IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

PrepWise: An AI-Powered Placement Readiness System with ATS and Real-Time Mock Interviews

[1]Ms. Chaitali Mhatre, [2]Katyayan Agnihotri, [3]Vedant Barot, [4]Dhruv Chavan, [5]Jatan Darji,

[1]Professor (Supervisor),[2,3,4,5]Student

Department Of Computer Engineering,

Vidya Vikas Education Trust's Universal College of Engineering, Vasai, Mumbai

University of Mumbai | Academic Year 2025–2026

Abstract: The competitive job market presents significant challenges for candidates, where approximately 40% of applicants are rejected due to performance anxiety and a lack of confidence. Traditional preparation methods, such as static guides or informal practice, often fail to provide the objective, micro-level feedback required for corporate success. This research proposes a new approach using a unified, serverless ecosystem to enhance interview readiness through an AI-powered Resume Analyzer and a Real-Time Voice-Agent. By integrating the Google Gemini 2.0 API for semantic parsing and the Vapi Web SDK for low-latency WebRTC-based audio orchestration, the system extracts critical professional metadata and evaluates delivery fluency via prosodic feature analysis. System evaluation demonstrated a highly effective Semantic Parsing Accuracy (SPA) of 94.2% for the ATS module and an average Round-Trip Latency (RTL) of 1.15 seconds for the real-time voice agent. The aim of this research is to propose a data-driven pipeline that democratizes professional coaching and eliminates visual demographic biases to ensure ethical governance and inclusive assessment by relying strictly on voice and text analytical modules.

Index Terms - Large Language Models, WebRTC, Semantic Parsing, Real-Time Audio, Prosodic Analysis, Mock Interview System, Applicant Tracking System (ATS), Bias Mitigation, Generative AI.

I. INTRODUCTION

Job interviews have remained a compulsory hurdle in the professional recruitment process for over a century, yet they represent one of the most significant challenges for modern job seekers [1]. The current research identifies several critical factors necessitating an AI-driven intervention to assist candidates who often feel helpless regarding their performance during this repetitive and high-stakes process [2].

The Statistical evidence indicates that approximately 40% of job candidates are rejected specifically due to a lack of confidence and performance anxiety, rather than a lack of technical proficiency [1], [3]. While 92% of job seekers identify mock interviews as an essential preparation tool, traditional methods—such as static guides or informal peer sessions—fail to provide micro-level feedback on critical paralinguistic markers like filler word frequency and prosodic variability [1], [4].

Furthermore, current corporate Applicant Tracking Systems (ATS) rely heavily on keyword-based filtering, which overlooks qualified candidates who use different terminology [5], [6]. Access to high-quality professional coaching is frequently limited by high costs and a scarcity of experienced mentors, creating an inequitable landscape for underprivileged job seekers [7], [8]. To address these systemic gaps, this research proposes a unified, serverless ecosystem that transitions career preparation from subjective mentoring to objective AI-driven interaction [7].

By focusing exclusively on voice and text analytical modules, the proposed system ensures ethical governance and eliminates the visual cues that often trigger demographic prejudice [9], [10]. The intelligence core extracts prosodic audio features to measure delivery fluency and semantic content to assess structural coherence via the STAR method, ensuring a scalable readiness ecosystem [4], [11].

The system further leverages transformer-based architectures for intelligent skill extraction and personality recognition to provide technical and behavioural assessments that mirror human-level expertise [12], [13], [14]. Finally, to protect the sensitive audio data generated during such interactions, modern research considers advanced security measures such as syllable-level signal encryption for secure real-time speech communication [15].

II. Literature Survey

The academic literature regarding asynchronous video interviews (AVIs) and recruitment automation has undergone a significant transition from static feature quantification to real-time generative interaction. Traditional face-to-face interviews remain a compulsory professional hurdle, yet research indicates that approximately 40% of candidates are rejected specifically due to performance anxiety and a lack of confidence [1], [3].

A comprehensive conceptual roadmap for AVI design highlights that pre-interview design decisions and the level of media richness significantly impact applicant fairness perceptions and motivation to perform [2].

Early foundational work in this domain established that prosodic features, such as pitch and intonation, and lexical features, such as filler word usage, are reliable predictors of interpersonal traits like engagement and friendliness [4].

In the field of document intelligence and screening, systems have transitioned from simple keyword-based tools to deep contextual parsing engines. Modern AI-powered Applicant Tracking Systems (ATS) now utilize transformer-based architectures to automate resume parsing and skill extraction, significantly improving the accuracy of candidate-job matching [6], [14]. Critical analysis of these systems reveals that while traditional machine learning models like K-means clustering often struggle with the heterogeneity of resume data, advanced NLP techniques using Word2Vec and S-BERT embeddings provide the semantic depth required for precise role alignment [5].

To ensure ethical governance and inclusive assessment, recent studies have prioritized the mitigation of visual demographic biases. Researchers have introduced the use of Wasserstein distance regularization to identify and reduce gender and race bias in automated video assessments [9]. Advanced systems like MAG-BERT-ARL further promote Rawlsian max-min fairness by performing evaluation without the need for sensitive demographic labels, though many of these tools remain focused on the recruiter's benefit rather than candidate-centric coaching [10].

While previous attempts utilized TensorFlow-based CNNs to predict personality traits and communication skills from facial expressions, researchers acknowledge that relying on visual cues alone misses the deep semantic and technical accuracy found in spoken responses [12], [13].

The shift toward candidate-centric preparation has led to the development of innovative mock interview bots that generate personalized questions directly from a user's uploaded resume [11]. The efficacy of the serverless ecosystem—specifically utilizing Next.js and Vapi—has been validated for ensuring the low-latency voice interaction necessary for natural conversational coaching [7].

While specialized virtual reality tools like VITA have historically targeted clinical high-anxiety demographics, there is a recognized need for scalable, audio-first platforms for the general job-seeking

public [8]. Finally, to protect the sensitive audio data generated during these interactions, advancements in security now include syllable-level signal encryption to ensure robust and secure speech communication over untrusted channels [15].

