Sprint 2

Multilingual Subtitle System

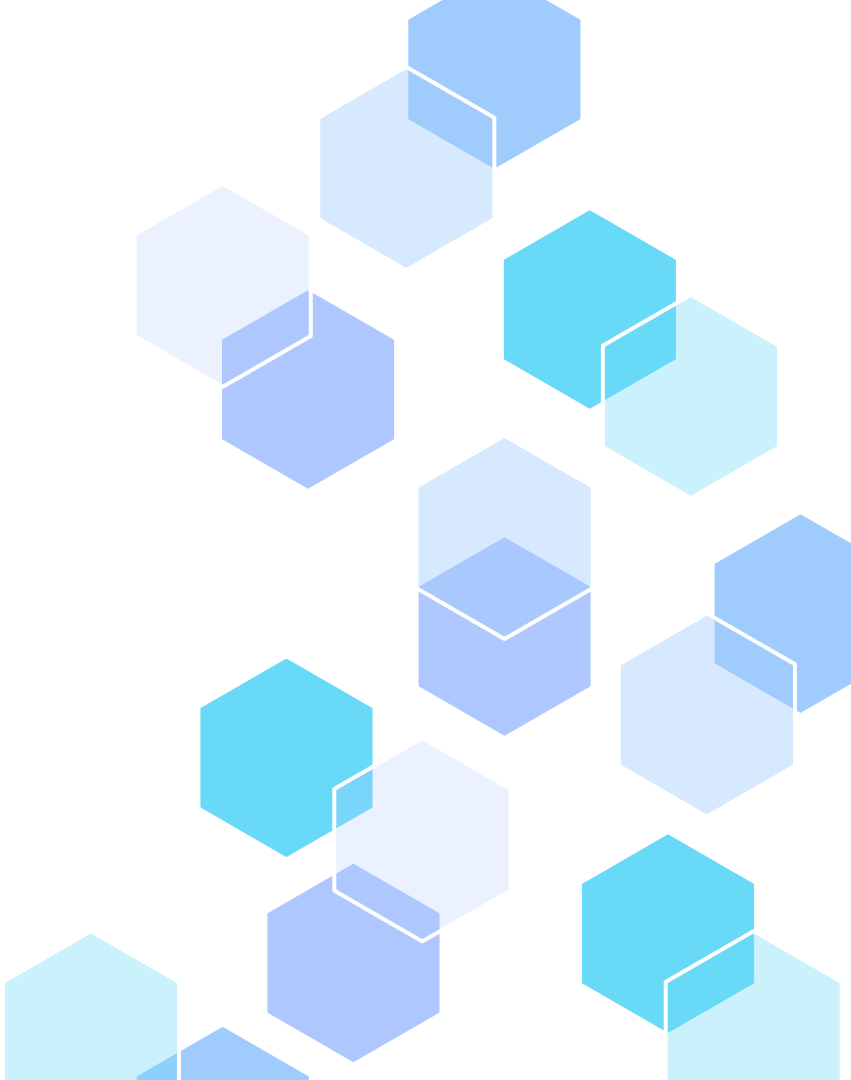
Welcome to my presentation, let’s dive deep into

Table of contents

01

Introduction

Why I chose this project?



Introduction

What

- I love movies!
- Sometimes movies lack subtitles

Why

- Absence of subtitles limits media access for non-native speakers
- Ensure everyone enjoys media regardless of the language spoken.

How

- Detects the audio language, transcribes it, and translates it into subtitles.
- Use Machine learning models to detect language spoken

