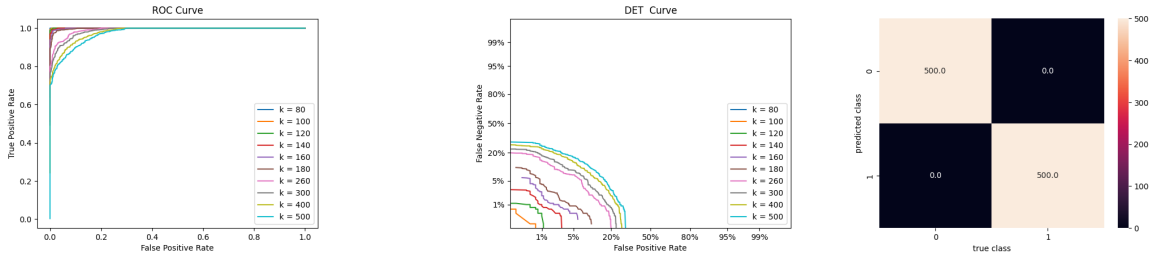
Assignment-4

K-Nearest Neighbour:-

1) Synthetic Image Data:-

a) **Without PCA and LDA :-** Accuracy is 100% for k 10 to 80 and there after accuracy drops.



(a) *ROCCurve* (b) *DETCurve* (c) *ConfusionMatrix*

Figure 1: ROC and DET curve

b) **With PCA and LDA :-**

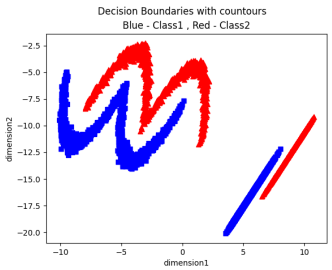
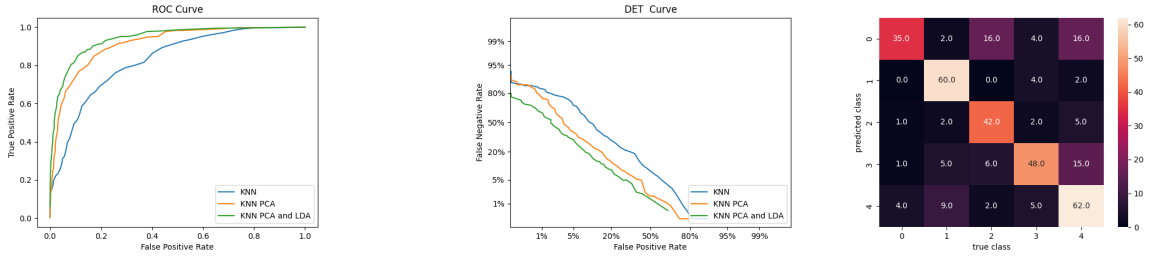


Figure 2: LDA Projection

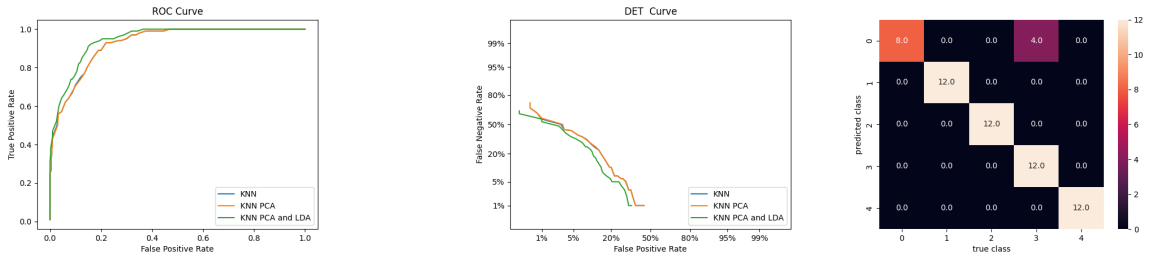
2) Real Image Data:- 1) 54% 2)PCA 70% 3) With PCA , LDA 72%



