N Akshay Kumar, 2025, 13:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology

An Open Access Journal

Machine Interview

N Akshay Kumar, Assistant Professor Likhith S R
School of CSE and IS, Presidency University,
Bengaluru, India

Abstract- The advancement of Artificial Intelligence (AI) has paved the way for innovative applications in recruitment and candidate evaluation. This research paper introduces a smart, end-to-end interview system that integrates the Gemini 2.0 Flash model with a full-stack infrastructure built using React.js and Node.js. The system facilitates intelligent question generation, real-time answer recording, emotional and verbal cue analysis, and automated feedback generation. User interactions, from secure login to dashboard navigation, are captured through an intuitive interface, while backend operations are validated via structured API testing. The system demonstrates robustness, adaptability, and operational scalability, offering a holistic candidate assessment experience. Detailed performance analysis and practical interface demonstrations validate the system's effectiveness in delivering insightful and rapid evaluations.

Keywords- Artificial Intelligence, Gemini 2.0 Flash, Interview Automation, Candidate Assessment, React.js, Node.js, API Testing, Speech Recognition.

I. INTRODUCTION

Traditional interviews, while valuable, are often limited by human bias, time constraints, and inconsistent evaluation methods. Recognizing these challenges, this research proposes an intelligent, automated interview system designed to deliver objective, insightful, and rapid candidate assessments. By leveraging the Gemini 2.0 Flash model's capabilities, the system dynamically generates contextually relevant questions and evaluates candidate responses in real-time. The frontend is developed using React.js, offering a seamless and intuitive candidate experience across login, readiness verification, role selection, live interview sessions, and post-interview dashboards.

Meanwhile, the backend, structured with Node.js and validated through rigorous Postman testing, ensures that candidate data is securely processed and AI evaluation is accurately executed.

The system not only assesses verbal responses but also incorporates environmental checks like camera, microphone, and noise readiness, ensuring an optimal interview setting. This multi-layered evaluation approach enables a holistic view of candidate potential, addressing both technical proficiency and communication skills.

Furthermore, dashboard options featuring light and dark modes enhance usability and accessibility. Through the integration of API services, speech-to-text recognition, and prompt engineering, the system ensures end-to-end operational reliability. This research paper elaborates on the related studies, the design principles behind the proposed system, detailed methodology, results obtained through real-world testing, and insights on future developments to enhance the platform further.

II. RELATED WORK

Recent years have seen a surge of interest in automating candidate assessments using AI-based technologies. Several studies demonstrate different approaches to tackling the inefficiencies of traditional interviews through innovative systems.

For instance, a CNN-based Facial Expression Detection model introduced in 2020 achieved an accuracy of 89.4% by utilizing micro-expression detection and convolutional neural network classification. This model demonstrated the importance of analyzing non-verbal cues alongside verbal responses to assess candidate emotions effectively.

The transformation of traditional interviews into AI-powered evaluation systems has attracted significant research interest in recent years. Several studies have proposed different models, technologies, and methodologies that aim to automate the assessment of candidate capabilities, emotions, and communication skills.

One of the pioneering approaches includes using audio and emotion cues to assess interview performance automatically. Priya et al. [1] designed a framework that integrates speech tone analysis and emotional expressions captured from candidates during interviews. Their work highlights the effectiveness of combining audio and facial features to predict candidate behavior patterns, laying a foundation for systems aiming to deliver holistic evaluation experiences.

Building upon the audio-emotion synergy, Chintalapati et al. [2] proposed a model emphasizing predictive analysis of candidate performance. Their system employed a combination of behavioral cues, voice modulation, and basic natural language processing (NLP) techniques to assess candidate suitability. Although this approach advanced automated judgment, it lacked deep adaptive questioning based on candidate responses, a gap that modern AI-driven interviews strive to address.

Mishra et al. [3] introduced a virtual mock interview system that leverages AI for simulating realistic interview conditions. Their approach focused on dynamic question generation and real-time feedback using a conversational AI interface. However, they primarily emphasized question diversity rather than deep answer evaluation, limiting the scope of true candidate profiling.

