

MICHIGAN STATE

UNIVERSITY

Beta Presentation

On-Premises ASR Pipeline for Michigan
English

The Capstone Experience

Team Michigan State University Linguistics

Eden Seo
Jacob Caurdy
Maria Irimie
Kyle Reinhart
Yichen Ding

Department of Computer Science and Engineering
Michigan State University

Spring 2022



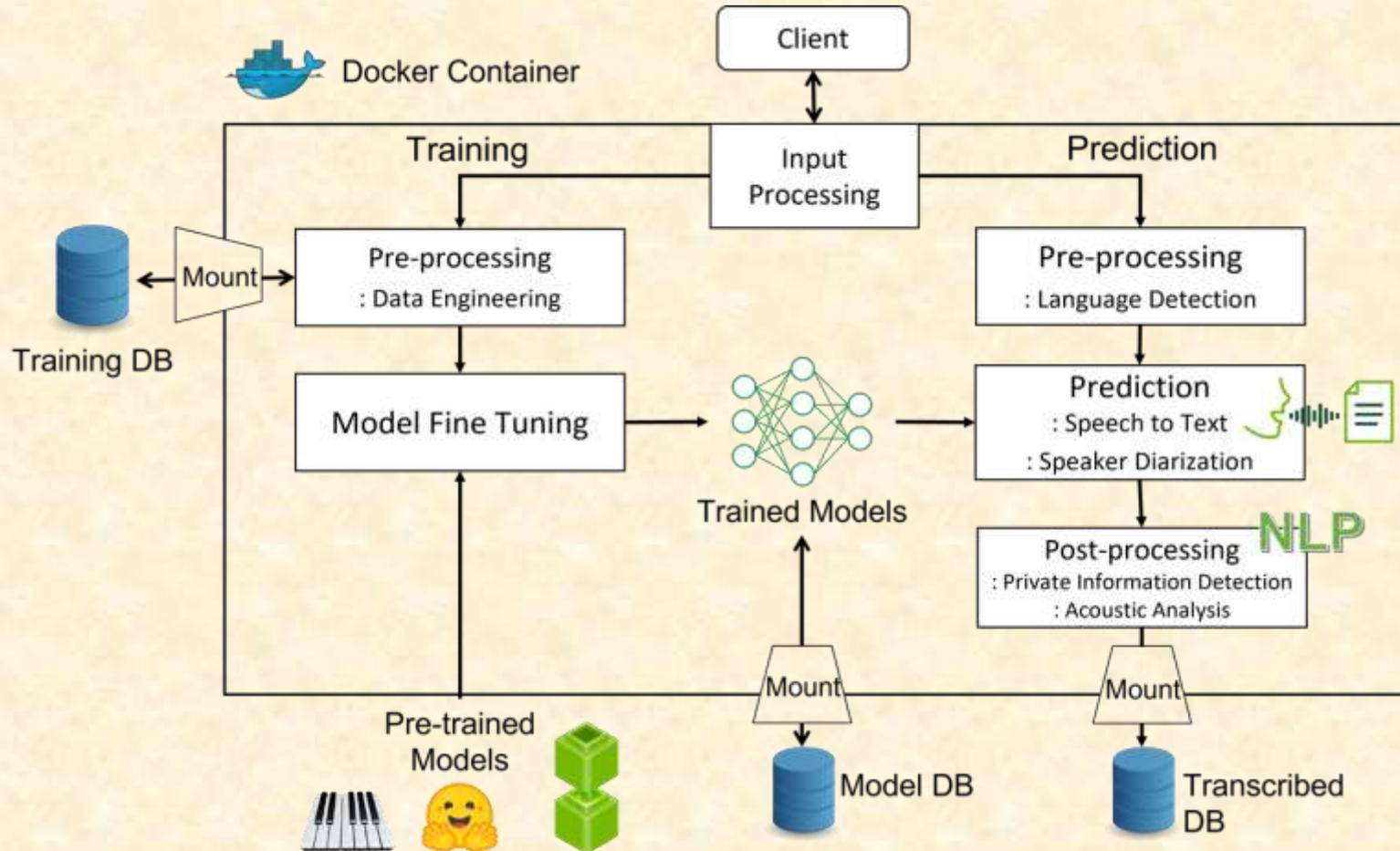
*From Students...
...to Professionals*

Project Overview

- Problem: The Department of Linguistics, Languages and Cultures' MI Diaries project needs to replace their outsourced speech-to-text software from Google to an in-house package
- Motivation: Privacy and monetary concerns
- Solution: All-in-one speech-to-text machine learning pipeline which takes in audio files and outputs time-aligned transcripts



System Architecture



Researchers will upload a .wav file and receive a well-timed transcript

MI Diaries



Upload

Train

Transcriptions

Drag and Drop or Select File

0short_audio.wav

▶ 0:00 / 1:31 — 🔊 ⋮

Transcription Time: 59.853 seconds

Diarization Time: 27.006 seconds

Average ASR Confidence: 0.982%

Start (sec.)	End (sec.)	Speaker	Transcript	Confidence
1.287	2.418	SPEAKER_00	APRIL TWENTY THIRD	0.9960397
2.992	13.859000000000002	SPEAKER_00	TODAY WAS PRETTY GOOD WORK IS GOING WELL BUT YOU OF EM SENT OUT AN EMAIL SINCE THE BEGINNING OF COVET UNTIL THE END OF THIS FISCAL	0.98890316
14.585	19.698	SPEAKER_00	O THEY'RE ASKING PEOPLE TO LAY OFF VOLUNTARILY OR FURLOUGH UNTIL JULY FIRST	0.9819576
20.12	28.827	SPEAKER_00	I TOLD AMY I CANNOT DO THIS BECAUSE RIGHT NOW I AM THE ONE BRINGING IN MONEY FOR THE FAMILY NO STIMULUS NO P P P NO	0.9895649