(a) *ROCCurve* (b) *DETCurve* (c) *ConfusionMatrix*

Figure 3: ROC and DET curve

3) Spoken Digit Data:-



(a) *ROCCurve* (b) *DETCurve* (c) *ConfusionMatrix*

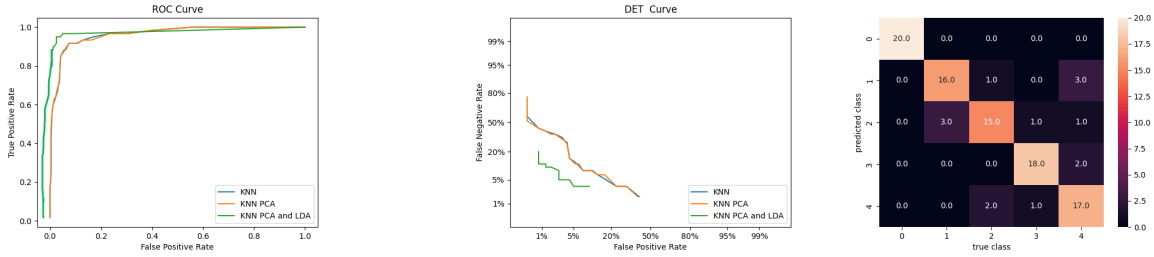
Figure 4: ROC and DET curves for k=100

1) KNN :- 95 %

2) PCA and then KNN :- 81.67 %

3) PCA then LDA and then KNN :- 88.33 %

4) Hand Written Data:-



(a) *ROCCurve* (b) *DETCurve* (c) *ConfusionMatrix*

Figure 5: ROC and DET curves for k=100

1) KNN :- 86.00 %

2) PCA and then KNN :- 82.00 %

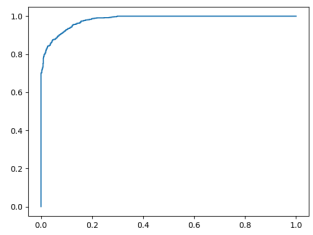
3) PCA then LDA and then KNN :- 82.00 %

In the above datasets , we are performing 1) KNN , 2) PCA and then KNN and 3) PCA then LDA and then KNN. For synthetic data set, PCA is not that much useful as there is no need of dimension reduction.