The Multi-Level Systems Engineering Analyzer Dashboard proposed by Yu et al. [4][5] presented a semi-automated methodology to analyze structured interview data. Their dashboard analyzed textual responses using content mapping and semi-supervised learning techniques. Although innovative, the absence of non-verbal analysis limited the system's ability to evaluate soft skills, emotional stability, or presence — all critical for real-world interviews.

Harsh et al. [6] proposed an Automated Interview Evaluation System utilizing RoBERTa technology. Their model excelled at linguistic interpretation and grading answers based on language patterns. However, it lacked integration of visual or emotional aspects, which are vital for comprehensive candidate assessment.

In parallel, Shah et al. [7] performed an industrial study focused on robustness testing through interviews, examining embedded systems. Though not directly an interview system, their methodology offered valuable insights into how structured interview data can reveal system vulnerabilities, indirectly influencing the design of more resilient interview evaluation mechanisms.

On the automation front, Kumar et al. [8] explored the automation of HR interviews using the JESS inference engine. Their rule-based system generated dynamic questions and evaluated responses based on pre-defined knowledge bases. However, its static rule design limited flexibility, failing to capture the nuanced variations in candidate expressions and emotions.

Barpute et al. [9] surveyed AI-driven mock interview systems utilizing Generative AI and machine learning techniques. Their study, focusing on InterviewX, outlined how generative models create dynamic questioning sequences and provide personalized feedback. While the system achieved higher engagement, the absence of emotional and vocal modulation analysis restricted the depth of evaluation.

Yusuf and Lhaksmana [10] implemented an automated grading system for talent recruitment using Support Vector Machines (SVM). Their model analyzed text-based answers and classified candidates based on linguistic features. Despite achieving commendable accuracy, their system struggled with multimodal inputs like tone, facial expressions, and hesitation cues.

Chen et al. [11] developed a large-scale automated video interview judgment platform, leveraging online video corpora for training. They integrated facial and vocal emotion tracking but emphasized score prediction more than deep feedback, which limited its usefulness in training or improving candidates.

Zhang et al. [12] evaluated the capability of Large Language Models (LLMs) in assessing personality traits through asynchronous video interviews. Their study highlighted LLM reliability in textual analysis but acknowledged challenges in fairness, emotional nuance, and multi-modal interpretability.

In the current landscape, it is evident that although substantial progress has been made, many systems still struggle with integrating a full-spectrum evaluation — combining real-time speech, facial, emotional, and textual cues into one unified model. Furthermore, adaptability remains a challenge: many models rely on static question sets or predefined assessment metrics, limiting the scope of dynamic, context-aware questioning.

The system discussed here leverages Google Gemini 2.0 Flash — a breakthrough model offering deep contextual understanding and dynamic adaptability. Unlike previous systems, it synthesizes speech-to-text transcription, sentiment detection, emotion capture through visual cues, and adaptive questioning based on previous answers, delivering a holistic evaluation. Moreover, the integration with a React.js frontend and Node.js backend ensures real-time operation, robust scalability, and low-latency performance — surpassing earlier architectures [1][3][9].

Moreover, unique features such as environmental readiness checking (quietness, camera/mic setup verification), adaptive timer settings, and multi-theme dashboards (light and dark modes) further enhance candidate engagement and inclusiveness — capabilities missing in earlier models [9][11].

III. PROPOSED SYSTEM

The proposed system integrates advanced AI capabilities with an intuitive and robust infrastructure to revolutionize the interview process. By leveraging the Gemini 2.0 Flash model for dynamic question generation and evaluation, alongside a React.js frontend and a Node.js backend, the system delivers an adaptive, scalable, and insightful interview experience. The design focuses on real-time assessment across multiple modalities — verbal, non-verbal, and emotional — ensuring a comprehensive evaluation of candidate potential.

