A PROJECT REPORT ON

SKILLSIGHT: AN AI-POWERED PLATFORM FOR SOFT SKILLS INTERVIEW & ASSESSMENT

SUBMITTED TO THE SAVITRIBAI PHULE PUNE UNIVERSITY,
PUNE IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE

OF

BACHELOR OF TECHNOLOGY (Computer Science And Business Systems)

SUBMITTED BY

ADITYA ANIL TIGHARE	Exam No :**RBT21CB049**
SANKET RAM RUPNAVAR	Exam No :**RBT21CB065**
PROSHANJEET DEEPAK CHANDA	Exam No :**RBT21CB006**
VEDANT VIJAY METE	Exam No:**RBTL22CB074**

Under the guidance Of
Prof. K.M. Patil

DEPARTMENT OF COMPUTER SCIENCE AND BUSINESS SYSTEMS

JSPM's RAJARSHI SHAHU COLLEGE OF ENGINEERING

TATHAWADE, PUNE 411033

(An Autonomous Institute Affiliated to Savitribai Phule Pune University,Pune)

2024-2025

ACKNOWLEDGEMENT

It is with a deep sense of gratitude that we would like to express our warm regards and sincere thanks to Prof. Kavita Patil, for her invaluable guidance, inspiration and whole hearted involvement during every stage of this project. Her experience, perception and thorough professional knowledge, has served as an indispensable source of guidance and hope. Her being available beyond the stipulated period of time for all kind of guidance and supervision and ever- willing attitude to help, has undoubtedly been the sole reason for the timely and successful completion of this project.

We would like to extend my sincere thanks to Dr. K. P. Moholkar Head of Department, Department of Computer Science And Business Systems, for her continual support and invaluable assistance rendered towards the presentation of this work. She was always there for suggestions and help in order to achieve this goal.

Finally, we are indebted to Dr. S. P. Bhosle, Director, JSPM's Rajarshi Shahu College of Engineering, Tathawade, Pune, for his encouragement and providing us with the opportunity to translate our thoughts into actions and for providing the institutional facilities to carry this work.

NAME OF THE STUDENTS

Mr. Aditya Anil Tighare	RBT21CB049
Mr. Sanket Ram Rupnavar	RBT21CB065
Mr. Proshanjeet Deepak Chanda	RBT21CB006
Mr. Vedant Vijay Mete	RBTL22CB074

ABSTRACT

SkillSight is an AI-driven mock interview and assessment platform focused on enhancing users' soft skills through a dynamic and immersive experience. With the increasing trend of companies conducting interviews virtually, it has become essential for job seekers to practice and adapt to online interview formats. SkillSight addresses this need by providing a realistic, simulation-based environment where users can practice interviews based on selected topics or job roles.

The platform intelligently generates a wide range of interview questions using advanced AI algorithms, ensuring diversity and relevance according to the user's chosen field. To offer a comprehensive evaluation, SkillSight integrates cutting-edge technologies like facial expression recognition and natural language processing (NLP). These technologies enable the system to assess crucial soft skill parameters such as communication effectiveness, emotional intelligence, reaction time, and nonverbal behavior including body language, eye contact, and facial cues.

After each interview session, users receive detailed, personalized feedback that highlights their strengths and identifies specific areas for improvement. The platform also provides actionable suggestions and recommends adaptive learning paths that users can follow to continuously enhance their skills over time. By focusing on both verbal and nonverbal aspects of communication, SkillSight ensures that users are better equipped to handle real-world interviews with confidence.
Moreover, the system fosters continuous learning by tracking user progress and dynamically adjusting the difficulty and focus areas of subsequent interview sessions. Through repeated practice and focused guidance, SkillSight empowers users to become more self-aware, responsive, and effective communicators. Ultimately, SkillSight aims to bridge the gap between technical competency and essential soft skills, helping users achieve their career goals and succeed in increasingly competitive professional environments.

Keywords: AI-based interview systems, Natural Language Processing, Facial emotions recognition, Soft skills, Adaptive learning paths.

TABLE OF CONTENTS

LIST OF ABBREVATIONS

ABBREVIATION	ILLUSTRATION
AI	Artificial Intelligence
DL	Deep Learning
CNN	Convolutional Neural Networks
KNN	K-Nearest Neighbors
SVM	Support Vector Machine
HOG	Histogram of Oriented Gradients
ML	Machine Learning
PyTorch	Python Torch (Deep Learning Library)
API	Application Programming Interface
NLP	Natural Language Processing
GPU	Graphics Processing Unit
TF	TensorFlow
BERT	Bidirectional Encoder Representations from Transformers

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 OVERVIEW

Skill Sight: Interview & Assessment Platform for Soft Skills we strive to be the bridge between potential and performance. AI job interview is an AI-powered mock job interview website that provides users with a dynamic and immersive experience with simulated job interviews in real-world scenarios.

Its main aim is closing the gap between theory and practice with the help of this platform, where a user will practice the most needed soft skills. Hence in job market full of competition, the essence of soft skills for success in professional life. SkillSight acts as a guide for anyone who wants to not only satisfy but exceed the expectations of potential employers by utilizing state of the art technology like facial recognition and natural language processing to measure a candidate's communication skills, emotional intelligence, and nonverbal behavior. This tool also prominent as it can analyze the response to varying interview questions and provide users with tips for improvement and tailored feedback. Such a practice not only improves their interview performance but facilitates continuous professional development.

With its ease of interface, personalized learning paths, and degree of privacy and security. Skillsight can help to boost people's employability and career prospects by changing the way people prepare for, tackle and overcome the challenges of giving job interviews.

1.2 MOTIVATION

Soft skills such as communication, problem-solving, and adaptability are critical for job success but are often less developed than technical abilities. A study by Harvard University, the Carnegie Foundation, and the Stanford Research Centre found that 75% of long-term job success is attributed to soft skills rather than technical knowledge. However, candidates often lack real-world opportunities to practice these skills, leading to underperformance in interviews. Virtual mock interviews offer a solution, helping job seekers build confidence, reduce anxiety, and strengthen their performance under pressure by blending hard and soft skills. For employers, traditional hiring processes are time-consuming and costly, often requiring 4 to 5 interview rounds to evaluate

candidates thoroughly. A study by SHRM highlights that a significant portion of hiring costs comes from these extended processes. By integrating a system to pre-assess soft skills, employers can streamline hiring, reduce turnover, and make better long-term hiring decisions, as poor soft skills often cause hiring failures.

1.3 PROBLEM DEFINITION

In the modern job market, soft skills such as communication, emotional intelligence, and adaptability have become as critical as technical expertise for career success, yet there remains a significant gap between the skills developed by candidates and those expected by employers. Traditional interview preparation methods fail to offer candidates practical, real-time feedback on these vital skills, resulting in underprepared job seekers and inefficient hiring processes. To address these challenges, the SkillSight platform proposes an AI-driven mock interview system that uses technologies like facial expression recognition, natural language processing, and behavioral analysis to simulate real-world interview scenarios. It aims to assess both verbal and nonverbal communication, provide personalized feedback, and create adaptive learning paths for continuous improvement. By bridging the gap between theoretical knowledge and practical application, SkillSight enhances employability, reduces interview anxiety, and enables employers to streamline recruitment by better evaluating candidates' soft skills early in the hiring process.

Hence, addressing this gap is crucial to improving candidate readiness and hiring efficiency in today's competitive job landscape.

1.4 OBJECTIVES

The primary objective of the SkillSight platform is to create a comprehensive AI-powered interview preparation system that enables users to enhance crucial soft skills, including communication, emotional intelligence, and nonverbal behavior. By simulating real-world interview environments, the platform seeks to bridge the gap between theoretical knowledge and practical application, offering candidates a dynamic space to practice and refine their interpersonal skills. It aims to leverage technologies such as facial expression recognition, natural language processing, and behavioral analysis to objectively assess user responses and provide detailed,

personalized feedback. Another key objective is to build adaptive learning paths tailored to individual user needs, helping them progressively strengthen weaker areas.

For employers, the system aspires to streamline recruitment processes by pre-evaluating candidates' soft skills, thus reducing the number of interview rounds and making better long-term hiring decisions. Additionally, SkillSight focuses on ensuring user privacy, offering a seamless, secure, and user-friendly experience that supports continuous learning and professional growth in an increasingly competitive job market.

1.5 SCOPE AND LIMITATION

- **Scope**

  SkillSight is designed to revolutionize the way job seekers prepare for interviews by offering a dynamic, AI-driven platform focused on the development of soft skills. The system simulates real-world interviews using technologies such as Natural Language Processing (NLP), facial expression recognition, and behavioral analysis. Users can practice interviews tailored to various domains, receive objective evaluations, and obtain personalized feedback and learning paths. The platform is especially valuable for students, fresh graduates, and professionals aiming to improve communication, emotional intelligence, and overall interview performance. Employers can also benefit from pre-assessed candidates, potentially reducing the number of interview rounds and making more informed hiring decisions. By bridging the gap between academic training and industry expectations, SkillSight aims to enhance employability and prepare users for the modern, virtual job market.

- **Limitation**

  Despite its innovative approach, the platform has certain limitations. AI models may misinterpret nuanced human expressions or cultural communication styles, leading to inaccurate evaluations. Additionally, while the system aims for fairness, there is potential for algorithmic bias in NLP and computer vision modules. Privacy concerns may arise from the handling of sensitive video and audio data, even with security measures in place. Furthermore, stable internet connectivity is essential, as real-time analysis can be disrupted by network issues.

Lastly, while SkillSight effectively simulates interviews by analyzing verbal and non-verbal cues in real time, it cannot fully replicate the unpredictability and psychological pressure of real-life human interactions. In an actual interview, candidates face dynamic conversations, unexpected follow-up questions, and real-time emotional responses from interviewers — factors that significantly influence performance. These situational elements, such as anxiety under pressure or adapting to interpersonal cues, are difficult to reproduce in a virtual environment, which can make the simulation feel less authentic or comprehensive despite its technical capabilities.

2 LITERATURE SURVEY

2.1 INTRODUCTION

The integration of Artificial Intelligence (AI) into recruitment processes has transformed how candidates are evaluated, especially in terms of soft skills. Traditional interview methods often struggle to objectively assess communication, emotional intelligence, and nonverbal cues, which are critical in today's competitive job market. To address this gap, AI-powered mock interview systems have emerged, offering scalable, consistent, and personalized assessments. Literature in this domain highlights the use of technologies such as Natural Language Processing (NLP), facial expression recognition, and sentiment analysis to evaluate candidate performance. Various platforms and studies have explored these innovations, yet many still lack holistic assessments that combine both verbal and nonverbal analysis. This chapter reviews significant research efforts and existing systems to understand their approaches, strengths, and shortcomings. The insights gained provide a foundation for the development of SkillSight—an advanced mock interview platform designed to overcome these limitations and offer a comprehensive soft skills evaluation experience.

2.2 HISTORICAL BACKGROUND

The evolution of interview systems began with traditional face-to-face interviews, where human judgment played a central role in evaluating candidates. However, this method was often prone to biases, lacked consistency, and was time-intensive. With the rise of digital transformation in the early 2000s, video conferencing tools became popular, allowing interviews to be conducted remotely. Despite this advancement, assessing soft skills remained subjective and limited. The integration of Artificial Intelligence marked a major turning point. Early AI-based systems primarily focused on keyword matching in resumes and automated question generation. Over time, advances in Natural Language Processing (NLP), computer vision, and machine learning enabled systems to analyze speech patterns, facial expressions, and body language. These innovations laid the groundwork for modern AI-driven interview platforms. As research evolved, focus shifted from technical assessments to holistic soft skills evaluation, setting the stage for solutions like SkillSight that provide real-time, multi-dimensional feedback to enhance interview readiness.

2.3 REVIEW OF LITERATURE SURVEY

2.3.1 Existing Solution

- **A Review of Artificial Intelligence Interview Systems: Techniques, Benefits, and Ethical Considerations, Journal of AI and Society.**

  In paper [1], article intends to give a detailed description of the definition of artificial intelligence interview systems. The brief outline of what the AI interview involves is explored to bring forth the principle factors and driving forces of the system. Various AI techniques such as natural language processing, machine learning, and computer vision utilized in the interview systems explained their roles in analyzing candidate responses, assessing non-verbal cues, and evaluating job fit. The mean benefits of the AI interview systems are also discussed, which as the aspects: increased efficiency, reduced-bias, objectivity, and scalability achieved. In addition, potential challenges and ethical implications of deploying AI based systems are considered as well. They include issues of data privacy, fairness of algorithms, and minimizing the human factor. Overall, the review article offers a general understanding of the topic and can be clearly positioned within the other works existing on this subject.

- **Development of a Deep Learning-Based Mock Interview System for Virtual Job Preparation, Proceedings of the International Conference on Smart Systems and AI Applications.**

  In paper [2], authors elaborate on the mock interview system being developed to allow the users to simulate interviews, specifically tailored towards the virtual interviews held by the companies. Using a deep learning algorithm, it offers feedback on everything from facial expressions and reaction time to speaking rates and grammatical execution. Face reading and grammar parsing are among the combination of multiple technologies applied by system. The system uses OpenCV Facial expression analysis which records user facial expressions while the interview is in progress and analyze it and store it in the specified database using a datasetimported in the system. This analysis may assist users who are nervous or anxious during an interview since interviewers look closely at

candidates' expressions. The facial expression recognition component of the system relies on the concept of deep learning methods.

- **Soft Skills and Employability: Employer Perspectives in a Competitive Job Market, International Journal of Human Resource Studies.**

  Perspective of Employers According to Al Asefer & Zainal Abidin [3], the soft skills are considered as the significant predictor of employability. In the current job market, employers deem soft skills such as communication, problem-solving, leadership, teamwork, and critical thinking as crucial for a graduate's success. Employers frequently stress these interpersonal skills as being equally if not more important than hard skills in terms of how well a candidate will perform on the job. But there is a wide-mismatch between the soft skills offered by higher education institutions and desired skills by employers, resulting in a skills-skill gap and hence the challenges of employability. However, many graduates remain unemployable, although academically qualified, because they lack these very high-end skills. This also highlights the need for universities to put more emphasis on soft skills training in addition to academics concurring with better preparation for students to step into the global workplace. The paper also points to employer discontent, as most of the new recruits were considered unemployable due to an absence of core soft skills, which speak to the increasing demand for such competencies in the competitive arena.

- **Generic Skills Component Approach: Rethinking the Soft-Hard Skill Divide, European Journal of Educational Psychology.**

  In their paper Jeremy Lamri and Todd Lubart [4] jointly explore the Generic Skills Component Approach, where they propose an integrated model that unifies the hard and soft skills. It challenges the historic divide between these skills, arguing both kinds include a mixture of similar underlying elements, including both knowledge and the active components of cognition, conation, affection and sense-perceptual motor ability. Unpacking skills into skills trees showcases not only the interdependencies of skills, but the context in which the skill is used which is crucial to understanding how someone is

able to checkbox a cert and not be prepared for real-world applications. The model could be used widely through various domains including education, workplace training and productivity, enabling deep understanding of the skill development process. It also applies Hilgard's Trilogy of Mind to discuss how cognitive, motivational and affective elements interweave in skill acquisition, ultimately questioning the traditional splitting of soft and hard skills. It emphasizes both types of skills are important in the workplace and that soft skills (such as empathy, emotional intelligence, and interpersonal skills) are at least as critical as hard skills (technical competencies) for many professions.

- **Facial Emotion Recognition Using Deep Learning: A Comparative Study on CNN Architectures, Journal of Computer Vision and Pattern Recognition.**

  In paper [5], the authors highlight the importance of emotions in effective communication and relationship building, emphasizing on technologies that can detect a person's emotions. The study uses the Fer2013 dataset and employs the Tensorflow framework, Keras library, and the Xception Architecture of CNN to train the model for facial emotion detection. The authors also discuss related work in the field, showcasing various approaches utilized for emotion detection, including Bayesian networks, speech segments, and multimodal affect analysis. The document presents the experimental evaluation of different models, including the Convolution Model, Inception Model, and Xception Model, for training, validation, and testing using the Fer2013 Dataset. The Xception Model outperformed the other models, achieving a training accuracy of 93.2% and a validation accuracy of 64.4%, making it the preferred model for real-time emotion detection.

- **Investigating Algorithmic Bias in NLP for Automated Hiring Systems, ACM Transactions on Fairness, Accountability, and Transparency.**

  Paper discusses about research by Rodriguez et al. [6] (2022) examined the presence of bias in natural language processing algorithms used to analyse candidate responses. They identified instances where the algorithms exhibited bias towards certain

demographic groups, emphasizing the need for continuous monitoring and improvement to mitigate algorithmic bias. Similarly, in paper [8] Smith and Johnson (2020) investigated the potential bias in computer vision algorithms analysing facial expressions and cautioned against the uncritical use of such algorithms in decision-making.

2.4 COMPARATIVE ANALYSIS

Title	Year	Author(s)	Idea	Technique	Results	Challenges
Unveiling the Potential of AI-Based Interview Systems	2023	Vaibhav Sharma et al.	Reviews AI interview systems, their components and benefits.	NLP, Machine Learning, Computer Vision	Highlights efficiency, reduced bias, objectivity, scalability.	Ethical concerns, data privacy, algorithm fairness.
Real-Time Mock Interview Using Deep Learning	2021	Sahil Temgire et al.	Develops a mock interview system simulating virtual interviews.	OpenCV, CNN, Deep Learning	Provides feedback on facial expressions, reaction time, grammar.	Limited anti-cheating measures, lacks detailed feedback storage.
Soft Skills and Graduates' Employability	2021	Al Asefer & Zainal Abidin	Emphasizes the importance of soft skills for job readiness.	Literature Review	Soft skills rated more important than hard skills by employers.	Mismatch between educational output and industry needs.
Reconciling Hard Skills and Soft Skills	2023	Jeremy Lamri & Todd Lubart	Proposes a unified model of hard and soft skills.	Generic Skills Component Model, Hilgard's Trilogy of Mind	Demonstrates interdependence of skill types for workplace use.	Challenging traditional skill classification systems.
Real-Time Emotion Detection Using Deep Learning	2021	Noel Jaymon et al.	Focuses on emotion detection through facial expressions.	FER2013 Dataset, TensorFlow, Keras, Xception CNN	Achieved 93.2% training, 64.4% validation accuracy.	Limited generalization, complex model tuning.

Addressing Bias in AI-Based Interview Systems	2022	Rodriguez et al.	Studies bias in NLP algorithms in AI interviews.	NLP Evaluation	Found bias against demographic groups in candidate analysis.	Need for continuous monitoring and mitigation mechanisms.
Bias in Computer Vision for Interviews	2020	Smith & Johnson	Investigates biases in facial expression analysis algorithms.	Computer Vision	Warns against uncritical use of biased facial analysis systems.	Algorithmic bias, lack of transparency.
Challenges in Unconstrained Face Recognition	2018	Guodong Guo & Na Zhang	Tests DL models under varying face image quality.	VGGFace, Light CNN, CenterLoss, FaceNet	Strong performance with good quality; significant drop with poor quality.	Image quality variation is a major hurdle.
Multi-Modal Emotion Recognition Database	2019	Tengfei Song et al.	Introduces physiological emotion dataset for recognition.	EEG, GSR, ECG, Respiration, ALSTM, k-NN, SVM	Developed public dataset for emotion recognition research.	High cost and complexity of multi-modal data collection.
Speech & Speaker Recognition for Smart Systems	2021	A. Mittal, M. Dua, S. Dua	Studies speech as a behavioral trait for AI systems.	MFCC, Deep Learning, Raw waveform analysis	Improved performance in speaker and speech recognition.	Contextual understanding limitations in noisy environments.

Table 2.1: Comparative Analysis

2.5 RESEARCH GAP AND LIMITATION

Despite significant advancements in AI-based interview systems, most existing solutions focus primarily on verbal analysis or static facial recognition, often neglecting real-time behavioral cues like eye contact, hand gestures, and emotional adaptability. Additionally, many platforms provide generic feedback, lacking personalized guidance for skill development. There is also limited integration of adaptive learning mechanisms that adjust to a user's unique progress. Furthermore, concerns around algorithmic bias, data privacy, and response interpretation still persist. These gaps highlight the need for a comprehensive system like SkillSight, which combines deep learning, behavioral analysis, and customized learning paths to offer more accurate and actionable assessments.

2.6 METHODOLOGICAL APPROACH

The methodological approach in existing literature typically involves the integration of AI technologies such as Natural Language Processing (NLP), computer vision, and deep learning models to evaluate candidate responses. Studies commonly employ datasets like FER2013 for facial emotion recognition and use models like CNNs, RNNs, or transformer-based architectures to analyze verbal and non-verbal cues. Many approaches rely on supervised learning with annotated datasets for training. However, while these methodologies demonstrate promising accuracy, they often overlook real-time adaptability, personalized feedback mechanisms, and multi-modal interaction—key areas addressed by the proposed SkillSight platform to ensure a more holistic interview assessment experience.

2.7 PROJECT PLAN

The **SkillSight: Interview & Assessment Platform for Soft Skills** was developed through a structured, phase-wise approach that allowed for focused progress across frontend development, backend architecture, AI integration, and system testing—ensuring an efficient and goal-driven implementation.

Project Phases

Phase	Timeline	Activities Performed
Requirement Analysis	1st Jan 2025 – 5th Jan 2025	Defined project scope, identified soft skill assessment needs, and finalized system objectives.
Research & Literature Survey	6th Jan 2025 – 10th Jan 2025	Studied emotion detection, NLP techniques, mock interview systems, and soft skill evaluation models.
System Design	11th Jan 2025 – 14th Jan 2025	Designed system architecture including video/audio capture flow, AI pipelines, and scoring logic.
Backend Development	15th Jan 2025 – 25th Jan 2025	Developed Flask APIs for video/audio input, emotion detection (OpenCV/Xception), and NLP-based response analysis.
Frontend Development	26th Jan 2025 – 4th Feb 2025	Built Web-based user interface for interview simulation, live input capture, and feedback display.
Integration & Testing	5th Feb 2025 – 10th Feb 2025	Integrated frontend and backend modules, performed unit and integration testing, and ensured real-time response accuracy.
Deployment & Documentation	11th Feb 2025 – 15th Feb 2025	Finalized the system setup, conducted final tests, prepared the project documentation and submitted the black book.

Table 2.2: Project Plan

3. PROBLEM STATEMENT

In the current employment landscape, the emphasis on soft skills—such as communication, emotional intelligence, adaptability, and interpersonal behavior—has become as crucial as technical proficiency. However, most educational systems and training environments continue to focus predominantly on hard skills, leaving a significant gap in the development and evaluation of soft skills. This mismatch results in a scenario where technically sound candidates struggle to perform effectively in real-world interviews due to inadequate interpersonal capabilities. As industries increasingly shift to virtual hiring processes, the need for a realistic and intelligent platform that helps users practice and improve their soft skills has never been more urgent.

Traditional interview preparation tools, like question banks or static mock interview sessions, often lack personalization, interactivity, and real-time feedback mechanisms. These systems do not assess non-verbal cues such as facial expressions, eye contact, tone of voice, or body language—factors that heavily influence interviewer perceptions. Additionally, the absence of tailored learning paths means users cannot identify specific weaknesses or receive actionable feedback, which is essential for iterative improvement. This leaves job seekers, especially fresh graduates and career changers, underprepared for the soft skill-intensive interview environments they encounter.

