Wilfredo Gao, 9195-5311

CIS4930 Individual Coding Assignment 3

Spring 2023

1. Problem Statement

The objective of this assignment was to test recognition of emotions using audio files. Speech recognition has become an increasingly important machine learning problem especially with tonal languages where the same phrases can have different meanings dependent on user emotion when said. As such, it is important for a machine learning model to properly predict the emotions based upon speech to accurately convey the users' messages. The implications of such a model are imperative for greater voice recognition technology that can be used in a variety of fields such as mental health AI where patients would convey their messages. The model built for this assignment implements the Librosa library and passes these files into machine learning models to predict emotion.

2. Data Preparation

The original data folder contains four folders each pertaining to a different emotion which are: Happy, Fear, Angry, and Sad. The folder then contains audio files pertaining to these emotional tones. The files were iterated through and split into both a testing and training subset relatively 70 to 30 split. After data separation, graphs were generated during the exploratory data analysis phase to see how the audio data was behaving. Then after examining the graphs and seeing appropriate behavior, the features were extracted into a test and training array that contained the relative features of each subset. The data was then passed into the model development phase.

3. Model Development

○ Model Training

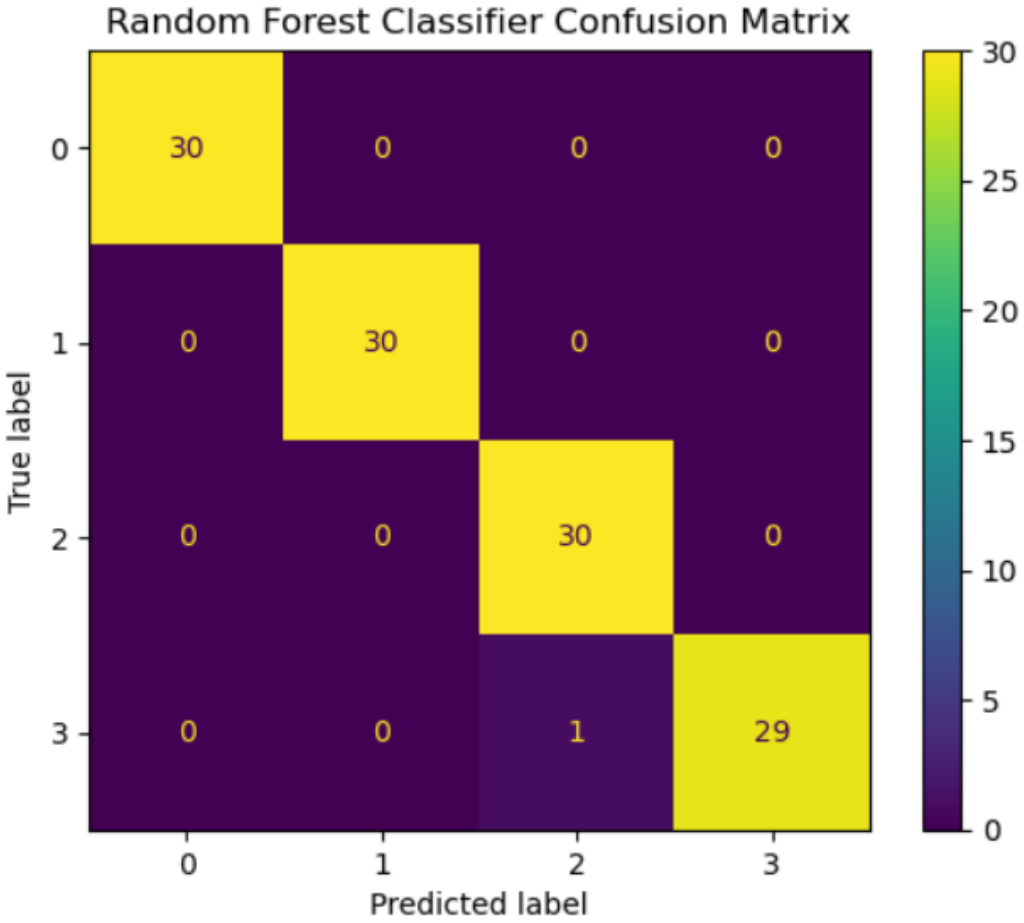
In the initial phase, the data was split into a training and testing subset for model development. These subsets were then extracted and implemented into two arrays: testFinal and trainFinal containing each relative dataset. For models, I selected Naives Bayes, Random Forest, Support Vector Machine, K-Nearest Neighbors, and Multi-Layer Perception for evaluation. Likewise, I experimented around with different parameters for model construction and settled upon those used in the code. I assigned Sad to be 1, Angry to be 2, Fear to be 3, and Happy to be 4.

