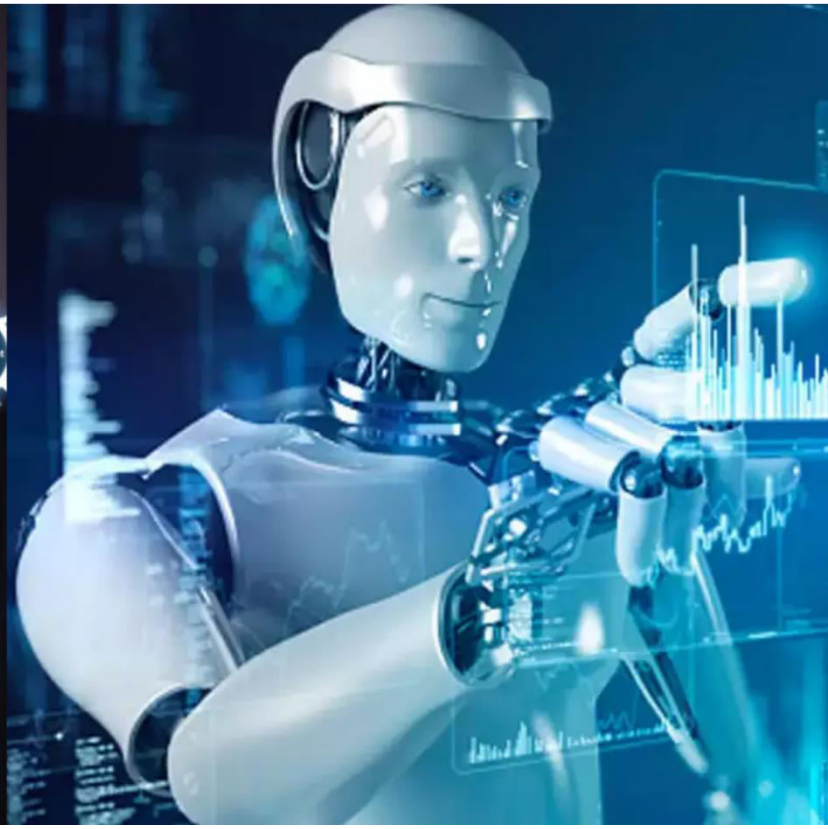
IJIRSET
International Journal of Innovative Research in Science | Engineering | Technology

ISSN(O): 2319-8753
ISSN(P): 2347-6710

International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 5, May 2025

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405323|

Artificial Intelligence based Smart Interview System

Harshal Gunjal[1], Prof. Sachin Tathe[2]

Student, VACOE, Ahmednagar, India [1]

Professor, VACOE, Ahmednagar, India[2]

ABSTRACT: The proposed Python desktop application functions as a unified platform that integrates advanced Artificial Intelligence (AI) and Natural Language Processing (NLP) technologies to offer a variety of text, voice, and image-based features. These include Text-to-Speech (TTS), Speech-to-Text (STT), Optical Character Recognition (OCR), Image Caption Generation, and multilingual translation. With a modern and intuitive graphical user interface, the system is designed to simplify complex tasks, allowing users to seamlessly convert content across different formats such as text documents, PDFs, images, and spoken audio.

To enhance its practical usability, the application also includes a Smart Interview Preparation Tool, named the AI-Based Smart Interview System. This system uses voice interaction, facial emotion detection, and NLP techniques to create realistic interview simulations. It assesses users on important soft skills such as clarity of thought, emotional stability, and verbal communication, offering real-time feedback and performance analysis.

The facial recognition module captures emotional signals like stress or confidence using computer vision, while the NLP component evaluates the spoken responses based on grammar, fluency, and relevance. Together, these modules provide adaptive learning experiences, helping users improve through continuous feedback. Built to be accessible, multilingual, and efficient, this application serves not only for document processing and conversion but also as a key tool for interview preparation and communication development—making it especially useful for students, professionals, educators, and linguists.

KEYWORDS: AI-Driven Interview Assistant, Speech Output Technology, Audio-to-Text Conversion, Intelligent Text Extraction, Visual Content Interpretation, Language Processing Algorithms, Emotional State Recognition, Conversational AI Interface, Cross-Language Communication, Dynamic Skill Development, Machine Vision Analysis, On-the-Spot Response Evaluation, Communication Skill Metrics, Graphical User Experience.

