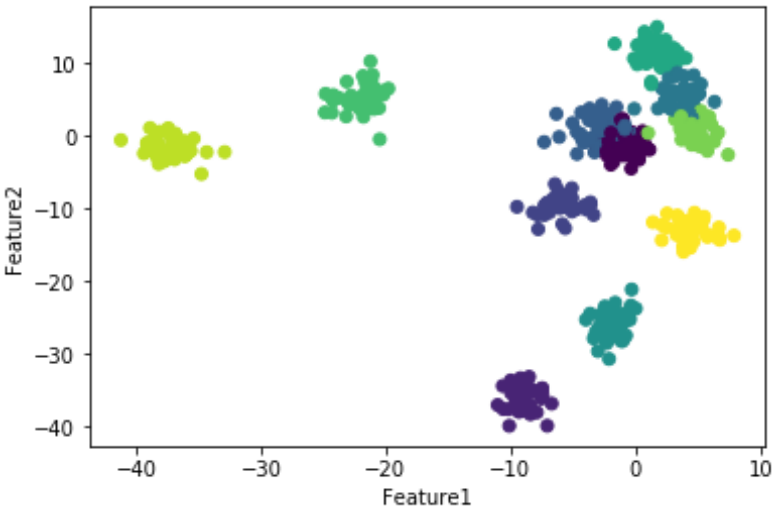
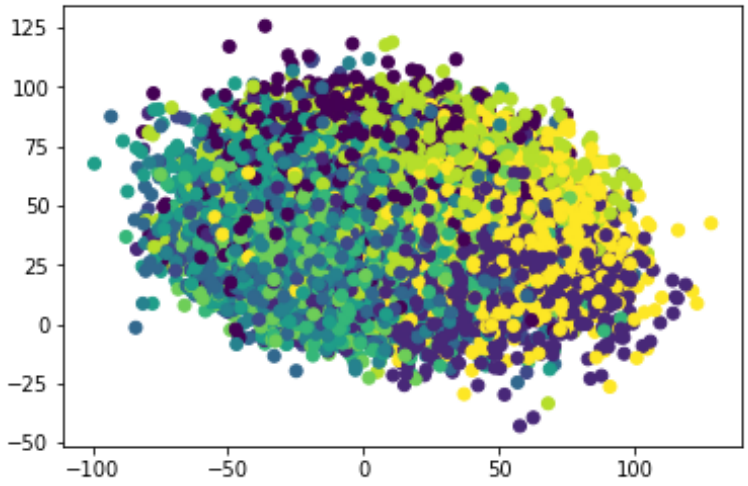
Ques1

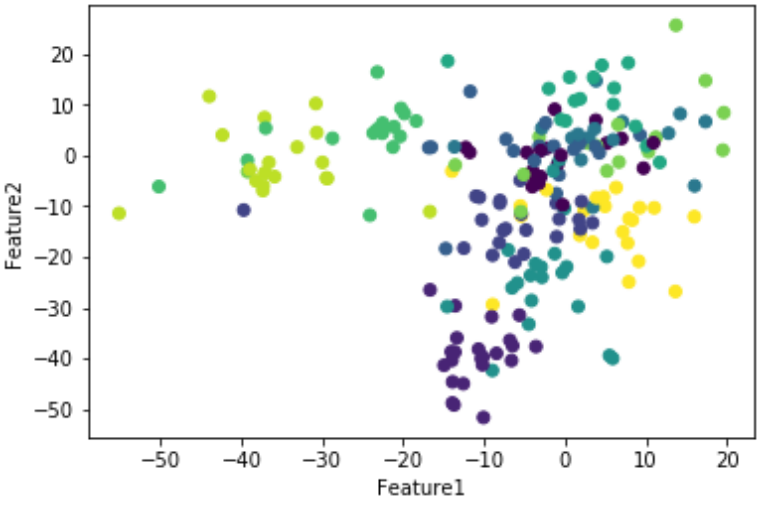
Data Visualizations(LDA Projected across best Projections)

