VISVESVARAYA TECHNOLOGICAL UNIVERSITY

JNANA SANGAMA, BELAGAVI – 590 018

An Internship Project Report on

"MUSIC GENRE CLASSIFICATION"

Submitted in partial fulfillment of the requirements as a part of the

AI/ML INTERNSHIP

(NASTECH)

For the award of degree of

Bachelor of Engineering
in
Computer Science and Engineering

Submitted by

KHUSHI PAI
1RN19CS065

MOONISAH BATOOL
1RN19CS081

Internship Project Coordinators

Dr. R Rajkumar
Associate Professor
Dept. of ISE, RNSIT

Mrs.Sunitha K
Assistant Professor
Dept. of ISE, RNSIT

Department of Computer Science and Engineering

RNS Institute of Technology

Channasandra, Dr. Vishnuvardhan Road, RR Nagar Post,
Bengaluru – 560 098

2021 -2022

RNS Institute of Technology

Channasandra, Dr. Vishnuvardhan Road, RR Nagar Post,

Bengaluru – 560 098

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that the mini project report entitled ***MUSIC GENRE CLASSIFICATION*** has been successfully completed by **KHUSHI PAI** bearing USN **1RN19CS065** and

MOONISAH BATOOL bearing USN **1RN19CS081**, presently VII semester students of **RNS Institute of Technology** in partial fulfillment of the requirements as a part of the ***AI/ML Internship (NASTECH)*** for the award of the degree of ***Bachelor of Engineering in Computer Science and Engineering*** under **Visvesvaraya Technological University, Belagavi** during academic year **2021 − 2022**. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the report and deposited in the departmental library. The mini project report has been approved as it satisfies the academic requirements as a part of Mobile.

Dr. R Rajkumar	**Mrs. Sunitha K**	**Dr. Suresh L**
Coordinator	Guide	Professor and HoD
Associate Professor	Assistant Professor	

External Viva

Name of the Examiners	**Signature with date**
1. _____	_____
2. _____	_____

ABSTRACT

The audio corpus available today on Internet and Digital Libraries is increasing rapidly in huge volume. We need to properly index them if we want to have access to these audio data. The search engines available in market also find it challenging to classify and retrieve the audio files relevant to the user's interest. In this project, we describe an automated classification system model for music genres. We firstly found good feature for each music genre. To obtain feature vectors for the classifiers from the GTZAN genre dataset, MFCC feature was used. K Neighbor classifier was trained and used to classify, each yielding varying degrees of accuracy in prediction.

ACKNOWLEDGMENT

The fulfillment and rapture that go with the fruitful finishing of any assignment would be inadequate without the specifying the people who made it conceivable, whose steady direction and support delegated the endeavors with success.

We would like to profoundly thank **Management** of **RNS Institute of Technology** for providing a healthy environment to carry out this AI/ML Internship Project.

We would like to express our thanks to our Principal **Dr. M K Venkatesha** for his support and for inspiring us toward the attainment of knowledge.

We wish to place on record our words of gratitude to **Dr. Kiran P,** Professor and Headof the Department, Computer Science and Engineering, for being the enzyme and master mindbehind our Mobile Application Development Laboratory with Mini Project Work.

We like to express our profound and cordial gratitude to my Internship Project Coordinators**, Dr. R Rajkumar,** Associate Professor, Department of Information Science and Engineering for their valuable guidance, constructive comments, continuous encouragement throughout the Mini Project Work and guidance in preparing report.

We would like to thank all other teaching and non-teaching staff of Information Science& Engineering who have directly or indirectly helped us to carry out the Mini Project Work.

Also, we would like to acknowledge and thank our parents who are source of inspirationand instrumental in carrying out this Mini Project Work.

KHUSHI PAI
USN:1RN19CS065

MOONISAH BATOOL
USN:1RN19CS081

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

1.1 ORGANIZATION/INDUSTRY

1.1.1 COMPANY PROFILE

NASTECH is formed with the purpose of bridging the gap between Academia and Industry. Nastech is one of the leading Global Certification and Training service providers for technical and management programs for educational institutions. We collaborate with educational institutes to understand their requirements and form a strategy in consultation with all stakeholders to fulfill those by skilling, reskilling and upskilling the students and faculties on new age skills and technologies.