From the employer's perspective, evaluating candidates' soft skills during hiring remains a time-consuming and subjective process. Organizations often need multiple interview rounds to assess behavioral and communication competencies, leading to inefficiencies and higher recruitment costs. Moreover, studies reveal that a large percentage of hiring failures stem from poor soft skills rather than technical shortcomings, highlighting the need for a more structured and objective pre-screening method. A system that can pre-assess these attributes would save time and help in making more informed hiring decisions.

Despite advancements in artificial intelligence and machine learning, existing interview preparation platforms fall short in several aspects. Many either focus solely on voice or textual

input without considering visual and behavioral data, or they provide generic feedback that lacks depth. Systems that attempt facial expression analysis often suffer from issues like model bias, privacy concerns, and limited real-world accuracy under varying video quality conditions. Furthermore, such platforms rarely offer adaptive learning mechanisms that evolve with the user's progress.

SkillSight aims to bridge these gaps by creating an AI-driven platform that not only simulates interviews but also evaluates candidates holistically—through voice, facial recognition, emotional analysis, and personalized feedback. However, developing such a comprehensive system introduces its own set of challenges, including the integration of multimodal data, bias mitigation, privacy safeguards, and maintaining system responsiveness. Addressing these challenges is essential to build a reliable, fair, and impactful interview preparation tool for the next generation of professionals.

4. Software Requirements Specification

4.1 Introduction

This section outlines the general context and purpose of the SkillSight – Interview and Assessment Platform, including its scope and the nature of its users.

4.1.1 Project Scope

SkillSight is an AI-driven interview and assessment platform designed to simulate real-world job interviews and evaluate users' soft skills in a dynamic, interactive manner. The system harnesses cutting-edge technologies in Artificial Intelligence (AI), Natural Language Processing (NLP), and Computer Vision to deliver personalized and comprehensive feedback to users. Key functionalities include:

- Real-time audio-video capture of user responses to simulate authentic interview scenarios.
- Facial expression and emotion recognition using deep learning models like Convolutional Neural Networks (CNN) to evaluate non-verbal cues.
- Natural Language Processing-based response analysis to assess clarity, coherence, and relevance of user answers using tools like spaCy and NLTK.
- Behavioral pattern analysis incorporating gestures, eye movement, and voice tone using OpenCV and sentiment analysis models.
- Personalized feedback generation through AI-driven evaluation metrics that highlight strengths and suggest areas for improvement.
- Secure and user-friendly web interface developed with React and Flask to ensure a seamless, intuitive experience.

The platform is designed to be modular, scalable, and domain-agnostic, making it suitable for individual job seekers, educational institutions, recruitment firms, and corporate training programs. SkillSight redefines how users prepare for interviews by integrating multimodal assessment and intelligent feedback, ultimately enhancing employability and soft skill proficiency across diverse user profiles.

4.1.2 User Classes and Characteristics

- **General Users**
    - Can register, log in, and participate in AI-simulated mock interviews.
    - Receive feedback on communication skills, emotional intelligence, and body language.
    - Require no technical background; the interface is designed to be intuitive and user-friendly.
    - Expect personalized, real-time assessments and recommendations for improvement.
- **Administrative Users**
    - Manage platform operations such as user access, interview content moderation, and performance analytics.
    - Monitor logs, usage statistics, and system health.
    - Require a basic understanding of AI components, user behavior metrics, and system workflow.
    - Can update question banks and manage feedback templates.
- **Developers/Maintainers**
    - Responsible for backend maintenance, AI model integration, and system upgrades.
    - Must be proficient in Python, React, Flask, OpenCV, and NLP libraries like spaCy or NLTK.
    - Handle bug fixes, performance optimization, and security updates.
    - Involved in deploying and refining machine learning models for facial expression and sentiment analysis.

4.2 Functional Requirements

This section defines the primary features and behaviour of the system in response to user and system-level inputs.

4.2.1 System Feature 1: Real-Time Interview Simulation

Description:

Users can participate in mock interviews where they respond to AI-generated questions through audio and video. The system captures responses in real-time for further analysis.

Functional Requirements:

- FR1.1: The system shall initiate webcam and microphone access when the interview begins.
- FR1.2: The system shall present dynamically generated questions based on selected topics.
- FR1.3: The system shall record audio and video responses during the interview session.

4.2.2 System Feature 2: Facial Expression and Emotion Analysis

Description:

The system analyzes user facial expressions and emotional states using computer vision and deep learning models to assess non-verbal communication.

Functional Requirements:

- FR2.1: The system shall use OpenCV and CNN-based models to detect facial landmarks and classify expressions.
- FR2.2: Emotion classification results shall be stored for feedback generation.
- FR2.3: The system shall map emotional responses to soft skill attributes such as confidence, stress, and empathy.

4.2.3 System Feature 3: NLP-Based Response Evaluation

Description:

User responses are evaluated using NLP techniques to assess clarity, relevance, and linguistic fluency.

Functional Requirements:

- FR3.1: The system shall transcribe spoken responses using a speech-to-text engine.
- FR3.2: The system shall analyze the transcribed text using NLP libraries like spaCy or NLTK.

- FR3.3: Evaluation metrics such as grammar correctness, sentiment, and coherence shall be generated and stored.

4.2.4 System Feature 4: Personalized Feedback Generation

Description:

The system generates detailed feedback based on verbal and non-verbal analysis to help users improve soft skills.

Functional Requirements:

- FR4.1: The system shall aggregate insights from both NLP and facial expression modules.
- FR4.2: A feedback report shall highlight strengths, weaknesses, and actionable suggestions.
- FR4.3: The feedback shall be tailored to the user's performance and interview topic.

4.3 External Interface Requirements

This section describes how the chatbot system interfaces with users, hardware, and software, including both the frontend and backend components.

4.3.1 User Interfaces

- **Frontend Application:**
  - Allows users to:
    - Register and log in securely.
    - Select interview topics and start mock interview sessions.
    - View real-time feedback and performance scores after each interview.
  - Features:
    - Intuitive UI with a dashboard for initiating interviews and tracking progress.
    - Live webcam and microphone integration for response capture.
    - Real-time response display and scoring.
    - Interview history panel with timestamps and performance metrics.
- **Admin Interface (optional future extension):**

- Dashboard for monitoring:
    - User activity and session logs.
    - AI performance metrics and error tracking.
    - System health and usage analytics.

4.3.2 Hardware Interfaces

- **Client-Side (User Device):**
    - Device with webcam and microphone.
    - Modern web browser (latest versions of Chrome, Firefox, Safari, Edge).
    - No specific hardware dependencies other than media capture capability.
- **Server-Side (Hosting Backend):**
    - Server with:
        - Python environment installed.
        - Flask and machine learning libraries (e.g., TensorFlow, OpenCV).
        - Access to external APIs for NLP and video analysis.
    - Optional GPU support for heavy model inference tasks such as real-time emotion detection.

4.3.3 Software Interfaces

- **Frontend ↔ Backend Communication:**
    - RESTful API endpoints (e.g., /start-interview, /analyze-response, /get-feedback).
    - Communication over HTTPS using JSON payloads.
- **Third-party Services:**
    - Google Generative AI API for:
        - Text transcription and NLP processing.
        - Contextual response evaluation.
    - Emotion detection models (e.g., Xception) integrated via OpenCV.
    - Sentiment analysis tools and speech-to-text APIs.
- **Python Libraries:**
    - OpenCV, SpeechRecognition, spaCy, NLTK, Flask, TensorFlow, dotenv, etc.

- For data processing, model inference, and server-side logic.

4.3.4 Communication Interfaces

- **Protocols:**
  - HTTP/HTTPS for secure communication between client and server.
  - TCP/IP stack for all backend communications and API calls.
- **Data Format:**
  - JSON used for all API request and response bodies.
  - Encoded video/audio data used internally for response capture and analysis.

4.4 Nonfunctional Requirements

This section outlines the overall system attributes such as performance, reliability, security, and maintainability.

4.4.1 Performance Requirements

- The system shall process and analyze user audio-video responses within 5–7 seconds under standard network and hardware conditions.
- Real-time facial expression detection shall produce inference results within 2 seconds per frame.
- NLP-based response analysis and feedback generation shall complete within 3–5 seconds per interview question.
- The system shall support up to 100 concurrent interview sessions without noticeable performance degradation.

4.4.2 Safety Requirements

- The system shall support:
  - Concurrent usage by multiple users through horizontal backend scaling.
  - Incremental analysis of individual interview sessions without needing full system reinitialization.

- User session crashes or interruptions shall automatically trigger recovery mechanisms and error logs.
- The system shall notify users of any technical or connectivity issues during interviews and log such incidents for administrative troubleshooting.

4.4.3 Security Requirements

- All user data (video, audio, text) must be transmitted over HTTPS to ensure data privacy.
- API keys and model endpoints shall be hidden via .env files and backend-protected routes.
- The system shall implement basic anti-spoofing measures (e.g., real-time face detection checks) to ensure authenticity.

4.4.4 Software Quality Attributes

- The codebase shall maintain a modular architecture with clear separation between:
  - Interview session handling
  - Audio-video capture and analysis
  - Natural language processing
  - Feedback and scoring modules
  - User interface and session history
- The platform shall be extensible to support future integration with other AI models (e.g., OpenAI, Amazon Comprehend, Emotion API).
- Logging and monitoring tools shall be integrated for easier debugging and system auditing.
- The frontend and backend components shall be independently deployable to facilitate microservices-based scaling.

4.5 System Requirements

This section outlines the necessary hardware and software configurations required to develop,

deploy, and run the chatbot system effectively.

4.5.1 Database Requirements

- **Local/Cloud File Storage**
  Used to temporarily store recorded audio-video responses, interview session logs, and system-generated feedback.
- **Relational Database (e.g., PostgreSQL/MySQL)**
  Required for user authentication, session tracking, feedback records, and admin dashboard data.

4.5.2 Software Requirements

- **Operating System:**
  Windows 10 or above / Linux (Ubuntu 20.04+ recommended)
- **Backend:**
  - Python 3.10+
  - Flask
  - OpenCV
  - TensorFlow / Keras (for facial emotion recognition)
  - spaCy / NLTK (for NLP)
  - SpeechRecognition (for audio-to-text transcription)
- **Frontend:**
  - HTML5 / CSS3
  - JavaScript
  - WebRTC (for real-time webcam and microphone access)
- **Development Tools:**
  - Visual Studio Code (VS Code)
  - GitHub / Git for version control
  - Postman (for API testing)
- **Browser Compatibility:**
  - Chrome, Firefox, Edge, Safari (latest versions)

4.5.3 Hardware Requirements

- **Processor:**
  Intel Core i5 or higher / AMD Ryzen 5 or equivalent
- **RAM:**
  Minimum 8 GB (16 GB recommended for smoother performance during real-time processing)
- **Storage:**
  - Minimum 100 MB for source code and system files
  - At least 1 GB of free space for user session data, cached videos, and logs
- **GPU (Optional but Recommended):**
  NVIDIA GPU (GTX 1050 or higher) for real-time emotion recognition and inference tasks using deep learning models

4.6 System Implementation Plan

The system implementation plan outlines the key stages involved in the development, deployment, and maintenance of the system. It covers the timeline, resources, risk management, and testing strategies necessary to ensure successful deployment.
The implementation process is broken down into several key phases:

1. **Phase 1: Requirements Gathering and Analysis**
   - **Objective:** Identify and document detailed functional and non-functional requirements from all relevant stakeholders including students, placement coordinators, and recruiters.
   - **Duration:** 4 weeks
   - **Tasks:**
     - Conduct interviews and surveys with end users (job seekers and educators).
     - Analyze competitor platforms to benchmark essential features.
     - Define project scope, key evaluation metrics (e.g., soft skill parameters), and system constraints (e.g., response time, accuracy).

2. **Phase 2: System Design**
    - **Objective:** Architect the system components including data pipelines, UI flows, and integration layers.
    - **Duration:** 4 weeks
    - **Tasks:**
        - Develop UML diagrams, flowcharts, and system wireframes for the platform.
        - Design backend architecture involving modules for emotion recognition, NLP, and feedback generation.
        - Finalize frontend frameworks and backend technologies (e.g., React, Flask, OpenCV, TensorFlow).
        - Plan the database schema and data storage mechanisms for interview logs and user feedback.
3. **Phase 3: Development**
    - **Objective:** Implement the core functionalities of the SkillSight platform as per the finalized design.
    - **Duration:** 12 weeks
    - **Tasks:**
        - **Backend Development:**
            - Build RESTful APIs for interview management, facial expression analysis, and feedback delivery.
            - Integrate NLP models and speech-to-text pipelines.
        - **Frontend Development:**
            - Develop a user-friendly UI with real-time video/audio streaming and session tracking.
            - Implement dashboards for users and admins.
        - **Integration:**
            - Ensure seamless connectivity between frontend, backend, and third-party APIs (e.g., AI models, cloud services).

4. **Phase 4: Testing and Debugging**
    - **Objective:** Validate the system's correctness, performance, and user experience.
    - **Duration:** 4 weeks
    - **Tasks:**
        - **Unit Testing:** Test each module including video capture, NLP, and feedback engines.
        - **Integration Testing:** Verify communication between frontend, backend, and AI services.
        - **Load Testing:** Simulate concurrent interviews to assess system stability and response times.
        - **User Acceptance Testing (UAT):** Conduct live tests with stakeholders and iterate based on feedback.
5. **Phase 5: Deployment**
    - **Objective:** Deploy the SkillSight system in a production-ready environment.
    - **Duration:** 2 weeks
    - **Tasks:**
        - Set up cloud infrastructure (e.g., AWS EC2/S3, Firebase) for hosting and data storage.
        - Configure security (HTTPS, CORS, API key handling) and scalability settings.
        - Integrate monitoring/logging tools (e.g., Prometheus, Sentry) for maintenance.
        - Conduct pre-launch validation and initiate go-live rollout.

4.7 System Design

4.7.1 System Architecture

The system is designed to simulate mock interviews for users by capturing audio-video responses, analyzing soft skills using AI techniques, and delivering personalized feedback and learning paths to enhance interview readiness and communication capabilities.

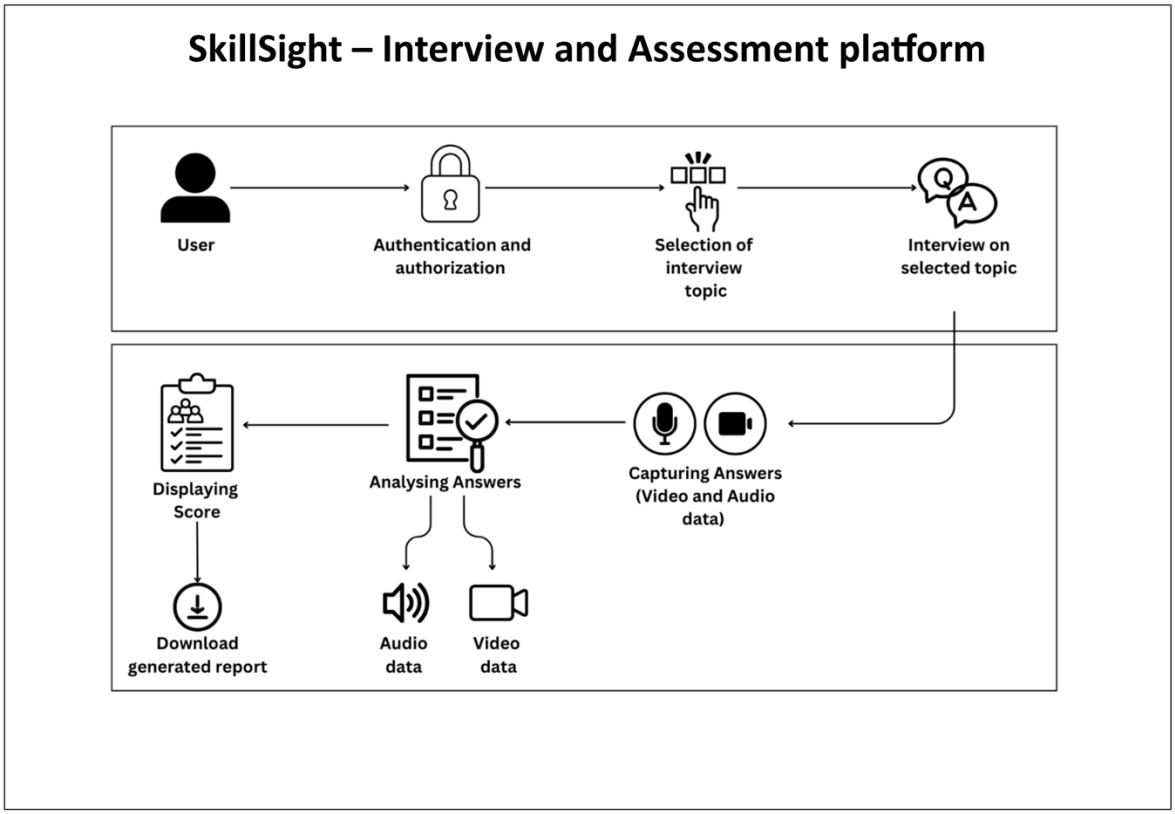


Fig4.1 System Architecture

1. **User Authentication and Topic Selection:**

   The user accesses the system through a secure authentication and authorization process. Once logged in, the user selects a desired interview topic to begin the simulation process.

2. **Interview Simulation:**

   The system initiates a virtual interview where the user is prompted with both topic-specific and general questions, simulating a real-world interview environment.

3. **Capturing Responses (Audio and Video Data):**

   During the interview, the user's responses are recorded via webcam and microphone. This raw media data (audio and video) forms the basis for subsequent analysis.

4. **Analyzing Answers:**

   The recorded audio and video data is processed through various AI modules. Audio is

transcribed and analyzed using NLP for linguistic quality, while video is analyzed using computer vision to evaluate facial expressions, tone, and body language.

5. **Score Display:**
   Based on the multimodal analysis, a performance score is generated and displayed to the user, indicating their effectiveness across several soft skill dimensions (e.g., confidence, clarity, emotional tone).
6. **Score Analysis for Improvement:**
   The system further analyzes the generated score to identify areas of improvement, isolating weak points in communication, non-verbal behavior, or response timing.
7. **Presenting Recommended Improvements:**
   Based on the analysis, the user receives detailed, actionable feedback including strengths and weaknesses in their responses, fostering a data-driven improvement process.
8. **Personalized and Adaptive Learning Path:**
   Leveraging a recommendation algorithm, the system curates a customized learning path tailored to the user's performance. This includes targeted exercises and training modules to reinforce weak areas.
9. **Suggesting Additional Content for Improvement:**
   The system concludes by offering supplementary resources such as tutorials, practice sessions, and articles aligned with the user's learning needs, thus supporting continuous soft skill development.

This workflow illustrates how SkillSight integrates user interaction, multimedia analysis, machine learning, and personalized learning to create an end-to-end intelligent interview preparation platform

