon authentication, users are redirected to a personalized dashboard where they can initiate interview sessions, view previous reports and track performance trends. This module ensures data integrity and session management.

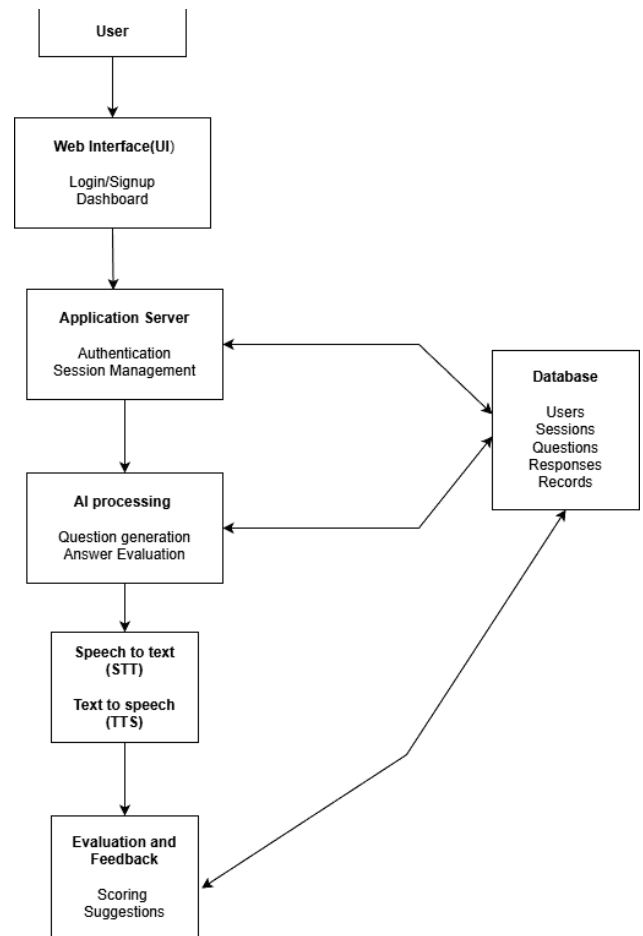


Fig. 1. System Architecture of the Proposed AI Mock Interview Platform

B. AI Question Generation and Voice Interface

Based on user-selected parameters (domain, job role, experience level), the system dynamically generates interview questions using a generative AI model. The questions are delivered either in text format or via a text-to-speech interface to simulate real interview conditions. This module enhances realism and contextual relevance.

C. Response Recording Module

User responses are captured in real time through text input or audio recording. Audio responses are securely stored and forwarded to the speech processing module. This component ensures accurate data capture and maintains user privacy through secure storage mechanisms.

D. Speech-to-Text and NLP Analysis Engine

The recorded audio is transcribed into text using a Speech-to-Text (STT) system. The transcribed response is processed using Natural Language Processing techniques to perform semantic similarity analysis, keyword relevance detection, and completeness evaluation. Unlike keyword-only systems, this module evaluates conceptual understanding rather than rigid answer matching.

E. Evaluation and Scoring Engine

The evaluation module applies a weighted scoring mechanism to assess:

- Content relevance
- Concept accuracy
- Answer completeness
- Communication clarity

The scoring output is normalized and forwarded to the feedback module.

F. Feedback and Report Generation

This module generates structured performance reports highlighting strengths, areas for improvement, and suggestions for enhancement. The results are stored in the database, enabling longitudinal performance tracking across multiple sessions.

V. METHODOLOGY

The proposed AI-driven Mock Interview Platform follows a structured workflow that simulates real-world interview conditions while enabling automated, objective evaluation. Initially, users authenticate and select interview parameters such as domain, job role, and experience level. Based on these inputs, the system generates domain-specific interview questions using a generative AI model.

User responses are captured either through text input or speech. For audio responses, a Speech-to-Text (STT) module converts speech into text for further processing. The responses are then analyzed using Natural Language Processing (NLP) techniques to assess semantic relevance, conceptual accuracy, completeness, and clarity of communication.

A weighted scoring mechanism calculates the final performance score based on predefined evaluation metrics. Finally, the system generates a structured feedback report highlighting strengths and areas for improvement and stores session data in the database for future performance tracking.

VI. TECHNOLOGY USED

The proposed AI-based Mock Interview Platform is implemented using modern web technologies and AI services to ensure scalability, performance, and modularity. The key technologies used in the system are summarized in Table II.

