CS185c Case Study

Understanding Artist's Music through Data

What makes an artist unique?

| Genre	Music Produced	Factor of Uniqueness
Artists usually have a certain style of music. Often it's genre specific and sometimes it's a blend of multiple genres	Music produced by an artist till date can be analysed to better understand the overall 'sound' of an artist.	The 'uniqueness' factor is what makes an artist stand out and get etched into history among the greats. This is a combination of multiple features of the artist's music.

Implementation

Resources

- Spotify API
  - Specifically used Spotipy API built for Python
- Spotify Playlists
- Sklearn
- Librosa

PLAYLIST

THIS Is: Kanye West

Yeezy taught me.

Created by: **Patryk Stanisz** • 121 songs, 8 hr 38 min

PLAY   FOLLOW   ⋯

Create a scraper that
allows reading the
data from playlists

Select features that
best represent the
artists' music

Perform Case Study
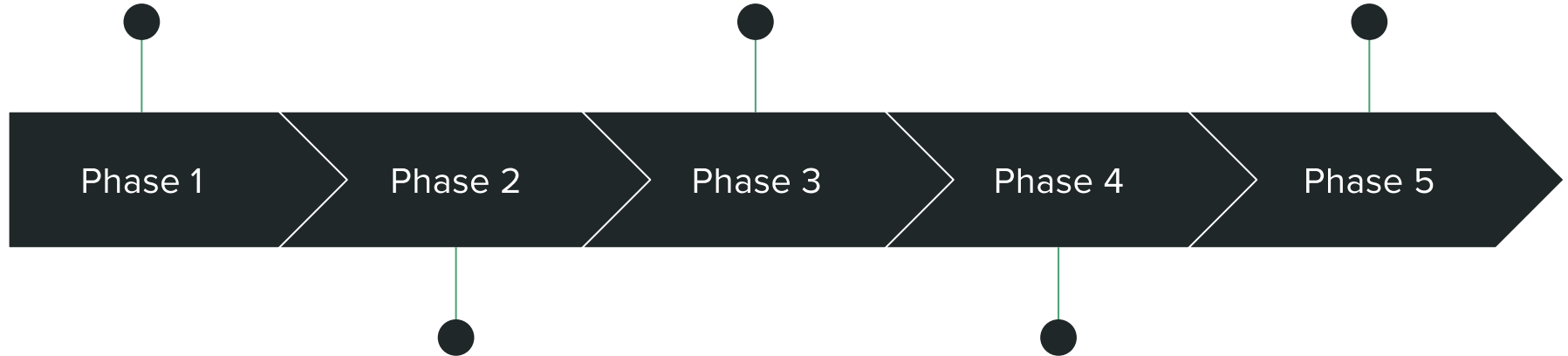
Phase 1	Phase 2	Phase 3	Phase 4	Phase 5

Extract data and create
data sets using the
scraper

Use One-Class SVM
for generating Contour
Plots for Entropy

Challenges deep-dive

Challenge 1	Challenge 2	Challenge 3

Data

What's the best way to extract data?

Features

What features best contribute to an artist's uniqueness?

Extracting Knowledge from Data

How can we represent and visualize the artist's uniqueness?

Challenge 1: Data Extraction

Spotify API

Spotify API allows developers to access their playlists and retrieve datasets for analysis

scraper.py

- Prompts user to create authorization token
- Gets audio features from given playlist
- Creates a CSV file with the scraped information

```python
import pandas as pd
import spotipy
import spotipy.util as util
token =
util.prompt_for_user_token("rayvo6",client_id='10ba56975e114851a531f5aa5e8ce549',client_secret='13cbdf2365d04eec89ccd109ba72fc99'
,redirect_uri='http://localhost/')

sp = spotipy.Spotify(auth=token)
sp.trace=False
playlist = sp.user_playlist("spotify", "spotify:user:spotify:playlist:37i9dQZF1DZ06evO3nMr04") #kaynewest


songs = playlist["tracks"]["items"]
ids = []
for i in range(len(songs)):
    ids.append(songs[i]["track"]["id"])

features = sp.audio_features(ids)
for i in range(len(songs)):
    features[i]["artists"] = (songs[i]["track"]["artists"][0]["name"])
    features[i]["name"] = (songs[i]["track"]["name"])
print (features[1])
cols_to_keep =
['artists','name','acousticness','danceability','energy','instrumentalness','liveness','loudness','mode','speechiness','tempo','v
alence','track_href']
df = pd.DataFrame(features)
df[cols_to_keep].to_csv('kanyealbums.csv',sep=';')
```

Challenge 2: Features Selection

We select the features that best describe the tracks for the artists and allow us to evaluate the artists' music better.
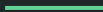
Feature Selection

- Acousticness - Measure of how acoustic a track is

- Danceability - Measure of how danceable a song is using a combination of tempo, rhythm stability, and beat strength

- Energy - Measure intensity and activity of a track

- Instrumentalness - Measures the absence of vocals in a track

- Liveness - Measure the presence of audience in the track

Feature Selection

- Loudness - Measures average loudness of a track

- Mode - Indicates modality of a track, 0 for minor and 1 for major

- Speechiness - Measures the presence of spoken words in a track

- Tempo - Measures tempo of a track in BPM

- Valence - Measure of musical positiveness in a track

Challenge 3: Technique used

One-Class SVM

Use contour plots plotting entropy of each track via One-Class SVM to visualize how much a song describes the artist's music style.
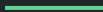
One-Class SVM

- We use the unsupervised learning algorithm, One-class SVM with non-linear kernel because it is mainly used for novelty detection
- Our goal is to identify whether certain songs are "artist songs" and this Machine Learning Model is one way to accomplish that
- Since this is a one-class SVM, all data is considered to be positive
  - For ex: assumes all songs come from Kanye West
- SVM tries to find a trend in the data

