Josh King

12/30/2025

INFO 698

Greg Chism

<div align="center">INFO 698 Project Plan</div>

<div align="center">**Language Translation Website**</div>

0.   Overview

This project aims to be an all in one multi-feature webpage derived from a large GitHub repository that will be able to take in audio and detect language, transcribe, and translate text to a couple languages. I plan to use some pre-trained models such as openai's Whisper(-small) for transcription and another model for translation while potentially building my own model such as an XGBoost or RandomForestClassifier for the language detection to decide which datasets to use from huggingface/Fluers. I hope to be able to have each model work together once individually created and in the end phase of the project put together a simple website that allows users to record audio directly to the site and have it become transcribed/captioned and then translated to a selected language with multiple options available. With more possible features with this projects potential if time allows.

1.   Initial Project Set up/Audio to Text – Language Transcription

<div align="right">**TIMELINE (1/2026-2/2026)**</div>

First step is creating the repository, get the imports needed for my virtual env, and solving any other initializing task that is needed such as research and what tools I will need to start this project before fully operating. Then once ready to start I will begin using a model to take in .wav files of audio to then to have it accurately transcribe it into text of the correct language. This initial phase will be a core data science competency for being able to think abstractly and decompose a large scale project to build from scratch.

2. **Language Detection Model – TIMELINE (2/2026-3/2026)**

Create a language detection model from scratch that will train on different language strings of text along with a language target variable to assign the correct language. This will be a classification model most likely using an XGBoost or RandomForestClassifier in order to give a probability of which language the text from the Step 1. Model transcription gives after being trained on huggingface/Fleurs data. This will be a pivotal model for later as it will be used to indicate which button datasets to use for the website when the next step of translation occurs and also gives me a chance to test a core competency of data science of being able to manipulate data to serve a purpose.

3. **Language Translation Model – TIMELINE (3/2026-4/2026)**

Make use of the final model to take the translated and language detected text from the models above and finally translate it to the desired language. This will require a pre-trained model and most likely some adjustments on my end to work smoothly for my final product. It showcases my ability to apply multiple complex models to work together and to be able to use powerful resources like huggingface to use.

4. **Website Creation + Functionality – TIMELINE (4/2026-5/2026)**

The final month of the semester I will create a basic website that will be able to use these models that were used and created to then be functional for users to try for themselves with a nice interface and no coding required. This will be a competency of using the skills of a data scientist to make multi-media functional art using the data and coding skills acquired throughout my degree.

FINALE

Overall this project should be challenging in many different aspects of using complex models, intensive data, and tying them all together to create a fully function webpage for users to translate direct audio to text. I am excited to put my skills to the test and make this work and perhaps one day further the capabilities of this project for future ideas!