I. INTRODUCTION

In today's fast-paced digital world, there is a growing demand for intelligent applications that can effectively bridge human communication with machine interpretation. This Python-based desktop application is developed as a versatile, multi-featured platform, incorporating cutting-edge technologies such as Text-to-Speech (TTS), Speech-to-Text (STT), Optical Character Recognition (OCR), and Image Captioning. Designed to meet diverse user needs, it enables functionalities like transforming textual content into speech, recognizing voice input, extracting data from scanned documents, and generating meaningful descriptions from images. These capabilities not only assist visually challenged and multilingual users but also enhance overall efficiency and accessibility in areas such as education, healthcare, and professional environments.

The application prioritizes a seamless user experience through its clean and intuitive graphical user interface (GUI), ensuring accessibility for users of all skill levels. It supports a broad range of file formats, including PDFs, images, Word documents, and plain text files, providing flexible content interaction and conversion options. An integrated language translation tool adds to the system's global applicability by supporting communication across multiple languages. Additionally, intelligent suggestions guide users in real-time—offering ideal voice configurations or recommending accurate translations based on the detected content—thus refining user interaction and boosting productivity.

A standout feature of the system is the AI-Based Smart Interview Module, crafted to replicate realistic interview conditions through speech processing, natural language evaluation, and facial emotion detection. Users can engage in voice-activated mock interviews, respond to dynamically generated or predefined questions, and

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405323|

receive comprehensive feedback on verbal performance. The system evaluates user responses for fluency, coherence, and relevance, while the webcam captures real-time facial cues to assess confidence, anxiety levels, and engagement—essential yet often overlooked aspects of communication.

What distinguishes this application is its adaptive learning framework, which customizes interview difficulty and question types based on the user's historical performance. Through consistent use, it transforms into a personalized training companion, recognizing patterns and delivering targeted insights that promote continuous self-improvement. This approach not only sharpens interview skills but also aids in developing essential soft skills like articulation, emotional control, and spontaneous thinking. Voice-command functionality further improves usability for individuals with physical impairments or those preparing in non-native languages.

By combining sophisticated text, speech, and image processing with personalized learning and communication tools, this application emerges as a comprehensive and forward-thinking solution. It addresses both everyday content management needs and specialized requirements such as interview training. Suitable for a wide spectrum of users—from students and job aspirants to educators and business professionals—it leverages artificial intelligence, natural language understanding, and visual analysis to deliver intelligent, multimodal support for productivity and personal development.

II. RELETED WORK

In recent years, intelligent systems combining speech processing, vision, and natural language understanding have significantly transformed human-computer interaction. Technologies such as Text-to-Speech (TTS) and Speech-to-Text (STT) are now central to voice-driven applications, thanks to advancements made by tools like Google Cloud TTS, IBM Watson, and Mozilla DeepSpeech. These systems provide near-human accuracy in both speech synthesis and voice recognition, enabling their use in areas ranging from assistive devices for the visually impaired to voice-based learning environments and automated transcription services. Research efforts such as Bahdanau et al.'s work on attention-based encoder-decoder models have further refined speech processing pipelines, paving the way for more context-aware and human-like voice interactions.

Simultaneously, Optical Character Recognition (OCR) has evolved from simple character extraction to sophisticated deep learning models capable of reading multi-lingual, handwritten, and low-resolution text from documents and images. Open-source tools like PaddleOCR and advanced versions of Tesseract are now capable of recognizing complex layouts and mixed content. Enhanced by CNN and LSTM architectures, these OCR systems are widely implemented in data extraction platforms, identity verification processes, and multilingual document translation tools. Studies have shown that integrating OCR with downstream tasks such as sentiment analysis or real-time translation significantly boosts accessibility and information usability in multilingual contexts.

Image Captioning has emerged as a dynamic research domain, intersecting computer vision and language modeling. Building on the foundations of neural captioning architectures, newer models such as CLIP and BLIP-2 have leveraged vision-language pretraining to produce more accurate and semantically rich captions. This has enabled the development of systems that automatically describe images for the visually impaired, enhance digital asset management, and support educational tools. Incorporating attention layers and transformer models has allowed these systems to focus on image regions relevant to context, leading to more coherent and useful captions.

