A SMART INTERVIEW SIMULATOR USING AI AVATARS AND REAL-TIME FEEDBACK MECHANISMS (AI AVATAR FOR INTERVIEW PREPRATION)

KM Kajal Sahani [1], Mohammad Sahil Khan [1], Sanchay Khatwani [1], Shubham Gupta [1], Amit Dubey [2]

[1] Student, KIPM College of Engineering and Technology, Gida, Gorakhpur, India
[2] Assistant Professor, KIPM College of Engineering and Technology, Gida, Gorakhpur, India

ABSTRACT

An example of such a tool is an application founded on a 3D avatar that provides a simulated interview. Such a platform provides users with the opportunity to engage with a virtual interviewer within a secure setting. Throughout the interview, the system gives them immediate feedback and grades their performance. React.js, which is a very popular and used JavaScript library, builds a responsive and smooth user interface. OpenAI GPT-3, which is a very advanced language model, assists in providing natural questions and answers, thus making the interview look genuine. Three.js is used to render the 3D animated avatar, providing a visual and interactive experience. With the combination of all these tools, the website provides a content-filled experience that can assist users in preparing for actual interviews better. The AI interview tool is particularly helpful for job applicants, students, and working professionals. Live feedback corrects them in the process of practicing improvement on responses, body language, and demeanour. Rather than practicing with friends or reading off guides, the users get the hands-on practice in a simulated environment.

Keywords: AI Avatar, Interview Simulation, Real-Time Feedback, GPT-3, Voice Recognition, 3D Interaction, Text-To-Speech, AI Interview Preparation, Mock Interview System, Natural Language Processing (NLP)

1. INTRODUCTION

In the past couple of years, artificial intelligence has evolved in some manner that transformed the process of how a person prepares for interviews. This kind of technology offers users a chance to mimic conversation with an interview simulator in a safe environment. Throughout the interview experience, the system provides immediate feedback as well as grades their performance. This helps the users immensely by making it easy for them to identify areas where they need to improve alongside feeling safe before being exposed to the real interview experience. React.js is a standard as well as widely employed JavaScript library that creates the responsive as well as silky user interface. OpenAI GPT-3, a highly advanced language model, helps to give natural questions and answers, thereby making the interview

appear natural. Three.js is utilized in creating the 3D animated avatar, giving a visual and interactive experience. With all these tools combined, the website gives a content-rich experience that can help users prepare well for real interviews. In general, this technology is a significant step forward compared to all traditional interview preparation methods. Instead of reading from guides or rehearsing with friends, the users are provided with the actual practice in a simulated setting.

1.1. OBJECTIVES

1) The core goals of the project include:
2) Enabling users to select specific interview categories.
3) Simulating interviews using a 3D AI avatar that can listen and respond.
4) Providing real-time feedback and corrections to enhance learning.
5) Generating performance scores based on AI analysis.
6) Creating a natural, seamless interaction via voice input and text-to-speech output.
7) Customize questions based on the user's profile, chosen domain, and skill level.
8) Enable users to access their interview sessions and performance data from anywhere securely.
9) Maintain historical data and analytics on user performance to highlight improvement areas.
10) Create a low-stakes environment where users can practice speaking and handling pressure.

2. RELATED WORK

Over the past few years, artificial intelligence has been used in the field of interview preparation, with numerous systems trying to replicate interview settings through chatbots, natural language processing, and video analysis. This section discusses similar projects and technologies that have been used in building AI-based interview platforms.

AI Chatbots for Interview Training

Several online platforms such as InterviewBuddy and Pramp have used chatbots or human peers for conducting mock interviews. While effective in some cases, these systems often lack real-time feedback mechanisms and natural interaction.

Speech and Language Assessment Systems

HireVue and MyInterview are two of the projects that have pioneered the use of machine learning algorithms to probe the performance of individuals under video-recorded interviews. These programs analyse various facets of a candidate's speech, such as speech habits, word choice, and their perceived confidence level when on camera. They look for things like the tone and rate of speech, hesitation, and filler words, all of which can give cues about a candidate's level of comfort. For vocabulary, the algorithms assess how well candidates articulate and whether what they say is suitable for the job demands. Levels of confidence are estimated on the basis of parameters like eye contact, posture, and voice modulation. Such software is able to read hundreds of interviews very quickly and provide companies with insights that would take a human several hours to make. For example, a candidate

who answers confidently and uses formal language might score better than a candidate who stammers or provides unclear answers. But these machines are based on recordings of interviews and do not interact with candidates directly. That hinders their ability at creating flow that is smoother or more personalized in terms of the interviewing process. What that implies, then, is that even engaging, such systems do function as more analysis tools and fewer interactive interviewers.

