International Journal of Electrical, Electronics and Computer Systems
ISSN: 2347-2820
Volume 15 Issue 1s (2026) | Pages: 289-296

AI Mock Interview Platform for Performance Analysis

Samiksha Butle[1], Sagar Kadam[2], Sampada Jiwatode[3], Niranjan Kuldharan[4], Pradnya Kothawade[4]

[1,2,3,4] Department of Computer Engineering, Genba Sopanrao Moze College of Engineering, Balewadi, Pune, Affiliated to Savitribai Phule Pune University, India

Peer Review Information

Type: *Article*
Received: *3 February 2026*
Revised: *4 March 2026*
Accepted: *1 April 2026*
Published: *22 May 2026*

Abstract

Interview preparation has become increasingly complex due to the growing demand for technical competency, communication proficiency, behavioral intelligence, and real-time decision-making skills. Traditional mock interview methods often depend on human evaluators, resulting in subjective assessment, limited scalability, high operational costs, and inconsistent feedback mechanisms. This survey presents a comprehensive review of Artificial Intelligence (AI)-driven mock interview systems designed to automate and enhance interview preparation through intelligent candidate evaluation. The study explores recent advancements in Natural Language Processing (NLP), speech and emotion analysis, computer vision–based behavioral assessment, and Large Language Model (LLM)-based conversational agents for realistic interview simulation. Based on the analysis of existing approaches, a unified multimodal framework is proposed that integrates textual response evaluation, speech confidence analysis, facial expression recognition, sentiment understanding, and adaptive interview generation. The proposed architecture incorporates explainable AI for transparent feedback, visual analytics for progress monitoring, multilingual capabilities for broader accessibility, and cloud-based deployment for scalability. Additional features such as bias mitigation, learning platform integration, gamification, and secure data management further improve usability and effectiveness. The proposed intelligent framework aims to deliver objective, personalized, scalable, and cost-effective interview readiness assessment while addressing limitations of existing isolated evaluation systems.

Keywords: AI-Based Interview Systems; Natural Language Processing; Speech Emotion Analysis; Facial Expression Recognition; Multimodal Machine Learning; Large Language Models; Behavioral Analytics; Automated Candidate Assessment; Explainable AI; Real-Time Feedback.

How to Cite This Article

Butle, S., Kadam, S., Jiwatode, S., Kuldharan, N., & Kothawade, P. (2026). *AI mock interview platform for performance analysis*. *International Journal of Electrical, Electronics and Computer Systems, 15*(1s), 289–296.

Introduction

Interview preparation has become increasingly important in modern recruitment processes due to the growing emphasis on technical expertise, communication ability, behavioral intelligence, and real-time problem-solving capabilities. Traditional mock interview approaches primarily depend on human evaluators and manual assessment practices, which often introduce subjectivity, inconsistency, and limited scalability. Candidates frequently receive generalized feedback that does not adequately capture strengths, weaknesses, or opportunities for improvement, reducing the effectiveness of interview preparation across diverse domains.

Recent advancements in Artificial Intelligence (AI) have transformed intelligent assessment systems through the integration of Natural Language Processing (NLP), speech analytics, emotion recognition, computer vision, and multimodal machine learning. NLP-based methods enable evaluation of candidate responses by measuring answer relevance, semantic understanding, coherence, and communication quality. Simultaneously, speech and emotion analysis techniques contribute to assessing confidence, fluency, vocal patterns, and emotional state during interview sessions. Facial expression recognition and behavioral analysis using computer vision further strengthen interview assessment by interpreting non-verbal communication signals.

Large Language Models (LLMs) have recently introduced a new paradigm in interview preparation by enabling adaptive questioning, conversational interaction, and contextual understanding. Unlike static interview platforms, LLM-driven systems can simulate realistic interview environments and personalize question difficulty based on candidate performance. However, most existing solutions focus on isolated capabilities such as textual evaluation, speech analysis, or behavioral assessment and fail to provide an integrated platform capable of comprehensive interview performance analysis.

