

Assignment 4 (Group 9)

Alankrit Shah
Nikhil Lohia
Srinivas Murali
Vidhi Patel

PHASE 1 : User dependent analysis

This is the training set from the features obtained which consists of 60% of the dataset. We train this dataset with 3 machines below. The remaining 40% of the dataset is the testing set from the features obtained. We test this dataset again using the 3 machines below. We use accuracy metrics such as Precision, Recall, F Score and ROC

a) Decision Tree

Closer the value to 1, higher is the accuracy metric. 1 being the ideal classifier.

Group	Precision	Recall	F Score	ROC (Area under the curve)
1	1.0000	0.8710	0.9310	0.9355
2	0.9688	0.9688	0.9688	0.9659
3	1.0000	0.9667	0.9831	0.9833
4	1.0000	0.9583	0.9787	0.9792
5	0.9143	0.9697	0.9412	0.9272
6	1.0000	0.9677	0.9836	0.9839
7	0.9630	1.0000	0.9811	0.9848
8	0.8929	0.9615	0.9259	0.9353
9	0.9655	1.0000	0.9825	0.9839
10	0.9630	1.0000	0.9811	0.9848
11	0.9565	0.9565	0.9565	0.9644
12	0.9000	1.0000	0.9474	0.9531

13	1.0000	1.0000	1.0000	1.0000
14	0.9667	1.0000	0.9831	0.9833
15	1.0000	1.0000	1.0000	1.0000
16	0.9615	0.8929	0.9259	0.9303
17	1.0000	1.0000	1.0000	1.0000
18	1.0000	0.9677	0.9836	0.9839
19	1.0000	1.0000	1.0000	1.0000
20	0.9688	1.0000	0.9841	0.9821
21	1.0000	1.0000	1.0000	1.0000
22	1.0000	1.0000	1.0000	1.0000
23	0.9688	0.9688	0.9688	0.9659
24	0.9655	0.8750	0.9180	0.9190
25	0.9643	0.9643	0.9643	0.9660
26	0.9583	0.9200	0.9388	0.9453
27	1.0000	0.9333	0.9655	0.9667
28	0.9412	1.0000	0.9697	0.9630
29	0.9677	0.9375	0.9524	0.9502
30	1.0000	0.9643	0.9818	0.9821
31	0.9355	1.0000	0.9667	0.9667
32	1.0000	1.0000	1.0000	1.0000
33	1.0000	0.9524	0.9756	0.9762

b) Support Vector Machine

Closer the value to 1, higher is the accuracy metric. 1 being the ideal classifier

Group	Precision	Recall	F Score	ROC (Area under the curve)
1	0.9688	0.9688	0.9688	0.9659

2	0.9655	1.0000	0.9825	0.9839
3	1.0000	1.0000	1.0000	1.0000
4	0.9655	0.9655	0.9655	0.9661
5	0.9286	0.8387	0.8814	0.8836
6	0.9630	0.8966	0.9286	0.9316
7	0.9643	1.0000	0.9818	0.9844
8	0.8929	1.0000	0.9434	0.9559
9	0.9706	1.0000	0.9851	0.9808
10	1.0000	1.0000	1.0000	1.0000
11	1.0000	1.0000	1.0000	1.0000
12	0.9355	1.0000	0.9667	0.9667
13	1.0000	1.0000	1.0000	1.0000
14	0.9355	1.0000	0.9667	0.9667
15	0.9655	1.0000	0.9825	0.9839
16	1.0000	1.0000	1.0000	1.0000
17	1.0000	1.0000	1.0000	1.0000
18	1.0000	1.0000	1.0000	1.0000
19	1.0000	1.0000	1.0000	1.0000
20	0.9600	1.0000	0.9796	0.9857
21	0.9688	1.0000	0.9841	0.9821
22	1.0000	1.0000	1.0000	1.0000
23	0.9667	1.0000	0.9831	0.9833
24	1.0000	0.9706	0.9851	0.9853
25	0.8387	1.0000	0.9123	0.9242
26	0.9667	1.0000	0.9831	0.9833
27	1.0000	0.9643	0.9818	0.9821
28	0.8710	1.0000	0.9310	0.9375

29	0.9375	0.9677	0.9524	0.9482
30	1.0000	1.0000	1.0000	1.0000
31	1.0000	0.9375	0.9677	0.9688
32	0.9667	1.0000	0.9831	0.9833
33	0.9167	1.0000	0.9565	0.9583

c) Neural Networks

Group	Precision	Recall	F Score	ROC (Area under the curve)
1	1.0000	0.9494	0.9644	0.9707
2	1.0000	0.9188	0.9484	1.0000
3	1.0000	1.0000	1.0000	1.0000
4	0.9677	0.9484	0.9484	0.9790
5	0.6667	0.8983	0.7565	0.7758
6	0.9118	1.0000	0.9348	0.9721
7	0.9429	0.9512	0.9374	0.9771
8	1.0000	0.6806	0.8033	0.9790
9	0.9000	0.9124	0.8969	0.9631
10	0.9412	1.0000	0.9503	1.0000
11	0.7500	0.5345	0.6189	0.7148
12	1.0000	0.8852	0.9302	0.9790
13	0.7500	0.9450	0.8269	0.9420
14	0.9706	1.0000	0.9654	0.9790
15	0.8966	0.7078	0.7840	0.8874
16	1.0000	0.9374	0.9582	1.0000
17	1.0000	0.9147	0.9462	1.0000

