



AI Based Interview Critique System: Interview Preparation Companion Using Deep Learning.

¹ Assistant Prof. Nirgide Shubhangi Vishal, ² Sayyed Arsh Aktharali, ³ Patil Paresh Narendra, ⁴ Raktate Shriraj Vikas, ⁵ Pathan Md Fazal Mushtaque

¹ Designation of 1st Author: Assistant Professor, ² Designation of 2nd Author: Student, ³ Designation of 2nd Author: Student, ⁴ Designation of 2nd Author: Student, ⁵ Designation of 2nd Author: Student.

1, Assistant Professor, Department Of Artificial Intelligence And Data Science, K. K. Wagh Institute Of Engineering Education And Research, City: Nashik, Country: India.

2,3,4,5, Students, Department Of Artificial Intelligence And Data Science, K. K. Wagh Institute Of Engineering Education And Research, City: Nashik, Country: India.

Abstract

In today's evolving recruitment landscape, the introduction of an innovative AI-powered mock interview evaluator revolutionizes candidate assessment while fostering skill enhancement. This system goes beyond traditional evaluation metrics, utilizing advanced technologies such as natural language processing and deep learning to gauge not only knowledge and skills but also emotional intelligence, confidence, and adaptability in real-time interviews. Unique features include a dynamic question adaptation mechanism tailored to individual responses, enabling comprehensive assessments aligned with specific job requirements. The system's feedback loop empowers candidates by providing constructive insights after each session, identifying strengths and areas for improvement to enhance their interview skills. Employing facial recognition, sentiment analysis, and speech recognition, it comprehensively evaluates both behavioral and technical competencies. Furthermore, it accommodates custom databases for domain-specific evaluations, ensuring alignment with an organization's unique needs. Compared to static evaluation methods, this AI-driven system demonstrates increased efficiency and effectiveness in candidate assessment. It stands as a pioneering approach, modernizing the interview process by offering a data-driven means to identify top

candidates while nurturing their skills through insightful feedback and analysis.

Keywords: Deep learning, Artificial Intelligence, Interview Evaluation System, Natural Language Processing.

1. INTRODUCTION

The Project aims to transform candidate assessment by introducing an AI-based mock interview evaluator. It combines speech recognition, facial recognition, emotion analysis, and knowledge assessment for a comprehensive evaluation, driven by the shortcomings of the current interview evaluations. The literature review emphasizes the value of interviews in hiring, the necessity of dynamic assessment techniques, and the growing influence of artificial intelligence in the recruiting process. With this project, we hope to close the gaps in traditional evaluations and establish new benchmarks for a more thorough and equitable interview assessment procedure.

1.1 PROJECT IDEA

By employing deep learning techniques, an AI-powered Interview Critique System functions as a customized interview assistant. Through the analysis of mock interview recordings, feedback, and customized improvement tactics, it helps people perform better in interviews. In order to provide thorough feedback on verbal and nonverbal communication, content delivery, and general interview abilities, the system combines natural language processing (NLP), sentiment analysis, and speech recognition.

1.2 MOTIVATION OF THE PROJECT

Interview preparation is critical in today's cutthroat job market, but many candidates struggle with confidence and insufficient feedback. AI shows up as a key answer to this problem. We address this need with our cutting-edge AI Interview Prep Companion, which provides individualized advice with a particular emphasis on helping applicants communicate their feelings in a productive way during distant interviews. Our platform provides candidates with the necessary skills to succeed in these virtual encounters, as online interviews grow increasingly common in today's employment environment. With the help of our AI-powered solution, people may improve their interviewing abilities, confidence, and ability to express emotions effectively. This will increase the likelihood that they will land desired jobs in their career aspirations.

1. **Paper:** Confidence Estimation for Speech Emotion Recognition Based on the Relationship Between Emotion Categories and Primitives.[1]

Author: Yang Li, Constantinos Papayiannis, Viktor Rozgic, Elizabeth Shriberg, Chao Wang
Dept. of Electrical and Computer Engineering, University of Rochester.

Description: The research paper delves into confidence estimation for Speech Emotion Recognition (SER), introducing a novel metric based on emotion primitives and categories' relationship. In order to predict confidence in emotion recognition, it suggests using a deep neural network called EmoConfidNet. The study examines current techniques for estimating confidence and highlights how data preprocessing can improve these techniques. Tested on the MSP-Podcast and IEMOCAP datasets, the suggested measure performs better than current methods in terms of F1-score and Unweighted Average Recall (UAR) with more extensive sample coverage. EmoConfidNet and the emotion classifier perform much better when data preprocessing is applied, such as filtering out conflicting emotion categories and primitives. The influence of techniques such as feature normalization, selection, and data augmentation on confidence estimation techniques is discussed. The study's contributions include the creation of the EmoConfidNet, a reliable confidence metric, and the focus on efficient data preprocessing methods for better speech emotion recognition confidence estimation.