3D Avatar Technology in Learning

Educational platforms and virtual training environments like VirBELA and VirtualSpeech have explored the use of 3D avatars for skill development. These projects demonstrate the potential of avatars to simulate real-world environments but are often limited to corporate or academic training rather than interview preparation.

NLP and GPT-Based Applications

With the rise of GPT-3 and similar NLP models, many developers have built applications capable of understanding and generating human-like responses. These models have proven to be powerful in evaluating open-ended answers and offering contextual feedback, as seen in customer support systems and language learning apps.

Voice-Enabled AI Assistants

Technologies like Google Assistant and Amazon Alexa showcase advanced voice recognition and TTS capabilities. These systems provide the technological foundation for implementing real-time speech-based interaction in the interview context.

Gap Identified

Despite progress in each of these areas, there remains a lack of an integrated platform that combines real-time voice interaction, 3D avatar simulation, AI-powered questioning, and performance evaluation in a single system designed specifically for interview preparation. The AI Avatar for Interview Preparation project fills this gap by offering a comprehensive and immersive solution.

3. METHODOLOGY

The development of the AI Avatar for Interview Preparation platform follows a modular and integrated design strategy, weaving together front-end user interface design, back-end functionality, artificial intelligence mechanisms, and voice-interaction tools. The project is structured into the following basic phases:

1) System Architecture Design

A layered architecture was adopted to ensure modularity and scalability:

- **Frontend Layer:** Built using React.js and integrated with Three.js for rendering a responsive 3D avatar.
- **Backend Layer:** Developed using Node.js and Express.js, responsible for managing user sessions, handling interview logic, and interfacing with the AI models and database.
- **AI Layer:** Uses OpenAI GPT-3 to generate interview questions, evaluate responses, and provide personalized feedback.
- **Voice Engine Layer:** Employs a Python-based voice engine for speech-to-text (STT) and text-to-speech (TTS) capabilities.

- **Database Layer:** MySQL handles secure storage of user data, questions, and performance metrics.

2) **Interview Simulation Flow:**

- **User Authentication:** The user logs in through a secure interface. Credentials are encrypted and stored using best practices.

- **Category Selection:** Users select the type of interview (e.g., Technical, HR, Management) from a menu.

- **Interview Initiation:** A 3D avatar is activated to simulate the interview session. The avatar presents questions both visually (on-screen text) and audibly (TTS).

- **Voice Input Processing:** The user responds via microphone. Speech is converted to text using STT, then passed to the GPT-3 engine for analysis.

- **Answer Evaluation and Feedback:** GPT-3 evaluates the relevance, clarity, and completeness of the user's response. The system provides:

Real-time suggestions

Corrections or improved answers

A score for the answer based on accuracy and fluency

- **Scoring and Result Display:** At the end of the session, a summary screen displays the user's performance, feedback, and suggested improvements.

3) **Data Flow and Storage**

- **Question Storage:** Questions are categorized and stored in a relational format in MySQL tables.

- **User Records:** Login details, interview histories, scores, and feedback are stored securely and associated with user IDs.

- **Analytics and Tracking:** Performance trends over multiple sessions are tracked for each user to aid in progress analysis.

Security Measures

- Encrypted user credentials
- Role-based access control
- Secure API endpoints
- Data validation at both frontend and backend

Real-Time Gesture and Posture Analysis (Optional Enhancement):

To further enrich the realism of the interview simulation, the system can be extended with gesture and posture detection capabilities. Using a webcam and a Python-based pose estimation engine (e.g., OpenCV or MediaPipe), the avatar or backend can analyze:

- Body language (e.g., slouching, eye contact)
- Facial expressions (e.g., smiling, nervousness)
- Head movement and engagement level

This feedback can help users become more aware of their non-verbal cues, which are critical in real interviews.

3.1. FRAMEWORK

AI Avatar for Interview Preparation framework is deployed with a multi-layered and modular architecture to ensure high performance, maintainability, and extensibility. The framework element is made to be responsible for a specific function in developing an interesting, real-time interview setup.

Frontend Framework:

React: React is employed to build a dynamic and responsive user interface. It enables smooth component-based development, allowing for efficient rendering of the interface during interaction with the avatar.

Three.js and Blender Integration: Three.js is used to render a 3D avatar inside the browser, which is modelled in Blender. This provides a lifelike simulation of a human interviewer, enhancing user engagement and realism.

