

Chayan AI: Real-Time Multimodal Interview Analysis Using NLP, Computer Vision, and Deep Learning for Candidate Evaluation

Pawan Joshi¹, Varun Tripathi², Swapnil Srivastava³, Usha Sharma⁴

¹UG Scholar, Dept. of IT, Babu Banarasi Das Institute of Tech. & Mang., Lucknow, India

²UG Scholar, Dept. of IT, Babu Banarasi Das Institute of Tech. & Mang., Lucknow, India

³UG Scholar, Dept. of IT, Babu Banarasi Das Institute of Tech. & Mang., Lucknow, India

⁴Associate professor, Dept. of IT, Babu Banarasi Das Institute of Tech. & Mang., Lucknow, India

Emails: pawanjoshii777@gmail.com¹, varuntrip99@gmail.com²,

Swapnilsrivastavaofficial484@gmail.com³, usha.sharma34@gmail.com⁴

Abstract

This paper presents Chayan AI, a next-generation approach to interview automation and candidate evaluation, designed to bring greater objectivity and efficiency into modern hiring processes. The proposed system integrates multiple artificial intelligence technologies, including Natural Language Processing (NLP), Machine Learning (ML), Computer Vision, and Generative AI, to create a unified platform capable of conducting and assessing interviews in a structured and data-driven manner.

Unlike traditional interview methods that rely heavily on human judgment, the proposed framework focuses on analyzing candidates through multiple dimensions such as communication skills, response quality, behavioral patterns, and facial expressions. It incorporates automated question generation, real-time speech and text analysis, and visual cues interpretation to provide a more comprehensive understanding of a candidate's performance. The system is also designed to adapt dynamically during interviews, ensuring a more personalized and interactive experience.

One of the key contributions of this work is the development of an integrated evaluation mechanism that minimizes bias while maintaining consistency across assessments. By combining linguistic analysis with visual and behavioral insights, the platform offers a balanced and fair evaluation model. In addition, the use of generative AI enables the system to create context-aware questions, making the interview process more flexible and realistic.

The effectiveness of Chayan AI is validated through performance metrics such as evaluation accuracy, consistency, and user experience. The results indicate a significant improvement in reducing manual effort and enhancing decision-making quality. The motivation behind this research arises from the growing need for scalable and unbiased recruitment solutions in today's fast-paced digital environment. Existing systems often operate in isolation and fail to capture the complete profile of a candidate.

This work addresses these challenges by proposing a comprehensive and intelligent solution that not only improves the efficiency of hiring processes but also lays the foundation for future advancements in AI-driven human resource management systems.

Keywords: AI Interview System, NLP, Machine Learning, Computer Vision, Generative AI, Automated Hiring, Candidate Evaluation, Bias Reduction, Human-Computer Interaction, Virtual Recruitment Systems

INTRODUCTION

In recent years, the hiring landscape has undergone a major transformation, driven by rapid

advancements in artificial intelligence and digital technologies. Organizations today receive an overwhelming number of applications for every job

opening, making it increasingly difficult to evaluate candidates efficiently and fairly. Traditional interview processes, which rely heavily on human judgment, often suffer from limitations such as unconscious bias, inconsistency, and time constraints. As a result, there is a growing demand for intelligent systems that can streamline recruitment while ensuring accuracy and fairness.

The integration of artificial intelligence into recruitment presents a promising opportunity to bridge the gap between human intuition and data-driven decision-making. Technologies such as Natural Language Processing (NLP), Machine Learning (ML), Computer Vision, and Generative AI have the potential to transform how interviews are conducted and evaluated. However, many existing solutions focus on isolated aspects such as resume screening or basic video analysis, which limits their effectiveness in providing a complete assessment of a candidate.

This paper introduces **Chayan AI**, a unified and intelligent platform designed to automate and enhance the interview and evaluation process. The proposed system combines multiple AI techniques to analyze candidates from different perspectives, including verbal communication, response quality, facial expressions, and behavioral patterns. By integrating automated question generation, real-time response analysis, and visual interpretation, the framework ensures a more comprehensive and structured evaluation process. Special emphasis is placed on creating a seamless and user-friendly experience, allowing both candidates and recruiters to interact with the system in a natural and efficient manner.