2. **Paper:** AI-Based mock interview evaluator: An emotion and confidence classifier model[2]

Author: Rubi Mandal, Pranav Lohar, Dhiraj Patil, Apurva Patil, Suvarna Wagh

Description: The research paper introduces an AI-driven mock interview evaluation system addressing interview challenges. It underscores the importance of candidate attributes and first impressions while stressing the role of interviews in the hiring process. It highlights the importance of practicing with mock interviews and emphasizes how important AI integration is. The review criticizes current systems for ignoring feelings and the pandemic-related shift toward online interviews. It supports a more thorough evaluation methodology in line with the innovation of the suggested system. The literature review, taken together with the changing nature of interviews, provides strong support for the AI-based mock interview system and acknowledges its potential to improve candidate preparation and evaluation.

3. **Paper:** An AI Mock-interview Platform for Interview Performance Analysis[3]

Author: Yi-Chi Chou, Felicia R. Wongso, Chun-Yen Chao, Han-Yen Yu

Description: The paper introduces an AI-driven Mock-Interview Platform (MIP) assessing interview performance and personality traits. It highlights areas for improvement even as it exhibits robust predictive models. It is imperative that AI models for voice, posture, and emotion analysis are improved; improvements such as eye gaze tracking could be made to increase accuracy. In order to provide more relevant interview simulations, customization for a variety of industries is emphasized. Tailored question banks and weighted assessments are recommended. By making these improvements, the platform's feedback quality for HR professionals and candidates could be greatly improved, which would help with customized interview preparation. The study highlights the potential for MIP's growth and stresses the need for more detailed analysis and industry-specific customization in order to maximize its usefulness and efficacy across a range of employment sectors.

4. **Paper:** Development of an AI-based interview system for remote hiring[4]

Author: B C Lee, B Y Kim

Description: The paper emphasizes the growing significance of securing top talent in the Fourth Industrial Revolution. It highlights the rise of AI-based interview systems, using AI and big data to assess candidates across various factors. Especially, it simulates the system in large Korean public enterprises and focuses on real-world implementation, showing high satisfaction because of increased efficiency and fairness. This strategy increases job opportunities while saving time and resources. Recruitment that is performance-based rather than qualification-based lowers social costs, fosters justice, and lessens bias. The AI interview model is an essential instrument for identifying outstanding applicants and streamlining the hiring procedure. Overall, the paper sheds light on how the field is changing and demonstrates how AI solutions improve efficiency and fairness while changing the hiring process to include performance-based evaluation.

5. **Paper:** AI-based Behavioural Analyser for Interviews/Viva[5]

Author: Dulmini Yashodha Dissanayake, Venuri Amalya, Raveen Dissanayaka, Lahiru Lakshan, Pradeepa Samarasinghe, Madhuka Nadeeshani, Prasad Samarasinghe.

Description: The paper introduces an AI-based Behavioral Analyzer tailored for interview and viva assessments, reflecting technology's influence on recruitment and academic evaluation. Understanding interviewee behavior continues to be difficult as virtual interviews become more popular. In order to gain insight into interviewee responses during virtual interactions, this

research uses machine-based analysis of nonverbal cues, such as emotions, eye movements, smiles, and head motions. It interprets these cues with over 85% accuracy using deep learning, showcasing the potential of technology to capture minute behavioral details. The study also investigates the use of a Random Forest model, which achieves over 75% accuracy in predicting the Big Five personality traits. In summary, this review highlights the importance of AI-driven behavioral analysis in interview environments, offering the potential to improve evaluation procedures and support data-driven choices for hiring and academic assessments.

6. **Paper:** An Intelligent System for Evaluation of Descriptive Answers[6]

Author: Vinal Bagaria, Mohit Badve, Manasi Beldar, Prof. Sunil Ghane

Description: The paper introduces a system for automated assessment of descriptive student answers, combining structural and conceptual analysis with language quality evaluation. It uses methods like keyword matching, similarity scoring, and concept graph generation, which are in line with the increasing interest in automated grading. The integration of NLP tools such as NLTK and spaCy reflects current developments in text analysis research. In line with earlier NLP and educational technology research, the emphasis on ontology learning and WordNet Synsets for synonym recognition recognizes the importance of semantic understanding in assessment systems. The future scope that has been outlined, which includes non-textual data handling and data analytics-driven educational insights, is in line with the overall trend in educational technology. Researchers and educators are looking for more and more ways to use data and technology to enhance teaching and assessment practices.

2 METHEDODOLOGY

2.1 METHODOLOGY

2.1.1 Face Recognition:

Image Acquisition: The system initiates face recognition by obtaining the candidate's photo, which is uploaded during the registration process. This photo serves as a reference point for comparison during the interview.

Image Pre-processing: The acquired image undergoes several pre-processing steps. Initially, it is converted from RGB to grayscale format to simplify further analysis. Subsequently, the system crops the image to a standard size, typically 50x50 pixels, for uniformity in the dataset.

Recognition: To execute the face recognition process, the system employs a Haar-cascade classifier, a machine learning-based approach that effectively detects objects in images. Leveraging the Face Recognition library, the system compares the live image captured during the interview with the registered image to authenticate the candidate.

