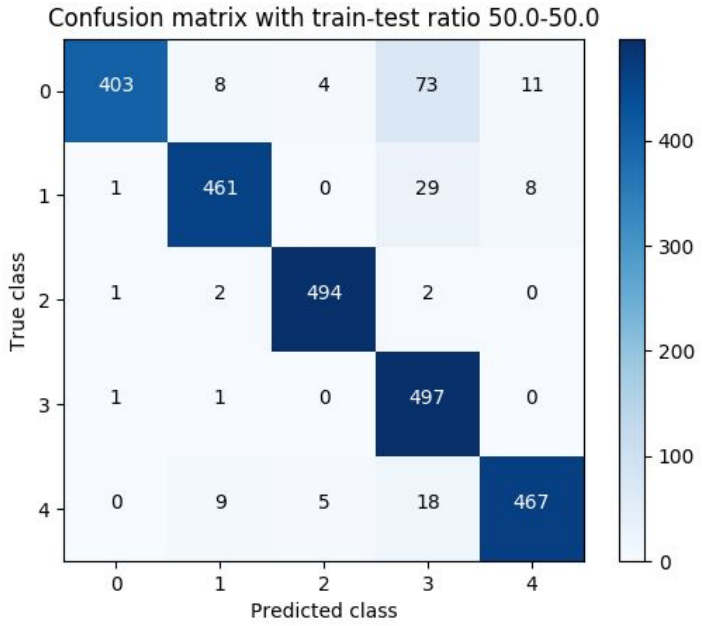
IR : Assignment 3

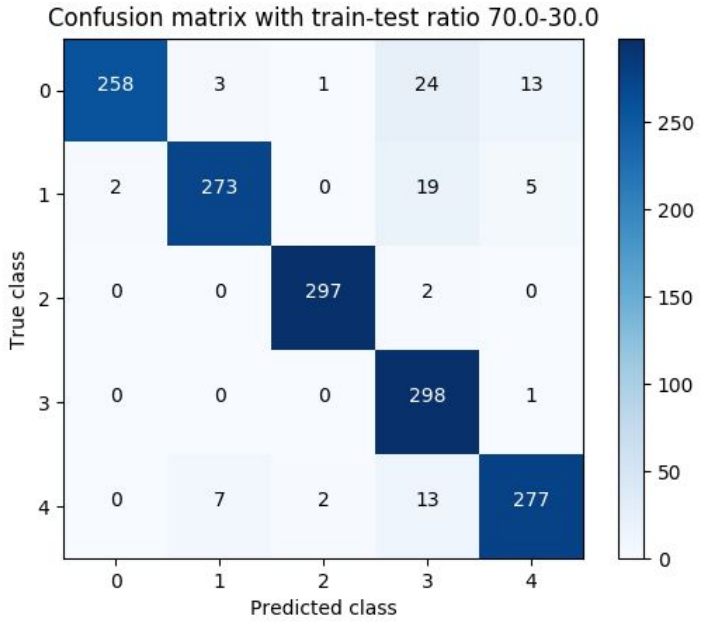
Harshvardhan Kalra
2014043

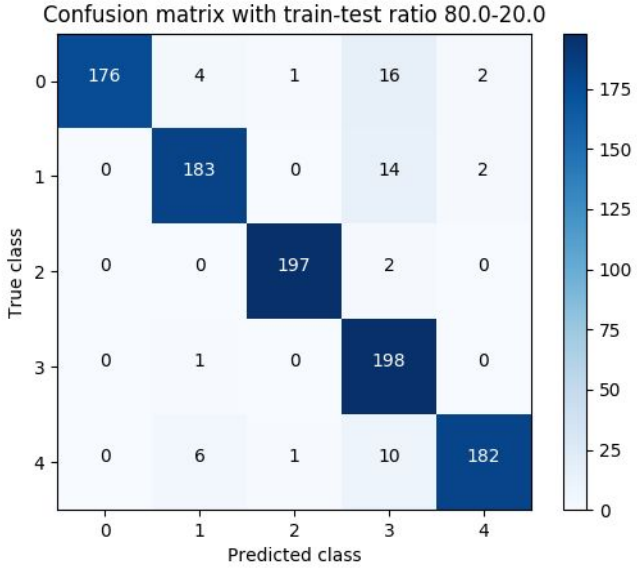
The accuracies decrease as the training to test set ratio is increased. This is because the model gets to see lesser training data and hence, fails to generalize over a larger set of unseen samples.

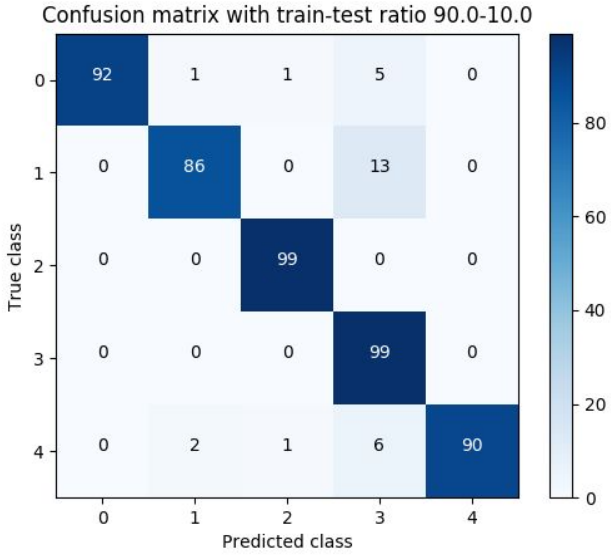
Train-test ratio	Accuracy(%)
90-10	94.1414141414
80-20	94.0703517588
70-30	93.8461538462
50-50	93.0661322645

The confusion matrices are attached below.


Confusion matrix with train-test ratio 50.0-50.0
True class
Predicted class
403 8 4 73 11
1 461 0 29 8
1 2 494 2 0
1 1 0 497 0
0 9 5 18 467
0 1 2 3 4
400 300 200 100 0


Confusion matrix with train-test ratio 70.0-30.0
True class
Predicted class
258 3 1 24 13
2 273 0 19 5
0 0 297 2 0
0 0 0 298 1
0 7 2 13 277
0 1 2 3 4
250 200 150 100 50 0


Confusion matrix with train-test ratio 80.0-20.0
True class
Predicted class
0 1 2 3 4
176 4 1 16 2
0 183 0 14 2
0 0 197 2 0
0 1 0 198 0
0 6 1 10 182
175 150 125 100 75 50 25 0


Confusion matrix with train-test ratio 90.0-10.0
True class
Predicted class
0 1 2 3 4
92 1 1 5 0
0 86 0 13 0
0 0 99 0 0
0 0 0 99 0
0 2 1 6 90
80 60 40 20 0