Ravanaboyine Saisri, 2026, 14:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Ai-Powered Interview Automation System for Real-Time Candidate Assessment

Ravanaboyine Saisri[1], N.Sandhya Rani[2]
[1]Student, Department of CSE,AI,ML, MAM Women's Engineering College, Kesanupalli(Narasaraopet).
[2]Assistant Professor, Department of CSE,AI,ML, MAM Women's Engineering College, Kesanupalli(Narasaraopet).

Abstract: Automated intelligent software agents may mimic human communication behaviours, allowing for more organic and interesting interactions with people, thanks to the fast development of conversational AI. With these new developments, it's possible to replace human interviewers with autonomous software agents that are smarter than humans, thereby automating the candidate interview process. Conversational AI allows machines to mimic human interviewers in many ways, including asking questions, understanding and evaluating responses, and starting dynamic discussions. Improving the efficiency of the whole interview process, this automation guarantees consistent and impartial assessment, which in turn leads to more effective and fair recruiting procedures. An AI-driven interview system that can evaluate candidates in real time is the focus of this research article, which intends to provide a thorough analysis of its design and implementation. Various artificial intelligence agents work together in this system to do things like choose questions based on predefined criteria, evaluate candidates' responses, analyse speech for signs of emotion and sentiment, and then combine these findings to give distinct scores for answers and emotions in the performance evaluation. Index Terms—Generative Pre-trained Transformers (GPT), AI-driven interviews, automated candidate assessments, multimodal emotion detection, and natural language processing (NLP).

Keywords: Artificial Intelligence (AI), Conversational AI, Automated Interview System, Candidate Assessment, Natural Language Processing (NLP), Generative Pre-trained Transformers, Emotion Detection, Sentiment Analysis, Multimodal Analysis, Intelligent Software Agents, Recruitment Automation, Human–Computer Interaction, Real-Time Evaluation, Machine Learning, Speech Analysis.

I. INTRODUCTION

With the help of AI, talent acquisition has become much better in the last several years [1]. Accurate and efficient applicant evaluation throughout the hiring process is more crucial than ever in today's competitive employment market. Analysing candidates using traditional interview methods has its limitations due to human subjectivity and time restrictions. However, by automating the interview process with AI-driven technologies, a more thorough evaluation of the applicant may be achieved. In order to provide a more objective, thorough, and scalable alternative for candidate evaluation, this study introduces a novel framework that makes use of sophisticated technology. Various artificial intelligence software agents, each with its own area of expertise, are integrated into the system to streamline the interview process. An interview's specialised question management agent is in charge of picking out predetermined questions that are both relevant and suitable given the environment. In order to provide a dynamic and responsive interview experience, the agent utilises GPT-4 and customised Natural Language Processing (NLP) models to adaptively pick questions depending on the candidate's past replies. Afterwards, a sentiment analysis agent records and examines the emotional components of the candidate's communication, while a response management agent records, processes, and stores the candidate's answer.

The capacity to identify emotions using both textual analysis and aural input is included into the system. Using this method, we can pick up on non-verbal signs of emotion, such tone, pitch, and speaking tempo, that could otherwise go unnoticed when analysing text alone. Emotional intelligence and interpersonal communication are examples of soft talents that are critically evaluated via this methodology. These agents

collaborate to give you a complete picture of the applicant by picking up on the tone and substance of their answer. To get a whole picture of the applicant, an all-encompassing assessment agent compiles all the information and findings from other agents. Scores, ratings, and other quantitative measures are part of an all-encompassing assessment, which also takes into account qualitative insights, such as feelings and opinions.

