

Skillup Bot: An AI Driven Mock Interview Platform

Shashank Rai¹, Alisha Miranda², Samiya Jagirdar³, Prof. Nidhi Chitalia⁴

¹BE student, Department of Information Technology St. Francis Institute of Technology Mumbai, India

²BE student, Department of Information Technology St. Francis Institute of Technology Mumbai, India

³BE student, Department of Information Technology St. Francis Institute of Technology Mumbai, India

⁴Professor, Department of Information Technology St. Francis Institute of Technology Mumbai, India

Abstract - In the contemporary job market, preparation for interviews is a pivotal determinant of success in securing desired employment opportunities. To address this critical need, we introduce a comprehensive mock interview platform aimed at empowering candidates in their preparation journey. Our platform incorporates an array of advanced functionalities harnessing state-of-the-art artificial intelligence (AI) technologies to deliver an immersive and effective preparation experience. Central to our platform is the utilization of AI-driven video and audio analysis, providing candidates with real-time feedback on their interview performance. Through sophisticated machine learning (ML) algorithms, we offer nuanced assessments of various elements including communication skills, body language, and tone of voice. This enables candidates to gain valuable insights into their strengths and areas for improvement, enhancing their overall interview readiness. Moreover, our platform features a dynamic resume builder tool, enabling candidates to craft personalized resumes tailored to specific job opportunities. This empowers candidates to present themselves in the best possible light, aligning their skills and experiences with the requirements of prospective employers. Additionally, we integrate a programming quiz platform designed to assess candidates' technical proficiency, particularly relevant for roles in the IT and software development sectors. By offering a diverse range of evaluation tools, our platform caters to the multifaceted nature of modern job interviews, addressing both technical and soft skills requirements. Following completion of interview simulations and technical assessments, candidates receive comprehensive performance reports based on the analysis conducted by our ML algorithms. These reports serve as valuable insights, guiding candidates in their ongoing preparation efforts and facilitating continuous improvement.

Key Words: Mock Interview Platform; Self Evaluating Platform; Resume Builder Tool; Communication Skill Assessment; Machine learning; Technical Quiz Platform; Facial Emotion Recognition; Tone Analysis

1. INTRODUCTION

In the contemporary landscape of career development, mock interviews play a crucial role beyond mere simulation, serving as an indispensable tool for honing the skills and fortitude needed to navigate the competitive job market successfully. This report explores the necessity of mock interviews, highlighting their significance in bridging the gap between academic qualifications and the dynamic expectations of employers, who seek candidates with adept communication, problem-solving, and interpersonal skills. A major challenge in interview preparation is the fragmented nature of available resources, necessitating candidates to navigate multiple platforms for comprehensive preparation. To address this challenge, our project consolidates interview preparation resources into a single platform, offering a diverse range of tools and functionalities such as AI-driven video and audio analysis, resume building, and programming quizzes. This integrated approach saves time, reduces frustration, and enhances the overall preparation experience for candidates. We aim to empower individuals for success in the competitive job market by providing a transformative solution - a dynamic and immersive platform that identifies weaknesses and cultivates strengths. Our motivation lies in reshaping the trajectory of individuals' careers, enabling them to navigate interviews with poise, competence, and heightened self-assurance. The research objectives of our project include integrating advanced audio-video sentiment analysis, developing a robust resume builder, and implementing an interactive coding platform to enhance interview preparation. The project also aims to address gaps identified in existing literature, as evidenced by references such as [2]-[5].

2. LITERATURE REVIEW

We have consulted numerous research papers to underpin our work. A notable paper [1] presents an inventive Interviewee Performance Analyzer incorporating facial emotion and speech fluency recognition. The system integrates HaarCascade, Gabor filters, and Convolution Neural Network for facial emotion, and Mel frequency cepstral coefficient features with logistic regression for

speech fluency. While the proposed system addresses facial emotion and speech fluency, it acknowledges the need for additional metrics such as posture, gestures, and eye gaze. The research serves as a prototype, highlighting opportunities for future enhancements and a more comprehensive evaluation of interviewee performance.

