



ClassifAI

Enhancing Education: Transforming Dialogue into Insights

John Nguyen, Taylor Griffin, Jaxon Hill,
Nagato Kadoya, John Henry Mejia

The Problem



Jaxon the Student



Taylor the Teacher

Imagine that you've just given a test....

My lectures were **very** detailed



Taylor the Teacher

I've **prepared** them well.

They've answered all of my **questions**

In reality....

I'd be surprised
if this was covered
**more than
once**

Lectures
were so
unengaging,
I fell asleep



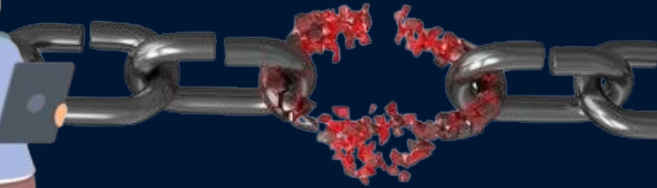
Jaxon the Student

I've **never** had to
think this **deeply**
In lectures

How Do We Fix This?



Jaxon the Student



Taylor the Teacher

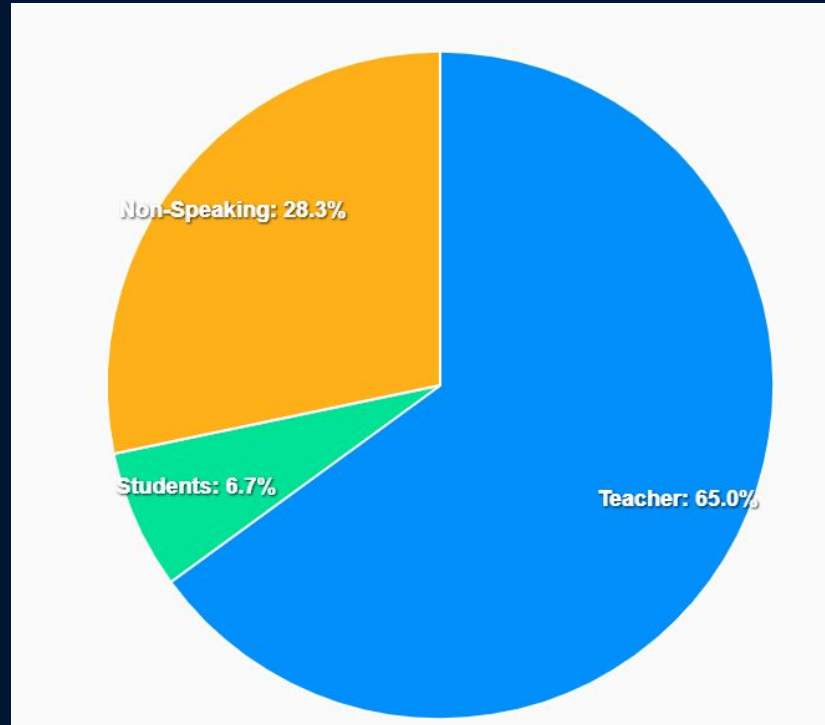
What if Taylor could:

See how deep his questions were



What if Taylor could:

See how much time he spent talking vs students



What if Taylor could:

See what each person said in the lecture

Full Transcript | Talking Distribution | Word Visualization | Summarization | Question Categorization | Question Distribution | Question Timeline | Collapsed Timeline

Full Transcript

Click on a cell to edit

Start Time	End Time	Speaker (editable)	Text (editable)
00:00:00	00:00:03	Main Speaker	Hey, I'm Jabril, and this is Crash Course AI.
00:00:03	00:00:06	Main Speaker	Today, we're going to try and teach John Greenbot something.
00:00:06	00:00:07	Main Speaker	Hey, John Greenbot.
00:00:08	00:00:10	Speaker 1	Hello, humanoid friend.
00:00:11	00:00:12	Main Speaker	Are you ready to learn?
00:00:12	00:00:14	Speaker 1	Hello, humanoid friend.
00:00:15	00:00:22	Main Speaker	As you can see, he has a lot of learning to do, which is the basic story of all artificial intelligence.

From our analysis, **Main Speaker** is the Teacher (based on greatest speaking time) and all other speakers are Students.
If this is not the case, please relabel the speakers in the "Full Transcript" box above to update this information.

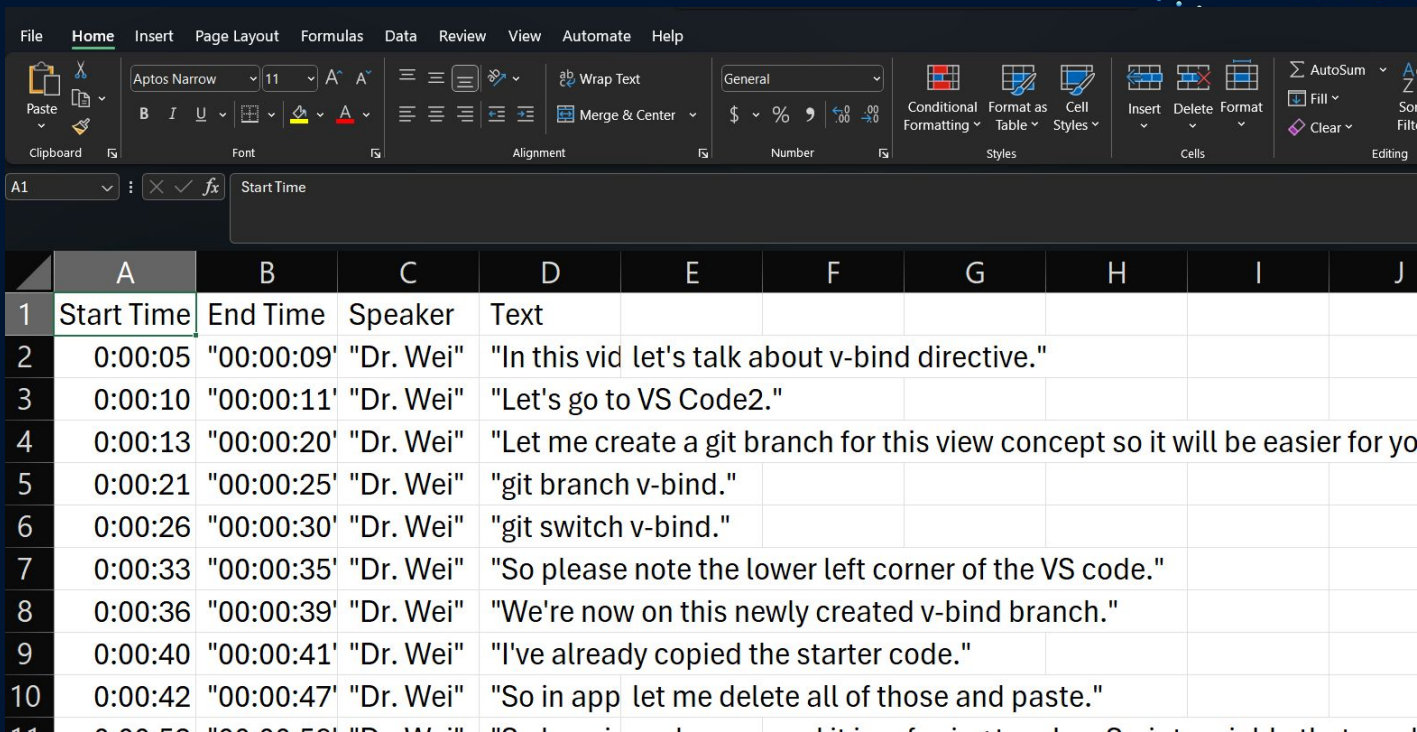
Or view a quick summary on what he taught