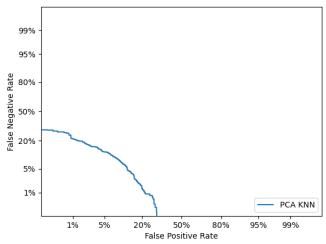
Logistic Regression:-

1) Synthetic Image Data:-

a) Without PCA and LDA :- Accuracy:-91.4 %



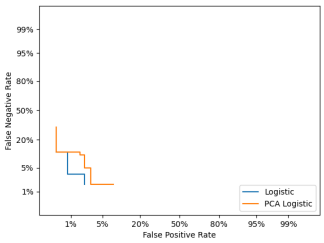
(a) *ROCCurve*



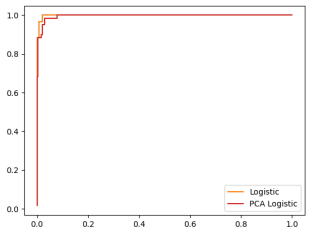
(b) *ConfusionMatrix*

Figure 1: ROC and DET curve

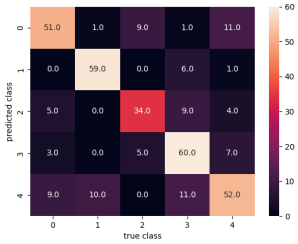
2) Real Image Data:-



(a) *DETCurve*



(b) *ROCCurve*



(c) *ConfusionMatrix*

Figure 2: ROC and DET curve

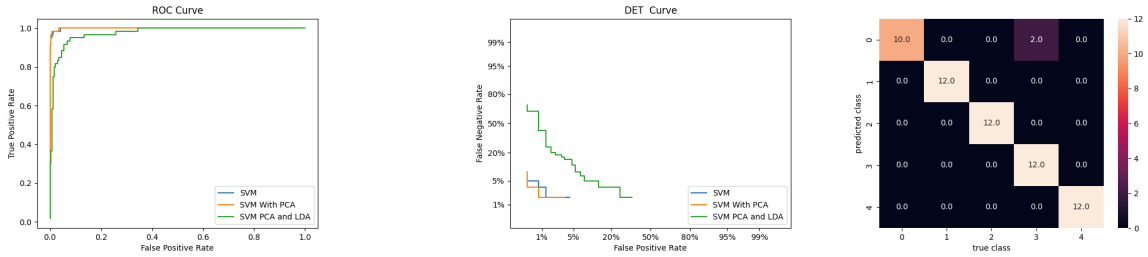
Accuracies:-

1) Logistic:- 73.56 %

2) PCA and then Logistic :- 72.12 %

3) PCA then LDA and then Logistic :- 75.85 %

3) Spoken Digit Data:-



(a) *ROCCurve* (b) *DETCurve* (c) *ConfusionMatrix*

Figure 3: ROC and DET curve

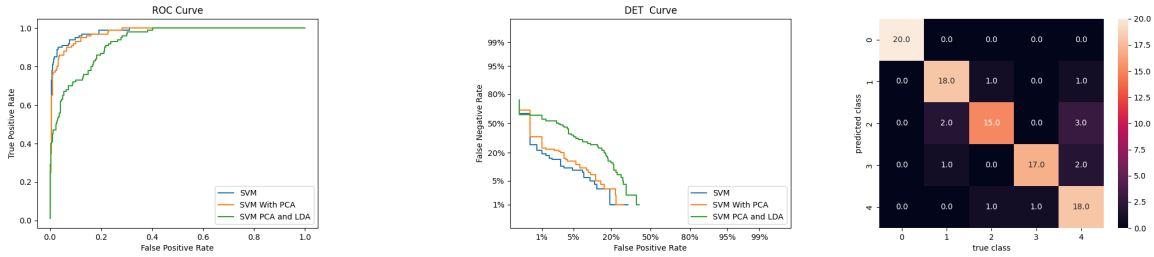
Accuracies:-

1) Logistic :- 96.67 %

2) PCA and then Logistic :- 96.67 %

3) PCA then LDA and then Logistic :- 85.0 %

4) Handwritten Data:-



(a) *ROCCurve* (b) *DETCurve* (c) *ConfusionMatrix*

Figure 4: ROC and DET curve

Accuracies:-

1) Logistic :- 88.0 %

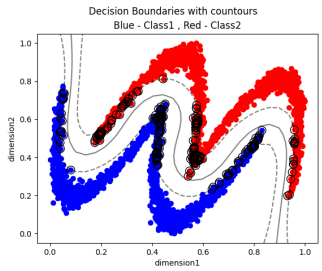
2) PCA and then Logistic :- 85.0 %

3) PCA then LDA and then Logistic :- 70.0 %

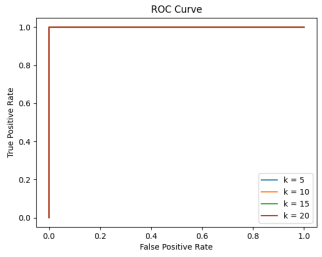
Support vector Machines:-

