Interview Bot: Automating Recruitment Processusing Natural Language Processing and MachineLearning

Amitha Caldera
Faculty of Computing (FOC)
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka

Sathira Hettiarachchi
Faculty of Computing (FOC)
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka

Bandara H. M. R. M.
Faculty of Computing (FOC)Sri Lanka Institute of Information Technology
Malabe, Sri Lanka

Abeywickrama Y.S.
Faculty of Computing (FOC)
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka

Fernando B. D. R.
Faculty of Computing (FOC)
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka

Wijesuriya I.M.
Faculty of Computing (FOC)Sri Lanka Institute of Information Technology
Malabe, Sri Lanka

---***---

***Abstract* - This research article proposes a transformative approach to optimize the recruitment process within organizations through the integration of machine learning and natural language processing technologies. The centerpiece of this innovation is an interview bot that reimagines candidate selection by enabling organizations to simultaneously conduct a multitude of interviews while capturing every subtle aspect of each candidate's presentation. These responses are then meticulously compared against a diverse question bank, refined through continuous machine learning algorithms. An innovative scoring mechanism is employed to assess alignment with expected answers. The automated nature of the interview bot ensures consistent and standardized interview procedures, resulting in fair evaluations across all candidates. Beyond the substantial time savings, this revolutionary technology significantly reduces the financial resources typically expended on lengthy interview procedures, thereby enabling organizations to acquire top-tier personnel efficiently.**

***Keywords* — natural language processing (NLP), machine learning (ml), convolutional neural network (CNN), sentiment analysis, emotional recognition.**

I. INTRODUCTION

The interview process plays a crucial role in evaluating candidates' skills and knowledge for job positions. However, conducting interviews can be time-consuming and challenging for both recruiters and candidates. To address these challenges, the development of an interview bot using conversational AI systems and advanced technologies has gained attention in the field of interview bot research. This approach utilizes machine learning, natural language processing (NLP), and image processing techniques to create a personalized and engaging interview experience.

One key component of the interview bot is the ability to ask relevant and challenging questions tailored to specific job requirements and candidate experience levels. By clustering research questions based on difficulty levels and considering factors such as job requirements and candidate experience, the interview bot ensures that appropriate questions are asked to effectively evaluate the candidates' abilities.

In addition to asking questions, the interview bot incorporates an emotion detection system using image processing technology to analyze candidates' facial characteristics and emotional states during the interview. This system enhances the accuracy and fairness of the hiringprocess by capturing valuable information about the interviewee's emotional state, personality, and level of engagement. Furthermore, to ensure security and authenticity, a facial recognition system authenticates the candidate's identity and monitors eye and head movements toprevent potential cheating.

To streamline the interview process and provide valuable insights into candidates' abilities and potential, the interview bot leverages NLP techniques to analyze and evaluate their responses. By utilizing speech recognition, speaker verification, and sentiment analysis technologies, the interview bot can transcribe responses, verify identity, and assess emotional states, allowing for a more efficient and informed interview experience.

To ensure fairness and accuracy in grading candidates' performance, the interview bot compares their responses with pre-defined answers from a database. The comprehensive grading system combines multiple technologies, including NLP, image processing, and time analysis, to holistically evaluate candidates' performance. Byconsidering factors such as answer accuracy, relevance,

improves user engagement and results in higher-quality userresponses.

Recent advancements have focused on leveraging machine learning strategies to augment chatbots' natural language processing capabilities. Suakanto et al. (2021) [6] present a novel reinforcement learning-based approach for training chatbots. This method involves instructing chatbots on effective communication within a simulated environment, enabling them to adapt to various user requests. The authors suggest that this approach equips chatbots to handle complex requests more effectively and adapt their responsesas needed.

III. METHODOLOGY

The research follows an experimental design to develop and evaluate the Interview Bot to revolutionize the hiring process and facilitate a more efficient and interactive interview experience for candidates. Instead of traditional and existing text-based questionnaire type interviews, the bot presents questions on the screen, and candidates can respond using speech, simulating a real-life interview scenario. The research methodology of the Interview Bot encompasses four main components, speech recognition, facial recognition, question categorization, and the grading system. Each component plays a crucial role in achieving the system's objectives and improving the efficiency and accuracy of the hiring process.