Summarization

Here is a concise and informative summary of the teacher's lecture: The teacher's lecture focused on the study of how minutes are spent teaching English in a middle school in Eastern Connecticut. The teacher, who is also a movie star, volunteered to participate in the study and explained the challenges of fitting the study into the class schedule. The class is currently studying Chapter 19 and will be writing examples of personification, simile, and setting. They will also be adding vocabulary words to their binders, including "thesis statement" and "denotation". The teacher explained that a thesis statement is a one-sentence summary of the main idea of a piece of writing, usually found in the introduction. She also defined denotation as the process of defining unknown words using a dictionary.

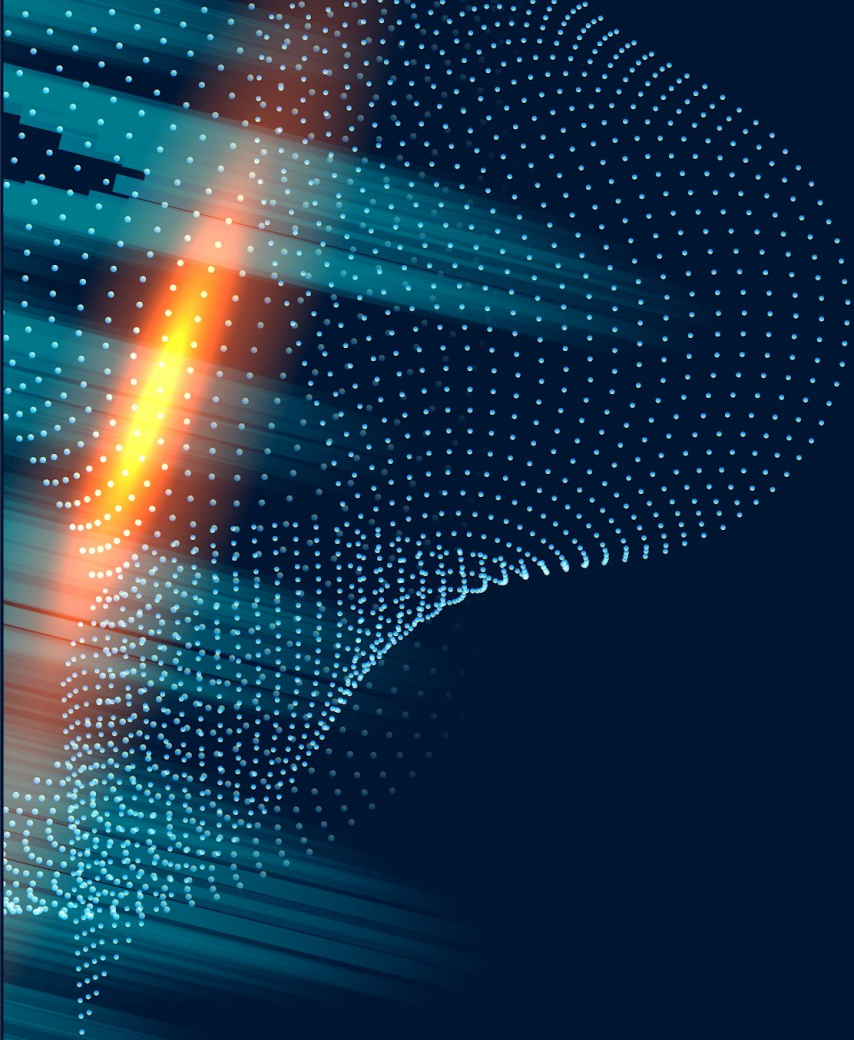
[Save & Download CSV](#)

Or even download these insights from anywhere, in many different formats



The image shows a screenshot of the Microsoft Excel interface. The ribbon is set to 'Home', and the active cell is A1, which contains the text 'Start Time'. Below the ribbon, a table is displayed with the following data:

	A	B	C	D	E	F	G	H	I	J
1	Start Time	End Time	Speaker	Text						
2	0:00:05	"00:00:09"	"Dr. Wei"	"In this vid let's talk about v-bind directive."						
3	0:00:10	"00:00:11"	"Dr. Wei"	"Let's go to VS Code2."						
4	0:00:13	"00:00:20"	"Dr. Wei"	"Let me create a git branch for this view concept so it will be easier for yo						
5	0:00:21	"00:00:25"	"Dr. Wei"	"git branch v-bind."						
6	0:00:26	"00:00:30"	"Dr. Wei"	"git switch v-bind."						
7	0:00:33	"00:00:35"	"Dr. Wei"	"So please note the lower left corner of the VS code."						
8	0:00:36	"00:00:39"	"Dr. Wei"	"We're now on this newly created v-bind branch."						
9	0:00:40	"00:00:41"	"Dr. Wei"	"I've already copied the starter code."						
10	0:00:42	"00:00:47"	"Dr. Wei"	"So in app let me delete all of those and paste."						
11	0:00:50	"00:00:50"	"Dr. Wei"	"So in app let me delete all of those and paste."						



**What if you
could do the
same?**

Introducing ClassifAI



Who

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

- **Question Level distribution** over time
- **Amount of Questions** asked

What

- ClassifAI is an online **video/audio analysis tool**
- Provides **transcripts** with **diarization**
- **Questions levels** based on **Costa's levels**
- **Summary** of the transcript

Why

ClassifAI is **faster** than Manual methods of classroom analysis.

In the **press** of a **button**, you can **activate** a suite of **services**, **saving** your most precious resource: **time**.

