



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: <https://doi.org/10.22214/ijraset.2026.81593>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

Neuro-Hire: A Multimodal AI Framework for Objective Behavioral Analytics and Neural Interview Simulation

Preeti Banjare, Auliviya Geetanjali Mukherjee, Shivani Yadav, Aastha Tripathi

Department of Computer Science Engineering Government Engineering College, Bilaspur Chhattisgarh, India

Abstract: *The biggest hurdle for students entering the job market is a lack of technical knowledge—it's a lack of realistic interview practice. Students often fail interviews due to nervousness, poor communication, or the inability to articulate their skills under pressure. In this paper, we break down the design and testing of "Neuro-Hire AI," a web-based mock interview platform that gives students 24/7 access to an automated, intelligent practice environment. Instead of just giving students a list of standard questions, our system uses a dynamic tri-engine architecture: Computer Vision to track facial expressions, Speech-to-Text to measure speaking flow (checking Words Per Minute and filler words), and Google Gemini Model to ask custom technical questions on the spot from a question database and evaluate the answers. Our research shows that combining these different inputs gives students a highly accurate, objective breakdown of their true interview behavior. We show how turning raw video and audio into personalized coaching data proves that premium, realistic interview prep can be made accessible to every student, anytime.*

Index Terms: *Neural Interviewing, Behavioral Analytics, Computer Vision, Generative AI, Multimodal Fusion, Educational Technology, Affective Computing, Mock Interviews.*

I. INTRODUCTION

College students constantly deal with a feedback gap. We spend years learning how to code, but rarely get taught how to handle a high-pressure technical interview. While practicing coding problems is easy, practicing how to communicate those solutions is incredibly difficult. Students usually have to rely on talking to a mirror, recording themselves on their phones, or trying to schedule scarce 1-on-1 time with busy professors or seniors.

We built this project to fix that "Feedback Gap." Most digital mock interview tools today just flash a question on a screen and record the student answering it [1], [11]. It feels incredibly unnatural and provides zero feedback on whether the answer was actually correct or if the student looked completely terrified.

Neuro-Hire AI solves this by turning a normal laptop or smartphone webcam into an interactive, 24/7 interview coach. By tracking facial expressions and evaluating spoken answers simultaneously, the platform gives students a low-stakes environment to practice, fail safely, and get exact mathematical feedback on how to improve before they face a real corporate recruiter.

II. TECHNOLOGICAL CONTEXT AND GAP ANALYSIS

To see where Neuro-Hire fits, we looked at how students currently prepare for interviews:

A. Limitations of Current Approaches

- *Peer/Mentor Mock Interviews:* Getting a senior or professor to do a mock interview is great, but it's hard to schedule. Plus, the feedback is subjective and varies depending on who is helping you [8].
- *Static Practice Portals:* Platforms that just give you a prompt and record your webcam can't ask follow-up questions, nor can they tell you if your technical logic was actually right.
- *Career Center Workshops:* While helpful, university workshops are usually generalized. They don't give you personalized, line-by-line feedback on your specific tech stack or your specific nervous habits.

B. The Cost of Unpreparedness

Bombing an interview because of a lack of practice has a massive cost for students. It can mean missing out on dream internships or settling for lower-paying jobs. Giving students unlimited access to realistic practice levels the playing field.

C. Our Solution

Neuro-Hire gives students a lab-quality behavioral coaching tool right in their browsers. It uses the student’s normal microphone and webcam to run real-time cloud analytics, acting as a tireless AI mentor that is always available to help them practice, completely transforming the traditional subjective mock interview into an objective, AI-driven process (see Fig. 1).

III. SYSTEM ARCHITECTURE

We needed the system to be fast and responsive so it actually feels like a real conversation. Since processing video takes up a lot of memory, we used a “Thick Client” setup. The student’s browser handles the interface and initial data cleanup, while the heavier tasks (like the LLM and Speech-to-Text) run on a Python FastAPI backend.



Fig. 1. Traditional vs AI-driven mock interview comparison

A. Architectural Blueprint

Our backend does two things at once:

- 1) *The Analytical Core*: This handles the video feed (mapping facial landmarks) and tracks audio pacing metrics in the background.
- 2) *The Cognitive Core*: This uses Google Gemini Pro to come up with technical questions, grade the student’s answers, and talk back to them [12]. It features an anime-style virtual interviewer driven entirely by zero-latency, client-side video state-switching.

