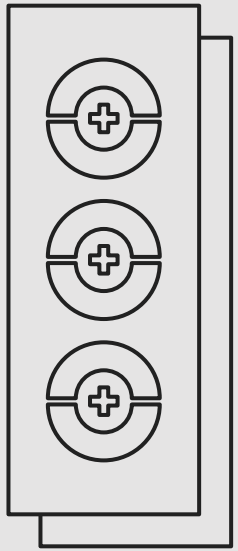
Now Featuring

Music genre classifier using Spotify audio features



Introduction

Motive:

- Genres today are getting more and more ambiguous.
- Want to create a model that helps classify genres based on spotify audio features
- Hope to help new artists get a grasp on how Spotify may categorize the genre of their music

Objective/Goals

- Create a classification model that will take in audio feature metrics and predict the genre.
- Objective is to be able to accurately predict songs genres

Workflow

Spotify

matplotlib

EDA → **BASELINE** → **REFINE** → **FINALIZE**

Dataset
90k → 50k

Used KNN, Logistic Regression, and Random Forest

Adjusting hyper-parameters

Chose to go with Random Forest

pandas

scikit learn

Results

KNN

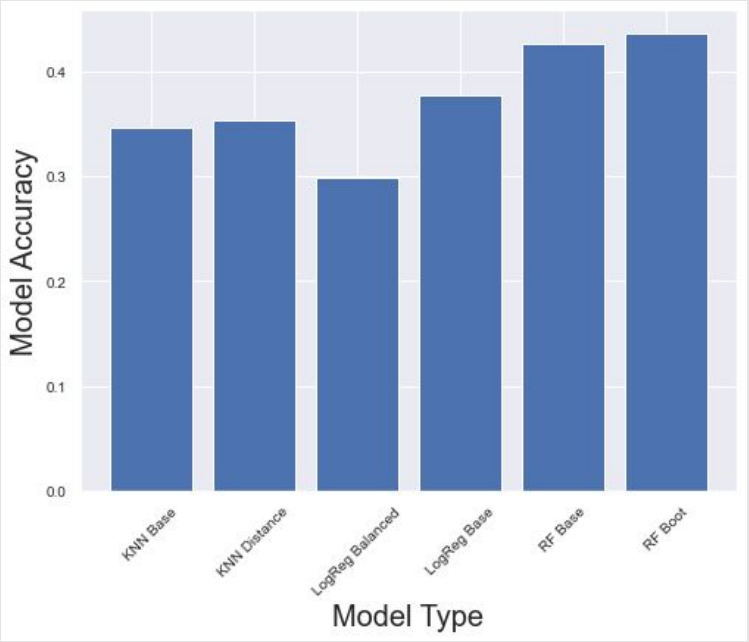
- Baseline accuracy = .347
- Distance Weighted = .354

Logistic Regression

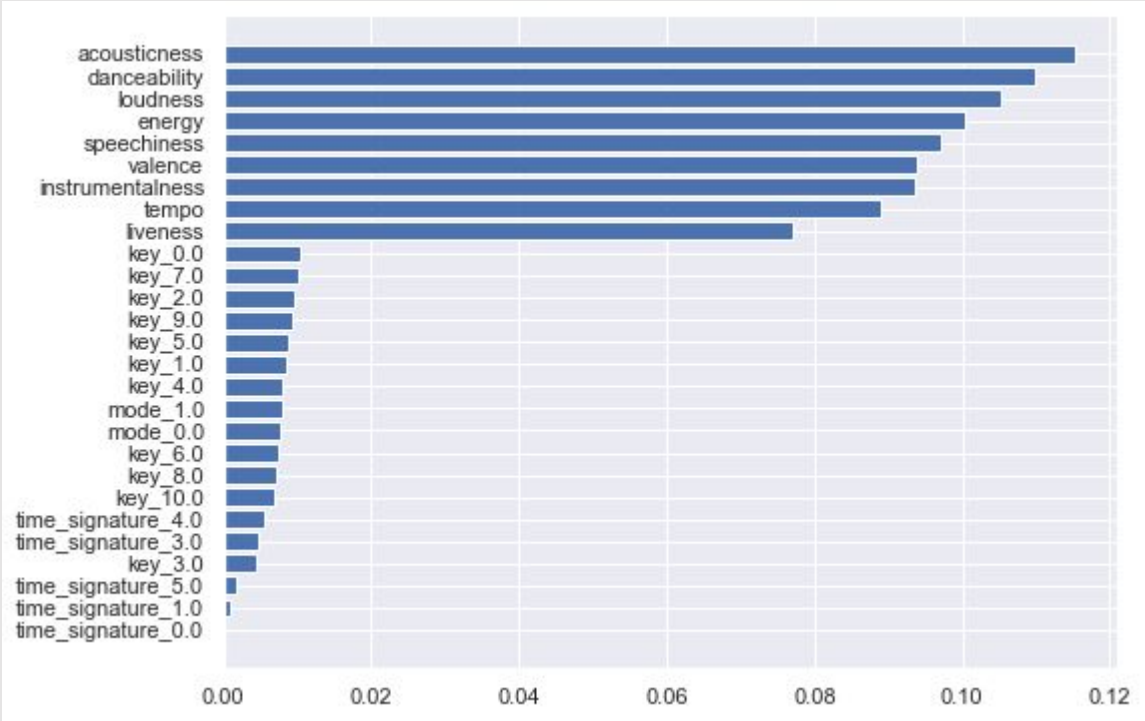
- Baseline accuracy = .377
- Balanced class weight = .299