Costa's Levels of Thinking

Medium Level - Processing

Compare, Contrast, Explain, Analyze

Example: What additional information is needed to solve this problem?



High Level - Applying

Evaluate, Generalize, Imagine, Predict

Example: What significance is this experiment to the subject you are learning?

Low Level - Gathering

Define, Identify, List

Example: What's the name of the compound H₂O?

Clients



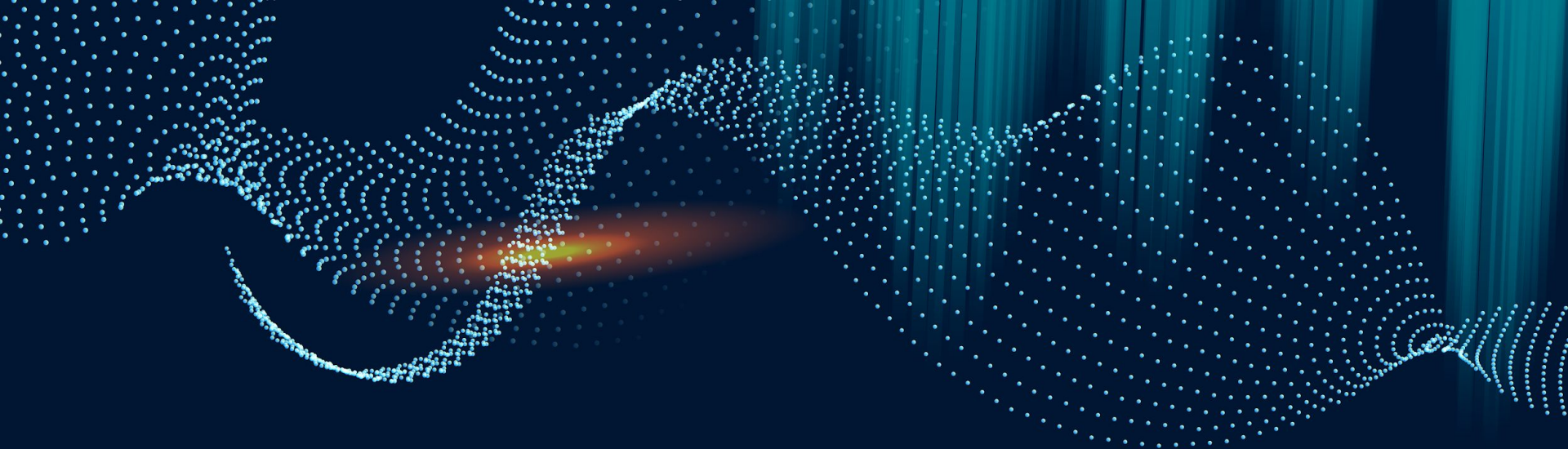
Dr. Michael
Faggella-Luby



Dr. Curby
Alexander



Dr. Liran Ma

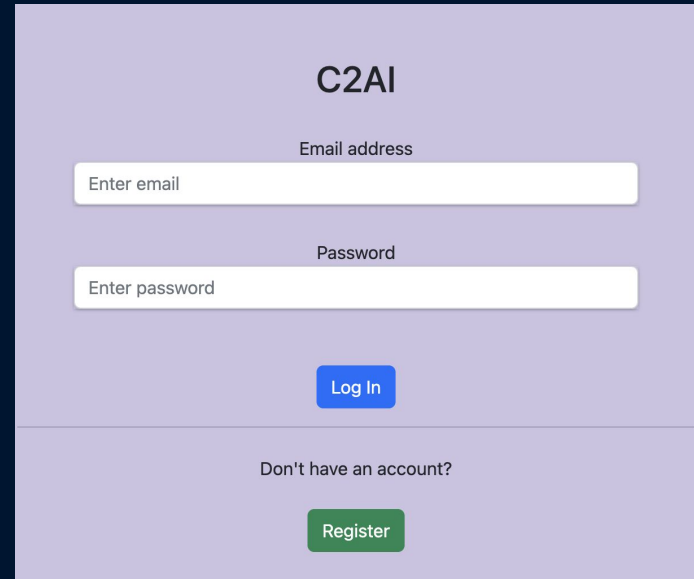


01

What We Did

Previous Iteration

- Costly API calls for every analysis
- No summarization
- Question categorization based on text keywords only
- Identifying only 6% of questions
- No Landing page



The image shows a login and registration interface for 'C2AI'. At the top center is the text 'C2AI'. Below it is the label 'Email address' followed by a white input field containing the placeholder text 'Enter email'. Underneath is the label 'Password' followed by a white input field containing the placeholder text 'Enter password'. A blue button labeled 'Log In' is positioned below the password field. At the bottom of the form, there is a link 'Don't have an account?' followed by a green button labeled 'Register'.