One-Class SVM

- Reduces the data into 2 dimensions using Principal Component Analysis
- The SVM Model is created and is fitted with the PCA reduced data
- SVM decides what songs are actually "artist songs"
- Cleans up the results of the SVM and forms a contour plot

```python
from sklearn.decomposition import PCA
pca = PCA(n_components=2)
pca.fit(x)
x_pca = pca.fit_transform(x) #PCA Dimension Reductio
```

```python
from sklearn import svm #Novelty Detection Using SVM
clf = svm.OneClassSVM(kernel="rbf", gamma=0.01) #Fit gamma to model size
clf.fit(x_pca)
```

```python
Z = clf.decision_function(np.c_[xx.ravel(), yy.ravel()]) #Decide which songs are classified
Z = Z.reshape(xx.shape)
plt.figure(1, figsize=(30, 20),)
plt.title("One Class SVM Model", fontsize= 30)
plt.contourf(xx, yy, Z, levels=np.linspace(Z.min(), 0, 7), cmap=plt.cm.Blues_r) #Contour Plo
a = plt.contour(xx, yy, Z, levels=[0], linewidths=2, colors='red')
plt.contourf(xx, yy, Z, levels=[0, Z.max()], colors='orange', alpha = 0.5) #Contour Plot Bas
```
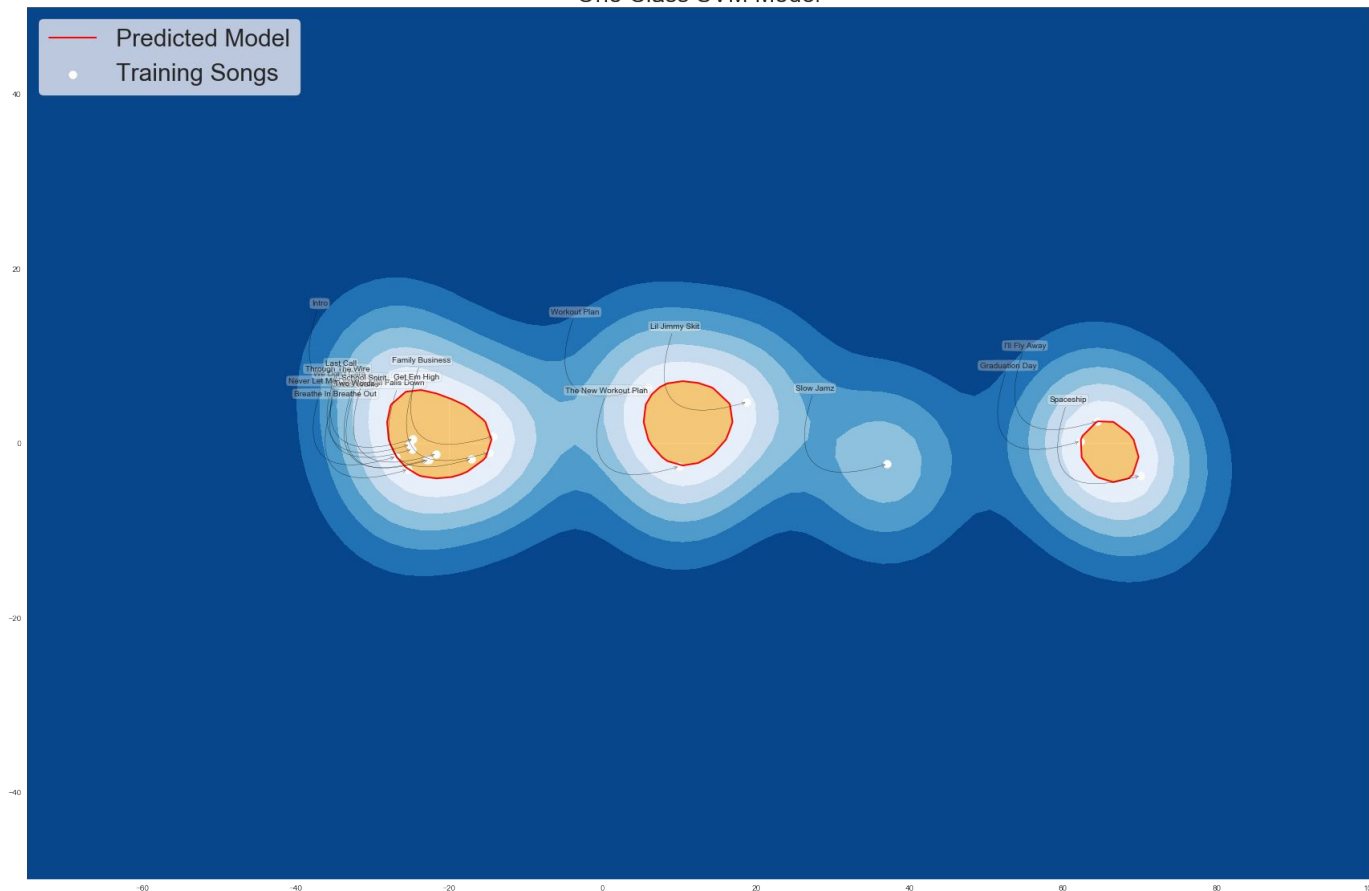
Case Study: Evolution of Kanye West

The greatest artist of all time has shown the ability to successfully change the sound of each album.

Can we map a model that represents Kanye West throughout his discography?

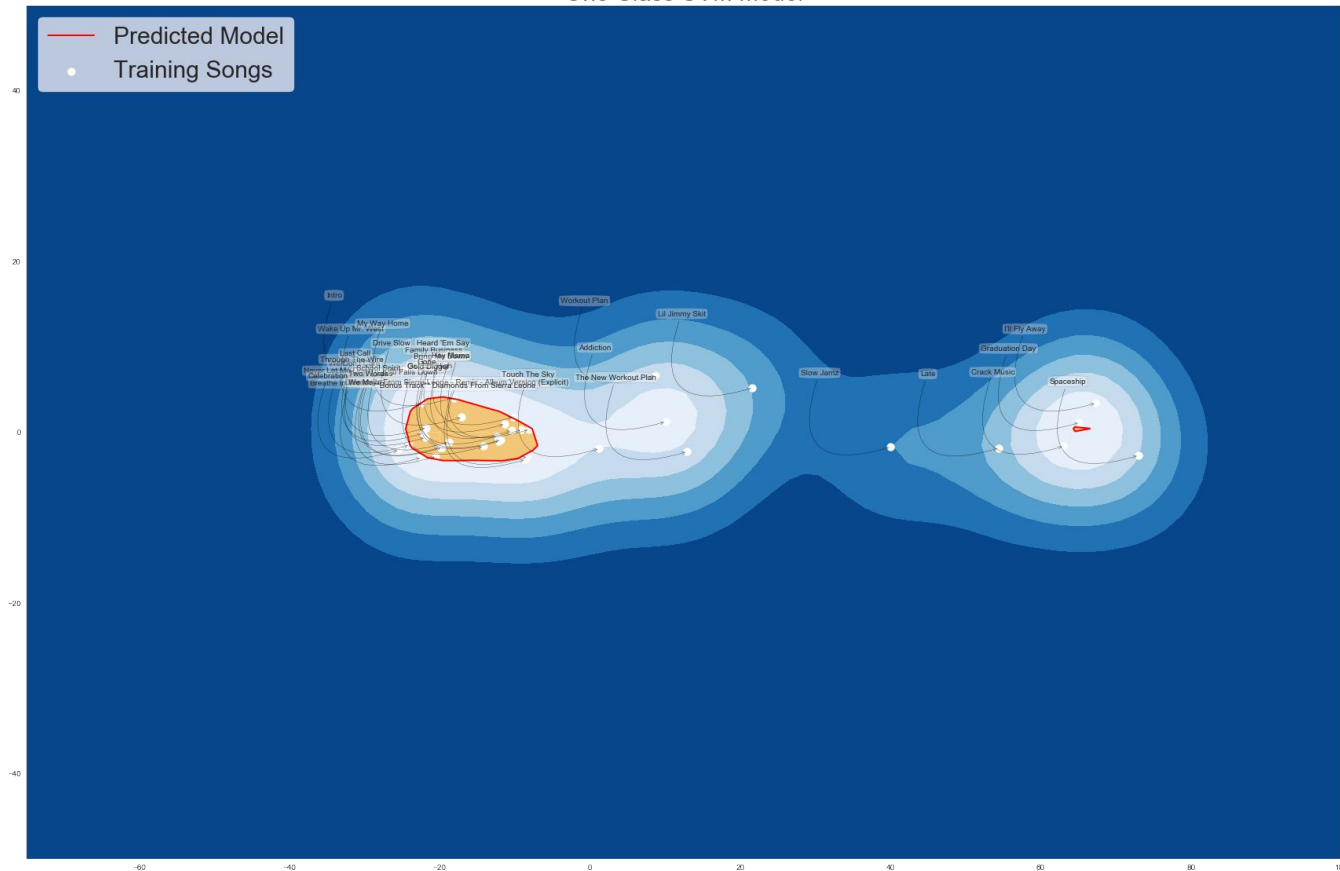# The College Dropout

One Class SVM Model

Observations

- CD features majority of "chipmunk soul" sampling
- Biggest cluster in model all have this sampling
- Outliers like Spaceship and I'll Fly Away have more traditional instrument backing

