Ai Based Interview Platform: Prepify

Prashant Itankar[1], Shubham Lone[2], Priyanshu Ingle[3], Chinmay Khairnar[4], Arya Kelkar[5]
[1]*Prof., Dept. of Computer Engineering, Mumbai University, Navi Mumbai, India*
[2,3,4,5]*Dept. of Computer Engineering, Mumbai University, Navi Mumbai, India*

***Abstract*—In this study we introduce Prepify, an AI-enabled system developed to modernize interview preparation through intelligent simulation. Conventional preparation techniques often lack measurable feedback and realistic interaction, limiting candidate improvement. Our Prepify platform generates role-specific questions based on user-provided resumes or selected domains. So, spoken responses are captured and transcribed using advanced speech recognition techniques. The transcribed content is analysed using natural language processing to understand meaning, structure, and clarity. Also, adaptive questioning mechanism allows dynamic follow-up interactions during the session. A real-time text-to-speech engine creates a conversational experience resembling actual interviews. Facial and eye movement tracking is implemented using Media Pipe Iris to evaluate engagement and detect cheating. The system produces a comprehensive performance report covering communication, confidence, and behavioural indicators. Overall, Prepify provides a scalable, data-driven solution for improving interview readiness and reducing candidate anxiety.**

***Index Terms*—Artificial Intelligence, Mock Interview System, Speech-to-Text Processing, Natural Language Processing (NLP), Computer Vision, Eye Tracking, Human-Computer Interaction, Machine Learning, Behavioural Analysis, Interview Simulation.**

I. INTRODUCTION

In today's competitive job market, candidates are required to demonstrate not only technical expertise but also effective communication skills and confidence during interviews. Interviews remain one of the most widely adopted methods for evaluating a candidate's suitability for a role [1]. However, many candidates struggle to perform well due to anxiety, lack of preparation, and limited exposure to realistic interview environments.

Traditional preparation methods, such as studying commonly asked questions or practicing with peers, provide limited benefits as they often lack structured and objective feedback [3]. These approaches fail to assess important aspects such as communication clarity, response quality, and behavioural attributes, which play a significant role in interview success.

Recent advancements in artificial intelligence have enabled the development of intelligent systems capable of enhancing human-computer interaction [2]. Technologies such as speech recognition and natural language processing allow systems to capture and analyse spoken responses with high accuracy [4]. These methods can evaluate linguistic quality, contextual relevance, and overall response effectiveness.

Furthermore, computer vision techniques have been widely used to analyze non-verbal cues, including facial expressions, eye movement, and head posture [5]. Such behavioral indicators are essential for assessing candidate confidence and engagement during interviews. Frameworks like MediaPipe enable real-time facial landmark detection and iris tracking, making it possible to monitor user behavior effectively [6].

This research proposes Prepify, an AI-driven interview preparation platform that integrates speech processing, natural language processing, and computer vision into a unified system. The platform generates domain-specific questions based on user input, analyzes spoken responses, and provides intelligent feedback. Additionally, it monitors candidate behavior using eye-tracking techniques to ensure engagement and authenticity.

By combining multiple AI technologies into a single framework, the proposed system aims to provide a realistic interview simulation environment, enabling candidates to improve their performance through structured feedback and data-driven insights.

II. LITERATURE REVIEW

Previous research has explored the use of AI in recruitment and interview systems, mainly focusing on automation and efficiency. While these systems improve hiring processes, they often overlook candidate preparation.

Speech recognition studies show that modern systems can accurately convert speech into text, making them useful for interactive applications. However, they do not evaluate the quality of responses.

Machine learning models have been used to score interview answers, but many lack real-time interaction and adaptability. Similarly, natural language processing techniques can analyze grammar and meaning, but they often ignore delivery and behavior.

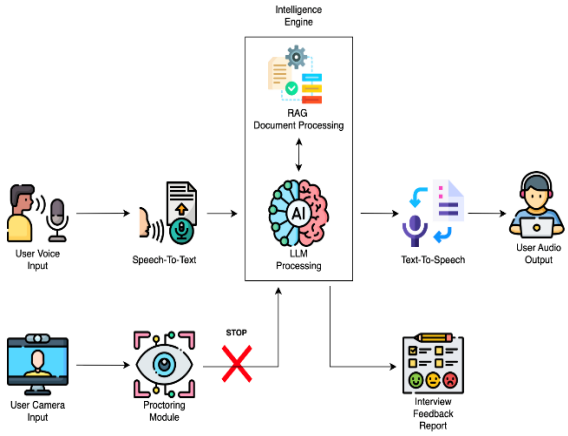
Computer vision research highlights the importance of non-verbal cues like facial expressions and eye movement in assessing confidence. However, these systems usually work independently without integrating speech analysis.

Most existing solutions focus on a single aspect—either speech, text, or behavior. There is a clear need for a unified system that combines all these features. Prepify fills this gap by integrating multiple AI technologies into one platform for a complete interview preparation experience.

