Innovation partner H2S HACK2SKILL

Media partner YOURSTORY

AI for Bharat Hackathon

Powered by aws

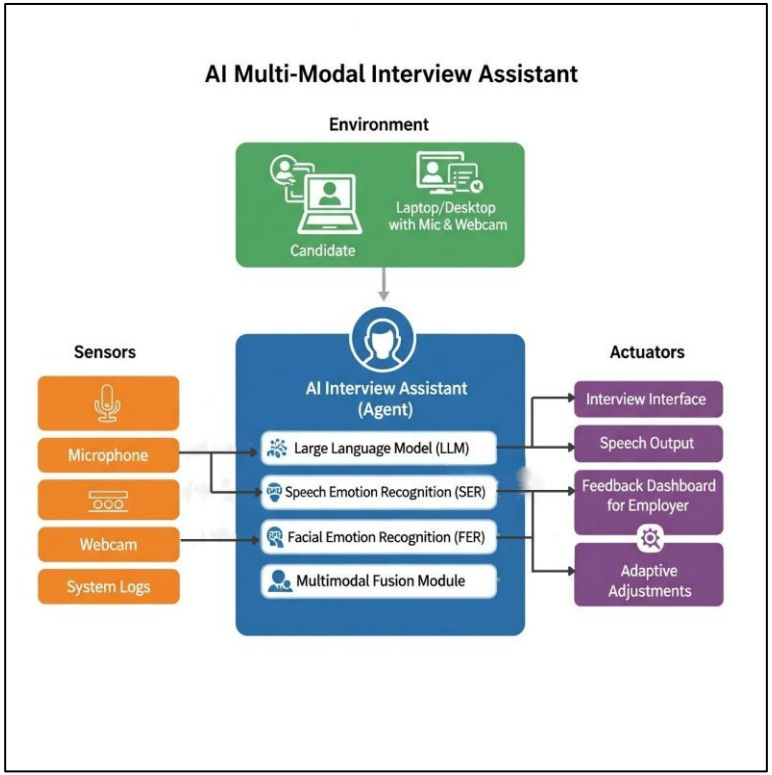
Team Name : Status Code 404

Team Leader Name : Saon Dey

Problem Statement : AI Multimodal Interview Assistant

Brief about the Idea:

The AI Multi-Modal Interview System functions as an intelligent agent operating within a mock or real interview environment. It interacts with the candidate using sensors such as a camera and microphone to capture facial expressions, eye movements, voice tone, and speech patterns. Through Speech Emotion Recognition (SER) and Facial Emotion Recognition (FER), the system analyzes emotions, confidence, and engagement levels. Based on this analysis, a Large Language Model (LLM) generates adaptive follow-up questions and adjusts the interview flow. It also provides feedback and detects distractions. This helps students perform better in real interviews.



Difference from Existing Solutions:

- Multimodal Integration: Unifies LLM, Speech Emotion Recognition (SER), and Facial Emotion Recognition (FER)
- Adaptive Questions: Real-time follow-up questions based on emotional and verbal cues, not static scripts

How the Proposed Solution Solves the Problem:

- Real-time Feedback: Detects nervousness, distractions, and weak points during mock interviews
- Personalized Practice: Adapts difficulty and question types based on performance patterns
- Actionable Insights: Detailed reports on confidence, emotional control, and content quality

