Proposal for Spotify Million song Data CSV

**Introduction:**

In today's digital age, music recommendation systems play a crucial role in providing personalized music recommendations to users based on their preferences and listening habits. With the abundance of music available on platforms like Spotify, it's essential to leverage data-driven approaches to enhance user experience and satisfaction. This proposal outlines the creation of a music recommendation system using the Spotify Millsong Data CSV, which contains key features such as Artist, Song, Link, and Text.

**Objective:**

The primary objective of this project is to develop a robust music recommendation system that accurately predicts songs that users may enjoy based on their listening history and preferences. By analyzing the Spotify Millsong Data CSV, we aim to extract valuable insights into user behavior and music preferences to enhance the recommendation algorithm's accuracy and effectiveness.

**Dataset Description**:

The Spotify Millsong Data CSV contains the following features:

Artist: The name of the artist(s) who performed the song.

Song: The title of the song.

Link: The URL link to the song on the Spotify platform.

Text: Additional textual information or metadata related to the song.

Methodology:

The development of the music recommendation system will involve the following steps:

Data Preprocessing: Cleaning and preprocessing the Spotify Millsong Data CSV to handle missing values, remove duplicates, and standardize the format of features.

Feature Engineering: Extracting relevant features from the dataset, such as artist information, song titles, and textual metadata, to create meaningful representations for modeling.

Model Development: Utilizing machine learning and/or deep learning techniques to train models that can effectively learn patterns and relationships between users and songs based on the available features.

Evaluation: Evaluating the performance of the recommendation system using appropriate metrics such as accuracy, precision, recall, and F1-score. Conducting cross-validation and testing on unseen data to ensure generalizability and robustness.

Deployment: Deploying the trained model as a web application or API that users can interact with to receive personalized music recommendations based on their preferences.

**Expected Outcome:**

The proposed music recommendation system is expected to provide users with accurate and relevant song recommendations based on their individual preferences and listening behavior. By leveraging the Spotify Millsong Data CSV and advanced machine learning techniques, we aim to enhance user satisfaction and engagement on the Spotify platform.

**Conclusion:**

In conclusion, the development of a music recommendation system using the Spotify Millsong Data CSV presents an exciting opportunity to leverage data-driven insights and advanced modeling techniques to enhance the music listening experience for users. By implementing this project, we can contribute to the advancement of personalized recommendation systems in the field of music streaming platforms.

**References:**

Spotify Millsong Data CSV

Research papers and articles on music recommendation systems

Online documentation and tutorials on machine learning and deep learning algorithms for recommendation systems.

**Dataset source:**

<https://www.kaggle.com/datasets/notshrirang/spotify-million-song-dataset?resource=download>

**Number of rows and columns:**

The dataset contains 4 columns and 57650 rows.

**Analysis questions**

1. How many rows and columns are there in the dataset?
2. What are the data types of each column?
3. Are there any missing values in the dataset?
4. How many unique artists are there in the dataset?
5. How many unique songs are there in the dataset?
6. What are the top 10 most frequent artists?
7. What are the top 10 most frequent songs?
8. What is the distribution of the length of the 'text' column?
9. Are there any outliers in the length of the 'text' column?
10. What are the average lengths of song titles?
11. What are the minimum lengths of song titles?
12. What are the maximum lengths of song titles?
13. Are there any correlations between the lengths of song titles and artist names?
14. How many songs does each artist have on average?