PrepAI: AI-Powered Interview Feedback System

A
Synopsis Report
Submitted in Minor II Synopsis Presentation
for the Degree
of
Bachelor Of Computer Science

by

Mansha Pradumn Soham

Under the supervision of

Dr Sandeep Chand Kumain

School of Computer Science
University of Petroleum and Energy Studies

Contents

Abstract

PrepAI is an AI-based platform that automates the interview preparation process for students through innovative approach and robust anti-cheating technology. The system combines multiple analytical components to deliver personalized feedback on interview performance. Through integrated modules for facial expression analysis, eye tracking, and personality assessment using the OCEAN model. PrepAI provides a general evaluation of both verbal and non-verbal communication skills. It provides a web-based interface through which organizations can carry out interview and analysis candidate's every minor movement, for the candidate, they can get real-time feedback of their interview. Its architecture contains specialized modules for speech-to-text conversion, natural language processing, and detailed report generation. The main distinguishing feature of PrepAI is its complete automation and accessibility for students. Once the question sets are uploaded, the system independently handles the whole interview process, from conducting sessions to generating full feedback. This way, professional interview preparation resources become available to a wider audience of students without compromising the uniformity of assessment.

This strategic focus allows PrepAI to act as an efficient tool for improving the interview skills of objective feedback-seeking students and for organizations seeking a clean interview for candidate selection.

1 Introduction

In today's competitive professional landscape, an efficient, objective, and data-driven interview process is crucial for both candidates and organizations. However, traditional interview methods come with significant challenges, such as scheduling constraints, interviewer subjectivity, inconsistent feedback, and resource-intensive execution. Additionally, candidates often struggle to receive detailed insights into their non-verbal communication, speech patterns, and personality traits, all of which are essential for making a strong impression during an interview. The lack of a standardized, data-driven approach further complicates the process, making it difficult for organizations to fairly and accurately assess candidates' true potential.

PrepAI offers a fully automated, AI-driven interview system that streamlines the entire process, ensuring both efficiency and fairness. By allowing organizations to upload their question sets, PrepAI autonomously manages interview assessments through eye tracking, face detection, voice analysis, and personality evaluation. This multi-dimensional approach helps evaluate a candidate's engagement levels, emotional responses, voice modulation, and personality traits, providing a holistic understanding of their interview performance. Advanced natural language processing (NLP) further enhances the assessment by analyzing content coherence, relevance, and fluency, ensuring a thorough and unbiased review of candidate responses.

By leveraging a multi-modal AI approach, PrepAI guarantees a structured, scalable, and objective evaluation system that eliminates human bias and enhances hiring efficiency. The primary goal of PrepAI is to offer organizations a seamless, end-to-end automated interview experience, while providing candidates with real-time performance feedback based on concrete, measurable parameters. Through automated reporting and comprehensive analysis, PrepAI transforms traditional interview assessments into a scalable, consistent, and insightful process. By bridging the gap between candidates and employers, PrepAI is poised to revolutionize the hiring landscape, ensuring that interviews are more accessible, data-driven, and effective in the ever-evolving job market.

1.1 Motivation

The development of PrepAI is driven by the need for a seamless, fully automated interview experience that eliminates traditional barriers in interview preparation. The current landscape presents challenges such as limited access to structured coaching, particularly for students from underprivileged backgrounds and

smaller institutions, where quality preparation often comes with high costs. Traditional mock interviews rely on subjective feedback that varies between interviewers, making it difficult for candidates to receive consistent, measurable progress insights.

Additionally, the logistical constraints of in-person mock interviews restrict practice opportunities, while most preparation methods focus primarily on verbal responses, neglecting crucial elements like non-verbal communication, voice modulation, and personality assessment.

PrepAI overcomes these limitations by providing an end-to-end automated interview solution, enabling organizations to upload question sets while the system autonomously evaluates eye tracking, facial expressions, voice analysis, and personality traits. By integrating data-driven evaluation and real-time feedback, PrepAI ensures a scalable, unbiased assessment process. Looking ahead, the platform aims to integrate with educational institutions, standardize AI-driven interview preparation, and revolutionize automated professional development, making high-quality interview training accessible to all.