Late Registration

One Class SVM Model

Observation

- Late Registration doesn't feature as much of the "chipmunk soul" sampling done in CD
- However the sampling is mostly instrumental/vocal work which fits well with the songs from CD
- Outliers include songs like Crack Music which has more of a jungle beat

Crack Music (Explicit)

Kanye West
The Game

Late Registration

Graduation

One Class SVM Model
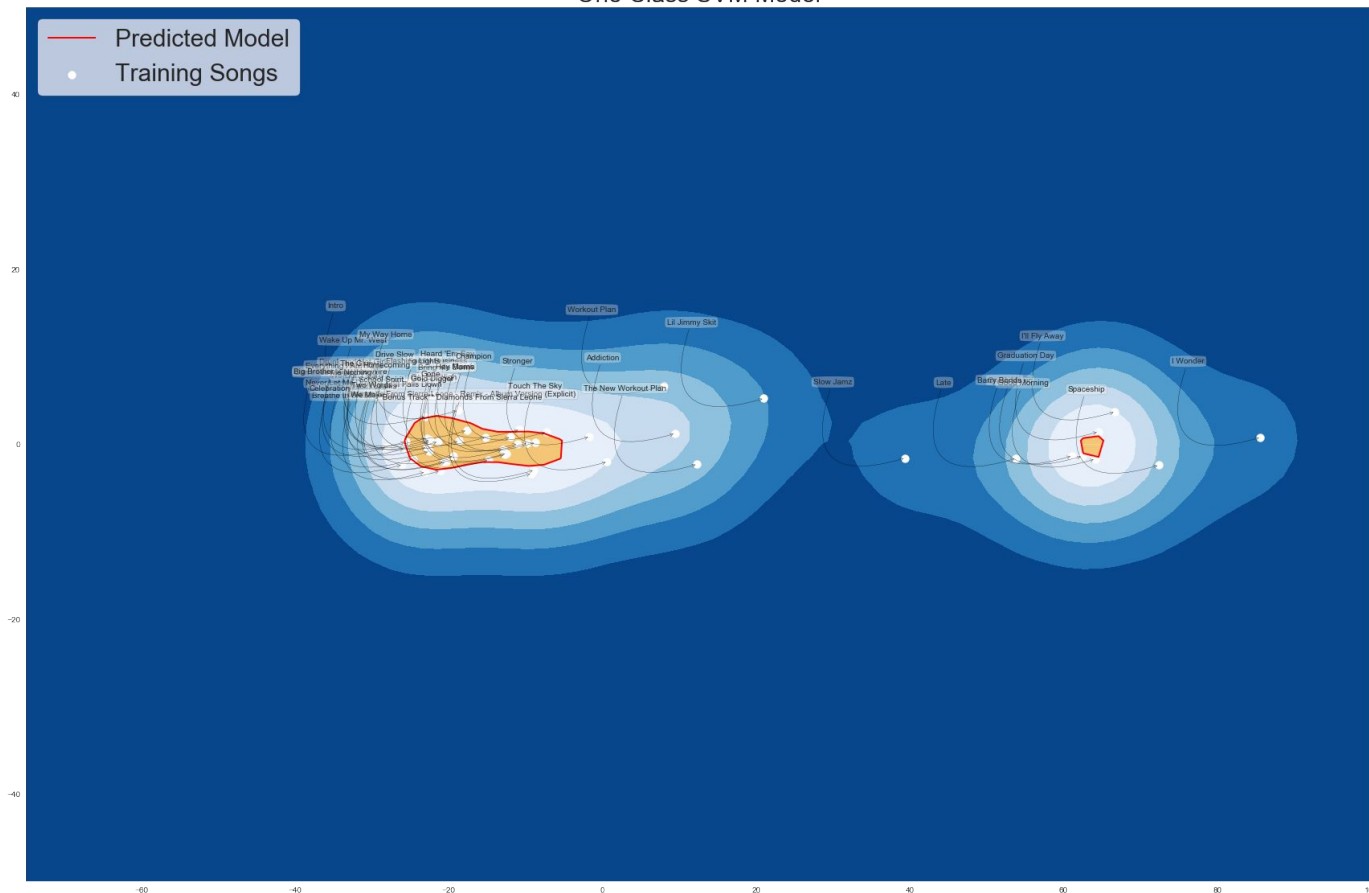
Observation

- Graduation was Kanye's album to play in stadiums, features lots of electronic sampling
- Fitted model is transitioning to less about the kind of sound in each song and more about style of beat and fitting into general qualities of hip-hop
- Outliers seem to all be songs that are slower in tempo and delivery of lyrics like Barry Bonds and Good Morning
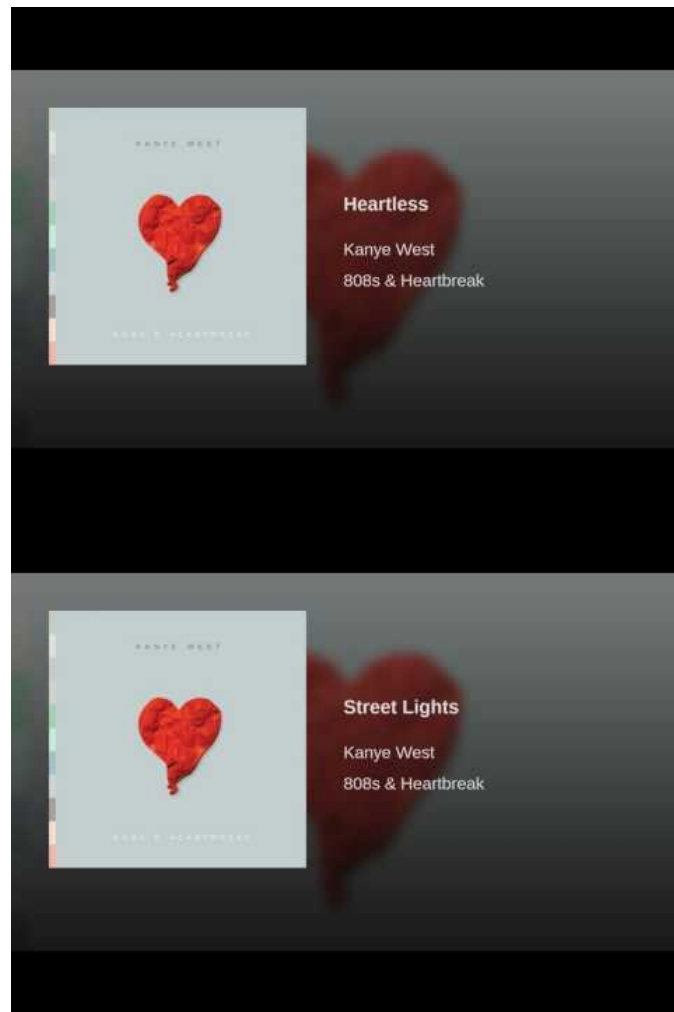

