iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79777

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Intraview - Virtual Human Interview System

Mrs. C. Maheswari[1], Kasba Sanjit[2], Nehal Abdul Gani[3], Borugadda Karthik[4]
[1]*Assistant Professor, Department of Computer Science and Engineering, Methodist College of Engineering and Technology, Hyderabad, Telangana, India.*
[2, 3, 4]*Student, Department of Computer Science and Engineering, Methodist College of Engineering and Technology, Hyderabad, Telangana, India.*

Abstract: ***Over the years, virtual interview tools have evolved significantly. What started as rigid and unchangeable platforms has now developed into systems that allow people to communicate more naturally, even in remote settings, while helping them build real-world skills. IntraView takes this idea a step further. It is an AI-based interview platform that analyzes a user's speech and provides feedback on aspects such as clarity, fluency, confidence, pacing, and overall effectiveness of communication. Unlike traditional tools that focus only on whether responses are correct, IntraView also considers behavioral cues such as hesitation, inconsistency, and loss of confidence, all of which play an important role in real interview scenarios. The system is designed for two main user groups. For students, it provides a safe environment where they can practice interviews multiple times, gradually reducing anxiety and improving confidence. It also highlights speaking patterns and communication habits, helping users understand where they need improvement. For organizations, the system offers an improvement over existing AI-based screening tools, which are often limited to text-based evaluation and lack real interaction. By analyzing conversational flow and behavioral patterns, IntraView enables more informed and reliable candidate assessment. The system focuses on speech-based interaction and behavioral evaluation. Currently, it primarily uses audio as the main input, but the architecture allows easy extension to features such as facial expression analysis and customizable interviewer behavior. The primary goal above everything else, is to make the interview process a natural experience. When you concentrate on your genuine method of communication, it not only calms your fears, allows you to gain self-assurance, but also makes preparing for an AI interview really useful.***

Keywords: Virtual Interview Systems, Behavioral Speech Analysis, Adaptive Dialogue System, Communication Assessment, Interview Preparation, Real-Time Interaction, AI-Based Evaluation, Emotion Recognition.

I. INTRODUCTION

The stakes are high in the contemporary era, particularly in highly competitive industries such as engineering and computing, which are constantly changing. Proficiency is not sufficient to ensure success in getting internships, employment, and careers. There is an increasing

requirement for students and graduates to have excellent communicative and presentation skills as demanded by the recruiters and professors alike. The interview process itself has become a difficult hurdle to overcome since it now requires a candidate to show the ability to think logically, maintain calm under pressure, and be able to express oneself confidently. As a result, there is a lot of emphasis on acquiring delicate interpersonal skills, such as vocal tone, speaking pace, eloquence, and flexibility in expression, which adds to the students' anxiety. The core issue, however, is that students rarely get opportunities to develop such qualities in settings that emulate the dynamics of a real-life situation such as an interview. Traditional interviewing preparation techniques fall short in addressing such issues. Group discussions, question banks, and peer reviews concentrate primarily on ensuring whether one's response to a question is accurate without considering their delivery process.

However, in the practical scenario, the interviewer looks for the smallest details such as the steadiness of one's voice, speech flow, and the absence of pauses and hesitations. Small pauses, tremble of one's voice, or nervousness can cost much since there would be no chance to give the desired impression. Without proper feedback, the candidate goes to an interview without being fully prepared, which leads to various mistakes due to not knowing what is expected. The lack of practice and structured feedback is likely to increase nervousness and uncertainty.

The same problem concerns organizations looking for good candidates among many others. To speed up the process, they employ automatic tools for testing candidates. These systems are restricted; therefore, they usually offer only the standard type of testing, like text-based test or voice-based. These tools cannot provide all the necessary data and cannot evaluate candidates' behavior and communication skills while interviewing.

Indeed, there is a demand for enhanced platforms that mimic interview scenarios and analyze communication abilities more thoroughly. IntraView meets these criteria. This is a platform powered by AI technology that offers a lifelike interactive experience and advanced speech analysis. How does it work? Simply put, you speak out your answers, and the software transcribes them using Whisper, a state-of-the-art speech-to-text model that accounts for diverse accents, speech characteristics, and audio artifacts. Next, the transcription enters the local large language model that creates appropriate interview questions. This allows for an organic flow of conversation as opposed to fixed Q&A sessions.

