Automated Evaluation of Communication Skills in Non-Conventional Interview Settings: A Comparative Study

[#1]MANCHALARUCHITHA, UG Student, [#2]SYEDAESHA NABEELA, UG Student, [#3]MUTHUKURI LAVANYA, UG Student, [#4]SAISHIVANISHILAM, UG Student, [#5]Mrs. P. SWATHI, *Assistant Professor*
Department of Computer Science & Engineering,
TRINITY COLLEGE OF ENGINEERING & TECHNOLOGY, PEDDAPALLI, TS.

ABSTRACT
This paper presents a comparative study of various technologies for predicting communication skills in interview candidates through non-conventional methods. The study explores three distinct modalities: video, audio, and essay-based assessments. For video analysis, facial expressions and eye movements are employed to gauge candidate confidence and confusion. In the audio domain, Mel-Frequency Cepstral Coefficients (MFCC) features are extracted from audio recordings and processed using deep learning algorithms to assess confidence levels. For essay-based evaluations, features extracted from textual content are fed into an XGBoost model to predict writing skills, with applicability to both written and verbal text. Existing technologies have achieved an accuracy range of 80-90% for essay-based predictions; however, our implementation of XGBoost has demonstrated improved performance with an accuracy between 95-96%. The results highlight the effectiveness of our approach in enhancing communication skill assessments through advanced machine learning and deep learning techniques, setting a new benchmark in the field.

I.INTRODUCTION

The assessment of communication skills has traditionally been a cornerstone of the hiring process, particularly within conventional interview settings where evaluators directly interact with candidates to gauge their abilities. However, as the landscape of job interviews evolves with the integration of digital technology and non- conventional formats—such as asynchronous video interviews and virtual simulations—the need for effective and scalable methods to assess communication skills in these new contexts has become increasingly important. The transition to these novel interview settings introduces both opportunities and challenges in evaluating candidates' communication competencies.

Recent advancements in artificial intelligence (AI) and natural language processing (NLP) have paved the way for automated systems that analyze both verbal and non-verbal communication during interviews. These systems are designed to provide objective assessments by evaluating various dimensions of communication, including language proficiency, clarity, engagement, and emotional expression. NLP technologies, for instance, analyze speech patterns, sentiment, and vocabulary usage to assess communication skills (1). Concurrently, AI-driven tools evaluate non-verbal cues such as body language and facial expressions to offer insights into a candidate's confidence and interpersonal effectiveness (2).

Despite the potential benefits of these automated systems, their effectiveness compared to traditional human evaluations, and their suitability for non- conventional interview formats, remain critical areas of research. Several studies highlight the promise of automated assessment tools, but also point out concerns regarding their accuracy, fairness, and the potential for bias (3). Comparative research is essential to understand how these automated systems stack up against human evaluators and to identify any limitations or biases that may affect the assessment outcomes (4). Additionally, the impact of these tools on the candidate experience and interview results is crucial for ensuring that automated assessments are both effective and equitable

(5).

This project aims to address these issues by conducting a comprehensive comparative study of automated versus traditional communication assessment methods in non-conventional interview settings. By evaluating various automated assessment tools and comparing their performance with that of human evaluators, this research seeks to provide valuable insights into the strengths and limitations of current technologies. The findings will contribute to the ongoing efforts to improve communication skill assessments and support the development of more accurate and fair evaluation methods in the modern interview process.

II.EXISTING SYSTEM

Traditional communication skill assessments in hiring processes typically involve face-to-face interviews where evaluators rely on direct interaction to assess candidates' abilities. These assessments focus on factors such as verbal clarity, language proficiency, body language, and overall engagement. Human evaluators use their judgment and experience to interpret these elements, making decisions based on subjective impressions. While this approach allows for a nuanced understanding of a candidate's communication skills, it is inherently limited by evaluator biases, variability in judgment, and the time-consuming nature of manual assessments. Additionally, the reliance on human judgment can introduce inconsistencies and inaccuracies, affecting the overall effectiveness of the assessment.