Encoding: Both the registered image and the live image are encoded to extract essential facial features, such as key points, contours, and structural components. These encoding are then compared to determine the similarity and verify the identity of the candidate.

Verification: Upon comparison, the system establishes the degree of similarity between the registered and live images, providing a measure of authentication based on predetermined thresholds. This process ensures the security and integrity of the interview process, preventing unauthorized access and impersonation.

2.1.2 Emotion Recognition(Video):

Deep Learning Model Training: The emotion recognition system analyzes candidates' facial reactions in real time during interviews using a 5-layered CNN model that has been extensively trained on a variety of emotional expressions.

Training the Model: Through frame segmentation, resizing, and intensity normalization, the system carefully preprocesses live interview video data to ensure standardized inputs for the CNN model and to enable accurate emotional recognition.

Pre-processing the Data: The 5-layered CNN thoroughly examines minute details, texture variations, and spatial relationships between important facial components by extracting complex features from facial frames. This allows it to provide nuanced insights into candidates' emotional states.

Feature Extraction: By utilizing its feature extraction capabilities, the CNN model tracks the temporal evolution of candidates' emotional responses during real-time interview scenarios to provide a comprehensive evaluation of their emotional engagement and adaptability.

2.1.3 Speech Recognition:

Data Preprocessing: The system uses a thorough approach that includes segmentation, normalization, and noise reduction for speech data preprocessing. The MFCC algorithm's integration makes it possible to extract important speech features like pitch, intensity, and spectral characteristics,

which improves the system's capacity to identify subtle emotional differences in the speech patterns of candidates.

LSTM Model Training: The system trains an advanced speech recognition model that can analyze speech dynamics and emotional fluctuations by making use of the Long Short-Term Memory (LSTM) architecture. The LSTM model offers a thorough evaluation of candidates' emotional state during interviews because it has been painstakingly trained to identify subtle variations in speech patterns, intonations, and tonal modulations.

Real-Time Speech Analysis: The system analyzes speech data from candidates in real-time, going into great detail and concentrating on important emotional cues like tone shifts, speech dynamics, and pitch changes. By utilizing cutting-edge signal processing methods and the extensive feature set of the MFCC algorithm, the system guarantees a detailed assessment of interviewees' emotional engagement and communication efficacy.

Model Validation: To guarantee accuracy and dependability in real-time speech recognition, the LSTM model is subjected to stringent validation processes. The model's capacity to identify emotional nuances and capture minute speech variations linked to a range of emotional states has been extensively tested. The LSTM model's output is consistent with categorical emotional classes, which makes it easier to assess candidates' emotional responses thoroughly when combined with the CNN model's output.

2.1.4 Knowledge Processing:

API Integration: The system seamlessly integrates an advanced API tailored for interacting with a large language model, enabling comprehensive document analysis and question generation. Leveraging the capabilities of the API, the system meticulously processes user-provided documents, extracting pertinent information and identifying key concepts and themes relevant to the interview domain. By harnessing the power of the large language model, the system ensures a sophisticated analysis of diverse textual data, facilitating an efficient generation of insightful and contextually relevant interview questions.

Dynamic Question Formulation: Leveraging the data insights extracted from the user-provided documents, the system dynamically formulates a diverse array of interview questions tailored to the specific requirements and intricacies of the job profile. By leveraging advanced natural language processing (NLP) techniques, the system generates questions that encompass a wide spectrum of topics and competencies, ensuring a comprehensive evaluation of the candidate's knowledge base and

problem-solving capabilities within the specified domain.

Semantic Coherence Analysis: The system executes a comprehensive semantic coherence analysis of the generated questions, ensuring their alignment with the contextual themes and subject-specific nuances derived from the user-provided documents. By validating the semantic coherence and contextual relevance of the questions, the system ensures the generation of a well-structured and comprehensive question set, tailored to comprehensively assess the candidate's knowledge proficiency and domain-specific expertise during the mock interview process.

2.1.5 Knowledge Processing Module:

Question Categorization: The system classifies questions into distinct categories to guide the evaluation process. These categories include conceptual, specific, and analytical questions. For instance, a conceptual question such as "explain a specific concept" demands a broader understanding, while a specific question like "define a particular concept" requires concise and precise answers. Analytical questions like "compare and contrast" necessitate in-depth analysis without strict adherence to the expected answer.

Question-Type Association: Each question is associated with a specific question tag that represents its nature. For example, question tags like "explain," "define," and "compare" align with conceptual, specific, and analytical questions, respectively. This alignment assists in determining the weighting for different evaluation strategies.

Text Preprocessing: The answers provided by candidates undergo preprocessing to ensure accurate evaluation. The system breaks down lengthy answers into meaningful sentences, improving computational efficiency.

Extractive Summarization: For excessively lengthy answers, an extractive summarization process is applied to condense the content, retaining the most relevant information. Gensim library in Python is used to perform this summarization.