What makes IntraView particularly unique is its ability to conduct comprehensive behavioral speech analysis. Contrary to other tools that concentrate on whether candidates respond correctly, it analyzes how candidates communicate, paying attention to such aspects as clarity, articulation, pace, speed, consistency, hesitations, and emotions. Therefore, even minor symptoms of nervousness or poor preparation can be identified and used to generate helpful feedback.

With this approach, individuals will be able to practice repeatedly in an environment that is devoid of the pressure associated with live interviews. As they use it, their levels of anxiety will diminish, and they will develop excellent interpersonal and intrapersonal skills and confidence. The system will serve as a reflective learning instrument where learners will identify their shortcomings, keep tabs on their improvement, and learn through experience.

In terms of organizational benefits, IntraView will bridge the gaps that have been experienced by other automated recruitment tools. It introduces the element of behavioral assessment as well as interaction and dialogue into the process, thereby offering recruiters detailed information on the emotional and mental well-being of applicants as well as their ability to communicate effectively and consistently. The result will be unbiased and accurate selection at the initial stages of the recruitment process.

The platform is modular, which means that new components may be introduced in response to technological advancements without causing disruptions to its operations. IntraView is adaptable, scalable, and future-oriented. At present, it integrates conversational artificial intelligence with speech-based behavioral assessment. However, it has the capacity to accommodate new features such as advanced emotional recognition and enhanced performance measurement.

IntraView's primary focus is on addressing the lack of correlation between testing technical knowledge and communicating capabilities. Due to its emphasis on simulation and feedback throughout the process, it introduces a better approach that allows for a more efficient preparation for interviews. This tool helps create confident candidates while providing organizations with a valuable model for the assessment of communication in practical terms. Overall, such efforts can serve as a foundation for further developments and improvements of virtual interviews to ensure their interactive, data-rich nature.

Over time, the tool allows for more in-depth information about user performance during the interview process to be gathered. Thanks to data storage on a per session basis and tracking of recurring activities, specific behavioral patterns can be identified for each particular user. As a result, over time, it becomes possible to compare one's performance across several rounds in order to see both problems that persist and those that have been resolved.

Changing the interview experience may be just as important as asking different questions. Currently, answers depend on the interview context. However, future interview versions will take a more specialized approach – whether for coders, executives, or high-level technologists. Such an approach will allow people to train under circumstances that reflect their actual experiences in conversation. The difficulty levels should also increase together with the increasing skills. Therefore, training will become more accurate and similar to a real interview experience.

It can be more beneficial for machines to receive smarter answers. Instead of yes or no, such information can be divided into subparts that include the quality of the content, communication, and consistent action. One such benefit would be a possibility to classify comments based on individual needs. Then, personal recommendations can be generated automatically depending on individual needs rather than providing vague guidelines. The more details are included, the more effective practice will become.

However, reducing response times still needs attention to make sure that the process runs smoothly. Due to the highly interactive nature of the program, any delay could ruin the experience completely. With optimized algorithms, better data transfer, and improved processing, more time will be saved in each process. The faster everything runs in the background, the easier it becomes to ensure that user intents are met properly. Unlike the original version, this modification does not require radical changes to start. Rather, it will be possible to adapt it to the classroom settings, where students would use the program alongside other lessons to monitor their progress and provide additional assistance. In this way, the program would become part of a new form of training combining machine guidance and human expertise. The biggest advantage of the program is its versatility and ability to switch between modes based on requirements. In addition to preparing people for job interviews, it will allow them to develop public speaking and conversational skills. It will be enough to adjust response rules slightly to create an all-in-one solution that would analyze speech and divide it into parts based on various criteria.

In addition to performing its current functions, the platform will also help people to learn about the evolution of their interviewing style. Through practicing and doing interviews regularly, people are bound to recognize some mistakes they may make during their performance and become consistent while speaking in interviews. It helps turn the system into an effective tool for practicing and learning at the same time.

An additional benefit of the proposed solution is its versatility. The fact that the system uses a modular structure makes it possible to easily expand its capabilities. For example, IntraView will enable preparing for interviews of all kinds. Moreover, over time, other features, such as giving more detailed feedback, handling interactions effectively, and communicating differently, may be implemented into the system without changing its basic functioning.

This will help the users develop an awareness about their speaking styles in different contexts, which will make their preparation more realistic and practical.

