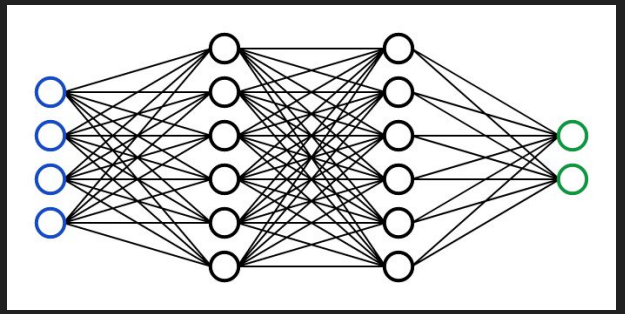
Audio Classification

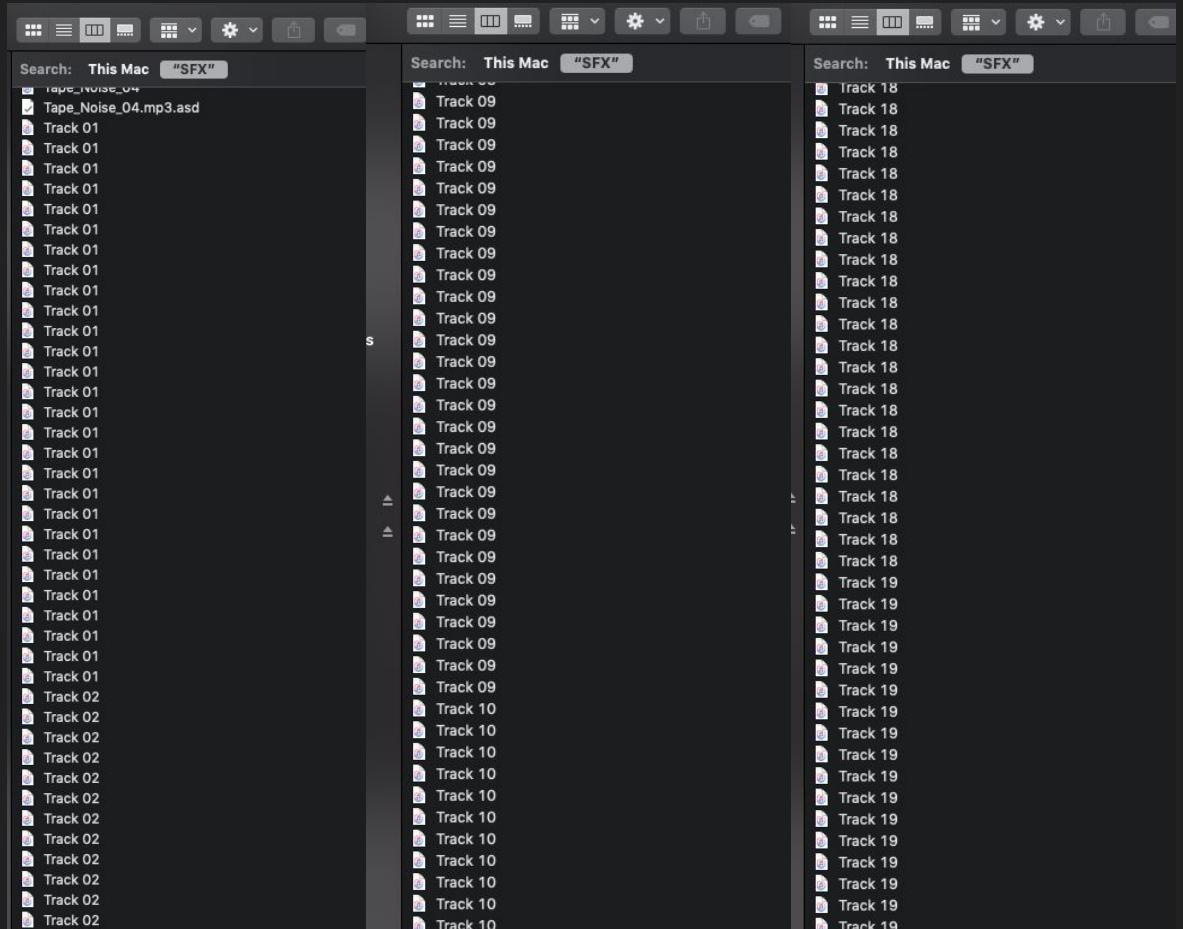
With Neural Networks



Why is Audio Classification Useful?