18	1.0000	1.0000	1.0000	1.0000
19	0.9677	0.9484	0.9484	0.9681
20	0.9667	1.0000	0.9634	0.9800
21	0.9412	0.9503	0.9361	0.9589
22	1.0000	0.8669	0.9200	0.9790
23	0.7750	0.9494	0.8439	0.9330
24	0.9583	0.7513	0.8348	0.9665
25	1.0000	0.9473	0.9634	1.0000
26	0.7895	0.9188	0.8400	0.9197
27	1.0000	0.7782	0.8675	1.0000
28	0.9189	1.0000	0.9386	0.9790
29	0.9722	0.9528	0.9528	0.9742
30	1.0000	1.0000	1.0000	1.0000
31	1.0000	0.9483	0.9639	1.0000
32	0.9705	0.9511	0.9511	0.9790
33	0.9062	1.0000	0.9318	1.0000

where,

$$\text{Precision} = \frac{\Sigma \text{True positive}}{\Sigma \text{Predicted condition positive}}$$

$$\text{Recall} = \frac{\Sigma \text{True positive}}{\Sigma \text{Condition positive}}$$

$$\text{Fscore} = 2 \cdot \frac{1}{\frac{1}{\text{recall}} + \frac{1}{\text{precision}}}$$

Receiver operating characteristic (ROC) = The ROC curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings.

PHASE 2 : User independent analysis

We have divided the user data based on users such that 10 users' data make the training data and the rest 23 users' data make up the testing data. We then trained the data using 3 machines and tested them against the testing data.

d) Decision Tree

Closer the value to 1, higher is the accuracy metric. 1 being the ideal classifier

Group	Precision	Recall	F Score	ROC (Area under the curve)
1	0.9459	1.0000	0.9722	0.9730
2	0.9730	0.9863	0.9796	0.9797
3	0.9865	1.0000	0.9932	0.9932
4	0.9459	1.0000	0.9722	0.9730
5	0.9730	1.0000	0.9863	0.9865
6	0.9324	1.0000	0.9650	0.9662
7	0.9595	1.0000	0.9793	0.9797
8	0.9459	0.9589	0.9524	0.9527
9	0.9595	0.9861	0.9726	0.9730
10	0.9189	1.0000	0.9577	0.9595
11	0.9730	1.0000	0.9863	0.9865
12	0.9865	1.0000	0.9932	0.9932
13	0.9595	0.9861	0.9726	0.9730
14	0.9459	0.9859	0.9655	0.9662
15	0.9189	0.9855	0.9510	0.9527
16	0.9054	1.0000	0.9504	0.9527
17	0.9459	0.9859	0.9655	0.9662
18	0.8514	0.9545	0.9000	0.9054
19	0.8919	0.9706	0.9296	0.9324

20	0.9730	0.9863	0.9796	0.9797
21	0.9595	0.9861	0.9726	0.9730
22	0.9865	1.0000	0.9932	0.9932
23	0.9483	1.0000	0.9735	0.9741

e) Support Vector Machine

Closer the value to 1, higher is the accuracy metric. 1 being the ideal classifier

Group	Precision	Recall	F Score	ROC (Area under the curve)
1	0.9865	1.0000	0.9932	0.9932
2	0.9730	0.9863	0.9796	0.9797
3	1.0000	1.0000	1.0000	1.0000
4	1.0000	1.0000	1.0000	1.0000
5	1.0000	1.0000	1.0000	1.0000
6	0.9459	0.9859	0.9655	0.9662
7	0.9865	0.9865	0.9865	0.9865
8	0.9865	1.0000	0.9932	0.9932
9	1.0000	1.0000	1.0000	1.0000
10	0.9730	1.0000	0.9863	0.9865
11	1.0000	1.0000	1.0000	1.0000
12	1.0000	1.0000	1.0000	1.0000
13	0.9865	0.9865	0.9865	0.9865
14	0.9595	0.9861	0.9726	0.9730
15	0.9595	0.9861	0.9726	0.9730
16	0.9865	1.0000	0.9932	0.9932
17	0.9459	0.9859	0.9655	0.9662
18	0.8514	1.0000	0.9197	0.9257

19	0.9459	0.9859	0.9655	0.9662
20	0.9865	0.9865	0.9865	0.9865
21	1.0000	0.9867	0.9933	0.9932
22	0.9459	1.0000	0.9722	0.9730
23	0.9310	1.0000	0.9643	0.9655

f) Neural Networks

Group	Precision	Recall	F Score	ROC (Area under the curve)
1	0.9658	1.0000	0.9728	0.9758
2	0.9561	1.0000	0.9679	0.9785
3	0.6582	0.7737	0.7113	0.7880
4	0.7236	0.9678	0.8280	0.8894
5	0.9011	1.0000	0.9389	0.9763
6	0.9090	0.7840	0.8419	0.9363
7	0.8147	0.9263	0.8669	0.9423
8	0.9561	1.0000	0.9679	0.9797
9	0.7467	0.3920	0.5141	0.7243
10	0.7231	0.9310	0.8140	0.8938
11	0.9678	0.9678	0.9678	0.9797
12	0.9119	0.8208	0.8639	0.9376
13	1.0000	0.9046	0.9408	0.9573
14	0.9408	0.8932	0.9164	0.9577
15	0.9446	1.0000	0.9620	0.9798
16	0.9678	0.9678	0.9678	0.9702
17	1.0000	0.9555	0.9676	1.0000

18	0.9428	0.9310	0.9369	0.9653
19	0.8602	0.9800	0.9162	0.9720
20	0.9441	0.9678	0.9558	0.9705
21	1.0000	1.0000	1.0000	1.0000
22	1.0000	0.8832	0.9291	0.9790
23	0.9558	0.9678	0.9617	0.9766

Sample ROC Curve

