ECE408 Spring 2023

# ENGLISH-LANGUAGE ACCENT CLASSIFICATION

McCormack Chew, Sydney Haupt, Yusuf Salim

# OBJECTIVE:

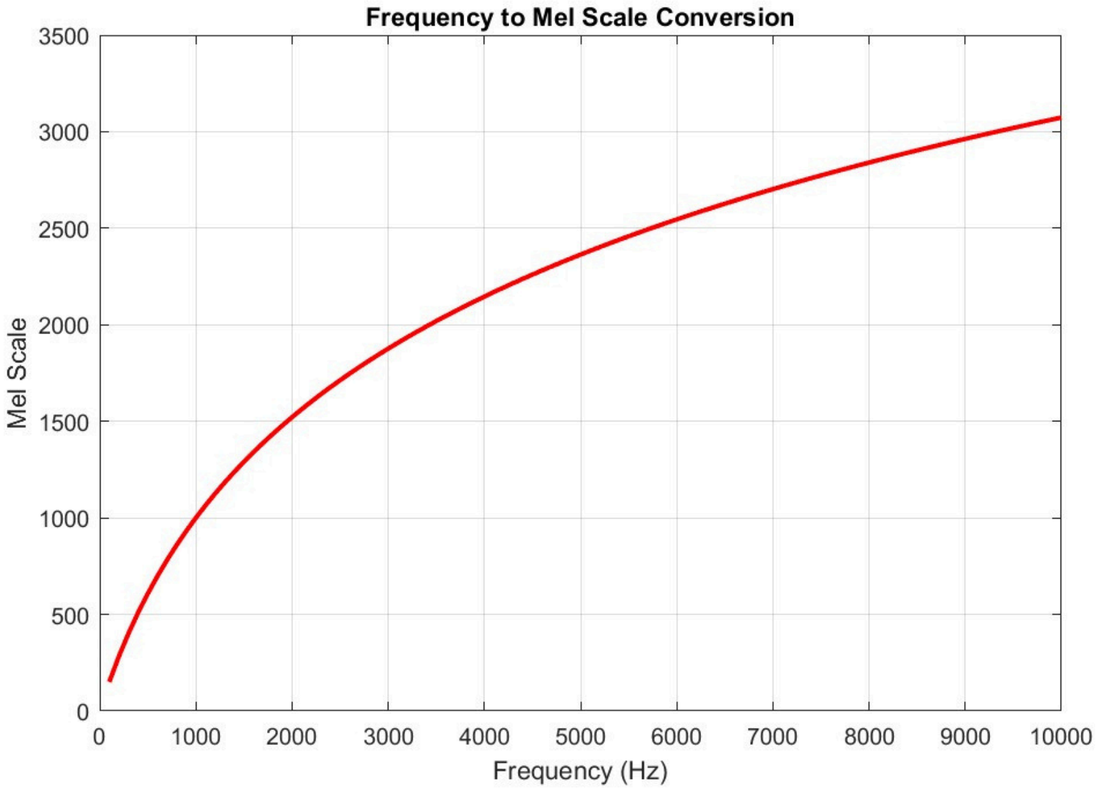
Detect and categorize English-language accents based on country of origin.