4.7.2 Data Flow Diagram:

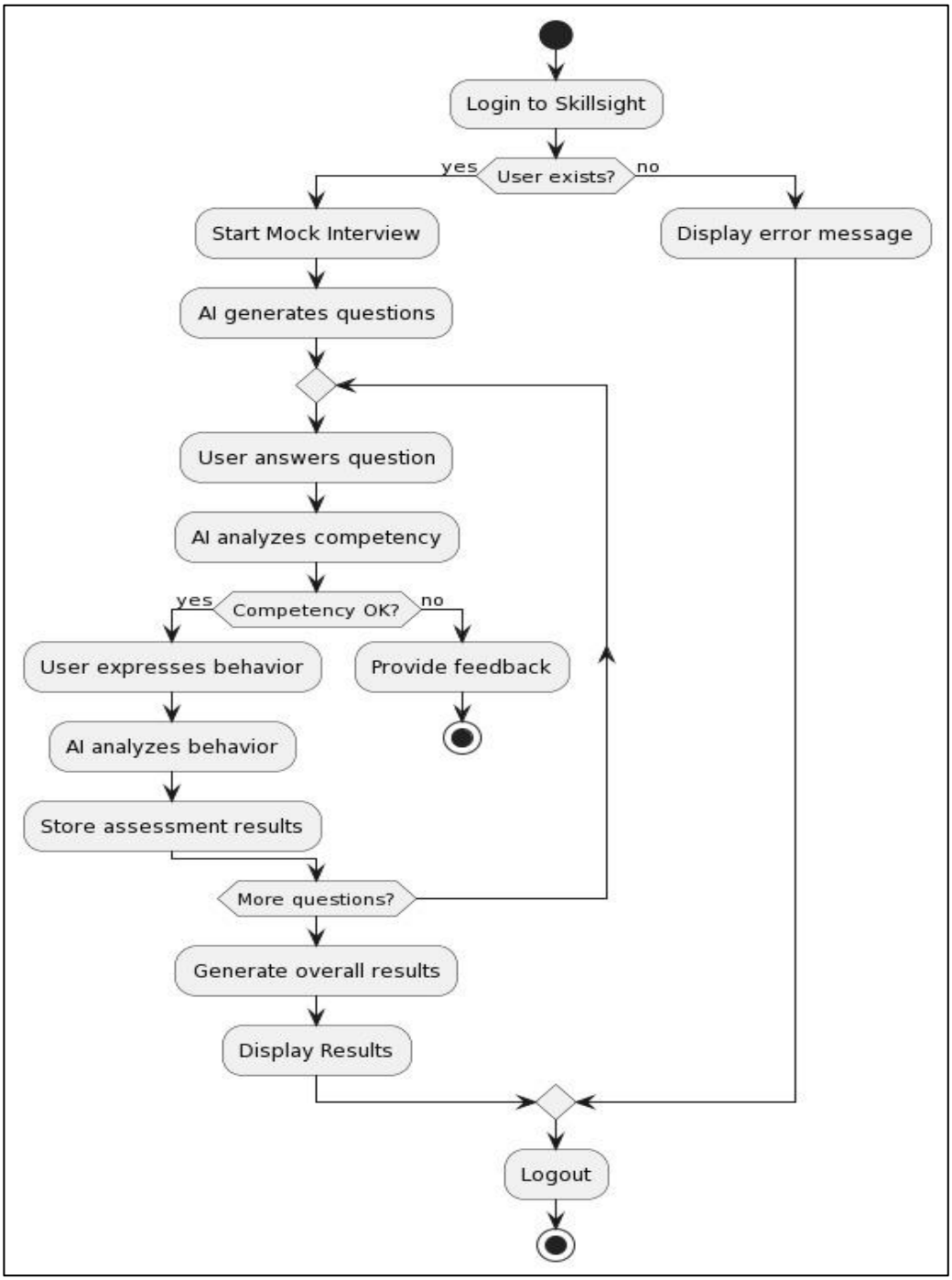


Figure 4.2 Data Flow Diagram

4.7.3 Entity Relationship Diagram

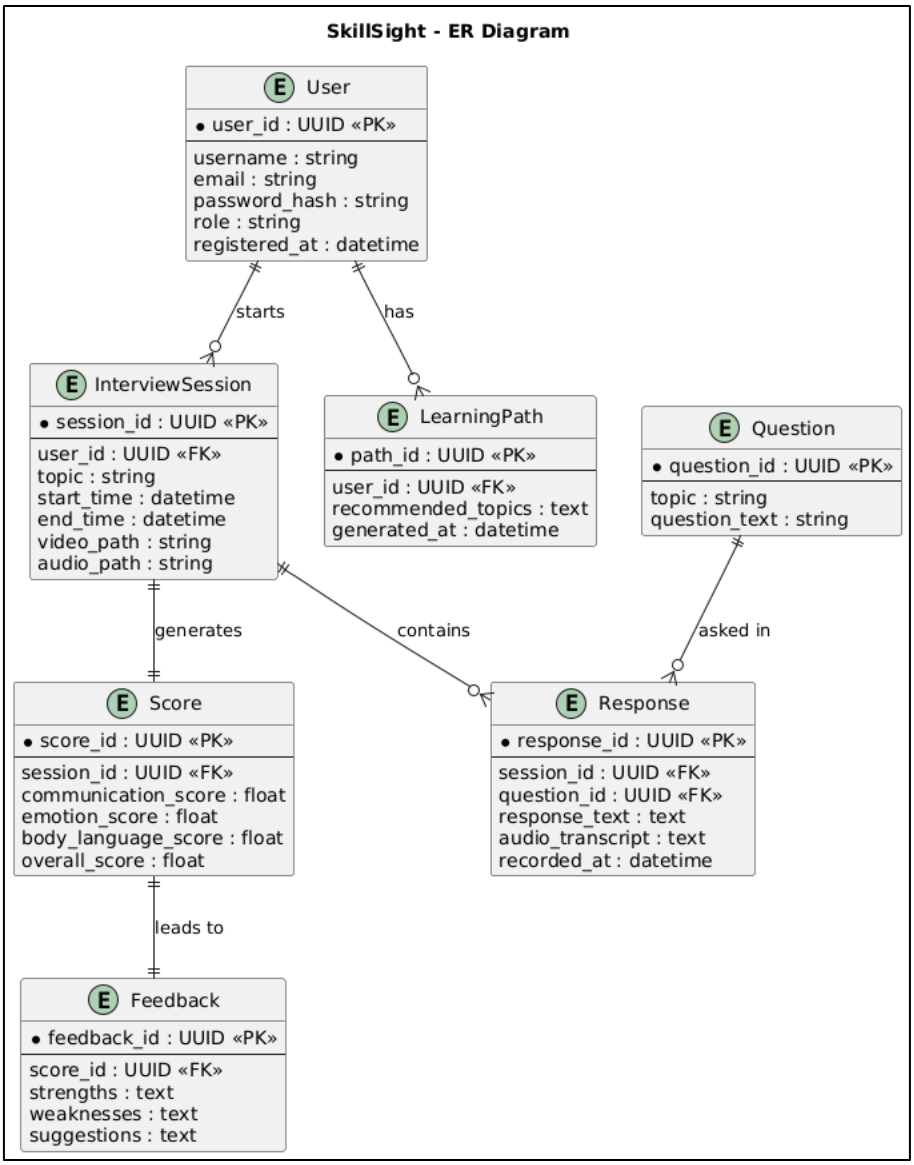


Figure 4.3 ER Diagram

4.7.4 UML Diagrams

1. Component Diagram:

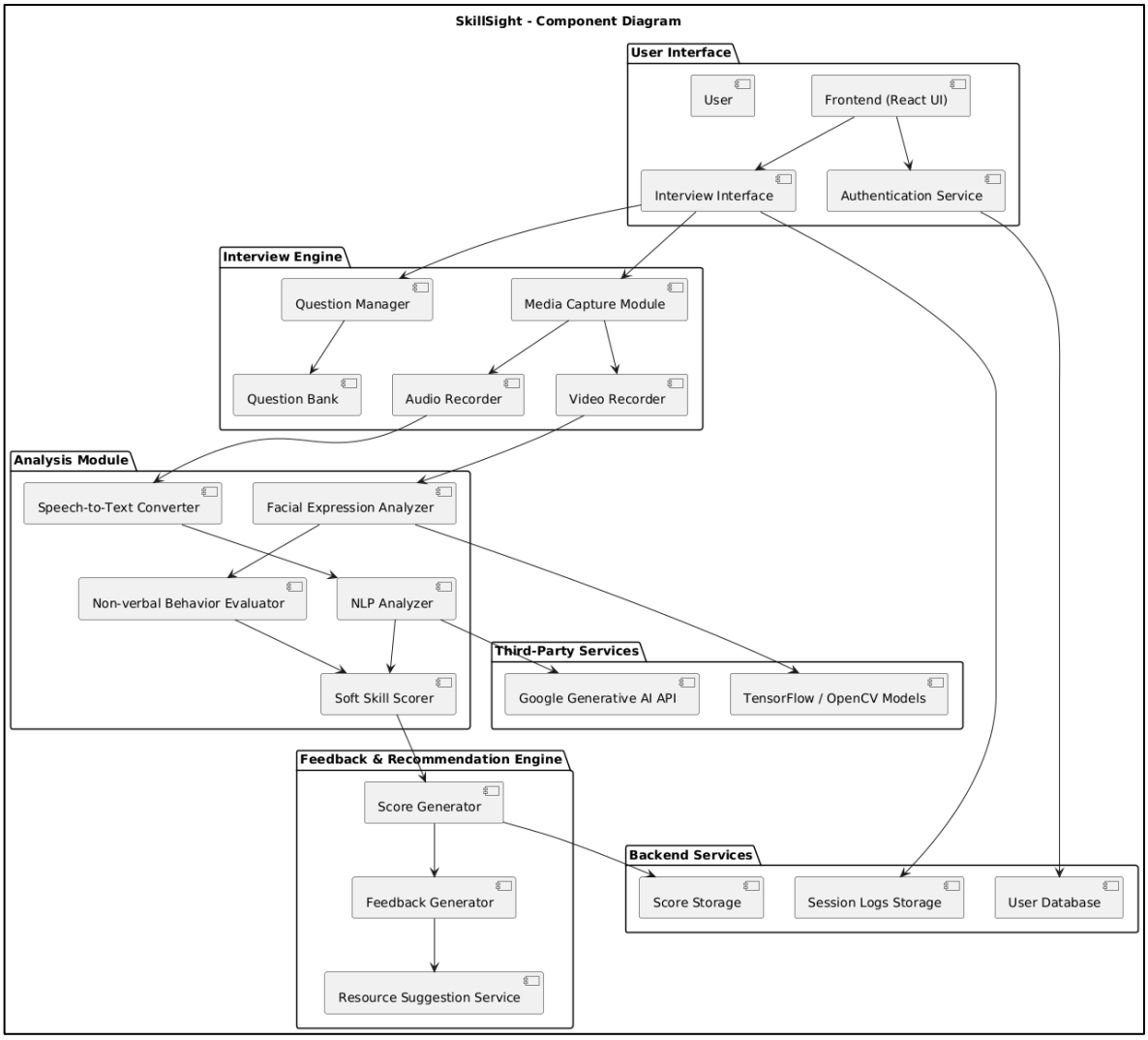


Figure 4.4: Component Diagram

2. Class Diagram:

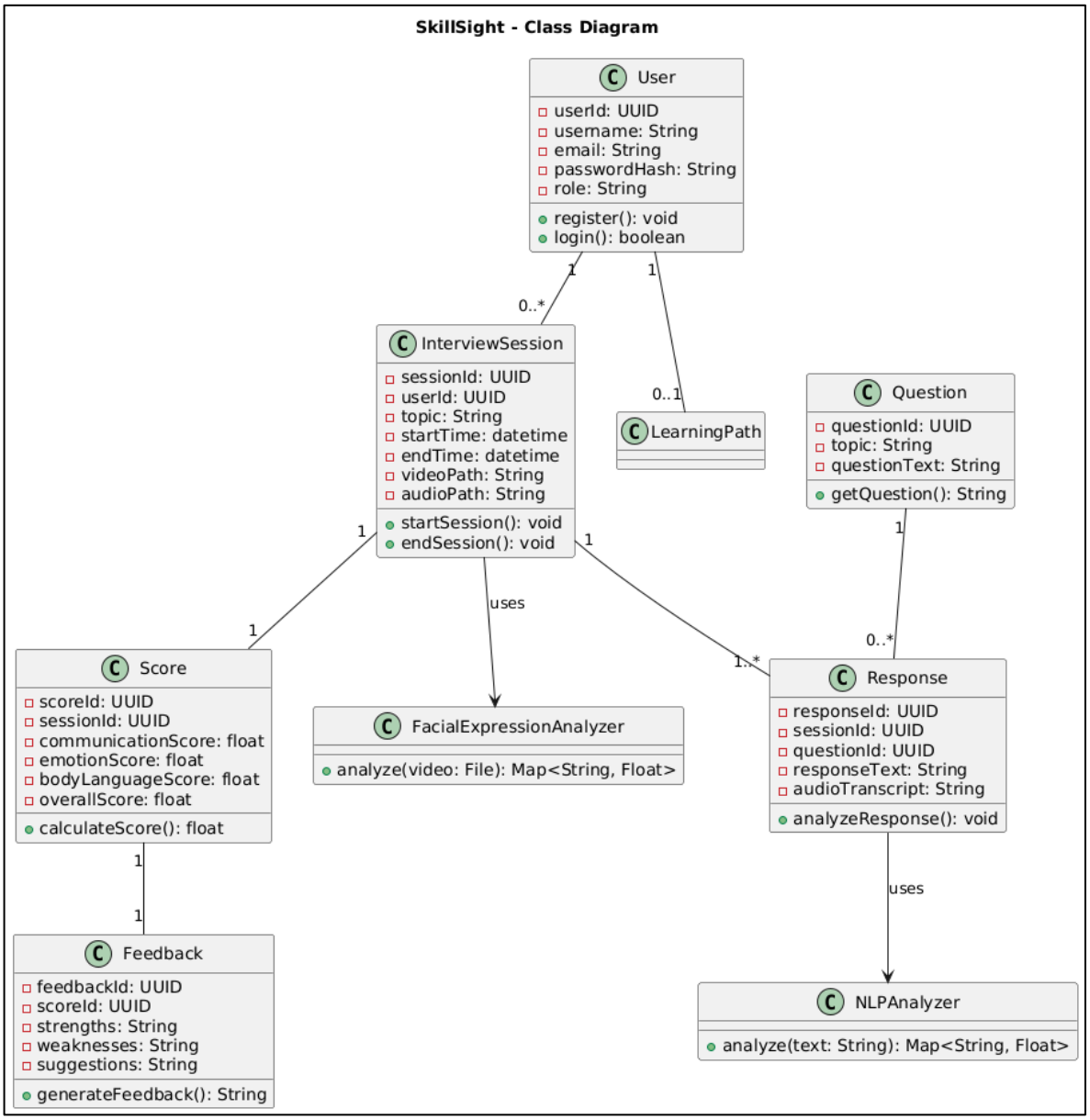


Figure 4.5: Class Diagram

3. Use Case Diagram:

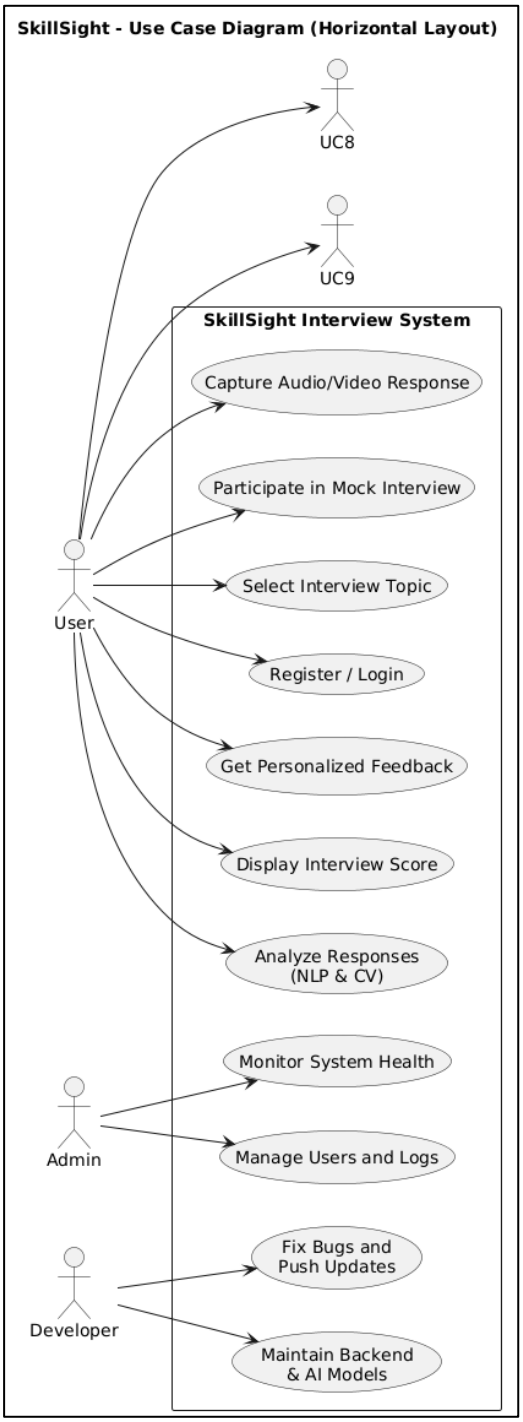


Figure 4.6: Use Case Diagram

4. Sequence Diagram

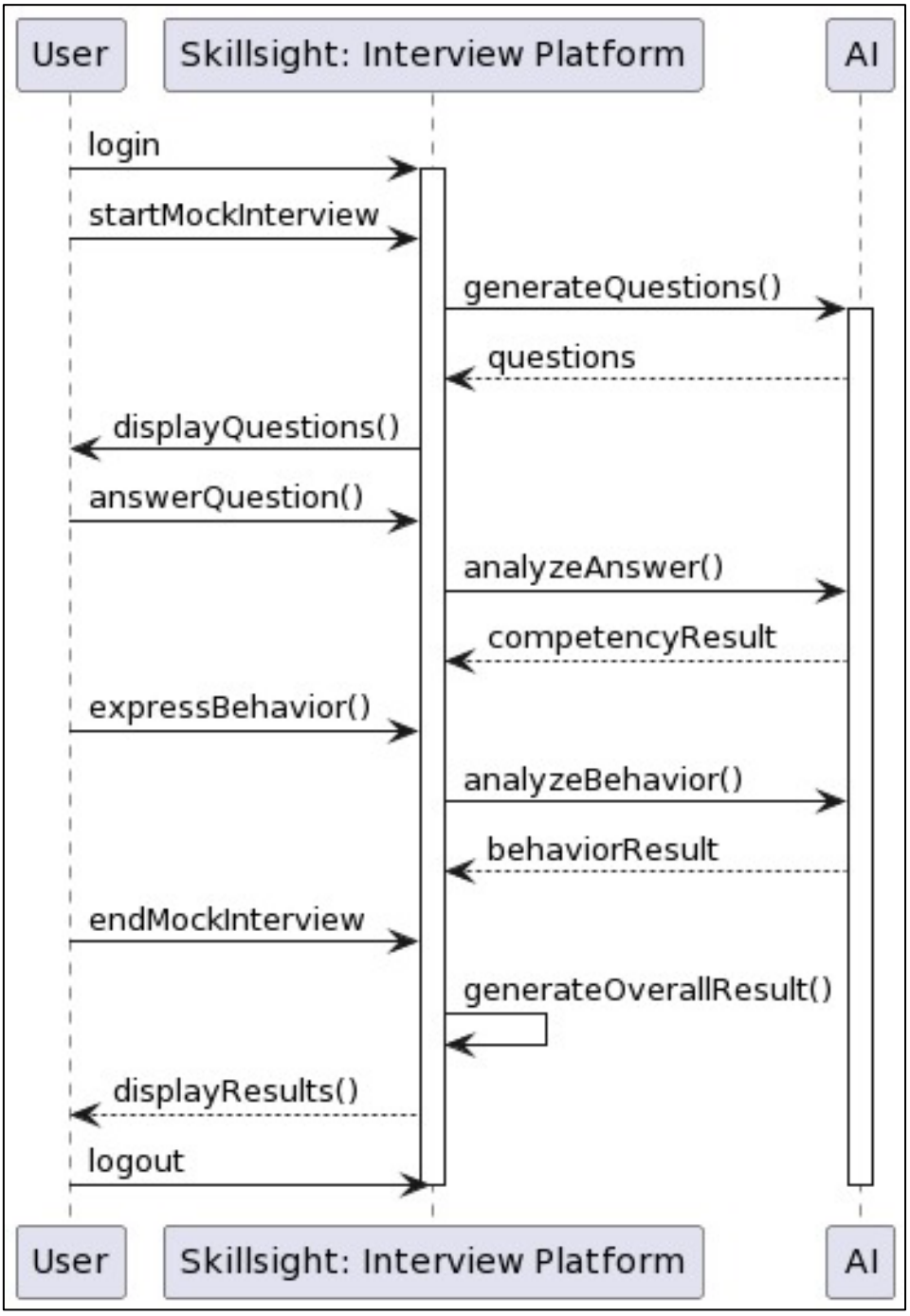


Figure 4.7: Sequence Diagram

5. PROJECT IMPLEMENTATION

This chapter provides a detailed explanation of the methodologies and processes used in implementing the interview system, it discusses the challenges faced during the development, how they were resolved, and a detailed breakdown of each module in the project. The approach adopted ensures that the interview system a seamless and engaging user experience, leveraging advanced NLP techniques and generative AI models.

5.1 OVERVIEW

The implementation of the SkillSight platform involves building an AI-powered mock interview system capable of evaluating a candidate's soft skills through real-time interaction. The architecture leverages advanced technologies such as Natural Language Processing (NLP), Computer Vision (CV), and machine learning models to simulate interviews, analyze user responses, and provide detailed feedback. Core functionalities include facial expression detection, sentiment analysis, speech-to-text conversion, and natural language understanding to assess parameters like communication ability, emotional intelligence, and non-verbal behaviour.

The system is composed of modular components that handle video/audio capture, response analysis, scoring, and recommendation generation. It integrates tools like OpenCV for visual analysis, spaCy for language processing, and TensorFlow-based models for emotion detection. Each interview session is processed to produce personalized scores and learning paths, guiding users toward targeted soft skill improvement. With its scalable design and interactive frontend, SkillSight is suitable for use by students, job seekers, educators, and recruitment platforms in real-world assessment scenarios.

5.2 LIST OF MODULES

1. **Module 1 - Frontend Implementation**

   The frontend, built using React, manages user interaction during mock interviews, renders AI-generated feedback, and provides a dynamic and intuitive experience.

- **Key Components:**

1. **Interview Interface**
   - The InterviewInterface component manages the real-time simulation.

- Users interact with webcam and mic-based prompts, with media captured via browser APIs.
- Inputs like selected interview topic and user video/audio streams are managed using useState and useEffect.

2. **Live Feedback Loader**
   - The LiveFeedbackLoader visually indicates that analysis is in progress.
   - Enhances UX during model inference and server processing.
3. **Session Tracker**
   - The SessionContext manages the application state.
   - Tracks:
     1) User's verbal and non-verbal responses
     2) Real-time feedback from backend
     3) Session history for review and reflection
   - Updates dynamically as new responses or feedback are received.
4. **Feedback Renderer**
   - The FeedbackRunner component ensures real-time rendering of feedback and scoring.
   - Uses useRef and useLayoutEffect to auto-scroll and maintain view consistency.

2. Module 2 - Backend Implementation

The backend, built with Flask, powers the core analysis engine: handling media input, AI models, NLP, and scoring.

- **Key Backend Features:**

1. **Audio-Video Capture & Preprocessing**
   - Captured media is received and parsed using Flask routes.
   - Audio is transcribed using speech-to-text engines (e.g., SpeechRecognition/Whisper).
   - Video is processed frame-by-frame via OpenCV for facial expression analysis.
2. **Facial Expression and Emotion Detection**
   - Uses CNN-based models (e.g., Xception) to extract emotion metrics like confidence, anxiety, and positivity.

- Output is scored and logged for feedback generation.

3. **Speech and NLP Analysis**
    - Transcribed text is analyzed using spaCy or NLTK for tone, grammar, coherence, and keyword relevance.
    - NLP outputs are scored to assess communication effectiveness.
4. **Soft Skill Scoring & Aggregation**
    - Inputs from video, audio, and NLP pipelines are fed into a scoring engine.
    - Generates scores for communication, emotional intelligence, non-verbal cues, and overall performance.
5. **Feedback and Recommendation Engine**
    - AI-generated insights are compiled into strengths, weaknesses, and suggestions.
    - Tailored learning paths are generated using a rule-based or ML-based recommender system.

3. Module 3 - Flask Routing and Integration

Workflow:

1. **Routing:**
    a. **/start-interview**
        - Initializes a session, triggers media capture, and sets session parameters.
    b. **/analyze-response**
        - Accepts user response data (video/audio), processes it via AI modules, and returns scores + feedback.
2. **API Flow:**
    a. **Frontend Request:**
        - React sends POST requests using Axios or Fetch API.
        - Payload includes video/audio blobs and session metadata.
    b. **Backend Response:**
        - Flask orchestrates the analysis pipeline and returns structured feedback in JSON format.

4. Module 4 - Backend Processing Pipeline

1. **Media Capture & Extraction**
    - **Input:** Webcam/mic data during mock interview sessions.
    - **Process:** Extracts and stores raw media for processing.
    - **Output:** Audio and video files.
2. **Facial Emotion Recognition**
    - **Input:** Video frames.
    - **Process:** OpenCV + CNN models classify expressions (e.g., happy, nervous, confused).
    - **Output:** Emotion scores.
3. **Speech-to-Text & NLP**
    - **Input:** Audio streams.
        - **Process:** Transcribe and analyze speech for fluency, sentiment, and keyword coverage.
        - **Output:** Linguistic performance metrics.
4. **Skill Scoring**
    - **Input:** Emotion and NLP outputs.
    - **Process:** Weight-based or ML-driven scoring model.
    - **Output:** Category-wise and overall performance score.
5. **Feedback Generation**
    - **Input:** Scoring data.
    - **Process:** Feedback templates are dynamically filled with personalized insights.
    - **Output:** Response string and improvement suggestions.
6. **Frontend Integration**
    - **Input:** JSON from backend.
    - **Process:** Chat interface updates with real-time feedback and visual cues.
    - **Output:** A smooth and responsive user experience.

5.3 TOOLS AND TECHNOLOGIES USED

Frontend Technologies:

- **JavaScript** (for dynamic and interactive features)
- **HTML/CSS** (for structure and styling)

Backend Technologies:

- **Flask** (for handling backend logic and API routing)
- **Python** (for backend development, AI model integration, and processing)
- **OpenCV** (for facial expression and gesture recognition via video data)
- **TensorFlow & Keras** (for deep learning-based emotion recognition models)
- **Xception CNN Model** (for real-time facial emotion detection)
- **Natural Language Toolkit (NLTK) / spaCy** (for natural language processing and semantic analysis)
- **Speech Recognition APIs** (for converting spoken responses into text for evaluation)
- **MediaPipe** (for analyzing body language and hand gestures)
- **GoogleGenerativeAIEmbeddings** (for embedding textual responses for analysis and feedback generation)

Development and Version Control:

- **Visual Studio Code (VS Code)** (for code development, debugging, and project organization)
- **GitHub** (for version control, team collaboration, and codebase management)

5.4 ALGORITHM DETAILS

Convolutional Neural Networks (CNNs) for Facial Emotion Recognition

A key algorithmic component in the SkillSight platform is the use of **Convolutional Neural Networks (CNNs)**, particularly the **Xception architecture**, for analyzing user facial expressions during mock interviews. CNNs are widely adopted in computer vision due to their ability to automatically learn hierarchical features from raw image data, making them ideal for tasks like facial expression classification.

