REPORT

Team members:

Dhwaniben Rameshbhai Kaneria

Best Results:

```
Number of training instances = 600
Number of training attributes = 20
Total number of nodes in the pruned tree = 53
Number of leaf nodes in the pruned tree = 27
Accuracy of the model on the training data = 74.0%

Number of validation instances = 2000
Number of validation attributes = 21
Accuracy of the model on the validation data = 70.1%

Number of testing instances = 2000
Number of testing attributes = 21
Accuracy of the model on the testing data = 70.35%
```

What was accomplished:

I succeeded in building an effective classifier using the ID3 algorithm that learns from the training data and attempts to classify new test data correctly. Pruning the decision tree helps in increasing the accuracy in some instances.

What was learned:

I learnt that ID3 algorithm splits a data set on the attribute with least entropy or the maximum information gain thereby, reducing the amount of randomness in the decision making process.

I learned the working of ID3 algorithm and how it helps in choosing the most effective attributes early, thereby, aiding in building the decision tree accordingly. I also observed that the algorithm learns and performs better when more training data is provided.

Results from 1 run of the program:

```
XO = 0 :
 | XM = 0 :
 |  | XF = 0 :
 |  |  | XB = 0 :
 |  |  |  | XG = 0 : 0
 |  |  |  | XG = 1 :
 |  |  |  |  | XD = 0 :
 |  |  |  |  |  | XS = 0 : 0
```

```
|   |   |   |   |   | XS = 1 :
|   |   |   |   |   |   | XC = 0 : 1
|   |   |   |   |   |   | XC = 1 :
|   |   |   |   |   |   |   | XH = 0 : 0
|   |   |   |   |   |   |   | XH = 1 : 1
|   |   |   |   | XD = 1 :
|   |   |   |   |   | XE = 0 : 0
|   |   |   |   |   | XE = 1 :
|   |   |   |   |   |   | XK = 0 : 0
|   |   |   |   |   |   | XK = 1 : 1
|   |   | XB = 1 :
|   |   |   | XD = 0 : 0
|   |   |   | XD = 1 :
|   |   |   |   | XI = 0 : 0
|   |   |   |   | XI = 1 :
|   |   |   |   |   | XG = 0 : 1
|   |   |   |   |   | XG = 1 : 0
|   | XF = 1 : 0
| XM = 1 :
|   | XB = 0 :
|   |   | XD = 0 :
|   |   |   | XG = 0 :
|   |   |   |   | XF = 0 : 0
|   |   |   |   | XF = 1 :
|   |   |   |   |   | XJ = 0 :
|   |   |   |   |   |   | XN = 0 : 1
|   |   |   |   |   |   | XN = 1 :
|   |   |   |   |   |   |   | XE = 0 :
|   |   |   |   |   |   |   |   | XK = 0 : 0
|   |   |   |   |   |   |   |   | XK = 1 : 1
|   |   |   |   |   |   |   | XE = 1 : 0
|   |   |   |   |   | XJ = 1 :
|   |   |   |   |   |   | XC = 0 :
|   |   |   |   |   |   |   | XT = 0 :
|   |   |   |   |   |   |   |   | XL = 0 :
|   |   |   |   |   |   |   |   |   | XE = 0 :
|   |   |   |   |   |   |   |   |   |   | XI = 0 : 0
|   |   |   |   |   |   |   |   |   |   | XI = 1 : 1
|   |   |   |   |   |   |   |   |   | XE = 1 : 1
|   |   |   |   |   |   |   |   | XL = 1 : 0
|   |   |   |   |   |   |   | XT = 1 : 1
|   |   |   |   |   |   | XC = 1 : 1
|   |   |   |   | XG = 1 :
|   |   |   |   |   | XU = 0 : 1
|   |   |   |   |   | XU = 1 :
|   |   |   |   |   |   | XI = 0 : 0
|   |   |   |   |   |   | XI = 1 : 1
|   |   |   | XD = 1 :
|   |   |   |   | XC = 0 :
|   |   |   |   |   | XF = 0 :
|   |   |   |   |   |   | XG = 0 : 0
|   |   |   |   |   |   | XG = 1 :
|   |   |   |   |   |   |   | XP = 0 :
|   |   |   |   |   |   |   |   | XS = 0 : 0
|   |   |   |   |   |   |   |   | XS = 1 : 1
```

```
  |  |  |  |  |  |  | XP = 1 : 0
  |  |  |  |  | XF = 1 :
  |  |  |  |  |  | XJ = 0 : 1
  |  |  |  |  |  | XJ = 1 :
  |  |  |  |  |  |  | XE = 0 :
