Capstone Project – Spotify Recommendation Engine

Table of Contents

Problem Statement

Song recommendation system using Spotify data. Create a recommender that suggests the songs based on selection genre and year. I used the Spotify music data available in Kaggle.com.

Dataset Information:

The **data.csv** contains 170653 rows and 19 columns.
The **genre_data.csv** contains 2973 rows and 14 columns
The **Year_data.csv** contains 100 rows and 14 columns

Problem Objective

The Objective of this project is to create a recommender that suggests songs to the lis teners based on song rating, genre and released year selection.

Data Description

data.csv contains all the information about songs released from the year 1921 up to 2 020. The dataset has all the relevant information to create a recommender. The reco mmender lines up the songs that has highest rating based on the genre selected by th e listeners.

genre_data.csv contains information such as genre and other features of the genres which helps us creating a recommender.

year_data.csv dataset has features year of release of the songs and based on the ye ar of release year other features aligned. We will refer the year of release of the songs another features to create a machine learning model.

Data Pre-processing Steps and Inspiration

As we will study the data and try to prepare the data before creating a model for bet ter performance of the model and accuracy of the recommender. I studied the data set and I found the dataset is clean and ready to produce desired results as we creat e recommendation system.

Choosing the Algorithm for the Project

I have chosen K-means algorithm to create recommendation system and PCA to red uce dimension of the dataset.

Motivation and Reasons for Choosing the Algorithm

As we know that recommending songs to listeners, we have to use different data poin ts based on genre and year information and how unique the genres, songs and audio features. I have used all these information to specify the interest of different listeners. T he genres that are mostly liked will be recommended to new subscribers.

Future Possibilities of the Project

The more data we will have on songs that are listened by users from Spotify will help th e model to be more accurate in predicting songs

References

https://www.kaggle.com/