29.705	34.363	SPEAKER_00	SHE SAID THERE IS NO PLAN YET AND THAT THEY STILLWILL LIKELY TAKE VOLUNTEERS FIRST	0.97181463
35.105	43.492	SPEAKER_00	THE PRESIDENT ANNOUNCED THAT THE THUNDERBIRDS AND BLUE ANGELS WILL FLY OVER CITIES AS A TRIBUTE TO THE HOSPITAL WORKERS THEY NEED THE	0.9854489
44.572	48.217000000000006	SPEAKER_00	NEW YORK CITY HAS FORTY FIVE PER CENT OF THE CASES AND DEATHS IN	0.98999745



Select which model and what model to use.

MI Diaries



Upload

Train

Transcriptions

Chose which aspect to retrain

Speech to Text



Select a Model or Upload a model

Model1



Drag and Drop or Select File

Speech Model set to Model1



Uploaded .wav files' transcripts are automatically stored in the archive

MI Diaries

LiLaC

Upload Train

Transcriptions

Transcript Archive

02/19/2022	audio_file_02192022.wav	transcript_02192022.csv	Show Transcript
------------	-------------------------	-------------------------	-----------------



View the transcript from previous uploads

MI Diaries



Upload Train

Transcriptions

Transcript Archive

Get Stored Data

04/04/2022 | 0hello_test.wav | transcript_4.csv | Show Transcript

transcript_4.csv

▶ 0:00 / 0:03 — 🔊 ⋮

Start (sec.)	End (sec.)	Speaker	Transcript	Confidence
1.895	2.789	SPEAKER_00	HELLO HILLO HILLO	0.9340012

04/04/2022 | 0short_audio.wav | transcript_3.csv | Show Transcript

04/04/2022 | 0hello_test.wav | transcript_2.csv | Show Transcript



What's left to do?

- Features
- Stretch Goals
- Other Tasks
 - HF Language Model improvements
 - Add ability to record files directly



Questions?

?

?

?

?

?

?

?

?

?

