Spotify Popularity Prediction

Robert Bull, Uriel Garcia, Ahsin Saleem, Ross Vrbanac
DS w207- Tanya Roosta
Dec. 12th, 2024

Table of Contents

1. Data Source
2. Research Question
3. Feature Overview
4. Data Processing
5. Genre 1 - Death Metal
6. Genre 2 - Hip Hop
7. Genre 3 - Salsa
8. Genre 4 - Piano
9. Conclusion

Data Source:

- Switched our original dataset from crop yield to Spotify popularity prediction.
- **New data:**
    - Kaggle Spotify 1 million tracks dataset.
    - Extracted from the Spotify using their API
    - https://www.kaggle.com/datasets/amitanshjoshi/spotify-1million-tracks/data
- **Key Highlights:**
    - Size: ~1 Million tracks.
    - Features: 19 musical and metadata features.
    - Artists: 61,445 unique artists.
    - Genres: 82 distinct genres.

Research Question:

What is the best performing model for classifying Spotify song popularity per genre, irrespective of artist?

Sub-Question: What features define each genre?

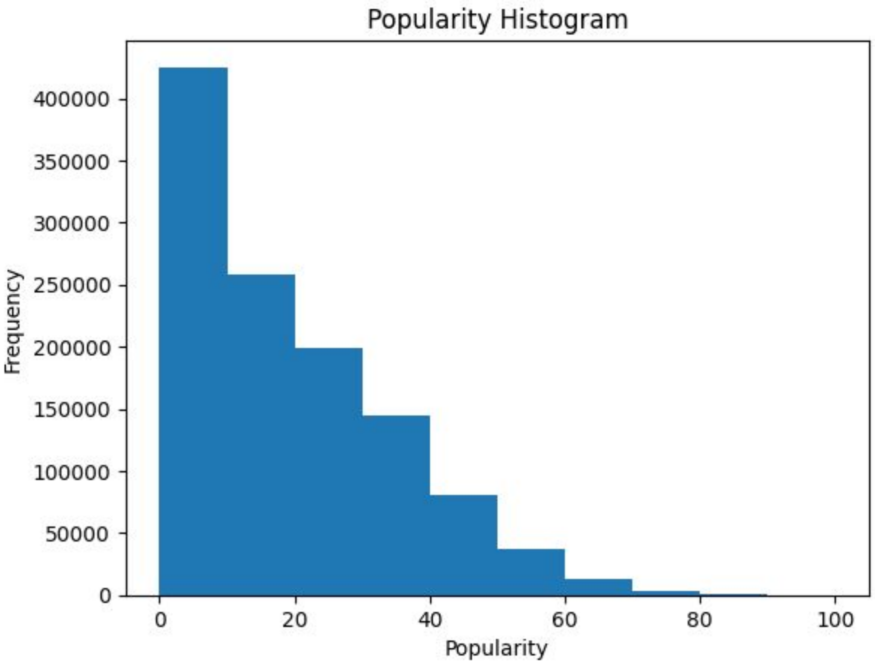
Sub-Question: How well can the popularity per song of different genres be predicted?

Features:

Audio Features	Description
Popularity	Track popularity (0 to 100)
Year	Year released (2000 to 2023)
Danceability	Track suitability for dancing (0.0 to 1.0)
Energy	The perceptual measure of intensity and activity (0.0 to 1.0)
Key	The key, the track is in (-1 to -11)
Loudness	Overall loudness of track in decibels (-60 to 0 dB)
Mode	Modality of the track (Major '1'/ Minor '0')
Speechiness	Presence of spoken words in the track
Acousticness	Confidence measure from 0 to 1 of whether the track is acoustic
Instrumentalness	Whether tracks contain vocals. (0.0 to 1.0)
Liveness	Presence of audience in the recording (0.0 – 1.0)
Valence	Musical positiveness (0.0 to 1.0)
Tempo	Tempo of the track in beats per minute (BPM)
Time_signature	Estimated time signature (3 to 7)
Duration_ms	Duration of track in milliseconds

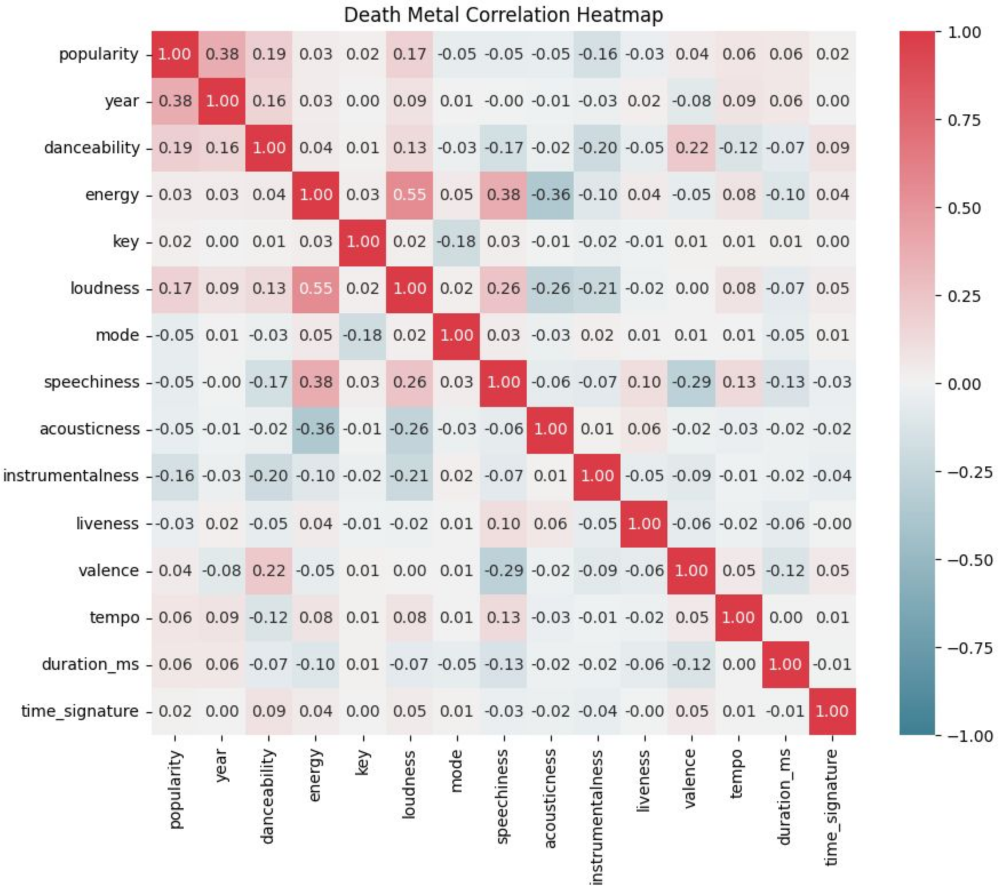
Data Processing

- Popularity separated into binary classes
  - 0 for popularity < median popularity per genre
  - 1 for popularity >= median popularity per genre
- Min-max standardization of numeric features
- One-hot encoding of categorical features
- 60/20/20 train/validate/test splits with shuffled data

