Capstone Proposal

Domain Background:

Automatic music classification/recognition is one such area which is being widely used in many commercial applications also like Shzam, Google Play, Sony Track ID, etc. All these applications have one thing in common that they aim to understand the semantic of the music rather than just curating the metadata out of it. In order to develop an advanced and intelligent music player there is a large semantic gap between audio signal processing and listeners' preference. Most of the cloud based music providers use collaborative filtering and sound meta data to recommend the next song to the listeners. But they could not fulfil the gap of listener's preference i.e genre, mood, lyrics, instrumentation, rhythm, music records time etc.

Thus, we aim to focus the major area for preference of a music listener is genre of music. Music genre classification is one such problem in MIR (Music Information Retrieval) which has been addressed by a lot of signal processing techniques along with standard machine learning algorithms in place. But as we have seen that the traditional signal processing techniques to extract features don't add much advantage when it comes to typical problems of clustering, classification because we might miss many important features at the time of feature extraction and selection. As Deep Learning using Convnets has proven its importance in image processing to understand the features from the image pixels. Similarly, we here aim to adopt a similar approach for audio processing.

Applications

1. Music Tag Recommender (Genre based tagging)
2. Personalized Song Recommender Systems (eg. spotify)
3. Music Teaching Applications to make the students understand, what genre are all about.
4. Automatic Meta Tagging of songs
5. Segmenting the genres in music players or in local PCs for organized storage.
6. Develop song indexing systems using genre as on the key to the hash or node to the btree.

Problem Statement

To develop an algorithm that can take a set of songs and assign genre to the particular song. Rather than mining features using traditional Audio Processing Techniques we will use Supervised Deep Learning using Convnets to understand both features and classify them efficiently into different genres eg blues, classical, country, disco, metal, pop, etc.

Data Set

We thus use the public dataset GTZAN which is a collection of 1000 songs each of 30 second duration and 10 tagged genres as blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae, and rock.

Solution

1. Extract the MFCCs
2. Develop a Deep Convoluted Neural Network
   a. Convolution with strided (valid or same)
   b. Pooling (Play with different Pooling Techniques)
   c. Fully Connected Layer

Commented [C1]:

Bench Mark Model

Tao Feng paper on Deep learning for music genre classification where he has implemented the 4 class classification model with training accuracy of 94 % and test accuracy of 61 % using restricted Boltzmann machine. Thus we aim to develop a fully 10 class classification system with some improvements.

Evaluation Metrics

1. Classification Accuracy
2. Confusion Matrix

Project Design

We will develop a RESTFul web service where one can POST an audio and get a probability distribution of all the genres.

Technologies to be used

1. Python
2. Flask
3. Tensorflow
4. pyAudio or YAAFEE or LabRosa for MFCCs and chroma extraction

Workflow:

1. Extract the features and metadata from the songs and store in NoSQL like Cassandra.
2. Read the features and tags from the database.
3. One Hot encoding of labels
4. Start with Just Fully Connected NN and then add few hidden layers and evaluate.
5. Add Convnet Layer(s) with different strides and pooling and evaluate.
6. After reaching a significant amount of accuracy we will finally write a flask application with following request
   a. Request-Header : application/json
   b. Method: POST
   c. Response: application/json
   d. Request-data:

```
{
        "song_url":"/webapp/data/music/*.mp3",
}
```

   e. Response-data:

```
[
{
        "genre":"blues",
        "score":0.8361
}
]
```

References

1. Music Genre Classification and Variance Comparison on Number of Genres By Miguel Francisco, miguelf@stanford.edu, Dong Myung Kim, dmk8265@stanford.edu

2. Recommending music on Spotify with deep learning
   (http://benanne.github.io/2014/08/05/spotify-cnns.html)
3. Tao Feng, Deep learning for music genre classification, University of Illinois.
4. Musical Genre Categorization Using Support Vector Machines By Shu Wang
   swang367@wisc.edu.
   http://homepages.cae.wisc.edu/~ece539/fall13/project/WangShu_rpt.pdf