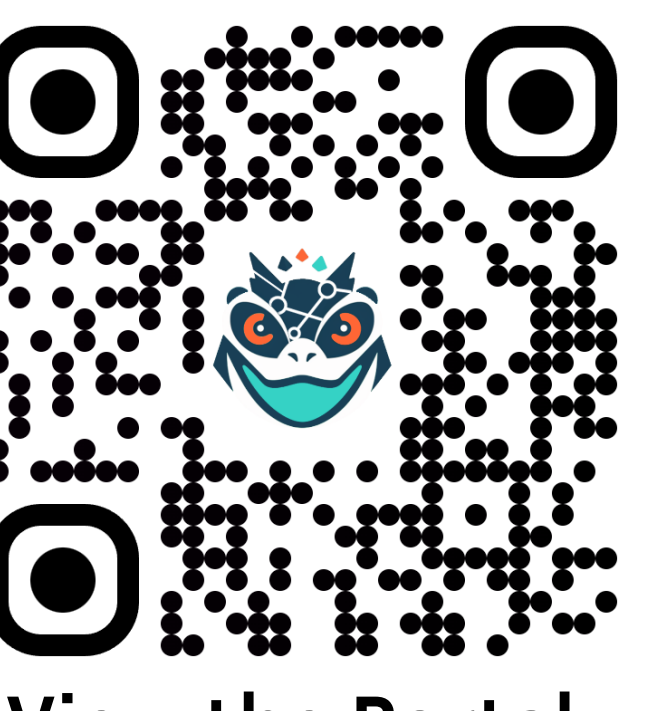




ClassifAI

Advancing Teacher-Student Interaction Analysis through Automated Speech Transcription and Question Classification

Authors: John Henry Mejía, John Nguyen, Taylor Griffin, Jaxon Hill, Nagato Kadoya



View the Portal: <https://classifai.tcu.edu>

Project Goals

What: ClassifAI is an online video/audio analysis tool built to assist teachers and educators in analyzing and understanding how their interaction with students translates into real learning methods. ClassifAI does this by providing transcripts of what was said, measuring how long it takes students to respond to questions, and evaluating the complexity of questions based on Costa's levels of thinking.

Why: Manual methods of analysis are time-consuming and may not be effective in quantifying metrics between teacher and student interactions.

Who: Instructors who seek more effective and efficient ways to analyze their teaching methodologies.

Previous Iteration:

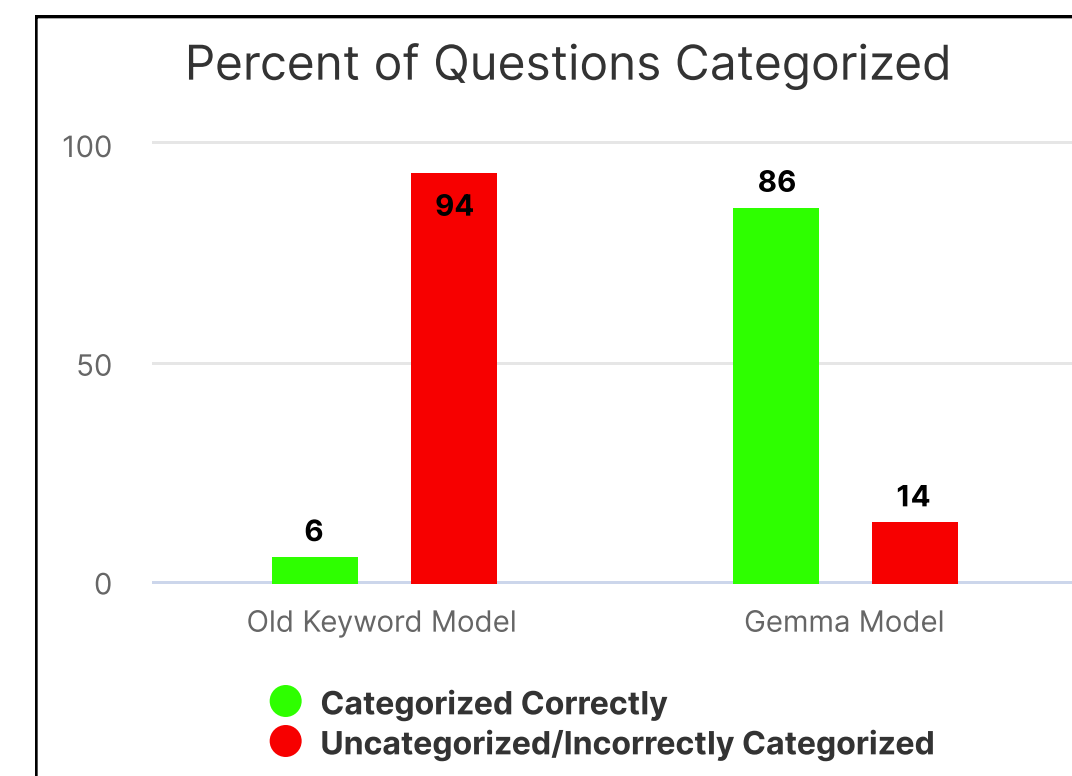
- Used Assembly AI to transcribe files (costly)
- Question Categorization was based on Bloom's Taxonomy
- No YouTube support or summarization
- Poor User Interface and User Experience (UI/UX)

Project Outcomes

Cost effectiveness- all processing is now done on TCU servers, while maintaining accuracy (100% reduction in month-over-month cost)

Automatic categorization according to Costa's Level of Reasoning.

