Stephanie Benavidez

DSC680-T301 (2241-1)

Applied Data Science

Week 4: Milestone 3– Final White Paper

Introduction

For this first project of this semester, I thought it would be neat to do a classification model on music genres. According to an article by Ashley King, “the top streaming genres of 2023 so far are as follows: R&B/Hip-Hop, Rock, Pop, Latin, Country, Dance/Electronic, World Music, Christian/Gospel, Classical, and Jazz” (King, 2023). Growing up, my family listened to about half of these genres, especially on long road trips. My family even would turn it into a game to see if we could guess what the title of the song was, and who sang it to kill time for a long drive. Doing these activities as a child has allowed me to enjoy all genres of music, though there may be some genres that I still do not like.

Data Explanation

With this project, I will be using audio files of ten different genres, and trying to run a KNN model to see if it can classify which genre the audio file is. According to (Music Classification: Beyond Supervised Learning, Towards Realworld Applications, 2023) , music classification “is a music information retrieval (MIR) task whose objective is the computational understanding of semantics.” Though there are four different music

classification tasks, this project will just be focused on genre classification. Some of the research questions that I would like to try to explore are as follows:

- Will the model be able to understand the quality of the audio to choose the correct genre?
- Are there enough good-quality audio files to predict the genre?
- Will the model allow you to know if one of the audio files is not of excellent quality, or is it giving an error that something is wrong?
- Can I get at least an 80% accuracy level within the KNN classification model used?
- Will the KNN Model choose random files?
- If the KNN model is choosing random audio files, does that change the accuracy level too?
- If the KNN model chooses random audio files, what is the range that it is set to predict?

Data Source

From research, there are existing GTZAN datasets that have the majority of the information that will be used within this project. One of those datasets, that I will be using is coming from Kaggle.com (GTZAN Dataset - Music Genre Classification, n.d.). In this dataset, it has over 10 genres of music, with 100 audio files per genre, and two CSV files. For these CSV Files, no cleaning or manipulation needed to be done. The only thing was to merge the two CSV files into one dataframe. These merged CSV and audio files will be used for the predictions in the KNN model. Since they are audio files,

there was no need to do any manipulation, or cleaning up due to damaging the audio file. The ten genres that the audio files are on are as follows: blues, classical, country, disco, hip hop, jazz, metal, pop, reggae, and rock. The libraries used within this project are as follows: import NumPy as np, import os, import pickle, import random, import operator, from collections import defaultdict, from python_speech_features import mfcc, import scipy.io.wavfile as wav, import librosa, import pandas as pd, import seaborn as sns, from matplotlib import pyplot as plt, from sklearn.model_selection import train_test_split, cross_val_score, GridSearchCV, from sklearn.ensemble import RandomForestClassifier, and from sklearn.metrics import confusion_matrix, accuracy_score, from sklearn.neighbors import KNeighborsClassifer, from sklearn. preprocessing import StandardScaler, and from sklearn.utils import shuffle.