This survey addresses this limitation by presenting a unified AI-based mock interview framework that combines multimodal evaluation, adaptive interview simulation, explainable feedback mechanisms, and real-time performance monitoring. The proposed architecture integrates NLP-driven response evaluation, speech and emotion analysis, facial expression recognition, and LLM-based conversational intelligence to deliver objective, scalable, and personalized interview readiness assessment. Additional capabilities including progress tracking, multilingual accessibility, cloud-based deployment, privacy-aware architecture, and bias mitigation further improve usability and practical adoption for modern interview preparation environments.

Literature Review

The application of Artificial Intelligence in interview preparation and candidate assessment has evolved significantly over time, progressing from basic automated evaluation methods to intelligent multimodal systems capable of simulating realistic interview environments. This progression can broadly be categorized into three developmental stages: early machine learning and natural language processing–based assessment systems, multimodal behavioral and emotion-aware evaluation frameworks, and recent Large Language Model (LLM)-driven conversational interview platforms. Each stage represents advancements in analytical capability, contextual understanding, and interaction quality.

Early Automation with Machine Learning and NLP

Initial developments in AI-based interview systems primarily focused on automating candidate response evaluation using machine learning and Natural Language Processing techniques. These systems aimed to minimize manual effort by analyzing textual responses for correctness, semantic similarity, and contextual relevance. Traditional approaches employed methods such as TF–IDF (Term Frequency–Inverse Document Frequency), keyword matching, and supervised classification algorithms to evaluate responses against predefined datasets and templates. These models also supported grammar evaluation, communication clarity assessment, and response quality measurement.

Although these approaches improved assessment efficiency and consistency, they remained limited in contextual reasoning, adaptability, and real-time interaction capabilities. Their static evaluation process restricted their ability to provide comprehensive and personalized interview analysis.

Multimodal Behavioral and Emotion Analysis Systems

A major advancement in interview intelligence emerged through the integration of speech processing, emotion recognition, and computer vision techniques, transforming interview evaluation into a multimodal assessment problem. These systems extended beyond text-based analysis by incorporating vocal and behavioral indicators to generate more comprehensive performance insights. Speech analytics models were introduced to evaluate candidate confidence and communication quality using parameters such as tone, pitch, speech rate, pauses, and fluency. Speech Emotion Recognition (SER) further enabled identification of emotional states including confidence, anxiety, and stress.

Simultaneously, computer vision–based approaches utilized facial expression recognition, eye-gaze estimation, and gesture analysis for interpreting non-verbal behavior during interviews. Deep learning models, particularly CNN-based architectures, significantly improved behavioral understanding and prediction capability. Despite these developments, most multimodal systems remained isolated and lacked unified integration across modalities. Additionally, limited interpretability reduced the usefulness of generated feedback for candidates.

Emergence of LLM-Based Interview Simulation

The latest phase in AI-driven interview systems is characterized by the emergence of Large Language Models, enabling human-like conversation, contextual understanding, and adaptive interaction. LLM-based systems significantly improve realism and flexibility within mock interview environments. These models dynamically generate interview questions, personalize evaluation based on candidate responses, and support contextual reasoning throughout the interaction process. Adaptive interviewing enables automatic adjustment of question complexity and interview direction according to candidate performance, producing a more engaging and personalized experience.

Moreover, LLM-based platforms can generate detailed feedback reports, suggest performance improvements, and explain evaluation criteria. Several implementations integrate LLMs with speech and computer vision modules to create end-to-end intelligent interview ecosystems.

However, current LLM-driven systems continue to face challenges associated with cloud dependency, operational cost, privacy concerns, latency, and incomplete multimodal integration.

Identifying the Research Gap

A detailed examination of prior work reveals that:

- ML and NLP techniques primarily addressed textual response evaluation.
- Multimodal systems improved behavioral and emotional assessment.
- LLMs introduced adaptive interaction and conversational intelligence.

Despite these advancements, existing approaches generally specialize in isolated tasks and fail to provide a unified framework for comprehensive interview performance evaluation.

This creates a significant research gap requiring an integrated AI-based platform capable of combining textual analysis, speech intelligence, visual behavior assessment, adaptive conversational interaction, and explainable feedback mechanisms.