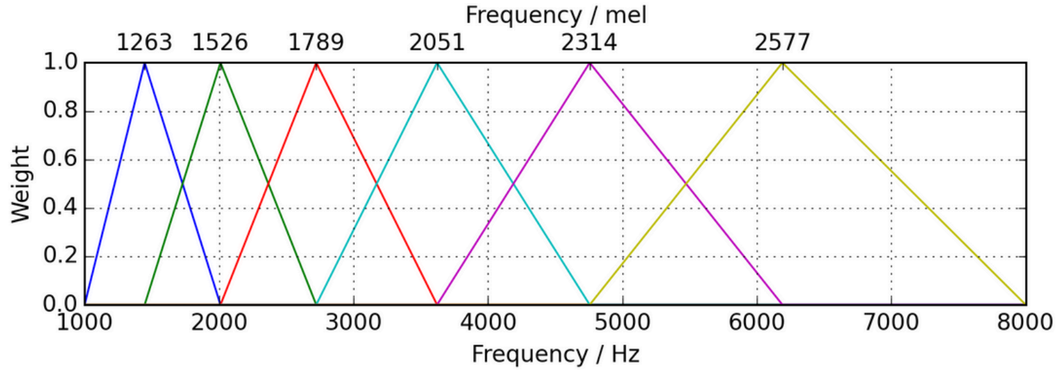
# Related Work

| *Dataset* | *Features* | *Approach* |
| --- | --- | --- |
| • ***UCI "Speaker Accent Recognition" Dataset***<br>• ***GMU "Speech Accent Archive"***<br>• ***Wildcat Corpus of Native- and Foreign-Accented English***<br>• ***INTERSPEECH 2016*** | ***"The MFCC feature was the best performing feature"***<br><br>*from: Features of speech audio for deep learning accent recognition by Singh, Pillay, Jembere* | • ***Binary vs Multiclass Classification***<br>• ***Elementary methods (kNN, SVM, Decision Tree)***<br>• ***Modern methods (DNN, CNN, RNN)*** |

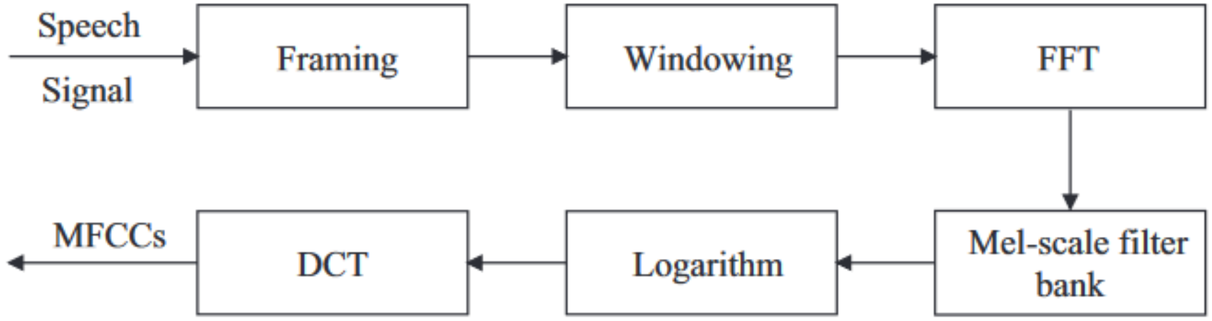
# Feature Extraction:

# Mel-Frequency Cepstral Coefficients

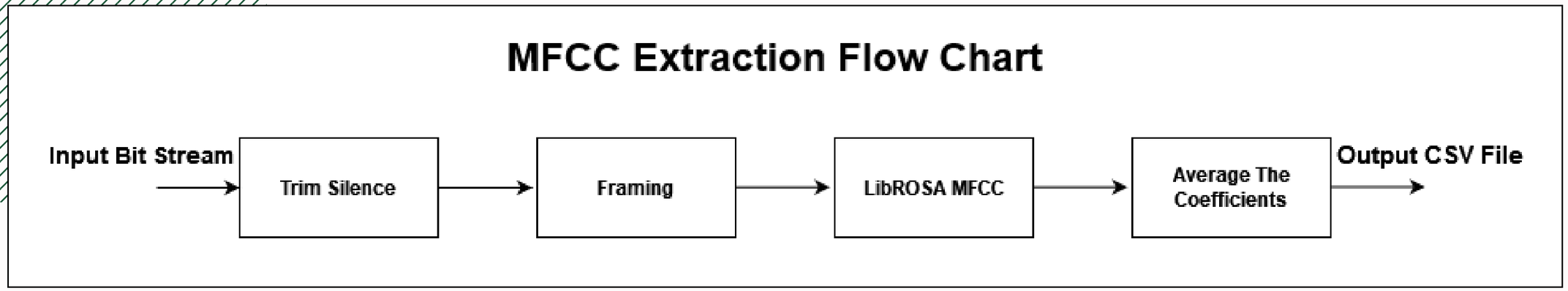


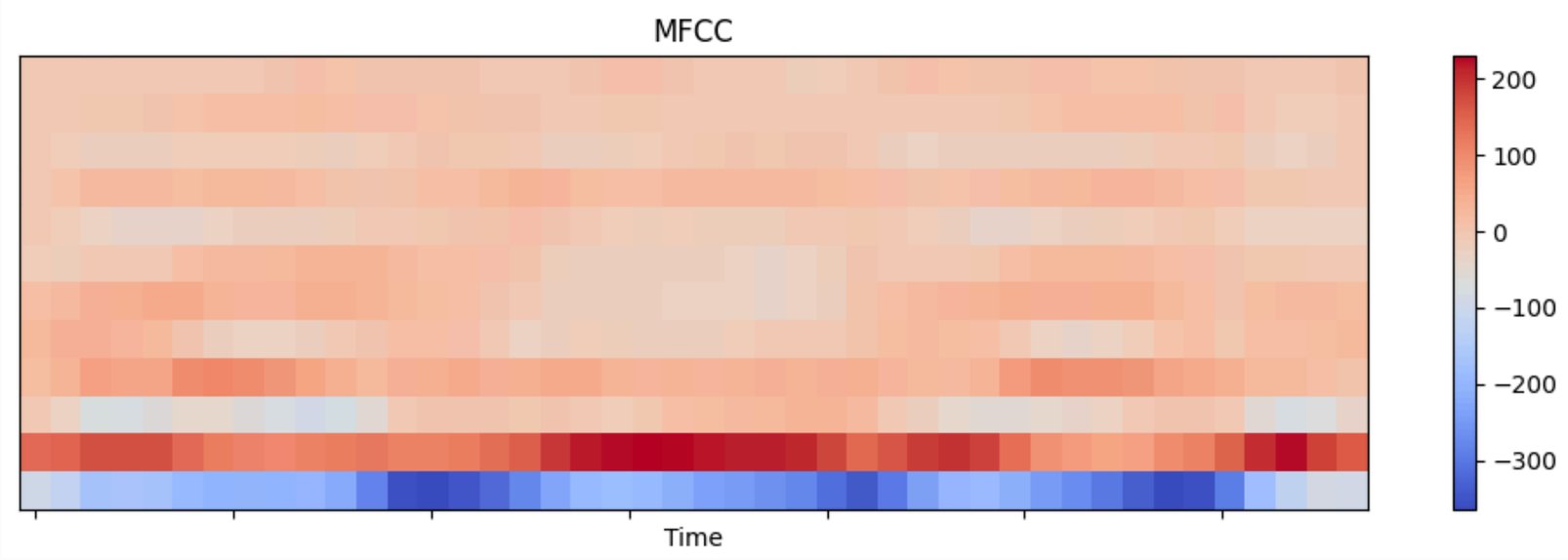




[1]




[2]