II. LITERATURE SURVEY

S. No.	Title of the Paper	Year	Objective	Methods Used	Key Findings
1	Emotion Recognition in Human–Computer Interaction	2001	To explore emotion detection for HCI	Feature engineering, early ML methods	Emotion-aware systems improve user experience and adaptability
2	Emotion Recognition From Speech: A Review	2012	To survey methods for SER	Review of ML, signal processing, prosody analysis	Overview of SER challenges, datasets, and model limitations
3	Using Regional Saliency for Speech Emotion Recognition	2017	To improve SER using feature saliency	Regional saliency, feature extraction	Saliency-based features enhance emotion-classification accuracy
4	A Lip Sync Expert Is All You Need for Speech-to-Lip Generation	2020	To generate accurate lip movements from speech	Lip-sync neural networks, audio-visual modeling	Achieves high-quality lip-sync in real-world conditions
5	Audio2Head: Audio-Driven One-Shot Talking-Head Generation	2021	To animate talking heads from a single image + audio	One-shot learning, audio-driven modeling	Generates realistic head motion from audio cues
6	An AI Mock-Interview Platform for Interview Performance Analysis	2022	To evaluate interview performance using AI analytics	Performance scoring, behavioral analysis, ML models	AI improves consistency and accuracy in interview performance evaluation
7	AI Mock-Interview Platform for Interview Performance Analysis	2022	To analyze performance using AI metrics	Behavioral metrics, ML analysis	AI improves fairness and accuracy in interview assessment
8	AI Powered Mock Interview System with Real-Time Voice and Emotion Analysis	2023	To provide real-time emotional + vocal feedback	SER, voice analytics	Improves self-awareness and emotional control in interviews
9	Using a Virtual Reality Interview Simulator to Explore Factors Influencing Behavior	2023	To study interview behaviors in VR	VR metrics, behavioral tracking	VR reveals deeper behavioral patterns under stress
10	Speech Emotion Recognition and Serious Games: Crowdsourcing Annotated Samples	2023	To collect SER data through game-based tasks	Serious games, crowdsourced SER labeling	Gamified SER improves data quantity and labeling quality

S. No.	Title of the Paper	Year	Objective	Methods Used	Key Findings
11	AI Powered Mock Interview System with Real-Time Voice and Emotion Analysis	2024	To simulate interviews with real-time voice/emotion feedback	Speech analysis, emotion detection, ML classification	Real-time cues help boost user confidence and response quality
13	Advancing Virtual Interviews: AI-Driven Facial Emotion Recognition for Better Recruitment	2024	To enhance interview assessment using facial emotions	Facial emotion recognition, computer vision	Facial cues improve prediction of candidate suitability
14	Avatar-Based Feedback in Job Interview Training Impacts Action Identities and Anxiety	2024	To reduce anxiety and improve interview skills via avatar feedback	Avatar-based feedback, affective computing	Avatar feedback reduces anxiety and improves performance readiness
15	Emotional Analysis of Candidates During Online Interviews	2024	To analyze emotional states during online interviews	Emotion classification, multimodal analysis	Emotional cues help predict candidate confidence and sincerity
16	AI-Driven Facial Emotion Recognition for Better Recruitment	2024	To evaluate candidates via facial emotion signals	FER models, deep learning	Enhances reliability of virtual interview evaluation
17	Avatar-Based Feedback in Job Interview Training	2024	To test feedback effects via avatars	Affective computing, avatar modeling	Avatar interaction reduces anxiety and enhances communication skills
18	NexInterview – AI-Driven Mock Interview Preparation Platform	2025	To provide AI-based personalized mock interview practice	NLP, scoring models, analytics	AI-based systems provide structured feedback and improve preparedness
19	AI-Driven Approach to Enhance Interview Performance Through Voice and Response Analysis	2025	To analyze voice and verbal delivery for performance scoring	Voice feature extraction, sentiment analysis	Voice patterns strongly correlate with interview success
20	Detection of Emotion Using Multi-Block Deep Learning in a Self-Management Interview App	2025	To detect user emotions in interview training apps	Multi-block deep learning, CNN models	Deep learning improves emotion recognition accuracy
21	Speech Emotion Recognition and Serious Games: Job Interview Simulator Case Study	2025	To integrate SER into gamified interview training	Speech emotion recognition, serious-games framework	Gamification increases engagement and emotional learning
22	Virtual Interview Simulator Leveraging AI/ML and Vision Technology	2025	To simulate interviews using AI and vision analytics	Face tracking, ML scoring, pattern analysis	Vision-based evaluation improves accuracy of feedback
23	Web-Based AI Mock Interview System Using ChatGPT with Real-Time Voice Interaction	2025	To create voice-interactive AI interview practice	ChatGPT, speech-to-text, NLP	Real-time AI conversation improves speaking fluency and relevance
24	AI Interview Preparation Web App	2025	To help users prepare for interviews using automation	NLP questioning, automated scoring	AI-based preparation increases user confidence and readiness
25	AI-Powered Mock Interview System for Automated Skill Assessment	2025	To measure skills using AI interview simulations	Skill-matching, ML scoring engine	Automated skill assessment reduces human evaluator bias

