PROJECT REPORT

Activity recognition with healthy older people using a battery less wearable sensor Data Set

GROUP NO: 16

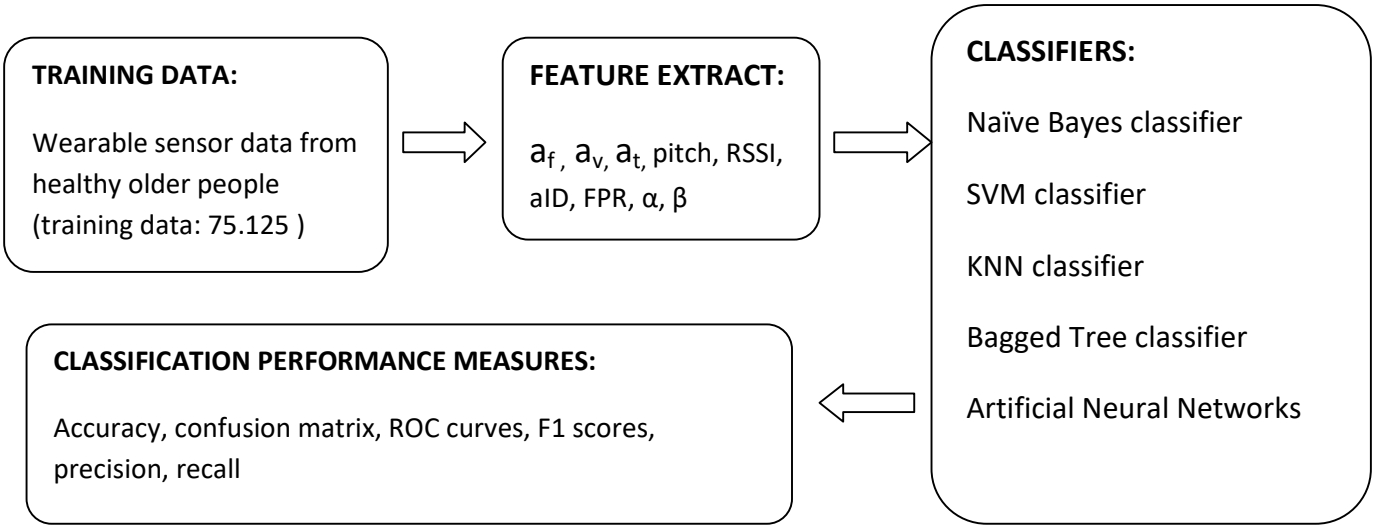
GROUP MEMBERS:

T.S.K.Charitha – S20170020242
T.Chandana –S20170020246
J.Lasya Kumari –S20170020209

INTRODUCTION:

Falls in hospitals are common, therefore strategies to minimize the impact of these events in older patients and needs to be examined. The system consists of a machine learning based activity classifier and gives an indicator that on an average the minimum duration for which a particular activity is performed by a person. Sequential motion data from 14 healthy older people aged 66 to 86 years old using a battery less, wearable sensor on top of their clothing for the recognition of activities in clinical environments.

CLASSIFICATION ARCHITECTURE:



FEATURES EXTRACTED:

Based on the correlation results between the acceleration attributes over the entire dataset frontal acceleration attribute was removed. The results are shown below:

Correlation of lateral and frontal accelerations= 0.4013

Correlation of lateral and vertical accelerations = 0.1600

Correlation of vertical and frontal accelerations = -0.9121

Features chosen:

FEATURES	DESCRIPTION
Frontal acceleration(a_f)	Left-right axis acceleration values in g.
Vertical acceleration(a_v)	Antero-posterior axis acceleration values in g.
Sine of body tilting angle(pitch)	Sine of body tilting angle(Θ) towards the front or back with respect to vertical in the mid sagittal plane, approximated as $\sin(\arctan(a_f/a_v))$.
Received power signal(RSSI)	The measured strength of theb received backscattered signal at RFID radar antenna, called received signal strength indicator(RSSI).
Resultant acceleration	Magnitude of acceleration vector given by $a_t = \sqrt{{a_f}^2 + {a_v}^2 + {a_I}^2}$.
Trunk yaw angle	Rotational angle from dorsoventral axis, approximated as: $\beta \approx \arctan(a_I/a_f)$.
Trunk roll angle	Tilting angle in the coronal plane, approximated as: $\alpha \approx \arctan(a_I/a_v)$.
Antenna collecting minimum power	ID of antenna (aID) with minimum received power in segment.
FPR	Frequency Phase Ratio (FPR) = $\frac{\Phi(t)}{F(t)}$, F represents frequency values in MHz of the channels used by the RFID reader to query the sensor; and RF phase (Φ), the phase angle between the transmitted RF carrier frequency F and the backscattered signal from the sensor.