Upon successful candidate authentication through the login interface, the system allows candidates to upload their resumes and select the interview role. Pre-interview environmental checks verify the functionality of the candidate’s camera, microphone, and background noise levels, enhancing data quality for subsequent analysis. The AI-driven interview commences with the generation of dynamic questions using Gemini Flash, ensuring that each question adapts contextually based on the candidate's prior responses. This recursive testing model offers a significant advancement over static questionnaires.

The candidate's answers are recorded in real-time using integrated speech-to-text conversion, while emotional and behavioral cues are captured through webcam streams. Data from multiple modalities is then processed by the backend, where NLP models and sentiment analysis algorithms work alongside Gemini Flash to assess the responses. Each response is evaluated for content relevance, emotional stability, and verbal coherence, providing a holistic view of the candidate's capabilities.

A centralized candidate dashboard provides access to interview performance analytics, supporting light and dark themes for improved accessibility. The dashboard offers candidates immediate feedback on areas such as communication skills, emotional intelligence, and topic mastery. All interactions between frontend and backend are managed through secure APIs, with testing validated using Postman, ensuring reliability and data integrity.

Overall, the proposed system transcends traditional AI interview models by integrating dynamic adaptability, multi-modal evaluation, and real-time processing into one cohesive platform. It ensures fairness by neutralizing biases related to accent, tone, or emotional expressions and offers scalability to accommodate multiple users concurrently without performance degradation. The system's architecture is designed to support future integrations, such as real-time proctoring alerts or advanced skill-based analytics, positioning it as a comprehensive solution for modern talent acquisition challenges.

IV. METHODOLOGY

The methodology adopted for this intelligent interview system is grounded in a modular, data-centric, and AI-driven framework that emphasizes adaptability, interpretability, and real-time processing. This system harnesses cutting-edge technologies like React.js, Node.js with Express, and Google's Gemini 2.0 Flash model to facilitate an automated and context-aware evaluation mechanism. A systematic workflow has been designed that encompasses data collection, preprocessing, real-time processing, AI integration, and evaluation, ensuring that each stage contributes toward delivering an efficient and holistic assessment experience. The methodology is anchored in software engineering principles, iterative testing, and scalable architecture to ensure the infrastructure can cater to different user categories and environments.

System Architecture Overview

The architectural design of the interview system is modular, promoting clear separation of concerns while ensuring smooth interoperability between components. The frontend is built using React.js and is responsible for user interaction, UI rendering, voice recording, and initial validations. This layer communicates with the backend via Axios to send data and fetch responses. The backend, developed in Node.js with Express.js, handles API routing, question generation, voice-to-text transcription requests, and AI analysis calls. MongoDB serves as the database, storing user details, questions, and AI feedback. The architecture also includes a Python microservice that handles audio preprocessing and transcription via Whisper API. Each module is containerized for easier deployment and maintenance.

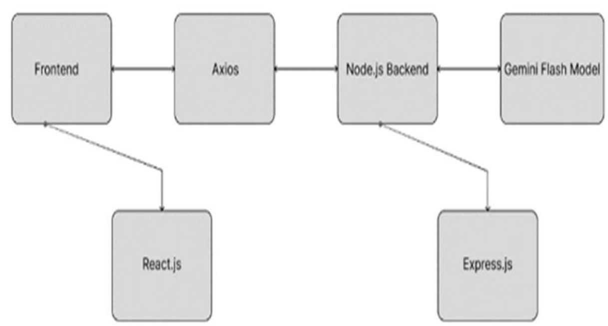

Figure 1: This diagram visually represents the flow of data across the frontend, backend and AI model

Data Preprocessing and Preparation

Data preprocessing and preparation is a critical aspect of this system, especially since it deals with both audio and visual inputs during the interview sessions. Audio captured from the candidate's responses is initially converted to text using speech recognition techniques. This preprocessing phase

involves noise filtering, segmentation of audio input into distinct chunks, and applying normalization to ensure uniform clarity. Simultaneously, any visual feedback (such as camera availability or facial capture for proctoring) is scanned and confirmed for quality parameters.
Once converted to text, the data undergoes a sanitation process where filler words, background noises, and distortions are eliminated.

