 A green circle with black lines in it

Description automatically generated

**Introduction:**

* Overview of the project's objective: Building a music recommendation system.
* Brief explanation of the methods used: Matrix Factorization, Cosine Similarity, and Collaborative Filtering.
* Purpose and significance of personalized music recommendations**.**

**About this dataset:**

Music is ubiquitous in today's world-almost everyone enjoys listening to music. With the rise of streaming platforms, the amount of music available has substantially increased. While users may seemingly benefit from this plethora of available music, at the same time, it has increasingly made it harder for users to explore new music and find songs they like. Personalized access to music libraries and music recommender systems aim to help users discover and retrieve music they like and enjoy.

This dataset is based on the subset of users in the #nowplaying dataset who publish their #nowplaying tweets via Spotify. In principle, the dataset holds users, their playlists and the tracks contained in these playlists.

**Data Acquisition and Preprocessing:**

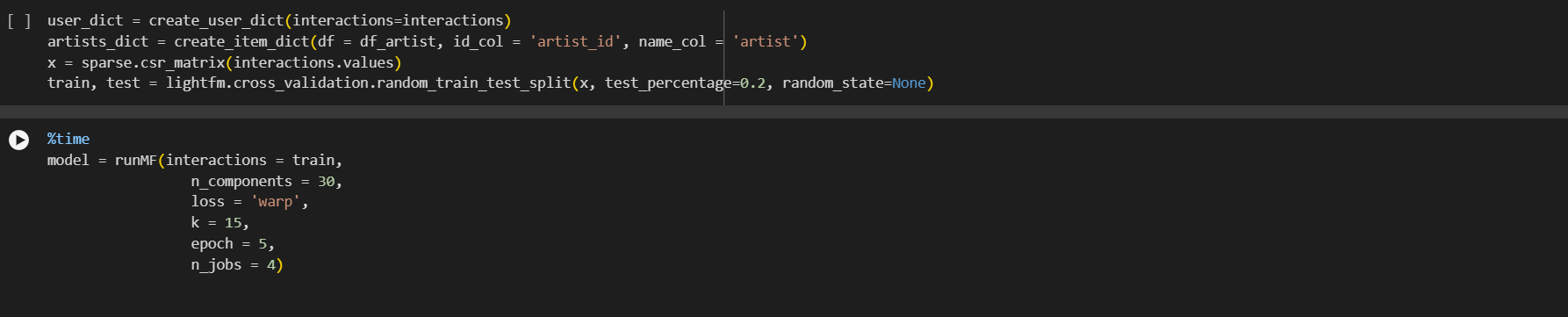
* Randomly selecting 50% of the rows from the Spotify dataset.
* Filtering and cleaning the dataset to ensure data quality.
* Aggregating data by artist and user, creating user-item interaction data.

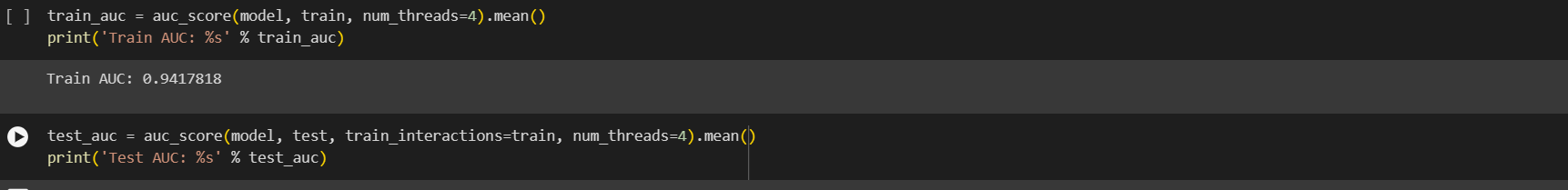
A screenshot of a computer

Description automatically generated

**Matrix Factorization Model:**

* Implementing the LightFM library for matrix factorization.
* Model training and evaluation using AUC and Precision at K metrics.





**Cosine Similarity for Recommendations:**

* Computing item-item similarity using cosine similarity.
* Function to recommend items based on cosine similarity.

**A screen shot of a computer program

Description automatically generated**

**Results and Analysis:**

* Train AUC: 0.9391825
* Test AUC: 0.9231451
* Train Precision: 0.3455297
* Test Precision: 0.1407859

A graph with a line

Description automatically generated**Visualizations:**

* Train vs Testing AUC
* **AUC Train and AUC Test Scores**

A blue and orange squares

Description automatically generated

* **Last the precision score**

A graph showing a green and red box

Description automatically generated

**Top Artists:**

A graph with blue and white bars

Description automatically generated

**Top tracks:**

A graph of a number of blue and white bars

Description automatically generated

**Team members:**

* **Name: Ahmad Talat ID: 221101084**
* **Name: Omar EmadEL-Din ID: 221101037**
* **Name: Maged Mohamed Beltagy ID: 221101048**