

Problem

The problem this capstone project aims at solving is predicting which songs a user will listen to. The problem is part of a Million Song Dataset Challenge on Kaggle.com.

For more info: Visit [Million Song Data Set Challenge](#)

Dataset

The dataset for the challenge can be directly obtained from the Kaggle.com. The data provided is relatively clean. The provided files contain:

- the official indexing of songs (note that indexing starts at 1);
- the official ordering of user IDs for your Kaggle submission;
- the visible half of the listening histories of the 110K evaluation users;
- the mapping from songs to tracks, more details below.

Approach

Collaborative filtering: a method of making automatic predictions (filtering) about the interests of a user by collecting preferences or taste information from many users (collaborating).

Deliverables

1. Code
2. Slide deck containing final report
3. Submission CSV (if applicable)