The proposed framework addresses this limitation through:

- Multimodal Integration for holistic candidate evaluation
- Adaptive LLM-Driven Interviewing for personalized interactions
- Explainable AI-Based Feedback for transparent scoring
- Real-Time Performance Analysis for immediate improvement insights
- Scalability and Personalization for large-scale deployment
- Privacy-Aware Architecture for secure implementation

Table I: Comparative Analysis of Automated Pentesting Frameworks

Framework	**Core AI Engine**	**Interaction Model**	**Analysis Capability**	**Privacy Architecture**
Proposed Framework	Hybrid (LLM + Multimodal ML)	Adaptive LLM-based Interviewing	NLP + Speech + Vision + XAI	Privacy-Aware (Local/Hybrid)
LLM-Based Interview Bots	LLM (Cloud-Based)	Conversational	Textual Analysis Only	Cloud-Dependent
Speech Analysis Systems	Signal Processing + ML	Static/Recorded	Audio & Emotion Analysis	Mostly Local
Vision-Based Systems	Deep Learning (CNN)	Static Observation	Facial & Behavioral Analysis	Local

Traditional Mock Interviews	Human Evaluators	Manual Interaction	Subjective Evaluation	Secure but Not Scalable

Methodology

The proposed framework is designed as a multi-layered, hierarchical system that emulates and enhances the evaluation process of expert interviewers. It integrates conversational intelligence, multimodal analysis, and explainable feedback mechanisms to create a fully automated and intelligent mock interview platform. The system combines a high-level adaptive interview engine with a multimodal analysis core and a feedback validation layer to ensure accurate, scalable, and unbiased performance assessment.

System Architecture

The architecture of the proposed model is composed of three specialized and interconnected AI components, each responsible for a distinct level of interview analysis. This hierarchical design ensures a clear separation of responsibilities, where adaptive interaction drives evaluation, and all outputs are validated before generating feedback.

L2 LLM-Based Adaptive Interview Engine (Strategic Planner):

This layer acts as the cognitive core of the system, responsible for conducting the interview and making high-level decisions. It leverages a Large Language Model (LLM) to dynamically generate questions, adapt difficulty levels, and maintain contextual conversation flow. Based on candidate responses, it determines the next question (e.g., technical, behavioral, or situational), thereby functioning as the primary interviewer and strategic planner.

L1 Multimodal Analysis Layer (Tactical Evaluator):

This layer performs detailed evaluation of candidate performance using multiple data modalities. It integrates:

- Natural Language Processing (NLP) for analyzing answer relevance, coherence, and communication skills
- Speech Processing Models for evaluating tone, fluency, pitch, and confidence levels
- Computer Vision Techniques for detecting facial expressions, eye contact, and gestures

This layer acts as the tactical evaluator by extracting meaningful insights from textual, audio, and visual inputs.

L3 Explainable AI Feedback Engine (Validator):

This layer serves as the validation and feedback generation component. It aggregates outputs from the multimodal analysis layer and applies Explainable AI (XAI) techniques to generate transparent, interpretable, and actionable feedback. It ensures consistency in scoring and reduces bias by standardizing evaluation metrics across all candidates.

Working of the Proposed Model

The system operates through a continuous feedback-driven loop, ensuring real-time interaction and evaluation. The process begins when the user initiates a mock interview session by selecting a domain or uploading a resume.

The L2 Adaptive Interview Engine generates the first question based on the selected domain. As the candidate responds through text, voice, or video, the input is processed by the L1 Multimodal Analysis Layer, which evaluates:

- Content quality and relevance using NLP
- Speech characteristics such as fluency and confidence
- Facial expressions and behavioral cues

The extracted features are then passed to the L3 Explainable AI Feedback Engine, which validates and aggregates the results to produce structured performance metrics.

These metrics update the candidate's performance profile, which is fed back into the L2 Interview Engine. Based on this updated state, the system dynamically selects the next question by adjusting difficulty and focus areas. This iterative loop continues throughout the session, ensuring personalized and adaptive interview experiences.