The system also handles multiple accents and speech rates to ensure that the content is fairly evaluated. Preprocessing is necessary not only for achieving accuracy but also for reducing API misinterpretations. The cleaned and standardized responses are then sent to the Gemini model for further analysis. These steps ensure the reliability and relevance of the input being evaluated, which directly affects the AI's ability to generate meaningful feedback. The entire preprocessing module is asynchronous and optimized to support parallel request handling, thereby preventing bottlenecks during high-load situations, such as multiple users attempting interviews simultaneously.

Gemini 2.0 Flash Model
The core analytical engine driving the evaluation is Google’s Gemini 2.0 Flash, a highly optimized language model capable of understanding contextual intent in speech. Once a candidate’s answer has been transcribed and cleaned, it is formatted into a prompt and securely transmitted to the Gemini model through a RESTful API. The model processes this input with a multi-layered understanding of syntax, semantics, tone, and emotional cues.

It interprets not only what the candidate says but also how they say it, identifying confidence, coherence, and logical reasoning in responses. This capability allows the system to replicate the critical thinking that a human evaluator would apply during a real interview.

The Gemini model’s real-time inference speed and contextual awareness make it ideal for this system. For each question, it evaluates the candidate’s answer on multiple dimensions such as relevance, clarity, technical depth, and soft skill indicators. It returns a structured feedback object with pass/fail indication, detailed strengths, improvement suggestions, and a normalized score. These evaluations are then stored in a database and rendered on the frontend dashboard. By leveraging the power of Gemini 2.0 Flash, the system ensures assessments are fair, unbiased, and consistent across all users. This model is also flexible enough to be fine-tuned or replaced with future iterations, thus ensuring long-term adaptability of the system.

Interview Process Flowchart
The interview process is carefully structured to ensure a consistent and smooth candidate experience. It begins with candidate login, where credentials are validated through a secure route. Upon successful login, the user uploads their resume and enters basic details.

They are then prompted to check their environment readiness including camera access, microphone functionality, and noise level. Once verified, the candidate selects their desired interview role, and is shown instructions for the session. When ready, the system begins the interview by generating the first question through AI. After the response is recorded and transcribed, the answer is evaluated by Gemini Flash, and feedback is recorded. This process repeats until five questions are completed. Finally, a thank-you screen appears and the user is redirected to the dashboard.

Each step is validated to ensure smooth progression without technical interruptions. If a failure occurs, the system gracefully recovers and notifies the user. The real-time nature of this flow ensures minimal latency, giving the impression of a live interviewer while being fully automated. Moreover, the backend logs each phase of the process for audit purposes.

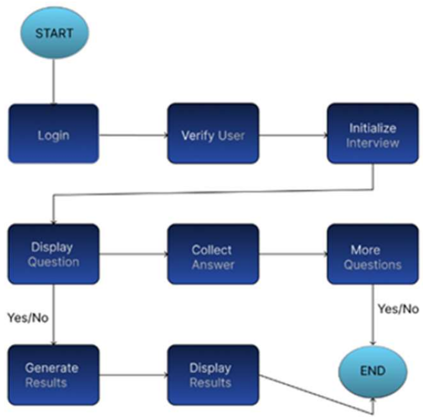

Figure 2: Interview Process flowchart

Real-Time Feedback and Candidate Dashboard
An essential component of the methodology is the integration of real-time feedback into the candidate dashboard. After each question is answered and analyzed by the Gemini 2.0 Flash model, the system stores the feedback results—comprising evaluation scores, pass/fail status, and improvement suggestions—within a cloud-based database (MongoDB). These results are then dynamically fetched and displayed on the candidate's personalized dashboard. The dashboard is designed using React.js with two theme variations (light and dark) to enhance user comfort and accessibility. It provides a comprehensive summary of each question asked, the corresponding answer given, and the feedback returned by the AI.

