Predicting Genres from Acoustic Features

Andric Brena
Maha Attique

Why genre recognition?

Scientific Question: How well can a model using Gaussian Naive Bayes be at multi-class classification? Does incorporating information gain improve the accuracy rate?

Motivation:

- Explore the complex nature of genres
- Analyze the ways in which large collections of music can be organized

What does the data look like?

- 3 separate csv files:
  - tracks.csv
    - Info on a track's id, title, language, genre(s), etc.
  - genres.csv
    - Info on the list of genres/subgenres
  - echonest.csv
    - Info on a track's auditory features, artist location, temporal features
- Focused on tracks with a single genre at the top level and their auditory features
  - Classes = 12 different genres
  - Features = acousticness, danceability, tempo, liveness, instrumentalness, energy, speechiness, valence

Preprocessing the Data

- Stored the contents of each file as a dataframe to filter and merge them based on:
  - Tracks with a single genre
  - The single genre is at the top level
  - The tracks contain auditory features
- Normalized the feature columns
- Split the resulting data into a train and test set

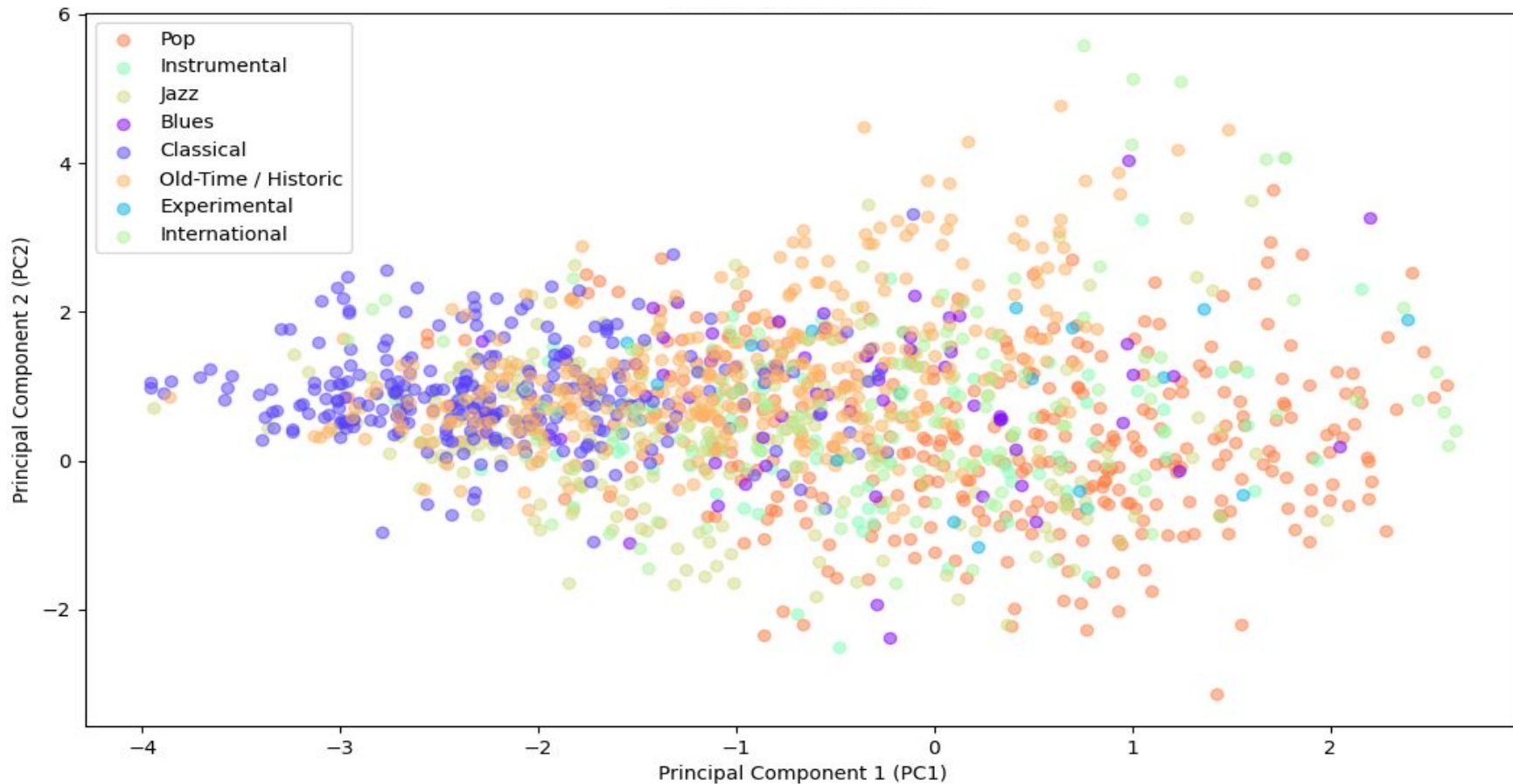|	track_id	genre_top	title	acousticness	...	liveness	speechiness	tempo	valence
0	2	Hip-Hop	Food	0.416675	...	0.177647	0.159310	165.922	0.576661
1	3	Hip-Hop	Electric Ave	0.374408	...	0.105880	0.461818	126.957	0.269240
2	5	Hip-Hop	This World	0.043567	...	0.373143	0.124595	100.260	0.621661
3	10	Pop	Freeway	0.951670	...	0.115474	0.032985	111.562	0.963590
4	134	Hip-Hop	Street Music	0.452217	...	0.096567	0.525519	114.290	0.894072
5264									
9413	42141	Experimental	Smoke In The Blues Bar	0.994404	...	0.316379	0.046528	80.870	0.089221
9414	42146	Experimental	Blues For Brownie	0.974165	...	0.120454	0.043112	106.959	0.851080

Methodologies

- PCA

- Gaussian Naive Bayes

- Feature selection through maximum importance (weight) in Gaussian Naive Bayes

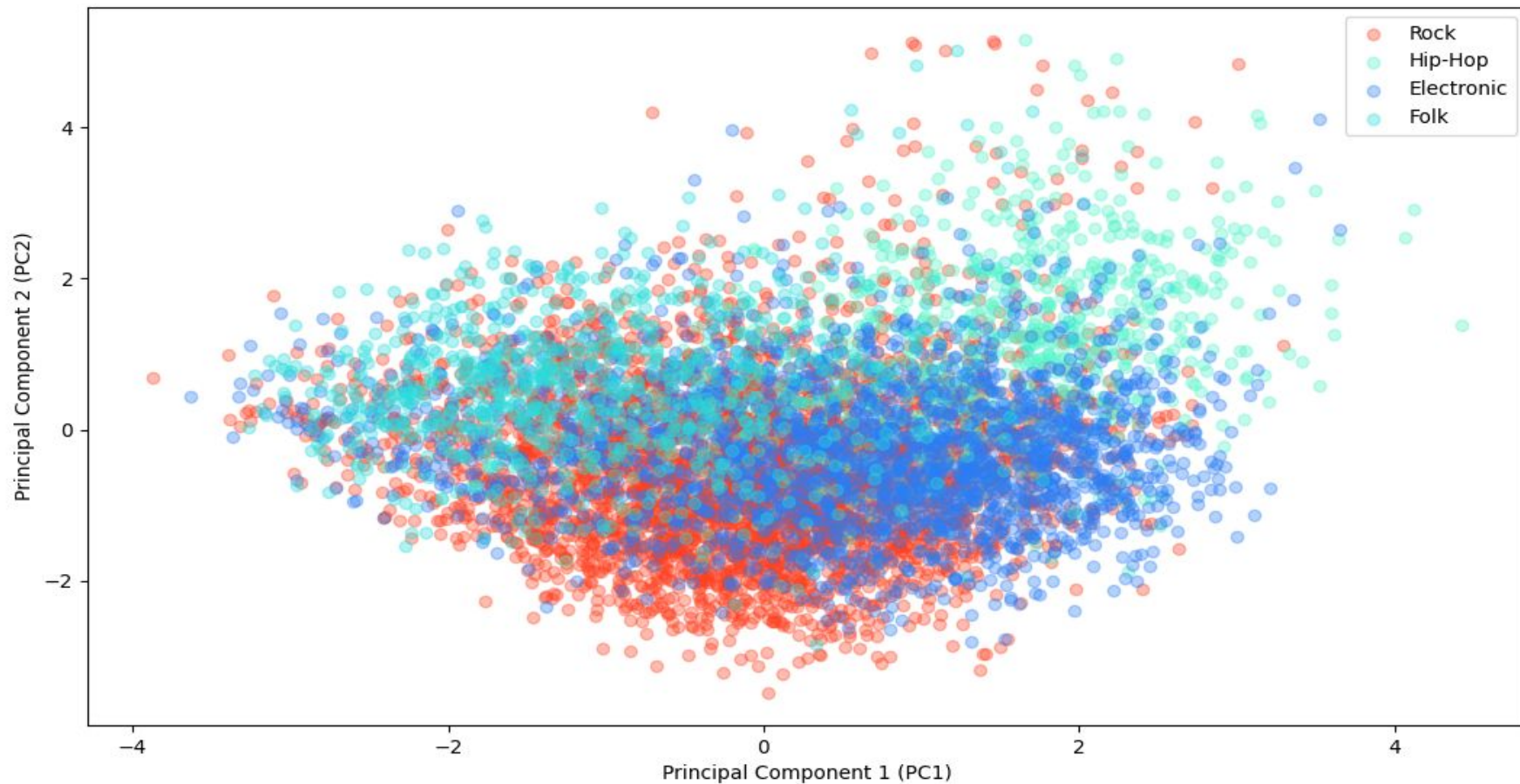- Feature selection through Information gain
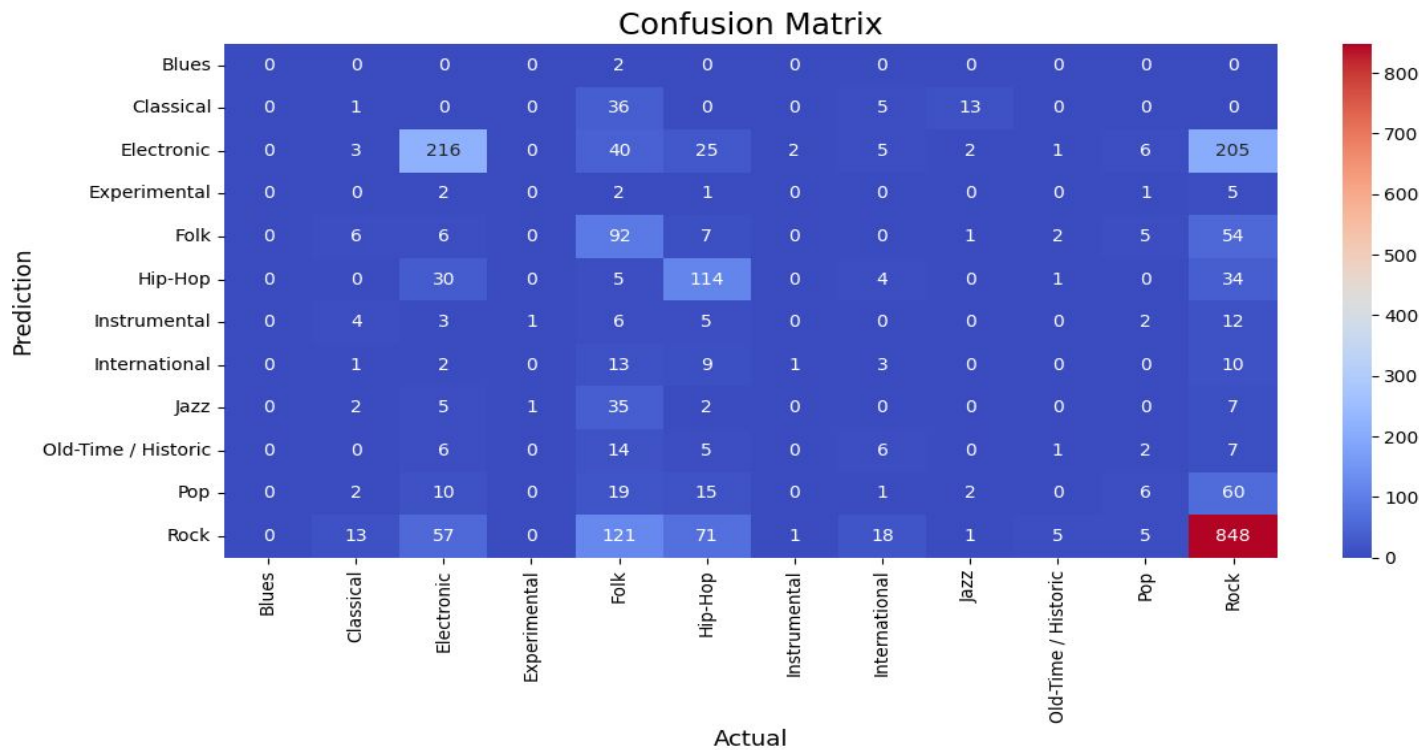
PCA Plot containing all 12 genres
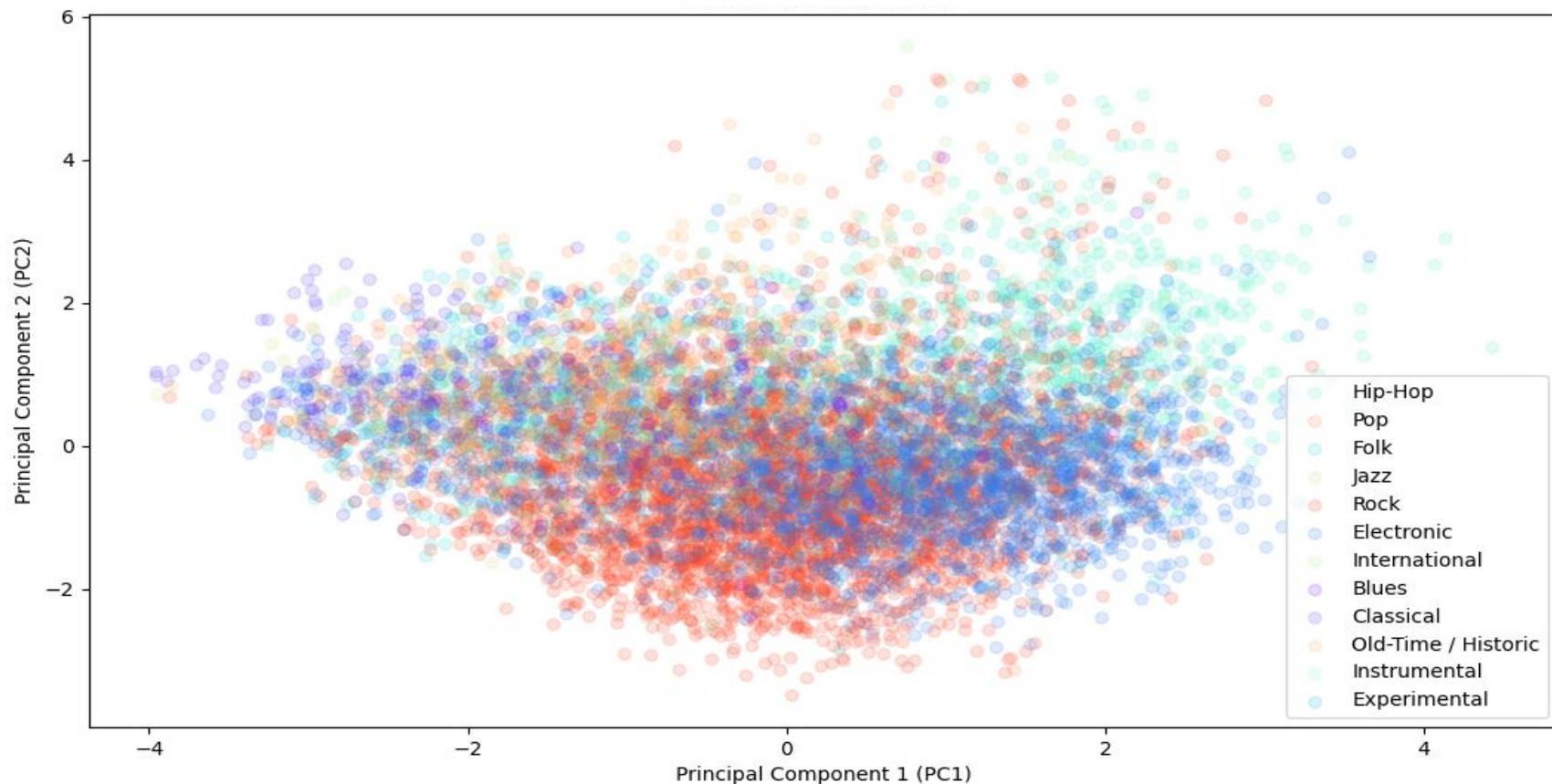
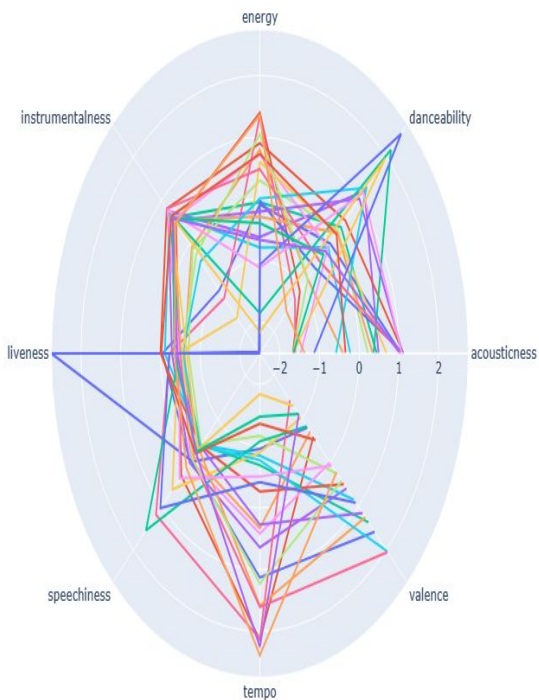# Breaking the PCA plot down

Breaking the PCA plot down

Gaussian Naive Bayes (across all acoustic features)

Confusion Matrix

Accuracy: **54%**

Highlights the imbalance in our dataset.

Radar Plots

Rock

Legend:
- Dance With Local Stars
- If I Can't Smoke & Swear I'm Fucked
- Obsessive
- Horse
- No Metal
- Why
- Track 02
- Take It With You
- Street Fighter 5
- What Makes a Plaza?
- Wizard's Castle
- Many More Melting- The When McGruff Was King
- Whistle while you work
- No Big Deal
- DeadTown
- Shadow of the Past
- Prance of the Ravening Eagle
- Hardcore Is Fuckin' Dead
- I just don't know
- Clapmosphere
- Positive
- This Is The Best
- Despair Squid + Ex-M3
- Where Is Everything
- Diving

Hip-Hop
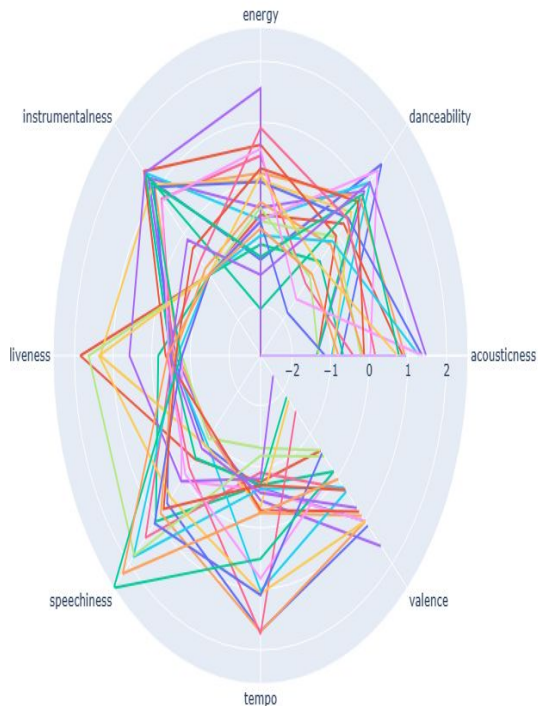
Legend:
- Go On, Take It Further
- I Won't Let You Fall
- Wild Man
- Two Grudges
- Out Of Here (Featuring Mic Jordan)
- Put it together
- Cold Nights (Get Your Money Right) (Featuring O.B. of M.O.B. Stars)
- True (TSP Remix)
- Happens to be
- Something To Do
- Tango de la muerte
- October 16th
- aint no thing
- In the Midst of Struggle
- Blockin ya on Blessings
- Inbetween
- 05 A-Trak & Lupe Fiasco feat. 6th Sense - Me & My Sneakers.mp3
- Just Got Chopped (Crack)
- Hood Hippie 2
- Rendezvouz feat. Mr. Nnaji - Produced by Comfort Fit
- Sweet Dreams
- Rapero mula
- Sloppy Joes
- Feel Me
- I Got You (Album Version)

Absolute Feature Importance based on Gaussian Naive Bayes
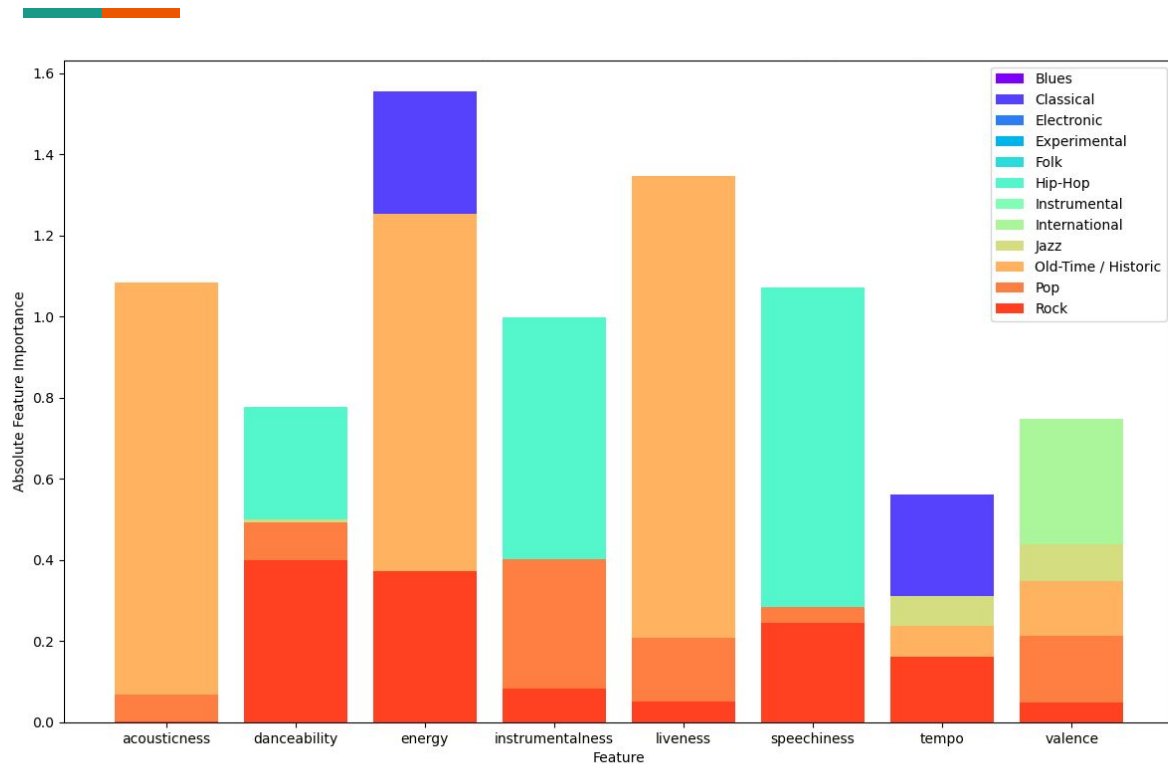
Each feature holds a different importance in predicting a specific genre!!!

Using a specific feature to predict a specific class

Genre: **Old-Time / Historic**

Most Important Feature = **liveness**

Accuracy achieved = 94.1%

Confusion Matrix

	Actual 0	Actual 1
Prediction 0	2208	105
Prediction 1	34	7
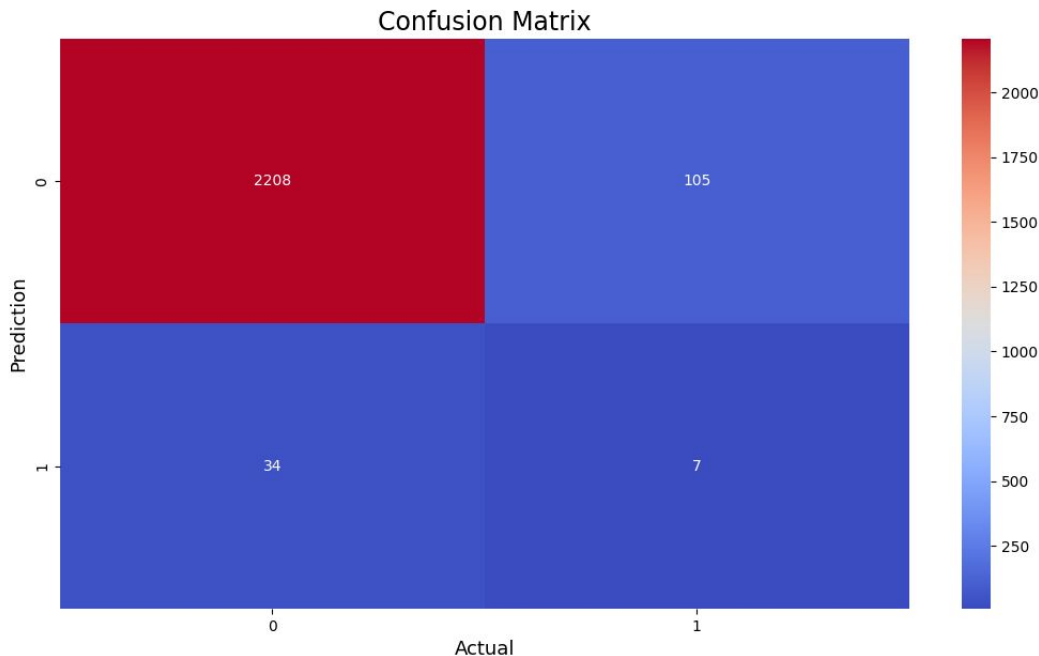
Genre predictability through most important feature

This process was heavily dependent on how **large the weight of the most important feature was for its class**.