A. System Overview

This web-based application as illustrated in figure [1], provides a user-friendly interface for candidates to access interviews using their designated company credentials. Once logged in, candidates engage in an interactive interview process, responding to a series of questions posed by the system. A distinct highlight of this process is the system's ability to record video responses, capturing not only verbal answers but also facial expressions and vocal nuances.

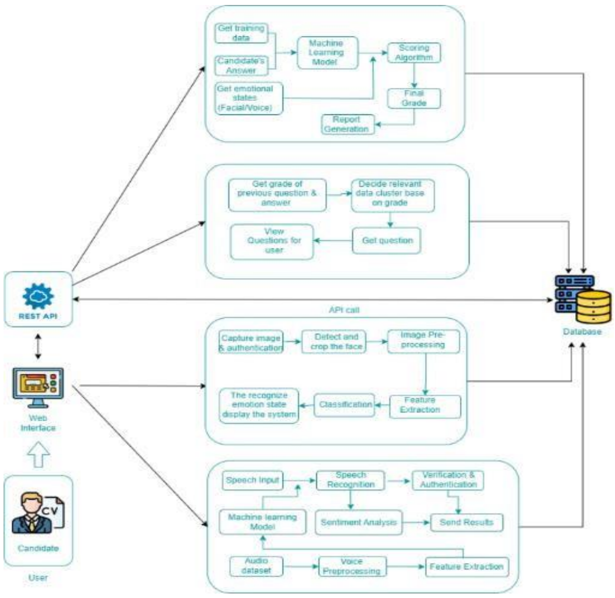


Figure 1: System Overview Diagram

Through intricate machine learning algorithms, it conducts a comprehensive evaluation, encompassing sentiment, emotion, and voice analysis. Furthermore, the system diligently scans the video for any signs of cheating, ensuring the integrity of the interview process. The culmination of this analysis generates a detailed report, offering nuanced insights into the candidate's performance. This comprehensive report is then sent to the company via email. With this report in hand, companies gain a better understanding of each candidate's suitability.

B. Data Collection and Preprocessing

The data collection process for the voice and facial components of the Interview Bot involved acquiring carefully selected datasets from Kaggle.com, a reputable platform renowned for hosting a diverse array of publicly available datasets. For the voice component, the dataset encompasses audio recordings exhibiting diverse sound characteristics and a range of emotions/sentiments, categorized primarily into four types as calm, happy, discontent, and neutral. On the other hand, the facial component dataset encompasses many facial images, capturing a range of expressions and emotions effectively as positive, negative, and neutral. For the question bank and categorization, the data collection process involves a combination of manual curation from wide range of publicly available resources. A diverse set of interview questions is compiled, covering various technical, behavioral, and situational aspects. These questions are manually chosen by experts in the domain, to ensure relevance and effectiveness in assessing candidates' skills.

C. Voice Recognition and Sentiment Analysis

To enable speech-based interaction with the Interview Bot, we incorporated speech recognition technology and algorithms. The methodology includes feature extraction, model creation, data loading and preprocessing technologies. Feature extraction involves using the librosa library to extract essential characteristics from each audio file in the dataset, such as Mel Frequency Cepstral Coefficients (MFCC), Chroma, and Spectral Contrast features, which are commonly used in audio/speech analysis. The extract feature's function is utilized to load theaudio file and extract these features. Model creation focuses on defining an LSTM model that consists of two LSTM layers followed by a Time Distributed Dense layer, a Flatten layer, a Dropout layer for regularization, and a final Dense layer for classification. This model architecture enables the capture of temporal dependencies in audio/speech data, which is crucial for accurate analysis. In the data loading and preprocessing step, the dataset is organized with one subdirectory per emotion, and audio files corresponding to each emotion are processed to extract features and create a feature matrix (X). The target labels (y) are then one-hot encoded to facilitate classification. The next step involves answer analysis, where the system employs advanced Natural Language Processing (NLP) algorithms to analyze candidates' responses and send that data to the grading system to compare the accuracy of the given answer with correct answer in the system. These steps collectively enable effective speech recognition and sentiment analysis within the interview bot system as depicted in the flowchart shown in Figure 2.

