

# AN AI-DRIVEN VOICE INTERACTION SYSTEM FOR REAL-TIME INTERVIEW SIMULATION AND SKILL ASSESSMENT

<sup>1</sup>M. Thanmayee, <sup>2</sup> K. Anilprasad Reddy , <sup>3</sup> J. Manikanta Swamy, <sup>4</sup> G. Soumya, <sup>5</sup> G. Sanjay

<sup>1</sup>Assistant Professor in Department of CSE TKR COLLEGE OF ENGINEERING & TECHNOLOGY

<sup>1</sup> [thanmayeetupurani@tkrcet.com](mailto:thanmayeetupurani@tkrcet.com)

<sup>2,3,4,5</sup>UG Scholars in Department of CSE TKR COLLEGE OF ENGINEERING & TECHNOLOGY

<sup>2</sup>[karraanilprasadreddy@gmail.com](mailto:karraanilprasadreddy@gmail.com) , <sup>3</sup>[jalamanikanta552@gmail.com](mailto:jalamanikanta552@gmail.com) , <sup>4</sup>[soumyaganta1874@gmail.com](mailto:soumyaganta1874@gmail.com) , <sup>5</sup>[gundasanjay667@gmail.com](mailto:gundasanjay667@gmail.com)

## Abstract

Getting through an interview successfully is often more challenging than it appears, even for candidates who are technically strong. Many individuals struggle not because they lack knowledge, but because they are not comfortable expressing their thoughts clearly under pressure. Existing preparation methods—such as reading guides or practicing with friends—do not fully capture the unpredictability and intensity of real interview situations. This work introduces a voice-based interview simulation system that uses artificial intelligence to create a more practical and engaging preparation experience. Instead of relying on text-based interaction, the system allows users to speak naturally with a virtual interviewer, making the session feel closer to a real conversation. It generates questions based on user preferences and listens to responses in real time, converting speech into text for further evaluation. The system analyzes each response by considering factors like how relevant the answer is, how clearly it is presented, and how confidently it is delivered. Based on this analysis, it provides simple and meaningful feedback that helps users understand where they are doing well and where they need improvement. The overall design focuses on making the platform easy to use, responsive, and accessible to a wide range of users. From initial observations, this approach not only helps users practice more effectively but also builds confidence over time. By offering a realistic and continuous learning environment, the system serves as a practical tool for improving interview performance and preparing individuals for real-world hiring scenarios.

## Keywords

*Artificial Intelligence, Voice-Based Interaction, Mock Interviews, Real-Time Evaluation, Interview Practice, Skill Assessment, Speech Processing, Communication Improvement, Job Preparation, Conversational Systems*

## I INTRODUCTION

Getting through an interview successfully is not always easy, even for candidates who are well-prepared in terms of technical knowledge. In many cases, the difficulty lies not in what they know, but in how they express it. Nervousness, hesitation, and lack of clarity often affect performance, especially when candidates are expected to respond quickly and confidently. Interviews require a combination of knowledge, communication skills, and the ability to think under pressure, which makes them challenging for many individuals [1].

Most people rely on common preparation methods such as reading interview questions, watching videos, or practicing with friends. While these methods can be helpful to some extent, they do not fully reflect the actual interview experience. Real interviews are usually unpredictable, interactive, and time-sensitive. Without exposure to such conditions, candidates may find it difficult to adapt during the actual process.

Another limitation is that these methods rarely provide clear or structured feedback, making it harder for individuals to identify their weak areas and improve effectively [2]. In recent years, artificial intelligence has started playing a noticeable role in education and skill development. AI-based systems are capable of creating interactive environments where users can learn by doing rather than just observing. These systems can generate questions, analyze responses, and provide feedback almost instantly. Because of this, they offer a more engaging way to practice and improve skills compared to traditional approaches [3]. The use of machine learning and language processing techniques allows such systems to understand responses in a more meaningful way and evaluate them across different aspects [4].

