PRML Assignment 1:

Human Activity Recognition using Smartphones

By:
Devendra Kumar - 13EC115
Karishma Mulani - 13EC124
Kevin D Souza - 13EC256
S Giritheja - 13EE141
(Group 10)

1. Introduction (About the Data Set):

This data consists of information which can be used to describe human activity using accelerometer, gyroscope meter already present in a smartphone. The experiments have been carried out with a group of 30 volunteers within an age bracket of 19-48 years. Each person performed six activities (WALKING, WALKING_UPSTAIRS, WALKING_DOWNSTAIRS, SITTING, STANDING, LAYING) wearing a smartphone (Samsung Galaxy S II) on the waist. Using its embedded accelerometer and gyroscope, we captured 3-axial linear acceleration and 3-axial angular velocity at a constant rate of 50Hz.
The information can be utilised to determine if a person is walking, walking up, walking down, sitting, standing or laying.

2. Data Visualisation:

The figure shown below is the plot of the two features with the maximum covariance. This was computed by taking the maximum covariance between the features from the covariance matrix of the data set.

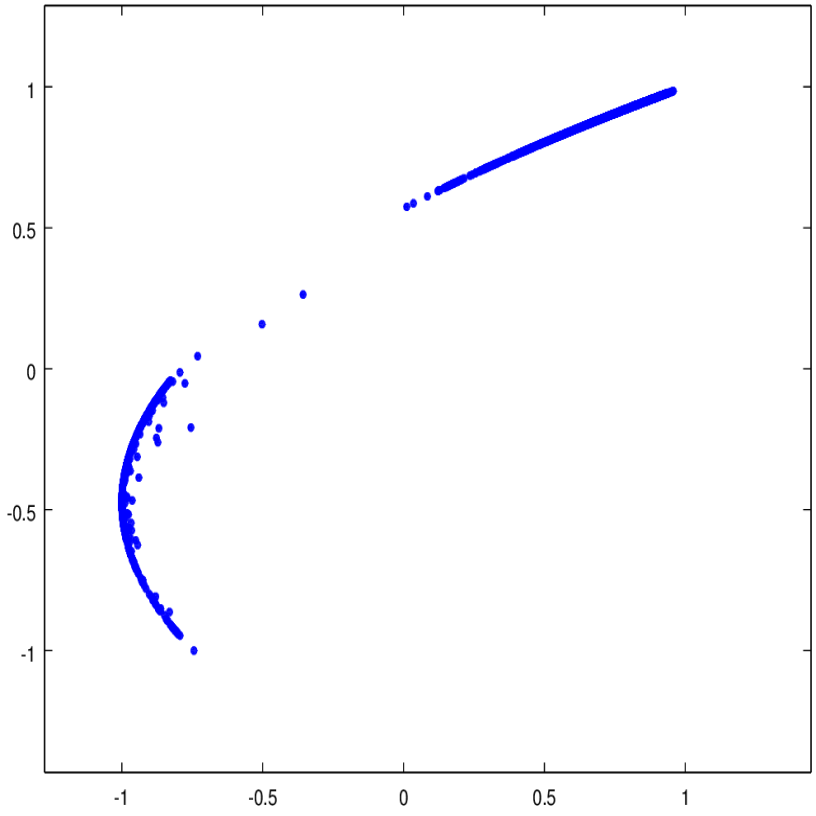

Figure 1: Plot of two features with Maximum Covariance tBodyAcc-mean()-Y and tGravityAcc-energy()-X

The following is the histogram of the feature with the least variance. We observed that most of our features are concentrated in particular values.

