Sml Assignment 3

Name: Pourav Surya

Roll No: 2021271

Applied PCA and got the reduced data with dimensions

```
Dimensions: (18623, 10)
```

Implemented Decision tree by comparing the gini index for each dimension and then comparing to find the best split and information gain and then taking that split.

Printed the Decision tree

```
X_0 <= (624.6639084076752+0j) ? 0.2983877586057319
 left:X_1 <= (179.67561986817168+0j) ? 0.3037225165600382
  left:X_0 <= (-202.45652113485986+0j) ? 0.14759057627204
    left:0j
    right:(2+0j)
  right:X_4 <= (-720.9898876115241+0j) ? 0.0891037058316576
    left:0j
    right:(2+0j)
 right:X_2 <= (231.62107754431915+0j) ? 0.06740047510098737
  left:X_7 <= (-575.7336812339543+0j) ? 0.0056004210280101545
    left:(2+0j)
    right:(1+0j)
  right:X_1 <= (-480.64812928740986+0j) ? 0.2659279778393353
    left:(1+0j)
    right:(2+0j)
```

Got accuracy for test dataset

```
0.6619002224340642
```

Class-wise accuracy

```
Accuracy for class 0: 0.32857142857142857
Accuracy for class 1: 0.9145374449339208
Accuracy for class 2: 0.7005813953488372
```

Implemented bagging

Classwise and total accuracy:

```
Total Accuracy: 0.6139180171591992
Class-wise Accuracy: {0: 0.6571428571428571, 1: 0.9409691629955947, 2: 0.2131782945736434}
```