The rapid evolution of machine learning and deep learning technologies has significantly influenced the development of intelligent recruitment systems. Earlier approaches primarily focused on keyword-based resume filtering or manual interview processes, which often failed to capture the true potential of candidates. More recent advancements have introduced video-based interviews and basic AI-driven analytics, but these systems still lack integration and adaptability. This highlights the need for a more holistic solution that brings together different AI capabilities into a single, cohesive framework.

The proposed system, **Chayan AI: AI-Based Interview and Evaluation Platform**, aims to address these challenges by providing an end-to-end

solution that combines adaptive interview generation with multimodal candidate analysis. The platform leverages NLP for understanding and evaluating responses, computer vision for analyzing facial cues and engagement, and machine learning models for generating unbiased evaluation scores. Additionally, generative AI enables the system to dynamically adjust interview questions based on candidate performance, creating a more interactive and realistic interview experience.

Furthermore, the current landscape of AI-driven recruitment systems can broadly be categorized into three areas: automated resume screening tools, video interview analysis systems, and AI-based assessment platforms. While each of these contributes valuable capabilities, they often operate independently and fail to provide a unified evaluation approach. This fragmentation limits their effectiveness and highlights the need for an integrated system that can deliver consistent and reliable results.

As advancements in artificial intelligence continue to reshape industries, systems like Chayan AI offer a practical pathway toward smarter and more efficient hiring processes. By combining multiple technologies into a single platform, the proposed framework not only enhances decision-making but also improves fairness and transparency in recruitment. This research contributes to the growing field of AI-driven human resource management and lays the groundwork for future innovations in intelligent hiring systems and digital workforce solutions.

LITERATURE REVIEW

Research on intelligent recruitment and automated interview systems can broadly be divided into three categories: (i) AI-based resume screening and candidate filtering systems, (ii) video interview analysis and behavioral assessment systems, and (iii) multimodal and adaptive AI-driven hiring platforms.

A. AI-Based Resume Screening and Candidate Filtering Systems

Significant progress has been made in automating the initial stages of recruitment through AI-based resume screening systems. These systems primarily use Natural Language Processing (NLP) and Machine Learning (ML) techniques to extract relevant information from resumes and match candidates with job requirements. Keyword-based filtering and semantic analysis have improved the speed and

efficiency of shortlisting candidates.

Advanced systems have introduced contextual understanding and ranking algorithms that go beyond simple keyword matching, enabling better alignment between candidate profiles and job descriptions. However, despite these improvements, such systems are limited in scope as they focus only on static data. They fail to evaluate critical aspects such as communication skills, problem-solving ability, and behavioral traits, which are essential for comprehensive candidate assessment.

B. Video Interview Analysis and Behavioral Assessment Systems

With the rise of remote hiring, video-based interview platforms have gained popularity. These systems incorporate computer vision and speech processing techniques to analyze facial expressions, eye contact, tone of voice, and speech patterns. They aim to provide insights into a candidate's confidence, engagement, and emotional state during an interview.

Recent advancements have improved facial recognition accuracy and enabled real-time analysis of non-verbal cues. Some systems also integrate speech-to-text conversion and basic sentiment analysis to evaluate candidate responses. While these approaches enhance traditional interviews, they often operate as standalone solutions and lack deeper contextual understanding. Additionally, concerns regarding transparency and fairness in automated behavioral analysis remain a challenge.

C. Multimodal and Adaptive AI-Based Interview Systems

More recent research has focused on combining multiple AI techniques to create intelligent and adaptive interview platforms. These systems integrate NLP for response evaluation, machine learning for scoring, and in some cases, generative AI for question generation. Adaptive questioning mechanisms allow interviews to adjust dynamically based on candidate performance, making the process more interactive and personalized.

Some platforms also attempt to combine resume analysis with video interview insights to create a more holistic evaluation. However, these systems are often limited by partial integration, where different components function independently rather than as a cohesive pipeline. Multimodal analysis is still in its early stages, and many systems struggle to balance accuracy, scalability, and fairness.

