ELL 784 : Introduction to Machine learning

Assignment 2:

Name: Mahadevanna.u

Entry number: 2021AMA2081

Assignment 2

Studying the effect of hyperparameters on binary classifier

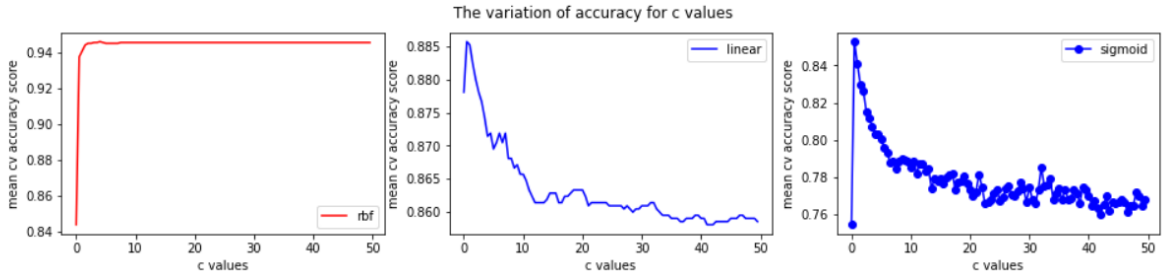
3 binary classifier cases of Two (2) and seven (7), eight(8) and nine(9) and (five(5) and six(6) were studied and their performance with different hyperparameters was tracked.

The observations are as follows,

- When values of **c** increased gradually, Training accuracy increased; However Test accuracy dropped. Maybe high values of c lead to overfitting of data.
- When very low values of C values were used, training error decreased but there was a slight rise in the testing accuracy.
- **Regarding kernels:** For binary classification sigmoid, rbf kernels worked better than the linear, polynomial kernels.

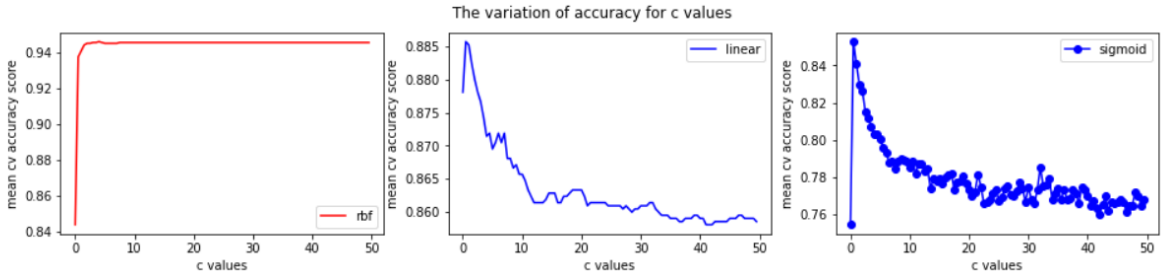
From the below figures we can clearly observe that accuracy increases till some value of c and there after accuracy drops with magnitude of C.

The variation of accuracy for the c value ranging from 0.01 to 20 was plotted in the figures below.



ACCURACY VARIATION WITH C VALUES FOR THE DIFFERENT KERNEL WITH 10 FEATURES

Observation: The accuracy of model with 10 features reached the max value of 99 , where as it's 98.5 for model with 25 features. Accuracy of model dropped quickly with c values for the dataset with 25 features. Overall, RBF kernel performed well for most of the datasets.