## MFCC Coefficients

# DATASET:

**Initial Dataset:**
**UCI “Speaker Accent Recognition Dataset”**

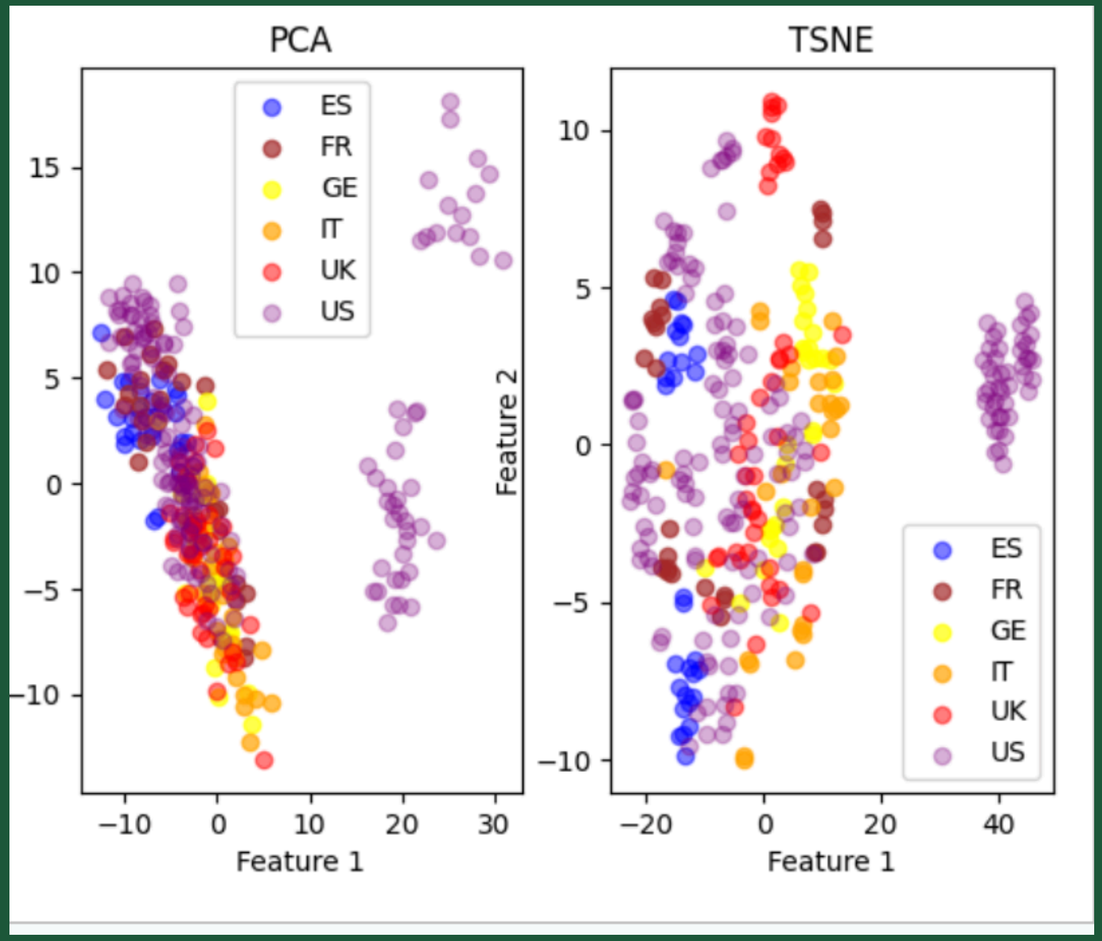
**Format:** .csv file
**Features:** 12 MFCCs
**Samples:** 329
**Countries:** Spain, France, Germany, Italy, United Kingdom, United States
**Issue:** Inability to add additional data to set

