


# Music Genre Classification

Phillip Wagner

# Project Summary

- Goal: Classify Songs by genre
- Steps
  - Data preparation
  - Feature extraction
  - Model training
  - Evaluation/Improvement

# Importing dataset



Create

Home

Competitions

Datasets

Models

Code

Discussions

Learn

More

Your Work

VIEWED

GTZAN Dataset - M...

Iris.csv

Search

741

New Notebook

Data Card

Code (221)

Discussion (17)

Suggestions (0)

## About Dataset

Context

Music. Experts have been trying for a long time to understand sound and what differentiates one song from another. How to visualize sound. What makes a tone different from another.

This data hopefully can give the opportunity to do just that.

Content

- genres original** - A collection of 10 genres with 100 audio files each, all having a length of 30 seconds (the famous GTZAN dataset, the MNIST of sounds)
- images original** - A visual representation for each audio file. One way to classify data is through neural networks. Because NNs (like CNN, what we will be using today) usually take in some sort of image representation, the audio files were converted to Mel Spectrograms to make this possible.
- 2 CSV files** - Containing features of the audio files. One file has for each song (30 seconds long) a mean and variance computed over multiple features that can be extracted from an audio file. The other file has the same structure, but the songs were split before into 3 seconds audio files (this way increasing 10 times the amount of data we feed into our classification models). *With data, more is always better.*

<input type="checkbox"/>	Name	Status	Date modified	Type	Size
<input type="checkbox"/>	blues	✓	5/8/2024 1:02 PM	File folder	
<input type="checkbox"/>	classical	✓	5/8/2024 1:02 PM	File folder	
<input type="checkbox"/>	country	↻	5/8/2024 1:02 PM	File folder	
<input type="checkbox"/>	disco	↻	5/8/2024 1:02 PM	File folder	
<input type="checkbox"/>	hiphop	↻	5/8/2024 1:02 PM	File folder	
<input type="checkbox"/>	jazz	↻	5/8/2024 1:02 PM	File folder	
<input type="checkbox"/>	metal	↻	5/8/2024 1:02 PM	File folder	
<input type="checkbox"/>	pop	↻	5/8/2024 1:03 PM	File folder	
<input type="checkbox"/>	reggae	↻	5/8/2024 1:03 PM	File folder	
<input type="checkbox"/>	rock	↻	5/8/2024 1:03 PM	File folder	

# Converting data format

- Convert MP3 to WAV files, because they can be directly read by the `scipy.io.wavfile` package

## MP3 to WAV Converter

CloudConvert converts your audio files online. Amongst many others, we support MP3, M4A, WAV and WMA. You can use the options to control audio quality and file size.

convert

MP3 ▼

to

WAV ▼



Select File



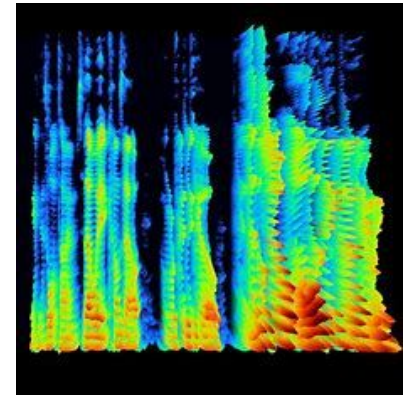
OPTIONS



Audio

# Fast Fourier Transform Feature Extraction

- Different genres often have characteristic frequency patterns, FFT serves as a critical tool for feature extraction.
- FFT is applied using `scipy.fft.fft`, function computes the spectrum of the signal, capturing the amplitude of each frequency component
- Reduces complexity to focus on most significant frequencies
- spectrogram is a visual representation of the frequencies that occur in a song
  - Shows intensity to use for classification



# Model Breakdown

- Logistic Regression Classifier -multiclass classification
- spectrogram is a visual representation of the frequencies that occur in a song
  - Shows intensity to use for classification
- model makes its genre prediction based on which genre has highest predicted probability for a given song, derived from the logistic function applied to the linear combination of FFT features.
- See accuracy with confusion matrix

```
[[26  1  2  0  0  2]
 [ 4  7  5  0  5  3]
 [ 1  2 14  2  8  3]
 [ 5  4  7  3  7  5]
 [ 0  0 10  2 10 12]
 [ 1  0  4  0 13 12]]
```