Organizing Sound FX Libraries

SFX Libraries can be enormous and very poorly labelled.



My personal SFX library is 413.52 GB and would not be considered a large SFX library

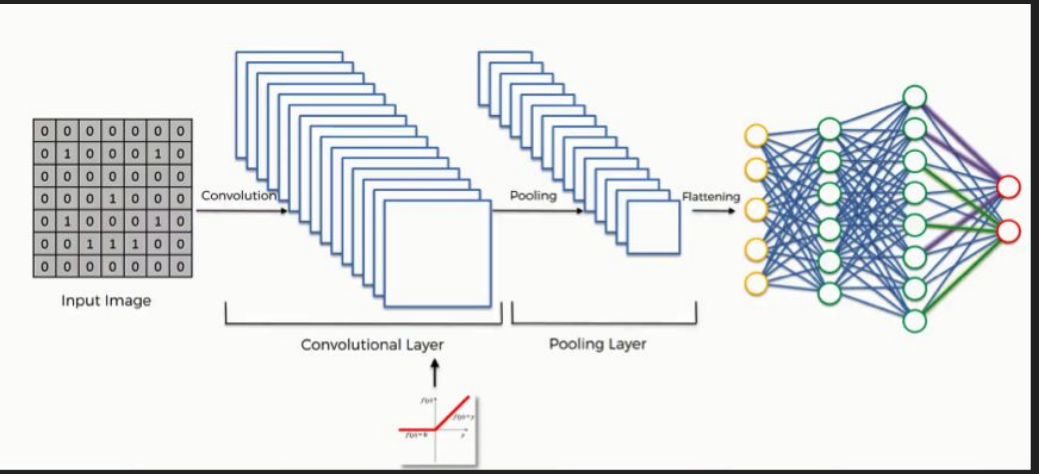
These are all titled Track 18

These have no meta data tags

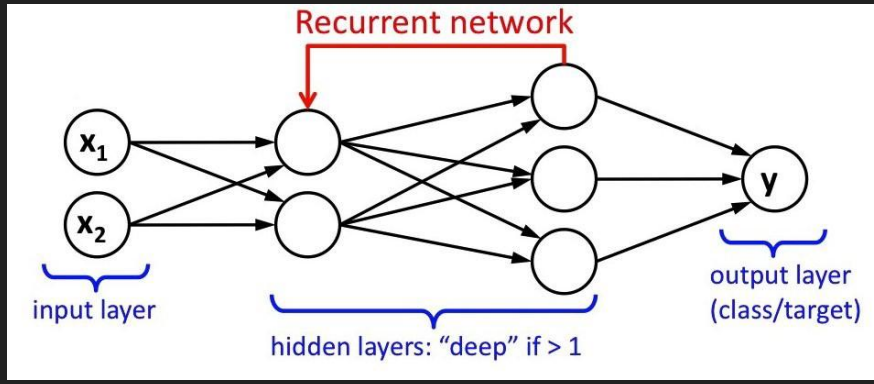
These are just a few examples of the audio files that have been sitting in my SFX library without anyway for them to come up in a search, simply wasting space.