Language precision, consistency, and conformity to task-specific competences are all evaluated by this agent using natural language processing methods. To maximise objectivity and minimise biases, a comprehensive knowledge of each applicant is used to make the final judgement. In order to give a thorough data-driven evaluation of the candidates' performance, the interview system can integrate with natural language processing (NLP) [2], convolutional neural networks (CNN) [3], recurrent neural networks (RNN) [4], and generative pre-trained transformers (GPT) [5]. This allows it to ask questions, score answers, and analyse speech for sentiment and emotion. The recruiting process stands to be drastically altered with the introduction of an AI-powered interview system. The technology improves overall efficiency, decreases expenses, and frees up human resources experts to concentrate on strategic decision-making instead of administrative activities by automating important portions of the interview process.

II. RELATED WORK

Automated interview evaluation was suggested by Priya et al. [6]. This system would evaluate audio and visual signals in order to quantify the behaviour of interviewees. Support Vector Machines were used for the classification of the auditory cues and face emotions. A smart interviewing application that automates the interview process using NLP and Deep Learning (DL) was suggested by Senarathne et al. [7]. To determine how well a candidate has done, the algorithm verifies their answers and makes predictions based on their accuracy. An automated computational framework was suggested for the purpose of recognising verbal and non-verbal behaviours that occur during employment interviews. Facial expressions, language, and prosodic elements are used by the system to predict interview evaluations [8].

To solve automation interview systems' fairness problems, Kim et al. [9] suggested a multimodal data approach. To strike a compromise between accuracy and fairness, the method included a regularisation term. By shortening the Wasserstein gap between vulnerable groups, the method further reduces bias. In order to categorise interview responses, Romadon et al. [10] analyses TF-IDF with word embeddings using Artificial Neural Networks (ANN). As TF-IDF reduces dimensionality, bias, and human mistakes more effectively than word embeddings, it has become the technique of choice for judging job interviews. Using natural language processing (NLP) capabilities, an interview chatbot was developed that could automatically produce questions and replies according to the talents shown on the résumé [11].

Researchers Pickard et al. looked at how various interview formats affected the amount of personal information that was divulged during in-person interviews. When compared to the human-like Embodied Conversational Agent (ECA) and the human interviewer mode with visible faces, the findings revealed that participants volunteered more sensitive material in the faceless Audio-only Computer Assisted Self Interview (ACASI) mode [12]. By examining interview transcripts, Yusuf et al. [13] investigated machine learning techniques for assessing cultural fit in job candidates. The research examined three different classifiers: SVM, Naive Bayes, and K-Nearest Neighbours (KNN). The results showed that SVM was the most successful solution for this job, consistently outperforming the other algorithms.

In their study, Fang et al. [14] investigated several natural language processing techniques for the purpose of symptom and quality of life effect identification in unstructured, qualitative interview data.

In terms of efficacy in identifying patient reported symptoms and affects, the research found that the Bidirectional Encoder Representations from Transformers (BERT) model typically demonstrated greater performance. Researchers Jiang et al. [15] looked studied how televideo interview data from voice, facial, linguistic, and cardiovascular modulation may be used to distinguish between people with mental problems. The results indicate that automated mental health evaluations may be made more accessible, scalable, and economical by using a multimodal approach. The significance of scientific methods in the classroom and evaluation is highlighted in the research [16]. It goes on to say that there's a lot of promise for using machine learning methods to evaluate science education programmes.

III. PROPOSED METHODOLOGY

The suggested method for creating an AI-powered interview system incorporates state-of-the-art tools including GPT, RNN, CNN, and natural language processing. By continuously modifying the asking approach in response to real-time monitoring of the applicant's emotional state answers, the system intends to automate and improve candidate assessment.

