C137 Project: Music Popularity Prediction

Matthew Cook, Michael Loturco, and Abigail Stone

1 Introduction

Music streaming is a growing area with a wealth of data, access to this data gives us the opportunity to try to predict popularity and growth of songs. We assume that popularity changes over time, and thus trying to estimate 'popularity' without considering time dimension will produce poorer model performance than predicting popularity while considering time.

Existing work has attempted to classify songs into buckets of popularity, 'low', 'medium' and 'high' or as 'hit song' and 'not hit song' by looking at extensive sets of features. Features of a track can range from lyrical(e.g. valence, mood) to metadata (e.g. artist name, album) information to actual audio features extracted from the audio file itself (e.g. danciness, energy). HitMusicNet uses a compressed and encoded set of these features as an input to a Convolutional Neural Network and outputs a predicted popularity from among three classes.

We propose augmenting the encoded data from HitMusicNet by joining the encoded data from each track with the time series data of stream counts of those songs over time. Then we propose using a Recurrent Neural Network to take in time-labeled vectors including the encoded data and stream count and output a regression for stream counts for the next timestep. We hypothesize that by considering stream counts over time, predictions for popularity will be more accurate.

2 Related Work

Martin-Gutierrez et. al [4] introduce the SpotGenTrack Popularity Dataset (SPD) and a deep learning architecture called HitMusicNet. SPD combines metadata and audio data from Spotify and corresponding lyrical information from Genius and then extracts features from those sources. The extracted features are then compressed and encoded through a network MusicAENet. Finally a CNN MusicPopNet consumes the encoded and compressed feature vector to produce a classification of popularity. Similar previous models are noted to be binary classification, 'hit' vs 'not hit' whereas HitMusicNet is a ternary classifier separating 'low', 'medium' and 'high' classes. The overall architecture is shown in 1.

Li et al [3] present LSTM Rolling Prediction Algorithm (LSTM-RPA), a popularity prediction algorithm based on a Recurrent Neural Network. LSTM-RPA breaks the long-term prediction into several short-term trend predictions. This work is an important example of using an RNN architecture to predict song popularity.

Araujo et al [1] present a popularity prediction method using Support Vector Machines; this method predicts whether a song will appear on the Spotify Top 50 Global charts. This work has similar goals to our project and will serve as an important reference for evaluating our model's accuracy.

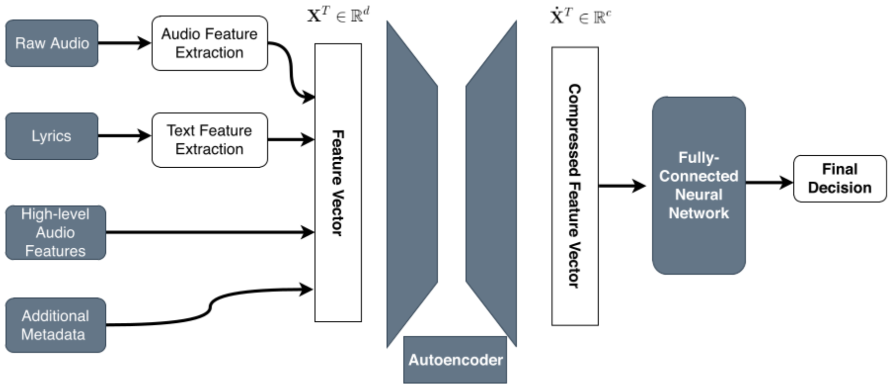

Figure 1: Model architecture presented by Martin-Gutierrez et al [4]. We intend to replace the fully-connected CNN with an RNN.

In their 2018 paper, Choi et al [2] compare audio pre-processing methods for deep neural networks. Their primary result demonstrates that many common preprocessing techniques are redundant; this conclusion will help us inform our pre-processing decisions for our project.

References

[1] Carlos Soares Araujo, Marco Cristo, and Rafael Giusti. Predicting music popularity on streaming platforms. In *Anais do Simpósio Brasileiro de Computação Musical (SBCM 2019)*, pages 141–148. Sociedade Brasileira de Computação - SBC.

[2] Keunwoo Choi, Gyorgy Fazekas, Mark Sandler, and Kyunghyun Cho. A comparison of audio signal preprocessing methods for deep neural networks on music tagging. In *2018 26th European Signal Processing Conference (EUSIPCO)*, pages 1870–1874. IEEE.

[3] Kun Li, Meng Li, Yanling Li, and Min Lin. LSTM-RPA: A simple but effective long sequence prediction algorithm for music popularity prediction.

[4] David Martin-Gutierrez, Gustavo Hernandez Penaloza, Alberto Belmonte-Hernandez, and Federico Alvarez Garcia. A multimodal end-to-end deep learning architecture for music popularity prediction. 8:39361–39374.