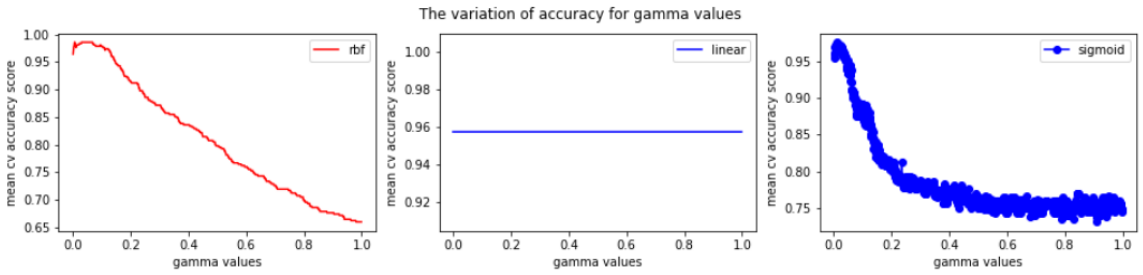
ACCURACY VARIATION WITH C VALUES FOR THE DIFFERENT KERNEL WITH 25 FEATURES

Assignment 2

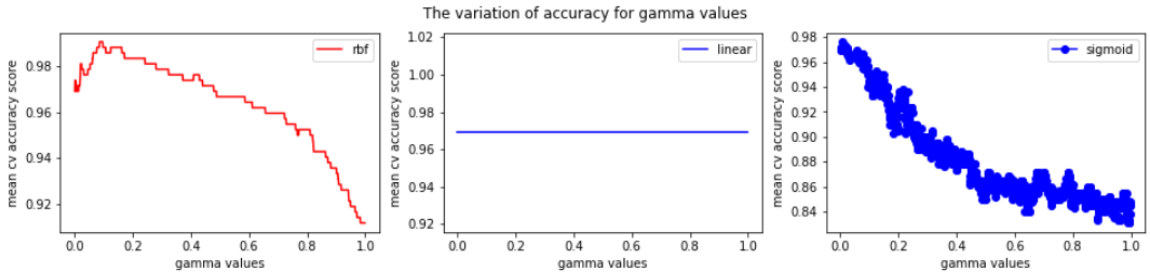
Again, accuracy was plotted against the gamma hyperparameter for three kernels rbf,linear,sigmoid.

Linear kernels showed no dependency on the gamma value which is evident in the figure below.

And for rbf kernel values, accuracy increased to a certain range and later dropped with increase in the gamma values. As gamma value increases , the data consideration range increases, which could hint the sign of overfitting of the data. So for all the kernels, testing accuracy decreased with gamma value after a certain range.



Variation of accuracy with gamma values with 10 features



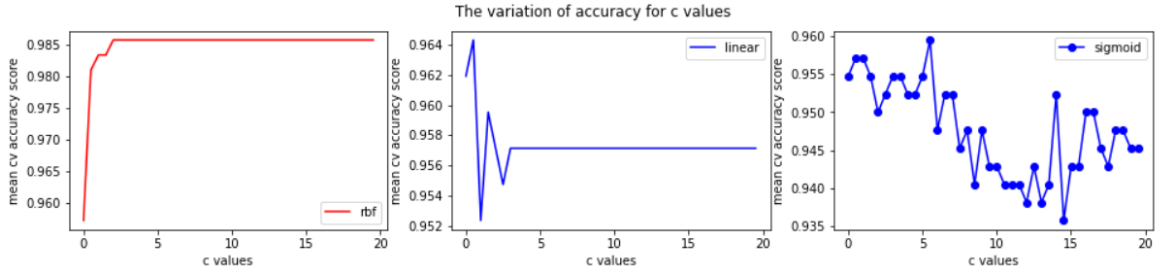
Variation of accuracy with gamma values with 25 features

Observations: as observed with c values, gamma value analysis followed the same trend. Availability of high number of features didn't improve the performance of the model; max accuracy found was lesser than of 10 features model. More prone to overfitting compared to model with less features.