○ Model Evaluation

I found that of all the models, Random Forest Classifier had the highest accuracy. This is likely because it was most suited for the data types in this data set.

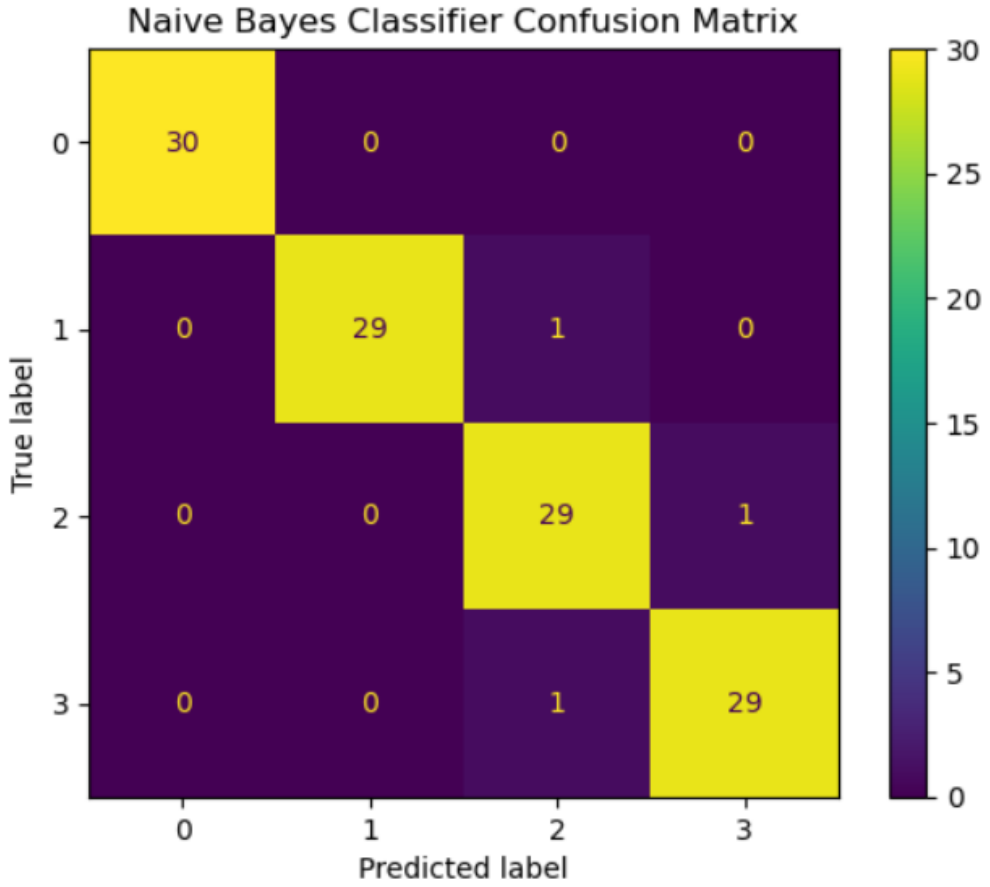
Random Forest Classifier Report

	precision	recall	f1-score	support
1	1.00	1.00	1.00	30
2	1.00	1.00	1.00	30
3	0.97	1.00	0.98	30
4	1.00	0.97	0.98	30
accuracy			0.99	120
macro avg	0.99	0.99	0.99	120
weighted avg	0.99	0.99	0.99	120