III. PROBLEM STATEMENT

The acquisition of communication skills necessary for an interview is complicated for both students and potential employees since most mock interview exercises are concentrated on testing the correctness of the response rather than its manner. It is hard to receive objective feedback from standard mock interview systems on such critical elements as tone, pace, articulateness, and even emotional stability in many cases. Therefore, the lack of feedback makes people go to their job interviews insecure and uncertain of their communication skills.

In addition, modern recruitment practices imply the usage of automated systems that help identify the level of qualifications of candidates through text messages or recordings. This technique does not provide any opportunity for conversational experience that implies some kind of adaptiveness. Consequently, it misses behavioral details since the process takes place under non-realistic conditions.

Hence, it is possible to suggest that there should be an AI-driven virtual interview platform that will imitate the situation of a genuine interview and assess behavior in a more efficient way. Such a platform needs to analyze speech recognition, generate interactions based on the content, as well as evaluate some vocal features to recognize behavioral strengths and weaknesses of the candidate.

IV. PROPOSED SYSTEM

IntraView is an advanced AI-based platform designed to conduct realistic interviews and provide structured feedback regarding your communication skills. While most tools evaluate only whether you give correct answers to questions asked during interviews, IntraView focuses on the quality of these responses and deviates from the conventional approach to mock interviews.

How does IntraView work? First of all, users should give their answers to interview questions verbally. Then, the software processes recorded voice by applying speech-to-text technology for transcription. Based on this transcript, the platform generates new questions using a language model in order to provide a natural conversation flow. As can be seen, unlike many existing tools where questions are predetermined, IntraView allows generating questions dynamically, depending on users' responses.

As was already mentioned above, IntraView features an advanced vocal analysis module aimed at evaluating the clarity, fluency, pace, and confidence of verbal communication. The key idea behind this feature consists of recognizing behavioral patterns that might have a negative impact on interview results and transforming them into recommendations that are actionable. From the perspective of design architecture, IntraView has a modular nature, implying that speech recognition, language processing, and emotion analysis modules function independently, yet complementing each other.

In essence, IntraView seeks to offer a comprehensive preparation process by combining artificial intelligence with behavioral analysis. IntraView offers a platform where users can engage in conversation and receive objective evaluations that will help them hone their communication skills and prepare for actual interviews.

V. METHODOLOGY

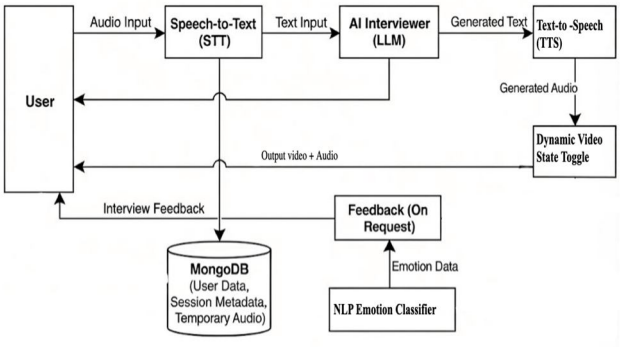


Fig – 1 System Architecture

IntraView emerges bit by bit, guided by definite decisions rather than speculation, developed with the intention of eliciting natural responses, facilitating natural dialogue, and giving appropriate feedback. Each step becomes integrated - needs analyzed first, design second, resources carefully selected, components assembled slowly, trials repeated numerous times. The pace is unhurried because

each layer responds to practical considerations rather than abstract principles. Ideas that are viable on paper should remain realistic during face-to-face conversation. Real communication requires preparedness - such preparedness emerges only when the process adheres to proper procedures, refusing to anticipate future steps

As illustrated by Figure 1, the system incorporates discrete components that interact within a coordinated sequence. Rather than relying on one huge unit, it divides into different processes - speech processing gives way to language comprehension, followed by behavior analysis. Designed for practical purposes, the system maintains its responsiveness while allowing for flexibility and development. While each component operates independently, their connection remains seamless enough to ensure smooth flow between inputs. Subsequent modifications will not require extensive changes throughout the rest of the components because the modules connect effectively.