A. Gathering Information Datasets from SAVEE[17], TESS[18], and CREMA-D [19] were used to train the audio emotion identification programme. The multi-modal dataset known as Surrey Audio-Visual Expressed Emotion (SAVEE) has four male actors who, in seven distinct emotional states, act out a total of fifteen lines. The goal of the two actresses recorded for the Toronto Emotional Speech Set (TESS) was to evoke seven distinct emotions via the use of 200 phrases. Audiovisual recordings of actors expressing various emotions are included in the Crowdsourced Emotional Multimodal Actors Dataset (CREMA-D). The International Survey on Emotion Antecedents and Reactions (ISEAR) dataset was used to train the model for emotion recognition from text [20]. The goal of ISEAR is to compare and contrast people's emotional reactions in various cultural and environmental settings. More than three thousand people from thirty-seven different nations filled out the survey, which asked them to describe how they felt in reaction to different stimuli. For the question and answer dataset, the interviewer and applicant were introduced to MedM CQA [21], a medical multiple-choice question answering system. More than 194,000 multiple-choice questions covering a wide range of medical topics are available in MedMCQA, along with detailed explanations of each answer. This material is crafted to tackle actual questions that may be found on the AIIMS and NEET PG entrance examinations. Section B: Interview System Design In Fig. 1, we can see the structure of an AI-powered interview system, with the applicant, interviewer, and backend components all interacted with.

As the interviewer asks questions, the applicant interacts with the technology in real-time. The interviewer retrieves sample questions from a database, and the technology records the candidate's voice replies for later review. The system's essential parts are: Conversational Server: The candidate, interviewer, and backend AI models all communicate and share data via this key hub. One option is Speech-to-Text (STT), which transcribes the candidate's voice replies. • Text-to-voice (TTS): This technology enables spoken communication with the applicant by converting text data into voice. It is used for questions and comments. • Session Management: Oversees the interview's status and surroundings Handles the streaming of audio data in real-time for emotion identification • APIs, or application programming interfaces, provide a means of connecting to various software agents. • The software agents known as QMA (III-C), RMA (III-D), MESAA (III-E), and CESA (III-F) manage the interview process as a whole. The central analytical component that integrates different AI capabilities to handle and analyse data is the AI engine. • Natural Language Processing Models: Review the candidate's recorded responses to extract recurring ideas. Include AI models that are specifically designed to analyse text and audio elements. • GPT: Help choose from a list of sample interview questions, revise the questions as needed,

and check candidates' answers for clarity and relevance to the situation.

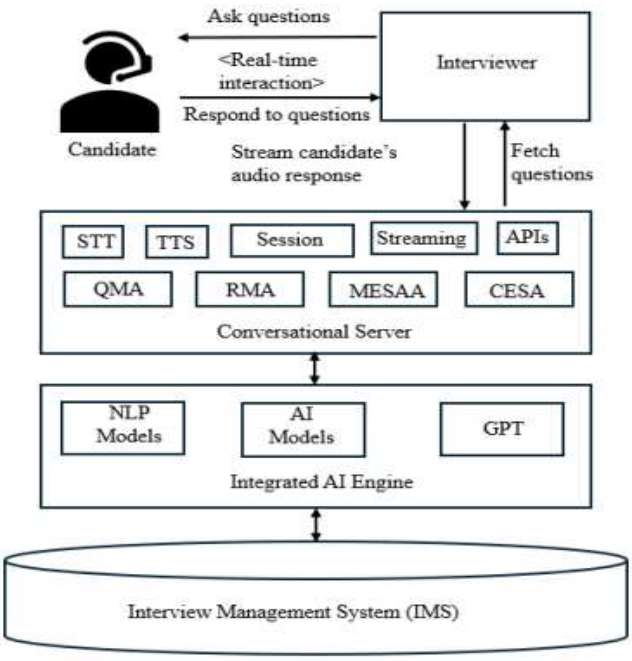

Fig. 1. Architecture of the Interview system

The Interview Management System (IMS) is the foundational system that stores and oversees all interview-related data. For effective index-based search, the document database keeps track of queries and answers. • Relational database: Stores analysed and aggregated data to provide detailed reports on candidate sentiment, performance, and emotions. • Vector database: Allows for retrieval with context by use of interview material that is embedded with semantics. This agent's job is to ask applicants questions in real time while taking into account their history, any established criteria, and any responses they may have given.