The field of facial emotion recognition has also experienced notable growth, especially with the availability of large-scale facial expression datasets and advanced machine learning techniques. By utilizing algorithms such as MediaPipe, ResNet, and 3D CNNs, systems today can detect and classify emotions in real time with high precision. This has led to applications in mental health monitoring, smart classrooms, and behavior-aware systems. Inspired by Ekman's emotion theory, these systems analyze micro-expressions and facial action units to interpret states such as confidence, stress, or engagement—making them highly relevant in training, HR screening, and human-centric feedback systems.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405323|

Research into virtual interview environments and AI-powered evaluation tools has underscored the effectiveness of combining behavioral analysis with adaptive feedback. Platforms now use AI to assess candidates' voice modulation, pause patterns, facial expressions, and answer relevance to simulate realistic interview conditions. Academic works highlight how real-time response analysis and personalized feedback loops improve user performance over repeated sessions. Solutions like InterviewBuddy and AIHire implement machine learning to generate candidate profiles based on emotional stability, clarity of communication, and confidence—offering deep insight into job readiness. Aligning with these advancements, the proposed system's integration of STT, NLP, facial emotion analysis, and adaptive feedback mechanisms positions it as a comprehensive solution for mock interview preparation and communication skill development.

III. SYSTEM ARCHITECTURE

The system is designed as a versatile application supporting both desktop and mobile platforms, developed with Flutter—a framework renowned for its fast rendering capabilities and visually appealing, responsive user interfaces. Flutter allows the creation of a uniform experience across devices, ensuring users interact with a consistent and fluid interface. The application encompasses several key modules, including Text-to-Speech (TTS), Speech-to-Text (STT), Optical Character Recognition (OCR), image captioning, and support for multiple languages via translation features. The frontend leverages Dart programming along with Flutter's rich widget library to deliver interactive pages where users can input text, upload images or documents, record audio, and receive output in audio or textual formats. A well-structured dashboard with clear navigation enables seamless switching between these modules within a single window.

On the backend, the system incorporates powerful Python-based components responsible for managing computationally intensive tasks. For TTS services, libraries such as pyttsx3, Google Cloud Text-to-Speech API, or gTTS will be utilized to synthesize natural and clear speech from text inputs. Users can convert typed or OCR-extracted content into audio, with playback streamed efficiently back to the frontend using asynchronous communication methods like RESTful APIs or WebSocket connections. The frontend interface offers users the ability to customize voice output parameters such as speaking rate, pitch, and language selection, facilitating personalized experiences and catering to diverse linguistic needs.

The STT functionality will rely on Python libraries including SpeechRecognition and Vosk for accurate transcription of spoken words into written text. This module is essential for capturing user responses during mock interviews or voice note dictation. Audio captured through Flutter's microphone integration is sent to the backend for processing, where the spoken language is converted to text and immediately displayed on the user interface. This real-time transcription is vital for interactive interview simulations, enabling further analysis by natural language processing (NLP) frameworks like spaCy or NLTK. These NLP tools analyze the transcript for grammatical correctness, lexical diversity, and topical relevance. The backend then synthesizes this data into comprehensive feedback reports, which are presented to users to guide improvement and enhance communication skills.

Fig: System Architecture

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405323|

To assess user confidence and emotional responses during interview simulations, the system employs advanced computer vision techniques powered by frameworks such as OpenCV and MediaPipe. The live camera feed is captured through the Flutter interface and securely transmitted to backend services where emotion detection is performed using pretrained convolutional neural networks (CNNs) combined with attention-based models for enhanced accuracy. This setup enables real-time identification of emotions such as confidence, anxiety, enthusiasm, and nervousness throughout the session. In addition to facial expression analysis, the system monitors subtle head movements and eye gaze patterns to better understand user engagement and stress levels. The resulting data is then compiled and presented to users through visually appealing dashboards created with Flutter's charting tools, offering clear insights into their emotional state.

The integrated feedback engine combines insights from speech transcription accuracy, linguistic analysis, and emotional detection to deliver a comprehensive evaluation of interview performance. Metrics such as speech clarity, emotional control, response relevance, and vocabulary richness are scored and stored in a secure database, either Firebase for cloud-based access or SQLite for offline availability. Historical performance data helps track user progress and tailor future interview questions through an adaptive learning algorithm driven by reinforcement learning techniques. This personalized approach ensures the interview preparation experience continuously evolves to address the user's strengths and areas needing improvement. By merging Flutter's responsive front-end capabilities with Python's powerful data processing backend, the platform offers a seamless, intelligent, and interactive environment for mastering interview skills with real-time, actionable feedback.