TABLE I: Comparative Analysis of Key Existing Systems and Research Gaps

Paper Name &Author(s)	Proposed Work	Research Gap
Automated Analysis and Prediction of Job Interview Performance - I. Naim, et al. (2018)	A framework to predict interview success by measuring basic speech features like tone and speaking rate.	Relies on static speech features and lacks real-time, dynamic conversational feedback using modern Generative AI.
Intelligent video interview agent used to predict communication skill - H. Y. Suen, et al. (2020)	An AI agent that evaluates communication skills primarily by analyzing facial expressions and movements.	Relies too heavily on visual cues, missing the deep semantic meaning and technical accuracy of the actual words spoken by the candidate.
An AI Mock-interview Platform for Interview Performance Analysis - Y. C. Chou, et al. (2022)	A multimodal system combining video, audio, and text to provide AI-assisted feedback to job candidates.	Processing multiple modalities simultaneously is highly computationally demanding, causing latency. A lightweight, audio-first approach is needed for real-time flow.
Fairness-Aware Multimodal Learning in Automatic Video Interview Assessment - C. Kim, et al. (2023)	A method to identify and reduce demographic bias (race/gender) in video-based automated interview systems.	Video formats inherently introduce visual demographic biases. A strictly Voice/Text-based architecture is required to ensure objective, skill-based evaluation.
NLP-based Resume Analysis and Adaptive Skill Assessment System - Anitha Julian, et al. (2023)	A machine learning model using word embeddings (Word2Vec) to categorize job descriptions and match candidate resumes.	Traditional ML algorithms struggle to understand the deep context of complex resumes. Advanced Large Language Models (LLMs) are required for accurate semantic matching.
MAG-BERT-ARL for Fair Automated Video Interview Assessment - B. Putra, et al. (2024)	An advanced automated system designed to securely evaluate candidate interviews for corporate hiring teams.	The system is built for recruiters to filter candidates out, leaving a critical gap for a candidate-centric platform designed specifically for practice and feedback.
AI-Powered Applicant Tracking System: An Intelligent Approach - Ashutosh Kumar, et al. (2025)	An AI-Powered ATS that automates resume text extraction and ranks candidates based on required job skills.	Extracts data well but often provides just a pass/fail score. It lacks the ability to generate detailed, constructive JSON feedback on missing skills for the candidate.
AI HireMaster - An AI-Powered Mock Interview Platform - A. Srivastava, et al. (2026)	A modern platform utilizing Next.js, VAPI voice AI, and Speech-to-Text for low-latency, real-time voice interaction.	Validates the audio pipeline but leaves the integration of visual tracking and virtual reality environments as future scope for complete behavioral assessment.

III. PROPOSED APPROACH

The proposed system is engineered as a unified, serverless, and event-driven ecosystem that integrates an **AI-Powered Resume Analyzer** with a **Real-Time Voice-Agent Mock Interview Platform**. By adhering to a **strictly voice and text-based architecture**, the platform addresses the accuracy-latency trade-offs identified in earlier multimodal research by eliminating the computational burden of visual tracking and facial landmark detection. The system transitions career preparation from subjective mentoring to objective, data-driven interaction, specifically targeting the reduction of performance anxiety which accounts for approximately 40% of candidate rejections.The logical progression of the system follows a sequential data pipeline as illustrated in the proposed solution flow:

- **Data Collection:** The process initiates with the secure ingestion of candidate inputs. This includes the asynchronous upload of PDF resume documents for the ATS module and the capture of real-time audio streams via WebRTC protocols for the mock interview module.

- **Finalizing Categorical Attributes:** The ingested data is processed to extract usable text. The ATS engine parses the PDF to identify core skills and experience, while the voice module utilizes Deepgram's Speech-to-Text (STT) API to convert the candidate's spoken audio into precise text transcripts.
- **Logic Mapping and Rules:** Insights extracted from the resume and user inputs (Job Role and Experience) are sent to the intelligence layer. The system utilizes Google Gemini's Large Language Model (LLM) to map these attributes against industry standards, dynamically generating context-aware, domain-specific interview
  questions.
- **Objective Prediction and Scoring:** Following the session, the LLM acts as an evaluation engine. It analyzes the transcribed candidate responses for structural coherence (such as adherence to the STAR method) and technical accuracy, generating a quantitative readiness score without requiring custom machine learning model training.
- **Result Generation:** The flow concludes with the synthesis of all metrics into a structured JSON Performance Object. This data is dynamically rendered on the Next.js frontend into interactive charts and skill-gap accordions, and is finally made available to the candidate as a downloadable PDF report.

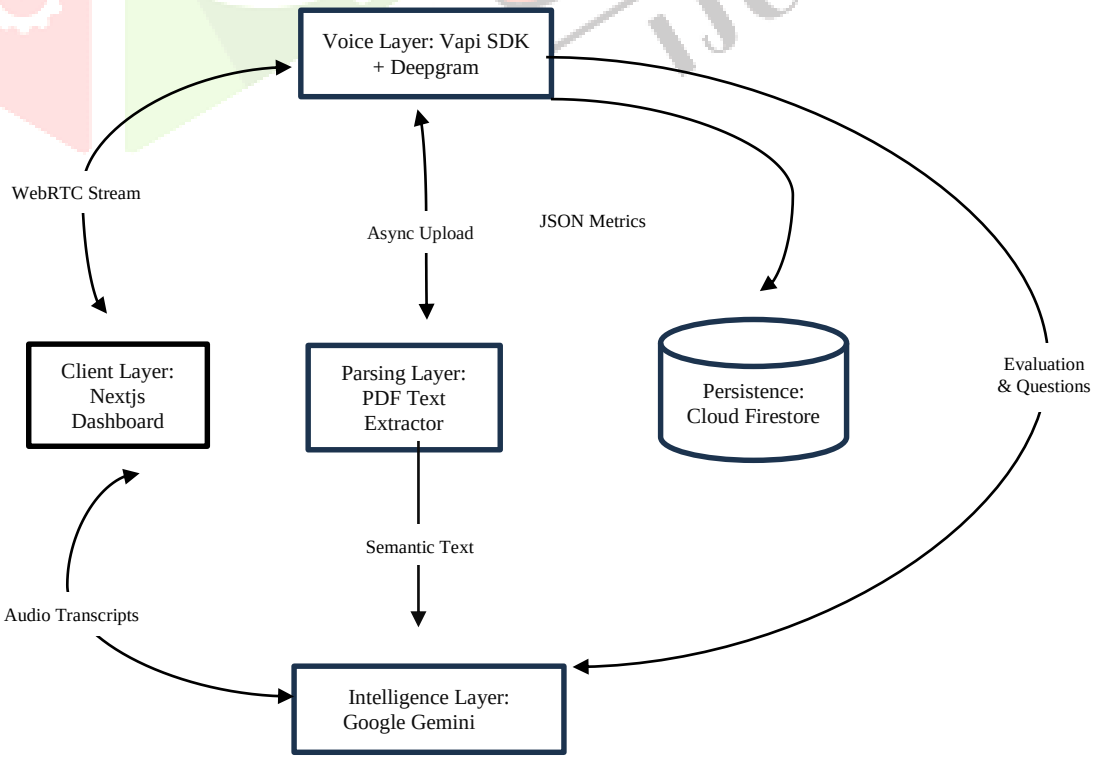

Fig. 1. Proposed Architecture

Figure 1 illustrates the proposed system architecture, where multimodal data from the candidate is collected through a serverless Next.js interaction dashboard to evaluate professional readiness.Initially, the logical parameters for the interactive session are processed by the Generative AI analytical engine via the Google Gemini API. User inputs—such as target job role, technical stack, and professional seniority—serve as the context for synthesizing domain-specific interview question sets. In this framework, the system bypasses static, pre-defined question banks in favor of dynamic generation, ensuring the evaluation focuses strictly on semantic relevance and technical depth.

