



HARIS RAMZAN
19I-2118

Section: A

Report

March 13, 2020

Question1:**What features did you use and why?**

Feature covariance and eigen decomposition technique is used to select the feature from the dataset. When two columns will have a positive value of covariance, they will be reduced to the single feature which will save our time and memory to train the model.

Question2:**What classification technique did you try?**

I have tried out multiple classification technique to train the model which includes SVM, KNN and LDA to check the accuracy of different model by doing train test split to get the idea which model is performing better. SVM showed good results.

Question3:**Which of the method (and for what parameters) showed best cross validation performance?**

SVM showed better results when kernel is changes from poly to rbf accuracy of model was around 72%. In the case of KNN the accuracy was decreased when the value of K is increased.

```
[Parallel(n_jobs=1)]: Done 144 out of 144 | elapsed: 15.5min finished
SVC(C=10, cache_size=200, class_weight=None, coef0=0.0,
    decision_function_shape='ovr', degree=3, gamma=0.01, kernel='rbf',
    max_iter=-1, probability=False, random_state=None, shrinking=True,
    tol=0.001, verbose=False)
```

Question4:**What test accuracy are you expecting?**

Around 70%