Barry Bonds

Kanye West
Lil Wayne

Graduation
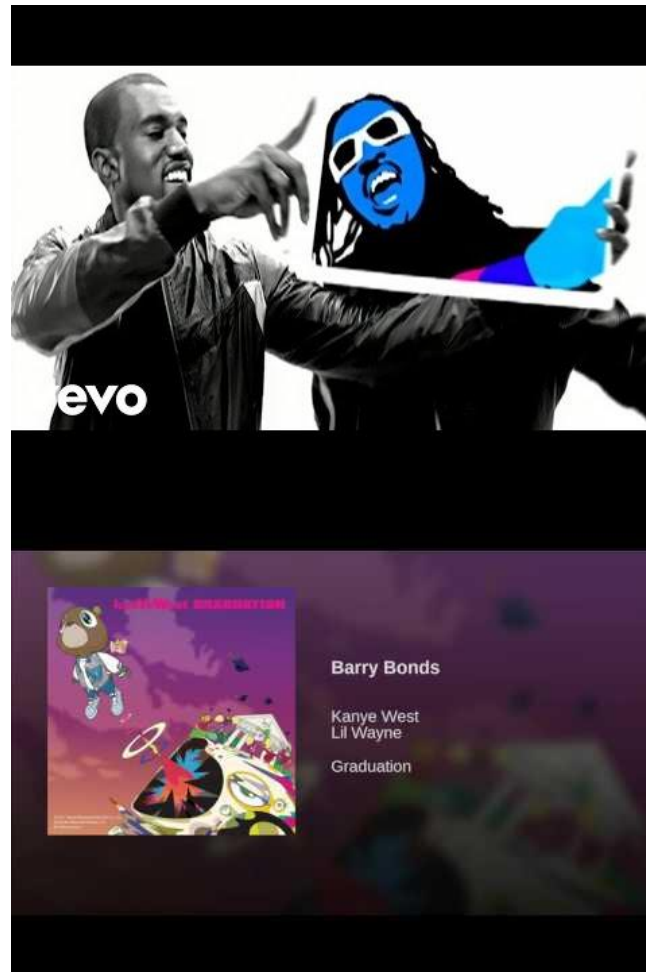
808s & Heartbreak

One Class SVM Model
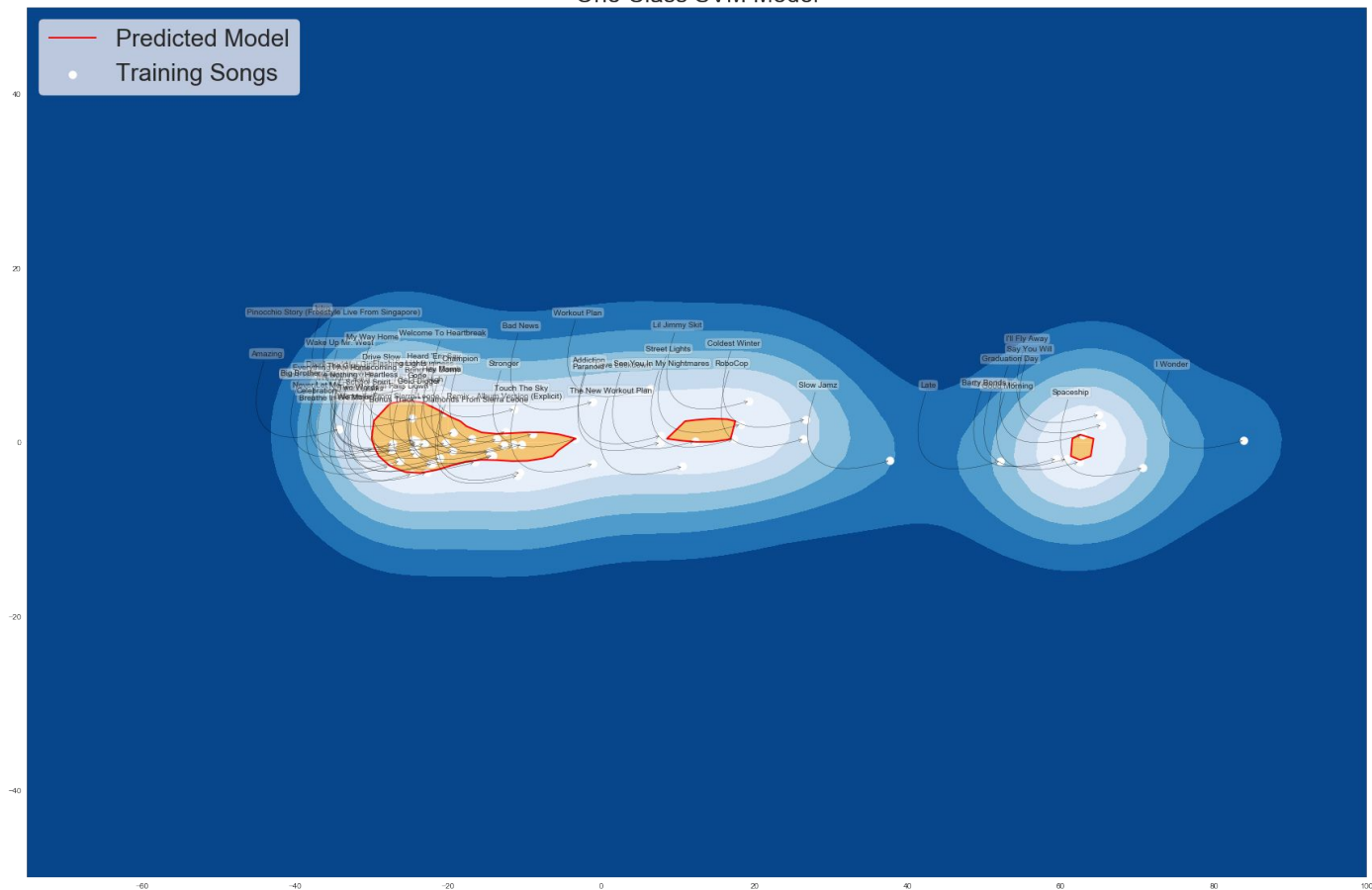
Observation

- 808s is a complete outlier from the last three albums
- Completely RnB with scattered electric sounds from famous Roland TR-808 and lots of Autotune
- Model creates a new modeled of area where majority of songs from 808s lie
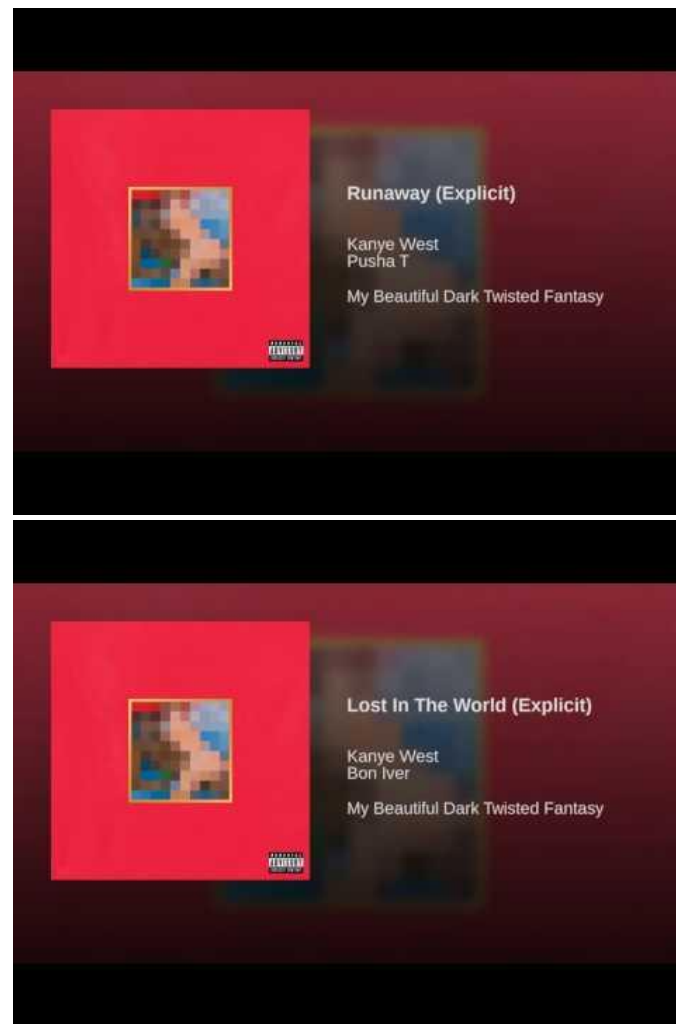- Songs like Heartless fit the original model

