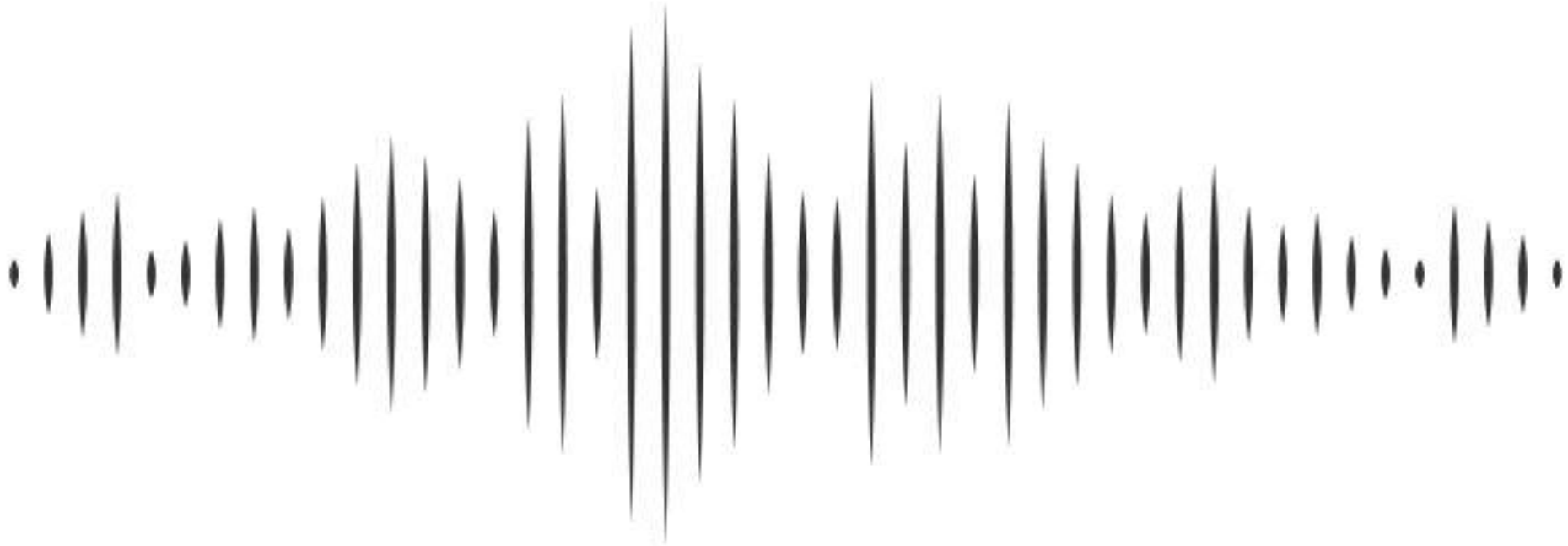
Aura: Voice-Based Thought Lag & Cognitive Load Detection System

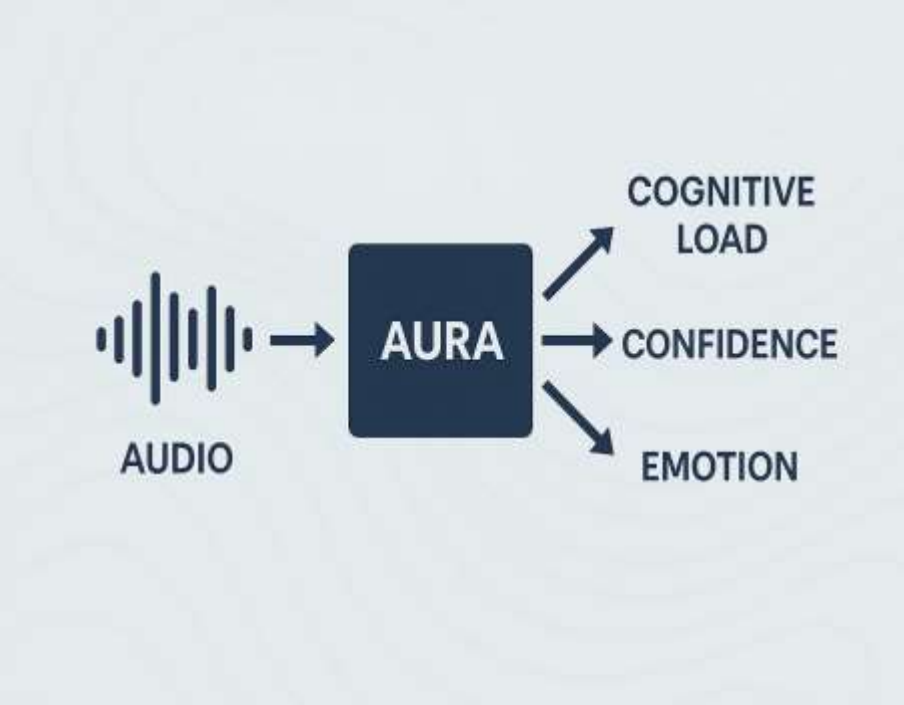
Team Name: VENUS
Team Members: Sandipan Majumder , Anik Kumar Ghosh , Sumit Roy
Tagline: *"Decoding hesitation. Detecting emotion. Delivering insight."*