Initially, the objective was breaking down general requirements into tangible tasks that involved making conversations seem realistic and providing feedback that relates to oral communication abilities. The initial step was real-time audio recording, which was succeeded by spoken words transcriptions, intelligent answers generation that takes context into account, and smooth conversation flow ensuring absence of any pauses. Simultaneously, consideration began focusing on voice quality evaluation, namely voice clarity, speed of uttering sounds, pauses, fluency, and confidence. Each technology works independently yet supplies its output to a single process that determines how components interact within an entire system.

The web-based interface provides control over the system operation and incorporates HTML, CSS, and JavaScript technologies used for user interactions through buttons, microphones, and visual displays. In turn, another component transfers information between modules responsible for data processing, managing concurrent actions. Components communicate with each other online in a way that guarantees small response times due to a specific architecture.

The spoken words are processed by an advanced technology based on Whisper which transcribes them accurately. Voices of any type or accent can be recognized regardless of noise in the environment. Transcribed text is transferred to another application which resides locally and which is a large-scale language model. This language model generates questions for future use and analyzes what was said in the course of a conversation to remember it. As the model reacts anew to every query made, conversations become more fluid because of the lack of automation.

Besides controlling the flow of the dialogue, technologies also identify emotions in the conversations using audio clues. How does it work? The clarity of speech, its rhythm, accuracy, tempo, and tone reveal patterns associated with certain emotions. For instance, there are patterns related to dominance, while other patterns reveal uncertainty. Small pauses are also taken into account when analyzing conversations in real time. All these factors help to draw conclusions concerning speaking patterns and the areas where users demonstrate great skills and require improvement respectively.

From the very beginning, the process of communication occurs in uninterrupted mode, much like real-time timing. Once the user says something in the application, this sound gets transmitted to the server immediately. First and foremost, this sound gets transformed into text by means of a speech decoder. Instantly thereafter, this text splits into two streams – one stream goes to the artificial intelligence processor that generates the response while the other is analyzed by an emotion analyzer to detect the appropriate mood. Next, the response is delivered to the screen for viewing within a few seconds. In parallel, the behavioral traits are silently recorded in the loop of machine learning. Throughout all these operations, previously received communication remains stored to form the basis for future actions without pauses. This entire mechanism persists till the conversation stops spontaneously

Each stage of development was realized by gradual trials and minor changes. Initially, the framework of interaction was created linking together modules of voice-to-text conversion and natural language processing. Then the behavioral pattern detector was added and tuned up to be efficient enough but not slow down the process.

The process of testing the performance of the system and its durability took place on several levels. The stability of separate elements and the accuracy of their performance was tested individually in each module. The integration of modules was also thoroughly tested; the analysis of each operation stage was conducted in order to make sure that all operations will be completed successfully. The speed of operations is another important aspect of performance, which was carefully analyzed. Delays between different stages were considered. Quality feedback was achieved only after several rounds of adjusting the system. The improvement of speed does not lead to loss of accuracy, since the parameters were adjusted based on the results of tests. The final version of the system demonstrates high stability and durability in the most complex situations and provides continuous conversation without failures.

VI. TECHNOLOGIES USED

The creation of IntraView involves integrating various technologies that function as a whole for conducting live chats, processing voices, as well as splitting messages. Every component is only acceptable if it is fast, scalable and effectively connects with all other components around it.

Developed with such fundamental web technologies as HTML, CSS, and JavaScript, the client side determines the interaction process of users. Via these elements, various interactions happen, such as voice recording and the ability to update screens instantly. The adaptive layout is flexible and easy to adopt to multiple devices and ensures clarity. Efficiency is vital during a real-time question answering.

There is a lightweight server-side architecture that acts as a bridge between the elements of the system. Unlike the other approaches, it operates by advancing the task and controlling the process of audio stream. Requests are handled sequentially via interconnected artificial intelligence components. The data of each session remains active.