Methods/Analysis

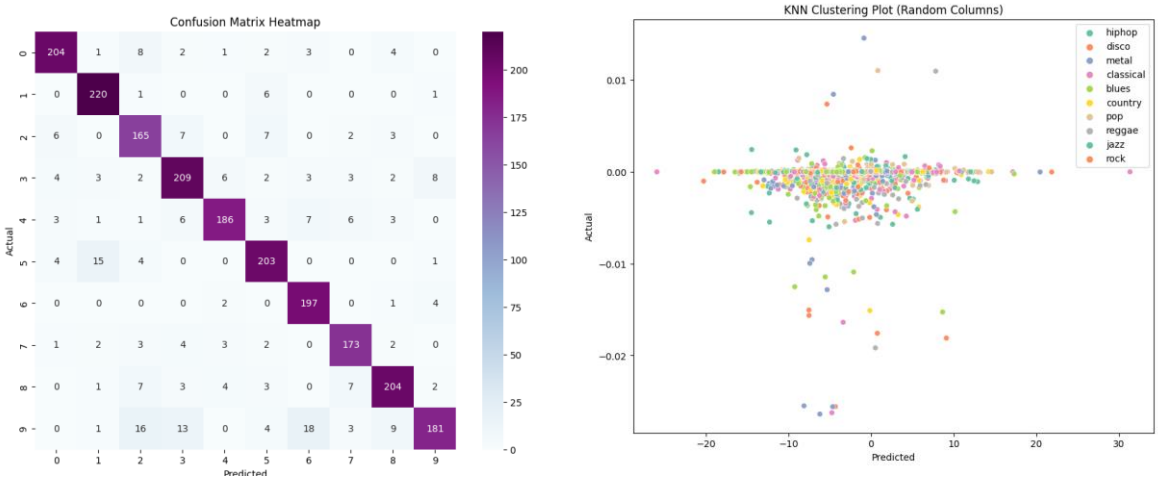
The analysis methods I have thought about using for this project are multiclass SVM, K-Means Clustering, K-Nearest neighbors, or a convolutional neural network. After starting this project, I have been able to see the analysis of KNN model predictions, and the K-Means Clustering to work from the KNN predictions given. Another analysis that was able to be done from the KNN predictions, is a confusion matrix heatmap, and K-Means clustering. In the final coding analysis, you will be able to see the genre predictions, confusion matrix heatmap, and K-Means Clustering. Currently, the KNN model can predict a range of 10 different audio files and give the

prediction, as shown in the screenshot below.

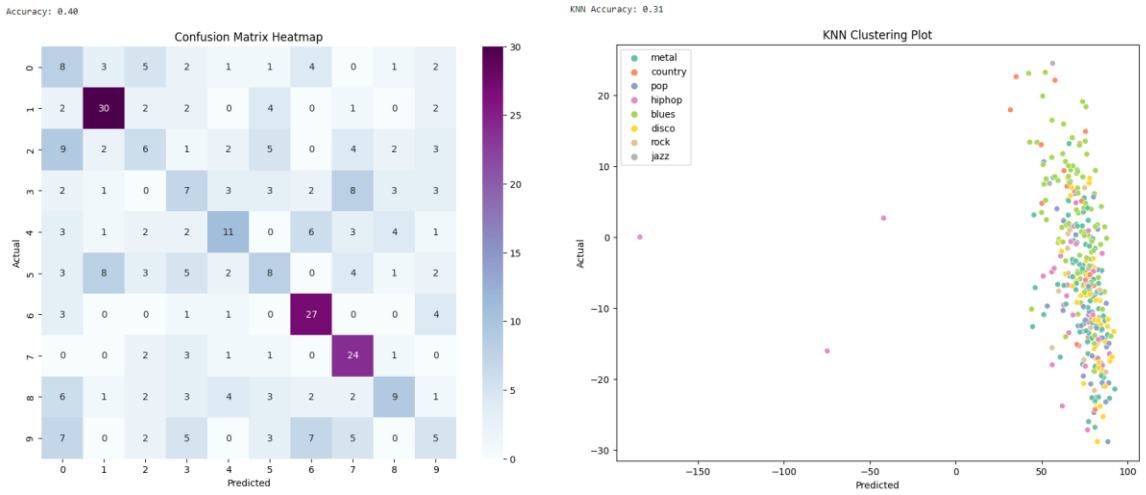
```
    ## print result
    print("Randomly selected audio file:", random_audio_file)
    print("Classified genre:", genre_name)
```

```
Randomly selected audio file: U:/School Homework/Fall 2023/DSC680-Applied Data Science/Project1/genres_original\pop\pop.00087.w
av
Classified genre: pop
Randomly selected audio file: U:/School Homework/Fall 2023/DSC680-Applied Data Science/Project1/genres_original\jazz\jazz.0004
7.wav
Classified genre: jazz
Randomly selected audio file: U:/School Homework/Fall 2023/DSC680-Applied Data Science/Project1/genres_original\classical\class
ical.00052.wav
Classified genre: classical
Randomly selected audio file: U:/School Homework/Fall 2023/DSC680-Applied Data Science/Project1/genres_original\country\countr
y.00010.wav
Classified genre: country
Randomly selected audio file: U:/School Homework/Fall 2023/DSC680-Applied Data Science/Project1/genres_original\metal\metal.000
80.wav
Classified genre: metal
Randomly selected audio file: U:/School Homework/Fall 2023/DSC680-Applied Data Science/Project1/genres_original\metal\metal.000
27.wav
Classified genre: metal
Randomly selected audio file: U:/School Homework/Fall 2023/DSC680-Applied Data Science/Project1/genres_original\jazz\jazz.0002
8.wav
Classified genre: jazz
Randomly selected audio file: U:/School Homework/Fall 2023/DSC680-Applied Data Science/Project1/genres_original\rock\rock.0003
2.wav
Classified genre: rock
Randomly selected audio file: U:/School Homework/Fall 2023/DSC680-Applied Data Science/Project1/genres_original\disco\disco.000
92.wav
Classified genre: disco
Randomly selected audio file: U:/School Homework/Fall 2023/DSC680-Applied Data Science/Project1/genres_original\disco\disco.000
04.wav
Classified genre: disco
```

