Speaker Accent Recognition

Prepared by

1. Mohamed Ehab Ali
2. Ameen Reda Ameen
3. Mahmoud Tarek Mohamed
4. John Adel Khaleel
5. John George Hana
6. Moustafa Kafafy Karam

PROJECT CASE

DATASET OVERVIEW

DATASET CONCLUSION

MODEL

CONCLUSIONS

Agenda

Having many **Audio Tracks** and you want to know the accent of each track, How to do that??

This project aims to classify an **Audio Track** to a specific accent by using Mel-frequency cepstral coefficients (MFCCS) of Audio Track.

Project Case

Dataset Overview

- The dataset is provided by : UCI Machine Learning Repository
- The dataset contains 329 rows and 13 columns.
- The dataset has 12 attribute columns and 1 label column.
- The 12 attribute columns obtained using MFCC on the original time domain soundtrack of the maximum 1s of reading of a word.
- The label column contains the six possible accents considered {ES, FR, GE, IT, UK, US}

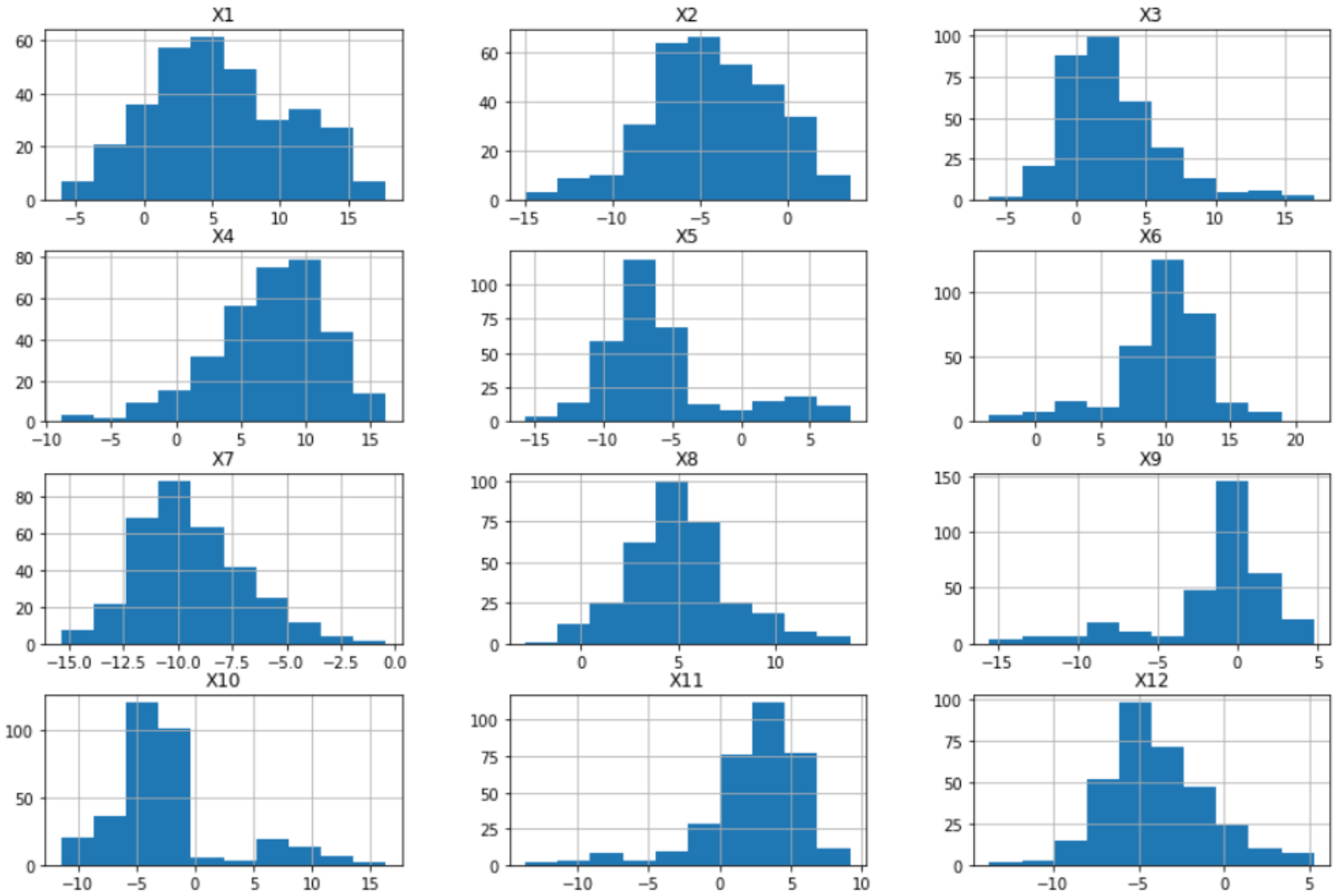
Dataset Conclusion

- The dataset is clean :
  - No Outliers
  - No dependencies
- But !!! :
  - The data is biased -> 165 US – 164 OTHERS
- So we decided to build two models based on two forms of data

Dataset Conclusion

Dataset Histogram :