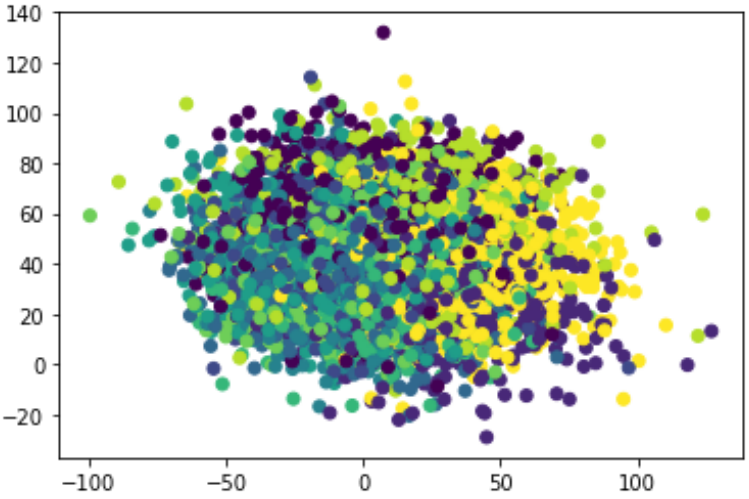
Training Dataset



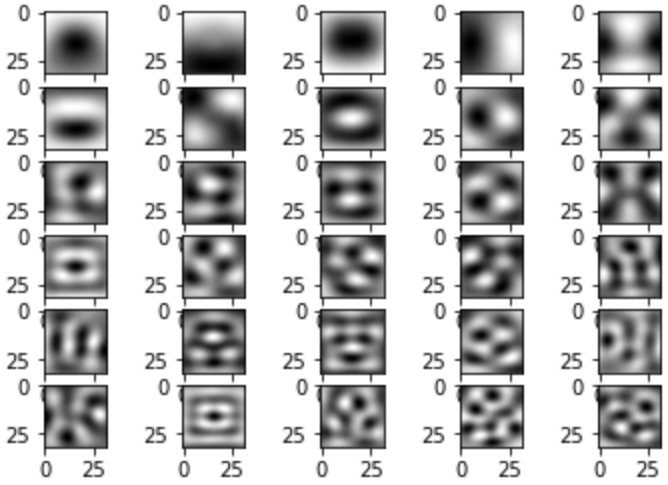


Test Dataset





Eigenvectors Visualization (Eig Energy<=95%)

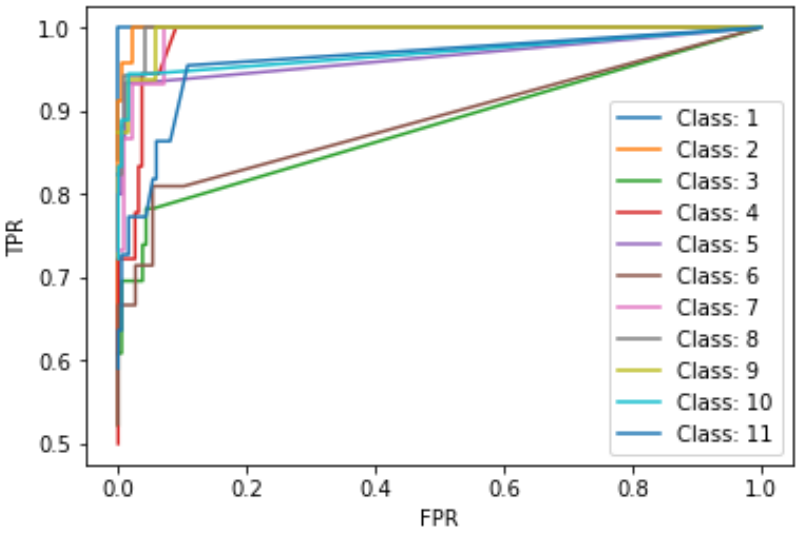


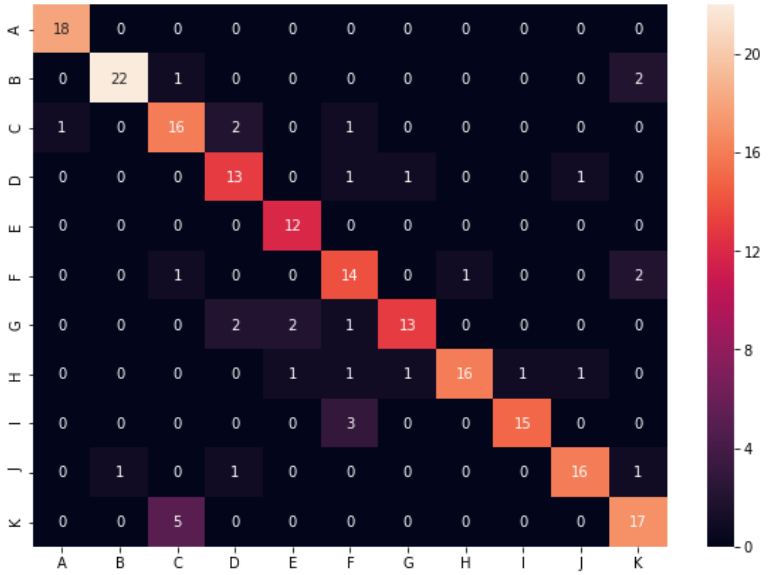
INITIAL ACCURACY: 70.04%

Dataset 1 ACCURACY TABLE(%)

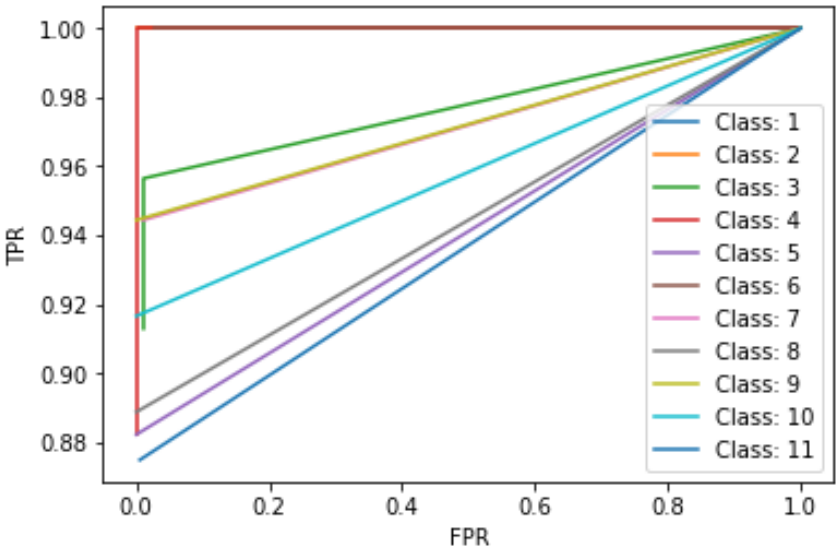
Configuration	Cross Validation				Test Dataset
	Mean	Standard Deviation	Best Model on Validation Dataset	Best Model on Test Data	
PCA(EE<70)	8.2	0.018	12.8	11.0	9.68
PCA(EE<90)	74.24	0.052	80.04	70.0	70.2
PCA(EE<95)	84.02	0.022	88.56	83.12	84.54
PCA(EE<99)	85.32	0.038	94.40	86.0	88.40
LDA	99.0	0.0042	100	94.28	94.20