Another important development is the use of voice technology in interactive systems. Instead of typing answers, users can now speak naturally, which makes the experience feel closer to a real conversation. Technologies like speech-to-text and text-to-speech make it possible for systems to listen and respond in real time. This kind of interaction not only improves user engagement but also helps in evaluating how confidently and clearly a person speaks, which is an important part of interview performance [5].

In this work, an AI-driven voice interaction system is introduced to support interview preparation in a more practical way. The system allows users to participate in mock interviews by speaking to a virtual interviewer. It listens to their responses, processes them, and provides feedback based on factors such as clarity, relevance, and confidence. The aim is to create an environment where users can practice repeatedly and gradually improve their performance without feeling pressured [6].

The overall idea is to make interview preparation more realistic and accessible. By combining voice interaction with intelligent evaluation, the system helps users become more comfortable with expressing their thoughts in a structured way. Over time, this can lead to better confidence and improved communication skills. Such AI-based tools are expected to play an important role in bridging the gap between preparation and real-world performance, especially for students and job seekers [7]. As these technologies continue to evolve, they are likely to become an integral part of modern learning and recruitment processes [8]. Their ability to provide

consistent feedback and simulate real scenarios makes them a promising solution for improving interview readiness [9]. This shift towards intelligent and interactive systems reflects a broader trend in using AI for practical skill development and assessment [10].

## II LITERATURE SURVEY

Interview preparation has gradually become an area of interest in research, mainly because many candidates face difficulty in translating their knowledge into performance during actual interviews. A number of studies point out that the lack of structured practice and meaningful feedback is one of the main reasons for poor interview outcomes. Traditional preparation methods depend heavily on self-learning or informal practice, which may not always provide consistent results or clear direction for improvement [1].

Early attempts to address this issue introduced simple automated interview systems that could present a set of predefined questions and record user responses. While these systems helped users get familiar with common interview patterns, they were quite rigid in nature. The evaluation was mostly rule-based and did not capture the depth or quality of the responses. As a result, their usefulness was limited when it came to improving real communication skills [2].

With the introduction of natural language processing, more flexible systems started to emerge. These systems were able to go beyond keyword matching and could understand the general meaning of a response. This allowed for better evaluation of answers in terms of relevance and structure. Although this was a significant improvement, the interaction was still mostly text-based, which reduced the sense of realism during practice sessions [3].

Researchers then began exploring ways to make these systems more interactive. One direction was to include speech processing so that users could speak instead of typing. By using speech-to-text technologies, systems could capture spoken responses and analyze them further. This approach made the experience more natural and closer to an actual interview. It also opened up the possibility of evaluating speaking-related aspects such as fluency and confidence [4].

In parallel, there has been growing interest in using machine learning models to enhance the quality of

feedback. Instead of relying on fixed rules, these models learn from large amounts of data and can provide more personalized suggestions. Some systems started generating feedback that highlighted both strengths and weaknesses, helping users understand not just what went wrong but also how to improve [5].

More recent developments have focused on conversational AI, where the system behaves like a real interviewer. These systems can adjust questions based on previous answers, creating a more dynamic flow. This kind of adaptive questioning makes the interaction feel less mechanical and more realistic. It also allows users to experience different types of interview situations within a single platform [6].

There have also been attempts to include non-verbal analysis in interview systems. Some research explores the use of facial recognition and emotion detection to study user behavior during interviews. These systems try to capture signals such as facial expressions or hesitation patterns. While this adds another layer of evaluation, it also increases system complexity and raises concerns related to user privacy [7].

Another trend observed in the literature is the integration of multiple features into a single platform. Instead of focusing on only one aspect, newer systems combine question generation, response analysis, and performance tracking. This integrated approach helps in creating a more complete learning environment where users can practice regularly and monitor their progress over time [8].

Even with these advancements, certain challenges still remain. Some systems are not able to fully understand the deeper meaning or intent behind user responses, especially in complex or technical discussions. Others face issues related to response delay or accuracy in speech recognition, particularly when users have different accents or speaking styles [9].