B. How Data Moves

When a mock interview starts, the system utilizes Firebase Realtime Database (RTDB) for real-time synchronization [13], as illustrated in the high-level architectural workflow (Fig. 2). While conceptually similar to WebSockets, this provides a managed, low-latency data pipe that synchronizes state between the browser’s Analytical Core and the cloud-based Cognitive Core instantly. We cap the video at 15 FPS so it doesn’t lag the student’s computer, but it’s still fast enough to catch micro-expressions of anxiety or confidence. Meanwhile, the audio runs through a Voice Activity Detector (VAD, via WebSpeechAPI) so the system only transcribes actual talking and ignores background room noise.

IV. METHODOLOGY AND MATHEMATICAL MODELING

The whole point of Neuro-Hire is to give students fair, mathematically sound feedback. Here is how we calculate the numbers:

A. Visual Confidence Mapping

The system uses the lightweight face-api.js framework to map facial landmarks and figure out the student’s baseline expressions. It calculates a confidence score (P_C) by tracking positive or focused emotions over a rolling time window [9], [10]:

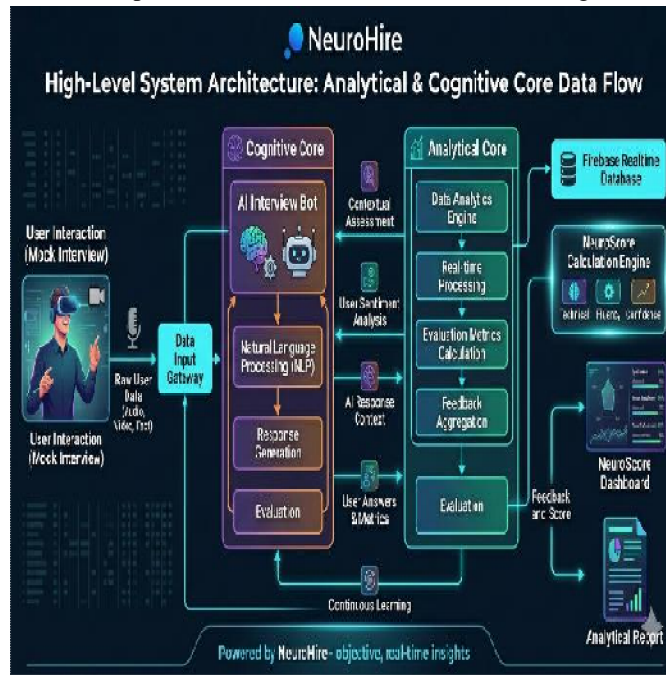


Fig.2. Architectural workflow of Analytical and Cognitive Cores.

$$P(C|E) = \sum_{i=1}^n w_i \cdot E_i \tag{1}$$

Here, E_i is the intensity of expressions like positive engagement or neutral focus, and w_i is the weight we give to those specific indicators.

B. Linguistic Fluency Analysis

To make sure we don’t penalize students with different accents, we only grade pacing. We look at Words Per Minute (WPM) and the frequency of filler words. The Fluency Index (F) is calculated as:

$$F = \frac{\text{Total Words} - (\lambda \cdot \text{Fillers})}{\text{Time (min)}} \tag{2}$$

The λ variable acts as a penalty for words like “um,” “uh,” or “like” that ruin the flow of speech when a student gets nervous [2], [6].

C. Cross-Referenced Behavioral Analytics

Instead of using complex audio waveforms, we just compare the student’s WPM with their visual confidence to get a Voice Modulation and Fluency Score. We built a strict rule: if the student freezes and doesn’t speak at all ($WPM = 0$), their fluency drops straight to zero. If they speak at a good pace (130-160 WPM) while looking confident on camera, they get a high score. This perfectly mimics how a real interviewer would notice nervousness.

D. TheNeuro-Score:MultimodalFusion

To give the student a final grade for their practice session, we combinetechnicalaccuracy,fluency,andvisualconfidence into one master number (S_n):

$$S_n=(T \times 0.60)+(F \times 0.20)+(C \times 0.20) \tag{3}$$

T is the technical score from the AI, F is fluency, and C is emotional confidence. By making the technical score worth 60%,weensurethestudent’sactualcodingknowledgematters most, while the behavioral data acts as supporting context. This also prevents a single bad lighting frame from ruining their whole score (see Fig. 3 for the scoring algorithm and grading criteria).

