CSE-489: SELECTED TOPICS IN DATA SCIENCE

Project

Submission Date:

MAY 28, 2024

Submitted to:

Dr. Alaa Hamdy

Submitted by:

Omar Mohamed Ibrahim Alsayed 19P7813
Zakaria Sobhy Abd-ElSalam Soliman Madkour 19P2676

Table of Contents

Task 1
Data 1
Approaches 2
1. Hidden Markov Model (HMM) 2
2. Long-short term memory (LSTM) 2
Results 2
1. HMM 2
2. LSTM 3

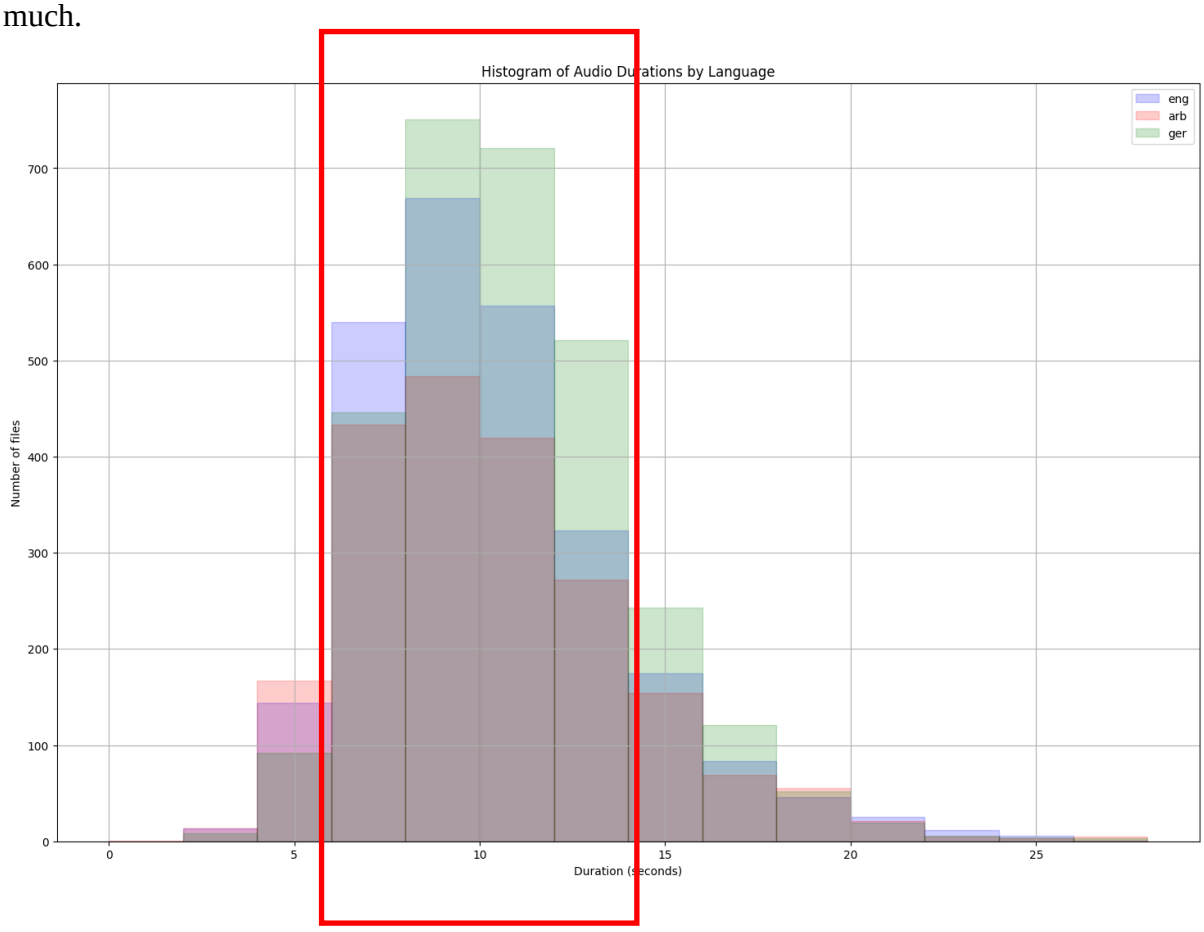
Task

This project is about spoken language identification where we are given an audio file with one speaker speaking in one of the designated language and the system will classify the language spoken in this recording.

Data

The data we are using is the Fleurs dataset. We only took three languages namely English, Arabic, and German. This dataset is already split into train, test, and validation sets. We used the preexisting split data. We need to unify the length of audio files input to the classifier. The solution was to pad small audio files to match large ones. But this introduced a problem to the models as the amount of padding was significantly large which caused the models to be biased towards the zeros of padding.

We filtered the audios by taking only files with durations ranging between 6-14 seconds which from the following histogram resembles nearly 70% of the dataset. This way when we pad 6s. audio files to match with 14s. the model behavior won't deteriorate very much.



Approaches

1. Hidden Markov Model (HMM)

We trained three HMMs one per each language and then for the classification we would run the audio file by the three models and the model with the highest probability would be the classification of this audio file. Each model had 10 hidden states and was trained on 50 iterations.

