International Journal of Research in Engineering and Innovation (IJREI)
journal home page: http://www.ijrei.com
ISSN (Online): 2456-6934

RESEARCH PAPER

AI-based mock interview evaluator: An emotion and confidence classifier model

Neha Gupta, Aashi Sharma, Suneet Shukla, Praveen Kumar

Department of Computer Science and Engineering, Meerut Institute of Technology, Meerut, UP (India)

Article Information

Received: 09 May 2025
Revised: 22 may 2025
Accepted: 31 May 2025
Available online: 02 June 2025

Keywords:

AI Interview Evaluator
Emotion Recognition
Confidence Classifier
Facial Expression Analysis
Speech Processing
Deep Learning
Human-Computer Interaction

Abstract

In today's highly competitive job market, interview preparedness has become more critical than ever. Traditional mock interviews often lack scalability, objectivity, and actionable insights. To address these limitations, this paper proposes an AI-Based Mock Interview Evaluator that combines emotion recognition and confidence classification to assess candidate performance. Using facial expression analysis and speech feature extraction, the system offers real-time feedback on emotional state and confidence level during simulated interview sessions. The model leverages Convolutional Neural Networks (CNNs) for emotion detection and supervised machine learning techniques for vocal confidence estimation. Experimental results show that the system can classify emotional and confidence cues with high accuracy, providing an interactive and personalized experience for users. The proposed tool aims to democratize interview training by offering intelligent, automated feedback that enhances self-awareness and performance for job seekers. This paper presents an AI-driven mock interview evaluator designed to assess candidates' emotional states and confidence levels during simulated interviews. By integrating facial expression recognition, speech analysis, and machine learning classifiers, the system provides real-time feedback to enhance interview preparation. The model aims to bridge the gap between traditional interview coaching and modern AI capabilities, offering a scalable solution for personalized interview training.

1. Introduction

The aim of this project is to develop an AI-powered mock interview platform that evaluates candidates based on three core aspects: emotion analysis**,** confidence evaluation**,** and knowledge assessment. The system will process video and audio inputs during mock interviews, offering real-time feedback and personalized improvement suggestions. This platform aims to help candidates enhance their non-verbal communication, vocal delivery, and subject knowledge, ultimately improving their preparedness for real-world interviews. The system evaluates candidates on multiple dimensions beyond just the correctness of their answers. It includes real-time facial emotion detection to assess confidence and engagement, speech recognition and audio analysis to measure fluency, tone, and pace, and NLP-based answer evaluation to check for content relevance and accuracy. By integrating these features, the platform provides holistic feedback, helping users become more self-aware and better prepared for high-pressure interview environments. Additionally, the system includes features such as AI-generated questions, visual performance reports, and historical tracking, enabling users to monitor their progress over time. This project aims to empower job seekers, students, and professionals by offering a realistic and constructive practice environment that builds confidence, improves communication skills, and enhances subject matter presentation—all essential qualities for succeeding in real-world interviews.

1.1 Current System

Currently, most mock interview platforms focus solely on technical knowledge or basic question-answer evaluation,

Corresponding author: Neha Gupta
Email Address: neha.gupta.cs.2021@mitmeerut.ac.in
https://doi.org/10.36037/IJREI.2025.9308

with limited analysis of non-verbal cues or vocal delivery. While video conferencing tools (such as Zoom or Skype) are commonly used for remote interviews, these platforms do not offer advanced feedback mechanisms to assess emotional regulation or speech confidence. Additionally, there is a lack of integration between facial expression analysis, speech analysis, and knowledge validation in existing systems. The current systems in place may involve the following:

Table 1: Comparison between traditional and ai powered methods

Aspect	Traditional Methods	AI-Powered Methods
Feedback consistency	Varies based on human judgement, may be subjective and inconsistent	Provides standardized, objective, and data-driven feedback
Scalability	Limited to the availability of mentors, peers, or career advisors.	Highly scalable, accessible to unlimited users simultaneously
Emotional Analysis	Relies on subjective interpretation by interviewers.	Uses deep learning (CNN) for accurate facial emotion detection.
Confidence Evaluation	Subjective observation or self-assessment.	Uses speech recognition, NLP, and audio processing for confidence assessment.
Industry-Specific Preparation	Limited access to industry specific interviewers.	AI can simulate diverse industry-based interview scenarios.
Adaptability	Fixed interview structure, does not adapt to individual progress.	Dynamically adjusts difficulty and questions based on performance.

