

MADE BY: OMAR
MOHAMED
ELGAMMAL
SUPERVISED BY:
ENG. A. ELTAHER &
ENG. M. TAMER

Spotify Recommendation system

1- Business Question

There is a dataset of analyzed songs (100 liked, 95 disliked).

What's required is to create a model that predicts if a song is liked or not.

2- Data Understanding

In this dataset there are 195 songs based on the dataset creator's taste of liked (mainly French rap) and disliked songs.

They contained audio features from spotify like: acousticness, danceability (how much the track is suitable for dancing), duration, energy, instrumentalness, key, liveness, loudness, mode, speechiness, tempo, time signature and valence.

And it was noticed that the creator likes songs that are loud and energetic.

A blue pen with a silver tip is resting on a document. The document features a bar chart with blue bars of varying heights. The pen is positioned diagonally across the frame, pointing towards the bottom right.

3- Data Preparation

In this step data will get prepared to be ready for using in the model.

The dataset didn't have any null or non-numerical values so no further actions were needed to set it up.

But it was needed for some values to be dropped of the data frame like: time signature and duration.

Also the data was split for training (80%) and testing (20%)

4- Model

For choosing a model for spotify recommendation system I tried to implement 2 methods from what I learned in this course and compare between them, K-Nearest Neighbor and Logistic Regression.

There was more accurate models like Random Forest. But I didn't want to use a model that I didn't take in this course.

5- Evaluation

It turned out that Logistic regression was more accurate than KNN in this system as shown in the notebook In the classification reports.

Classification Report: (Logistic Regression)

	precision	recall	f1-score
0	0.81	0.93	0.87
1	0.96	0.88	0.92
accuracy			0.90
macro avg	0.88	0.90	0.89
weighted avg	0.90	0.90	0.90

Classification Report: (K-Nearest Neighbor)

	precision	recall	f1-score
0	0.76	0.93	0.84
1	0.95	0.84	0.89
accuracy			0.87
macro avg	0.86	0.88	0.87
weighted avg	0.89	0.87	0.87