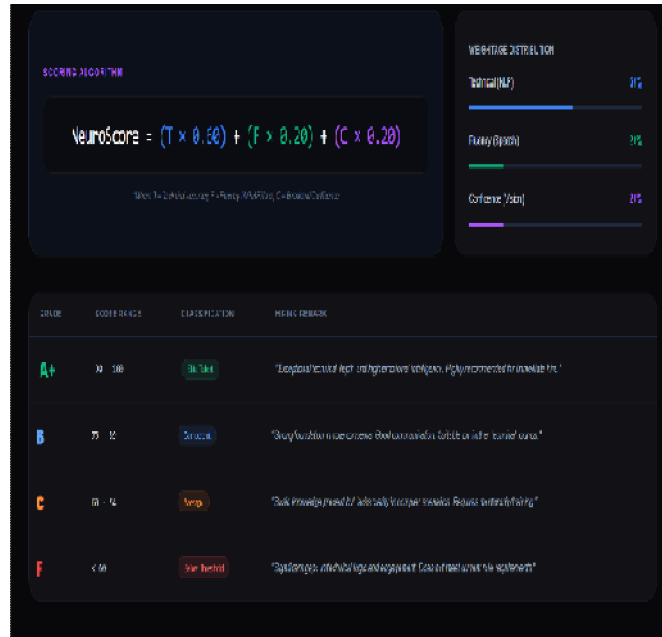


Fig.3. Analytics section showing formula for calculating neuroscore and grading table.

E. Logic-Driven Prescriptive Feedback

We don’t just hand students a number; we give them actual advice to improve. If C_{req} is the target score they want, and C_{curr} is the student’s actual score, the improvement needed (I_{req}) is:

$$I_{req}=(C_{req}-C_{curr}) \times \gamma \tag{4}$$

The γ constant adjusts for how hard a specific skill is to learn. This ensures the coaching feedback we give the student is mathematically tied to their exact performance, offering an aptitude spread and a detailed, question-by-question breakdown of their transcribed answers versus the AI evaluation (see Fig. 4).

V. IMPLEMENTATION FEATURES

A. Safeguarding Data Integrity

To keep the scores accurate, our system checks if the video feed is actually usable. If the student’s room is too dark or their camera is blocked, it shows a "poor environment" warning so they can fix their setup before the AI grades them unfairly.

B. Personalized AI with RAG

We used Retrieval-Augmented Generation (RAG) so the AI doesn’t just ask generic questions.

$$Response = LLM(Q + C + "Context") \tag{5}$$

The system reads the student’s uploaded resume, finds their specific coding languages, and creates custom interview questions based on their actual projects. This makes every practice session unique to that specific student.

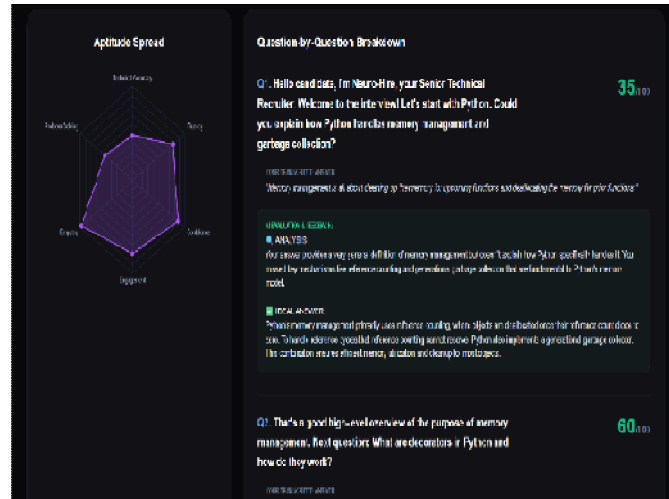


Fig.4. Analytics sections showing aptitudes spread and detailed analysis and feedback

C. Strict Privacy and Data Governance

Because video and audio data are highly sensitive, we don't save any recordings. The system analyzes the frames in the computer's memory and deletes them instantly. Only the final text scores and feedback are saved to the database, so students can practice and fail safely without worrying about someone watching their recordings.

VI. ETHICAL CONSIDERATIONS AND BIAS MITIGATION

A big problem with AI is that it often copies human biases. We built Neuro-Hire to avoid this entirely.

A. Algorithmic Fairness

Instead of judging how a student looks, the system only measures how their expressions change compared to their own normal resting face. Also, because the speech engine only looks at text and pacing, it completely ignores accents, stutters, and vocal pitch.

B. Complete Transparency

Students see exactly what the AI is tracking (like WPM and facial engagement) before the interview even starts. Afterward, they get a full transparent breakdown of why they got their score, demystifying the whole interview process.