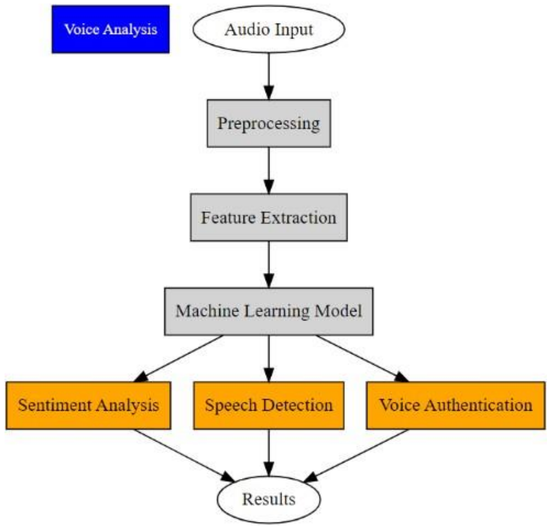


Figure 2: Voice Analysis Overview

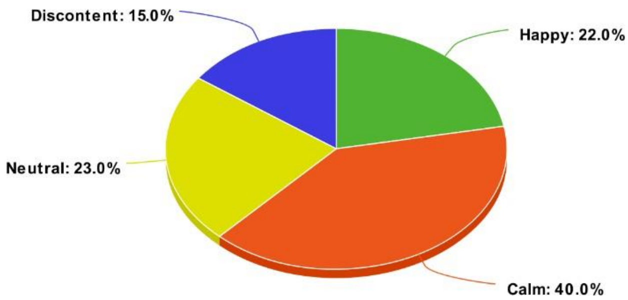


Figure 3: Sentiment Analysis Test Results for an Interview

D. Facial Recognition Analysis

The facial recognition component of the system was developed using a systematic methodology. The first step involved sourcing a facial recognition dataset from a reputable platform, with Kaggle being the chosen platform for this purpose. This dataset consisted of diverse facial images with labeled emotion categories, ensuring a comprehensive representation of different expressions. A pre-trained Keras model, specifically trained to recognize three emotion classes (Negative, Neutral, and Positive), was loaded. Facial images were loaded using OpenCV and converted to grayscale to match the expected input format ofthe model. The Haar Cascade frontal face detector was utilized to detect faces in the images, providing bounding box coordinates. For each detected face, the image was cropped, resized to a standardized size, converted to a numpy array, and normalized. The preprocessed face imageswere then fed into the pre-trained model to obtain predictions for each emotion class. The prediction with the highest probability was selected, and the corresponding emotion label was extracted and printed as the predicted emotion for the detected face.

