

Predicting Music Genres with Audio Features

Team Members: Kimberly Simmonds, Rohan Giri,
Maxine Gardner, Kush Patel, Akash Barathan

Presentation Outline

01

Introduction

02

Data Overview

03

Model Evaluation

04

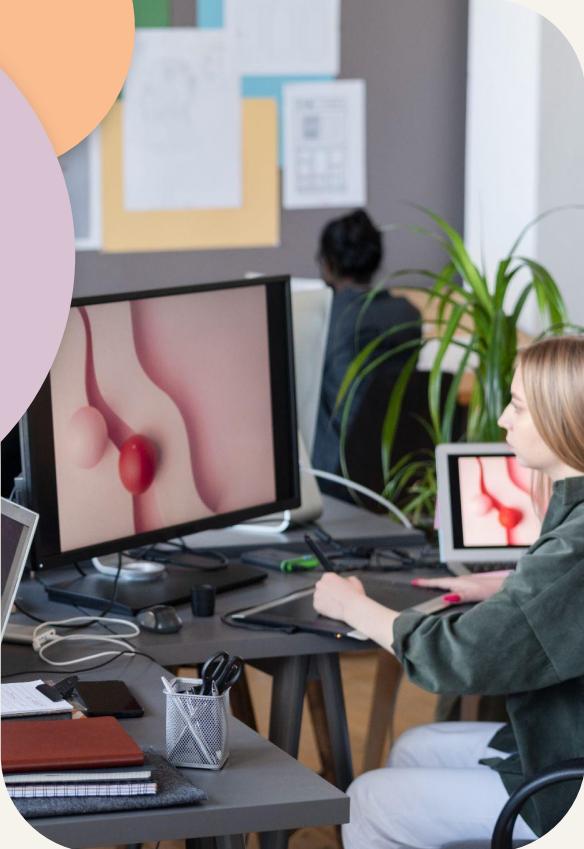
Application to
Modern Dataset

05

Results

06

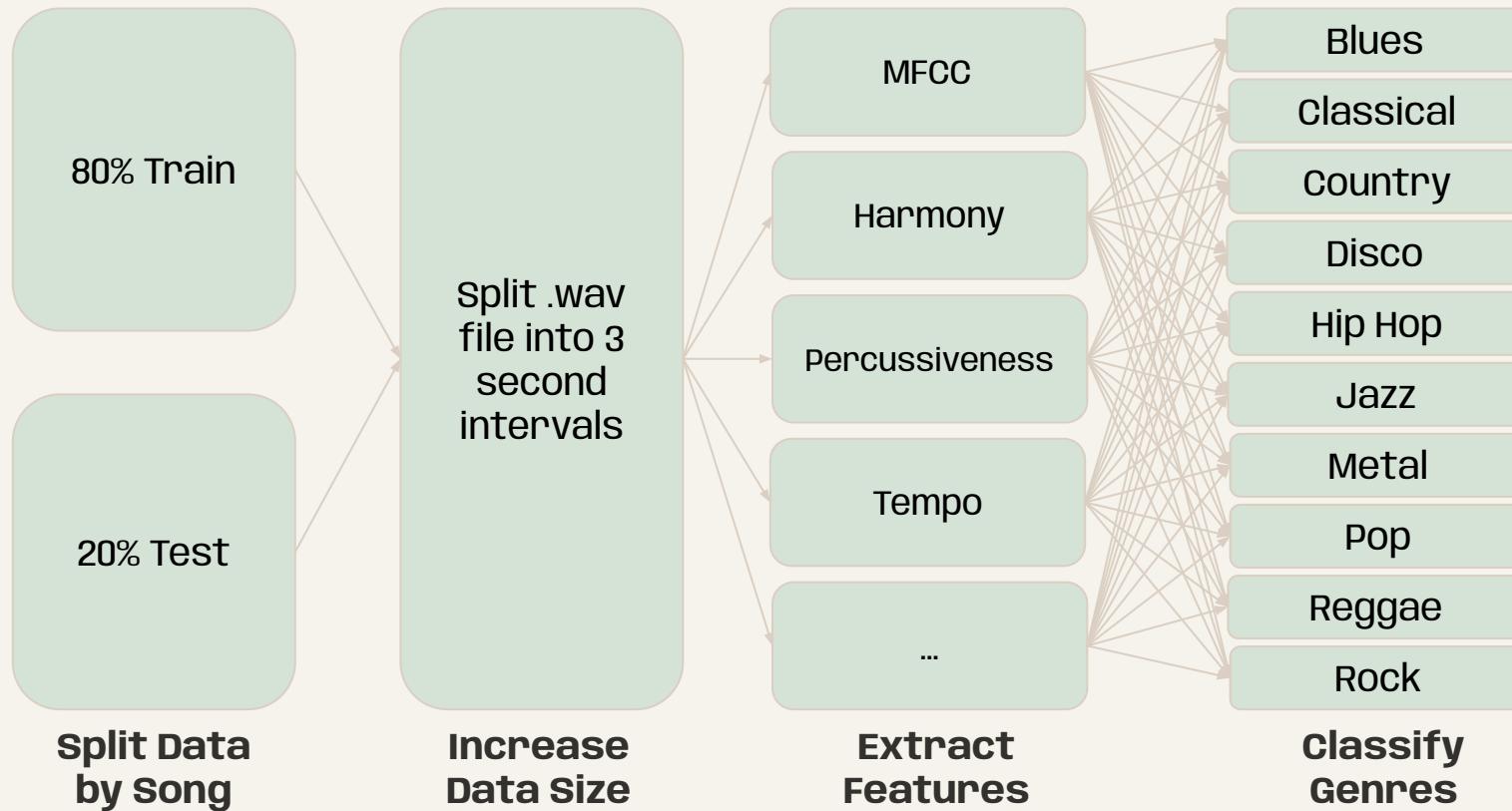
Future Scope &
Conclusion



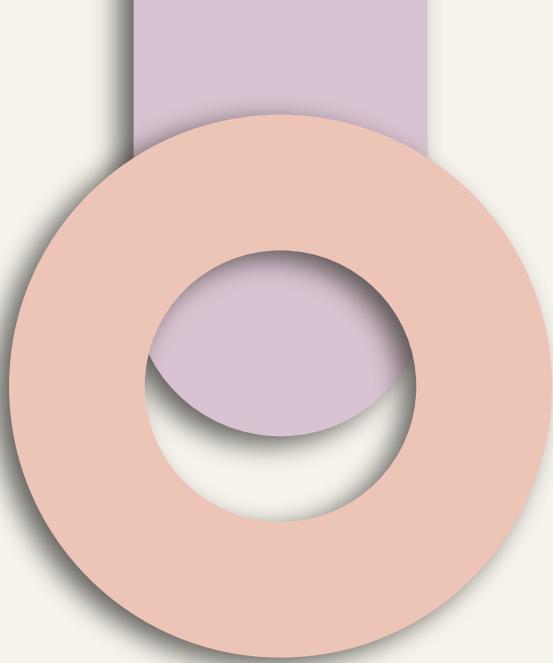
Introduction

Our project explores the intersection of music and artificial intelligence, classifying various music tracks into ten genres and using features such as tempo, harmony, and rhythm extracted using Librosa. Through combining machine learning algorithms with deep learning techniques, we aim to accurately categorize music tracks into their respective genres.

Workflow Diagram



Data Overview

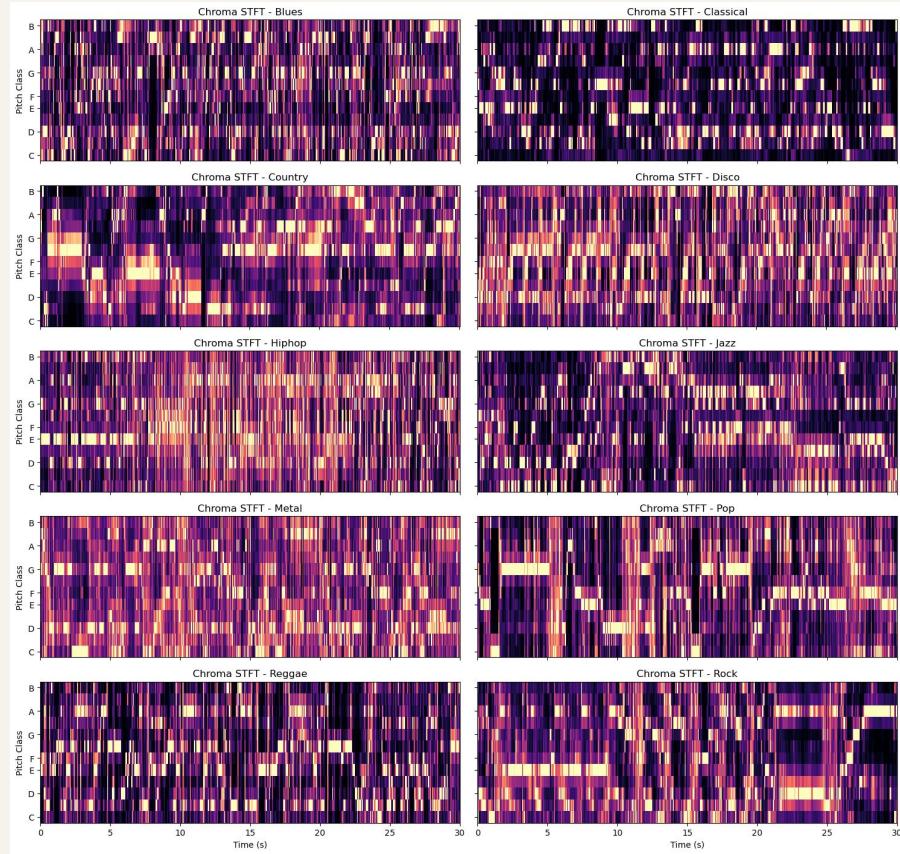


Data Pre-processing

<u>Dataset</u>	GTZAN Dataset - Kaggle 100 30-second audio clips from 10 different music genres up to 2001
<u>Feature Engineering</u>	187 musical features extracted from audio segments using Librosa
<u>Youtube Data</u>	Scraped 20 additional songs for each genre from Youtube starting 2002+

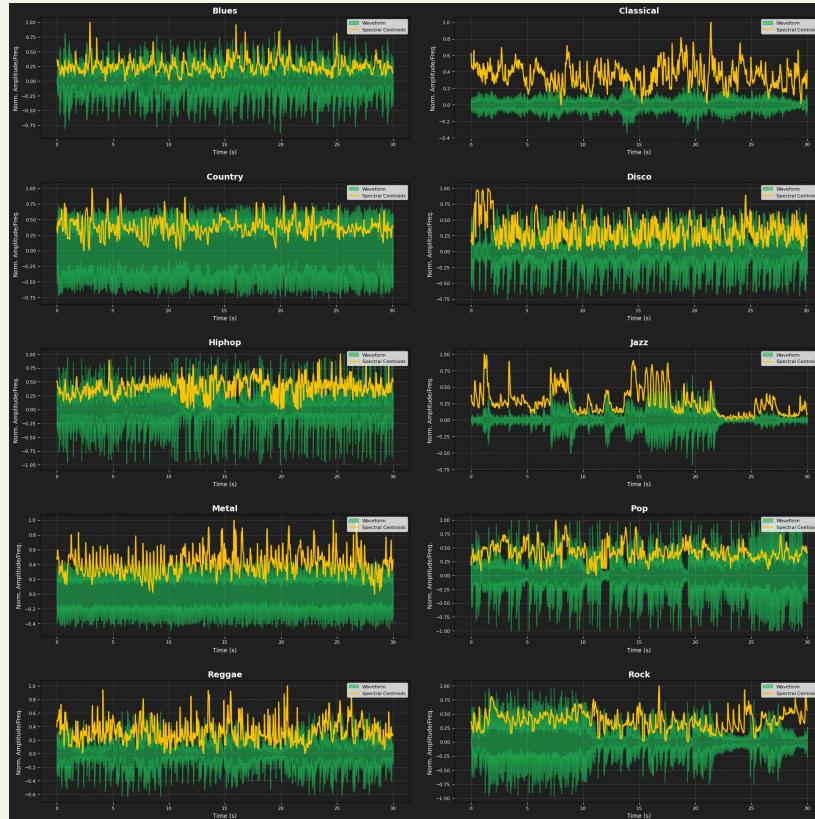
Feature Extraction: Chroma STFT

The Chroma STFT is a feature representation that uses a Short-Time Fourier Transform (STFT) to extract energy distribution over the 12 possible pitches.



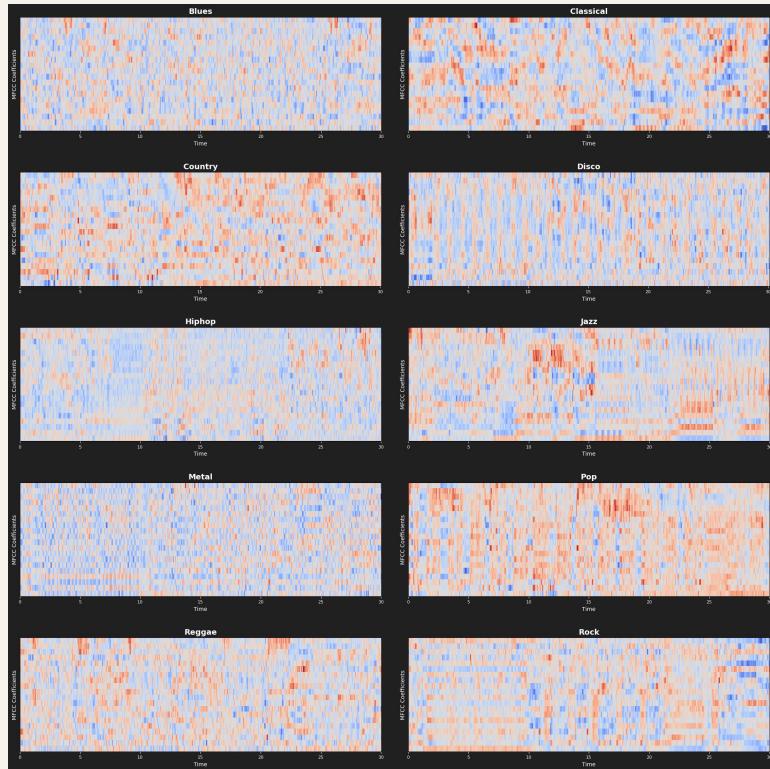
Feature Extraction: Spectral Centroid

The Spectral centroid measures the “tonal character” or brightness in the sound over time. A higher spectral frequency represents a brighter sound, while a lower spectral frequency represents a darker sound.



Feature Extraction: MFCC

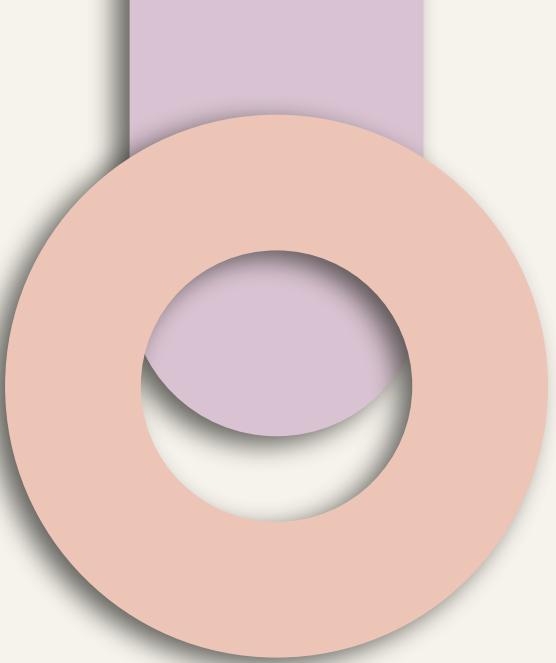
Mel Frequency Cepstral Coefficients (MFCCs) represents the short-term power spectrum of the sound which can be used to describe the timbre of the sound. The Mel scale aligns with how humans perceive pitch and frequency. This makes MFCCs excellent at capturing perceptual characteristics of audio signals, which plays a critical role in distinguishing musical genres.



Feature Extraction: More

- chroma_cqt
- chroma_cens
- Tonnetz
- Rms
- Spectral_bandwidth
- Spectral_rolloff
- Zero_crossing_rate
- Harmony
- Percussive
- Tempo

Model Evaluation



Baseline Models

Model	KNN (n=5)	Random Forest	Decision Tree
Accuracy	0.672	0.716	0.518
Precision	0.675	0.705	0.526
Recall	0.685	0.723	0.534
ROC-AUC	0.902	0.949	0.740

Convolutional Neural Network

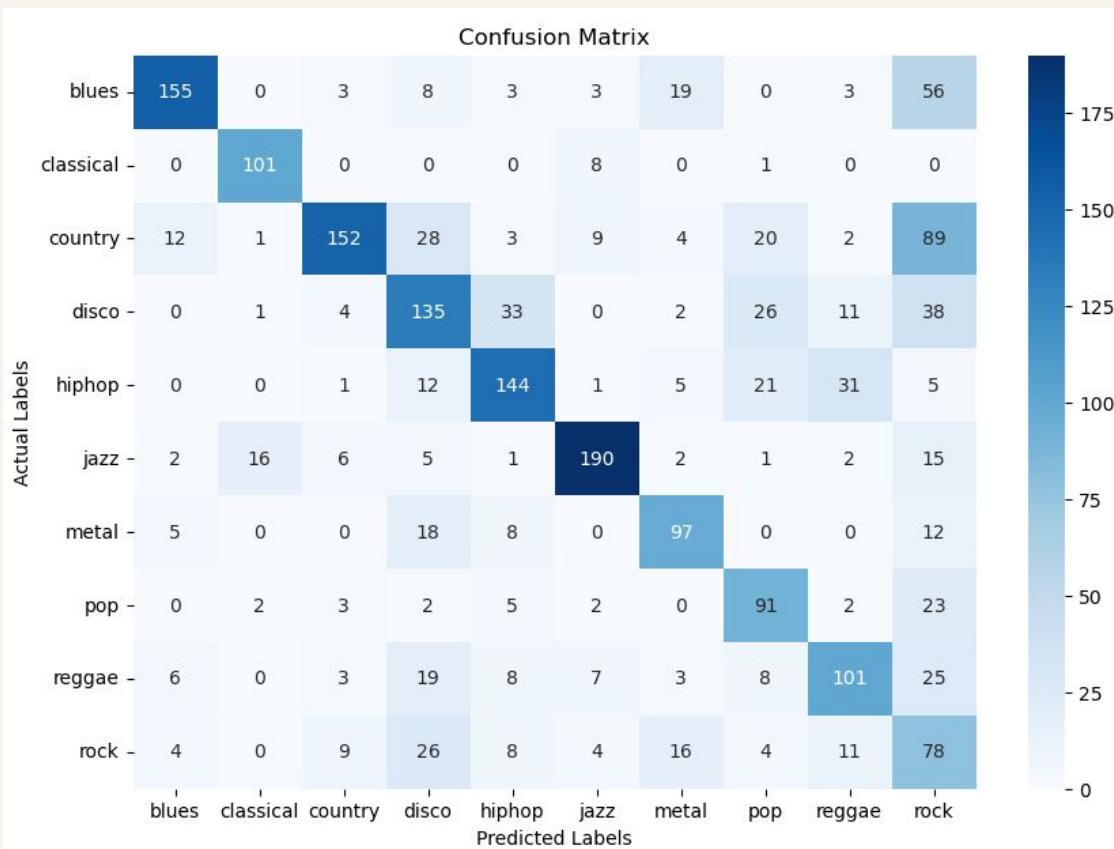
Overall

Accuracy: 0.622

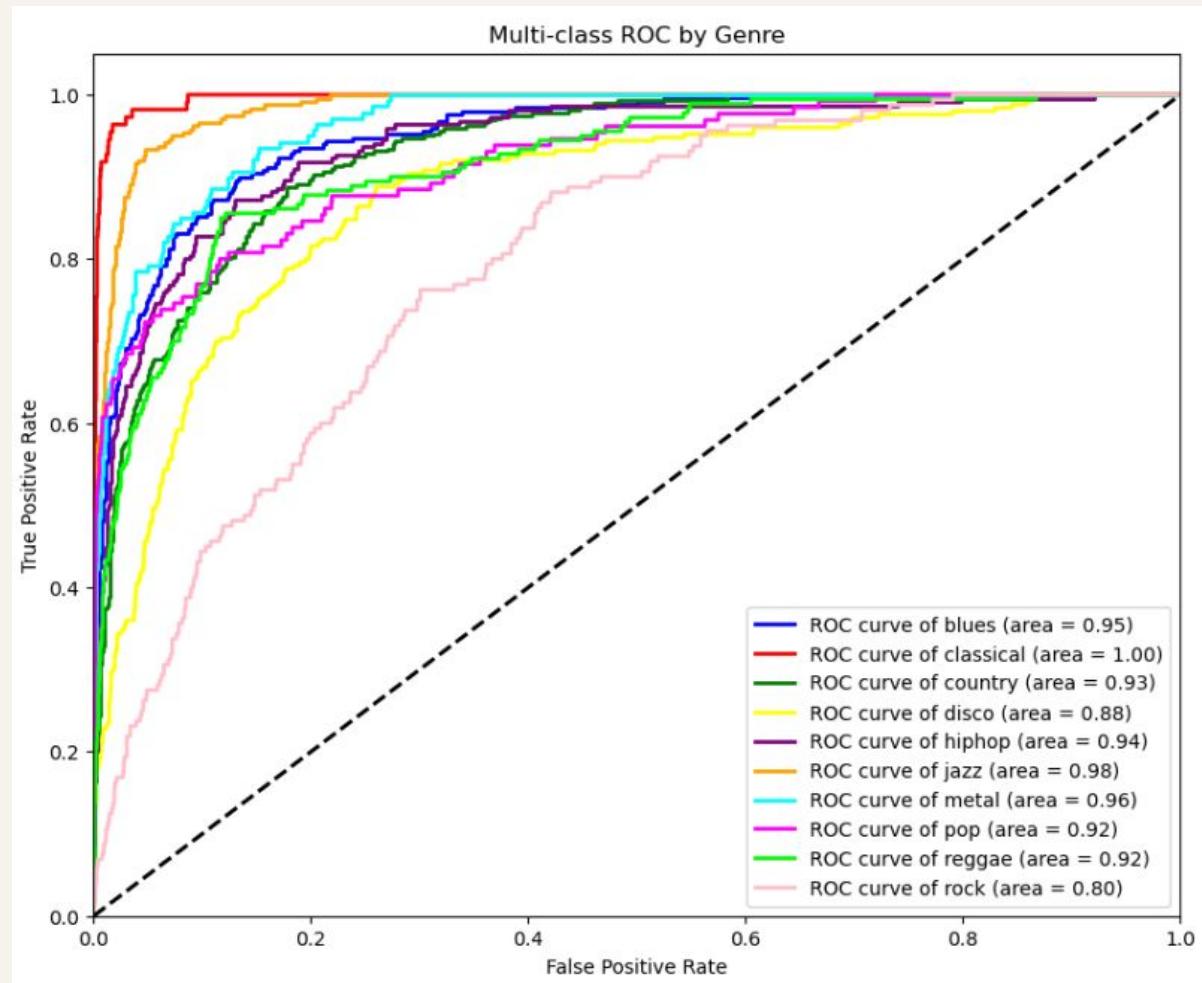
Precision: 0.660

Recall: 0.644

ROC-AUC: 0.93



<u>Genre</u>	<u>ROC AUC</u>
Blues	0.95
Classical	1.00
Country	0.93
Disco	0.88
Hiphop	0.94
Jazz	0.98
Metal	0.96
Pop	0.92
Reggae	0.92
Rock	0.80



Best Performing Model

XGBoost

HyperOpt Tuned Model Parameters:

n_estimators: 1187
min_child_weight: 5
max_depth: 15
learning_rate: 0.048

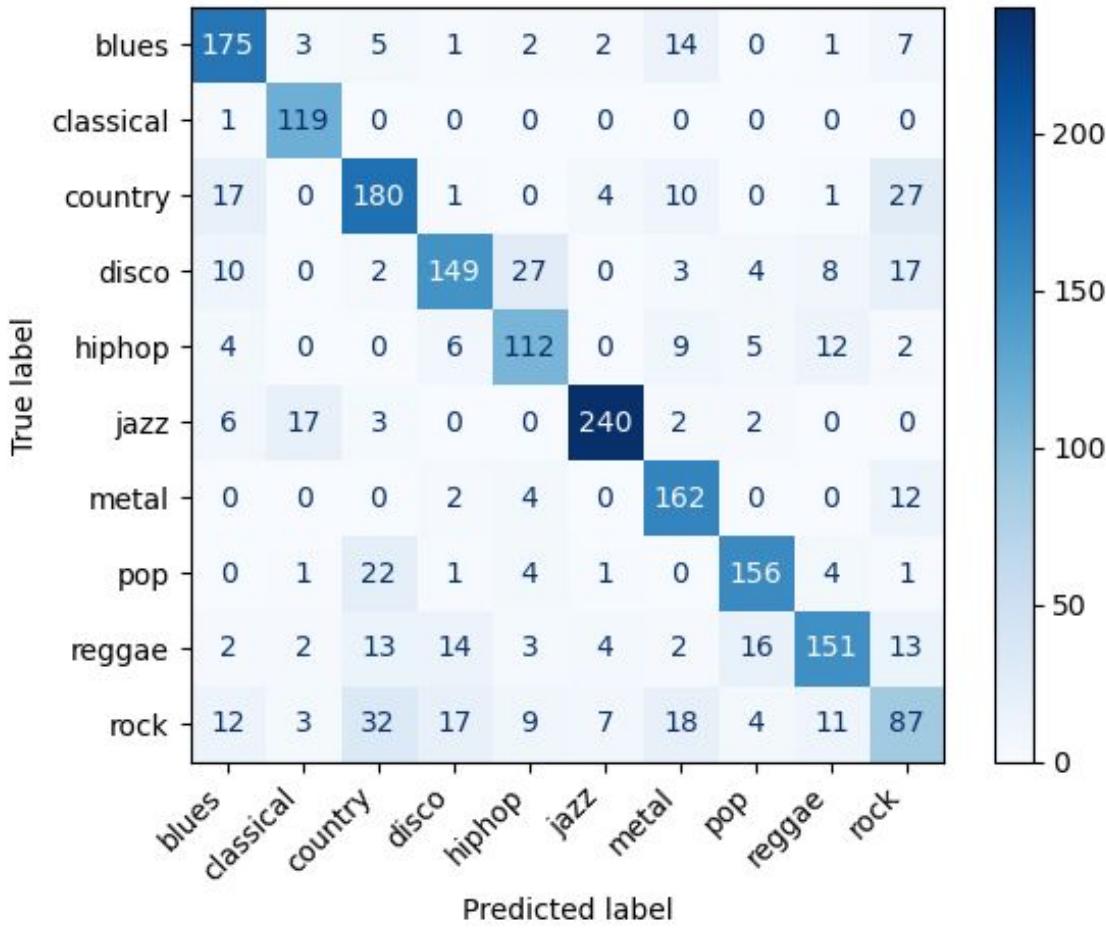
Overall

Accuracy: 0.765
Precision: 0.760
Recall: 0.773
ROC-AUC: 0.965

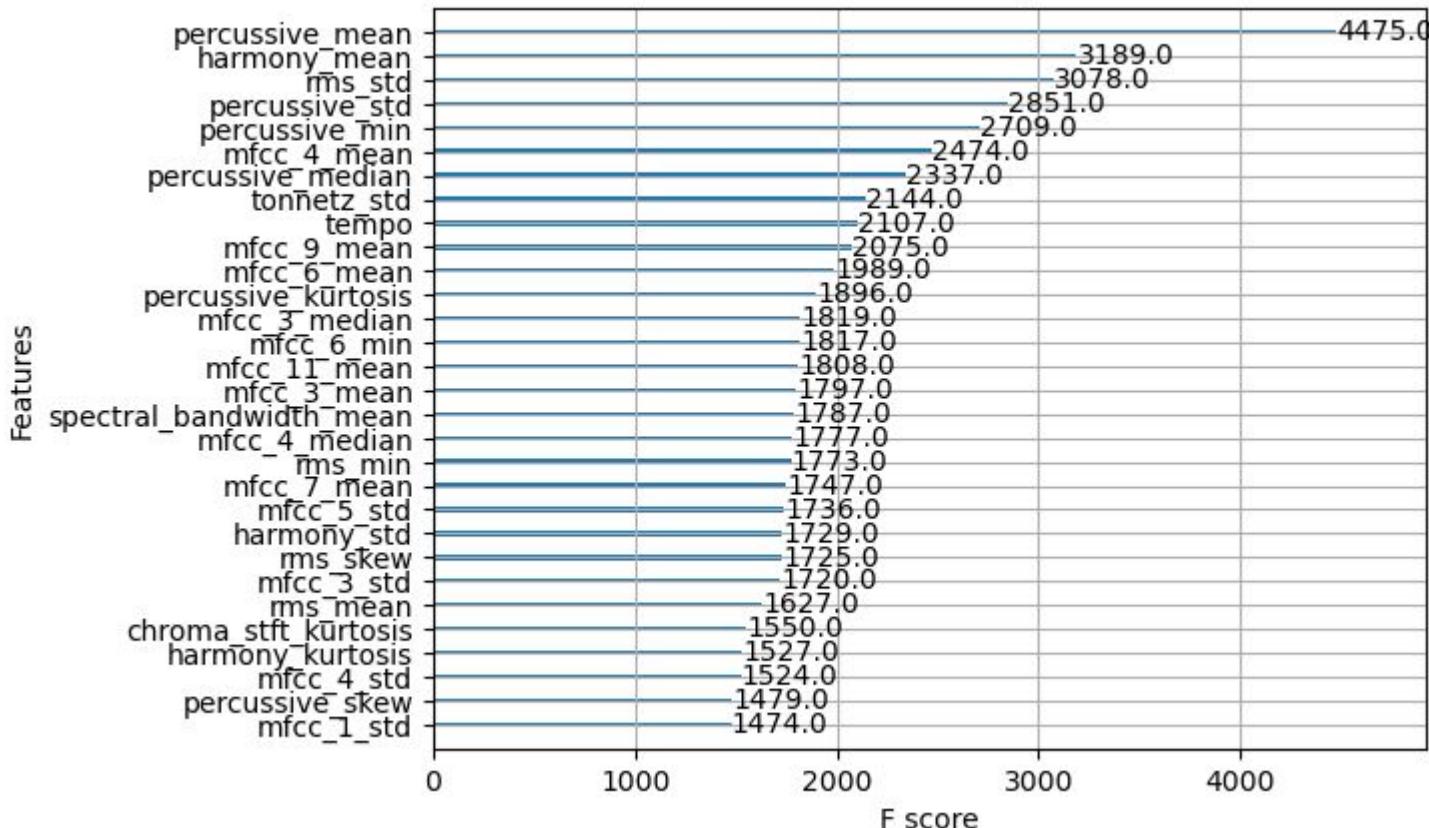
<u>Genre</u>	<u>Accuracy</u>	<u>Precision</u>
Blues	0.77	0.83
Classical	0.82	0.99
Country	0.70	0.75
Disco	0.78	0.68
Hiphop	0.70	0.75
Jazz	0.93	0.89
Metal	0.74	0.90
Pop	0.83	0.82
Reggae	0.80	0.69
Rock	0.52	0.43

* Train accuracy/precision was 100%

XGBoost Confusion Matrix



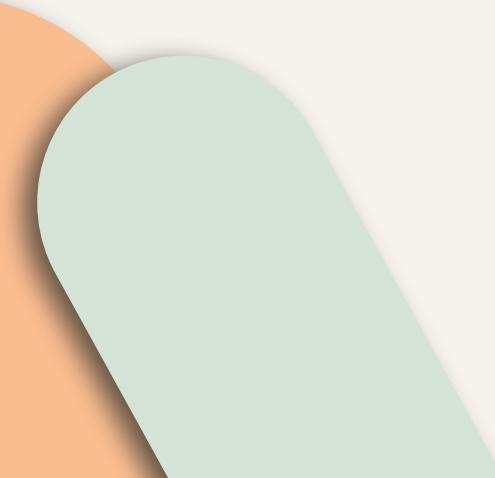
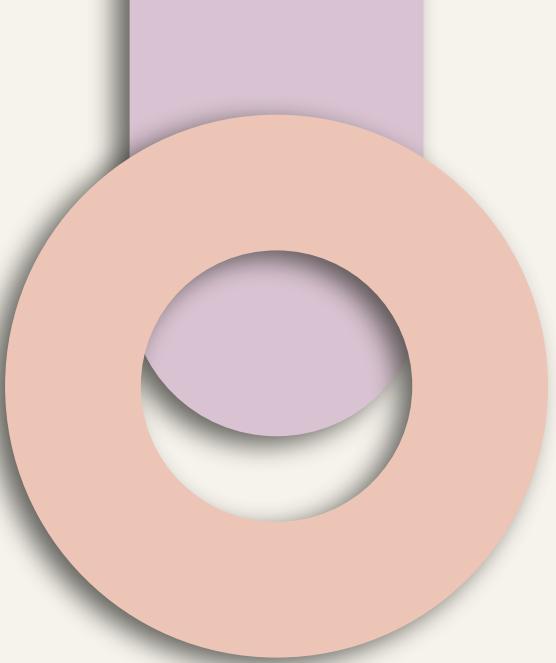
XGBoost Feature Importance



Results

Models	Accuracy	Precision	Recall	ROC-AUC
KNN	0.672	0.675	0.685	0.902
Decision Tree	0.518	0.526	0.534	0.740
Random Forest	0.716	0.706	0.723	0.949
XGBOOST	0.765	0.760	0.773	0.965
CNN	0.622	0.660	0.644	0.93

what about
newer music?





0:56 / 4:23



Ed Sheeran - Shape of You (Official Music Video)



Ed Sheeran



55.5M subscribers

Subscribe



33M



1.2K



Share



Download



6,379,014,191 views 7 years ago #divide #EdSheeran #ShapeOfYou

The official music video for Ed Sheeran - Shape Of You

Subtract, the new album, out 05.05.2023. Pre-order: es.lnk.to/subtract

more

Apply Best Model to Modern Data

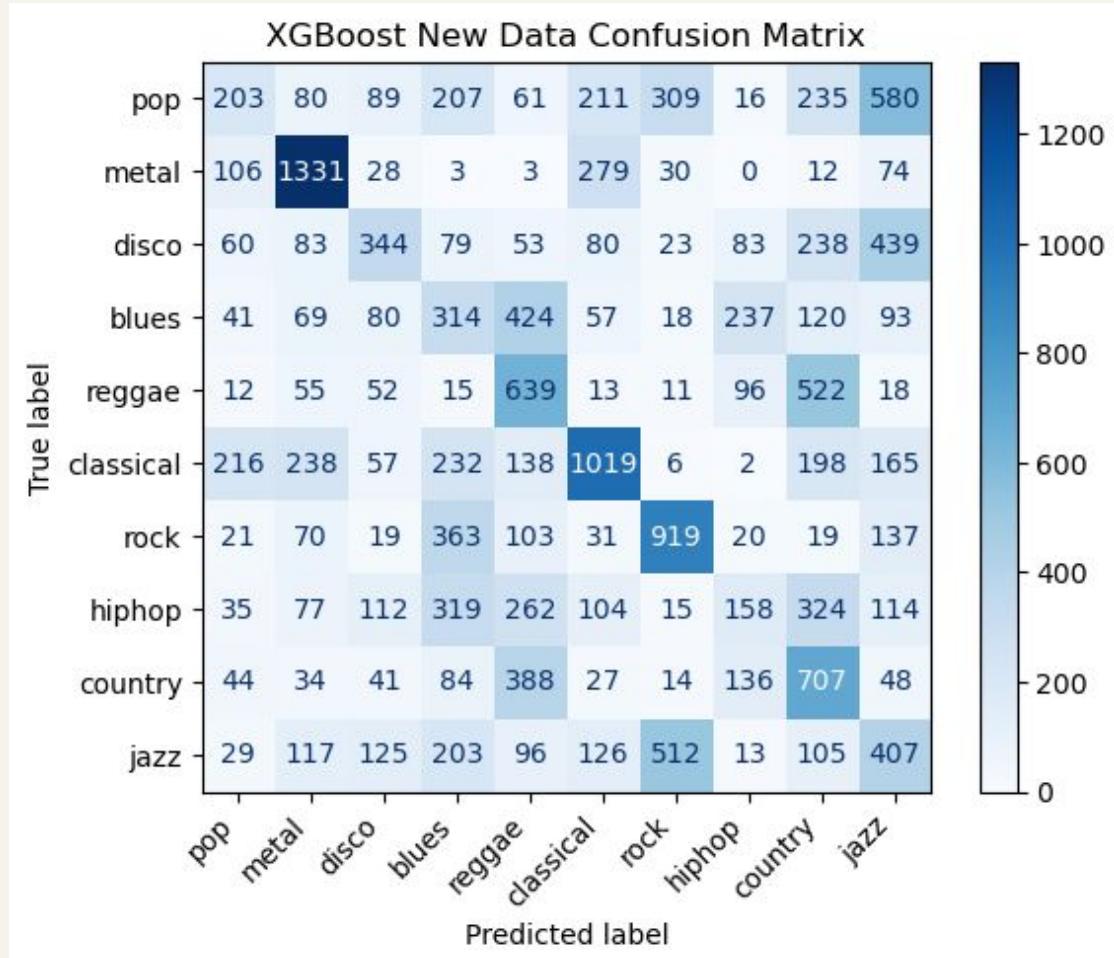
<u>Genre</u>	<u>Accuracy</u>	<u>Precision</u>
Blues	0.26	0.10
Classical	0.62	0.71
Country	0.36	0.23
Disco	0.17	0.22
Hiphop	0.29	0.45
Jazz	0.52	0.45
Metal	0.49	0.54
Pop	0.21	0.10
Reggae	0.29	0.46
Rock	0.20	0.23

Accuracy	0.356
Precision	0.342
Recall	0.350
F1 Score	0.334

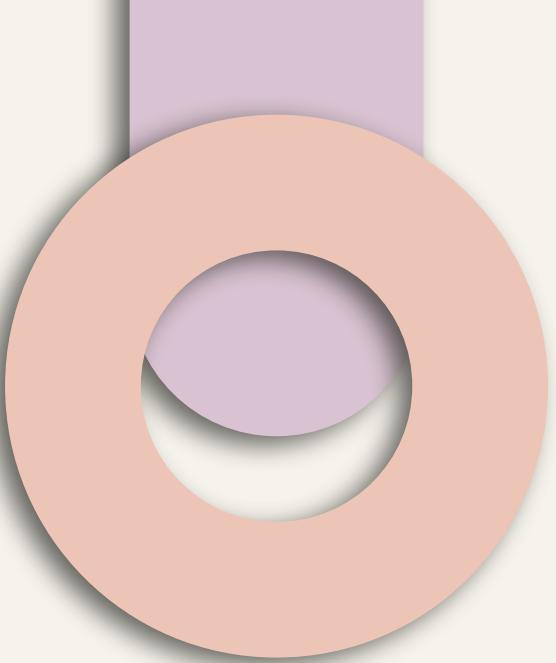
The overall accuracy was 0.356 Precision 0.342
Recall 0.350 F1 Score 0.334

<u>Genre</u>	Blues	Classical	Country	Disco	Hiphop	Jazz	Metal	Pop	Reggae	Rock
<u>Accuracy</u>	0.26	0.62	0.36	0.17	0.29	0.52	0.49	0.21	0.29	0.20
<u>Precision</u>	0.10	0.71	0.23	0.22	0.45	0.45	0.54	0.10	0.46	0.23

Data Drift



Future Scope



Applications

Music Recommendation Systems:

Enhance music streaming services with improved genre-based recommendations.

Digital Music Libraries Organization:

Automate and refine the categorization of large music collections.

Music Production Assistance:

Help producers align their tracks with genre-specific characteristics.

Enhancements

Real-Time Classification:

Optimize the model for immediate music genre classification in live settings.

Classify into Multiple Genres:

Many songs don't fit into exactly one bucket. Creating a binary classifier for each genre could be beneficial.

Global Scale Adaptation:

Expand the model to recognize and classify a diverse range of global music genres.

Conclusion

Importance of Model Selection:

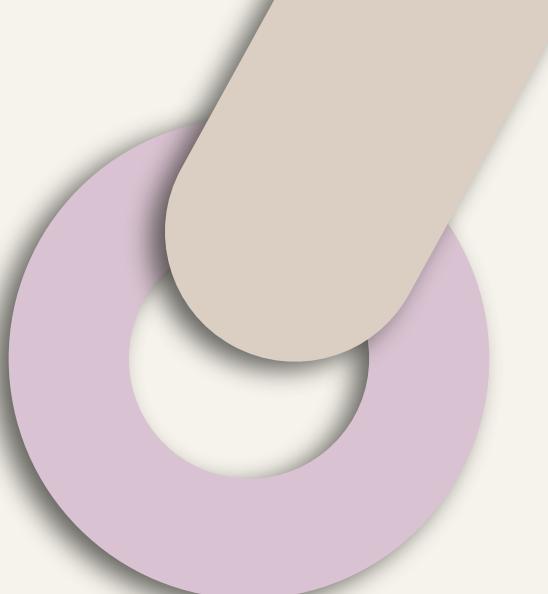
- Evaluated tradeoffs between different models and observed that ensemble models can excel at classifying items with many features

Insights Regarding Audio Data:

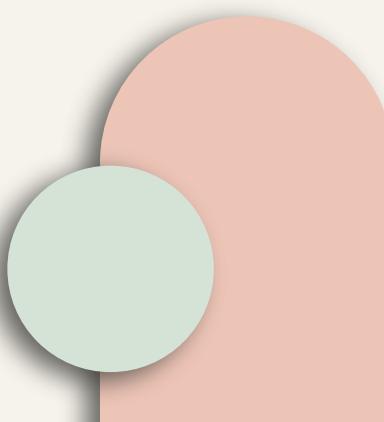
- Music genre classification requires models that are able to capture non-linear patterns and adapt to overlapping features

Model Optimization:

- Learned how to balance the challenges of computation time, hyperparameter tuning, and the complexities of our dataset to achieve optimal model performance



Thanks!



Appendix/Resources

- <https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification>
- <https://towardsdatascience.com/music-genre-classification-with-python-c714d032f0d8>
- <https://towardsdatascience.com/extract-features-of-music-75a3f9bc265d>
- <https://medium.com/bisa-ai/music-genre-classification-using-convolutional-neural-network-7109508ced47>
- <https://towardsdatascience.com/music-genre-classification-using-a-divide-conquer-crnn-2ff1cf49859f>
- <https://arxiv.org/pdf/1712.08370v1>