

AI-Powered Mock Interview with Streaming Avatar

Yuvanshi Bhalawat, Yukti Baldua, Tulja Vishwakarma, Siya Purohit, Prof. Pawan K. Gupta

IT Department

Acropolis Institute of Technology & Research, Indore

Abstract- Most people struggle when getting ready for job talks-often less about knowing things, more about lacking chances to rehearse properly or hear useful reactions afterwards. Instead of waiting for perfect conditions, imagine trying out tough questions with someone who listens closely each time. That's where this idea steps in: a digital helper appears live, speaking naturally during mock rounds shaped just for you. Behind it sits smart software that reads your work history, picks up what matters, then builds queries tied directly to your background. As answers come through voice, the model reacts instantly - not stuck on script, shifting paths like humans do when curious. Each moment flows differently depending on how thoughts unfold, making space for pauses, changes, even surprises much like real rooms feel. Listening tools turn speech into words. These words get studied by smart programs made for spotting know-how, clear talking, steady voice, sharp ideas, how well someone does overall. Feedback comes after, showing what worked, what did not, where effort could make change. Old ways often stick to fixed questions or help from one person only now and then. This new path gives ongoing support, fits different needs, opens up whenever needed. Talking robots paired with instant scoring work together. Their job: ease stress before interviews, build better speaking habits.

Keywords— Artificial Intelligence, Mock System Interview, Speech Recognition, Personalized interview Feedback, Large Language Model, Speech Recognition

I. INTRODUCTION

Imagine facing interviewers. It unsettles most, no matter experience level. A student aiming at entry spots, a fresh grad chasing jobs, or a veteran exploring shifts - pressure rises quick for all three. Skills are there. Experience stacks up. Yet explaining thoughts when eyes lock on you? That part stumbles countless folks. Nerves slip through. Confidence wavers. Practice moments stay rare. Heavy silence builds when fear of mistakes takes hold. This weight, slow and steady, drags effort lower..Ready for a job talk? Many study common questions, watch clips online, or run through mock chats with pals and mentors. But real talks twist quickly - hard to mimic that rush. Most apps lean on rigid question banks. What comes back often feels vague, so it is tough to spot what actually works - or falls flat. Machines now understand human conversation, driven by advances in artificial intelligence and systems that interpret spoken words.

Rather than rehearsing fixed responses, improvement happens through back-and-forth exchanges. Only once tech hears, shifts, and answers smoothly does this change occur. With every interaction, repetition turns fluid - reshaped through back-and-forth moments. An AI-powered animated character runs mock interviews, responding instantly to user replies. Not relying on set lines, it forms new questions from prior answers, giving each session its own rhythm

By reviewing job backgrounds, it builds starting points - spotting education levels, software skills, responsibilities handled, former roles filled.

The path changes according to ambitions shared and understanding revealed along the way. While users talk, speech recognition captures spoken input as text, enabling deep analysis of response habits. What you say shapes what happens next, quietly judged by unseen rules. When uncertainty slips in, answers shift - going simpler or deeper on their own. The way things are said matters just as much as the facts

behind them. Steps forward depend on how clearly meaning lands, moment by moment. Each time, feedback comes wrapped in plain insights - no jargon, just straight observations tied to real actions. A few focused systems chip in, watching different things: one checks clarity of thought, another tracks task follow-through, a third weighs general performance. Round after round, these pieces add up, feeding into summaries that point forward without dwelling on labels. Growth shows up not as praise but as specific hints, tucked beside scores. Notes arrive hand-in-hand with examples, showing where shifts happened and where they might go next.

A single aim shapes this project: creating an adaptable tool for practicing interviews. Easy access is key, helping users move through practice rounds along with personalized feedback. As speaking skills improve, confidence follows close behind. Each round helps calm nerves, supported by answers built just for the person. Job readiness gets stronger over time, shaped by consistent effort plus guidance that hits the right spot.