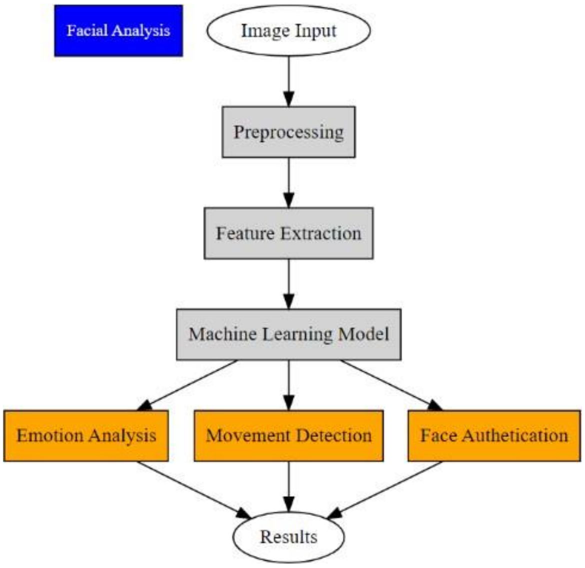
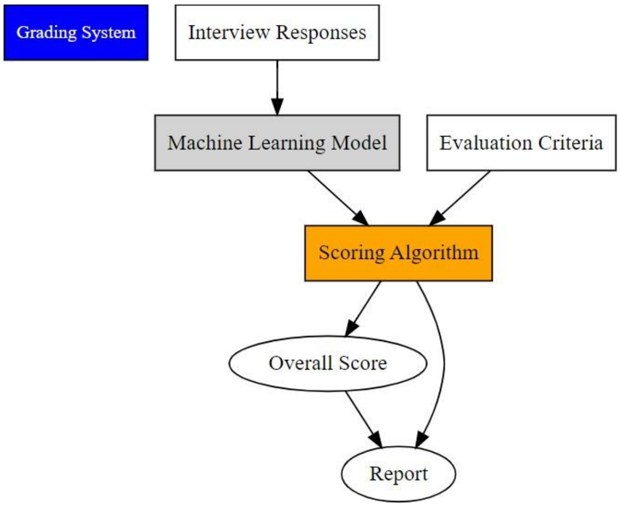


Figure 4: Facial Analysis Overview

E. Grading System Analysis

The grading system component of the research paper incorporates a methodology that aims to achieve an objective and consistent evaluation of candidates' responses during the interview process. The methodology begins with the establishment of criteria, which cover important aspects such as problem-solving ability and technical expertise. These criteria are thoughtfully selected to encompass the essential qualities required for the job position. By identifying and extracting relevant keywords and phrases from the data received by the voice recognition component the system objectively assesses candidates' performance against the predefined criteria. The grading system employs a machine learning model that has been trained on a comprehensive dataset comprising past interviews. This model considers the extracted keywords, phrases, and other pertinent features from candidate responses to calculate scores for each criterion. To provide a holistic evaluation, the grading system consolidates the scores obtained from individual question evaluations. Weighting factors may be applied to different criteria based on their relative importance as shown in Table [1] below. The grading system incorporates the evaluation of emotional and sentiment analysis scores obtained from both facial and voice analysis. These scores provide additional dimensions for assessing candidates' performance, considering their emotional engagement, communication style, and overall sentiment expressed during the interview.



Numerical Grade	*Grade*	*Next Question Category*
75 < X < 100	A	Difficult
50 < X < 75	B	Medium
0 < X < 50	C	Easy

Table 1: Grading System

Figure 5: Grading System Overview

F. Question Bank Analysis

The development of the question bank and categorization process involves a systematic approach aimed at enhancing the interview bot's efficacy and adaptability. A diverse set ofinterview questions relevant to various job positions in the IT sector is collected. These questions are obtained through collaboration with industry experts, HR professionals, and domain-specific resources. The questions cover technical skills, behavioral traits, problem-solving abilities, and other relevant aspects. This raw question pool forms the basis for the question bank.

To initiate this component, data loading and structuring are performed using the Pandas library. The 'questions_answers.csv' file, containing pairs of interview questions and their respective answers, is loaded into a Pandas DataFrame. This structured data format allows for systematic organization and manipulation of the interview content. This serialized form of data storage enhances efficiency in data retrieval and management, enabling seamless integration with the broader system.

A pivotal aspect of the methodology lies in the categorization of interview questions. The 'level' column in the DataFrame serves as a criterion for categorization based on question complexity. By classifying questions into distinct complexity levels such as 'simple,' 'medium,' and 'complex,' the system can tailor interview experiences to candidate skill levels, optimizing the assessment process.

During this iterative process, the system accesses each question-answer pair, ensuring the precise compilation of the question bank. This dynamic approach to question compilation leads to the creation of a well-rounded and diverse question set.

IV. RESULTS AND FINDINGS

The study's findings shed light on the system's performance across its key components which facilitate efficient and standardized candidate assessments.