The Hidden Data in Conversation

- In high-stakes environments - like job interviews, therapy sessions, or educational assessments - how we speak is as important as what we say.
- Critical indicators of a person's mental state - such as cognitive load, stress, and lack of confidence - are conveyed through subtle vocal cues.
- Today, analyzing these cues is either:
  - Subjective: Relies on human observation, which is prone to bias and inconsistency.
  - Inaccessible: Requires specialized, expensive hardware like EEG or galvanic skin response sensors, limiting its use.
- The core problem: We lack an objective, scalable, and accessible way to quantify the rich psychological data embedded in the human voice.

An AI Platform for Objective Speech Analysis

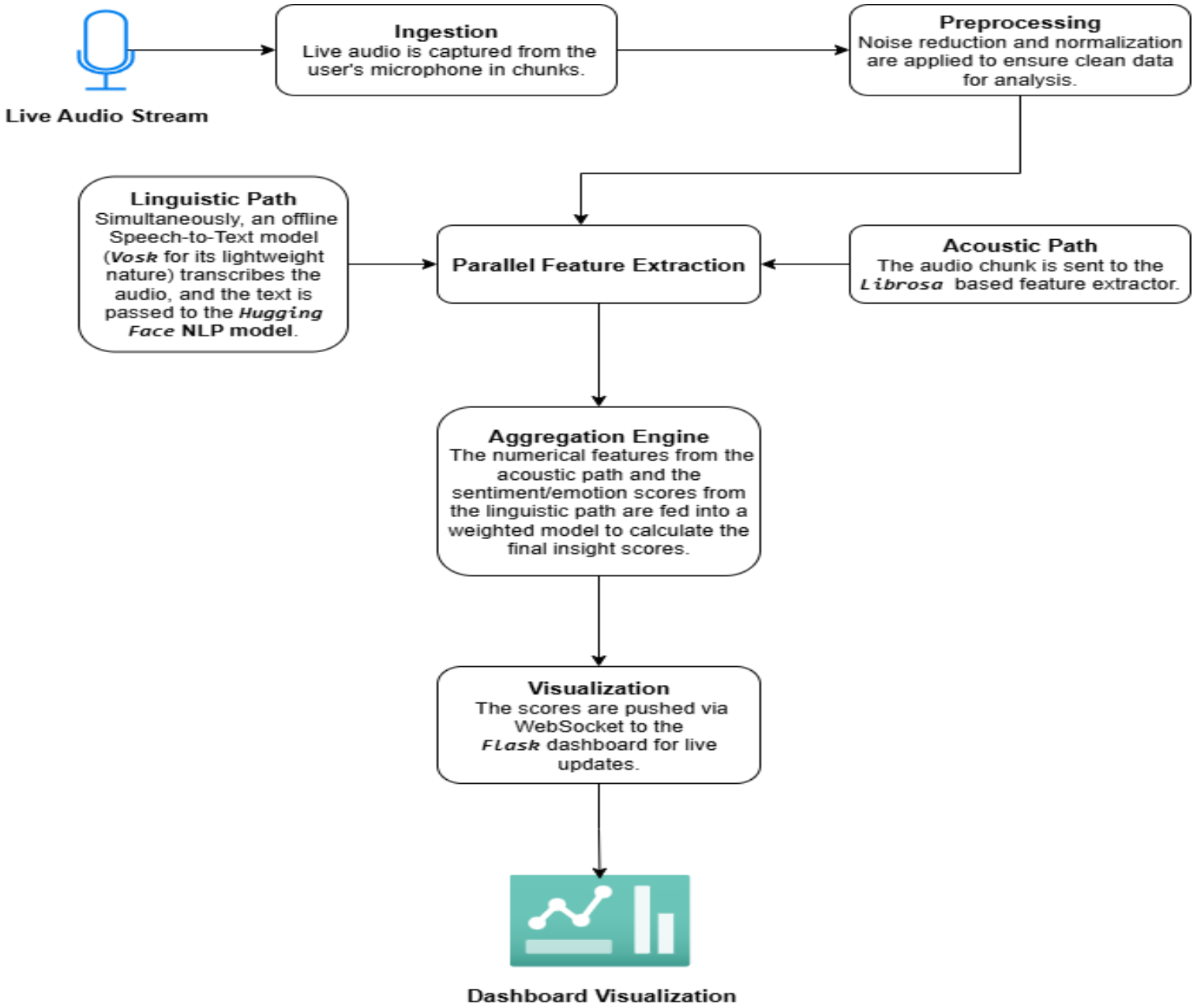


- Aura is a software-only solution that ingests raw audio and translates it into actionable insights in real-time.
- It operates on a dual-analysis pathway:
    1. **Acoustic Analysis**: Measures the physical properties of speech. It quantifies hesitation markers ('ums', 'ahs'), pitch variation (jitter/shimmer), speech rate, and response latency.
    2. **Linguistic Analysis**: Understands the content and sentiment of what is being said using Natural Language Processing (NLP).
- **The Output**: A unified dashboard that provides a clear, data-driven score for cognitive load, confidence, and emotional tone, helping users understand their communication patterns objectively.

