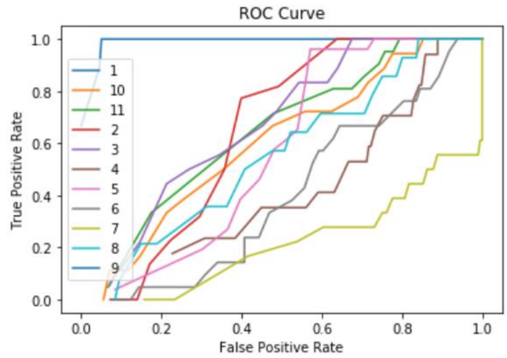
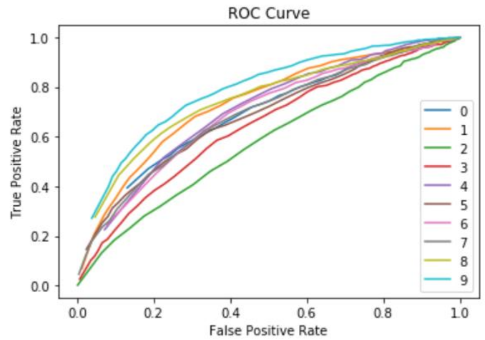
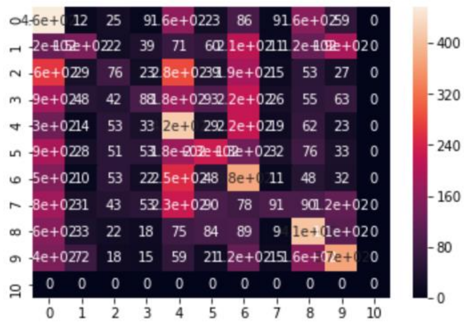
Classify the dataset as it is and use the same classification algorithm

When given data is split in ratio 70:30 ratio then 70% is training and 30% is testing then

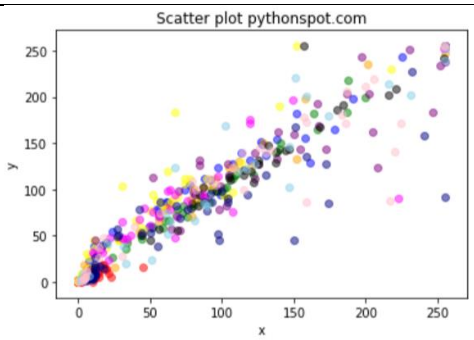
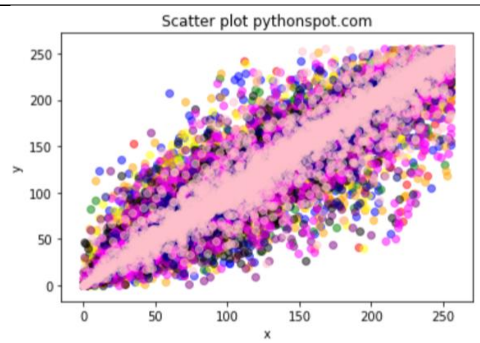
- Accuracy

