

# Smart AI Interviewer and Resume Analyzer

Pooja Vachkal                      Vishal Chole                      Gaurav Padol                      Samarth Kawane Dept.                      Omkar Kasar  
Dept. of Computer Engg.    Dept. of Computer Engg.    Dept. of Computer Engg.                      of Computer Engg.                      Dept. of Computer Engg.  
JSCOE, Hadapsar Pune,    JSCOE, Hadapsar Pune,    JSCOE, Hadapsar Pune,    JSCOE, Hadapsar Pune,    JSCOE, Hadapsar Pune,  
India                                      India                                      India                                      India                                      India

**Abstract**—Smart Interview & Resume Analyzer (IndusAI) is an AI-powered platform designed to help candidates prepare for industrial job interviews and improve their resumes. The system analyses interview recordings using speech-to-text, natural language processing (NLP), and voice feature extraction to evaluate communication clarity, confidence, and tone. AI models provide sentiment and filler-word analysis, producing a confidence score along with detailed, actionable feedback. A complementary Resume Analyzer module assesses resume structure, relevance, and industry keyword usage for Applicant Tracking System (ATS) compliance. The system automatically generates a professional PDF report highlighting candidate strengths, weaknesses, and improvement suggestions, effectively bridging the gap between academic preparation and real-world hiring expectations.

**Keywords**—Smart Interview Analyzer, Resume Analyzer, Automatic Speech Recognition, NLP, Confidence Score, ATS Compliance, Whisper ASR, BERT, Sentence-BERT, MediaPipe, Job Placement, AI Coaching.

## I. INTRODUCTION

Communication is central to professional success, yet most candidates receive little structured feedback after job interviews. In today's competitive hiring landscape, applicants struggle to objectively evaluate their performance across communication clarity, confidence, and semantic content quality.

Traditional resume formats frequently fail to meet industry standards or pass Applicant Tracking Systems (ATS), which silently filter candidates based on keyword matching before any human recruiter reviews the application, creating an invisible barrier that disproportionately affects students and fresh graduates.

The Smart AI Interviewer and Resume Analyzer (IndusAI) is an intelligent, web-accessible AI coaching platform that bridges this gap. The system simulates mock interviews, evaluates spoken communication across multiple dimensions, assesses resume ATS compliance, and produces detailed actionable feedback in a downloadable PDF report.

IndusAI combines Automatic Speech Recognition (Whisper ASR), deep learning NLP (BERT, Sentence-BERT), voice prosody analysis (openSMILE), optional non-verbal analysis (MediaPipe), and ATS-aware resume evaluation into a single unified multimodal pipeline.

The platform is domain-agnostic and multilingual, suitable for students, graduates, and professionals across disciplines. It requires only a web browser, eliminating hardware dependencies of VR-based simulators and democratising access to quality career coaching.

## II. LITERATURE REVIEW

Recent research in automated interview analysis and resume screening has grown rapidly, driven by advances in transformer-based NLP, multimodal learning, and speech

processing. IndusAI builds on this work by integrating these advances into a unified practical platform.

**Nagasawa et al. (2024)** [IEEE Trans. Affective Computing] proposed an adaptive interview strategy based on speaking willingness recognition for interview robots, showing that real-time affect monitoring can dynamically guide pacing. This directly informs IndusAI's confidence scoring and feedback generation.

**Artiran et al. (2022)** [IEEE Trans. Neural Syst. Rehabil. Eng.] measured social gaze modulation in Autism Spectrum Condition via virtual reality interviews, motivating IndusAI's optional non-verbal gaze and posture analysis using MediaPipe.

**Ashrafi et al. (2023)** [IEEE Access] proposed resume-based career recommendation for rapidly evolving job markets. Their keyword-matching and career-fit analysis techniques directly inspired IndusAI's ATS compliance and resume relevance evaluation modules.

**Stoev et al. (2025)** [IEEE Access] demonstrated that BERT embeddings approach expert linguistic feature sets in automated interview classification, motivating IndusAI's transformer-based semantic content scoring.