Database Storage: The preprocessed answers, along with corresponding questions, question types, important keywords, and allocated marks, are stored in a database for streamlined and efficient retrieval.

Keyword and Similarity Score: The presence of important keywords provided by the candidate in their answers. Computing cosine similarity scores between expected answers and candidates'

responses.

Language Score: To assess writing and language proficiency, the system uses language-check libraries in Python. It identifies and quantifies grammar and spelling mistakes in candidates’ answers, and the language score is calculated accordingly.

Fuzzy String Matching Score: Utilizing fuzzy string matching, the system gauges the approximate similarity between candidates’ and expected answers. This technique accounts for partial matches, providing approximate matching scores

Concept Graph Score: While many previous systems require predefined concept graphs, the proposed system generates concept graphs automatically. These graphs represent the structure of both expected and candidate answers, which are then compared for evaluation purposes.

Synonyms and Ontology Learning: The system accounts for synonyms and related terms present in the answers. Synonyms of keywords in the expected answer are extracted using WordNet Synsets, ensuring comprehensive evaluation by considering a broader vocabulary.

Final Score Assignment: Question types are used to determine the weightage assigned to each evaluation strategy. For instance, conceptual questions might assign higher weight to the concept graph score. The final cumulative weighted score is calculated for each answer using a predefined formula, as depicted in Equations (1) and (2).

$$FS = W_{ks}S_{ks} + W_lS_l + W_fS_f + W_cS_c \tag{1}$$

$$W_{ks} + W_l + W_f + W_c = 1 \tag{2}$$

FS	Final Score
W_{ks}	Weight assigned to Keyword and Similarity Score
S_{ks}	Keyword and Similarity Score
W_l	Weight assigned to Language Score
S_l	Language Score
W_f	Weight assigned to Fuzzy String Matching Score
S_f	Fuzzy String Matching Score
W_c	Weight assigned to Concept Graph Score
S_c	Concept Graph Score

Table: Parameters

DATA SET

DATASET1:

MED4 consists of synchronously recorded signals of participants' EEG, photoplethysmography, speech and facial images when they were influenced by video stimuli designed to induce happy, sad, angry and neutral emotions. The experiment was performed with 32 participants in two environment conditions, a research lab with natural noises and an anechoic chamber.

DATASET2:

The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) is a publicly available dataset of emotional speech and song recordings. The dataset contains 7,356 recordings from 24 professional actors (12 female and 12 male), vocalizing two lexically-matched statements in a neutral North American accent. Speech emotions include calm, happy, sad, angry, fearful, surprise, and disgust expressions, and song emotions include calm, happy, sad, angry, and fearful emotions. Each expression is produced at two levels of emotional intensity (normal, strong), with an additional neutral expression. All conditions are available in three modality formats: audio-only (16bit, 48kHz .wav), video-only (MPEG-4), and full-AV (MPEG-4).