D. Research Gap and Motivation

The reviewed literature highlights three major but largely disconnected research directions: resume-based screening systems, video interview analysis tools, and partially integrated AI interview platforms. While each approach contributes valuable capabilities, several limitations persist. Most existing systems focus on specific aspects of the hiring process rather than providing a complete end-to-end solution.

In particular, many systems lack:

- A unified framework that combines textual, visual, and behavioral analysis
- Real-time adaptive interview mechanisms powered by generative AI
- Consistent and unbiased evaluation models across different candidates

Furthermore, current solutions often fail to provide transparency in decision-making and may still inherit biases present in training data. The absence of a fully integrated, intelligent system limits their effectiveness in real-world recruitment scenarios.

Therefore, there is a clear research gap: no existing approach offers a comprehensive, scalable, and intelligent platform that integrates NLP, Computer Vision, Machine Learning, and Generative AI into a single cohesive system for automated interview conduction and evaluation.

The proposed system, **Chayan AI**, aims to address this gap by delivering a unified, multimodal, and adaptive interview platform that ensures fair, accurate, and data-driven hiring decisions.

PROPOSED SYSTEM AND METHODOLOGY

The proposed **Chayan AI** system integrates multiple artificial intelligence techniques into a unified platform for automated interview conduction and candidate evaluation. Unlike conventional recruitment systems that focus on isolated tasks such as resume screening or basic video interviews, this system performs **end-to-end interview management, multimodal analysis, and intelligent evaluation** within a single framework. The architecture is designed around interconnected modules coordinated through a central decision engine that ensures consistency, scalability, and real-time processing.

A. System Architecture Overview

The architecture comprises three major interconnected modules:

- (1) **AI Interview Generation Module**,
- (2) **Multimodal Candidate Analysis Module**, and
- (3) **Evaluation and Reporting Module**.

These modules operate through a unified web-based interface following a **client-server architecture**, where the backend handles AI models, data processing, and evaluation logic, while the frontend ensures seamless interaction for candidates and recruiters.

The system operates in four primary phases:

- (i) **Input Phase** — candidates register, upload resumes, and initiate the interview session.
- (ii) **Interaction Phase** — the system conducts interviews using dynamically generated questions.
- (iii) **Analysis Phase** — responses are processed using NLP, speech analysis, and computer vision techniques.
- (iv) **Evaluation Phase** — machine learning models generate scores, rankings, and detailed reports.

B. AI Interview Generation Module (Adaptive Questioning)

This module is responsible for conducting intelligent and interactive interviews. It uses **Generative AI models** to create context-aware questions based on job roles, candidate profiles, and previous responses.

Text prompts and candidate inputs are processed using NLP techniques to understand intent and relevance. Based on this, the system dynamically adjusts the difficulty level and direction of questioning, ensuring a personalized interview experience.

Key features include:

- Dynamic question generation
- Context-aware follow-up questions
- Domain-specific interview customization
- Real-time interaction handling

This module eliminates the need for manual interviewers in early screening stages while maintaining a natural conversational flow. In addition to adaptive questioning, the module incorporates a feedback-driven learning mechanism that continuously refines the interview process based on candidate interactions and historical data. By analyzing response patterns, response time, and confidence indicators, the system can identify

knowledge gaps and adjust subsequent questions accordingly. This not only improves the accuracy of assessment but also ensures that candidates are evaluated across a balanced range of competencies. Furthermore, the module supports multilingual interaction and can tailor questions to different experience levels, making it suitable for diverse recruitment scenarios. This intelligent adaptability enhances both fairness and efficiency, creating a more engaging and realistic interview environment.

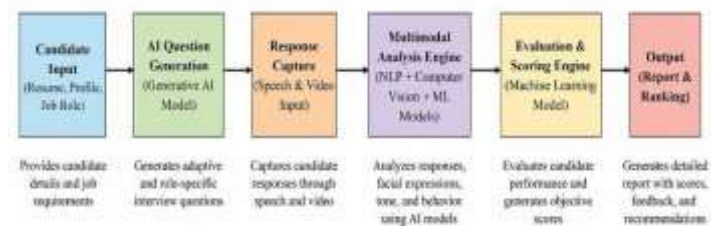


Fig.1. Architecture Diagram of Chayan AI

C. Multimodal Candidate Analysis Module

This module acts as the core intelligence layer, analyzing candidates across multiple dimensions using different AI techniques.