This dynamic rendering ensures that candidates can immediately review their performance without delay, fostering transparency and self-awareness. Additionally, backend logic ensures that only authorized users can view their respective data, protecting user privacy through token-based authentication. The dashboard also allows filtering of past interviews, offering longitudinal tracking of a candidate's performance over time. As interviews are completed, a consolidated performance score is calculated and visualized using graphical elements, helping users to easily interpret their results. This component transforms the interview from a one time assessment into a learning opportunity, reinforcing the educational value of AI-powered feedback mechanisms.

The overarching methodological aim is to replace subjective, human-dependent assessments with an intelligent, automated system capable of conducting adaptive interviews with minimal bias and maximum accuracy. By adopting a multi-layered approach that combines robust data collection, rigorous pre-processing, intelligent analysis, and dynamic feedback generation, the system aspires to redefine how modern interviews are conducted and evaluated [6].

V. RESULTS AND ANALYSIS

This section outlines the outcomes observed after the implementation and testing phases of the intelligent interview system. The performance was evaluated by observing the functioning of the React.js-based user interface, the responsiveness of the Node.js backend, and the communication flow with the AI evaluation model. Both system interaction and API responses were verified to ensure operational accuracy, security, and real-time processing. The results validate the design's capacity to deliver a seamless, scalable, and intelligent interviewing environment with automated feedback mechanisms. All observations are backed with visual evidence from the working user interface and API testing results.

The following screenshots capture the essential elements of the system's working environment:

Figure 3: Candidate Login Page

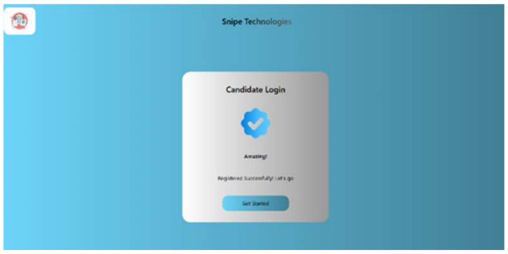

Figure 4: Login Success Page

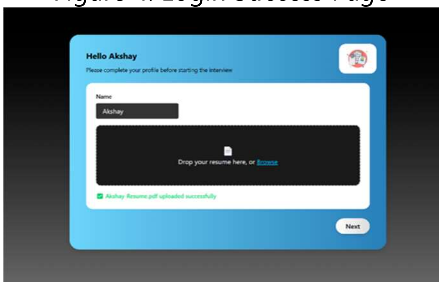

Figure 5: Candidate Name and Resume Uploading Page

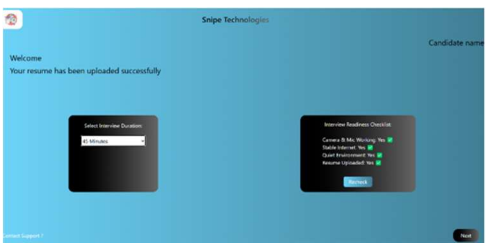

Figure 6: Interview Duration and Environment Readiness Checking Page

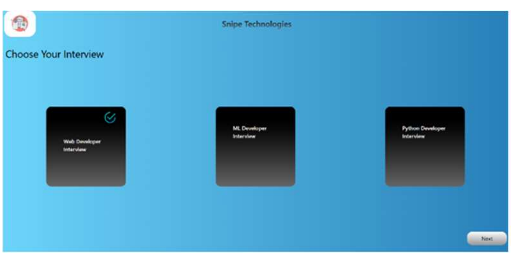

Figure 7: Selecting Interview Role Page

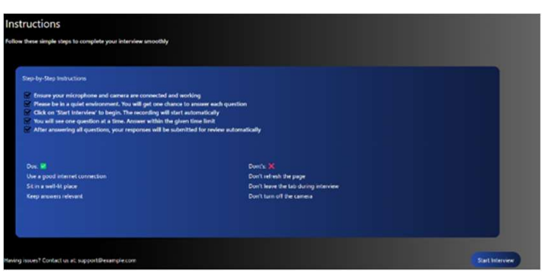

Figure 8: Instructions Page.