III. METHODOLOGY

The proposed Prepify system follows a multi-module architecture integrating speech processing, natural language processing, and computer vision to simulate a realistic interview environment. Users begin by registering on the platform and uploading their resumes or selecting a job role. Relevant keywords such as skills and experience are extracted to tailor the interview. A question generation module retrieves domain-specific questions from a structured database based on the candidate's profile.During the interview, candidates respond verbally, and their audio input is captured and converted into text using a speech-to-text module [2]. The transcribed responses are then analyzed using NLP techniques to evaluate relevance, grammar, and coherence [4]. Based on this evaluation, the system dynamically generates adaptive follow-up questions, ensuring an interactive experience. Responses are also converted back into audio using a text-to-speech module for real-time interaction.In parallel, computer vision techniques monitor candidate behavior using MediaPipe Iris tracking to detect eye movement and facial landmarks [6]. Behavioral cues such as gaze deviation are recorded for further analysis. After the session, the system generates a comprehensive performance report, providing structured feedback on communication skills, confidence, and areas for improvement.

IV.ARCHITECTURE DIAGRAM



V. SPECIFICATIONS

The Prepify system is designed as an integrated AI-based interview preparation platform. It combines multiple technologies to simulate a realistic interview environment. The architecture consists of speech processing, NLP, computer vision, and backend modules. Each component is responsible for a specific functionality within the system. The speech recognition module captures candidate responses in audio form. It converts spoken input into textual data using speech-to-text techniques [2]. This ensures accurate transcription for further processing and analysis. The natural language processing module evaluates the textual responses. It analyzes semantic relevance, grammar, and contextual understanding [4]. The module also enables intelligent feedback and follow-up question generation. A text-to-speech module is used to convert system responses into audio output. This allows real-time conversational interaction with the virtual interviewer. Audio streaming is optimized to ensure minimal latency during communication. The system integrates a

computer vision module for behavioral analysis. MediaPipe-based iris tracking is used to monitor eye movement [6]. Facial landmark detection helps evaluate engagement and attentiveness. Behavioral data is recorded to identify irregular or suspicious activity. The backend system manages user authentication and session control. It stores interview data, responses, and performance history securely. This data is used to generate personalized feedback and recommendations. The platform is designed for scalability and efficient real-time processing. All modules are integrated to ensure smooth and seamless interaction. Overall, the system provides a comprehensive and data-driven interview solution.

VI. PROPOSED FRAMEWORK

The proposed Prepify framework is designed as a multi-stage process integrating various AI technologies. It aims to simulate a realistic and interactive interview environment. The framework consists of six major stages for systematic execution. The first stage involves user authentication and profile setup. Candidates register and upload resumes or select a target job role. Relevant information such as skills and experience are extracted. The second stage is the question generation module. Domain-specific interview questions are retrieved from a structured database. This ensures that the interview is tailored to the candidate's profile. The third stage involves capturing candidate responses. Audio input is recorded through a microphone interface. Speech-to-text processing converts audio into textual format [2]. The fourth stage is the AI evaluation model. Natural language processing techniques analyze response quality [4]. The system evaluates relevance, clarity, and completeness of answers. Based on evaluation, adaptive follow-up questions are generated. The fifth stage involves text-to-speech interaction. System responses are converted into audio using a TTS engine. Real-time audio streaming ensures smooth conversational flow. Simultaneously, computer vision monitors candidate behavior. MediaPipe Iris tracking detects eye movement and engagement [6]. The final stage generates a detailed performance report. The report highlights strengths, weaknesses, and improvement areas.

VII. RESULTS AND DISCUSSION

The proposed Prepify system was evaluated through multiple mock interview sessions. Participants from different technical backgrounds were included in the evaluation process. The system successfully simulated a realistic and interactive interview environment. Speech recognition performance was observed to be consistently accurate under normal conditions. The speech-to-text module effectively transcribed candidate responses [2]. Minor inaccuracies were observed only in cases of unclear audio input. The NLP-based evaluation model analyzed responses with satisfactory accuracy. It effectively assessed relevance, grammatical structure, and contextual meaning [4]. The system was able to generate meaningful and adaptive follow-up questions. This contributed to creating a dynamic and engaging interview experience.

