DS 340: Music Recommendation Using Unsupervised Deep Learning

Kyle Salitrik | Tomoki Takasawa

January 29, 2018

Project Description & Purpose

For our project we intend to create a way to find similarities between songs in a user's music library without using provided labels. This unsupervised approach will hopefully be able to provide the likelihood that a user will enjoy a song based solely on the music itself. The purpose of being genre, artist, and popularity agnostic in decision making for the recommendations is to introduce the users to new music that they may not normally enjoy.

The intended application would be to scan a user's music library as training data and then scan a wide library of songs to find ones with features similar to the user's current library with a tunable width of variability. These songs will be recommended to the user to listen to. Ideally the user may be able to respond whether or not they actually enjoyed the song in order to adjust future recommendations.

Who Will Benefit

Artists, users, and music providers would be able to benefit from successful results of our application. Not taking into account the popularity of a song or artist would allow new or unknown artists to be discovered. As genres are not considered – only the music itself – users will be able to benefit by expanding their music tastes.

Frameworks to be Used

We plan to use following open-source libraries and data sets will be used to accomplish our task

- LibROSA
- Keras
- TensorFlow
- LabROSA One-Million Song Data Set

Milestones

Week 6

We plan to accomplish followings by the end of Week 6

- Collecting data

- Designing the architecture of the project

- Extracting some of the features
  1. frequency
     - frequency band
     - frequency power
  2. repetition
     - rhythm
     - lyrics

Week 9

- Extracting all the features
    1. tempo
    2. musical keys

- Potentially applying existing Cocktail Party Effect networks
    1. masking noises (other instruments)

INITIAL RESOURCES

REFERENCES

Ernst Dietrich Munz. Psychotherapie in der Psychiatrie. *Nervenheilkunde*, 36(10):800–805, 2017. ISSN 07221541. doi: 10.1007/s13398-014-0173-7.2.

Juhan Nam, Jorge Herrera, and Kyogu Lee. A Deep Bag-of-Features Model for Music Auto-Tagging. pages 1–9, 2015. URL http://arxiv.org/abs/1508.04999.

Jordi Pons, Olga Slizovskaia, Rong Gong, Emilia Gómez, and Xavier Serra. Timbre Analysis of Music Audio Signals with Convolutional Neural Networks. 2017. URL http://arxiv.org/abs/1703.06697.

Andrew J R Simpson, Gerard Roma, and Mark D. Plumbley. Deep karaoke: Extracting vocals from musical mixtures using a convolutional deep neural network. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9237:429–436, 2015. ISSN 16113349. doi: 10.1007/978-3-319-22482-4_50.

Aaron van den Oord, Sander Dieleman, and Benjamin Schrauwen. Deep content-based music recommendation. *Electronics and Information Systems department (ELIS)*, page 9, 2013. ISSN 10495258. doi: 10.1109/MMUL.2011.34.van. URL http://papers.nips.cc/paper/5004-deep-content-based-music-recommendation.pdf.

Shi Zhong. Efficient online spherical k-means clustering. *Proceedings of the International Joint Conference on Neural Networks*, 5:3180–3185, 2005. ISSN 14673045. doi: 10.1109/IJCNN.2005.1556436.