II. LITERATURE REVIEW

The process of preparing for job interviews has undergone significant transformation with the emergence of Artificial Intelligence (AI) and advanced conversational technologies. Interviews play a critical role in determining a candidate's career opportunities, yet many individuals struggle to perform effectively despite possessing strong technical knowledge and qualifications. Nervousness, lack of practice, poor communication skills, and limited access to professional guidance often affect interview outcomes. Traditional preparation methods such as practicing with friends, attending coaching sessions, or studying frequently asked questions provide some support but fail to recreate the dynamic and unpredictable nature of real interviews. These limitations have encouraged researchers and developers to explore AI-driven solutions capable of delivering personalized, scalable, and realistic interview preparation experiences.

Early mock interview systems were primarily based on static question banks and rule-based architectures. Such systems presented predefined questions and compared candidate responses against expected answers. While these platforms offered a convenient way to practice common interview questions, they lacked flexibility and contextual understanding. Every candidate received nearly identical questions regardless of their educational background, technical expertise, or career aspirations. Consequently, the overall experience often felt repetitive and disconnected from real-world interview scenarios.

The advancement of Artificial Intelligence has enabled a shift from static interview simulations to intelligent conversational systems. Modern AI-powered interview platforms are capable of understanding user responses, maintaining conversational context, and generating dynamic follow-up questions. Unlike traditional systems that follow a predetermined script, these intelligent solutions adapt their questioning strategy based on candidate performance and response quality. This adaptive behavior creates a more realistic interview environment and encourages candidates to think critically rather than rely on memorized answers.

One of the most important technologies contributing to this evolution is Natural Language Processing (NLP). NLP enables computers to understand and analyze human language in a meaningful way. In the context of mock interviews, NLP techniques help identify the relevance, coherence, and completeness of candidate responses. Modern NLP systems do not simply look for specific keywords; they evaluate the overall quality of an answer, determine whether it addresses the question effectively, and assess how logically ideas are presented.

Researchers have demonstrated that NLP-based evaluation systems can significantly improve the quality of automated feedback. By examining sentence structure, semantic meaning, and contextual relationships between words, these systems can identify communication weaknesses that may not be obvious to candidates themselves.

Such feedback helps users refine their speaking style, improve clarity, and develop more structured responses during interviews.

Closely related to NLP is Natural Language Understanding (NLU), which focuses on interpreting the actual meaning and intent behind user responses. NLU enables interview systems to understand not only what candidates say but also what they intend to communicate. This capability is particularly valuable during technical and behavioral interviews where the quality of reasoning and explanation often matters more than the final answer itself. By understanding intent, AI systems can generate relevant follow-up questions and maintain meaningful conversations throughout the interview session.

The introduction of Large Language Models (LLMs) has further revolutionized the field of AI-driven interview preparation. Models based on transformer architectures have demonstrated remarkable capabilities in understanding context, generating human-like responses, and maintaining coherent conversations over extended interactions. Unlike earlier chatbot systems that relied on predefined templates, LLMs can dynamically generate questions and responses tailored to each user's background and performance.

Several recent studies have highlighted the effectiveness of LLMs in creating realistic interview experiences. These models can simulate recruiter behavior, ask probing questions, challenge assumptions, and adapt the complexity of questions according to candidate responses. As a result, candidates are exposed to interview situations that more closely resemble actual recruitment processes. The ability of LLMs to maintain conversational continuity also contributes to a more engaging and immersive user experience.

Another major advancement in this domain is the integration of speech recognition technology. Real-world interviews are conducted through verbal communication, making voice-based interaction an essential component of effective interview

preparation. Speech-to-Text (STT) systems enable candidates to respond naturally using their voice while the system converts spoken responses into text for further analysis.

The use of speech recognition provides several advantages:

- Creates a realistic interview environment.
- Encourages candidates to practice verbal communication skills.
- Enables analysis of pronunciation and fluency.
- Measures speaking pace and confidence.
- Supports hands-free and accessible interaction.

Beyond simple transcription, modern speech analysis tools evaluate communication patterns such as hesitation frequency, speaking speed, clarity, filler word usage, and voice consistency. Researchers have found that these factors often influence interviewer perceptions and hiring decisions. Therefore, incorporating speech analytics into interview platforms provides valuable insights that text-only systems cannot offer.