Milestones Completed



IN HOUSE HOSTING

AI Models are now hosted locally

Reducing costs
Run Faster



MORE FEATURES

Diarization,
question categorization,
summarization

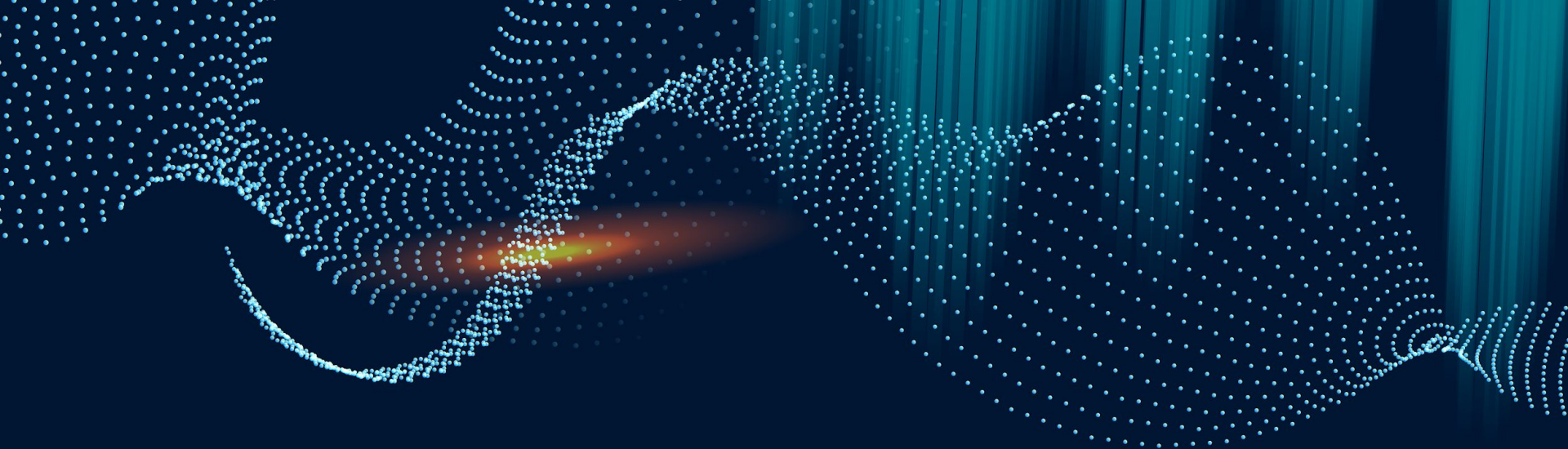
Detailed analysis



REDESIGN

Redesigned frontend
experience,
Landing page

Better UI/UX



02

The System



System Architecture

React: for Website look and function



React JS + Vite
Frontend written in React and built with Vite



D3.js
Visualization

Vite: build react code and make it readable by browser

Express.js

Backend, Handles API calls to GPU server



MongoDB

Stores reports, exported files, and audio

- Audio File
- Video File
- YouTube Link



Flask

Request Processing



Redis

Queueing of GPU-intensive tasks



WhisperX + NeMo
Transcription + Diarization of Audio Files



Meta LLaMA 3
Meta's LLaMA 3 - 8B is used to categorize questions and summarize transcripts



Each Service runs in its own Docker container to maintain independence

TCU-Hosted Nginx Web Server

TCU-Hosted GPU Server

System Architecture



Log In/ Sign Up



React JS + Vite
Frontend written in React and built with Vite



D3.js
Visualization

AWS Cognito
Handles Login



Express.js
Backend, Handles API calls to GPU server



MongoDB
Stores reports, exported files, and audio

Responsible for Login and sign up



Flask
Request Processing



Redis
Queueing of GPU-intensive tasks

- **Transcript**
- **Category of Questions**
- **Summary**

WhisperX + NeMo
Transcription + Diarization of Audio Files

Meta LLaMA 3
Meta's LLaMA 3 - 8B is used to categorize questions and summarize transcripts

Each Service runs in its own Docker container to maintain independence

TCU-Hosted Nginx Web Server

TCU-Hosted GPU Server

System Architecture



Upload for analysis



React JS + Vite
Frontend written in React and built with Vite



D3.js
Visualization



AWS Cognito
Handles Login

ex

Express.js
Backend, Handles API calls to GPU server



MongoDB
Stores reports, exported files, and audio

Middleman,
talks to
Database and
GPU Server

- Audio File
- Video File
- YouTube Link

- Transcript
- Category of Questions
- Summary

Flink
Request Processing



Redis
Queueing of GPU-intensive tasks



WhisperX + NeMo
Transcription +
Diarization of Audio Files



Meta LLaMA 3
Meta's LLaMA 3 - 8B is used
to categorize questions and
summarize transcripts



Each Service runs in its own Docker container to maintain independence

TCU-Hosted Nginx Web Server

TCU-Hosted GPU Server

System Architecture



Upload for analysis



Analyze

Creates Report

 
React JS + Vite
Frontend written in React and built with Vite


D3.js
Visualization


AWS Cognito
Handles Login

 
Express.js
Backend, Handles API calls to GPU server


MongoDB
Stores reports, exported files, and audio

- Audio File
- Video File
- YouTube Link

 **YouTube**



Flask
Request Processing



Redis
Queueing of GPU-intensive tasks



WhisperX + NeMo
Transcription +
Diarization of Audio Files



Meta LLaMA 3
Meta's LLaMA 3 - 8B is used to categorize questions and summarize transcripts