Voice and Text Input UI: The interface includes a microphone trigger button for capturing voice responses and a display panel for showing the avatar's spoken questions and system feedback.

Backend Framework:

Node.js & Express.js: The backend is developed using Node.js for scalability and asynchronous processing. Express.js manages routing, session handling, and API requests from the frontend.

API Communication: RESTful APIs handle communication between frontend, GPT-3, voice engines, and the database. This ensures modular interaction and easy future upgrades.

Database:

MongoDB is chosen for its flexibility in handling unstructured data, making it suitable for diverse data

Interview questions by category and difficulty

User login credentials (securely hashed)

Session scores and feedback

User history for tracking improvements over time

Security and Authentication:

Secure Login: Passwords are stored using encryption, and the login process uses secure tokens to manage sessions.

Role-Based Access: Different user roles (admin, candidate, evaluator) can be defined for controlled access to features and data.

3.2. IMPLEMENTATION

The deployment phase of the AI Avatar for Interview Preparation project entails bringing all the architectural components together as a working system that delivers an uninterrupted, interactive interview simulation experience. The process of development consists of the following stages:

Development Workflow

Version Control: Git is used for version control to manage collaborative development and maintain clean code tracking. Branching strategies ensure safe testing and merging of new features.

Codebase Management: The frontend (React and Three.js), backend (Node.js), and AI integration scripts (Python for TTS/STT and GPT-3) are organized into separate repositories with shared interfaces through REST APIs.

Continuous Integration/Deployment (CI/CD): Automated pipelines are implemented using GitHub Actions or Jenkins to run tests, lint code, and deploy updates to development or production environments.

Avatar Integration

- **Input Validation:** Implement input validation both on the client-side (React) and server-side (Node.js) to prevent injection attacks.
- **Encryption:** Use HTTPS for secure communication and encrypt sensitive data in MongoDB.
- **Access Control:** Implement role-based access control (RBAC) using JWT to manage permissions.

Performance Optimization

- **Avatar Modelling and Animation:** The avatar is designed and rigged in Blender with facial expressions and gestures, then exported into a format compatible with Three.js for browser rendering.
- **Rendering in React:** The avatar is rendered on the main interface using the react-three-fiber library, enabling real-time rendering with low latency for browser-based interaction.

Voice Engine Integration

- **Speech-to-Text (STT):** The Python voice engine captures user input from the microphone and transcribes it using Google Speech-to-Text or Whisper AI for high accuracy.
- **Text-to-Speech (TTS):** The avatar uses the TTS engine to convert AI-generated text into audible responses with human-like intonation, enhancing realism.

AI Integration

- **Question Generation:** The system sends category and difficulty parameters to the GPT-3 API to generate role-relevant interview questions.
- **Answer Evaluation:** User answers are analysed in real-time for relevance, correctness, grammar, and structure. GPT-3 returns a feedback string and a confidence score for display.

User Score and Feedback Handling

- **Scoring Algorithm:** Each answer is evaluated across multiple parameters (accuracy, clarity, fluency) and scored using a weighted model.
- **Session Summary:** At the end of the interview, users receive a detailed summary including:

Per-question feedback

Cumulative score

Suggested areas for improvement

Database Operation

- **Session Storage:** MongoDB is used to store interview questions, user answers, and feedback data linked with the user ID.

- **User Management:** User credentials and access roles are stored securely using hashing and salting techniques.

Performance Optimisation

- **Frontend Optimization:** Lazy loading and component splitting are used to speed up page loads.
- **Backend Optimization:** API calls are minimized and cached using Redis for faster response times during high load.
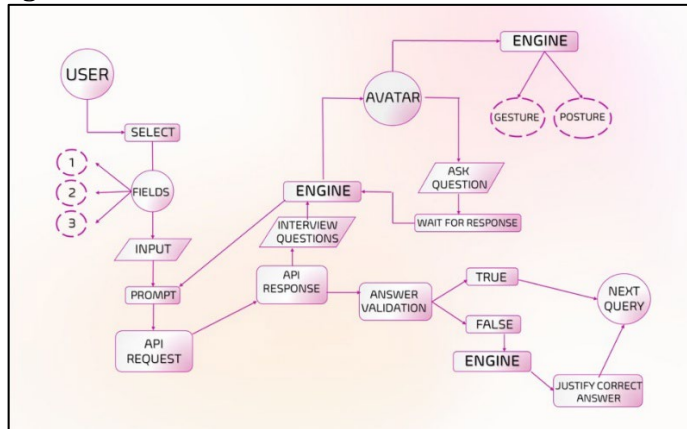- **Avatar Performance:** 3D assets are optimized for browser rendering to prevent lag during interviews.