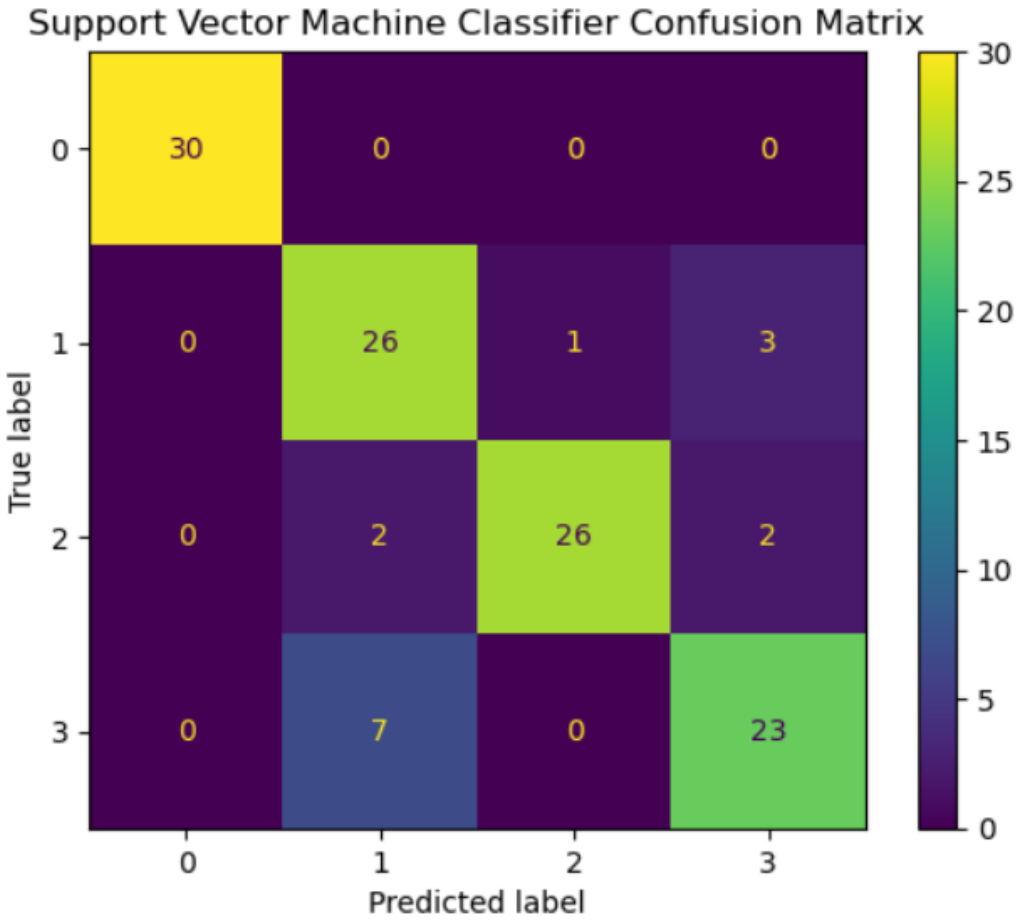
Naive Bayes Classifier Report

	precision	recall	f1-score	support
1	1.00	1.00	1.00	30
2	1.00	0.97	0.98	30
3	0.94	0.97	0.95	30
4	0.97	0.97	0.97	30
accuracy			0.97	120
macro avg	0.98	0.98	0.98	120
weighted avg	0.98	0.97	0.98	120



