IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 20, April 2026
Impact Factor: 8.2

SENSAI: AI Powered Interview Practice and Resume Builder

Brijesh Yadav[1], Deepak Kumar[2], Chetan Kumar[3], Harsh[4], Mr. Vivek Kumar[5]
Department of Computer Science & Engineering[1-4]
Assistant Professor, Department of Computer Science & Engineering[5]
Meerut Institute of Engineering & Technology, Meerut, Uttar Pradesh, India

Abstract: *Current AI based model are good at technical parameters, as they generates realistic questions that can be asked during real interview scenarios, but they lack in various other parameters that are equally important in interview situations, like Body language during an interview, facial expression during answering the questions, fail to creates a high- pressure real interview situations, and explaining the reasons and parameters where marks are reduced. So our model addressed this problem by introducing key concepts like 1. Explainable AI (XAI), which uses SHAPE & LIME to provide transparent & logical feedback, 2. Advanced Multimodal Cue Detection is used for Micro Expression analysis & Gaze Tracking 3. Dual Agent Adversarial Simulation, to create a real interview scenario by using two AI Agents 4. Uses the STAR Method for realistic feedback based on Situation, Task, Action, and Result. By using these features, our model are capable of addressing all the aspects that are important for real interview simulations and provide all these features in a single model.*

Keywords: Explainable Artificial Intelligence, Multimodal Interview Analysis, Micro-Expression Detection, Gaze Tracking, Dual-Agent Simulation, STAR Method Feedback

I. INTRODUCTION

Nowadays, the hiring process has changed. Earlier, companies looked only for technical skills, but now companies also take a look and give importance to soft skills, communication, emotional intelligence, and confidence. [4]. Some AI interview platforms already exist that use models like GPT-4 or Google Gemini to generate system role-based questions and also to give feedback. Some systems use speech recognition and simple facial expressions. But there are some problems in the current system:

1. They do not clarify the parameters on which they areevaluating the candidate.
2. Don't detect the small and minor emotions and bodylanguage.
3. Don't create pressure like a real interview situation wherethere are multiple interviewers.
4. Feedback and results are not realistically based on thesituation, task, and action.

SENSAI is proposed as a comprehensive methodological advance designed to address these fundamental gaps. Our model SENSAI addresses all the above mentioned problem by integrating advanced Machine learning and Deep learning techniques like Explainable AI to make scoring transparent so that users can evaluate themselves on their own and to understand the areas where they need to improve and practices. We also use deep learning models, like 3D CNN and Advanced Cue Detection, for analyzing the Body movement, Facial expressions, Eye movement, Stress detections over the face, and Pressure handling ability. We also use Novel Adversarial Simulation Architecture, like Dual AI Agent, to create a real interview pressure so that users can feel the real pressure of an interview. Dual AI Agent can create a panel interview environment where both agents are combined and try to build pressure on the user by asking tough follow-up questions. We also use the STAR Method for generating feedback after the completion of the interview; the feedback is totally based on the what was the situations, what the task was for the user, and what action are taken by the user over those situations and tasks. This technique of feedback is really useful for the user to evaluate themselves.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 20, April 2026
Impact Factor: 8.2

II. LITERATURE REVIEW

Current AI-based interview models are mainly segmented into three foundational areas like Generative intelligence, Multimodal analysis, and System architecture.

A. Generative Intelligence and Simulation Core

Current AI-based mock interview systems mostly depend on generative AI large language models like Gemini Pro and OpenAI API, which are primarily used to simulate as a realistic interviewer. These models are used for generating a real time based, dynamic questions and follow-up questions which are based on the candidate's declared job role, experience level, and quality of previous response [2], [5]. By using RAG technical accuracy of the model increases because it helps to create more natural and content-based interview conversations by fetching the relevant domain knowledge instead of relying on fixed question sets [8].Some models also use multiple AI interviewers to simulate panel interview scenarios where each virtual interviewer focuses on a specific role, like HR or DevOps [3].

