Koushik Sahu
118CS0597
Machine Learning Laboratory – 7
2nd November 2021
Lab-7 Report

Checking for overfitting: to check for overfitting we need to check if the training and validation accuracy are close or not. A model that has not overfitted will give similar accuracy for both.

Preprunning technique used: I used the maximum depth method to preprune the decision tree. In this technique we maintain a depth parameter in dfs and whenever we encounter a situation where the dfs algorithm is trying to branch out the decision tree farther than the maximum depth we convert it into a leaf node with a prediction.

Postprunning technique used: For postprunning we start from the leaf nodes and go up. At each non-leaf node we convert the node into leaf by associating the class that appears the most times in the tree. If the accuracy goes up we prune at this node, else we leave the tree as it was before and continue going up.

Output:
Full decision tree:

```
Sex = Male
        Class = Third
                Age = Adult
                        Prediction: No

                Age = Child
                        Prediction: No

        Class = Crew
                Prediction: No

        Class = First
                Age = Adult
                        Prediction: No

                Age = Child
                        Prediction: Yes

        Class = Second
                Age = Adult
                        Prediction: No

                Age = Child
                        Prediction: Yes

Sex = Female
        Class = Third
                Age = Adult
                        Prediction: No

                Age = Child
                        Prediction: No

        Class = Second
                Age = Adult
                        Prediction: Yes

                Age = Child
                        Prediction: Yes

        Class = First
                Age = Adult
                        Prediction: Yes

                Age = Child
                        Prediction: Yes

        Class = Crew
                Prediction: Yes
```

```
For the full decision tree:
Training Accuracy: 0.7846590909090909
Validation Accuracy: 0.8140589569160998
```

Koushik Sahu **118CS0597**

For preprunned decision tree:

```
Sex = Male
        Class = Third
                Prediction: No

        Class = Crew
                Prediction: No

        Class = First
                Prediction: No

        Class = Second
                Prediction: No

Sex = Female
        Class = Third
                Prediction: No

        Class = Second
                Prediction: Yes

        Class = First
                Prediction: Yes

        Class = Crew
                Prediction: Yes
```

```
After preprunning:
Training Accuracy: 0.7767045454545455
Validation Accuracy: 0.8095238095238095
```

For postprunned decision tree:

```
Sex = Male
        Class = Third
                Prediction: No

        Class = Crew
                Prediction: No

        Class = First
                Age = Adult
                        Prediction: No

                Age = Child
                        Prediction: Yes

        Class = Second
                Age = Adult
                        Prediction: No

                Age = Child
                        Prediction: Yes

Sex = Female
        Class = Third
                Prediction: No

        Class = Second
                Prediction: Yes

        Class = First
                Prediction: Yes

        Class = Crew
                Prediction: Yes
```

```
After post prunning:
Training Accuracy: 0.7846590909090909
Validation Accuracy: 0.8140589569160998
```