ARCHITECTURAL DESIGN

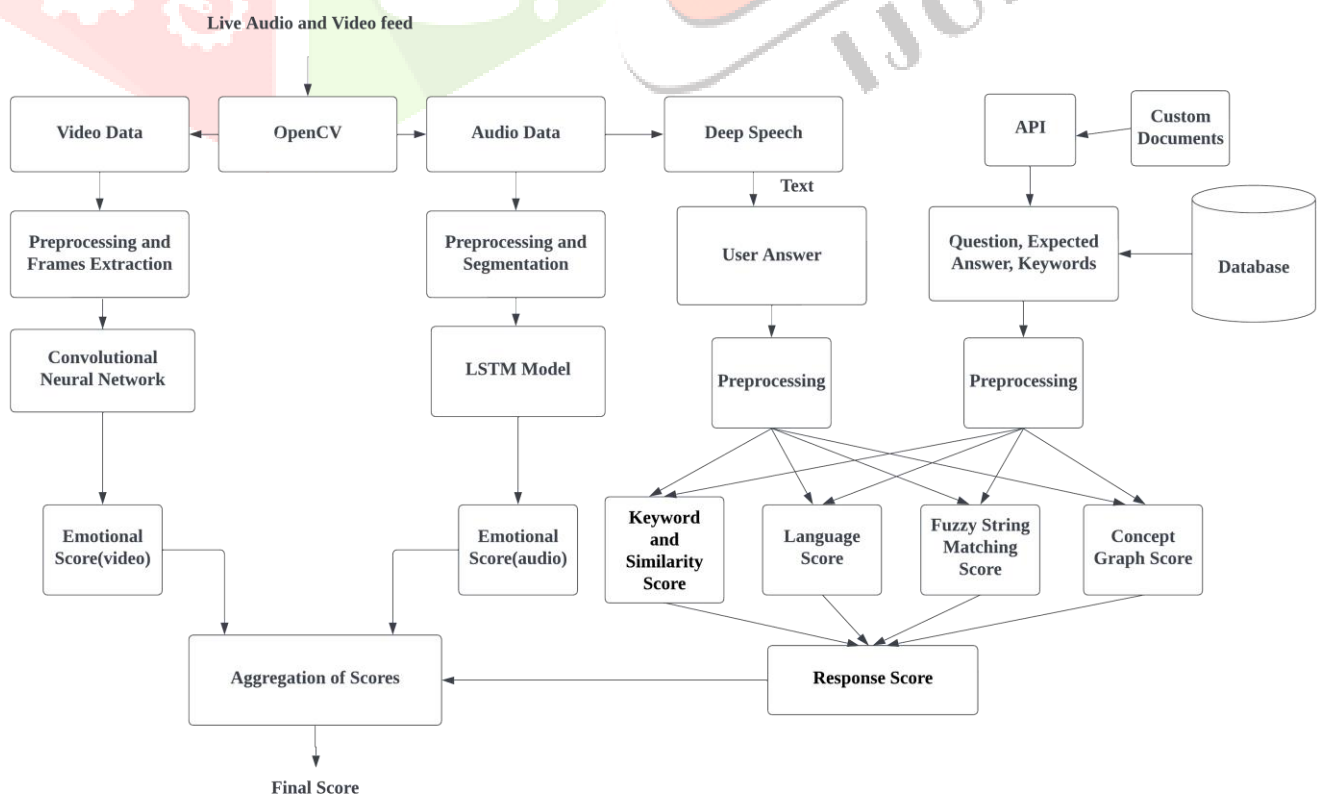


Figure: Architecture

The architecture diagram presents a thorough system intended for interview assessment. Using OpenCV for video and Deep Speech for audio, it uses real-time audio and video feeds. Convolutional Neural Networks are used in this system to predict emotional scores and extract frames from video footage. An LSTM model is used for both emotional scoring and pre-processing of audio data. The system includes user response segmentation, question-answer matching database creation, and text analysis using API for customized documents. To evaluate responses, it combines a number of scoring methods, including fuzzy string matching, concept graph scoring, language scoring, and keyword matching. In the end, this architecture provides a comprehensive assessment of interviewee performance by combining scores from various sources to produce a final score.

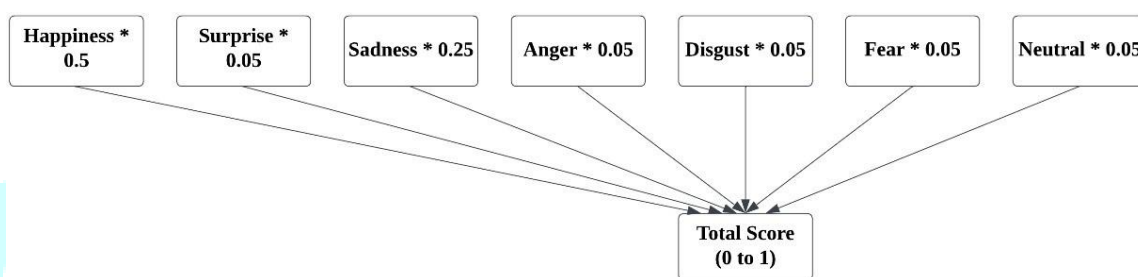


Figure 5.2: Emotion Score

The emotional scoring model quantifies different emotional states' impact by assigning weighted values. The model allocates varying scores to emotions, with happiness holding the most influence at 0.5 and sadness at 0.25. It calculates the total score (ranging from 0 to 1) by summing the weighted values of happiness, surprise, sadness, anger, disgust, fear, and neutrality. This scoring method assesses individuals' emotional states during interviews, providing an aggregated score representing their emotional disposition.

TECHNOLOGY USED

1. Facial Recognition:

- (a) Haar-cascade classifier
- (b) Face Recognition library
- (c) Image pre-processing techniques

