Spotify feat. Logistic Regression - Popularity, Nothing Else Matters

Sebastian Hoffmann
Matrikelnummer 5954377
sebastian.hoffmann@student.uni-tuebingen.de

Yannick Streicher
Matrikelnummer 5331817
yannick.streicher@student.uni-tuebingen.de

Abstract

What is the musical taste of the world? With the recent rise and global pervasiveness of music streaming services, such as Spotify, Deezer, or Apple Music, answering this question has become tractable. For this, we plan to analyze a subset of 1.2 million songs scraped from Spotify. However, this dataset lacks crucial information about popularity. Thus, an important step of our work is to augment the dataset further by querying the official Spotify REST API for a randomly sampled subset of the data. Besides a birds-eye overview of the musical landscape, e.g. distribution of genres, we want to identify common musical properties shared by popular songs, and likewise, very unpopular songs, using logistic regression. Such properties can be, for instance, tempo, mode, or key.