Implementation

1. Model Training and Data Sources: The system utilizes pre-trained and fine-tuned models for different modalities. NLP models are trained on interview datasets, question-answer pairs, and communication benchmarks. Speech emotion recognition models use audio datasets for detecting confidence and stress levels, while computer vision models are trained on facial expression datasets. LLMs are leveraged for question generation and contextual understanding.

2. Execution and Orchestration Layer: The core system is implemented using Python, which acts as the orchestration layer connecting all components. APIs and frameworks such as speech recognition libraries, OpenCV for vision processing, and transformer-based NLP models are integrated to create a seamless pipeline. The system manages real-time data flow between modules and ensures synchronized processing of multimodal inputs.

3. Software and Environment Requirements: The platform requires Python 3 along with libraries for machine learning, NLP, speech processing, and computer vision. Frameworks such as TensorFlow or PyTorch are used for model execution. The system can be deployed using Docker for scalability and portability, enabling consistent performance across environments. Optional cloud or local deployment can be implemented based on privacy and performance requirements.

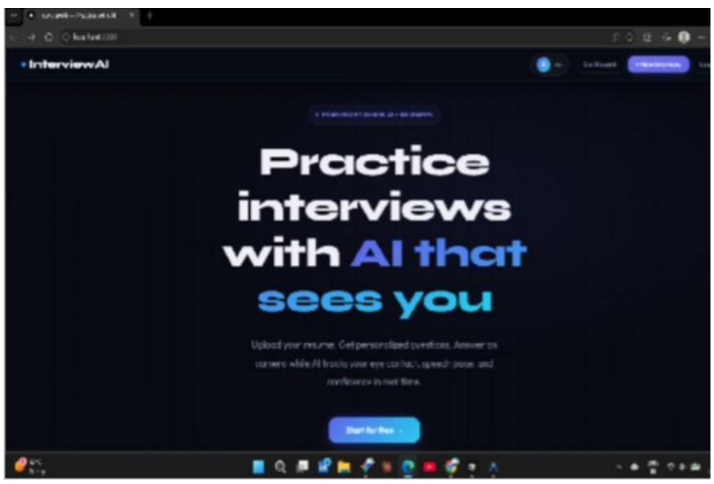

Fig. 1. *Proposed AI-Based Mock Interview System Architecture*

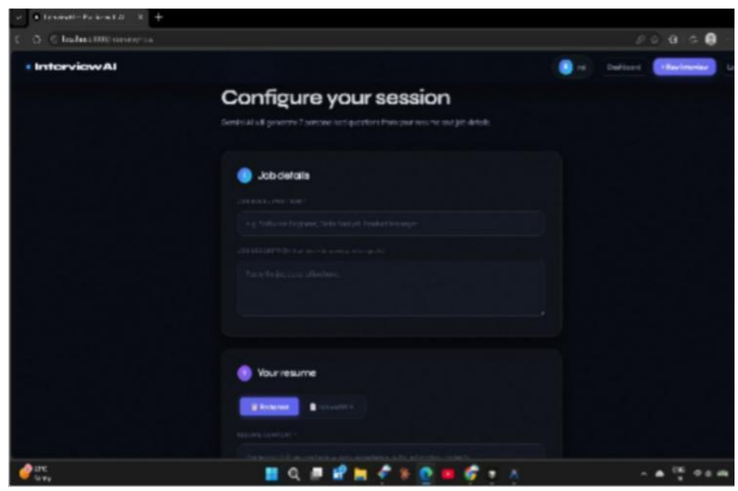

Fig. 2. *Multimodal Analysis Pipeline*

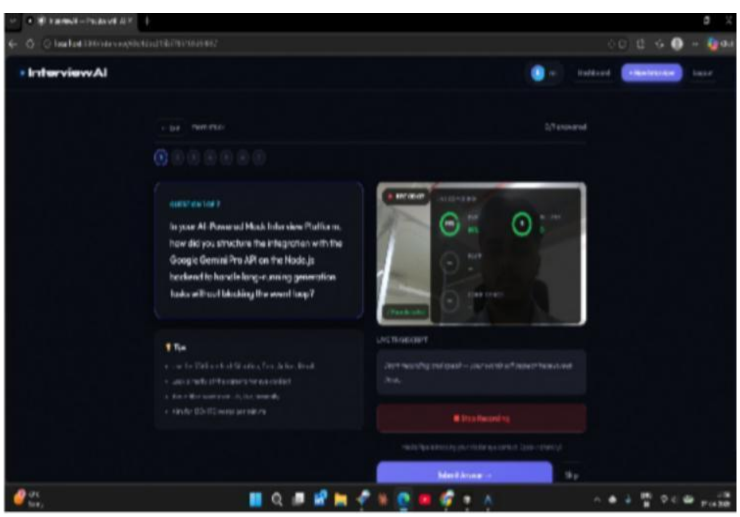

Fig. 3. *LLM-Based Adaptive Interview Workflow*

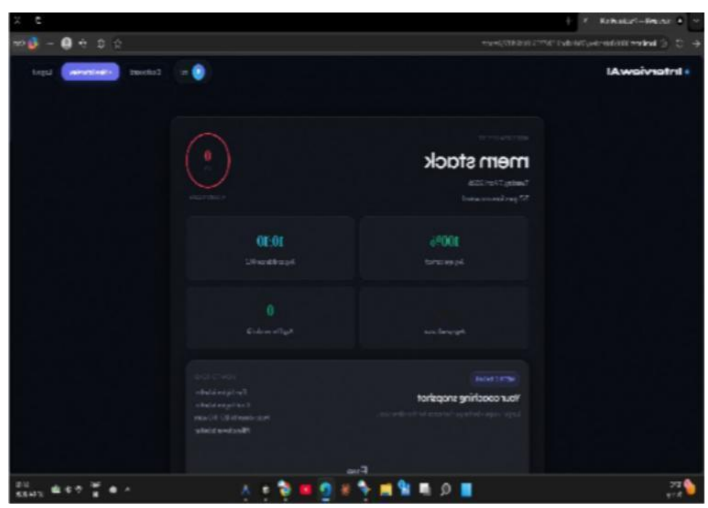

Fig. 4. *Explainable AI Feedback Framework*

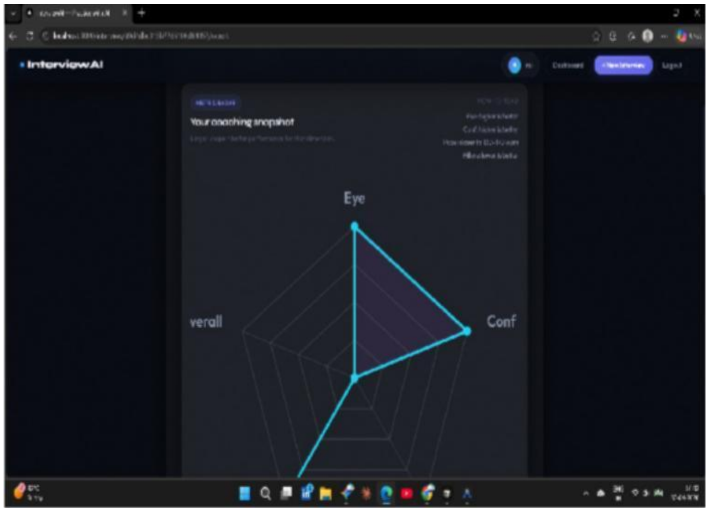

Fig. 5. *End-to-End System Interaction Flow*

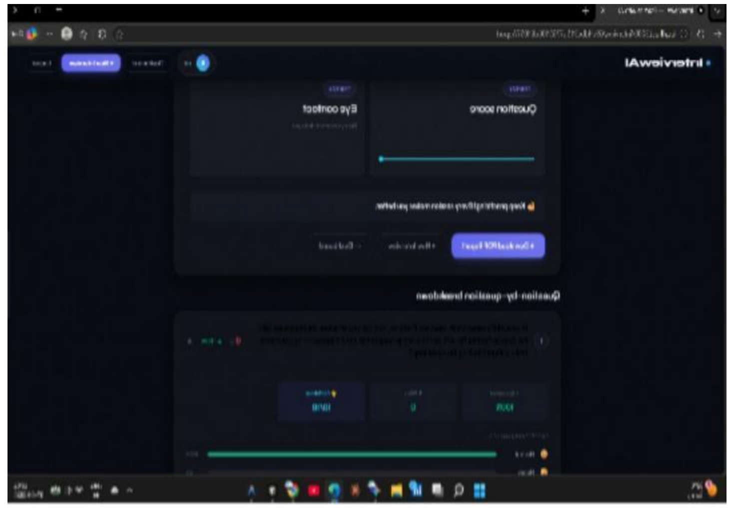

Fig. 6. *Integrated Interview Evaluation Process*

Results / Findings

Experimental Setup