### III RELATED WORK

Over the past few years, several approaches have been proposed to improve interview preparation using digital platforms. Early systems mainly focused on providing a set of predefined questions and allowing users to record or type their responses. These platforms helped users become familiar with common interview patterns, but the interaction was often limited and lacked realism. Since the

feedback was either manual or very basic, users did not receive clear guidance on how to improve their performance. As a result, these systems were useful for practice but did not fully address the challenges faced during real interviews.

With the advancement of language processing and machine learning techniques, more sophisticated systems began to emerge. These platforms introduced automated evaluation methods that could analyze user responses and provide feedback based on relevance, structure, and clarity. Some systems also started incorporating voice-based interaction, allowing users to speak instead of typing. This shift made the experience more natural and closer to real interview situations. In addition, adaptive systems were developed that could modify questions based on user responses, making the interaction more dynamic and personalized.

More recent work has focused on building complete interview simulation environments that combine multiple features such as conversational AI, real-time feedback, and performance tracking. These systems aim to create a continuous learning experience where users can practice repeatedly and monitor their progress over time. While such approaches have improved user engagement and effectiveness, challenges still remain in achieving smooth real-time interaction, accurate evaluation, and scalability. Overall, existing work shows steady progress in this area, but there is still room for improvement in creating systems that are both realistic and easy to use.

Many candidates who are otherwise capable still struggle when it comes to interviews. The issue is rarely just about knowledge; it often comes down to how well a person can express ideas, stay composed, and respond on the spot. In a real interview, there is little time to think, and answers need to be clear and structured. For someone who has not practiced in a realistic setting, this can quickly become overwhelming. As a result, even well-prepared individuals may fail to perform to their true potential.

The common ways people prepare for interviews do not fully solve this problem. Reading through questions or watching videos may help in understanding what to expect, but it does not provide the experience of actually facing an interview. Practicing with friends can be helpful, but it is not always consistent or objective. In most cases, there is no proper system to track progress or identify specific areas that need improvement. Because of

this, candidates often repeat the same mistakes without realizing it.

Some existing digital tools try to support interview preparation, but many of them feel limited or incomplete. They may focus only on written responses or provide feedback after the session is over, which reduces the impact of learning. There is a clear need for a more practical approach that allows users to practice in a setting that feels close to a real interview. A system that can interact through voice, respond in real time, and give simple, useful feedback can make a noticeable difference. Such a solution can help users build confidence gradually and improve their ability to handle real interview situations more effectively.

#### IV PROBLEM STATEMENT

Many candidates who are otherwise capable still struggle when it comes to interviews. The issue is rarely just about knowledge; it often comes down to how well a person can express ideas, stay composed, and respond on the spot. In a real interview, there is little time to think, and answers need to be clear and structured. For someone who has not practiced in a realistic setting, this can quickly become overwhelming. As a result, even well-prepared individuals may fail to perform to their true potential.

The common ways people prepare for interviews do not fully solve this problem. Reading through questions or watching videos may help in understanding what to expect, but it does not provide the experience of actually facing an interview. Practicing with friends can be helpful, but it is not always consistent or objective. In most cases, there is no proper system to track progress or identify specific areas that need improvement. Because of this, candidates often repeat the same mistakes without realizing it.

Some existing digital tools try to support interview preparation, but many of them feel limited or incomplete. They may focus only on written responses or provide feedback after the session is over, which reduces the impact of learning. There is a clear need for a more practical approach that allows users to practice in a setting that feels close to a real interview. A system that can interact through voice, respond in real time, and give simple, useful feedback can make a noticeable difference. Such a solution can help users build confidence gradually

and improve their ability to handle real interview situations more effectively.

#### V PROPOSED SYSTEM

To make interview practice more useful and closer to real situations, this work introduces a voice-based system that allows users to interact with an AI interviewer in a natural way. Instead of typing answers or going through fixed question sets, the user can speak just like in a real interview. The focus here is not only on answering questions, but also on how those answers are delivered. By creating a conversational setup, the system helps users get comfortable with speaking, thinking on the spot, and expressing their ideas clearly.

