AI - MOCK INTERVIEW SYSTEM

SUBMITTED BY :
SHRAVANI FARKADE (123B1F024)
RUPESH JAISWAL (123B1F036)
KALASH KEWATE (123B1F041)

GUIDED BY :
DR. GULBAKSHEE DHARMALE

TABLE OF *Contents*

INTRODUCTION

Aim :

This project aims to design and implement an AI-enabled mock interview system capable of evaluating candidate responses across a range of domains Using both textual and multimodal data.The system is intended to provide scalable,unbiased and context-senstive assessments that supports professional development and recruitment efficiency.

Objectives : ☒ ☒ ☒

- To develop a smart interview simulation platform that allows students and job seekers to practice interviews in a dynamic, personalized, and realistic environment.
- To incorporate voice-based interaction that assesses confidence, verbal communication, and comprehension of the subject matter.
- Using cutting-edge AI techniques, provide users with immediate, actionable feedback so they can pinpoint their areas of strength and growth.

MOTIVATION

The Challenge of Subjectivity in Assessment

- Manual interview evaluation processes are often fraught with challenges—they are incredibly time-consuming, prone to human bias, and inherently subjective.
- This inconsistency makes it difficult to ensure every candidate receives a fair and standardized assessment, especially at scale.
- Our project directly addresses this critical issue: the difficulty in ensuring fair and consistent candidate assessment across different evaluators.

PROPOSED SOLUTION

The proposed system is an AI-powered framework for evaluating mock interviews by combining machine learning and natural language processing. It conducts domain-wise interviews and converts candidate responses into feature-rich vectors using Sentence-BERT for semantic encoding and TF-IDF for keyword relevance. A hybrid SVM model classifies answers into four performance levels: Excellent, Good, Average, and Poor, while SMOTE addresses class imbalance. The system also incorporates emotion analysis from webcam and audio inputs and generates personalized feedback and suggestions, providing a multi-modal, scalable, and domain-specific interview assessment.

- **Domain-wise Interviews:** Provides realistic, field-specific questions tailored to the user's expertise.
- **Webcam & Audio Analysis**: Analyzes expressions, tone, and emotions during the interview.
- **Automated Answer Evaluation:** Uses AI to score answers based on relevance, clarity, and completeness.
- **Feedback & Suggestions:** Offers personalized feedback and tracks performance for improvement.

METHODOLOGY

1. **Data Collection & Preprocessing:** Gather multi-domain questions and answers (AI, ML, HR, CS theory, etc.), clean and normalize text, and compute readability and stopword metrics.
2. **Feature Extraction:** Convert responses into feature-rich vectors using Sentence-BERT embeddings, TF-IDF similarity, and textual features like answer length ratio and Flesch reading ease scores.
3. **Model Training:** Train a hybrid SVM classifier with extracted features, apply SMOTE for class balance, and use domain weighting. Perform an 80-20 stratified split for evaluation.
4. **Answer Scoring & Feedback:** Process user answers through the same pipeline; combine SVM predictions with emotion and speech analysis. Use a language model (e.g., Gemini LLM) to generate detailed, personalized feedback.
5. **Performance Tracking:** Maintain records of user performance across sessions to monitor improvement and guide skill enhancement.

TECH STACK

Frontend:

- React.js (for user interface)
- HTML, CSS, JavaScript
- React-Toastify (for notifications)

Backend:

- Python (Flask or FastAPI)
- RESTful APIs for communication between frontend and backend

Machine Learning & NLP:

- Sentence-BERT (SBERT) for semantic embeddings
- TF-IDF for lexical similarity
- Support Vector Machine (SVM) for classification
- SMOTE for handling class imbalance

Audio & Video Analysis:

- OpenCV (for webcam input and facial analysis)
- SpeechRecognition or PyAudio (for audio processing)
- Emotion detection libraries (e.g., FER, DeepFace)