The processing pipeline includes:

- (i) **Speech and Text Analysis** — candidate responses are converted from speech to text using ASR models and evaluated using NLP for grammar, relevance, keyword matching, and sentiment.
- (ii) **Facial Expression Analysis** — computer vision models detect facial landmarks, eye movement, and expressions to assess confidence, engagement, and emotional state.
- (iii) **Behavioral Pattern Recognition** — combines verbal and non-verbal cues to identify hesitation, stress levels, and communication effectiveness.
- (iv) **Feature Fusion** — outputs from all modalities are combined into a unified feature representation for accurate evaluation.

This multimodal approach ensures a **holistic understanding of candidate performance**, going beyond traditional evaluation methods.

D. Evaluation and Decision-Making Module

This module generates final candidate scores and insights based on analyzed data. Machine learning models process the extracted features to produce

objective and consistent evaluations.

The evaluation pipeline includes:

- Performance scoring based on predefined metrics
- Candidate ranking and comparison
- Bias minimization through standardized models
- Automated feedback generation

The system ensures fairness by relying on data-driven decision-making rather than subjective human judgment.

E. Automated Reporting and Visualization Module

This module generates detailed reports summarizing candidate performance. The reports include:

- Communication skill analysis
- Technical competency evaluation
- Behavioral insights
- Overall performance score and ranking

The system presents results through dashboards and visualizations, enabling recruiters to make quick and informed decisions. In addition to static reports, the module incorporates interactive analytics and comparative insights to support deeper decision-making. Recruiters can visualize candidate performance trends through graphs, score distributions, and benchmarking against predefined criteria or peer groups. The system also highlights strengths, weaknesses, and improvement areas using explainable AI techniques, making the evaluation process more transparent and interpretable. Furthermore, reports can be exported in multiple formats and securely shared across hiring teams, ensuring seamless collaboration and efficient recruitment workflows.

F. Technology Stack

Table I summarizes the key technologies used in the proposed system.

Table I. Technology Stack of SyncVoice AI

Category	Tools / Frameworks
Frontend	React.js, HTML, CSS, JavaScript
Backend	Python (Flask / FastAPI), Node.js
NLP	SpaCy, NLTK, Transformers

Machine Learning	Scikit-learn, TensorFlow
Computer Vision	OpenCV, MediaPipe
Generative AI	GPT-based APIs
Database	MongoDB / Firebase
Deployment	Docker, Cloud (AWS / GCP)