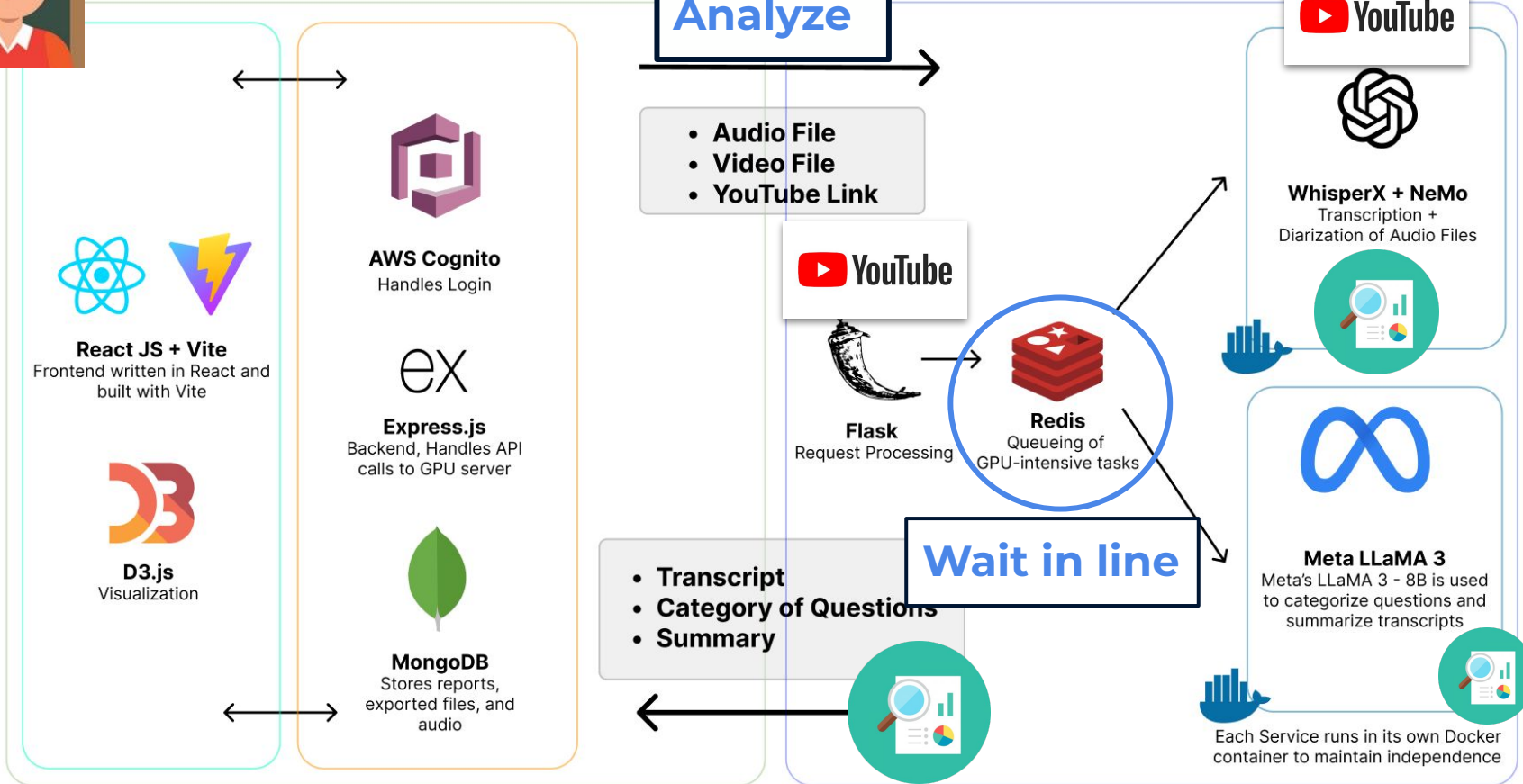


Each Service runs in its own Docker container to maintain independence

TCU-Hosted Nginx Web Server

TCU-Hosted GPU Server

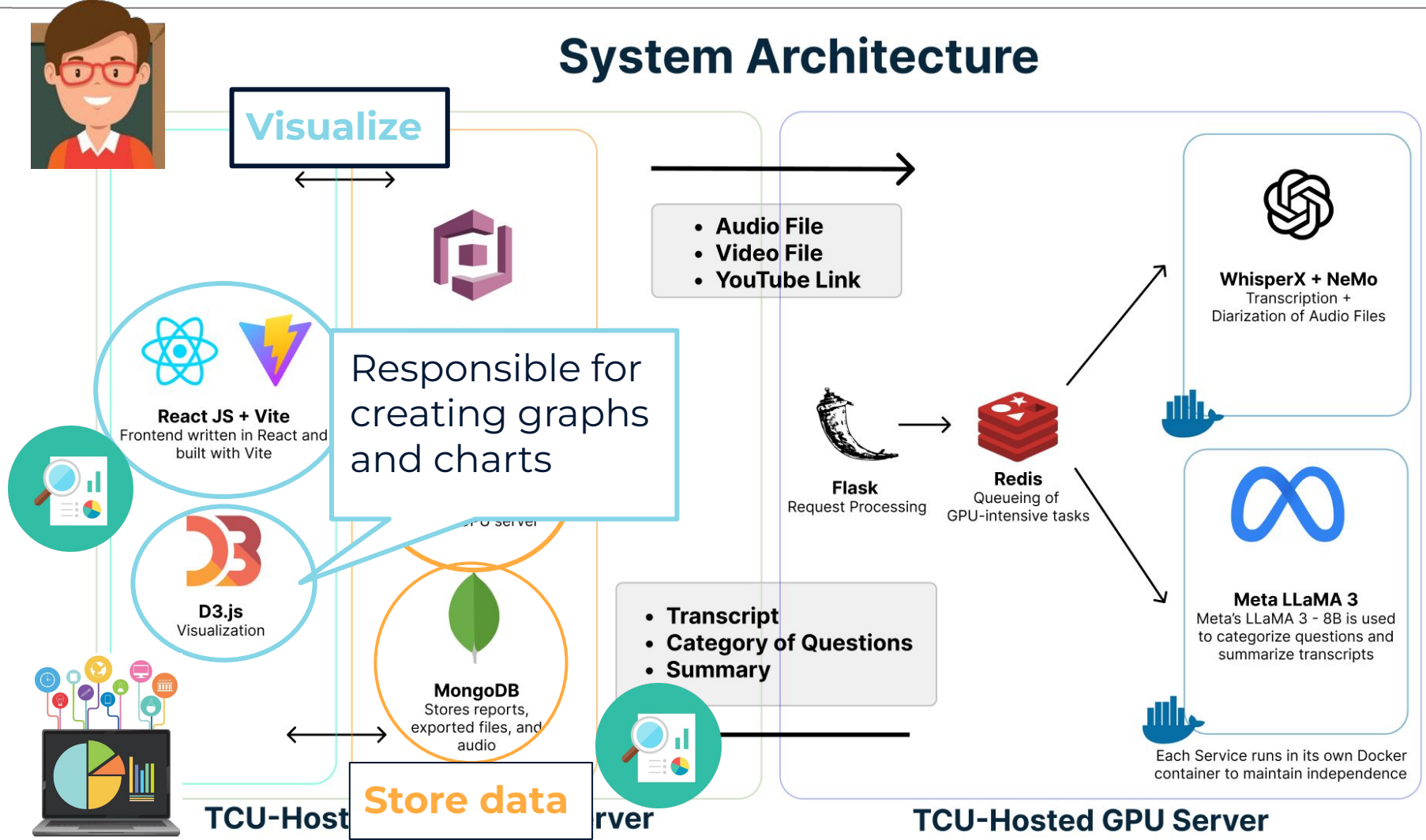
System Architecture



TCU-Hosted Nginx Web Server

TCU-Hosted GPU Server

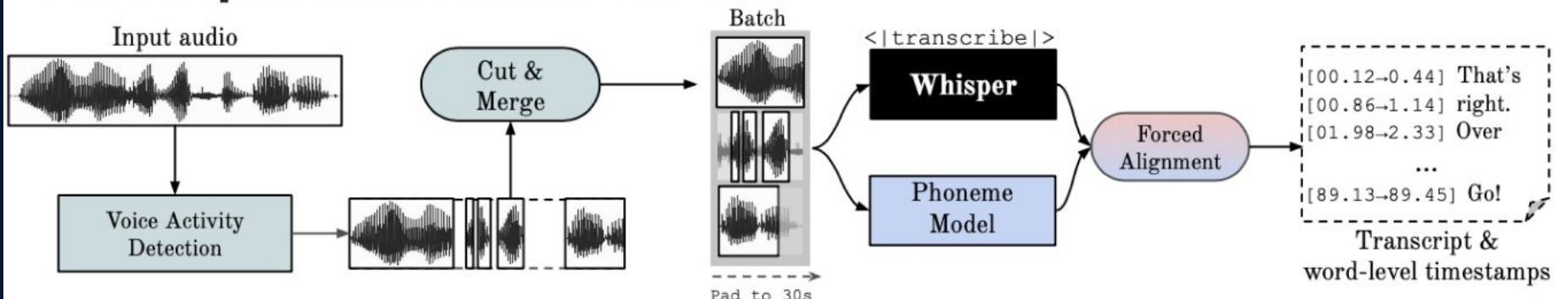
System Architecture



Powered by AI - Transcription

- OpenAI's Whisper + WhisperX (shown in diagram) transcribes the audio.
- Pyannote's speaker diarization extracts speaker embeddings and finds the speaker of each segment, in parallel.

