# GSHARP- Milestone Report

**Satvik Upadhyay**
ECE-4424 Machine learning
Virginia Tech, Blacksburg, VA
satvik@vt.edu

**Ujjwal Jain**
ECE-4424 Machine learning
Virginia Tech, Blacksburg, VA
ujjwalj7@vt.edu

## Abstract

GSharp is an application which is intended for aspiring musicians. GSharp takes a song as an input and identifies the genre of the song.

## 1 Introduction

GSharp is a music genre identifier which uses various machine learning concepts. It is implemented as an application which takes a Spotify playlist and a song number. Based on this data, it identifies whether the input song is a Rock, Hip-Hop (rap), or a jazz song. It requires the URL of the song library.

## 2 General formatting instructions

Currently, we have made an application where we can input the song in the GUI application. This should be a Spotify playlist URL with a song number. We are currently working on making the KNN classifier to identify music genres. We chose to move this way as we believe it would be better to get the simplest things done first. We are also working on making a more interactive way to input a song.

We are still working on this project to complete the reach goal that we had. We are discussing how we are going to isolate all the chords played by an instrument. This is the most challenging part of this project and we are trying to take our time because accuracy is the most important to us. We are discussing various optimization algorithms we could use to identify notes correctly and quickly. We are also discussing how we could split the dataset of musical notes into training and testing data efficiently.

### 2.1 Music genre Identifier

For this, we are using several Spotify libraries consisting of songs from different genres. This is primarily to train our classifier to recognize Rock, Jazz, and Hip-Hop. We are extracting several key features to achieve this goal. These features include: Acousticness, Instrumentalness, Danceability, Energy, Liveness, Loudness, Speechiness, Valence, Tempo, Key. 300 songs from each genre are used to train the classifier and rate the genres on these categories. After this, we tried another playlist of 30 songs each genre. And we used 15 random songs from each genre. We were unable to train the classifier with more songs since Spotify Apis prohibit usage over 100 songs. As a result, we had to create 9 playlists of 100 songs each to get at least 900 songs for training.

### 2.2 Note Identifier (Stretch goal; Currently underway)

This part helps in identifying all the notes played by a single instrument. Currently we are only discussing all the notes played in a solo (guitar or piano). We will later add a method to isolate instruments and identify the notes played by those instruments in the song.

### 2.3 Dataset used

1. spotify:playlist:6Sc9vDK1evql8o1mmjeVVf
2. spotify:playlist:6dTyRbNI6RAmXkj2EtFR1u
3. spotify:playlist:5ZgtvAEkmUklwUN4AgRFQM
4. spotify:playlist:0fGex3lz2Mb9SD8DnfF4Vt
5. spotify:playlist:70YgcjjKvHNbwH7yjBZxFN
6. spotify:playlist:47sWyvykENqVu5HfVY6TWw
7. spotify:playlist:7M9OSpMghzeP8Baoqr7pPY
8. spotify:playlist:6NYyTnXuUJi2nrBopdTaf3
9. spotify:playlist:2AsJXGbLUKtaQisJWzV28i
10. spotify:playlist:2Fr363VeYWOAfQwMPrTdEL (Rock Test)
11. spotify:playlist:37i9dQZF1DX7QOv5kjbU68 (Rap Test)
12. spotify:playlist:37i9dQZF1DXbITWG1ZJKYt (Jazz Test)

## 3 Results

We used about songs from each genre and we noticed that our classifier works well with Hip-Hop music. It identifies it correctly approximately 90% of the time. It also never misclassifies any other genre as Hip-Hop. We believe this stems from the fact that hip-Hop music is a relatively new form of music. As a result, it has many distinctive key features.

Also, Rock music was very easily misclassified. It was about 80% accurate with the playlist which we chose for testing. We believe this playlist contained music from many sub-divisions of Rock music. We wanted to add variation since Rock music is a very broad genre. Some of its key features were very similar to jazz, as a result, it reported those songs as jazz. We believe we needed a playlist of at least 5000 songs for Rock. But, unfortunately, Spotify's limits made it difficult for us to achieve this goal.