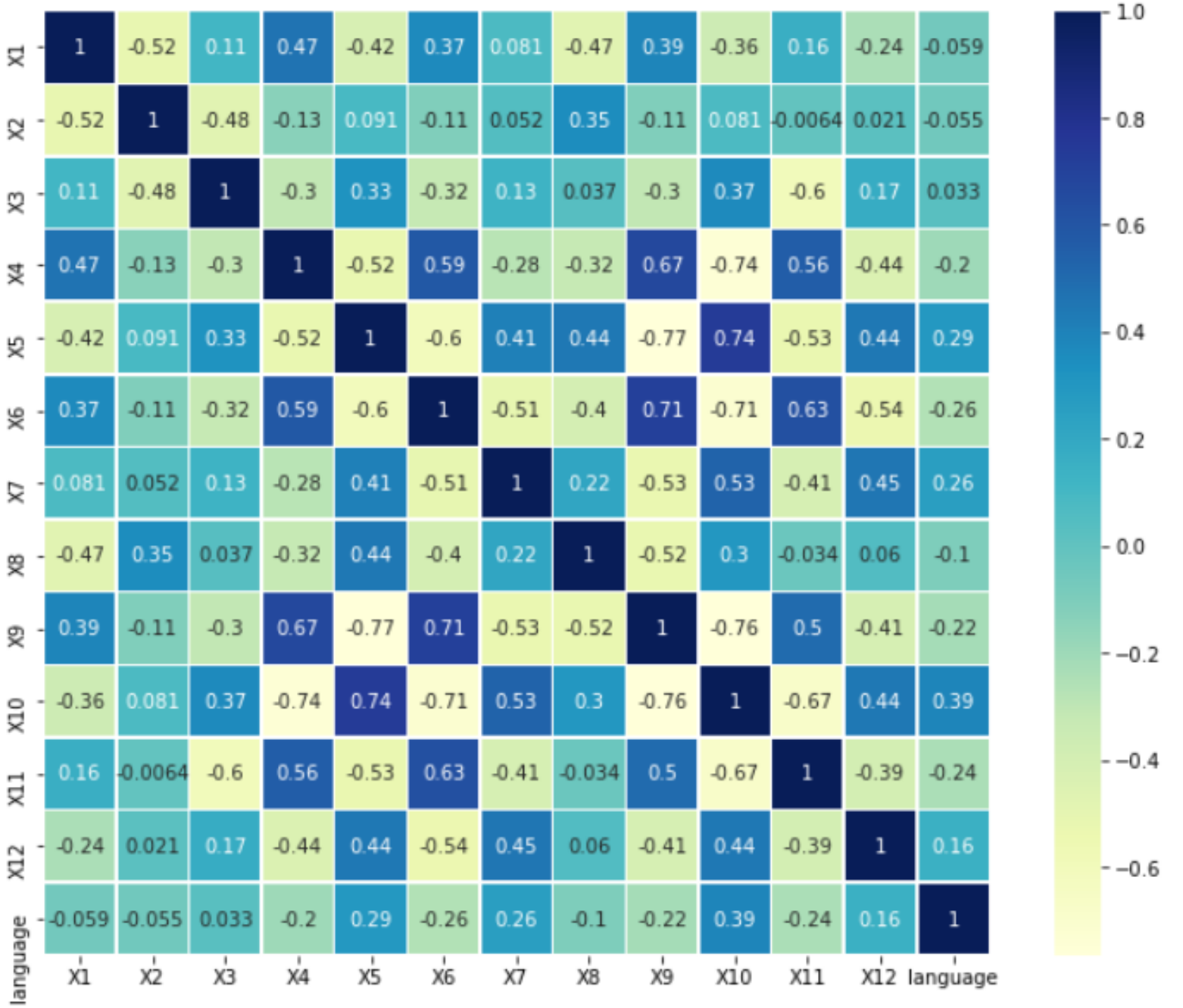
- No Outliers