Transcription and Diarization



Categorization & Summarization

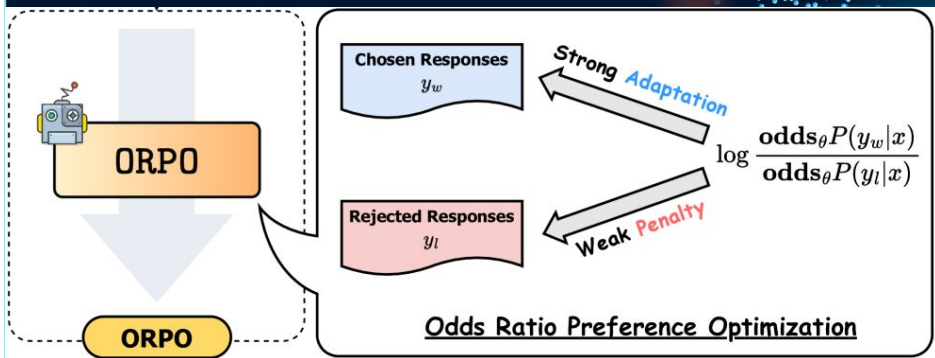
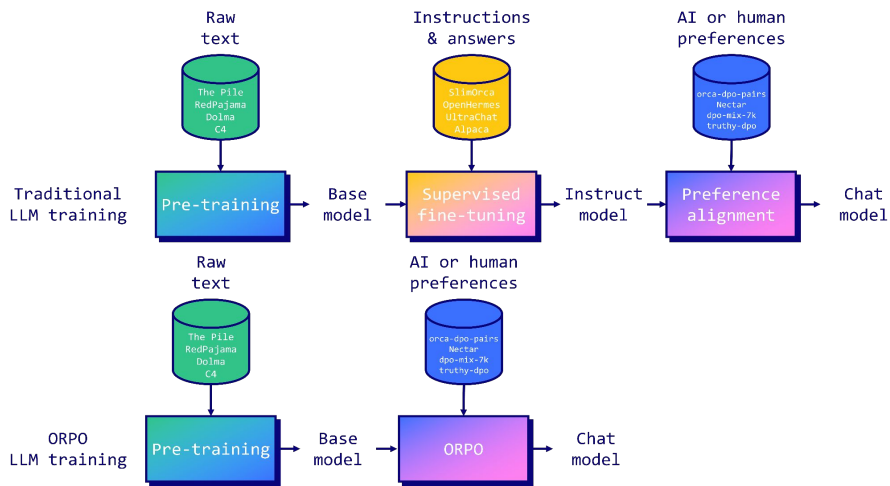
Meta LLaMA 3 8B-IT

Released April 18th, 2024

Best in its class for size

	Meta Llama 3 8B	Gemma 7B - It Measured	Mistral 7B Instruct Measured
MMLU 5-shot	68.4	53.3	58.4
GPQA 0-shot	34.2	21.4	26.3
HumanEval 0-shot	62.2	30.5	36.6
GSM-8K 8-shot, CoT	79.6	30.6	39.9
MATH 4-shot, CoT	30.0	12.2	11.0

Enhancing Categorization Performance with ORPO Fine-Tuning



ORPO Dataset

Prompt:

Categorize this according to Costa's level of reasoning:
What is the definition of a trapezoid?

Chosen (correct) response:

Level 1

Rejected response:

Level 3

Dataset was around 600 examples from looking at client's English classroom/AVID data, and used an 80/20 Train/Test split. Training data was augmented to around 4000 examples.



Other Enhancements and Insights

- **Question without context:**

- How do we do that?
 - Level 1?

- **Question with context:**

- Last class, we discussed web design. The challenge now is figuring out the most effective way to deploy. How do we do that?
 - This is processing... Level 2!

- **Adding more context** with the question significantly improved the model's ability to find the level of reasoning
- Language Models are **Few Shot Learners**

Model Performance



- After fine-tuning, the model was significantly more likely to pick the correct answer over the rejected answer
- Accuracy rose from 70% (base LLaMA 3-IT model) to 76.1% on Test set data
 - 8.6% increase over base model
- Dataset & fine-tuned model are available on huggingface for other researchers to work on classification, etc.

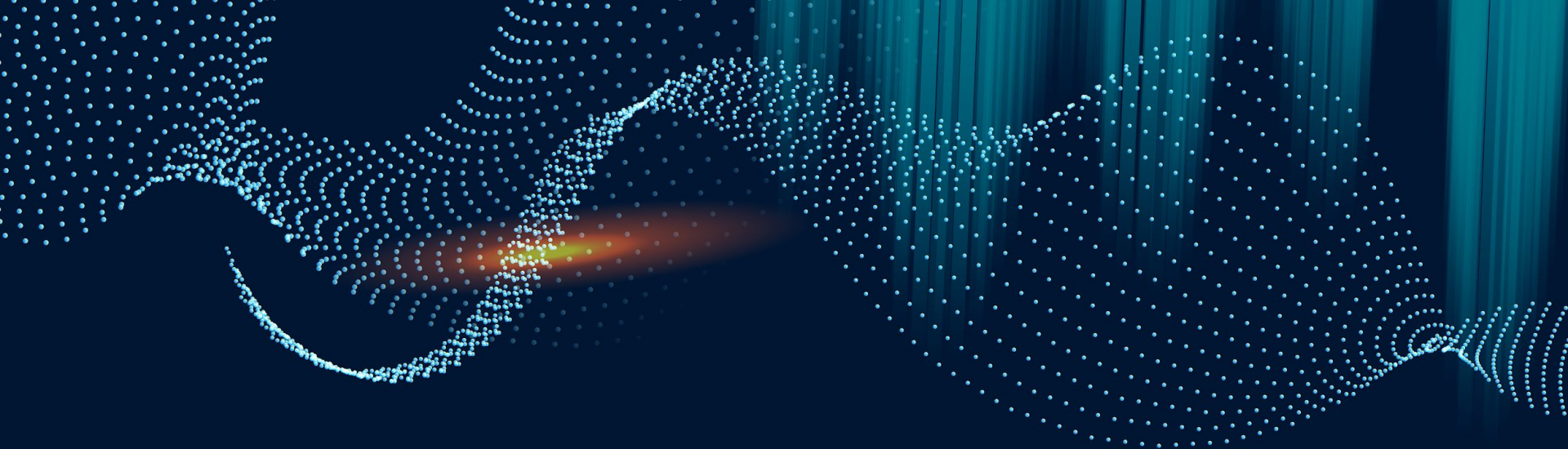
Summarization

Summarization is done through the base LLaMA model.