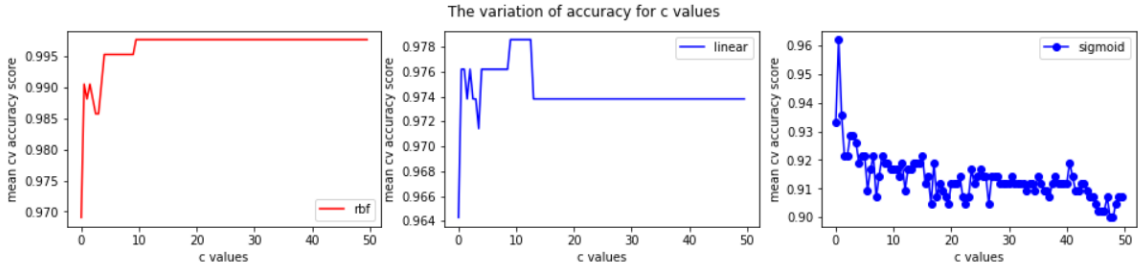
Assignment 2

Part 2: Binary classifiers

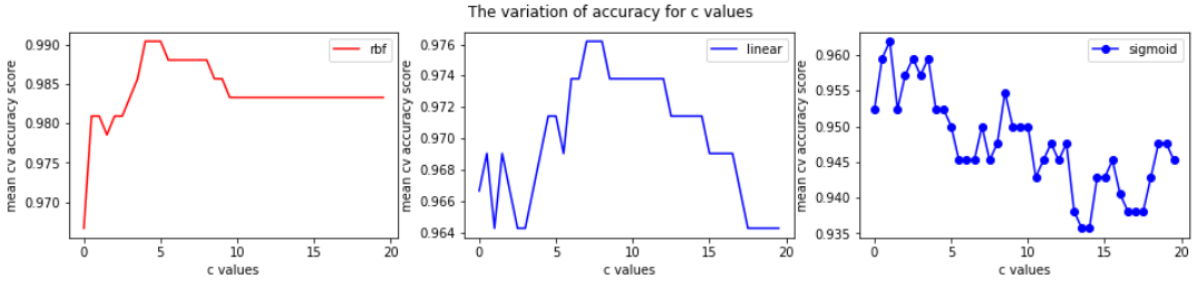
For binary classifiers various classes were tried with svm and results were analysed.



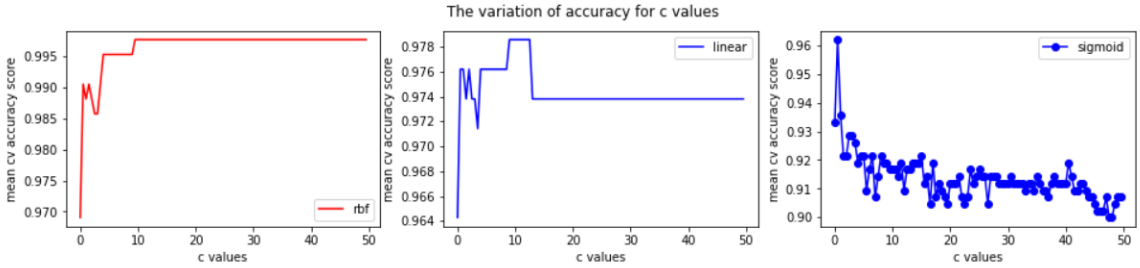
Classes 8 and 9 accuracy plot with 10 features



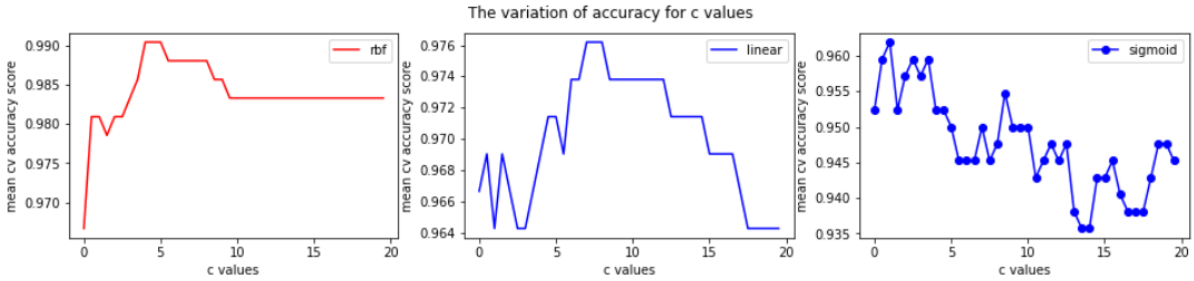
Classes 8 and 9 accuracy plot with 25 features