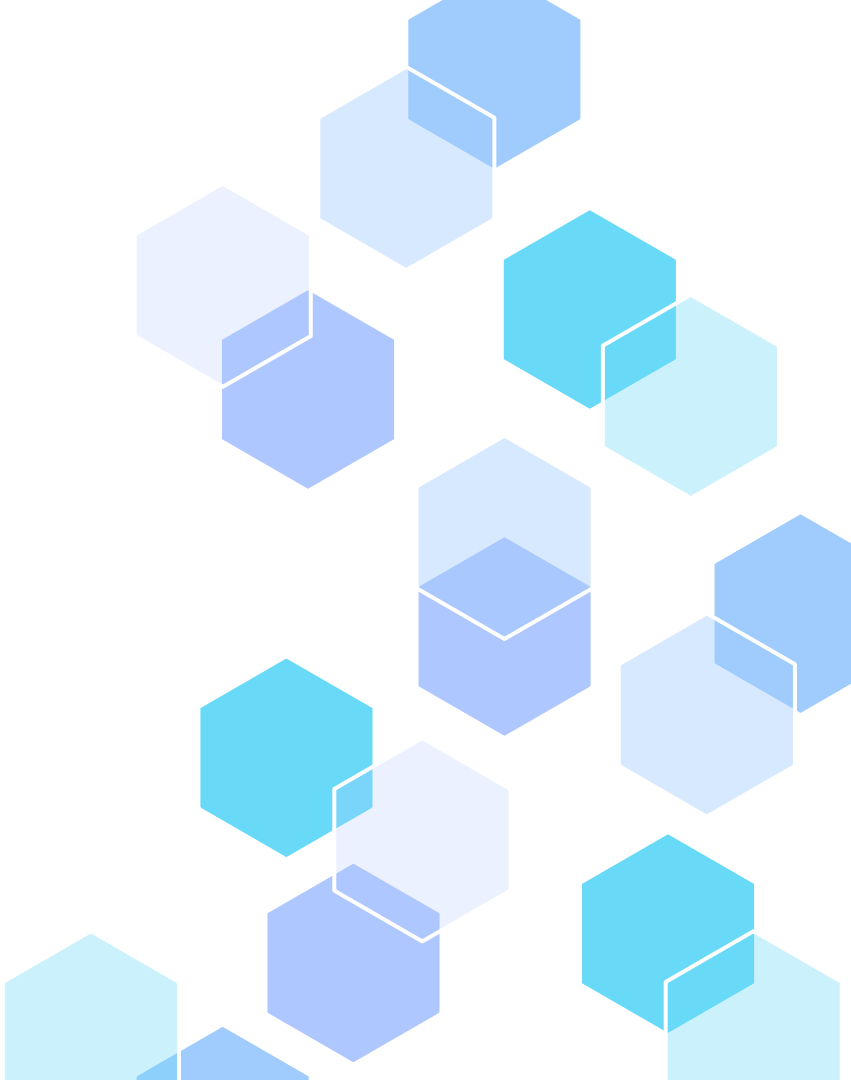
Potential Impact:

- Enhances accessibility for non-native speakers and the hearing impaired.
- Facilitates language learning and understanding.

