

Mock Interview Platform with a Real Time AI Voice Agent

Ayush Modanval¹, Ankitha Mohan², Nabeel Ahmed Khan³, Dr. Payal Koolwal⁴

¹Computer Science and Engineering, Bangalore Technological Institute, Bangalore, India

²Computer Science and Engineering, Bangalore Technological Institute, Bangalore, India

³Computer Science and Engineering, Bangalore Technological Institute, Bangalore, India

⁴Computer Science and Engineering, Bangalore Technological Institute, Bangalore, India

Abstract - This project, "Mock Interview Platform with a Real-Time AI Voice Agent," is an AI-powered system designed to simulate realistic job interviews. It uses Next.js for the front-end, VAPI for real-time voice interactions, and Google Gemini for natural language understanding. The platform provides a dynamic, voice-driven experience where an AI agent asks adaptive questions and evaluates a candidate's responses. Key features include real-time voice processing, sentiment analysis, and customized questions. The system provides detailed performance reports and feedback to help users improve. Its goal is to automate and enhance the interview process for job seekers and recruiters by offering an efficient, personalized, and bias-reducing solution.

Key Words: Mock interview, next.js, VAPI, Gemini, AI agent, real-time voice processing, sentiment analysis, and customized questions, performance reports and feedback.

1. INTRODUCTION

Mock Interview Platform with a Real-Time AI Voice Agent aims to solve the inefficiencies of traditional interview methods. It's a web-based system that uses AI to simulate realistic, voice-based interviews. The platform uses Next.js for the frontend, VAPI for speech processing, and Google Gemini for natural language understanding and adaptive questioning. This system is designed to help candidates prepare with realistic practice and unbiased feedback, while also giving recruiters a more consistent, scalable, and efficient way to screen applicants.

2. OVERVIEW

The platform is a cutting-edge **AI-powered interview simulation** that uses **Next.js** for the frontend, **VAPI** for real-time voice interaction, and **Google Gemini** for its core intelligence. Unlike traditional tools, it engages users in **dynamic, voice-based conversations**, generating **context-aware and adaptive questions**. The system analyzes speech and sentiment to provide **instant feedback and performance reports** for both technical and behavioral interviews. This not only helps candidates prepare effectively but also provides organizations with a standardized, unbiased tool for efficient candidate screening.

3. PROBLEM STATEMENT

Traditional interview processes are largely manual, time-intensive, and prone to subjectivity, which reduces their effectiveness for both candidates and recruiters. Candidates often lack access to realistic practice environments and structured feedback, limiting the development of communication skills, confidence, and domain-specific knowledge. Recruiters, meanwhile, contend with interviewer bias, inconsistent evaluation standards, and the high cost of repetitive early-stage screening, all of which compromise fairness, reliability, and

scalability in hiring. These inefficiencies underscore the need for an AI-driven, real-time, and adaptive interview platform that can simulate human-like interactions, deliver objective performance assessments, and generate actionable insights to enhance candidate preparedness while streamlining recruitment workflows and improving organizational outcomes.

4. AIM

The primary objective of this project is to create an **AI-powered, web-based mock interview platform** that uses a **real-time voice agent**. It aims to solve the inefficiencies of traditional interview methods by providing a more efficient, interactive, and intelligent solution. The platform uses **Next.js**, **VAPI**, and **Google Gemini** to simulate realistic conversations and offer **immediate, constructive feedback** to candidates. The goal is to both enhance a candidate's preparation and provide recruiters with a consistent and unbiased screening tool.

5. OBJECTIVES

The primary objective of this project is to design and implement an AI-powered interview platform that provides real-time, voice-based mock interview experiences for candidates and recruiters. The specific objectives are as follows:

- **Simulation** – To replicate domain-specific, real-time mock interviews using an AI voice agent.
- **Evaluation** – To analyze user responses for communication skills, confidence, and content relevance.
- **Feedback** – To provide personalized, automated insights and performance reports for continuous improvement.
- **Efficiency** – To reduce reliance on human interviewers in early-stage screening, saving time and resources.
- **Usability** – To ensure a seamless, scalable, and user-friendly interface for diverse users.
- **Reliability** – To guarantee consistent and accurate query handling and system performance.
- **Analytics** – To apply real-time speech recognition and NLP techniques for fluency, sentiment, accuracy, and confidence analysis.
- **Versatility** – To simulate realistic interview scenarios across technical, HR, and behavioral domains