CLASSIFICATION IMPLEMENTATION:

Gaussian SVM Classifiers: Since the data is non linear we have chosen the Gaussian SVM as basic classifier.

KNN Classifiers: KNN being most simple algorithm and also does not assume anything about data in prior and the model structure is completely determined by data KNN was taken as classifier.

Bagged Tree Classifiers: By observing data we can understand that variance is high so we have taken bagged trees as a classifier. Also the advantages of bagged trees classification model is the ability to handle higher dimensionality data very well we can say beforehand that bagged trees will give good accuracy.

Naive Bayes Classifiers: Naive Bayes is classification algorithm that is suitable for multi-class classification. It assumes that every feature is independent and since we have removed correlated features and so our assumption is that the features are independent and hence Naive Bayes should give very good accuracy.also naïve bayes performs well in categorical data compared to regression.

ANN: It was taken to know how neural networks will classify the data compared to the other chosen classifiers.

Results including ROC curves, confusion matrix, F1 scores:

FINE GAUSSIAN:

confmat2(TRAIN DATA) =

5458	0	0	0
0	2067	0	0
0	0	10575	0
10	0	208	352

accuracy_FineGaussian = 98.5536

confmat2test =

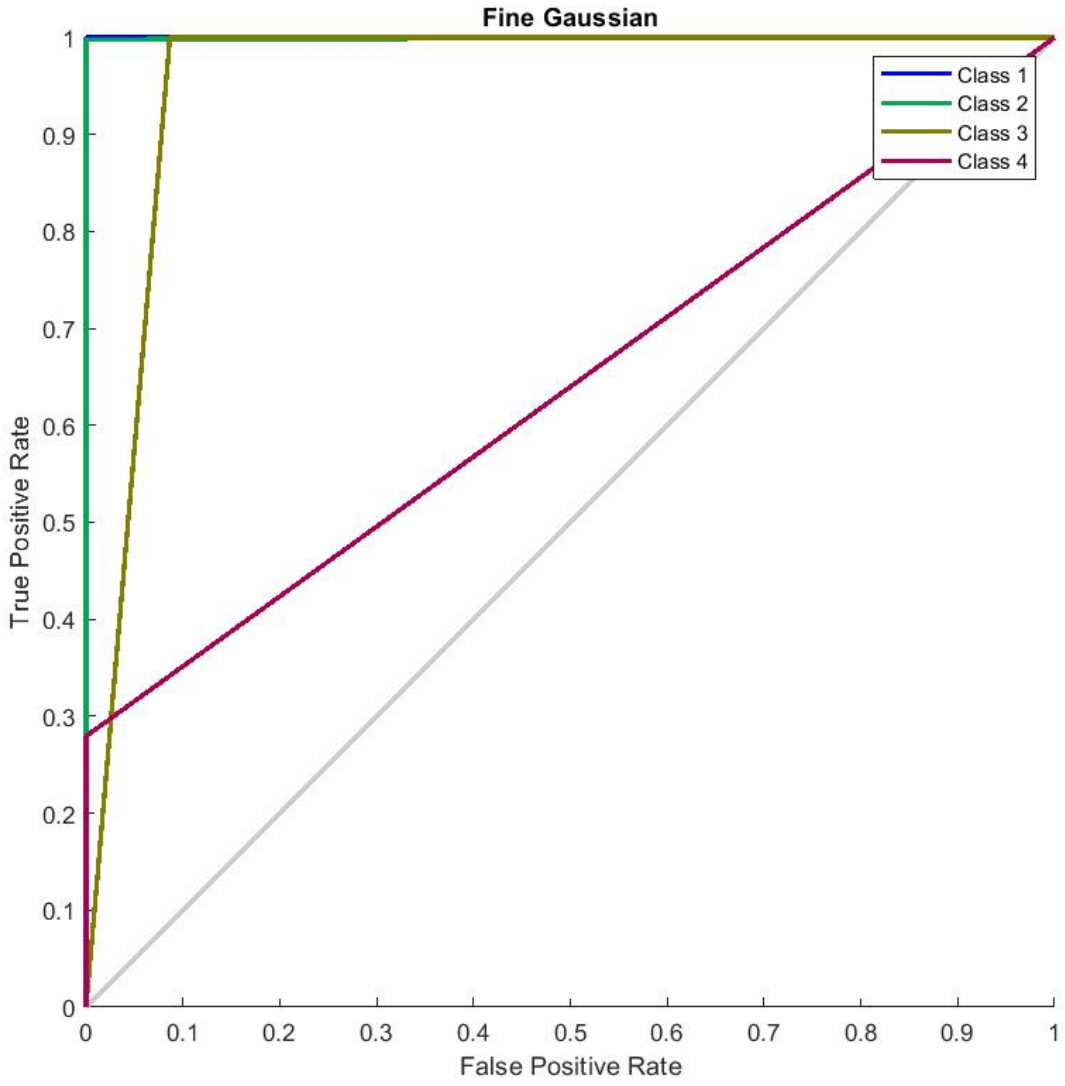
5042	9	6594	0
533	1589	1264	0
2	0	32643	0
52	1	1168	269

