CS145
Group Project

Abstract

The project's aim is to analyze key features on popular songs to see what kind of elements affects song popularity. The Million Songs Dataset will serve as the beginning point for this project proposal's analysis of song popularity. Using the dataset, we plan to identify intriguing connections between popularity and various musical characteristics. The research will investigate musical elements like the song's key, the chosen lyrics, and other significant elements. The project will use the dataset in addition to the Spotify APIs and the MusixMatch API to obtain pertinent data about the songs.

To learn more about the popularity of songs and their related qualities, we can pick from a variety of machine learning models and data mining techniques to be applied. To group songs with similar features together, one strategy might be to utilize clustering techniques like K-means or hierarchical clustering.

Another strategy would be to create models that can forecast the popularity of songs based on their attributes using decision trees or random forests. This can be used to determine which characteristics are most crucial in determining a song's popularity and to forecast the success of upcoming or new songs.

This project will ultimately try to show what elements might influence song popularity through data science techniques.

Schedule

Week 7

Gather data and analyze key points and elements that we can utilize for our learning models. Decide which models we should use for such data.

Week 8

Implement model solutions to try to analyze the collected data and try out the implementation and fine tune model and parameters to get satisfying results

Week 9

Finish the study of the data and complete the report and presentation.

Week 10

Presentation on Wednesday 16:10

Team name:

Team Whatever

Members

Min Choi	UID: 904951987
Aritra Mullick	UID: 105317282
James Youn	UID: 505399332
Angus Wu	UID: 205344893