Abigail Villacorta

aev64

9 Dec, 2024

Github: <https://github.com/aev25/Genre-Detection>

Final Report

Project Definition:

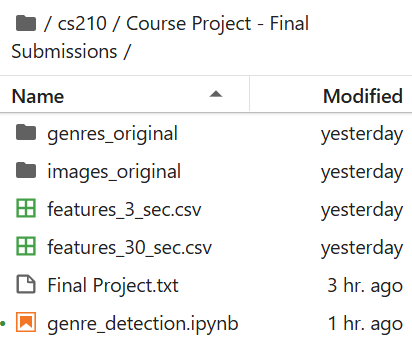
* This will try to classify songs by genre by using audio feature data. Applications like collection of content, specific music recommendations, and digital library management are supported by accurate genre classification. The ability of traditional classification methods to tell the difference between unique sound features is limited because they frequently rely on metadata like artist or song. This presents an analysis based on data to understand the unique characteristics that identify genres using audio features.
* The strategic aspects are:
* Gathering and selecting data: Use Python to prepare audio datasets for analysis.
* Transforming data: creating a data collection suitable for machine learning by removing significant characteristics from audio streams.
* Data management: Using SQL to store organized audio feature data for quick access and analysis.
* Machine Learning: Building and improving genre prediction models.
* These components relate to the lectures in the course, which include maintaining datasets with databases and Python, collecting real-world datasets, and analyzing data to find patterns.

Novelty and Importance:

* The constantly growing scale of streaming music data requires advanced techniques for suggestion and management. Compared to standard metadata-based techniques, audio-based classification offers deeper insights, improving music discovery and playlist creation.
* This project is both practically important and personally interesting because it blends a personal interest in music with the useful application of data analysis tools.
* Commonly found issues include insufficient data, inaccurate genre classifications created by users, and the lack of standard methods to categorize music that are unclear or cross-genre. Both accuracy and manageability are limited by these restrictions.
* The significance of auditory characteristics including tempo, pitch, and spectral characteristics in genre classification has been recognized by Music Information Retrieval (MIR) research. Our project explores useful applications in a database-driven environment while building on known strategies by including these ideas.

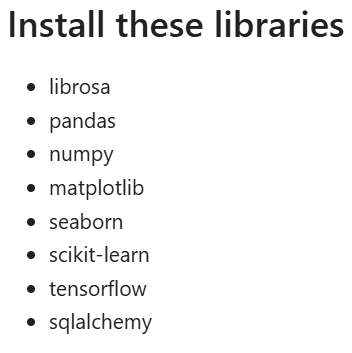
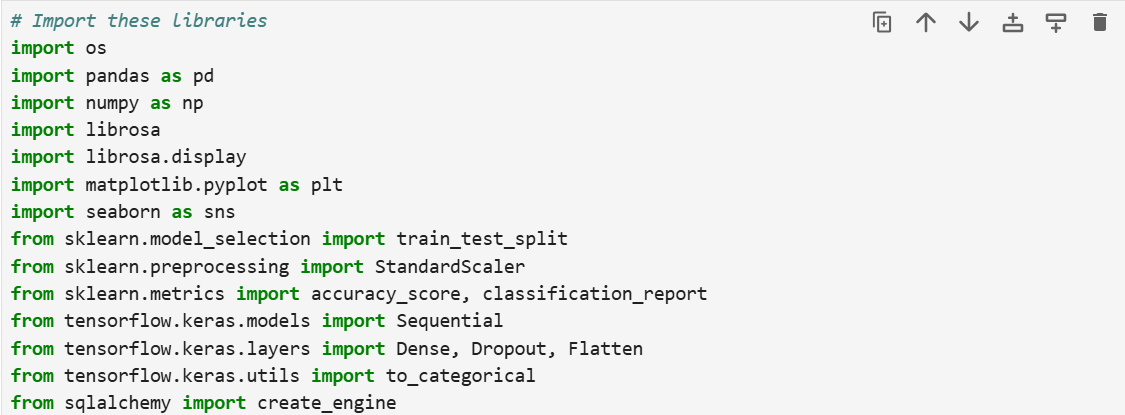
Progress and Contribution:

* Data Used:
* GTZAN Dataset - Music Genre Classification: from kaggle. <https://www.kaggle.com/datasets/andradaolteanu/gtzan-dataset-music-genre-classification/data>
* The dataset was downloaded directly after getting it straight from Kaggle. Python modules such as Librosa for audio feature extraction and Pandas for metadata handling were used to process and select the audio files.