Large weight for the most important acoustic feature:

Genre: **Hip-Hop**

Most Important Feature = **speechiness**

Accuracy achieved = 91.5%

Small weight for the most important acoustic feature:

Genre: **Rock**

Most Important Feature = **danceability**

Accuracy achieved = 65.6%

Information gain across all features

Confusion Matrix

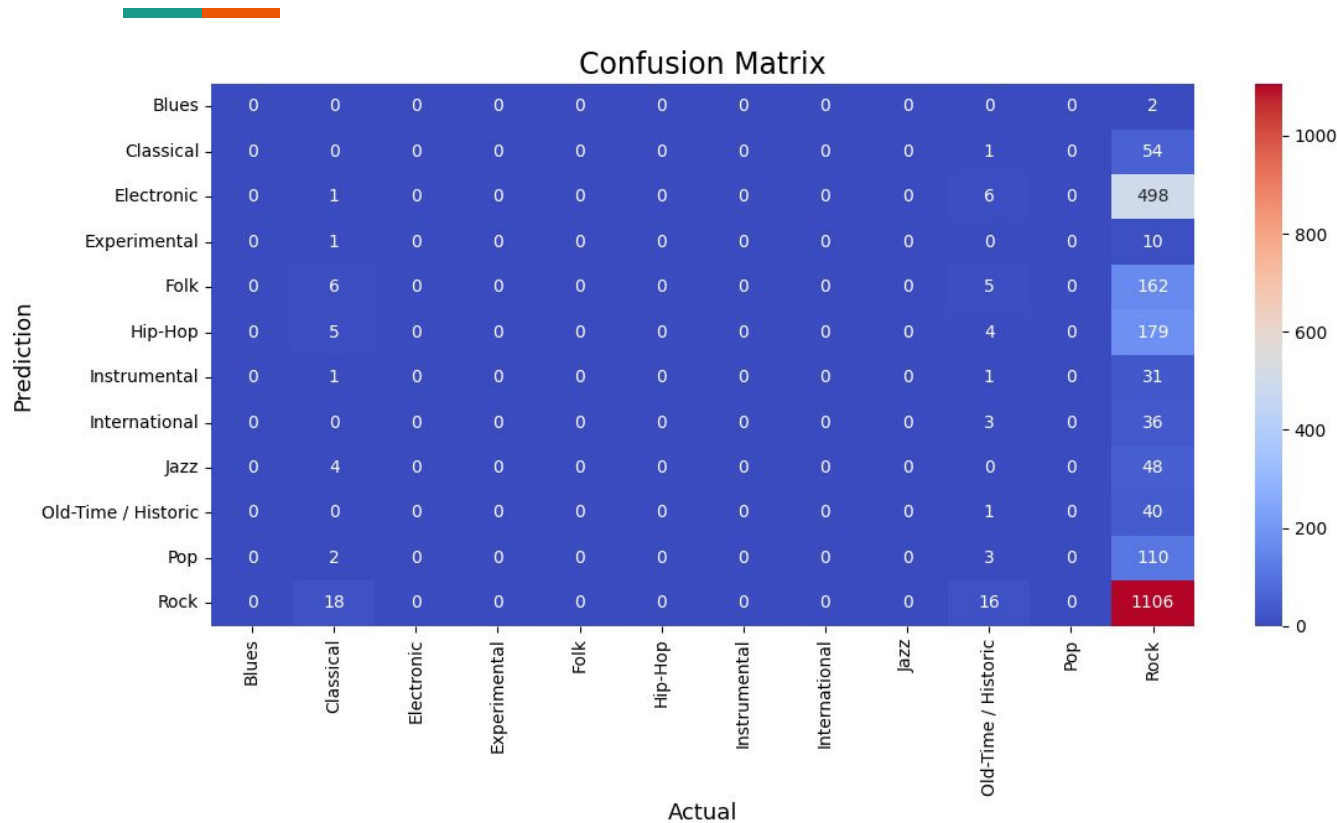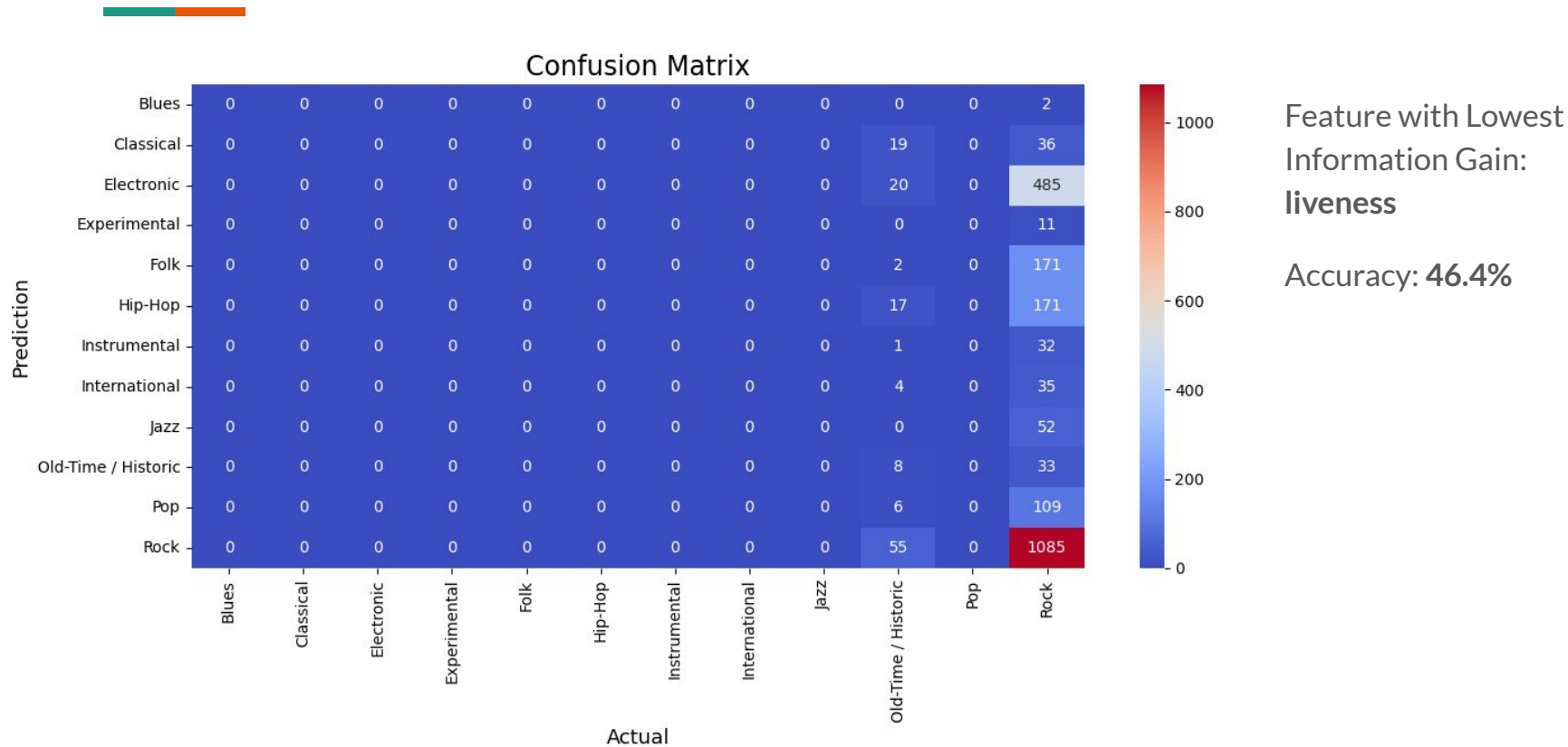|	Blues	Classical	Electronic	Experimental	Folk	Hip-Hop	Instrumental	International	Jazz	Old-Time / Historic	Pop	Rock
Blues	0	0	0	0	0	0	0	0	0	0	0	2
Classical	0	0	0	0	0	0	0	0	0	1	0	54
Electronic	0	1	0	0	0	0	0	0	0	6	0	498
Experimental	0	1	0	0	0	0	0	0	0	0	0	10
Folk	0	6	0	0	0	0	0	0	0	5	0	162
Hip-Hop	0	5	0	0	0	0	0	0	0	4	0	179
Instrumental	0	1	0	0	0	0	0	0	0	1	0	31
International	0	0	0	0	0	0	0	0	0	3	0	36
Jazz	0	4	0	0	0	0	0	0	0	0	0	48
Old-Time / Historic	0	0	0	0	0	0	0	0	0	1	0	40
Pop	0	2	0	0	0	0	0	0	0	3	0	110
Rock	0	18	0	0	0	0	0	0	0	16	0	1106

Prediction (y-axis) — Actual (x-axis)

Feature with Highest Information Gain: **acousticness**

Accuracy: **47%**

Information gain across all features

Confusion Matrix

Feature with Lowest Information Gain: **liveness**

Accuracy: **46.4%**

Conclusions

- Gaussian Naive Bayes is a viable method for music genre recognition.

- Feature selection highlights the ability for some features to better predict a genre than others.

- The acoustic features expressed a level of subjectivity.

Looking forward….

- Address bias towards rock, hip-hop, electronic, and folk tracks:

  - Using a more balanced data set that represents all genres equally

- Expanding/changing the set of auditory features:

  - Work with raw music files and use a different set of acoustic features to classify the genre (harmony, timbre etc). Maybe employ sentiment analysis for the lyrics to finetune the classification.

- Classify into sub-genres like the tracks in the original data set

- Conduct a user lab study, where we compare our models classifications to actual users who classify the music based on personal opinion.