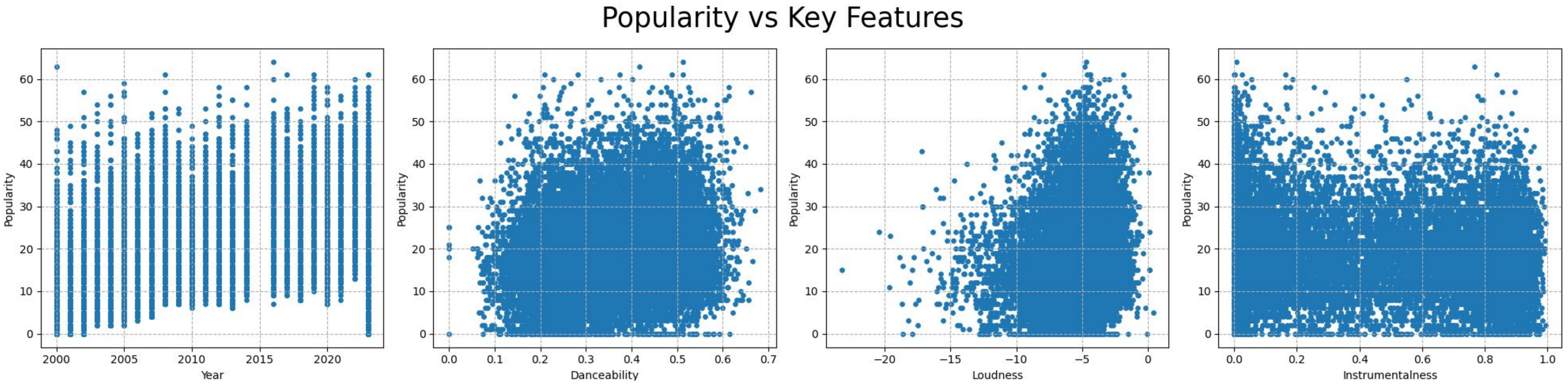


Death Metal

Death Metal: EDA

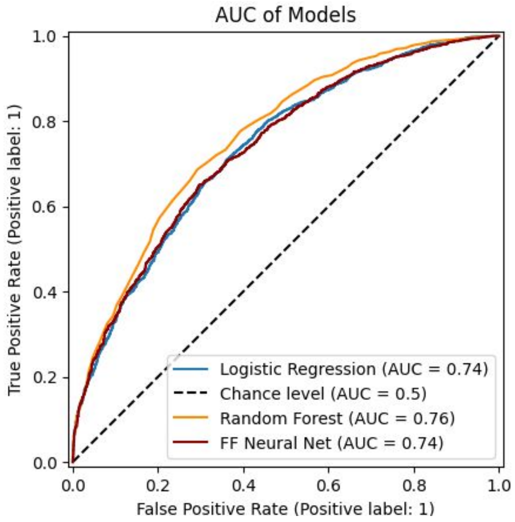


Death Metal: EDA



Death Metal: Modeling

Model	Accuracy	Precision	AUC Score
Baseline	0.52	0.52	0.5
Logistic Regression	0.67	0.67	0.74
Random Forest Classifier	0.69	0.69	0.76
FF Neural Net	0.67	0.68	0.74

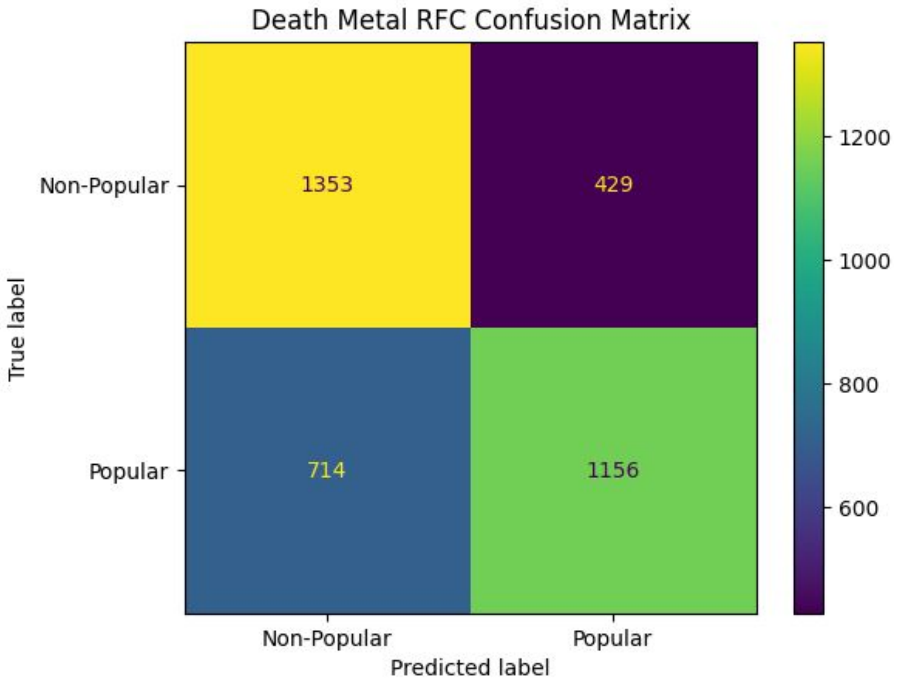


- Comparisons on validation set
- Baseline - majority class classifier
- Logistic Regression - L2 penalization
- Random Forest Classifier - no maximum depth, splits based on Gini impurity
- FF Neural Net - 4 hidden layers of sizes [204, 204, 102, 102], 0.1 dropout layer for each, and tuned hyperparameters

Death Metal: Results (Random Forest)

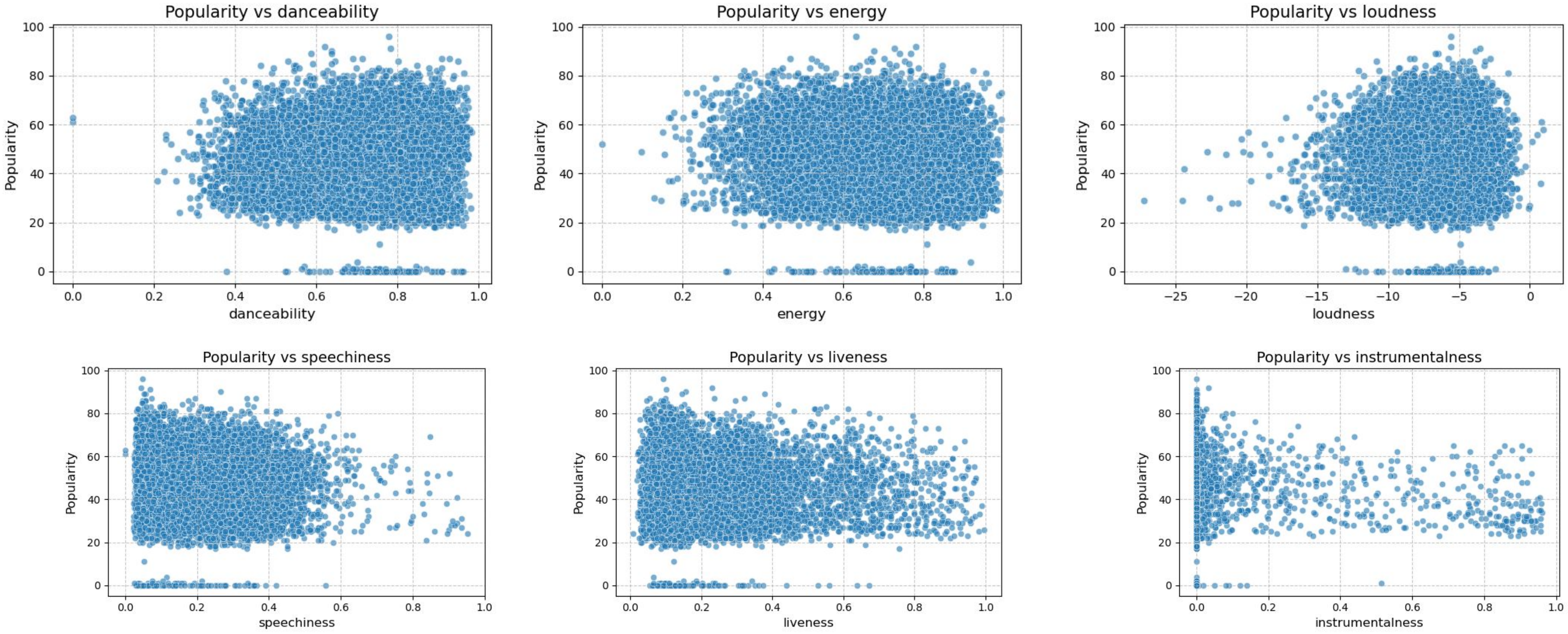
	Feature	Feature Importance
0	danceability	0.083296
1	loudness	0.081751
2	instrumentalness	0.076436
3	duration_ms	0.072134
4	tempo	0.069519
5	speechiness	0.069384
6	acousticness	0.068290
7	valence	0.068242
8	liveness	0.065552
9	energy	0.063458

```
Death Metal RFC Test Accuracy:
0.7009423503325942
Death Metal RFC Test Precision:
0.7036852589641435
Death Metal RFC Test AUC Score:
0.7622558858969235
```



