Decision Tree & Adaboost

CSE 472: Machine Learning Sessional

Submitted By:
Tanjim Bin Faruk
1505082

Submitted To:
Md. Shariful Islam Bhuyan
Atif Hasan Rahman
Tahmid Hasan

March 23, 2020

Department Of Computer Science and Engineering
Bangladesh University of Engineering and Technology
(BUET)

1 Decision Tree

1.1 Telco Dataset

Performance Measure	Training	Testing
Accuracy	98.52680156194533	75.1596877217885
True Positive Rate	95.39473684210526	53.29512893982808
True Negative Rate	99.68400583373845	82.35849056603773
Positive Predictive Value	99.11141490088859	49.865951742627345
False Discovery Rate	0.8885850991114164	50.134048257372655
F1 Score	97.21756620851491	51.52354570637119

Decision Tree Performance on Telco Dataset

1.2 Adult Dataset

Performance Measure	Training	Testing
Accuracy	90.4333404993704	83.11528775873718
True Positive Rate	72.15916337201888	55.20020800832033
True Negative Rate	96.22977346278317	91.74909529553679
Positive Predictive Value	85.85735963581183	67.41822800889172
False Discovery Rate	14.142640364188164	32.58177199110829
F1 Score	78.41452428799113	60.700500357398134

Decision Tree Performance on Adult Dataset

1.3 Credit Card Dataset

Performance Measure	Training	Testing
Accuracy	99.9604994623538	99.9403110845827
True Positive Rate	76.5625	72.22222222222221
True Negative Rate	100.0	99.99296443522003
Positive Predictive Value	100.0	95.1219512195122
False Discovery Rate	0.0	4.878048780487809
F1 Score	86.72566371681415	82.10526315789473

Decision Tree Performance on Credit Card Dataset

2 AdaBoost

2.1 Telco Dataset

Number of Boosting Rounds	Training	Testing
5	77.10330138445154	79.347054648687
10	79.30422435214768	80.90844570617459
15	79.71246006389777	80.69552874378992
20	80.24494142705005	81.12136266855926

AdaBoost Accuracy on Telco Dataset

2.2 Adult Dataset

Number of Boosting Rounds	Training	Testing
5	82.00608089432143	82.17554204287207
10	83.92555511194374	84.0550334746023
15	84.3094499554682	84.1041705054972
20	84.70870059273365	84.4911246237946

AdaBoost Accuracy on Adult Dataset

2.3 Credit Card Dataset

Number of Boosting Rounds	Training	Testing
5	99.86657596172836	99.86482216214318
10	99.86657596172836	99.86482216214318
15	99.88413175623779	99.89115550718023
20	99.88457065110052	99.88939995084442

AdaBoost Accuracy on Credit Card Dataset