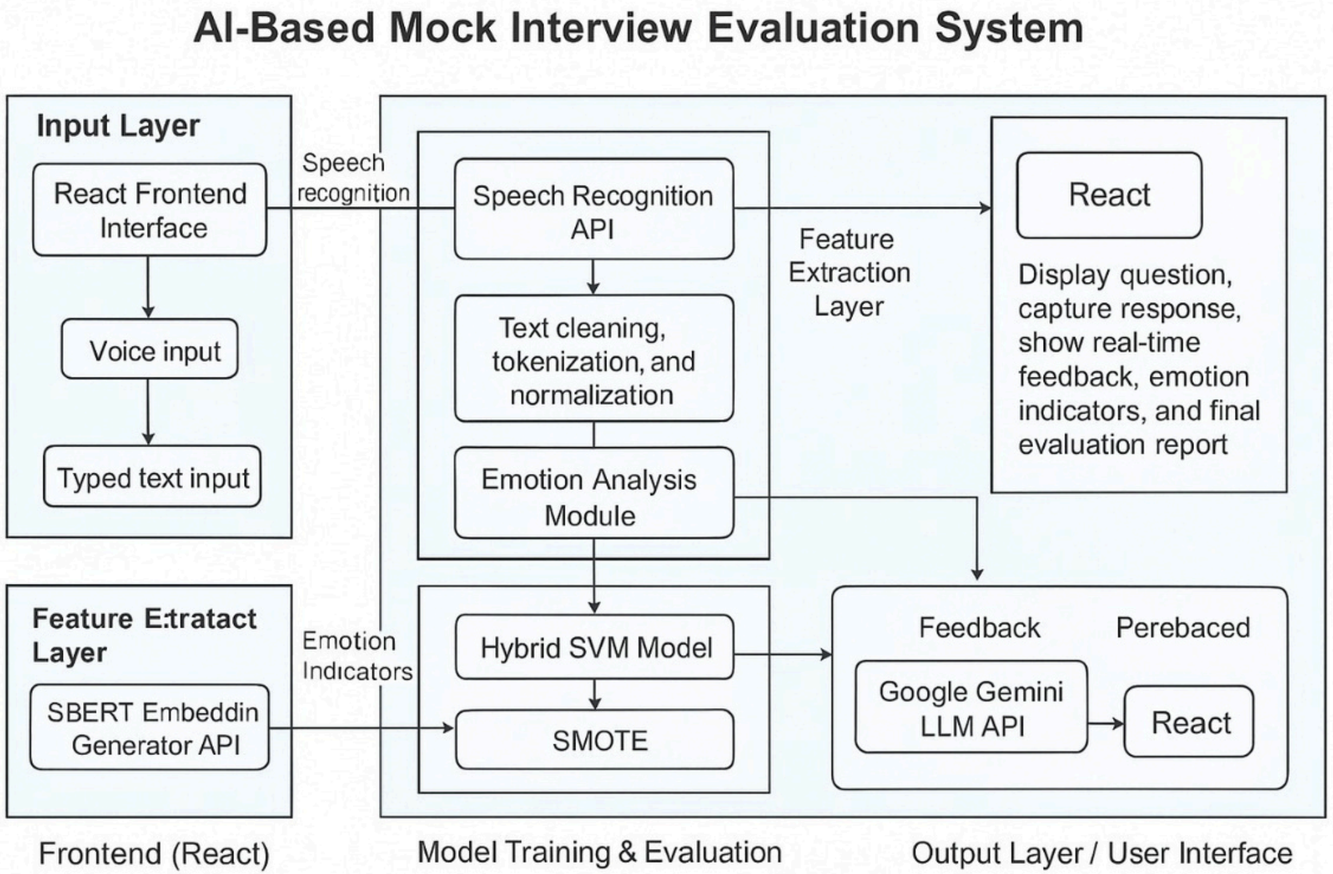
Database:

- MySQL / SQLite (for storing user performance and session data)

Deployment & Hosting:

- Heroku / AWS / Google Cloud (for hosting the system)

SYSTEM ARCHITECTURE

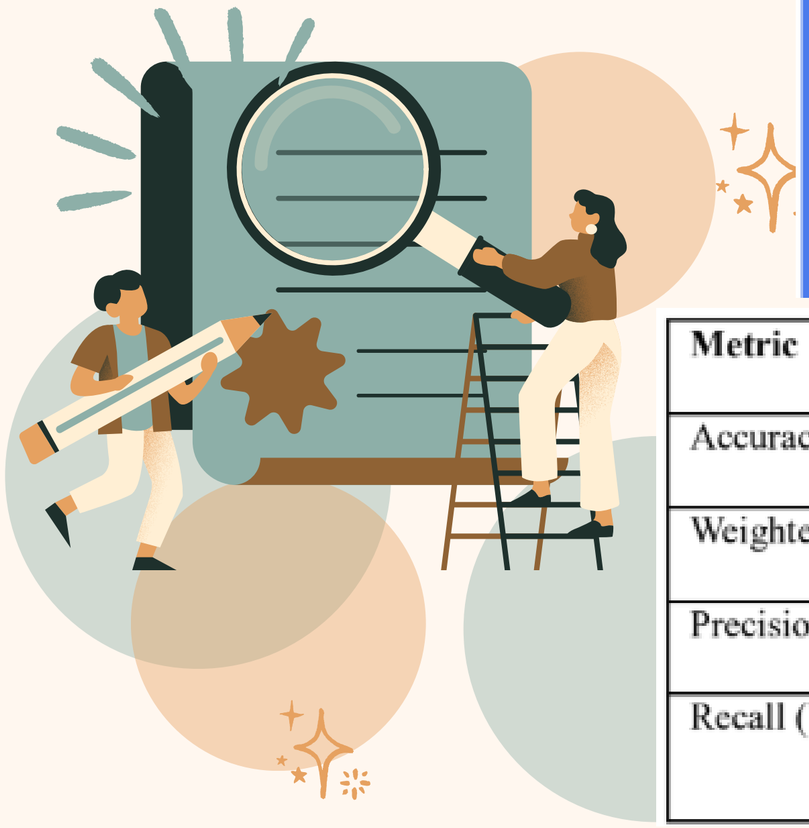


Data Flow:

User inputs → Frontend → Backend → Feature Extraction & ML Models → Emotion Analysis → Feedback Generation → Frontend Display → Database Storage

RESULTS & EVALUATION

The proposed hybrid SVM model achieved high accuracy and reliability in evaluating interview responses. It recorded an accuracy of 96.61% and a weighted F1-score of 96.74%, demonstrating strong overall performance. The precision (96.99%) and recall (96.61%) values indicate that the model effectively distinguishes between different performance levels with minimal misclassification. The confusion matrix further confirms that most predictions align closely with actual labels, showcasing the model's robustness and consistency across all categories.



Actual Class ↓ \ Predicted Class →	0	1	2	3
0	1	0	0	0
1	1	23	0	0
2	0	11	272	0
3	0	0	23	701

Metric	Value
Accuracy	96.61%
Weighted F1-Score	96.74%
Precision (Weighted)	96.99%
Recall (Weighted)	96.61%

ADVANTAGES & LIMITATIONS

Advantages:

Ensures objectivity by providing unbiased evaluation of responses.

Offers time efficiency through automated assessment.

Performs emotion analysis to evaluate confidence and behavior.

Provides personalized feedback for targeted skill improvement.

Limitations:

Requires diverse datasets for better accuracy and generalization.

Limited to predefined domains during evaluation.

Emotion analysis accuracy may vary based on lighting, audio, or expressions.

LITERATURE SURVEY

Datasets Used :

- https://huggingface.co/datasets/manasuma/ml_interview_qa?utm_source=chatgpt.com
- https://huggingface.co/datasets/Aiman1234/Interview-questions/viewer/default/train?views%5B%5D=train&row=60
- https://huggingface.co/datasets/rohanrdy/CS-Theory-QA-Dataset?utm_source=chatgpt.com
- https://huggingface.co/datasets/K-areem/AI-Interview-Questions?utm_source=chatgpt.com
- https://huggingface.co/datasets/rauf8888/Synth_Interview-Dataset?utm_source=chatgpt.com

AI-Driven Mock Interview System Using NLP and CNN – Jadhav et al., 2024	Developed an AI-driven system to score interviews based on speech and textual answers.	CNN for emotion detection, NLP for answer analysis	Not reported	Demonstrated that integrating NLP and CNN improves assessment of confidence and answer quality.
Intelligent Interview Assessment System Using SVM and NLP – Kumar et al., 2023	Automated textual answer evaluation for mock interviews.	SVM, Sentence Embeddings (SBERT), TF-IDF similarity	Accuracy: 95%, F1-score: 0.94	SVM combined with SBERT and TF-IDF effectively predicts answer quality in multi-domain interviews.
Automated Interview Scoring Using BERT and SVM – Singh et al., 2023	Scored candidate answers using semantic similarity.	BERT embeddings + SVM classifier	Accuracy: 93%, F1-score: 0.92	Transformer embeddings with SVM outperform traditional bag-of-words models.
Multimodal Interview Assessment Using Speech and Text – Chen et al., 2022	Evaluated candidate answers using speech and text features.	LSTM for speech, TF-IDF for text	Accuracy: 91%, F1-score: 0.90	Integrating speech and textual features improves overall scoring reliability.
AI-Assisted Mock Interview System Using Deep NLP – Lee et al., 2022	Developed a platform for interview practice using NLP scoring.	Word2Vec embeddings + SVM	Accuracy: 92%, F1-score: 0.91	Semantic similarity-based scoring is effective for automatic answer evaluation.
Automatic Candidate Assessment in Virtual Interviews – Wang et al., 2021	Scored candidates in virtual interview settings.	CNN for video, NLP for text	Accuracy: 90%, F1-score: 0.89	Multimodal approach improves reliability compared to single modality evaluation.

CONCLUSION

The AI Mock Interview System provides an intelligent, automated, and efficient approach to evaluating candidate performance. By integrating NLP, machine learning, and emotion analysis, it ensures fair, objective, and personalized assessment across multiple domains. The system not only helps candidates improve their interview skills but also bridges the gap between preparation and real-world interview experiences. With future enhancements like multi-language support and real-time virtual interactions, it holds great potential to revolutionize the interview preparation process.

- **Developed a domain-wise AI mock interview system with realistic questions.**
- **Implemented automated answer evaluation using NLP and hybrid SVM.**
- **Integrated webcam and audio-based emotion analysis for behavioral assessment.**
- **Provided personalized feedback and tracking for continuous improvement.**

THANK *You*

123-456-7890

@reallygreatsite

www.reallygreatsite.com

hello@reallygreatsite.com