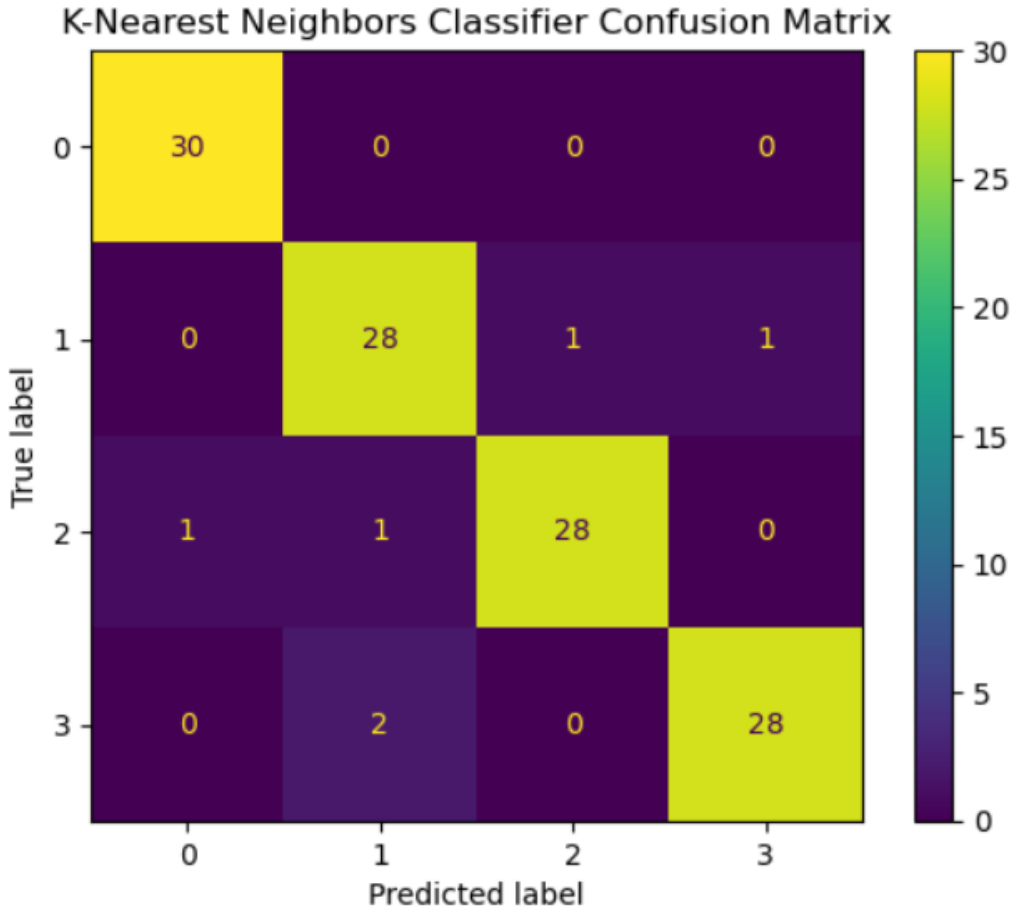
Support Vector Machine Classifier Report

	precision	recall	f1-score	support
1	1.00	1.00	1.00	30
2	0.74	0.87	0.80	30
3	0.96	0.87	0.91	30
4	0.82	0.77	0.79	30
accuracy			0.88	120
macro avg	0.88	0.88	0.88	120
weighted avg	0.88	0.88	0.88	120