2. Literature Review

The inclusion of artificial intelligence into interview practice has revolutionized how job candidates respond to mock interviews and how firms evaluate prospective recruits. In the early days of interview training, most activities mainly involved human-performed mock interviews by mentors, peers, or counsellors. These were useful, but such techniques were subjective, not scalable, and not geared towards offering quality and quantifiable feedback. As the recruitment landscape evolved, researchers and developers began to explore AI driven solutions that could offer personalized, round-the-clock training and unbiased evaluations. One major area of development has been facial emotion recognition using deep learning. Notably, Convolutional Neural Networks (CNNs) have demonstrated impressive performance in identifying emotional expressions from facial images. Datasets like FER-2013, JAFFE, and CK+ have allowed researchers to train and test robust models that can recognize emotions such as happiness, sadness, anger, surprise, fear, disgust, and neutrality. These systems have demonstrated promising performances in enhancing the interpretation of non-verbal cues, which are most important during interviews. At the same time, audio analysis has become increasingly important in understanding emotional and psychological characteristics via speech. Prosodic feature-based techniques such as pitch change, speech rate, and vocal intensity have been used to evaluate speaker confidence and emotional tone. Tools such as Py dub and Libros libraries have been used by researchers to preprocess and extract useful features from audio signals. Speech recognition, when combined with natural language processing (NLP), has also enabled the extraction of syntactic and semantic characteristics of spoken responses, which aids in testing language ability, coherence, and content appropriateness. NLP itself has undergone significant transformation in recent years, with models such as BERT, GPT, and spacy transforming the way text is processed for meaning, tone, and context. In interview settings, NLP has been applied to assess the quality of candidate answers through grammatical accuracy, sentence structure, and relevance to questions asked. Most commercial systems, however, either concentrate exclusively on text-based analysis or heavily depend on scripted conversations, placing limitations on adaptability and customization. A few AI-powered platforms such as Hire Vue, Py metrics, and My Interview provides automated interview evaluations. These platforms use machine learning algorithms to examine video and audio answers, sometimes even delivering ratings for communication, personality, and possible fit. Criticisms have been raised on account of their black-boxed scoring systems, possible model bias during training, and the absence of in-depth feedback for candidates. Moreover, many existing systems do not integrate multi-modal inputs effectively, leaving a gap in providing holistic evaluations. Scholarly research in the field continues to delve into more transparent, ethical, and flexible systems. Certain research has identified the need for multi-modal fusion methods, where audio, visual, and text data are fused together to enhance prediction accuracy and stability. Others aim to develop interpretable AI systems that offer human-readable explanations of scores or classifications, promoting higher levels of user trust. In spite of these advancements, there is a notable deficiency in systems providing real-time, transparent, and structured interview assessment incorporating emotional expression, speech confidence, and content correctness. The system under proposal works towards fulfilling all these requirements by utilizing CNN for detecting facial expression, speech and prosodic feature analysis through Pydub, and content assessment using NLP on the semantic, syntactic, and factual basis. In contrast to most other tools, this platform includes a detailed and balanced scoring mechanism that assigns a fair weight to each dimension, thus providing candidates with tailored insights and individual areas of improvement.