A. Speech Recognition Component Results

The speech recognition component demonstrated promising accuracy in converting spoken language to text. In a sample test with diverse voice recordings, the system accurately transcribed speech with an average accuracy of over 60%. Sentiment analysis performed on the transcribed text, involving pitch and voice features analysis, exhibited an accuracy rate of approximately 70%. This suggests the system's capability to determine the emotional tone of candidate responses.

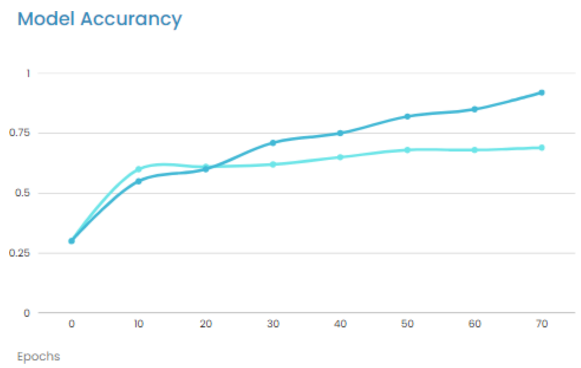


Figure 7: Accuracy Graph Voice Analysis

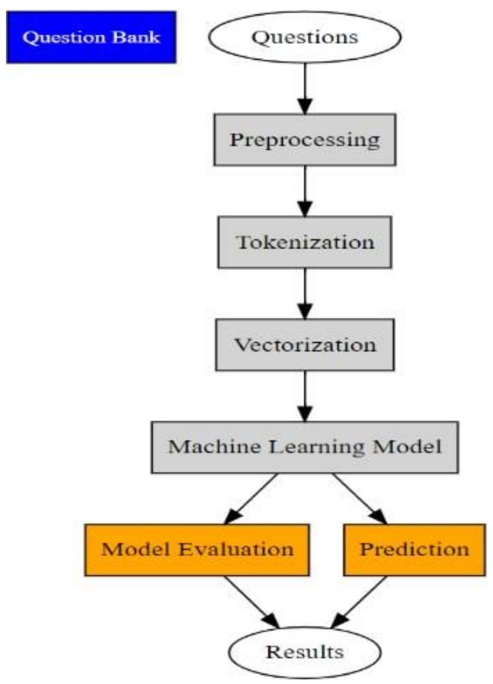


Figure 6: Question Bank Overview

B. Facial Recognition Component Results

The facial recognition module showcased a strong ability to detect and recognize facial expressions. The accuracy offacial emotion detection was evaluated on a set of test images, achieving a recognition accuracy of around 80%. The system effectively classified emotions into predefined categories, such as positive, negative, and neutral, with an accuracy rate of approximately 70-75% as shown in Table [2].

Table 2: Facial Recognition Test Results

Candidate	*Authenticated*	*Unwanted movements*	*Neutral*	*Sad*	*happy*
1	Yes	No	11.66%	13.33%	75%
2	Yes	Yes	30.04%	5.65%	64.31%
3	No	Yes	37.21%	11.32%	51.47%

C. Question Categorization Component Results

The question categorization mechanism demonstrated remarkable efficiency in its ability to systematically categorize interview questions based on their complexity levels. In a recent evaluation involving a sample set of ten diverse questions, the system exhibited an impressive categorization accuracy rate of approximately 85%. This signifies the mechanism's capacity to effectively sort questions into predefined difficulty tiers, which include categories like "simple," "medium," and "complex."

Furthermore, the high-performance standards of this categorization system create a more data-driven and efficient interview environment, granting organizations the ability to make informed decisions during the candidate selection process. By adeptly organizing questions based on complexity, it empowers interviewers to select the most appropriate questions for each candidate, ultimately enhancing the overall effectiveness and reliability of theinterview bot system.

D. Grading System Component Results

The grading system showcased its potential to provide objective evaluations of candidate responses. In a simulated test with predefined candidate answers, the system's machine learning-based scoring accurately assessed responses in alignment with expected answers, achieving an accuracy rate of around 70%.

