Song Recommendations

Shohum Boker
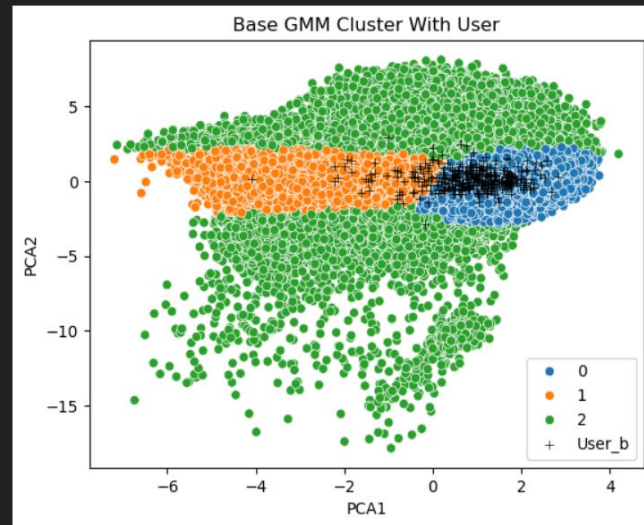
Business Question

- Given over 1 million songs and a user's input of like songs, can we recommend new songs that are not on this list?

- Can we also do it quickly, incase we need to train and produce new data as soon as possible?
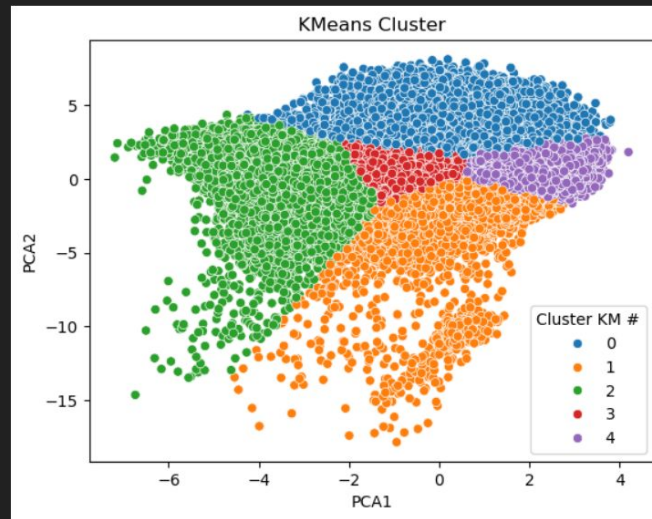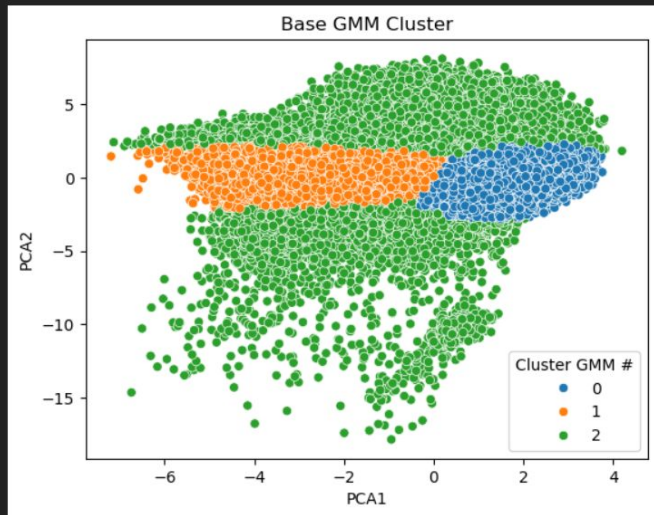
Approach: Overview

After get the metrics from Spotify:

1. First we use clustering as a means to divvy up the data.

2. Then we find the users liked songs and the clusters they belong to.

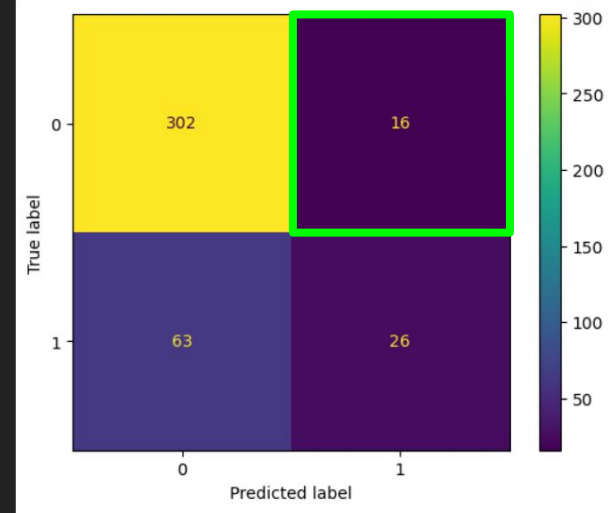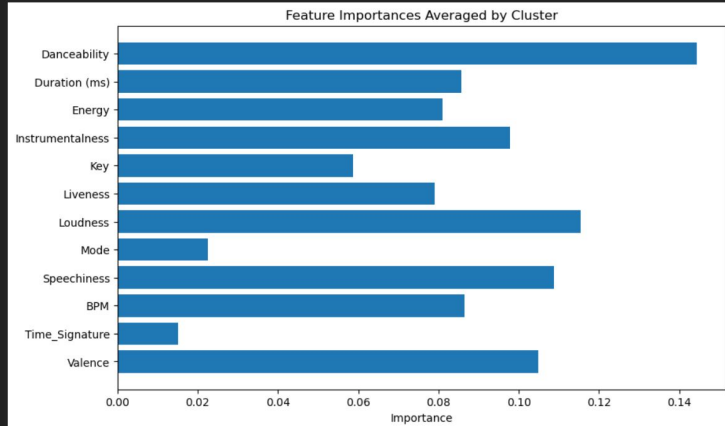3. Lastly we model data proportional to the user's presence within each cluster.

Base GMM Cluster With User
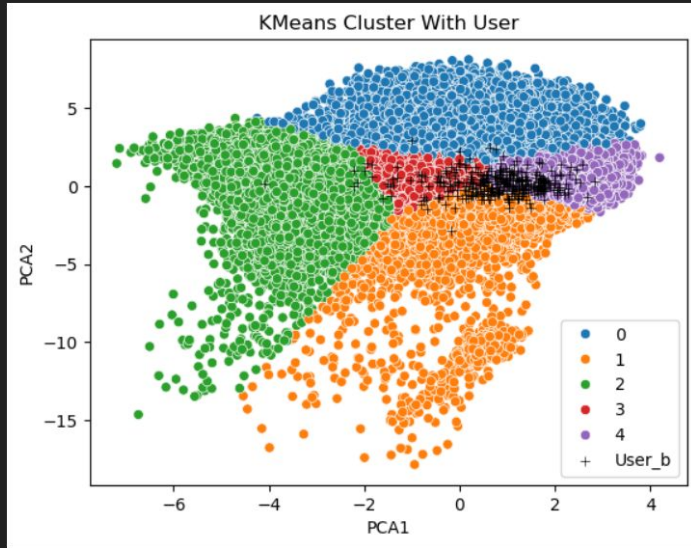
Approach: Clustering

Approach: Random Forest Classifier

Using a Random Forest has class imbalance, not good for smaller user input

Feature Importances Averaged by Cluster

On the bright, side we can see the Feature importance

Approach: K-neighbors

Using PCAs, K-neighbors is able to get over 300 recommendations.

Because we used PCAs, we cannot find any meaning behind the model's choices as related to the metrics from spotify

Trade Offs

K-neighbors

- Able to get a nice amount of recommendations

- ????How????

Random Forest Classifier

- Hard to get recommendation with a small user data set

- We can see how it learns

What's Next

- Testing via users and getting feedback

- Get a hardware to hyper parametrize

- Using two user's input of like songs, can we recommend a list that both users will regardless of size.

QnA