My Beautiful Dark Twisted Fantasy
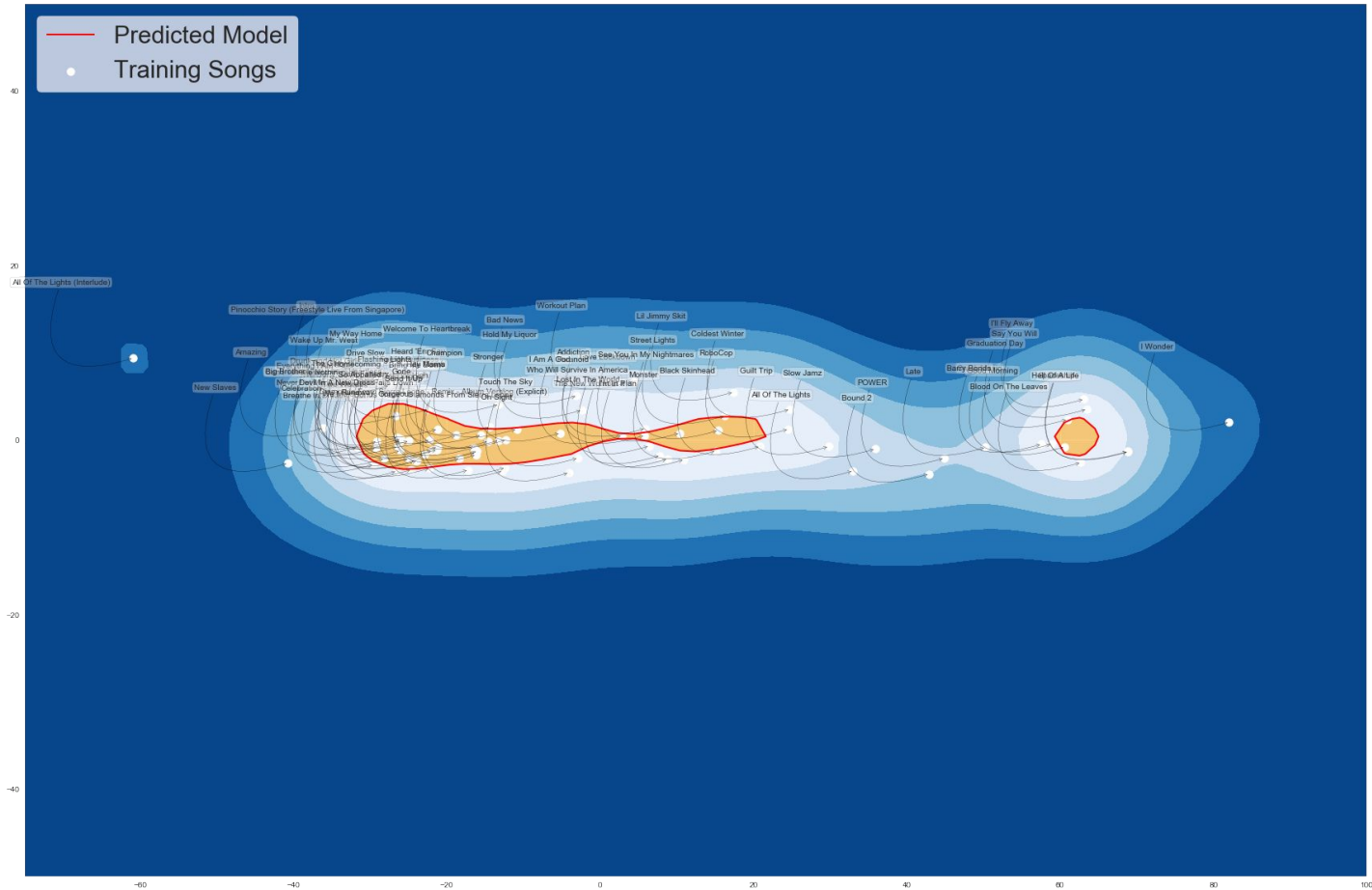

One Class SVM Model
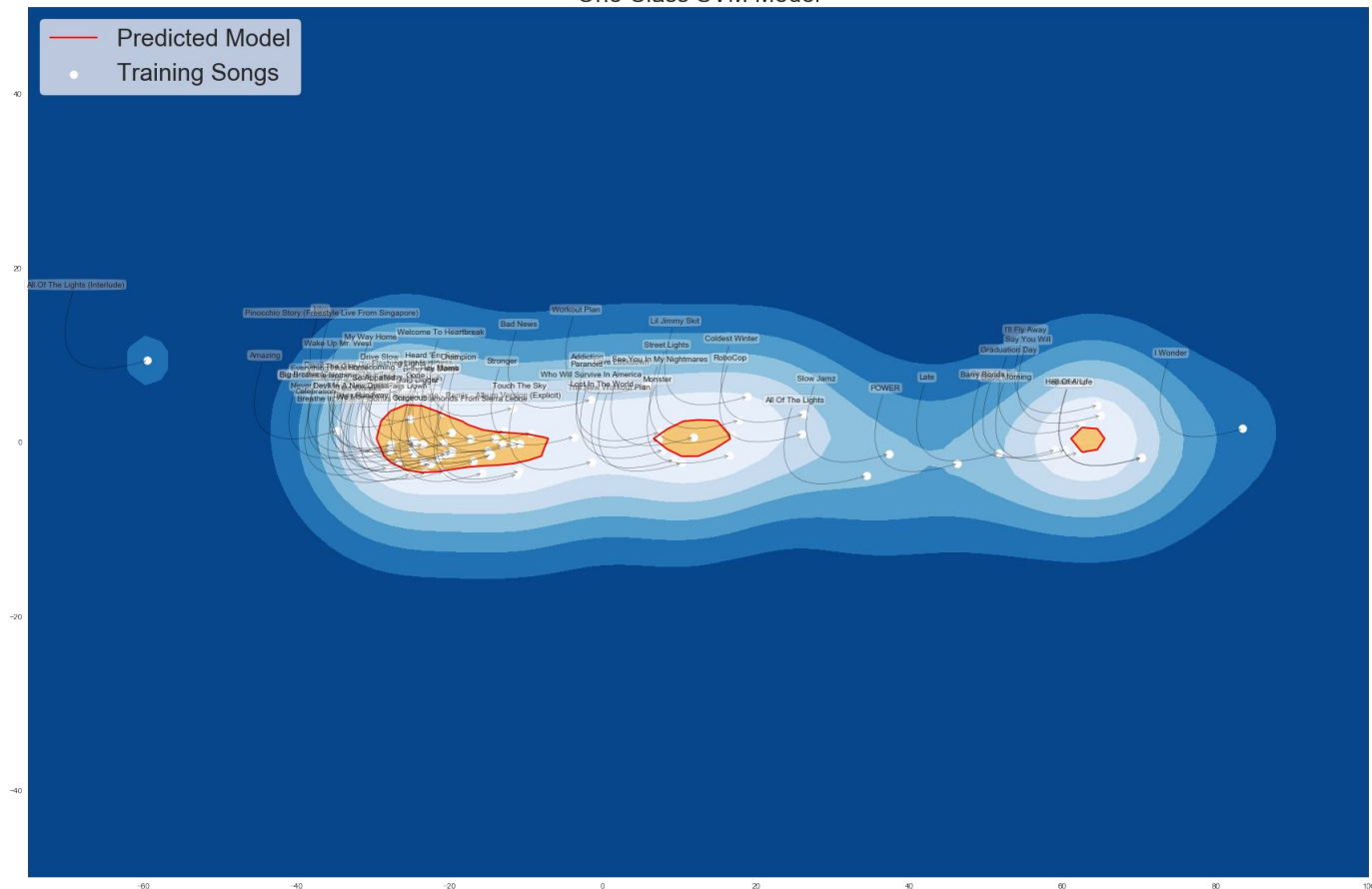
Observation

- MBDTF is a conglomerate of all the albums before it
- Features a large variety of techniques and sounds
- Majority of songs fit the model because of Kanye's return to hip-hop
- All of the Lights(Interlude) is a music-only piece done by an orchestra

Yeezus

One Class SVM Model

Observation

- Perhaps his most polarizing album, Yeezus is a stripped down industrial collection of minimalist and brash sounds