Download the genres\_original folder, images\_original folder, features\_3\_sec.csv, and features\_30\_sec.csv. Just as displayed on the left.

* Python libraries, including Librosa, were gathered to extract features like tempo, rhythm, and spectral contrast. For machine learning, the selected features were cleaned up and defined. The original study was conducted using simple classifiers like Support Vector Machines (SVM) and K-Nearest Neighbors (KNN). Convolutional Neural Networks (CNNs) were used to effectively process sequential audio data by making use of their ability to identify complex patterns.

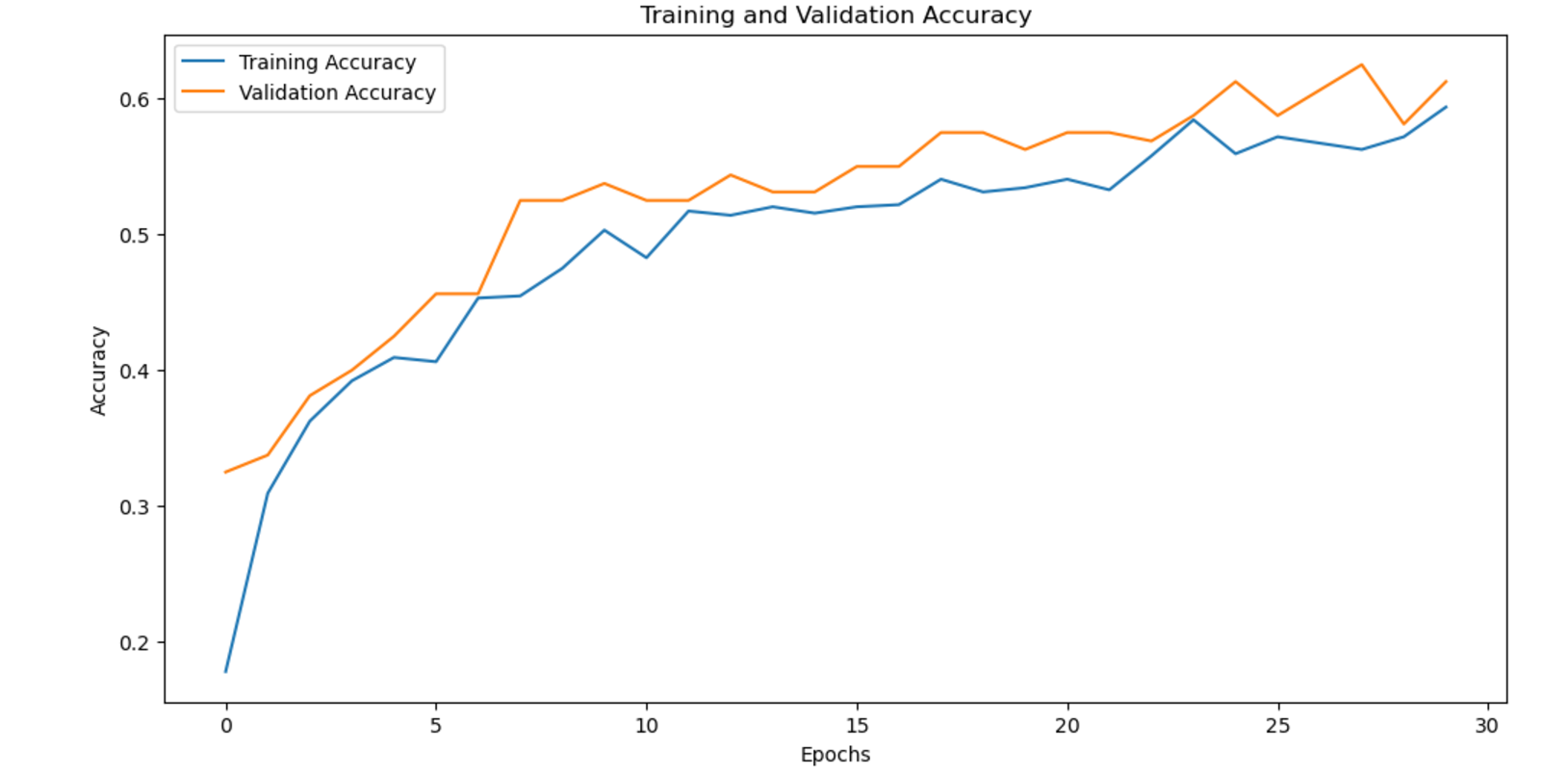
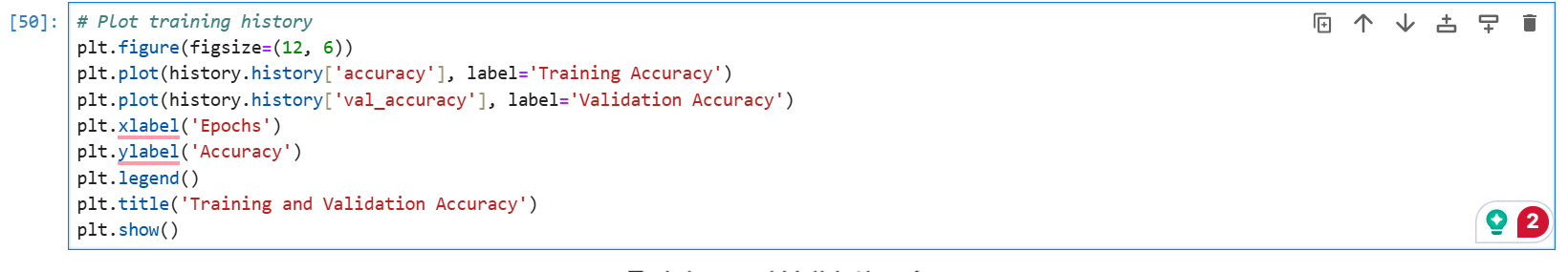
* A database with a relational structure, SQL, was used to store the changed and selected datasets for practical maintenance and analysis. The characteristics obtained were used to train the algorithms and determine accuracy, precision, and recalls being used to decide how well the predictions performed.
* MFCCs and other spectral properties were crucial to genre classification. The results proved the idea that audio feature-based categorization is more accurate and trustworthy than metadata-reliant techniques.
* Advantages and Limitations:
* Advantages: This method provides more flexibility and manages uncertain songs more skillfully by concentrating on audio aspects.
* Limitations: Among the challenges were the calculating requirements for the overlapping genres, which lowered accuracy.