6. MOTIVATION

The "Mock Interview Platform with a Real-Time AI Voice Agent" was developed to address the significant inefficiencies and biases in the traditional hiring process. It solves a dual problem: candidates often lack access to realistic practice and personalized feedback, and recruiters struggle with time-consuming, subjective, and inconsistent evaluations. The platform uses AI, VAPI, and Google Gemini to create a lifelike interview simulation that offers instant, objective feedback. This approach helps candidates improve their skills while providing

organizations with a more efficient, consistent, and unbiased tool for screening applicants.

7. APPLICATION

- **For Candidates:** The platform helps them prepare for interviews by providing realistic, voice-based mock interviews with instant, personalized feedback to identify and improve weaknesses.
- **For Recruiters & HR:** It acts as an efficient initial screening tool, ensuring consistent and unbiased assessments while providing deep insights into candidate skills.
- **For Educational Institutions:** The system serves as a scalable training tool to bridge the academic-to-job readiness gap and allows educators to track student performance.

8. SYSTEM ARCHITECTURE

8.1 Project Overview

The platform is designed to simulate real-world interview experiences by integrating **AI-driven voice interaction** with personalized evaluation. It helps candidates practice interviews anytime, anywhere, and provides feedback on communication, confidence, and technical knowledge.

Key technologies:

- **Frontend:** Next.js (for building the user interface and handling routing, SSR, and performance optimization).
- **Voice Agent:** VAPI (Voice Agent API) for real-time AI-powered voice interactions.
- **AI Intelligence:** Google Gemini (for natural language understanding, contextual Q&A, and response evaluation).
- **Backend & Storage:** Firebase (for authentication, database, file storage, and hosting).

8.2 System Architecture:

a) Frontend (Next.js)

- UI/UX Layer
 - Candidate Dashboard → Displays available mock interview types (HR, Technical, Behavioral).
 - Interview Room → Provides real-time voice chat with the AI agent.
 - Feedback Reports → Shows detailed analysis of performance.
- SSR/SSG → Faster loading, SEO optimization for platform.
- API Integration → Connects with Firebase (for user data) and VAPI (for voice agent).
- b) AI Voice Agent Layer (VAPI + Gemini)
 - Voice Processing
 - VAPI handles speech-to-text (STT) and text-to-speech (TTS) conversion.
 - Candidate speaks naturally; VAPI converts audio into structured text.
 - Contextual AI Responses
 - The transcribed text is passed to Google Gemini.
 - Gemini interprets intent, interview context, and generates appropriate responses/questions.
 - Feedback Capture
 - AI records speech tone, hesitation, filler words, and sentiment.
 - These features are logged for later analysis.
- Backchanneling
 - Provides real-time conversational flow (e.g., “Okay, tell me more,” “Interesting, can you elaborate?”).
- c) Backend (Firebase)
 - Authentication
 - Firebase Auth (Email/Password, Google Sign-In).

- Manages different roles: Candidate, Admin, Recruiter (optional).
- Database (Firestore)
 - Stores user profiles, mock interview history, feedback reports, and AI evaluation scores.
- File Storage (Firebase Storage)
 - Saves audio recordings and transcripts of mock interviews.
- Hosting (Firebase Hosting)
 - Serves the Next.js frontend and API endpoints.

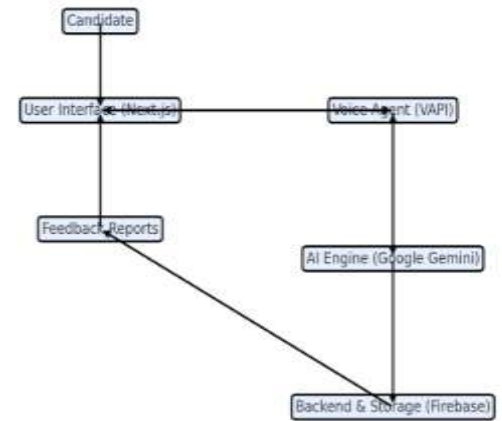


Fig 1: Simple Workflow Model

8.3 Detailed Workflow

1. User Authentication & Setup

- Candidate signs up/logs in.
- Selects desired mock interview type (Technical, HR, Behavioral, Case Study).

