AI-DRIVEN INTERVIEW PREPARATION SYSTEM

Kumara Sundari , Gopika S, Tania Stacy G and Praveena I
Assistant Professor[1], Final Year[2]
Department of Information Technology,
St.Joseph College of Engineering,
Sriperumbudur, Chennai-602117

***Abstract*—** The AI-driven Interview Preparation System is a web-based platform designed to simulate real interview environments and enhance a candidate's readiness through intelligent analysis. Upon accessing the system, users are presented with a series of interview-based questions displayed on the webpage. Simultaneously, the system activates the user's webcam to capture real-time facial expressions and emotional responses using a deep learning model based on the InceptionV3 algorithm. This model processes visual cues to detect emotional states or hesitation during the interview. The system not only records the user's responses but also evaluates them in the context of the detected emotions, providing a holistic performance review. This innovative integration of emotion recognition and answer validation mimics real-world interview conditions, helping users to identify areas of improvement in both verbal and non-verbal communication. The solution aims to build user confidence and improve interview outcomes by offering personalized, making it a powerful tool in modern interview preparation.

I. INTRODUCTION

In today's competitive job market, interview performance plays a crucial role in determining a candidate's success. While many individuals prepare by practicing questions, few platforms address both verbal and non-verbal aspects of communication.

Non-verbal cues such as facial expressions, body language, and emotional responses can greatly influence an interviewer's perception. The AI-driven Interview Preparation System aims to bridge this gap by simulating a real-time interview environment. This web-based tool leverages deep learning and computer vision techniques to monitor users' facial expressions while they answer typical interview questions.

A webcam is activated during the session to capture emotional responses, such as anxiety, confusion, or confidence. These responses are analyzed using an InceptionV3-based emotion recognition model. At the same time, the system evaluates the content and delivery of the answers provided. This dual-layer analysis offers users a holistic review of their performance. The system ultimately helps users enhance their communication skills and emotional control during interviews.

II. EXISTING AND PROPOSING SYSTEM

The existing interview preparation systems primarily focus on multiple-choice questions or basic mock interviews. Some platforms provide recorded questions that users can answer verbally, with limited feedback. These systems often rely on keyword matching to evaluate answers, which lacks depth and understanding of content.

Non-verbal communication analysis is usually not included or is minimal. In some premium platforms, live coaching or human evaluation may be available, but it is expensive and time-consuming. Emotion detection, if implemented, is typically isolated from the actual answer evaluation. This leads to fragmented feedback.

Most tools are static and do not adapt to the user's performance. They are unable to offer personalized improvement strategies. As a result, users fail to gain holistic feedback on their preparation.

III. SYSTEM STUDY

System study is a detailed analysis of how feasible, cost-effective, and practical the proposed project will be. It helps evaluate whether the Emotion-Aware AI Interview Preparation System can be successfully developed and implemented. The study covers technical feasibility, economic feasibility, and operational feasibility to ensure the project is achievable, affordable, and easy to operate.

A.Technical Feasibility

The system leverages existing, well-supported technologies including deep learning (InceptionV3), real-time computer vision, and natural language processing. These technologies are mature, with available libraries like TensorFlow and OpenCV enabling rapid development and deployment. The platform is browser-based, ensuring accessibility across devices with a webcam and internet connection.

B.Economic Feasibility

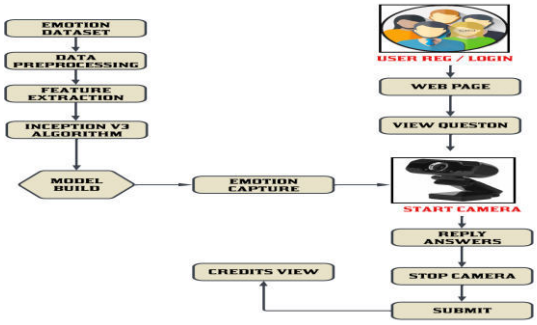
Compared to human-led mock interviews or coaching, the system is cost-effective. Once deployed, it requires minimal recurring cost apart from maintenance. It removes the need for expensive personal coaching by providing automated feedback, making high-quality interview training accessible to a wider audience.

C. Operational Feasibility

The platform is easy to use with a clean, interactive web interface. Users simply access it through a web link, answer questions, and receive intelligent feedback. No specialized training is needed to operate it. It can scale to serve students, job seekers, and professionals with minimal effort from administrators..

IV. ARCHITECTURE DIAGRAM

The architecture of the Emotion-Aware AI Interview Preparation System is designed to deliver real-time, intelligent, and personalized interview simulation experiences. It consists of tightly integrated modules for question delivery, facial emotion detection, verbal response analysis, and feedback generation—all connected through a web-based interface.