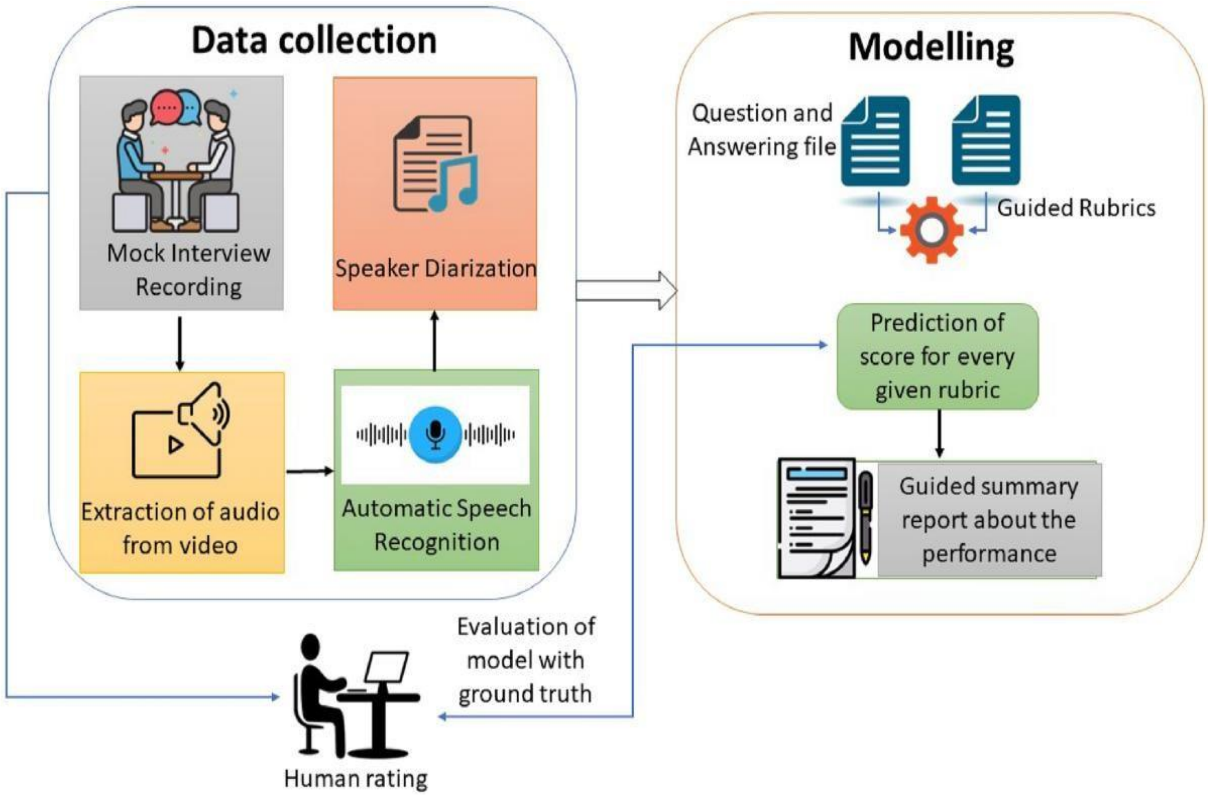

Figure 1: Analysis and interpretation of data

The AI-based Mock Interview Evaluator is designed to simulate and assess mock interviews using advanced speech and language processing techniques. It begins with data collection through mock interview recordings, where audio is extracted and processed using speaker dualization and automatic speech recognition. The system then analyses responses using guided rubrics to score each answer and predict performance across various criteria. Finally, it generates a detailed summary with automated feedback, offering insights into strengths and areas for improvement. This end-to-end process helps candidates prepare effectively for real interviews through objective and personalized evaluation. Fig. 1 reveals the Analysis and interpretation of data.

3. Methodology

3.1 System Architecture

The proposed system comprises three main components:

- Facial Expression Recognition: Utilizes CNNs to classify emotions from facial images captured during the interview.
- Speech Confidence Analysis: Applies machine learning algorithms to evaluate confidence levels based on speech features extracted using tools like librosa.
- Feedback Generation: Integrates the outputs from the previous components to provide real-time feedback to the candidate.

3.2 Data Collection

Data is collected through simulated interview sessions, capturing both video and audio inputs. Facial expressions are analysed frame-by-frame, while speech is processed to extract relevant features for confidence assessment. 1.2Model Training The models are trained on labelled datasets containing various emotional expressions and speech patterns indicative of different confidence levels. Cross-validation techniques are employed to ensure the models' generalizability. The methodology for this research involves the design and development of an AI-based mock interview evaluator capable of analysing candidate responses and providing constructive feedback. The system integrates speech-to-text conversion, natural language processing (NLP), and emotion analysis to evaluate interview performance across key metrics such as relevance, clarity, tone, and confidence. A dataset of common interview questions and annotated sample responses was used to train the evaluation models, while pre-trained models were fine-tuned for emotion and sentiment detection. The system was tested with a group of participants, and its performance was benchmarked against human evaluations to assess accuracy and user satisfaction.