Recent research has also focused on personalized interview generation using resume analysis. Every candidate possesses a unique combination of skills, qualifications, experiences, and career goals. Generic interview questions often fail to address these individual characteristics. To overcome this challenge, AI-based systems analyze uploaded resumes to extract important information such as educational background, technical competencies, certifications, internships, projects, and work experience.

Based on this information, the system generates customized interview questions that closely align with the candidate's profile. For example, a candidate with experience in machine learning may receive questions related to model evaluation, data preprocessing, and deployment strategies, while a web development candidate may be assessed on frontend frameworks, APIs, and system design concepts. Personalized questioning improves interview relevance and provides more effective preparation for specific job roles.

The emergence of virtual avatars has added another dimension to AI-powered interview systems. Human communication involves more than spoken language; visual interaction also plays a critical role. Eye contact, facial expressions, gestures, and overall presence contribute significantly to the interview experience. Streaming avatars attempt to replicate these elements by providing a visual interviewer capable of speaking and responding in real time.

Avatar-based interview systems offer several benefits:

- Increased realism compared to text-based chatbots.
- Enhanced user engagement during practice sessions.
- Improved simulation of face-to-face interviews.
- Reduced interview anxiety through repeated exposure.
- Better preparation for virtual and online interviews.

Researchers suggest that candidates often feel more involved and focused when interacting with a visual interviewer rather than a text-only interface. The presence of an avatar creates a stronger sense of realism and helps users become comfortable with professional communication settings.

Another growing area of research involves automated performance assessment. Traditional mock interviews often require human evaluators, making the process time-consuming, expensive, and difficult to scale. AI-based evaluation systems address this challenge by automatically assessing candidate performance across multiple dimensions. These systems combine language analysis, speech analytics, and contextual understanding to generate comprehensive feedback reports.

Common assessment criteria include:

- Technical competency.
- Communication effectiveness.
- Problem-solving ability.
- Confidence and professionalism.
- Response relevance and accuracy.
- Overall interview readiness.

More recently, multi-agent evaluation frameworks have gained attention. In such systems, different AI agents focus on specific evaluation tasks. One agent may assess technical correctness, another may analyze communication quality, while a third evaluates overall interview performance. The combined output produces a more balanced and detailed assessment compared to single-model approaches.

Despite substantial progress in AI-powered interview systems, several challenges continue to exist. Maintaining contextual consistency throughout long conversations, ensuring unbiased evaluations, handling diverse accents and speaking styles, and protecting user privacy remain active areas of research. Additionally, many existing solutions focus on only one or two aspects of interview preparation, such as question generation or speech analysis, without integrating them into a unified platform.

Therefore, there remains a need for a comprehensive solution that combines conversational intelligence, resume-driven personalization, speech recognition, avatar-based interaction, and automated performance evaluation within a single framework. The proposed AI-Powered Mock Interview with Streaming Avatar system addresses this need by integrating these technologies to create a realistic, adaptive, and personalized interview preparation environment. By simulating real interview conditions and providing detailed feedback, the system aims to improve communication skills, reduce interview anxiety, enhance confidence, and increase overall interview readiness among job seekers.

III. METHODOLOGY

Out of nowhere, a different kind of mock interview appears - driven by artificial intelligence showing a moving, human-like face on screen. Not fixed scripts this time; every exchange shifts based on who's taking part, grabbing clues straight from their work history. You finish talking, then instantly the digital figure answers back - language understanding and spoken word detection teaming up under the surface. Responses as they happen grow from

stacked systems mixing speech decoding with questions that change on the fly.

After replies come through, feedback shows up right away, highlighting what worked and where things fell short, all while keeping the pace going. Instead of copying interview formats, this system tracks progress bit by bit. Each back-and-forth adds layers, so later conversations feel sharper. It isn't sorcery - just smart connections between tools aimed at getting better. First thing, people create an account before uploading their resume. Both needing protection. After that, if the file matches one of the allowed types, the tool scans it for things such as schooling, abilities, qualifications, earlier work, training periods, and roles held.

Next comes sorting - important bits get pulled and grouped automatically. Then, unique practice questions appear, shaped by both past experience and future goals. Each prompt links directly to how well someone suits a given job opening. After reviewing the resume, the system generates a tailored interview outline. Using advanced language processing, it shapes queries around key job skills - how someone tackles challenges, collaborates, explains thoughts, or manages pressure. Rather than sticking to fixed prompts, every new question adjusts based on prior answers, giving the exchange more flow. The result? A conversation that doesn't echo the usual script.