accuracy_FineGaussian_test = 72.8522

precision = 0.9955

recall = 0.8194

F1-score = 0.8989



FINE KNN:

confmat3 =

```
        5458         0         0         0
           0      2067         0         0
           0         0     10575         0
           0         0         0       570
```

accuracy_FineKNN = 100

confmat3test =

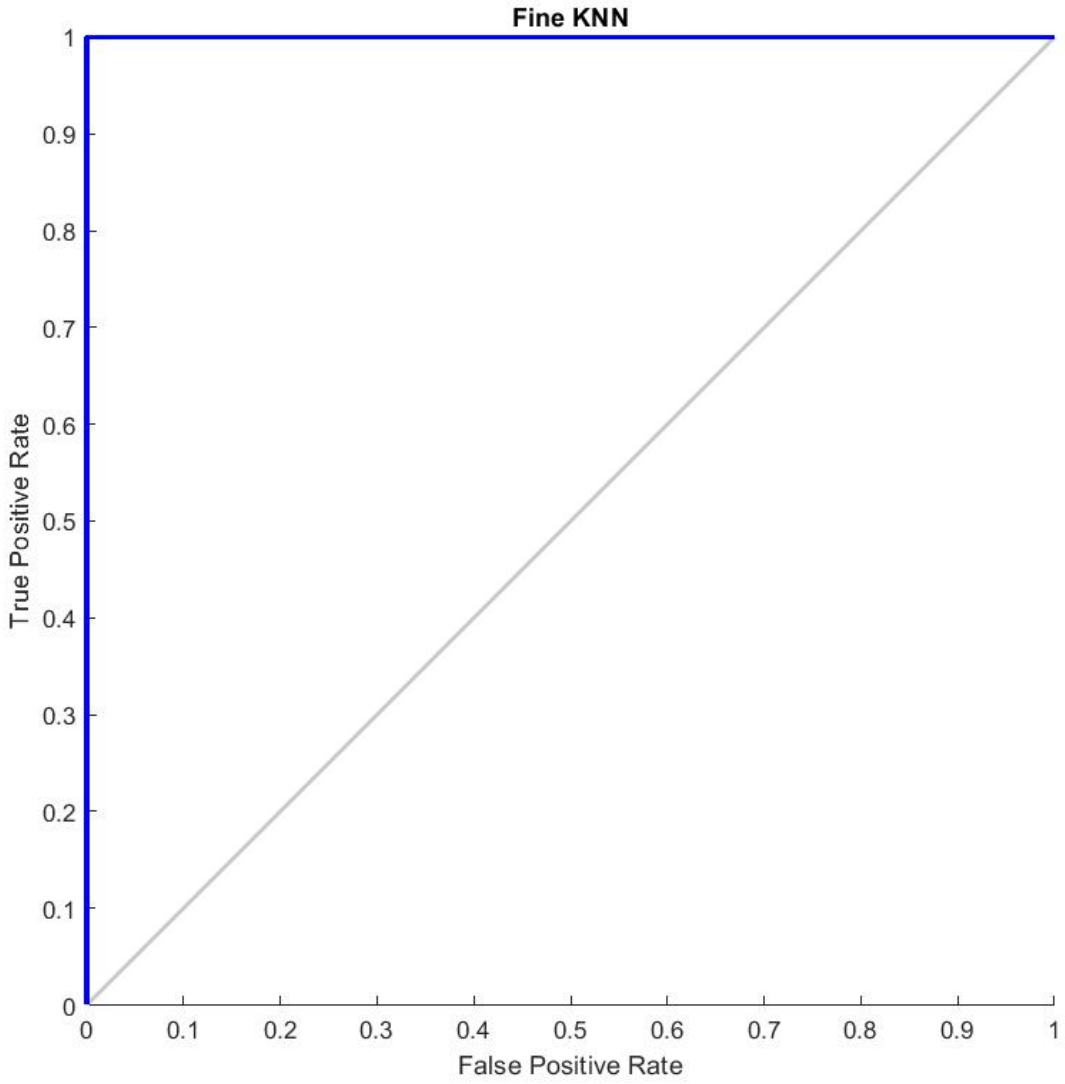
```
       10420       490       102       633
         399      2920         0        67
        9125        58     23460         2
         507       102         1       880
```

accuracy_FineKNN_test = 88.0216

precision = 0.5691

recall = 0.6217

F1-score = 0.5942



NAÏVE BAYES:

confmat8 =

4950 293 45 170

467 1506 0 94

0 4 10571 0

213 25 6 326

accuracy_NaiveBayes = 96.6193

confmat8test =

10299 495 259 592

947 2256 0 183

6 3 32613 23

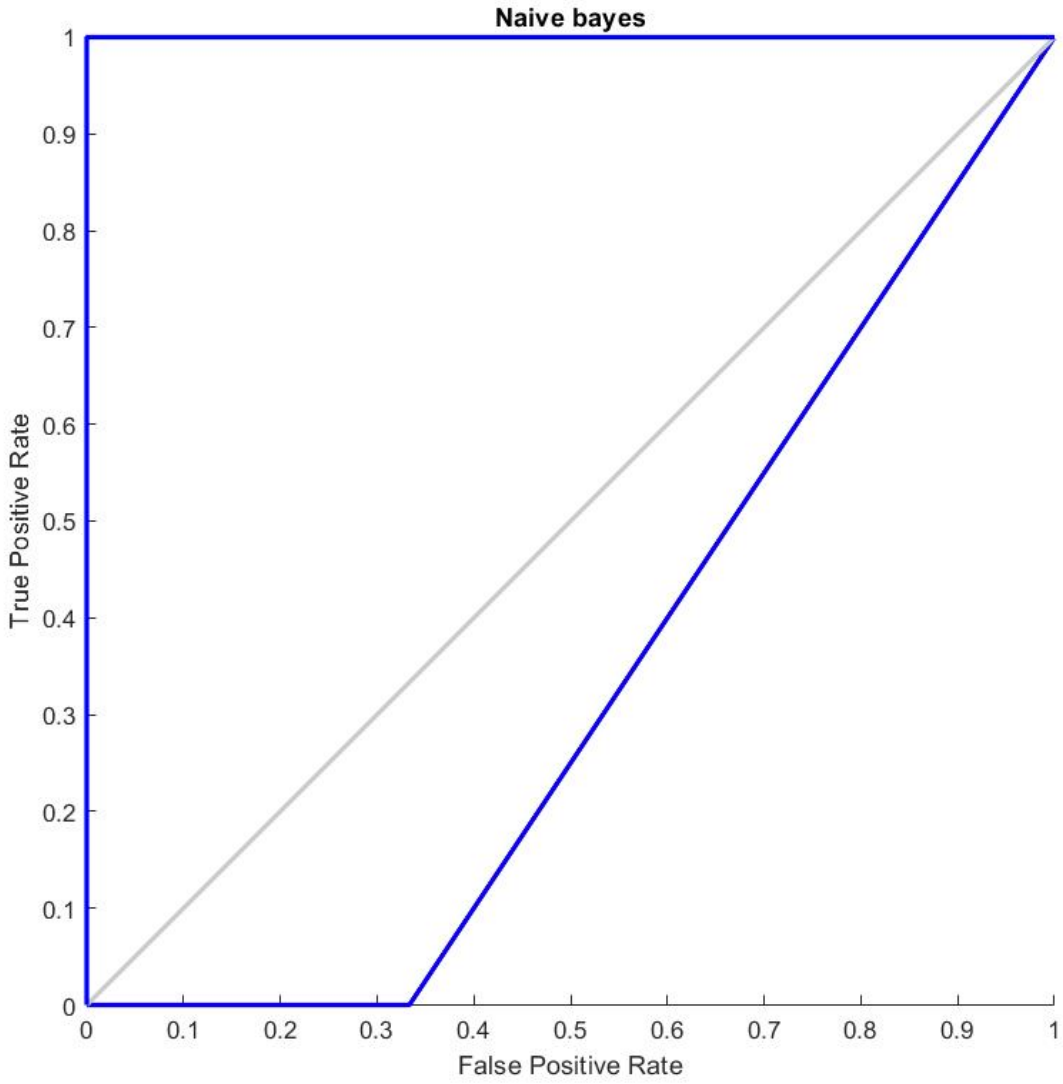
623 82 0 785

accuracy_NaiveBayes_test = 93.4650

precision = 0.6674

recall = 0.6550

F1-score = 0.6612



BAGGED TREES:

confmat9 =

```
        1124           0           0           1
           0         553           0           0
           0           0        9570           0
           0           0           0         229
```