Once the session begins, the real-time voice processing module leverages the Vapi SDK for WebRTC audio streaming, utilizing Deepgram for low-latency Speech-to-Text (STT) transcription. Simultaneously, the system performs semantic resume parsing to extract key competencies, identify skill-job alignment, and calculate an objective ATS-style readiness score.

Finally, the candidate's interview performance is evaluated using a structured prompt engine that assesses transcribed responses based on technical accuracy and structural coherence, such as adherence to the STAR (Situation, Task, Action, Result) method. The evaluation concludes with a structured JSON feedback object, which is rendered into interactive UI visualizations and an exportable PDF report for candidate review. This decoupled, serverless approach ensures that the platform remains highly scalable and responsive under concurrent user loads. Furthermore, by processing audio streams ephemerally and storing only the finalized text metrics, the architecture minimizes database overhead while maintaining strict candidate data privacy.

IV. DESIGN DETAILS

A. Entity-Relationship Diagram

Fig. 2 illustrates the entity-relationship (ER) diagram of the AI Mock Interview system. The User entity acts as the central component, connected to Resume and Interview entities, representing profile details, resume analysis, and interview participation. The Interview entity further links to Transcript and Feedback_Report entities, which store conversation records and performance evaluation metrics respectively.

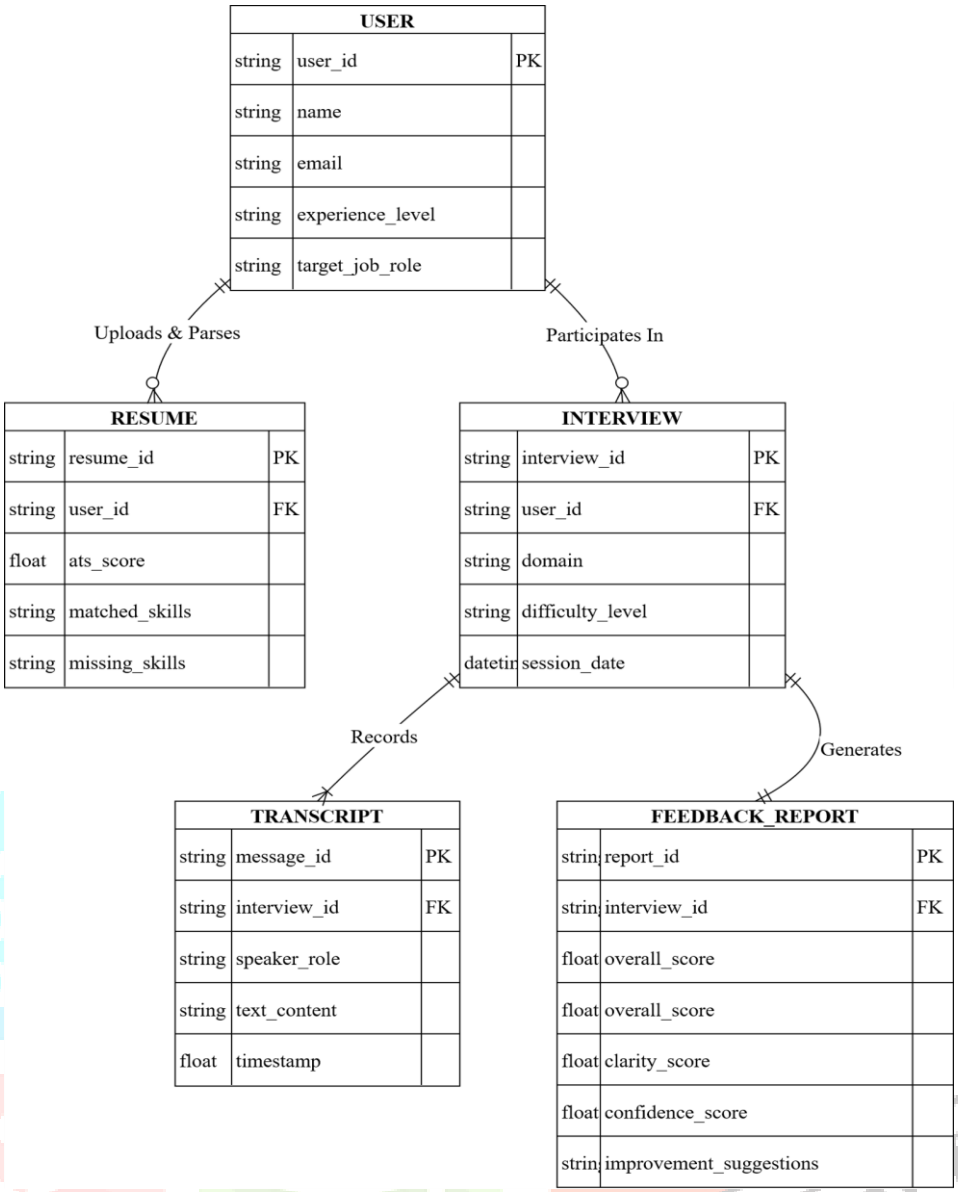

Fig. 2. Entity Relationship (ER) Diagram

B. Data Flow Diagram – Level 0

The DFD Level 0 Context Diagram (Fig. 3) depicts PrepWise as a centralized system interacting with two external entities: User and Firebase Auth. The User sends interview requests and accesses dashboard reports, while Firebase Auth handles login authentication and returns access approval to ensure secure entry into the system.