1.2 Project Objectives

Project Objective I: Develop a fully autonomous interview simulation system:

1. The system should function without human intervention, conducting end-to-end interviews autonomously.
2. AI-driven automation ensures a seamless experience for candidates, mimicking real-world interviews.
3. The platform should support various interview formats like technical, behavioural, and case-based interviews.

Project Objective II: Create seamless question delivery and response recording mechanisms:

1. The system should automatically present questions in a structured format based on the candidate's profile and performance.
2. It must record answers in multiple formats (text, audio, and video) to allow in-depth AI-driven analysis.
3. Ensures secure and lag-free recording, preventing data loss or technical glitches.

Project Objective III: Implement real-time feedback generation capabilities:

1. AI should analyse candidate responses in real-time and provide immediate feedback on performance.
2. Instant alerts should be generated if suspicious activities (looking away, external voices, or distractions) are detected.
3. Feedback should be structured to include technical correctness, communication clarity, and behavioural insights.

Project Objective IV: Develop robust voice processing for speech pattern analysis:

1. Assess clarity, confidence, and emotional tone to understand communication effectiveness.
2. Identify signs of nervousness, hesitation, or inconsistency in student.

Project Objective V: Integrate OCEAN personality model assessment:

1. Evaluate Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism based on behavioral and verbal cues.

2. Generate insights into a candidate's problem-solving approach, teamwork ability, and leadership potential.

3. Helps candidates understand their strengths and weaknesses in personality-driven traits.

Project Objective VI: Create clear and comprehensible feedback reports:

1. Generate graphical and textual reports summarizing interview performance.

2. Use visual indicators like progress bars, scores, and heat maps for easy interpretation.

3. Provide actionable feedback with specific improvement areas for candidates to work on.

2 Literature Review

I. In "Best Practices in Eye Tracking Research" (2020), Carter and Luke provide a comprehensive guide for utilizing eye-tracking technology across various disciplines. They begin with an overview of ocular anatomy and physiology, detailing the mechanics of eye movements. The authors then discuss the types of research questions that eye tracking can address, emphasizing its applicability in fields such as psychology, marketing, and human-computer interaction. They explain the operational principles of eye-tracking devices, the nature of data they produce, and offer guidance on selecting appropriate equipment and measures. The paper also highlights the importance of proper calibration, participant positioning, and data interpretation to ensure valid and reliable results. This resource serves as a valuable reference for researchers aiming to implement eye-tracking methodologies effectively. [1]

II. In "A Real-Time Face Tracker" (1996), Yang and Waibel introduce a system capable of tracking human faces in real-time. Their approach combines image processing techniques with predictive modeling to detect and follow facial features dynamically. The system utilizes a camera to capture images, processes these to identify facial regions, and employs motion estimation algorithms to predict and track face movements. This work addresses challenges in face detection and tracking, such as variability in human appearance and motion, and lays foundational principles for subsequent developments in real-time facial recognition technologies. [2]

III. Essa, Darrell, and Pentland's "Tracking Facial Motion" (1994) presents a method for capturing and analyzing facial movements. They propose a system that models facial expressions by tracking the motion of facial features, considering the non-rigid and articulated nature of facial anatomy. The technique involves using computer vision to monitor skin movements and muscle activations, enabling the synthesis of realistic facial animations. This research contributes to the fields of computer graphics and human-computer interaction by providing insights into the dynamics of facial expressions and their computational modeling. [3]

IV. Ryumina et al., in "Gated Siamese Fusion Network Based on Multimodal Deep and Hand-Crafted Features for Personality Traits Assessment" (2024), propose a novel framework for evaluating personality traits. Their approach integrates multimodal data, combining deep learning features with hand-crafted ones, within a Gated Siamese Fusion Network architecture. This fusion allows the model to capture complex patterns across different data modalities, enhancing the accuracy of personality assessment. The study demonstrates the effectiveness of this method in analyzing human behavior, with potential applications in psychology, human resources, and social computing. [4]

