Dhruv Verma 2017046

Homework-2: **Report**

(MFCC + SVM)

1. Clean Data

	precision	recall	f1-score	support
Class 0	0.86	0.90	0.88	260
Class 1	0.80	0.74	0.77	230
Class 2	0.72	0.78	0.75	236
Class 3	0.82	0.78	0.80	248
Class 4	0.90	0.86	0.88	280
Class 5	0.77	0.80	0.78	242
Class 6	0.86	0.89	0.88	262
Class 7	0.83	0.80	0.82	263
Class 8	0.82	0.80	0.81	243
Class 9	0.75	0.77	0.76	230
accuracy			**0.81**	2494
macro avg	0.81	0.81	0.81	2494
weighted avg	**0.82**	**0.81**	**0.81**	**2494**

2. Augmented Data

	precision	recall	f1-score	support
Class 0	0.94	0.88	0.91	260
Class 1	0.77	0.80	0.78	230
Class 2	0.74	0.82	0.78	236
Class 3	0.84	0.79	0.82	248
Class 4	0.90	0.89	0.90	280
Class 5	0.80	0.82	0.81	242
Class 6	0.88	0.87	0.87	262
Class 7	0.86	0.87	0.87	263
Class 8	0.87	0.88	0.87	243
Class 9	0.80	0.77	0.78	230
accuracy			**0.84**	2494
macro avg	0.84	0.84	0.84	2494
weighted avg	**0.84**	**0.84**	**0.84**	**2494**

Dhruv Verma 2017046

(Spectrogram + SVM)

1. **Clean Data**

	precision	recall	f1-score	support
Class 0	0.77	0.41	0.53	260
Class 1	0.66	0.51	0.58	230
Class 2	0.68	0.44	0.53	236
Class 3	0.51	0.41	0.45	248
Class 4	0.80	0.38	0.52	280
Class 5	0.56	0.25	0.35	242
Class 6	0.22	0.93	0.35	262
Class 7	0.72	0.41	0.53	263
Class 8	0.65	0.40	0.49	243
Class 9	0.65	0.45	0.53	230
accuracy			**0.46**	2494
macro avg	0.62	0.46	0.49	2494
weighted avg	**0.62**	**0.46**	**0.49**	**2494**

2. **Augmented Data**

	precision	recall	f1-score	support
Class 0	0.87	0.67	0.75	260
Class 1	0.74	0.68	0.71	230
Class 2	0.69	0.72	0.70	236
Class 3	0.68	0.69	0.68	248
Class 4	0.85	0.67	0.75	280
Class 5	0.78	0.55	0.65	242
Class 6	0.55	0.89	0.68	262
Class 7	0.79	0.73	0.76	263
Class 8	0.65	0.83	0.73	243
Class 9	0.72	0.68	0.70	230
accuracy			**0.71**	2494
macro avg	0.73	0.71	0.71	2494
weighted avg	**0.73**	**0.71**	**0.71**	**2494**

Dhruv Verma 2017046

Classification Pipeline

Audio Data: Data imported through **librosa** library

Data Augmentation: Noise is added to the audio signal according to the following formula

$$Noisy\ Signal(t) = Signal(t) + \alpha \times Noise_i(t)$$
$$where\ i = randomSelect\ (Noise\ Database),\ \alpha = 0.1$$

Feature Extraction: Time-frequency features like **Spectrograms** and **MFCC** extracted.

PCA: Principal Component Analysis was used to reduce the dimensionality of the feature vector produced by flattening the Spectrogram and MFCC matrices. The value of 'k' was chosen so as to ensure 'explained_variation_ratio' ~ 0.9 (We don't end up losing a lot of info.)

SVM Classifier: Support Vector Machine with **rbf** kernel used. The hyperparameters ('C' and 'Gamma') were tuned using GridSearchCV. **Sklearn** library was used to implement SVM.

Analysis of Results

- The accuracy results are a clear signal of MFCC being a better representative feature for the audio signals.
- The data is balanced, hence we see similar values for accuracy and weighted-F1 score.
- Data augmentation makes the algorithm more robust towards the noise, hence we can see increase in performance metrics.
- The spectrogram is quite sensitive to noise and hence, with data augmentation shows significant improvement in results.