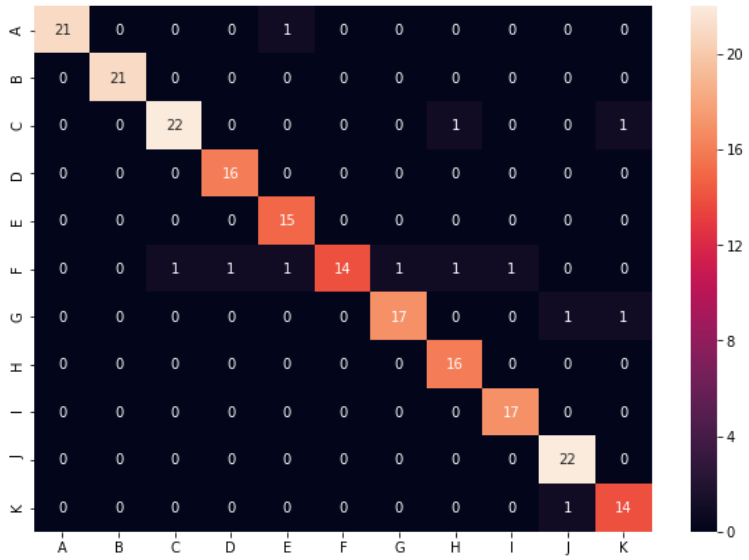
PCA





LDA



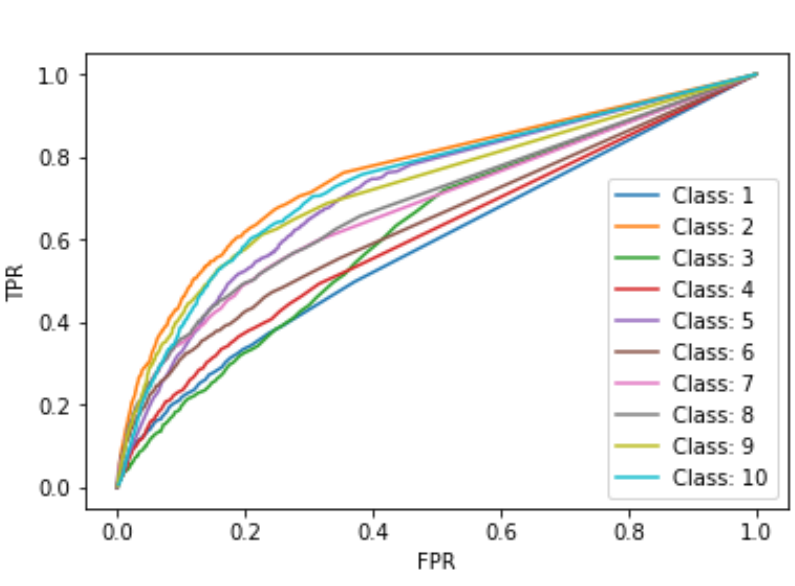


INITIAL ACCURACY: 26.64

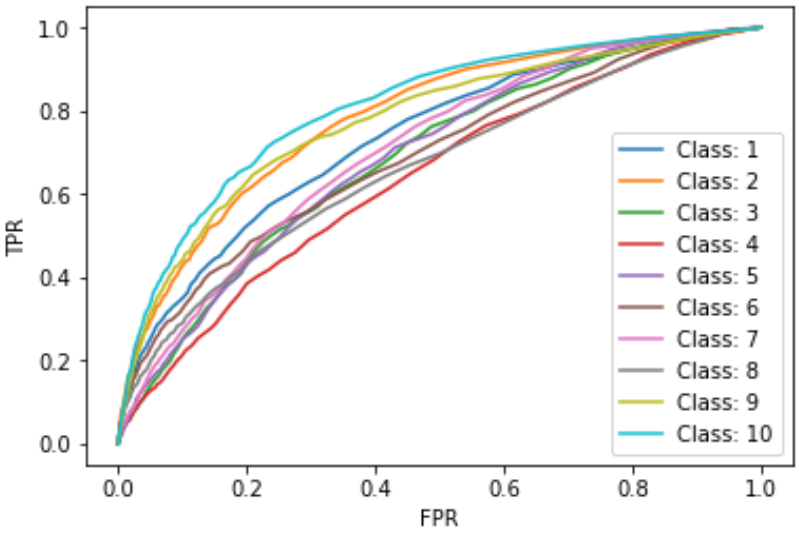
Dataset 2 ACCURACY TABLE(%)