Advantages of CNNs Over Traditional Vision Methods

Traditional image processing techniques often rely on hand-crafted features that may not generalize well across different lighting, facial orientations, or individuals. CNNs, in contrast, learn discriminative features directly from training data and are robust to these variations.
The **Xception model**, a depthwise separable CNN, achieves high accuracy while being computationally efficient — crucial for real-time inference.

Key Components of the CNN-based Emotion Recognition System

1. **Input Preprocessing and Frame Extraction**:
   Video frames are extracted in real-time from the user's webcam feed. Each frame is resized, normalized, and converted to grayscale before being passed to the CNN model.
2. **Face Detection and Cropping**:
   Using **OpenCV** and **MediaPipe**, the system detects facial landmarks and crops the region of interest (ROI) containing the face for analysis. This ensures the model focuses on expressive regions such as eyes, eyebrows, and mouth.
3. **Xception Model for Feature Extraction**:
   The preprocessed face image is fed into a pre-trained **Xception CNN**, which extracts deep feature representations and classifies the expression into categories like happy, sad, angry, neutral, etc. These emotions are later correlated with performance metrics such as confidence or nervousness.
4. **Emotion Classification and Confidence Scoring**:
   The output layer of the model provides a softmax probability distribution over emotion classes. The predicted emotion and its confidence score are logged and used for behavioral analysis and feedback.
5. **Temporal Aggregation for Video Analysis**:
   Since emotions are dynamic, the model aggregates predictions across frames using a sliding window or majority voting approach to provide a stable emotion profile throughout the interview session.

Natural Language Processing (NLP) and Semantic Analysis Algorithms

For textual and spoken responses, the system uses **transformer-based language models** and **traditional NLP pipelines** (e.g., with **spaCy** or **NLTK**) to:

- Evaluate grammatical accuracy and clarity,
- Detect sentiment and tone,
- Extract semantic coherence and topical relevance.

These outputs are combined with emotion analysis to generate comprehensive interview feedback.

Importance of CNNs and NLP in SkillSight

Together, CNNs for facial expression recognition and NLP for verbal response evaluation form the backbone of SkillSight's assessment capabilities. They enable:

1. Real-time monitoring of candidate behavior and mood,
2. Insightful analysis of verbal and non-verbal communication,
3. Objective scoring across communication, emotional intelligence, and soft skills,
4. Personalized, data-driven feedback and learning suggestions.

These algorithmic components ensure that SkillSight provides a robust and intelligent interview simulation environment that mimics real-world conditions while supporting professional development.

5.5 SCREEN SHOT OF PROJECT EXECUTION

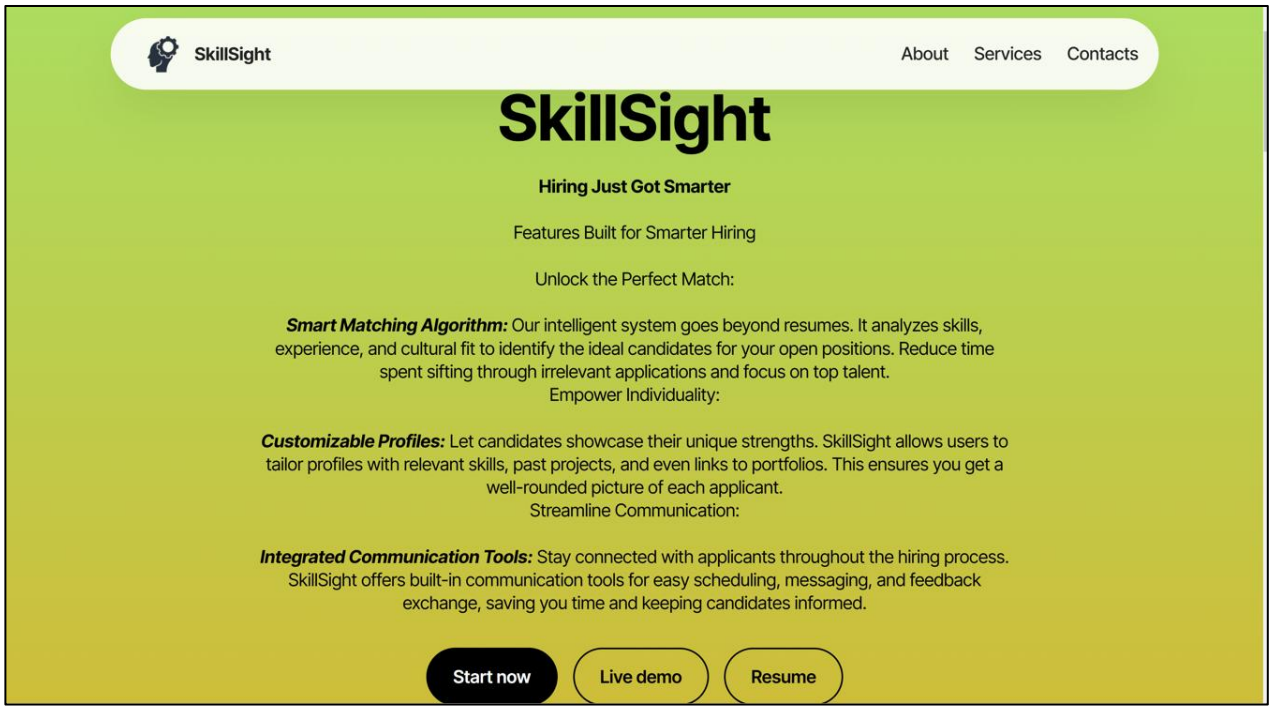


Figure 5.1: Home Page

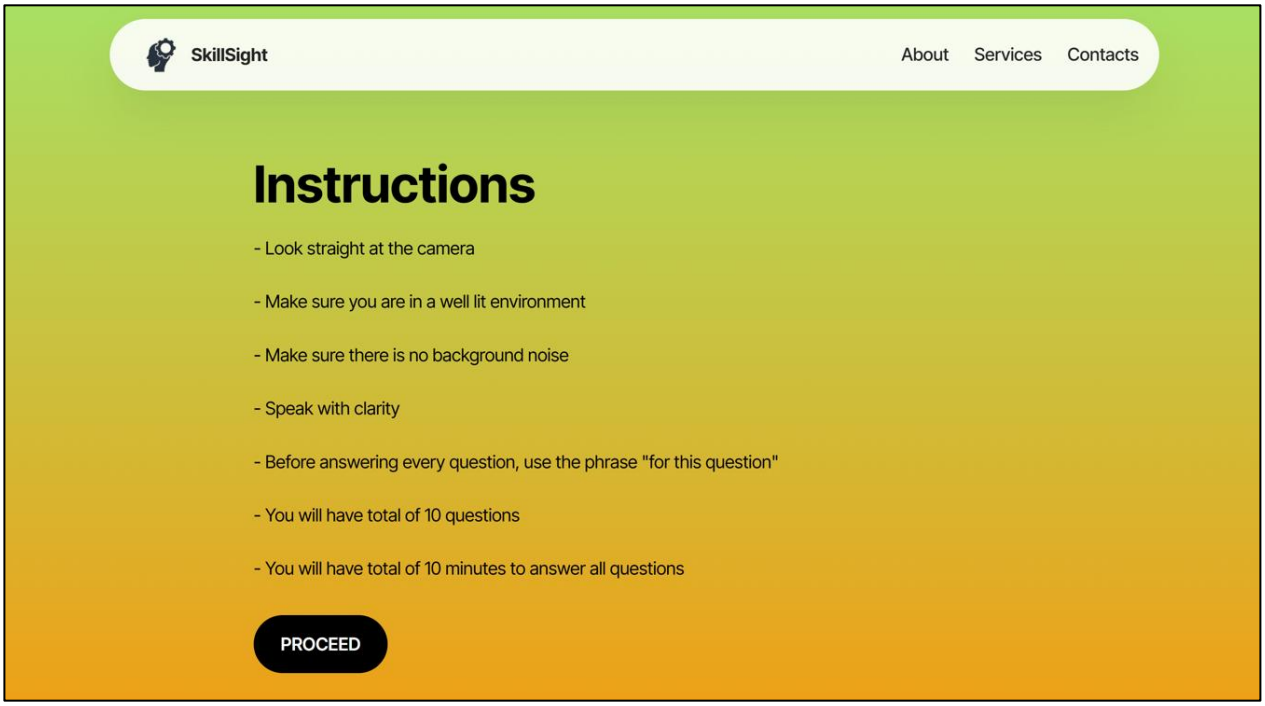


Figure 5.2: Instructions Page

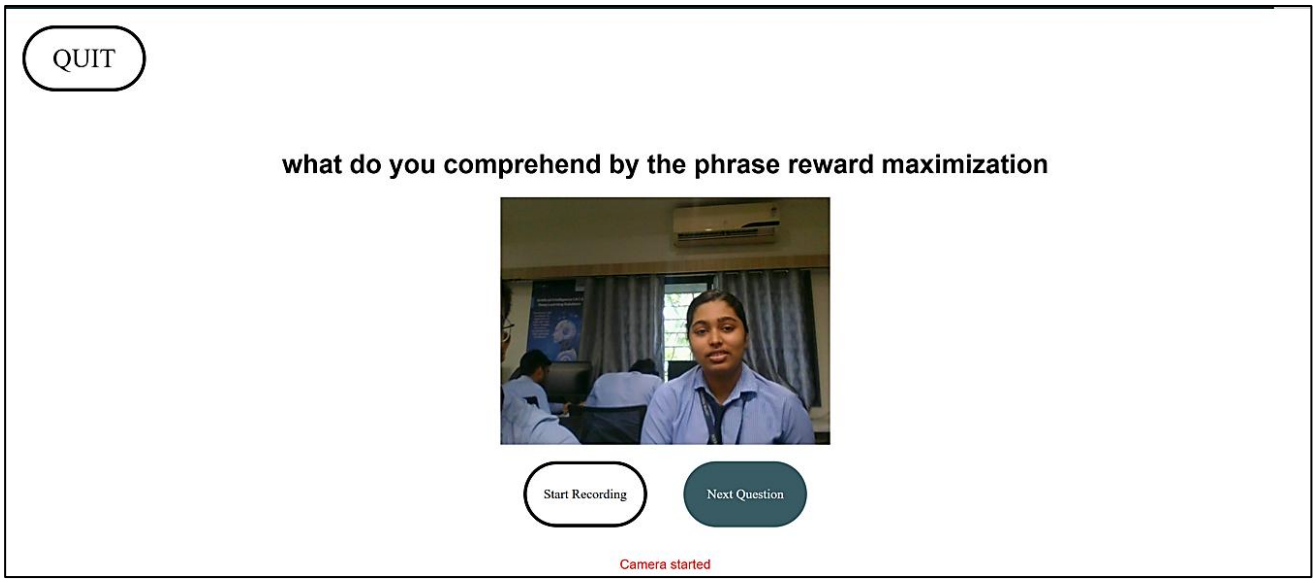


Figure 5.3: Interview Page

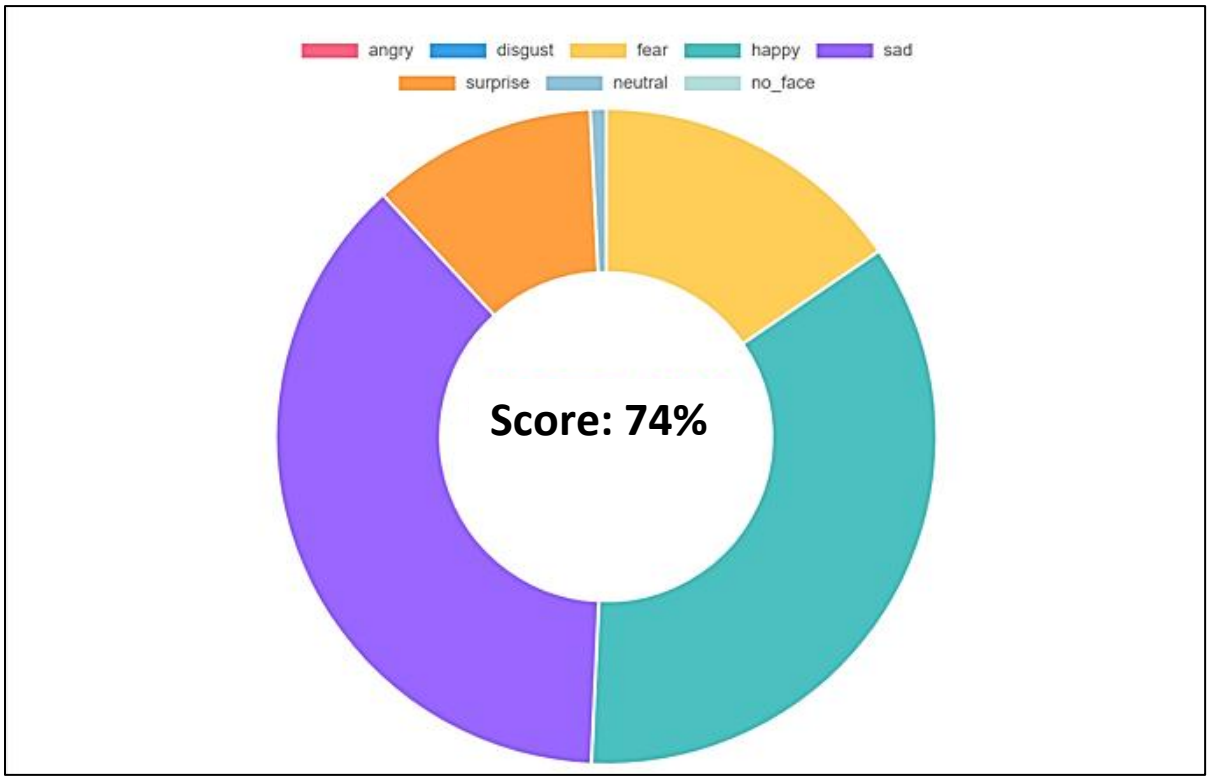


Figure 5.3: Test Interview Score

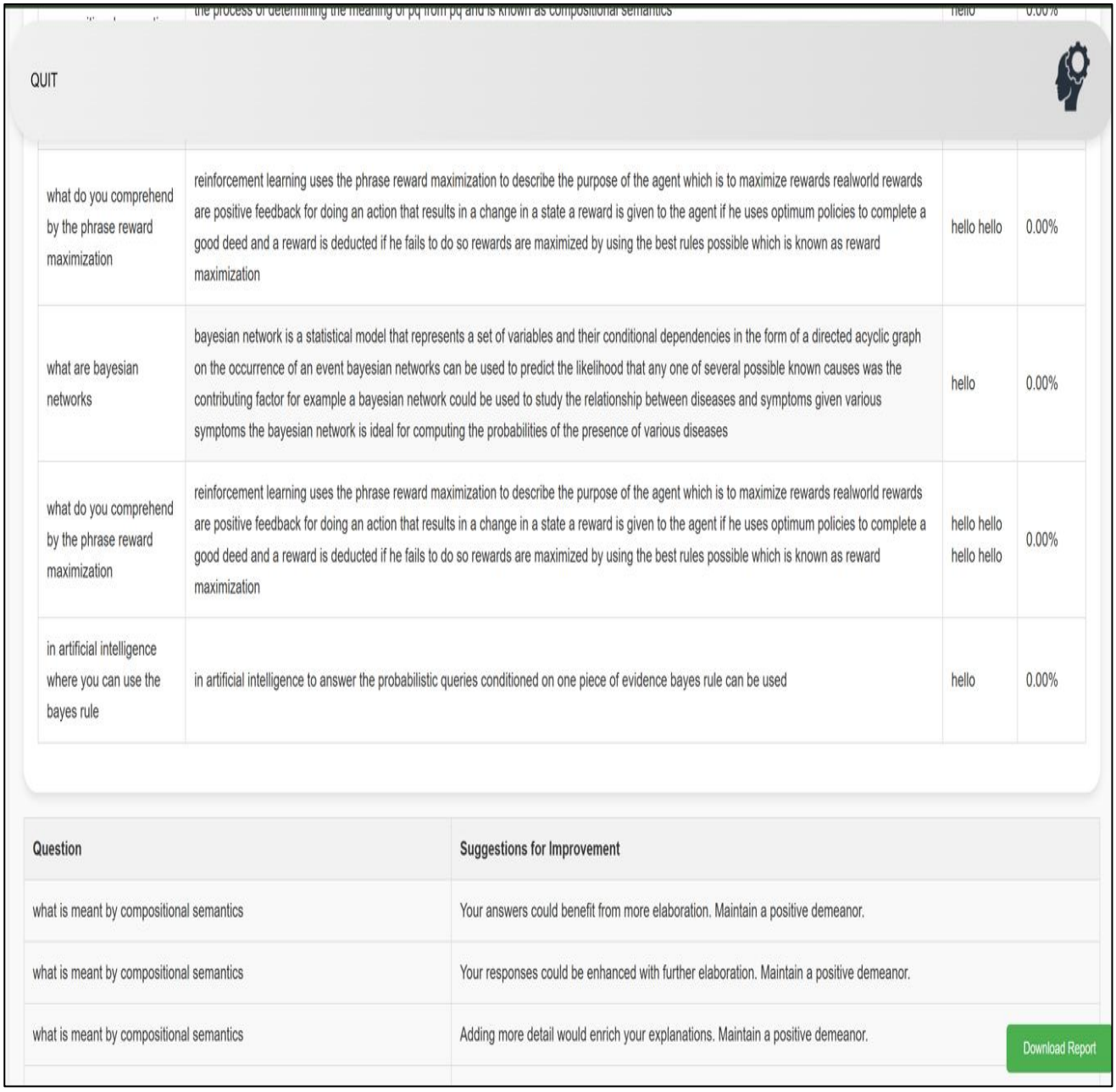


Figure 5.5: Result Page

6. SOFTWARE TESTING

6.1 TYPES OF TESTING:

Testing was a critical part of the project to ensure the system's robustness, accuracy, and performance. The following methodologies were implemented:

1. Unit Testing

Unit testing focused on verifying the functionality of individual components of the chatbot backend and frontend.

1. **Backend Tests:**
    - **Facial Emotion Detection:**
      Tested the facial expression recognition module using the Xception model to ensure correct classification across different emotional states (e.g., happy, sad, angry, neutral) using sample video frames.
    - **Speech-to-Text Conversion:**
      Validated the speech recognition module to confirm accurate transcription of user responses across various accents and audio qualities.
    - **Sentiment and Tone Analysis:**
      Assessed the sentiment analysis logic to verify accurate classification of responses into positive, neutral, or negative tones.
    - **NLP Evaluation:**
      Tested the NLP pipeline for coherence, relevance, and grammar analysis of interview responses, ensuring consistent scoring output.
    - **Score Aggregation Logic:**
      Ensured that the scoring system correctly aggregated facial, verbal, and content-based evaluations into final user performance metrics.

2. **Frontend Tests:**

- **Interview Screen Interface:**
  Verified functionality of the real-time interview screen including webcam, microphone access, and question display.
- **Feedback Display Module:**
  Tested the feedback interface to confirm that users received performance breakdowns, improvement suggestions, and scores accurately after completing the mock interview.
- **Media Stream Handling:**
  Ensured reliable initialization and teardown of webcam and audio streams during interview sessions.
- **Navigation and State Management:**
  Tested route transitions and component states across different interview phases (login → interview → feedback).

2. Integration Testing

Integration testing ensured that the different components of the interview system worked seamlessly together.

1. **Frontend-Backend Communication:**
   - Verified that RESTful API calls initiated from the React frontend (via Axios) were correctly handled by Flask backend routes (e.g., /start-interview, /analyze-response, /get-feedback) and returned JSON responses with appropriate status codes and data structures.
   - Ensured webcam and microphone data captured on the frontend were successfully transmitted and processed by the backend for emotion and voice analysis.

2. **End-to-End Workflow:**
   - Simulated complete user sessions, including logging in, selecting interview topics, participating in the mock interview, receiving emotion and speech analysis, and obtaining personalized feedback.

- Confirmed data consistency and synchronization across multiple modules such as video capture, NLP processing, and feedback generation.
- Validated session handling and state persistence between frontend views and backend data flow.

3. **Facial and Speech Analysis Integration:**
   - Verified that audio and video streams from the frontend were properly received by the backend for concurrent facial emotion recognition and speech-to-text conversion.
   - Ensured accurate merging of audio-visual insights (e.g., emotion + sentiment) for cohesive feedback output.
4. **NLP and Scoring Engine Integration:**
   - Tested integration between the NLP analysis module and scoring logic to ensure that semantic, syntactic, and emotional metrics were accurately interpreted and translated into performance scores.
   - Confirmed that AI-generated improvement suggestions were based on combined input from speech, facial, and textual assessments.

3. Functional Testing

Functional testing validated the SkillSight's features against project requirements.

1. **Interview Question Generation:**
   - Ensured that that the system generated appropriate and relevant interview questions based on the selected topic.
2. **Real-Time Video and Audio Capture:**
   - Validated the system's ability to access the webcam and microphone, capturing high-quality audio and video streams during live interview sessions without interruptions or lag.
3. **Facial Emotion Recognition:**
   - Verified that facial expressions were correctly detected and classified in real-time using the Xception CNN model, and corresponding emotion metrics were accurately recorded for analysis.
4. **Speech-to-Text Transcription:**

- Confirmed that spoken responses were accurately transcribed to text using the integrated speech recognition module across different accents and noise conditions.

5. **NLP-Based Answer Evaluation:**
   - Tested the system's ability to evaluate the content of user responses for clarity, relevance, and coherence using natural language processing algorithms.
6. **Feedback and Scoring Generation:**
   - Verified that the system provided personalized feedback, including performance scores, emotion analysis, and improvement suggestions, immediately after the interview session.
7. **User Interface Functionality:**
   - Checked the frontend components such as interview interface, and feedback page to confirm seamless navigation, responsive design, and real-time data updates.

4. Usability Testing

Usability testing focused on the SkillSight's user interface and experience.

1. **Interface Design:**
   - Evaluated the simplicity, clarity, and navigability of the SkillSight platform's user interface, including the login screen, topic selection, interview session layout, and feedback dashboard to ensure a smooth user experience, even for non-technical users.
2. **Loading Times:**
   - Assessed system responsiveness, including the time taken to load interview questions, initiate webcam/audio streams, and generate post-interview feedback to ensure minimal wait times and uninterrupted user flow.
3. **User Feedback:**
   - Conducted usability sessions with a group of beta testers (students and job seekers) to gather qualitative feedback on overall system usability, perceived performance, and interface comfort.

- Feedback was used to refine design elements such as button placements, color themes, question visibility, and feedback clarity for better user satisfaction.

5. Security Testing

Security testing ensured that user data and interactions were protected.

1. **File Upload Validation:**
   - Ensured that only valid media input (e.g., video/audio from webcam and microphone) was accepted by the system, with strict MIME-type checking and size limits to prevent injection of malicious content or unauthorized file formats.
2. **Data Privacy:**
   - Verified that all sensitive user data—including video recordings, speech transcriptions, and evaluation results—were either encrypted at rest or transmitted securely via HTTPS to maintain user confidentiality and comply with data protection standards.