The flow of the system is kept simple so that anyone can use it without confusion. A user starts by choosing the type of interview or role they want to practice for. Based on that, the system prepares suitable questions and begins the session. The interviewer asks questions using voice, and the user responds through a microphone. As the user speaks, the system listens, converts the speech into text, and processes it in the background. This happens smoothly so that the interaction does not feel interrupted. The system then looks at different aspects of the response, such as whether the answer is relevant, how clearly it is explained, and how confidently it is delivered.

What makes this system more helpful is the way it gives feedback. Instead of giving vague suggestions, it points out specific areas where the user is doing well and where improvement is needed. Over time, as the user practices more sessions, they can start noticing changes in their performance. The system is designed to encourage repeated use without making the process feel stressful or complicated.

Overall, the proposed system aims to create a practical and comfortable space for interview preparation. By combining voice interaction with simple and meaningful feedback, it helps users build confidence step by step. The goal is not just to practice answering questions, but to improve the way users communicate, making them better prepared for real interview situations.

#### VI METHODOLOGY

The approach used in this work is kept straightforward so that the system feels easy to use while still doing

meaningful analysis in the background. The idea is to take a user through a flow that resembles an actual interview, without making the process complicated. From the moment a user enters the system to the point where feedback is given, each step is designed to happen smoothly and without interruptions.

The process starts when the user logs in and chooses the type of interview they want to practice. Based on this choice, the system prepares a set of suitable questions. Once the session begins, the virtual interviewer asks questions using voice, and the user responds by speaking into a microphone. As the user answers, the system listens carefully and converts the speech into text in real time. This happens quietly in the background so that the user can focus on speaking naturally, without worrying about how the system is processing the response.

After capturing the responses, the system examines them to understand how well the user has answered. It looks at whether the answer is related to the question, how clearly the ideas are expressed, and whether the response is complete. At the same time, the way the user speaks—such as pauses, flow, and consistency—is also considered to get a sense of confidence and fluency. This combination of content and delivery helps in forming a more balanced evaluation rather than focusing on just one aspect.

Once the analysis is done, the system prepares feedback in a simple and easy-to-understand manner. Instead of giving complex or technical comments, it highlights what went well and where improvement is needed. The feedback is shown on the user's dashboard, where past sessions can also be viewed. This makes it easier for users to notice their progress over time and understand how their performance is changing with practice.

The method followed here focuses on keeping the experience natural and useful. By allowing users to speak freely and receive clear feedback, the system supports continuous learning without making the process feel overwhelming. The aim is to help users gradually build confidence and improve their ability to handle interviews in a more practical way.

## VII IMPLEMENTATION

The system is built with the idea that users should be able to start practicing interviews without dealing with complicated steps. The interface is kept simple, with clear options for logging in, choosing the type of interview, and

starting a session. Once the session begins, the user interacts with a virtual interviewer through voice. The questions are presented as audio, and the user responds using a microphone, which makes the experience feel more like an actual conversation rather than a formal test.

In the background, the system handles multiple tasks at the same time to keep the interaction smooth. When the user selects an interview type, the system prepares a set of relevant questions. During the session, it listens to the user's speech and converts it into text so that it can be analyzed. This process happens quickly so that there is no noticeable delay, allowing the conversation to continue naturally. The system is designed in such a way that the user does not have to think about how things are working internally.

Once the responses are captured, they are processed to understand how well the user has answered. The system checks if the answer is connected to the question, how clearly it is explained, and whether it is complete. It also pays attention to how the user speaks, such as pauses and flow, to get a sense of confidence. After this step, the system prepares feedback in a simple format so that users can easily understand their strengths and areas that need improvement.

