Predicting Spotify Track Popularity

Project Overview- This project aims to build a machine learning model that can predict how popular a song will be on Spotify. The dataset includes details about various Spotify tracks, such as song names, albums, artists, release dates, and popularity scores, along with social media engagement metrics like likes and shares. By analyzing these factors, the model helps identify what makes a song popular. This information can be useful for artists, music promoters, and fans to understand trends in music.

My Contribution- In this project, I focused on building and testing the main predictive model in Google Colab. My work included:

1. and Preparing the Data- Loaded the dataset and handled any missing information to make it ready for analysis.Used one-hot encoding to convert text data, like song and album names, into a format that the model could understand. Split the data into a training set (to teach the model) and a test set (to check its accuracy), with 80% of the data used for training and 20% for testing.
2. Building and Testing the Model- Chose a regression model that works well for predicting numbers, like popularity scores.Tested the model’s performance using cross-validation to make sure it was accurate and reliable. Measured the model’s accuracy using statistical tests, checking how well it predicted popularity compared to actual data. Saved the final version of the model as a `.joblib` file so it could be easily reused.
3. Using GitHub for Collaboration- Uploaded my code to GitHub, making it easy for the team to review and track changes.Documented the steps clearly in the code to help other team members understand my work.

**Proof of My Work**:- My work can be found in the `mini\_project\_final.ipynb` file and is documented in the GitHub repository at [GitHub Repository URL]. This includes all the steps for loading data, building the model, and testing it.

ITBIN-2110-0111

D A Udani