

Predictive Analysis of Song Ranking and Sales

Team Members:

1. Rohit Vasudevan
2. Giridhar Channe Gowda
3. Muralidhar Gopinath

What we have done so far:

- We have collected the data from the Billboard and Spotify data sets. The attributes collected from the dataset are:
 - Date of the chart release
 - Title of the song
 - Artist name
 - Weeks on top of the chart
 - Peak position of the song
 - Previous position of the song
- The obtained data from the dataset is stored in a CSV file.
- We have tried to visualize the data, by collecting the data over a particular duration (eg: a few weeks) and tried to classify and join the data with respect to its attributes. The data collected is stored into a CSV file.

Challenges Encountered:

- The data obtained after visualization is hard to collate, and we've encountered a few challenges due to the various attributes present in the dataset and we hope to resolve this by the next checkpoint.
- Have to filter the noise present in the Million Song Dataset and the Billboard and Spotify data set, due to the multiple

attributes present in the dataset. (Data preprocessing technique)

- Have to implement the data mining techniques, and calculation of sales by the next checkpoint.

Google Drive Link:

https://drive.google.com/open?id=1lQKzIX0N7joA8l7OrT8-mR9D_rddXnAi