02

Techinques

Extraction, EDA, Preprocessing



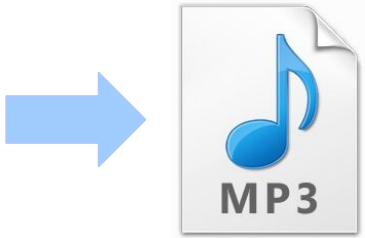
Extraction

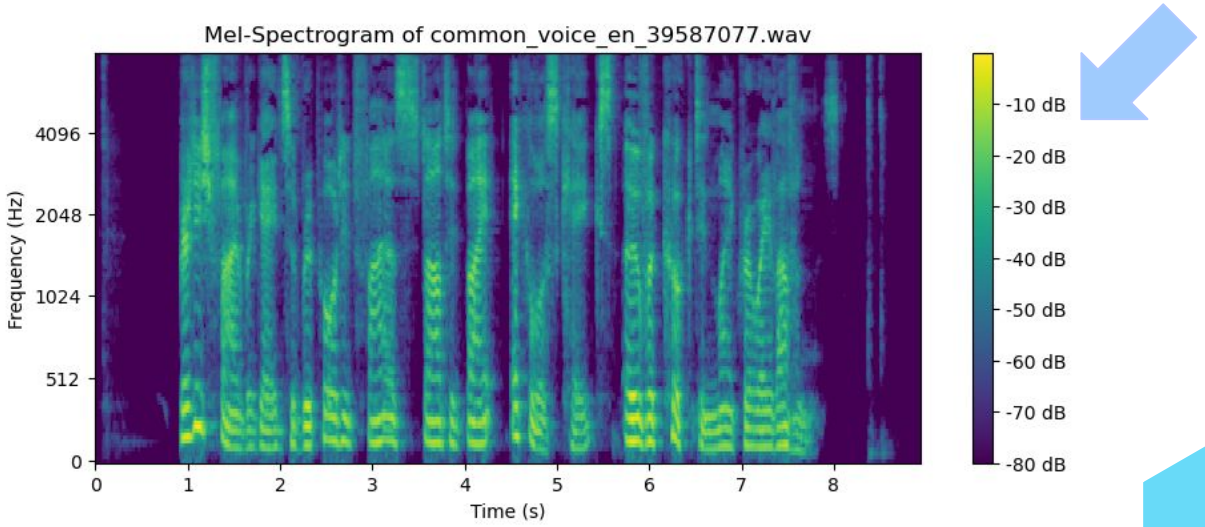
- Extracted audio files for Chinese, English, Spanish, Arabic
- Converted to .wav for better audio quality
- Converted .wav to spectrograms for models
- Resized spectrograms to 128x128
- Flattened pixels from spectrogram to feed models

Data collection techniques

Common Voice

moz://a



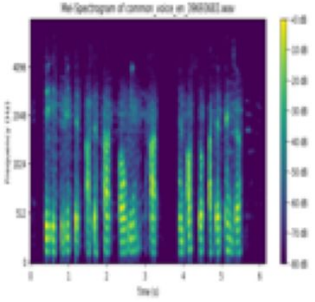


EDA (Exploratory Data Analysis)

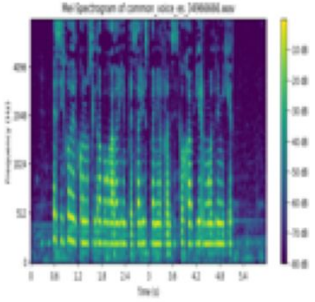
- Reshaped images
- Made figures to showcase uniqueness of each language
- Made separate figures for each language on their average pixel values

EDA (Exploratory Data Analysis)

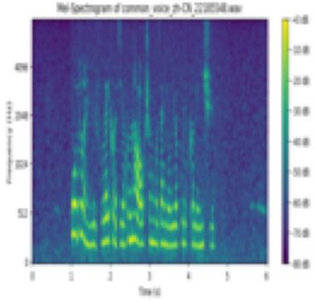
english (First Image)



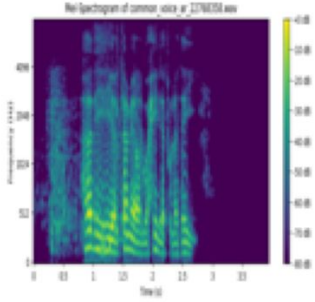
spanish (First Image)



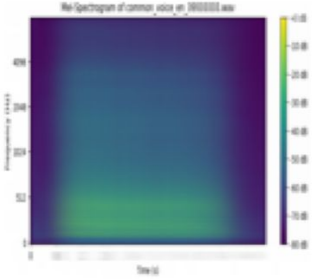
chinese (First Image)



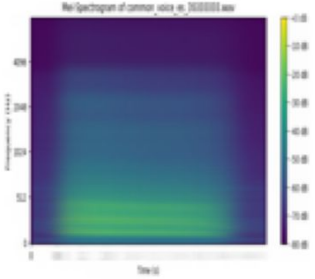
arabic (First Image)



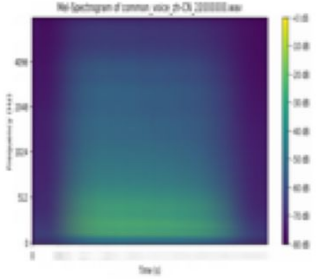
english (Mean Image)