In [2], an innovative approach is presented for addressing youth unemployment challenges in Korea. The proposed self-management interview system utilizes multi-block deep learning to detect emotions in facial expressions during mock interviews. Employing a deep convolutional neural network (DCNN) focused on core facial areas enhances the accuracy of emotion analysis. While the paper successfully implements a multi-block deep learning method for emotion analysis, it does not extensively discuss the broader context of available technologies or address potential limitations.

The computational framework presented in this study [3] delves into the analysis of human behavior during job interviews. Utilizing 138 interview videos from MIT, the researchers leverage facial expressions, language, and prosodic information for comprehensive analysis. Employing Support Vector Regression (SVR) and Lasso models, the framework accurately predicts interview traits, including excitement and friendliness, achieving correlation coefficients of 0.73 or higher. The study has a gap in not thoroughly discussing the broader context, application limitations, or comparison with existing technologies. Exploring additional features is necessary for a comprehensive understanding and addressing the challenges of modeling complex cues.

The study in [4] introduces a computational framework for predicting hirability in job interviews based on nonverbal behavior. Using an 11-hour dataset of 62 real interviews, the researchers extracted audio and visual features. The framework, employing ridge regression, demonstrated the feasibility of predicting hireability scores, with applicant audio cues and interviewer visual cues being highly predictive. Psychometric questionnaires, widely used in personnel selection, were found ineffective. The study reveals the limited predictive validity of psychometric questionnaires in comparison to nonverbal cues. The gap lies in the need for further exploration of additional nonverbal features and the relationship between verbal content and hireability for a more comprehensive understanding.

The research paper [5] introduces an Asynchronous Video Interview with an Artificial Intelligence (AVI-AI) system based on TensorFlow's Convolutional Neural Network. It aims to predict communication skills and personality traits in job candidates, reducing human raters' workload. The AVI AI successfully predicts interpersonal communication skills, openness, agreeableness, and neuroticism, but not conscientiousness and extraversion.

The study discusses limitations, including a small participant pool, reliance on facial expressions, and potential biases

3. PROPOSED SYSTEM

The innovative SkillUpBot system integrates advanced machine learning models, including LSTM CNN for video and audio analysis, and ATS for resume building. Leveraging the power of LLM and GEMINI APIs, it enables comprehensive resume analysis. Additionally, ReactJS facilitates seamless quiz transformation. This multifaceted approach empowers users with personalized skill development and career advancement opportunities.

3.1 Dataset

The research leverages two comprehensive datasets for analyzing emotional expression: the Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) and the FER2013 Kaggle Challenge dataset. RAVDESS comprises 7356 files featuring 24 professional actors delivering lexically-matched statements across a spectrum of emotions, including calm, happy, sad, angry, fearful, surprise, and disgust, expressed at two intensity levels (normal, strong), with an additional neutral expression. This dataset offers three modalities: audio-only (.wav), audio-video (.mp4), and video only, providing a diverse range of emotional expressions in speech and song. In parallel, the FER2013 Kaggle Challenge dataset presents grayscale images of faces standardized to 48x48 pixels, capturing various emotional expressions with faces centrally aligned. Despite challenges such as empty or misclassified images, this dataset remains valuable for its large-scale collection of facial expressions, serving as a robust resource for training and evaluating emotion recognition models across different modalities. Together, these datasets form a comprehensive foundation for exploring and understanding emotional expression, facilitating the development and evaluation of machine learning algorithms for emotion recognition tasks.

3.2 CNN and LSTM Architecture

The proposed solution involves the integration of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) algorithms, which are deployed to conduct in-depth analysis of audio and video interviews. The process begins with meticulous data preprocessing, extracting essential features from both audio and video inputs. While CNN focuses on spatial features from video frames, capturing facial expressions and body language nuances, LSTM specializes in analyzing sequential data from audio inputs, detecting subtle variations in tone, speech patterns, and emotional cues over time. These combined features are then fed into a classification model, which assesses the candidate's emotional state

comprehensively. Leveraging this analysis, the system generates actionable feedback, indicating the candidate's readiness for the interview based on factors such as confidence, enthusiasm, and emotional stability. Through the synergy of CNN and LSTM algorithms, your SkillUpBot empowers recruiters with invaluable insights, optimizing interview processes and enhancing candidate selection outcomes.