Random Forest

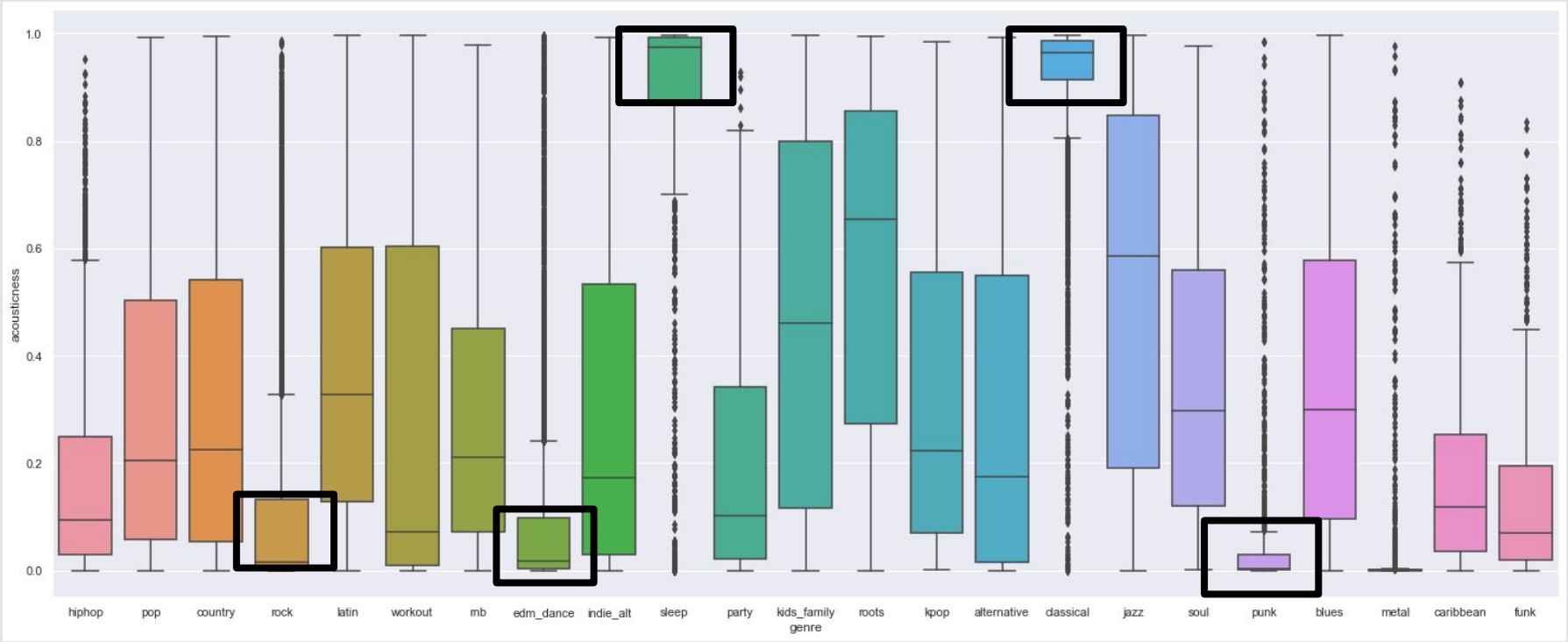
- Baseline accuracy = .427
- With Bootstrapping =.437



Feature Importance



Feature Importance(Acousticness)



Performance

Where is it good?

