

CS1675 - Assignment 10

Zachary M. Mattis

April 10, 2019

I. Problem 1 - Feature/Input Ranking

a. Fisher_score(x, y)

Rank	Index	Fisher Score
1	48	0.3192
2	25	0.2140
3	21	0.1910
4	70	0.1892
5	65	0.1693
6	40	0.1673
7	29	0.1650
8	19	0.1402
9	57	0.1255
10	20	0.1212
11	24	0.0995
12	30	0.0950
13	12	0.0858
14	47	0.0846
15	61	0.0607
16	10	0.0579
17	34	0.0527
18	27	0.0462
19	39	0.0461
20	41	0.0422

Table 1: Fisher Score, FeatureSelectionData.txt

b. AUROC_score(x, y)

Rank	Index	AUROC Score
1	25	0.7340
2	48	0.7133
3	40	0.6887
4	29	0.6837
5	21	0.6833
6	67	0.6730
7	70	0.6707
8	11	0.6695
9	47	0.6661
10	65	0.6620
11	12	0.6459
12	24	0.6432
13	39	0.6412
14	6	0.6383
15	19	0.6315
16	57	0.6270
17	20	0.6280
18	34	0.6174
19	5	0.6168
20	14	0.6090

Table 2: AUROC Score, FeatureSelectionData.txt

There are a total of 15 shared dimensions between the two result sets. They are not in the same order, but they are relatively close. This is expected as the dimensions that have the greatest predictive power should be very similar across different interpretative algorithms.

II. Problem 2 - Bagging of Classifiers

a. Bagging

T	Training	Testing
1	0.1046	0.2164
2	0.0900	0.1486
3	0.0558	0.1657
4	0.0523	0.1443
5	0.0369	0.1643
6	0.0342	0.1236
7	0.0258	0.1507
8	0.0242	0.1357
9	0.0200	0.1429
10	0.0181	0.1393

Table 3: Bagging SVM

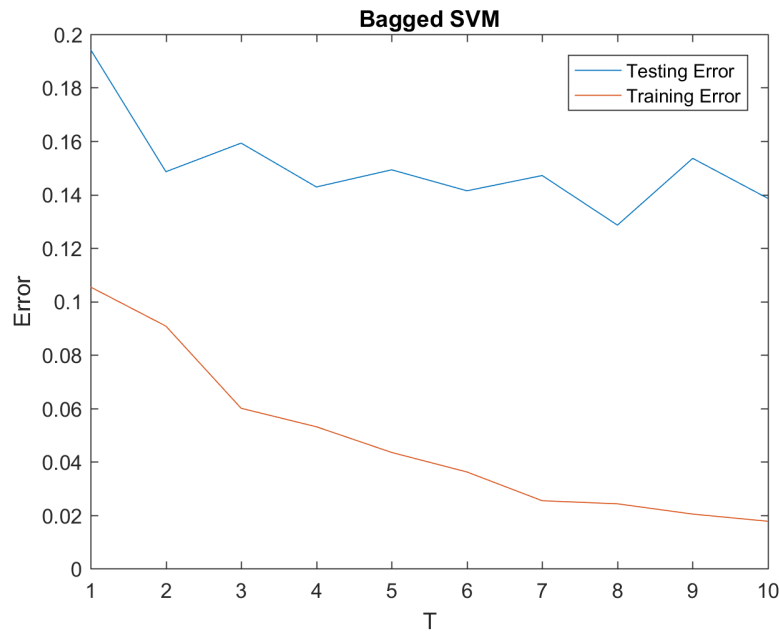


Figure 1: Bagged SVM

An observation of the bagged SVM plot demonstrates a steady decreasing of the training error as the number of models increases. However, the testing error appears to approach a constant of ~ 0.14 .

b. Decision Tree

T	Training	Testing
1	0.1396	0.2521
2	0.1235	0.1771
3	0.0781	0.2243
4	0.0723	0.1543
5	0.0423	0.1886
6	0.0546	0.1443
7	0.0292	0.1671
8	0.0365	0.1464
9	0.0200	0.1443
10	0.0335	0.1343

Table 4: Decision Tree



Figure 2: Decision Tree

An observation of the Decision Tree plot yields similar results to the bagging SVM above. This graph demonstrates a generally steady decreasing of the training error as the number of models increases. However, the testing error appears to approach a constant of ~ 0.14 .