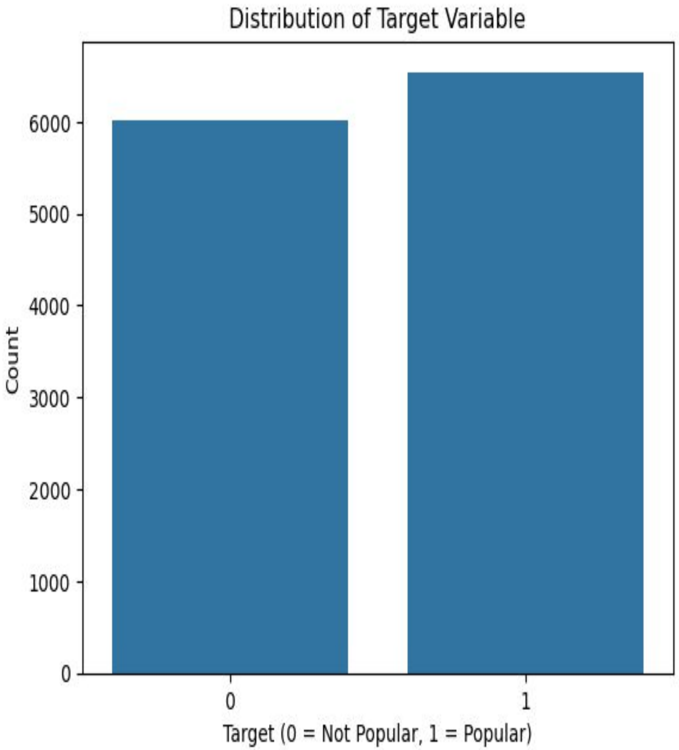
Hip-Hop

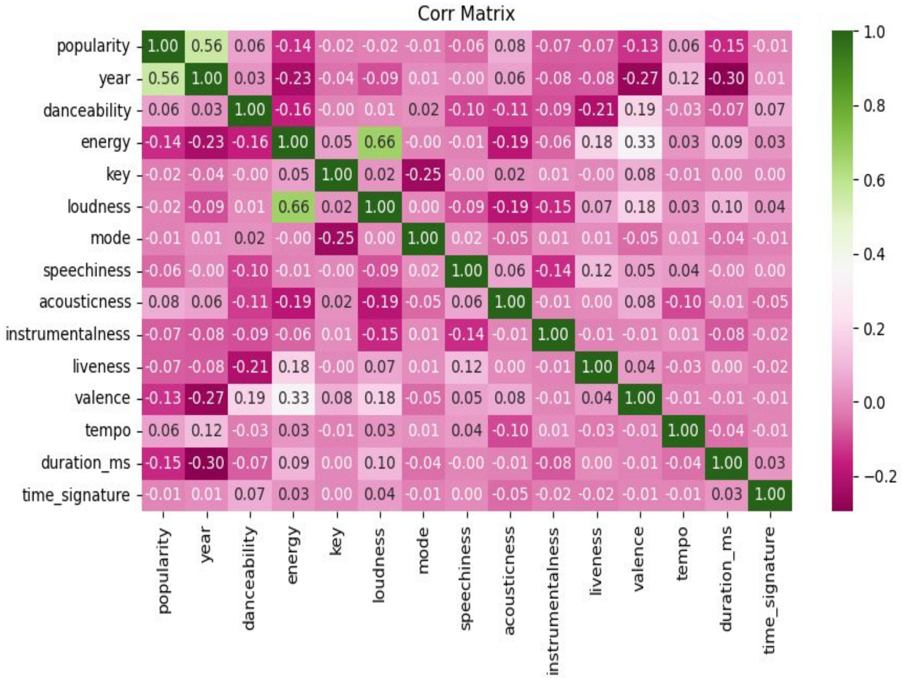
Hip-Hop: EDA

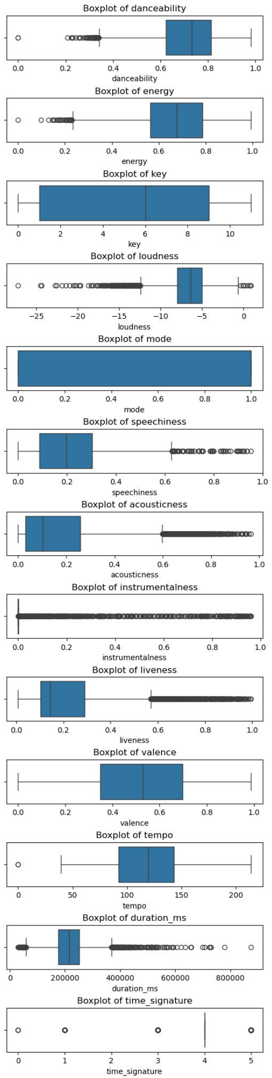


Popularity vs danceability
Popularity
danceability
Popularity vs energy
Popularity
energy
Popularity vs loudness
Popularity
loudness
Popularity vs speechiness
Popularity
speechiness
Popularity vs liveness
Popularity
liveness
Popularity vs instrumentalness
Popularity
instrumentalness

Hip-Hop: EDA

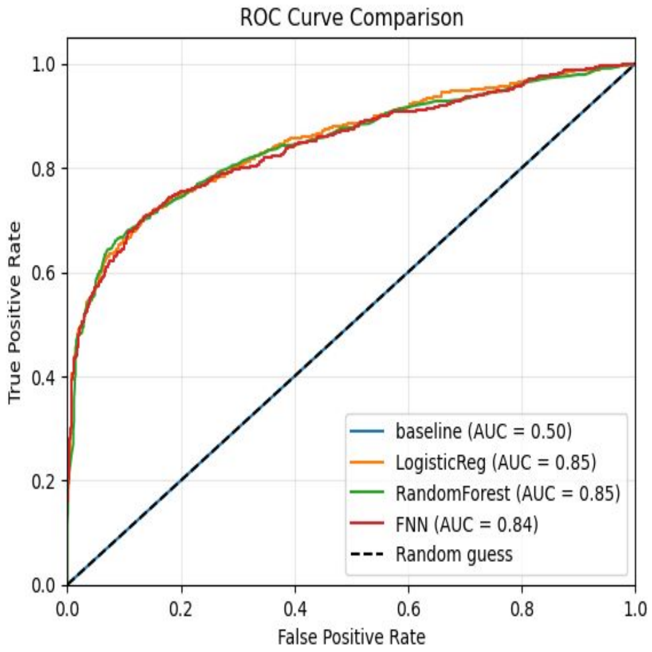

Distribution of Target Variable
Count
Target (0 = Not Popular, 1 = Popular)
0
1
0
1000
2000
3000
4000
5000
6000


