# Report

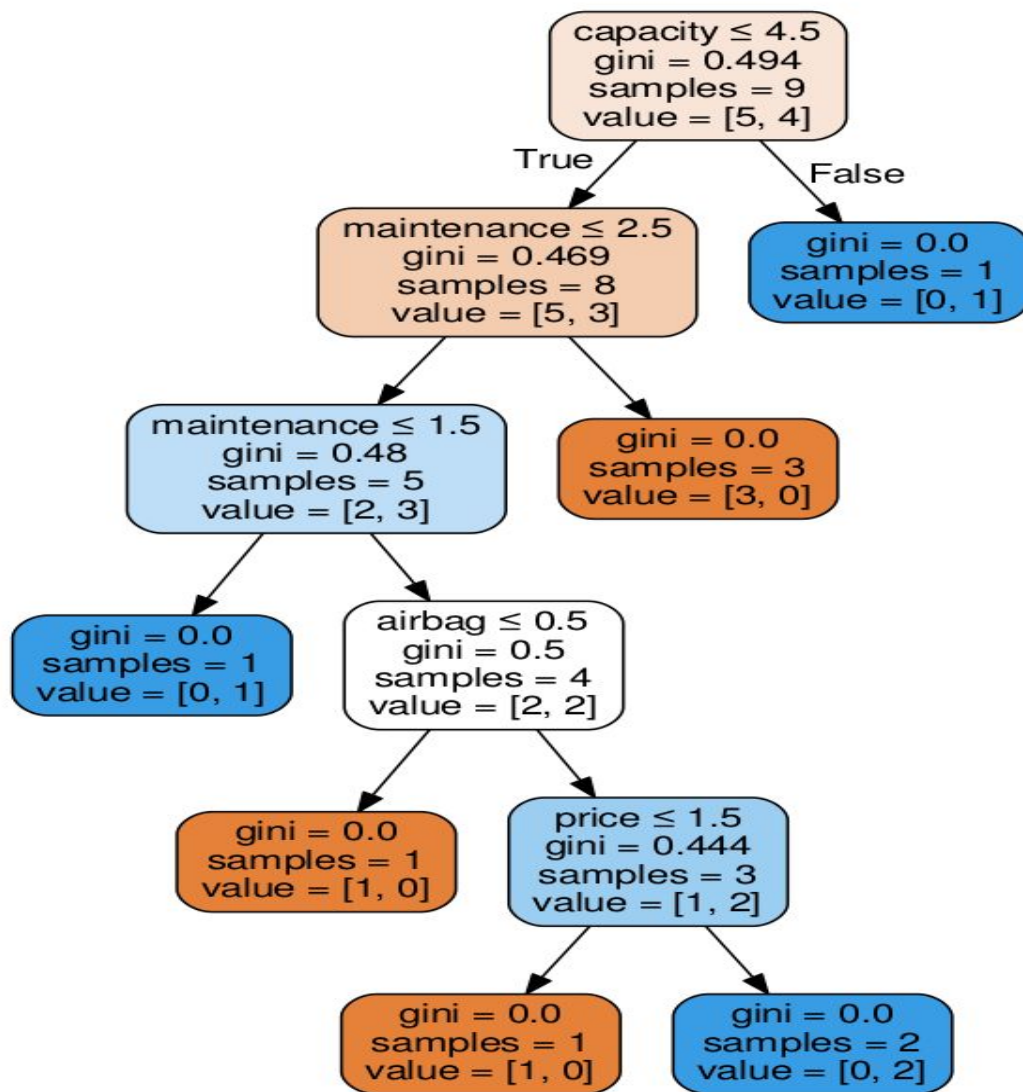