- Majority of songs don't fit the model but since the data is all positive, it reshapes it to fit with the sounds from 808s
- Hold My Liquor and Black Skinhead are closest to the fit model
- Bound 2 is surprisingly not despite its chipmunk sampling
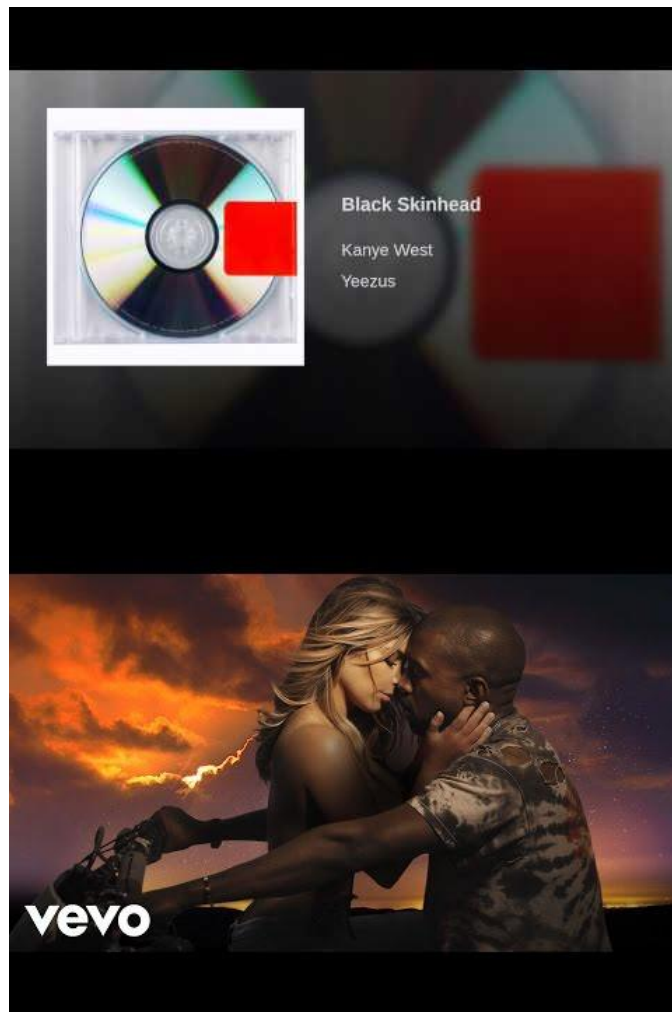
The Life of Pablo

One Class SVM Model

Predicted Model
Training Songs
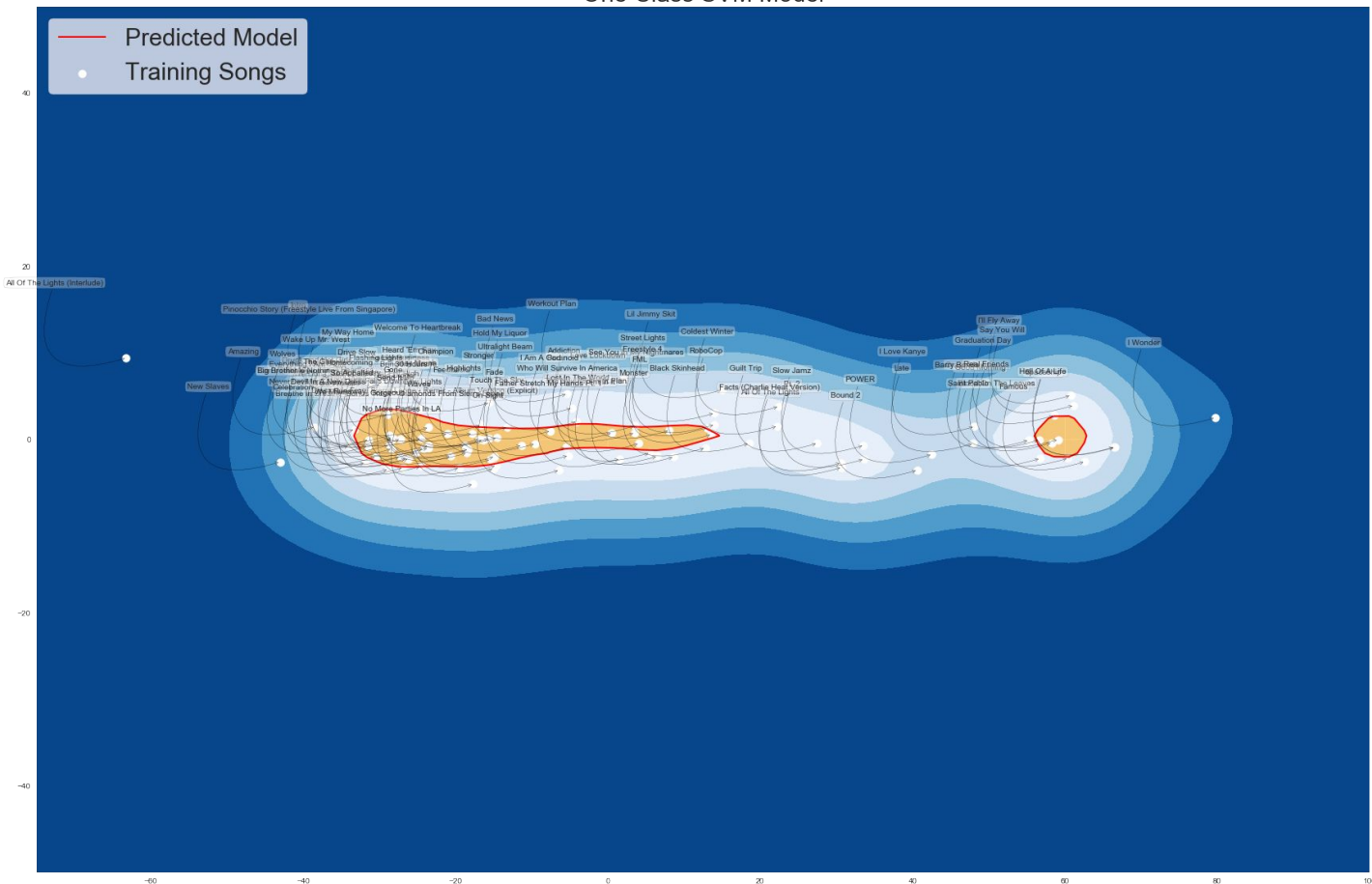
Observation

This is not album of the year. This is album of the life.
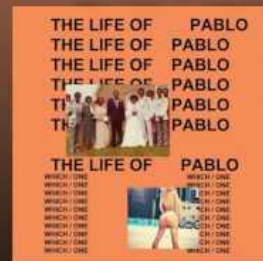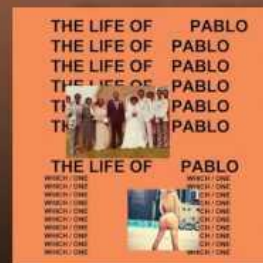
Case Study:
Does Greta Van Fleet sound too much like Led Zeppelin?