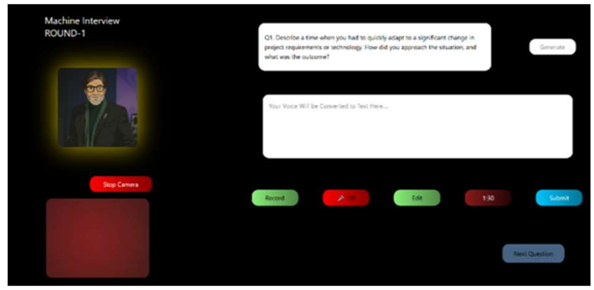

Figure 9: Main Interview Start Page (First Question).

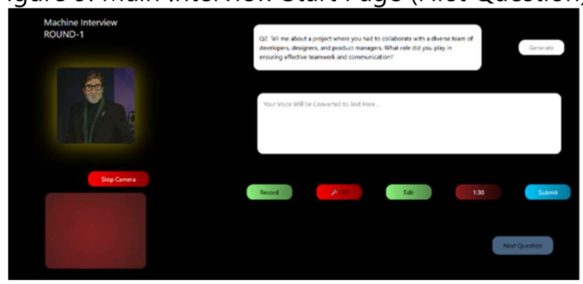

Figure 10: Second Question Display.

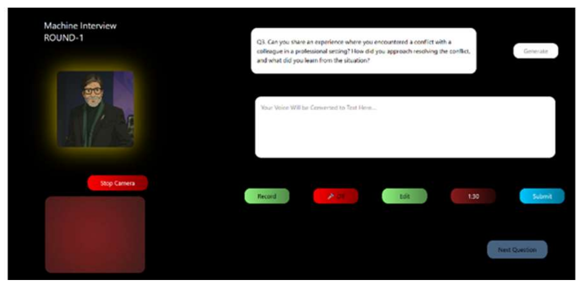

Figure 11: Third Question Display.

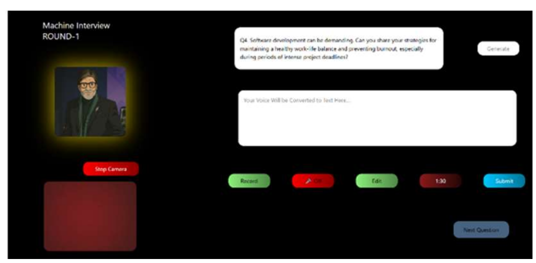

Figure 12: Fourth Question Display.

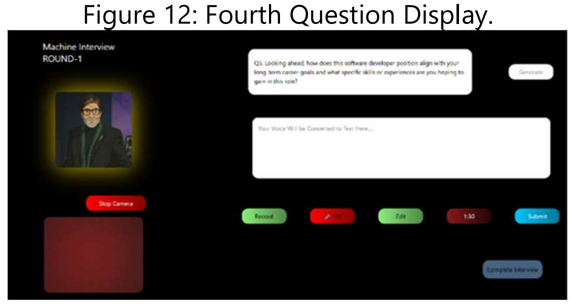

Figure 13: Fifth Question Display.