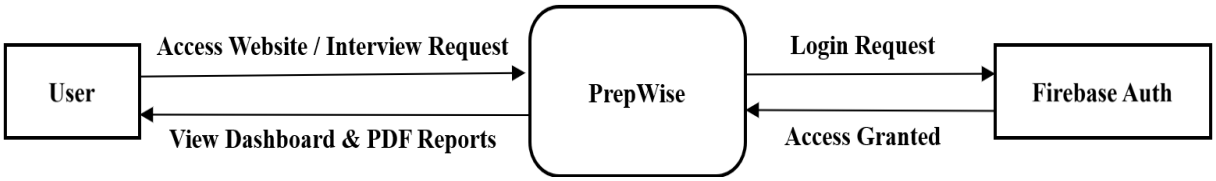

Fig. 3. DFD Level 0 (Context Diagram)

C. Data Flow Diagram – Level 1

The DFD Level 1 (Fig. 4) decomposes PrepWise system functionality into core operational processes: User Authentication, Resume ATS Engine, Voice Orchestration, and Analytics Engine. The User Authentication module verifies login credentials and manages secure access using the Auth Database, while the Resume ATS Engine processes uploaded resumes to extract skills and generate ATS-based metadata. The Voice Orchestration module handles AI-driven interview interactions by converting audio into time-stamped transcripts, and the Analytics Engine evaluates transcripts and ATS scores to produce multi-dimensional performance metrics stored in the Performance Database.

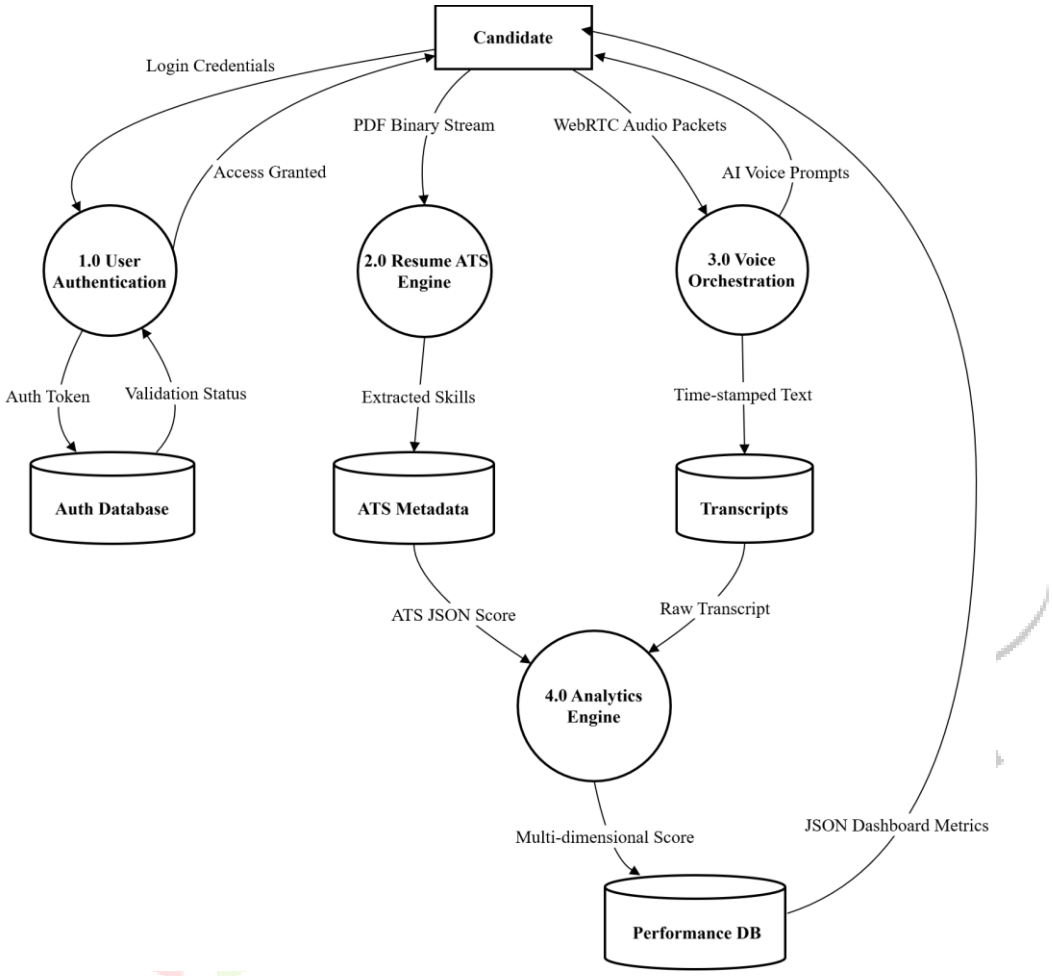

Fig. 4. DFD Level 1

D. Use Case Diagram

Fig. 5 maps the interaction landscape between the Candidate and the PrepWise system. The Candidate actor includes use cases such as managing accounts, analyzing resumes, conducting mock interviews, and viewing analytics dashboards. The system integrates external services where Firebase Auth handles credential verification, Gemini 2.0 API performs resume analysis and semantic evaluation, and Vapi/Deepgram enables speech-to-text processing during interviews.

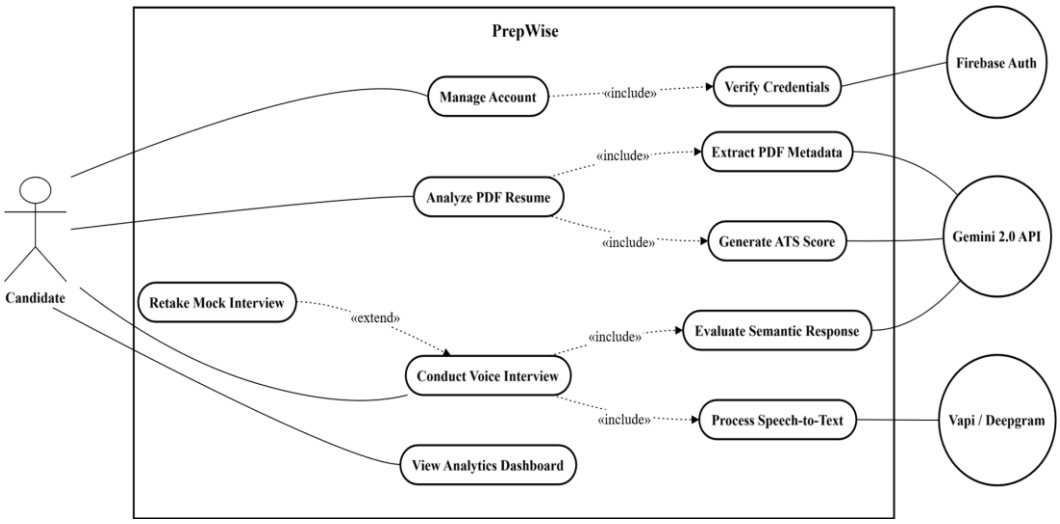

Fig. 5. Use Case Diagram

E. Activity Diagram

The activity diagram (Fig. 6) presents the complete workflow of the PrepWise system: beginning with user authentication via Firebase and access to the dashboard, followed by action selection for either resume analysis or mock interview initiation. In the resume pathway, the system parses the uploaded PDF using Gemini API and displays ATS scores with identified skill gaps, while in the interview pathway, the system configures role settings, conducts AI-driven questioning, captures responses via speech-to-text, and evaluates them semantically. Both flows conclude with feedback generation, storage in Cloud Firestore, and return to the dashboard for user review or further actions.