**GMU “Speech Accent Archive” Subset**

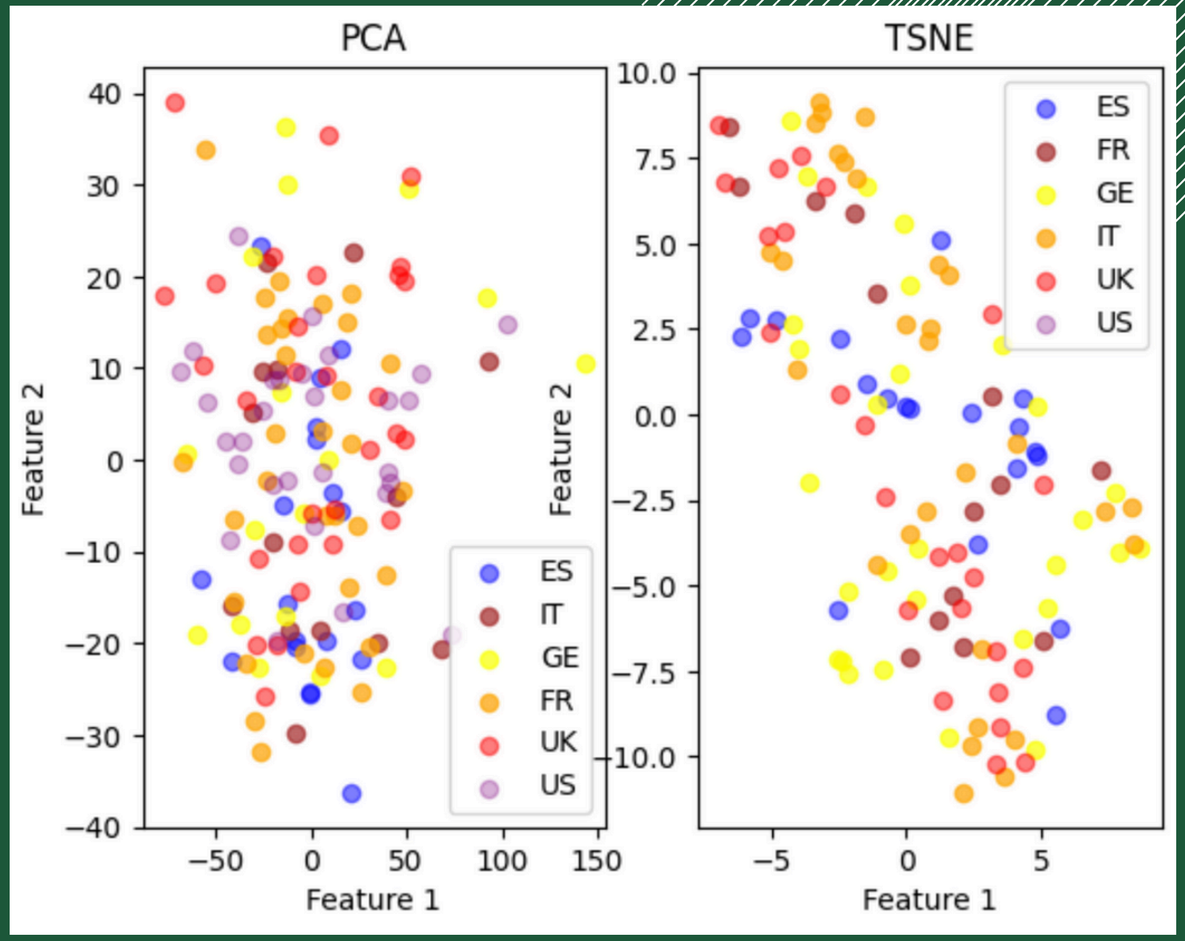
**Format:** .mp3 files
**Samples:** 222/2140
**Countries:** Spain, France, Germany, Italy, United Kingdom, United States
**Issue:** Not all samples actually accented

**GMU “Speech Accent Archive” Subsubset**

- ★ Removed Unaccented Samples (Samples: 141)
- ★ Equalized Classes
- Normalized MFCCs
- Found Mean and Standard Deviation of MFCCs




UCI DATASET




REDUCED GM DATASET

# CLASSIFICATION METHODS

k-Nearest Neighbors (kNN)

Support Vector Machine (SVM)

Multilayer Perceptron (MLP)

Convolutional Neural Network (CNN)

Recurrent Neural Network (LSTM)

# PERFORMANCE

GMU dataset, Unaccented Samples Removed, Classes Equalized, MFCC features

| | kNN | SVM | MLP | CNN |
|---|---|---|---|---|
| Train Set Accuracy | 100% | 93.52% | 100% | 100% |
| Test Set Accuracy | 45.65% | 52.08% | 52.08% | 53.24% |