To dynamically adjust queries and ensure they are contextually appropriate, the agent uses proprietary NLP models and GPT. Using real-time analysis of responses and specified criteria, this agent also regulates the interview flow by selecting when to move on to the next question. Elastic search is a document-based database that stores the questions and answers from the MedMCQA dataset. It is a distributed search and analytics engine. Quick and efficient data retrieval is possible, and it scales well. See Figure for a list of the pertinent fields kept in the document database. 2. The database's full-text search features are perfect for the interview system, which requires questions to be chosen and found using criteria like relevancy, contextual similarity, or keywords. Table I lists some typical situations and things that need to be addressed throughout the interview process in order to keep the interviewer-candidate discussion going strong.

TABLE I
CANDIDATE'S RESPONSE SCENARIOS AND IMPLEMENTATION STRATEGIES

#	Scenario	Action	Implementation strategies
1	Providing correct answers	Acknowledgment, next question, more in-depth questions	Custom NLP: Managed with contextually relevant keyword
2	Providing partially correct answers	Encouragement, guidance to complete the answers	GPT: Adapt system response based on candidate's response or choose additional questions
3	Incorrect answers	Correction, clarification, next question	GPT: Generate detailed and contextually relevant responses
4	Completely different answers	Rephrase the question, reiterate the question	Custom NLP: Measure semantic similarity between the response and the question
5	Correct answers but for a different question	Clarification, acknowledge and redirect to original question	Custom NLP: Detect mismatches using semantic similarity analysis, topic modeling, and keyword matching
6	Asks for clarification or repeats the question	Rephrase the question, offer additional context to the question	Custom NLP: Refine question wording with additional context
7	Unable to answer or cannot provide a response	Encouragement to respond, move on to another question or topic	Custom NLP: Supportive feedback and transitioning to next topic or question
8	Provides vague or general answers	Clarification, reiterate the question	GPT: Interpret vague response, provide contextually appropriate feedback and prompt for more details

To deal with different situations, a mix of GPT-4 and bespoke NLP models was used. For simple tasks like keyword extraction, rule-based processing, and predetermined answer creation, custom NLP models were used. The ability of GPT to generate new languages is unnecessary for these jobs. Performance, customisation, data protection, and scalability are some of the other considerations while deciding between GPT and bespoke NLP models.

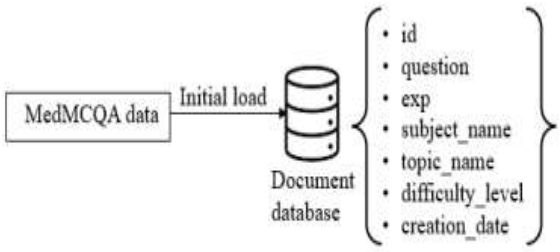

Fig. 2. Data Ingestion and Storage Schema of MedMCQA Dataset

IV. RESPONSE MANAGEMENT AGENT (RMA)

Capturing, processing, and storing the candidate's replies is done by the response management agent. The system's ability to correctly capture responses, link them to the corresponding questions, and get the data ready for analysis relies on this agent. The information is stored in an RDBMS like PostgreSQL, and the pertinent fields are shown in Figure 3.