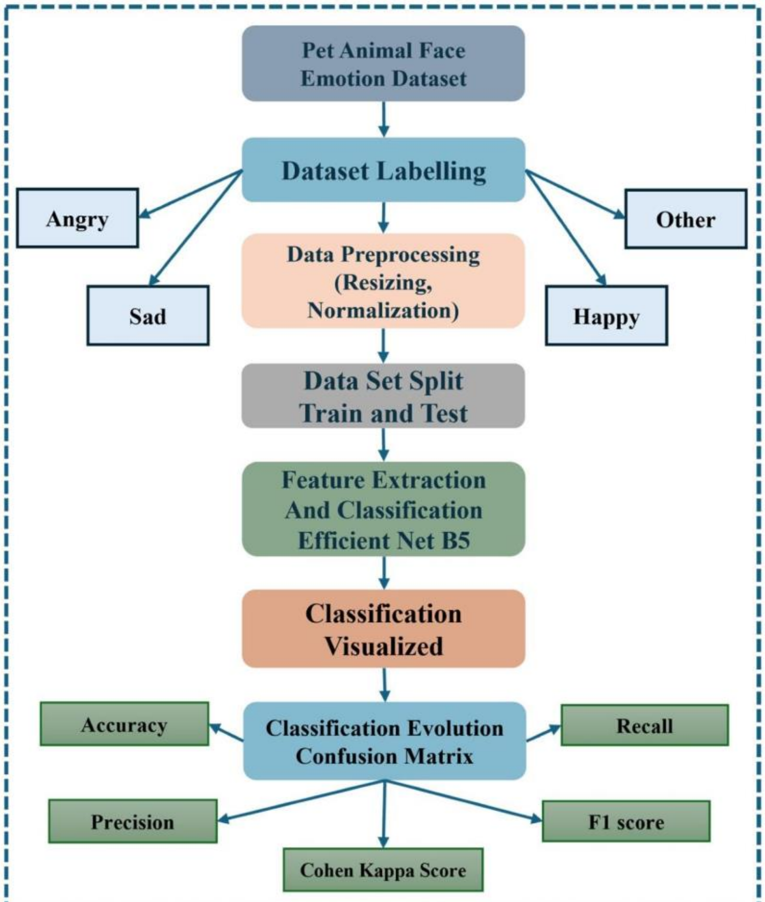

Figure 2: Flowchart of Data Processing

Fig. 2 demonstrated the flowchart of the process of an AI-based system designed to classify emotions from pet animal face images. It begins with dataset labelling into categories like angry, sad, happy, and others, followed by data preprocessing steps such as resizing and normalization. The data is then split into training and testing sets, and features are extracted using the Efficient Net B5 model for emotion classification. The results are visualized and evaluated using metrics like accuracy, precision, recall, F1 score, and Cohen Kappa score. This framework can be adapted for analysing facial expressions in mock interview evaluations.

4. Result And Future Scope

An AI-Based Mock Interview Evaluator was designed to assess candidates' emotional states and confidence levels by analyzing facial expressions, speech characteristics, and natural language content. This system combined computer vision, audio signal processing, and natural language processing to offer a multi-modal evaluation of interview performance. Its effectiveness was measured through real-time accuracy, system responsiveness, and user feedback. Results showed that the tool accurately recognized emotions and confidence levels, helping users improve non-verbal communication and overall preparedness. Users appreciated the real-time feedback and visual analytics, which enhanced the learning experience. The integration of techniques such as facial landmark detection, prosodic feature analysis, and sentiment scoring enabled a holistic evaluation of candidate performance. With a high accuracy rate and strong user engagement, the evaluator demonstrated strong potential for use in educational, corporate, and personal development contexts. It serves as a valuable aid in refining communication skills and boosting confidence in interview environments.