- We went from approximately 6% of questions categorized via keywords to all questions categorized with a fine-tuned Gemma, resulting in 86% accurate categorization of questions via Gemma.



Accuracy was determined by manually reviewing Costa's level of thinking for 4 days of a middle school English class, with question level data from our client Dr. Faggella-Luby. Using keywords only identified 23 of 523 questions, and the fine-tuned Gemma model identified 445/523 questions correctly, with the other 87 questions being incorrect.

Automatic summarization of transcripts!

Mobile Optimization

YouTube uploads added

UI Enhanced

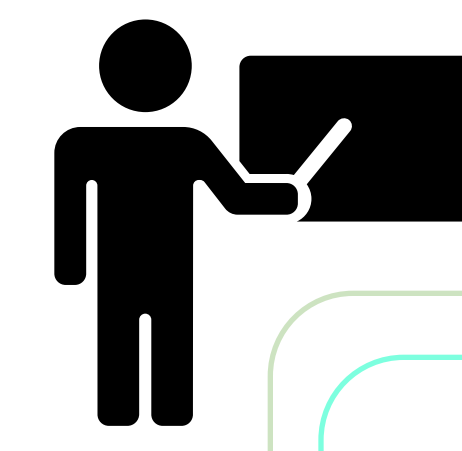
Possible Improvements for Next Iteration

- Develop more accurate models of question categorization and summarization
- Identify heatmap of confusion, and a sentiment analysis of the transcript- identify problematic times.
- Add a student-teacher network graph to identify which people interact with each other most often.

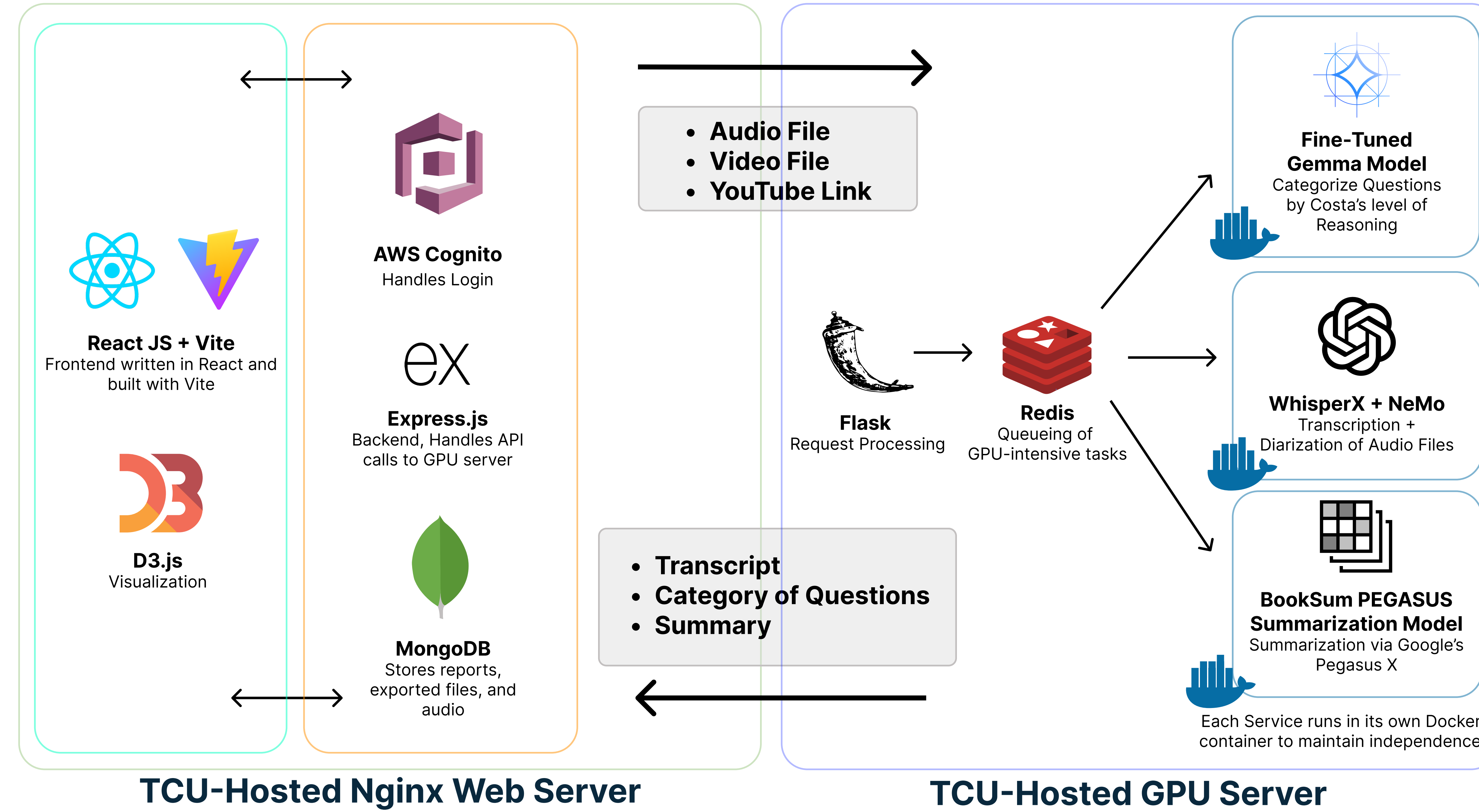
Acknowledgments

- Dr. Bingyang Wei, for being our faculty advisor
- Our Clients:
 - Dr. Liran Ma
 - Dr. Alexander Richard
 - Dr. Michael Faggella-Luby
- Dr. Michael Denkowski, for advising us on our NLP models