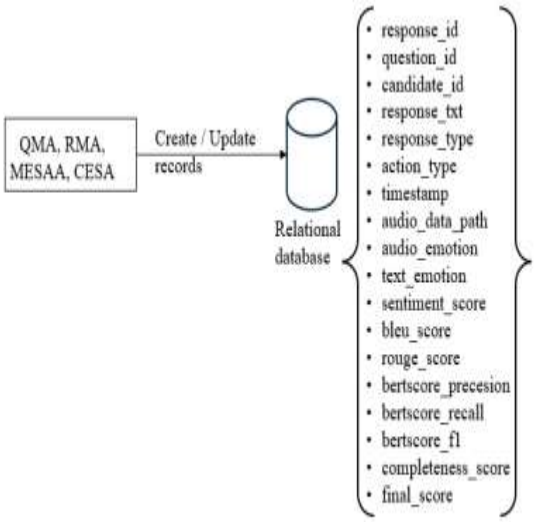

Fig. 3. Data Management and Record Maintenance in the Interview System

Mesa Al-Sabah In-depth study of a candidate's sentimental and emotional manifestations throughout the interview process is the goal of MESAA. MESSA analyses voice parameters including pitch, loudness, tone, and speech rate using CNN and other deep learning algorithms. The candidate's emotional state and degree of involvement may be better understood with the use of this analysis, which helps detect emotions such as happiness, sadness, anger, and worry. In order to identify certain feelings and thoughts expressed in text, MESAA employs a Bidirectional Long Short-Term Memory (BiLSTM) network. Section F. C. ESA–Comprehensive Evaluation and Scoring Agent CESA compiles and analyses the varied interview data gathered by different specialised agents.

The main purpose of CESA is to score candidates accurately and comprehensively while also providing an accurate, fair, and all-encompassing evaluation of each individual. CESA assesses the quality of answers using a variety of measures. Criteria for judging the quality of responses: When assessing the candidate's performance, there aren't many important criteria to consider in terms of answer quality and relevancy. Various parts of the answer may be better understood with the help of the following metrics. • Evaluation that is bilingual Understudy (BLEU): BLEU measures how similar a candidate's response is to a reference answer in terms of words and n-grams. The more close the BLEU score is to the reference, the more likely it is that the candidate's response will utilise phrases and words that are similar to the reference. Metrics developed for the remember-Oriented Understudy for Gisting Evaluation (ROUGE) provide an emphasis on candidates' ability to remember and include relevant information in their answers.

Two ROUGE variations were developed to evaluate the response's coherence and structure: ROUGE-1, which evaluates the overlap of unigrams, and ROUGE-L, which examines the longest common sequence. In the interview system, ROUGE scores assist evaluate whether applicants cover important ideas or concepts, and this measure is helpful for summarising. • BERTScore: This contextual embedding technique compares candidates' responses to reference answers based on their semantic similarity. When evaluating replies, BERTScore considers more than just word matches, unlike BLEU and ROUGE. By comparing two texts with one serving

as the "ground truth" and the other as the "candidate," we may get a score between zero and one that represents the degree to which the two texts are comparable. A score for completeness is determined by using Alg. 1. A strong and multi-faceted assessment of candidate replies is produced by integrating these indicators.

Algorithm 1 Calculating Completeness Score

```
Split the candidate answer into list of individual sentences.
Split the actual answer into list of individual sentences.
Initialize coverage to 0.
for each sentence in candidate answer list do
    Initialize sentence coverage as an empty list.
    for each answer in actual answer list do
        Calculate SIMILARITY for answer & sentence.
        Append SIMILARITY to sentence coverage.
    end for
    Add maximum value in sentence coverage to coverage value.
end for
Divide coverage by length of candidate sentence list & save as completeness score.
```

V. EXPERIMENTAL RESULTS

On Google Colab, the BiLSTM and CNN models were trained using NVIDIA T4 Tensor Core GPUs. In this research, we recorded every possible measure for training and assessment. Part A: BiLSTM for Text-Based Emotion and Sentiment Analysis Precision: After 14 iterations, the model was able to reach a general precision of 70% on the training dataset. On the other hand, the validation set shows a 60% accuracy rate, which is considered moderate for text-based emotion classification.

