Indian Institute of Technology Delhi

ELL784 Introduction to Machine Learning: Assignment 3

Deepak: 2019MT10685

Pre-processing the data

The training data provided consists of 7352 rows and 563 columns, i.e, 562 features and 1 column that contains the classification label that includes the labels ['LAYING', 'SITTING', 'STANDING', 'WALKING', 'WALKING_DOWNSTAIRS', 'WALKING_UPSTAIRS']

Handling missing values or duplicates

The data did not contain any missing values or duplicates.

Normalisation and Feature Scaling

It is essential to normalise or scale the data before feeding it into the SVM because without scaling, large values can introduce bias in the results. Thus the features are usually scaled down to the range [-1,1] or [0,1].
We used the normalise method from the preprocessing library of sklearn and applied MinMax scaling on top of it.Thus, the range off feature values was first brought down to [-1,1] and then to [0,1].

Preparing Validation Set

We have split the training data and used 80% for training the model and remaining 20% for validation

Various Kernels in SVM

Four kernels were tried out on the training data, Linear, RBF, Polynomial, and Sigmoid.

Linear Kernel

Linear Kernel is used when the data is Linearly separable.This kernel comes the hyper-parameter C that governs regularisation. Changing C in multiples of 10, it was observed that accuracy is best when C = 100.
Best Accuracy Score obtained on testing data is 0.9566

Classification Report on Test Data:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	537
1	0.97	0.89	0.93	491
2	0.91	0.98	0.94	532
3	0.92	0.99	0.96	496
4	1.00	0.94	0.97	420
5	0.96	0.93	0.94	471
accuracy			0.96	2947
macro avg	0.96	0.96	0.96	2947
weighted avg	0.96	0.96	0.96	2947

Figure 1: *Classification Report for Linear Kernel*

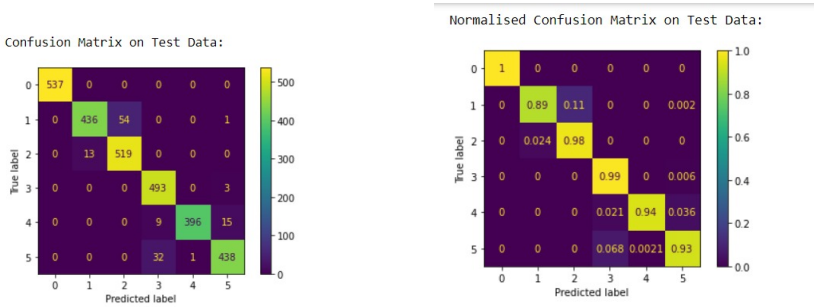

Figure 2: *Confusion Matrix for Linear Kernel*

RBF Kernel

RBF or Radial Basis Functions were also used as a kernel. This Kernel is used to perform transformation, when there is no prior knowledge about data. On tuning the hyper parameters: gamma (governs curvature of the separation boundary) and C, relatively better performance was observed at C = 100 and gamma = 0.02.
Best Accuracy Score obtained on testing data is 0.9552

Classification Report on Test Data:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	537
1	0.98	0.88	0.92	491
2	0.90	0.98	0.94	532
3	0.91	1.00	0.95	496
4	1.00	0.94	0.97	420
5	0.96	0.93	0.95	471
accuracy			0.96	2947
macro avg	0.96	0.95	0.95	2947
weighted avg	0.96	0.96	0.96	2947

Figure 3: *Classification Report for RBF kernel*

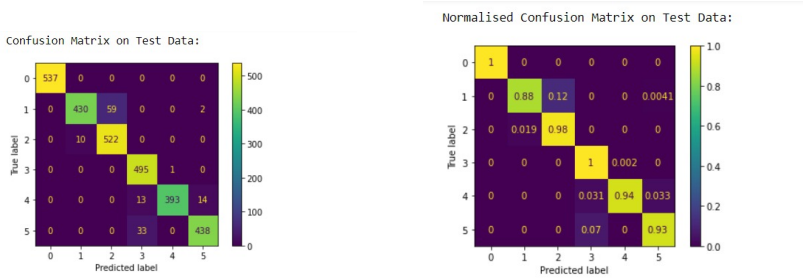

Figure 4: *Confusion Matrix for RBF Kernel*

Polynomial Kernel

Polynomial kernel represents the similarity of vectors in training set of data in a feature space over polynomials of the original variables used in kernel. Polynomial kernel comes with an extra parameter of degree. On tuning the hyper parameters: gamma, degree and C, relatively better performance was observed at gamma=0.03, degree=3 and C=1.
Best Accuracy Score obtained on testing data is 0.9545

