***ELL409***

***Assignment 2 Report***

***-Rajdeep Das***

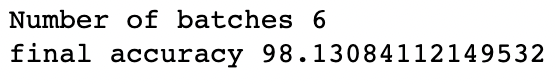
***2019MT10718***

***PART 1 A***

Binary Classification:

Classes chosen 0 and 1

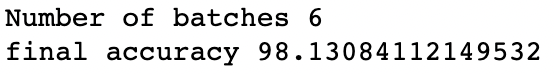
Cross validation performed using 6 batches gave output as 98.43 percent accuracy for all 25 features using default parameter settings (i.e radial basis function kernel).



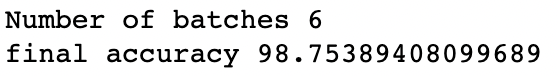
Parameter tuning:

-Changing kernel type:

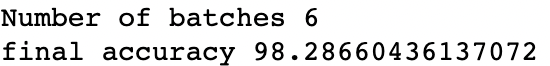
Linear kernel gives result as



Polynomial kernel gives result as



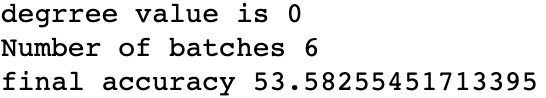
Sigmoid kernel gives result as

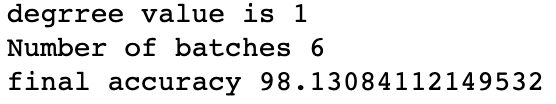


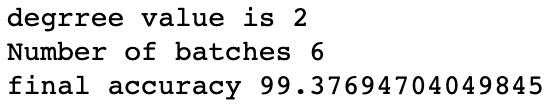
So we see that polynomial kernel gives the best yet accuracy on 5-batch cross validation by 98.75 percent.

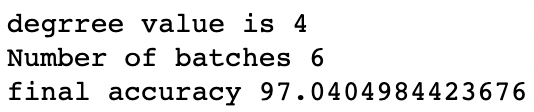


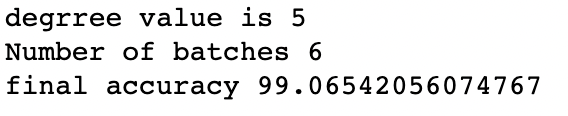
-changing degree (default value is 3):

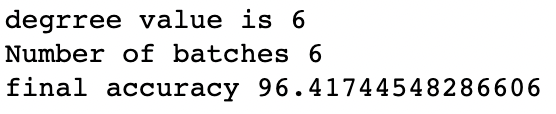










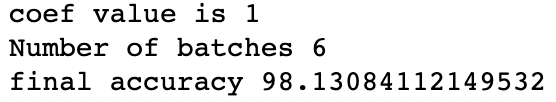


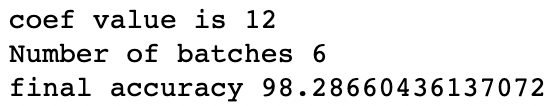
We see that the best result is obtained for a degree of 2 with 99.37 percent accuracy.

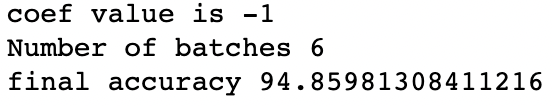
-changing gamma(default 0):

By changing gamma it was observed that there was no significant improvement in the accuracy and the values were more or less same.

-changing coef0 (default is 0):



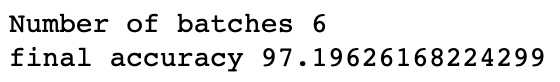




It was seen that accuracy was the best for the default value of 0 only.

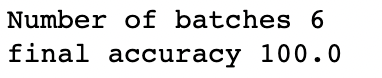
->Using 10 features instead of 25:

Under the same parameters the accuracy drops to 97.19% when the number of features are reduced.



