**Project Overview:**  
Our project aims to develop an AI-driven software system that identifies and distinguishes between different speakers at the start of a conversation, lesson, or meeting. The software will analyze the speakers’ voices and assign a unique identity to each participant. Additionally, the system will generate an individual report for each participant, which includes emotion analysis based on tone and content, a summary of key points said by the participant, and an analysis of the elements in the conversation that triggered specific emotions.

**1. What is the problem you are trying to solve?**  
In multi-person conversations, it’s often challenging to track who is speaking, what each person contributed, and how they emotionally responded to the discussion. Moreover it is very useful to analyze the emotional effect of the responses in the conversation. This is especially relevant for meetings, lessons, or negotiations where it is essential to understand participants’ emotional responses, summarize key contributions, and identify emotional triggers.

**2. Describe briefly, in high level, your presumed solution**Our solution is an AI-powered software that, at the start of a conversation, prompts each participant to say a sentence for voice identification purposes. The system then tracks and separates each participant’s contributions throughout the conversation. It uses Natural Language Processing (NLP) and Speech Emotion Recognition (SER) to analyze both the content and tone of each participant’s speech. At the end of the conversation, the system generates a personalized report for each participant, summarizing their key contributions, emotional trends, and identifying moments that led to emotional shifts. Additionally, we will observe the emotional effect of responses in a conversation.

**3. Are there other approaches?**  
Yes, there are several existing approaches that attempt to address parts of this problem. For example:

- Voice Assistants (like Alexa, Google Assistant):These can identify voices but do not generate participant-specific reports.  
- Transcription Tools (like Otter.ai, Rev):These convert speech to text but do not distinguish speakers' emotions or analyze emotional triggers.  
- Emotion Analysis Tools (like Affectiva, Microsoft’s Azure Cognitive Services): These analyze emotions but are not speaker-specific or linked to summaries of spoken content.  
 Our approach integrates and expands on these existing solutions by combining speaker identification, speech-to-text transcription, emotion analysis, and personalized reporting into one comprehensive system.

**4. Who are the expected users of the application?**

- Teachers & Educators: For tracking student engagement and emotional responses during lessons.  
- Business Professionals: For summarizing key points in meetings and analyzing emotional responses to negotiations or presentations.  
- Therapists & Counselors: For tracking emotional shifts and participant involvement in group therapy sessions.  
- Market Researchers: For analyzing customer sentiment in focus groups or interviews.

**5. What will be the main features and flows of the (different) user(s)?**

1. Voice Registration: At the start of the session, each participant records a short sentence for voice identification.  
2. Speaker Tracking: The system identifies who is speaking and tracks their contributions.  
3. Speech-to-Text Conversion: Converts all spoken words into text in real-time.  
4. Emotion Detection: Tracks emotional states of each participant throughout the conversation.  
5. Trigger Analysis: Identifies which points in the conversation triggered emotional shifts.  
6. Personalized Reports: At the end of the session, each participant receives a detailed report that includes:  
 - Key statements they made.  
 - A timeline of their emotional shifts.  
 - An analysis of key moments that triggered these shifts.

**6. Are there any external dependencies?**

- Speech-to-Text API: For converting spoken language to text (e.g., Google Speech-to-Text, IBM Watson).  
- Emotion Recognition API: For tracking and analyzing emotions from voice data (e.g., Affectiva, Microsoft’s Emotion API).  
- Cloud Services/Storage: To store the recorded sessions and participant reports.  
- AI/ML Models: Custom models for speaker recognition, NLP for text analysis, and emotion analysis.

**Objectives:**

- To create a system that can identify and distinguish participants in a multi-person conversation.  
- To track emotions during conversations and provide analysis on emotional triggers.  
- To generate individual participant reports that summarize key contributions and emotional responses.

- Characterize the emotional effect of responses in a conversation

**Key Features:**

- Speaker Identification: Voice-based speaker recognition at the start of the session.  
- Emotion Analysis: Real-time emotion tracking for each speaker.  
- Personalized Reports: Participant-specific summaries at the end of each session.  
- Trigger Analysis: Identification of conversation elements that led to emotional shifts.

**Target Audience:**

- Teachers & Educators  
- Business Professionals  
- Therapists & Counselors  
- Market Researchers

**Technology Stack:**

-Programming Languages: Python (for ML/AI models), JavaScript (for front-end development), SQL (for databases).  
- APIs/Services: Google Speech-to-Text, Microsoft’s Emotion API, IBM Watson (for NLP).  
- Cloud Storage: AWS S3 or Google Cloud Storage for recording and report storage.  
- AI/ML Tools: TensorFlow or PyTorch for machine learning models for speaker and emotion recognition.

**Implementation Approach:**

- Agile Methodology: Development will be done iteratively, with sprints focused on key features.  
- Data Collection & Training: Collect diverse voice data to train speaker and emotion recognition models.  
- System Integration: Combine speech-to-text, speaker recognition, and emotion tracking into one pipeline.  
- Report Generation: Develop a system to generate PDF/HTML reports for each participant at the end of the session.

**Challenges Faced:**

- Voice Identification: Ensuring accurate voice identification, especially with overlapping voices.  
- Emotion Analysis: Making sure emotional analysis is accurate and context-aware.  
- Data Privacy: Ensuring participant data is secure and meets privacy standards.

**Impact and Benefits:**

- Improved Understanding: Insightful reports help users understand participant engagement, emotional states, and key contributions.  
- Efficiency: Automates the process of note-taking, emotional analysis, and participant tracking.  
- Better Decision Making: Helps identify emotional triggers, improving negotiations and discussions.

- Stakeholder management via emotional intelligence : Ability to monitor the emotional effect of our responses in a conversation may be a way to increase the emotional intelligence of the participants

**Future Scope:**

- Language Support: Adding support for multiple languages.  
- Custom Reports: Allowing users to customize the structure of participant reports.  
- Integration: Integration with existing business tools like Microsoft Teams or Zoom.

**Key Takeaways:**

- The system combines speaker identification, NLP, and emotion analysis into one tool.  
- The system’s reports provide valuable insights for users in education, business, and mental health.  
- Emphasis on privacy and ethical use of data.

**Demonstration or Showcase:**

- A live demonstration where multiple people register their voices, participate in a brief conversation, and receive their personalized reports.

**Team and Contributors:**

- Team Members:

Amal Amsis

Yarden Daniel

Yaniv Zamir