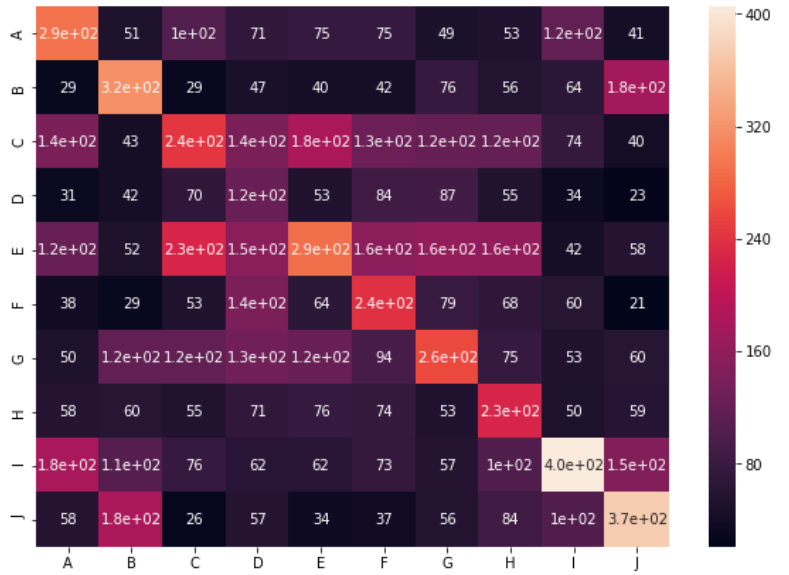
Configuration	Cross Validation				Test
	Mean	Standard Deviation	Best Model on Validation Dataset	Best Model on Test Data	
PCA(EE<70)	29.56	0.00474	29.68	29.96	29.89
PCA(EE<90)	28.88	0.0080	29.54	28.87	28.82
PCA(EE<95)	27.44	0.0056	28.64	27.72	27.84
PCA(EE<99)	27.84	0.0082	29.14	28.22	28.16
LDA	38.48	0.0060	35.48	32.62	30.52

PCA



LDA

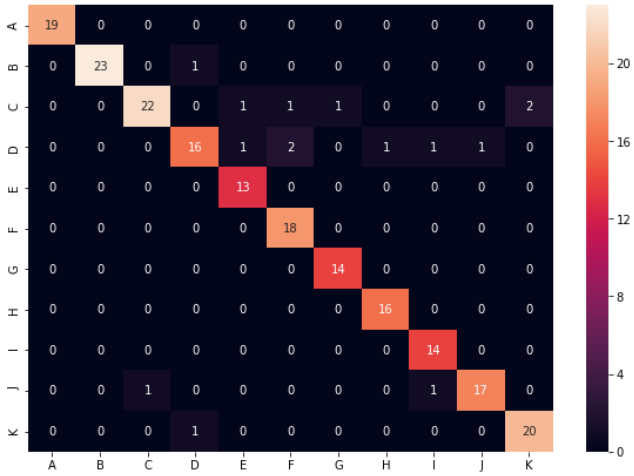
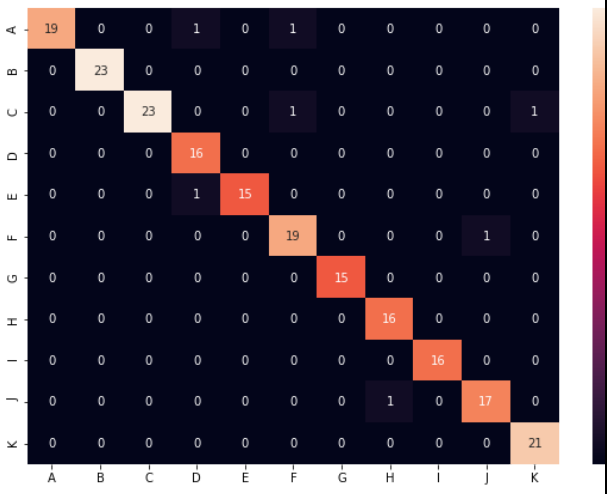
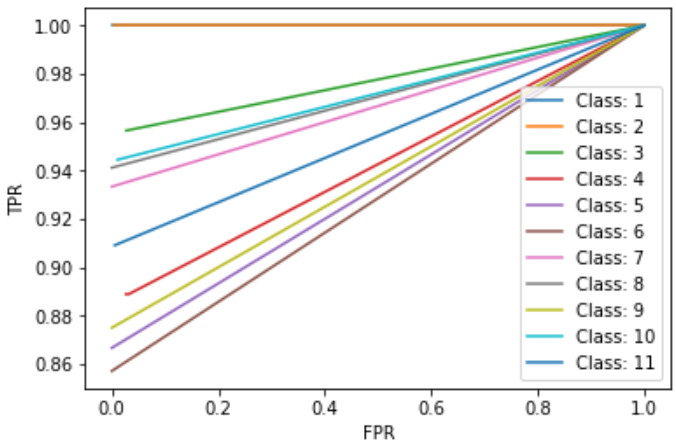
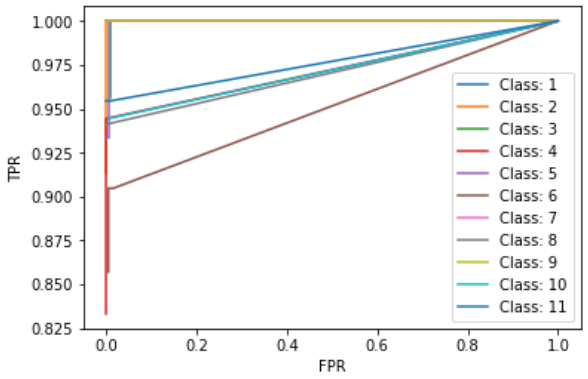