1) Synthetic Image Data:-

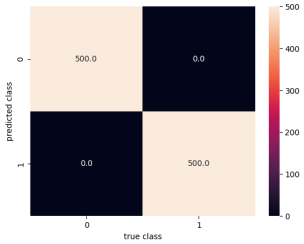
a) Without PCA and LDA :-



(a) *DecisionBoundary*



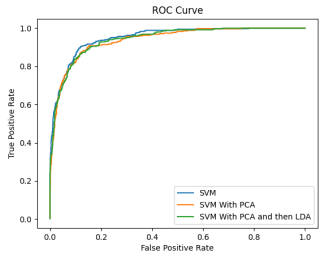
(b) *ROCCurve*



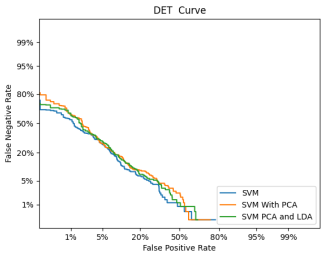
(c) *ConfusionMatrix*

Figure 1: ROC and DET curve

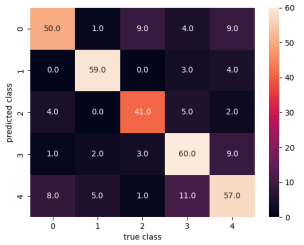
2) Real Image Data:-



(a) *ROCCurve*



(b) *DETCurve*



(c) *ConfusionMatrixforSVM*

Figure 2: ROC and DET curve

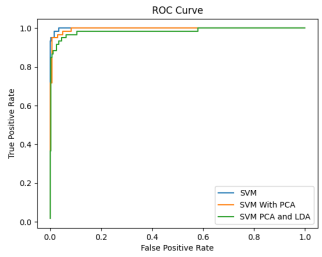
Accuracies:-

1) SVM :- 76.72 %

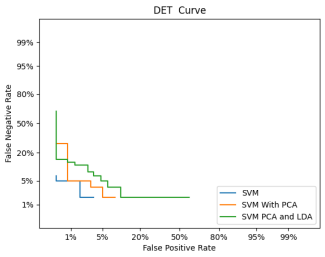
2) PCA and then SVM :- 75.57 %

3) PCA then LDA and then SVM :- 73.27 %

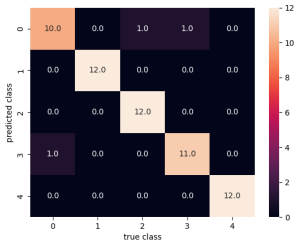
3) Spoken Digit Data:-



(a) *ROCCurve*



(b) *DETCurve*



(c) *ConfusionMatrixforSVM*

Figure 3: ROC and DET curve

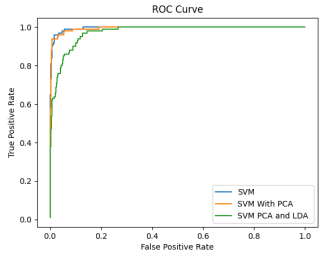
Accuracies:-

1) SVM :- 95.0 %

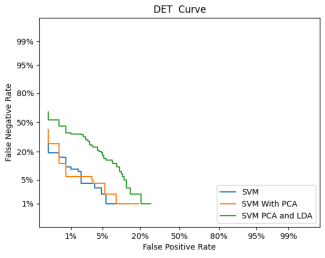
2) PCA and then SVM :- 95.0 %

3) PCA then LDA and then SVM :- 90.0 %

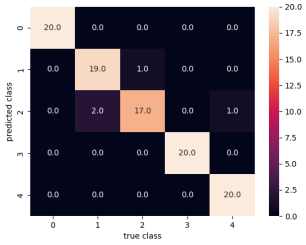
4) Handwritten Data:-



(a) *ROCCurve*



(b) *DETCurve*



(c) *ConfMatrixforPCAandSVM*

Figure 4: ROC and DET curve

Accuracies:-

1) SVM :- 93.0 %

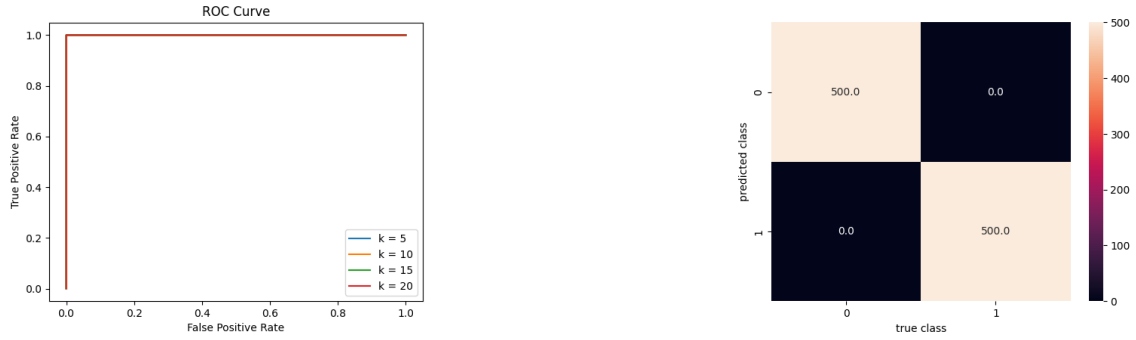