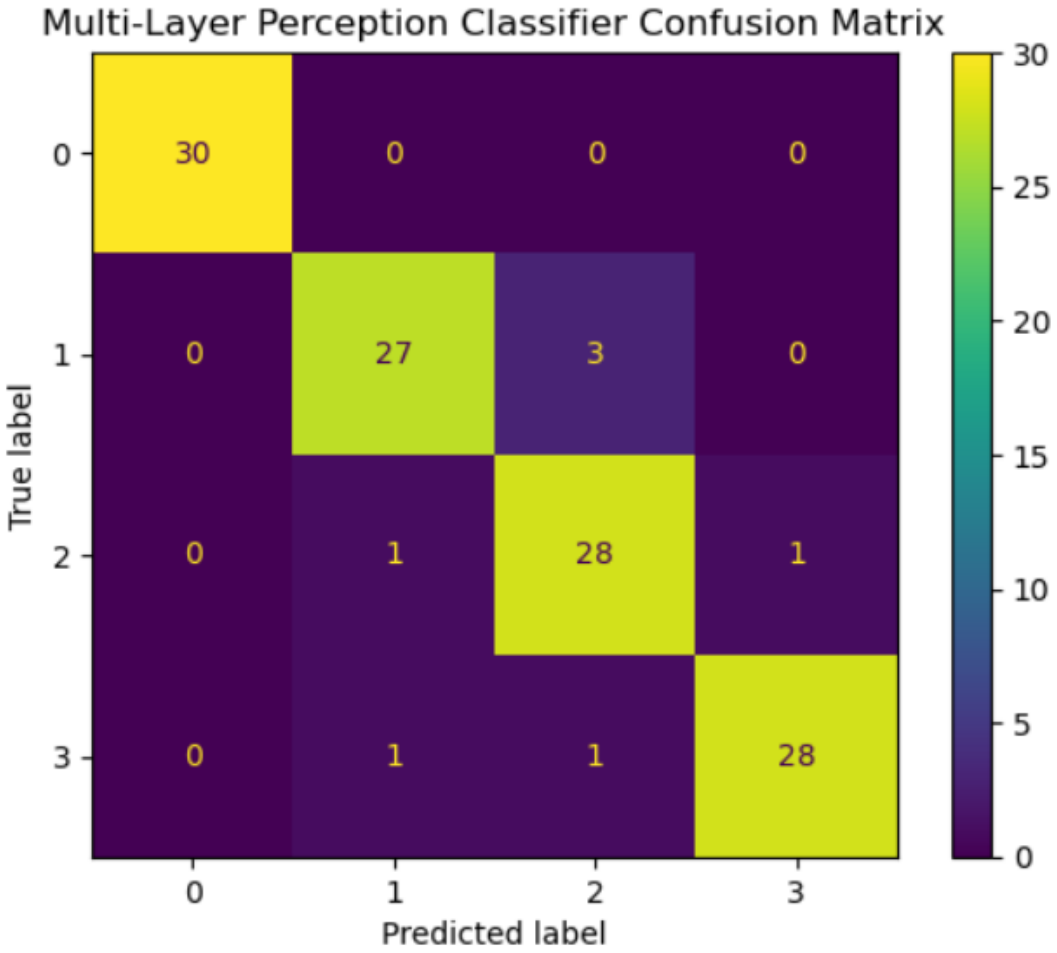
K-Nearest Neighbors Classifier Report

	precision	recall	f1-score	support
1	0.97	1.00	0.98	30
2	0.90	0.93	0.92	30
3	0.97	0.93	0.95	30
4	0.97	0.93	0.95	30
accuracy			0.95	120
macro avg	0.95	0.95	0.95	120
weighted avg	0.95	0.95	0.95	120



Multi-Layer Perception Classifier Report

	precision	recall	f1-score	support
1	1.00	1.00	1.00	30
2	0.93	0.90	0.92	30
3	0.88	0.93	0.90	30
4	0.97	0.93	0.95	30
accuracy			0.94	120
macro avg	0.94	0.94	0.94	120
weighted avg	0.94	0.94	0.94	120



4. Discussion

○ *My model performed better than I thought it would with very high accuracy. I think this is because the dataset was appropriately and accurately split for the model to train on with the 70-30 split.*

○ *The challenges I met during data preparation was how to properly split the dataset and post processing features. The experience was sort of learning as I go and fixing any errors that would occur.*

○ *I found this assignment to be really interesting as this was the first time I was working on an audio modality. The concept was fun to learn about in lecture and getting the opportunity to test it out during this assignment was unique. It was an interesting experience to go about feature extraction of the audio files and the complexity of the process. I think this form of modality is extremely important to consider for building machine learning models. Thank you for the opportunity and the awesome semester! Wishing you all the best!*

5. Appendix

- *Github link: https://github.com/wilfredogao/CIS4930-Assignment03*