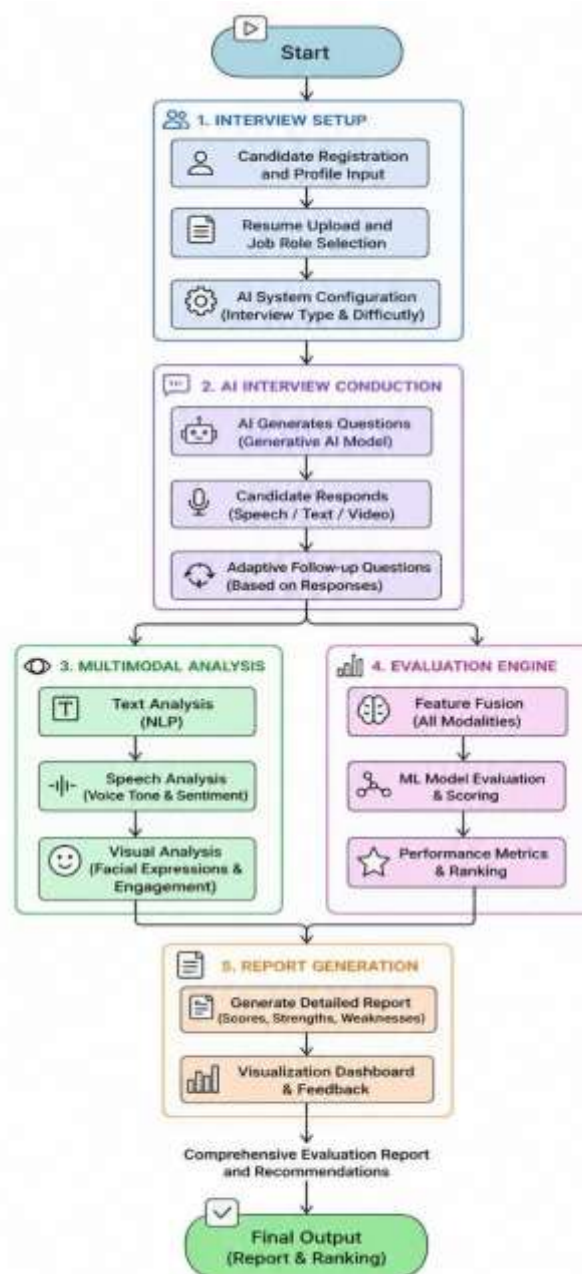


Fig.2. Workflow structure of CHAYAN AI

CONCLUSION

This work presented Chayan AI, an intelligent interview and candidate evaluation platform that integrates advanced artificial intelligence techniques to modernize and streamline the recruitment process. By combining Natural Language Processing, Machine Learning, Computer Vision, and Generative AI within a unified framework, the proposed system effectively bridges the gap between traditional human-driven interviews and fully automated, data-driven hiring systems.

The developed platform enables end-to-end interview automation, from adaptive question generation to multimodal candidate analysis and final evaluation. By analyzing both verbal and non-verbal cues, the system provides a more comprehensive understanding of candidate performance compared to conventional methods. Experimental observations indicate improved consistency in evaluation, reduced bias, faster processing time, and enhanced user experience for both recruiters and candidates. The integration of generative AI allows dynamic and context-aware interviews, while machine learning models ensure objective and reliable scoring.

The automated evaluation and reporting pipeline eliminates the need for manual assessment, significantly reducing recruiter workload and improving decision-making efficiency. Leveraging modern technologies such as React-based interfaces, Python backend frameworks, and scalable cloud infrastructure, the system is modular, extensible, and adaptable to various recruitment scenarios across industries.

Despite its advantages, certain limitations remain. The accuracy of evaluation depends on the quality and diversity of training data, and real-time video analysis may vary across different hardware environments. Additionally, ethical considerations such as data privacy, fairness, and transparency must be carefully addressed through secure data handling, bias mitigation techniques, and explainable AI models.

Looking ahead, future enhancements may include integration with job portals and professional networks, advanced emotion and personality analysis, multilingual interview support, and real-time cheating detection mechanisms. Incorporating explainable AI and continuous learning models can further improve trust and adaptability in the system.

In summary, **Chayan AI** successfully demonstrates how a unified, multimodal AI framework can transform traditional hiring practices into a more scientific, unbiased, and efficient process. The

proposed system lays a strong foundation for next-generation intelligent recruitment platforms, enabling smarter decision-making and more equitable opportunities in the evolving digital workforce.

ACKNOWLEDGEMENT

The author would like to express sincere gratitude to Babu Banarasi-Das Institute of Technology and Management (BBDITM), Lucknow, Department of **Information Technology & Engineering**, for providing the academic support, research environment, and technical resources essential for developing this project. The infrastructure and guidance offered by the institute played a crucial role in the successful completion of this work on intelligent audiovisual synthesis and talking-face generation.

The author also extends heartfelt thanks to the faculty members and fellow students whose valuable feedback, constructive suggestions, and continuous encouragement greatly improved the quality, clarity, and practical relevance of this research. Their support was instrumental in refining the integration of deep-learning models with real-world multimedia applications.

The author conveys special appreciation to Dr. Usha Sharma, Associate Professor, Department of IT, BBDITM, for her consistent guidance, insightful recommendations, and academic mentorship throughout the duration of this project. Her direction was fundamental in shaping the methodology, strengthening the technical framework, and ensuring the successful execution of this research.