accuracy_baggedtrees = 99.9913

confmat9test =

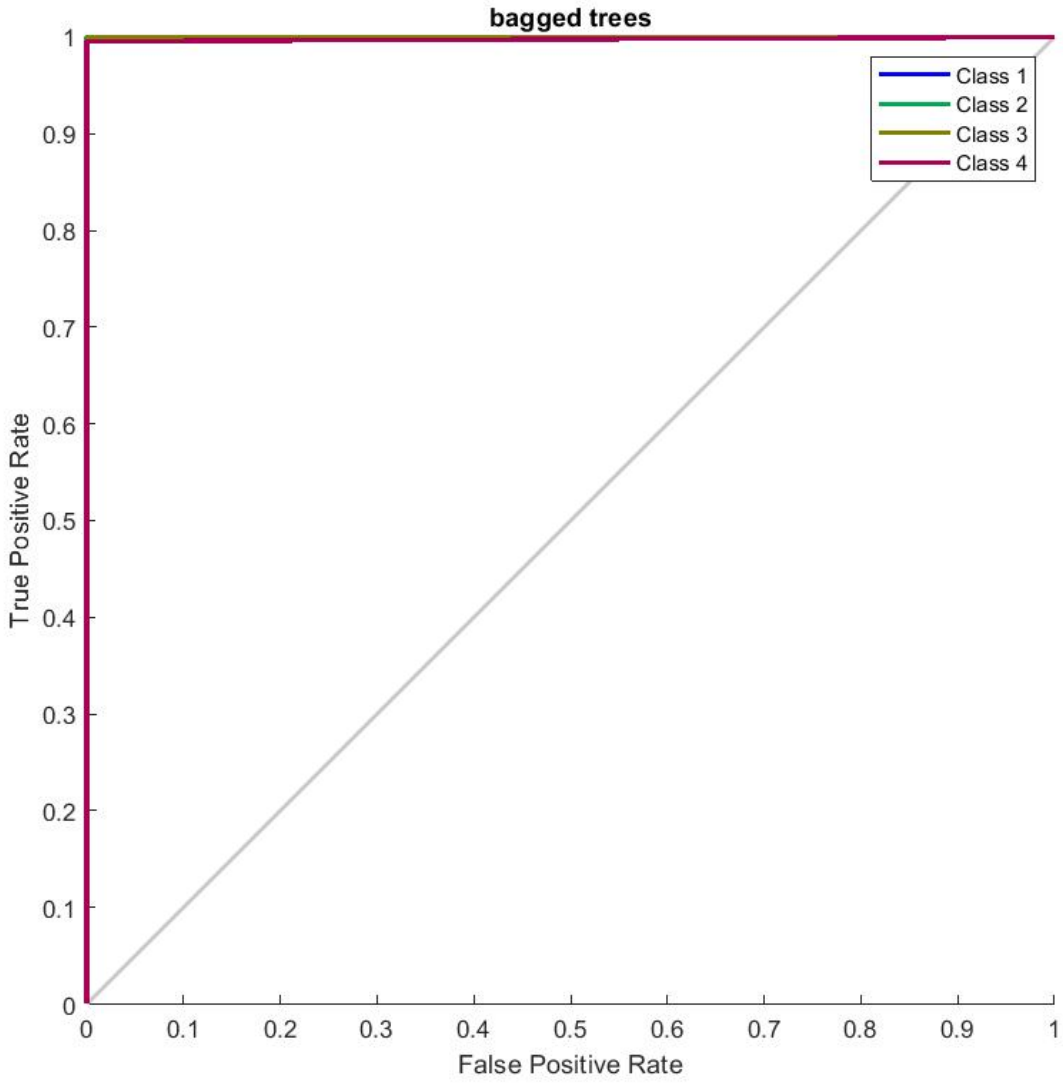
```
        5458           0           0           0
           1        2066           0           0
           0           0       10575           0
           2           0           0         568
```