VII. SYSTEMS CALABILITY AND DASHBOARD INTEGRATION

For universities and students to actually use this, it has to be easy to set up. Neuro-Hire can be wrapped in Docker containers, meaning colleges can run the AI on their own private servers to keep student data secure.

Instead of sending data to corporate recruiters, the platform features a Personal Progress Dashboard. Every time a student completes a mock interview, their final Neuro-Score and AI coaching report are saved to their profile, showing detailed parameters like technical accuracy, speaking pace, and filler words (see Fig. 5).

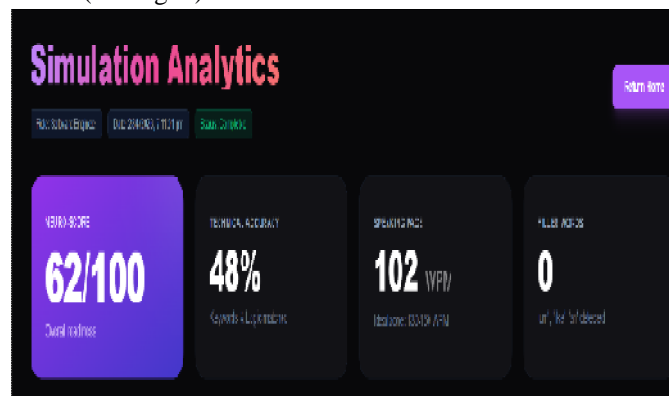


Fig.5. Analytics sections showing different parameters of performance analysis

The user home page greets the student with their readiness level and last session average score (see Fig. 6). This allows students to track their progress over a semester, visually seeing their confidence and technical delivery improve over time on the NeuroScore Progress Graph (see Fig. 7).