The young band has puzzled critics across the internet with their extremely 'Zeppelin-esque' music.

We use our technique to compare the two artists and check how closely their music resembles each other.
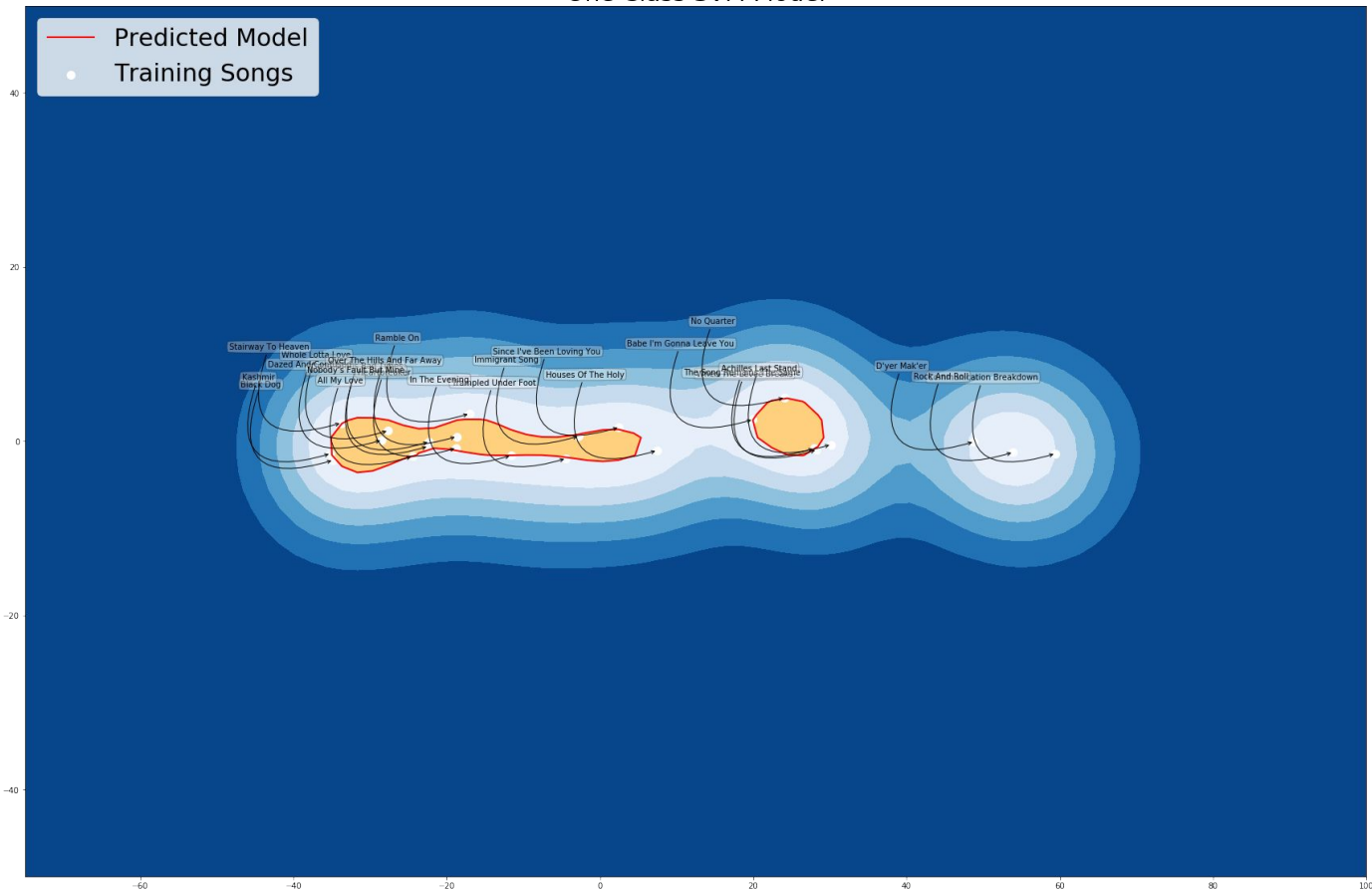
Led Zeppelin: Mothership

One of Led Zeppelin's best compiled albums that consist of their most original songs.

Led Zeppelin: Mothership

One Class SVM Model

Observation

- This album serves as a great metric for Led Zeppelin's music as it is a well curated compilation of their most famous and iconic work.
- This album contains the tracks that gave Led Zeppelin their memorable sound that was a confluence of Robert Plant's vocals, Jimmy Page's guitar, John Bonham's drums and John Paul Jones' bass. The factors that would be notable are the levels of instrumentalness, Acousticness, Danceability and Energy
- With tracks like 'Heartbreaker', 'Black Dog', 'Stairway to Heaven', 'Immigrant Song' and 'Whole Lotta Love'. Led Zeppelin had a well established sound that made them one of the most historic bands in the world.
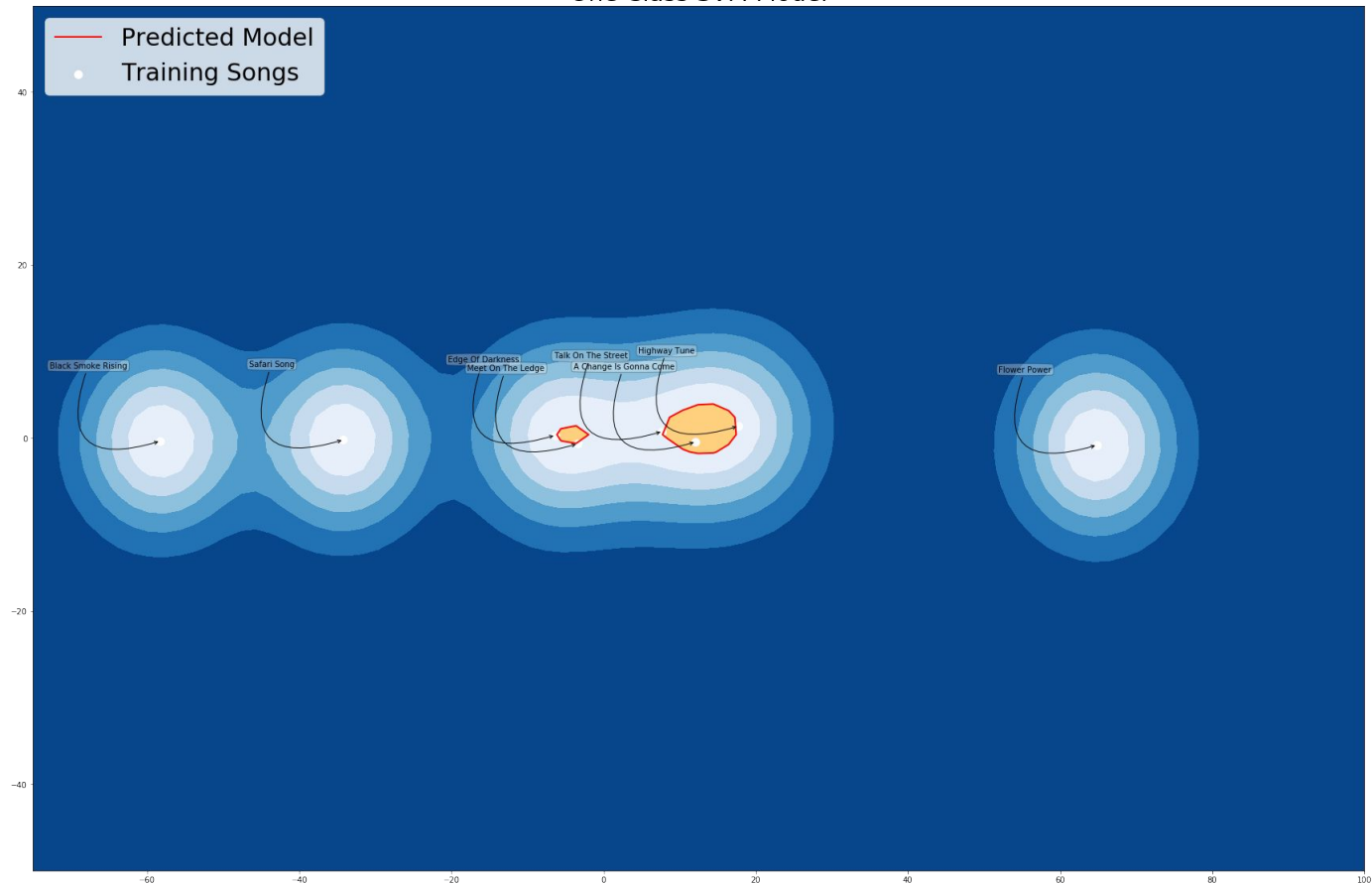
Greta Van Fleet