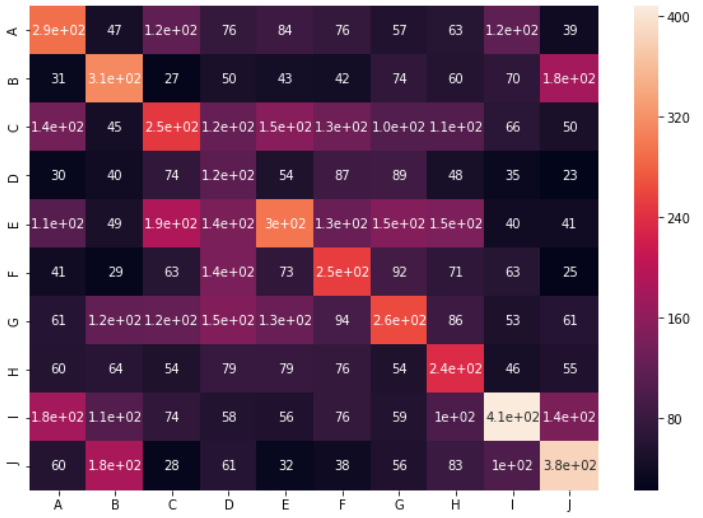
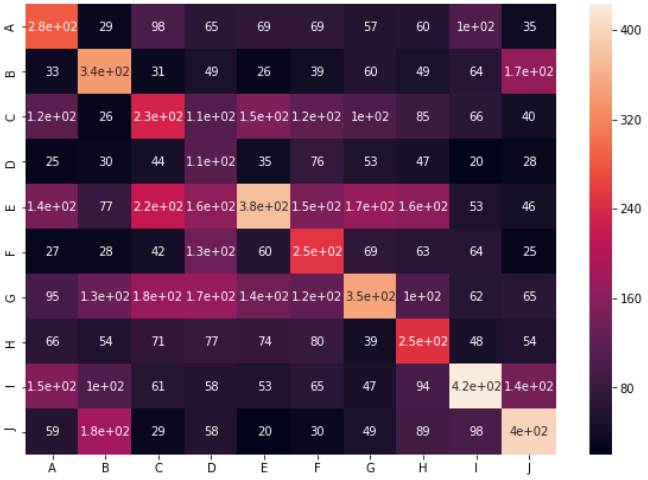
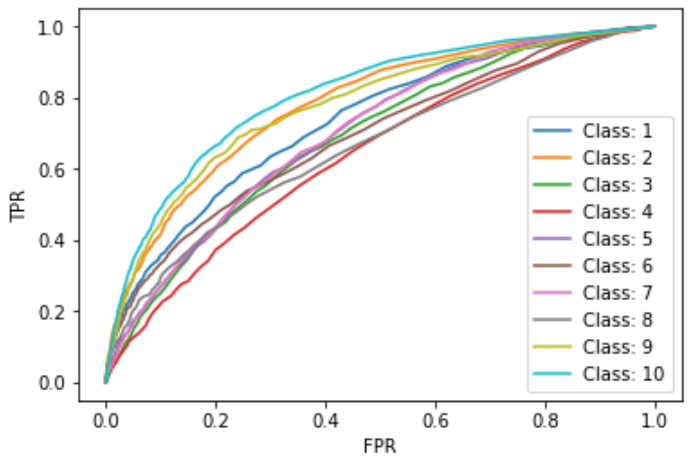
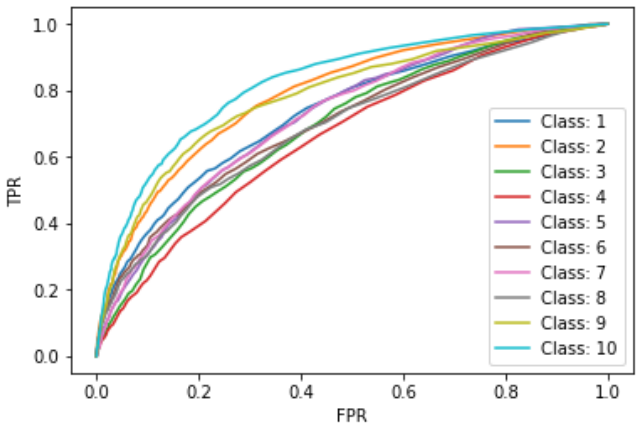
Accuracy on Dataset 1 after PCA or LDA is more than Dataset 2:

1. **Huge difference in the number of features even after PCA (conserving 95% EE).**
2. **High randomness/difference (between class variance) among instances of different classes.**
3. **Initial image size in Dataset 2 is smaller than Dataset 1.**