Neural Networks for Audio Classification

Convolutional Neural Network



Recurrent Neural Network



Trained on 9,457 audio files corresponding to 41 classes

```
['Acoustic_guitar', 'Applause', 'Bark', 'Bass_drum', 'Burping_or_eructation', 'Bus', 'Cello', 'Chime', 'Clarinet', 'C
omputer_keyboard', 'Cough', 'Cowbell', 'Double_bass', 'Drawer_open_or_close', 'Electric_piano', 'Fart', 'Finger_snapp
ing', 'Fireworks', 'Flute', 'Glockenspiel', 'Gong', 'Gunshot_or_gunfire', 'Harmonica', 'Hi-hat', 'Keys_jangling', 'Kn
ock', 'Laughter', 'Meow', 'Microwave_oven', 'Oboe', 'Saxophone', 'Scissors', 'Shatter', 'Snare_drum', 'Squeak', 'Tamb
ourine', 'Tearing', 'Telephone', 'Trumpet', 'Violin_or_fiddle', 'Writing']
```

```
training convolutional model
Train on 1009704 samples, validate on 112190 sam
Epoch 1/1
1009704/1009704 [==============================]
val_acc: 0.6766
```

```
training recurrent model
Train on 1009704 samples, validate on 112190 sam
Epoch 1/1
1009704/1009704 [==============================]
val_acc: 0.6754
```

Cello

Predictions

CNN →

```
Meow
Laughter
Applause
```

RNN →

```
Applause
Cello
Violin_or_fiddle
```

Laughter

Predictions

CNN →

```
Laughter
Cough
Telephone
```

RNN →

```
Laughter
Cough
Writing
```

Bus

Predictions

CNN →

```
Gong
Scissors
Computer_keyboard
```

RNN →

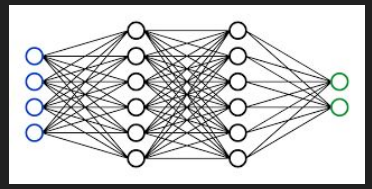
```
Drawer_open_or_close
Scissors
Writing
```

IBM Developer Model Asset Exchange: Audio Classifier

This repository contains code to instantiate and deploy an audio classification model. This model recognizes a signed 16-bit PCM wav file as an input, generates embeddings, applies PCA transformation/quantization, uses the embeddings as an input to a multi-attention classifier and outputs top 5 class predictions and probabilities as output. The model currently supports 527 classes which are part of the Audioset Ontology. The classes and the label_ids can be found in class_labels_indices.csv. The model was trained on AudioSet as described in the paper 'Multi-level Attention Model for Weakly Supervised Audio Classification' by Yu et al.

Trained on the google audio dataset:

2,084,320 YouTube Videos with 527 classes



Classification	Prob
Vehicle	.63
Whir	.47
Engine	.38
Rumble	.17
Earthquake	.06

Threshold = 0.25

Sound FX Library

Model Running on Docker Image

Classifications

Metadata Tags

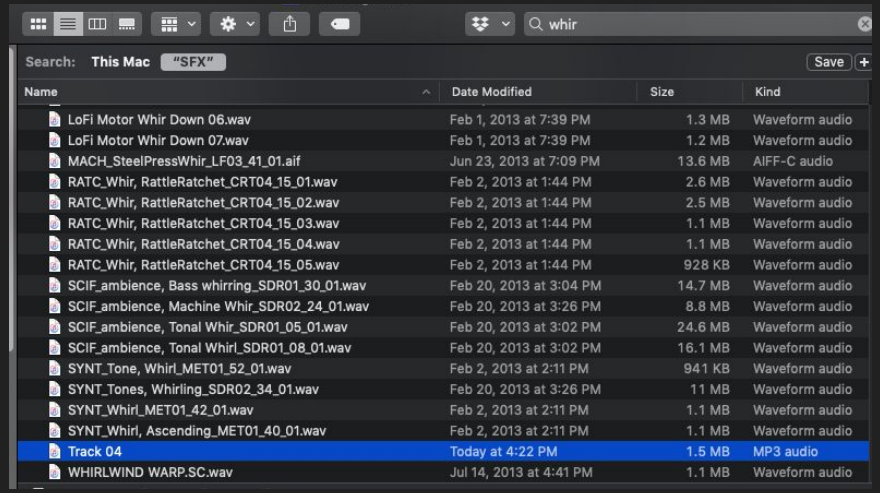
```
sending Track 24.wav to API
Adding these tags: Animal, Dog, Domestic animals, pets, Bow-wow, Bark
```

