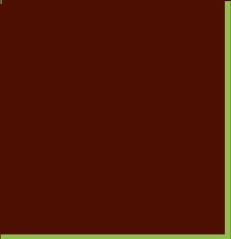# Music Genre Classification

Submitted by:

Suraj Kumar - 193050003
Yadnyesh Patil - 193050067
Rajershi Gupta - 193059002
Aman Kumar Singh -193050022

Topic and Motivation

- **Improving accuracy of existing implementation of genre classification models.**
- A large number of companies out there work on the music genre classification problem like spotify, shazam etc.
- The problem is really challenging since it involves complex features.

Prior work

- Most influential work on genre classification using ML technique was pioneered by Tzanetakis and Cook

- Companies like spotify and Shazam constantly work in the music recommendation and detection industry, of which genre classification is a primary operation

Requirements

- **Dateset** : GTZAN genre dataset from the MARYSAS website

- **Language and framework :** Python on Jupyter notebook

- **Python Libraries** : Numpy, Pandas, Scipy, Scikit-learn, Librosa

- **System** : Any normal system with at least 2 GB RAM and the capability to run Jupyter

Implementation of existing code
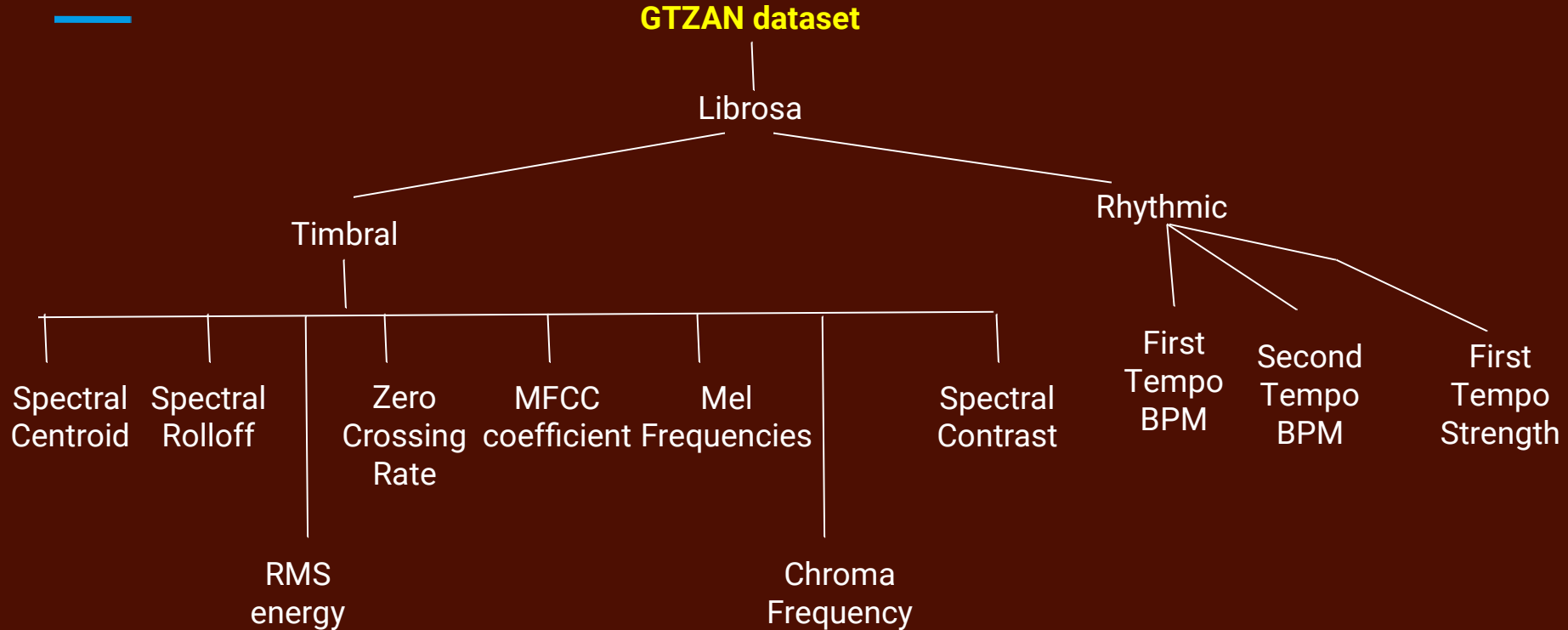
Classifiers and hyperparameters

Classifier	Hyperparameters
Decision Tree	Min_sample_split = 20
Random Forest	Min_sample_split = 10
KNN	Default
Logistic Classifier	Max_iter = 15,000
Linear SVM	Max_iter = 10,000
Naive Bayes Bernoulli	Default
Naive Bayes Gaussian	Default

Features Description

Spectral Centroid	”centre of mass” for a sound. Blues--center, Rock--end
Spectral Rolloff	measure of the shape of the signal. It represents the frequency below which a specified percentage of the total spectral energy
MFCC coefficients	Differentiates between same rhythm, pitch audio
Mel spectogram	Differentiates between same rhythm, pitch audio
RMS energy	Detecting new environments
Zero Crossing Rate	the rate at which the signal changes from positive to negative or back
Spectral Contrast	Helps to detect noise in samples.
Rhythmic features	the timing, beat, and tempo of the song, differentiates between 2 different rhythm audio

Combinations tried

	8 features	contrast	Mfcc-coeff	Mfcc-mel	Chroma	Rythm
table-1.	✓	X	✓	X	X	✓
table-2.	✓	X	7	X	X	✓
table-3.	✓	✓	5	5	✓	X
table-4.	✓	X	X	X	X	✓
table-5.	✓	✓	4	4	X	✓
table-6.	✓	✓	5	5	✓	✓
table-7.	X	✓	5	5	6	✓

Results

	Train Accuracy	Test Accuracy
Decision Tree	76.35	45
Random Forest	93.45	62
KNN	63.15	43
Logistic Regression	67.70	57
Support Vector	47.39	41
Naive Bayes Bernoulli	19.97	24
Naive Bayes Gaussian	54.16	46

Train accuracy	Test Accuracy
73	52
92	60
49	27
51	47
21	22
20	23
42	43

Thank You