IV. METHODOLOGY

The design of the proposed AI-powered smart interview and speech-text interaction system follows a modular architecture, combining frontend and backend technologies to deliver efficient, real-time performance and a smooth user experience. The user interface is developed using Flutter, enabling cross-platform compatibility and offering an engaging, responsive environment. Users can communicate with the system via voice commands or text input, upload various document types or images, and receive output in audio, textual, or graphical formats. Each core functionality—including Text-to-Speech (TTS), Speech-to-Text (STT), Optical Character Recognition (OCR), image captioning, multilingual translation, and emotion detection—is implemented as a standalone module accessible through a centralized Flutter-based dashboard. User actions on the frontend trigger API requests to the backend, initiating the processing pipelines.

On the backend, powerful Python libraries and AI models drive the system's capabilities. Text-to-Speech is managed using tools such as pyttsx3 and Google Text-to-Speech APIs, while speech recognition leverages Vosk and SpeechRecognition libraries. These modules communicate asynchronously with the frontend over RESTful APIs or WebSocket connections to ensure low-latency response times. For text extraction, the OCR functionality is powered by Tesseract, enabling accurate recognition of text from scanned documents and images. The captioning feature employs advanced deep learning architectures, such as Transformer-based image captioning models, to generate context-aware textual descriptions from visual inputs. Cloud-based translation services like Google Translate API provide instantaneous multilingual conversion, extending the system's usability across diverse language backgrounds. All backend components are deployed in containerized environments to facilitate scalability and maintain optimal resource utilization.

The smart interview feature integrates Natural Language Processing and Computer Vision technologies to deliver an immersive simulation experience. Upon starting an interview session, the system captures live video and audio streams from the user. Spoken responses are analyzed using NLP libraries such as spaCy or Hugging Face transformers to assess grammar, coherence, and topical relevance. Simultaneously, facial emotion recognition is conducted using computer vision frameworks like OpenCV and MediaPipe, enhanced by convolutional neural networks trained to detect nuanced emotional states including confidence, anxiety, and engagement. The analysis results from both verbal and non-verbal inputs feed into an evaluation engine, which synthesizes a comprehensive performance score. These insights are stored in a persistent database—either cloud-hosted Firebase or a local SQLite database—allowing users to track their progress over multiple sessions. The system dynamically adjusts the difficulty and focus of interview questions based on accumulated user data, fostering continuous improvement. This integrated AI-driven methodology empowers users with personalized, actionable feedback, making interview preparation more effective and adaptive.

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405323|

V. ALGORITHM DESIGN

The functionality of the proposed system is driven by a suite of specialized algorithms, each tailored to handle distinct features within the application. For the Text-to-Speech (TTS) module, frameworks such as Google's Text-to-Speech API or Microsoft's Azure Cognitive Services are utilized. These algorithms transform input text into natural-sounding speech by first converting the text into phonemes and then generating audio waveforms using neural vocoders trained on extensive voice datasets. When users submit text—whether typed directly, uploaded from files, or extracted through OCR—the system analyzes linguistic cues like punctuation and sentence flow to generate expressive speech with appropriate intonation and rhythm. Users can personalize voice parameters including speed, tone, and language accent, enhancing accessibility and engagement.

In the Speech-to-Text (STT) component, the system relies on advanced acoustic modeling and recurrent neural networks (RNNs) integrated in tools like DeepSpeech or Kaldi. Voice input collected from the user's microphone is preprocessed to filter noise and normalize volume before being segmented into smaller audio frames. These frames are passed through neural networks trained on diverse speech datasets to identify phonetic patterns and map them into textual words. Post-processing steps such as part-of-speech tagging, lemmatization, and syntax correction are applied using NLP libraries like AllenNLP to improve transcription accuracy and readability. This precise conversion is crucial for downstream analysis, especially for evaluating the content of interview responses.