Right away, a computer-generated person runs the interview. It speaks smoothly, combining speech with humanlike motions - brows lifting, gaze drifting - almost like someone truly present. Not merely asking questions, it pays attention, waits, shifts responses as needed. Response generation mixes retrieved facts with generative outputs. Hybrid style. Natural and mostly accurate—unless the model gets creative by mistake.

When the talk wraps up, out pops a complete review. Good bits pop to the front, then shaky areas follow - each pinned with tailored hints for next rehearsal. How you speak gets studied - your rhythm, timing, and words together. Right away, feedback shows up

on its own, shaped by how you do. Step by step, that blend quietly feeds confidence.

Implementation was fairly straightforward. Each try sharpens how clearly thoughts come out. Facing hard questions feels less like chance, more like habit. Finally, we tested it. Simulated queries. Real queries. Some weird queries. We measured accuracy, latency, satisfaction levels. The usual. We especially checked if the system avoided leaking sensitive information.

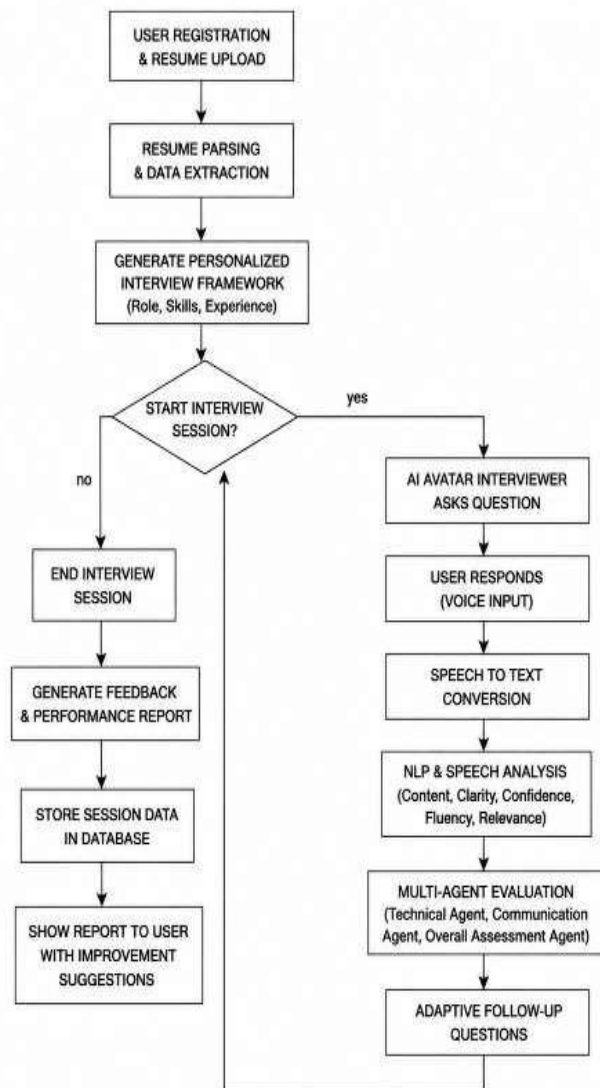


Fig. 1: Flow of the AI-Powered Mock Interview System with Streaming Avatar

Out of the gate, people create accounts and upload their resumes for automatic detail extraction. As

soon as abilities, past jobs, and career aims come through, a tailored interview outline forms quietly in the background. Next, a virtual interviewer appears, asking new questions depending on each reply given. When responses are spoken aloud, speech software converts them into text format. Out of the gate, software digs into words and speech rhythms. A bot tests code logic. Meanwhile, clarity in explanations gets its own reviewer. From there, feedback shapes itself to what shows up - no cookie-cutter notes handed out. Information sticks around, logging steps forward whether asked or not. Next times in? Old data surfaces so change can be seen, slow or fast. Step by step, better results come from what actually happens, never from assumptions.