If the transcript is too long, we break it down into 'chunks' and summarize those

Summarization

This lecture was about modeling human problem-solving, specifically symbolic integration. The teacher discussed how humans solve problems by applying transformations to make them simpler, and then using a table of integrals to find the solution. The teacher introduced the concept of "problem reduction" and showed how to apply safe transformations to simplify a problem. These transformations include taking out constants, using the sum of integrals, and dividing polynomials. The teacher also discussed the importance of understanding the problem-solving process in order to develop a skill, and how to represent knowledge in a program.



03

Demo

A Look At Our Progress!

<https://classifai.tcu.edu/>



User creates, views, edits account information

Name*

Email*

School

Grade Level*

State*

Zip Code

Password*

Password Reset

Password must be 8 characters with uppercase, lowercase, numbers, and symbols.

Email Address*

John Nguyen
Teacher
TCU

About Me

Full Name

Email

School

Grade Level

State

Zip Code

Language-Agnostic Transcription

ClassifAI supports transcription in multiple languages, even mixed languages.

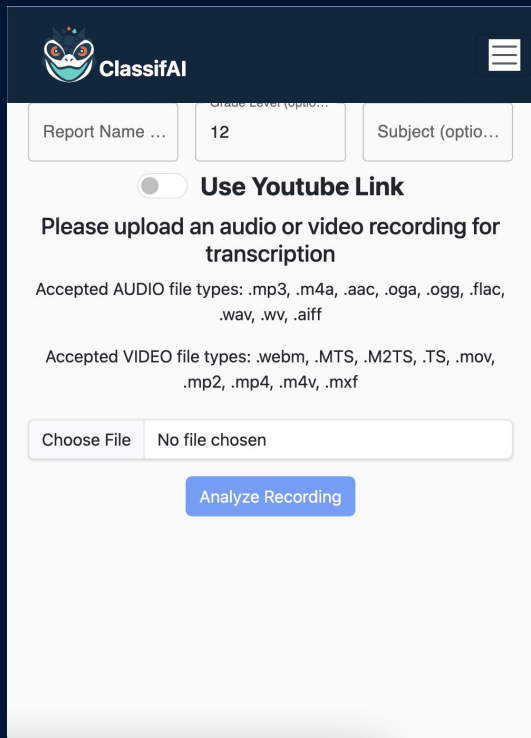
NOTE: Question categorization is trained on English questions*

Click on a cell to edit			
Start Time	End Time	Speaker (editable)	
00:00:06	00:00:07	Main Speaker	Good morning.
00:00:07	00:00:14	Main Speaker	Buenos días.
00:00:18	00:00:18	Main Speaker	Good afternoon.
00:00:21	00:00:25	Main Speaker	Buenas tardes.
00:00:29	00:00:30	Main Speaker	Good night.
00:00:32	00:00:33	Main Speaker	Buenas noches.

Full Transcript	
Click on a cell to edit	
Speaker (editable)	Text (editable)
Speaker 1	大家好,我是王颖,来自香港中文大学汉语语言学及语言习得专业。我今天教授的内容是程度补语,教授对象是中高级汉语学习者。好,我们开始上课。同学们好。老师好。非常好。今天我们上课的时间有点早。李一同学,你今天几点钟起床的
Speaker 2	?
Speaker 1	七点。七点就起了。那么李一同学,我要问你,你今天起床起得很早吗?
Speaker 2	我今天起床起得很早非
Speaker 1	常好李懿同学今天起床起得很早大家一起来一遍
Speaker 2	李

Whisper supports: Afrikaans, Arabic, Armenian, Azerbaijani, Belarusian, Bosnian, Bulgarian, Catalan, Chinese, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, Galician, German, Greek, Hebrew, Hindi, Hungarian, Icelandic, Indonesian, Italian, Japanese, Kannada, Kazakh, Korean, Latvian, Lithuanian, Macedonian, Malay, Marathi, Maori, Nepali, Norwegian, Persian, Polish, Portuguese, Romanian, Russian, Serbian, Slovak, Slovenian, Spanish, Swahili, Swedish, Tagalog, Tamil, Thai, Turkish, Ukrainian, Urdu, Vietnamese, and Welsh.

Mobile Friendly



ClassifAI

Report Name ... 12 Subject (optio...

Use Youtube Link

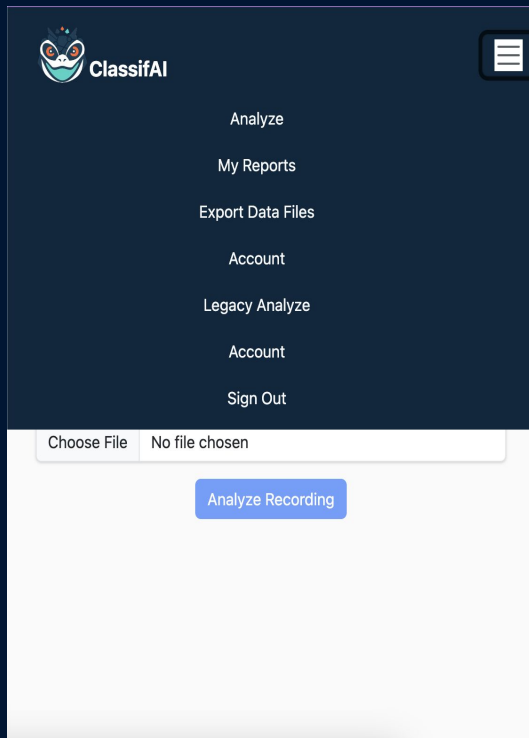
Please upload an audio or video recording for transcription

Accepted AUDIO file types: .mp3, .m4a, .aac, .oga, .ogg, .flac, .wav, .wv, .aiff

Accepted VIDEO file types: .webm, .MTS, .M2TS, .TS, .mov, .mp2, .mp4, .m4v, .mxf

Choose File No file chosen

Analyze Recording

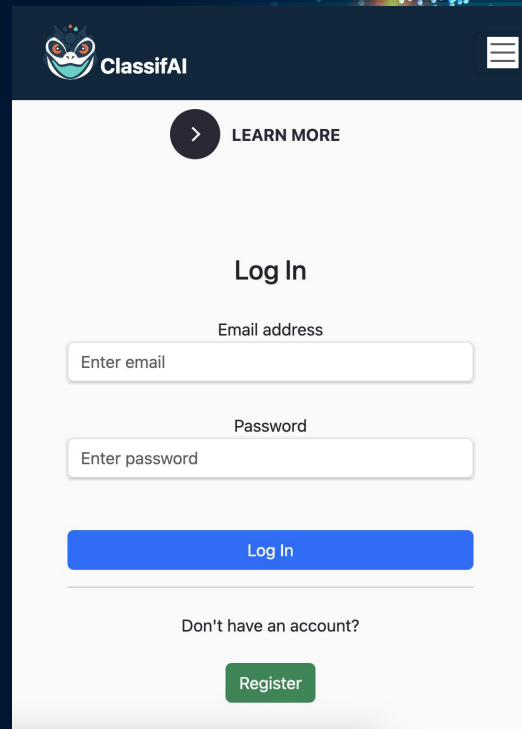


ClassifAI

- Analyze
- My Reports
- Export Data Files
- Account
- Legacy Analyze
- Account
- Sign Out