2. Interview Initialization

- Next.js frontend establishes connection with VAPI for real-time voice.
- Gemini is primed with interview context (e.g., “Act as a software engineering interviewer for a backend developer role”).

3. Real-Time Interview Flow

- Candidate speaks → Audio captured → VAPI converts speech to text.
- Gemini processes the question/answer → Generates next question or follow-up.
- VAPI converts Gemini’s text back into speech → Candidate hears the interviewer.
- This cycle repeats until interview ends.

4. Data Collection During Interview

- Audio recordings saved in Firebase Storage.
- Transcripts stored in Firestore.
- Metrics logged: speaking speed, clarity, filler words, tone sentiment.

5. Post-Interview Analysis

- Gemini analyzes transcript for content quality:
 - Technical accuracy.
 - Depth of explanation.
 - Structure of answers.
- AI evaluates communication skills:
 - Confidence (based on speech patterns).
 - Clarity & coherence.
 - Sentiment/emotional stability.
- A feedback report is generated and stored in Firebase.

6. Feedback Report Delivery

- Candidate dashboard displays:
 - Strengths & Weaknesses.

- Score Breakdown (Technical Knowledge, Communication, Confidence).
- Recommendations for improvement.
- Option to replay interview audio and read transcript.

8.4 Additional Features (Optional Enhancements)

- Adaptive Questioning:
 - Gemini adjusts difficulty based on user's performance during the interview.
- Gamification:
 - Badges, progress levels, and streaks for motivation.
- Recruiter/Admin Panel:
 - Recruiters can upload custom interview question sets.
 - Admins can monitor system performance and analytics.
- Multilingual Support:
 - VAPI + Gemini allows interviews in multiple languages.

8.5 Security & Privacy Considerations

- All audio and transcripts stored securely in Firebase with access control.
- GDPR/Privacy compliance for candidate data.
- Option for candidate to delete interview history.

8.6 Theoretical Tech Stack Justification

- Next.js → Best for building scalable, responsive web apps with SSR.
- VAPI → Enables real-time, natural voice-based interaction (key differentiator).
- Google Gemini → Provides advanced NLP and contextual AI reasoning.
- Firebase → Reduces backend complexity, offers reliable cloud database, authentication, and hosting.

Workflow Summary

- User Login → Candidate signs in via Firebase Authentication.
- Interview Selection → Chooses type (HR, Technical, Behavioral).
- Interview Start →
 - Voice captured → VAPI converts speech to text.
 - Gemini processes and generates interview questions/responses.
 - VAPI converts AI text back to speech → Candidate hears interviewer.
 - Interaction continues in real time.
- Data Logging → Audio, transcripts, and interaction metrics saved in Firebase.
- Post-Interview Analysis →
 - Gemini evaluates answers (accuracy, structure, depth).
 - Communication and confidence assessed (tone, filler words, speed).
- Feedback Delivery → Report with strengths, weaknesses, scores, and recommendations shown on candidate dashboard.

Key Characteristics

- Real-time Voice Interaction → AI interviewer speaks and listens naturally.
- AI-driven Evaluation → Gemini analyzes both technical and soft skills.
- Personalized Feedback → Detailed reports for performance improvement.

- Scalable Cloud Backend → Firebase handles auth, storage, and hosting.
- Adaptive Interviewing → Questions adjust based on candidate's performance.
- Data Security → Controlled access and option to delete interview history.
- User-Centric Design → Accessible via a responsive Next.js web interface.