Whisper transcribes the spoken language to text irrespective of how quickly people speak or how strong their accents are. The process is also resilient to the influence of background noises since the software adapts in real-time to the speaker's needs during the interviews. The program runs directly on the user's device and relies on a sophisticated language algorithm based on technology such as Llama to provide the appropriate prompts and generate conversation flow. The script does not rely on prepared questions but rather dynamically shapes each query depending on the context. It also adjusts the prompts in response to conversational deviations for smoother dialogue continuity.

Based on the analyzed audio file, a special tool evaluates the way an individual speaks. It assesses vocal qualities instead of analyzing the words used by a person. As a result, emotions and feelings revealed through pitch, rhythm, confidence expressed through the pace, fluency indicated through pausing, clarity determined by sharpness, and tone shifts are identified.

In addition to request handling, the system requires a database, such as MongoDB, to keep records of users, active sessions, and actions taken during the process. As a result, thanks to the database, quick access is ensured, which allows forming responses based on previous interactions. All this technology allows for running live interviews and practices within IntraView that scale depending on the needs of users, being capable of providing automatic responses to requests. Speech recognition interacts with chat intelligence as well as user behavior analysis; all of them are integrated into one system.

VII. FUTURE ENCHANCEMENTS

Presently, IntraView serves excellently as a developing environment for artificial intelligence-driven interviews and communication assessments. However, some features can be improved to enhance immersion and effectiveness

Facial or hand gesture analysis conducted by a camera as one speaks. Rather than only listening to the interviewee, the software will monitor glances, smirk faces, and even shrugging shoulders. This is because a quick look away can mean hesitation. On the other hand, nodding indicates agreement. Both body language indicators contribute greatly to evaluating the interview skills. Whereas voice intonations show one side of the story, body language tells another side that voice alone cannot tell.

The improvement of conversation management can lead to further enhancements. The current responses are quite good but future models may perform better when handling multi-response situations. Also, follow-ups after user queries can be easily accomplished in the next models. It is highly possible to achieve advanced comprehension in specialized areas such as job interviews

As model accuracy and data size improve, emotion recognition becomes better. Subtle changes in language become more visible. That way, feedback regarding speaking behavior becomes more accurate

The scope can be widened when more languages are involved. This opens up opportunities for different users. The interviews conducted in one's own language flow freely once the process is integrated. It comes naturally if there are varied linguistic requirements in one region. It becomes easier if there are differences due to more choices available

One level up may involve more detailed feedback features. Feedback in form of progress report may be provided regarding performance within multiple sessions. It would indicate the areas where improvement is needed. Progress becomes easier if it becomes visible with each passing day. Skills develop not in jumps but in glimpses

Slowly but surely, IntraView will start feeling more realistic due to improved visuals and interaction capabilities. Over time, better analysis functions will emerge, making it easier for users to see their performance trends. Device compatibility will be achieved by simplifying the interface and supporting multiple platforms. All these improvements are small but accumulate over time, allowing users to experience a practice session almost as realistic as an interview itself. The objective remains consistent throughout: create a tool that easily integrates into the preparation process.

VIII. CONCLUSION

Imagine a platform known as IntraView. The platform makes use of artificial intelligence technology to create simulated job interviews that other practice tools cannot offer. Unlike other platforms, the system is not only able to check whether the answers provided are correct but also to listen keenly. The speech is then converted to text format as the interaction continues. Emotional analysis comes next as the tone, pace, and pauses are considered. The process is therefore more interactive, and the participants feel as if they are having an actual conversation rather than an assessment. The clarity with which the message is conveyed becomes equally important as its content.

The integration of different components of artificial intelligence is seen in a particular setting. Since the process involves both interacting and observing at the same time, the individual is able to conduct job interviews repeatedly without losing track of their progress. There is thus continuous feedback on the style of communication, including its fluency, rhythm, and presence.

