DATA MINING IN MUSIC

Petr Kadlec · Ondrej Kužlík · Klára Petrovičová · Tomáš Repák · Samuel Šuľan

Introduction

Nowadays, music information retrieval (MIR) is a very popular and growing field with many real-world applications. Many people use apps for automatic songs recognition daily. But even more untrivial tasks such as genre recognition, personalized recommender systems or music generations are part of the research in MIR. The biggest challenges are high dimensional representation and feature extraction. In our project, we summarize some of the most interesting applications of MIR and examine several state-of-the-art machine learning approaches.

Emotion Recognition

Queries for production music (stock music) contain quite a lot of emotional terms. This creates a need for the music industry to categorize large-scale online music collections. This need is naturally helping music emotion recognition field to expand.
At the same time explanations from the prediction model could be used as parameters in the fields of recommender systems or music generation. The most common methods used are Convolutional Neural Networks, Long Short-Term Memory Recurrent Neural Network.

Composer Classification

Machine learning models which recognize a composer based on the musical scores or the audio files are extensively studied in the music information retrieval community. The major challenges in this field are learning from limited datasets and sufficient data representation. Many approaches have been taken both in the classification of musical scores - by Convolutional Neural Networks, as well as in recognizing the composer by audio files - by Decision Trees, SVM, logistic regression and Neural Networks.

Music source separation

Musical instrument identification is a crucial problem in music information retrieval. Without it, we would not be able to: search music by instrument, recognize musical genres, or make accurate music transcription. Music source separation algorithms have also been successfully used in predominant pitch tracking, accompaniment generation for Karaoke or singer identification. Modern techniques involve methods such as "deep clustering" or "deep convolutional neural networks".

Automatic music transcription

Automated creation of human readable representation of audio wave inputs is a challenging problem which has implications in many parts of the music industry, from visualization of music content through music search and recommendation to helping improve music creation. Current approaches are based on the usage of both long short-term memory networks and non-negative matrix factorization.

Music generation

Automatic music generation is a complicated task related to MIR. Nevertheless, there are already existing algorithms, which are able to produce plausible music. Since music is a temporal domain, models like Recurrent Neural Networks, LSTMs, or Gated Recurrent Units are usually used. While these models are widely employed, the problem of music generation has not been solved yet.