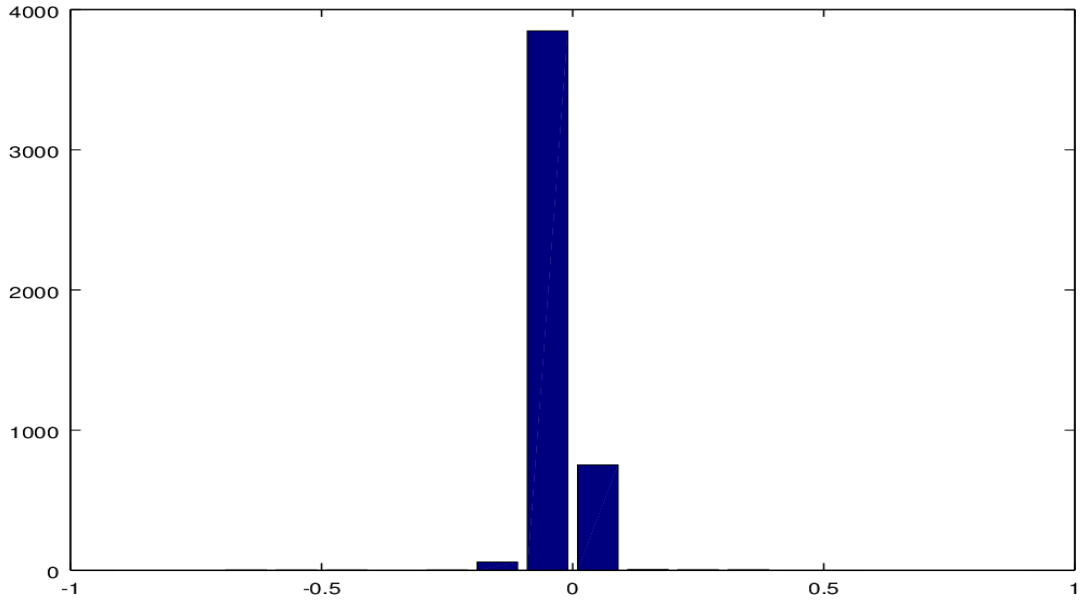

Figure 2: Histogram of feature with Minimum Variance

We performed a Prinicipal Component Analysis and got the following scatter plot. The class LAYING is in cyan and can be clearly distinguished from the classes WALKING(Blue), WALKING DOWNSTAIRS(Green) and WALKING UPSTAIRS(Red).