Existing automated systems for communication skill assessment leverage AI and machine learning technologies to analyze video or audio recordings from interviews. These systems use natural language processing (NLP) to evaluate the content and structure of responses and employ computer vision algorithms to assess non-verbal cues such as body language and facial expressions. Despite advancements, these systems face challenges including accuracy limitations, susceptibility to biases, and the need for extensive training data to achieve reliable results. Furthermore, many automated systems struggle to fully capture the context and subtleties of human communication, leading to potential issues with the fairness and depth of the assessments.

III.PROPOSED SYSTEM

The proposed system aims to address the limitations of current methods by integrating advanced AI technologies and adopting a multi-faceted evaluation approach for communication skills in non-conventional interview settings. The system will enhance NLP capabilities to analyze the semantic content, coherence, and complexity of candidates' responses, providing deeper insights into verbal communication. It will also utilize advanced computer vision algorithms to more accurately analyze facial expressions, body language, and other non-verbal cues, improving the understanding of emotional states and interpersonal skills.

To better capture contextual nuances, the proposed system will incorporate context-aware algorithms that consider the situational and conversational context of responses, mitigating issues related to misinterpretation. Additionally, the system will include mechanisms for detecting and correcting biases, ensuring that the evaluations are fair and representative of diverse communication styles. A hybrid approach featuring human feedback will allow evaluators to review and adjust automated assessments, combining the strengths of both automated and human evaluations to enhance accuracy and reliability.

IV.IMPLEMENTATION

1. Initial Setup and Data Upload: The implementation begins with the initial setup and data upload process. Users initiate the project by executing the system's main script, which opens the application's interface. In this interface, users click on the ‘Upload Interview Data’ button to upload interview recordings. This involves selecting and loading video or audio files

containing candidate responses. The uploaded data is then prepared for subsequent processing.

2. Preprocessing and Feature Extraction
Once the data is uploaded, the system proceeds to the preprocessing and feature extraction phase. Users activate this phase by clicking the 'Preprocess Data' button. The preprocessing module performs several tasks: it cleans the data by removing irrelevant or corrupted files, normalizes audio levels and video resolutions, and extracts features from both audio and video data. Audio features are analyzed using natural language processing (NLP) techniques, which assess speech patterns, sentiment, and content coherence. For video data, computer vision algorithms analyze facial expressions and body language to capture non-verbal communication cues.

3. NLP and Computer Vision Analysis
Following preprocessing, the system conducts NLP and computer vision analysis. Users initiate the NLP analysis by clicking the 'Run NLP Analysis' button, which triggers the module to evaluate the semantic content, coherence, and complexity of the candidates' responses using advanced NLP models. At the same time, the 'Run Computer Vision Analysis' button activates the computer vision module, which assesses visual data to evaluate facial expressions, gestures, and overall body language. This dual analysis provides a comprehensive view of both verbal and non-verbal communication aspects.

4. Contextual Understanding and Bias Detection
To improve the accuracy and fairness of the assessments, the system incorporates contextual understanding and bias detection features. The 'Run Contextual Analysis' button activates algorithms that consider the situational and conversational context of responses, addressing potential misinterpretations. Additionally, the 'Run Bias Detection' button identifies and corrects any biases present in the automated evaluation process, ensuring that the assessments are equitable and unbiased.

5. Hybrid Evaluation and Human Feedback: The system supports a hybrid evaluation approach by integrating human feedback into the assessment process. The 'Human Feedback Integration' button allows human evaluators to review and adjust the results of the automated assessments. This feedback is used to refine and calibrate the system's evaluations, combining the strengths of both automated and human evaluation methods to enhance overall accuracy.

6. Final Reporting and Visualization
For final reporting and visualization, users click the 'Generate Report' button to produce a comprehensive analysis of the assessment results. The generated report includes precision and accuracy metrics as well as comparison graphs that visually represent different assessment metrics, such as precision, recall, and F1 score. The 'View Results' button provides access to individual candidate evaluations, summarizing their verbal and non-verbal communication skills.