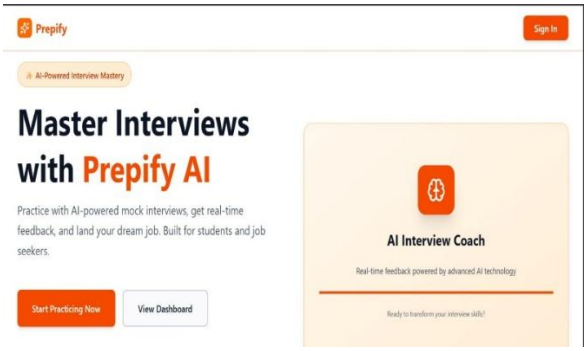


Figure 1: Home Page

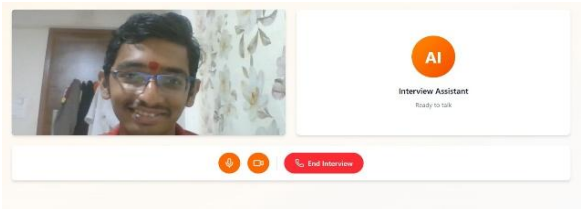


Figure 2: Live Interview Interface

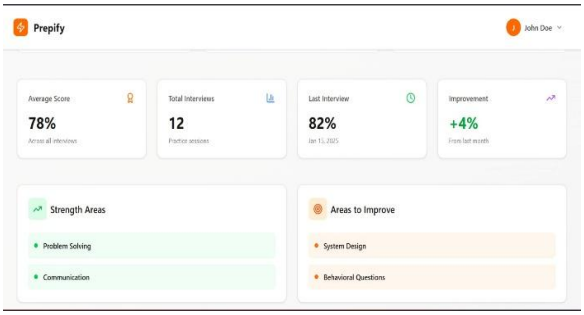


Figure 3: Candidate Dashboard

The text-to-speech module provided smooth and natural audio interaction. Low latency ensured uninterrupted communication during the session. The integration of computer vision enhanced behavioural analysis capabilities. MediaPipe Iris tracking successfully monitored eye movement in real time [6]. The system detected patterns such as gaze deviation and reduced attention. These behavioral indicators were useful in evaluating candidate engagement. The final performance report provided structured and detailed feedback. It included metrics related to communication, confidence, and response accuracy. Participants reported improved awareness of their strengths and weaknesses. The system also helped reduce interview anxiety through repeated practice. However, performance may vary depending on audio quality and lighting conditions. Further improvements can enhance accuracy and adaptability of the system. Overall, the results demonstrate the effectiveness of the proposed framework.

VIII. CONCLUSION

This research introduced Prepify, an AI-driven platform designed to enhance interview preparation through an interactive and realistic simulation environment. The system integrates speech processing, natural language processing, and computer vision to evaluate both verbal responses and non-verbal behavioral cues. By providing structured feedback, adaptive questioning, and real-time interaction, the platform enables candidates to gain deeper insights into their performance and improve key skills such as communication, clarity, and confidence. The inclusion of behavioral analysis further strengthens the evaluation by capturing engagement and attentiveness during the interview process. The system not only addresses the limitations of traditional preparation methods but also offers a scalable and data-driven approach to learning. Experimental observations indicate that the platform effectively improves user awareness and reduces interview anxiety through continuous practice. Overall, Prepify demonstrates a practical and intelligent solution for modern interview preparation, with strong potential for future enhancements through advanced AI models, emotion recognition, and personalized learning mechanisms.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the institution and faculty members for their continuous support, guidance, and encouragement throughout the development of this research work. Their valuable insights and suggestions played a crucial role in improving the quality and direction of the project. The authors also extend their appreciation to all participants who contributed to testing and evaluating the system, providing meaningful feedback that helped refine the platform. We also acknowledge the technical resources and infrastructure provided, which enabled the successful implementation of the proposed system. Special thanks are extended to peers and colleagues for their cooperation and support during the research process, which contributed to overcoming challenges and improving the overall outcome of this work.

REFERENCES

[1] R. Mandal, P. Lohar, D. Patil, A. Patil *et al.*, “AI-based mock interview evaluator: An emotion and confidence classifier model,” in *Proc. International Conference on Intelligent Systems*, 2023.

[2] Y. C. Chou, F. R. Wongso, and C. Y. Chao *et al.*, “An AI mock-interview platform for interview performance analysis,” in *Proc. 10th International Conference on Intelligent Systems*, 2022.

[3] M. Krithika, A. Jamuna, and S. Indhulega, “AI-based mock interview application,” in *Proc. 3rd International Conference on Intelligent Systems*, 2025.

[4] P. J. Uppalapati, M. Dabbiru, and V. R. Kasukurthi, “AI-driven mock interview assessment: Leveraging generative language models for automated evaluation,” *International Journal of Intelligent Systems*, 2025.

[5] R. Patel and P. Singh, “Computer vision techniques for facial behavior analysis,” *International Journal of Computer Vision Applications*, vol. 30, no. 1, pp. 61–70, 2023.

[6] Google AI Research, “MediaPipe documentation: Real-time face and iris tracking framework,” 2024.

[7] D. Jurafsky and J. H. Martin, *Speech and Language Processing*, 3rd ed. Pearson, 2023.
[8] Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
[9] Vaswani *et al.*, "Attention is all you need," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.
[10] T. Baltrušaitis, P. Robinson, and L.-P. Morency, "OpenFace: An open-source facial behavior analysis toolkit," in *Proc. IEEE Winter Conf. on Applications of Computer Vision (WACV)*, 2016.
[11] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[12] J. Devlin *et al.*, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, 2019.