Classes chosen 3 and 7:

Cross validation performed using 6 batches gave output as 100% accuracy for all 25 features using default parameter settings (i.e radial basis function kernel and degree as 3 ,coef0 as 0 and constant as 1).

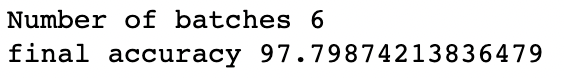


Parameter tuning:

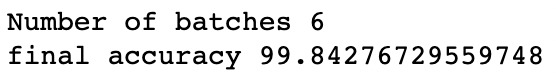
Although it gives a 100% cross validation accuracy we still perform parameter tuning for experimental purposes.

-Changing kernel type:

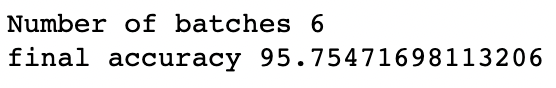
Linear kernel gives result as



Polynomial kernel gives result as



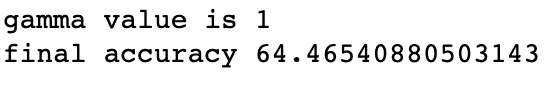
Sigmoid kernel gives result as

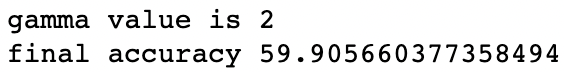


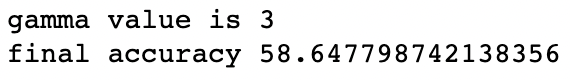
So we see that rbf kernel gives the best accuracy on 5-batch cross validation at 100 percent.



-changing gamma(default 1/number of features):



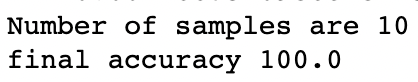




We see that accuracy decreases with value of gamma so will keep it default value

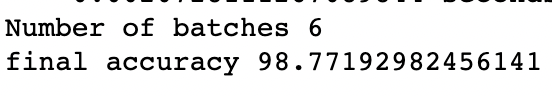
->Using 10 features instead of 25:

Under the same parameters the accuracy stays the same when the number of features are reduced (might suggest some redundancy of features).



Classes chosen 2 and 9

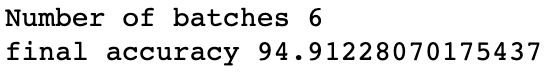
Cross validation performed using 6 batches gave output as 98.77 percent accuracy for all 25 features using default parameter settings (i.e radial basis function kernel).



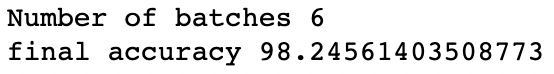
Parameter tuning:

-Changing kernel type:

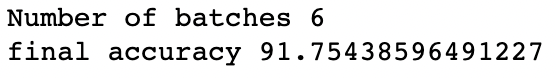
Linear kernel gives result as



Polynomial kernel gives result as



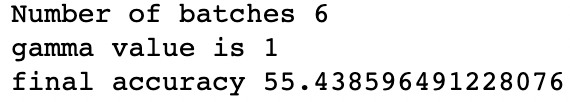
Sigmoid kernel gives result as



So we see that rbf kernel gives the best yet accuracy on 5-batch cross validation by 98.77 percent.



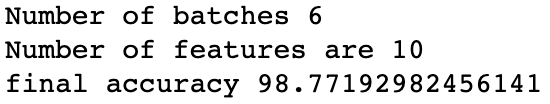
-changing gamma(default 0):



By changing gamma further it was observed that there was no significant change in the accuracy and the values were more or less same. So default of 0 is optimal.

->Using 10 features instead of 25:

Under the same parameters the accuracy stays the same when the number of features are reduced (might suggest some redundancy of features).