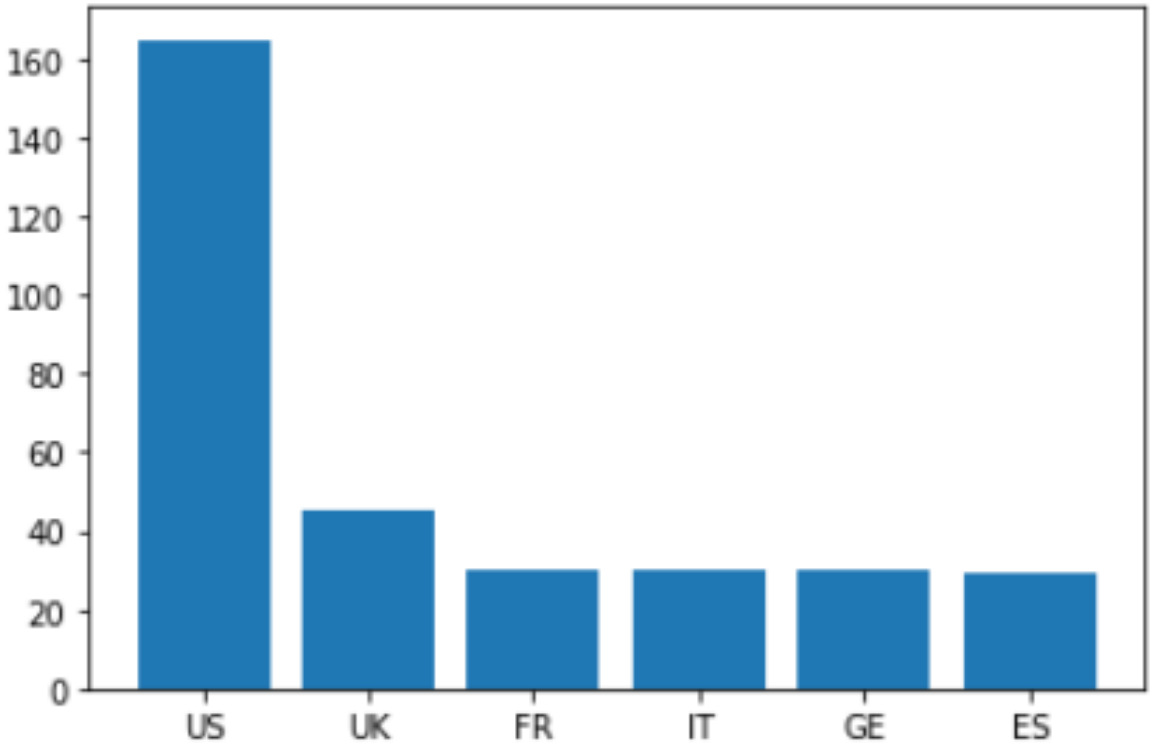
Dataset Conclusion