Corr Matrix
popularity 1.00 0.56 0.06 -0.14 -0.02 -0.02 -0.01 -0.06 0.08 -0.07 -0.07 -0.13 0.06 -0.15 -0.01
year 0.56 1.00 0.03 -0.23 -0.04 -0.09 0.01 -0.00 0.06 -0.08 -0.08 -0.27 0.12 -0.30 0.01
danceability 0.06 0.03 1.00 -0.16 -0.00 0.01 0.02 -0.10 -0.11 -0.09 -0.21 0.19 -0.03 -0.07 0.07
energy -0.14 -0.23 -0.16 1.00 0.05 0.66 -0.00 -0.01 -0.19 -0.06 0.18 0.33 0.03 0.09 0.03
key -0.02 -0.04 -0.00 0.05 1.00 0.02 -0.25 -0.00 0.02 0.01 -0.00 0.08 -0.01 0.00 0.00
loudness -0.02 -0.09 0.01 0.66 0.02 1.00 0.00 -0.09 -0.19 -0.15 0.07 0.18 0.03 0.10 0.04
mode -0.01 0.01 0.02 -0.00 -0.25 0.00 1.00 0.02 -0.05 0.01 0.01 -0.05 0.01 -0.04 -0.01
speechiness -0.06 -0.00 -0.10 -0.01 -0.00 -0.09 0.02 1.00 0.06 -0.14 0.12 0.05 0.04 -0.00 0.00
acousticness 0.08 0.06 -0.11 -0.19 0.02 -0.19 -0.05 0.06 1.00 -0.01 0.00 0.08 -0.10 -0.01 -0.05
instrumentalness -0.07 -0.08 -0.09 -0.06 0.01 -0.15 0.01 -0.14 -0.01 1.00 -0.01 -0.01 0.01 -0.08 -0.02
liveness -0.07 -0.08 -0.21 0.18 -0.00 0.07 0.01 0.12 0.00 -0.01 1.00 0.04 -0.03 0.00 -0.02
valence -0.13 -0.27 0.19 0.33 0.08 0.18 -0.05 0.05 0.08 -0.01 0.04 1.00 -0.01 -0.01 -0.01
tempo 0.06 0.12 -0.03 0.03 -0.01 0.03 0.01 0.04 -0.10 0.01 -0.03 -0.01 1.00 -0.04 -0.01
duration_ms -0.15 -0.30 -0.07 0.09 0.00 0.10 -0.04 -0.00 -0.01 -0.08 0.00 -0.01 -0.04 1.00 0.03
time_signature -0.01 0.01 0.07 0.03 0.00 0.04 -0.01 0.00 -0.05 -0.02 -0.02 -0.01 -0.01 0.03 1.00
popularity
year
danceability
energy
key
loudness
mode
speechiness
acousticness
instrumentalness
liveness
valence
tempo
duration_ms
time_signature
1.0
0.8
0.6
0.4
0.2
0.0
−0.2


Boxplot of danceability
danceability
Boxplot of energy
energy
Boxplot of key
key
Boxplot of loudness
loudness
Boxplot of mode
mode
Boxplot of speechiness
speechiness
Boxplot of acousticness
acousticness
Boxplot of instrumentalness
instrumentalness
Boxplot of liveness
liveness
Boxplot of valence
valence
Boxplot of tempo
tempo
Boxplot of duration_ms
duration_ms
Boxplot of time_signature
time_signature

Hip-Hop: Modeling

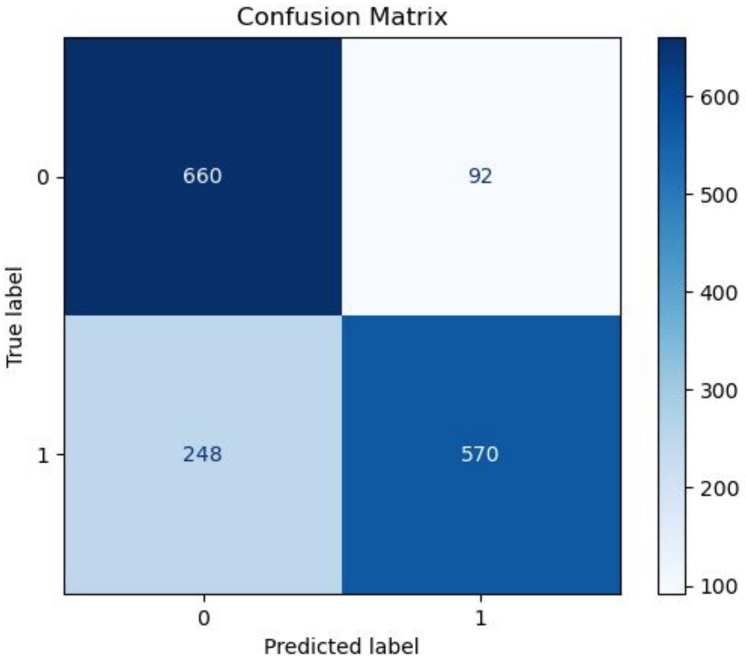
Model	Accuracy	Precision	AUC Score
Baseline	0.52	0.52	0.5
Logistic Regression	0.78	0.83	0.85
Random Forest Classifier	0.78	0.85	0.85
FF Neural Net	0.77	0.81	0.84



- Comparisons on validation set
- Baseline - majority class classifier
- Logistic Regression - L1 penalization
- Random Forest Classifier - no maximum depth, splits based on Gini impurity
- FF Neural Net - 4 hidden layers of sizes [256, 128, 64, 32], 0.4 dropout layer for each, and tuned hyperparameters

Hip-Hop: Results (Random Forest)

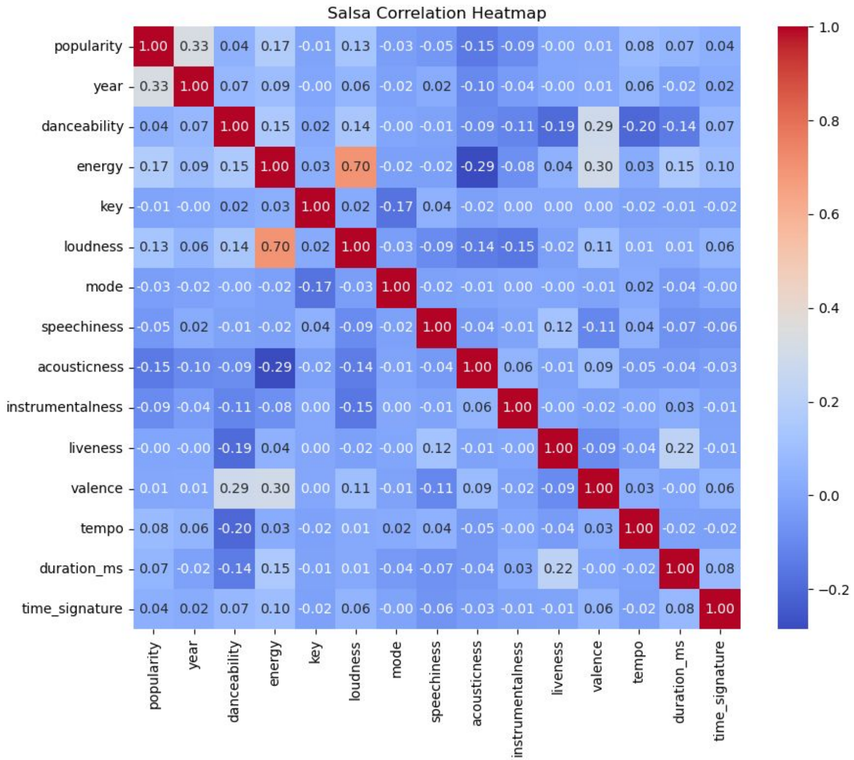
	Feature	Coefficient
8	duration_ms	0.075731
6	valence	0.061783
1	energy	0.061680
4	acousticness	0.060087
7	tempo	0.059545
3	speechiness	0.058403
0	danceability	0.058317
5	liveness	0.057761
2	loudness	0.057374
30	year_2022	0.050634