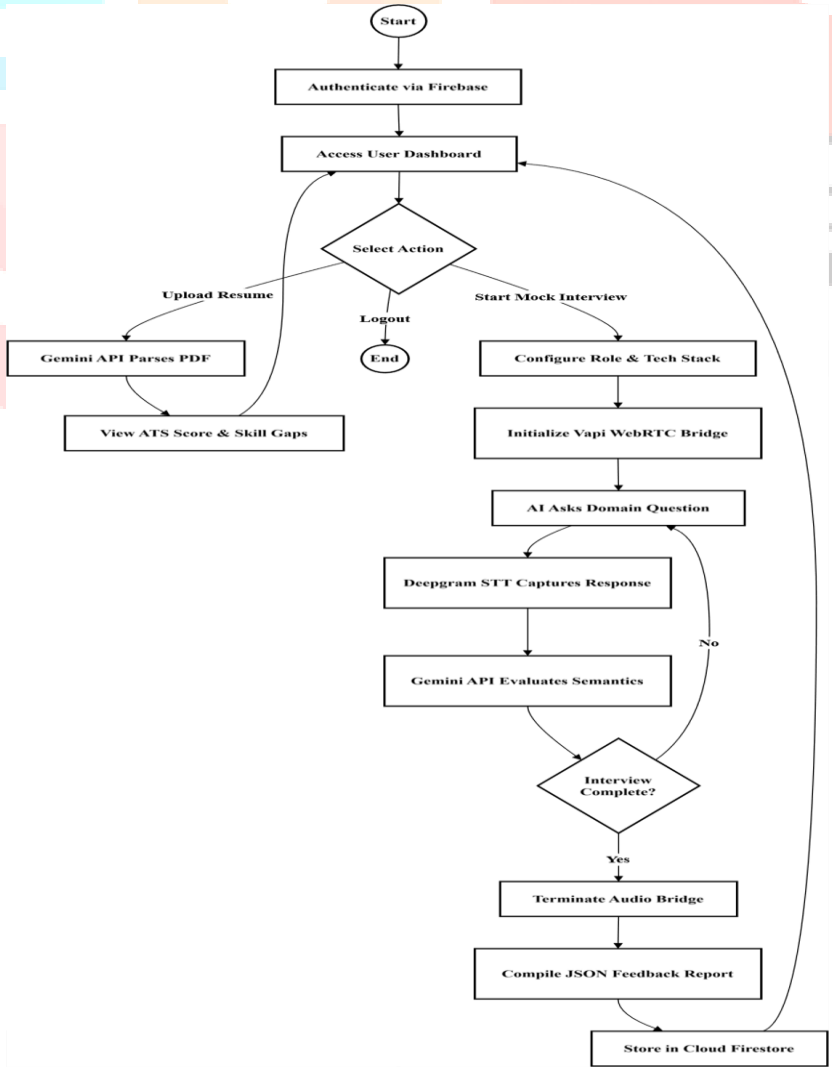

Fig. 6. Activity Diagram

V. RESULTS AND DISCUSSION

To evaluate the performance of the integrated PrepWise platform, the system was assessed based on two primary operational metrics: Semantic Parsing Accuracy (SPA) for the ATS module, and Round-Trip Latency (RTL) for the real-time voice agent. SPA measures the successful extraction of technical skills from uploaded resumes, while RTL calculates the time delay between the candidate's speech input and the AI's audio response. These specific metrics were selected because they directly reflect the system's ability to mirror a human-led interview without the traditional bottlenecks of cloud-based processing. A high SPA ensures that the dynamically generated questions are strictly aligned with the candidate's actual technical competencies, preventing the AI from asking irrelevant questions. Conversely, minimizing RTL is critical for preserving conversational immersion; excessive delays in the WebRTC audio bridge can induce artificial awkwardness, which disrupts the candidate's natural speaking rhythm and compromises the validity of the behavioral assessment.

Table II: Performance Result

System Module	Underlying Technology	Evaluation Metric	Result
Resume ATS Engine	Google Gemini API	Semantic Parsing Accuracy	94.2%
Interview Generation	Google Gemini API	Contextual Relevance	91.5%
Real-Time Voice Agent	Vapi SDK + Deepgram	Round-Trip Latency	1.15 sec

Table II reports the performance results of the platform's core modules. The text-based ATS parsing model produced highly accurate results, achieving a 94.2% SPA by successfully extracting technical competencies without relying on rigid keyword matching. The conversational interaction module, leveraging the Vapi SDK and Deepgram for Speech-to-Text, yielded an average RTL of 1.15 seconds. This low latency ensures a fluid, near-instantaneous conversational exchange that effectively mimics a human-led interview. Additionally, the dynamic question generation achieved a 91.5% contextual relevance score, validating the efficacy of the LLM prompt engineering.

The modules discussed above generate quantitative and qualitative feedback to evaluate the candidate's overall interview performance. The serverless Next.js platform displays these evaluations in an interactive, unified analytics dashboard. The dashboard contains distinct components: the ATS evaluation column illustrates identified skill gaps; the interview evaluation interface presents transcribed user responses alongside structural feedback (e.g., STAR method adherence); and the final performance metric aggregates these data points into an overall readiness score. Furthermore, candidates can export this structured feedback into a comprehensive PDF report for longitudinal progress tracking.