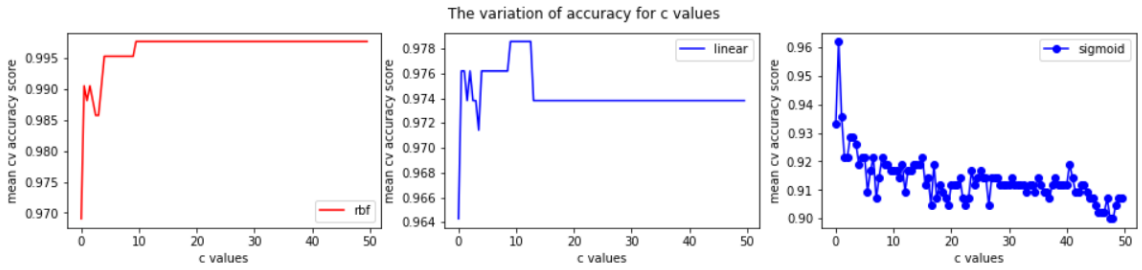
Classes 2 and 7 with 10 features



Assignment 2



Classes 5 and 6 accuracy plot with 10 features



Classes 5 and 6 accuracy plot with 25 features

Observations on binary classifiers:

- **It was found that different classes needed different hyperparameters for their best performance.**
- **Test accuracy varied significantly on based on the channels.**
- **Rbf kernel worked the best with most of the classes.**
- **Classes 5 and 6 performed compare to other classes.**
- **When 10 features model was used, good accuracy was obtained than using 25 features and hyperparamters were different for different features.**
- **Higher values of C were required for full dataset model to reach good accuracy.**
- **For classes like 2 and 5, which have lot of differences required few features(10) to achieve good accuracy, as it performed poorly with 25 features.**

Assignment 2

Sl.no	Model description	Kernel	C value	Gamma value
1	Binary classifier+classes (2,7)	rbf	100	0.01
2	Binary classifier+classes (8,9)	rbf	11	.1
3	Binary classifier+classes (5,6)	rbf	1	.1
4	Binary classifier+classes (4,8)	Linear	1	1
5	Binary classifier+classes (2,7) +full features	sigmoid	12	.1
6	Binary classifier+classes (8,9) +full features	rbf	11	0.01
7	Binary classifier+classes (8,4) +full features	rbf	1	.1

From the table, above we can see how hyperparameters changed with the classes we are training.

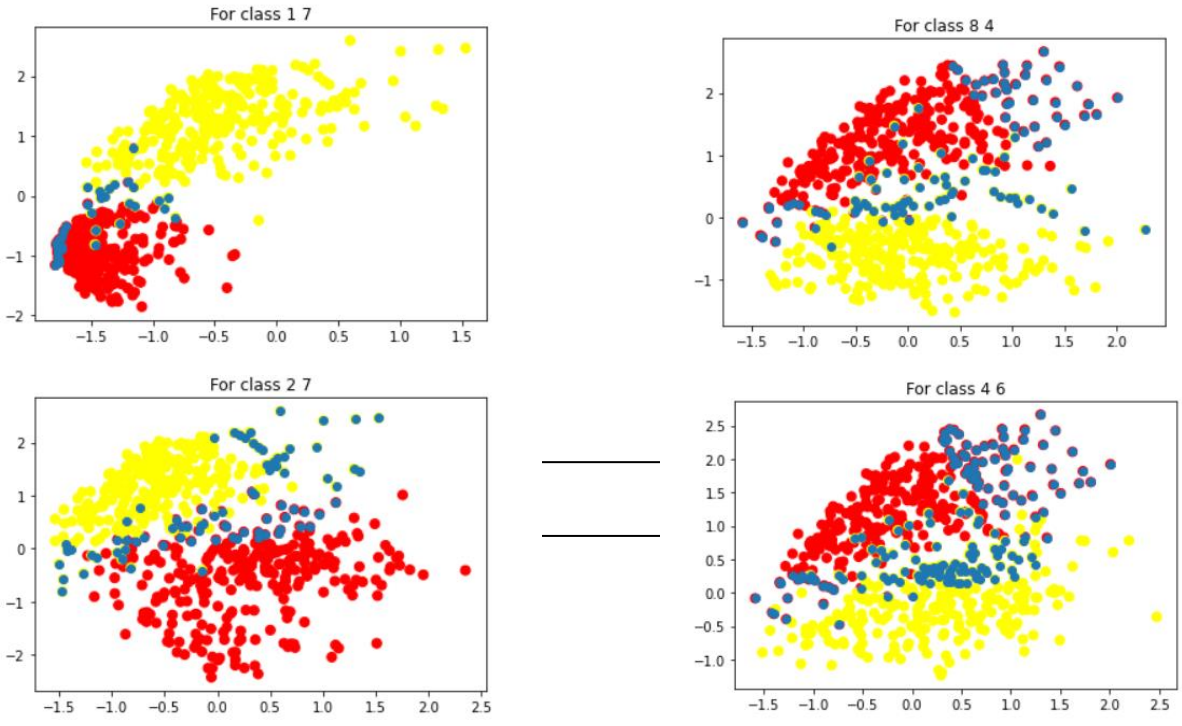
Most of the cases, rbf kernel performed well.