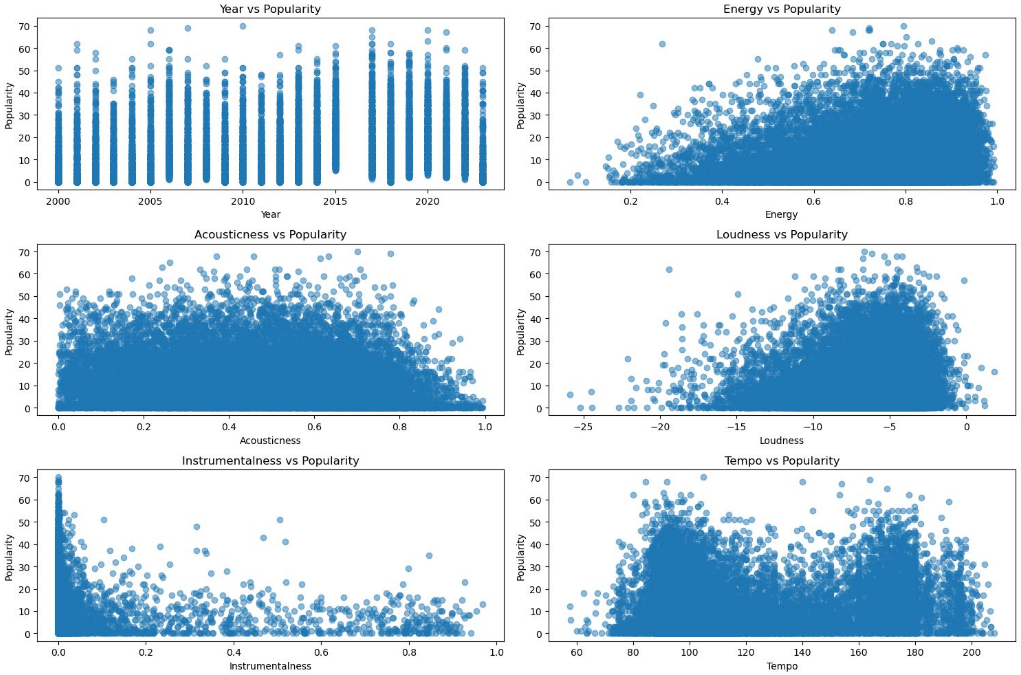


Final Results:
- Test Accuracy: 0.78
- Test Precision: 0.86
- Test_AUC: 0.78

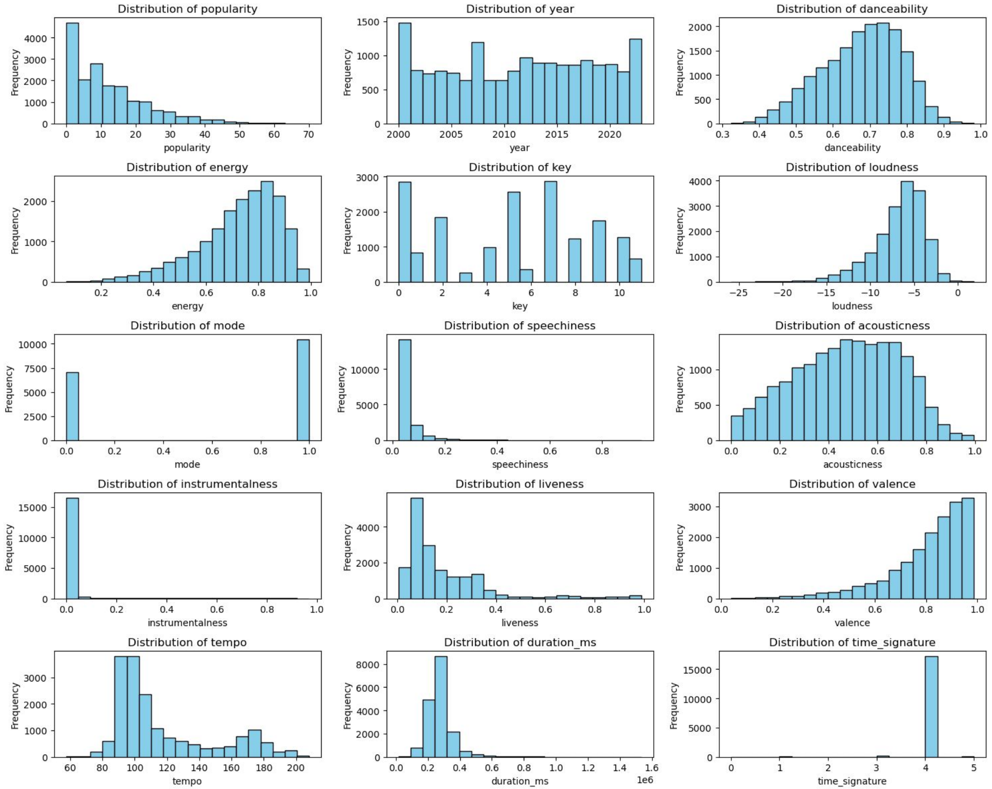
Salsa

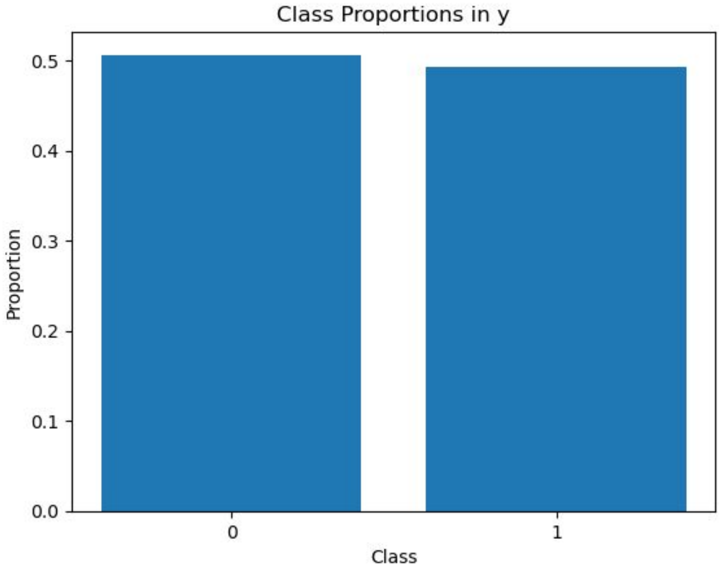
Salsa: EDA


Salsa Correlation Heatmap
popularity 1.00 0.33 0.04 0.17 -0.01 0.13 -0.03 -0.05 -0.15 -0.09 -0.00 0.01 0.08 0.07 0.04
year 0.33 1.00 0.07 0.09 -0.00 0.06 -0.02 0.02 -0.10 -0.04 -0.00 0.01 0.06 -0.02 0.02
danceability 0.04 0.07 1.00 0.15 0.02 0.14 -0.00 -0.01 -0.09 -0.11 -0.19 0.29 -0.20 -0.14 0.07
energy 0.17 0.09 0.15 1.00 0.03 0.70 -0.02 -0.02 -0.29 -0.08 0.04 0.30 0.03 0.15 0.10
key -0.01 -0.00 0.02 0.03 1.00 0.02 -0.17 0.04 -0.02 0.00 0.00 0.00 -0.02 -0.01 -0.02
loudness 0.13 0.06 0.14 0.70 0.02 1.00 -0.03 -0.09 -0.14 -0.15 -0.02 0.11 0.01 0.01 0.06
mode -0.03 -0.02 -0.00 -0.02 -0.17 -0.03 1.00 -0.02 -0.01 0.00 -0.00 -0.01 0.02 -0.04 -0.00
speechiness -0.05 0.02 -0.01 -0.02 0.04 -0.09 -0.02 1.00 -0.04 -0.01 0.12 -0.11 0.04 -0.07 -0.06
acousticness -0.15 -0.10 -0.09 -0.29 -0.02 -0.14 -0.01 -0.04 1.00 0.06 -0.01 0.09 -0.05 -0.04 -0.03
instrumentalness -0.09 -0.04 -0.11 -0.08 0.00 -0.15 0.00 -0.01 0.06 1.00 -0.00 -0.02 -0.00 0.03 -0.01
liveness -0.00 -0.00 -0.19 0.04 0.00 -0.02 -0.00 0.12 -0.01 -0.00 1.00 -0.09 -0.04 0.22 -0.01
valence 0.01 0.01 0.29 0.30 0.00 0.11 -0.01 -0.11 0.09 -0.02 -0.09 1.00 0.03 -0.00 0.06
tempo 0.08 0.06 -0.20 0.03 -0.02 0.01 0.02 0.04 -0.05 -0.00 -0.04 0.03 1.00 -0.02 -0.02
duration_ms 0.07 -0.02 -0.14 0.15 -0.01 0.01 -0.04 -0.07 -0.04 0.03 0.22 -0.00 -0.02 1.00 0.08
time_signature 0.04 0.02 0.07 0.10 -0.02 0.06 -0.00 -0.06 -0.03 -0.01 -0.01 0.06 -0.02 0.08 1.00
popularity year danceability energy key loudness mode speechiness acousticness instrumentalness liveness valence tempo duration_ms time_signature
1.0 0.8 0.6 0.4 0.2 0.0 −0.2