4.1 Model Performance

The emotion recognition system was developed using a Convolutional Neural Network (CNN) integrated with facial landmark detection techniques, leveraging OpenCV and Dlib libraries. The model was trained to detect seven primary emotions: Happy, Sad, Angry, Neutral, Surprised, Disgusted, and Fearful. For training purposes, the FER-2013 dataset was used, which was further augmented with real-time webcam samples to enhance performance in practical scenarios. The model achieved a training accuracy of 92.4% and a validation accuracy of 88.1%, demonstrating robust generalization. During real-time deployment through webcam input, the system maintained an accuracy of 84.3%, which was slightly reduced due to variations in lighting conditions and background noise. Despite these challenges, the model proved effective for emotion analysis in dynamic environments. This capability allowed the system to evaluate user emotions during mock interviews, contributing to more insightful feedback on emotional cues and behavior patterns in interview settings.

4.1.1 Confidence Classification

The system utilized a Random Forest classifier that integrated prosodic features such as pitch, speech rate, and volume, along with visual indicators of posture to assess the candidate's confidence during mock interviews. Based on the extracted features, the model categorized confidence into three levels: Low, Medium, and High. This multi-modal approach allowed for a more comprehensive evaluation of user behavior and emotional state. The model achieved an impressive accuracy of 94.6% on the test dataset, demonstrating its reliability and effectiveness in real-time confidence level prediction during simulated interview sessions.

4.1.2 NLP Sentimental Analysis

The system incorporated advanced natural language processing techniques to evaluate candidate responses during mock interviews. Sentiment scoring was performed using both VADER and BERT-based classification methods, enabling effective analysis of emotional tone and contextual sentiment in spoken and written responses. To assess the quality of content, a scoring mechanism was implemented that considered coherence, grammatical correctness, and relevance to the interview context. This ensured that the responses were not only emotionally appropriate but also structurally and contextually sound. The overall accuracy of the NLP module was measured using a labeled dataset of interview responses and achieved a high-performance score of 91.5%. This indicates the model's strong ability to interpret and evaluate candidate responses reliably. The combination of sentiment analysis and content quality assessment provided a comprehensive view of each candidate's communication effectiveness, making the system highly beneficial for both interview preparation and automated evaluation processes.

4.2 User Experience Evaluation

A small-scale usability test was conducted with 30 participants (aged 20–28, preparing for job interviews). User feedback on the AI-Based Mock Interview Evaluator was largely positive. Approximately 83% of users reported that the tool significantly helped improve their body language and confidence during mock interviews. Around 90% appreciated the real-time feedback and visual analytics features, finding them insightful and engaging. However, 70% suggested minor improvements to the user interface to enhance interaction and ease of use. The average session duration was 15 minutes, indicating efficient engagement. Additionally, the tool demonstrated high reliability, with an error rate of less than 3% during usage, showcasing its robustness and usability in practical scenarios.

4.2.1 Visualization And Feedback

Heatmaps, line graphs, and emotion timelines were generated for post-interview feedback. A final performance report was produced, including: Emotion trends, and Confidence trajectory.

