**CSE 6363 MACHINE LEARNING**

**HOMEWORK-3**

**A REPORT ON THE IMPLEMENTATION OF KNN, PCA, LDA, KERNEL SVM FOR FACE RECOGNITION**

Submitted by,

Vivek Arvind Balaji

UTA ID: 1001108300

**Introduction:**

Face Recognition is one of the most challenging and interesting classification problems in machine learning. Face recognition is a multiclass classification problem where the classifier is trained with the image vectors, weight vectors, etc., and the class of a new image is predicted. This project focuses on face recognition problem with respect to 5-fold cross validation, dimensionality reduction based on benchmark algorithms like K – Nearest Neighbor, Principal Component Analysis, Linear Discriminant Analysis and Kernel Support Vector Machine.

**Implementation:**

The following are the steps that were followed during the implementation of KNN, PCA, LDA and Kernel SVM for face recognition.

* Due to the implementation constraints, the dataset is stored in two formats., one in which image files are within the respective folders (this is in a folder named ‘data’) and the second one in which the image files are present in a single folder (this is a folder named ‘Dataset’).
* Every task in this project is trained and tested using 5 – Fold Cross Validation.
* Task 1 – KNN is implemented with respect to PCA. The images are centerized and the Eigen values are plotted out.
* Task 2 – The images are resized from 112x92 to 56x46 and Task 1 is repeated.
* Task 3 – PCA in task 1 is replaced by LDA and Task 1 is repeated.
* Task 4 – Task 3 is repeated but both PCA and LDA are applied in sequence to reduce the image dimensionality.
* Task 5 – Kernel SVM is implemented using 5 – fold cross validation. I have used the polynomial kernel, since it is the optimum kernel for face recognition problem.
* Task 6 – Task 5 is repeated but the image dimensionality is first reduced using PCA.

**System Requirements:**

1. Matlab R2014a or higher.
2. Optimization Toolbox.

**Execution Instructions:**

1. Extract the code package.
2. Open Matlab.
3. There must be three folders in the zip file. The first folder ‘code’ contains all the executable Matlab files. The other two folders are the two formats of data sets, ‘data’ and ‘Dataset’ where the data is present in the respective folders and in the same folder respectively.
4. Before executing the files, paste the URI of the folder ‘Dataset’ in the file ‘Vivek\_LDA.m’ and paste the URI of the folder ‘data’ in all other files.
5. Execute the project by beginning with the file name ‘Task1\_Vivek\_PCA\_KNN.m’ and then continuing with the files in order of their Task order. There must be six files namely ‘Task1\_Vivek\_PCA\_KNN.m’, ‘Task2\_Vivek\_KNN\_resize.m’, ‘Task3\_Vivek\_LDA\_KNN.m’, ‘Task4\_Vivek\_PCA\_LDA\_KNN.m’,‘Task5\_Vivek\_KERNEL\_SVM’and‘Task6\_Vivek\_PCA\_KERNEL\_SVM.m’.
6. The other files are the functions which will be invoked by other functions. Those need not be executed separately.
7. The tasks involving PCA and LDA will take a long time to get executed. So other tasks can be executed first if needed.
8. Execute each file by clicking Run on the toolbar in the matlab window.