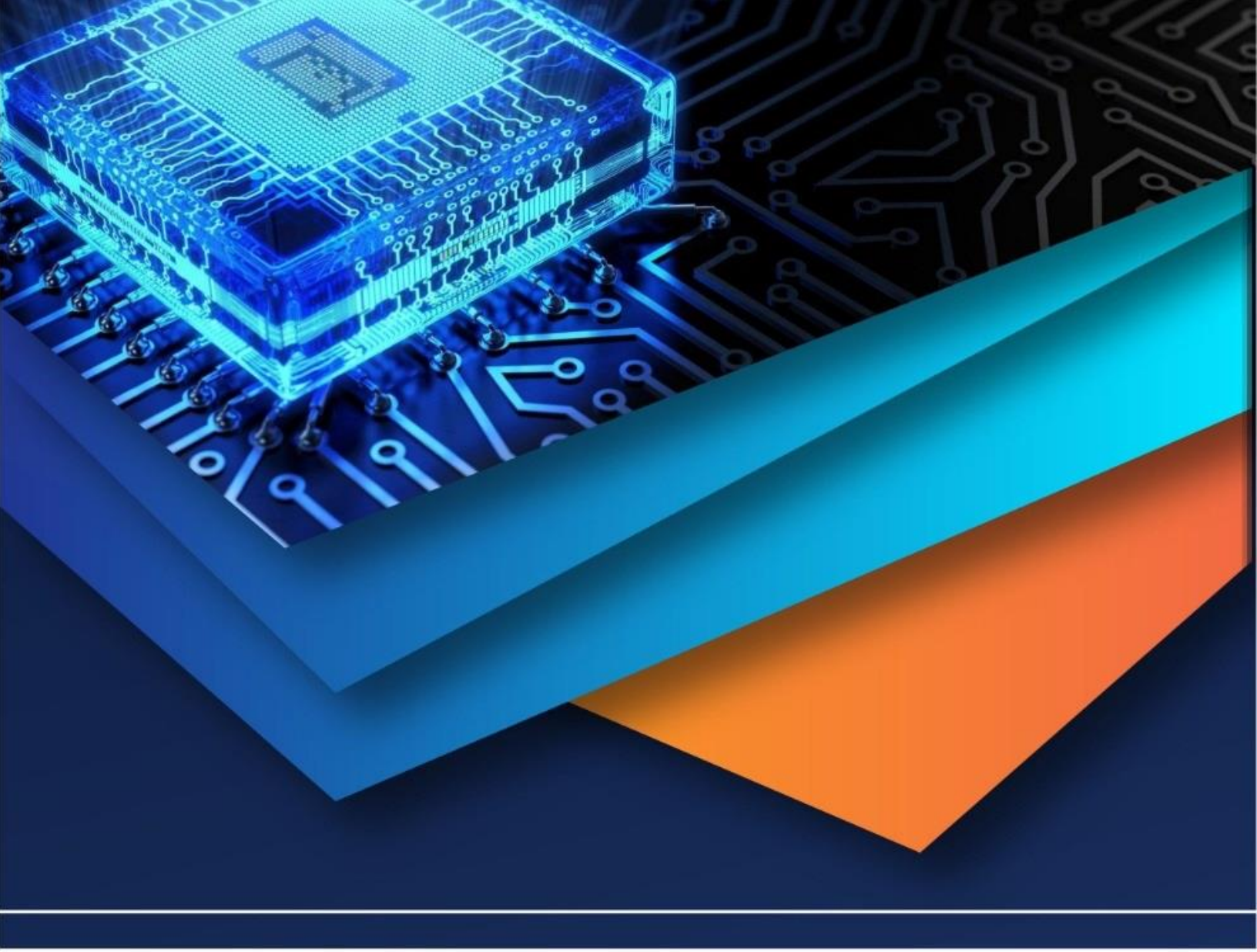
Benefit number one involves the flexibility involved in its construction as it evolves according to recruitment needs, leading to more reliable screening in the first round. It can detect behavioral traits, which other conventional methods ignore, thus making decisions more informed when screening candidates. The most obvious advantage involves improved reliability in conducting these checks due to the ability to pick subtle signs that would be overlooked otherwise

The next advantage highlights a new approach in measuring abilities, taking into account the importance of verbal communication. In place of traditional tests, this process involves engaging job applicants through discussion in addition to gathering useful data for analysis purposes. As a result, job candidates are able to improve their communication skills, as well as receive additional feedback on suitability.

