MUSIC POPULARITY PREDICTOR

...

Sean Mccormick, Ryan Power, Sam Springer, and Jenna Dean

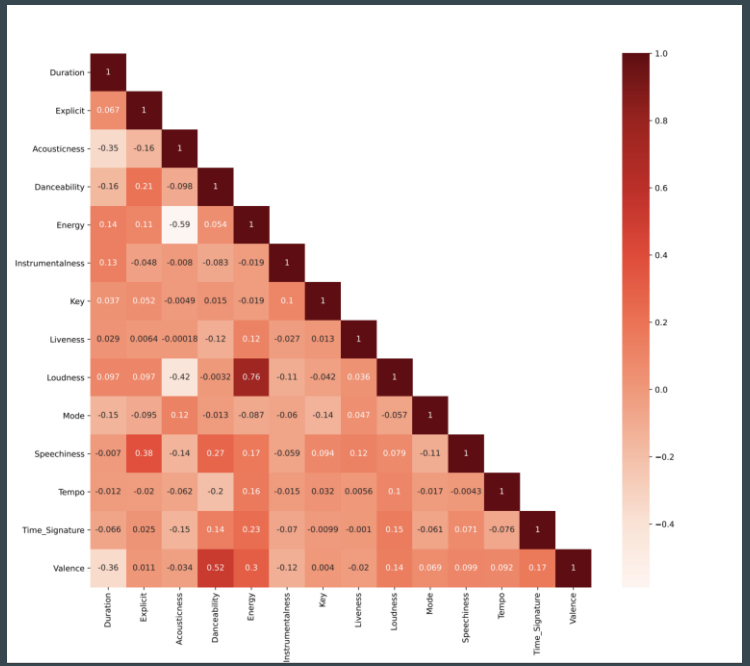
FINAL PRESENTATION

QUESTIONS SOUGHT TO ANSWER

- Discover the song characteristics most likely to contribute to a song’s or artist’s popularity.
- Predict song or artist popularity based on song or artist attributes.
- Predict song or artist popularity based on popular genres.

Data Preparation Work/Techniques Applied

- Data Cleaning
  - Remove outlier attributes
  - Cleansed dates
  - Filtered data
- Data Preprocessing
  - Dimensionality reduction
    - Find correlations between attributes
  - Normalization
- Data Integration
  - Call Spotify API using Python script
  - Using Python Dataframes>csv



TOOLS USED

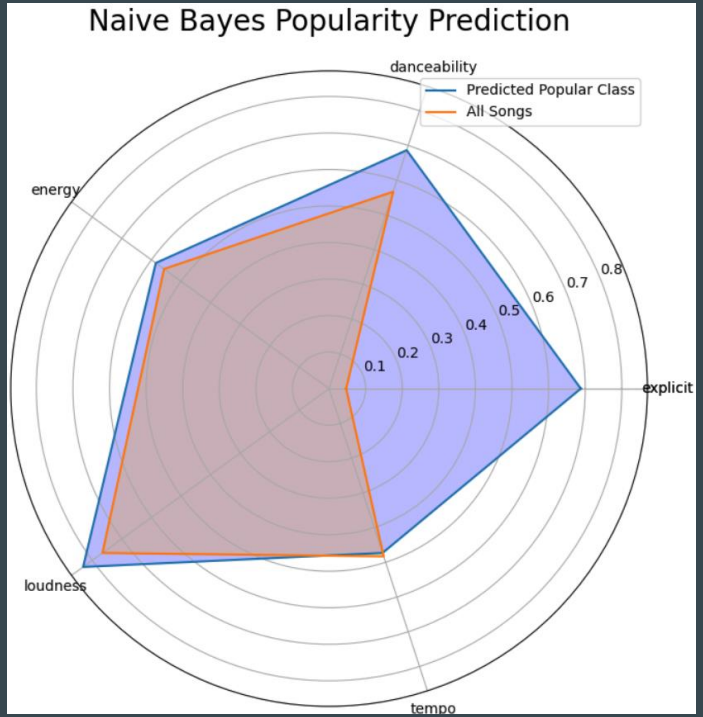
- Spotify API
- Python scripting/dataframe
- Sklearn
- Seaborn

TOOLS USED (cont.)

- Kaggle and Spotify API to supply the data:
  - Kaggle dataset supplied over 500,000 data objects, each of which represent an unique artist or song along with several attributes
  - Spotify API + python script to generate song/artist attributes that are placed in dataframes in order to analyze and manipulate the data
- NumPy, Pandas and Mlxtend - Apriori Python libraries

