Results for K Nearest Neighbors Algorithm and Boosting Technique

Aditi Helekar

March 12, 2017

Contents:

1. K Nearest Neighbors Results:

1.1 Digit Recognition Dataset

```
Reading dataset.
(3822, 65)
(2293, 64)
(2293, 1)
(1529, 64)
Sample 1:::
[0 2 3 ..., 2 9 4]
Accuracy:  0.984303466318
Sample 2:::
[0 2 3 ..., 2 9 4]
Accuracy:  0.967952910399
Sample 3:::
[0 2 3 ..., 2 9 4]
Accuracy:  0.985611510791
Sample 4:::
[0 2 3 ..., 2 9 4]
Accuracy:  0.985611510791
Sample 5:::
[0 2 3 ..., 2 9 4]
Accuracy:  0.982341399608
Testing::::
[1 1 3 ..., 1 9 8]
Accuracy:  0.978841870824
```

1.2 Amazon Dataset

KNN Amazon Dataset:

1. 10 testing samples
   - Output: [5 5 5 5 5 5 5 4 5 5]
   - Accuracy: 0.7
   - Neighbors: 10

2. 100 testing samples
   - Output: [5 5 5 5 5 5 5 4 5 5 4 5 5 5 5 5 5 5 5 5 5 5 5 1 4 5 4 5 4 5 5 5 5 5 5 5 5
     5 5 5 5 4 5 5 5 5 4 5 4 5 5 5 5 3 5 5 5 5 4 5 5 5 5 5 5 3 5 5 5 5 5 5 5 5
     5 5 5 5 5 5 5 4 5 5 4 5 5 5 5 5 5 5 4 5 5 5 5 5 5 5]
   - Accuracy: 0.51
   - Neighbors: 10

3. 100 testing samples
   - Output: [5 5 5 5 5 3 5 5 5 5 4 4 5 5 5 5 5 5 5 2 5 5 5 1 4 5 4 5 5 5 5 5 2 4 5 5 5
     5 4 5 5 3 5 5 5 5 4 5 4 5 4 5 5 1 5 5 5 1 5 5 5 5 5 5 5 3 3 5 5 5 5 5 5 5
     5 5 5 5 4 5 5 4 5 5 4 5 5 5 5 5 4 5 4 5 5 5 5 5 5 5]
   - Accuracy: 0.48
   - Neighbors: 5

4. 100 testing samples

Output: [5 5 5 5 5 5 5 5 5 5 4 5 5 5 5 5 5 5 4 5 5 5 5 4 5 4 5 5 5 5 5 5 5 5 5
5 5 5 5 4 5 5 5 5 5 5 4 5 5 5 5 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5 5 5 5 5 5 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5]

Accuracy: 0.55

Neighbors: 20

5. 100 testing samples

Output: [5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5]

Accuracy: 0.61

Neighbors: 50

6. 100 testing samples

Output: [5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5]

Accuracy: 0.61

Neighbors: 50 and weights='distance',algorithm='ball_tree',leaf_size=50

7. 100 testing samples

Output: [5 5 5 5 5 3 5 5 5 5 4 4 5 5 5 5 5 5 5 2 5 5 5 1 5 5 4 5 5 5 5 5 2 5 5 5 5
5 4 5 5 3 5 5 5 5 4 5 4 5 4 5 5 1 5 5 5 2 5 5 5 5 5 5 5 3 3 5 5 5 5 5 5 5 5
5 5 5 5 4 5 5 4 5 5 4 5 5 5 5 5 4 5 4 5 5 5 5 5 5 5]

Accuracy: 0.5

Neighbors: 5 and weights='distance',algorithm='ball_tree',leaf_size=50

8. 150 testing samples

Output: [5 5 5 5 5 5 5 4 5 5 4 5 5 5 5 5 5 5 5 5 5 5 5 1 4 5 4 5 4 5 5 5 5 5 5 5
5 5 5 5 4 5 5 5 5 4 5 4 5 5 5 5 3 5 5 5 5 5 5 5 5 5 5 3 5 5 5 5 5 5 5 5
5 5 5 5 5 5 5 4 5 5 4 5 5 5 5 5 5 5 4 5 5 5 5 5 5 5 4 5 5 5 5 5 5 5 1 5 5
5 5 4 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 4 5 5 5 5 5 4 5 5 5 5 5 5 5 5 5 5 5
5 5]

Accuracy: 0.5

Neighbors: 10 and weights='distance',algorithm='auto',leaf_size=50

9. 150 testing samples

Output: [5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5]

Accuracy: 0.58

Neighbors: 50 and weights='distance',algorithm='auto',leaf_size=100

10. 150 testing samples

Output: [5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
5 5]

Accuracy: 0.58

Neighbors: 100 and weights='distance',algorithm='auto',leaf_size=50

2. AdaBoost Technique

2.1 Digit Recognition Dataset

```
Reading dataset.
(3822, 65)
(2293, 64)
(2293, 1)
(1529, 64)
Sample 1
Accuracy:  0.534990189666
Sample 2
Accuracy:  0.535644211903
Sample 3
Accuracy:  0.933943754088
Sample 4
Accuracy:  0.97580117724
Sample 5
Accuracy:  0.986919555265
Testing Data::
(3822, 64)
(1796, 64)
(3822, 1)
(1796, 1)
Test Sample 1
Accuracy:  0.919821826281
Test Sample 2
Accuracy:  0.953786191537
Test Sample 3
Accuracy:  0.964922048998
```

2.1 Amazon Dataset

1. Accuracy: 0.578066812705

   max_depth=2,n_estimators=100,learning_rate=1

2. Accuracy: 0.572590361446

   max_depth=2,n_estimators=500,learning_rate=1

3. Accuracy: 0.579052573932

   max_depth=2,n_estimators=50,learning_rate=1