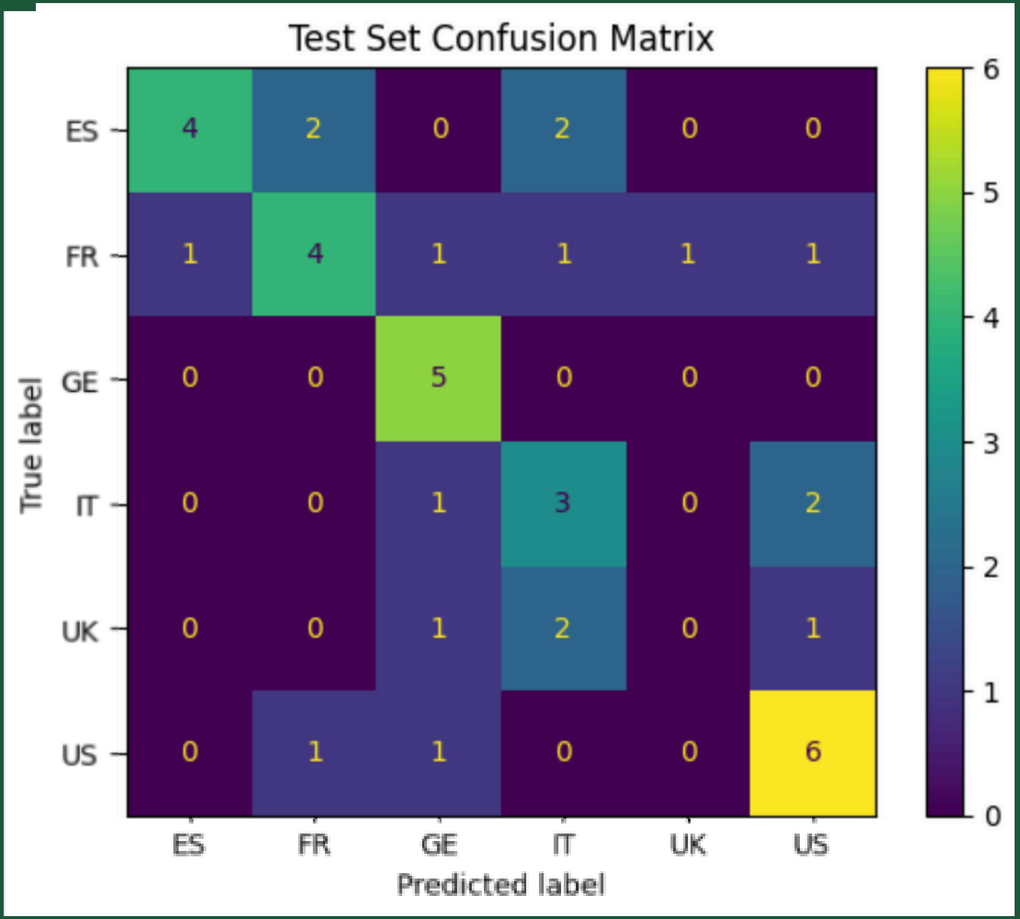
UCI Dataset

| | kNN | SVM | MLP | CNN |
|---|---|---|---|---|
| Train Set Accuracy | 100% | 95.51% | 100% | 100% |
| Test Set Accuracy | 80.17% | 88.19% | 88.25% | 87.24% |