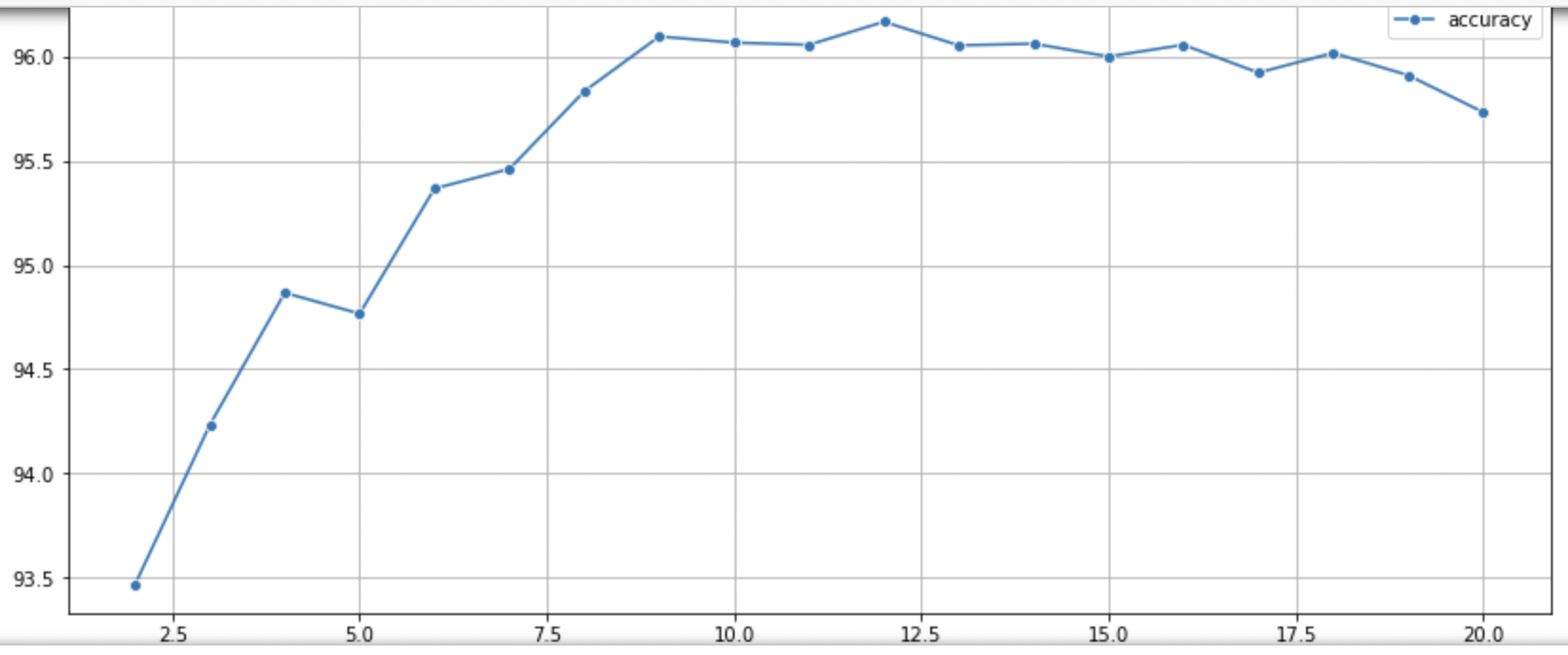


Observations:

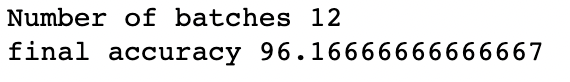
Hyperparameters like kernel function and other constants are seen to depend on the pair of classes chosen and not independent of them. Also in some classes some features are found to be redundant as reducing from 25 to 10 retains the accuracy rate.

Multiclass Classification:

Accuracy plotted as a function of cross validation batches and it is seen that maximum cross validation accuracy comes for 12 batches.



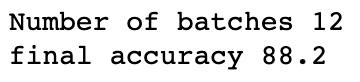
Cross validation performed using 12 batches gave output as 90.73 percent accuracy for all 25 features using default parameter settings (i.e radial basis function kernel).