2. Emotion Recognition (from video data):

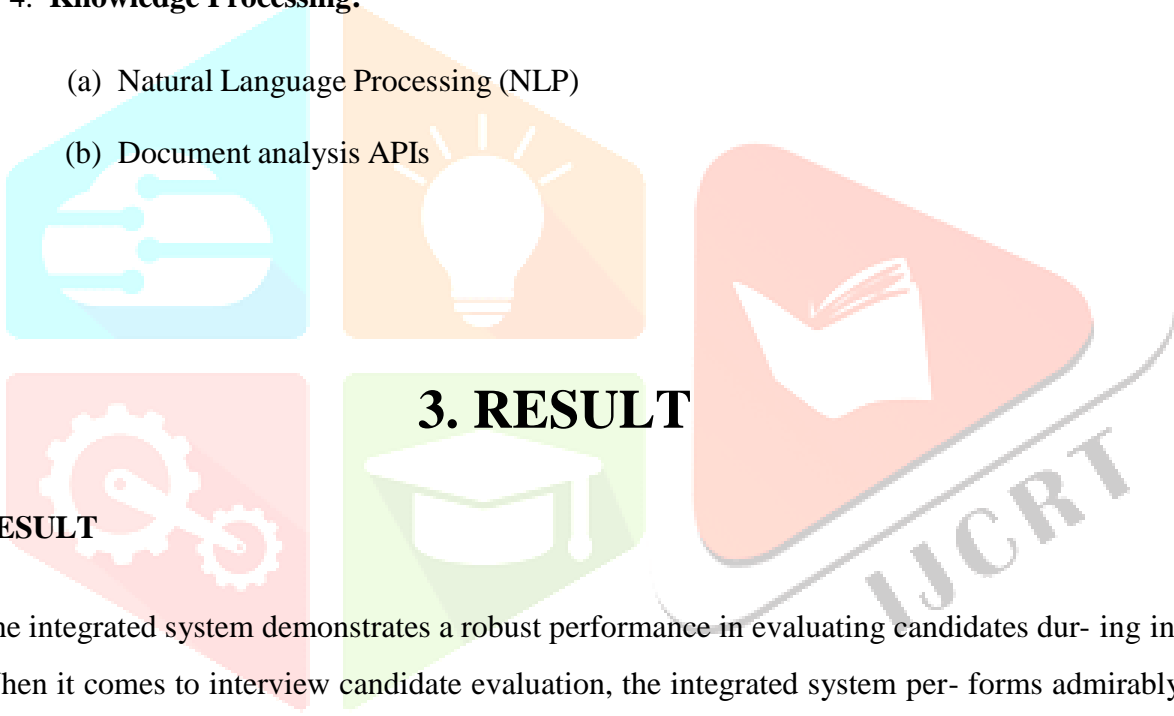
- (a) Convolutional Neural Network (CNN)
- (b) Deep learning for emotional analysis
- (c) Diverse dataset for emotional expressions

3. Speech Recognition:

- (a) Long Short-Term Memory (LSTM)
- (b) Mel-Frequency Cepstral Coefficients (MFCC)
- (c) Signal processing techniques

4. Knowledge Processing:

- (a) Natural Language Processing (NLP)
- (b) Document analysis APIs



RESULT

3. RESULT

The integrated system demonstrates a robust performance in evaluating candidates during interviews. When it comes to interview candidate evaluation, the integrated system performs admirably. With a high degree of accuracy and dependability, it guarantees a thorough evaluation of candidates' facial features, speech patterns, and emotional expressions from video data. With user-provided documents, the system efficiently creates questions that are contextually relevant and cover a broad range of subjects. In the assessment stage, candidates' responses are carefully examined, taking into account conceptual comprehension, language ability, and content accuracy. The comprehensive assessment process guarantees a comprehensive evaluation of applicants, exhibiting their expertise and subject-matter expertise.