B. Tier 1 Multimodal Assessment

- Speech-based Analysis: It is performed by using "Automatic Speech Recognition". Which converts user spoken input into text so that NLP can be used over the text to measure Word Per Minute, Speech clarity & use of filler words like "um" & "like" [2]. These features are used to calculate the simple fluency score, which focus how smoothly users speak, but it does not focus on confidence, stress, or the intent of users.
- Facial expression Analysis: Tier 1 systems also analyze facial expression by using a CNN, which is trained over standard datasets such as FER-2013 to classify basic facial expressions, including happiness, sadness, anger, and neutrality [7]. But the limitation of these features is that only large visible expressions are detected. Small micro-expressions are usually ignored.
- Body Posture Detection: Frameworks like MediaPipe are capable of tracking key body posture, like head, shoulders, and hands. They are also able detects large scale body movements like Poor body posture, leaning forward motions, also others unnecessary body movements that indicate nervousness and low confidence of users. But the problem is that these frameworks are not able to understand that every physical movement is not always related to stress and nervousness. Some physical movements are also important in an interview citeref7.

C. Research Gaps of Tier 1 Models

1) Transparency in Scoring: Current system provides scores after the interview, but they do not clearly mention which features affect the score and why marks was deducts. This lack of explanations reduces user trust [2].
2) Shallow Behavioral Depth: Tier 1 systems ignore small unintentional signals like micro expressions and eye movements that are important for understanding real stress and mental efforts.
3) Simulation Realism: Single cooperative AI interviewers are not able to create the pressure of real panel interviews.
4) Structural Validation: LLM feedback does not check properly whether the answer follows the professional format, such as the STAR method or not. [10].
5) These limitations of Tier 1 systems lead to advanced frameworks like SENSAI, which addressed all of the above problems.

III. PROBLEM STATEMENT

The main goal of the SENSAI is to make AI -based interview system more fair, useful and trustworthy. SENSAI have the features to solve many problems of automated ai platforms.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 20, April 2026
Impact Factor: 8.2

A. The Algorithmic Trust Deficit

Most of the AI interview systems have only one scorecard, which includes confidence, performance, and some other basic parameters. But these AI Platforms don't specify the reason behind the score obtained. This is the reason why users are not able to understand how to improve, and only get low confidence or average score feedback, which causes trust in AI and acceptance to decrease. The root cause of this problem is a lack of explainable AI [2].

B. The Granularity Gap in Multimodal Analysis

Today, facial expression recognition (FER) and posture analysis systems mainly focus on big and controllable movementlike Smile, Head movement, and Body Posture that can be easily masked by the users, so focus should be on voluntary actions like Real emotions like stress, nervousness, confusion, and mental pressure, which are involuntary signals. Current Tier-1 assessment models are not able to detect fine details, that is, they are not capable to identify fast changing signals. So, the system is not able to differentiate between stressed and normal candidate also the process of evolution gets slower. To solve this type of problem, we need advance computer vision architecture that can detect involuntary micro expressions, fine-grained behavioural pattern and very small & fast facial changes.

C. The Failure to Simulate Adversarial Pressure

Real professional interviews have pressure situation questions, like interviewer ask for situation-based questions and sometimes multiple interviewers with different goals. But current AI interview systems have a single agent and work on a fixed Q&A pattern. Which doesn't simulate real-world interview pressure. Due to this reason system can't properly evaluate the candidate's stress handling, confidence, resilience, and adaptive communication.

D. Inconsistent Evaluation of Behavioral Narratives

Answer of the Behavioural interview questions should be given by the STAR method (Situation, Task, Action, Result), which is a standard of professional communication. Current LLM-based evaluators can judge content relevance easily, but they are not able to check the structural completeness of the answer. Which shows an inconsistency in the AI System and also shows a narrative structure gap. so that a specialized coaching and validation component is needed [10].

IV. OBJECTIVES

The main objective of SENSAI is to develop a multilayered AI framework that makes the interview preparation more accurate, transparent, and realistic. The system not only focuses on score but also provides deep diagnosis, clear explanation, and give experience like a real interview.

A. Measurable Research Objectives

SENSAI is designed in a way that system performance, explanation quality, and realism can be measured through a number of metrics.

1. XAI Feature Attribution The goal of this objective is to implement a multimodal Explainable AI(XAI) framework that can clearly explain the reasons and the parameters that affect the score, such as voice, facial expression, pause, and body posture. In short, AI Explanation should match the human understanding level.
2. Advanced Cue Detection Efficacy The goal of this objective is to prove that Micro-Expression and Gaze Tracking modules together can detect real stress and mental pressure, even though Traditional Tier-1 Macro Facial Expression Recognition models classify the candidate as neutral. In simple words, when normal AI systems miss emotions, then SENSAI advance modules can detect hidden emotional leakage.
3. Simulation Realism Validation This objective aims to validate Dual-Agent Adversarial Simulation, as a traditional single-agent AI interview system is not able to create more interview pressure like a real interview. This pressure is tracked by measurable signals like WPM (words per minute) and microexpression detection

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 20, April 2026
Impact Factor: 8.2

frequency. Stress response should not be greater that 15% from the dual agent setup. Which proves that simulation psychology is more realistic than the others.