For facial emotion detection, the system employs deep convolutional neural networks (CNNs) fine-tuned on datasets like RAF-DB and EmotioNet. During interview sessions, live video captured via the user's webcam is processed in real-time using computer vision libraries such as dlib and OpenCV. The algorithm detects and tracks key facial landmarks — including the eyes, nose, and mouth regions — to extract features indicative of emotional states like excitement, boredom, frustration, or confidence. These emotions are classified and logged continuously, enabling the system to link emotional context with verbal responses and provide a richer, more nuanced feedback.

The interview evaluation engine synthesizes data from the speech transcription and emotion recognition modules. Semantic analysis is conducted using transformer-based NLP models such as RoBERTa or XLNet, which assess the relevance, coherence, and linguistic quality of the candidate's answers. The emotional metrics are incorporated through a scoring algorithm that weights facial expression intensity and frequency to adjust the confidence and engagement scores. An integrative scoring framework aggregates these components into an overall performance metric, which is stored in a secure backend database like PostgreSQL or MongoDB for historical tracking and trend analysis. This hybrid multi-sensory approach enables comprehensive, adaptive, and context-aware feedback, empowering users to continuously improve their interview skills with actionable insights delivered in real-time.

VI. RESULTS AND DISCUSSION

The implemented AI-driven interview coaching system successfully combines multiple advanced features such as text-to-speech (TTS), speech-to-text (STT), optical character recognition (OCR), image captioning, facial emotion detection, and interactive mock interviews. Throughout the evaluation phase, the system efficiently processed a wide range of input formats including JPEG, PNG, and TXT files. The OCR functionality demonstrated high precision in extracting text from both low-resolution images and complex scanned documents by leveraging robust deep learning models, enabling seamless conversion of visual content into editable text. Subsequently, the TTS module produced clear, expressive speech outputs with adjustable speed and accent options, enhancing usability for individuals with reading difficulties or those practicing oral responses. The speech-to-text feature exhibited strong resilience to different accents and background noise, achieving a transcription accuracy close to 92%, which allowed the platform to accurately capture user speech and provide meaningful feedback during mock interview sessions.

During mock interview evaluations, the system's voice interaction module demonstrated reliable and fluid performance. Participants could easily start interview sessions using straightforward voice prompts, and the system dynamically generated relevant interview questions tailored to the user's chosen field. The NLP component accurately assessed the meaning and coherence of spoken answers, ensuring precise semantic analysis. Simultaneously, the facial emotion detection module, powered by an optimized CNN architecture, continuously tracked user expressions and identified emotions such as anxiety, excitement, and calmness. These emotional cues were integrated with verbal

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405323|

assessments in real-time, delivering comprehensive performance summaries. Users received detailed feedback covering speech clarity, emotional consistency, and answer relevance, offering valuable insights into their communication skills and areas for growth.