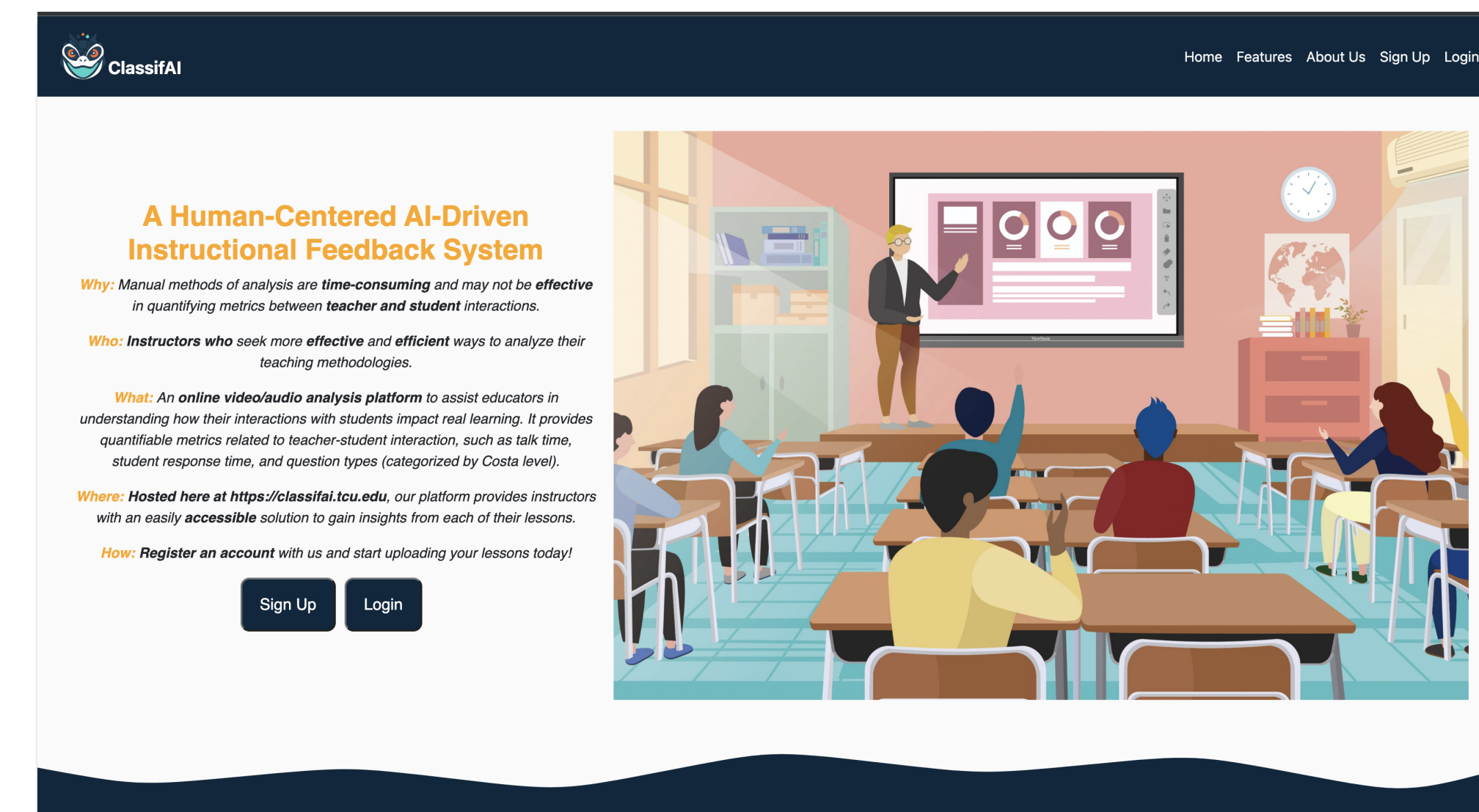
ClassifAI is an online video/audio analysis tool that assists educators by quantifying different metrics about their teacher-student interaction. Instructors can view metrics such as the time the teacher talks during a lesson, the timeline of questions asked and their Costa's level of Inquiry, the response time of students to those questions, and other data points such as an automatic summary of the classroom audio.