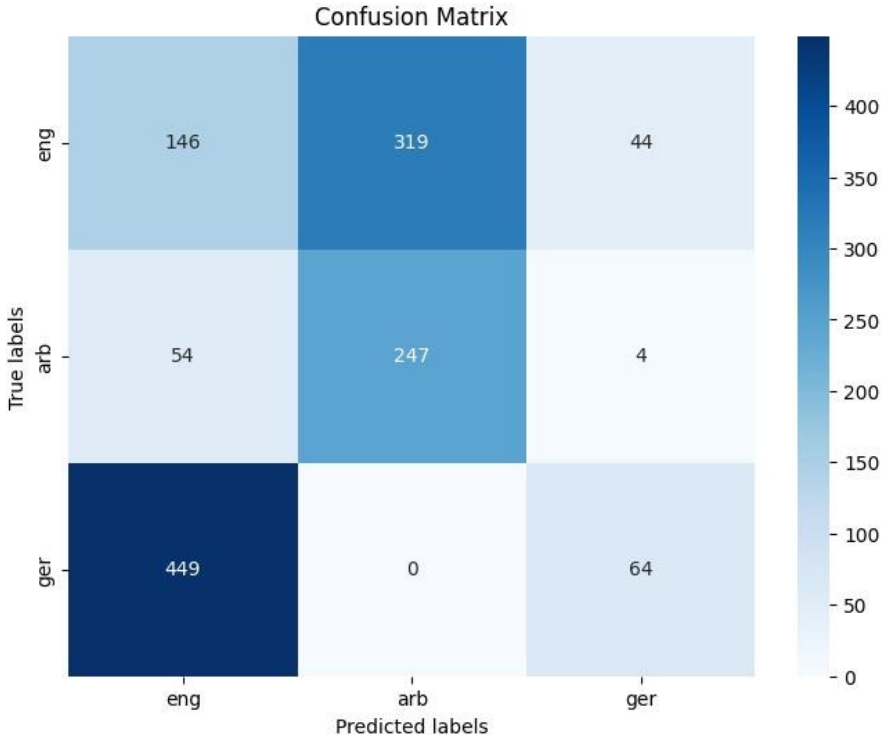
2. Long-short term memory (LSTM)

We trained a recurrent neural network with the LSTM acting as our temporal capturing mechanism. Then the hidden state of the last time step was feed to a fully-connected layer that outputs the class probability. The LSTM had 5 layers, dropout of 0.5 and hidden state vector of size equals 512.

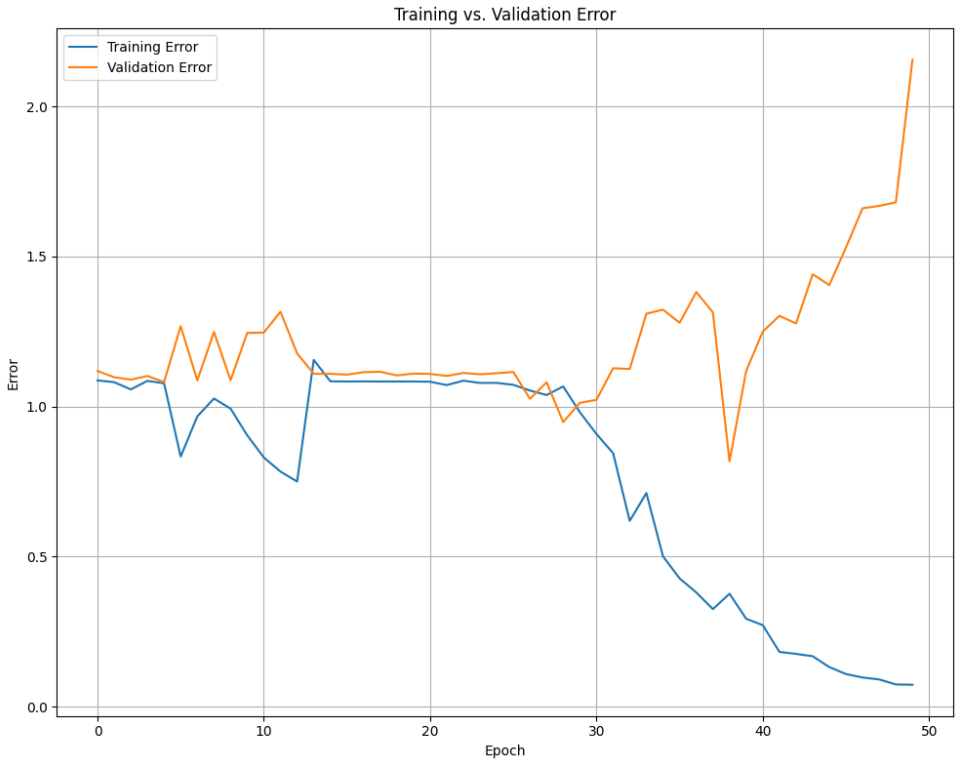
Results

1. HMM

Test accuracy of 34.44%



2. LSTM



Test accuracy of 71.29%