***Machine Learning Homework 2***

Topic: Linear SVM

|  |  |
| --- | --- |
| Name: | Abu Ayub Ansari Syed |
| UTA ID: | 1001297330 |
| Course ID: | CSE 6363 |
| Email: | abuayubansari.syed@mavs.uta.edu |

**Implementation**:

Formatting the data:

I have used the previously formatted data i.e. from images to csv file. I have also added the python file that does the same. The complete data is in the *attfull.csv* file. I have chosen the first five elements from each class into the train matrix. During the future steps we use this for the testing as well. I have chosen the last five elements from each class into a test matrix. During the future steps we use this for the training as well.

Training the classifier:

I have used the train/test matrix as a parameter to the *trainSVM* function. This function creates the required parameters for the quadratic programming language i.e. **quadprog**. The solution **α** is used to calculate the w and w0. These forms the 40 linear boundaries to all the classes.

Testing the classifier:

Use the other matrix, to serve as the training data. The samples are tested one by one, by substituting them to the linear equation. We take the max out of the all the 40 comparisons are classify them to the respective class. We them compare with the actual class and count the correct predictions.

Accuracy:

Based on the count from the test, calculate accuracy with respect to the samples was tested. I run the similar steps by interchanging the train and test data. Final Accuracy is the average of the two.

**Execution:**

* Program code is in **LinearSVM.m**
* Keep the data in the same folder as the matlab file.
* The data locations are hard coded.
* Coded in MATLAB- R2016b, using optimization tool for Quad Prog function.
* Straight forward execution of the matlab file is sufficient.

**Results**:

The accuracy achieved is 94.7500%.

The files in the same folder have few experimental results. The files are as follows:

Attfull.csv: Full data of the images into the matrix form.

z.csv: Class data for the train and test datasets.

results\_fulldata\_normalised.pdf: Contains the MATLAB API result console of the program execution.

Att\_faces(Folder): actual images for Dataset.

Imagedatatoxlsheet.py: Python file needed to convert the Images to pixel data in CSV file.