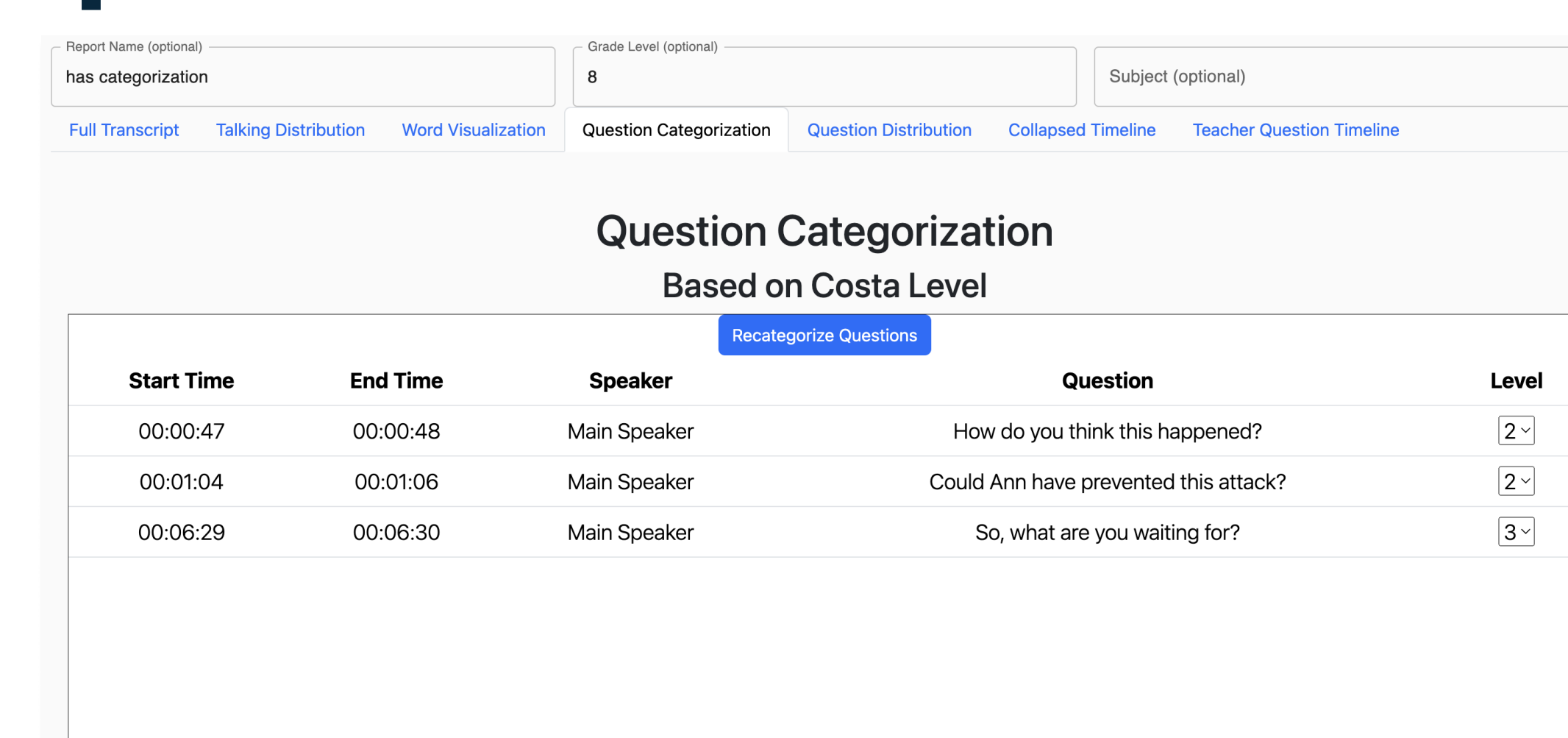
System Architecture



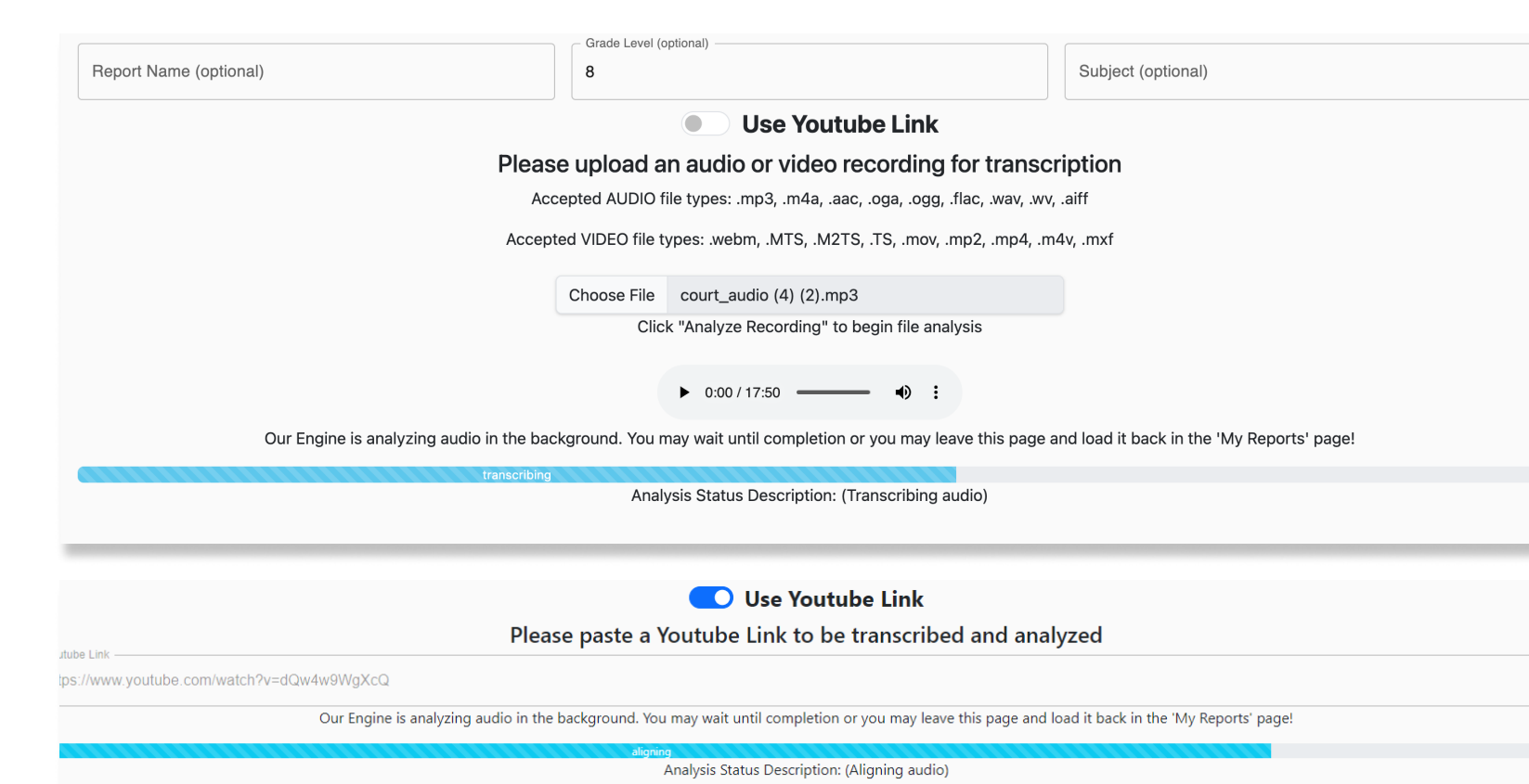
User Interface/Experience



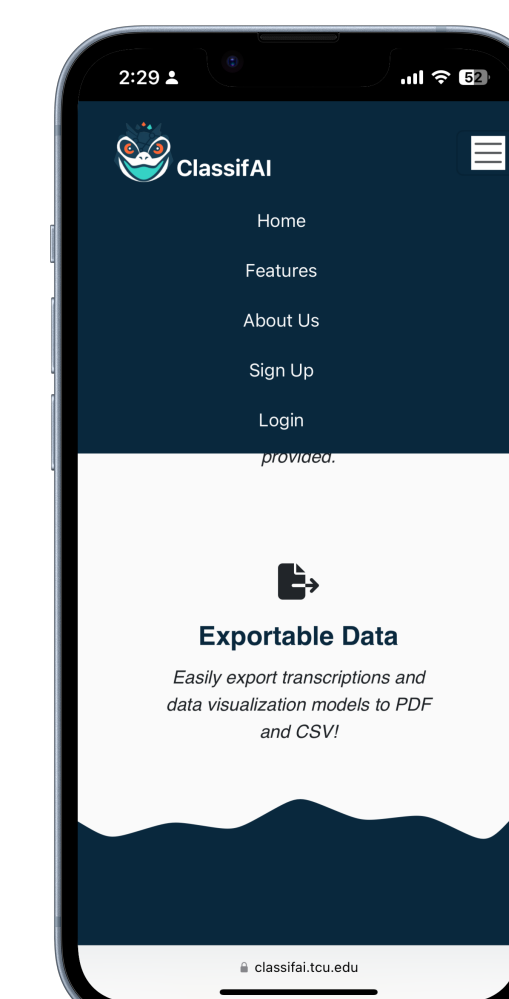
Landing page



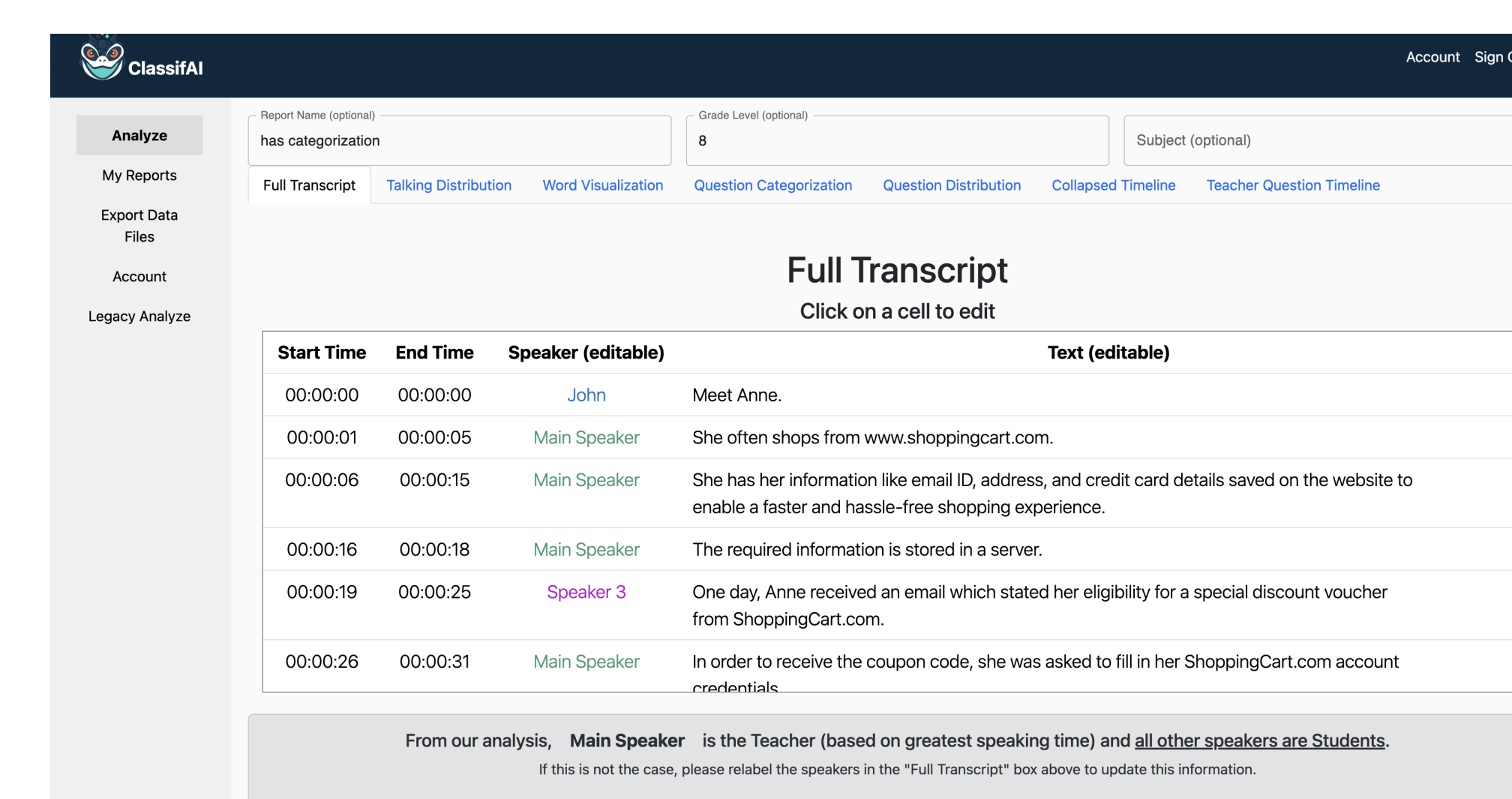
Question Categorization



Transcribing View / Progress bar



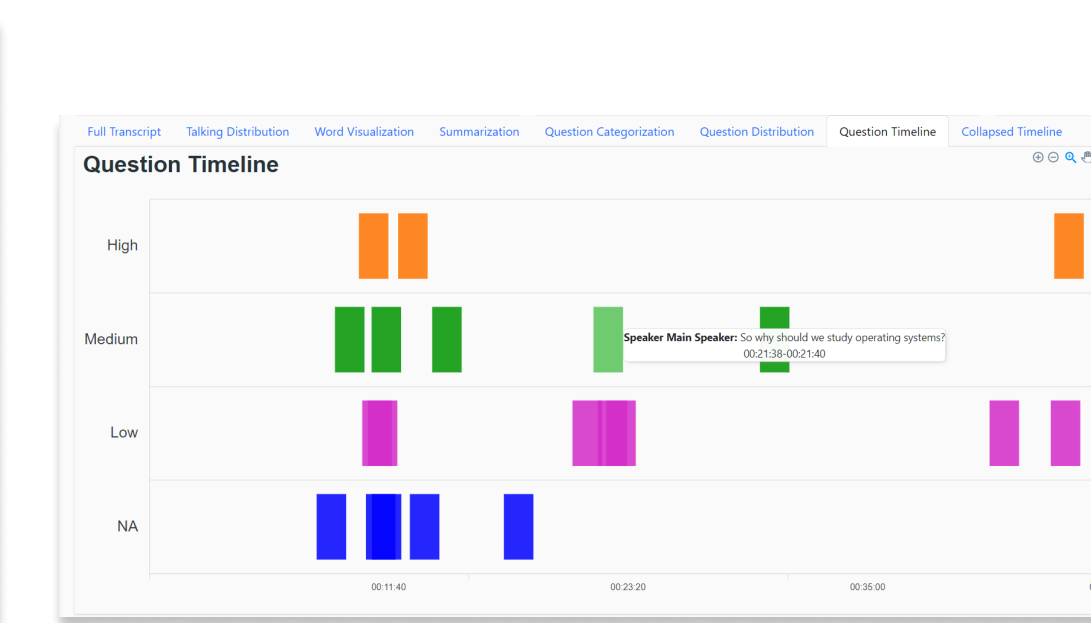
Mobile View



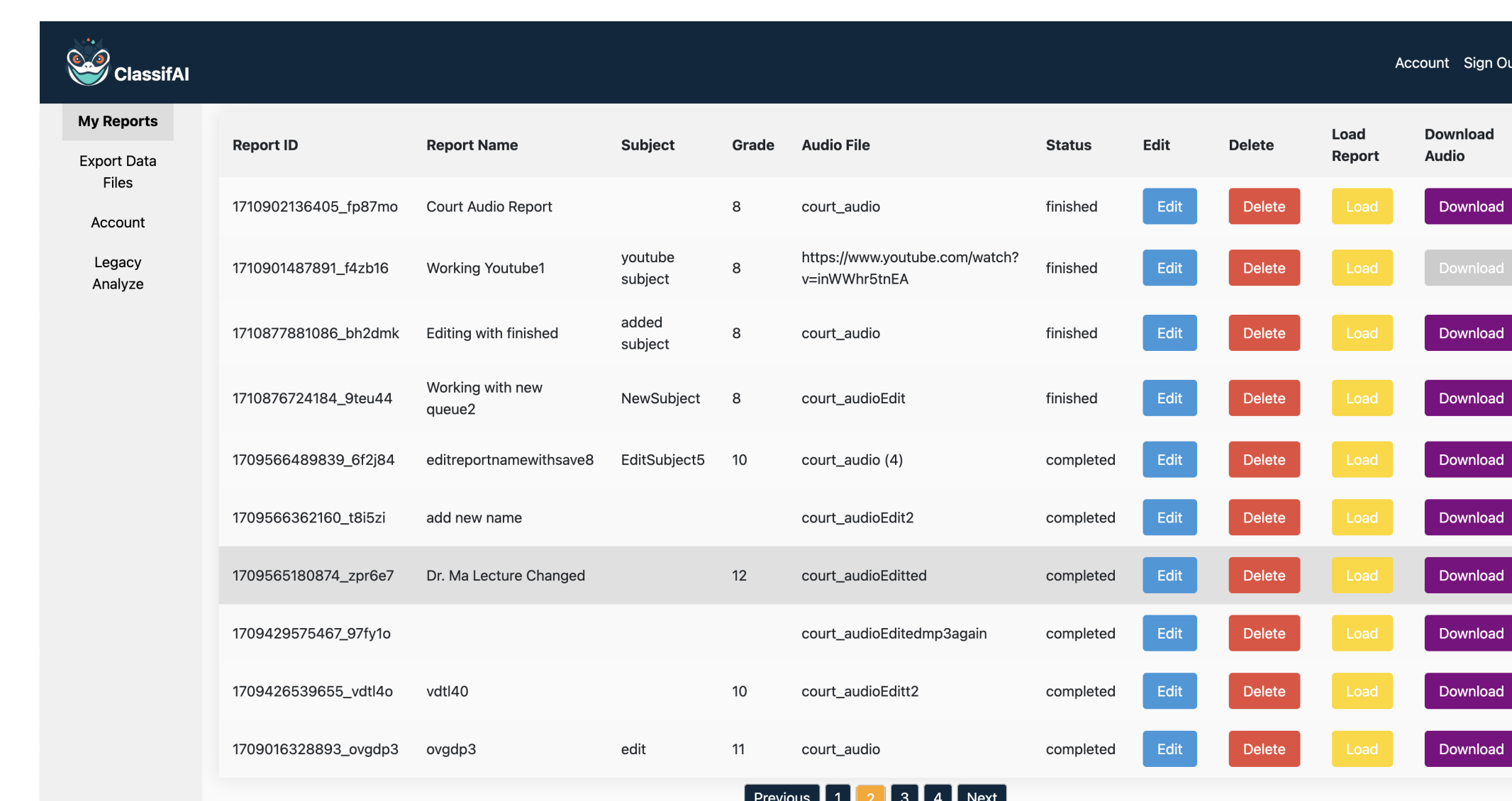
Transcript View



Word Visualization



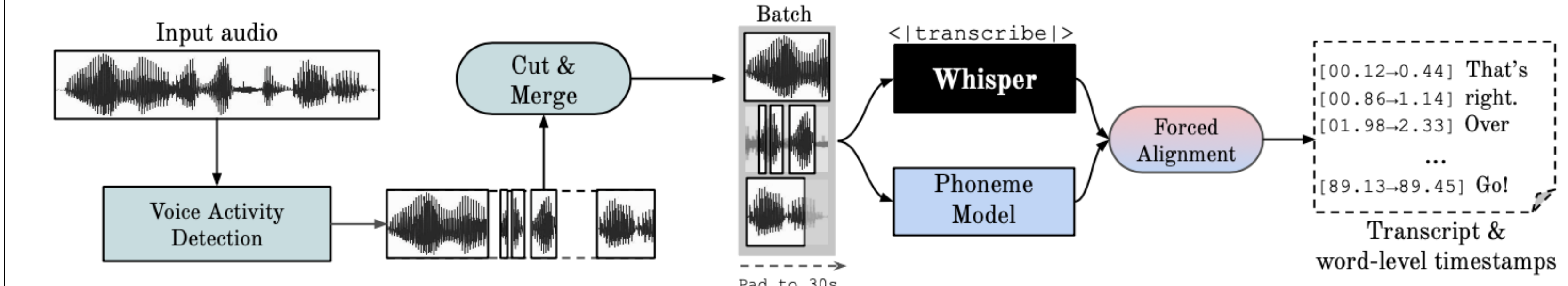
Level Distribution



Report File Management

Artificial Intelligence

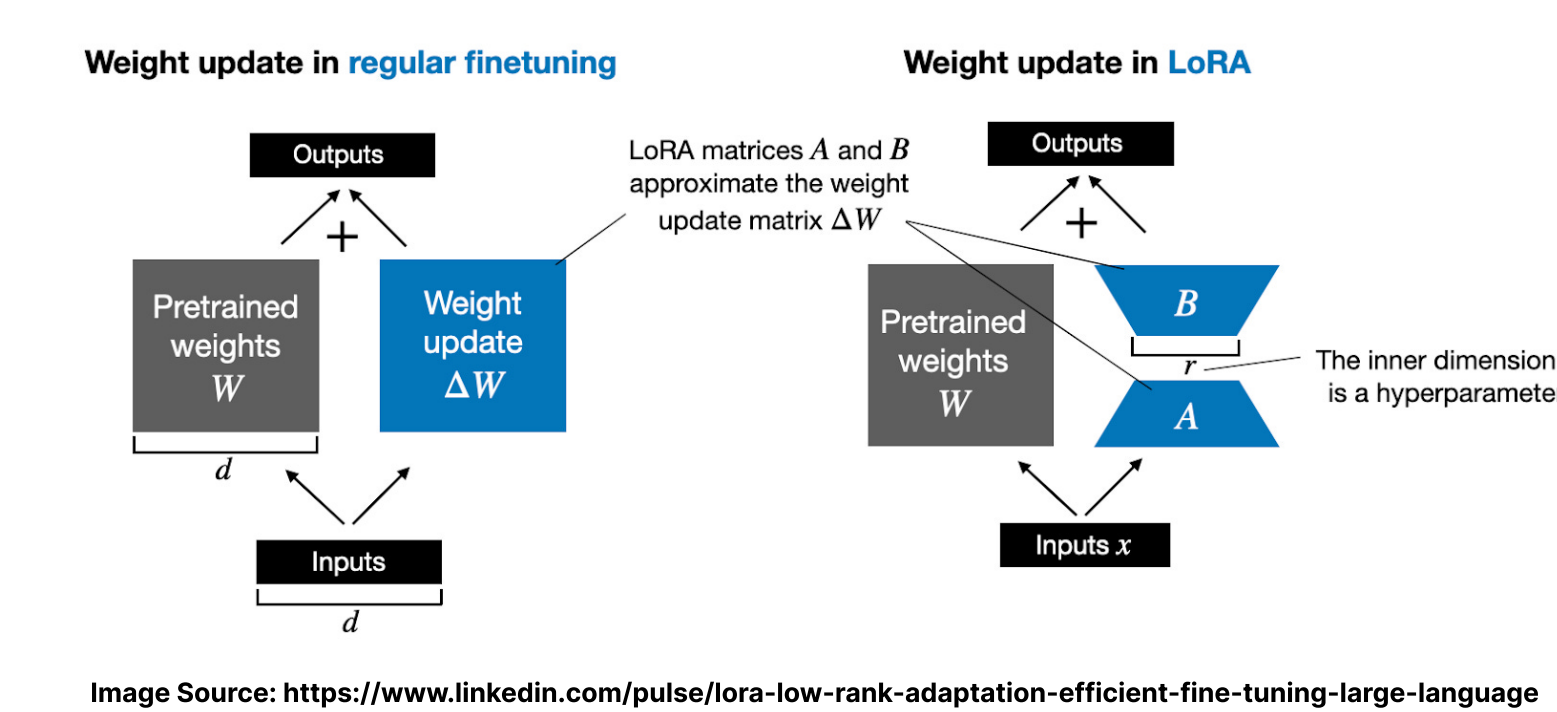