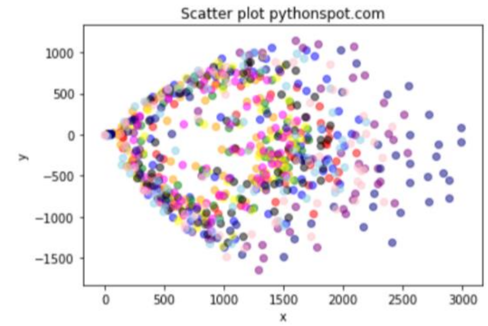
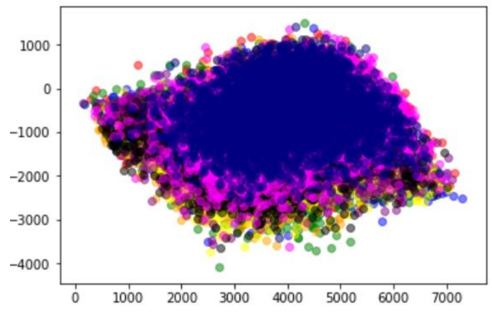
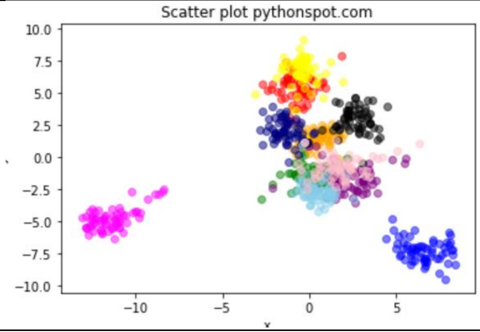
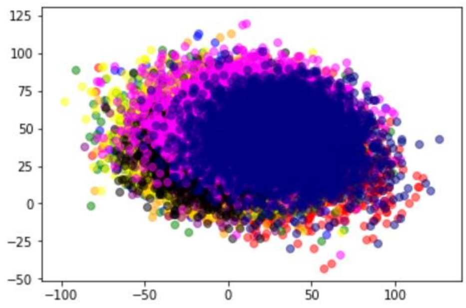
	Accuracy
Dataset1	0.7116
Dataset2	0.2662

- ROC Curve and Confusion Matrix

<table>
<tr><td></td><td>ROC</td><td>Confusion Matrix</td></tr>
<tr><td>Dataaset 1</td><td>
</td><td></td></tr>
<tr><td>Dataset2</td><td>
</td><td></td></tr>
</table>

Project your data on the new projection

<table>
<tr><td></td><td>DataSet1</td><td>DataSet2</td></tr>
<tr><td>Original Data</td><td>
</td><td>
</td></tr>
</table>

New Project Data Using PCA		
New Project data Using LDA		

Performing LDA:

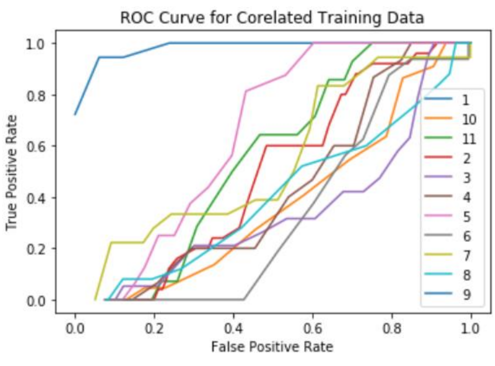
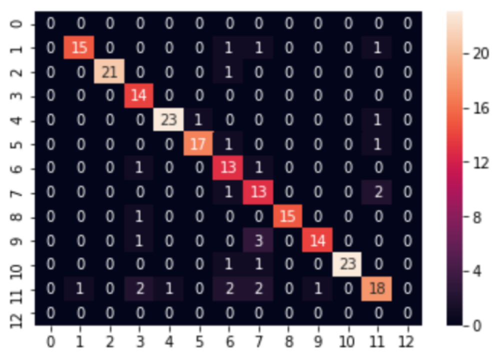
- Training Accuracy and Testing Accuracy

