Proposal for Final Project in Machine Learning: Music Genre Classification using Convolutional Neural Networks

Daniel Bick (s3145697), Stella Tsoutsouri (s3904210),
Shubham Jinde(s3993914), Federico Ferlito (s2936860)

December 10, 2019

1 Proposal

This research will be concerned with *Music Genre Classification*. The training data is provided by the GTZAN dataset, which consists of 100 song excerpts for each of the ten represented classes of 30 seconds each in *.wav* format. As a pre-processing step, spectograms will be created for each of the contained songs. Afterwards, a spectogram is partitioned into equally sized chunks from left to right, i.e. into chunks representing the evolution of the spectogram representation of a song over time. Sliding a convolutional filter over the partitions from left to right will result in a stream of incoming features to a Recurrent Neural Network (RNN) performing the classification task. Inspiration comes from [1] and [2], as well as multiple unpublished sources. [1] [2] [3] Working in the music domain and on time series data will be new to all group members. If time allows, further techniques for classification will be explored.

References

[1] Fady Medhat, David Chesmore, and John Robinson. Music genre classification using masked conditional neural networks. *arXiv preprint arXiv:1802.06432*, 2018.

[2] Zain Nasrullah and Yue Zhao. Music artist classification with convolutional recurrent neural networks. *arXiv preprint arXiv:1901.04555*, 2019.

[1] http://cs231n.stanford.edu/reports/2017/pdfs/22.pdf
[2] http://cs229.stanford.edu/proj2016/report/BurlinCremeLenain-MusicGenreClassification-report.pdf
[3] http://cs229.stanford.edu/proj2018/report/21.pdf