CSE472 (Machine Learning Sessional)

Assignment 1: Decision Tree and AdaBoost for Classification

Report on Performance Evaluation

Submitted by

Md Salman Shamil
Student ID: 1505021

Dataset 1: Telco Customer Churn

Decision Tree:

Performance measure	Training	Test
Accuracy	0.8067	0.7977
True positive rate (sensitivity, recall, hit rate)	0.6109	0.6117
True negative rate (specificity)	0.8785	0.8611
Positive predictive value (precision)	0.6482	0.6
False discovery rate	0.3518	0.4
F1 score	0.6290	0.6058

Adaboost:

Number of boosting rounds	Training	Testing
5	0.7774	0.7679
10	0.7934	0.7942
15	0.7943	0.7977
20	0.7978	0.8013

Dataset 2: Adult Dataset

Decision Tree:

Performance measure	Training	Test
Accuracy	0.8758	0.8417
True positive rate (sensitivity, recall, hit rate)	0.6475	0.5705
True negative rate (specificity)	0.9483	0.9256
Positive predictive value (precision)	0.7990	0.7034
False discovery rate	0.2010	0.2966
F1 score	0.7153	0.6300

Adaboost:

Number of boosting rounds	Training	Testing
5	0.8198	0.8232
10	0.8404	0.8426
15	0.8441	0.8459
20	0.8449	0.8480

Dataset 3: Credit Card

Decision Tree:

Performance measure	Training	Test
Accuracy	0.9959	0.9944
True positive rate (sensitivity, recall, hit rate)	0.8413	0.8105
True negative rate (specificity)	0.9997	0.9988
Positive predictive value (precision)	0.9853	0.9390
False discovery rate	0.0147	0.0610
F1 score	0.9076	0.8701

Adaboost:

Number of boosting rounds	Training	Testing
5	0.9930	0.9919
10	0.9930	0.9919
15	0.9930	0.9919
20	0.9930	0.9919