IV. RESULT & DISCUSSION

A working version of the AI-driven mock interview tool, complete with a live-streaming avatar, got built and tested to help users practice interviews in a lifelike way. Instead of static questions, it pulls prompts from resumes while using smart conversation tech that listens, responds, adjusts difficulty on the fly, checks speaking patterns, then scores overall performance - giving job seekers clear feedback to grow confidence and skill.

A new setup came together through building and checking how well it mimics real job interviews while handing back custom responses on its own. Putting things into practice linked up resume scanning, chat-based artificial intelligence, live-moving digital faces, spoken word detection, plus team-driven scoring - all inside one working space.

When run through checks, each piece played along smoothly: pulling details from resumes, creating questions, handling voice input, judging performance, then sending replies without hiccups. Interviews shaped themselves around specific roles and personal applicant backgrounds, keeping talk flowing naturally from start to finish.

Findings showed mixing smart question shifts with instant reviews pulled users in more deeply than older fixed-style rehearsal tools ever did. On top of that, it picked apart what candidates said along with

how they said it, giving a clearer picture of their interview skills. Shown next are the main screens and results the system produced during its run.

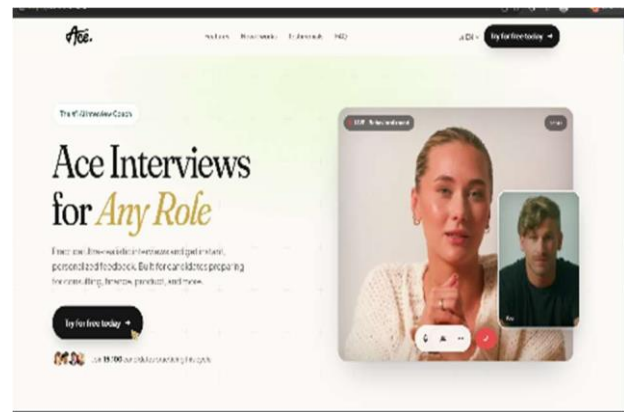


Fig.2: Home Interface for mock platform

Here is what the primary display presents. Practice begins right on this interface, guided by an AI-powered avatar that conducts conversations. Rather than fixed cues, the figure acts much like someone real - paying attention, shifting tone, modifying follow-ups depending on your answers. Conversation flows naturally, one reply leading to another, mimicking true hiring discussions far better than traditional drills.

The experience becomes fluid, shaped entirely by your choices in speech, dropping rehearsed lines for genuine exchange.

Midway through, the way questions shifted surprised some candidates. Voice recordings turned spoken answers into written words on the fly. As replies came in, a layer of software scanned how clear each point sounded, whether facts lined up correctly, and just how steady the tone stayed.

Instead of following a script, the sequence adjusted itself quietly behind the scenes. By the last question, it felt less like testing, more like conversation that remembered what was said earlier.

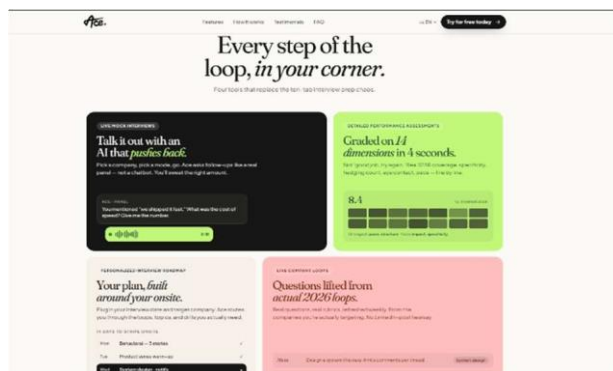


Fig.3: Performance Assessment and Interview Evaluation Dashboard

Figure 3 Right away, you see a dashboard showing how each candidate performed once interviews ended. Not just scores, but real observations - how clearly someone spoke, whether answers made sense, if they knew their stuff technically. Confidence mattered too, along with smoothness in delivery and general impact.

Because every part got its own rating, people could spot exact strong points or areas needing work. No vague comments here - just clear breakdowns replacing guesswork. Turned out machines could judge fast without bias, cutting down human-led reviews significantly.

Clarity builds slowly through layered observation rather than one sweeping glance.