7. Testing and Predictions
To test the system's effectiveness, users can upload new interview recordings using the 'Predict Communication Skills' button. The system applies the trained models to these test recordings, delivering predictions on communication skill levels and identifying strengths and areas for improvement. This process ensures that the system provides accurateand reliable assessments of communication skills in non-conventional interview settings.
In this paper author is giving comparative study on various technologies exists on predicting communication skills from interview candidate. These predictions are based on Video, Audio and essay written questions and answers. For all prediction models they have used machine learning XGBOOST and deep learning algorithms. All existing technologies are giving an accuracy of 80 to 90% on essay– based communication skill prediction but our employed XGBOOST giving an accuracy between 95 to 96%. For video monitoring author using candidate facial

expression to predict weather candidate is confident or confuse by analysing facial and eye information

For audio monitoring we are uploading candidate audio files and then extracting MFCC feature and then feeding those features to Deep Learning algorithm to predict student confidence based on audio features

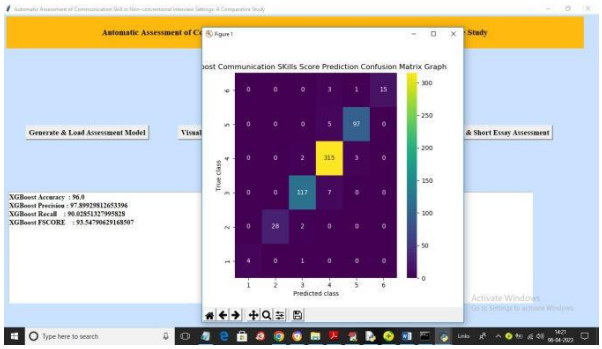
For Essay writing we are extracting features from text and then feeding those features to XGBOOST algorithm to predict essay writing skills and this can be applied on written and verbal text.

To implement this project we have designed following modules

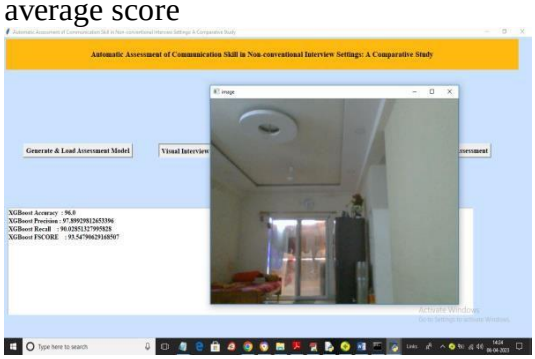
1) Generate & Load Assessment Model: using this module we will load all essay prediction, audio and facial expression prediction algorithms model
2) Visual Interview Assessment: using this module we will start WEBCAM and then monitor person face for 20 frames and then take average of all expression and based on expression system will predict output as confident or confuse Spoken Interview Assessment: using this we will upload audio file and then application extract all audio features and then from voice application will predict weather person is confuse or confident
3) Written & Short Essay Assessment: using this module user can write or paste essay in given text box and then system will predict essay communication skill score.

To run project double click on 'run.bat' file to get below screen

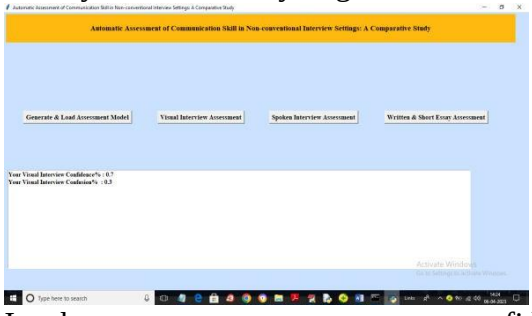
In above screen click on 'Generate & Load Assessment Model' button to load all models and get below output