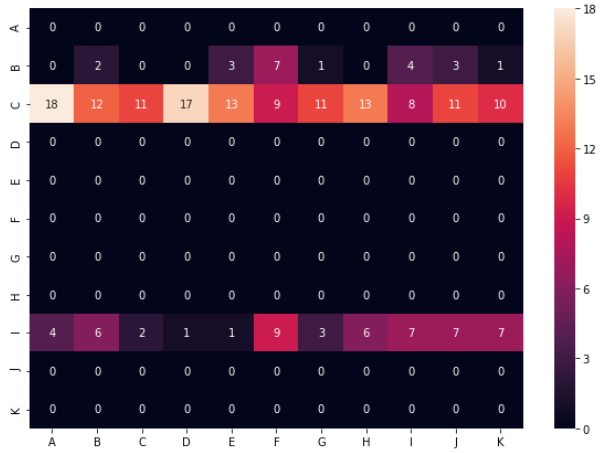
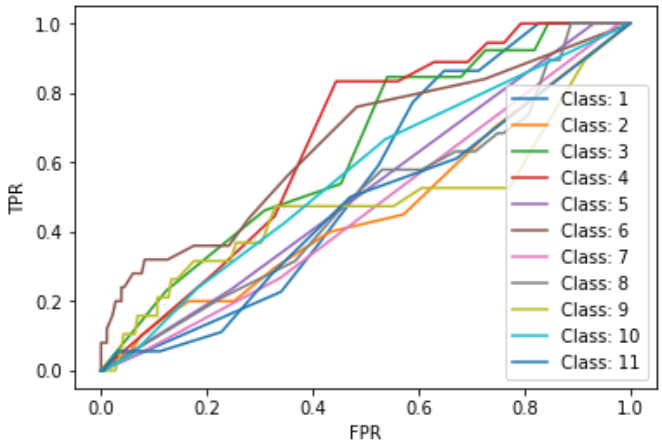
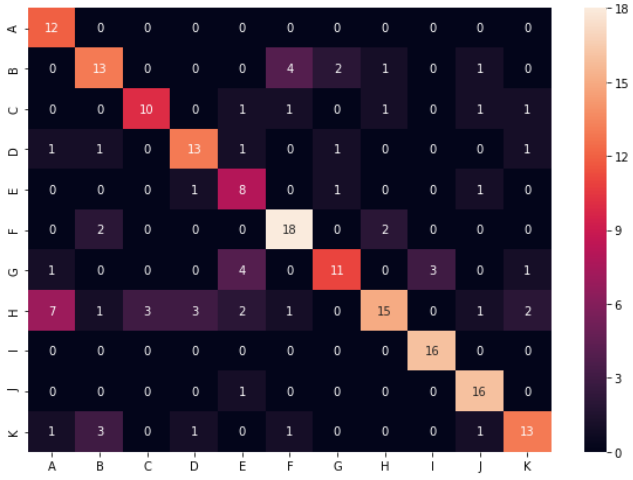
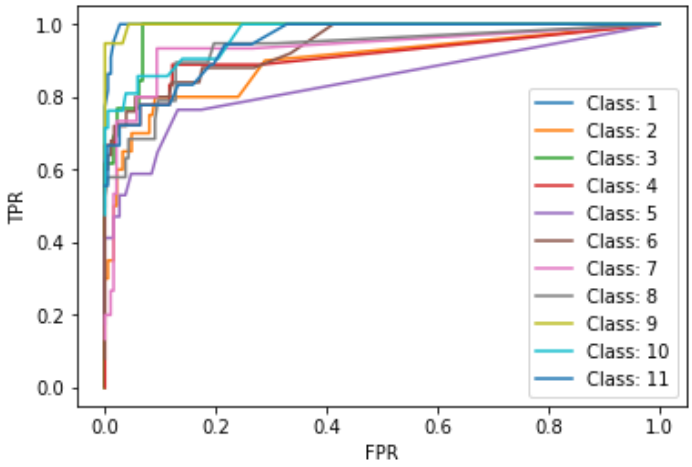
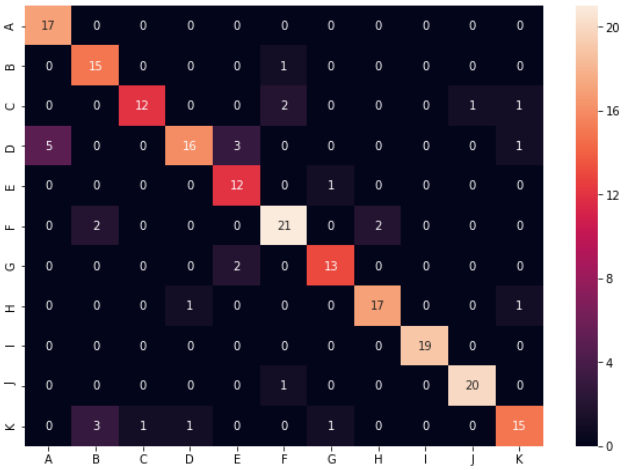
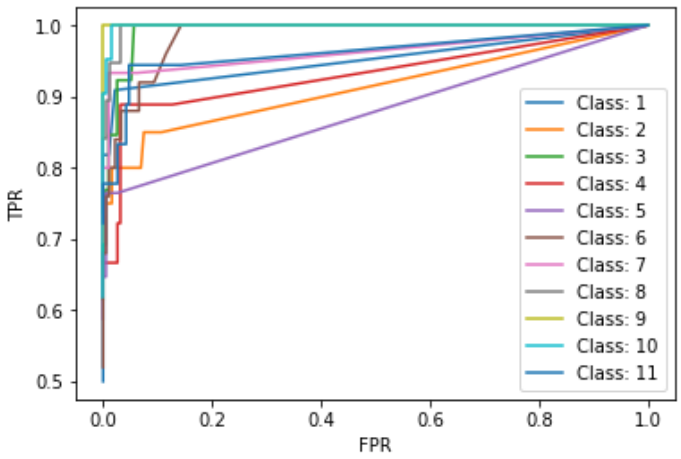
Dataset 1