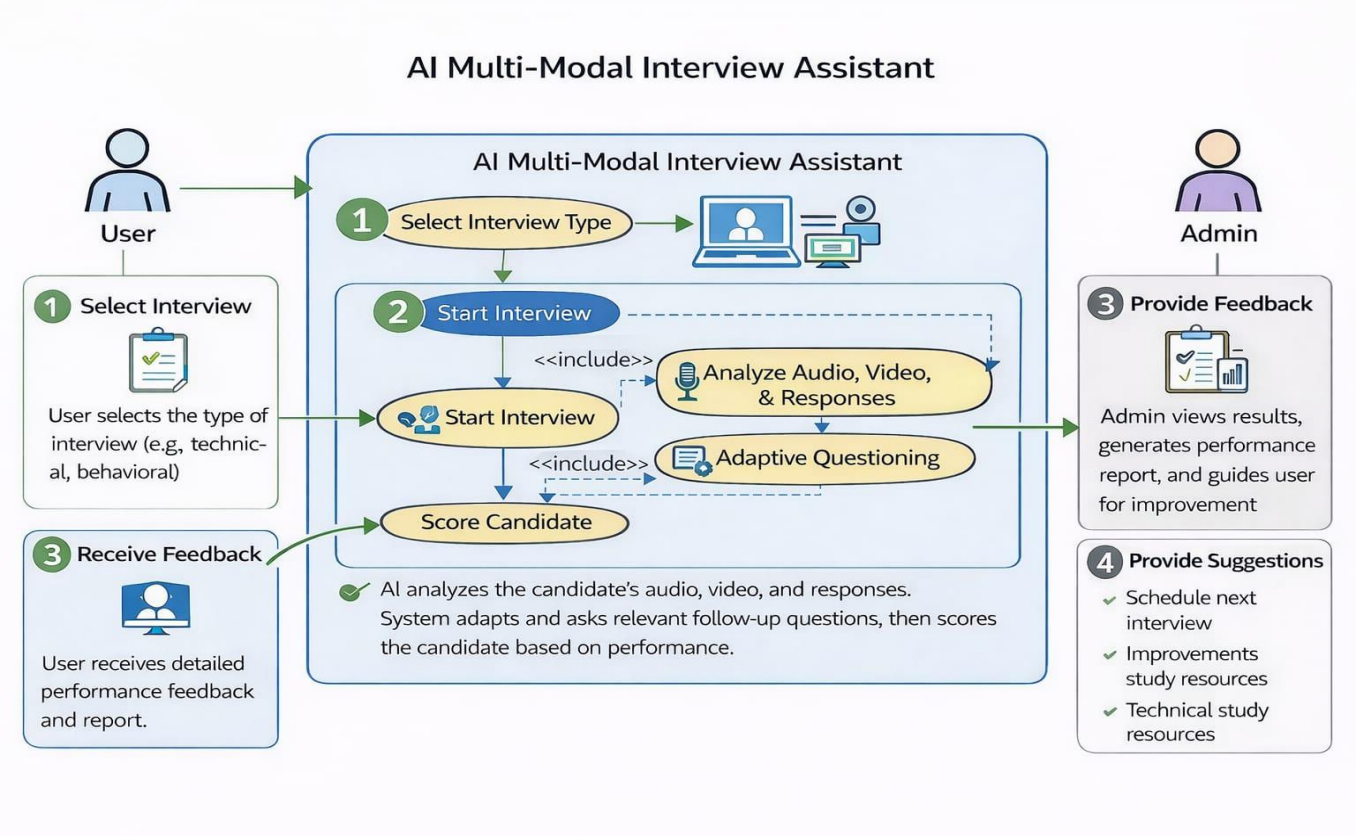
USP of the proposed solution:

- Comprehensive Candidate Evaluation – Analyzes voice tone, facial expressions, and answer content together to provide a complete understanding of the candidate's performance.
- Adaptive & Engaging Interviews – Dynamically adjusts questions in real time based on candidate responses, keeping the interview relevant and interactive.
- Actionable Feedback & Insights – Provides detailed performance reports and analytics, helping employers make better hiring decisions and candidates improve for future interviews.