Choose File No file chosen

Analyze Recording



ClassifAI

[LEARN MORE](#)

Log In

Email address

Enter email

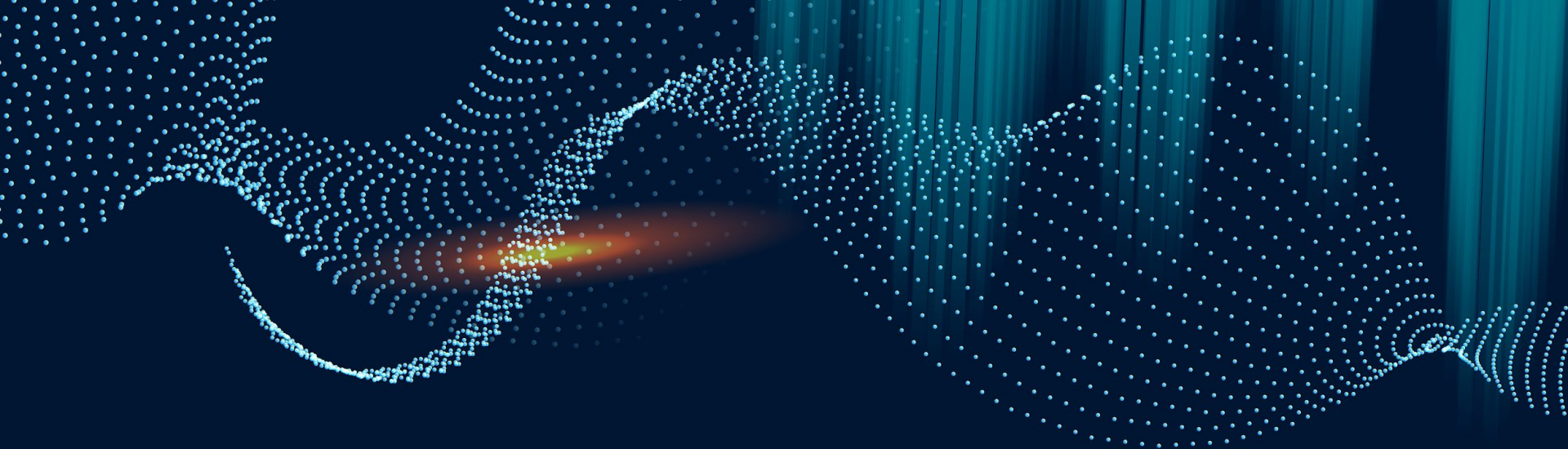
Password

Enter password

Log In

Don't have an account?

Register



04

Future

Future Iteration Improvements

AI

Accuracy

Question Categorization model is still inconclusive and off at times. Confidence level

Mobile

Mobile Views

Tables and charts are still not optimized for mobile views

Speed

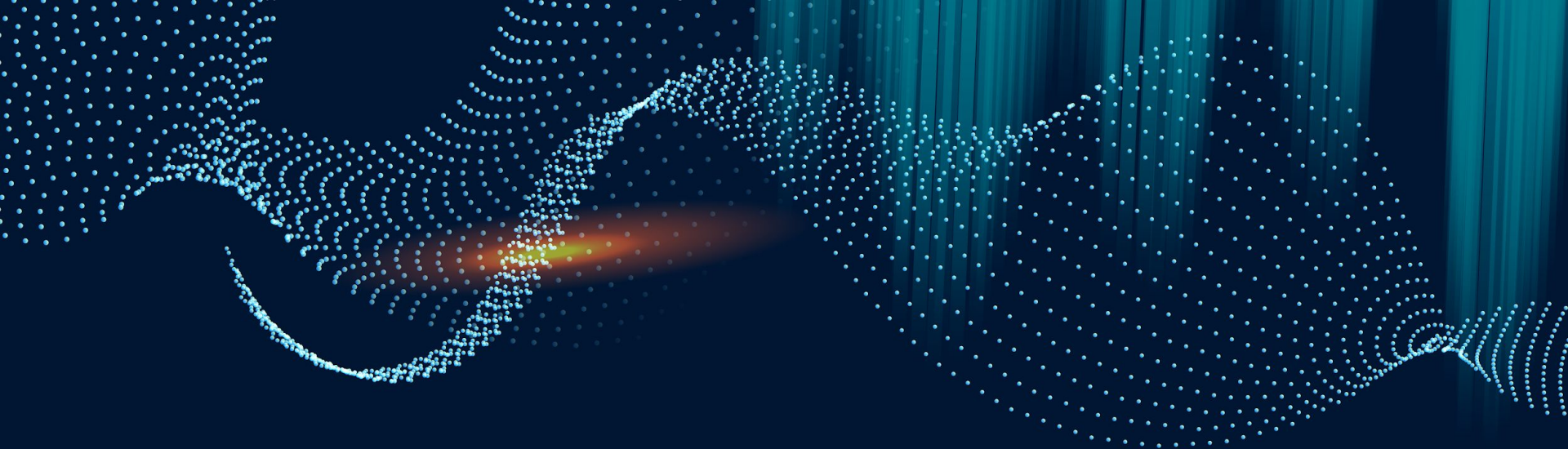
Analysis Speed

Long recordings take several minutes to analyze

Insights

New Insights

Points of confusion. Network graph to show interactions among speakers



05

**Lessons
Learned**



Client

Minimize The Gap

What they want vs what we think they want



Frontend



Backend



Acknowledgements

- **TCU Computer Science Department,**
for funding this project
- **Our Clients,**
for continued support & regular feedback
- **Dr. Bingyang Wei & Dr. Liran Ma,**
for being our comp sci faculty advisors
- **Dr. Michael Denkowski,**
for advising us on our NLP models





Thank you!

ClassifAI Team

<https://classifai.tcu.edu/>