V. ALGORITHMS USED

1) Natural Language Processing (NLP)

Natural Language Processing (NLP) is a field that combines computer science, artificial intelligence and language studies. It helps computers understand, process and create human language in a way that makes sense and is useful. With the growing amount of text data from social media, websites and other sources, NLP is becoming a key tool to gain insights and automate tasks like analyzing text or translating languages.

2) InceptionV3

At the heart of this project lies an advanced emotion detection algorithm powered by the InceptionV3 deep learning model. InceptionV3 is a convolutional neural network known for its high accuracy in image classification tasks. The system captures real-time video frames from the user's webcam during a mock interview. Each frame is processed and passed through the InceptionV3 model to classify facial expressions into emotional states like happy, sad, neutral, or anxious. This emotion data is timestamped and correlated with the user's spoken response. The algorithm ensures emotions are detected precisely during key moments of the interview. It enables the system to identify patterns such as repeated nervousness or loss of confidence. This emotion classification forms a vital part of the performance review. The algorithm is trained on diverse facial datasets to handle varied expressions across demographics. Its speed and accuracy make it ideal for real-time emotion recognition in web-based systems.

VI. LIST OF MODULES

- Emotion Detection Algorithm
- Intelligent Feedback Generation System
- Algorithmic Flow of the Proposed System
- Real-Time Web-Based AI Integration

A. Algorithmic Flow of the Proposed System:

At the heart of this project lies an advanced emotion detection algorithm powered by the InceptionV3 deep learning model. InceptionV3 is a convolutional neural network known for its high accuracy in image classification tasks. The system captures real-time video frames from the user's webcam during a mock interview. Each frame is processed and passed through the InceptionV3 model to classify facial expressions into emotional states like happy, sad, neutral, or anxious. This emotion data is timestamped and correlated with the user's spoken response. The algorithm ensures emotions are detected precisely during key moments of the interview. It enables the system to identify patterns such as repeated nervousness or loss of confidence. This emotion classification forms a vital part of the performance review. The algorithm is trained on diverse facial datasets to handle varied expressions across demographics. Its speed and accuracy make it ideal for real-time emotion recognition in web-based systems.

B. Intelligent Feedback Generation System : The feedback generation system is designed to offer users a detailed review of their performance, combining both emotional and verbal analysis. After detecting emotions using InceptionV3, the system evaluates the quality of the user's responses using a response validation module. This module checks the clarity, completeness, and relevance of each answer. It also tracks hesitation points or inconsistent emotions. These two streams of data—verbal and emotional—are synchronized to generate insightful, personalized feedback. The user receives suggestions on how to improve composure, confidence, and answer quality. The system flags areas where emotional control was lost or where answers lacked structure. This feedback is presented in a report format at the end of each session. It mimics the kind of feedback a professional coach might provide, making it a valuable self-learning tool. This intelligent mechanism helps users become more self-aware and better prepared for real interviews.

C. Algorithmic Flow of the Proposed System: The functioning of the proposed system follows a step-by-step algorithm that integrates both computer vision and natural language processing. The process begins with the system prompting the user with interview questions on a web interface. Simultaneously, the user's webcam is activated to capture facial expressions. These facial images are sent through the pre-trained InceptionV3 model to identify the emotional state of the user in real-time. Meanwhile, the user's spoken or typed response is captured and processed by the response evaluation module. This module checks for answer relevance, fluency, and logical structure. The results from both modules are merged based on time markers to assess emotional stability during each response. Finally, the system generates a detailed performance report including emotional trends, answer quality scores, and improvement suggestions. This complete algorithm helps simulate a real interview setting with AI-driven evaluation. It ensures that both content and delivery are assessed accurately for better user preparation..

D. Real-Time Web-Based AI Integration: This system demonstrates the power of integrating artificial intelligence into a user-friendly web application. The platform operates entirely online, using a browser-based interface for accessibility and convenience. Users don't need to install any software—they simply access the system via a web link. Behind the scenes, the webcam feed is processed using real-time APIs and computer vision tools. The InceptionV3 emotion model runs in the backend, classifying expressions as soon as they're captured. At the same time, user responses are transcribed or evaluated live, depending on the input method. The web interface remains interactive, smooth, and intuitive, encouraging user engagement. Real-time performance feedback is made possible by this seamless integration of

frontend and backend components. This makes the system scalable and responsive across devices. It serves as a powerful example of how modern web technologies can be combined with AI to deliver intelligent learning tools.