Winners of the Best New Artist Award 2017. The young band's music has been critiqued for sounding like a lost B-Side of one of Led Zeppelin's sessions.

Greta Van Fleet

One Class SVM Model

Observation

- The band's vocals are extremely similar to that of Robert Plant's and the overall sound definitely is reminiscent of Led Zeppelin's iconic sound. The guitar riffs sound very inspired by Jimmy Page's playing style. The reason the album sounds like a lost B-side of Led Zeppelin's work is because it has different components of Led Zeppelin's style in each track. It all comes together when heard as whole.
- The band shares the same genre as Led Zeppelin's and also has extremely similar mean values for their attributes in the dataset. The band does have a higher mean **energy** and a lower mean **instrumentalness** considering the fact that the band is quite young and does not share the complexity of Led Zeppelin's instrumental pieces such as Stairway to Heaven.
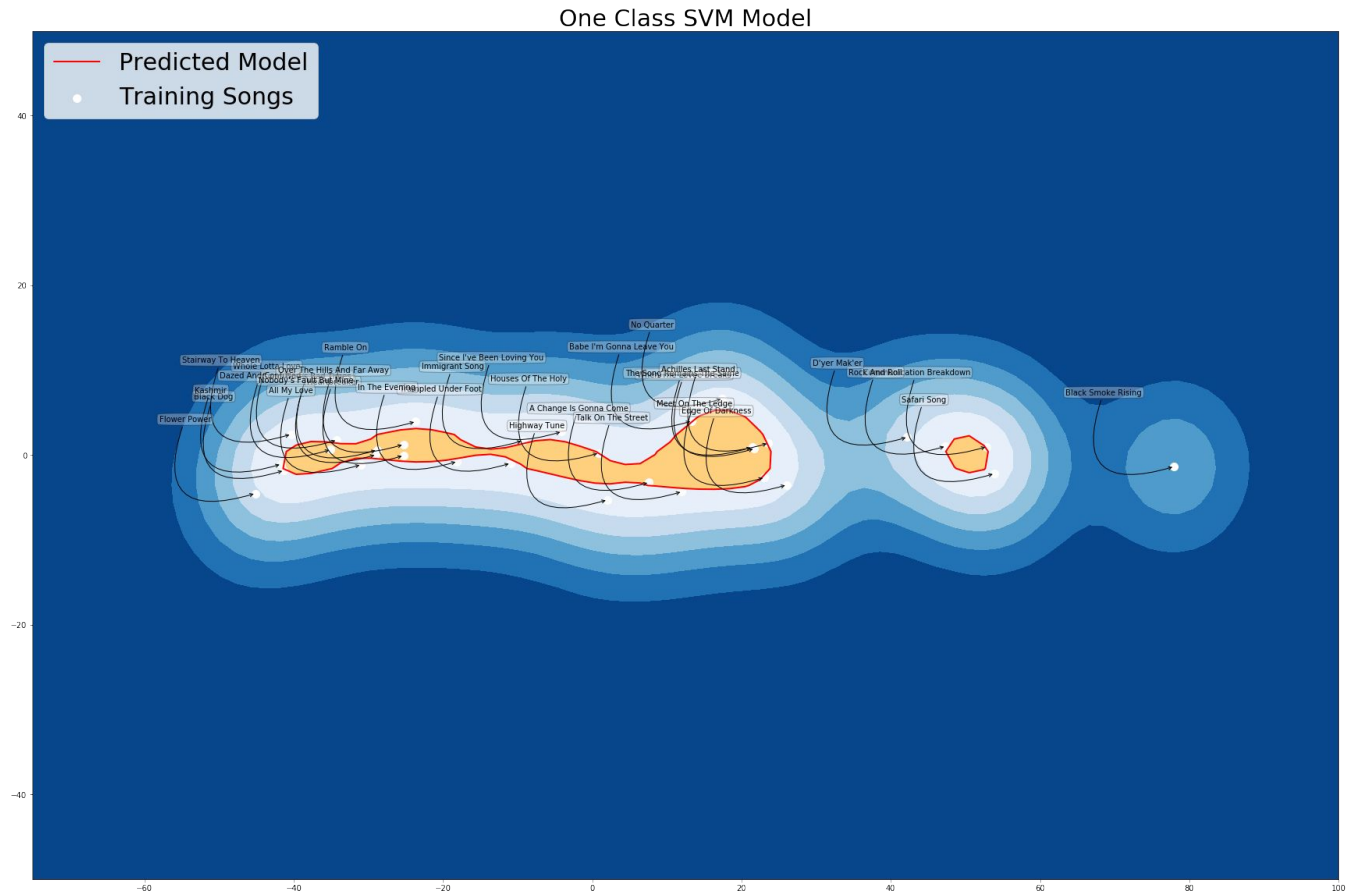
The Verdict.

Greta Van Fleet do speak about Led Zeppelin being a major influence on their music, but how closely do their tracks compare to each other.

The Verdict

One Class SVM Model

Observation

- All but one track are are very closely positioned to each Led Zeppelin's tracks. 'Black Smoke Rising', the band's first EP was released in 2017 that introduced the band to the world. Their album From the Fires was their first and only EP till date and consisted of their heavily Led Zeppelin influenced tracks.
- Looking at the plot, it is safe to say that their music is very closely positioned to Led Zeppelin's work. Their uniqueness among today's bands is surprisingly the result of a strong influence of the playing style of Led Zeppelin. 'Safari Song' positions itself very close to Zeppelin's 'Communication Breakdown' and 'Rock and Roll'.