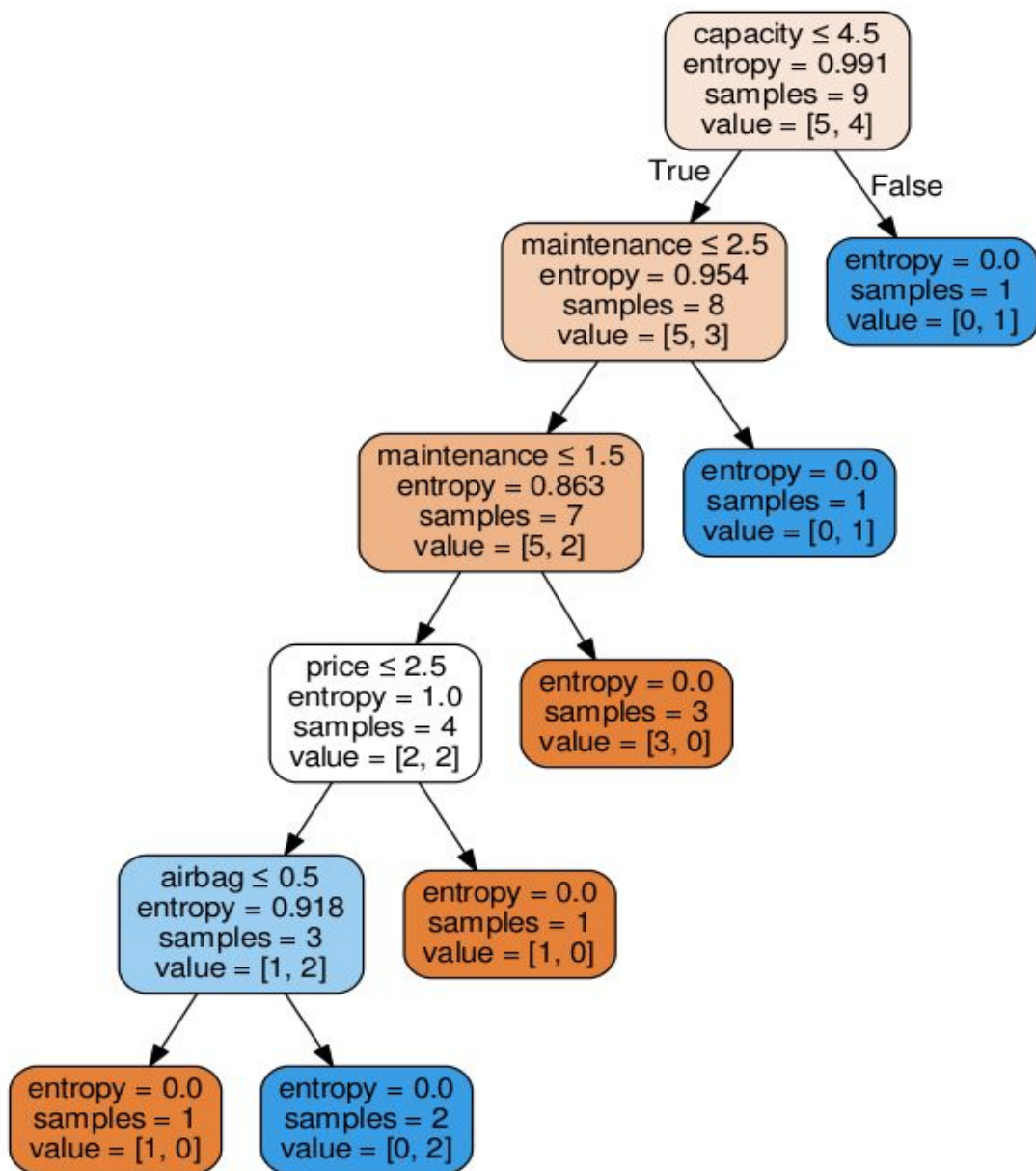**Part 1**