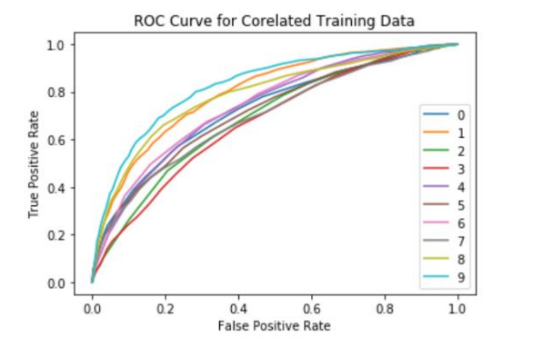
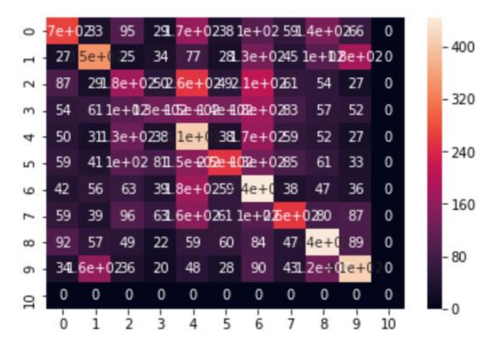
	Training Accuracy Fold1	Training Accuracy Fold2	Training Accuracy Fold3	Training Accuracy Fold4	Training Accuracy Fold5	Testing Accuracy
Dataset1	0.86	0.83	0.8	0.88	0.81	0.865
Dataset2	32.5	31.13	32.21	31.16	32.13	31.36

- Mean and Standard Deviation : 0.836 and 0.003

	Mean	Standard Deviation
Dataset1	0.836	0.003
Dataset2	0.31826	0.005695823030958737

- ROC and Confusion Matrix

	ROC	Confusion Matrix
Dataset 1		

Dataset 2		

<u>Analysis of LDA :</u>

1. For Face Data when normal classifier then accuracy is 71% while when we apply LDA along with K-fold then we get best model and the test data is tested the accuracy is 86.5% because LDA reduces the dimension of features to (k-1) where k is number of classes and perform best projection so accuracy increases after applying classifier.
2. For Cifar data since the features are not good so when we apply normal classifier the accuracy is 256% while applying LDA with K-fold then accuracy becomes 31%.The reason is same as stated above.
3. When comparing performance of LDA on face-data and cifar, it performs well on face-data because the features of cifar data are not so good as compared to face-data. As the projected data for cifar data is so overlapping between the classes so misclassification rate for LDA will be more. It will be hard to maximize the distance between the classes.

(f). Perform PCA on the same data by preserving 95% eigen energy and report the mean accuracy and standard deviation over the 5 folds. Use the best model, to classify the test set and plot the ROC curve and confusion matrix. How does the results on the 2 databases differ and why?

Performing PCA with 95% eigen-energy :

- Accuracy

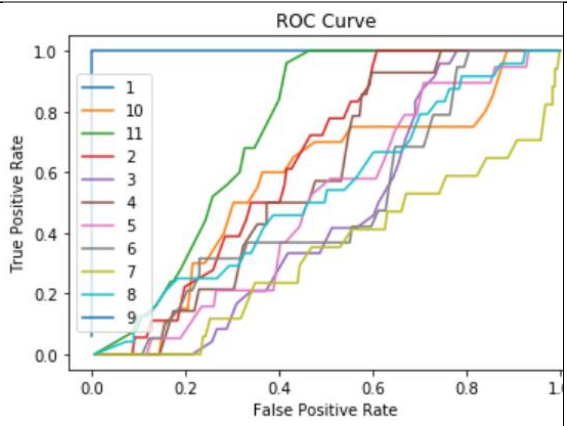
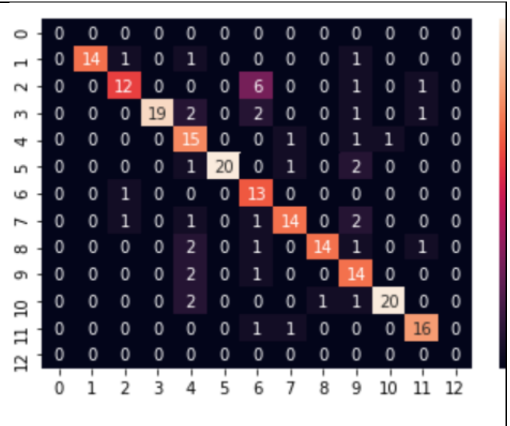
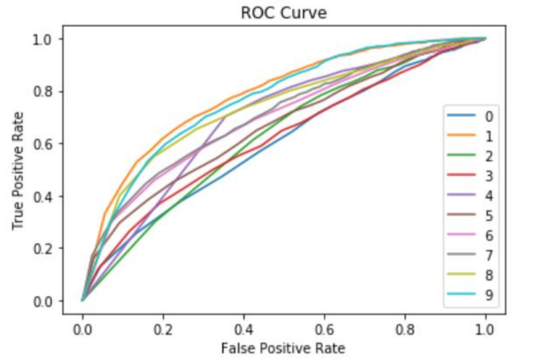
	Training Accuracy Fold1	Training Accuracy Fold2	Training Accuracy Fold3	Training Accuracy Fold4	Training Accuracy Fold5	Testing Accuracy
Face-data	0.72	0.8	0.79	0.83	0.82	0.7953
cifar	0.2803	0.271	0.2777	0.2651	0.2752	0.2775