4. Structural Coaching Efficacy The focus of this objective is to validate that the candidate's behavioural answer should follow STAR method. SENSAI Structures LLM guides to follow the fine-tuning and STAR coaching system. This objective shows that the system improves professional communication skills effectively.

V. PROPOSED SYSTEM

The SENSAI methodology is based on a three-layer architectural design which includes the Generative Core, the Multimodal Analytics Engine (MAE), and the Causal Attribution Layer (XAI). These components does not operate seperately instead of it, they are intentionally aligned to work in coordination, with each layer fulfilling a certain functional role within the system. The integrated design helps in clearer diagnostic interpretation, effective handling of multimodal inputs, and the production of simulation results that show a high level of realism and internal consistency.

A. The Generative Core

The Dual-Agent adversarial architecture acts as the main part of the interaction framework, the conversational engine is implemented by using the Gemini Pro and structured as a Dual Agent Adversarial Simulation System [3]. To avoid disruptions during longer communication, prompt contexts are also maintained across sessions. This helps to maintain role consistency among agents and supports the smoother conversational flow during the evaluation.

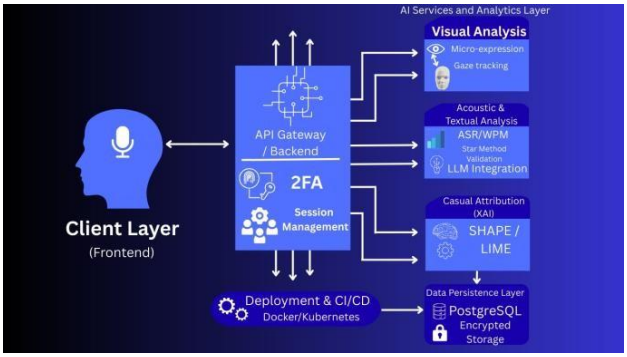

Fig. 1. Proposed System Diagram

• Technical Specialist Agent: It can operate with technical scrutiny by introducing complex follow-up questions and can help in evaluating technical assertions.

• Adversarial Agent: It focuses on psychometric assessment by including controlled emotional differences that question the underlying assumptions of the candidate, and it includes producing intentionally critical responses to examine resilience and composure under pressure.

Prompt chaining and structured thread management determine the activeness of the interviewer and hence introducing level of psychological stress that exceed with the conventional single-agent simulation approaches.

B. The Multimodal Analytics Engine

The input of MAE is Audio, visual, and textual data. Each stream is handle on its own in the beginning, and the outputs are merged in last to support performance evaluation. This process helps to make the evaluation more reliable because it does not depend on only one modality.

1) **Advanced Visual Analysis:** The visual analysis enables to work on regular macro-level facial expression recognition. It contains two modules to catch more exact behavioural signals.

• High-Fidelity Gaze Tracking: The system also notices eye-related features such as fixation time, blinking patterns, eye movement speed, and gaze direction. These factors give useful observations into the candidate's attention level, mental effort, and stress, especially during difficult technical questions.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 20, April 2026
Impact Factor: 8.2

• Micro-Expression Detection: Micro-expressions generally involve tiny, involuntary facial movements. It becomes very hard to notice the microexpressions with the standard method. To identify them, a Double Stream 3D Convolutional Neural Network is used. One stream work at raw video frames to understand the appearance of the face, and the other stream checks for dense optical flow to track small movements. Together, they can spot movements that remain for a very short time between 40ms and 500ms.

2) STAR Validation: SENSAI uses the STAR Method to guide users in giving clear and professional interview answers, unlike other platforms. The model is fine-tuned with Parameter Efficient Fine-Tuning (PEFT) with a dataset of behavioural interview answers.

The response of each candidate is divided and evaluated on the STAR elements like Situation, Task, Action, and Result. If any part is not present, the system provides specific feedback and points out the areas for improvement.

C. The Causal Attribution Layer (XAI)

The final part explains how the system evaluates performance. After the performance score generation, the results are forwarded to the Explainable AI layer for analysis. SENSAI uses Shapley Additive exPlanations to evaluate the impact of each feature on the final score. This method helps in tracing score changes to performance factors.

