

# PROJECT DRAFT

**NAME:** MANSI PATIL

**CWID:** A20549858

**Title:** SONG/ARTIST RECOMMENDATION SYSTEM USING SPARK STREAMING.

## **PROBLEM STATEMENT / INTRODUCTION:**

My aim is to develop a Song/Artist Recommendation System using Spark Streaming in order to improve the user experience. Users will find new artists and songs that suit their musical interests due to the system's evaluation of their listening history and personal recommendations.

## **OVERVIEW OF THE PROJECT:**

The project leverages Apache Spark, PySpark, PySpark SQL, FindSpark, and MLlib to create an interactive, scalable, and effective recommendation system. By employing distributed stream processing, the system can handle real-time data ingestion, preprocessing, and recommendation generation.

## **PROJECT PLAN:**

### Importing Dataset:

To import and handle huge data sets effectively, using Spark.

### Loading Data:

Use Spark Streaming to continuously process and consume data in real time.

### Data Exploration:

Use PySpark SQL to explore and understand the dataset and gain knowledge about user preferences and current trends.

### Splitting Data:

Effectively train and assess the recommendation model by dividing the dataset into training and testing sets.

### Artist/Song Evaluation:

Evaluate the model's performance by assessing its ability to predict user preferences accurately.

### Model Construction:

Implement collaborative filtering algorithms from MLlib to construct a robust recommendation model.

### Recommendation Generation:

Develop a mechanism to generate personalized recommendations for users based on their listening history and preferences.

## **RESULTS:**

THE PROJECT AIMS TO ACHIEVE THE FOLLOWING GOALS:

- Improved user experience with specific suggestions for songs/artists.
- By doing the analysis and recommendation using hive and spark, each user is recommended with a unique playlist based on their interest.
- This also provides the information for the user like what/at are the most popular songs, who are all the artist, which artist is liked by the people through which he gets to know about the current trend.

## **REFERENCE:**

- <https://www.analyticsvidhya.com/blog/2021/06/spotify-recommendation-system-using-pyspark-and-kafka-streaming/>
- <https://induraj2020.medium.com/recommendation-system-using-pyspark-kafka-and-sparkstreaming3eb36b3c3df0>
- [https://ijaem.net/issue\\_dcp/Music%20Recommender%20System%20Based%20on%20Collaborative%20Filtering.pdf](https://ijaem.net/issue_dcp/Music%20Recommender%20System%20Based%20on%20Collaborative%20Filtering.pdf)
- <https://notebook.community/abhinandanramesh/Music-Recommendation-system/recommender>