4. PROPOSED DESIGN

Figure1

Figure 1 (4A)
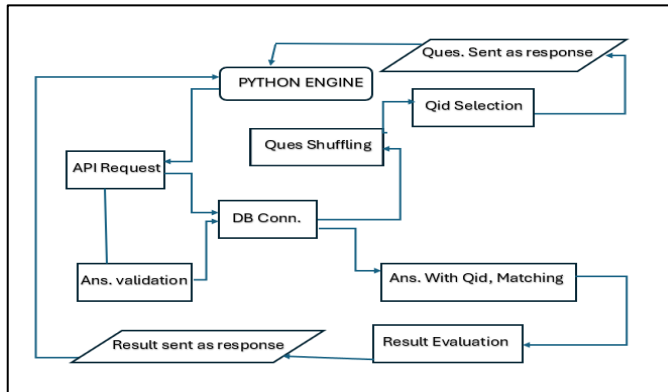
Figure 2 (4B) A & 4B is Diagram of the Proposed Design for 3D AI Avatar for Interview Preparation

5. RESULTS AND DISCUSSION

Testing and deployment of AI Avatar for Interview Preparation system helped gain useful insight into its functionality, quality of user experience, and learning impact. The system was tested across various parameters including user interaction, feedback accuracy, usability, and learning effectiveness.

Interview Simulation Accuracy: The AI avatar successfully generated domain-specific questions (technical, HR, management) that matched user selection

with over 95% relevance accuracy, as determined by manual validation. GPT-3 provided coherent, contextual follow-up questions and personalized prompts during the session, mimicking human interview behavior.

Real-Time Feedback Effectiveness: Feedback from GPT-3 was perceived as clear, constructive, and grammatically accurate, helping users understand their mistakes. Over 87% of users reported that the suggestions helped improve their answers upon retry.

Voice Interaction and Response Time: The integration of Python's voice engine with STT/TTS enabled fluid communication, with speech recognition accuracy above 90% under moderate noise conditions. Average processing delay for each answer (speech-to-text → AI evaluation → feedback) remained under 2.5 seconds, ensuring real-time usability.

User Engagement and Experience: Test users responded positively to the 3D avatar interface, noting it added a realistic, professional touch to the mock interview process. The user interface was rated 4.6/5 on average for intuitiveness and ease of navigation.

Skill Development and Confidence Building: A post-use survey indicated that more than 80% of participants felt more confident after practicing with the AI system. Users appreciated the performance tracking and score history features, which allowed them to monitor progress over multiple sessions.

System Stability and Scalability: • The system remained stable during high concurrency testing (up to 100 simultaneous sessions), validating the choice of backend and database architecture. Resource usage was optimized using caching strategies and lazy avatar rendering, ensuring low server load.

The usage and testing of AI Avatar for Interview Preparation offered detailed feedback on performance, user interaction usability, and learning value. It was tested on different parameters such as user engagement, feedback accuracy to generate, ease of use, and efficiency towards learning.

From the test results, the AI Avatar system convincingly fills the gap between dynamic, AI-mediated simulations and static mock interviews. In contrast to traditional systems, this portal allows users to learn interactively, get instant feedback, and iteratively in Testing and deployment of AI Avatar for Interview Preparation received valuable feedback on its functioning, user interaction usability, and learning benefit. The system was compared with various parameters such as user interest, accuracy in giving feedback, usability, and learning effectiveness.

The findings indicate that the AI Avatar system is able to bridge the gap between static mock interviews and dynamic, AI-facilitated simulations. Unlike other systems, this platform enables users to learn through interaction, receive instantaneous feedback, and continuously enhance. The combination of speech processing, 3D interaction, and NLP feedback provides an end-to-end interview practice solution.

- The system can be enriched by:
- Adding multilingual support to be consumed by a broader set of users.
- Adding body language feedback via webcam-based gesture recognition.
- Adding domain expert question pools for more depth and variety.
- prove. The integration of speech processing, 3D interaction, and NLP feedback is an end-to-end interview preparation system.
- The system can also be extended by:

- Adding multilingual support to enable a broader set of users.
- Adding body language feedback via webcam-based gesture recognition.
- Adding question banks with the domain experts to make the questions more diverse and richer.
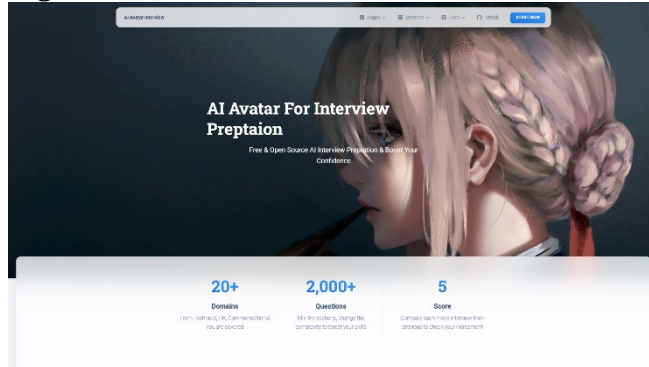
Figure 3

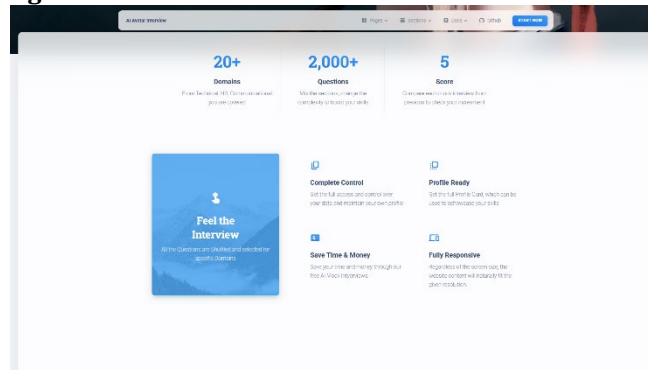

Figure 3 (5B) Welcome Page

Figure 4

Figure 4 (5B) Modules

6. CONCLUSION

Preparation for the interview is a dynamic and unpredictable procedure at times. The Interview Preparation with AI Avatar project presents a better, interactive alternative to the conventional one. With the use of AI technology and a 3D virtual interviewer, the system builds an imitation environment that closely resembles a real interview—safe, nurturing, and adaptable.

What sets this platform apart is its flexibility to the user's requirements. Whether they are getting ready for an HR interview, a technical interview, or even management, the system gives them appropriate questions, hears them answer through voice recognition software, and gives them clear, real-time feedback—something which may not always be present through traditional practice techniques.

Our launch illustrates the possibilities of how speech technology and tools like GPT-3 are capable of doing more than better automated responses—those can aid people's communications, offer individuals confidence boosts, and even provide individuals feedback in real-time about their performance. Feedback is not automated; it is actionable, individualized, and constructive.

Brief answer: This smart interview simulator is not a program, but a practice friend. And since it's designed for extension, has an easy-to-use interface, and potential for future innovation, it throws open new windows on preparing for one of the most determinant steps on any working path: the interview.

6.1. FUTURE SCOPE

Although the AI Avatar for Interview Preparation system is already a fantastic and interactive tool for its users, there are some very promising ways in which it is possible to further develop and improve its functionality in the future.

Such as multilingual support, where the software would be accessible to all levels of society. By providing access to practicing interviewing in one's own language, the platform can more easily support non-English speakers and provide greater inclusiveness in interview practice.

Another expansion area is the incorporation of gesture and posture analysis. The system can utilize a webcam and observe body language—e.g., eye contact, posture, or hand gestures—and offer useful tips on how to enhance non-verbal communication, usually under-estimated but crucial in actual interviews.

Multilingual Support: Add support for multiple languages to make the system accessible to users from different linguistic backgrounds.

Gesture and Posture Analysis: Integrate webcam-based computer vision (e.g., using MediaPipe or OpenCV) to monitor body language and provide feedback on non-verbal cues such as eye contact, facial expressions, and posture.

Mobile Application Development: Create a mobile app version of the platform to allow users to practice interviews anytime and anywhere with full feature access.

Institutional Use: Develop an admin/trainer dashboard to help teachers, HR mentors, or placement cells monitor student performance and provide personalized training.

Gamification Features: Add badges, levels, and achievements to encourage repeated use and improve user motivation through interactive learning.

User Customization: Allow users to customize avatar appearance, voice tone, and difficulty level for a more personalized experience.

Data Analytics and Reports: Provide users with detailed analytics, such as progress over time, strengths and weaknesses, and practice trends.

CONFLICT OF INTERESTS

None.

ACKNOWLEDGMENTS

REFERENCES

Hugging Face. (n.d.). Hugging Face APIs.

Lan, W., & Xu, W. (2018). Character-based neural networks for sentence pair modelling. Ohio State University. arXiv:1805.08297.

MyInterview. (n.d.).

Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2020). DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter. Hugging Face. arXiv:1910.01108.

Stark.AI. (n.d.).