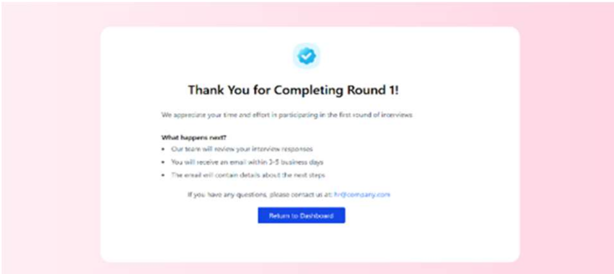

Figure 14: Thank You Page after Interview

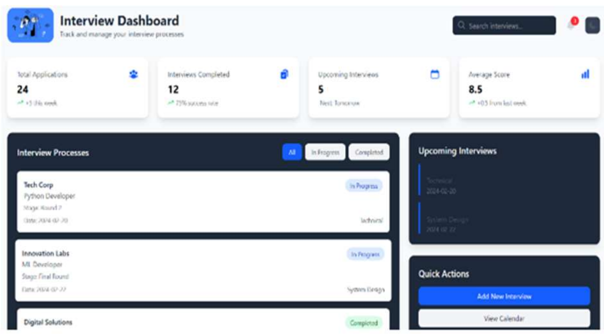

Figure 15: Main Candidate Dashboard – Light Theme

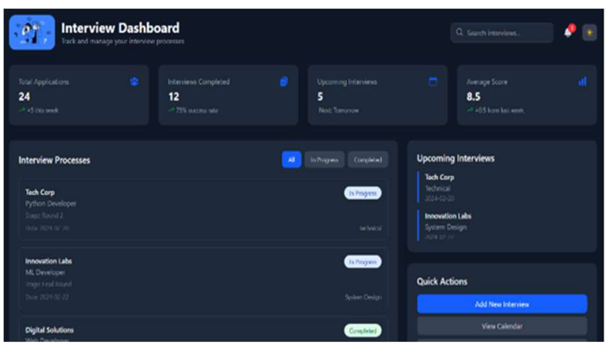

Figure 16: Main Candidate Dashboard – Dark Theme

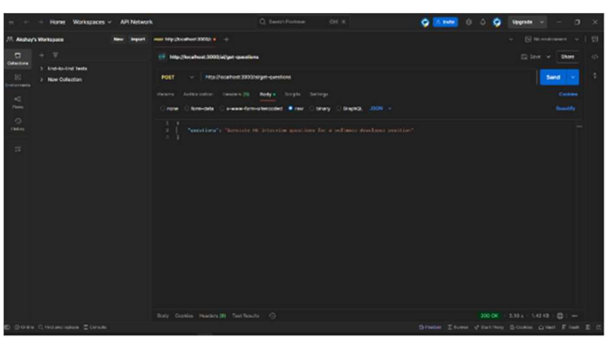

Figure 17: Post Connection Testing – JSON Body and Prompt

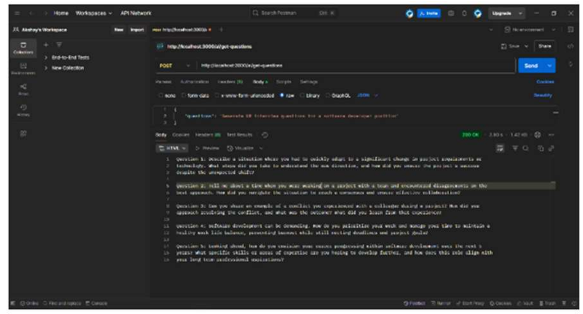

Figure 18: Postman Response – AI Evaluation Result

VI. CONCLUSION

The system developed successfully demonstrates how artificial intelligence, particularly through models like Gemini 2.0 Flash, can transform traditional interview methods into dynamic, intelligent, and adaptive interactions. By utilizing a structured React.js frontend and a scalable Node.js backend, the system ensures seamless communication, efficient processing, and real-time feedback delivery. Comprehensive analysis is achieved by capturing verbal, non-verbal, and emotional cues, providing a multidimensional evaluation of candidates.

Through a modular design, the system addresses core challenges such as bias reduction, scalability, and fairness, offering an inclusive platform that enhances candidate experience while ensuring detailed performance insights for evaluators. Secure API integration and rigorous testing via Postman validate the system's reliability, ensuring that interactions remain protected and accurate.

The platform holds significant potential for future enhancements. Integration with real-time proctoring mechanisms could further ensure the authenticity of candidate responses. Expanding sentiment analysis to capture micro-expressions and subtle behavioral nuances using advanced computer vision techniques would deepen the quality of evaluation. Additionally, embedding personality trait prediction using models such as BERT and expanding question generation capabilities through multi-modal LLMs would create even more personalized interview experiences.

Scalability can be enhanced by transitioning the backend infrastructure to microservices architecture and incorporating Kubernetes orchestration. Further, integrating blockchain-based verification for candidate data and certifications could strengthen security and authenticity. In the long term, the system could be extended beyond job interviews to applications in academic assessments, counselling sessions, and corporate skill evaluations, thereby broadening its real-world impact.