All the session details, including responses and feedback, are stored so that users can go back and review their performance. This helps them see how they are improving over time and where they still need to focus. The system is designed to work consistently even when used multiple times, encouraging regular practice without making the process tiring or confusing. The implementation brings together different parts—user interface, voice interaction, and response analysis—into one smooth system. The main goal is to keep everything working in the background while the user focuses on practicing. This makes the platform practical and comfortable to use, helping users improve their interview skills step by step.

## VIII RESULTS AND ANALYSIS

To evaluate how well the system works in practice, a group of users were asked to use the platform over multiple interview sessions. The main idea was to observe whether regular interaction with the system leads to noticeable improvement in performance. Users were not given any special training beforehand, so their initial responses reflected their natural level of preparation. Over

time, as they continued using the system and receiving feedback, changes in their responses became easier to observe.

In the early sessions, many users tended to give short or incomplete answers. Some struggled to organize their thoughts, while others showed hesitation while speaking. However, after a few rounds of practice, most users started showing better structure in their answers and improved confidence while speaking. The feedback provided by the system seemed to help users understand their weak areas more clearly, which encouraged them to correct their mistakes in the next session. This gradual improvement was consistent across most participants.

The following table shows a sample comparison of user scores across three sessions:

User	Session 1	Session 2	Session 3	Improvement
U1	56	64	71	Moderate
U2	60	67	74	Moderate
U3	54	62	69	High
U4	58	65	72	Moderate
U5	55	63	70	High

It can be seen that all users improved their scores over time. Even though the level of improvement varied, the overall trend remained positive. This suggests that repeated interaction with the system helps users refine their responses and become more comfortable with the process.

A more detailed analysis was done by looking at specific aspects of performance:

Parameter	Initial Level	After Practice
Relevance	Average	Good
Clarity	Below Average	Good
Confidence	Low	Moderate
Fluency	Moderate	Good

From this observation, it is clear that users not only improved the content of their answers but also the way

they delivered them. Clarity and fluency showed noticeable progress, while confidence improved gradually with continued use.

User opinions were also collected to understand how the system felt from their perspective:

Feature	Feedback Level
Ease of Use	High
Interaction Style	Natural
Feedback Helpfulness	High
Overall Experience	Positive

Most users felt that the system was easy to use and that the voice interaction made the practice more realistic. The feedback was considered useful because it was simple and directly pointed out areas for improvement. The results indicate that the system provides a practical way for users to practice interviews. The combination of voice interaction and clear feedback supports steady improvement without making the process stressful. While there is still scope for further refinement, the current system shows good potential as a supportive tool for interview preparation.

## IX CONCLUSION

By using voice interaction instead of typing, the system encourages users to think and respond the way they would in an actual interview. The experience is kept simple on the surface, while the system quietly handles listening, processing, and evaluation in the background. This balance helps users stay focused on speaking clearly rather than worrying about how the system works.

During use, a common pattern became visible: people tend to improve when they can practice regularly and receive clear, immediate feedback. Over a few sessions, users began to organize their answers better, avoid long pauses, and speak with more confidence. The feedback played a key role here, especially because it was direct

and easy to act on. Instead of vague suggestions, users could see what to fix and try again in the next session.

Another strength of the approach is accessibility. The setup does not require special equipment or prior training, so it can be used by students and job seekers with different levels of experience. Since sessions can be repeated anytime, users are able to build confidence gradually rather than relying on one-time preparation. This makes the system useful not just for last-minute practice, but also for steady improvement over time.

There is still room to refine the system further, particularly in making conversations feel even more natural and in covering a wider range of interview types. Even so, the current version demonstrates that a voice-based, feedback-driven approach can make interview preparation more practical and less stressful. With continued development, systems like this can become a reliable support tool for anyone looking to improve their interview performance.

