Assignment 3 (data mining)  
 USER DEPENDENT ANALYSIS

Kishore Madhava Muruganandan(ID:1212457608)

Manish Tummala (ID: 1213350526)

Manjusha Ravindranath(ID:1212109143)

Nikhil Singh(ID:1213158633)

Venkatakrishnan Karthikeyan (ID:1213195345)

# 1.Introduction

The feature matrix that we got from Assignment 2 was multiplied with the PCA output and the resultant was used as input for the machines in Assignment 3. The machines used were –

1. Decision Tree
2. Support Vector Machine
3. Neural Network

Based on the predicted labels for each file of each user, we calculated the accuracy metrics of Precision, recall and F1 score.

2. Implementation

2.1 Machine 1

Decision trees

Input Predictor: GLX, GLY, GLZ, GRX, GRY, GRZ are the gyroscope PCA features obtained in Assignment 2.

**Training vs Testing Ratio**: 60/40

Used fitctree function in Matlab which has this syntax:-

Mdl7 = fitctree(T1, T2,'CrossVal','on');

**Dimensions of feature matrix:**

Training Data:

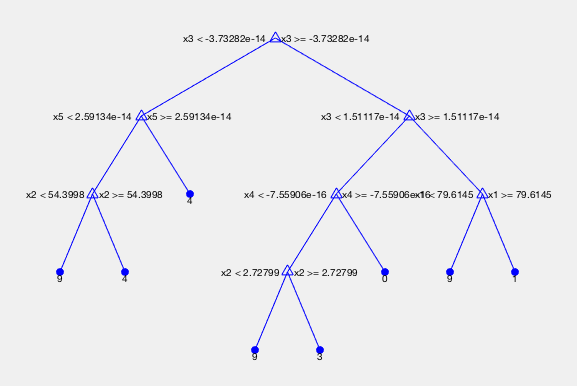
For training, the 60% data of 5 users are taken.

Testing Data:

For testing purpose, the rest 40% of 5 users data is taken.

T1- input with training data of 5 users excluding labels

T2-input with labels of training data of 5 users



**Accuracy –**

For User1 – 63.4%.

For User2 – 66.6%

For User3 – 84.8%

For User4 – 68.9%

For User5 – 88.4%

2.2 Machine 2

SVM

Input Predictor: GLX, GLY, GLZ, GRX, GRY, GRZ are the gyroscope PCA features obtained in Assignment 2.

**Training vs Testing Ratio**: 60/40

Used fitctree function in Matlab which has this syntax:-

Mdl7 = fitceoc(T1, T2,'CrossVal','on');

**Dimensions of feature matrix:**

Training Data:

For training the 60% data of 5 users are taken.

Testing Data:

For testing purpose, the rest of 40% of users data is taken.

T1- input with training data of 5 users excluding labels

T2-input with labels of training data of 5 users

**Accuracy –**

For User1 – 63.4%.

For User2 – 66.6%

For User3 – 84.8%

For User4 – 68.9%

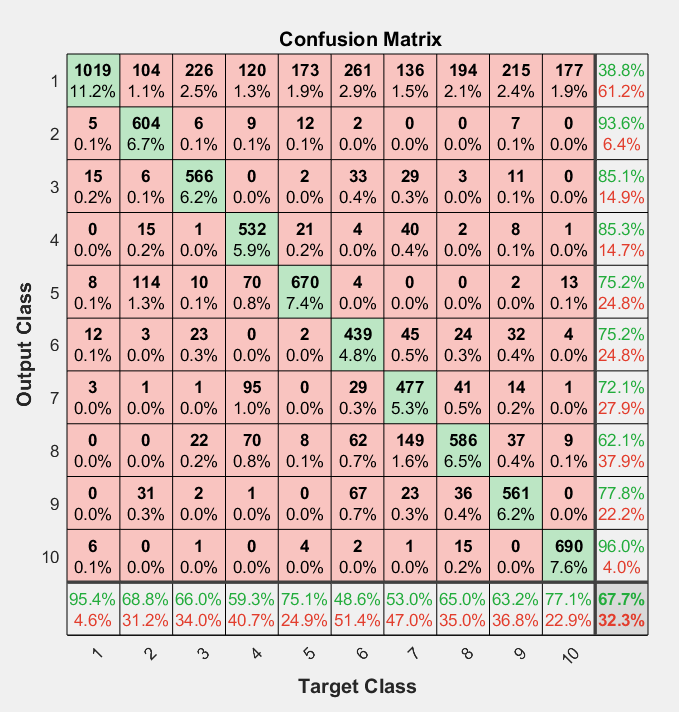
For User5 – 88.4%

**2.3 Machine 3**

Neural Networks

Input Predictor: GLX, GLY, GLZ, GRX, GRY, GRZ are the gyroscope PCA features obtained in Assignment 2.

**Training vs Testing Ratio**: 60/40



Used the neural network toolbox in matlab which can be started with the command:

nnstart

**Dimensions of feature matrix:**

Training Data:

For training the 60% data of 5 users are taken.

Testing Data:

For testing purpose, the rest 40% of 5 users data is taken.

Input data was the training data from 5 users without the labels.

Target data was a matrix with 0s and 1s, with 1s present in the position of the correct class.

25 hidden neurons were used to get the best classification.

**Accuracy –**

For User1 – 40.4%.

For User2 – 66.6%

For User3 – 84.8%

For User4 – 68.9%

For User5 - 65.1%