Hip-hop:

- DNA - Kendrick Lamar correctly classified
- Acousticness =.023
- Danceability - .638
- Energy = .523

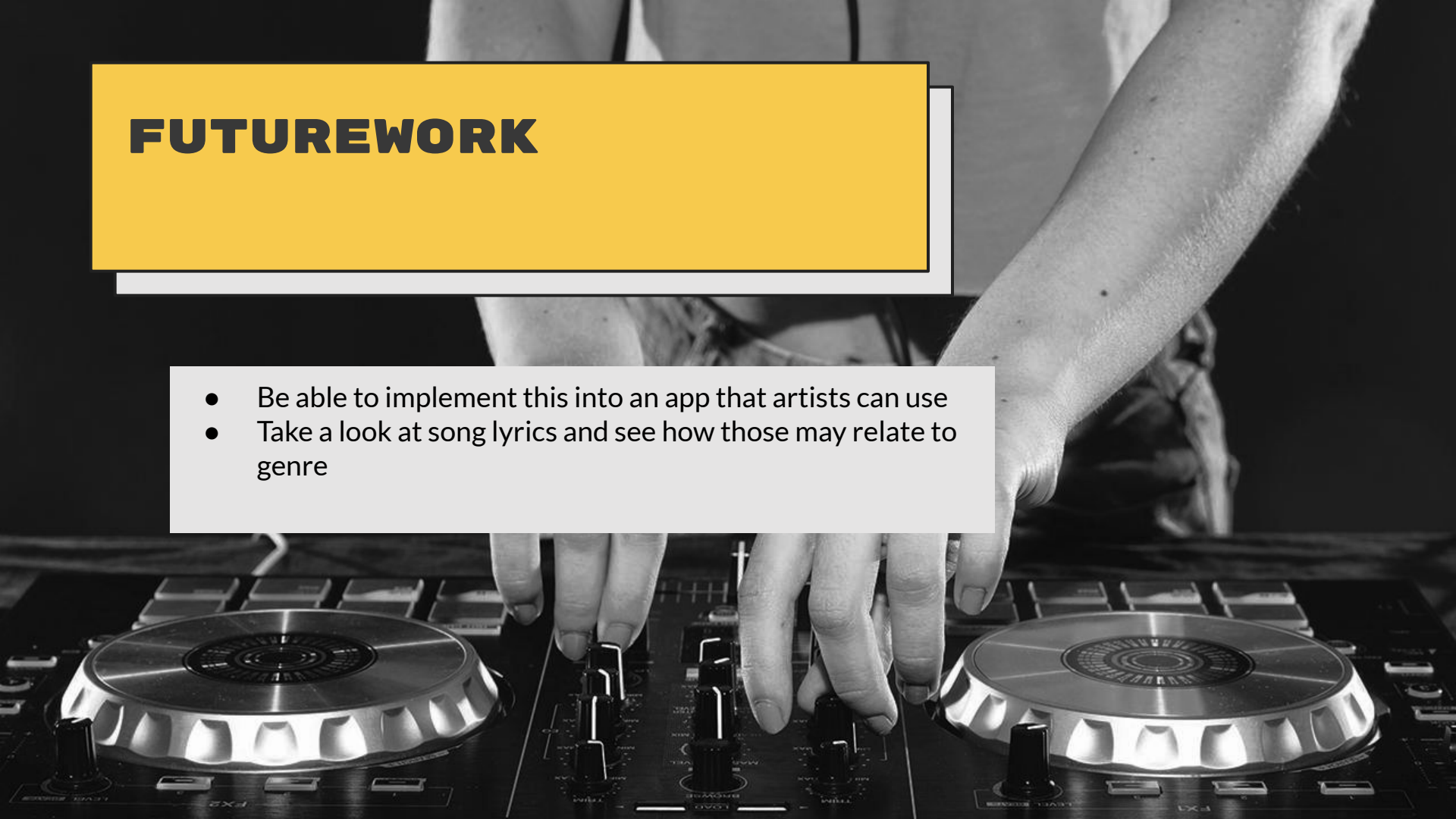
Where is it bad?

Country:

- Had over 1,500 False positives

Conclusion

- This turned out to be a difficult problem to solve with a large number of classes
- Model faced the most confusion with with Country and Classical (False Positives)
- Did better with Hip-hop genre



FUTUREWORK

- Be able to implement this into an app that artists can use
- Take a look at song lyrics and see how those may relate to genre