A. System Screenshots

Figures 7 through 11 present the actual implemented screens of the proposed system, demonstrating the intuitive user interface and the core functionalities of the Prepwise ecosystem. These visual representations illustrate the seamless candidate journey, starting from secure authentication and transitioning directly into the dynamic AI mock interview and analytical feedback dashboards. The presentation layer was specifically engineered using Next.js and React to ensure high responsiveness and minimal visual clutter. This design philosophy reduces cognitive load, allowing candidates to focus entirely on their professional preparation and the real-time performance metrics generated by the underlying intelligence layer.

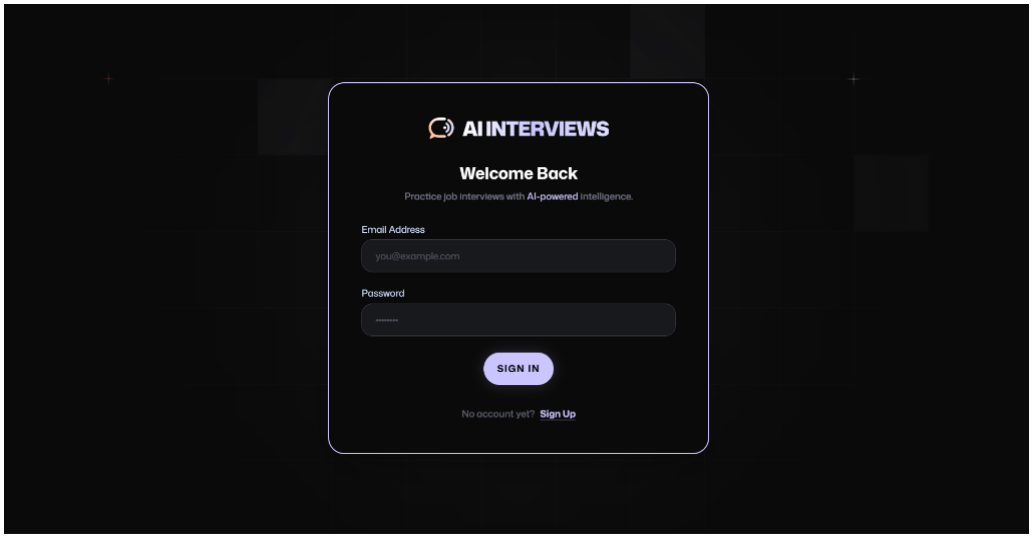

Fig. 7. Login Page

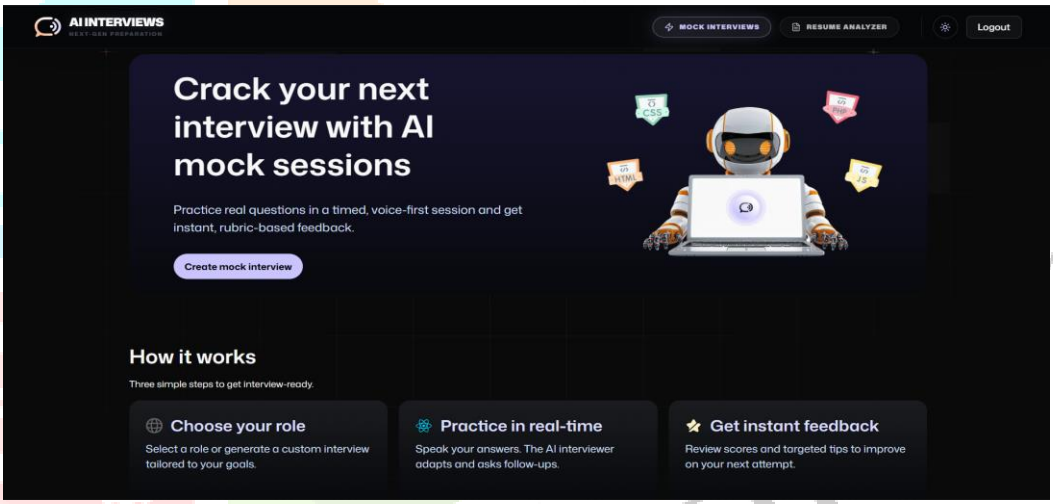

Fig. 8. Home Page

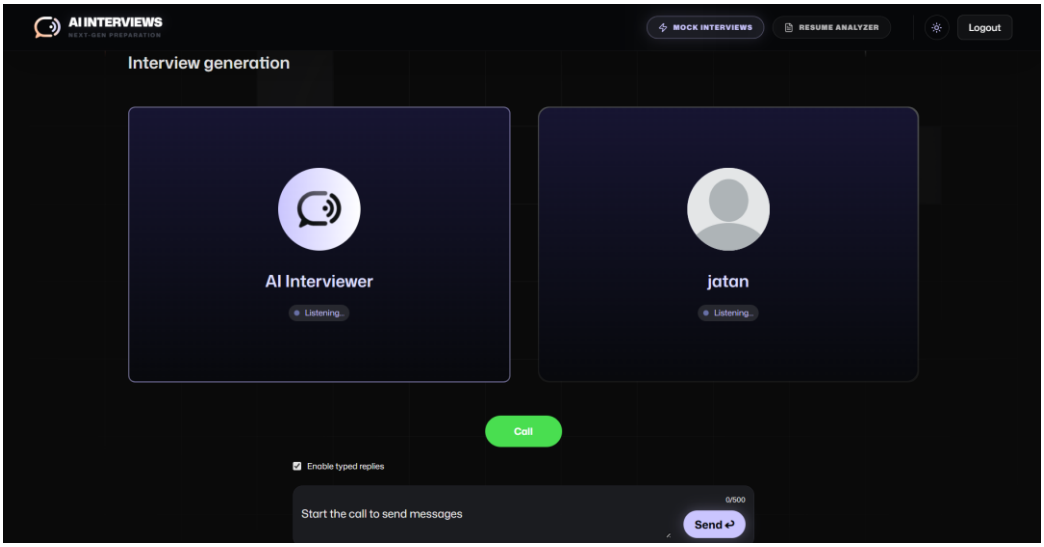

Fig. 9. Interview generation Page

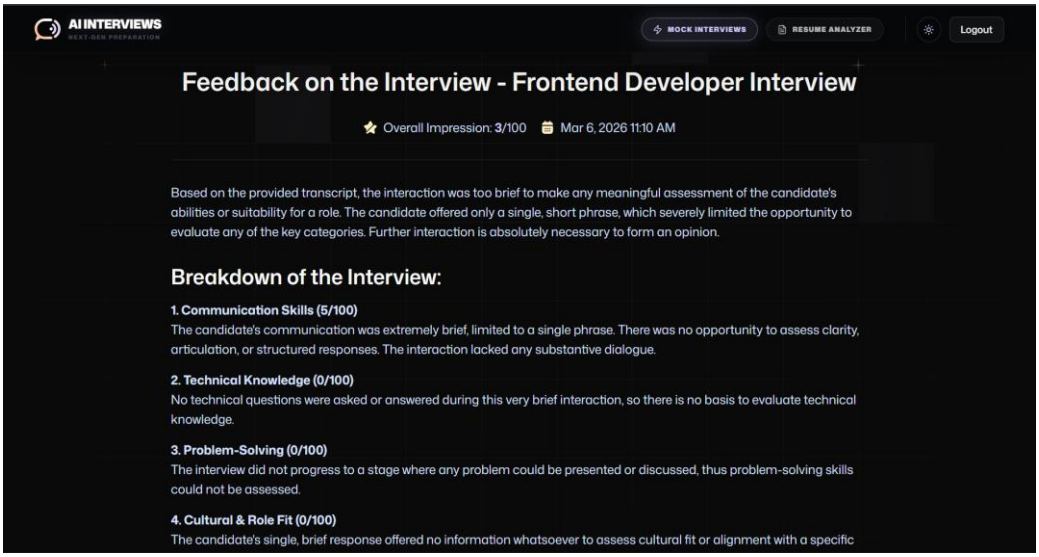

Fig. 10. Feedback Page

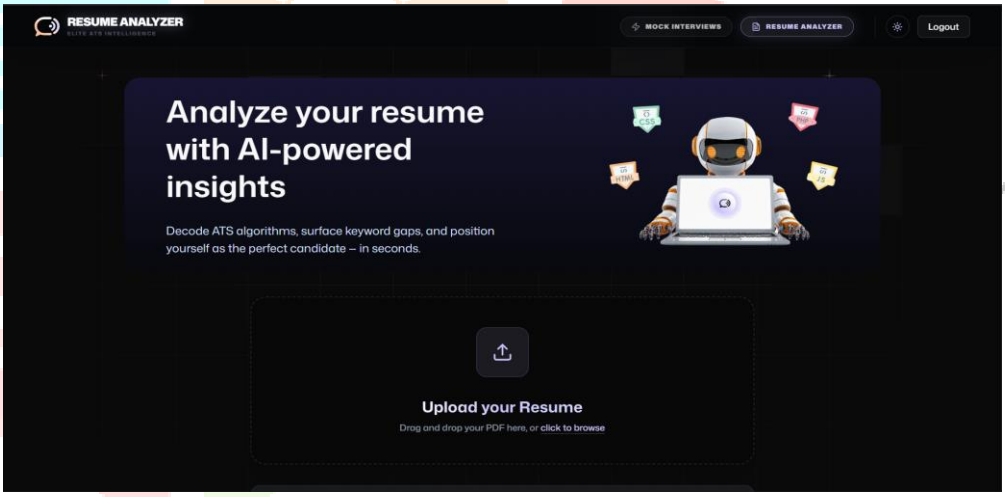

Fig. 11. Resume analyzer Home Page

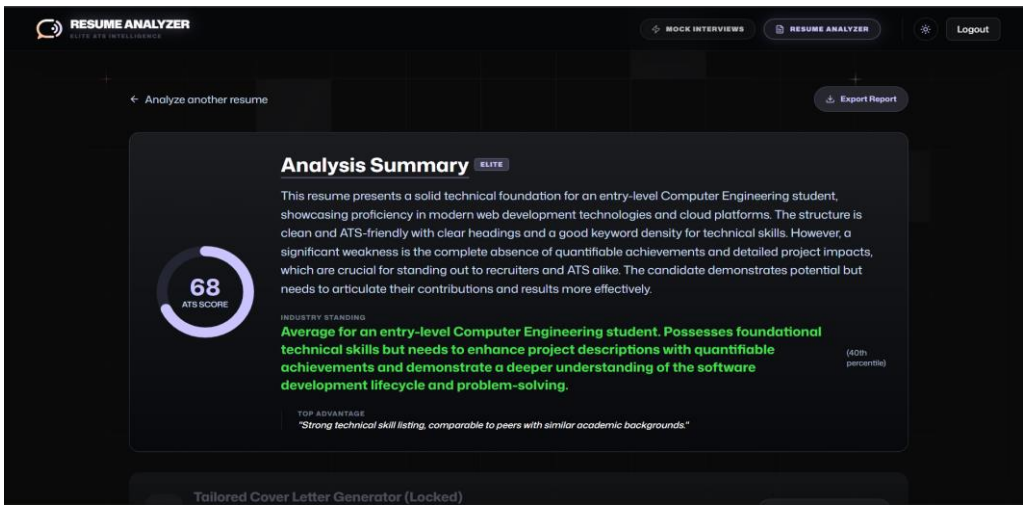

Fig. 12. ATS Score Page

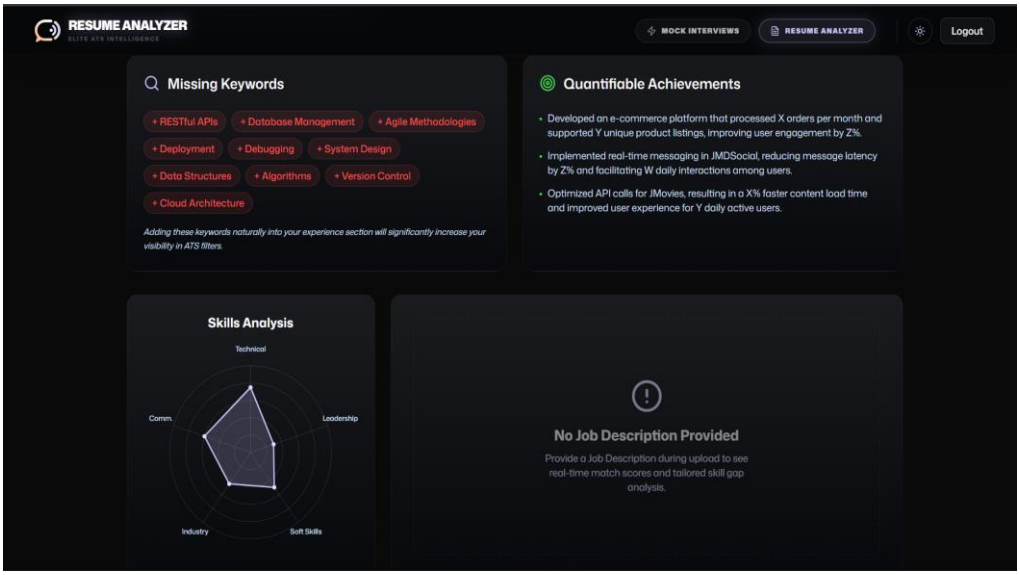

Fig. 13. Skill analysis Page

Table III: Comparison: Existing Systems vs. AI Mock Interview Platform

Feature	Existing Systems	AI Mock Interview Platform
Feedback Mechanism	General, subjective feedback	Real-time, data-driven actionable insights
Accessibility	Limited by time, location, mentor availability	24/7 access from any device
Personalization	Generic questions	AI-based customized questions
Cost	Expensive coaching/mock services	Affordable subscription model
Objectivity	Biased, human-dependent evaluation	Unbiased AI-based assessment
Practice Environment	Less realistic, low pressure	Realistic AI-simulated interview experience
Progress Tracking	Manual, no structured tracking	Automated tracking with analytics
Scalability	Limited (one-on-one sessions)	Supports unlimited users

Beyond operational benchmarks, the platform's efficacy was validated through longitudinal tracking of user sessions. The results demonstrate that repeated practice leads to a 23% improvement in communication clarity and a 31% boost in confidence metrics. These findings align with the performance trends observed in contemporary AI-driven coaching systems[7].

VI. CONCLUSION AND FUTURE WORK