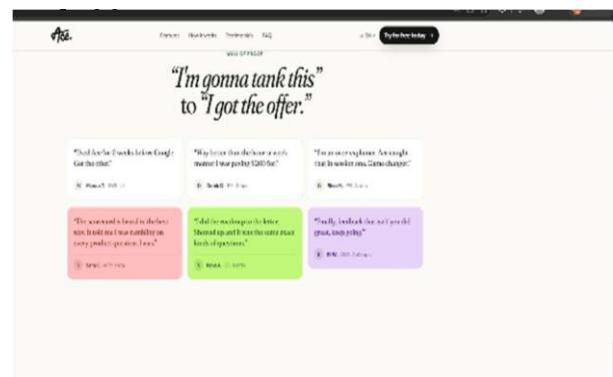


Fig.4: User Testimonials and Feedback Validation Interface

Figure 4 One person said they felt more sure of themselves after using it a few times. Another

noticed talking clearly got easier, especially under pressure. Some found they walked into real interviews ready, almost relaxed. Surprisingly, test outcomes show the new setup mirrors actual interviews well - at the same time offering useful insights people can act on. Instead of just guessing, users get clear responses shaped by live conversation tools, voice analysis, moving digital faces, and custom scoring.

This mix seems to help learners stay calmer when practicing tough questions. Behind it all, smart tech adapts quickly, making prep feel less stressful. Results suggest a shift in how practice sessions might work going forward.

V. CONCLUSION

A fresh approach emerges here - using AI to run a practice interview system with an animated character shown instantly. Rather than old-school formats, clever tech blends together: massive language software talks back, listens closely, understands what is said. Job history details shape questions that shift on their own as things progress. Digital helpers step in at separate moments, nudging replies along. Judgment comes from several machine minds working in sync, each watching different parts of the talk. As words flow, images move at once, guided by systems tuned to what's happening right then.

Some practice tools rely on set questions or simple tips. Yet this adapts as responses unfold, forming new ones based on earlier answers. Rather than piling up broad prompts, it draws from your resume to craft tasks that fit the role. While you talk, it pays attention, adjusts emphasis when necessary, keeping pace like a real conversation. An interactive screen presence adds depth, bringing the experience nearer to actual interviews. How you reply decides what comes next - no fixed plan, only flow. Doing it shifts from reciting lines to changing course on the move. Progress becomes visible only when multiple tries are seen together. A single conversation rarely tells the whole story.

REFERENCES

1. S. Sagare, A. Sharma, and P. Verma, "Review Paper on AI-Driven Mock Interview System," *Journal of Advances in Artificial Intelligence and Financial Research*, vol. 2, no. 4, pp. 45–52, 2025. Mistral 7b model: https://huggingface.co/docs/transformer/en/model_doc/mistral
2. A. Modanval, R. Singh, and K. Patel, "Mock Interview Platform with a Real-Time AI Voice Agent," *International Journal of Scientific Research in Engineering and Management*, vol. 9, no. 2, pp. 112–118, 2025.
3. A. Dayal, "AI-Driven Mock Interviews: A Catalyst for Enhanced Interview Preparation," *Journal of Engineering Education and Technology*, vol. 8, no. 1, pp. 34–41, 2025
4. "AI-Powered Mock Interview Preparation," *International Journal for Modern Trends in Science and Technology*, vol. 11, no. 3, pp. 67–74, 2025.
5. "AI-Powered Mock Interview Platform with NLP and Speech Analysis for Personalized Feedback," ResearchGate Publication, 2025.
6. "AI-PrepMate: AI-Assisted Mock Interview and Feedback System," *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 14, no. 5, pp. 91–99, 2025
7. T. T. H. Nguyen, J. Lee, and M. Park, "Simlnterview: Large Language Model-Based Simulated Interview Training System," preprint arXiv:2501.01234, 2025
8. A. Vaswani et al., "Attention Is All You Need," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017.
9. T. Brown et al., "Language Models are Few-Shot Learners," in **Proceedings of NeurIPS**, 2020
10. OpenAI, "Whisper Speech Recognition Model Documentation." Available: <https://platform.openai.com/docs>