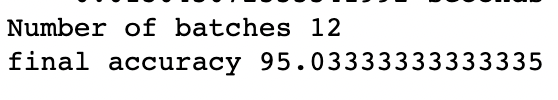
Parameter tuning:

-Changing kernel type:

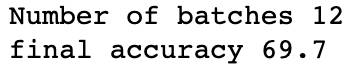
Linear kernel gives result as



Polynomial kernel gives result as



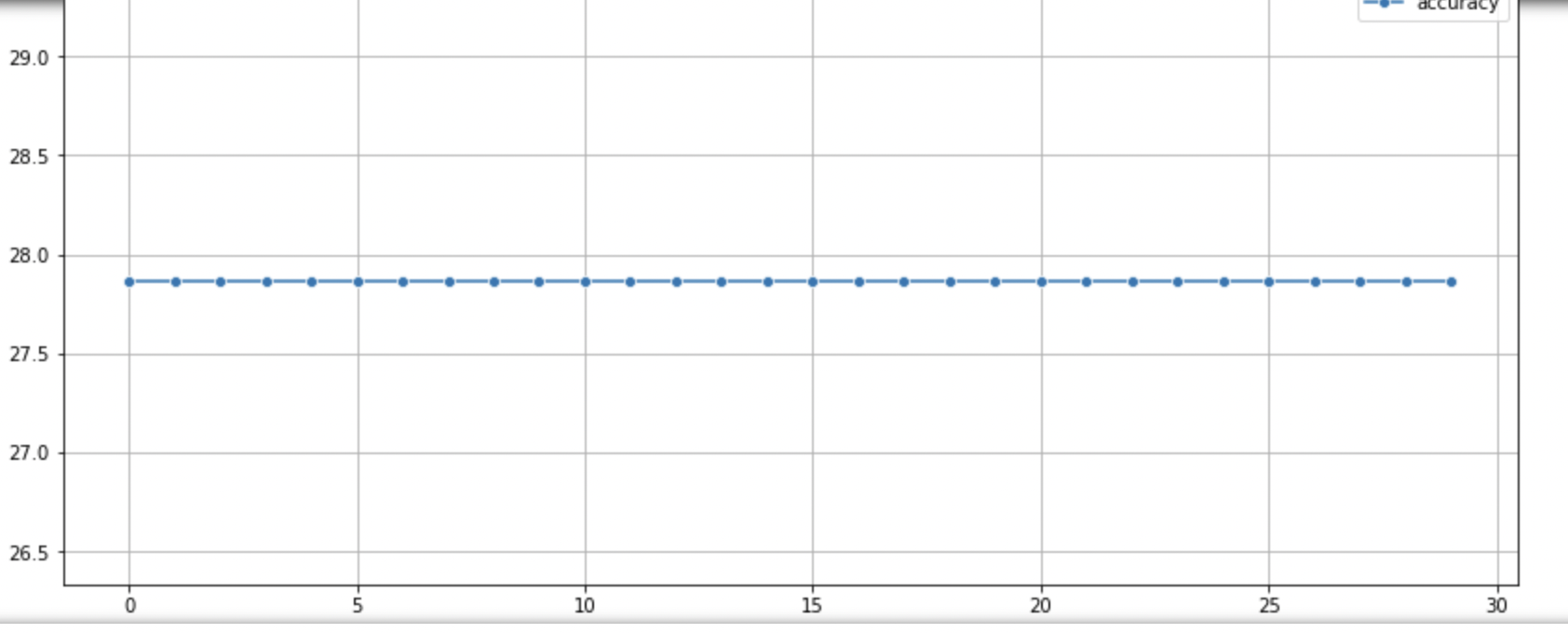
Sigmoid kernel gives result as



So we see that rbf kernel gives the best yet accuracy on 5-batch cross validation by 96.16 percent.



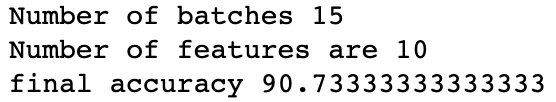
-changing gamma(default 0):



By changing gamma it was observed that there was no significant increase in the accuracy and the values were more or less same at 27.8%. So default of 0 is optimal.