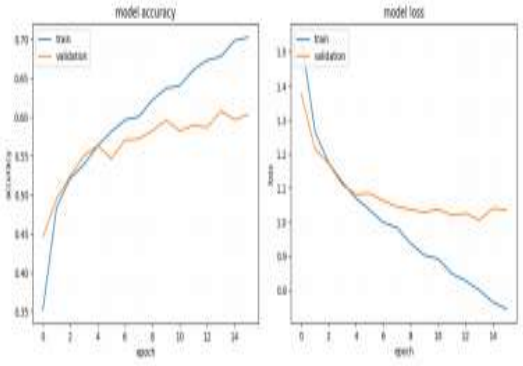

Fig. 4. Text Emotion Recognition Model Training and Validation Accuracy

Class 3 (Joy) has the best recall (0.76) and F1-score (0.70), indicating that the model does a good job of recognising occurrences of this class, according to precision, recall, and F1-score. It is challenging to accurately identify all occurrences of Class 4 (neutral) due to its poorer recall (0.47) and F1-score (0.52). Anger, contempt, fear, sorrow, and surprise are represented by classes 0, 1, 2, and 6, correspondingly.

	precision	recall	f1-score	support
0	0.57	0.66	0.61	228
1	0.58	0.50	0.54	208
2	0.65	0.59	0.62	235
3	0.65	0.76	0.70	224
4	0.59	0.47	0.52	190
5	0.62	0.54	0.58	208
6	0.59	0.71	0.64	202
accuracy			0.61	1495
macro avg	0.61	0.60	0.60	1495
weighted avg	0.61	0.61	0.60	1495

Fig. 5. Precision, Recall and F1-Score for Emotion Detection from Text

Section B. Convolutional Neural Networks for Detecting Emotions in Audio Precision: On the training dataset, the model attained a total accuracy of 95%, whereas on the validation dataset, it only managed 67%.

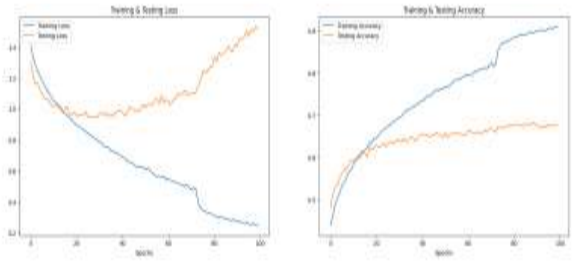

Fig. 6. Audio Emotion Recognition Model Training and Validation Accuracy

With an F1-score of 0.86, the "surprise" class offers the best mix of recall and accuracy. Despite having a moderate recall (0.65), the class "disgust" has the lowest accuracy (0.57) and F1-score (0.61), indicating that the model has difficulty accurately predicting disgust.

	precision	recall	f1-score	support
angry	0.79	0.75	0.77	1099
calm	0.74	0.82	0.78	130
disgust	0.57	0.65	0.61	1098
fear	0.68	0.61	0.64	1082
happy	0.64	0.64	0.64	1122
neutral	0.67	0.62	0.64	1001
sad	0.65	0.68	0.67	1138
surprise	0.83	0.90	0.86	495
accuracy			0.68	7165
macro avg	0.70	0.71	0.70	7165
weighted avg	0.68	0.68	0.68	7165

Fig. 7. Precision, Recall and F1-Score for Emotion Detection from Audio

C. Provide measures for evaluating quality Emotional intelligence and topic knowledge are two distinct criteria that are evaluated independently. Score for emotions: Scores for emotions are determined by classifying anticipated and unexpected feelings according to the criteria laid forth in Table. II. On a scale from 0 to 1, each emotion is given an own score. Expected and favourable feelings are given higher ratings, whereas less desirable emotions are given lower marks. The greatest score goes to happiness, while the lowest goes to rage. We use Eqs. 1 and 2 to get the final emotion score for each question after capturing them from voice and text.

$$\text{Emotion Score} = \frac{\text{Voice Emotion Score} + \text{Text Emotion Score}}{2} \quad (1)$$

$$\text{Average Emotion Score} = \frac{\sum_{i=1}^{n} \text{Emotion Score}_i}{n} \quad (2)$$

Number of questions is denoted by n. • Emotion Scorei represents the score for the i-th question in terms of emotions. • By adding together all of the scores and then dividing by the total number of questions, the equation determines the average emotion score. The overall knowledge score for response assessment is determined by adding together all of the individual scores in a weighted manner, as shown in Eq. 3.