3. **API Security:**
   - Conducted vulnerability assessments on Flask API endpoints to prevent unauthorized access, injection attacks, and improper data exposure.
   - Implemented authentication checks and input sanitization across all routes, especially those handling session data, emotion analysis, and scoring feedback.

6. Performance Testing

Performance testing was conducted to ensure the chatbot's reliability under varying loads.

1. **Load Testing:**
   - Simulated concurrent user sessions where multiple users participated in mock interviews, uploaded audio/video data, and requested feedback simultaneously to evaluate server throughput and response consistency under normal operational loads.
2. **Stress Testing:**

- Assessed system behaviour under extreme conditions by processing long interview sessions, high-resolution video feeds, and large volumes of audio data to ensure the backend could handle performance degradation gracefully without crashing.

3. **Latency Analysis:**
    - Measured system response times across different stages—question rendering, real-time video analysis, speech-to-text conversion, and feedback generation—to ensure that total latency remained within acceptable bounds for a smooth user experience.

7. Regression Testing

Regression testing ensured that updates and fixes did not break existing functionality.

1. **Continuous Integration:**
    - Used automated test scripts integrated with the CI pipeline (e.g., GitHub Actions) to verify critical functionalities—such as user login, interview flow, and feedback generation—after every code update to ensure that new changes did not break existing features.
2. **Feature Revalidation:**
    - Re-tested core modules including facial emotion recognition, speech transcription, NLP-based scoring, and adaptive feedback after each new deployment or bug fix to ensure previously working functionalities remained stable and accurate.
3. **UI Consistency Checks:**
    - Validated those changes to frontend components (e.g., layout updates, new interface elements) did not impact existing views or user navigation, maintaining a consistent and error-free user experience across sessions.

8. Scalability Testing

Scalability testing assessed the SkillSight's ability to handle increasing user demand.

1. **Concurrent Users:**
    - Tested the system’s ability to handle a high number of simultaneous users by simulating multiple concurrent interview sessions with live audio/video processing

and real-time feedback generation, ensuring stable performance under increased load.

2. **Database Optimization:**
   - Verified that **FAISS-based semantic search** maintained low-latency retrieval times even as the volume of embedded data grew, confirming that vector indexing and retrieval scaled efficiently with larger user datasets and response histories.
3. **Cloud Resource Utilization (Optional):**
   - Monitored system scalability in a cloud-hosted environment by evaluating CPU, memory, and bandwidth usage during peak loads, ensuring that the architecture could be scaled horizontally to accommodate growing user demand.

6.2 TEST CASES AND TEST RESULTS

1. Test Cases and Results for Unit Testing

Test Case	Expected Outcome	Status
Facial expression detection classifies a smiling face.	Detects and classifies the emotion as "happy."	Pass
Speech-to-text module transcribes clear audio.	Converts spoken input into accurate text transcription.	Pass
NLP module evaluates sentence for clarity and grammar.	Identifies grammar issues and rates clarity score correctly.	Pass
Scoring engine computes performance metrics.	Generates correct scores based on input from NLP and emotion analysis.	Pass

Table 6.1: Test Cases and Results for Unit Testing

2. Test Cases and Results for Integration Testing

Test Case	Expected Outcome	Status
Test if the chatbot can access and retrieve data from uploaded PDF files.	Uploaded files should be processed, and relevant content should be used for responses.	Pass
Test if user authentication integrates with session management.	Users should be able to log in and maintain session throughout the chat session.	Pass
Test if video and audio streams are successfully received and processed by backend services.	Video and audio data should be correctly captured and analyzed by facial and speech recognition modules.	Pass
Test integration of emotion analysis with scoring module.	Detected emotions should influence final feedback and scoring.	Pass

Table 6.2: Test Cases and Results for Integration Testing

3. Test Cases and Results for Functional Testing

Test Case	Expected Outcome	Status
Test if facial emotion recognition triggers during the interview session.	The system should detect and classify user emotions in real time.	Pass
Test if the user receives personalized feedback after the mock interview.	The feedback should reflect user performance, including suggestions and a score.	Pass
Test if interview question generation aligns with selected topic.	Topic-specific questions should be generated without duplication or irrelevance.	Pass

Table 6.3: Test Cases and Results for Functional Testing

4. Test Cases and Results for Usability Testing

Test Case	Expected Outcomes	Status
Test if error messages are clear and informative.	Error messages should guide users to resolve issues or understand system behaviour.	Pass
Test if users can easily navigate between different stages (login, topic selection, interview, feedback).	Navigation should be intuitive, with clear labels and smooth transitions between screens.	Pass
Test if users can understand and interpret feedback reports.	Feedback, including scores and suggestions, should be clearly displayed and easy to comprehend.	Pass
Test if icons and buttons are appropriately placed and labelled.	All interactive elements should be clearly marked and accessible for first-time users.	Pass

Table 6.4: Test Cases and Results for Usability Testing

5. Test Cases and Results for Security Testing

Test Case	Expected Outcome	Status
Test if user-uploaded files are securely stored and deleted after the session.	Files should be stored securely and removed immediately after session completion.	Pass
Test if inputs are sanitized to prevent injection attacks.	Inputs should be sanitized to prevent malicious commands or code injection.	Pass
Test if session tokens are encrypted.	Session tokens should be encrypted to protect user sessions.	Pass

Test if unauthorized users are blocked from accessing restricted endpoints.	System should deny access and return appropriate error codes (e.g., 401/403).	Pass
Test if HTTPS is enforced for all data transmission.	All communications between frontend and backend should occur over HTTPS.	Pass
Test if access to video and audio streams is restricted.	Only authenticated sessions should access camera and microphone resources.	Pass

Table 6.5: Test Cases and Results for Security Testing

6. Test Cases and Results for Performance Testing

Test Case	**Expected Outcome**	**Status**
Test response time for video and audio analysis.	Emotion and speech processing should complete within 2–3 seconds after input submission.	Pass
Test file processing speed for a 20 MB video file.	The system should analyze and process the video within 10 seconds.	Pass
Test feedback generation time after interview completion.	Feedback should be generated and displayed within 5 seconds post-interview.	Pass
Test system behavior under peak load.	The application should remain stable and responsive when subjected to maximum traffic.	Pass

Table 6.6: Test Cases and Results for Performance Testing

7. Test Cases and Results for Regression Testing

Test Case	Expected Outcomes	Status
Test if updates to the interview analysis module do not break existing features.	All core functionalities (emotion detection, scoring, feedback generation) should continue to work as intended after updates.	Pass
Test if resolved bugs remain fixed after future code changes.	Previously fixed issues (e.g., UI glitches, score calculation errors) should not reappear after subsequent deployments.	Pass
Test if new feature integrations (e.g., adaptive learning paths) work alongside existing functionalities.	New features should operate without interfering with the existing workflow or degrading performance.	Pass

Table 6.7: Test Cases and Results for Regression Testing

8. Test Cases and Results for Scalability Testing

Test Case	Expected Outcomes	Status
Test facial emotion analysis for 100 video inputs in batch.	The system should process all videos within a reasonable time frame (e.g., <10 seconds per video) without memory overload.	Pass
Test audio transcription on 100 recorded audio clips.	The speech-to-text engine should maintain accuracy and respond within acceptable time (<3 seconds per clip).	Pass
Test feedback generation for 100 interview sessions back-to-back.	The system should consistently generate complete feedback reports with no failures or timeouts.	Pass

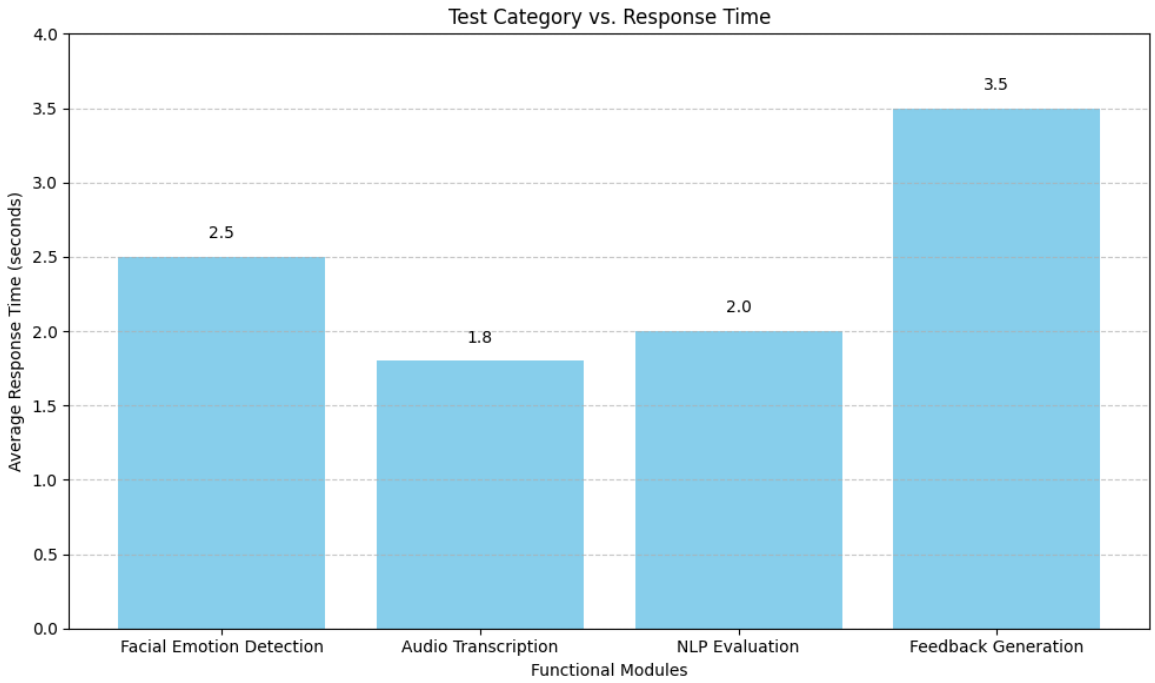
Table 6.8: Test Cases and Results for Scalability Testing

7. RESULTS AND DISCUSSION

7.1 SUMMARY:

The testing and evaluation of the **SkillSight: Interview & Assessment Platform for Soft Skills** project demonstrate that the system performs effectively across functional, performance, usability, and scalability dimensions. The results reflect high reliability, efficient processing, and responsiveness in both controlled and concurrent-use environments.

Graph: Test Category vs. Response Time



The bar chart illustrates the average response times (in seconds) of key functional modules in the **SkillSight** platform, highlighting the system's real-time processing capability.

- **Facial Emotion Detection** recorded an average response time of **2.5 seconds**, reflecting the time taken to analyze user expressions from video input using the Xception CNN model.
- **Audio Transcription** was the fastest module, averaging **1.8 seconds** per input. This efficiency indicates robust speech recognition performance, even under varying audio conditions.

- **NLP Evaluation**, responsible for analyzing the clarity, coherence, and relevance of responses, averaged **2.0 seconds**, indicating reliable language model processing.
- **Feedback Generation** took the longest time, with an average of **3.5 seconds**. This step includes aggregating emotional, verbal, and textual evaluations to produce a personalized assessment and learning suggestions.

Mathematical Analysis

1. Mean Response Time

The overall mean response time across the four modules is:

$$\text{Mean} = \frac{2.5 + 1.8 + 2.0 + 3.5}{4} = \frac{9.8}{4} = 2.45 \text{ seconds}$$

This confirms that the average system component responds in under **2.5 seconds**, well within real-time performance standards.

2. Standard Deviation of Response Times

To assess consistency in performance:

$$\sigma = \sqrt{\frac{(2.5 - 2.45)^2 + (1.8 - 2.45)^2 + (2.0 - 2.45)^2 + (3.5 - 2.45)^2}{4}}$$

$$\sigma = \sqrt{\frac{(0.0025 + 0.4225 + 0.2025 + 1.1025)}{4}} = \sqrt{\frac{1.73}{4}} = \sqrt{0.4325} \approx 0.66 \text{ seconds}$$

A standard deviation of **0.66 seconds** indicates moderate variability, mostly driven by the higher time of feedback generation.

3. Response Time Range

$$\text{Range} = \text{Max} - \text{Min} = 3.5 - 1.8 = 1.7 \text{ seconds}$$

The maximum difference between the fastest and slowest module is **1.7 seconds**, with all values staying within acceptable bounds.

Feature	Existing Systems	Proposed System (SkillSight)
Technology Stack	Traditional computer vision and rule-based NLP (e.g., SIFT, SURF, HOG, basic sentiment analysis)	Advanced AI: Deep learning (CNN for vision), NLP with transformer-based models, sentiment analysis
Accuracy	70–85% based on dataset quality and limited contextual understanding	85–90% accuracy using large-scale datasets, multimodal analysis, and continuous model tuning
Response Time	3–5 seconds per response/image due to sequential processing	Real-time or < 3 seconds with parallelized model inference and optimized algorithms
User Interface (UI/UX)	Basic, often text-heavy with limited user feedback and customization	Modern, interactive UI with real-time feedback, visual indicators, and guided interview paths
Feedback Mechanism	Generic, score-based feedback without detailed improvement suggestions	Personalized feedback highlighting strengths, weaknesses, and suggesting adaptive learning paths
Soft Skills Evaluation	Focused mainly on verbal responses or basic scoring	Comprehensive evaluation of verbal, facial, and behavioural cues for holistic soft skills analysis

Table 7.1: Comparison Table

7.2 CHALLENGES FACED:

During the implementation of the SkillSight, several challenges were encountered:

1. **Real-Time Emotion and Speech Processing**
    - **Challenge:** Capturing and analyzing user facial expressions and voice in real-time without lag or data loss.
    - **Solution:** Optimized OpenCV-based facial detection and used lightweight audio transcription libraries to ensure synchronized and low-latency processing.
2. **Facial Emotion Recognition Accuracy**
    - **Challenge:** Accurately detecting emotions across different facial structures, lighting conditions, and video quality.

- **Solution:** Implemented the Xception CNN model trained on FER2013 dataset, improving detection across varied real-world inputs.

3. **Speech-to-Text Accuracy in Noisy Environments**
    - **Challenge:** Recognizing spoken input accurately when background noise or unclear speech is present.
    - **Solution:** Preprocessing techniques like noise filtering and speech enhancement were applied before transcription, improving the quality of converted text.
4. **Feedback Personalization Logic**
    - **Challenge:** Generating meaningful, user-specific feedback that reflects performance across multiple criteria.
    - **Solution:** Developed a modular scoring system combining verbal content, emotional data, and behavioural cues to generate detailed, personalized recommendations.
5. **UI Responsiveness During Media Capture**
    - **Challenge:** Maintaining a smooth and responsive interface while simultaneously handling webcam and microphone input.
    - **Solution:** React components were optimized to handle asynchronous updates, and UI loaders were added to keep the user informed during heavy processing.
6. **Browser Compatibility for Media Access**
    - **Challenge:** Ensuring webcam and microphone access worked consistently across different browsers and devices.
    - **Solution:** Used standardized browser APIs and included fallback handlers to handle permission denials or unsupported features.
7. **Team Collaboration and Code Synchronization**
    - **Challenge:** Coordinating efforts across different modules (frontend, backend, AI) among team members.
    - **Solution:** GitHub was used for version control with structured pull requests and branching strategies to ensure modular and conflict-free development.
8. **Managing Performance on Mid-Range Systems**
    - **Challenge:** Ensuring smooth performance during testing and interviews on machines with average hardware specs.

- **Solution:** Reduced model complexity where possible and introduced preprocessing steps to reduce real-time computational load.

7.3 FUTURE WORK:

The development of the trainable chatbot opens up several avenues for future enhancements:

1. Multilingual Interview Support

- **Goal:** Expand the platform to support mock interviews in multiple languages.
- **Approach:** Integrating multilingual speech recognition and emotion detection models will enable users to practice interviews in their native languages, improving accessibility and global reach.

2. Emotion Recognition Enhancement

- **Goal:** Improve the accuracy and depth of emotion classification.
- **Approach:** Incorporate advanced deep learning models (e.g., attention-based CNNs or transformer models) and train them on diverse, large-scale datasets to detect subtle and compound emotions (e.g., nervousness, confidence).

3. Real-Time AI Feedback During Interviews

- **Goal:** Provide live feedback to users while they are answering questions.
- **Approach:** Use asynchronous processing and lightweight models to display basic cues like “eye contact,” “confidence level,” or “speech clarity” during the session, allowing users to self-correct in real time.

4. Personalized Progress Tracking Dashboard

- **Goal:** Help users monitor their growth over multiple sessions.
- **Approach:** Create user-specific dashboards that visualize progress across various soft skills (e.g., communication, emotional intelligence) using session histories and performance trends.

5. Integration of Additional Soft Skill Metrics

- **Goal:** Broaden the platform’s evaluation scope to include leadership, teamwork, and adaptability.

- **Approach:** Develop behavioural rubrics and machine learning models trained on labelled interview datasets to infer these additional traits from voice tone, response style, and body language.

6. Mobile Application Development

- **Goal:** Make the platform accessible on smartphones and tablets.
- **Approach:** Build a responsive mobile app that includes core features like mock interview sessions, performance feedback, and learning path suggestions, optimized for mobile processing.

7. Integration with Learning Platforms

- **Goal:** Provide a seamless transition from feedback to learning.
- **Approach:** Integrate the system with online learning platforms (e.g., Coursera, LinkedIn Learning) to offer tailored course recommendations based on user performance.

7.4 POSSIBLE IMPROVEMENT:

While the chatbot system is functional, there are several improvements that could be made:

1. Broader Emotion Detection Capabilities

- **Improvement:** Current facial emotion recognition is limited to basic emotions (e.g., happy, sad, angry).
- **Solution:** Upgrade the emotion detection model to support complex and blended emotions such as nervousness, confusion, and confidence using more diverse training datasets and hybrid deep learning architectures.

2. Improved Audio Clarity and Noise Filtering

- **Improvement:** Background noise can affect the accuracy of speech-to-text transcription.
- **Solution:** Implement more advanced noise suppression techniques or integrate pre-processing filters to enhance voice clarity and transcription accuracy.

3. Faster Feedback Generation

- **Improvement:** Post-interview feedback generation may take several seconds, which affects user experience.

- **Solution:** Optimize the scoring pipeline by reducing model size or implementing lightweight inference techniques for faster processing.

4. Adaptive Interview Complexity

- **Improvement:** All users currently receive the same level of interview difficulty.
- **Solution:** Introduce adaptive questioning based on user performance (e.g., increasing complexity if the user performs well), using dynamic interview generation algorithms.

5. Enhanced UI Feedback Elements

- **Improvement:** The interface could provide more real-time visual cues (e.g., eye contact indicator, timer alerts).
- **Solution:** Include intuitive frontend components such as live progress bars, emotion icons, or speech clarity meters for more engaging and informative interactions.

6. Cross-Platform Optimization

- **Improvement:** Performance may vary slightly across devices and browsers.
- **Solution:** Further optimize frontend components and media access handling to ensure consistent behavior and smooth functionality on all major browsers and platforms.

7. User Guidance and Tutorial Features

- **Improvement:** First-time users may struggle to understand how to interact with the system effectively.
- **Solution:** Add an onboarding tutorial or interactive walkthrough to guide users through setting up the camera/mic and navigating interview features.

APPENDIX A: SURVEY PAPER AND DETAILS

A Comprehensive Analysis of Interview Systems: Introducing Skill Sight - Interview & Assessment Platform for Soft Skills

Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024)

10 Pages
Posted: 21 Dec 2024

Aditya Tighare (https://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=7275638)
Rajarshi Shahu College of Engineering, Pune

Vedant Mete (https://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=7275644)
Rajarshi Shahu College of Engineering, Pune

Proshanjeet Chanda (https://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=7275648)
Rajarshi Shahu College of Engineering, Pune

Sanket Rupnavar (https://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=7275652)
Rajarshi Shahu College of Engineering, Pune

Kavita Patil (https://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=7275655)
Rajarshi Shahu College of Engineering, Pune

Date Written: November 15, 2024

Abstract

SkillSight refers to Interview & Assessment Platform for Soft Skills is a AI-driven mock job interview platform offering a dynamic and immersive experience, simulating real-world interviews to refine crucial soft skills. Currently most of the companies are conducting interviews virtually through online mode. So, it is necessary to develop a system where users can practice for these interviews. With a focus on practical application of chosen interview-topic, system generates diverse interview questions and utilize cutting-edge technologies like facial expressions recognition and natural language processing to assess on various evaluation parameters such as communication skills, emotional intelligence, reaction time and nonverbal cues(body-language). By providing personalized feedback, improvement suggestions and adaptive learning paths, platform help users to excel in interviews and to grow continuously.

A Comprehensive Analysis of Interview Systems: Introducing Skill Sight - Interview & Assessment Platform for Soft Skills

[1]Aditya Tighare, [2]Vedant Mete, [3]Proshanjeet Chanda, [4]Sanket Rupnavar, [5]Kavita Patil

[1, 2, 3, 4, 5]Computer Science and Business Systems, Rajarshi Shahu College of Engineering, Pune, Maharashtra, India.

[1]adityatighare2002@gmail.com, [2]vedantmete@gmail.com, [3]prashanjeetchanda07@gmail.com, [4]sanketrupnavar007@gmail.com, [5]kavitapatil.06@gmail.com

Abstract

SkillSight refers to Interview & Assessment Platform for Soft Skills is a AI-driven mock job interview platform offering a dynamic and immersive experience, simulating real-world interviews to refine crucial soft skills. Currently most of the companies are conducting interviews virtually through online mode. So, it is necessary to develop a system where users can practice for these interviews. With a focus on practical application of chosen interview-topic, system generates diverse interview questions and utilize cutting-edge technologies like facial expressions recognition and natural language processing to assess on various evaluation parameters such as communication skills, emotional intelligence, reaction time and nonverbal cues(body-language). By providing personalized feedback, improvement suggestions and adaptive learning paths, platform help users to excel in interviews and to grow continuously.

Keywords: AI-based interview systems, Natural Language Processing, Facial emotions recognition, Soft skills, Adaptive learning paths.

1. Introduction

"Skill Sight: Interview & Assessment Platform for Soft Skills" we strive to be the bridge between potential and performance. AiJobInterview is an AI-powered mock job interview website that provides users with a dynamic and immersive experience with simulated job interviews in real-world scenarios. Its main aim is closing the gap between theory and practice with the help of this platform, where a user will practice the most needed soft skills. Hence in job market full of competition, the essence of soft skills for success in professional life. SkillSight acts as a guide for anyone who wants to not only satisfy but exceed the expectations of potential employers by utilizing state of the art technology like facial recognition and natural language processing to measure a candidate's communication skills, emotional intelligence, and nonverbal behavior. This tool also prominent as it can analyse the response to varying interview questions and provide users with tips for improvement and tailored feedback. Such a practice not only improves their interview performance but facilitates continuous professional development. With its ease of interface, personalized learning paths, and degree of privacy and security. Skillsight can help to boost people's employability and career prospects by changing the way people prepare for, tackle and overcome the challenges of giving job interviews.

2. Motivation

2.1 For Job Seekers- Practicing Interviews to Improve Soft Skills

Soft skills like communication, problem-solving and adaptability are also important for success in a job setting and are often less developed than hard skills like technical knowledge.