*LSTM not applicable for these datasets

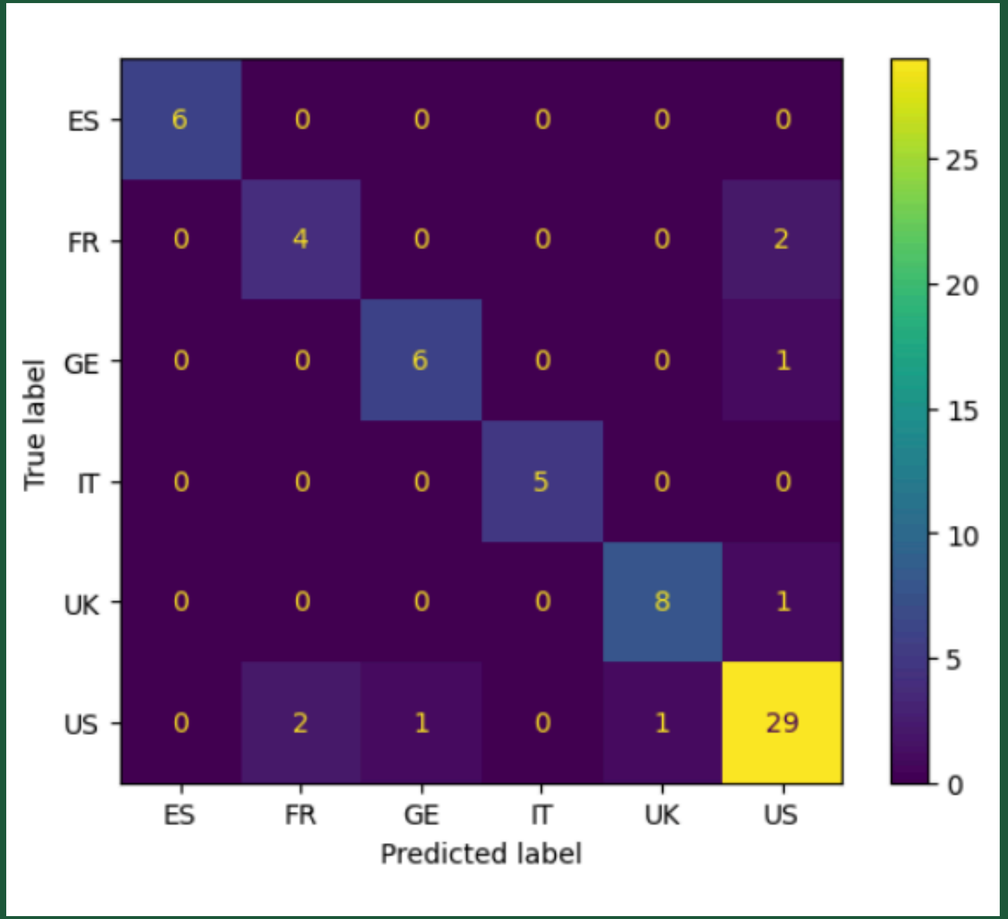
# SVM

## GMU Dataset




C = 2, gamma = 0.1, kernel = rbf

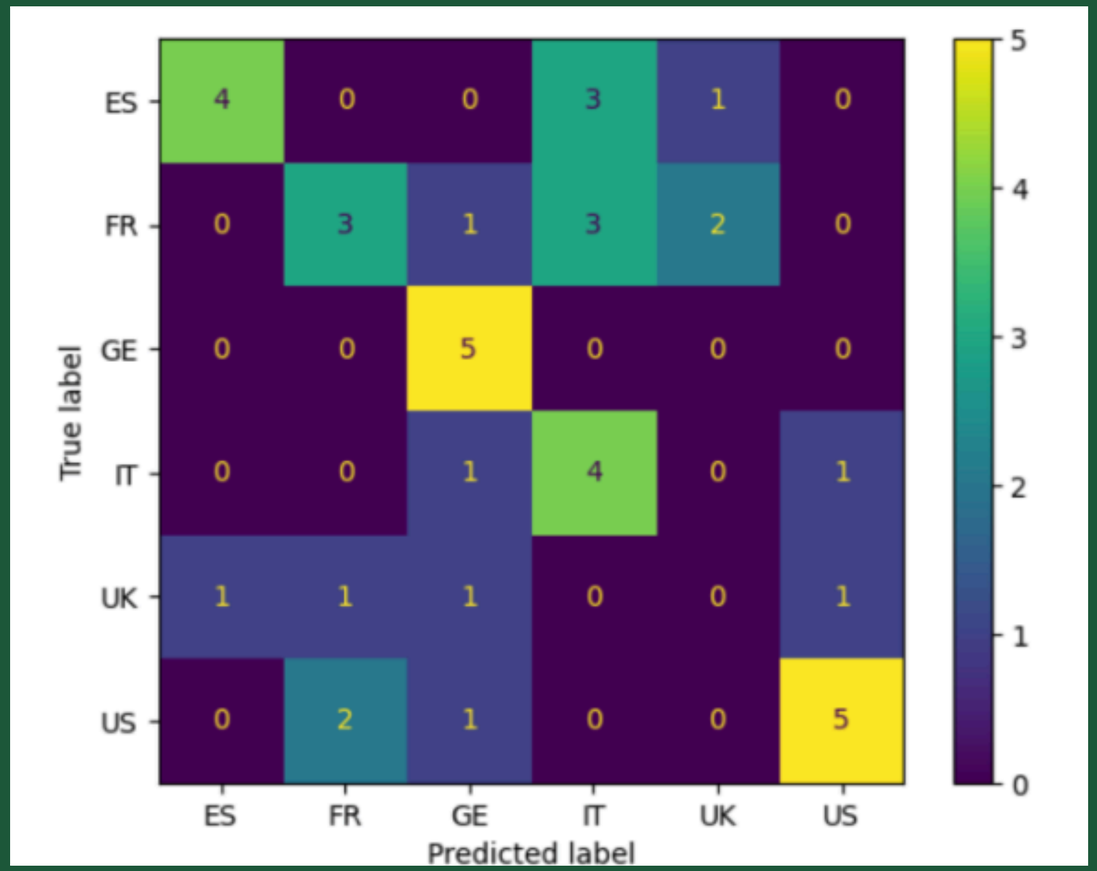
## UCI Dataset




C = 5, gamma = 0.1, kernel = rbf

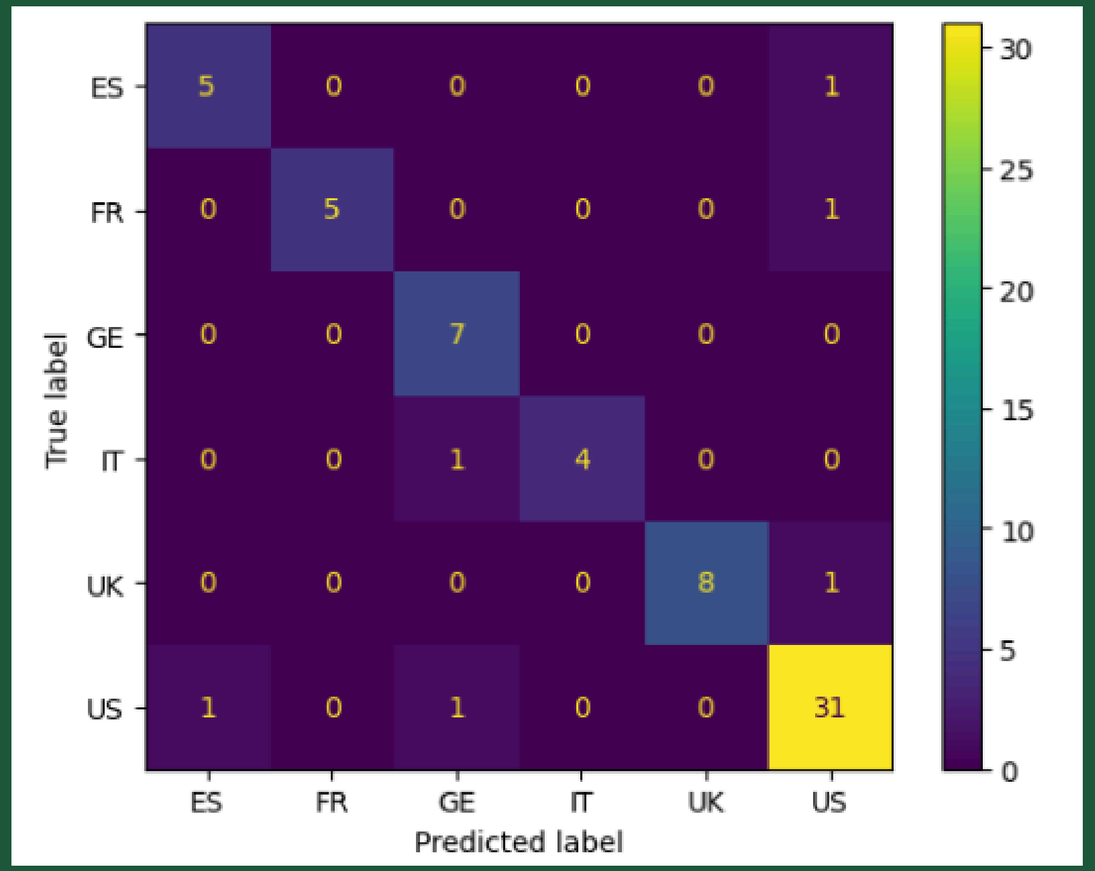
# MLP

## GMU Dataset




hidden layer size = (20,20), learning rate = 0.01

## UCI Dataset




hidden layer size = (300,), learning rate = 0.001

# Conclusions

- Importance of feature extraction
- Impact of class imbalance on performance
- Quality of dataset impacts performance
- **Notable results with french and german samples**

# THANK YOU!

## QUESTIONS?

# REFERENCES


[1] "Mel filter bank generation," pyfilterbank Documentation. [Online]. Available: https://siggigue.github.io/pyfilterbank/melbank.html. [Accessed: 06-May-2024].

[2]P. Mahesha and D. S. Vinod, "LP-Hillbert transform based MFCC for effective discrimination of stuttering dysfluencies," 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 2017, pp. 2561-2565, doi: 10.1109/WiSPNET.2017.8300225.