Year vs Popularity
Year
Popularity
Energy vs Popularity
Energy
Acousticness vs Popularity
Acousticness
Loudness vs Popularity
Loudness
Instrumentalness vs Popularity
Instrumentalness
Tempo vs Popularity
Tempo

Salsa: EDA (part 2)

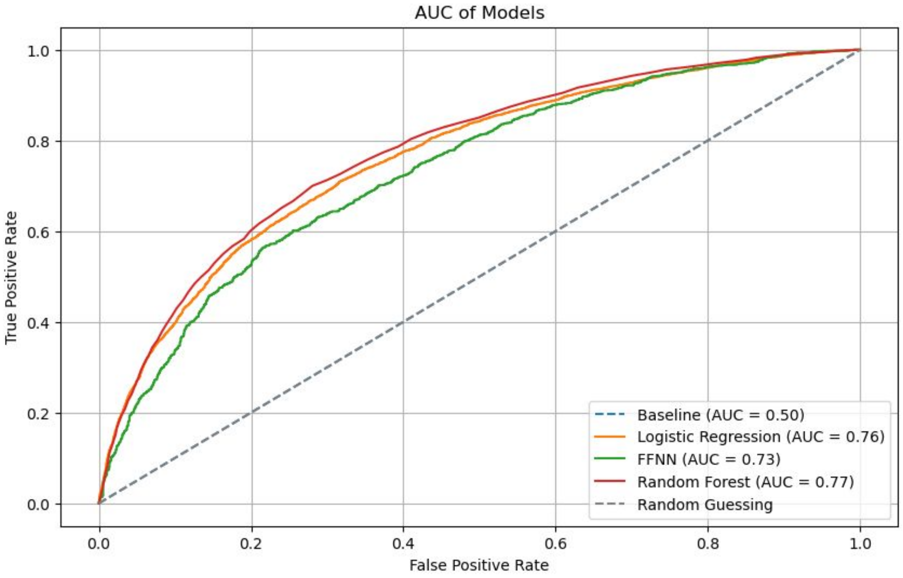




Salsa: Modeling

- Comparisons on validation set
- Baseline - majority class classifier
- Logistic Regression - L1 penalization
- Random Forest Classifier - no maximum depth, splits based on Gini impurity
- FF Neural Net - 2 hidden layers of sizes [128, 64], 0.3 dropout layer and tuned hyperparameters

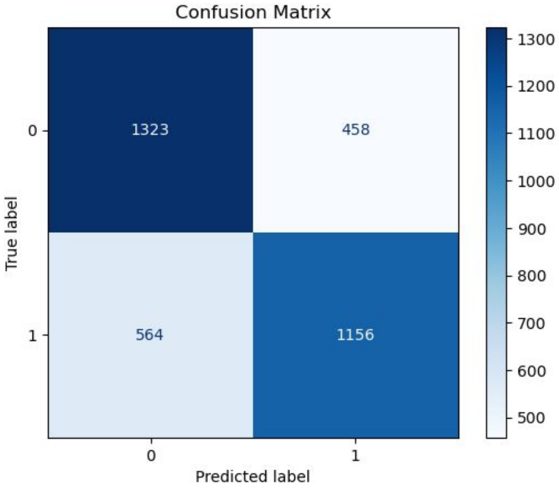
Model	Accuracy	Precision	AUC
Baseline	0.50	0.50	0.50
Logistic Regression	0.70	0.70	0.76
FNN	0.67	0.68	0.73
Random Forest	0.70	0.71	0.77

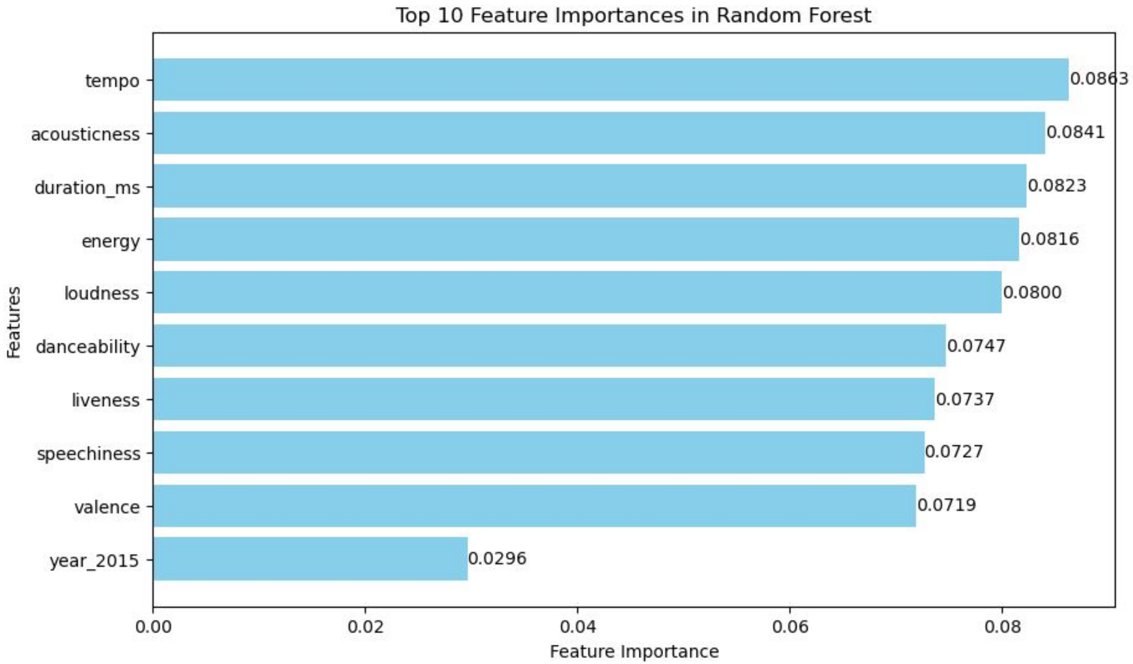


Salsa: Results(Random Forest)

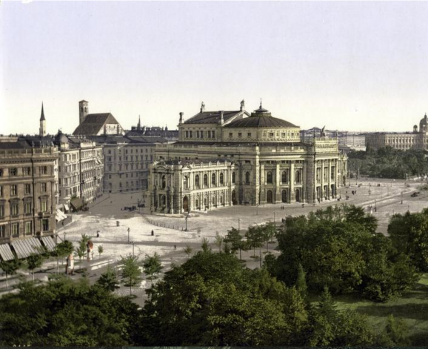
Final Results:

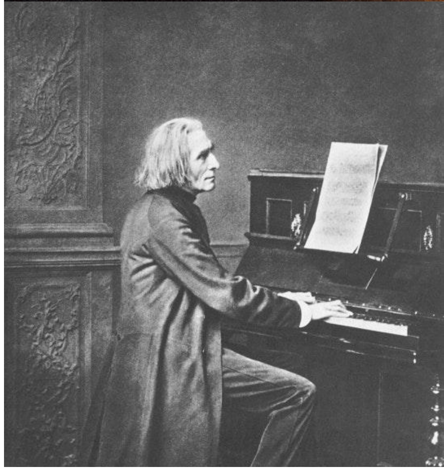
- Test Accuracy: 0.70
- Test Precision: 0.71
- Test AUC Score: 0.77