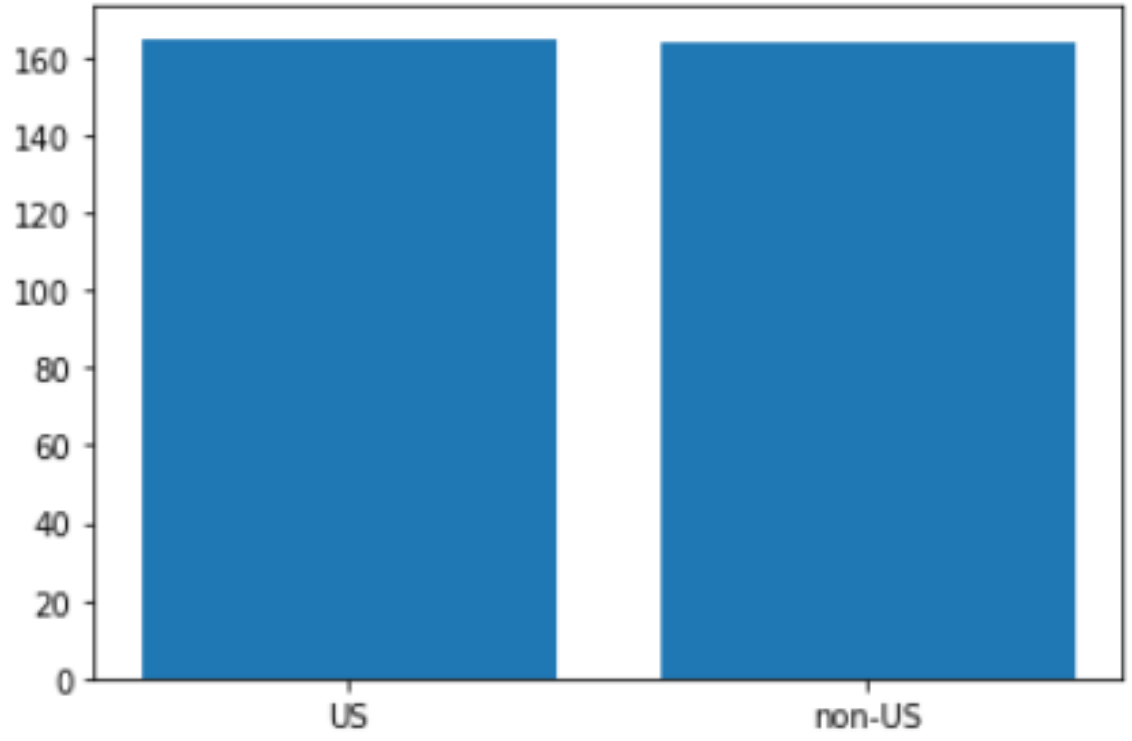
Dataset Histogram :