$$\begin{aligned}\text{Final Score} = {} & 0.1 \times \text{BLEU} \\ & + 0.2 \times \text{Completeness} \\ & + 0.3 \times \text{ROUGE-L F1} \\ & + 0.4 \times \text{BERTScore}\end{aligned} \quad (3)$$

In order to conduct this experiment, we randomly chose a 4-line response from the dataset and generated 5 potential replies. potential respondent's response 1 offered a solution that was essentially a paraphrase of the original, with only the opening sentence remaining. Candidate 2's response consisted of two lines of paraphrased text.

TABLE II
CLASSIFICATION AND SCORING OF EXPECTED VS. UNEXPECTED EMOTIONS

Interview response	Emotion	Score
Expected emotion	Fear	0.5
	Surprise	0.6
	Neutral	0.8
	Happy	1
Unexpected emotion	Anger	0.1
	Disgust	0.2
	Sad	0.4

according to the first response, and the same holds true for the subsequent two candidates. The response given by the fifth contender was a carbon duplicate of the first. Each candidate's score was determined independently and then added together to get the final

score. Figure 8 shows the graph displaying the findings. A knowledge score of 0.45 or above is deemed a valid response in the plotted graph; answers between 0.2 and 0.45 are deemed incomplete and need more explanation.

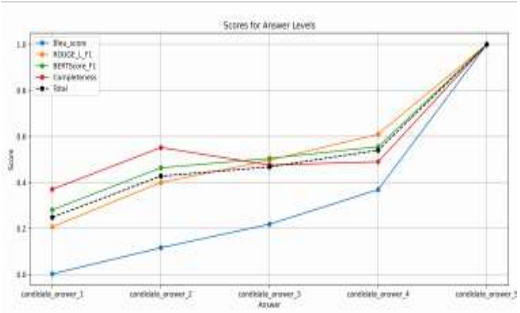

Fig. 8. Metrics of Scores at Various Answer Levels

VI. CONCLUSION

The automated interview system uses artificial intelligence and machine learning to efficiently evaluate candidates by detecting emotions from audio and text, analysing verbal responses, and responding to conversations. By combining key components including QMA, RHA, MESAA, and CESA, the system offers a comprehensive assessment of applicants. Future updates will allow the system to detect and respond to candidates' nonverbal clues, facial expressions, and body language in real-time. Additionally, testing and real-world deployment across varied candidate pools and recruiting procedures will be the focus of future effort. Finally, applicant assessment and hiring have taken a giant leap forward with the advent of the AI-powered automated interview system.