Classification Report on Test Data:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	537
1	0.96	0.88	0.92	491
2	0.90	0.97	0.93	532
3	0.92	1.00	0.96	496
4	0.99	0.93	0.96	420
5	0.97	0.94	0.95	471
accuracy			0.95	2947
macro avg	0.96	0.95	0.95	2947
weighted avg	0.96	0.95	0.95	2947

Figure 5: *Classification Report for Polynomial Kernel*

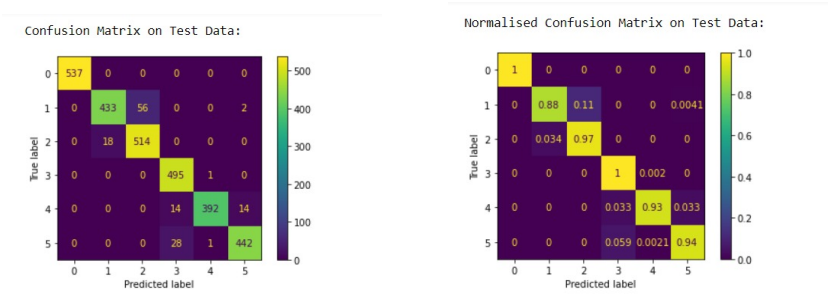

Figure 6: *Confusion Matrix for Polynomial Kernel*

Sigmoid Kernel

Sigmoid kernel is equivalent to a two-layer, perceptron model of neural network, which is used as activation function for artificial neurons.Sigmoid kernel works on the parameters C and gamma. The sigmoid kernel provides poor accuracy scores when compared to other kernels except when C = 100 and gamma = 0.001 at which the score is comparable to the rest.
Best Accuracy Score obtained on testing data is 0.9491

Classification Report on Test Data:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	537
1	0.94	0.88	0.91	491
2	0.90	0.95	0.92	532
3	0.93	1.00	0.96	496
4	0.99	0.92	0.96	420
5	0.95	0.94	0.94	471
accuracy			0.95	2947
macro avg	0.95	0.95	0.95	2947
weighted avg	0.95	0.95	0.95	2947

Figure 7: *Classification Report for Sigmoid Kernel*

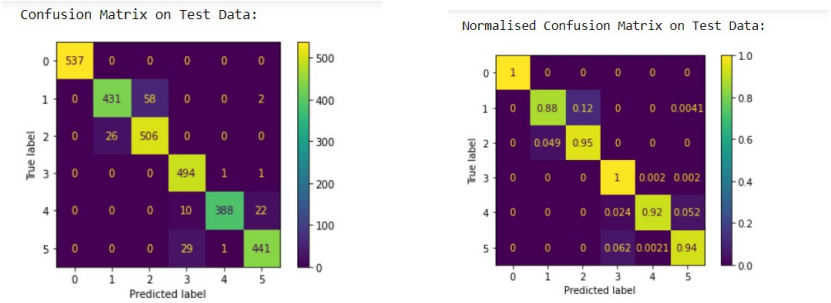

Figure 8: *Confusion Matrix for Sigmoid Kernel*

Comparison Between various Kernels used

The first three kernels give almost equal performances with Linear Kernel performing slightly better compared to the other two. A possible explanation for better results using Linear can be the fact that training data is linearly separable. Also, RBF is a commonly used kernel, used when no prior information is known about the data. However, ultimately it is the data that decides which kernel suits it the best.

Model Evaluation

The accuracy score on test data was our prime performance indicator. Accuracy scores on training data and validation scores were also evaluated. We have also compared the results by plotting the confusion matrix and computing the precision, recall, f1 scores for different trials.