V. In "OCEAN-AI Framework with EmoFormer Cross-Hemiface Attention Approach for Personality Traits Assessment" (2024), Ryumina and colleagues introduce the OCEAN-AI framework, which employs the EmoFormer model incorporating a cross-hemiface attention mechanism. This design focuses on analyzing facial expressions by considering the asymmetry between the two halves of the face, thereby capturing subtle emotional cues. The framework aims to assess personality traits more accurately by leveraging these nuanced facial expressions, contributing to advancements in affective computing and automated personality analysis. [5]

VI. Shapiro's article "Speaker Differentiation Using Deep Learning" (2021) discusses the application of deep learning techniques in distinguishing between different speakers. The author explores various neural network architectures and feature extraction methods to improve speaker recognition systems. The study highlights the challenges in speaker differentiation, such as variability in speech patterns and background noise, and presents solutions leveraging deep learning to enhance system robustness and accuracy. This work is pertinent to fields like speech recognition, security, and human-computer interaction. [6]

3 Problem Statement

The traditional interview process is time-consuming, requiring manual intervention at multiple stages, from question preparation to candidate evaluation. Organizations do face challenges related to consistency, the reduction of bias, and the efficient analysis of responses with regard to modality such as speech, facial expressions, or personality traits.

To overcome such issues, PrepAI seeks to deliver a fully automated interview process without any gaps wherein an organization can upload their question set and the system itself handles the rest of the procedure. This solution integrates a combination of eye-tracking, face detection, voice processing, and personality assessments to allow for real-time analysis of the response provided by the candidate. The assessment is supplemented by an LLM, while the automated report generation module summarizes insights for recruiters. PrepAI applies AI-driven multi-modal analysis in order to ensure objective measurements, efficiency improvement, and scalability in an interview solution.

4 Methodology for Data Collection and Model Preparation

To develop an AI-driven automated interview system, PrepAI follows a structured methodology that involves data collection, model selection, and fine-tuning across multiple modalities, including eye tracking, facial recognition, voice analysis, and language processing.

4.1 Site Selection

The system is designed to function in diverse environments, including remote and proctored setups, ensuring accessibility to students and organizations worldwide. The AI models must generalize well across different lighting conditions, backgrounds, and noise levels.

4.2 Data Collection and Model Fine-Tuning

To enhance the accuracy and robustness of PrepAI, data collection is performed across multiple domains:

4.2.1 Video Data Collection

Video data is crucial for eye tracking, face detection, and personality assessment. The following dataset is utilized:

ChaLearn Dataset (UAB Barcelona)[1]: The ChaLearn dataset is widely used for facial expression and personality trait recognition. It contains annotated images and videos capturing various facial expressions, angles, and movements. This dataset helps fine-tune our face detection and personality analysis models by training them to recognize subtle facial cues and behavioral patterns.

4.2.2 Voice Data Collection

Speech processing plays a key role in interview assessments by converting spoken responses into text and analyzing vocal characteristics. The following dataset is used:

Indian Accent Dataset (Kaggle)[2]: This dataset contains audio samples of Indian-accented English speech, allowing our system to fine-tune speech recognition models for improved transcription accuracy. By leveraging diverse linguistic variations, we ensure PrepAI's adaptability to different regional accents.

4.2.3 Model Selection and Fine-Tuning

1. **Eye Tracking:** The `eye.py` module continuously monitors gaze behavior, detecting if the candidate looks away from the screen. If detected three times, a warning is issued, and repeated violations lead to termination. 2. **Face Detection:** The `face.py` module tracks head position and facial angles. If the candidate's face is angled away for longer than a predefined threshold, a warning is triggered. Additionally, it detects multiple faces in the background to prevent impersonation. 3. **OCEAN Personality Model:** Implemented in `OCEAN.py`, this model is trained using the ChaLearn dataset to assess personality traits from facial expressions. 4. **Speech Recognition:** The `transcript.py` module leverages a fine-tuned speech recognition model based on the Indian Accent Dataset to convert spoken responses into text with high accuracy. 5. **LLM Processing:** The `llm_add.py` module applies natural language processing techniques to analyze responses and generate AI-driven feedback.