CLASSIFICATION/CLUSTERING APPLIED

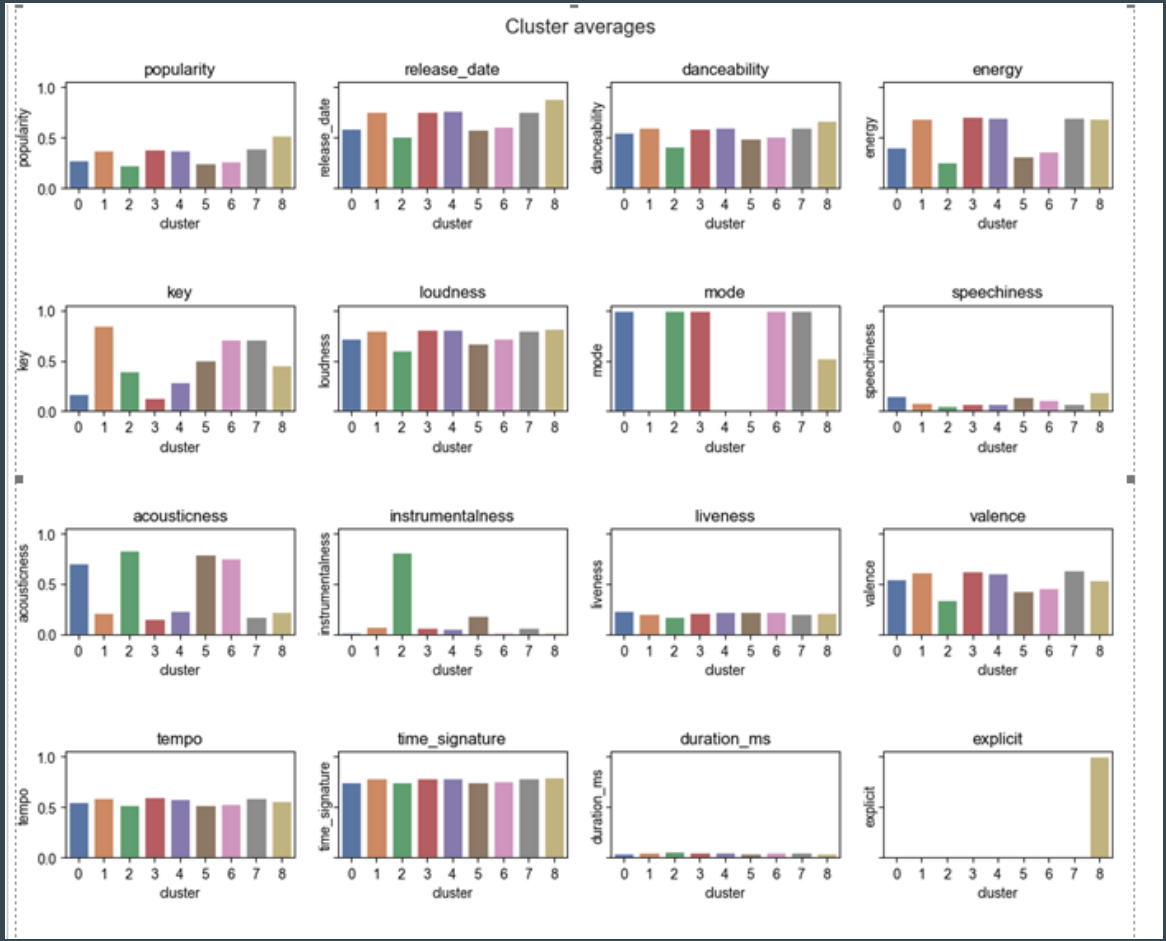
- K-Nearest Neighbors
- Random Forest
- Logistic Regression
- Naive Bayes
- K-Means Clustering
- Apriori



Apriori Results

Recommended Artist	Support
Taylor Swift	0.944444
Ariana Grande	0.777778
Lady Gaga	0.722222
Harry Styles	0.722222
Miley Cyrus	0.666667
Little Mix	0.611111
Sia	0.555556
Halsey	0.5
Katy Perry	0.5
Lorde	0.444444
Ava Max	0.444444

	support	itemsets
0	0.213836	(dance pop)
1	0.132075	(hip hop)
2	0.113208	(melodic rap)
3	0.465409	(pop)
4	0.176101	(pop dance)
5	0.220126	(pop rap)
6	0.182390	(post-teen pop)
7	0.245283	(rap)
8	0.176101	(trap)
9	0.207547	(dance pop, pop)
10	0.157233	(dance pop, pop dance)
11	0.138365	(dance pop, post-teen pop)
12	0.113208	(hip hop, rap)
13	0.176101	(pop, pop dance)
14	0.182390	(pop, post-teen pop)
15	0.113208	(post-teen pop, pop dance)
16	0.150943	(pop rap, rap)
17	0.113208	(pop rap, trap)
18	0.150943	(rap, trap)
19	0.157233	(dance pop, pop, pop dance)
20	0.138365	(dance pop, pop, post-teen pop)
21	0.113208	(dance pop, post-teen pop, pop dance)
22	0.113208	(pop, post-teen pop, pop dance)
23	0.100629	(pop rap, rap, trap)
24	0.113208	(dance pop, pop, post-teen pop, pop dance)


Cluster averages
popularity
release_date
danceability
energy
key
loudness
mode
speechiness
acousticness
instrumentalness
liveness
valence
tempo
time_signature
duration_ms
explicit
cluster
0.0
0.5
1.0
0 1 2 3 4 5 6 7 8

KNOWLEDGE GAINED

- Loudness and Energy are highly correlated with popularity of music;
- Popular music is associated with the following attributes: (1) energy; (2) loudness; (3) explicit; and (4) danceability;
- The genres of dance, pop, and rap appear the most frequently in the itemsets.
- Each song’s genre is important for predicting; however, we did not have access to it.

APPLICATION OF KNOWLEDGE

- With the use of the predictive music popularity models, the music industry can identify acoustic attributes that are correlated with music popularity.
  - Create songs that include acoustic attributes correlated with popularity, i.e., loudness and energy.
  - Sign artists that have characteristics associated with the acoustic attributes.
- The models can also be extended to create song recommendations based upon songs that contain similar attribute values.