4 DIAGRAMS

Class Diagram

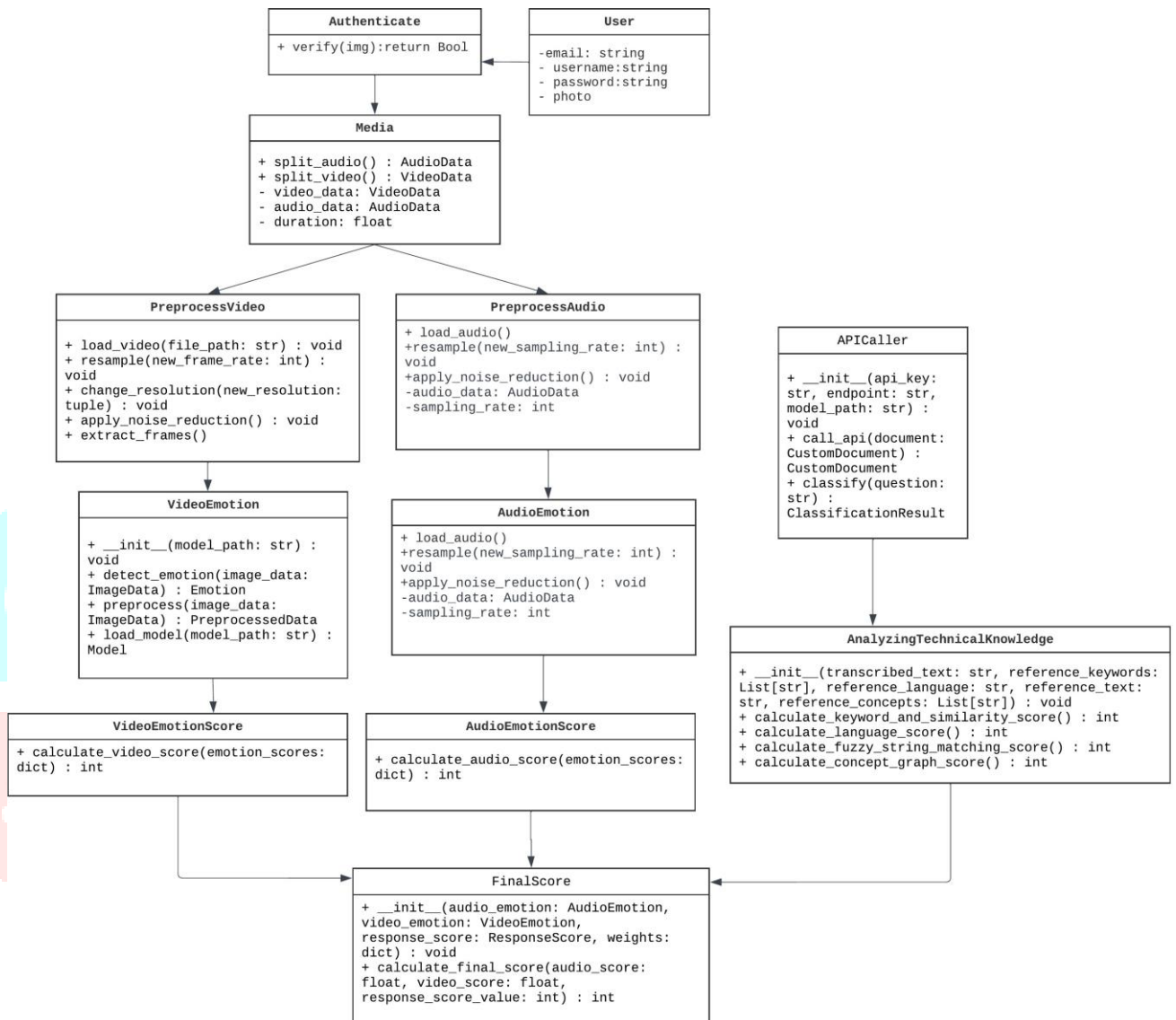


Figure: Class Diagram

The various functionalities for the assessment project are encapsulated in the classes that are indicated in the class diagram. Classes like "PreprocessVideo" and "PreprocessAudio" take care of particular data preprocessing tasks, while "Media" manages audio and video data. "AudioEmotion" and "VideoEmotion" handle extracting emotions from audio and video data, while "APICaller" manages API communication. Emotional scores are calculated by the "VideoEmotionScore" and "AudioEmotionScore" classes. "AnalyzingTechnicalKnowledge" evaluates

technical responses by processing them. Lastly, using weighted values for each component, the "FinalScore" class combines the multiple scores—audio, video, and response—to calculate the final assessment score.