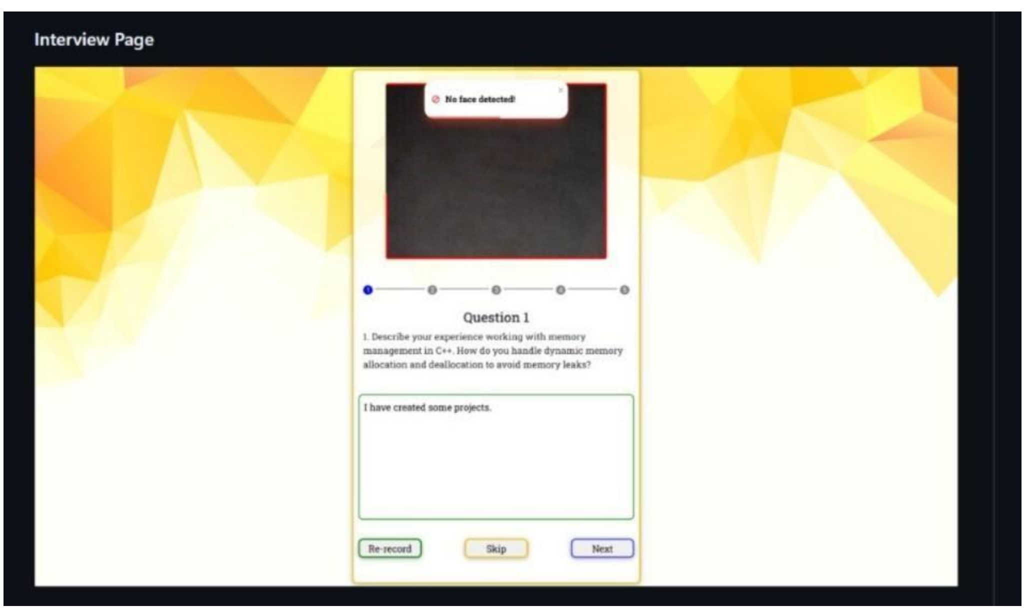

Figure 3: When no face is detected

5. Discussion

AI-powered mock interview systems offer several key advantages that enhance the recruitment and training process. One of the primary benefits is scalability. These systems can evaluate multiple candidates simultaneously, significantly reducing the time, cost, and logistical effort required compared to traditional in-person mock interviews. This is especially useful for institutions or companies managing large pools of candidates. Another strength is objectivity. Unlike human interviewers, who may be influenced by unconscious bias or fatigue, AI systems provide consistent and data-driven evaluations. This ensures a fairer assessment process for all candidates. Additionally, AI platforms offer a high degree of personalization by analyzing each candidate's performance and generating specific, actionable feedback. This helps individuals focus on improving their weaknesses—such as communication, confidence, or body language—and enhances their overall readiness for real interviews. Despite these advantages, some limitations remain. AI models rely heavily on the availability and quality of training data; poor or limited datasets can result in inaccurate assessments. Moreover, interpreting complex emotional cues, such as sarcasm, nervousness, or subtle non-verbal behavior, continues to be a challenge for most AI systems. Therefore, while AI tools are powerful aids, they are most effective when complemented by human oversight.