According to a study performed by Harvard University, the Carnegie Foundation, and the Stanford Research Centre [4], on the other hand, 75% of job success over a lifetime is due to mastery of "soft skills" and people skills—not technical skills. However, candidates often find themselves with limited chances to actually practice these skills with an IRL human, putting them at a loss when it comes time for an interview.

This ensures job seekers can consistently vicariously repeat these interpersonal skills to reduce communication errors without worrying about a job behind the wheel, building this skill set slowly but surely. Candidates who get interview practice help develop their confidence, lower their anxiousness during their 'real-world' interviews in future, and help them present stronger performance under stress [17]. This not only gives you a competitive advantage but allows you as a job seeker to blend hard and soft skills.

2.2 Employers- Reducing Interview Rounds and Enhancing Efficiency

For employers, traditional recruitment processes can be time-consuming, especially when evaluating soft skills. A study by the Society for Human Resource Management (SHRM)[19] highlighted that significant portion of cost of hiring is spent on conducting multiple interview rounds. On average, companies conduct 4 to 5 interview rounds before making a hiring decision, which increases costs and prolongs the hiring process [4]. By integrating a virtual mock interview system focused on both soft and hard skills, employers can streamline their hiring processes.

Moreover, pre-assessed soft skills can reduce turnover. According to a study by LinkedIn, 89% of recruiters say that when a new hire doesn't work out, it's often due to poor soft skills, not lack of technical ability [18]. By ensuring that candidates' soft skills are vetted early, employers can make better long-term hiring decisions.

3. Related works

In paper [1], article intends to give a detailed description of the definition of artificial intelligence interview systems. The brief outline of what the AI interview involves is explored to bring forth the principle factors and driving forces of the system. Various AI techniques such as natural language processing, machine learning, and computer vision utilized in the interview systems explained their roles in analyzing candidate responses, assessing non-verbal cues, and evaluating job fit. The mean benefits of the AI interview systems are also discussed, which as the aspects: increased efficiency, reduced-bias, objectivity, and scalability achieved. In addition, potential challenges and ethical implications of deploying AI based systems are considered as well. They include issues of data privacy, fairness of algorithms, and minimizing the human factor. Overall, the review article offers a general understanding of the topic and can be clearly positioned within the other works existing on this subject.

In paper [2], authors elaborate on the mock interview system being developed to allow the users to simulate interviews, specifically tailored towards the virtual interviews held by the companies. Using a deep learning algorithm, it offers feedback on everything from facial expressions and reaction time to speaking rates and grammatical execution. Face reading and grammar parsing are among the combination of multiple technologies applied by system. The system uses OpenCV Facial expression analysis which records user facial expressions while the interview is in progress and analyze it and store it in the specified database using a dataset

imported in the system. This analysis may assist users who are nervous or anxious during an interview since interviewers look closely at candidates' expressions. The facial expression recognition component of the system relies on the concept of deep learning methods (convolutional neural network) as well.

Perspective of Employers According to Al Asefer & Zainal Abidin [3], the soft skills are considered as the significant predictor of employability. In the current job market, employers deem soft skills such as communication, problem-solving, leadership, teamwork, and critical thinking as crucial for a graduate's success. Employers frequently stress these interpersonal skills as being equally if not more important than hard skills in terms of how well a candidate will perform on the job. But there is a wide-mismatch between the soft skills offered by higher education institutions and desired skills by employers, resulting in a skills-skill gap and hence the challenges of employability. However, many graduates remain unemployable, although academically qualified, because they lack these very high-end skills. This also highlights the need for universities to put more emphasis on soft skills training in addition to academics concurring with better preparation for students to step into the global workplace. The paper also points to employer discontent, as most of the new recruits were considered unemployable due to an absence of core soft skills, which speak to the increasing demand for such competencies in the competitive arena.

In their paper Jeremy Lamri and Todd Lubart [4] jointly explore the Generic Skills Component Approach, where they propose an integrated model that unifies the hard and soft skills. It challenges the historic divide between these skills, arguing both kinds include a mixture of similar underlying elements, including both knowledge and the active components of cognition, conation, affection and sense-perceptual motor ability. Unpacking skills into skills trees showcases not only the interdependencies of skills, but the context in which the skill is used which is crucial to understanding how someone is able to checkbox a cert and not be prepared for real-world applications. The model could be used widely through various domains including education, workplace training and productivity, enabling deep understanding of the skill development process. It also applies Hilgard's Trilogy of Mind to discuss how cognitive, motivational and affective elements interweave in skill acquisition, ultimately questioning the traditional splitting of soft and hard skills. It emphasizes both types of skills are important in the workplace and that soft skills (such as empathy, emotional intelligence, and interpersonal skills) are at least as critical as hard skills (technical competencies) for many professions.

In paper [5], the authors highlight the importance of emotions in effective communication and relationship building, emphasizing on technologies that can detect a person's emotions. The study uses the Fer2013 dataset and employs the Tensorflow framework, Keras library, and the Xception Architecture of CNN to train the model for facial emotion detection. The authors also discuss related work in the field, showcasing various approaches utilized for emotion detection, including Bayesian networks, speech segments, and multimodal affect analysis. The document presents the experimental evaluation of different models, including the Convolution Model, Inception Model, and Xception Model, for training, validation, and testing using the Fer2013 Dataset. The Xception Model outperformed the other models, achieving a training accuracy of 93.2% and a validation accuracy of 64.4%, making it the preferred model for real-time emotion detection.

Paper discusses about research by Rodriguez et al. [6] (2022) examined the presence of bias in natural language processing algorithms used to analyse candidate responses. They identified

instances where the algorithms exhibited bias towards certain demographic groups, emphasizing the need for continuous monitoring and improvement to mitigate algorithmic bias. Similarly, in paper [8] Smith and Johnson (2020) investigated the potential bias in computer vision algorithms analysing facial expressions and cautioned against the uncritical use of such algorithms in decision-making. In paper [7], the author discusses about deep learning methods for unconstrained face recognition when there are variations in face image quality. Four deep learning models (VGGFace, Light CNN, CenterLoss, FaceNet) are evaluated on their ability to perform face identification and verification across different quality levels (high, middle, low). The performance is measured on two face datasets (IJB-A, FaceScrub) that have been partitioned into quality levels. The results show that while the models perform well when matching middle to high quality images, their performance drops significantly when matching low quality to high quality images. This suggests that face image quality variation remains a grand challenge, even for state-of-the-art deep learning methods, in unconstrained face recognition.

Tengfei Song et al. [8], A multi-modal physiological emotion database was developed using EEG, GSR, respiration, and ECG signals, aiming to facilitate emotion recognition research. To elicit six discrete emotions and neutral emotion, 28 video clips were selected from over 1,500 options. Emotion classification was achieved using methods like SVM, k-NN, and ALSTM. This database is publicly accessible for further research and development in the field.Mittal, A., Dua, M., Dua, S. [9] discusses how speech incorporates both context and speaker-specific information, making it a valuable human behavioral trait for smart systems. It examines the influence of speech-driven devices, such as those based on speech and speaker recognition, which use various feature extraction and machine learning (ML) techniques. Classical feature extraction methods, particularly Mel Frequency Cepstral Coefficients (MFCC), are highlighted for their effectiveness in speech recognition. The chapter also emphasizes the impact of deep learning (DL), which enhances performance by processing raw waveform data directly for front-end analysis Table 1.

Table 1. Evaluation of Interview Platforms on Key Parameters

Platform	Customization	AI-based Analytics	User Experience
HireVue [10]	High	Advanced	Intuitive
Codility [11]	Moderate	Basic	Good
SparkHire[12]	High	Basic	Easy
Virtuated [13]	High	Advanced	Clean
Outmatch [14]	Moderate	Advanced	Smooth
Vervoe [15]	High	Basic	Intuitive

4. Research Methodology

In developing the SkillSight platform, we use a research methodology that encompasses deep dives into the target audience, multipronged objectives, and lofty goals. That will include managing resources, constraints, planning and strategies, execution, monitoring and progressively improving over time results. Thus, development and amplification of the SkillSight A. The target audience are job seekers:

Equipping job-seekers with the skills and confidence to smash their actual job interviews, leading to higher employability and long-term career success. Becoming the best versions of themselves.

4.1 Assessment

Mimicking what an actual job looks like, a wide variety of interview questions offers a full degree of practice within time limits measuring a range of skills. Responses were assessed for complexity, clarity and relevance.

4.2 Competency Analysis

Sophisticated natural language processing algorithms detected the degree of communicative competence and problem-solving skills through content, the coherence and relevance of responses. Candidates' capabilities were thoroughly assessed.

4.3 Behavioural Analysis

Both verbal and nonverbal elements of interaction are evaluated through the application of cutting-edge technologies including computer vision, sentiment analysis of tone, facial expressions, eye movements and hand gestures. A deeper examination of communication styles aids in generating more accurate projections of interview outcomes.

4.4 Personalized Feedback

Detailed yet personalized feedback is given to candidates, spotlighting strengths and weaknesses noticed during their performances while also offering tailored recommendations for progressing their skills.

4.5 Objective Evaluation

As an AI assistant conducts the interviews, evaluations are performed impartially without prejudices that can unintentionally colour human judgments, ensuring all users receive consistent and unprejudiced appraisals of their performances Table 2.

Table 2. Description of Related Systems/Papers and Gap/Voids

Related System/Paper Name	Description	Technologies Used	Gap/Voids
Google Interview Warmup [16]	Interview tool that analyses responses based on voice and written/typed data	Speech-to-text converter (Transcriber), NLP	1) System only based on voice data, no use of facial or other verbal cues are not considered 2) Interview Warmup was designed specifically for Google career certificate learners, so it has question sets more related to their certificates
Unveiling the Potential of AI: A Comprehensive Review of AI Based Interview Systems [1]	Proposed interview tool analyses user responses based on voice data, facial expression, reaction time, and behavioural patterns	Facial expression recognition, CNN, Speech-to-text, GSL	1) Quality of the answers to the interview questions, decided only on the basis of the non-verbal parameters and competency analysis of answers are seem to be neglected completely. 2) No use of body language parameters for interview analysis like hand gestures, body movement, etc.
RealTimeMock Interview using Deep Learning [2]	Interview System designed for increased efficiency, reduced bias, enhanced objectivity, and improved scalability	NLP,TensorFlow, Keras,Xception, Computer Vision	1) The summary mentions feedback but didn't mention about detailed observations of the interviewee from the attended interview. 2) No provision for anti-cheating and fairness measures.

5. Proposed System

The breakdown of architecture for the interview assessment system with distinct components from the initial interaction to the final recommendation for improvement is given as follows

- **User Authentication and Authorization:** Users initiate the system by logging in, which ensures secure access to the platform. This step validates user identity, ensuring only authorized users can proceed further.

- **Selection of Interview Topic:** After authentication, users are prompted to select a specific interview topic based on their preferences. This could range from technical to soft skills-based interview scenarios. The system provides customized questions based on the chosen topic Fig 1.

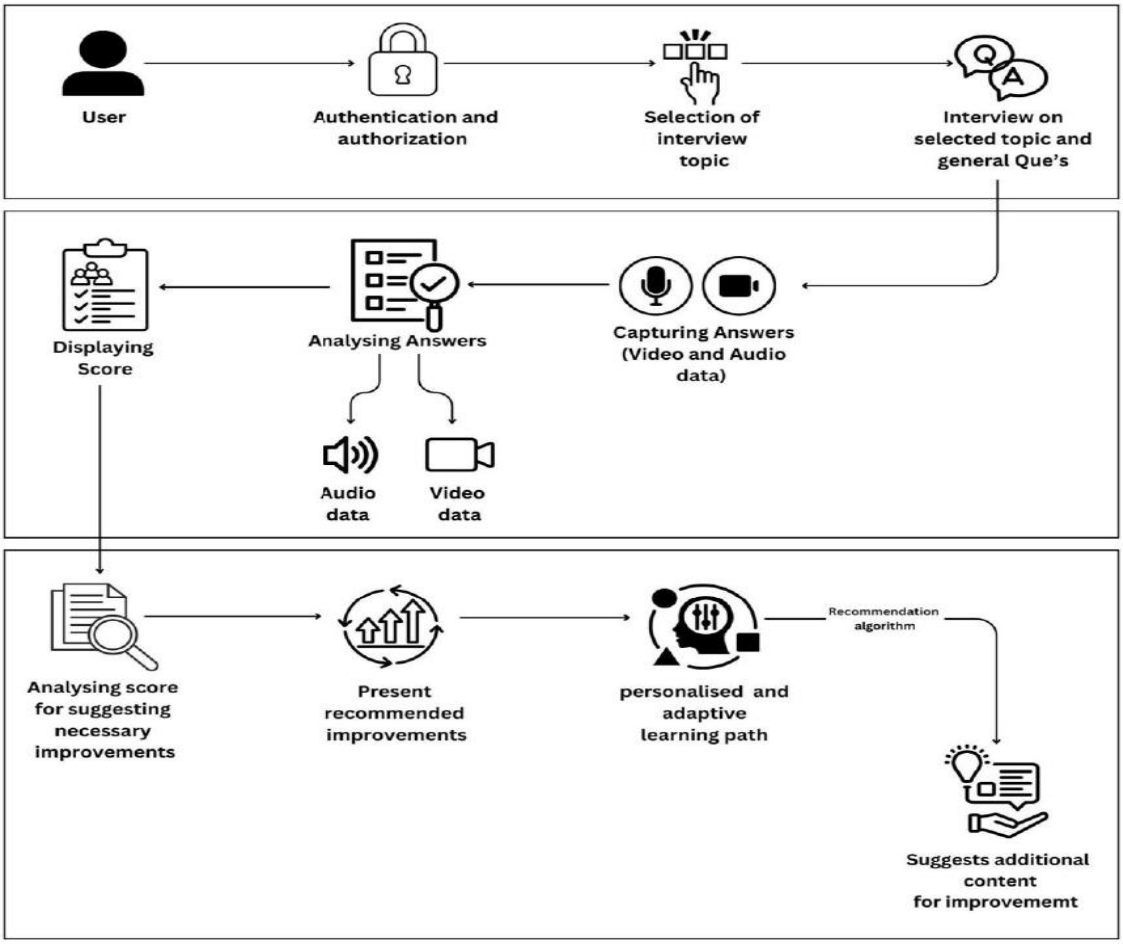


Figure 1. Proposed System Architecture

- **Interview on Selected Topic and General Questions**: The user is presented with both general and topic-specific interview questions. The platform simulates an interview environment where users respond to questions in real-time.
- **Capturing On-Board Answers (Video and Audio Data)**: When the user comes before responding, the system captures audio and video data, which are used for a detailed evaluation. The consideration of both verbal and non-verbal aspects, such as tone, facial expressions, and body language, is then considered.

- **Assessing Output:** The recorded video and sound information is transferred to an assessment module to determine how effectively the task was learned according to established categories. This element is based on both the content of answers (i.e. correct, clear) and non-verbal communication (e.g. eye contact, confidence).
- **POST: Displays score:** A score will be calculated and displayed based on the analyzed data. The score is indicative of the candidate's performance on various factors in the interview such as communication skills, technical knowledge, behavioral responses etc.
- **None of the above;** Next, once the score appears the system analyzes it to provide you with suggestions on improvements. Here lies the importance of personalized feedback to the users guiding them on the skills to work on.
- **Outline Suggest Changes:** This is a part where suggestions will be made keeping in view the analysis. From communication to concrete technical suggestions, etc.
- **Personalized and Adaptive Learning Path:** In a bid to help users learn better, the system generates a personalized learning path. This path is personalized, and is adapted by the recommendation algorithm to the user's weaknesses, allowing them to gradually improve their abilities.
- **Suggest More Resources:** The last component is a follow up with resources to help the user improve in any areas they may want to — links to tutorials, articles, and practice exercises.

6. Limitations and Challenges

Enhancing analysis of soft skills in interview preparation tools using audiovisual data Four studies have been conducted, from which two studies focused on enhancing the annotation quality [2, 3], whereas two studies focused on improving the quality of the assigned annotation labels based on the facial expressions to gain the emotional intelligence and communication [4]. Here are some examples of issues you want to consider as you use it effectively (although there may be potential privacy issues here with correct interpretation).

Looking ahead, AI-based evaluation in the interview preparation tool beneficial as well as challenging in this journey. Despite improvements to scaling, objectivity, and efficiency, AI might struggle with rare but rich, in-depth human interactions 6. That's why it's important to balance automation with human judgement. The key is measuring prejudices, sticking to ethics, and acknowledging the unsurpassed human element in assessment.

Even after having the strongest Security measures, there might still be privacy issues with sensitive information related to job interview.

Yet connectivity problems are critical to solve for the interview prep platform to succeed. Introducing features like offline mode, caching mechanisms, and data compression can improve user experience when connectivity is lower.

7. Conclusion

"SkillSight: Interview & Assessment Platform for Soft-Skills" innovatively cover the critical connection between potential and performance. With AI-powered mock job interviews as well as innovative technologies such as face emotions detection and natural language processing, the platform allows users to have a live experience. It provides jobseekers(interviewee) with personalized feedback, which hones their communication and problem-solving skills. Though there may be issues with scalability and privacy, SkillSight represents a new paradigm for career development. By continuously enhancing user experience, personalizing learning routes, and focusing on user safety, the system transforms the landscape of job interviews, making it a breakthrough in self-development and employability.

References

1. Vaibhav Sharma, Virendra Kumar Nishad, Shubham Chaurasia, Nitin Raj, Karishma Devi," Unveiling the Potential of Ai: A Comprehensive Review of Ai-Based Interview Systems", International Journal of Research and Analytical Reviews (IJRAR), May 2023, Volume 10, Issue 2
2. Sahil Temgire, Akash Butte, Rohan Patil, Varun Nanekar, Shivganga Gavhane," Real Time Mock Interview Using Deep Learning", International Journal of Engineering Research & Technology (IJERT) Published By: Www.Ijert. Org Vol. 10 Issue 05, May-2021
3. Al Asefer M., Zainal Abidin, N.S.," Soft Skills and Graduates' employability In The 21st Century from Employers' Perspectives: A Review of Literature", International Journal of Infrastructure Research and Management Vol. 9 (2), December 2021
4. Lamri, Jeremy, And Todd Lubart. 2023." Reconciling Hard Skills And Soft Skills In A Common Framework: The Generic Skills Component Approach", Journal Of Intelligence 11: 107
5. Noel Jaymon, Sushma Nagdeote, Aayush Yadav, Ryan Rodrigues," Real Time Emotion Detection Using Deep Learning", 2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT) — 978-1-7281-5791-7/20 2021 IEEE — Doi: 10.1109/Icaect49130.2021.9392584
6. Rodriguez, M. D., Garcia, E., & Srinivasan, S. (2022)." Addressing Bias in Ai-Based Interview Systems: A Case Study of Natural Language Processing Algorithms." Proceedings of the ACM on Human-Computer Interaction.
7. Guodong Guo, Na Zhang, "What Is the Challenge for Deep Learning in Unconstrained Face Recognition?" 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (Fg 2018).
8. Tengfe Song, Wenming Zheng, Cheng Lu, Yuanzong, Xilei Zhang1and Zhen Cui" Mped: A Multi-Modal Physiological Emotion Database for Discrete Emotion Recognition" IEEE Access, Vol. 7, Pp. 12177–12191, 2019
9. Mittal, A., Dua, M., Dua, S. (2021). Classical and Deep Learning Data Processing Techniques for Speech and Speaker Recognitions. Signals and Communication Technology. Springer, Cham.
10. Hirevue - Video Interviewing and Recruitment Solutions- Human Resources Company
11. Codility - Online Coding Tests Technical Interviews Platform
12. Sparkhire - Platform for Interview and ATS
13. Vidcruiter - Hiring Platform
14. Outmatch - Hiring Solutions Platform

15. Vervoe - Platform for Skills Assessments
16. Interview Warmup - Grow with Google
17. Jayaram, S., Munge, W., Adamson, B., Sorrell, D., Jain, N. (Eds) Bridging the Skills Gap. Technical and Vocational Education and Training: Issues, Concerns and Prospects, Vol 26. Springer, Cham.
18. Measuring Soft Skills & Life Skills in International Youth Development Programs: A Review and Inventory of Tools", United States Agency for International Development (Usaid).
19. Society for Human Resource Management - Professional Human Resources Membership Association.
20. Du, S., Xie, C. (2020). Paradoxes of Artificial Intelligence in Consumer Markets: Ethical Challenges Opportunities. Journal of Business Research, 129.

Certificate of Presentation

ICOFE- 603-4

CERTIFICATE

of Presentation on 22 & 23 October 2024

This Certificate is Proudly Presented To

SANKET RUPNAVAR

A COMPREHENSIVE ANALYSIS OF INTERVIEW SYSTEMS: INTRODUCING SKILLSIGHT- INTERVIEW & ASSESSMENT PLATFORM FOR SOFT SKILLS

in Conference of Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024) Organized by Department of Computer Science and Engineering at K.S.R. College of Engineering, Namakkal, Tamil Nadu, India and Debre Tabor University, Ethiopia

We appreciate his work and contributions.

Dr. V. Sharmila
HoD /CSE,
K.S.R. College of Engineering
Nammakal, Tamil Nadu.

Prof.(Dr.) Bhupendra Kumar
Accounting & Finance,
International Relation Officer,
Debre Tabor University, Ethiopia.

ICOFE- 603-3

CERTIFICATE

of Presentation on 22 & 23 October 2024

This Certificate is Proudly Presented To

PROSHANJEET CHANDA

A COMPREHENSIVE ANALYSIS OF INTERVIEW SYSTEMS: INTRODUCING SKILLSIGHT- INTERVIEW & ASSESSMENT PLATFORM FOR SOFT SKILLS

in Conference of Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024) Organized by Department of Computer Science and Engineering at K.S.R. College of Engineering, Namakkal, Tamil Nadu, India and Debre Tabor University, Ethiopia

We appreciate his work and contributions.

Dr. V. Sharmila
HoD /CSE,
K.S.R. College of Engineering
Nammakal, Tamil Nadu.

Prof.(Dr.) Bhupendra Kumar
Accounting & Finance,
International Relation Officer,
Debre Tabor University, Ethiopia.

ICOFE- 603-1

CERTIFICATE

of Presentation on 22 & 23 October 2024

This Certificate is Proudly Presented To

ADITYA TIGHARE

A COMPREHENSIVE ANALYSIS OF INTERVIEW SYSTEMS: INTRODUCING SKILLSIGHT- INTERVIEW & ASSESSMENT PLATFORM FOR SOFT SKILLS

in Conference of Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024) Organized by Department of Computer Science and Engineering at K.S.R. College of Engineering, Namakkal, Tamil Nadu, India and Debre Tabor University, Ethiopia

We appreciate his work and contributions.

Dr. V. Sharmila
HoD /CSE,
K.S.R. College of Engineering
Nammakal, Tamil Nadu.

Prof.(Dr.) Bhupendra Kumar
Accounting & Finance,
International Relation Officer,
Debre Tabor University, Ethiopia.

ICOFE- 603-2

CERTIFICATE

of Presentation on 22 & 23 October 2024

This Certificate is Proudly Presented To

VEDANT METE

A COMPREHENSIVE ANALYSIS OF INTERVIEW SYSTEMS: INTRODUCING SKILLSIGHT- INTERVIEW & ASSESSMENT PLATFORM FOR SOFT SKILLS

in Conference of Proceedings of the 3rd International Conference on Optimization Techniques in the Field of Engineering (ICOFE-2024) Organized by Department of Computer Science and Engineering at K.S.R. College of Engineering, Namakkal, Tamil Nadu, India and Debre Tabor University, Ethiopia

We appreciate his work and contributions.