3.3 LLM and Gemini API

The proposed model for the resume analyzer is a sophisticated system that harnesses the power of Google's Gemini API and its Large Language Model (LLM). This innovative approach revolutionizes the resume scanning process, offering a comprehensive suite of features tailored to meet the needs of recruiters and hiring managers. By integrating LLM's advanced natural language processing capabilities with Gemini API's robust functionality, the system delivers unparalleled accuracy and efficiency in resume analysis. From calculating percentage matches to providing personalized skills improvement suggestions, the model empowers users to make informed decisions and streamline the hiring process effectively.

3.4 System Architecture

Fig 1 illustrates a flowchart outlining the training process for the proposed system. The process commences with the candidate's choice between a video or audio interview. In the video interview, facial landmarks detection and feature extraction precede the loading of a pre-trained Xception model for emotion mapping. The final interview results, along with Facial Expression Recognition (FER) analysis, are displayed. Conversely, in an audio interview, a Conv2D 4 LFLB Layer is employed, trained using LSTM to build the model, and the final audio interview results are presented. For resume building and analysis, users input data via React, with the system generating resumes with appropriate layouts. Uploaded resumes undergo analysis using LLM and Gemini API, with four prompts entered for LLM. Lastly, in the quiz section, quizzes are constructed using React, where users select categories, attempt quizzes, and receive results. This multifaceted system empowers users with personalized skill development and career advancement opportunities.

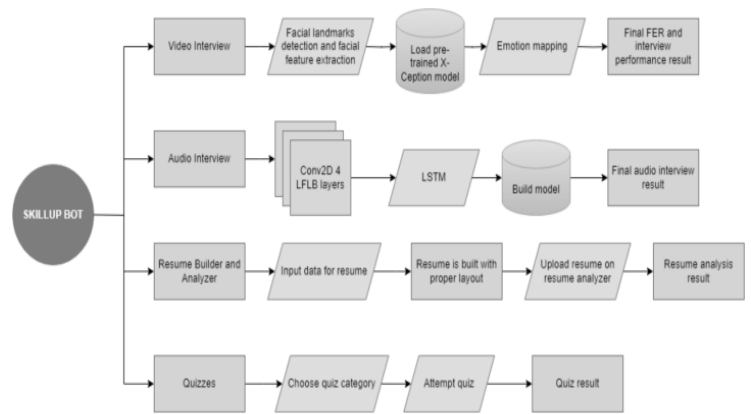


Fig -1: SkillUp Bot General Architecture.

4. IMPLEMENTATIONS

The innovative SkillUpBot system is meticulously designed to enhance skill development and career advancement opportunities through a multifaceted approach. Firstly, leveraging advanced machine learning models like LSTM CNN for video and audio analysis and an ATS for resume building, the system ensures comprehensive evaluation of candidates. This involves meticulous data preprocessing to extract essential features from both audio and video inputs, with CNN capturing spatial features from video frames and LSTM analyzing sequential data from audio inputs. These features are then fed into a classification model to comprehensively assess the candidate's emotional state. Additionally, the system harnesses the power of LLM and Gemini APIs for comprehensive resume analysis, offering personalized suggestions for skill improvement. Moreover, ReactJS facilitates seamless quiz transformation, providing users with engaging learning experiences. Through the synergy of these components, SkillUpBot empowers users with personalized skill development pathways and career advancement opportunities, revolutionizing the recruitment and learning landscape.