Transcription and Diarization



- Audio vocals are isolated for greater speaker embedding accuracy.
- OpenAI's Whisper + WhisperX transcribes the audio (see diagram).
- Pyannote's speaker diarization extracts speaker embeddings and finds the speaker of each segment, in parallel.
- Embeddings are aligned with WhisperX for word-level speaker identification.
- Punctuation models fine-tune alignment, compensating for minor time discrepancies and restoring lost punctuation.

Categorization- Gemma Fine-Tuning via PEFT and LoRA

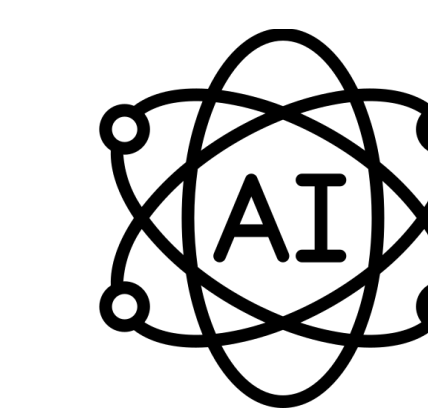
- We categorize questions according to their Costa's level of inquiry (1/Gathering, 2/ Processing, 3/Applying).
- Categorization is done by fine-tuning the Gemma-7B-Instruct model via Parameter Efficient Fine Tuning (PEFT) on sample questions and their corresponding level.
- The training data is derived from our client's classroom, along with AVID's "Costa's Levels of Thinking and Questioning" resource (which provides question examples for different subjects and different levels).
- We employ the LoRA (Low Rank Adaptation) technique, which modifies only a small subset of the model's weights during