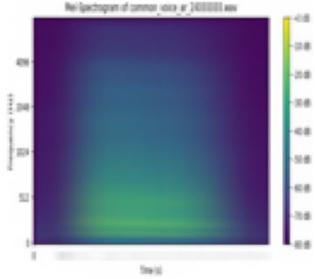
spanish (Mean Image)



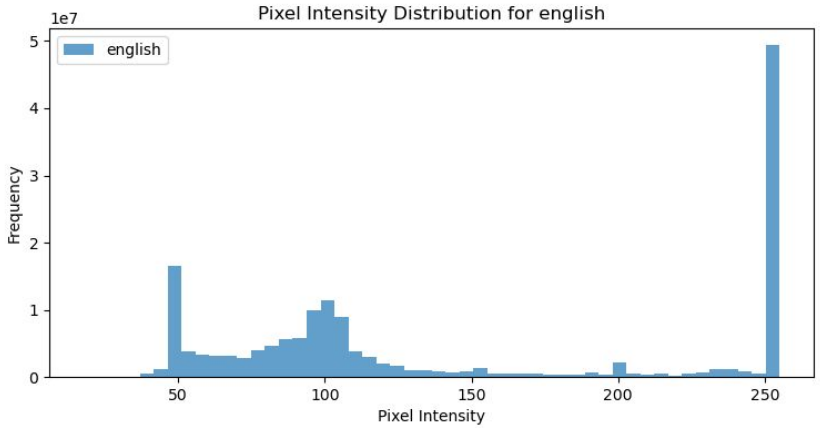
chinese (Mean Image)

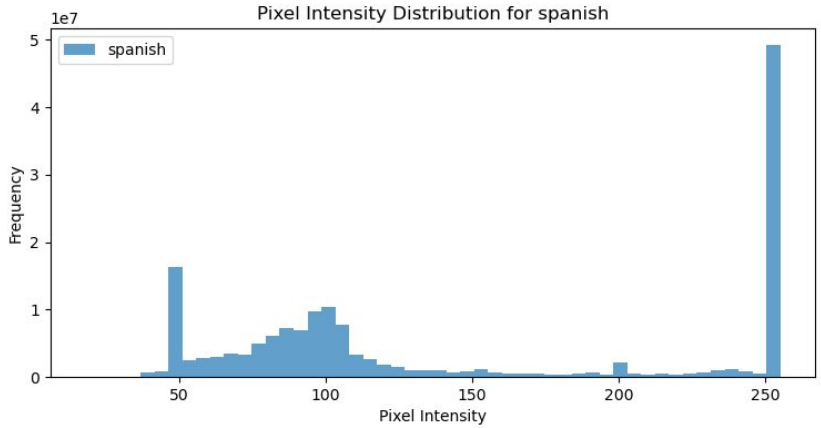


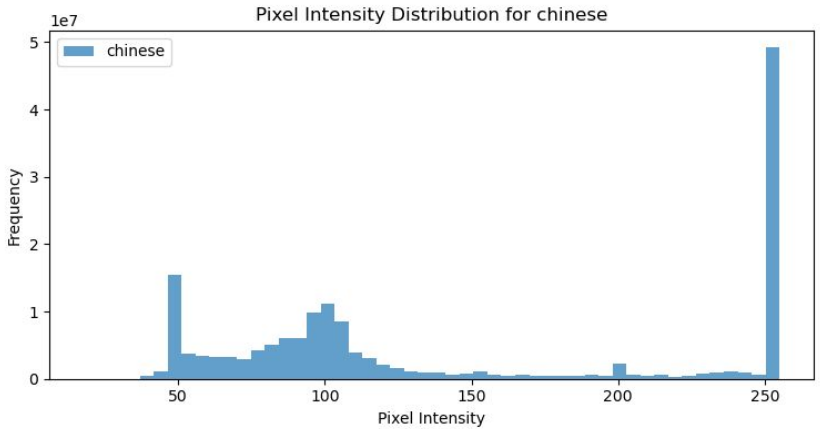
arabic (Mean Image)

