Audio Tagging Challenge(New)
Group Proposal
Madison Turano, Jingjing Xu, Jingyi Xiao

Since the last dataset we selected had 41 categories, we decided to use a different audio tagging dataset instead.

Problem Selection

- Tagging/classifying audio recordings
- We thought we would learn more from using audio data because we've never worked with audio data.

Dataset

- This audio dataset is quite large as well, the data is about 6 GB. It consists of solo recordings of 8 musical instruments (classes): clarinet, distorted electric guitar, female singer, flute, piano, tenor saxophone, trumpet, and violin. All audio samples have a single label. Each clips is 3 seconds long.
- It is split into 3 sets: train, validation, and test, and has a total of 21572 clips.

Deep Network Framework

- We will use convolution network or LSTM network, as we are analyzing the similarities between audio samples.
- We decide to use Pytorch, since we want to practice it, and gain more experience with it.

Reference Materials

- A Survey of Audio-Based Music Classification and Annotation
- Getting Started with Audio Data Analysis using Deep Learning (with case study)
- Neural Network Design, Chapter 25: Case Study 3: Pattern Recognition

Performance Metrics

- Accuracy both in terms of overall and each category.
- Confusion matrix.

- Time used.

Rough Schedule

- Apr. 20, Rough Network designed;
- Apr. 24, Network designed, and group report draft, start working on individual report, create GitHub repo;
- Apr. 26, Exploring different optimizer and functions, with final modification on the report, also finish individual report.