Jazz music had a relatively good success rate of about 85%. We believe this was the case since Jazz also has some very key components which separate it from the other genres. It is very different from Rap but a little similar to Rock. We believe this problem may have emerged since Rock music needed a larger dataset. As a result, the identifier was not trained well.

## 4 Future Goals

We are looking to create some more playlists to obtain at least 5000 songs from each genre. We believe this is essential for broad categories like Rock. We are also looking to use algorithms like Gradient Descent Algorithm to find the appropriate key features. Also, we are looking to implement some more genres to the program.

We are also hoping to implement some optimization algorithms to correctly identify the notes. We are also discussing testing these algorithms since we have 3 separate machine learning algorithms. Note identifier and Song identifier are interlinked. As a result, we could have cascading errors which stem from imperfection in one algorithm. We are hoping to identify all the notes correctly in a solo guitar song.

## 5 Contributions

Table1: contributions from each member

| Task | Member |
|---|---|
| Gui | Satvik |
| Key feature extraction | Ujjwal |
| Implementing KNN | Satvik |
| Results | Satvik, Ujjwal |
| Optimization | Satvik, Ujjwal |

## 6 GUI for GSHARP

The following image depicts the Gui environment for the project. There are different options which depicts the tools.

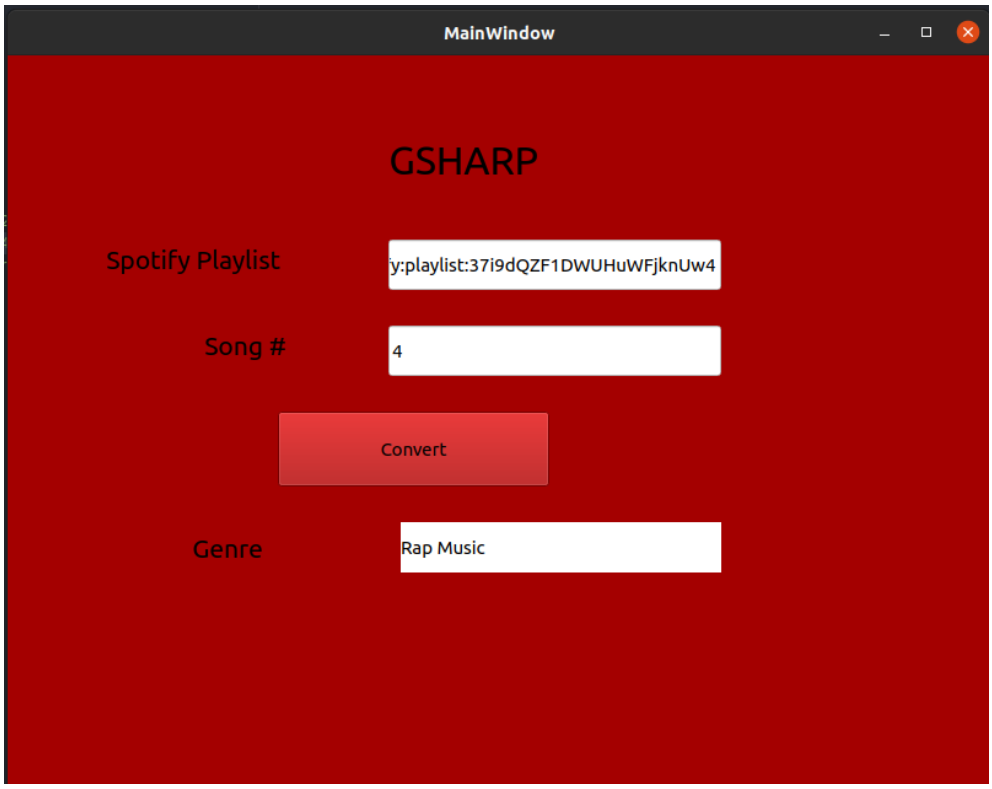


Figure1: Gui for GSharp showing a Rap songs being correctly identified.