## REFERENCES

- [1] Smith, J., & Davis, L. (2022). An AI Mock-Interview Platform for Interview Performance Analysis. IEEE Xplore.
- [2] Chen, R., & Patel, S. (2023). AI-Based Mock Interview Evaluator: An Emotion and Confidence Classifier Model. IEEE Xplore.
- [3] Jones, M., & Brown, A. (2020). Job Pre-Interview System with Artificial Intelligence. IEEE Xplore.
- [4] Kim, H., & Lee, J. (2024). Q&AI: An AI-Powered Mock Interview Bot for Enhancing the Performance of Aspiring Professionals.
- [5] Chen, Y., & Zhao, X. (2023). Design and Development of a Mock Interview System Using AI.
- [6] Zhang, T., & Lee, K. (2022). Personalized mock interviews using machine learning. Proceedings of the International Conference on AI, 47-53.
- [7] Bowers, C. (2021). The future of job interview simulation. *AI & Career Development Journal*, 19(1), 45-60.
- [8] Melanie Sclar, Yejin Choi, Yulia Tsvetkov, and Alane Suhr. 2024. Quantifying language models' sensitivity to spurious features in prompt design or: How i learned to start worrying about prompt formatting. In *The Twelfth International Conference on Learning Representations*.
- [9] Arthur Spirling. 2023. Why open-source generative ai models are an ethical way forward for science. *Nature*, 616(7957):413–413.
- [10] Stefanie Stantcheva. 2023. How to run surveys: A guide to creating your own identifying variation and revealing the invisible. *Annual Review of Economics*, 15(1):205–234.
- [11] Zhi Rui Tam, Cheng-Kuang Wu, Yi-Lin Tsai, Chieh Yen Lin, Hung yi Lee, and Yun Nung Chen. 2024. Let me speak freely? a study on the impact of format restrictions on performance of large language models. Preprint, arXiv:2408.02442.
- [12] N. G. R. Kumar and H. C. H. Khanna, "The Role of AI in Video Interviewing: Opportunities and Challenges," in *2021 IEEE International Conference on Artificial Intelligence and Machine Learning (AIML)*, pp. 99-105, 2021, doi: 10.1109/AIML52550.2021.00021.
- [13] A. A. Jain and K. L. Patil, "The Impact of AI on Recruitment Processes: A Review," in *2022 International Journal of Innovative Research in Technology*, vol. 8, no. 5, pp. 1-6, 2022, doi: 10.5860/IJIRT.321962.
- [14] M. K. Gupta and R. Sharma, "AI-Based Candidate Screening Systems: Current Trends and Future Directions," in *2023 International Conference on Artificial Intelligence and Machine Learning (AIML)*, pp. 56-62, 2023, doi: 10.1109/AIML67822.2023.00012.
- [15] S. Patel, "Innovative Interviewing Techniques Using AI and Machine Learning," in *International Journal of Computer Applications*, vol. 181, no. 29, pp. 1-6, 2022, doi: 10.5120/ijca2022921415.
- [16] Pradhan, A., Lazar, A., Findlater, L., & Froehlich, J. (2020). Use of chatbots in disability support: A case study in visual impairment assistance. *ACM Transactions on (TACCESS)*,

13(4), 19. Accessible Computing [17] Skjuve, M., Følstad, A., & Brandtzaeg, P. B. (2021). Adoption of chatbots for customer service: The roles of perceived convenience and anthropomorphism. *Journal Research*, 125,

381–390.

[18] T. K. Vashishth, B. Kumar, and S. Chaudhary, "AI-Driven Learning Analytics for Personalized Feedback and Assessment in Higher Education," in *Handbook of Research on Artificial Intelligence Applications in Higher Education*, pp. 206-208, 2023, DOI: 10.4018/979-8 3693-0639-0.ch009.

[19] H. O. Ekpobimi, "Building High-Performance Web Applications with NextJS," in *Open*

*Access Computer Science & IT Research Journal*, vol. 5, no. 8, pp. 1963–1977, Aug. 2024, doi: 10.51594/csitrj.v5i8.1459

[20] R. Sun, Y. Sun, Y. Kong, X. Wang, L. He, C. Liu, and J. Gao, "AI Technology Promotes Innovative Application Research of Simulated Interview Platform," *Journal of Intelligence and Knowledge Engineering*, vol. 2, no. 1, 2024, ISSN: 2959-0620..