For the confusion matrix and K-Mean Clustering based on the merged CSV, the results are as follows:



For the confusion matrix and K-Mean Clustering based on the audio file, the results are as follows:



As you can see the accuracy for the audio files is above the plots. The accuracy score changes with each random shuffle when the model is run, and the confusion matrix and K-Mean Clustering changes too. These results happen to the merged CSV file as well.

Ethical Assessment

The ethical considerations to consider are the audio files not having excellent quality, therefore, not allowing the project to reach the potential that is needed. However, there can be a workaround within the coding to allow it to bypass an audio file that is erroring out. There may also not be enough resources available to have the proper audio files to have a successful project. This semester's time is shortened to four weeks and may not produce the outcomes desired.

Lastly, some of the ethical questions asked during this project were:

- Is there anything wrong with any of the audio files?
- If so, does the coding allow you to see what is wrong with the file?
- Is there anything else that was trying to be accomplished but couldn't?

Challenges/Limitations/Assumptions

For the challenges or issues that may be faced within this project, the main concern is the data. Data always plays a huge role in the outcome of the project. Secondly, is to make sure your coding is worded properly to allow the audio files to be read, and the model to predict correctly. Being new to some of these models, the coding can be a little tricky, not allowing the desired results to be formulated. Lastly, a lack of resources can be difficult in any project, which also makes it difficult to get the desired results.

Some of the limitations of this project may relate to the amount of time allotted. Throughout this program, we have been given a 10-to-12-week time frame to complete one project. Whereas, with this class, we have four weeks to complete 3 projects. Which is a lot shorter time frame. Depending on how an individual does with time, they may try to rush the results, not allowing them to get the results desired.

Recommendations

The recommendations that I have for myself are to get the K-Means clustering plot coded and double-check all of the coding I have so far. Make any adjustments as needed so that the predictions, confusion matrix, and plots are presentable. If there was more time allotted to this project, something else I would add is for the model to be able to print the image of the audio file, or how the audio file looks in a PNG image with the genre original file. However, with this project, my main focus was to get the model to predict the correct genre of the audio file chosen.

Conclusion

In conclusion, the KNN classification model has been able to successfully look at 10 different audio files and predict the genres. Based on the predictions and accuracies for the model, it plotted a Confusion Matrix heatmap and K-Means Clustering plot. For the research and ethical questions asked throughout this project, the answers are as follows:

- Will the model be able to understand the quality of the audio to choose the correct genre? If the quality of the audio file is good, then yes it will be able to choose the correct genre.
- Are there enough good-quality audio files to predict the genre? Yes, there are over 1,000 audio files from ten different genres.
- Will the model allow you to know if one of the audio files is not of excellent quality, or is it giving an error that something is wrong? Yes, I have written an error reading exception in the coding that bypasses the bad file.
- Can I get at least an 80% accuracy level within the KNN classification model used? Since the KNN model is set on choosing random audio files, the accuracy level changes with those random files.
- Will the KNN Model choose random files? Yes, it currently is set to choose 10 random audio files each time the model is run.
- If the KNN model is choosing random audio files, does that change the accuracy level too? Yes.
- If the KNN model chooses random audio files, what is the range that it is set to predict? Currently, the range is set to 10.