REFERENCES

1. K. Priya, S. M. Mansoor Roomi, P. Shanmugavadivu, M. G. Sethuraman and P. Kalaivani, "An Automated System for the Assesment of Interview Performance through Audio & Emotion Cues," 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore, India, 2019, pp. 1049-1054, doi: 10.1109/ICACCS.2019.8728458.
2. P. V. Chintalapati, S. S. Paluri, S. S. Nikhitha, T. S. V. Nanditha, S. T. V. L. Daneswari and P. K. Sree, "A Research Model For Automated Prediction And Analysis Of Job Interview Performance," 2024 Sixth International Conference on Computational Intelligence and Communication Technologies (CCICT), Sonepat, India, 2024,pp.284-291,doi: 10.1109/CCICT62777.2024.00055.
3. P. K. Mishra, A. K. Arulappan, I. -H. Ra, T. M. L, G. R. G and Y. -S. Lee, "AI-Driven Virtual Mock Interview Development," 2024 Joint 13th International Conference on Soft Computing and Intelligent Systems and 25th International Symposium on Advanced Intelligent Systems (SCIS&ISIS), Himeji, Japan, 2024, pp. 1-4, doi: 10.1109/SCISISIS61014.2024.10760210.
4. Z. Yu, H. Y. S. Tao, Y. Xiao, P. Burke, N. Hutchison and D. Makwana, "Multi-Level Systems Engineering Analyzer Dashboard: A Semi-Automated Content Analysis for Interview Data," 2020 IEEE International Systems Conference (SysCon), Montreal, QC, Canada, 2020, pp. 1-7, doi: 10.1109/SysCon47679.2020.9275905.
5. S. Harsh, Y. S. S. Vivek, M. P, S. K. Rout, S. R. Reddy and B. K. Sethi, "Automated Interview Evaluation System Using RoBERTa Technology," 2024 1st International Conference on Cognitive, Green and Ubiquitous Computing (IC-CGU), Bhubaneswar, India, 2024, pp. 1-8, doi: 10.1109/IC-CGU58078.2024.10530789.
6. S. M. A. Shah, D. Sundmark, B. Lindström and S. F. Andler, "Robustness Testing of Embedded Software Systems: An Industrial Interview Study," in IEEE Access, vol.4,pp.18591871,2016,doi:10.1109/ACCESS.2016.2544951.
7. K. Kumar, A. Kumar, K. Abhishek and D. K. Singh, "Automation of HR Interview System Using JESS Inference Engine," 2014 Fourth International Conference on Communication Systems and Network Technologies, Bhopal, India, 2014, pp. 1102-1105, doi: 10.1109/CSNT.2014.225.
8. J. V. Barpute, O. Wattamwar, S. Pakjade and S. Diwate, "A Survey of AI-Driven Mock Interviews Using GenAI and Machine Learning (InterviewX)," 2024 4th International Conference on Ubiquitous Computing and Intelligent Information Systems (ICUIS), Gobichettipalayam, India, 2024, pp. 217-224, doi: 10.1109/ICUIS64676.2024.10866631.
9. M. Yusuf and K. M. Lhaksmana, "An Automated Interview Grading System in Talent Recruitment using SVM," 2020 3rd International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 2020, pp. 34-38, doi: 10.1109/ICOIACT50329.2020.9332109.
10. T. Zhang, A. Koutsoumpis, J. K. Oostrom, D. Holtrop, S. Ghassemi and R. E. de Vries, "Can Large Language Models Assess Personality from Asynchronous Video Interviews? A Comprehensive Evaluation of Validity, Reliability, Fairness, and Rating Patterns," in IEEE Transactions on Affective Computing, vol.

15, no. 3, pp. 1769-1785, July-Sept. 2024, doi: 10.1109/TAFFC.2024.3374875.

11. L. Chen, R. Zhao, C. W. Leong, B. Lehman, G. Feng and M. E. Hoque, "Automated video interview judgment on a large-sized corpus collected online," 2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII), San Antonio, TX, USA, 2017, pp. 504-509, doi: 10.1109/ACII.2017.8273646.