Changes After Proposal:

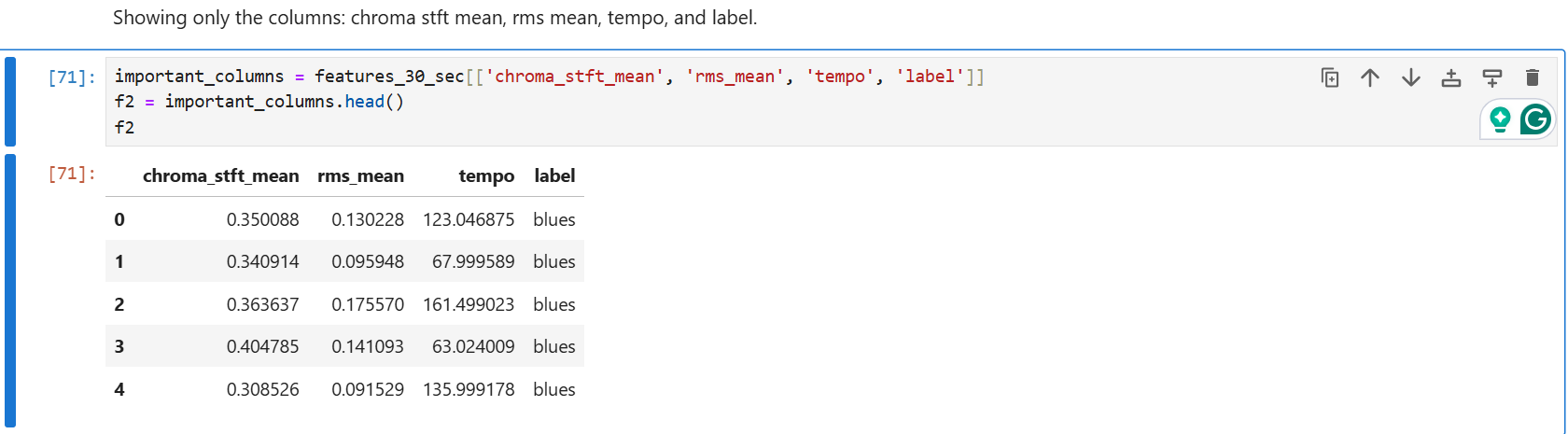
* At first, the project considered merging several databases, but concentrating only on GTZAN made things easier and maintained data format consistency. During feature extraction, other features were added, like spectral departure, which improved the accuracy of the model. CNN training and separating features took longer to process than expected. These improvements were necessary to strike a balance between performance and accuracy.