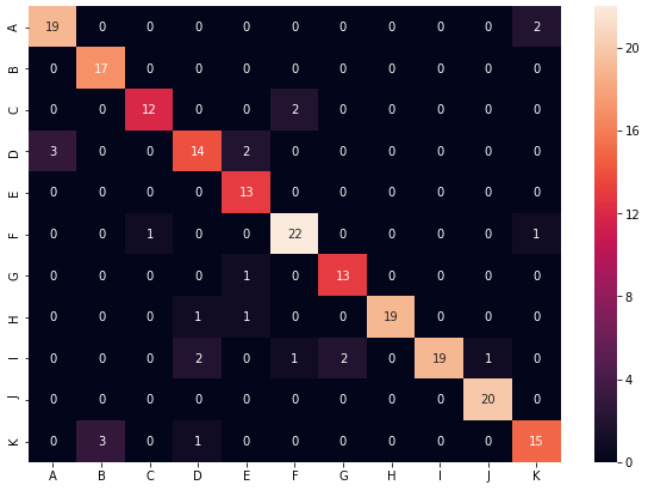
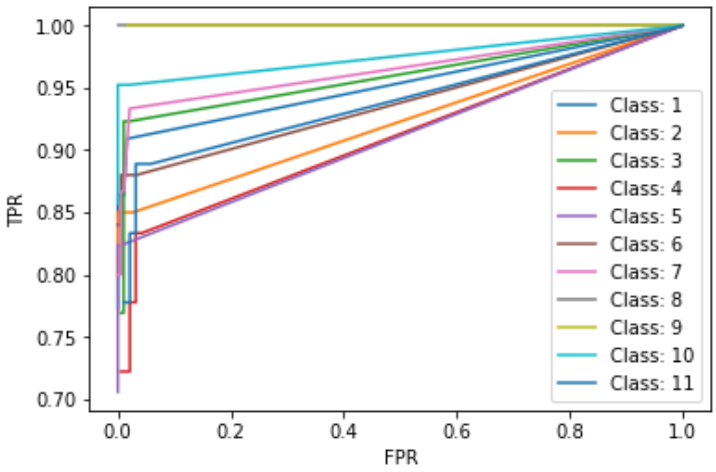
Measure of Analysis	PCA on LDA Projected Data (Eig Eng<=.99)	LDA on PCA Projected Data
Confusion Matrix		
ROC Curve		
Accuracy/ No. of Features	**92.76** **94** **PCA is not a classification method, hence PCA on LDA does not increase the accuracy much than only applying LDA.**	**96.62** **10** **LDA can be used as a classifier because it increases the separability(clustering) among the classes hence accuracy increases.**

Dataset 2

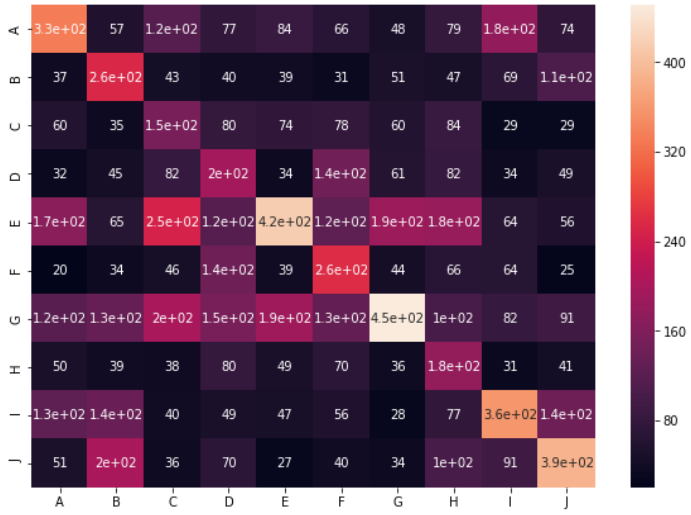
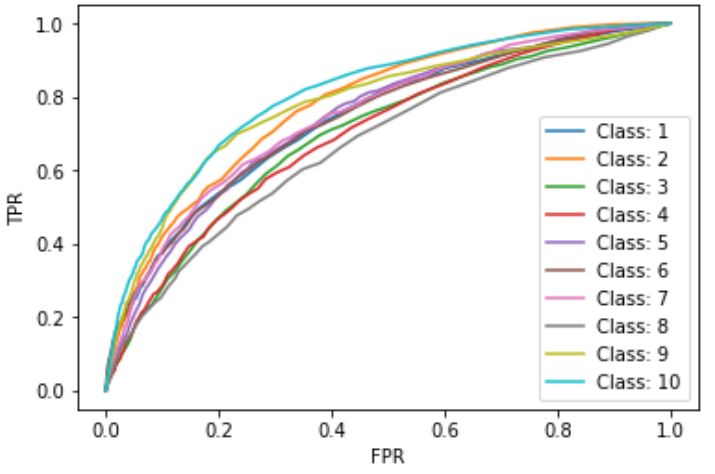
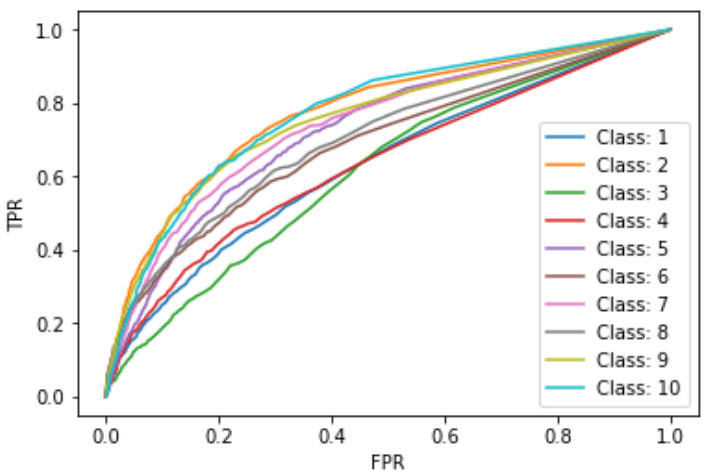
Measure of Analysis	PCA on LDA Projected Data (Eig Eng<=.99)	LDA on PCA Projected Data
Confusion Matrix		
ROC Curve		
Accuracy/ No. of Features	**29.11** **8** **PCA is not a classification method, hence PCA on LDA does not increase (adds to) the accuracy much than already only applying LDA.**	**30.68** **9** **LDA can be used as a classifier because it increases the separability(clustering) among the classes hence accuracy increases.**

