

matsubara
(PROJECT OUTLINE)

Alexander Stradnic



Final Year Project on the Use of
Machine Learning for Creating Playlists

University College Cork

1 Introduction

My project is on creating playlists using Recommender Systems, focusing on Maximum Likelihood Inverse Reinforcement Learning. The aim is to produce playlists that listeners would enjoy, based on playlists created and passed to the algorithm beforehand.

Inverse Reinforcement Learning(IRL): The playlists or paths are defined by "experts". IRL differs from standard Reinforcement Learning by not including the weights of the nodes (songs) in the paths, thus leaving it up to the AI to determine the way in which songs were added to each playlist and how they are related to each other in the given playlist, and there are multiple approaches to carrying out Inverse Reinforcement Learning.

2 Datasets and Processing

The data was acquired the Million Song Dataset, based on the MusicBrainz repository. Processing was done on the dataset running in SQLite. The processing itself is done using *Python*, with the packages *scikit-learn* and *pandas*.

3 Testing

The main MLIRL algorithm will be tested and compared against other machine learning algorithms, such as other types of reinforcement learning, as well as simpler algorithms based on pseudorandom techniques and similarity comparisons.

4 Ranking

The ranking of the algorithms is planned to be done by survey, using a website. The website front-end will pass user choices to the back-end, which will handle the processing and return the results. Users will be asked to choose genres or artists of their liking, then choose one of three options :

- Ask each algorithm to create a new playlist
- Choose one song, and ask each algorithm to create a playlist starting from that song

- Choose multiple songs, and ask each to extend the playlist

The results will then be returned to the user, who will then rank each playlist by their own personal preference. The rankings will then be stored and analysed.

5 Future Expansion

1. More analysis of the stored rankings would increase the dataset of desired playlists and paths for the improvement of the algorithms.
2. Graphing of results would show the performance of the algorithms in a more clear manner, as well as highlight potential trends.
3. Possible integration with music playing sites such as *Spotify*, *Apple Music*, *Youtube Music*, *Tidal*, *etc.* would allow a user to take a generated playlist and add it to their library.