Data Visualization:

This code create a line plot that visualizes how the accuracy of the model changes over each epoch during training and validation.

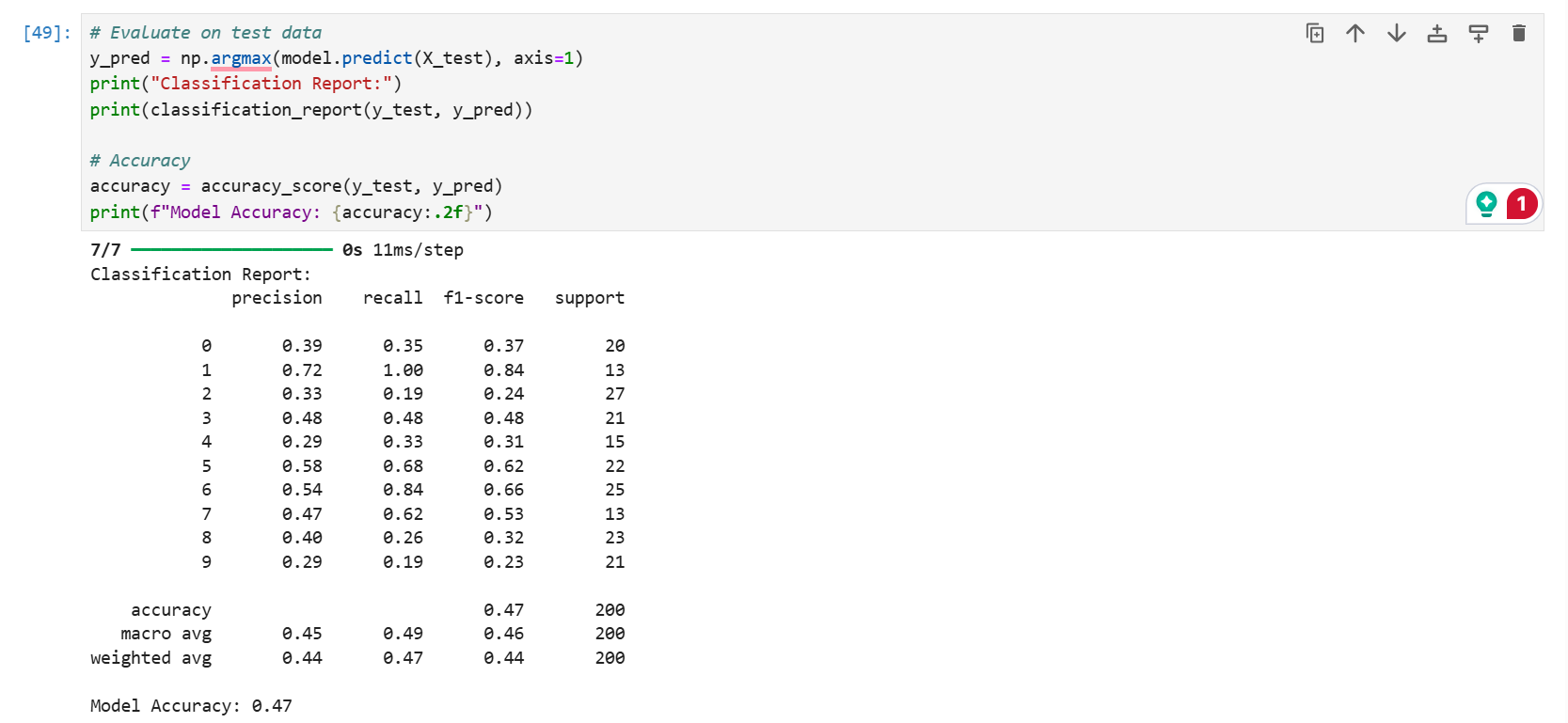


Data Preprocessing:



Machine Learning Model:

This code evaluates the trained model’s performance on the test dataset. It also provides detailed metrics to measure the model’s ability to predict data that has not been shown before.



CNN Training and Results:

This code shows a plot of the training and validation accuracy over epochs for the CNN model. This also includes the final accuracy and loss values.