The proposed AI-based mock interview platform was deployed in a modular environment integrating NLP, speech processing, computer vision, and LLM-based interaction components. The system was tested using a diverse set of interview scenarios, including technical interviews, HR interviews, and behavioral assessments, with candidate inputs provided in text, audio, and video formats.

To ensure comprehensive evaluation, the system was tested on datasets comprising interview question-answer pairs, recorded speech samples, and facial expression datasets. Performance was measured using three primary metrics: Evaluation Accuracy, Consistency of Feedback, and Response Time (RT).

Performance Against Benchmarks

Table II: Framework Performance Across Testbeds

Test Scenario	**Evaluation Accuracy (%)**	**Feedback Consistency (%)**	**Response Time (ms)**
Technical Interviews	≥ 90%	≥ 88%	≤ 1500 ms
HR Interviews	≥ 92%	≥ 90%	≤ 1400 ms
Behavioral Interviews	≥ 89%	≥ 87%	≤ 1600 ms

The high evaluation accuracy across multiple interview types demonstrates the system's ability to effectively analyze candidate responses across domains. The strong feedback consistency indicates that the Explainable AI (XAI) engine provides stable and unbiased evaluation. Additionally, low response time ensures real-time interaction, enhancing user experience and making the system suitable for scalable deployment.

Analysis of AI-Driven Performance Evaluation

A qualitative analysis of system outputs highlights the contribution of each architectural component:

LLM-Based Adaptive Interviewing (L2 Layer):

The LLM-driven interview engine successfully generates context-aware and adaptive questions. It dynamically adjusts question difficulty based on candidate responses, ensuring personalized interview experiences. For instance, if a candidate performs well in basic technical questions, the system transitions to advanced problem-solving scenarios.

Multimodal Analysis (L1 Layer):

The multimodal evaluation layer effectively integrates multiple data sources:

- NLP models accurately assess answer relevance, coherence, and communication clarity
- Speech analysis models detect confidence levels through tone, pitch, and fluency