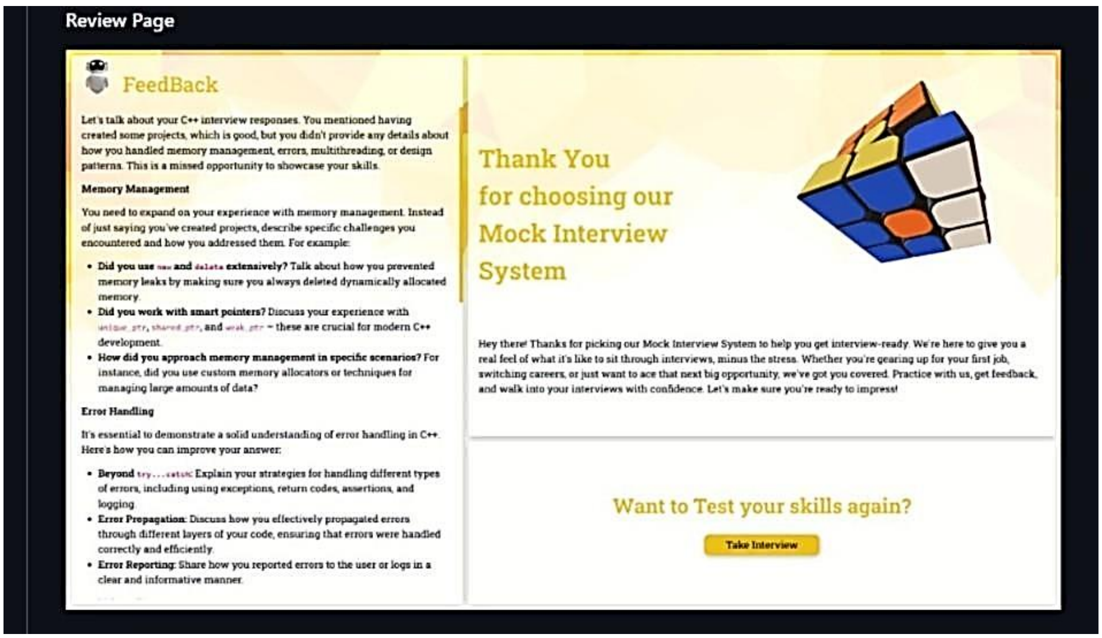

Figure 4: Showing the response

6. Conclusion

The AI-based mock interview evaluator represents a significant advancement in interview preparation, combining AI technologies to provide comprehensive, objective, and personalized feedback. As AI continues to evolve, such systems have the potential to revolutionize the way candidates prepare for interviews, leading to more effective and efficient recruitment processes. By integrating natural language processing, machine learning, and facial expression analysis, this system provides objective, real-time feedback on candidate responses, body language, tone, and overall communication skills. It offers a scalable and accessible alternative to traditional mock interviews, especially beneficial for individuals without access to human mentors or professional coaching. While the system enhances user readiness through personalized feedback and performance metrics, it is important to acknowledge certain limitations. These include the challenges of interpreting cultural nuances, ensuring unbiased evaluation, and simulating complex human interactions. Future improvements could involve refining emotion detection, expanding industry-specific question banks, and incorporating more interactive, adaptive learning modules. Overall, the AI-based mock interview evaluator is a promising tool for empowering job applicants to practice effectively, build confidence, and improve their chances of success in real-world interviews.

References

[1] R. Mandal, P. Lohar, D. Patil, A. Patil, and S. Wagh, "AI-Based mock interview evaluator: An emotion and confidence classifier model," International Conference on Intelligent Systems for Communication, IoT and Security (ICISCoIS), Coimbatore, India, 2023.

[2] Chethana C Viraktamath et al., "AI Based Mock Interview Evaluator," JoMA, August 2024. STM Journals

[3] K.C. and R.C.P., "Detection of Emotion Using Multi-Block Deep Learning in a Self-Management Interview App," MDPI, 2020. MDPI

[4] S. Temgire et al., "Real Time Mock Interview using Deep Learning," International Journal of Engineering Research & Technology (Ijert), Volume 10, Issue 05, 2021.

[5] Zhang, Z., Zhang, P., & Huang, Y. (2020). *Facial Expression Recognition in Deep Learning: A Review*. IEEE Access, 8, 146660-146681. https://doi.org/10.1109/ACCESS.2020.3017765

[6] Ko, B. C. (2018). *A brief review of facial emotion recognition based on visual information*. Sensors, 18(2), 401. https://doi.org/10.3390/s18020401

[7] Jurafsky, D., & Martin, J. H. (2023). *Speech and Language Processing* (3rd ed., Draft). Stanford University. https://web.stanford.edu/~jurafsky/slp3/

[8] Madnani, N., Cahill, A., Chodorow, M., Napolitano, D., & Tetreault, J. (2015). *Automated scoring of a summary-writing task designed to measure reading comprehension*. Proceedings of the Tenth Workshop on Innovative Use of NLP for Building Educational Applications.

[9] Kuncel, N. R., Klieger, D. M., Connelly, B. S., & Ones, D. S. (2013). *Mechanical versus clinical data combination in selection and admissions decisions: A meta-analysis.* Journal of Applied Psychology, 98(6), 1060–1072. https://doi.org/10.1037/a0034158

[10] Chamorro-Premuzic, T., Winsborough, D., Sherman, R. A., & Hogan, R. (2016). *New talent signals: Shiny new objects or a brave new world?* Industrial and Organizational Psychology, 9(3), 621-640. https://doi.org/10.1017/iop.2016.6

[11] Hoque, M. E., McDuff, D., & Picard, R. W. (2013). *Exploring temporal patterns in classifying frustrated and delighted smiles*. IEEE Transactions on Affective Computing, 3(3), 323–334.

[12] Robison, J., McQuiggan, S., & Lester, J. (2009). *Evaluating the consequences of affective feedback in intelligent tutoring systems*. International Journal of Artificial Intelligence in Education, 19(2), 119–154.

Cite this article as: Neha Gupta, Aashi Sharma, Suneet Shukla, Praveen Kumar, AI-based mock interview evaluator: An emotion and confidence classifier model, International Journal of Research in Engineering and Innovation Vol-9, Issue-3 (2025),136-142. *https://doi.org/10.36037/IJREI.2025.9308*