2) PCA and then SVM :- 96.0 %

3) PCA then LDA and then SVM :- 81.0 %

Artificial Neural Networks:-

1) Synthetic Image Data:-

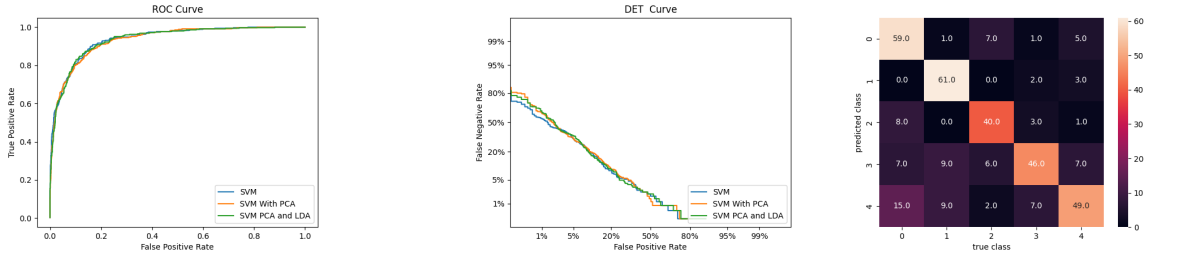
a) Without PCA and LDA :- Accuracy :- 100%



(a) *ROCCurve* (b) *ConfusionMatrix*

Figure 5: ROC and DET curve

2) Real Image Data:-



(a) *ROCCurve* (b) *DETCurve* (c) *ConfusionMatrixforSVM*

Figure 6: ROC and DET curve

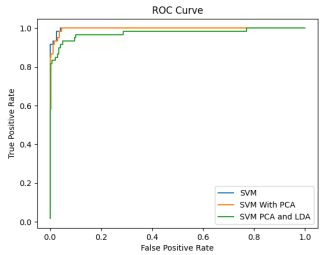
Accuracies:-

1) ANN :- 73.27 %

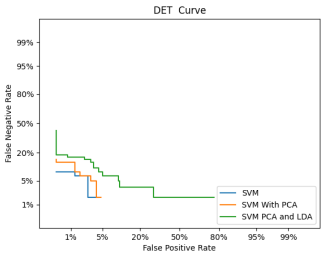
2) PCA and then ANN :- 72.12 %

3) PCA then LDA and then ANN :- 71.83 %

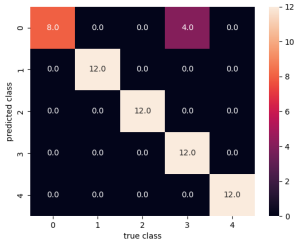
3) Spoken Digit Data:-



(a) *ROCCurve*



(b) *DETCurve*



(c) *ConfusionMatrixforSVM*

Figure 7: ROC and DET curve

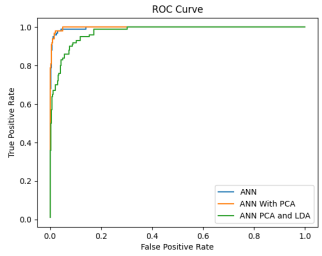
Accuracies:-

1) ANN :- 95.0 %

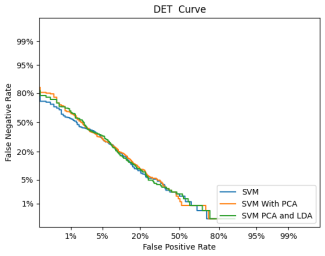
2) PCA and then ANN :- 93.33 %

3) PCA then LDA and then ANN :- 86.67 %

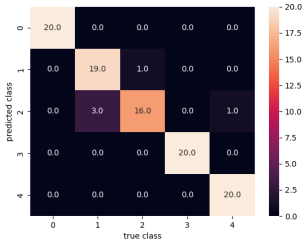
4) Handwritten Data:-



(a) *ROCCurve*



(b) *DETCurve*



(c) *ConfMatrixforPCAandSVM*

Figure 8: ROC and DET curve

Accuracies:-

1) ANN :- 95.0 %

2) PCA and then ANN :- 95.0 %

3) PCA then LDA and then ANN :- 79.0 %