REFERENCES

[1] Y. Chou et al., “An AI Mock-interview Platform for Interview Performance Analysis,” IEEE Conference on Smart Computing, pp. 233–240, 2022.
[2] S. Kumar and R. Smith, “AI Powered Mock Interview System with Real-Time Voice and Emotion Analysis,” International Journal of Novel Research and Development (IJNRD), vol. 9, no. 2, pp. 210–215, 2024.
[3] D. Miller and A. Patrizio, “Using a Virtual Reality Interview Simulator to Explore Factors Affecting Performance,” Virtual Reality Journal (Springer), vol. 29, no. 3, pp. 421–434, 2024.
[4] H. Zhang et al., “Advancing Virtual Interviews: AI-Driven Facial Emotion Recognition for Better Recruitment,” IEEE Transactions on Affective Computing, vol. 15, no. 1, pp. 310–320, 2024.
[5] J. W. Burgoon et al., “Avatar-Based Feedback in Job Interview Training Impacts Action Identities and Anxiety,” IEEE Transactions on Affective Computing, vol. 15, no. 2, pp. 487–496, 2024.
[6] H. Lee and P. Kaur, “Emotional Analysis of Candidates During Online Interviews,” Proceedings of the Conference on Multimedia Systems (CMS), pp. 312–318, 2024.
[7] Dr. Prakash S et al., “NexInterview - AI-Driven Mock Interview Preparation Platform,” International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), vol. 5, no. 7, pp. 279–287, 2025.
[8] R. K. Verma and M. Singh, “An AI-Driven Approach to Enhance Interview Performance through Voice and Response Analysis,” Journal of Information Systems and Engineering Management (JISEM), vol. 9, no. 3, pp. 78–85, 2025.
[9] K. Choi and Y. Lee, “Detection of Emotion Using Multi-Block Deep Learning in a Self-Management Interview App,” Applied Sciences (MDPI), vol. 15, no. 1, pp. 1–14, 2025.
[10] R. Silva and E. Pereira, “Speech Emotion Recognition and Serious Games: A Job Interview Simulator Case Study,” Information (MDPI), vol. 16, no. 4, pp. 210–222, 2025.
[11] K. R. Prajwal, R. Mukhopadhyay, V. P. Namboodiri, and C. V. Jawahar, “A Lip Sync Expert Is All You Need for Speech to Lip Generation in the Wild,” Proc. ACM MM, 2020.
[12] Z. Aldeneh and E. M. Provost, “Using regional saliency for speech emotion recognition,” ICASSP, 2017, pp. 2741–2745.
[13] Shashidhar G. Koolagudi and K. S. Rao, “Emotion recognition from speech: A review,” Int. J. Speech Technology, vol. 15, no. 2, pp. 99–117, 2012.
[14] Suzhen Wang, “Audio2Head: Audio-driven One-shot Talking-head Generation with Natural Head Motion,” https://doi.org/10.24963/ijcai.2021/152
[15] R. Cowie et al., “Emotion recognition in human–computer interaction,” IEEE Signal Processing Magazine, vol. 18, no. 1, pp. 32–80, 2001.
[16] Sridevi R, Nithyabharathi S, “Virtual interview simulator: leveraging AI/ML and vision technology,” https://doi.org/10.22214/ijraset.2025.66680
[17] Minu Hwang, Insung Kwack, and Ugun Won, “Web-Based AI Mock Interview System Using ChatGPT with Real-Time Voice Interaction,” Engineering and Technology Journal, vol. 10, no. 10, 2025.
[18] Dr. Rais Abdul Hamid Khan, Garima Narendra Khatri, Aarti Ganesh More, and Utkarsha Pramod Mahajan, “AI Interview Preparation Web App,” IJNRD, vol. 10, no. 11, pp. 1–5, 2025.
[19] Vijayant Verma, Rana Padwar, Apurwa Chandrakar, Khushi Jaiswal, and Palak Mishra, “AI-Powered Mock Interview System for Automated Skill Assessment,” IJRASET, vol. 13, no. 11, 2025.
[20] Y. Chou et al., “An AI mock-interview platform for interview performance analysis,” SMARTCOMP, 2022.
[21] Sonu Khapekar, Srishti Bothara et al., “AI powered mock interview system with real-time voice and emotion analysis,” Int. J. of Novel Research and Development, vol. 10.
[22] D. Miller and A. Patrizio, “Using a virtual reality interview simulator to explore factors influencing people’s behavior,” Virtual Reality.

[23] Rohini Mehta et al., "Advancing virtual interviews: AI-driven facial emotion recognition for better recruitment," IEEE Trans. Affective Computing, vol. 15, no. 1, pp. 310–320, 2024.
[24] J. W. Burgoon et al., "Avatar-based feedback in job interview training impacts action identities and anxiety," 10.1109/TAFFC.2024.3363835.
[25] Eleni Siamtanidou et al., "Speech Emotion Recognition and Serious Games: An Entertaining Approach for Crowdsourcing Annotated Samples," https://doi.org/10.3390/info16030238



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY