MUSIC GENRE CLASSIFICATION

Eric Heidbreder | DSIR-720

Problem Statement

PROBLEM STATEMENT

Using features derived from an audio source, can a classification model predict the genre of a 30-second audio clip with high enough accuracy to organize songs into their respective genres automatically?

WHO CARES?

Classifying genre is important for music **distribution and streaming platforms**. Automating the process would save time and money.

Assigning genre properly allows **listeners** to find new artists they might like, and, in turn, helps **musicians** connect with new audiences via genre-based playlists. Everyone's happy, everyone wins!

The Dataset

9365 Songs

- 30-second audio samples pulled from Spotify.
- Spotify API gives access to:
  - General song metadata
  - Pre-engineered features

Issues with the data?

- Genre connected to artist instead of song
- One song could have multiple genres
- Some songs have no 30-sec preview
- Some track names contain non-English characters
- Unbalanced classes

10 Genres

- Bootstrapped 500 songs per genre
- Some genres are closely related
- Created a subset of 5 general genres

What are the genres?

5-genre set

- Classical
- Rock
- Rap
- R&B
- Progressive Bluegrass

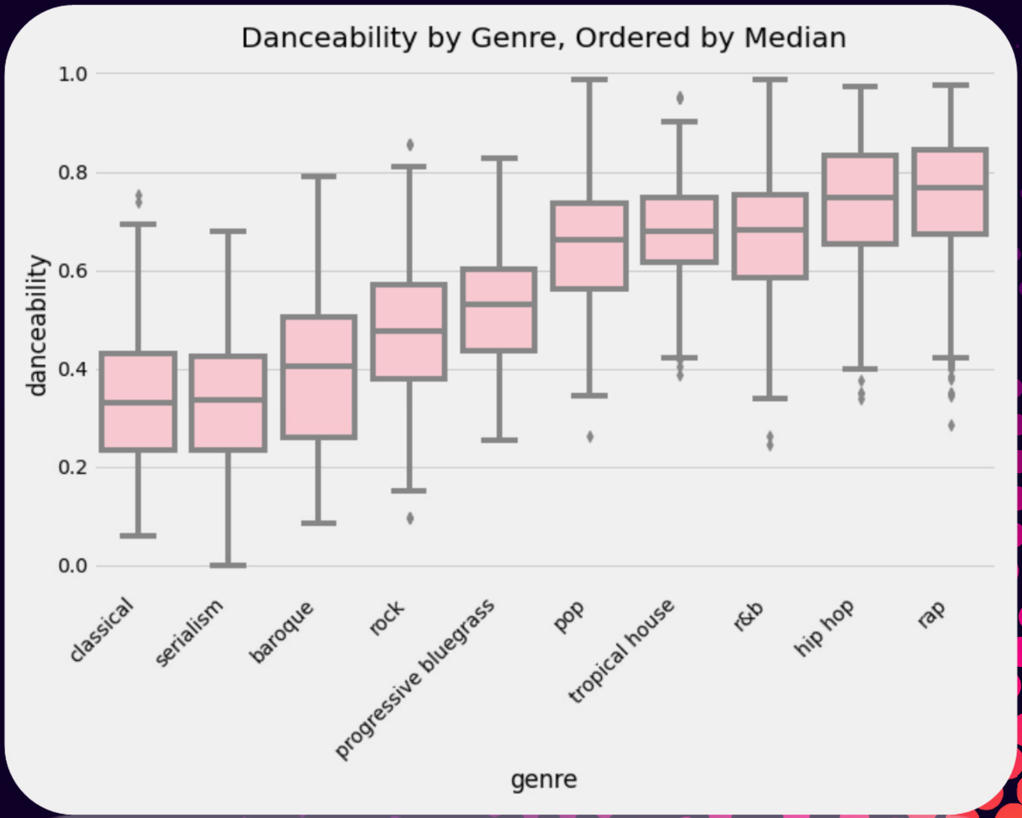
10-genre set (orig. 5 plus)

- Baroque
- Serialism
- Hip Hop
- Pop
- Tropical House

EDA

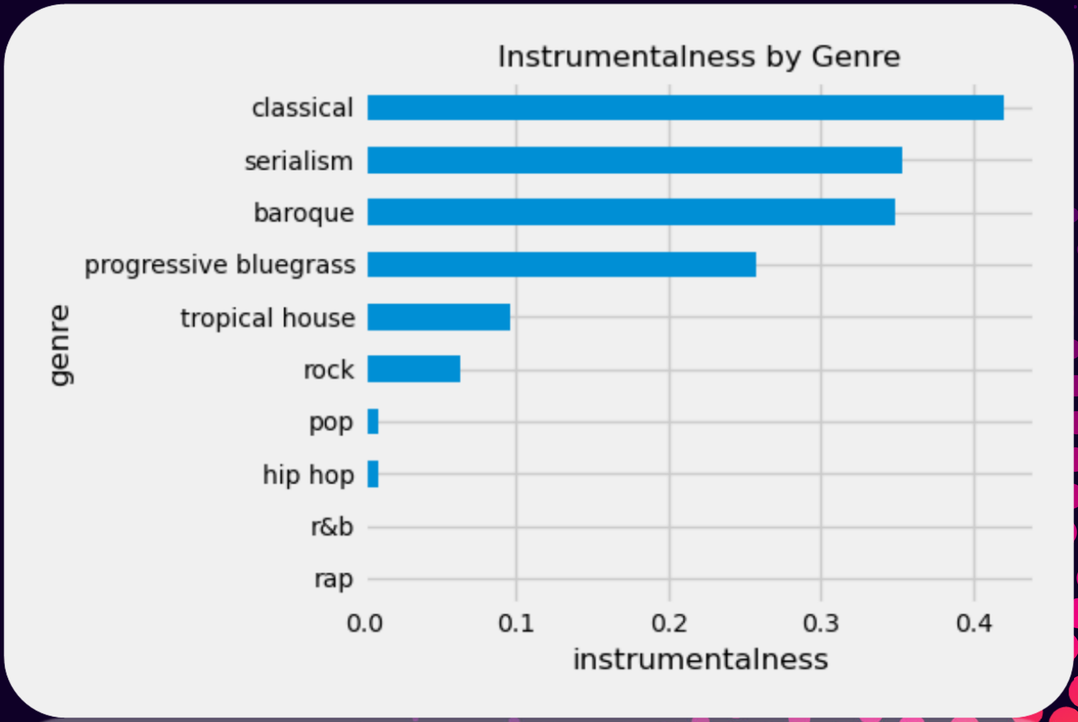
Danceability

Pre-made feature from Spotify combining many different rhythmic aspects of a song.



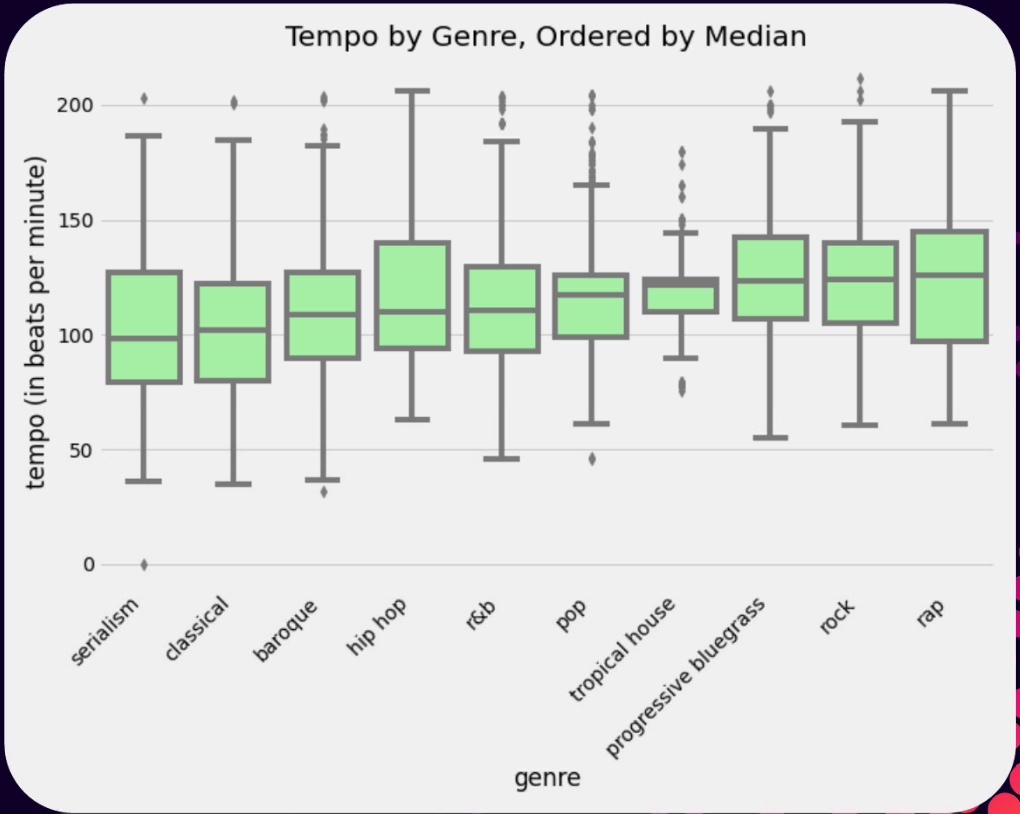
Instrumentalness

Pre-made feature from Spotify rating how much of the song is purely instrumental.

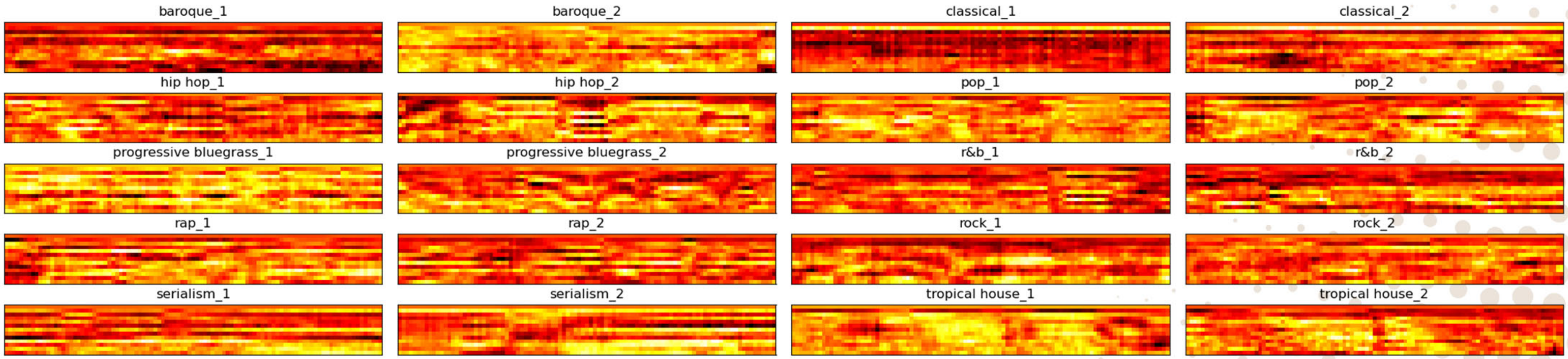


Tempo

Pre-made feature from Spotify showing the speed of the song.



