

PROGRAMMING ASSIGNMENT 4

ENSEMBLE METHODS

By

Thota, Sai Abhishek – SXT210056

Gauri, Rahul – RXG200002

Part A – We have executed the above bagging function to get confusion matrices of maximum depth $d = 3,5$ and bag size ($k = 10,20$).

The screenshots of the matrices are below

Max

Depth: 3

Bag Size:

10

	Classifier Positive	Classifier Negative
Actual Positive	120.0	96.0
Actual Negative	120.0	96.0

Max

Depth: 3

Bag Size:

20

	Classifier Positive	Classifier Negative
Actual Positive	108.0	108.0
Actual Negative	60.0	156.0

Max
Depth: 5
Bag Size:
10

	Classifier Positive	Classifier Negative
Actual Positive	120.0	96.0
Actual Negative	72.0	144.0

Max
Depth: 5
Bag Size:
20

	Classifier Positive	Classifier Negative
Actual Positive	120.0	96.0
Actual Negative	120.0	96.0

Part 2 - For Boosting - We have executed the above bagging function to get confusion matrices of maximum depth $d = 1, 2$ and bag size ($k = 20, 40$).

Max

Depth: 1

Bag

Size: 20

	Classifier Positive	Classifier Negative
Actual Positive	168.0	48.0
Actual Negative	96.0	120.0

Max

Depth: 1

Bag

Size: 40

	Classifier Positive	Classifier Negative
Actual Positive	144.0	72.0
Actual Negative	84.0	132.0

Max
Depth: 2
Bag
Size: 20

	Classifier Positive	Classifier Negative
Actual Positive	162.0	54.0
Actual Negative	108.0	108.0

Max
Depth: 2
Bag
Size: 40

	Classifier Positive	Classifier Negative
Actual Positive	156.0	60.0
Actual Negative	24.0	192.0

Part C – scikit implementation

We have used scikit library to implement bagging and adaboost implementation.

For bagging using Scikit implementation

Max

Depth: 3

Bag

Size: 10

	Classifier Positive	Classifier Negative
Actual Positive	204.0	12.0
Actual Negative	6.0	210.0

Max

Depth: 5

Bag

Size: 10

	Classifier Positive	Classifier Negative
Actual Positive	216.0	0.0
Actual Negative	3.0	213.0

Max
Depth: 3
Bag
Size: 20

	Classifier Positive	Classifier Negative
Actual Positive	200.0	16.0
Actual Negative	14.0	202.0

Max
Depth: 5
Bag
Size: 20

	Classifier Positive	Classifier Negative
Actual Positive	205.0	11.0
Actual Negative	0.0	216.0

Adaboost implementation

Max

Depth: 1

Bag

Size: 20

	Classifier Positive	Classifier Negative
Actual Positive	141.0	75.0
Actual Negative	53.0	163.0

Max

Depth: 2

Bag

Size: 20

	Classifier Positive	Classifier Negative
Actual Positive	216.0	0.0
Actual Negative	0.0	216.0

Max
Depth: 1
Bag
Size: 40

	Classifier Positive	Classifier Negative
Actual Positive	140.0	76.0
Actual Negative	57.0	159.0

Max
Depth: 2
Bag
Size: 40

	Classifier Positive	Classifier Negative
Actual Positive	216.0	0.0
Actual Negative	0.0	216.0

We can observe that scikit implementation is better because it can categorize positives and negatives with less error. It has higher TPR and TNR.