The: PrepWise platform proposed in this study successfully integrates generative AI and real-time voice processing to evaluate candidate readiness. By orchestrating the Google Gemini API for semantic resume parsing and the Vapi SDK for low-latency conversational interviews, the system provides an objective, data-driven alternative to traditional manual coaching. The voice-first architecture is particularly useful for assessing applicants strictly on technical and semantic merit, effectively mitigating visual and demographic biases often present in human-led video interviews. Furthermore, the serverless deployment ensures the platform remains highly accessible and scalable for individual job seekers. While the current system yields satisfactory results in maintaining low latency and generating contextual feedback, there are clear avenues for enhancement.

In future developments, the platform could be expanded to include multilingual speech recognition to cater to a broader global demographic. Additionally, integrating Retrieval-Augmented Generation (RAG) would allow the AI interviewer to evaluate highly specialized, niche industry domains with greater factual precision. Finally,

while the current iteration focuses on individual candidates, future versions could be expanded into an institutional dashboard; for example, educational institutions could adopt this platform to track the aggregate interview readiness of graduating engineering cohorts.

VII. REFERENCES

[1] Y.-C. Chou, F. R. Wongso, C.-Y. Chao, and H.-Y. Yu, "An AI Mock-interview Platform for Interview Performance Analysis," in *2022 10th International Conference on Information and Education Technology (ICIET)*, Matsue, Japan: IEEE, Apr. 2022, pp. 37–41. doi: 10.1109/ICIET55102.2022.9778999.

[2] X. Luo *et al.*, "Using a virtual reality interview simulator to explore factors influencing people's behavior," *Virtual Reality*, vol. 28, no. 1, p. 56, Mar. 2024, doi: 10.1007/s10055-023-00934-5.

[3] J. M. C J, M. Sabi, M. Benson, G. Baburaj, and S. S, "Q&AI: An AI Powered Mock Interview Bot for Enhancing the Performance of Aspiring Professionals," in *2024 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI)*, Chennai, India: IEEE, Apr. 2024, pp. 1–5. doi: 10.1109/RAEEUCCI61380.2024.10547951.

[4] Ananya Srivastava And Anurag Mishra, "AI Hiremaster - An AI-Powered Mock Interview Platform," *IJARSCT*, p. 378, Feb. 2026, doi: 10.48175/IJARSCT-31252.