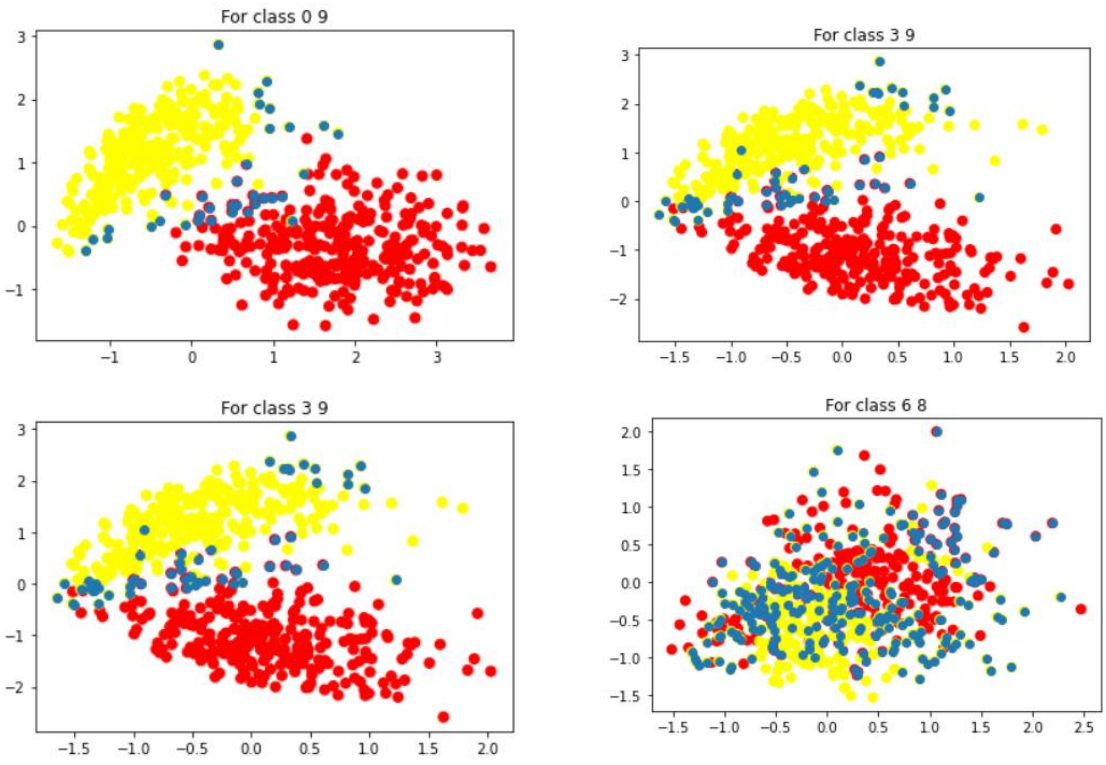
Next question to be answered: Do you consistently get the best results for the same hyperparameters settings, or does it vary a lot depending on on which pair of classes you're looking at?

Answer: It's not consistent, best parameters were changing as the classes trained were changed.