**Radford et al. (2023)** [Proc. ICML] introduced Whisper, a transformer ASR trained on 680,000 hours of multilingual audio achieving state-of-the-art transcription. IndusAI uses Whisper for speech-to-text with word-level timestamps for filler-word and speech-rate analysis.

**Reimers & Gurevych (2019)** [Proc. EMNLP] extended BERT with Siamese networks for efficient sentence embeddings. IndusAI applies Sentence-BERT cosine similarity to score candidate answer relevance against expert reference responses.

## III. PROPOSED METHODOLOGY

IndusAI is structured around six primary processing modules. The pipeline accepts an interview video/audio file

and an optional resume document, aggregating all sub-scores into a final Confidence Score and an auto-generated PDF feedback report.

#### A. Input Modu

Accepts interview video/audio files (MP4, WAV, MP3) and optional resume documents (PDF, DOCX). Raw audio is extracted for speech processing and documents are forwarded to the resume analysis pipeline.

#### B. Speech and Audio Analys

Whisper ASR transcribes the recording with word-level timestamps. openSMILE extracts prosodic and spectral audio features for voice confidence sub-scoring. Filler-word frequency and speech rate are derived from the transcript and timestamps.

#### C. Text and NLP Analys

BERT-based models perform sentiment classification and contextual relevance checking. Sentence-BERT generates dense sentence embeddings enabling cosine similarity scoring between candidate answers and reference responses, yielding a semantic content score.

#### D. Non-Verbal Analysis (Optiona

MediaPipe performs real-time face detection, pose estimation, and gaze-direction tracking. Eye contact frequency and posture stability contribute a non-verbal sub-score when video input is provided.

#### E. Resume Analysis Modu

The system parses uploaded resumes and evaluates structure, section completeness, and industry keyword density. ATS simulation checks formatting and keyword alignment against a target job description to predict screening pass/fail likelihood.

#### F. Confidence Score and Repo

All raw features are normalised to [0,1]. Sub-scores  $S_{voice}$ ,  $S_{fluency}$ , and  $S_{content}$  are combined as:  $C = 100 \times (b1 \cdot S_{voice} + b2 \cdot S_{fluency} + b3 \cdot S_{content} + b4 \cdot S_{nv} - Penalties)$ , where  $b1+b2+b3+b4 = 1$ . Score labels: **Very Confident** ( $C \geq 75$ ), **Moderate** ( $50 \leq C < 75$ ), **Needs Improvement** ( $C < 50$ ). A PDF report with scores, charts, and improvement suggestions is auto-generated and immediately downloadable.

### IV. SYSTEM ARCHITECTURE & FEATURES

IndusAI follows a modular, layered architecture for end-to-end interview intelligence, divided into six functional layers covering input, multimodal processing, scoring, output generation, reporting, and deployment scalability.

#### A. Presentation Lay

A responsive web front-end allows candidates to upload recordings and resumes, initiate mock sessions, and view or download results. The interface is designed for clarity and accessibility across desktop and mobile browsers.

#### B. Input Acquisition Lay

This layer accepts video/audio and resume files, extracts raw audio streams for speech processing, and forwards documents to the resume analysis pipeline, supporting MP4, WAV, MP3, PDF, and DOCX formats.

#### C. Multimodal Processing Lay

Speech, text, audio prosody, and optional visual streams are processed in parallel by dedicated sub-modules: Whisper, BERT, Sentence-BERT, openSMILE, and MediaPipe. Results are aggregated into normalised feature vectors for downstream scoring.

#### D. Scoring and Feedback Lay

The confidence score engine combines sub-scores with configurable weights. Penalties for excessive filler words, low ATS compliance, and poor eye contact are applied before final score normalisation and label assignment.

#### E. Report Generation Lay

A templated PDF report is compiled from all scoring outputs, visual charts, and natural-language improvement suggestions generated by the NLP modules.

#### F. Scalability and Deployme

The platform is domain-agnostic and multilingual via Whisper ASR. Future additions include live job-portal API integration, multi-language resume analysis, expanded non-verbal analytics, and an institutional cohort analytics dashboard.

### V. GAP ANALYSIS

A comparative study was conducted between IndusAI and traditional interview preparation methods including human mock interviewers, generic feedback apps, and manual resume reviews. The analysis shows that traditional methods are subjective and non-scalable, while IndusAI delivers objective, automated, multimodal, and instantly available assessment.

