TITLE??

Problem and Motivation

* Identifying the unique country a song is popular in using only musical properties. No contextual information is used. It is a single-label, multiclass classification problem.
* We want to see if the music tastes of people from different countries can be distinguished. Being able to predict the countries a song would be popular in, the industry can identify the markets to target for promotional activities. It potentially may uncover market – music type pairs which has high probability of being a good match but not utilized for promotion planning.

Why deep learning

* Complex relations including non-linearities and interactions.
* Other ML approaches that allow for complexity, such as Boosting Trees, are not appropriate as they do account for sequential nature of the data.
* Other ML approaches do not naturally extend to the multiclass cases???
* Abundance of song data allows us to train complex deep learning models.

Goal

* Proposing a competitive prediction model that can achieve the above-mentioned task. Additionally, discovering the important aspects of the learning task – how sampling rate and sampling window selection affects the performance.
* The inference part, such as identifying new market – music type pairs, is out of the scope of this research. However, a good prediction model would help understanding the underlying mechanism and pave the way for the inference.

Outcome

* Report it shortly in intro

Previous work and positioning

Methodology

* In order to represent the differences in segment durations in our input data, we sampled segments at a rate proportional to their durations. We also expected higher sample rates to capture this relation more precisely because of rounding to integers.
* Adequate information to reproduce the analysis
* Make the github repo public and add the link in the paper. ( with a read me file for the pipeline to reproduce)
* Explain why you chose the metric
* Specified architecture based on intuitions about music
  + Music is listened to sequentially (LSTM layer)
  + Perception is retroactively shaped by later information (Backward LSTM layer)
  + Musical properties interact in non-linear ways (Dense layer)
  + Output needed to be a classification (Dense layer with softmax)
* Wanted to know how duration (data quantity) interacted with sample rates (data quality) to improve classification
  + Performed a grid search

Results

* Higher sample rate hurt the performance. We attribute this to the fact that it gets harder for LSTM to remember what the state of the system was like a few segments ago because a segment consists of a higher number of time steps.
  + Higher sample rates in this methodology may have introduced redundant information with the only real value coming from difference in the temporal dimension
* Compare against baseline

Discussion

* Future work
* Success, failure, takeaway, unexpected outcome

Possible tables

* Country list with number of observations
* Descriptive stats
* Histogram of segment durations
* Tensorboard plots
* Confusion matrix