- Is there anything wrong with any of the audio files? From the error reading exception, each time I ran the model, the only file that seemed to be corrupt was a jazz audio file.
- If so, does the coding allow you to see what is wrong with the file? Yes, the screenshot below shows what is wrong with that audio file.

```
Error reading U:/School Homework/Fall 2023/DSC680-Applied Data Science/Project1/genres_original\jazz\jazz.00054.wav: File forma
t b'\xcb\x15\x1e\x16' not understood. Only 'RIFF' and 'RIFX' supported.
```

- Is there anything else that was trying to be accomplished but couldn't? Yes, after the model randomly chose the audio files, I wanted to get the PNG file to pull in together. Unfortunately, I was not able to.

Though I have been able to answer the questions that I have thought about through this project, there is more to do to get better results. From the expectations from each project, I have come to learn that there is always room for improvement, and the results could get better each time with more time.

References

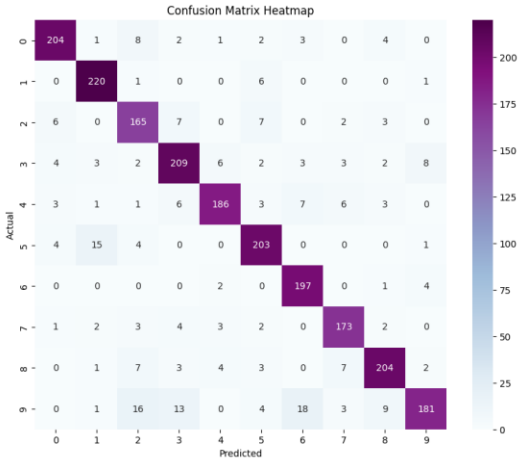
GTZAN Dataset - Music Genre Classification. (n.d.). Retrieved from https://www.kaggle.com: https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification

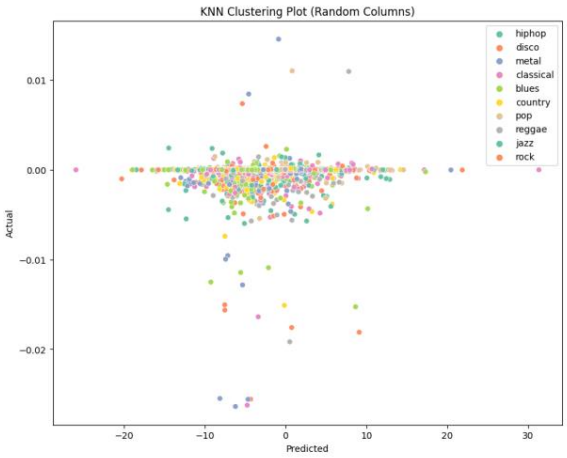
King, A. (2023, July 22). *Top Streaming Genres of 2023 So Far & Consumption Habits.* Retrieved from https://www.digitalmusicnews.com: https://www.digitalmusicnews.com/2023/07/22/top-streaming-genres-of-2023-so-far-and-consumption-habits/

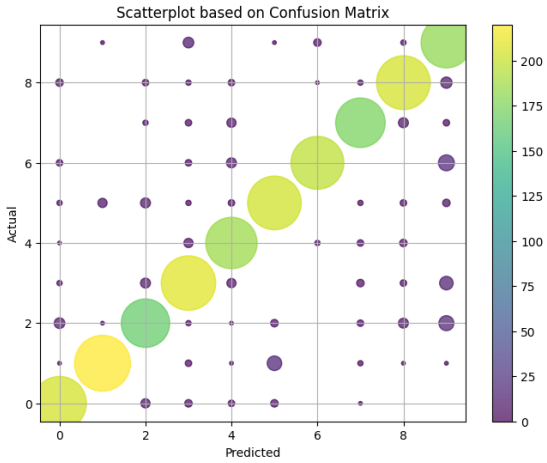
Music Classification: Beyond Supervised Learning, Towards Real-world Applications. (2023). Retrieved from https://music-classification.github.io: https://music-classification.github.io/tutorial/part1_intro/what-is-music-classification.html

Appendix

1: Plots from merged CSV Datasets







2: Correlation Matrixes for Audio Files