Component Diagram

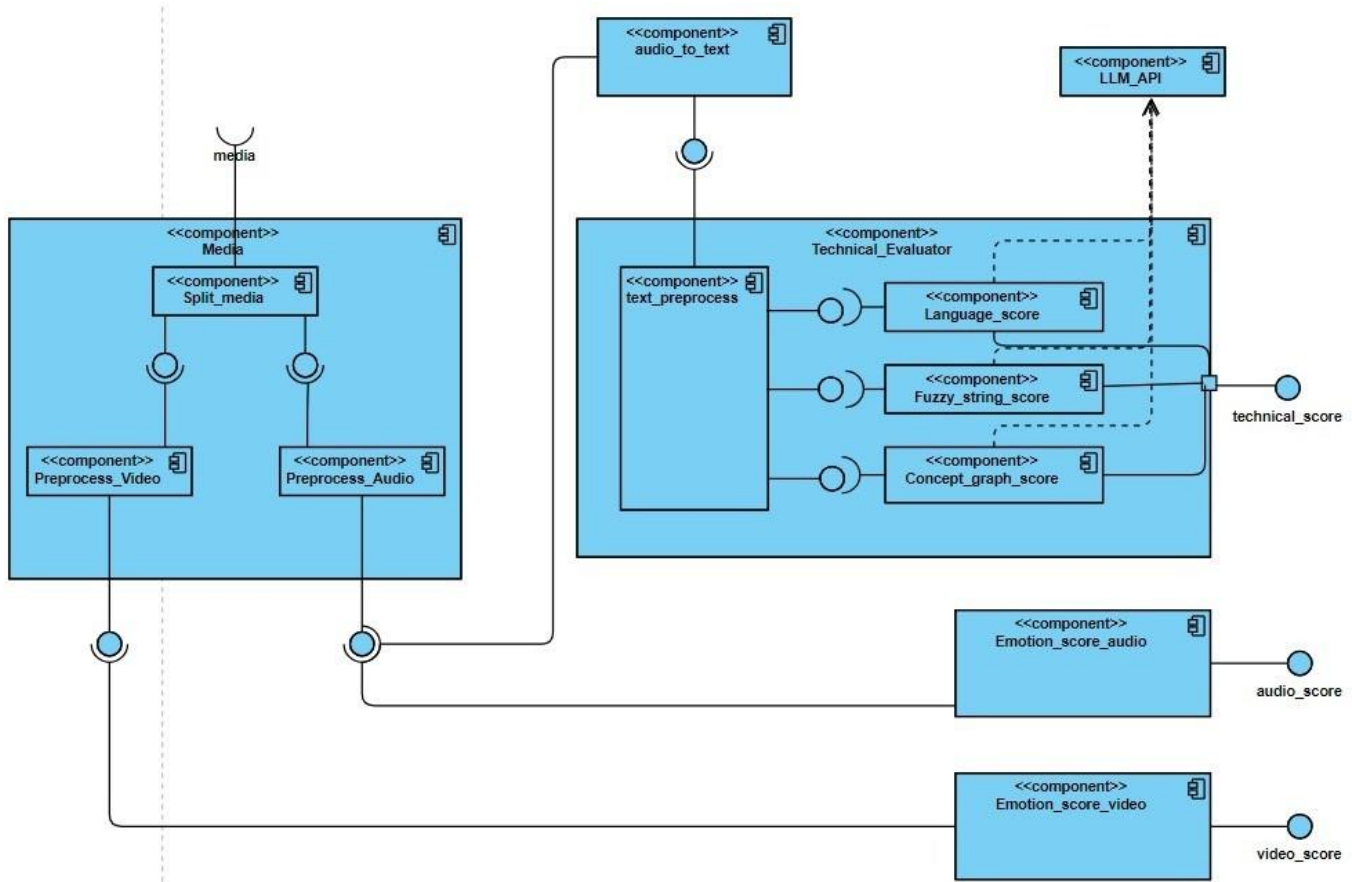


Figure : Component Diagram

In the diagram of components The diagram’s listed components each have a different pur- pose within the project. "Media" deals with various media formats. Video data is organized using "Split media" and "Preprocess Video". Audio data, managed by "Preprocess Audio". "text preprocess" and "audio to text" take care of text preparation and audio-to-text conver- sion, respectively. "Technical Evaluator" evaluates technical skill. The evaluation metrics "Language score," "Fuzzy string score," "LLM API," "Concept graph score" assess language, _strings, cognitive understanding, and API. The main topic of "Emotion score audio" and "Emo- tion score video". The terms "technical score," "audio score," "video score" aggregate differ- ent aspects of evaluation. By combining, these elements simplify the evaluation process for a thorough assessment by assessing linguistic, emotional, and technical competencies.

5 SUMMARY AND CONCLUSION

The summary presents the main findings of the research and describes the elements and goals of the AI-based mock interview evaluator. The system's potential to transform interview assessments is discussed in the Conclusion, which also addresses its drawbacks and opens the door to a more thorough and equitable evaluation procedure.

SUMMARY:

This project introduces an AI-based mock interview evaluator aiming to comprehensively assess candidates' abilities and support their skill development. It uses deep learning architectures, machine learning models, and facial expressions to assess candidates' knowledge-based answers, speech patterns, and emotional cues in real time. Furthermore, the system adjusts interview questions in response to candidate answers, allowing for a customized evaluation that is in line with particular job requirements. Important features include sentiment analysis for emotional assessment, facial recognition for identity verification, and behavioral and technical competency evaluation.

CONCLUSION:

The proposed AI-based mock interview evaluator represents a significant stride in modernizing the interview process. Through the integration of various cutting-edge technologies, it offers a thorough assessment method for determining candidate suitability and encouraging skill advancement. A comprehensive evaluation of candidates is made possible by the integration of facial, emotional, speech recognition, and knowledge processing modules. The adaptive question generation adjusts assessments to match job requirements based on candidate answers. By giving candidates thorough, individualized feedback, this system improves the interview process by enabling them to recognize their strengths and potential areas for development. This project proposal shows the potential to greatly help organizations find the most qualified candidates, even though it is not a research study. This creative method presents itself as a viable means of identifying qualified applicants as well as developing their abilities via feedback, which will ultimately strengthen the hiring procedure.

REFERENCES

- [1] Yang Li, Constantinos Papayiannis, Viktor Rozgic, Elizabeth Shriberg, Chao Wang Dept. of Electrical and Computer Engineering, University of Rochester. *"CONFIDENCE ESTIMATION FOR SPEECH EMOTION RECOGNITION BASED ON THE RELATIONSHIP BETWEEN EMOTION CATEGORIES AND PRIMITIVES."*, IEEE Publication, 2023.
- [2] Rubi Mandal, Pranav Lohar, Dhiraj Patil, Apurva Patil, Suvarna Wagh, *"AI-Based mock interview evaluator: An emotion and confidence classifier model"*, IEEE Publication, 2023.
- [3] Yi-Chi Chou, Felicia R. Wongso, Chun-Yen Chao, Han-Yen Yu , *"An AI Mock-interview Platform for Interview Performance Analysis"*, IEEE Publication, 2022.
- [4] B C Lee, B Y Kim, *"Development of an AI-based interview system for remote hiring"*, IJARET Publication, 2021.
- [5] Dulmini Yashodha Dissanayake, Venuri Amalya, Raveen Dissanayaka, Lahiru Lakshan, Pradeepa Samarasinghe, Madhuka Nadeeshani, Prasad Samarasinghe. *"AI-based Behavioural Analyser for Interviews/Viva"*, Springer, 2021.
- [6] Vinal Bagaria, Mohit Badve, Manasi Beldar, Prof. Sunil Ghane, *"An Intelligent System for Evaluation of Descriptive Answers"*, Springer, 2020.