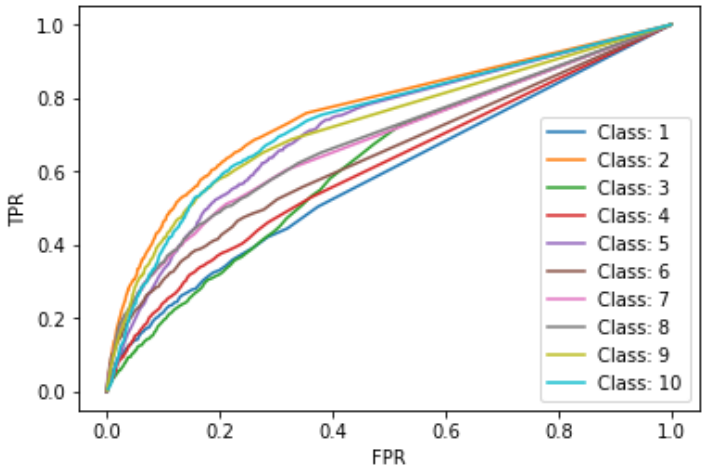
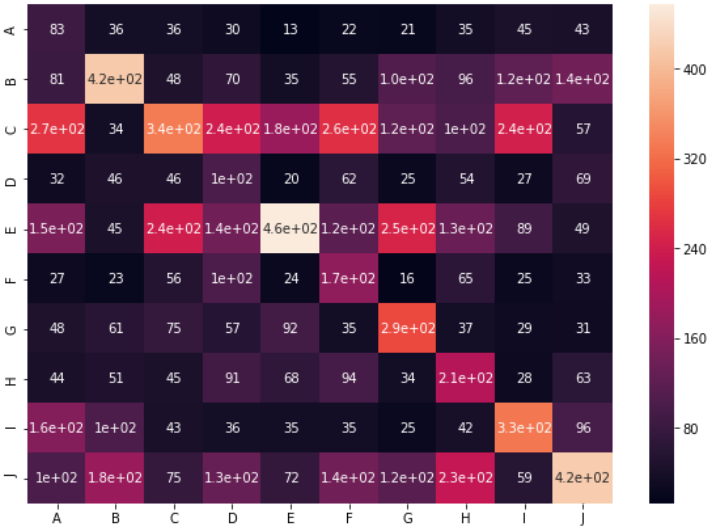
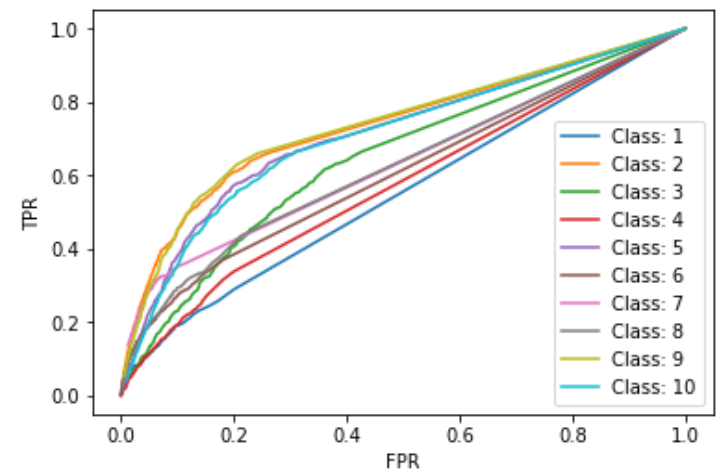
Dataset 1

Eigen Energy	Confusion Matrix	ROC Curve
70 1		
90 13		
95 30		

99 94		

Dataset 2

Eigen Energy	Confusion Matrix	ROC Curve
70 11		
90 76		

95 **159**		
99 **435**		

Ques 2

BAGGING ACCURACY TABLE(%)

Configuration (10 Weak Learners)	WL Mean	WL Standard Deviation	Accuracy on Training Data	Accuracy on Test Data
Bagging	15.38	0.0054	19.46	19.28
Bagging(Min-Max Normalization)	15.30	0.0046	27.66	26.66
Bagging(Z-Score Normalization)	15.26	0.0061	27.52	25.74
Bagging(Tanh Normalization)	15.24	0.0070	27.82	26.18

Normalization equi-weights the prediction contribution of each weak classifier to final prediction so uneven/unusual prediction do not get an uneven say in the final prediction and we get the most common/popular prediction and not the most uneven one.

BOOSTING ACCURACY TABLE(%)

Boosting Configuration	Cross Validation				Mean Error of all Learners	Test
	Mean	Standard Deviation	Best Model on Validation Dataset	Best Model on Test Dataset		
10 Weak Learners	43.52	0.036	46.25	45.90	86.76	42.04
100 Weak Learners	56.90	0.016	46.25	45.92	91.26	50.50
250 Weak Learners	60.84	0.022	44.14	43.12	89.58	54.12
500 Weak Learners	62.36	0.016	43.34	42.28	89.78	60.12