4.3 File Structure

```
/PrepAI

 main.py                # CLI testing code to integrate all components
 eye.py                 # Eye detection module
 face.py                # Face detection module
 OCEAN.py               # OCEAN personality model implementation
 voice.py               # Voice processing module
 llm_add.py             # LLM processing module
 report.py              # Report generation module
 transcript.py          # Converts speech to text

 templates              # Frontend templates folder
    index.html         # Frontend HTML file
```

[1] https://access.archive-ouverte.unige.ch/

[2] https://www.kaggle.com/datasets/polly42rose/indian-accent-dataset

```
    styles.css          # CSS styles for the frontend
    script.js           # JavaScript for frontend interactivity

 assets                  # Folder for assets like text files
    ques.txt            # Text file with questions

 temp                    # Temporary folder for intermediate files
     response1.txt       # Temporary response file
```

This methodology ensures a structured and efficient approach to data collection, model training, and system implementation, making PrepAI a robust and scalable AI interview assessment tool.

5 System Requirements

To ensure the seamless functioning of **PrepAI**, the system must meet the following Python-based requirements across different components:

5.1 Python Version

Python 3.8+ (Recommended: Python 3.10)

5.2 Required Libraries & Dependencies

5.2.1 Core Dependencies

- **numpy** – Numerical computations
- **pandas** – Data handling and analysis
- **matplotlib / seaborn** – Data visualization
- **joblib** – Model serialization

5.2.2 Computer Vision & Face Detection

- **opencv-python** – Image processing
- **dlib** – Facial landmark detection
- **mediapipe** – Face and eye tracking

5.2.3 Speech & Voice Processing

- **speechrecognition** – Speech-to-text conversion
- **pyaudio** – Real-time audio processing
- **librosa** – Audio analysis
- **transformers** – Pre-trained LLM for voice-based analysis

5.2.4 NLP & Large Language Model Processing

- **transformers** – Hugging Face models for LLM integration
- **torch** – PyTorch backend for deep learning models
- **sentence-transformers** – Sentence embedding generation

5.2.5 Personality Assessment (OCEAN Model)

- **scikit-learn** – Machine learning models
- **tensorflow / torch** – Neural network-based model implementation

5.2.6 Report Generation

- **reportlab** – PDF generation
- **jinja2** – HTML templating for structured reports

5.2.7 Backend & Web Application

- **flask / fastapi** – Web server for handling requests
- **jinja2** – HTML rendering
- **gunicorn** – Deployment in production environments

5.3 System Requirements

5.3.1 Hardware Requirements

- **CPU**: Minimum 4-core processor (Recommended: 8-core for parallel processing)
- **RAM**: Minimum 8GB (Recommended: 16GB for handling large models)
- **GPU**: Optional but recommended (NVIDIA CUDA support for deep learning models)

5.3.2 Storage Requirements

- Minimum **10GB** free space (for storing models, logs, and temporary files)
- SSD recommended for faster data processing

5.3.3 OS Compatibility

- Windows 10 / 11
- Ubuntu 20.04+
- macOS Monterey+ (with support for TensorFlow and PyTorch)

6 SWOT Analysis

A SWOT (Strengths, Weaknesses, Opportunities, and Threats) analysis is essential for evaluating PrepAI's current position in the market. This analysis helps in identifying the system's internal advantages and limitations, as well as external factors that present growth opportunities or risks. Understanding these aspects allows for strategic improvements, ensuring PrepAI remains competitive and aligned with industry needs.

6.1 Strengths (Internal Advantages)

- **Fully Automated** – No human intervention required.
- **Multi-Modal Analysis** – Evaluates facial expressions, voice, and personality.
- **Real-Time Feedback** – Instant AI-driven assessment.
- **Scalable & Accessible** – Available to students anytime, anywhere.

6.2 Weaknesses (Internal Limitations)

- **Limited Focus on Organizations** – Currently student-centric.
- **AI Accuracy Dependence** – Errors in facial, voice, or NLP analysis.
- **Internet & Hardware Requirements** – Needs a stable setup.

6.3 Opportunities (Growth Potential)

- **Integration with Hiring Platforms** – Collaboration with LinkedIn, Naukri, etc.
- **AI-Based Career Coaching** – Personalized interview feedback.
- **Multi-Language Support** – Expanding to global markets.
- **Cloud & Mobile Expansion** – Making PrepAI available on all devices.