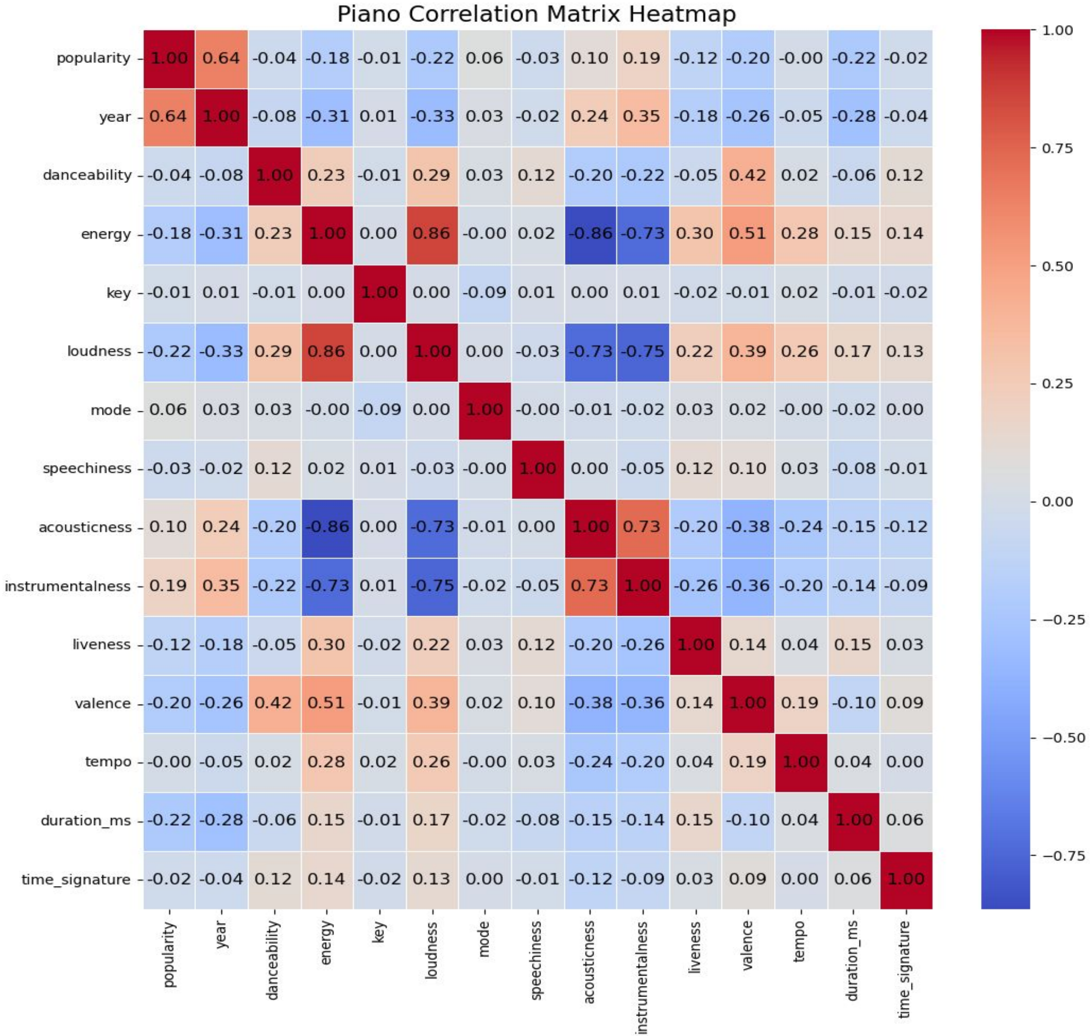


Piano



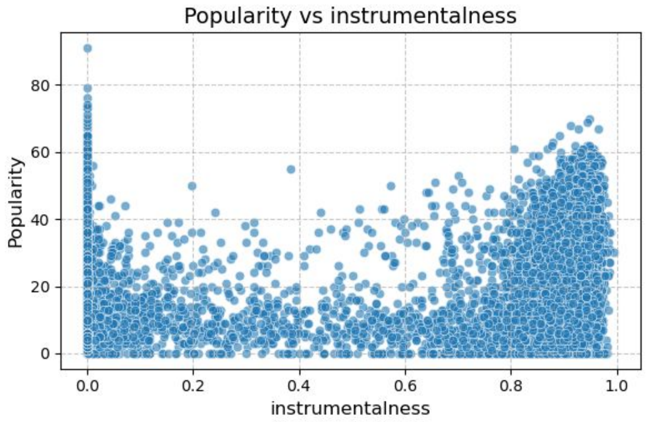


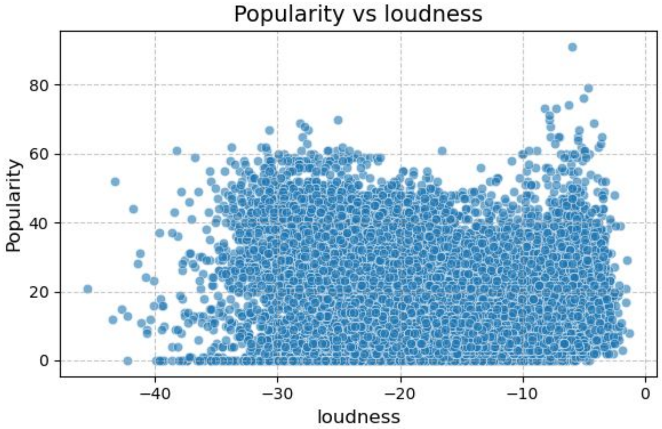
Piano: EDA

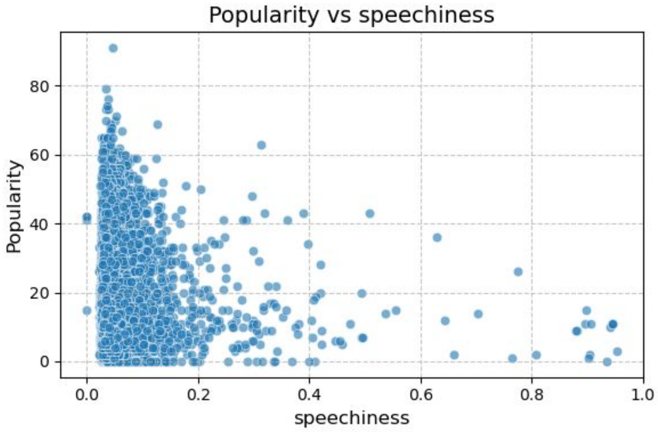


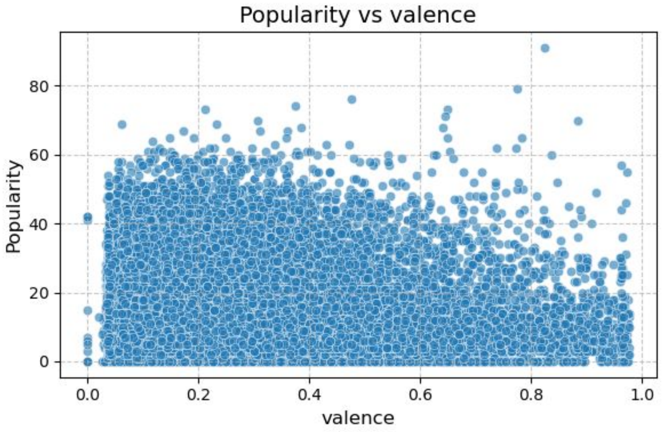
Piano Correlation Matrix Heatmap
popularity 1.00 0.64 -0.04 -0.18 -0.01 -0.22 0.06 -0.03 0.10 0.19 -0.12 -0.20 -0.00 -0.22 -0.02
year 0.64 1.00 -0.08 -0.31 0.01 -0.33 0.03 -0.02 0.24 0.35 -0.18 -0.26 -0.05 -0.28 -0.04
danceability -0.04 -0.08 1.00 0.23 -0.01 0.29 0.03 0.12 -0.20 -0.22 -0.05 0.42 0.02 -0.06 0.12
energy -0.18 -0.31 0.23 1.00 0.00 0.86 -0.00 0.02 -0.86 -0.73 0.30 0.51 0.28 0.15 0.14
key -0.01 0.01 -0.01 0.00 1.00 0.00 -0.09 0.01 0.00 0.01 -0.02 -0.01 0.02 -0.01 -0.02
loudness -0.22 -0.33 0.29 0.86 0.00 1.00 0.00 -0.03 -0.73 -0.75 0.22 0.39 0.26 0.17 0.13
mode 0.06 0.03 0.03 -0.00 -0.09 0.00 1.00 -0.00 -0.01 -0.02 0.03 0.02 -0.00 -0.02 0.00
speechiness -0.03 -0.02 0.12 0.02 0.01 -0.03 -0.00 1.00 0.00 -0.05 0.12 0.10 0.03 -0.08 -0.01
acousticness 0.10 0.24 -0.20 -0.86 0.00 -0.73 -0.01 0.00 1.00 0.73 -0.20 -0.38 -0.24 -0.15 -0.12
instrumentalness 0.19 0.35 -0.22 -0.73 0.01 -0.75 -0.02 -0.05 0.73 1.00 -0.26 -0.36 -0.20 -0.14 -0.09
liveness -0.12 -0.18 -0.05 0.30 -0.02 0.22 0.03 0.12 -0.20 -0.26 1.00 0.14 0.04 0.15 0.03
valence -0.20 -0.26 0.42 0.51 -0.01 0.39 0.02 0.10 -0.38 -0.36 0.14 1.00 0.19 -0.10 0.09
tempo -0.00 -0.05 0.02 0.28 0.02 0.26 -0.00 0.03 -0.24 -0.20 0.04 0.19 1.00 0.04 0.00
duration_ms -0.22 -0.28 -0.06 0.15 -0.01 0.17 -0.02 -0.08 -0.15 -0.14 0.15 -0.10 0.04 1.00 0.06
time_signature -0.02 -0.04 0.12 0.14 -0.02 0.13 0.00 -0.01 -0.12 -0.09 0.03 0.09 0.00 0.06 1.00
popularity year danceability energy key loudness mode speechiness acousticness instrumentalness liveness valence tempo duration_ms time_signature
1.00 0.75 0.50 0.25 0.00 −0.25 −0.50 −0.75