1.1.2 DOMAIN/TECHNOLOGY

The domain chosen for our project is AI/ML. Machine learning, the fundamental driver of AI, is possible through algorithms that can learn themselves from data and identify patterns to make predictions and achieve your predefined goals, rather than blindly following detailed programmed instructions, like in traditional computer programming. This technology allows the machine to perceive, learn, reason and communicate through observation of data, like a child that grows up and acquires knowledge from examples. Machines also have the advantage of not being limited by our inherent biological limitations. With machine learning, manufacturing companies have increased production capacity up to 20%, while lowering material consumption rates by 4%.

Nowadays, the revolutionary AI technology evolved from rule-based expert systems to machine learning and more advanced subcomponents such as deep learning (learning representations instead of tasks), artificial neural networks (inspired by animal brains) and reinforcement learning (virtual agents rewarded if they made good decisions).

The AI can master the complexity of the intertwining industrial processes to enhance the whole flow of production instead of isolated processes. This enormous cognitive capacity gives the AI the ability to consider the spatial organization of plants and the timing constraintsof live production. Another key advantage is the capability of AI algorithms to think

probabilistically, with all the subtlety this allows in edge cases, instead of traditional rule based methods that require rigid theories and a full comprehension of problems.

1.1.3 Department

R.N.Shetty Institute of Technology (RNSIT) established in the year 2001, is the brain-child ofthe Group Chairman, Dr. R. N. Shetty. The Murudeshwar Group of Companies headed by Sri.
R. N. Shetty is a leading player in many industries viz construction, manufacturing, hotel, automobile, power & IT services and education. The group has contributed significantly to thefield of education. A number of educational institutions are run by the

R. N. Shetty Trust, RNSIT being one amongst them. With a continuous desire to provide quality education to the society, the group has established RNSIT, an institution to nourish and produce the best of engineering talents in the country. RNSIT is one of the best and top accredited engineering colleges in Bengaluru.

1.2 PROBLEM STATEMENT

1.2.1 Existing System and their Limitations

Music is made with a variety of components like frequency, tones, pitch etc. All these components when combined in a mixed ratio form a unique audio file. It's very time-consuming and inefficient to decide the genre by just listening to the audio file while there are many other factors to consider.

1.2.2 Proposed Solution

Music Feature Extraction coupled with machine learning algorithms can be used to implement the model which helps in classifying the audio file to a particular music genre.

1.2.3 Program formulation

Music genre classification uses specific music extraction features and functions to systematically identify, extract, quantify, and study affective components of the music file and come up with genre.

Chapter 2

REQUIREMENT ANALYSIS, TOOLS &TECHNOLOGIES

2.1 Hardware and Software Requirements

2.1.1 Hardware Requirements:

- Processor : Any Processor above 500 MHz
- RAM: 512Mb
- Hard Disk: 4GB or more

2.1.2 Software Requirements:

- Operating System: Windows 10 or above
- IDE: Visual Studio Code

2.2 Tools/Languages/Platforms

- Python, MFCC, sklearn Module, Scipy, Streamlit Numpy library

CHAPTER 3

3.1 Problem Statement

The main aim is to create a machine learning model, which classifies music samples into different genres. It aims to predict the genre using an audio signal as its input.

The idea behind this project is to see how to handle sound files in python, compute sound and audio features from them, run Machine Learning Algorithms on them, and see the results.

The objective of automating the music classification is to make the selection of songs quick and less cumbersome. If one has to manually classify the songs or music, one has to listen to a whole lot of songs and then select the genre. This is not only time-consuming but also difficult. Automating music classification can help to find valuable data such as trends, popular genres, and artists easily. Determining music genres is the very first step towards this direction.

Description of Dataset

For this project, the dataset that we will be working with is GTZAN Genre Classification dataset which consists of 1,000 audio tracks, each 30 seconds long. It contains 10 genres, each represented by 100 tracks.

The 10 genres are as follows:

- Blues
- Classical
- Country
- Disco
- Hip-hop
- Jazz
- Metal
- Pop
- Reggae
- Rock

	filename	length	chroma_stft_mean	chroma_stft_var	rms_mean	rms_var	spectral_centroid_mean	spectral_centroid_var	spectral_bandwidth_mean
0	blues.00000.0.wav	66149	0.335406	0.091048	0.130405	0.003521	1773.065032	167541.630869	1972.744388
1	blues.00000.1.wav	66149	0.343065	0.086147	0.112699	0.001450	1816.693777	90525.690866	2010.051501
2	blues.00000.2.wav	66149	0.346815	0.092243	0.132003	0.004620	1788.539719	111407.437613	2084.565132
3	blues.00000.3.wav	66149	0.363639	0.086856	0.132565	0.002448	1655.289045	111952.284517	1960.039988
4	blues.00000.4.wav	66149	0.335579	0.088129	0.143289	0.001701	1630.656199	79667.267654	1948.503884