Dr. V. Sharmila
HoD /CSE,
K.S.R. College of Engineering
Nammakal, Tamil Nadu.

Prof.(Dr.) Bhupendra Kumar
Accounting & Finance,
International Relation Officer,
Debre Tabor University, Ethiopia.

APPENDIX B: IMPLEMENTATION PAPER

SkillSight: An AI-Powered Platform for Soft Skills Interview & Assessment

1. Aditya Tighare, Computer Science and Business Systems, Rajarshi Shahu College of Engineering, Pune, India, adityatighare2002@gmail.com
2. Vedant Mete, Computer Science and Business Systems, Rajarshi Shahu College of Engineering, Pune, India, vedantmete@gmail.com
3. Proshanjeet Chanda, Computer Science and Business Systems, Rajarshi Shahu College of Engineering, Pune, India, prashanjeetchanda07@gmail.com
4. Sanket Rupnavar, Computer Science and Business Systems, Rajarshi Shahu College of Engineering, Pune, India, sanketrupnavar007@gmail.com
5. Kavita Patil, Computer Science and Business Systems, Rajarshi Shahu College of Engineering, Pune, India, kavitapatil.06@gmail.com

***Abstract*—The integration of Artificial Intelligence (AI) into recruitment processes has changed traditional hiring methodologies, enabling more subtle assessments of candidates' competencies. This paper presents an AI-driven mock interview system that evaluates both technical proficiencies and soft skills through multimodal analysis. The system employs computer vision techniques to analyze facial expressions and gestures, utilizing datasets like FER+ for emotion recognition. Natural Language Processing (NLP) models, such as BERT, are fine-tuned to assess the semantic relevance and clarity of candidates' responses. Additionally, computational paralinguistics extract vocal features—intonation, pitch, and speaking rate—to infer affective states and communication effectiveness. By integrating these modalities, the system provides comprehensive feedback, translating complex model outputs into actionable insights, thereby facilitating targeted improvements in interviewee performance. This approach not only enhances the objectivity and efficiency of candidate evaluations but also addresses ethical considerations by ensuring interpretability and transparency in AI-driven assessments.**

***Index Terms*—Artificial Intelligence, Mock Interview System, Multimodal Analysis, Facial Expression Recognition, Natural Language Processing, BERT, Computational Paralinguistics, Soft Skills Assessment.**

I. Introduction

In today's competitive job market, the ability to excel in interviews hinges not only on technical expertise but also on mastery of soft skills such as communication, emotional intelligence, and adaptability. Despite their critical importance, traditional interview preparation methods often fall short in addressing these aspects, offering limited scope for personalized feedback and non-verbal communication analysis.

Identify applicable funding agency here. If none, delete this.

To bridge this gap, SkillSight emerges as an innovative, AI-powered platform designed to revolutionize how individuals prepare for interviews. By simulating real-world scenarios, the platform integrates advanced technologies such as Natural Language Processing (NLP), facial expression recognition, and behavioral analysis to assess candidates comprehensively. It provides users with tailored feedback on verbal and non-verbal performance metrics, empowering them to refine their soft skills effectively.

This paper delves into the practical implementation of SkillSight, detailing its system architecture, core methodologies, and technological underpinnings. From real-time video and audio data capture to adaptive learning recommendations, SkillSight leverages state-of-the-art AI solutions to create an immersive and impactful interview practice environment. By addressing challenges such as scalability, data privacy, and real-time processing, the platform demonstrates its potential as a transformative tool for enhancing employability and professional growth.

II. Problem Statement

In the modern job market, excelling in interviews requires a balance of technical knowledge and soft skills such as effective communication, emotional intelligence, and adaptability. However, traditional interview preparation methods fall short in addressing critical aspects of soft skills development. Existing systems often lack personalized feedback, fail to analyze non-verbal communication such as facial expressions and gestures, and do not provide a comprehensive, immersive simulation of real-world interview scenarios.

Additionally, with the growing prevalence of virtual interviews, job seekers face challenges in presenting themselves effectively online, where non-verbal cues play a significant

role in creating a positive impression. Fresh graduates and professionals alike struggle to refine these skills due to the absence of platforms that integrate AI-driven insights and dynamic, adaptive learning.

Employers also face inefficiencies in traditional hiring processes, where assessing candidates' soft skills often requires multiple interview rounds. This increases hiring costs and fails to identify candidates who might struggle due to underdeveloped interpersonal abilities, leading to higher turnover rates.

SkillSight aims to address these challenges by providing an AI-powered solution that offers real-time, data-driven feedback on verbal and non-verbal communication, enabling job seekers to refine their soft skills effectively while reducing inefficiencies for employers.

III. Literature Survey

With the increasing adoption of AI-driven solutions in employability enhancement, various systems have emerged focusing on mock interviews, candidate evaluation, and skill development. These systems utilize advances in deep learning, natural language processing, and multimodal analysis to assess both the technical and non-technical aspects of candidate performance. This section presents a critical review of related works relevant to the proposed system.

Vaibhav Sharma et al. [1] provide a comprehensive review of AI-based interview systems, highlighting the emergence of intelligent systems capable of automating interview evaluations through NLP, facial analysis, and sentiment detection. The review also underlines the need for integrating soft skill evaluation as a complementary metric to traditional answer correctness.

Sahil Temgire et al. [2] proposed a real-time mock interview system using deep learning techniques. Their system focuses primarily on textual analysis and basic emotion recognition from audio input, laying foundational work in AI-based interactive interviews. However, it lacks an adaptive learning component and deeper multimodal fusion.

Al Asefer and Zainal Abidin [3] emphasize the significance of soft skills such as communication, emotional intelligence, and adaptability in enhancing employability. Their findings justify the inclusion of soft skill analysis in automated interview platforms, aligning with the core objective of our system.

Lamri and Lubart [4] introduce a framework for reconciling hard and soft skills under a unified competency model. Their work informs the design of composite scoring systems that evaluate candidates not only on the basis of correct answers but also their delivery, behavior, and emotional state.

Noel Jaymon et al. [5] presented a system for real-time emotion detection using deep learning. Their work on facial expression recognition using CNNs has influenced the facial emotion sub-module in our architecture, particularly in leveraging FER+ and AffectNet datasets for emotion classification.

Rodriguez et al. [6] raise ethical concerns about AI-based interview systems, especially in terms of bias in NLP algorithms. Their study reinforces the need for fairness-aware design in candidate evaluation engines, motivating our use of interpretable features and explainable AI modules.

Guodong Guo and Na Zhang [7] explore the challenges in deep learning for face recognition in uncontrolled environments, providing important guidelines on model robustness and generalization—critical in real-time video interviews under diverse user settings.

Tengfe Song et al. [8] contribute a multimodal physiological emotion dataset that supports the development of systems capable of discrete emotion recognition using video and audio. This inspires our integrated emotion and prosody analysis for richer soft skill inference.

Mittal et al. [9] review classical and deep learning methods for speech and speaker recognition. Their study informs the design of the acoustic processing pipeline within our system, particularly the use of spectrogram-based CNNs and transformer-based acoustic encoders.

In addition to research literature, commercial platforms like HireVue [10] and Google's Interview Warmup [11] showcase industry interest in automated interviews. While these platforms provide large-scale interview simulation, they often lack open scientific evaluation metrics and adaptive feedback mechanisms, which our system aims to provide.

Jayaram et al. [12] edited Bridging the Skills Gap, focusing on addressing the disconnect between industry needs and the skills of job seekers. The volume explores technical and vocational education frameworks to improve training programs and better align them with the evolving job market.

Moreover, policy and training institutions such as the USAID [13] and SHRM [14] have stressed the importance of quantifiable soft skill training. Our system aligns with this vision by offering data-driven, personalized improvement suggestions based on soft skill deficiencies.

Mollahosseini et al. [15] introduced AffectNet, a large-scale dataset for facial expression recognition in real-world settings, annotated for both categorical emotions and continuous dimensions like valence and arousal. The system uses pre-trained CNN models based on AffectNet to detect emotions from video frames during mock interviews. Unlike AffectNet's focus on emotion detection, the current system integrates emotion, gesture, and prosodic analysis for a comprehensive soft skills evaluation.

Devlin et al. [16] developed BERT, a model that enables deep bidirectional context understanding in language, widely used for tasks like sentiment analysis. The system incorporates fine-tuned BERT to assess the accuracy, clarity, and relevance of interviewee responses, adding multi-dimensional evaluation by integrating audio-visual cues and feedback generation.

Schuller and Batliner [17] highlighted the role of vocal features such as pitch, prosody, and speaking rate in interpreting emotional states. The system uses these features,

extracted through MFCCs and pitch analysis, to infer traits like clarity, nervousness, and confidence, enhancing soft skill assessment.

Lipton [18] discussed the importance of model interpretability and user-centric explanations in deep learning. The system includes interpretable feedback modules, translating complex model outputs into actionable advice, such as "Speak with more energy" or "Avoid filler words," to improve usability and educational value.

IV. System Overview

The architecture of the proposed mock interview system is designed to facilitate end-to-end automation of interview simulation, response evaluation, and personalized feedback delivery, with a specific emphasis on soft skill assessment and answer accuracy. The system comprises three primary functional layers: User Interaction, Response Acquisition and Analysis, and Feedback and Personalization.

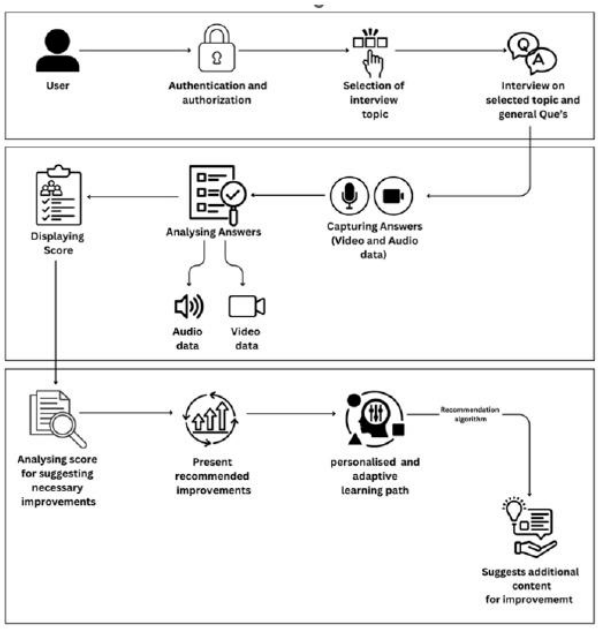


Fig. 1. System Architecture

A. User Interaction Layer

This layer handles the initial engagement with the system. The process initiates with secure authentication and authorization of users, ensuring access control and data integrity. Once authenticated, users proceed to select an interview topic from a predefined set of domains. Subsequently, the system conducts an automated interview session, presenting a mix of domain-specific and general behavioral questions.

B. Response Acquisition and Analysis Layer

During the interview, the system leverages integrated media modules to capture multimodal data, specifically video and audio responses from the user. This data is passed to the analysis module, which incorporates speech recognition and computer vision algorithms to extract and analyze verbal and non-verbal cues. Key soft skill indicators such as tone, pitch, speech clarity, facial expressions, eye contact, and body language are evaluated in conjunction with semantic accuracy of the responses.

The analysis engine computes a composite score, representing the user's performance across communication, confidence, domain knowledge, and behavioral metrics. This score is visualized and presented to the user in a structured format.

C. Feedback and Personalization Layer

Following score generation, the system performs a diagnostic analysis to identify strengths and deficiencies. Based on the evaluation, a recommendation engine suggests targeted improvements and presents a personalized, adaptive learning path. This path is dynamically generated using content-based filtering and rule-based logic that align with the user's specific improvement areas.

To reinforce learning, the system also suggests supplementary resources—such as curated videos, reading materials, and practice exercises—tailored to the user's weak zones. This closed feedback loop ensures continuous self-improvement, mimicking real-world interview preparation.

V. Methodology

This section outlines the detailed methodological pipeline used in the implementation of the mock interview system. The system is modular, and each module contributes to the goal of simulating realistic interview conditions, analyzing both verbal and non-verbal performance, and recommending personalized improvements.

Users interact with the system via a secure web-based interface. Upon successful authentication, they proceed to select a domain from a predefined dataset D=d1,d2,...,dn. Based on the selected topic, the system initiates an interview simulation by generating a sequence of questions Q=q1,q2,...,qk comprising both technical and behavioral types:

1. Data Acquisition Module:

During the interview, real-time audio-visual data is captured through the user's device camera and microphone.

Audio Stream a(t): Time-series of spoken audio input. Video Stream v(t): Frame-wise visual data including facial expressions and gestures.

Both a(t)) and v(t) are timestamp-synchronized and stored as tuples D=(a1,v1),...,(ak,vk) for each question-response pair.

2. Answer Analysis Module:

This module performs multi-modal analysis using signal processing and machine learning models. The analysis includes:

A. Video data analysis:

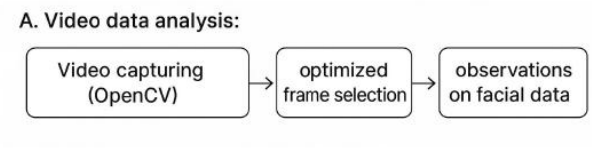


Fig. 2. video analysis module

To extract non-verbal behavioral indicators, the system employs a computer vision pipeline that processes the user's video stream in real time. The procedure involves the following stages:

i. Video Capturing: The system captures continuous video data using the OpenCV library. The video stream is divided into frames vi=f1,f2,...,fn, which are then subject to optimized frame selection strategies to reduce computational redundancy while preserving expressive variance.

ii. Frame Selection: An adaptive frame selection technique is employed to retain frames with significant motion or facial variation. This improves the relevance and accuracy of downstream inference.

iii. Facial Expression and Gesture Recognition: Selected frames are passed through a pre-trained Convolutional Neural Network (CNN) model fcnn(fi) trained on FER+ datasets. The model outputs a vector of probabilities across emotion classes:

$$E_i = f_{\text{CNN}}(f_i) = [p_{\text{happy}}, p_{\text{sad}}, p_{\text{angry}}, \ldots, p_{\text{neutral}}] \quad (1)$$

The highest scoring emotion is selected as the primary facial state of the frame. In parallel, gesture attributes such as head nods, hand movements are inferred using pose estimation techniques.

B. Audio data analysis::

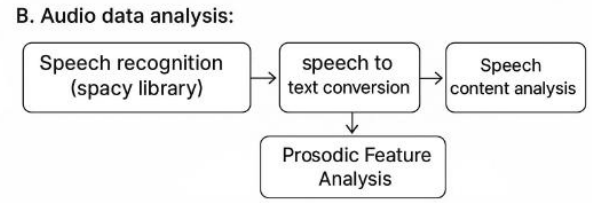


Fig. 3. audio analysis module

i. Speech Content Analysis:

We use an Automatic Speech Recognition (ASR) model $f_{\text{ASR}}(\alpha_i) \rightarrow T_i$ to convert audio α_i into a text transcript T_i. This transcript is semantically evaluated against a ground-truth vector G_i using cosine similarity:

$$\text{Accuracy}_i = \cos(T_i, G_i) = \frac{T_i \cdot G_i}{\|T_i\| \|G_i\|} \quad (2)$$

ii. Prosodic Feature Analysis:

Features such as pitch (f_0), speech rate, energy, and pauses are extracted from the speech signal. A prosodic score is computed using a weighted linear combination:

$$\text{ProsodicScore} = \alpha_1 \cdot \text{PitchVariance} + \alpha_2 \cdot \text{EnergyMean} - \alpha_3 \cdot \text{SilenceDuration} \quad (3)$$

3. Aggregated Soft Skill Scoring:

A weighted fusion of audio and video insights gives an overall soft skill score:

$$\text{SoftSkill} = \beta_1 \cdot \text{ProsodicScore} + \beta_2 \cdot \text{GestureScore} + \beta_3 \cdot \text{FacialEmoti} \quad (4)$$

4. Report Generation: Scoring and Evaluation Engine The final score for each response is computed as:

$$\text{FinalScore}_i = \gamma_1 \cdot \text{Accuracy}_i + \gamma_2 \cdot \text{SoftSkill}_i \quad (5)$$

This score is normalized to a scale of 0–100 and presented to the user through an intuitive dashboard. The dashboard visualizes:

- Individual question scores
- Communication metrics
- Confidence heatmaps over time

5. Feedback and Recommendation Module:

The score vector S=s1,...,sk is passed to a recommendation algorithm fR(S), which performs:

- Error diagnosis: Identify below-threshold scores.
- Learning path generation: Construct a directed acyclic graph (DAG) of learning objectives where each node represents a micro-skill (e.g., "maintain eye contact", "answer concisely").

$$L = f_R(S) = \text{DAG}(M, E) \quad (6)$$

Where M are micro-skills and E are learning dependencies. Content resources C=c1,...,cn are mapped to M using a relevance score computed by semantic similarity between skill descriptions and content metadata

This methodology ensures a seamless flow from user interaction to actionable feedback, enabling users to refine their soft skills effectively.

CONCLUSION

The proposed system implements a multimodal framework for soft skill evaluation during mock interviews by integrating computer vision, speech processing, and natural language understanding techniques. Facial expressions are analyzed using pre-trained CNN model, while vocal attributes such as pitch, energy, and speaking rate are used to infer prosodic and paralinguistic features. Additionally, semantic content of interview responses is assessed using fine-tuned BERT models. These components collectively enable emotion detection, gesture interpretation, and response quality analysis. Interpretable feedback modules convert these analyses into user-friendly suggestions, offering an effective and scalable approach for soft skill development.

FUTURE WORK

Future work will focus on expanding the system's dataset to include more diverse user profiles for improved generalization. Enhancements will also include refining the synchronization between audio-visual modalities and improving the temporal resolution of gesture and emotion tracking. Further development of the feedback module will aim to incorporate more nuanced performance metrics and adaptive feedback strategies based on user history. Additionally, continuous evaluation through user studies will be conducted to assess system effectiveness and usability in real-world training environments.

ACKNOWLEDGMENT

The authors would like to express their gratitude to JSPM Rajarshi Shahu College Of Engineering , Tathawade for supporting this research. Special thanks go to Kavita Patil for their valuable feedback, guidance, and insightful discussions, which greatly contributed to shaping the scope and direction of this survey.

We also thank the broader research community for their foundational work and open discussions, which have provided a robust base for this paper. Finally, we acknowledge the use of resources, tools which were instrumental in conducting this comprehensive review.

REFERENCES

[1] V. Sharma, V. K. Nishad, S. Chaurasia, N. Raj, and K. Devi, "Unveiling the Potential of AI: A Comprehensive Review of AI-Based Interview Systems," *Int. J. Res. Anal. Rev. (IJRAR)*, vol. 10, no. 2, May 2023.

[2] S. Temgire, A. Butte, R. Patil, V. Nanekar, and S. Gavhane, "Real Time Mock Interview using Deep Learning," *Int. J. Eng. Res. Technol. (IJERT)*, vol. 10, no. 05, May 2021.

[3] M. A. Al Asefer and N. S. Zainal Abidin, "Soft Skills and Graduates' Employability in the 21st Century from Employers' Perspectives: A Review of Literature," *Int. J. Infrastruct. Res. Manag.*, vol. 9, no. 2, Dec. 2021.

[4] J. Lamri and T. Lubart, "Reconciling Hard Skills and Soft Skills in a Common Framework: The Generic Skills Component Approach," *J. Intell.*, vol. 11, p. 107, 2023.

[5] N. Jaymon, S. Nagdeote, A. Yadav, and R. Rodrigues, "Real Time Emotion Detection Using Deep Learning," in *Proc. Int. Conf. Adv. Elect., Comput., Commun. Sustain. Technol. (ICAECT)*, 2021, doi: 10.1109/ICAECT49130.2021.9392584.

[6] M. D. Rodriguez, E. Garcia, and S. Srinivasan, "Addressing Bias in AI-Based Interview Systems: A Case Study of Natural Language Processing Algorithms," *Proc. ACM Hum.-Comput. Interact.*, 2022.

[7] G. Guo and N. Zhang, "What is the Challenge for Deep Learning in Unconstrained Face Recognition?" in *Proc. 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, 2018.

[8] T. Song, W. Zheng, C. Lu, Y. Yuanzong, X. Zhang, and Z. Cui, "MPED: A Multi-Modal Physiological Emotion Database for Discrete Emotion Recognition," *IEEE Access*, vol. 7, pp. 12177–12191, 2019.

[9] A. Mittal, M. Dua, and S. Dua, "Classical and Deep Learning Data Processing Techniques for Speech and Speaker Recognitions," in *Signals and Communication Technology*, Springer, Cham, 2021.

[10] HireVue, "Video Interviewing and Recruitment Solutions." [Online]. Available: https://www.hirevue.com/

[11] Interview Warmup - Grow with Google. [Online]. Available: https://grow.google/interview-warmup/

[12] S. Jayaram, W. Munge, B. Adamson, D. Sorrell, and N. Jain (eds.), Bridging the Skills Gap, Tech. and Voc. Educ. and Training, vol. 26, Springer, Cham.

[13] United States Agency for International Development (USAID), "Measuring Soft Skills Life Skills in International Youth Development Programs: A Review and Inventory of Tools."

[14] Society for Human Resource Management. [Online]. Available: https://www.shrm.org/

[15] A. Mollahosseini, B. Hasani, and M. H. Mahoor, "AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild," *IEEE Trans. Affect. Comput.*, 2017.

[16] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proc. NAACL*, 2019.

[17] B. Schuller and A. Batliner, *Computational Paralinguistics: Emotion, Affect and Personality in Speech and Language Processing*, Wiley, 2014.

[18] Z. C. Lipton, "The Mythos of Model Interpretability," *Commun. ACM*, 2018.

APPENDIX C: PLAGIARISM REPORT

Survey Paper

Plagiarism Checker X - Report

Originality Assessment

10%

Overall Similarity

Date: May 9, 2025
Matches: 347 / 3534 words
Sources: 15

Remarks: Low similarity detected, consider making necessary changes if needed.

Verify Report:
Scan this QR Code

v 9.0.3 - WML 4
FILE - CHAPTER 1 - JEET.DOCX

A Comprehensive Analysis of Interview Systems: Introducing Skill Sight - Interview & Assessment Platform for Soft Skills

1Aditya Tighare, 2Vedant Mete, 3Proshanjeet Chanda, 4Sanket Rupnavar, 5Kavita Patil

1,2,3,4,5Computer Science and Business Systems, Rajarshi Shahu College of Engineering, Pune, India.

1adityatighare2002@gmail.com, 2vedantmete@gmail.com, 3prashanjeetchanda07@gmail.com, 4sanketrupnavar007@gmail.com, 5kavitapatil.06@gmail.com

Abstract

SkillSight: Interview & Assessment Platform for Soft Skills is a AI-driven mock job interview platform [1] offering a dynamic and immersive experience, simulating real-world interviews to refine crucial soft skills. Currently most of the companies are conducting interviews virtually through online mode. So, it is necessary to develop a system where users can practice for these interviews. With a focus on practical application of chosen interview-topic, system generates diverse interview questions and utilize cutting-edge technologies like facial expressions recognition and natural language processing to assess on various evaluation parameters such as communication skills, emotional intelligence, reaction time and nonverbal cues(body- language). By providing personalized feedback, improvement suggestions and adaptive learning paths, platform help users to excel in interviews and to grow continuously.