Piano: EDA


Popularity vs instrumentalness
Popularity
instrumentalness

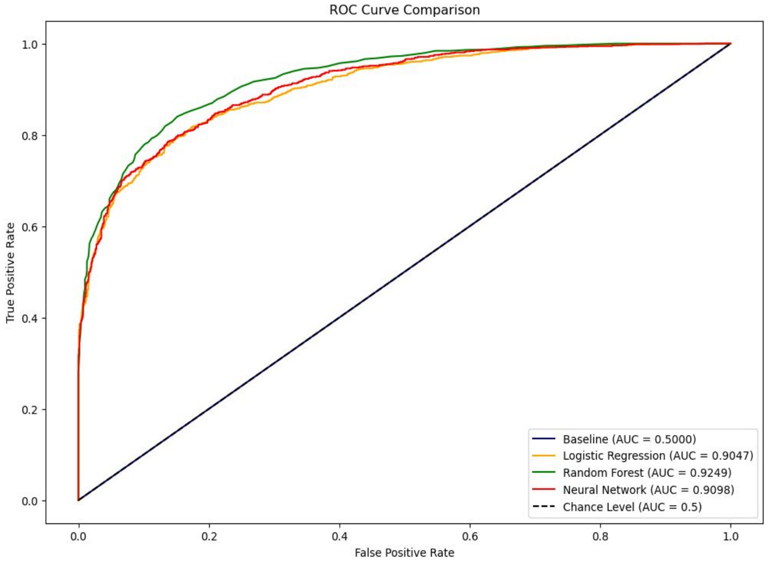

Popularity vs loudness
Popularity
loudness


Popularity vs speechiness
Popularity
speechiness


Popularity vs valence
Popularity
valence

Piano: Modeling

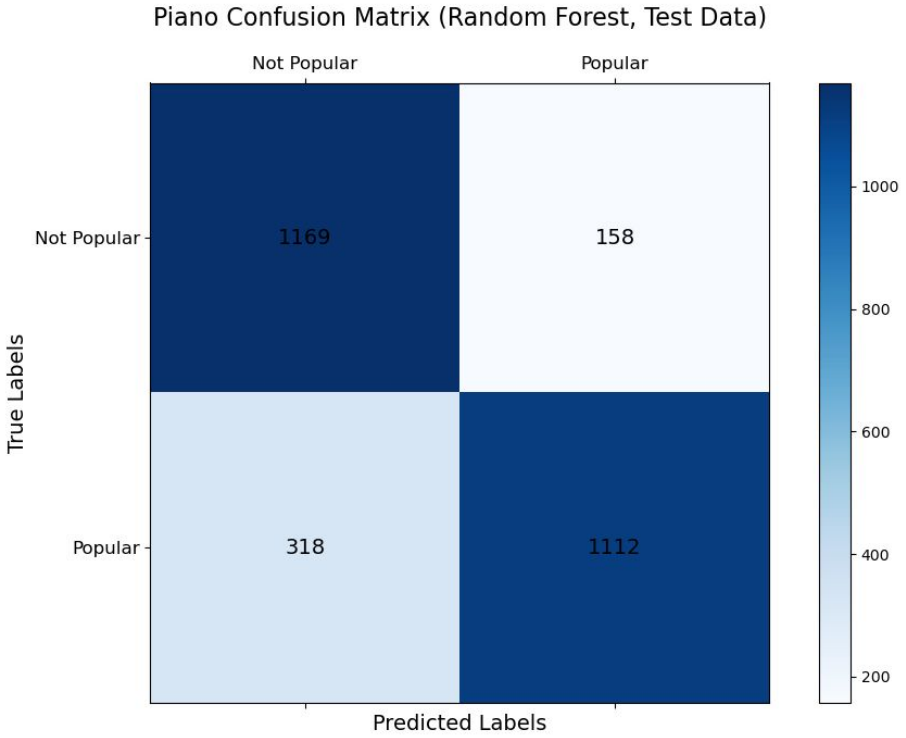
Model	Accuracy	Precision	AUC
Baseline	0.52	0.52	0.50
Logistic Regression	0.82	0.87	0.90
Random Forest Classifier	0.84	0.88	0.92
FF Neural Net	0.82	0.84	0.91



- Comparisons on validation set
- Baseline - majority class classifier
- Logistic Regression - L1 penalization
- Random Forest Classifier - no maximum depth, splits based on Gini impurity
- FF Neural Net - 2 hidden layers of sizes [64, 64], 0.2 dropout layer for each, and tuned hyperparameters. Tuned.

Piano: Results (Random Forest)

	Feature	Importance
0	instrumentalness	0.072622
1	loudness	0.063618
2	duration_ms	0.063003
3	valence	0.060654
4	energy	0.060022
5	acousticness	0.056234
6	liveness	0.047553
7	speechiness	0.045470
8	tempo	0.044045
9	danceability	0.043561
10	year_2020	0.043095

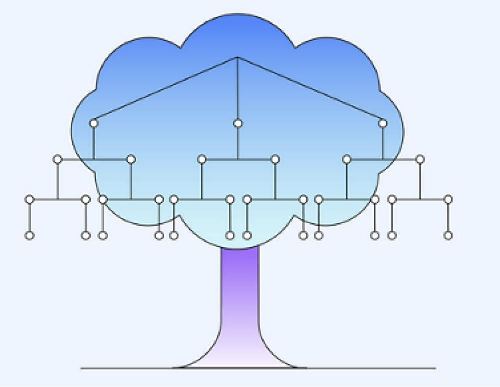


Final Results:

- Test Accuracy: 0.83
- Test Precision: 0.88
- Test AUC: 0.92

Conclusion

- Random Forest performed the best on all genres
    - Handles irrelevant features well



- Feature Importance varied across genres
    - Loudness varied significantly

Thank you for listening to this presentation today!

Contributions

Name	Presentation Contribution	Code Contribution
Robert Bull	Data Features, Data Processing, Death Metal	All modeling and EDA with respect to the Death Metal genre
Uriel Garcia	Motivation, Salsa	All modeling and EDA with respect to the Salsa genre
Ahsin Saleem	Introduction, Hip-Hop	All modeling and EDA with respect to the Hip-Hop genre
Ross Vrbanac	Conclusion, Piano	All modeling and EDA with respect to the Piano genre

Github

https://github.com/rfbull/mids-w207-final-project