5 rows × 60 columns

Figure 3.1 Description of dataset

3.2 Algorithm

K Nearest Neighbour Algorithm

K-Nearest Neighbour, in short “KNN”. It is a supervised machine learning algorithm. The calculation can be utilized to tackle both classification and regression problem statements. The quantity of closest neighbours to another obscure variable that must be predicted or classified is indicated by the symbol 'K'. Its aim is to find every one of the nearest neighbours around a new unknown data point to sort out what class it has a place with. It's a distance-based methodology. KNN calculates the distance from all focuses nearby the obscure data and filters out the ones with the briefest distances to it. Therefore, it's frequently alluded to as a distance-based calculation. To accurately arrange the outcomes, we should initially determine the value of K (Number of Nearest Neighbours).

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

Figure 3.2 KNN Formula

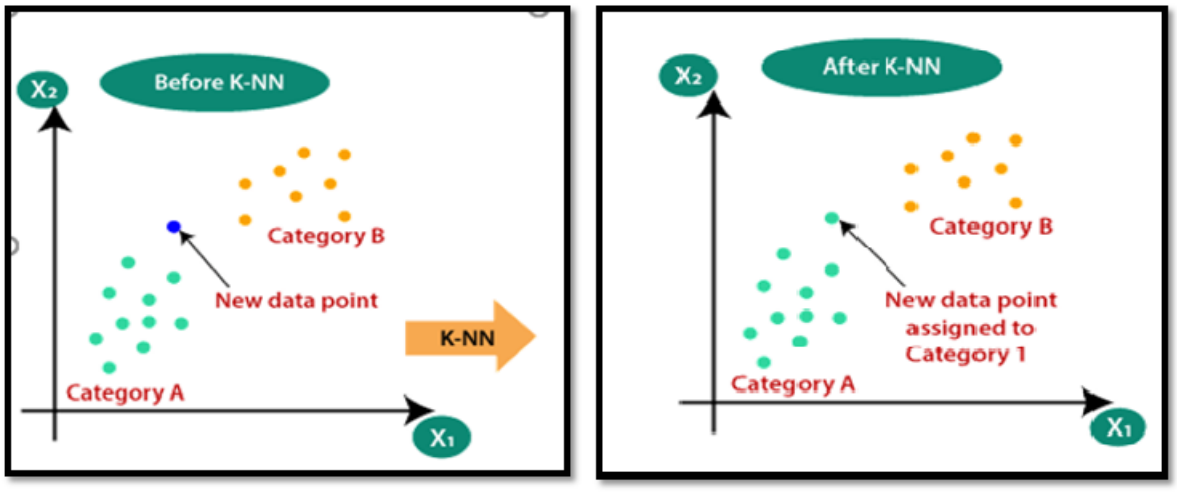

Figure 3.3 K Nearest Neighbours

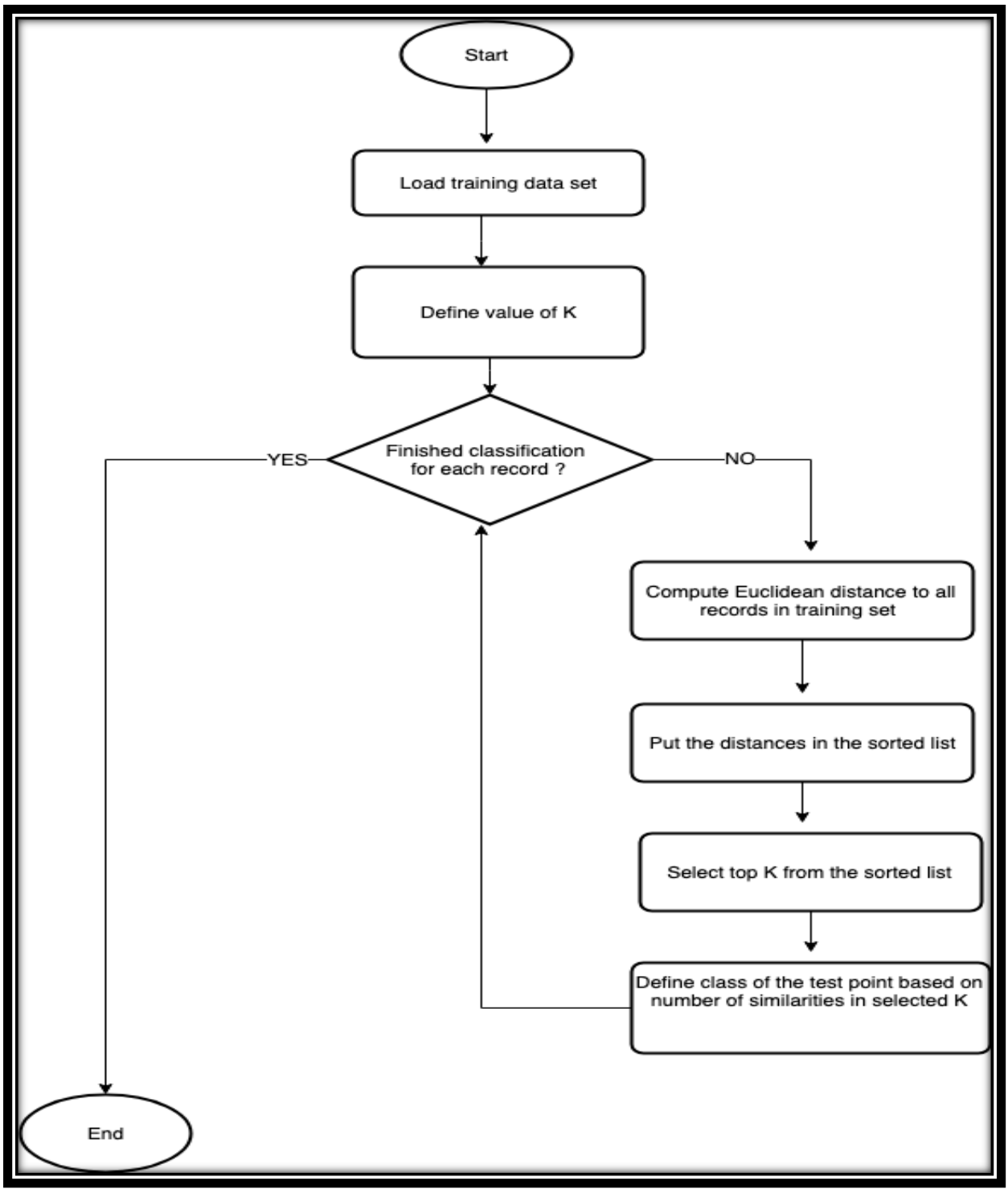

Fig 3.4 Working of KNN

WORKING OF MFCC