Keywords: AI-based interview systems, Natural Language Processing, Facial emotions recognition, Soft skills, Adaptive learning paths.

1. Introduction

“Skill Sight: Interview & Assessment Platform for Soft Skills” we strive to be the bridge between potential and performance. AiJobInterview is an AI-powered mock job interview

website that provides users with a dynamic and immersive experience with simulated job interviews in real-world scenarios. Its main aim is closing the gap between theory and practice with the help of this platform, where a user will practice the most needed soft skills. Hence in job market full of competition, the essence of soft skills for success in professional life. SkillSight acts as a guide for anyone who wants to not only satisfy but exceed the expectations of potential employers by utilizing state of the art technology like facial recognition and natural language processing to measure a candidate's communication skills, emotional intelligence, and nonverbal behavior. This tool also prominent as it can analyse the response to varying interview questions and provide users with tips for improvement and tailored feedback. Such a practice not only improves their interview performance but facilitates continuous professional development. With its ease of interface, personalized learning paths, and degree of privacy and security. Skillsight can help to boost people's employability and career prospects by changing the way people prepare for, tackle and overcome the challenges of giving job interviews.

2. Motivation

2.1 For Job Seekers- Practicing Interviews to Improve Soft Skills

Soft skills like communication, problem-solving and adaptability are also important for success in a job setting and are often less developed than hard skills like technical knowledge.

According to a study performed by Harvard University, the Carnegie Foundation, and the Stanford Research Centre [4], on the other hand, 75% of job success over a lifetime is due to mastery of "soft skills" and people skills—not technical skills. However, candidates often find themselves with limited chances to actually practice these skills with an IRL human, putting them at a loss when it comes time for an interview.

This ensures job seekers can consistently vicariously repeat these interpersonal skills to reduce communication errors without worrying about a job behind the wheel, building this

skill set slowly but surely. Candidates who get interview practice help develop their confidence, lower their anxiousness during their 'real-world' interviews in future, and help them present stronger performance under stress [17]. This not only gives you a competitive advantage but allows you as a job seeker to blend hard and soft skills.

2.2 Employers- Reducing Interview Rounds and Enhancing Efficiency

For employers, traditional recruitment processes can be time-consuming, especially when evaluating soft skills. A study by the Society for Human Resource Management (SHRM)[19] highlighted that significant portion of cost of hiring is spent on conducting multiple interview rounds. On average, companies conduct 4 to 5 interview rounds before making a hiring decision, which increases costs and prolongs the hiring process [4]. By integrating a virtual mock interview system focused on both soft and hard skills, employers can streamline their hiring processes.

Moreover, pre-assessed soft skills can reduce turnover. According to a study by LinkedIn, 89% of recruiters say that when a new hire doesn't work out, it's often due to poor soft skills, not lack of technical ability [18]. By ensuring that candidates' soft skills are vetted early, employers can make better long-term hiring decisions.

3. Related works

In paper [1], article intends to give a detailed description of the definition of artificial intelligence interview systems. The brief outline of what the AI interview involves is explored to bring forth the principle factors and driving forces of the system. Various AI techniques such as natural language processing, machine learning, and computer vision utilized in the interview systems explained their roles in analyzing candidate responses, assessing non-verbal cues, and evaluating job fit. The mean benefits of the AI interview systems are also discussed, which as the aspects: increased efficiency, reduced-bias, objectivity, and scalability achieved. In addition, potential challenges and ethical implications of deploying AI based systems are considered as well. They include issues of data privacy, fairness of algorithms, and minimizing the human factor. Overall, the review

Plagiarism Checker X - Report

Originality Assessment

7%

Overall Similarity

Date: May 9, 2025
Matches: 200 / 3070 words
Sources: 20

Remarks: Low similarity detected, consider making necessary changes if needed.

Verify Report:
Scan this QR Code

v 9.0.3 - WML 4
FILE - SKILLSIGHT_IMPLEMENTATION_PAPER (2).PDF

SkillSight: An AI-Powered Platform for Soft Skills Interview & Assessment 1. Aditya Tighare, Computer Science and Business Systems, Rajarshi Shahu College of Engineering, Pune, India, adityatighare2002@gmail.com 2. Vedant Mete, Computer Science and Business Systems, Rajarshi Shahu College of Engineering, Pune, India, vedantmete@gmail.com 3. Proshanjeet Chanda, Computer Science and Business Systems, Rajarshi Shahu College of Engineering, Pune, India, prashanjeetchanda07@gmail.com 4. Sanket Rupnavar, 1 Computer Science and Business Systems, Rajarshi Shahu College of Engineering, Pune, India, sanketrupnavar007@gmail.com 5. Kavita Patil, Computer Science and Business Systems, Rajarshi Shahu College of Engineering, Pune, India, kavitapatil.06@gmail.com

Abstract—The integration of Artificial Intelligence (AI) into recruitment processes has changed traditional hiring methodologies, enabling more subtle assessments of candidates' competencies. This paper presents an AI-driven mock interview system that evaluates both technical proficiencies and soft skills through multimodal analysis. The system employs computer vision techniques to analyze facial expressions and gestures, utilizing datasets like FER+ for emotion recognition. 6 Natural Language Processing (NLP) models, such as BERT, are finetuned to assess the semantic relevance and clarity of candidates' responses. Additionally, computational paralinguistics extract vocal features—intonation, pitch, and speaking rate—to infer affective states and communication effectiveness. By integrating these modalities, the system provides comprehensive feedback, translating complex model outputs into actionable insights, thereby facilitating targeted improvements in interviewee performance. This approach not only enhances the objectivity and efficiency of candidate evaluations but also addresses ethical considerations by ensuring interpretability and transparency in AI-driven assessments.

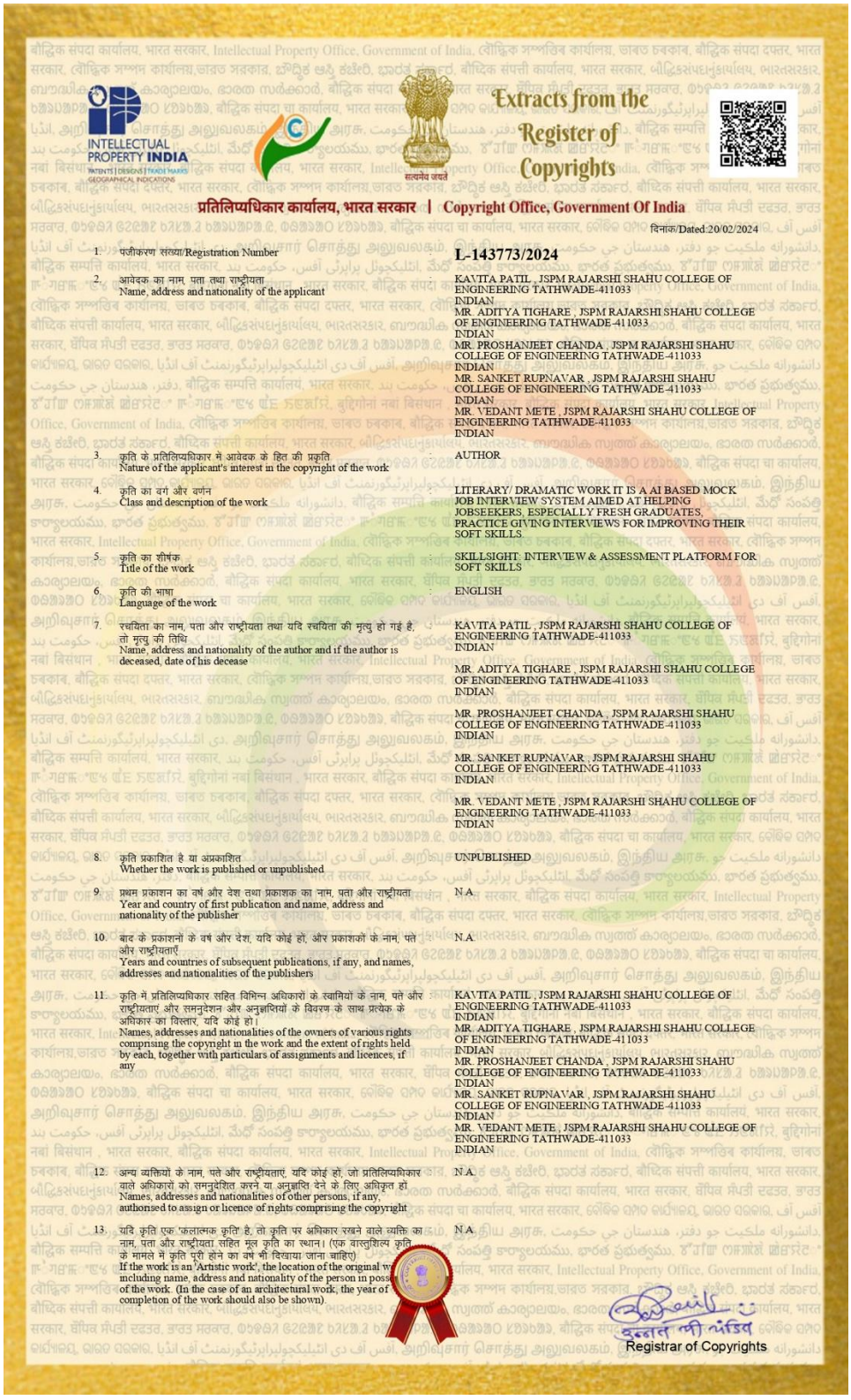
Index Terms—Artificial Intelligence, Mock Interview System, Multimodal Analysis, Facial Expression 2 Recognition, Natural Language Processing, BERT, Computational Paralinguistics, Soft Skills Assessment. I. INTRODUCTION In today's competitive job market, the ability to excel in interviews hinges not only on technical expertise but also on

mastery of soft skills such as communication, emotional intelligence, and adaptability. Despite their critical importance, traditional interview preparation methods often fall short in addressing these aspects, offering limited scope for personalized feedback and non-verbal communication analysis. To bridge this gap, SkillSight emerges as an innovative, AIpowered platform designed to revolutionize how individuals prepare for interviews. By simulating real-world scenarios, the platform integrates advanced technologies such as [6] Natural Language Processing (NLP), facial expression recognition, and behavioral analysis to assess candidates comprehensively. It provides users with tailored [3] feedback on verbal and non-verbal performance metrics, empowering them to refine their soft skills effectively. This paper delves into the practical implementation of SkillSight, detailing its system architecture, core methodologies, and technological underpinnings. From real-time video and audio data capture to adaptive learning recommendations, SkillSight leverages state-of-the-art AI solutions to create an immersive and impactful interview practice environment. By addressing challenges such as scalability, data privacy, and real-time processing, the platform demonstrates its potential as a transformative tool for enhancing employability and professional growth. II. PROBLEM STATEMENT In the modern job market, excelling in interviews requires a balance of technical knowledge and soft skills such as effective communication, emotional intelligence, and adaptability. However, traditional interview preparation methods fall short in addressing critical aspects of soft skills development. Existing systems often lack personalized feedback, fail to analyze nonverbal communication such as [11] facial expressions and gestures, and do not provide a comprehensive, immersive simulation of real-world interview scenarios. Additionally, with the growing prevalence of virtual interviews, job seekers face challenges in presenting themselves effectively online, where non-verbal cues play a significant

role in creating a positive impression. Fresh graduates and professionals alike struggle to refine these skills due to the absence of platforms that integrate AI-driven insights and dynamic, adaptive learning. Employers also face inefficiencies in traditional hiring

processes, where assessing candidates' soft skills often requires multiple interview rounds. This increases hiring costs and fails to identify candidates who might struggle due to underdeveloped interpersonal abilities, leading to higher turnover rates. SkillSight aims to address these challenges by providing an AI-powered solution that offers real-time, data-driven feedback on verbal and non-verbal communication, enabling job seekers to refine their soft skills effectively while reducing inefficiencies for employers. III. LITERATURE SURVEY With the increasing adoption of AI-driven solutions in employability enhancement, various systems have emerged focusing on mock interviews, candidate evaluation, and skill development. These systems utilize advances in deep learning, natural language processing, and multimodal analysis to assess both the technical and non-technical aspects of candidate performance. This section presents a critical review of related works relevant to the proposed system. Vaibhav Sharma et al. [1] provide a comprehensive review of AI-based interview systems, highlighting the emergence of intelligent systems capable of automating interview evaluations through NLP, facial analysis, and sentiment detection. The review also underlines the need for integrating soft skill evaluation as a complementary metric to traditional answer correctness. Sahil Temgire et al. [2] proposed a real-time mock interview system using deep learning techniques. Their system focuses primarily on textual analysis and basic emotion recognition from audio input, laying foundational work in AI-based interactive interviews. However, it lacks an adaptive learning component and deeper multimodal fusion. Al Asefer and Zainal Abidin [3] emphasize the significance of soft skills such as communication, emotional intelligence, and adaptability in enhancing employability. Their findings justify the inclusion of soft skill analysis in automated interview platforms, aligning with the core objective of our system. Lamri and Lubart [4] introduce a framework for reconciling hard and soft skills under a unified competency model. Their work informs the design of composite scoring systems that evaluate candidates not only on the basis of correct answers but also their delivery, behavior, and emotional state. Noel Jaymon et al. [5] presented a system for real-time emotion detection using deep learning. Their work on facial expression recognition using

APPENDIX D: COPYRIGHT CERTIFICATE



INTELLECTUAL PROPERTY INDIA
PATENTS | DESIGNS | TRADE MARKS | GEOGRAPHICAL INDICATIONS

सत्यमेव जयते

Extracts from the Register of Copyrights

प्रतिलिप्यधिकार कार्यालय, भारत सरकार | Copyright Office, Government Of India

दिनांक/Dated:20/02/2024

1.	पंजीकरण संख्या/Registration Number	:	**L-143773/2024**
2.	आवेदक का नाम, पता तथा राष्ट्रीयता Name, address and nationality of the applicant	:	KAVITA PATIL , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN MR. ADITYA TIGHARE , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN MR. PROSHANJEET CHANDA , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN MR. SANKET RUPNAVAR , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN MR. VEDANT METE , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN
3.	कृति के प्रतिलिप्यधिकार में आवेदक के हित की प्रकृति Nature of the applicant's interest in the copyright of the work	:	AUTHOR
4.	कृति का वर्ग और वर्णन Class and description of the work	:	LITERARY/ DRAMATIC WORK.IT IS A AI BASED MOCK JOB INTERVIEW SYSTEM AIMED AT HELPING JOBSEEKERS, ESPECIALLY FRESH GRADUATES, PRACTICE GIVING INTERVIEWS FOR IMPROVING THEIR SOFT SKILLS.
5.	कृति का शीर्षक Title of the work	:	SKILLSIGHT: INTERVIEW & ASSESSMENT PLATFORM FOR SOFT SKILLS
6.	कृति की भाषा Language of the work	:	ENGLISH
7.	रचयिता का नाम, पता और राष्ट्रीयता तथा यदि रचयिता की मृत्यु हो गई है, तो मृत्यु की तिथि Name, address and nationality of the author and if the author is deceased, date of his decease	:	KAVITA PATIL , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN MR. ADITYA TIGHARE , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN MR. PROSHANJEET CHANDA , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN MR. SANKET RUPNAVAR , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN MR. VEDANT METE , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN
8.	कृति प्रकाशित है या अप्रकाशित Whether the work is published or unpublished	:	UNPUBLISHED
9.	प्रथम प्रकाशन का वर्ष और देश तथा प्रकाशक का नाम, पता और राष्ट्रीयता Year and country of first publication and name, address and nationality of the publisher	:	N.A.
10.	बाद के प्रकाशनों के वर्ष और देश, यदि कोई हों, और प्रकाशकों के नाम, पते और राष्ट्रीयताएँ Years and countries of subsequent publications, if any, and names, addresses and nationalities of the publishers	:	N.A.
11.	कृति में प्रतिलिप्यधिकार सहित विभिन्न अधिकारों के स्वामियों के नाम, पते और राष्ट्रीयताएं और समनुदेशन और अनुज्ञप्तियों के विवरण के साथ प्रत्येक के अधिकार का विस्तार, यदि कोई हो। Names, addresses and nationalities of the owners of various rights comprising the copyright in the work and the extent of rights held by each, together with particulars of assignments and licences, if any	:	KAVITA PATIL , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN MR. ADITYA TIGHARE , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN MR. PROSHANJEET CHANDA , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN MR. SANKET RUPNAVAR , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN MR. VEDANT METE , JSPM RAJARSHI SHAHU COLLEGE OF ENGINEERING TATHWADE-411033 INDIAN
12.	अन्य व्यक्तियों के नाम, पते और राष्ट्रीयताएं, यदि कोई हों, जो प्रतिलिप्यधिकार वाले अधिकारों को समनुदेशित करने या अनुज्ञप्ति देने के लिए अधिकृत हों Names, addresses and nationalities of other persons, if any, authorised to assign or licence of rights comprising the copyright	:	N.A.
13.	यदि कृति एक 'कलात्मक कृति' है, तो कृति पर अधिकार रखने वाले व्यक्ति का नाम, पता और राष्ट्रीयता सहित मूल कृति का स्थान। (एक वास्तुशिल्प कृति के मामले में कृति पूरी होने का वर्ष भी दिखाया जाना चाहिए) If the work is an 'Artistic work', the location of the original w... including name, address and nationality of the person in poss... of the work. (In the case of an architectural work, the year of completion of the work should also be shown)	:	N.A.

दस्तावेज़ की जाँच
Registrar of Copyrights

14.	यदि कृति एक 'कलात्मक कृति' है जो किसी भी माल या सेवाओं के संबंध में उपयोग की जाती है या उपयोग किए जाने में सक्षम है, तो आवेदन में प्रतिलिप्यधिकार अधिनियम, 1957 की धारा 45 की उप–धारा (i) के प्रावधान के अनुसार व्यापार चिह्न रजिस्ट्रार से प्रमाणन शामिल होना चाहिए। If the work is an 'Artistic work' which is used or capable of being used in relation to any goods or services, the application should include a certification from the Registrar of Trade Marks in terms of the provision to Sub-Section (i) of Section 45 of the Copyright Act, 1957.	:	N.A.
15.	यदि कृति एक 'कलात्मक कृति' है, तो क्या यह डिजाइन अधिनियम 2000 के अंतर्गत पंजीकृत है? यदि हां, तो विवरण दें। If the work is an 'Artistic work', whether it is registered under the Designs Act 2000, if yes give details.	:	N.A.
16.	यदि कृति एक 'कलात्मक कृति' है, जो डिजाइन अधिनियम 2000 के तहत एक डिजाइन के रूप में पंजीकृत होने में सक्षम है, तो क्या यह औद्योगिक प्रक्रिया के माध्यम से किसी वस्तु पर प्रयुक्त की गई है और यदि हाँ, तो इसे कितनी बार पुनरुत्पादित किया गया है? If the work is an 'Artistic work', capable of being registered as a design under the Designs Act 2000.whether it has been applied to an article though an industrial process and ,if yes ,the number of times it is reproduced.	:	N.A.
17.	टिप्पणी, यदि कोई हो/Remarks, if any	:	

डायरी संख्या/Diary Number: 110/2024-CO/L

आवेदन की तिथि/Date of Application: 02/01/2024

प्राप्ति की तिथि/Date of Receipt: 02/01/2024

Registrar of Copyrights

REFERENCES

[1] Sharma, V., Nishad, V. K., Chaurasia, S., Raj, N., & Devi, K. (2023). *Unveiling the Potential of AI: A Comprehensive Review of AI-Based Interview Systems*. *International Journal of Research and Analytical Reviews (IJRAR)*, 10(2), May.

[2] Temgire, S., Butte, A., Patil, R., Nanekar, V., & Gavhane, S. (2021). *Real-Time Mock Interview Using Deep Learning*. *International Journal of Engineering Research & Technology (IJERT)*, 10(5), May.

[3] Al Asefer, M., & Zainal Abidin, N. S. (2021). *Soft Skills and Graduates' Employability in the 21st Century from Employers' Perspectives: A Review of Literature*. *International Journal of Infrastructure Research and Management*, 9(2), December.

[4] Lamri, J., & Lubart, T. (2023). *Reconciling Hard Skills and Soft Skills in a Common Framework: The Generic Skills Component Approach*. *Journal of Intelligence*, 11(107).

[5] Jaymon, N., Nagdeote, S., Yadav, A., & Rodrigues, R. (2021). *Real-Time Emotion Detection Using Deep Learning*. In *2021 International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)* (pp. 1–6). IEEE. https://doi.org/10.1109/ICAECT49130.2021.9392584

[6] Rodriguez, M. D., Garcia, E., & Srinivasan, S. (2022). *Addressing Bias in AI-Based Interview Systems: A Case Study of Natural Language Processing Algorithms*. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW1), Article 70.

[7] Guo, G., & Zhang, N. (2018). *What Is the Challenge for Deep Learning in Unconstrained Face Recognition?* In *13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)* (pp. 970–976). IEEE.

[8] Song, T., Zheng, W., Lu, C., Zhang, Y., Zhang, X., & Cui, Z. (2019). *MPED: A Multi-Modal Physiological Emotion Database for Discrete Emotion Recognition*. *IEEE Access*, 7, 12177–12191.

[9] Mittal, A., Dua, M., & Dua, S. (2021). *Classical and Deep Learning Data Processing Techniques for Speech and Speaker Recognitions*. In *Signals and Communication Technology*. Springer, Cham.

[10] HireVue. (n.d.). *Video Interviewing and Recruitment Solutions*. Retrieved from https://www.hirevue.com

[11] Codility. (n.d.). *Online Coding Tests and Technical Interviews Platform*. Retrieved from https://www.codility.com

[12] SparkHire. (n.d.). *Video Interviewing Platform and ATS Integration*. Retrieved from https://www.sparkhire.com

[13] VidCruiter. (n.d.). *Hiring Platform for Video Interviews and Recruitment*. Retrieved from https://www.vidcruiter.com

[14] Outmatch. (n.d.). *Hiring Solutions for Talent Assessment*. Retrieved from https://outmatch.com

[15] Vervoe. (n.d.). *Skills-Based Hiring and Assessment Platform*. Retrieved from https://vervoe.com

[16] Grow with Google. (n.d.). *Interview Warmup Tool*. Retrieved from https://grow.google/interview-warmup/

[17] Jayaram, S., Munge, W., Adamson, B., Sorrell, D., Jain, N. (Eds). *Bridging the Skills Gap*. Technical and Vocational Education and Training: Issues, Concerns and Prospects, Vol 26. Springer, Cham.

[18] *Measuring Soft Skills & Life Skills in International Youth Development Programs: A Review and Inventory of Tools*, United States Agency for International Development (USAID).

[19] *Society for Human Resource Management* – Professional Human Resources Membership Association.

[20] Du, S., Xie, C. (2020). *Paradoxes of Artificial Intelligence in Consumer Markets: Ethical Challenges and Opportunities*. *Journal of Business Research*, 129.

[21] McStay, A. (2020). *Emotional AI: The Rise of Empathic Media*. *SAGE Publications Ltd*.