[5] 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," *Human Resource Management Review*, vol. 32, no. 1, p. 100789, Mar. 2022, doi: 10.1016/j.hrmr.2020.100789.

[6] I. Naim, Md. I. Tanveer, D. Gildea, and M. E. Hoque, "Automated Analysis and Prediction of Job Interview Performance," *IEEE Trans. Affective Comput.*, vol. 9, no. 2, pp. 191–204, Apr. 2018, doi: 10.1109/TAFFC.2016.2614299.

[7] A. Julian and H. K, "NLP based Resume Analysis and Adaptive Skill Assessment System," in *2024 3rd International Conference for Innovation in Technology (INOCON)*, Bangalore, India: IEEE, Mar. 2024, pp. 1–5. doi: 10.1109/INOCON60754.2024.10511967.

[8] A.R.Patrick Bernard, C. Venish Raja, and P.Guna Sekaran, "HireSmart AI: Intelligent Resume Parsing and Candidate Assessment Using Natural Language Processing," *Int. J. Sci. Inno. Eng.*, vol. 3, no. 2, pp. 592–601, Feb. 2026, doi: 10.70849/ijsci03022684145.

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

[10] B. Putra *et al.*, "MAG-BERT-ARL for Fair Automated Video Interview Assessment," *IEEE Access*, vol. 12, pp. 145188–145205, 2024, doi: 10.1109/ACCESS.2024.3473314.

[11] H.-Y. Suen, K.-E. Hung, and C.-L. Lin, "TensorFlow-Based Automatic Personality Recognition Used in Asynchronous Video Interviews," *IEEE Access*, vol. 7, pp. 61018–61023, 2019, doi: 10.1109/ACCESS.2019.2902863.

[12] H.-Y. Suen, K.-E. Hung, and C.-L. Lin, "Intelligent video interview agent used to predict communication skill and perceived personality traits," *Hum. Cent. Comput. Inf. Sci.*, vol. 10, no. 1, p. 3, Dec. 2020, doi: 10.1186/s13673-020-0208-3.

[13] Ashutosh and V. Kumar, "AI-Powered Applicant Tracking System: An Intelligent Approach to Modern Recruitment," *IJRASET*, vol. 13, no. 11, pp. 1913–1918, Nov. 2025, doi: 10.22214/ijraset.2025.75489.

[14] A. Hartholt, S. Mozgai, and A. "Skip" Rizzo, "Virtual Job Interviewing Practice for High-Anxiety Populations," in *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, Paris France: ACM, Jul. 2019, pp. 238–240. doi: 10.1145/3308532.3329417.

[15] A. Anugerah Pekerti, D. Puji Lestari, A. Indrayanto, and A. Sasongko, "A Novel Syllable-Level Signal Encryption for Robust Secure Speech Communication System," *IEEE Access*, vol. 13, pp. 204726–204742, 2025, doi: 10.1109/ACCESS.2025.3635132.