These are a set of short term power spectrum characteristics of audio files. It models the characteristics of human voice. We are taking into consideration 13 coefficients as the part of the final feature vector. The method to implement this feature vector is shown in figure 3.7

The five features extracted are –

1. Mel Frequency Cepstral Coefficients
2. Chroma Frequencies
3. Spectral Centroid
4. Spectral Roll-Off
5. Zero-crossing rate

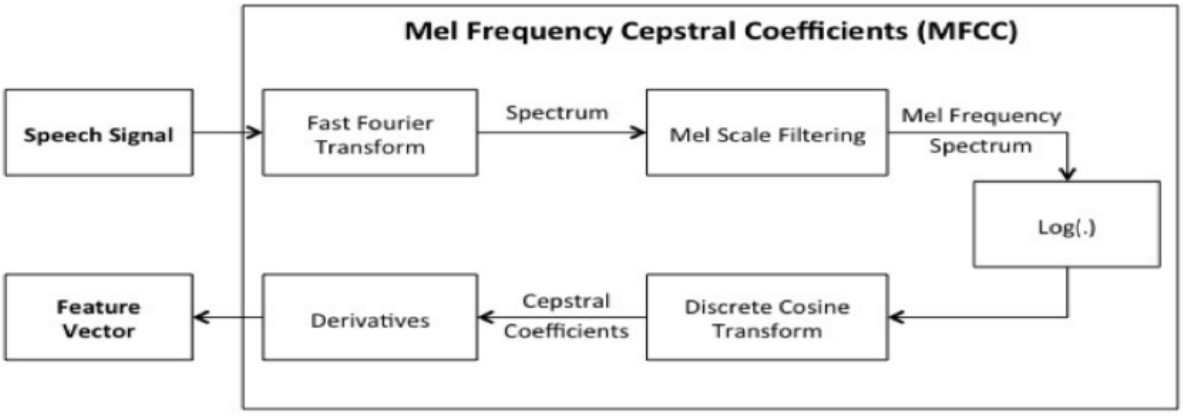

Figure 3.5 Feature Extraction using MFCCs

3.3 LIBRARIES

- Pandas
- Numpy
- Scipy
- Streamlit
- Sklearn
- Mfcc

proj.py > ...

```
import numpy as np
import pandas as pd
from python_speech_features import mfcc
import scipy.io.wavfile as wav
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
from scipy.stats import uniform, randint
import sklearn.metrics as skm
import sklearn.model_selection as skms
import sklearn.preprocessing as skp
from sklearn.linear_model import LogisticRegression
import sklearn.ensemble as ske
import streamlit as st
from pprint import pprint
import random
import librosa, IPython
import librosa.display as lplt
seed = 12
np.random.seed(seed)
```