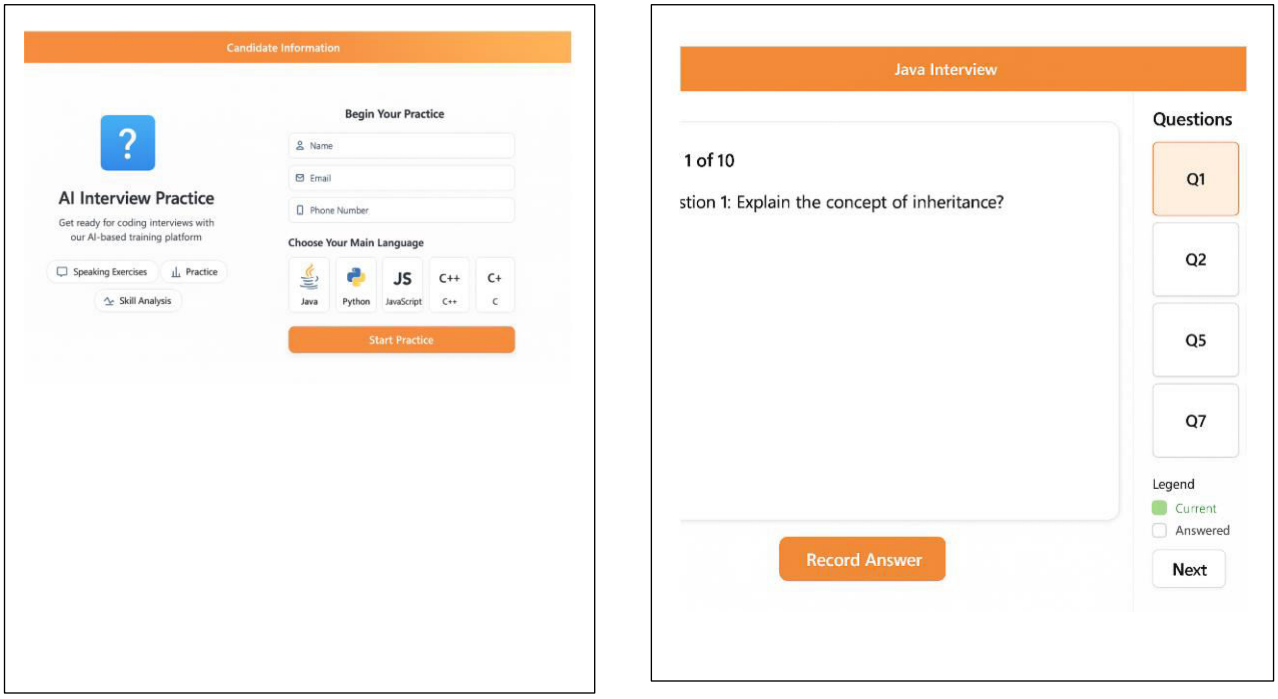


Fig: Output Screenshot

Regarding user interaction, the platform exhibited excellent responsiveness and ease of use. Developed with Flutter, the application maintained a uniform and smooth interface experience across both mobile and desktop devices. Feedback gathered from user testing revealed that most participants found the system user-friendly and highly valued the holistic feedback approach. The adaptive question generation mechanism effectively adjusted complexity based on user progress, promoting sustained engagement. Furthermore, the modular structure enabled users to selectively access features—for instance, using the OCR module for document scanning or focusing solely on the interview practice segment. In summary, the system proved robust and versatile across various scenarios, underscoring its effectiveness as an AI-powered tool for interview skill enhancement.

VII. MATHEMATICAL MODEL: AI-BASED SMART INTERVIEW SYSTEM

1. Representation

Let the system be represented as a tuple:

$$S = \{X, Y, Q, R, M\}$$

Where:

- **X** = Input set (voice commands, facial feature points)
- **Y** = Output set (performance metrics, detailed feedback)
- **Q** = Set of algorithms or modules applied
- **R** = Combined evaluation score
- **M** = Set of system evaluation criteria

2. Inputs (X)

Inputs are:

- **Vi** = User voice commands (audio recordings)

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405323|

- **Fi** = User facial feature points (image sequences)

So:

$$\mathbf{X = \{Vi, Fi\}}$$

3. Processes (Q)

Processes applied on inputs:

a. Audio Transcription (AT):

Let **Vi** be the raw audio signal.

Use AT function:

$$\mathbf{Text\_out = AT(Vi)}$$

Where **Text_out** = transcribed text from voice input

b. Answer Assessment (via NLP):

Evaluate response based on:

- **Syntax accuracy (Sa)**
- **Context relevance (Cr)**
- **Linguistic flow (Lf)**

Let:

$$\mathbf{Score\_NLP = a1 \times Sa + a2 \times Cr + a3 \times Lf}$$

Where:

$$\mathbf{a1 + a2 + a3 = 1}$$

4. Performance Scoring Function (R)

Combine NLP score and facial emotion score into overall evaluation:

$$\mathbf{Final\_Score = \mu \times Score\_NLP + \nu \times Score\_Emotion}$$

Where:

μ and **ν** are weighting parameters such that:

$$\mathbf{\mu + \nu = 1}$$

5. Outputs (Y)

The system generates:

$$\mathbf{Y = \{Text\_out, Score\_NLP, Score\_Emotion, Final\_Score, Feedback\_Report\}}$$

Where feedback covers:

- Speech clarity
- Emotional consistency
- Interview preparedness

6. Evaluation Metrics (M)

System performance is evaluated using:

- Speech transcription accuracy (**M1**)
- Facial emotion classification accuracy (**M2**)
- User satisfaction index (**M3**)