TABLE I  
 GAP ANALYSIS: TRADITIONAL VS. INDUSAI

Metric	Traditional Method	IndusAI
Feedback Mode	Manual / subjective	Automated AI scoring
Speech Analysis	Not available	Whisper ASR + prosody
Resume Check	Manual HR review	ATS compliance scan
Non-Verbal Cues	Human observation only	MediaPipe gaze & pose
PDF Report	Verbal / informal notes	Auto-generated PDF
Availability	Scheduled sessions only	On-demand, 24/7
Scalability	Limited by human capacity	Web-based, unlimited

### VI. CONCLUSION

IndusAI successfully presents an intelligent, multimodal interview analysis and resume evaluation platform targeting

industrial job placement. By combining Whisper ASR, BERT, Sentence-BERT, openSMILE, MediaPipe, ATS resume assessment, and automated PDF reporting, the system delivers a holistic and objective view of candidate job-readiness.

*ACM MM*, 2010, pp. 1459–1462.

The project eliminates dependence on expensive human coaching and subjective feedback, improving preparation quality by addressing verbal communication, semantic content, voice confidence, non-verbal cues, and resume ATS compliance in a single unified workflow. IndusAI promotes equal access to quality interview coaching for students, fresh graduates, and professionals regardless of geography or financial means.

Future work will expand non-verbal analysis, add multi-language resume support, integrate live job-portal APIs, develop an institutional cohort dashboard, and conduct large-scale longitudinal user studies to validate long-term placement outcomes.

#### ACKNOWLEDGMENT

The authors sincerely thank their project guide Pooja Vachkal, faculty members, and the Department of Computer Engineering at Jayawantrao Sawant College of Engineering, Hadapsar, Pune (Savitribai Phule Pune University, Academic Year 2025–26) for their guidance and support throughout the development of IndusAI. They also thank classmates and family for constant motivation.

#### REFERENCES

- [1] F. Nagasawa, S. Okada, T. Ishihara, and K. Nitta, "Adaptive Interview Strategy Based on Interviewees' Speaking Willingness Recognition for Interview Robots," *IEEE Trans. Affective Comput.*, vol. 15, no. 2, pp. 230–242, Feb. 2024.
- [2] S. Artiran, R. Ravisankar, S. Luo, L. Chukoskie, and P. Cosman, "Measuring Social Modulation of Gaze in Autism Spectrum Condition With Virtual Reality Interviews," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 30, pp. 2373–2385, Sept. 2022.
- [3] S. Artiran, P. S. Bedmutha, and P. Cosman, "Analysis of Gaze, Head Orientation, and Joint Attention in Autism With Triadic VR Interviews," *Frontiers Virtual Reality*, vol. 5, pp. 1–13, Mar. 2023.
- [4] S. Ashrafi, B. Majidi, E. Akhtarkavan, and S. H. R. Hajiagha, "Efficient Resume-Based Re-Education for Career Recommendation in Rapidly Evolving Job Markets," *IEEE Access*, vol. 11, pp. 124350–124367, Nov. 2023.
- [5] T. Stoev, E. Flemming, B. Strauss, K. Petrowski, C. Spitzer, and K. Yordanova, "Towards Automated Classification of Adult Attachment Interviews in German Language Using the BERT Language Model," *IEEE Access*, vol. 13, pp. 155305–155320, Sept. 2025.
- [6] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust Speech Recognition via Large-Scale Weak Supervision," in *Proc. ICML*, 2023, pp. 28492–28518.
- [7] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proc. NAACL*, 2019, pp. 4171–4186.
- [8] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence Embeddings Using Siamese BERT-Networks," in *Proc. EMNLP*, 2019, pp. 3982–3992.
- [9] C. Lugaresi et al., "MediaPipe: A Framework for Building Perception Pipelines," *arXiv*, 2019, arXiv:1906.08172.
- [10] F. Eyben, M. Wöllmer, and B. Schuller, "openSMILE – The Munich Versatile and Fast Open-Source Audio Feature Extractor," in *Proc.*