- Mean and Standard Deviation

	Mean	Standard Deviation
Face-data	0.792	0.03867815921162743
cifar	0.27386	0.0053466

- ROC and Confusion Matrix :

	ROC	Standard Deviation

faceda ta		
cifar		

<u>Analysis of PCA :</u>

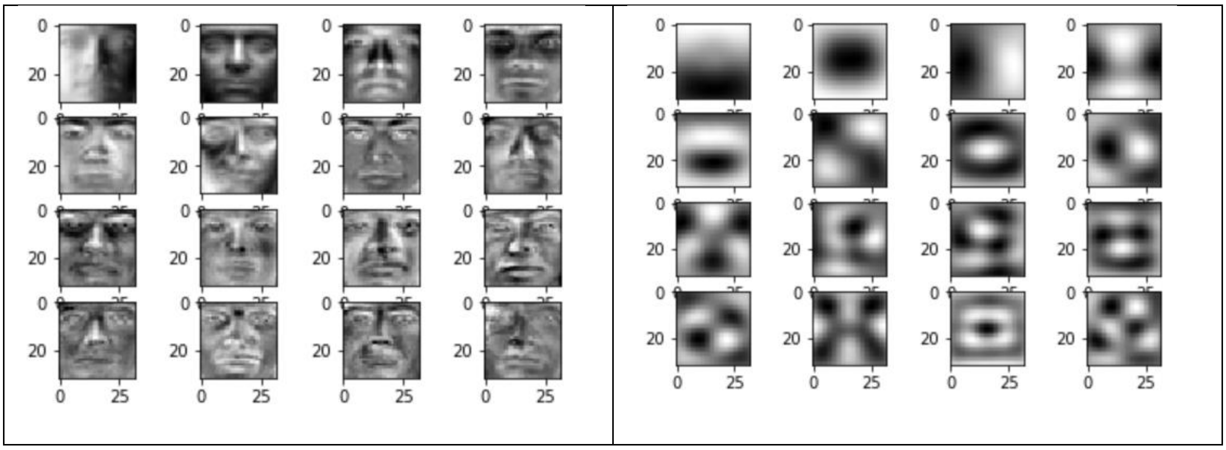
When we apply PCA with 95% eigen energy, the feature space is reduces to 32 features for face-data and 160 features for cifar data. For face-data the PCA is performing well as compared to cifar data since original features of cifar data are not good so what we observe from above information is that the accuracy is still same and there is no effect of dimension reduction as there is no loss of information.PCA can't separate the classes from each other. From above observation we can infer that the features are corelated for cifar data but not for face-data so PCA is performing well on face-data as compared to cifar data.

Compare and analyze the results obtained by PCA and LDA.

The LDA performs well on both the dataset that is accuracy of LDA is more as compared to PCA. This is because LDA is supervised which takes class labels into consideration and always reduces upto (k-1)dimension reduction where k is number of classes while PCA is unsupervised learning as it reduces dimension based on eigen energy. More the eigen energy the number of features will be more. Since PCA finds the direction of maximal variance while LDA tries to find feature subspace which will maximize the class separability.