Kernels remained same as radial basis funcition('rbf') and gamma was also found consistent with value of 0.01, but C value was changing significantly with the type of classes chose for training.

Data visualization with first 2 features





Observation: for different binary classes, data visualization is done using first 2 features of the data.

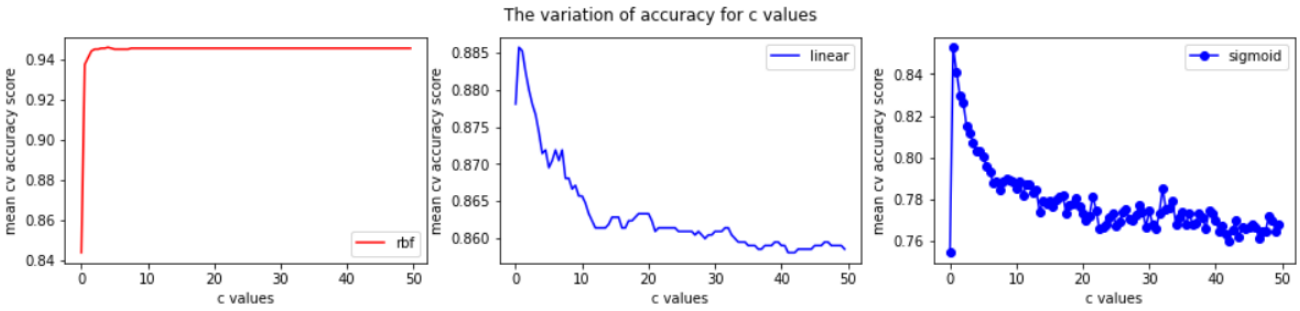
Along with the data support vectors are plotted after training the model.

We can clearly see that for the classes 6 and 8, most of the data points seems to be overlapping. Support vectors are not clearly distinguishable in that figure. Here, clearly more number of dimensions are required to separate the data.

For classes 8 and 4, we can clearly see a good gap between the classes and support vectors are not overlapping as we saw in the previous case. For these kind of models two features would be enough in the case of frugal condition.

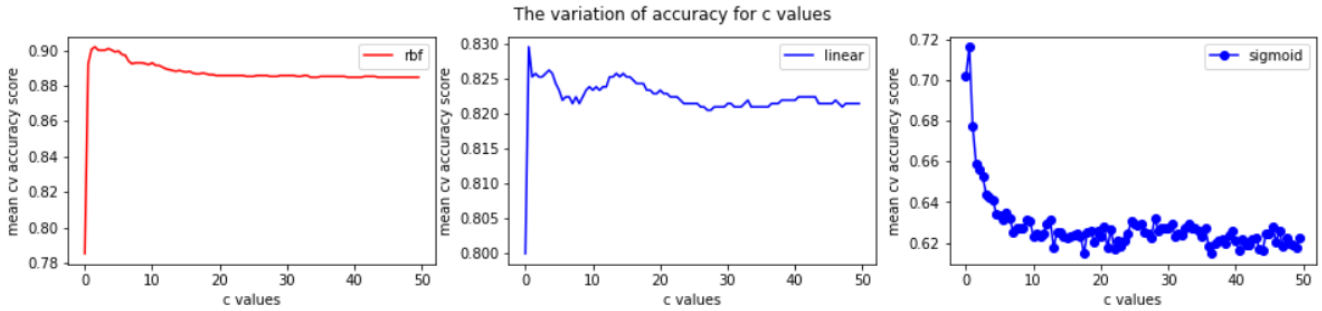
Question 1: **Last segment**

Here multiclass classifier was trained using all the classes and results are analysed.



Variation of accuracy of the multi class classifier with 10 features

Assignment 2



Variation of accuracy of the multi class classifier with 25 features

Observations: Test accuracy clearly dropped when we trained the model with data set of 25 features compared to previous dataset with 10 features. Initially model reached the highest accuracy of 95%, where as model's accuracy dropped to 90 percent, when full dataset was used to train the model. C value remained almost same as of the binary classifier. It also followed the same trend of binary classifier in variation of accuracy with c values.

Assignment 2

Assignment 2

Assignment 2