accuracy_baggedtrees_test = 91.9877

precision = 0.6090

recall = 0.6517

F1-score = 0.6297



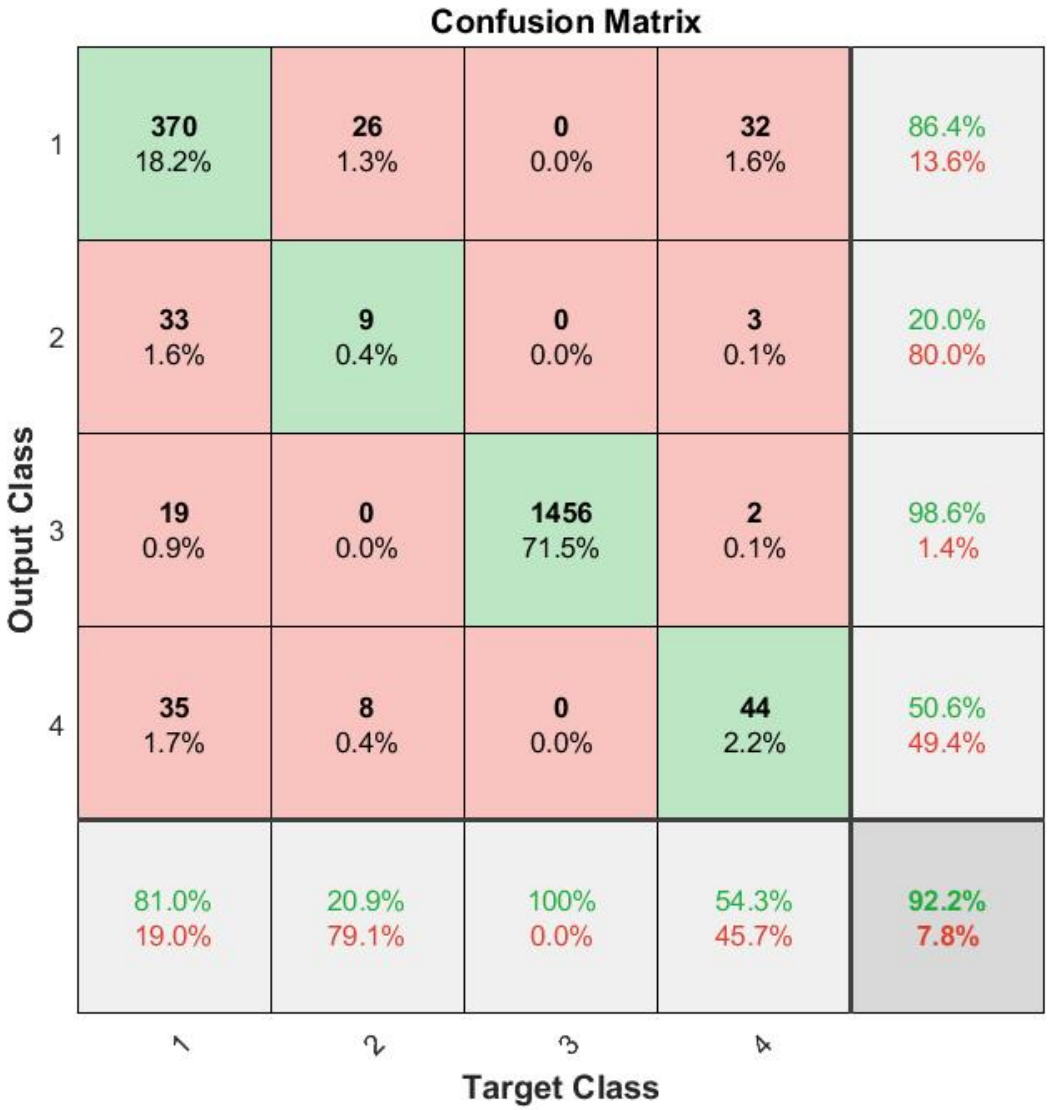
ARTIFICIAL NEURAL NETWORKS:

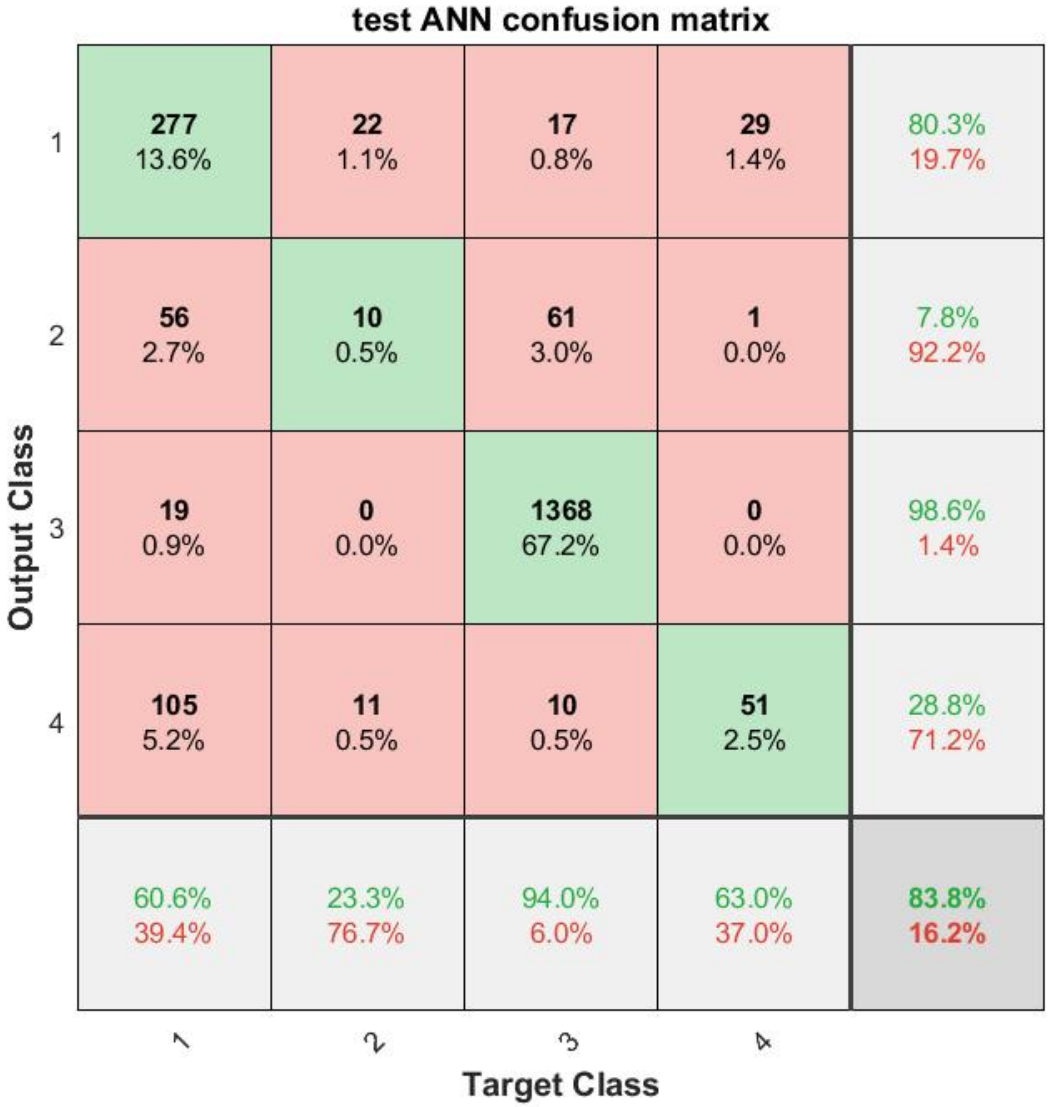
precision = 0.5389

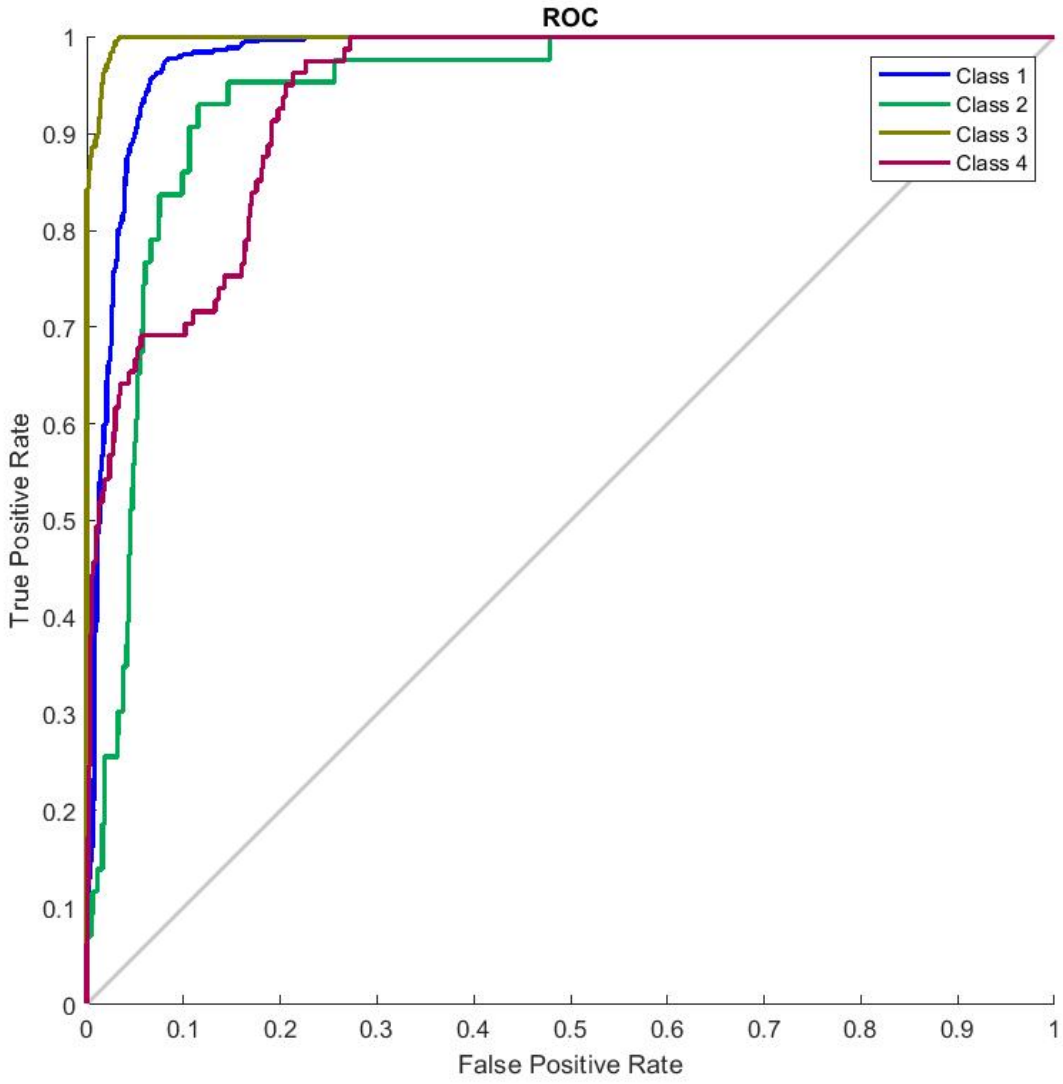
recall = 0.6020

F1-score = 0.5687

Confusion matrix of ANN




test ANN confusion matrix
Output Class
1
2
3
4
277
13.6%
22
1.1%
17
0.8%
29
1.4%
80.3%
19.7%
56
2.7%
10
0.5%
61
3.0%
1
0.0%
7.8%
92.2%
19
0.9%
0
0.0%
1368
67.2%
0
0.0%
98.6%
1.4%
105
5.2%
11
0.5%
10
0.5%
51
2.5%
28.8%
71.2%
60.6%
39.4%
23.3%
76.7%
94.0%
6.0%
63.0%
37.0%
83.8%
16.2%
1
2
3
4
Target Class



Main inference and other results from the study including comparison with other classifiers and previous methods in literature:

Classification done until now finds the activity of an old person at the instant of time when the sensor gives the attribute values, but for clinical purposes we have to know for how long on an average a

particular activity is performed, so we have also found those values. The results being the minimum time in general for a person to

(i) Sit on bed is N1 = 1.0000
(ii) Sit on the chair is N2 = 8.2500

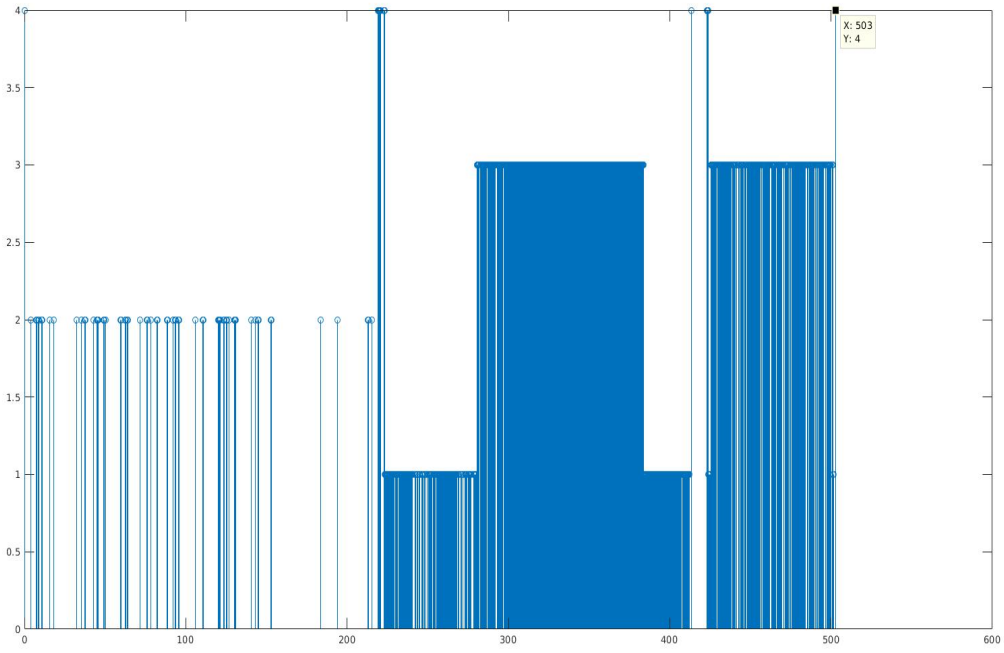
(iii) Lying on bed is N3 = 18.8750

(iv) Ambulating N4 = 1.2500

As the data is collected once by 4 antennas and then by 3 antennas so the interval between the data collection is not constant and can be in the range of 0.1 s to 1.3s.

You may verify these values of the minimum time for a person to perform a given activity.

PLOT OF ACTIVITY OF A TEST PERSON:



According to the literature the results when used Conditional random fields as the classifier:

Recall	60.50 ± 2.12
Precision	64.74 ± 2.26
F1-score	63.78 ± 1.87

According to the literature accuracies when the following classifiers where used

Using KNN : 87.0216

Using random forest:92.8402

Random forest being an extension of bagged trees has given better accuracy than bagged trees but naive bayes gave better accuracy than random forest.

CONCLUSION:

As expected SVM didn't give better results than other selected classifiers.KNN though is a good classifier suffered from over-fitting even if we have increased the number of nearest neighbors. Naive Bayes is best classifier as expected since we have removed the correlated features and so the final features are independent. Bagged trees was the second best classifier. ANN couldn't even classify as good as KNN.