6.4 Threats (External Risks)

- **Competition** – Similar AI interview tools exist (e.g., HireVue).
- **Data Privacy Concerns** – Handling sensitive user data.
- **AI Bias Risks** – Potential unfair assessments due to model limitations.
- **User Trust Issues** – Candidates may prefer human feedback.

7 Conclusions, Practical Applications and Scope for Future Works

7.1 Conclusions

PrepAI provides a fully automated and seamless interview experience for students. By integrating eye and face detection, voice processing, personality assessment (OCEAN model), LLM-based evaluation, and speech-to-text conversion, the system ensures a comprehensive candidate assessment.

The project eliminates human bias and manual intervention, making interviews more accessible and scalable for students. The entire process, from question delivery to report generation, is automated, demonstrating AI's potential in recruitment and education. Prep AI's modular approach allows efficient data processing, ensuring adaptability for various interview formats.

7.2 Practical Applications

PrepAI enhances student skill development by improving communication and confidence while also providing AI-driven hiring solutions that automate early-stage recruitment processes. It enables fully remote and unbiased interviews, enhancing accessibility for candidates. Additionally, AI-generated feedback supports continuous training and assessment, helping students refine their performance over time. With personalized interview preparation, Prep AI tailors interview questions and feedback to individual needs, ensuring a more effective and customized experience.

University Mock Interviews: Students can practice real-world interview scenarios with AI-based feedback.

Campus Recruitment: Automates the first-round candidate screening, saving time and resources.

E-Learning & Skill Assessment: Can be integrated into e-learning platforms to track student progress.

Government & Competitive Exams: AI-driven interview rounds for various recruitment processes.

Corporate HR & Talent Acquisition: Streamlines hiring evaluations while reducing costs.
Psychological & Behavioral Analysis: Prep AI can analyze behavior, confidence, and personality traits.

7.3 Scope for Future Research

Donec vehicula augue eu neque.Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris.

- Organization-Side Dashboard: A feature-rich portal for customizing interviews and accessing analytics.
- AI-Based Feedback & Coaching: Real-time feedback with performance improvement recommendations.
- Multimodal Emotion Detection: Analyzing facial expressions, voice tone, and body language.
- Integration with Hiring Platforms: Collaboration with job portals like LinkedIn, Naukri, and Indeed.
- Multi-Language Support: Expanding Prep AI to support multiple languages for global accessibility.
- Cloud-Based AI Processing: Scaling to handle high-volume assessments with cloud integration.
- Bias Mitigation Ethical AI: Ensuring fairness in AI by reducing biases related to gender, accent, and background.
- Real-Time Proctoring: Advanced cheat detection mechanisms using AI-driven monitoring.
- Future developments will focus on enhancing adaptability, efficiency, and personalization in AI-driven interview systems. As technology advances, Prep AI can redefine automated interviews, bridging gaps between education and employment while making assessments more inclusive and intelligent.

References

[1] B. T. Carter and S. G. Luke, "Best practices in eye tracking research," *Pattern Recognition Letters*, vol. 136, pp. 50–56, 2020.

[2] J. Yang and A. Waibel, "A real-time face tracker," in *Proceedings Third IEEE Workshop on Applications of Computer Vision. WACV'96*, 1996, pp. 142–147.

[3] I. Essa, T. Darrell, and A. Pentland, "Tracking facial motion," in *Proceedings of 1994 IEEE Workshop on Motion of Non-rigid and Articulated Objects*, 1994, pp. 36–42.

[4] E. Ryumina, M. Markitantov, D. Ryumin, and A. Karpov, "Ocean-ai framework with emoformer cross-hemiface attention approach for personality traits assessment," *Expert Systems with Applications*, vol. 239, p. 122441, 2024.

[5] ——, "Gated Siamese Fusion Network based on Multimodal Deep and Hand-Crafted Features for Personality Traits Assessment," *Pattern Recognition Letters*, vol. 185, pp. 45–51, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167865524002071

[6] D. Shapiro, "Speaker differentiation using deep learning," *Towards Data Science*, December 2021. [Online]. Available: https://towardsdatascience.com/speaker-differentiation-using-deep-learning-68b2dede498f?gi=ca7a52c127bf