VII. SCOPE FOR FUTURE DEVELOPMENT

In the future, the Hospital Navigation Voice Chatbot system can be significantly enhanced to provide more advanced and user-friendly features. One major enhancement could include integrating multilingual voice support to cater to a diverse range of users from different linguistic backgrounds. Additionally, the system could be extended to include indoor GPS or Bluetooth-based location tracking for more accurate real-time navigation within complex hospital layouts.Integration with hospital databases can allow the chatbot to provide appointment reminders, real-time doctor availability, and emergency response guidance. Moreover, incorporating facial recognition and personalized user profiles could help in delivering customized services to returning patients.Future versions may also support wearable device integration, enabling hands-free voice navigation and accessibility for physically challenged individuals. These enhancements would collectively improve patient experience, hospital efficiency, and overall accessibility.

VIII.CONCLUSION

The AI-driven Interview Preparation System offers a modern solution to a long-standing problem in interview readiness. By combining emotion detection with answer evaluation, it goes beyond traditional preparation tools.

The use of InceptionV3 ensures accurate emotional state classification, while verbal analysis ensures the quality of communication is assessed. This system empowers users with meaningful, actionable insights.

It fosters confidence, improves communication, and prepares candidates for real-world scenarios. The platform represents a step forward in intelligent career training. With its adaptability and AI capabilities, it stands out as an innovative tool in the job preparation landscape.

IX. REFERENCE

[1]. F. Nagasawa, S. Okada, T. Ishihara and K. Nitta, "Adaptive Interview Strategy Based on Interviewees' Speaking Willingness Recognition for Interview Robots," in IEEE Transactions on Affective Computing, vol. 15, no. 3, pp. 942-957, July-Sept. 2024, doi: 10.1109/TAFFC.2023.3309640.

[2]. C. Kim, J. Choi, J. Yoon, D. Yoo and W. Lee, "Fairness-Aware Multimodal Learning in Automatic Video Interview Assessment," in IEEE Access, vol. 11, pp. 122677-122693, 2023, doi: 10.1109/ACCESS.2023.3325891.

[3]. M. Baslyman, "Digital Transformation From the Industry Perspective: Definitions, Goals, Conceptual Model, and Processes," in IEEE Access, vol. 10, pp. 42961-42970, 2022, doi: 10.1109/ACCESS.2022.3166937.

[4]. S. Ryan, C. Nadal and G. Doherty, "Integrating Fairness in the Software Design Process: An Interview Study With HCI and ML Experts," in IEEE Access, vol. 11, pp. 29296-29313, 2023, doi: 10.1109/ACCESS.2023.3260639.

[5]. D. Ceneda, A. Arleo, T. Gschwandtner and S. Miksch, "Show Me Your Face: Towards an Automated Method to Provide Timely Guidance in Visual Analytics," in IEEE Transactions on Visualization and Computer Graphics, vol. 28, no. 12, pp. 4570-4581, 1 Dec. 2022, doi: 10.1109/TVCG.2021.3094870.

[6]. E.-R. Lukacik, J. S. Bourdage, and N. Roulin, ''Into the void: A conceptual model and research agenda for the design and use of asynchronous video interviews,'' Hum. Resource Manage. Rev., vol. 32, no. 1, Mar. 2022, Art. no. 100789. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1053482220300620

[7]. HireVue. About the Company: Leadership & Ceo: Hirevue. Accessed: May 29, 2024. [Online]. Available: https://www.hirevue.com/about

[8]. A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, ''Robust speech recognition via large-scale weak supervision,'' in Proc. Int. Conf. Mach. Learn., vol. 202, A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, and J. Scarlett, Eds., Honolulu,

HI, USA, 2023, pp. 28492–28518. [Online]. Available: https://proceedings.mlr.press/v202/radford23a.html

[9]. H. J. Escalante, H. Kaya, A. A. Salah, S. Escalera, Y. Güçlütürk, U. Güçlütürk, X. Baró, I. Guyon, J. C. S. Júnior, M. Madadi, S. Ayache, E. Viegas, F. Gürpinar, A. S. Wicaksana, C. C. S. Liem, M. A. J. van Gerven, and R. van Lier, ‘‘Explaining first impressions: Modeling, recognizing, and

[10]. L. Chen, G. Feng, C. W. Leong, B. Lehman, M. P. Martin-Raugh, H. Kell, C. M. Lee, and S. Yoon, ‘‘Automated scoring of interview videos using Doc2Vec multimodal feature extraction paradigm,’’ in Proc. 18th ACM Int. Conf. Multimodal Interact., Y. I. Nakano, E. André, T. Nishida, L. Morency, C. Busso, and C. Pelachaud, Eds., Tokyo, Japan, Nov. 2016, pp. 161–168, doi: 10.1145/2993148.2993203