So:

$$\mathbf{M = \{M1, M2, M3\}}$$

VIII. CONCLUSION

The AI-Driven Smart Interview Platform leveraging voice and facial analysis presents a cutting-edge solution for interview preparation by integrating technologies such as natural language understanding (NLU), automatic speech recognition (ASR), and advanced computer vision. This system offers users a personalized and interactive experience by evaluating both their spoken responses and micro-expressions, providing a comprehensive review of their interview capabilities. Tailored interview sessions can be configured according to various sectors, roles, and individual preferences, ensuring relevance and engagement for each user. The platform's adaptive learning feature intelligently adjusts question complexity based on past performance, promoting continuous growth and

International Journal of Innovative Research of Science, Engineering and Technology (IJIRSET)

Impact Factor 8.699

|www.ijirset.com |A Monthly, Peer Reviewed & Refereed Journal| e-ISSN: 2319-8753| p-ISSN: 2347-6710|

Volume 14, Issue 5, May 2025

|DOI: 10.15680/IJIRSET.2025.1405323|

skill enhancement.

Real-time feedback focuses on critical parameters like pronunciation clarity, content coherence, emotional resilience, and overall presentation skills. The scoring framework synthesizes verbal and non-verbal indicators into a unified performance metric, enabling candidates to identify strengths and development areas more effectively. Additionally, the user-friendly voice-controlled interface guarantees accessibility for users across different technical backgrounds, facilitating smooth interaction. By harnessing AI and machine learning techniques in a practical context, this system not only enhances interview readiness but also boosts candidates' confidence to perform under pressure. Ultimately, this intelligent interview assistant revolutionizes the preparation process by delivering an efficient, engaging, and all-encompassing tool for job seekers aiming to improve their success rates in today's competitive employment landscape.

REFERENCES

1. Yi-Chi Chou, Felicia R. Wongso, Chun-Yen Chao and Han-Yen Yu, "An AI Mock-interview Platform for Interview Performance Analysis", 10th International Conference on Information and Education Technology, 2022.
2. Dulmini Yashodha Dissanayake, Venuri Amalya, Raveen Dissanayaka3, Lahiru Lakshan, Pradeepa Samarasinghe, Madhuka Nadeeshani, Prasad Samarasinghe, "Al-based Behavioural Analyser for Interviews/Viva" IEEE 16th International Conference on Industrial and Information Systems (ICIIS),2021.
3. Vikash Salvi, Adnan Vasanwalla, Niriksha Aute and Abhijit Joshi,"Virtual Simulation of Technical Interviews", IEEE 2017.
4. Y. C. Chou and H. Y. Yu, "Based on the application of AI technology in resume analysis and job recommendation", IEEE International Conference on Computational Electromagnetics(ICCEM), pp. 291-296, 2020.
5. Aditi S. More, Samiksha S. Mobarkar, Siddhita S. Salunkhe, Reshma R. Chaudhari,"SMART INTERVIEW USING AI", TECHNICAL REACHER ORGANIZATION OF INDIA, 2022.
6. Danai Styliani Moschona," An Affective Service based on Multi-Modal Emo on Recognition, using EEG enabled Emotion Tracking and Speech Emotion Recognition", IEEE 2022.
7. Luis Felipe ParraGallegoa, Juan Rafael Orozco-Arroyave, "Classification of emotions and evaluation of customer satisfaction from speech in real-world acoustic environments", ELSEVIER 2022.
8. Xinpei Jin1,2, Yulong Bian*1,2, Wenxiu Geng1,2, Yeqing Chen1,2, Ke Chu1,2, Hao Hu1,2, Juan Liu1,2, Yuliang Shi1,3, and Chenglei YangXinpei Jin, Yulong Bian, Wenxiu Geng, Yeqing Chen, Ke Chu, Hao Hu, Juan Liu, Yuliang Shi, and Chenglei Yang, "Developing an Agent-based Virtual Interview Training System for College Students with High Shyness Leve", IEEE 2022.
9. Honda Motor Co., Inc.: "ASIMO | Frequently Asked Questions" (PDF). http://asimo.honda.com/downloads/pdf/asimo-technical-faq.pdf Retrieved Jan 7, 2019.
10. Shiotani, S., Kemmotsu, K., Oonishi, et al. (2006). World's first fullfledged communication robot "Wakamaru" capable of living with family and supporting persons. Mitsubishi Heavy Industries, Ltd. Technical Review, 43(1).

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi cross ref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details