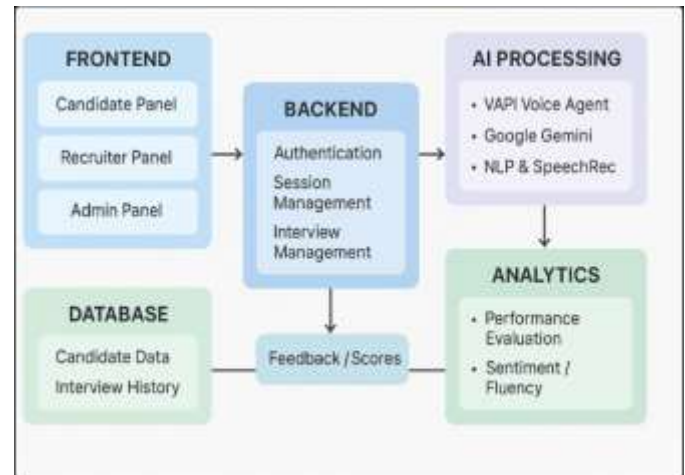


Fig 2: System Architecture

9. RESULTS AND DISCUSSION

9.1 Results

The implementation of the Mock Interview Platform with a Real-Time AI Voice Agent successfully demonstrates how artificial intelligence, natural language processing, and cloud technologies can be combined to simulate real interview environments. The system was able to:

- Enable Seamless Real-Time Interaction – Through VAPI, candidates experienced fluid two-way communication, closely resembling real interviewer behavior. Speech-to-text and text-to-speech conversions functioned effectively, ensuring minimal delay during conversations.
- Provide Intelligent Questioning – Google Gemini adapted to user responses by generating contextually relevant follow-up questions, making the interview process dynamic rather than pre-scripted.
- Capture and Store Comprehensive Data – Firebase handled authentication, session management, audio storage, and transcript saving efficiently, providing a secure and scalable backend.
- Generate Detailed Feedback Reports – Post-interview reports highlighted technical accuracy, communication style, confidence indicators, and areas of improvement, offering personalized insights for each candidate.
- Support Multiple Interview Modes – HR, Technical, and Behavioral interviews were supported, giving candidates a variety of practice scenarios.

9.2 Discussion

The results highlight the potential of AI-driven platforms to transform traditional mock interview practices:

- Enhanced Accessibility and Scalability
 - Unlike human-led mock interviews, this platform is available anytime, anywhere, and can serve unlimited candidates simultaneously.
 - It reduces dependency on professional interviewers and lowers the overall cost of preparation.

- Realism of the Interview Experience
 - The real-time voice interaction provided by VAPI created an immersive, human-like experience. Candidates reported that the natural conversational flow improved their ability to respond under pressure, similar to real interviews.
 - The adaptive questioning mechanism ensured interviews did not feel repetitive but instead matched the candidate's level of knowledge and confidence.
- Comprehensive Skill Evaluation
 - Traditional practice often focuses only on technical accuracy. In contrast, this platform assessed both technical competence and soft skills, including tone, pacing, and communication clarity.
 - This dual assessment provides candidates with a holistic understanding of their performance.
- Data-Driven Improvement
 - The structured feedback reports allowed candidates to identify specific weaknesses (e.g., overuse of filler words, lack of detail in answers).
 - Continuous use of the platform supports iterative learning, enabling measurable progress over time.
- Limitations and Considerations
 - Latency Sensitivity: Although interaction was mostly smooth, occasional delays in speech processing could affect flow.
 - Contextual Depth: While Gemini provided intelligent questioning, domain-specific interviews (e.g., advanced technical coding rounds) might require additional dataset fine-tuning.
 - Human Nuance: Subtle non-verbal cues (body language, facial expressions) were not captured, which are often crucial in real interviews.
- Future Enhancements
 - Integration of emotion detection through voice sentiment analysis for deeper feedback.
 - Support for multilingual interviews to expand accessibility globally.
 - Adding recruiter dashboards where professionals can design and monitor custom interview scenarios.

Conclusion of Discussion

The project establishes that AI-powered mock interview platforms are not just theoretical but practical, effective, and scalable. By combining Next.js, VAPI, Google Gemini, and Firebase, the system bridges the gap between academic preparation and real-world interviewing. It empowers candidates with realistic practice, actionable feedback, and confidence-building opportunities, making it a valuable tool for modern job seekers.

10. CONCLUSIONS

The Mock Interview Platform with a Real-Time AI Voice Agent demonstrates how modern AI and cloud technologies can transform the way candidates prepare for interviews. By leveraging Next.js for an interactive and scalable user interface, VAPI for natural and fluid voice-based interactions, Google Gemini for intelligent question generation and response evaluation, and Firebase for reliable authentication, storage, and hosting, the project achieves its goal of delivering a realistic, accessible, and data-driven mock interview experience.