Mel Frequency Cepstral Coefficient (MFCC)



Modeling

Metrics and Baseline Model

Metric

- Accuracy

Baseline Model

- 5-genres: 20% Accuracy
- 10-genres: 10% Accuracy

A Tale of Two Feature Sets

Model #1

Trained on Spotify's pre-made features only

Model #2

Trained on my own extracted features only

What features did we create?

- Volume/Loudness
  - Energy
  - Root Mean Squared Energy
- Power across frequency spectrum
  - MFCC

What models did we consider?

SVC

- Good training/testing accuracy
- Takes seconds to fit model
- CON: Does not scale well to large datasets

Convolutional Neural Net

- Similar training/testing accuracy as SVC
- Large dataset (>3 GB)
- CON: Can't compare Spotify features to newly-extracted features.

What models did we consider?

SVC

- Good training/testing accuracy
- Takes seconds to fit model
- CON: Does not scale well to large datasets

Convolutional Neural Net

- Similar training/testing accuracy as SVC
- Large dataset (>3 GB)
- CON: Can't compare Spotify features to newly-extracted features.

Model Comparison:
Spotify features vs. my features

# of Genres	Acc (Spotify features)	Acc (my features)
5	78.4%	84.2%
10	50.6%	56.5%

5.85%

Mean Improvement when compared to model trained on Spotify features

10-Genre Classifier Breakdown

Pop

Genre predicted correctly *least often* by both models (Spotify: 14.4%, Extracted: 17.6%)

Progressive Bluegrass

Genre predicted correctly *most often* by both models (Spotify: 77.6%, Extracted: 84%)

Pop Quiz

(or maybe Tropical House Quiz)

Digging in - why is our model confused?

- What genre would you classify this as?
  - Pop
  - Classical
  - Tropical House

Digging in - why is our model confused?

- What genre would you classify this as?
  - Pop → Predicted Genre
  - Classical
  - Tropical House → Correct Genre!

Drums not as heavy as other Tropical House tracks?

Digging in - why is our model confused?

- What genre would you classify this as?
  - Rock
  - Progressive Bluegrass
  - Tropical House

Digging in - why is our model confused?

- What genre would you classify this as?
  - Rock → Correct Genre!
  - Progressive Bluegrass
  - Tropical House → Predicted Genre

Dynamic Contrast and "Drops" in Tropical House Music

Digging in - why is our model confused?

- What genre would you classify this as?
  - Rap
  - Tropical House
  - Hip Hop

Digging in - why is our model confused?

- What genre would you classify this as?
  - Rap → Correct Genre!
  - Tropical House → Predicted Genre
  - Hip Hop

Bad 30-second draw, the hook of this Rap song seems to be inspired by Tropical House genre

Conclusions

CONCLUSIONS

Good things

Our extracted-features model is outperforming the baseline as well as a similar model trained on Spotify's features.

When predicting wrong, the models tend to predict related genres.

Bad things

~56% accuracy is not good enough to automatically assign genre at a large scale. We'd wind up with a lot of misclassified genres and confused listeners.

What I learned

Sub-genres are not very well defined at the sonic level. There are likely other factors that go into the creation of sub-genre. (Artist location, lyrical content, time period, instrumentation)

Future Steps

- Manually assign genres per song
- Instrument extraction
- Unsupervised learning for more accurate genres

SOURCES

- Presentation template by [SlidesCarnival]
- Photographs by [Unsplash]

1. https://musicinformationretrieval.com/energy.html
2. https://musicinformationretrieval.com/mfcc.html
3. Sahidullah, Md.; Saha, Goutam (May 2012). "Design, analysis and experimental evaluation of block-based transformation in MFCC computation for speaker recognition". Speech Communication. 54 (4): 543–565. doi:10.1016/j.specom.2011.11.004.
4. Adams, Seth. "DSP Background – Deep Learning for Audio Classification". https://www.youtube.com/watch?v=Z7YM-HAz-IY&list=PLhA3b2k8R3t2Ng1WW_7MiXeh1pfQJQi_P

THANKS!

Any questions?