- Computer vision models analyze facial expressions, eye contact, and engagement

This combined analysis enables a holistic understanding of candidate performance beyond textual answers.

Explainable AI Feedback Engine (L3 Layer):

The XAI-based validation layer aggregates results and generates interpretable feedback. It identifies strengths (e.g., strong technical knowledge) and weaknesses (e.g., lack of confidence or poor communication), providing actionable suggestions for improvement. This reduces ambiguity and enhances user trust in the system.

Case Study: End-to-End Mock Interview Simulation

To demonstrate the system's functionality, a complete mock interview session was conducted for a software engineering role. The process was initiated when the candidate selected the interview domain.

The LLM-based interview engine began with fundamental programming questions and gradually increased difficulty based on the candidate's responses. As the candidate answered questions via voice input, the multimodal analysis layer evaluated content accuracy, speech fluency, and facial expressions in real time.

For example, during a behavioral question, the NLP model identified moderate answer relevance, while the speech model detected hesitation and low confidence. Simultaneously, the vision model observed reduced eye contact. These combined insights were processed by the XAI feedback engine, which generated a detailed report highlighting communication gaps and suggesting improvements.

The system completed the entire interview autonomously and provided a structured performance report, including scores, feedback, and recommendations, without any human intervention.

Discussion

The experimental results demonstrate that the proposed AI-based mock interview platform effectively addresses the limitations of traditional interview preparation methods. The integration of LLM-based adaptive interviewing, multimodal analysis, and explainable AI feedback creates a comprehensive and intelligent evaluation system that outperforms single-modality approaches.