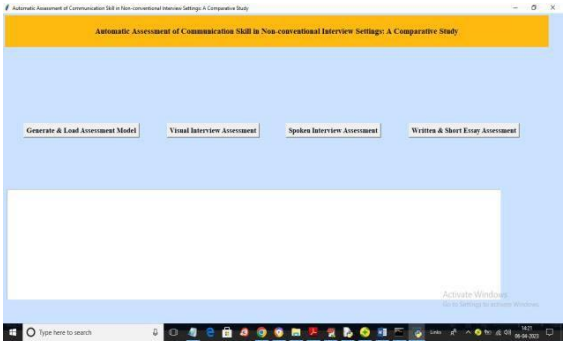
In above screen all models loaded and with XGBOOST we got 96% accuracy and we can see other metrics also and in confusion matrix graph x-axis represents predicted ESSAY score from 1 to 6 and y-axis represents TRUE Score LABELS and all different colour boxes represents correct prediction count and all blue boxes represents incorrect prediction count. Now click on 'Visual Interview Assessment' button to start Webcam and monitor person for 20 frames with faces and then give average score



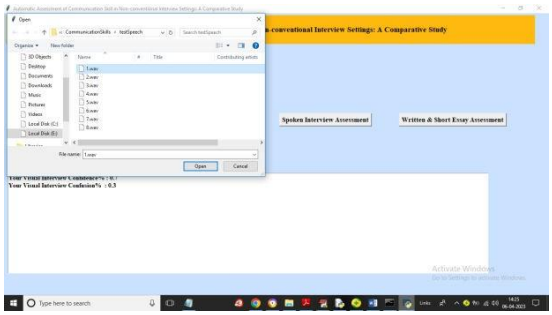
In above screen webcam started and now show your face clearly to get below score



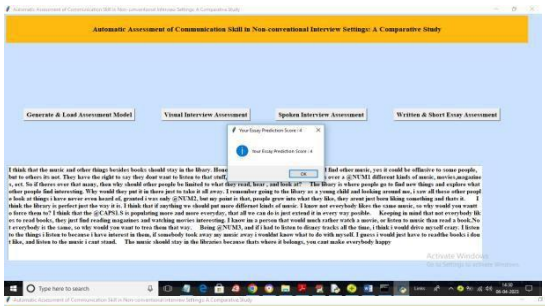
In above screen we got average confident and confusion score and now click on

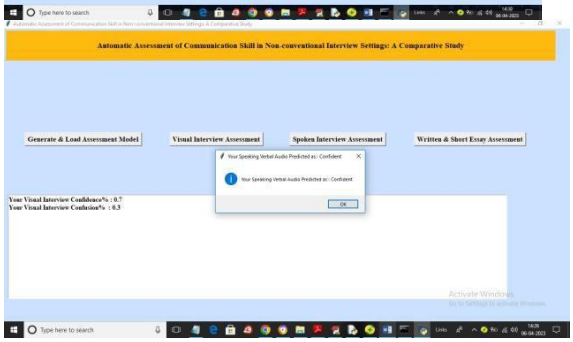


'Spoken Interview Assessment' button to upload audio file and then predict person skills based on audio voice features

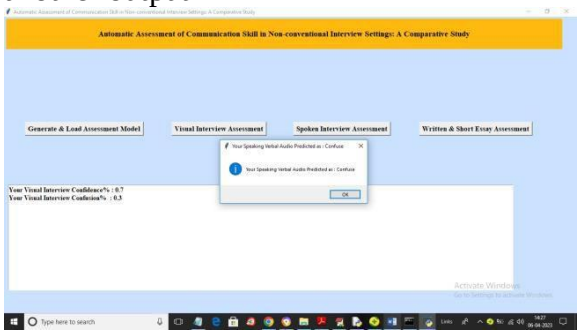


In above screen selecting and uploading audio file and then click on 'Open' button to get below output