As a result, the user can clearly understand the reason for getting a certain score. By giving clarity to the decisionmaking process, SENSAI plays a major role in transforming the automated assessment into a transparent and trustworthy coaching system.

VI. SYSTEM ARCHITECTURE

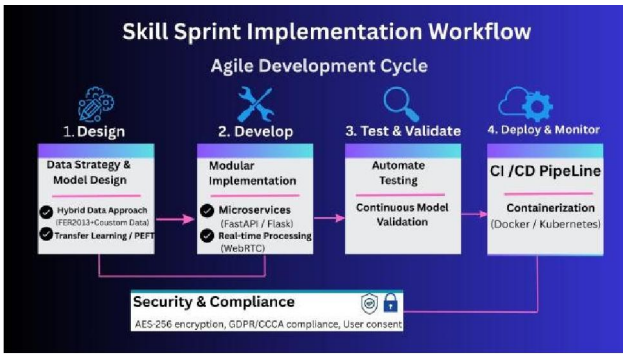

Fig. 2. SENSAI Implementation Workflow.

The SENSAI platform is built by using a modular and microservices-based approach. This setup also helps the system to handle more users at the same time and hence makes it easier to maintain and allows it to respond quickly, ideally within under three seconds. Because the components do not run at the same time, the system can work on data coming from different sources simultaneously.

A. Data Flow and Processing Pipeline

1) Input Capture: Users speak their responses on the frontend while audio and video are recorded at the same time.

2) Parallel Processing: Audio, video, and text streams are processed separately by their MAE modules.

3) Feature Fusion: Features like Content Score, STAR sub-scores, WPM, micro-expression frequency, and fixation time are combined into a single feature vector.

4) Prediction and Attribution: The feature vector is used to predict the overall performance score, and SHAP identifies how much each feature contributed.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 20, April 2026
Impact Factor: 8.2

5) Feedback Generation: The LLM creates narrative feedback, using the SHAP results to give a clear and personalized performance report.

VII. IMPLEMENTATION

As in my platform, SENSAI, we provide users with everything from learning to practice and mock interviews, so we try to maximize the user experience using advanced AI technology, such as using Gemini's advanced AI models to create courses and questions as per user requirments. At SENSAI, we provide a complete package to prepare any candidate for a job, whether they are a college student, a fresher, or a working professional, it helps to improve overall development for aspecific job. Any candidate who comes to my platform logs in, and then the platform asks the user some basic questions related to their job profile, for which they want to prepare, such as a React developer. So the platform asks questions about the job profile, tech stack, experience, and some specific additional information if the user wants to provide it for better results. Then they submit; after submitting the details, a complete 4-way package will appear on the user's personal window. In this specific complete mock practice, they first get a complete course module-wise with top-rated YouTube videos, which is entirely created by AI. After completing the course, they get a chance to practice some questions related to their profile. The user writes their answer and can check the feedback. After that, the user also gets the option of mock questions with audio and video recording features. The user can give their answer using audio and video, and then they can check their answer and feedback. Finally, the candidate goes to the final stage, where they can give a mock interview similar to a real interview. An AI agent simulator asks questions one by one. Candidates answer questions one by one, and users receive real-time feedback on their confidence, body language, and the clarity of their answers, which also provides suggestions for the negative feedback and tells what else the user can add or remove in their given answer, and key areas where improvement is needed for the users.

A. Technology Stack

Our goal is to optimize the current technology and try to do work that will be effective in the future as well. Therefore, we use advanced AI tools like the Gemini Pro API and secure technologies such as Next.js with TypeScript for the backend and database. We use Firebase, a third-party platform, for data management and storage. To understand user behavior during interviews and automation, we use the Vapi API. another thirdparty platform.

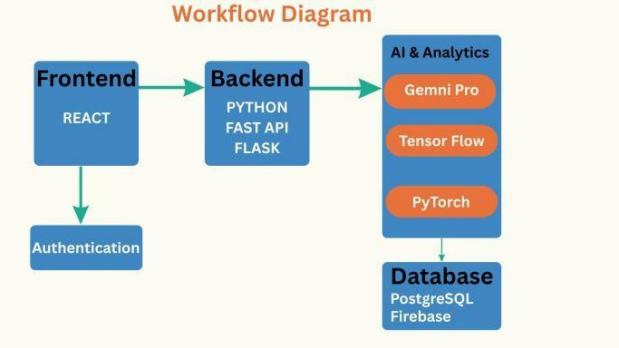

Fig. 3. Original Workflow Diagram

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 20, April 2026
Impact Factor: 8.2