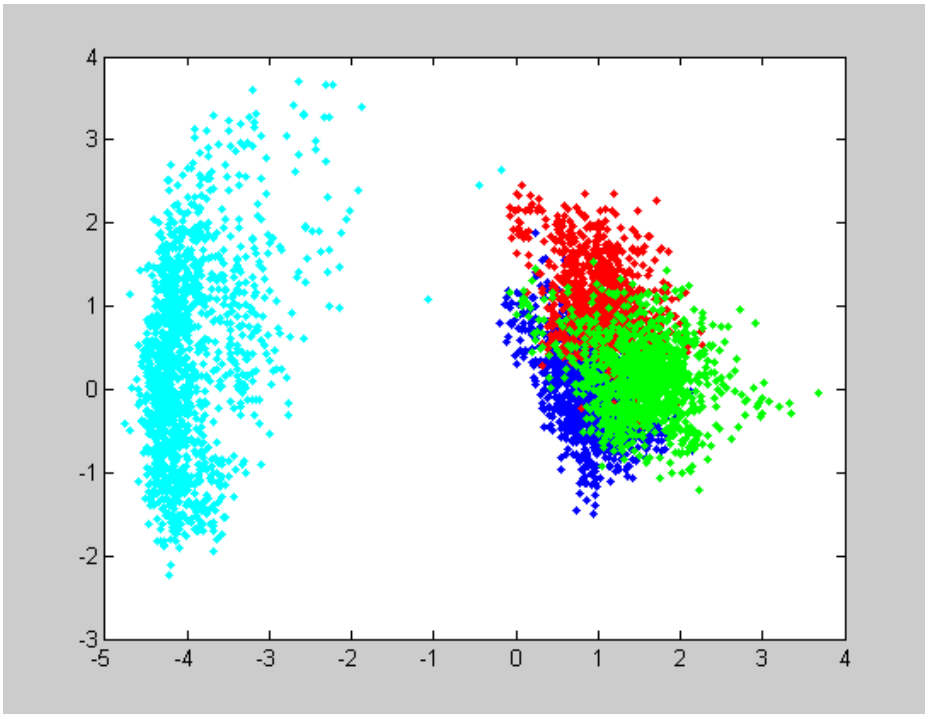

Figure 3: Plot of Individual classes

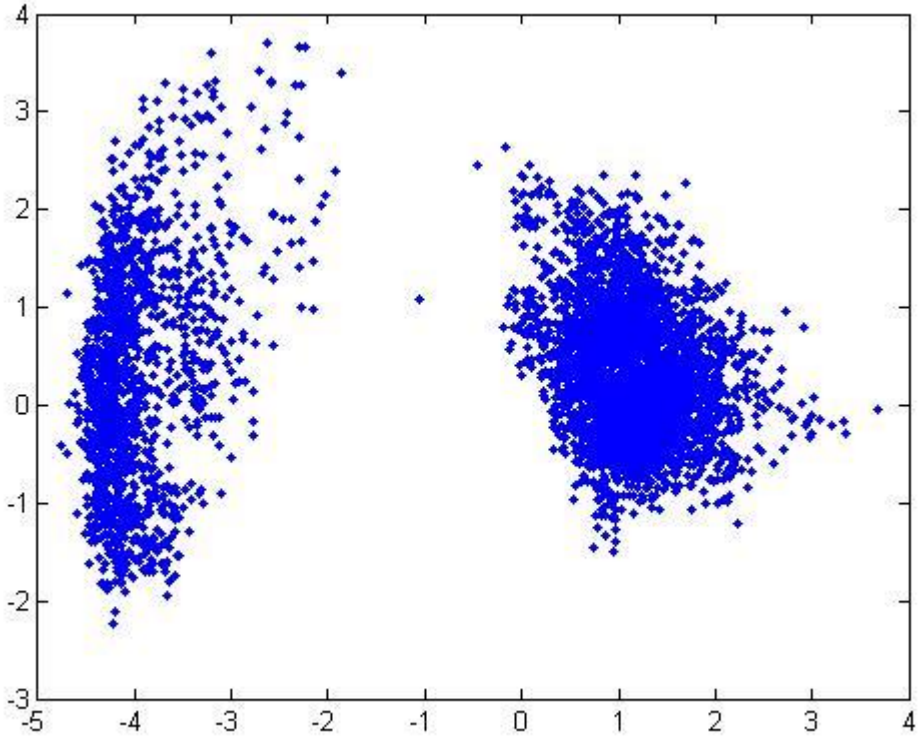

Figure 4: Plot of PCA by Matlab function

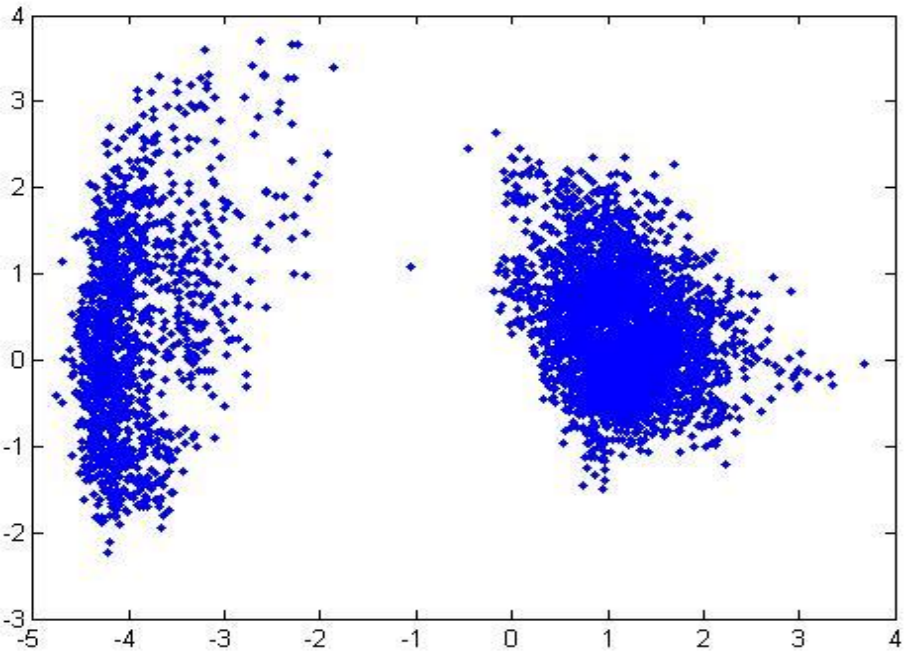

Figure 5: PCA obtained by projecting data on the first two Eigen vectors of the Covariance Matrix