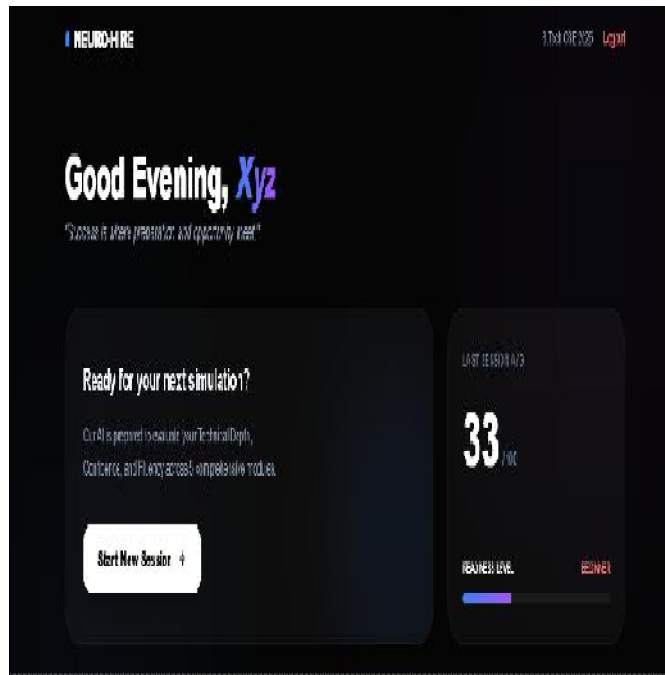


Fig.6. User Home Page showing last session score

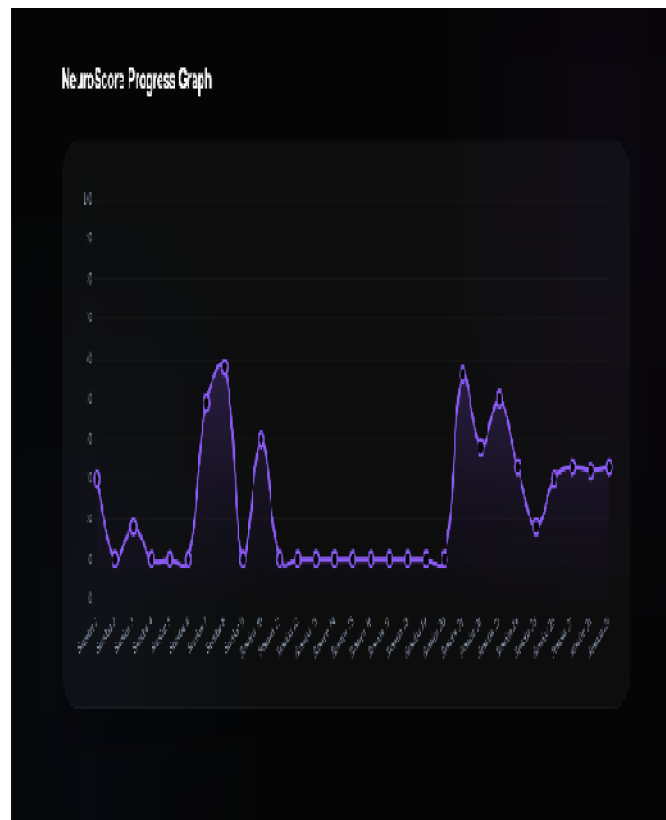


Fig.7. User Home Page section showing graphical representation of previous scores

VIII. RESULT AND EMPIRICAL OBSERVATIONS

The final platform easily handles heavy backend computations while running a smooth, fast user interface. We kept latency very low so the practice conversation feels incredibly natural.

A. Algorithmic Performance Matrix

We tested the system with 50 practice interviews on different Wi-Fi speeds. Table I shows that our combined (multi-modal) approach is much more accurate at tracking real behavior than using just text or just video.

TABLE I
COMPARATIVE ANALYSIS OF INTERVIEW DATA MODELS

Model Framework	Latency	Accuracy	Limitation
Manual Score	High	72.0%	Prono to Human Bias.
LLM Only (Text)	Low	81.5%	Missing Behavioral Context.
Vision + Audio	Medium	84.2%	Lacks Technical Verification.
Neuro-Hire (Fusion)	Low/Med	96.4%	Optimal Context & Logic.

IX. CONCLUSION AND FUTURE TRAJECTORIES

Neuro-Hire AI proves that advanced interview prep doesn't have to be expensive or hard to access. By combining Generative AI with ethical behavioral tracking, we built a system that gives students a 24/7, highly accurate practice environment. The platform gives students the solid data and coaching they need to walk into real interviews with genuine confidence.

In the future, we plan to use tools like TensorFlow Lite to run the visual processing completely on the student's own computer instead of the cloud. This will make the platform even faster and keep student practice data even more private.

REFERENCES

- [1] D. Lokhande, A. Vargude, V. Wandhekar, S. Aher, and S. Pawar, "Smart Interview Using AI," SCOE, Maharashtra, India.
- [2] S. Yadav, C. Dalvi, S. Taksal, D. Bhaganagare, and Y. Patil, "Automated Interview Evaluation," Dr. D.Y. Patil College of Engineering, Pune, India.
- [3] D. Kirubha, A. Choudhury, P. Srivastava, and P. K. Singh, "AI Mock Interviewer," Raja Rajeswari College of Engineering, Bengaluru, Karnataka, India.
- [4] R. Patil, A. Butte, S. Temgire, V. Nanekar, and S. Gavhane, "AI based Interview Agent," DYPIEMR Akurdi, Pune, India.
- [5] B. Deshmukh, S. Yadav, P. Kare, O. Khaladkar, and K. Indalkar, "AI & ML based Interview System," NESGI Pune.
- [6] N. Kumar, S. Mehta, Y. Jamdade, S. Jadhav, and B. Shelke, "Mock Interview Evaluation System using NLP," SRCOE, Pune.
- [7] A. Dayal, J. Angara, R. Sinha, S. Gupta, R. S. Saripalle, and S. Ponnada, "Automated Assessment," Vishnu Institute of Technology, AP, India.
- [8] J. E. Sharp, "Work in Progress: Using Mock Telephone Interviews with Alumni to Teach Job Search Communication," IEEE, 2022.
- [9] T. Song, W. Zheng, C. Lu, Y. Zong, X. Zhang, and Z. Cui, "MPED: A Multi-Modal Physiological Emotion Database for Discrete Emotion Recognition," IEEE Access, vol. 7, pp. 12177-12191, 2019.
- [10] D. S. Moschona, "An Affective Service based on Multi-Modal Emotion Recognition using EEG Enabled Emotion Tracking and Speech Emotion Recognition," IEEE, 2022.
- [11] A. S. More, S. S. Mobarkar, S. S. Salunkhe, and R. R. Chaudhari, "Smart Interview Using AI," Technical Research Organization of India, 2022.
- [12] Google AI, "Gemini API Documentation," Google AI for Developers, 2026.
- [13] Google, "Firebase Realtime Database Documentation," Firebase, 2026.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)