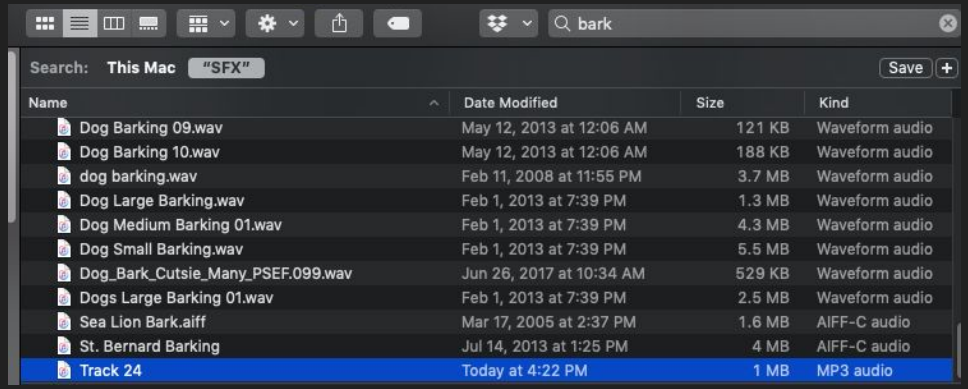
```
sending Track 35.wav to API
Adding these tags: Trumpet, Brass instrument, Music, Musical instrument
```

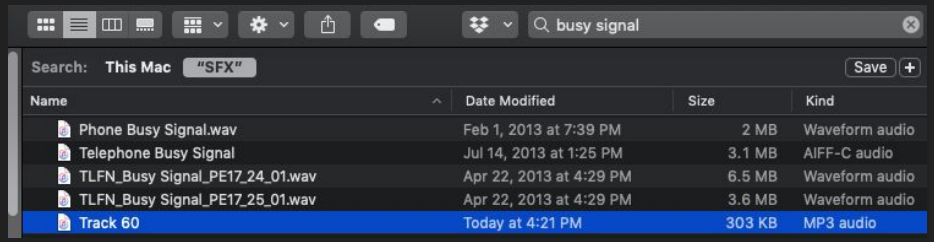
```
sending Track 60.wav to API
Adding these tags: Telephone, Busy signal, Music
```

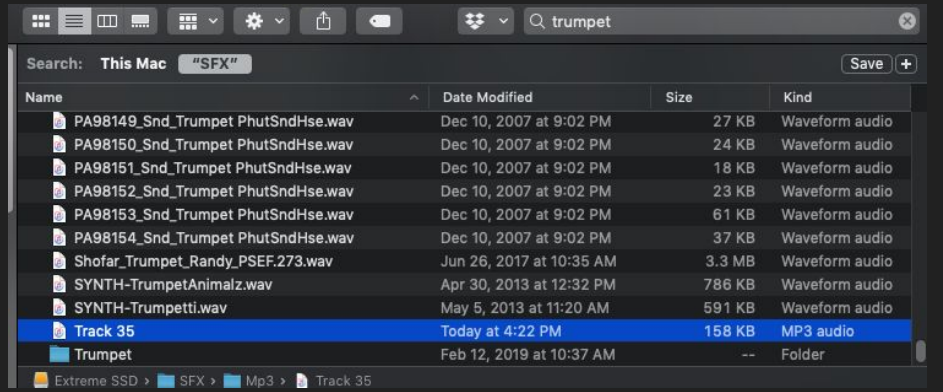
```
sending Track 04 4.wav to API
Adding these tags: Vehicle, Whir
```

Sample Searches









Dylan Case

dylancase@gmail.com
310-701-5537

dylancase Changed Title — Latest commit 90cefb3 1 minute ago

ClassifyAndTag.py	Added Script	20 hours ago
README.md	Changed Title	1 minute ago
models.py	Added Models	20 hours ago

README.md

Audio Classification and Metadata Tags

My Final Capstone Project from the Galvanize Data Science Immersive

Goals

Having been a recording engineer prior to enrolling in Galvanize's Data Science Immersive, I really wanted to use my new found data science skills to solve a problem I often encountered in the Post Production Audio world, poorly organized and labelled sound effect libraries. I set out to solve that problem by creating an audio classification model to help organize and tag Sound FX directories/libraries.

GitHub

github.com/dylancase/

linkedin.com/in/dylancase/