  **Sklearn**



Decision Tree made using Gini index

Gini_index (Root Node) = 0.494

Attribute chosen at root = 'capacity'

Labels generated (Test Data) = [Yes, Yes]

Test set accuracy = 100%

Decision Tree made using Information Gain

Entropy (Root Node) = 0.991

Gain (Root Node) = .461            [.991 - (.954*(5/9) + 0*(4/9)]

Labels generated (Test Data) = [Yes, Yes]

Test set accuracy = 100%

## Own implementation

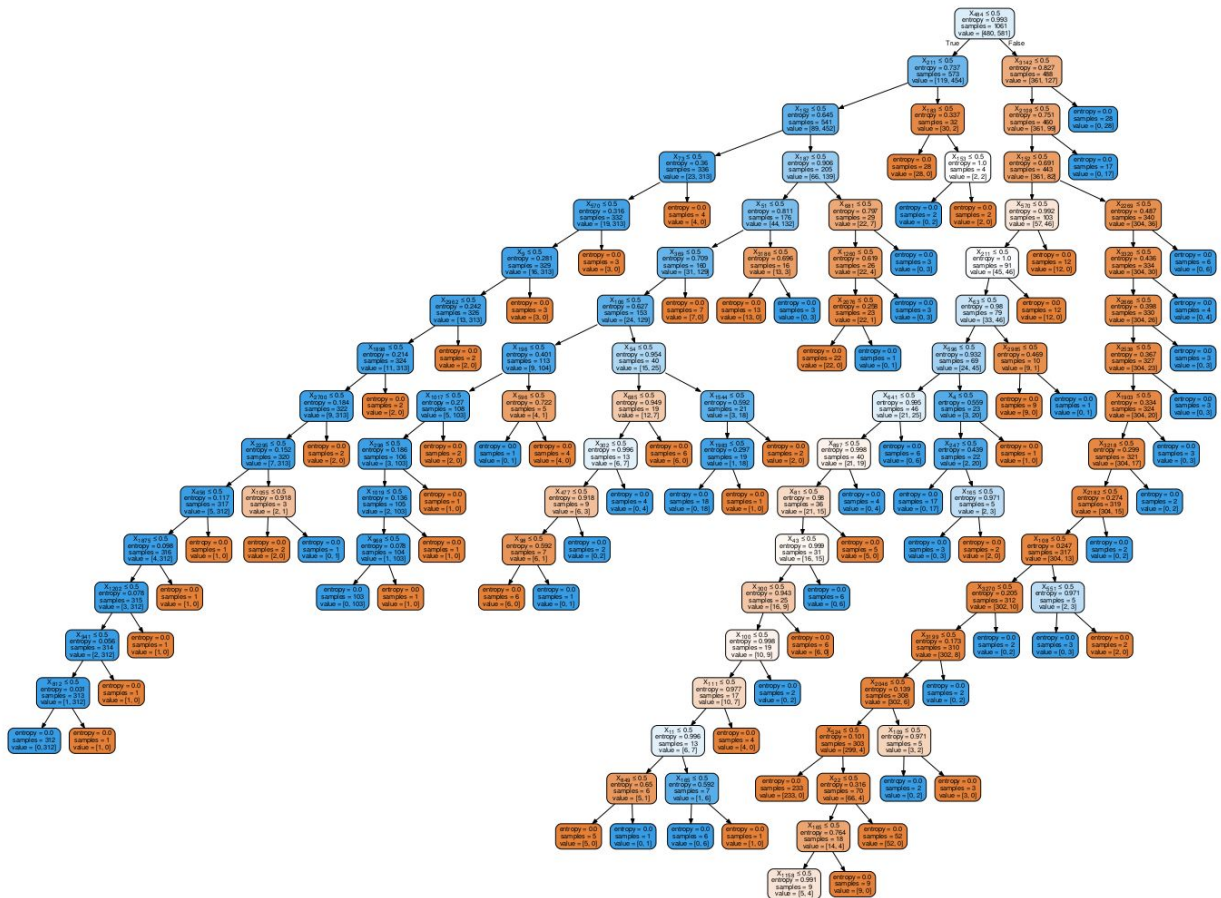Gini_index (Root Node) = 0.4938
Attribute chosen at root = 'airbag'

Labels generated (Test Data) = [Yes, Yes]
Test set accuracy = 100%

## Part B
### Sklearn

Entropy (Root Node) = 0.993

Information gain = .993 - (.737*(480/1061)) - (.827*(581/1061))

= .2067

Predictions on test set:

[1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
 1 1 2 1 1 1 2 1 2 1 2 2 2 1 1 1 1 1 2 1 1 2 1 1 1 1 2 2 1 1 1 1 1 1 1 2
 1 2 2 2 2 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 2 1 1 1 1 1 2 1 1 1 1 1
 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 2 2 1 1 1 1 1 1 2
 1 1 1 2 2 2 1 1 1 1 1 1 1 1 1 1 1 1 2 2 1 1 1 1 1 1 2 1 1 1 1 2 1 1 1 2
 2 2 1 1 2 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 1 1 1 1 2 2 1 1 1 1
 1 1 1 2 2 2 1 1 2 1 1 1 1 1 1 1 1 1 2 2 1 1 1 1 2 2 1 2 1 1 2 2 1 1 1 1 1
 1 2 1 1 1 1 1 1 1 1 2 2 1 2 2 1 1 1 1 1 1 1 2 2 1 2 2 1 1 1 1 1 1 1 1
 1 1 2 2 1 1 2 2 2 1 2 1 2 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 1 2 1 2 2 2 2 1
 1 2 2 2 2 2 2 2 2 2 2 2 1 1 2 2 2 2 1 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2
 2 1 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 1 2 2 2
 2 2 2 2 1 2 2 1 2 1 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 1 2 2 1 2 2 2 2 2
 1 1 2 2 2 1 2 2 2 2 1 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
 2 2 1 2 1 2 2 2 2 2 2 2 2 2 2 2 1 2 1 2 2 2 1 2 2 2 2 2 2 2 2 1 2 2 2 2 2
 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 1 2 2 2 1 2 2 2 2 2 2 1 1 2
 2 2 2 2 2 2 1 2 2 2 2 2 2 1 2 2 2 1 2 2 1 1 2 2 2 2 2 2 1 2 2 2 2 2 2 2
 2 2 2 2 1 1 2 1 2 2 1 2 2 2 2 2 2 2 2 1 2 2 1 2 2 2 2 2 2 2 2 2 2 1 2 2
 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1
 2 2 2 1 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2
 1 2 2 2]

Test Accuracy =  0.8316831683168316