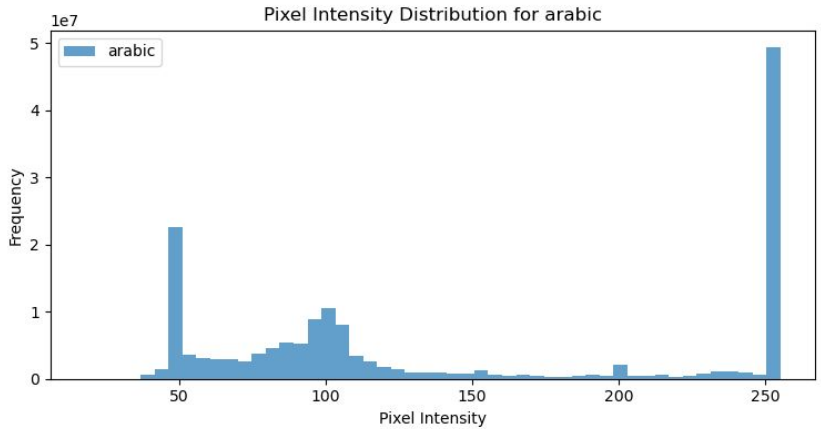


EDA (Exploratory Data Analysis)









Preprocessing

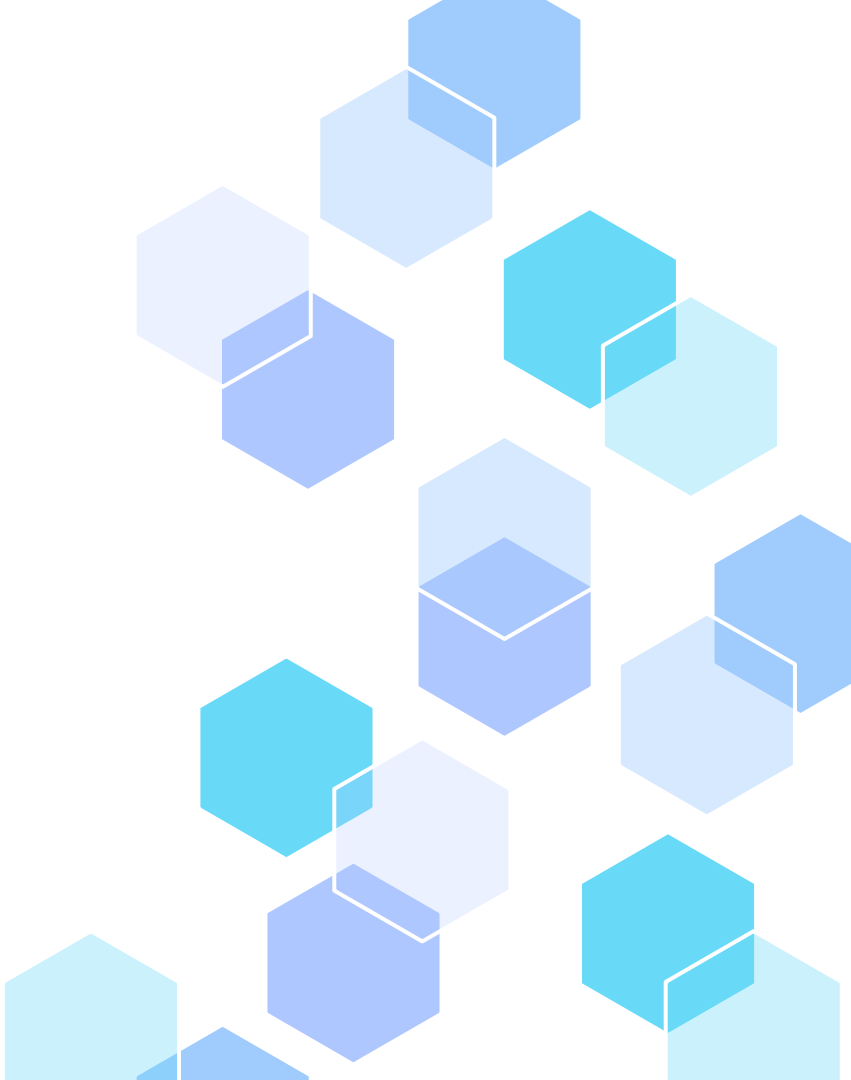
- Normalized the flattened images with standard scaling
- Reduced dimensionality with PCA
- Label encoded my languages

But this is all mambo jumbo... all you need to this that this is done before adding my data to my models to make it more accurate.