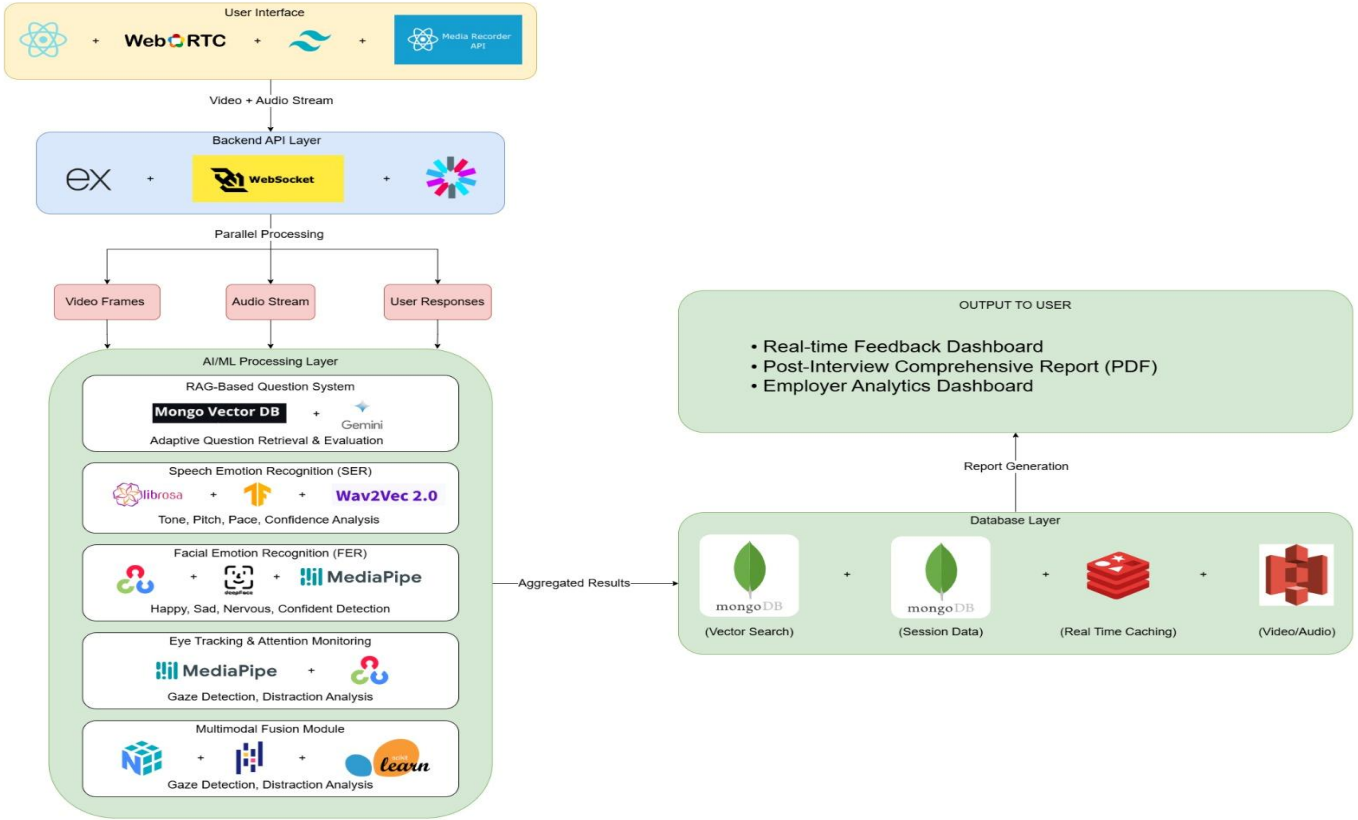
Features Offered by the Solution

1. **AI-Based Interview Question Generation**
   - Generates adaptive technical & HR questions
2. **Facial Emotion Detection (FER)**
   - Analyzes candidate's emotional state
3. **Speech Emotion Recognition (SER)**
   - Analyzes tones & voice modulation
4. **Eye Contact & Attention Tracking**
   - Measures eye contact & focus
5. **Response Behavior Analysis**
   - Tracks response time and typing speed
6. **Performance Scoring System**
   - Provides confidence & communication scores
7. **Real-Time Feedback**
   - Gives instant performance insights
8. **Session Recording & Storage**
   - Saves interview video and audio
9. **Detailed Analytics Dashboard**
   - Visual performance reports & metrics
10. **Secure & Scalable Architecture**
    - Ensures data privacy and scalability

Use-case diagram:



Architecture diagram of the proposed solution :



Technologies to be used in the solution :

FRONTEND:
- React.js
- Tailwind CSS
- TypeScript
- WebRTC
- MediaRecorder API
- WebSocket Client

BACKEND:
- Express.js
- Node.js
- WebSocket(WS)
- JWT Authentication
- RESTful API

DATABASE & STORAGE :
- MongoDB Atlas-
  - Vector Search
  - Session Data
- Redis-
  - Caching
  - Real-time Data
- AWS S3-
  - Video/Audio Storage
  - PDF Reports

AI/ML:
- MongoDB Vector Search
- Google Gemini API-
  - gemini-embedding-001
  - gemini-2.0-flashlite
- TensorFlow
- PyTorch
- librosa
- OpenCV
- DeepFace
- MediaPipe Face Mesh
- wav2vec 2.0
- NumPy
- Pandas
- Scikit-learn

Innovation partner H2S HACK2SKILL | Media partner YOURSTORY

AI for Bharat Hackathon

Powered by aws

Thank You