The covariance matrix of this dimensionally reduced data had its off-diagonal elements as zero as expected. The covariance matrix was:

5.7882 0.0000
0.0000 0.7084

3. BAYES CLASSIFICATION:

We assumed that our distribution for each of the class was a Multivaritate Gaussian distribution so we found out the conditional probability using the mvnpdf() function in Matlab and multiplied with the prior probabilities so as to make a Bayes' desicion for classification. The figure below shows the 2D plot of the decision boundry.

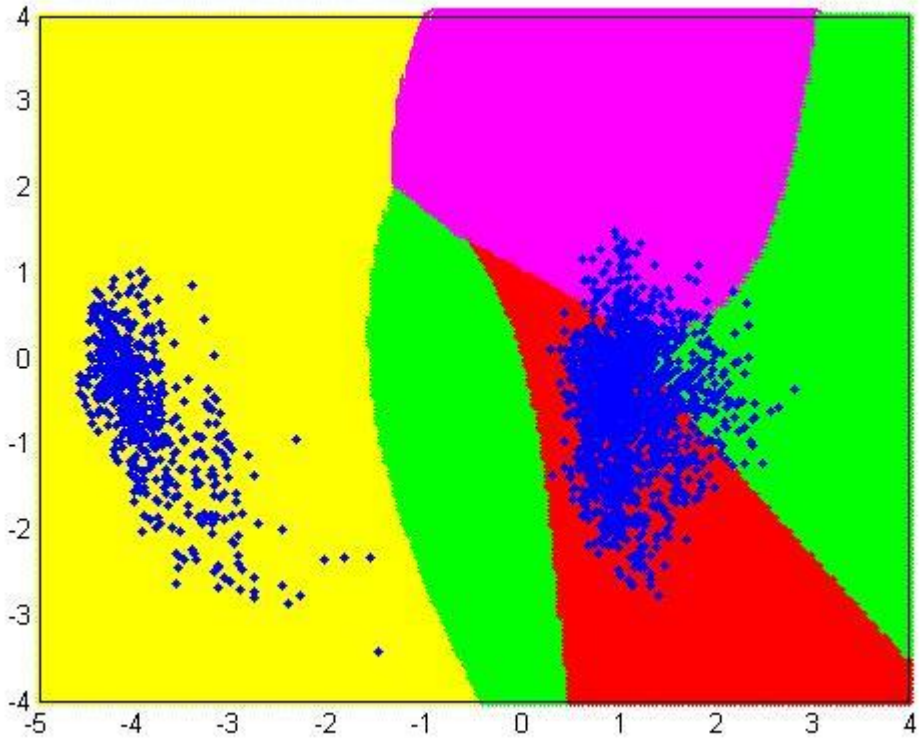

Figure 6: Plot of Decision Boundary

Observations:

1. The number of features should be optimal.

2. Gaussian Model is not best fit for the given data therefore find a better model to describe the dataset.

3. Scale the input data appropriately to improve precision and avoid computation problems such as singularity.

4. Dimensionality reduction results in slightly smaller efficiency even though the computation complexity is reduced.

5. The results of PCA function by Matlab and our own function using covariance matrix and Eigen vectors gave the same result.

Results:

1. The classifier worked with an efficiency of about 76% without dimensionality reduction.

2. The classifier worked with an efficiency of about 56% with dimensionality reduction.

Problems faced:

1. Due to too many features the determinant of Covariance matrix became zero.

2. The covariance matrices of some classes were not Positive definite matrices.

Approach Taken:

1. Reduced the number of features.

2. Removed the Classes whose covariance matrix were not positive definite