fine-tuning and approximates weights. This approach offers efficiency advantages over traditional full-model fine-tuning. The model was trained for 2 epochs.



Transcript Summarization

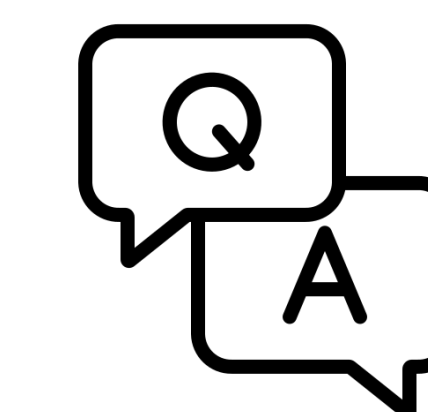
- Summarization is done using Google's pre-trained PEGASUS (Pre-training with Extracted Gap-sentences for Abstractive Summarization) model, specialized in giving a summary of arbitrary topics.
- Created by fine-tuning Google's Pegasus-x-large model on the kmfoda/ booksum dataset for eight epochs.

Features



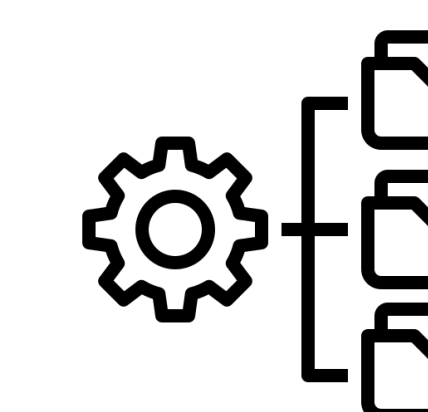
Speech Recognition and Diarization

Extremely fast performance transcription through optimized speech recognition and diarization algorithms (~80 sec for a 45 minute video)



Question Categorization and Analysis

All transcripts undergo a thorough analysis to categorize the questions and evaluate instructional effectiveness. Detailed data visualization is provided.



Exportable and Manageable Data

Easily export transcriptions and data visualization models to PDF and CSV.



Transcript Summarization

Transcripts are summarized to extract relevant information for ease of access and note taking.