TABLE II. TECHNOLOGY USED IN SYSTEM IMPLEMENTATION

Component	Technology Used	Purpose
Frontend	Next.js	User Interface & Client Interaction
Backend	Next.js API Routes	Server-side Logic & API Handling
ORM	Drizzle ORM	Database Query Management
Database	Neon (PostgreSQL)	Secure Cloud Data Storage
AI Model	Google Gemini API	Question Generation & Semantic Analysis
Speech Processing	Speech-to-Text API	Audio Transcription

The system utilizes Next.js for both frontend rendering and backend API handling, enabling seamless full-stack development. Drizzle ORM is used for efficient and type-safe database interactions with the Neon PostgreSQL cloud database, ensuring secure and scalable data management. The Google Gemini API is integrated for dynamic interview question generation and semantic evaluation of user responses. Speech-to-text services are used to convert spoken responses into text for uniform NLP processing.

VII. SCORING ALGORITHM AND EVALUATION MODEL

The proposed AI-based Mock Interview Platform is implemented using modern web technologies and AI services to ensure scalability, performance, and modularity. A. Evaluation Parameters

Each response is evaluated based on the following parameters:

1. Relevance (R): Measures how closely the response aligns with the interview question using semantic similarity analysis BERT [10].
2. Keyword Match (K): Evaluates the presence of essential domain-specific keywords.
3. Clarity (C): Assesses sentence structure, coherence and communication quality.
4. Completeness (M): Determines whether the response fully addresses all aspects of the question.

Each parameter is normalized on a scale of 0 to 1.

B. Weighted Scoring Formula

The final score is computed using a weighted linear combination model:

$$\text{Score} = (R \times 0.4) + (K \times 0.2) + (C \times 0.2) + (M \times 0.2)$$

Where:

R = Relevance Score

K = Keyword Match Score

C = Clarity Score

M = Completeness Score

The final score is scaled to a 0–10 range:

$$\text{Final Score} = \text{Score} \times 10$$

C. Algorithmic Workflow

The evaluation process follows the steps below:

1. Convert speech response to text.
2. Generate reference answer or expected concept set using Gemini API.
3. Compute semantic similarity between the user response and the reference answer [11].
4. Extract domain-specific keywords and compute keyword match ratio.

5. Evaluate clarity using sentence structure and coherence analysis.
6. Calculate the weighted score using the defined formula.
7. Generate structured feedback based on score thresholds.

D. Performance Categorization

The key technologies used in the system are summarized in Table III.

TABLE III. PERFORMANCE CATEGORIZATION

Score Range	Performance Level
8-10	Excellent
6-7.9	Good
4-5.9	Average
Below 4	Needs Improvement

VIII. RESULTS AND PERFORMANCE EVALUATION

The proposed AI-based Mock Interview Platform was tested with 40 undergraduate students across three technical domains: Web Development, Data Structures and Machine Learning. Each participant completed at least 3 interview sessions.

A. User Performance Improvement

Performance improvement was measured by comparing the average scores from the first and third interview attempts.

- Average Initial Score: 5.84 / 10
- Average Final Score: 7.46 / 10
- Improvement Percentage: 27.7%

Approximately 75% of users demonstrated measurable improvement in answer structuring and conceptual clarity after repeated sessions.

B. User Satisfaction Rating

A post-session feedback survey was conducted using a 5-point rating scale.

- Overall Satisfaction Rating: 4.2 / 5
- 82% users reported increased confidence.
- 78% users found feedback actionable and clear.

C. System Performance Metrics

The system's technical performance was evaluated based on response time and processing efficiency.

- Average Question Generation Time: 1.8 seconds
- Average Speech-to-Text Processing Time: 2.3 seconds
- Average Feedback Generation Time: 2.7 seconds
- Total Average System Response Time: 6.8 seconds

The system demonstrated stable and acceptable response times under normal usage conditions. The detailed quantitative evaluation metrics of the system are summarized in Table IV.

TABLE IV. QUANTITATIVE PERFORMANCE EVALUATION

Metric	Value
Number of Users Tested	40
Average Initial Score	5.84
Average Final Score	7.46
Improvement Percentage	27.7%
User Satisfaction Rating	4.2 / 5
Average Response Time	6.8 seconds

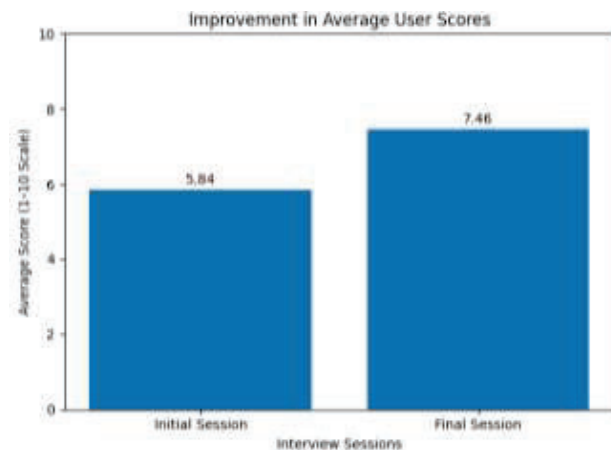


Fig. 2: Improvement in Average User Scores Across Sessions

As illustrated in Fig. 2, the average performance scores increased significantly after repeated practice sessions.