- No Dependencies



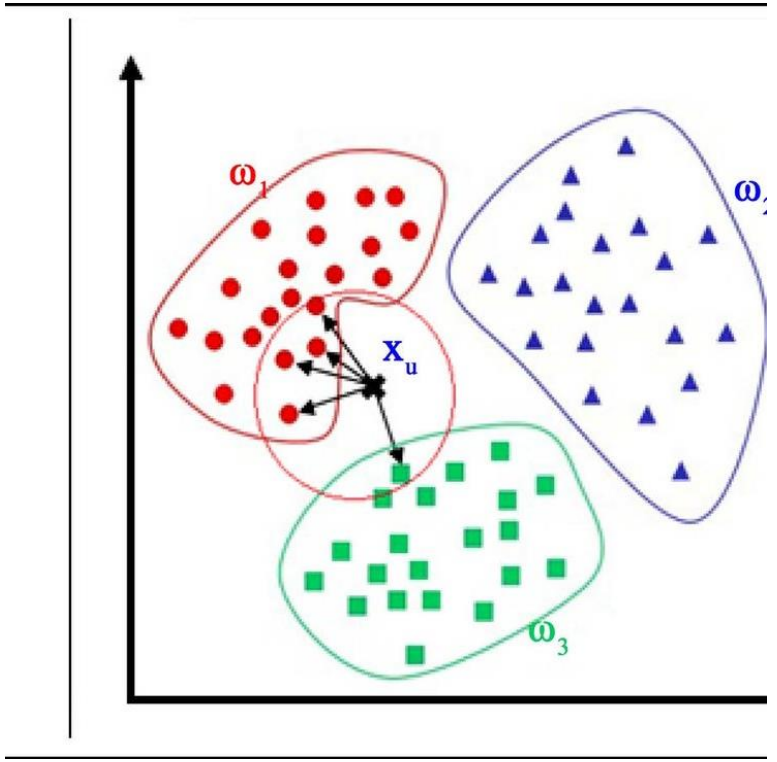
Dataset Conclusion





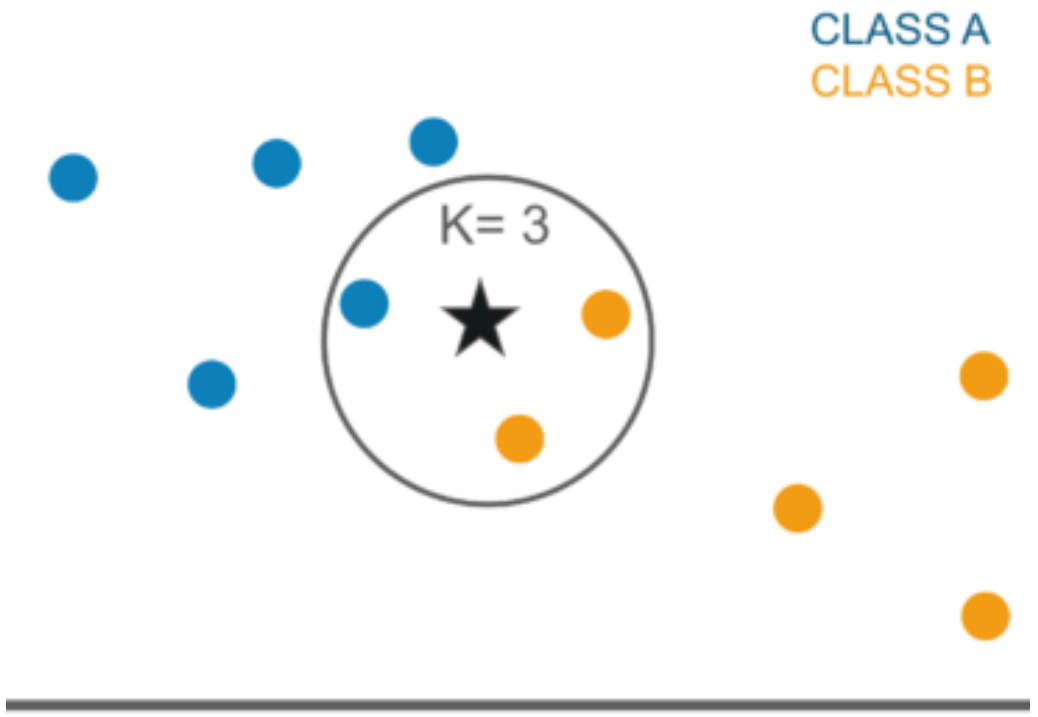
Model (First Model)

- We Trained a K-NN model on the audio track with six possible accents (ES, FR, GE, IT, UK, US).
- We obtained 86% accuracy rate and 14% error rate.



Model (Second Model)

- We Trained a K-NN model on the audio track with two categories (US / non-US).
- We obtained 97% accuracy rate and 3% error rate.



Conclusions

By comparison of ML algoritms,K-nearest neighboors yields the highest average test-accuracy.

We used two models (K-NN) to solve the biased data problem one on the biased data and the other on the data with no bias.

Thank you