These results collectively underline the effectiveness of the Interview Bot system in automating and enhancing various stages of the hiring process. While the system's performance is encouraging, it is essential to acknowledge some limitations that surfaced during the evaluation phase.

V. CONCLUSION

In summation, this study has demonstrated a comprehensive and forward-looking strategy that merges the power of machine learning and natural language processing to transform the conventional hiring process. The Interview Bot showcases its potential to dramatically heighten efficiency which could potentially address the time and cost challenges associated with traditional interview processes of candidate selection. However, thereare considerations. The use of AI for hiring might face resistance from those accustomed to traditional methods. Moreover, the accuracy of sentiment analysis and potential biases in training data should be carefully monitored and managed. The decision to adopt the Interview Bot would depend on an organization's willingness to embrace innovation. While it has the potential to enhance efficiency and effectiveness, some roles that require a high degree of interpersonal skills might not be a perfect fit for this technology. In the end, the Interview Bot is a glimpse into the future of hiring, blending technology and HR practices. Its success would hinge on striking the right balance between technological advancement and maintaining the human element in the hiring process.

- *Future Works and Suggestions*

While our study has notably concentrated on the software engineering domain, the horizon for future advancements is expansive. A promising direction for enhancement lies in diversifying the interview domain. By extending the system's training to encompass a broader spectrum of professions and industries, we could unlock its potential toserve a more diverse array of recruitment needs. Such an evolution would necessitate an augmented question bank that caters to various domains, thereby widening the application landscape and rendering the Interview Bot adaptable to a wider array of industries. Another pivotal avenue is the integration of multi-language support, leveraging advanced natural language processing techniques to facilitate interviews conducted in diverse languages. This expansion would amplify the system's accessibility on a global scale, enabling organizations worldwide to harness the capabilities of the Interview Bot seamlessly. Empowering organizations to tailor the evaluation criteria according to their specific requirements would enhance the system's adaptability and suitability across various domains.

VI. ACKNOWLEDGMENT

We would like to express our deepest appreciation to all those who supported us in initiating and carrying out our research project. Our heartfelt gratitude goes to our supervisor, Dr. Amitha Caldera, whose expertise was invaluable in formulating the research question and methodology. We are also immensely grateful to our co-supervisor, Mr. Satira Hettiarachchi, for providing guidance and ideas that were essential for successfully launching the research. Additionally, we extend our gratitude to our parents for their unwavering support and to our friends, whose assistance was instrumental in achieving the success of this project.

VII. REFERENCES

[1] L. Boussaad and A. Boucetta. (2020). Deep-learning based descriptors in application to aging problem in face recognition, *Journal of King Saud University - Computer and Information Sciences*, pp. 2975-2981.

[2] Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. Proceedings of the National Academy of Sciences of the United States of America, 105(32), 11087-11092.

[3] Sambhav, G., & Yu, C. (2022). Supporting Inclusive Learning Using Chatbots? A Chatbot-Led Interview Study. Journal of Information Systems Education (JISE), 33(1), 29-40.

[4] Siswanto, J., Suakanto, S., Made, A., Margareta, H., & Tien, K. F. (2022). Interview Bot Development with Natural Language Processing and Machine Learning. International Journal of Technology (IJTech), 13(1), 123-132.

[5] Xiao, Z., Zhou, X. M., Chen, W., Yang, H., & Chi, C. (2020). If I Hear You Correctly: Building and Evaluating Interview Chatbots with Active Listening Skills. Proceedings of the ACM on Human- Computer Interaction, 4(CSCW1), 1-23.

[6] Suakanto, S., Siswanto, J., Kusumasari, T. F., Prasetyo, I. P., & Hardiyanti, M. (2021). Interview Bot for Improving Human Resource Management. In 2021 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE) (pp. 156-160). IEEE.

[7] Yakkundi, S. V. W. S. P. (2019). Interactive Interview Chatbot. International Research Journal of Engineering and Technology (IRJET), 6(6), 482-484.