Figure 3.6 libraries

3.4 PSEUDOCODE - Reading and understanding the data

1) Read the CSV file

```
df =pd.read_csv("C:/Users/SyedMoimn/Desktop/InternshipProject/archive (1)/Data/features_3_sec.csv")
```

2) Mapping labels to indices, Shuffle the data set and drop unwanted columns

```
# map labels to index
label_index = dict()
index_label = dict()
for i, x in enumerate(df.label.unique()):
    label_index[x] = i
    index_label[i] = x

# shuffle samples
df_shuffle = df.sample(frac=1, random_state=seed).reset_index(drop=True)

# remove irrelevant columns
df_shuffle.drop(['filename', 'length'], axis=1, inplace=True)
df_y = df_shuffle.pop('label')
df_X = df_shuffle
```

Chapter 4

OBSERVATION AND RESULTS

4.1 Training and Testing

- **Train Test split**

The Dataset has been split for training and testing by considering 80% data for training and 20% data for testing.

```
# split into train dev and test
X_train, X_test, y_train, y_test= skms.train_test_split(df_X, df_y, train_size=0.7, random_state=seed, stratify=df_y)
```

- **Training KNN model**

```
from sklearn.neighbors import KNeighborsClassifier
classifier = KNeighborsClassifier(n_neighbors=5)
classifier.fit(X_train, y_train)
```

- **Training Loading Trained Model**

```
import joblib
joblib.dump(classifier,'classifier')
model = joblib.load('classifier')
```

- **Extracting Features Using MFCC**

```
audio, sample_rate = librosa.load(filename, res_type='kaiser_fast')
mfccs_features = librosa.feature.mfcc(y=audio, sr=sample_rate, n_mfcc=57)
mfccs_scaled_features = np.mean(mfccs_features.T,axis=0)
mfccs_scaled_features=mfccs_scaled_features.reshape(1,-1)
```

4.2 Results & Snapshots

Figure 4.2.1.Front end using Streamlit

Figure 4.2.2 Model predicting a real world audio file

Chapter 5

CONCLUSION AND FUTURE ENHANCEMENT

5.1 Conclusion

- We implemented the KNN machine learning technique to classify music genres using GTZAN dataset.
- The model's test accuracy is 87%.
- In real application, a new music track can turn into features the same way as we mentioned, and applying our machine learning model we can predict its genre.
- To further improve the accuracy, we definitely need more music data to train our model and consequently the prediction accuracy is sure to increase

5.2 Future Enhancement

- Gathering more data for genres with less data currently to balance data distribution
- Model ensembling: combining classifiers by voting or averaging to improve performance
- Feature refining: add other musically relevant features for better classification results
- Real application: input new music tracks and transform them into features the same way as we mentioned, and apply our machine learning models to predict its genre.

Chapter 6

REFERENCE

[1] G. Tzanetakis, P. Cook, ―Musical genre classification of audio signals‖, IEEE Transactions on Speech and Audio Processing, Vol. 10, Issue 5, July 2002.

[2] Chandsheng Xu, Mc Maddage, Xi Shao, Fang Cao, and Qi Tan, ―Musical genre classification using support vector machines‖, IEEE Proceedings of International Conference of Acoustics, Speech, and Signal Processing, Vol. 5, pp. V-429-32, 2003.

[3] N. Scaringella, G. Zoia, and D. Mlynek, ―Automatic genre classification of music content: a survey‖, IEEE Signal Processing Magazine, Vol. 23, Issue 2, pp. 133–141, 2006.

[4] Jan Wülfing and Martin Riedmiller, ―Unsupervised learning of local features for music classification‖ ISMIR, pp. 139–144, 2012.

[5] Sox.sourceforge.net. Sox - sound exchange— homepage, 2015.

[6] https://ijcert.org/ems/ijcert_papers/V4I206.pdf

[7] http://marsyas.info/downloads/datasets.html