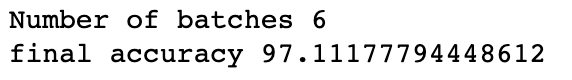
->Using 10 features instead of 25:

Under the same parameters the accuracy drops by 6% when the number of features are reduced as removal of important features leads to fraudulent decision boundaries.

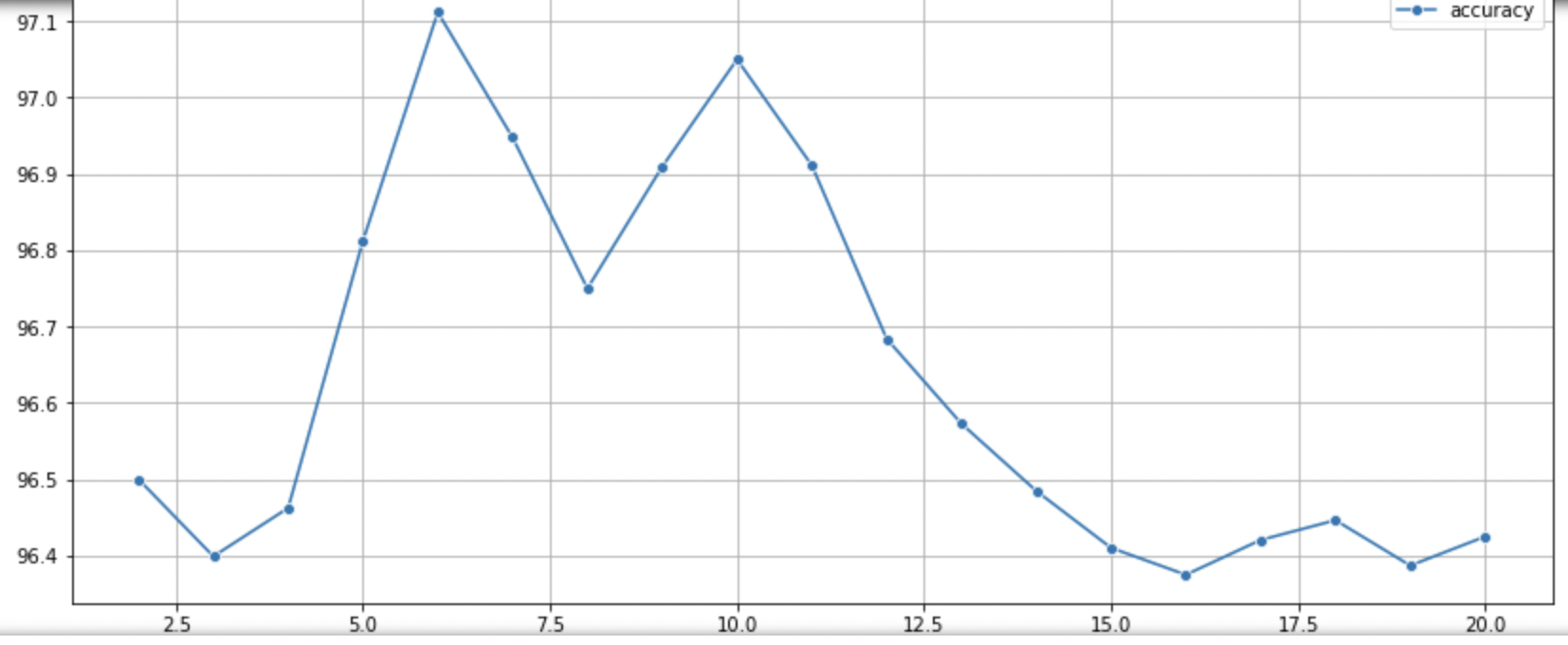


Part 2

Libsvm default fit gives this result on 6 batch cross validation



Highest cross validation accuracy reached at 6 batches



Only polynomial kernel comes close to the accuracy reached by the rbf kernel but it is still lesser.

The results were uploaded using this model.

X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X-X