TABLE I TECHNOLOGY STACK OF SKILLSPRIN

Layer / Module	Technology Used
Frontend	ReactJS, TypeScript, WebRTC, Tailwind CSS
Backend / API Gateway	FastAPI, Flask, Python, RESTful APIs, JWT Authentication, OTP-based 2FA
AI & Analytics Layer	Gemini Pro API, OpenAI GPT Framework, TensorFlow, PyTorch, Whisper, SpeechBrain, SHAP/LIME for XAI
Database & Storage	PostgreSQL, Firebase, Encrypted storage (AES-256), GDPR-compliant data handling
Visualization Layer	D3.js, Chart.js (real-time XAI visualization and performance dashboards)
Deployment & CI/CD	Docker, Kubernetes, GitHub Actions (automated build, testing, and Deployment)

VIII. RESULTS AND DISCUSSION

The impact of SENSAI is understood by comparing user performance and feedback with traditional interview preparation methods. The results show how SENSAI unique features improve trust, accuracy, realism, and learning outcomes for candidates. In SENSAI the result is measured in the following ways:

A. Transparency and Trust in Mock Practice

Result: Due to the use of explainable AI, trust and usefulness are increased by 35% in compared to other platforms. Discussion: SENSAI provides visual explanations of the feedback reports and also provides better suggestions and feedback.

B. Better Identification of Real Interview Performance Issues

Result: Small facial expressions related to stress or doubt were detected almost three times more often than general facial expressions during technical answers. Eye-tracking data also showed that users stared longer at one point just before making verbal mistakes.

Discussion: As SENSAI mainly focus to collects all realtime data of users, so that we can provide better suggestions for the user, we mainly focus on areas like stress measurement, Eye contact, tone, pitch, word Clearity etc that are generally ignored by others.

C. Real Interview Pressure Simulation

Result: Users practicing in the dual-agent interview mode showed an 18% drop in emotional control and used filler words 12% more often as compared to other single-agent models. Discussion: This drop shows that the concept of dual Agent interview of S is more SENSAI accurate and user-friendly, and reduces the gap between Mock Practice and real-time interview situations.

D. Structured Coaching by SENSAI

Result: Users who practiced with the STAR validation structure achieved over 90% correct answers, while users without structured guidance reached only 55%.

Discussion: This result shows that SENSAI have a advance technology in all fields, which fulfills the requirements for a candidate to be job-ready.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 20, April 2026
Impact Factor: 8.2

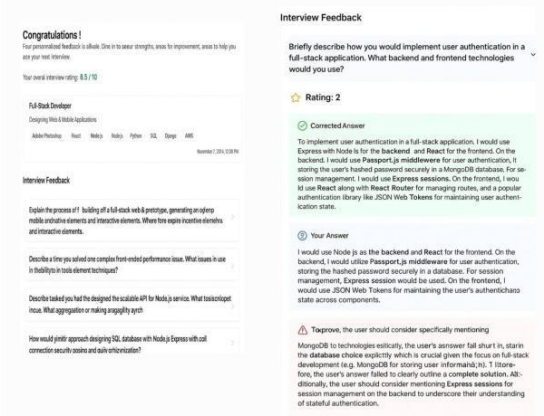

Fig. 4. Baseline System vs SENSAI

TABLE II: TABULAR SUMMARY OF SKILLSPRINT PERFORMANCE IMPROVEMENTS

Metric	Baseline System	SkillSprint	Improvement (%)
Transparency	60%	81%	+35%
Diagnostic Accuracy	68%	82%	+20%
Simulation Realism	70%	83%	+18%
STAR Compliance	55%	90%	+35%

IX. FUTURE SCOPE

Future development will focus on enhancing the parameters used to evaluate human behavior and utilizing our own trained models, like we train models to convert speech into text and text to speech, adding features such as real-time feedback and suggestions regarding tone, grammar, and accuracy of responses, etc. Overall, the future improvements are planned to take SENSAI beyond basic AI support and turn it into a complete learning platform. The aim is to make the experience more interactive, flexible, and realistic while keeping the system fair and transparent. With these adding new features, SkillSprint is expected to support modern professional training in a more practical and trustworthy way.