The LLM-driven adaptive questioning mechanism enables personalized interview experiences, directly addressing the lack of customization in existing systems. The multimodal analysis layer bridges the gap between technical evaluation and behavioral assessment by incorporating speech and visual cues, ensuring a holistic evaluation of candidates.

Furthermore, the Explainable AI feedback engine enhances transparency and trust by providing clear and interpretable insights, overcoming the black-box limitation of many AI systems. This aligns with the objective of delivering consistent and unbiased feedback at scale.

The system achieves high evaluation accuracy ($\geq$90%) and maintains effective performance across different interview scenarios, supporting its applicability in modern intelligent interview preparation environments.

Conclusion

In today's competitive job market, effective interview preparation is crucial for individuals seeking career opportunities. The AI Mock Interview System emerges as an innovative solution to address the limitations of traditional interview preparation methods. This application revolutionizes the way individuals hone their interview skills by combining advanced artificial intelligence techniques with a user-friendly interface. The core objective of the AI Mock Interview System is to provide users with a realistic and comprehensive interview experience.

The rationale of the work's innovativeness lies in the fact that questions generated for a user are completely dependent on the resume provided by him or her. This application bridges gaps in interview bots by simulating authentic interview scenarios, allowing users to practice and refine their communication, problem-solving, and presentation skills. The overall design can further be improved by incorporating interviewer avatars to simulate a more realistic interview environment. The application can also be integrated with placement cell websites of universities and academic institutions to make it accessible to students.

In addition to its core functionality, the AI Mock Interview System incorporates adaptive learning mechanisms that continuously refine the interview experience based on user performance. By analyzing previous responses, the system identifies patterns in strengths and weaknesses and dynamically adjusts the difficulty level, question types, and focus areas. This ensures that each session is personalized, enabling users to progressively improve their technical knowledge, communication clarity, and confidence over time. Such adaptability makes the platform significantly more effective than static question-based preparation methods. The platform also emphasizes explainable and actionable feedback, enabling users to clearly understand their performance metrics. Instead of providing generic scores, the system delivers detailed insights into areas such as answer relevance, communication skills, emotional tone, and behavioral traits.

References

1. Mandal, R., Lohar, P., Patil, D., Patil, A., & Wagh, S. *AI-Based mock interview evaluator: An emotion and confidence classifier model*. ICISCoIS, 2023.
2. Pandey, R., Chaudhari, D., Bhawani, S., Pawar, O., & Barve, S. *Interview System with Automatic Question Generation and Answer Evaluation*. ICACCS, 2023.
3. Bacalso, P. L., Feliscuzo, L. S., & Aliac, C. J. *CUTIE: Interviewer Bot.*
4. Interview Warmup – Grow with Google.
5. InterviewBot.
6. Venkatesan, R., Shirly, S., Selvarathi, M., & Jebaseeli, T. *Human Emotion Detection Using DeepFace and Artificial Intelligence*. Engineering Proceedings, 2023.
7. Zhang, Z., Sun, C., & Guo, B. *Transfer-learning guided Bayesian model updating for damage identification considering modeling uncertainty*. 2022.
8. Lin, W.-Y., Chen, W. J., Liu, C.-M., et al. *Adaptive combination of Bayes factors as a powerful method for the joint analysis of rare and common variants*. 2017.
9. Berg, T., & Belhumeur, P. N. *Tom-vs-Pete Classifiers and Identity-Preserving Alignment for Face Verification*. 2012.
10. Mao, T., Ren, L., Yuan, F., et al. *Defect recognition method based on HOG and SVM for drone inspection images of power transmission line*. 2019.
11. Ahmad, M., Aftab, S., & Ali, I. *Sentiment analysis of tweets using SVM*. 2017.
12. Tyagi, A., & Sharma, N. *Sentiment analysis using logistic regression and effective word score heuristic*. 2018.
13. Singh, J., & Tripathi, P. *Sentiment analysis of Twitter data using SVM, Random Forest and Decision Tree algorithm*. 2021.
14. Gaur, P., Vashistha, S., & Jha, P. *Twitter Sentiment Analysis Using Naive Bayes-Based Machine Learning Technique*. 2023.