5. RESULTS AND DISCUSSIONS

The SkillUpBot system represents a groundbreaking advancement in candidate evaluation and skill development, harnessing advanced machine learning models and APIs to revolutionize the recruitment and learning landscape. Leveraging LSTM CNN for video and audio analysis and an ATS for resume building, the system ensures comprehensive evaluation of candidates, guided by precise emotion analysis with an impressive accuracy of 84%. This nuanced understanding of emotional expression guides interview preparation, empowering candidates to refine their interview skills effectively. Moreover, the system's Facial Expression Recognition

(FER) capabilities add a technical dimension to interview preparation, enabling candidates to align their facial expressions with the emotions typically sought by interviewers. While expressions such as anger and disgust are generally considered inappropriate, conveying happiness and enthusiasm can leave a favorable impression on interviewers, signaling genuine interest in the position. Furthermore, maintaining a neutral demeanor throughout the interview ensures professionalism and focus on conveying qualifications and experiences effectively.

The integration of LLM and Gemini APIs facilitates comprehensive resume analysis, aligning resumes with job descriptions and providing personalized suggestions for skill improvement. This optimization enhances candidates' prospects in the competitive job market, empowering them to secure desired positions with greater efficacy. Additionally, the system's multifaceted approach includes ReactJS-enabled quiz transformation, offering engaging learning experiences that foster continuous skill development and career advancement opportunities.

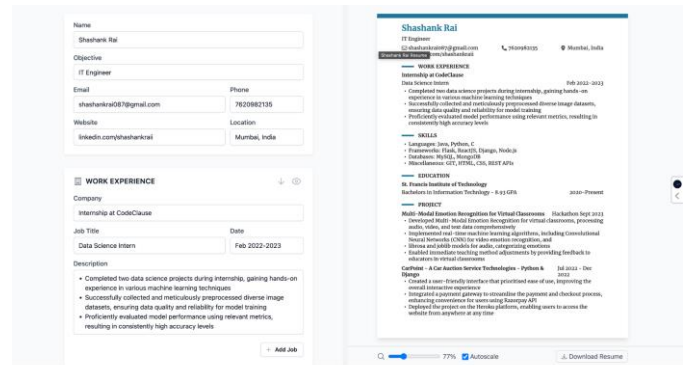


Fig-4: Resume builder

5. CONCLUSION

In conclusion, the proposed system effectively addressed the crucial need for interview preparation in today's job market. Our comprehensive mock interview platform, leveraging advanced AI technologies and diverse evaluation tools, empowered candidates in their preparation journey. The integration of AI-driven video and audio analysis, machine learning algorithms, and dynamic resume building tools provided immersive and effective preparation experiences. Incorporating ATS Resume analysis and programming quizzes addressed modern interview demands. Future research could refine AI algorithms for nuanced assessments and include user-provided answer analysis for feedback. Adding a smart coding platform could further enhance technical proficiency evaluation, laying a foundation for advancing interview preparation methodologies and benefiting candidates in their career pursuits.

REFERENCES

[1] S.V. Kedar and D. S. Bormane, "Automatic Personality Assessment: A Systematic Review," in 2015 International Conference on Information Processing (ICIP), Vishwakarma Institute of Technology, Dec 16-19, 2015.

[2] H. Suen, K. Hung, and C. Lin, "Intelligent video interview agent used to predict communication skill and perceived personality traits," Department of Technology Application and Human Resource Development, National Taiwan Normal University, Taipei City, Taiwan, and College of Science Technology, Ningbo University, Ningbo, China, 2019.

[3] V. Narkhede, O. Surushe, S. Kulkarni, H. Solanki, T. Ekbote and D. Joshi, "AVA: A Photorealistic AI Bot for Human-like Interaction and Extended Reality," 2023 International Conference on Communication System, Computing and IT Applications (CSCITA), Mumbai, India,