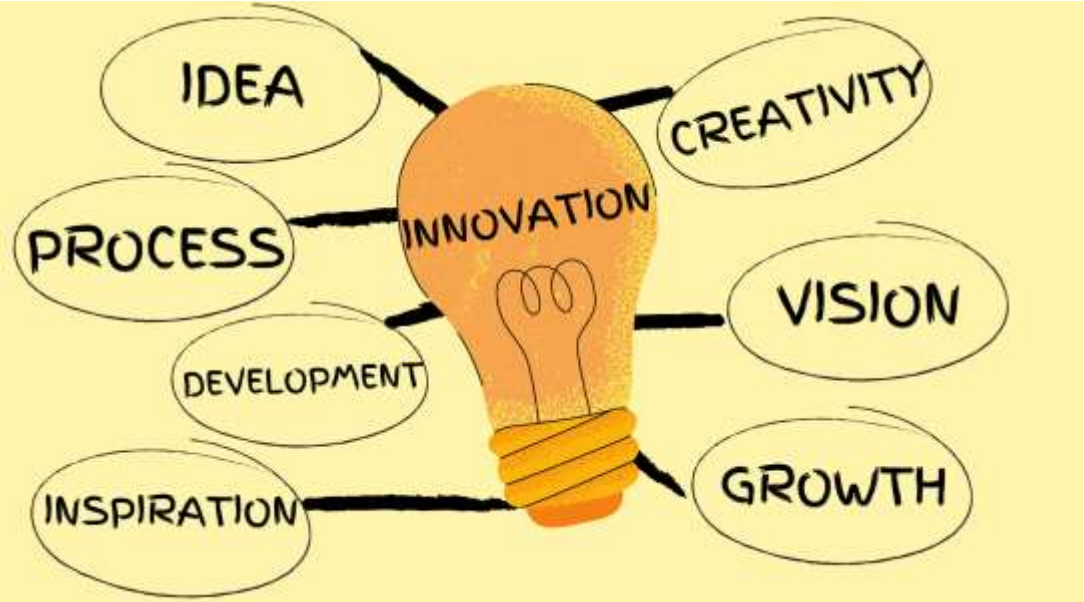
Built with Proven, High-Performance Tools

- Audio Capture & Processing:
  - ***PyAudio***: For real-time, cross-platform audio stream capture directly from a microphone.
  - ***Librosa***: A powerful Python library for robust acoustic feature extraction (e.g., MFCCs, Chroma, Pitch, Jitter, Shimmer).
- Backend & API:
  - ***FastAPI***: Chosen for its high performance and automatic API documentation, ideal for a real-time system.
- AI & Machine Learning:
  - ***Hugging Face Transformers*** :Utilizes a tuned **Distilbert** model for classifying sentiment and emotion efficiently, offering balance between accuracy and computing cost.
  - ***Scikit-learn***: For creating the final aggregation model, a simple regression or classifier that combines all extracted features into a final score.
- Frontend & Visualization:
  - ***Flask***: Easily developing an interactive and real time-based dashboard i.e. an MVP.

Architecture & Workflow Diagram



Unique Features & Innovation



- **Hybrid Vocal Biomarker Model (Our Core Innovation)**: Unlike tools that only analyze what you say (NLP) or how you say it (acoustics), our strength lies in the fusion of both data streams. This provides a holistic and more accurate picture of the speaker's cognitive state.
- **Hardware-Agnostic & Accessible**: Runs on any standard laptop or phone. This democratizes access to communication analytics, fulfilling a key social impact goal.
- **Real-Time Feedback Loop**: The system is designed for sub-second latency, making it practical for live coaching and self-assessment, not just post-analysis.
- **Explainable AI:** Or dashboard gives reason instead of giving scores. For example, "Confidence score is low due to high hesitation marker frequency and pitch instability."

Hackathon Prototype to Real-World Impact

Impact

Reducing bias in interviews by providing an objective layer of candidate assessment.

Aiding therapists by tracking patient progress and detecting early signs of distress between sessions.

Providing students data driven feedbacks to improve their speaking and presentation skills.

Future Scope

Evolve into an "AI Agent" that provides prescriptive feedback ("Try speaking slower and pausing at the end of your sentences"). (This hits the **Agentic AI** bonus).

Integrate video analysis to capture facial expressions and body language for a complete communication profile.

Develop a commercial API for platforms like Zoom, Google Meet, and telehealth portals.