Visualise and analyse the eigenvectors obtained using PCA

DataSet1	DataSet2



From above observation, the eigen vectors for face-data are more clear as compared to eigen-vectors of cifar data because the original features of cifar data is also not good. From above observation we can infer that as we go down the eigen faces is getting blur it means that the eigen vector that is obtained is not important.More the value of eigen vector more information it will provide.

<u>LDA on the PCA :</u>

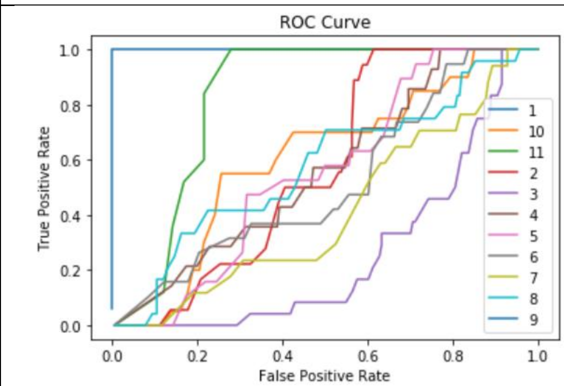
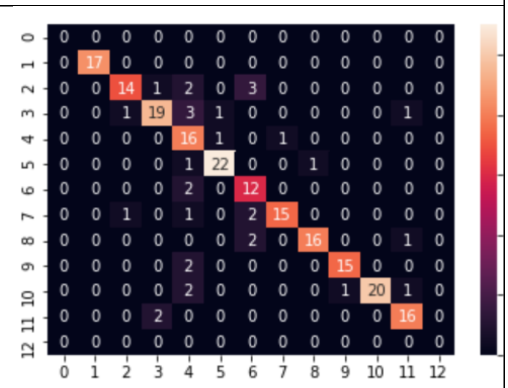
- Accuracy :

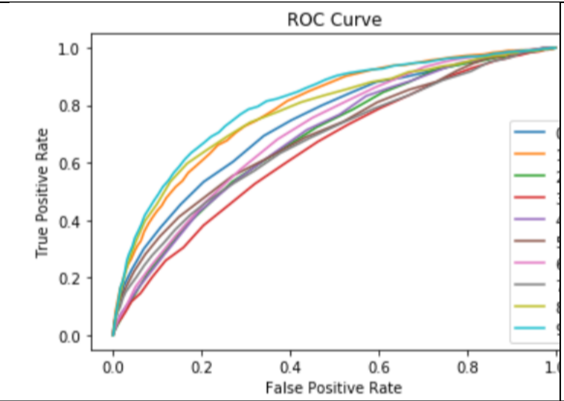
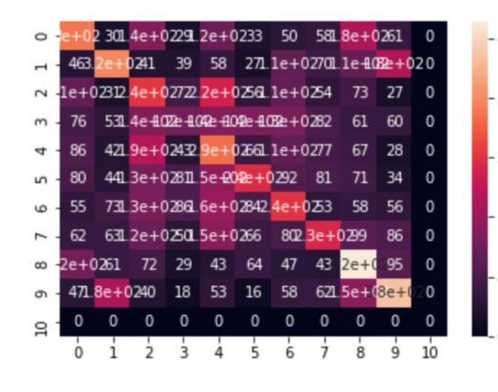
	Training Accuracy Fold1	Training Accuracy Fold2	Training Accuracy Fold3	Training Accuracy Fold4	Training Accuracy Fold5	Testing Accuracy
Face-data	0.8	0.87	0.86	0.88	0.9	0.8465
Cifar-data	0.325	0.3113	0.3221	0.3116	0.3213	0.3136

- Mean and Accuracy : 0.8619999999999999 and 0.03370459909270543

	Mean	
Face-data	0.8619999999999999	0.03370459909270543
Cifar-data	0.3491000000000001	0.004687856653098512