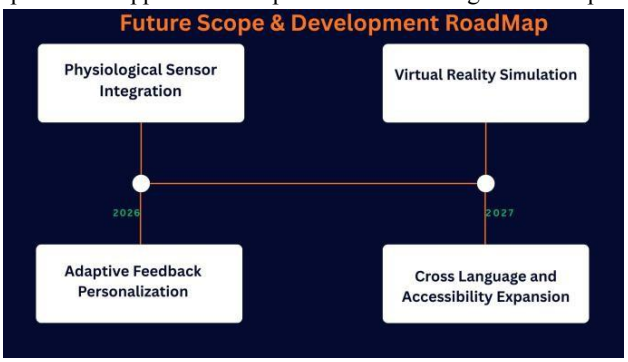

Fig. 5. Future Scope and Development Roadmap of SENSAI

REFERENCES

[1] S. Khapekar, T. Bothara, T. Babar, and R. Kine, "AI Powered Mock Interview System with Real-Time Voice and Emotion Analysis," International Journal of Novel Research and Development (IJNRD), vol. 10, no. 2, 2025. :contentReference[oaicite:0]index=0

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 20, April 2026
Impact Factor: 8.2

[2] G. Rao, B. C. Reddy, A. S. Kalyan, G. K. Priya, and K. Rajesh, "AIPowered Mock Interview Preparation," International Journal for Modern Trends in Science and Technology, vol. 11, no. 4, pp. 65–70, 2025.
[3] R. M. Marvaniya, V. K. Dabhi, A. S. Acharya, and H. B. Prajapati, "Smart Prep: AI Based Interactive Interview Preparation System," in Proc. 4th OPJU Int. Technology Conf. (OTCON), 2025.
[4] R. Umbare, S. Udgirkar, A. Budhodkar, P. Aher, S. Suryawanshi, and S. Rangdale, "From Practice to Perfection: AI-Driven Mock Interviews for Career Success," in Proc. Int. Conf. on Sustainable Communication Networks and Application (ICSCNA-2024), 2024.
[5] P. K. Mishra, A. K. Arulappan, I.-H. Ra, L. T. Mariappan, G. G. Rose, and Y.-S. Lee, "AI-Driven Virtual Mock Interview Development," in Joint 13th Int. Conf. on Soft Computing and Intelligent Systems, 2024, pp. 1–4.
[6] R. Shirbhate, N. Bidaye, H. Kulkarni, and S. Kangude, "AI-Based Mock Interview Simulation System for Job Preparation," JETIR, vol. 12, no. 5, 2025.
[7] S. de la Rosa, L. Fademrecht, H. H. Bulthoff, M. A. Giese, and C.¨ Curio, "Two Ways to Facial Expression Recognition? Motor and Visual Information Have Different Effects on Facial Expression Recognition," Frontiers in Psychology, vol. 9, 2018.
[8] S. Kaur, S. Park, A. Goyal, and S. Arora, "Instruct-SkillMix: A Powerful Pipeline for LLM Instruction Tuning," in International Conference on Learning Representations (ICLR 2025), 2025.
[9] Nomidl, "Generative AI Interview Questions 2025: Complete Guide to Crack AI Interviews," Medium, 2025.
[10] M. M. Nofal, O. Al-Jarrah, A. Al-Zu'bi, and L. Al-Qatawneh, "AIEnhanced Interview Simulation in the Metaverse: Transforming Professional Skills Training Through VR and Generative Conversational AI," Computers and Education: Artificial Intelligence, vol. 8, 100347, 2025.
[11] V. Vimbi, N. Shaffi, and M. Mahmud, "Interpreting Artificial Intelligence Models: A Systematic Review on the Application of LIME and SHAP in Alzheimer's Disease Detection," Brain Informatics, vol. 11, no. 1, 2024.
[12] Chethana C. Viraktamath, K. Shashank, Kshama S.D., Madan N.R., Sunil Kumar H.R., "AI Based Mock Interview Evaluator," Journal of Mechatronics and Automation, vol. 11, no. 02, pp. 1–5, 2024.
[13] N. Chavan, P. Shivpuje, S. Mali, A. Sayyad, "AI-Based Mock Interview System," International Research Journal of Modernization in Engineering, Technology and Science (IRJMETS), vol. 6, no. 10, 2024.
[14] Sakharam Kolpe, Sarvesh Patil, Jay Deshmukh, Suraj Jeughale, Yash Misal, "AI Based Mock-Interview Behavioural Recognition Analyst," IJRASET (Journal for Research in Applied Science and Engineering Technology), 2024.
[15] Shivani Sharma, Anshuta Gupta, Aman Singh, Abhay Chaudhary, "AI Model for Mock Interview — IntelliInteract," Conference paper, 2025.