REFERENCES

1. J. Attupuram, P. Sequeira, and A. H. Sequeira, "Talent Acquisition Process in a Multinational Company: A Case Study," Management of Innovation e-Journal, CMBO, Dec. 24, 2015.
2. T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent Trends in Deep Learning Based Natural Language Processing [Review Article]," IEEE Computational Intelligence Magazine, vol. 13, no. 3, pp. 55-75, Aug. 2018.
3. O. Abdel-Hamid, A.-R. Mohamed, H. Jiang, L. Deng, G. Penn, and D. Yu, "Convolutional Neural Networks for Speech Recognition," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 22, no. 10, pp. 1533-1545, Oct. 2014.
4. G. Xu, Y. Meng, X. Qiu, Z. Yu, and X. Wu, "Sentiment Analysis of Comment Texts Based on BiLSTM," IEEE Access, vol. 7, pp. 51522- 51532, 2019.
5. T. Wu et al., "A Brief Overview of ChatGPT: The History, Status Quo and Potential Future Development," IEEE/CAA Journal of Automatica Sinica, vol. 10, no. 5, pp. 1122-1136, May 2023.
6. K. Priya, S. M. Mansoor Roomi, P. Shanmugavadivu, M. G. Sethuraman, and P. Kalaivani, "An Automated System for the Assessment of Interview Performance through Audio & Emotion Cues," 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore, India, 2019, pp. 1049-1054.
7. P. Senarathne, M. Silva, A. Methmini, D. Kavinda, and S. Thelijjagoda, "Automate Traditional Interviewing Process Using Natural Language Processing and Machine Learning," 2021 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2021, pp. 1-6.
8. I. Naim, M. I. Tanveer, D. Gildea, and M. E. Hoque, "Automated Analysis and Prediction of Job Interview Performance," IEEE Transactions on Affective Computing, vol. 9, no. 2, pp. 191-204, April-June 2018.
9. C. Kim, J. Choi, J. Yoon, D. Yoo, and W. Lee, "Fairness-Aware Multimodal Learning in Automatic Video Interview Assessment," in IEEE Access.
10. A. W. Romadon, K. M. Lhaksmana, I. Kurniawan, and D. Richasdy, "Analyzing TF-IDF and Word Embedding for Implementing Automation in Job Interview Grading," 2020 8th International Conference on Information and Communication Technology (ICoICT), Yogyakarta, Indonesia, 2020, pp. 1-4.
11. R. Pandey, D. Chaudhari, S. Bhawani, O. Pawar, and S. Barve, "Interview Bot with Automatic Question Generation and Answer Evaluation," 2023 9th International Conference on Advanced Computing and

Communication Systems (ICACCS), Coimbatore, India, 2023, pp. 1279- 1286.
12. M. D. Pickard and C. A. Roster, "Using computer automated systems to conduct personal interviews: Does the mere presence of a human face inhibit disclosure?," Computers in Human Behavior, vol. 105, 2020, Art. no. 106197.
13. M. Yusuf and K. M. Lhaksmana, "An Automated Interview Grading System in Talent Recruitment using SVM," 2020 3rd International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 2020, pp. 34-38.
14. C. Fang, N. Markuzon, N. Patel, and J.-D. Rueda, "Natural Language Processing for Automated Classification of Qualitative Data From Interviews of Patients With Cancer," Value in Health, vol. 25, no. 12, pp. 1995-2002, 2022.
15. Z. Jiang et al., "Multimodal Mental Health Digital Biomarker Analysis From Remote Interviews Using Facial, Vocal, Linguistic, and Cardiovascular Patterns," IEEE Journal of Biomedical and Health Informatics, vol. 28, no. 3, pp. 1680-1691, Mar. 2024.
16. E. P. Beggrow, M. Ha, and R. H. Nehm, "Assessing Scientific Practices Using Machine-Learning Methods: How Closely Do They Match Clinical Interview Performance?," J. Sci. Educ. Technol., vol. 23, no. 2, pp. 160-182, 2014.
17. S. Haq, P. J. Jackson, and J. Edge, "Speaker-dependent audio-visual emotion recognition," in Proc. AVSP, vol. 2009, pp. 53-58, 2009.
18. N. Neubauer and K. Dupuis, "Toronto Emotional Speech Set (TESS)," Aging and Communication Lab, Univ. of Toronto, 2011, Scholars Portal Dataverse.
19. H. Cao, D. G. Cooper, M. K. Keutmann, R. C. Gur, A. Nenkova, and R. Verma, "CREMA-D: Crowd-sourced Emotional Multimodal Actors Dataset," IEEE Transactions on Affective Computing, vol. 5, no. 4, pp. 377-390, Oct.-Dec. 2014.
20. K. R. Scherer and H. G. Wallbott, "International Survey on Emotion Antecedents and Reactions (ISEAR) [Data set]," Swiss Center for Affective Sciences, Univ. of Geneva, 1997.
21. A. Pal, L. K. Umapathi, and M. Sankarasubbu, "Medmcqa: A largescale multi-subject multi-choice dataset for medical domain question answering," in Conf. Health, Inference, and Learning, PMLR, 2022.