Logistic Regression with Adaboost

1605057: Zahin Ahmed

Performance:
Best Logistic regression only learners:

Dataset 1: alpha = 0.1, max_iter = 1000, error = 0.5

Performance measure	Training	Test
Accuracy	80.32	80.97
True positive rate (recall)	0.53	0.53
True negative rate (specificity)	0.9	0.91
Positive predictive value (precision)	0.66	0.697
False discovery rate	0.34	0.30
F1 score	0.59	0.60

Dataset 2: alpha = 0.1, max_iter = 1000, error = 0.5

Performance measure	Training	Test
Accuracy	82.54	82.84
True positive rate (recall)	0.48	0.49
True negative rate (specificity)	0.93	0.93
Positive predictive value (precision)	0.70	0.70
False discovery rate	0.3	0.30
F1 score	0.57	0.57

Dataset 3: alpha = 0.1, max_iter = 1000, error = 0.5

Performance measure	Training	Test
Accuracy	80.78	78.16
True positive rate (recall)	0	0
True negative rate (specificity)	0.99	0.99
Positive predictive value (precision)	0	0
False discovery rate	0	0
F1 score	0	0

Adaboost:

Dataset 1: alpha = 0.1, max_iter = 1000, cutoff = 0.5

Number of boosting rounds	Training	Test
5	79.71	80.13
10	80.19	80.34
15	80.12	80.62
20	80.10	80.23

Dataset 2: alpha = 0.1, max_iter = 1000, cutoff = 0.5

Number of boosting rounds	Training	Test
5	83.30	83.39
10	83.23	83.20
15	83.40	83.23
20	83.63	83.61

Dataset 3: alpha = 0.1, max_iter = 1000, cutoff = 0.5

Number of boosting rounds	Training	Test
5	96.39	95.59
10	96.83	96.39
15	96.99	96..19
20	97.09	95.79

How to run the script :

To perform logistic regression on the training dataset of dataset 1: uncomment line: 307 . To check the performance uncomment: 308, 429

To perform logistic regression on the training dataset of dataset 2 uncomment line: 316 . To check the performance uncomment: 317, 431

To perform logistic regression on the training dataset of dataset 3 uncomment line 321 . To check the performance uncomment: 322, 433

To test the performance on the test data: comment out the previous lines and then uncomment (312, 430) for dataset 1, uncomment (317, 432) for dataset 2, uncomment (316, 434) for dataset 3

To perform the adaboost implementation, lines 402-405 should be uncommented and the previous lines should be commented out.