03

Models

Cool machine learning models!



Model Info

I used 3 models (Logistic reg, Random Forest, XGBoost)

Logistic Regression

- Accuracy: 97%, no overfitting
- 98% Precision & 98% Recall

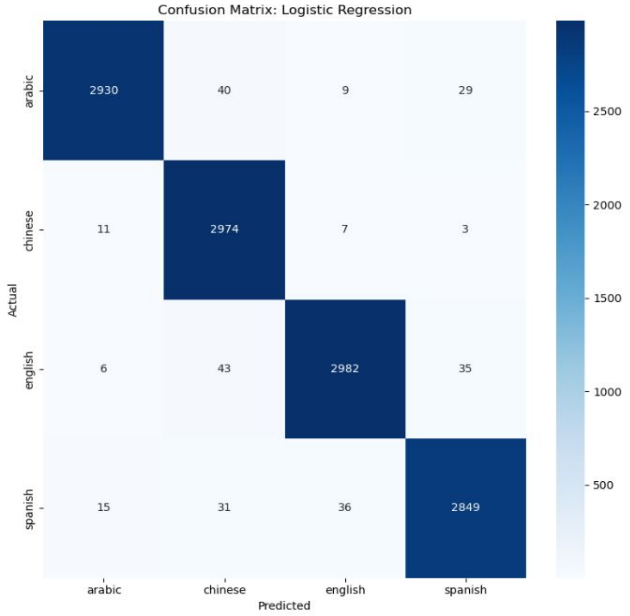
Random Forest

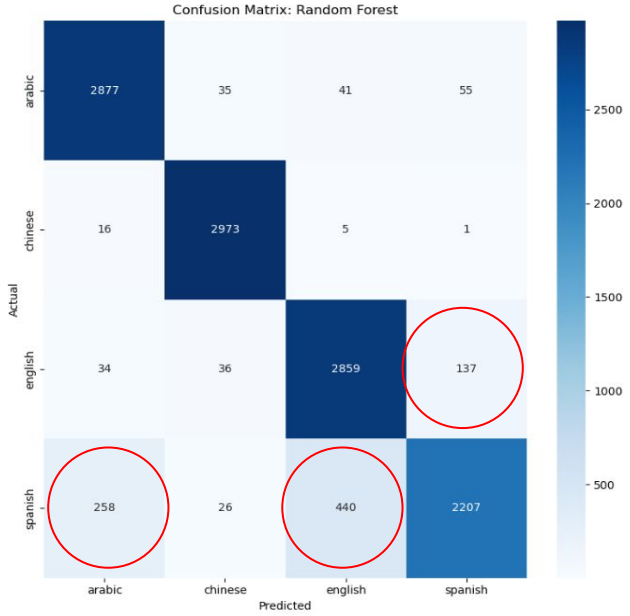
- Accuracy: 90%, overfitting ~10% difference
- 91% Precision & 91% Recall

