Project Proposal

* \*\*Course:\*\* AI102 – Artificial Intelligence and Python
* \*\*Project Title:\*\* Spotify Song Data Dashboard: Interactive Analysis & Prediction
* \*\*Group Members:\*\*

• Hassan Raza – 2024229

• Zara Gul – 2024679

# 1. Project Description

This project aims to explore and analyze a large Spotify dataset consisting of 10,000 songs. Using Python and Streamlit, we have developed an interactive web-based dashboard that offers visual insights into song performance metrics such as streams, chart rankings, and artist frequency. Additionally, we integrate basic machine learning to predict whether a song will reach the Top 10 on Spotify charts.

# 2. Objectives

• Perform exploratory data analysis (EDA) on the Spotify dataset.

• Build a user-friendly dashboard to visualize patterns and trends.

• Apply classification models (Logistic Regression, Random Forest) to predict Top 10 chart success.

• Visualize results using pie charts, scatter plots, confusion matrices, and genre-wise histograms.

• Enable filtering by artist and interactive data exploration.

# 3. Methodology

• Dataset: Cleaned and preprocessed Spotify dataset with key attributes like Total Streams, Peak Position, Days, and Top 10 (xTimes).

• Tools: Python, Pandas, Matplotlib, Scikit-learn, Streamlit.

• ML Models: Logistic Regression and Random Forest used to predict the binary classification outcome (Top 10 or not).

• Visualization: Scatter plots for feature relationships, confusion matrix for model performance, pie chart for Top 10 frequency, and bar chart for genre distribution.

# 4. Features of the Dashboard

• Artist Filter: Explore data for any specific artist interactively.

• Classification Report: Accuracy, confusion matrix, and performance report for model evaluation.

• Visual Plots:

- Scatter plot: Days vs Total Streams

- Pie chart: Frequency of songs in Top 10

• Interactive Sliders and Tables: For peak stream filtering and data preview.

# 5. Expected Outcomes

• A complete data analysis app that demonstrates proficiency in Python, data visualization, and basic machine learning.

• Meaningful insights into what factors contribute to a song's streaming success.

• A functional tool to assist in music trend analysis.

# 6. Deliverables

• Python Streamlit dashboard script.

• Processed Spotify dataset.

• Final presentation with insights and visuals.

• Full report detailing methodology and findings.

# 7. Dashboard sketch

