PROGRAMMING ASSIGNMENT 2

(Submitted By: AXM190041, PXN190012)

Bagging

Confusion matrices for depths 3, 5 and bag sizes 10, 20

```
Bagging
Max depth is 3 Bag size is 10
Train Error is 3.58%.
Test Error is 4.23%.
Confusion Matrix
[[ 815   29]
 [  57 1130]]
Max depth is 3 Bag size is 20
Train Error is 3.58%.
Test Error is 4.23%.
Confusion Matrix
[[ 815   29]
 [  57 1130]]
Max depth is 5 Bag size is 10
Train Error is 0.07%.
Test Error is 0.20%.
Confusion Matrix
[[ 844    0]
 [   4 1183]]
Max depth is 5 Bag size is 20
Train Error is 0.07%.
Test Error is 0.20%.
Confusion Matrix
[[ 844    0]
 [   4 1183]]
```

Boosting

Confusion matrices for depths 3, 5 and bag sizes 10, 20

```
Boosting
Max depth is 1 Bag size is 20
Train Error is 11.80%.
Test Error is 12.65%.
Confusion Matrix
[[793  51]
 [206 981]]
Max depth is 1 Bag size is 40
Train Error is 11.80%.
Test Error is 12.65%.
Confusion Matrix
[[793  51]
 [206 981]]
Max depth is 2 Bag size is 20
Train Error is 9.49%.
Test Error is 10.14%.
Confusion Matrix
[[844   0]
 [206 981]]
Max depth is 2 Bag size is 40
Train Error is 9.49%.
Test Error is 10.14%.
Confusion Matrix
[[844   0]
 [206 981]]
```

Scikit-learn Bagging

Confusion matrices for depths 3, 5 and bag sizes 10, 20

```
Scikit-learn Bagging
Max depth is 3 Bag size is 10
Train accuracy is 0.9647136057771213
Test accuracy is 0.9561792220580995
Confusion Matrix
[[1102   85]
 [   4  840]]
Max depth is 3 Bag size is 20
Train accuracy is 0.9647136057771213
Test accuracy is 0.9561792220580995
Confusion Matrix
[[1102   85]
 [   4  840]]
Max depth is 5 Bag size is 10
Train accuracy is 0.9881831610044313
Test accuracy is 0.9881831610044313
Confusion Matrix
[[1187    0]
 [  24  820]]
Max depth is 5 Bag size is 20
Train accuracy is 0.9881831610044313
Test accuracy is 0.9881831610044313
Confusion Matrix
[[1187    0]
 [  24  820]]
```

Scikit-learn Boosting:

Confusion matrices for depths 3, 5 and bag sizes 10, 20

```
Scikit-learn Boosting
Max depth is 1 Bag size is 20
Train accuracy is 0.8819957328081405
Test accuracy is 0.8734613490891187
Confusion Matrix
[[981 206]
 [ 51 793]]
Max depth is 1 Bag size is 40
Train accuracy is 0.8819957328081405
Test accuracy is 0.8734613490891187
Confusion Matrix
[[981 206]
 [ 51 793]]
Max depth is 2 Bag size is 20
Train accuracy is 0.9451829968816675
Test accuracy is 0.9320531757754801
Confusion Matrix
[[1061  126]
 [  12  832]]
Max depth is 2 Bag size is 40
Train accuracy is 0.9451829968816675
Test accuracy is 0.9320531757754801
Confusion Matrix
[[1061  126]
 [  12  832]]
```
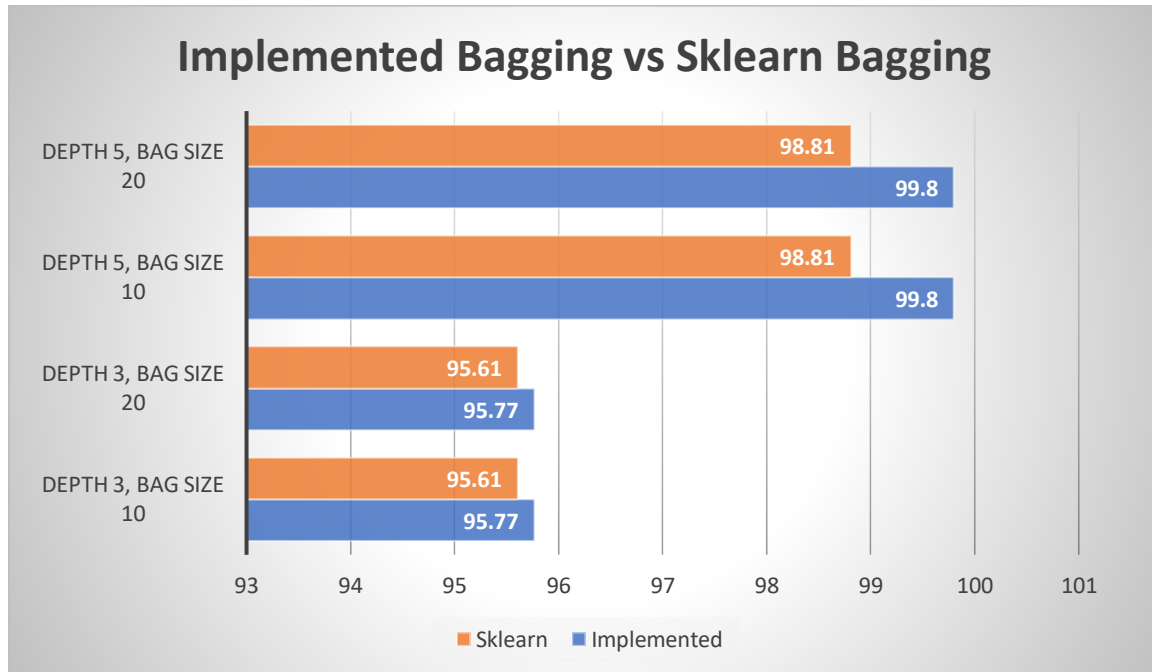
Performance

Testing accuracy of our implemented bagging is slightly better than Scikit-learn bagging for all 4 combinations.

Implemented Bagging vs Sklearn Bagging

Configuration	Sklearn	Implemented
DEPTH 5, BAG SIZE 20	98.81	99.8
DEPTH 5, BAG SIZE 10	98.81	99.8
DEPTH 3, BAG SIZE 20	95.61	95.77
DEPTH 3, BAG SIZE 10	95.61	95.77

Testing accuracy of Scikit-learn is better than implemented boosting when depth is 5. But when depth is 3 both have similar performance.

Implemented Boosting vs Sklearn Boosting

Configuration	Sklearn	Implemented
DEPTH 5, BAG SIZE 20	93.2	89.86
DEPTH 5, BAG SIZE 10	93.2	89.86
DEPTH 3, BAG SIZE 20	87.34	87.35
DEPTH 3, BAG SIZE 10	87.34	87.35