ML Finals Project Proposal

Team

Eliyahu Greenblatt - 316302934
Arad Ben Menashe - 207083353

Problem + Datasets

1. Given a song's features (like bpm, energy, loudness and so on...), what are the song's genres?
    - Here we will pick one of the following datasets: https://www.kaggle.com/datasets/amitanshjoshi/spotify-1million-tracks
    - https://www.kaggle.com/datasets/joebeachcapital/30000-spotify-songs
    - We will try to predict a song's genre based off of the features in the track data set.
    - We will analyse the data and the relationship within the data to delete some rows we think are not interesting.
2. Given a song file (.wav/mp3), what are the song's genres?
    - We have for that the following datasets:
    - https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification
    - https://www.kaggle.com/datasets/yash9439/emotify-emotion-classificaiton-in-songs (This dataset features of the songs are classification of emotions in songs but we will use it for our purpose)
    - Note we can also combine the datasets after converting one of the formats to the other one.

* We would like to hear your opinion on what you think is more interesting/better.

Classifiers

We will choose 4 of the following tools:

1. Random Forests.
2. Decision Trees.
3. KNN.
4. Adaboost
5. Neural Networks.