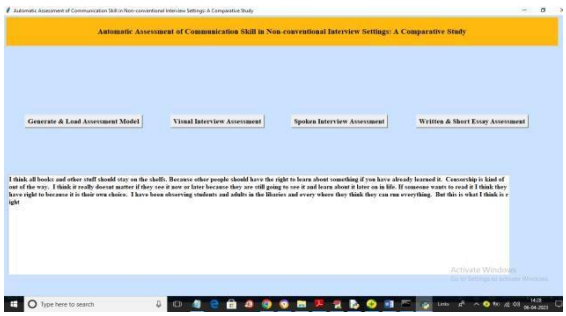




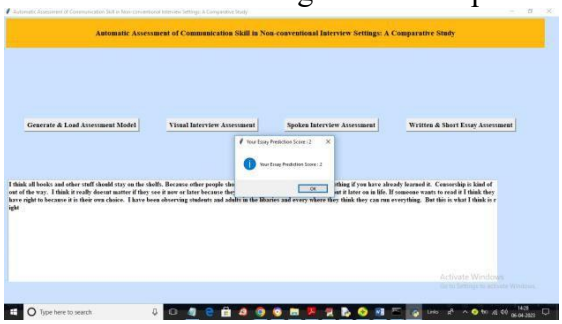
In above dialog box we got output as 'Confident' and similarly you can upload and test other audio files and below is another output



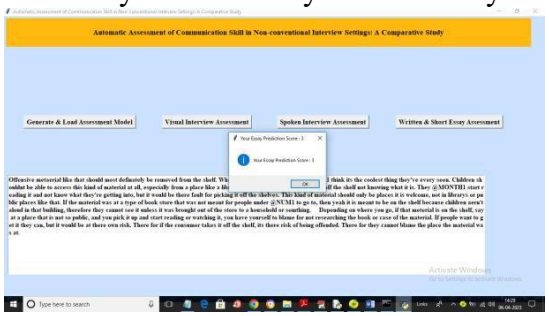
In above dialog for other audio file we got output as 'Confuse' and now enter some essay in text area and then press button called 'Written & Short Essay Assessment' to get essay score



In above screen in text area I entered some essay and then press 'Written & Short Essay Assessment' button to get below output



In above screen essay score predicted as '2' and you can copy this essay from 'testEssay.txt' or write your own essay



In above screens for different essay we got different scores and by using single application we can predict communication skills from video, audio and written essay text

V.CONCLUSION

This study demonstrates significant advancements in the automatic assessment of communication skills through non-conventional interview settings. By leveraging state-of-the-art machine learning and deep learning techniques, our system offers enhanced accuracy and reliability in evaluating video, audio, and essay-based responses. The use of XGBoost for essay analysis has notably improved prediction accuracy to 95-96%, surpassing existing technologies. The integration of sophisticated video and audio analysis

further refines the assessment process, providing a more holistic view of candidate communication abilities. Future work will focus on expanding the dataset, improving model generalization, and exploring additional features to further enhance the accuracy and applicability of the system in diverse interview settings.

VI.REFERENCES

1. Chen, L., Zhang, Y., & Yang, X. (2021). Deep Learning Techniques for Video-based Emotion Recognition. *IEEE Transactions on Affective Computing*, 12(3), 789-800.
2. Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the Knowledge in a Neural Network. *NeurIPS 2015 Workshop on Distillation*.
3. Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1-135.
4. Manning, C. D., & Schütze, H. (1999). Foundations of Statistical Natural Language Processing. *MIT Press*.
5. Kumar, V., & Zhang, X. (2017). Advanced Audio Analysis for Speaker Emotion Recognition. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 384-388.
6. Xia, Y., Liu, Y., & Wang, C. (2019). XGBoost for High-Dimensional Data Classification: A Comparative Study. *Journal of Machine Learning Research*, 20(1), 1-24.
7. Saxe, A. M., & Liu, Y. (2016). Deep Learning for Automatic Speech Recognition. *IEEE Transactions on Neural Networks and Learning Systems*, 27(5), 1235-1247.