REFERENCES

- [1] Sharma, P., & Gupta, N., 2025, *AI-Based Automated Interview Systems for Smart Recruitment*, International Journal of Engineering Research & Technology (IJERT), Volume 14, Issue 07.
- [2] Wang, X., Li, Y., & Chen, Z., 2025, *Intelligent Candidate Evaluation Using Multimodal Machine Learning*, arXiv preprint arXiv:2503.11234.
- [3] Kumar, A., Singh, R., & Verma, S., 2025, *AI-Driven Hiring Platforms with Bias Reduction Techniques*, IEEE International Conference on Smart Computing Systems, pp. 210-218.
- [4] Dutta, S., et al., 2025, *Automated Interview Analysis Using NLP and Computer Vision*, IEEE

Conference on Artificial Intelligence Applications (WACV), 2025.

[5] Liu, Y., Fang, Y., & Lin, Z., 2024, *Multimodal AI Systems for Behavioral Analysis in Recruitment*, arXiv preprint arXiv:2409.05221.

[6] Jang, Y., et al., 2024, *Joint Speech and Visual Analysis for Intelligent Human Interaction Systems*, Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024.

[7] Sun, Y., et al., 2024, *AI-Based Human Behaviour Understanding Using Multimodal Data*, IEEE Access, Vol. 12, pp. 57288–57301.

[8] Chen, S., et al., 2024, *Speech and Facial Expression Analysis for Automated Interviews*, arXiv preprint arXiv:2401.09951.

[9] Yaman, D., et al., 2024, *Video-Based Candidate Assessment with Deep Learning Models*, European Conference on Computer Vision (ECCV), Springer, 2024.

[10] Biadisy, F., et al., 2024, *Cross-Lingual Speech Understanding for AI Systems*, arXiv preprint arXiv:2409.13910.

[11] Choi, J., et al., 2024, *Unified Audio-Visual Representation Learning for Human Interaction Analysis*, CVPR 2024.

[12] Ghosh, S., et al., 2024, *Multimodal Deep Learning for Human Behavior Prediction*, Applied Intelligence, Vol. 54, pp. 4507–4524.

[13] Reddy, N. B., et al., 2023, *Machine Learning Approaches in Recruitment Systems*, E3S Web of Conferences, Vol. 430.

[14] Milis, G., et al., 2023, *Neural Conversational Models for Automated Interviews*, arXiv preprint arXiv:2312.06613.

[15] Brown, T., et al., 2023, *Generative AI Models for Human-Computer Interaction*, Proceedings of ICCV 2023.

[16] Zhang, W., et al., 2023, *Deep Learning-Based Behavioral Analysis Systems*, IEEE Conference Proceedings, 2023.

[17] Lu, J., et al., 2022, *Speech Processing and NLP Techniques for AI Applications*, ICASSP 2022, IEEE.

[18] Tamboli, S., Raut, P., & Kawane, A., 2022, *A Review Paper on AI-Based Text Processing Systems*, ICOET 2022.

[19] Song, H. K., et al., 2022, *Multilingual AI Systems*

for Human Interaction Analysis, CVPR 2022, pp. 21425–21430.

[20] Ren, Y., et al., 2021, *FastSpeech 2: Fast and High-Quality End-to-End Speech Processing*, arXiv preprint.

[21] Kim, J., et al., 2020, *Glow-TTS: Generative Flow for Speech Processing*, Advances in Neural Information Processing Systems, Vol. 33.

[22] Radford, A., et al., 2021, *Learning Transferable AI Models from Natural Language Supervision*, ICML, PMLR.

[23] Prajwal, K. R., et al., 2020, *Deep Learning Models for Speech and Visual Synchronization*, ACM Multimedia Conference, pp. 484–492.

[24] Ramani, A., et al., 2020, *Automatic Subtitle Generation Using AI*, ICACCS 2020, pp. 132–135.

[25] Jamaludin, A., et al., 2019, *Speech2Vid: Talking Face Generation from Speech Audio*, International Journal of Computer Vision, Vol. 127.

[26] Shen, J., et al., 2018, *Neural Text-to-Speech Synthesis Using WaveNet*, IEEE Conference, pp. 4779–4783.