IX. LIMITATIONS AND FUTURE WORK

Despite its effectiveness, the proposed AI-based mock interview platform has certain limitations. The system primarily evaluates the textual and semantic content of responses and does not fully capture nonverbal aspects, such as body language, facial expressions, and emotional cues, that human interviewers typically assess. In addition, the platform depends on stable internet connectivity, as AI processing and speech recognition services require real-time cloud communication, and network instability may affect response time or session continuity. AI-based evaluation models may

also exhibit bias depending on training data and prompt design. Although structured scoring mechanisms are used to improve objectivity, complete elimination of algorithmic bias remains a challenge.

Future enhancements can address these limitations by incorporating advanced AI and multimodal analysis techniques. Potential improvements include integrating facial expression detection and speech emotion recognition to evaluate candidates' behavioral traits, resume-based question generation for personalized interviews, and multilingual NLP models to support interviews across multiple languages. Furthermore, integration with institutional placement management systems could enable direct performance tracking and recruiter access, while adaptive interview mechanisms based on reinforcement learning may dynamically adjust question difficulty according to candidate performance. These advancements would transform the platform into a more comprehensive and intelligent interview assessment system.

X. CONCLUSION

This paper presented the design and implementation of an AI-driven Mock Interview Platform aimed at enhancing interview readiness through automated simulation, semantic evaluation, and structured feedback generation. The system integrates generative AI for dynamic question creation, speech-to-text processing for response transcription, and Natural Language Processing techniques for objective semantic assessment. A weighted scoring mechanism ensures consistent and measurable performance evaluation.

Experimental testing with undergraduate students demonstrated measurable improvement in interview performance and positive user satisfaction ratings, validating the effectiveness of the proposed approach. The modular and scalable architecture enables cost-effective deployment in academic institutions and training environments.

The proposed system bridges the gap between theoretical learning and real-world recruitment expectations by

providing accessible, data-driven, and continuous interview practice. By leveraging artificial intelligence for automated assessment, the platform represents a significant step toward intelligent employability enhancement systems in modern education and recruitment ecosystems.

REFERENCES

- [1] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. Pearson Education, 2010.
- [2] A. Vaswani et al., "Attention Is All You Need," in *Proc. 31st Conf. Neural Information Processing Systems (NeurIPS)*, 2017, pp. 5998–6008.
- [3] D. Jurafsky and J. H. Martin, *Speech and Language Processing*, 2nd ed. Pearson Education, 2019.
- [4] Google Careers, "Interview Warmup," 2023. [Online]. Available: <https://grow.google/certificates/interview-warmup/>
- [5] M. Poria, E. Cambria, and A. Hussain, "A Review of Affective Computing: From Unimodal Analysis to Multimodal Fusion," *Information Fusion*, vol. 37, pp. 98–125, 2017.
- [6] S. Latif, R. Rana, S. Younis, J. Qadir, and J. Epps, "Direct Modeling of Speech Emotion from Raw Speech Using Deep Learning: A Survey," *IEEE Access*, vol. 7, pp. 130–145, 2019.
- [7] P. Pecuchova, L. Benko & M. Drlik, "Automated Grading of Open-Ended Questions in Higher Education Using GenAI Models," *International Journal of Artificial Intelligence in Education*, vol. 35, pp. 3813–3846, 2025.
- [8] J. Atkinson & D. Palma, "An LLM-Based Hybrid Approach for Enhanced Automated Essay Scoring," *Scientific Reports*, vol. 15, art. 14551, 2025.
- [9] Y. Tay, M. Dehghani, and D. Metzler, "ColBERT: Efficient and Effective Passage Search via Contextualized Late Interaction over BERT," in *Proc. ACM SIGIR Conference on Research and Development in Information Retrieval*, 2020.
- [10] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proc. NAACL-HLT*, 2019, pp. 4171–4186.
- [11] .Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," in *Proc. EMNLP*, 2019.
- [12] M. Mohler, R. Bunescu, and R. Mihalcea, "Learning to Grade Short Answer Questions using Semantic Similarity Measures and Dependency Graph Alignments," in *Proc. ACL*, 2011.