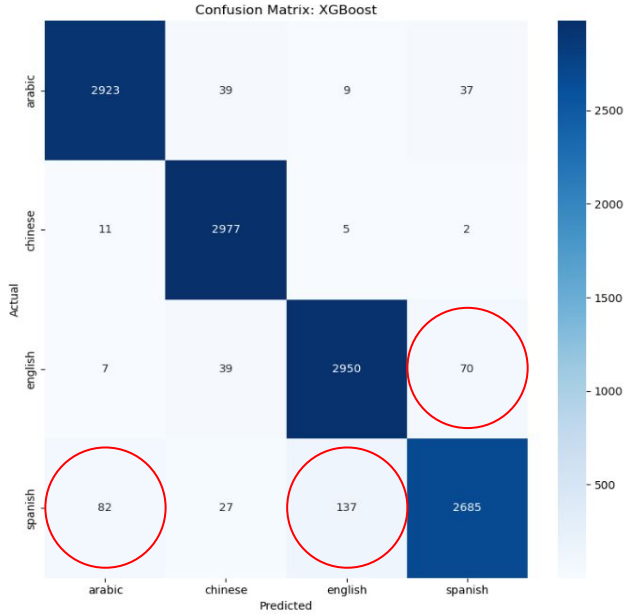
XGBoost

- Accuracy: 96%, no overfitting
- 96% Precision & 96% Recall

Linear Regression is the winner!



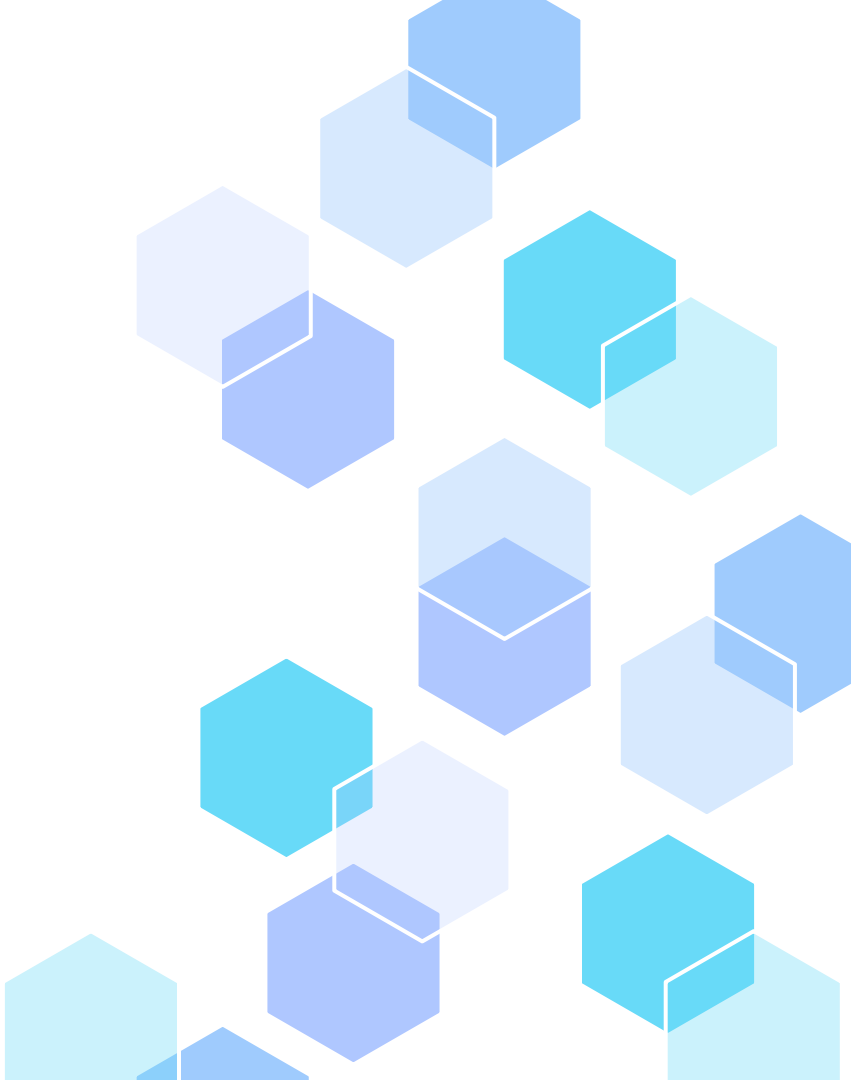




04

Next Steps

Advanced Models, Design, Functionality



Coming Soon...

Adv Modeling

CNN: Improve prediction accuracy

Whisper AI: Used for transcribing and translation

Functionality

Interface: Create a user friendly front end

Subtitles: Customizable subtitle options

Deployment

Cloud: Deploy on the cloud (AWS)

Sharing: Release to the public!

Any Questions?