- ROC and Confusion Matrix

<table>
<tr><td></td><td>ROC</td><td>Confusion Matrix</td></tr>
<tr><td>Face
-
data</td><td>
</td><td></td></tr>
</table>

Cifar - data		

PCA on the LDA :

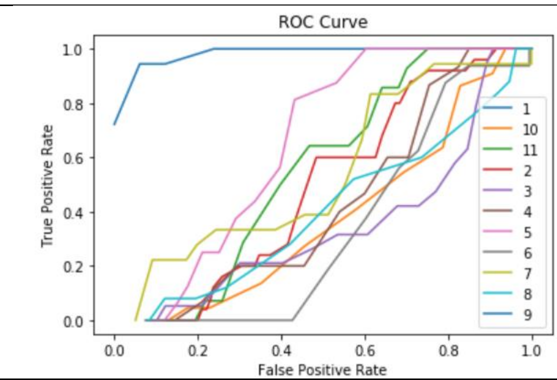
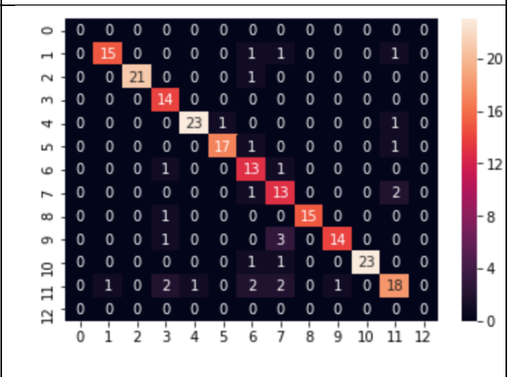
- Accuracy

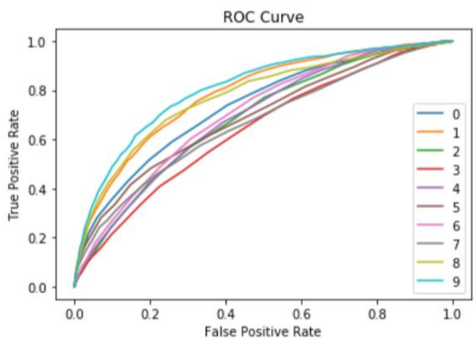
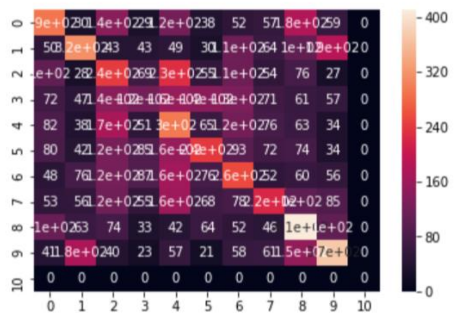
	Training Accuracy Fold1	Training Accuracy Fold2	Training Accuracy Fold3	Training Accuracy Fold4	Training Accuracy Fold5	Testing Accuracy
Face-data	0.86	0.8	0.78	0.85	0.81	0.8034
Cifar-data	0.3545	0.3416	0.356	0.3429	0.3486	0.2773

- Mean and Accuracy : 0.8034 and 0.030066592756745798

	Mean	Standard Deviation
Face-data	0.8034	0.030066592756745798
Cifar-data	0.34872	0.005847871407614902

- ROC and Confusion Matrix :

	ROC	Confusion Matrix
Face - data		

Cifar - data		

Comparison of LDA on the PCA and PCA on the LDA :

From above observation we can infer that LDA on the PCA is performing well than PCA on LDA because if we first apply PCA with high eigen energy then useful features will be obtained and all noisy information will be removed and when we apply LDA on PCA projected data then it will be easier to classify the data for LDA. While on the other hand if we first apply LDA then number of features will be reduced to (k-1) where k is number of classes then we apply PCA then again some features will be loose so as number of features are less than may be PCA is performing well on LDA projected data.