One of the most significant outcomes of this project is its ability to replicate the pressure and flow of real interviews while remaining fully automated and available on demand. Unlike traditional mock interview methods that require human interviewers and are limited by cost and availability, this

platform ensures scalability, affordability, and accessibility for a diverse range of candidates.

The detailed feedback reports generated by the system not only highlight technical knowledge but also address communication style, confidence, and clarity, aspects often overlooked in conventional practice. This holistic evaluation empowers candidates to understand their strengths and areas for improvement, leading to progressive learning and measurable performance gains over time.

Furthermore, the adaptive questioning mechanism and real-time voice simulation contribute to an immersive learning environment that goes beyond static, text-based preparation platforms. While challenges remain—such as occasional latency, limitations in capturing non-verbal cues, and the need for domain-specific fine-tuning—the results underscore the potential of AI-driven platforms to bridge the gap between academic learning and real-world professional expectations.

In conclusion, this project successfully establishes a scalable, intelligent, and user-centric mock interview solution. It not only enhances the interview preparation journey for candidates but also lays the foundation for future advancements, such as multilingual support, recruiter-customized interviews, and integration of emotional intelligence metrics. With further development, this platform can evolve into a powerful tool that significantly impacts how individuals prepare for and succeed in competitive job markets.

ACKNOWLEDGEMENT

We would like to sincerely thank our faculty advisors and mentors for their continuous guidance, support, and valuable feedback throughout the course of this project. Their expertise and insights have been instrumental in shaping our research approach, methodology, and overall system design. We are also grateful to our colleagues and teammates who contributed to data collection, annotation, model training, and testing, helping us develop a robust and reliable theft detection system.

We would like to acknowledge the valuable contributions of the open-source community and the developers of modern frameworks and tools such as Next.js, VAPI, Google Gemini, and Firebase, which provided a strong foundation for implementing and integrating the core functionalities of our project. Their continuous innovation and documentation enabled us to build an effective, scalable, and interactive system. Finally, we extend our gratitude to all individuals who supported the conceptual validation and shared their feedback during testing, offering practical insights that helped refine the platform and enhance its overall effectiveness in real-world scenarios.

REFERENCES

- [1] S. M. Patil, K. V. Shinde, B. G. Vakare, and S. S. Dunbale, "AI Powered Mock Interview Platform," *International Journal of Innovative Research in Management, Engineering and Technology (IJIRMP)*, vol. 13, no. 1, Jan.–Feb. 2025.
- [2] F. S. Ghodichor, P. Dandage, A. Jadhav, A. Shahane, and P. Mohite, "AI-POWERED MOCK INTERVIEW EVALUATOR," *Journal of Emerging Technologies and Innovative Research (JETIR)*, vol. 11, no. 11, pp. a931–a933, Nov. 2024.
- [3] J. V. Barpute, O. Wattamwar, S. Pakjade, and S. Diwate, "A Survey of AI-Driven Mock Interviews using GenAI and

Machine Learning (InterviewX),” in Proc. 4th Int. Conf. on Ubiquitous Computing and Intelligent Information Systems (ICUIS), 2024, pp. 217–219.

[4] R. Umbare, A. Budhodkar, S. Suryawanshi, S. Udgirkar, P. Aher, and S. Rangdale, “From Practice to Perfection: AI-Driven Mock Interviews for Career Success,” in Proc. Int. Conf. on Sustainable Communication Networks and Applications (ICSCNA), 2024, pp. 1250–1252.

[5] P. K. Mishra, T. M. L., A. K. Arulappan, G. R. G., I.-H. Ra, and Y.-S. Lee, “AI-Driven Virtual Mock Interview Development,” in Proc. Joint 13th Int. Conf. on Soft Computing and Intelligent Systems & 25th Int. Symp. on Advanced Intelligent Systems (SCIS & ISIS), 2024.

[6] J. M. C. J., G. Baburaj, M. Sabi, S. S., and M. Benson, “Q&AI: An AI powered Mock Interview Bot for Enhancing the Performance of Aspiring Professionals,” in Proc. Int. Conf. on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI), 2024.