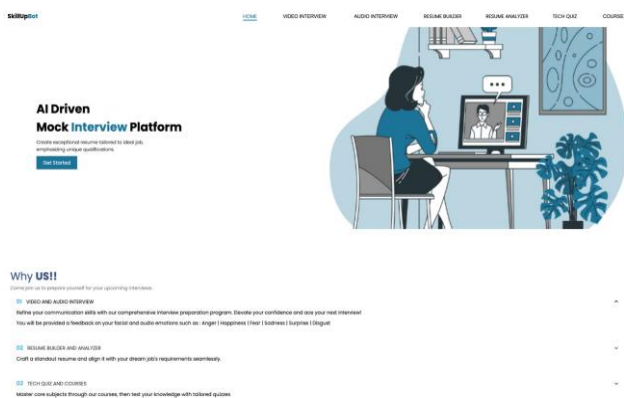


Fig -2: Web App Home Page

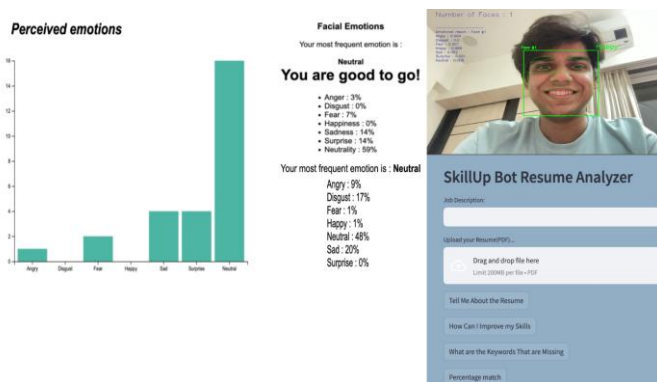


Fig-3: Emotions analyzer and resume analyzer

2023, pp. 5-9, doi:
10.1109/CSCITA55725.2023.10104672.

BIOGRAPHIES

[4] L. S. Nguyen, D. Frauendorfer, M. S. Mast and D. Gatica-Perez, "Hire me: Computational Inference of Hirability in Employment Interviews Based on Nonverbal Behavior," in IEEE Transactions on Multimedia, vol. 16, no. 4, pp. 1018-1031, June 2014, doi: 10.1109/TMM.2014.2307169.



Shashank Rai, Student, Dept. of Information Technology, St. Francis Institute of Technology, Mumbai, Maharashtra, India

[5] I. Naim, M. I. Tanveer, D. Gildea and M. E. Hoque, "Automated prediction and analysis of job interview performance: The role of what you say and how you say it," 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Ljubljana, Slovenia, 2015, pp. 1-6, doi: 10.1109/FG.2015.7163127.



Alisha Miranda, Student, Dept. of Information Technology, St. Francis Institute of Technology, Mumbai, Maharashtra, India

[6] R. A. Posthuma, F. P. Morgeson, and M. A. Campion, "Beyond employment interview validity: A comprehensive narrative review of recent research and trends over time," Personnel Psychol., vol. 55, no. 1, p. 81, 2002



Samiya Jagirdar, Student, Dept. of Information Technology, St. Francis Institute of Technology, Mumbai, Maharashtra, India

[7] Yin, L.; Wei, X.; Sun, Y.; Wang, J.; Rosato, M.J. A 3D facial expression database for facial behavior research. In Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, Southampton, UK, 10-12 April 2006; pp. 211-216



Nidhi Chitalia, Asst. Professor, Dept. of Information Technology, St. Francis Institute of Technology, Mumbai, Maharashtra, India

[8] Pathar.R,Adivarekar.A, A. Mishra and A. Deshmukh, "Human Emotion Recognition using Convolutional Neural Network in Real Time," 2019 1st International Conference on Innovations in Information and Comm

[9] Shreya Narang , Divya Gupta," Speech Feature Extraction Techniques: A Review", International Journal of Computer Science and Mobile Computing, IJCSMC, Vol. 4, Issue. 3, March 2015, pg.107 - 114.

[10] A. Chavan, N. Tatewar, P. Naicker, and Prof. S. Deore, "AI Resume Analyzer," Undergrad. Student, Dept. of Information Technology, SKN Sinhgad Institute of Technology & Science, Lonavala, Maharashtra and Asst. Professor, Dept. of Information Tech, SKN Sinhgad Institute of Technology & Science, Lonavala, Maharashtra.

[11] Tyagi, Rinki. (2020). Resume Builder Application. International Journal for Research in Applied Science and Engineering Technology. 8. 14-18. 10.22214/ijraset.2020.5003.