Assignment 1

Decision Tree and AdaBoost for Classification Introduction

Mohammed Latif Siddiq | 1505069

Telco-Customer-Churn

DECISION TREE

Performance Measure	Training	Testing
Accuracy	93.6812211572595	71.82398864442867
True positive rate	0.8830645161290323	0.4645669291338583
True negative rate	0.9561022672455378	0.4645669291338583
Positive predict value	0.8783422459893048	0.4783783783783784
False discovery rate	0.12165775401069519	0.5216216216216216
F1 score	0.8806970509383377	0.47137150466045274

ADABOOST

Number of Boosting Round	Training	Testing
5	77.36954206602769	78.28246983676367
10	78.64749733759318	79.4180269694819
15	79.67696130635427	79.27608232789213
20	79.28647497337593	78.49538679914833

Adult Dataset

DECISION TREE

Performance Measure	Training	Testing
Accuracy	90.86330272411782	80.81813156440022
True positive rate	0.7682693533988012	0.5423816952678107
True negative rate	0.9531553398058252	0.890390028146361
Positive predict value	0.8387635756056808	0.6048129892722528
False discovery rate	0.16123642439431912	0.39518701072774715
F1 score	0.16123642439431912	0.571898560657985

ADABOOST

Number of Boosting Round	Training	Testing
5	82.00608089432143	82.17554204287205
10	83.29596756856363	83.50838400589645
15	83.79656644451951	83.84620109329894
20	84.10368231933909	84.12873902094466

Credit Card Dataset

DECISION TREE

Performance Measure	Training	Testing
Accuracy	99.97278851851038	99.92977774656788
True positive rate	0.8609112709832134	0.68
True negative rate	0.9999824120161106	0.9997187406613111
Positive predict value	0.9999824120161106	0.7611940298507462
False discovery rate	0.011019283746556474	0.23880597014925373
F1 score	0.9205128205128205	0.7183098591549296

ADABOOST

Number of Boosting Round	Training	Testing
5	99.90432091992363	99.89993328885924
10	99.90432091992363	99.89993328885924
15	99.90432091992363	99.89993328885924
20	99.90432091992363	99.90344440153085