  |  |  |  |  |  |  |  | XG = 0 :
  |  |  |  |  |  |  |  |  | XI = 0 : 1
  |  |  |  |  |  |  |  |  | XI = 1 : 0
  |  |  |  |  |  |  |  | XG = 1 : 0
  |  |  |  |  |  |  | XE = 1 :
  |  |  |  |  |  |  |  | XT = 0 :
  |  |  |  |  |  |  |  |  | XG = 0 : 1
  |  |  |  |  |  |  |  |  | XG = 1 : 0
  |  |  |  |  |  |  |  | XT = 1 : 1
  |  |  |  | XC = 1 : 0
  |  | XB = 1 :
  |  |  | XI = 0 : 0
  |  |  | XI = 1 :
  |  |  |  | XC = 0 :
  |  |  |  |  | XK = 0 :
  |  |  |  |  |  | XP = 0 : 1
  |  |  |  |  |  | XP = 1 :
  |  |  |  |  |  |  | XS = 0 :
  |  |  |  |  |  |  |  | XG = 0 : 1
  |  |  |  |  |  |  |  | XG = 1 :
  |  |  |  |  |  |  |  |  | XF = 0 : 0
  |  |  |  |  |  |  |  |  | XF = 1 : 1
  |  |  |  |  |  |  | XS = 1 : 0
  |  |  |  |  |  | XK = 1 : 0
  |  |  |  | XC = 1 : 0
XO = 1 :
  | XI = 0 :
  |  | XM = 0 :
  |  |  | XQ = 0 :
  |  |  |  | XF = 0 :
  |  |  |  |  | XH = 0 :
  |  |  |  |  |  | XB = 0 : 0
  |  |  |  |  |  | XB = 1 :
  |  |  |  |  |  |  | XC = 0 : 1
  |  |  |  |  |  |  | XC = 1 : 0
  |  |  |  |  | XH = 1 : 1
  |  |  |  | XF = 1 : 0
  |  |  | XQ = 1 :
  |  |  |  | XJ = 0 :
  |  |  |  |  | XN = 0 :
  |  |  |  |  |  | XP = 0 : 1
  |  |  |  |  |  | XP = 1 :
  |  |  |  |  |  |  | XB = 0 :
  |  |  |  |  |  |  |  | XF = 0 : 0
  |  |  |  |  |  |  |  | XF = 1 : 1
  |  |  |  |  |  |  | XB = 1 : 0
  |  |  |  |  | XN = 1 : 0
  |  |  |  | XJ = 1 :
  |  |  |  |  | XL = 0 :
  |  |  |  |  |  | XH = 0 : 0
```

```
|   |   |   |   |   | XH = 1 :
|   |   |   |   |   |   | XK = 0 :
|   |   |   |   |   |   |   | XU = 0 : 1
|   |   |   |   |   |   |   | XU = 1 : 0
|   |   |   |   |   |   | XK = 1 : 1
|   |   |   |   | XL = 1 : 1
|   | XM = 1 :
|   |   | XQ = 0 :
|   |   |   | XF = 0 :
|   |   |   |   | XL = 0 :
|   |   |   |   |   | XC = 0 : 1
|   |   |   |   |   | XC = 1 :
|   |   |   |   |   |   | XH = 0 : 1
|   |   |   |   |   |   | XH = 1 :
|   |   |   |   |   |   |   | XU = 0 :
|   |   |   |   |   |   |   |   | XB = 0 :
|   |   |   |   |   |   |   |   |   | XD = 0 : 1
|   |   |   |   |   |   |   |   |   | XD = 1 : 0
|   |   |   |   |   |   |   |   | XB = 1 : 0
|   |   |   |   |   |   |   |   | XU = 1 : 1
|   |   |   |   | XL = 1 :
|   |   |   |   |   | XC = 0 :
|   |   |   |   |   |   | XB = 0 : 0
|   |   |   |   |   |   | XB = 1 :
|   |   |   |   |   |   |   | XP = 0 : 0
|   |   |   |   |   |   |   | XP = 1 : 1
|   |   |   |   |   | XC = 1 : 1
|   |   |   | XF = 1 : 0
|   |   | XQ = 1 : 0
| XI = 1 :
|   | XT = 0 :
|   |   | XH = 0 :
|   |   |   | XP = 0 :
|   |   |   |   | XF = 0 : 0
|   |   |   |   | XF = 1 :
|   |   |   |   |   | XQ = 0 :
|   |   |   |   |   |   | XK = 0 : 1
|   |   |   |   |   |   | XK = 1 :
|   |   |   |   |   |   |   | XC = 0 : 0
|   |   |   |   |   |   |   | XC = 1 : 1
|   |   |   |   |   | XQ = 1 :
|   |   |   |   |   |   | XK = 0 : 0
|   |   |   |   |   |   | XK = 1 : 1
|   |   |   |   | XP = 1 :
|   |   |   |   |   | XS = 0 :
|   |   |   |   |   |   | XD = 0 :
|   |   |   |   |   |   |   | XC = 0 :
|   |   |   |   |   |   |   |   | XJ = 0 :
|   |   |   |   |   |   |   |   |   | XN = 0 : 0
|   |   |   |   |   |   |   |   |   | XN = 1 : 1
|   |   |   |   |   |   |   |   | XJ = 1 :
|   |   |   |   |   |   |   |   |   | XB = 0 : 0
|   |   |   |   |   |   |   |   |   | XB = 1 :
|   |   |   |   |   |   |   |   |   |   | XG = 0 : 1
|   |   |   |   |   |   |   |   |   |   | XG = 1 : 0
```

```
|   |   |   |   |   |   | XC = 1 : 0
|   |   |   |   |   | XD = 1 :
|   |   |   |   |   |   | XM = 0 :
|   |   |   |   |   |   |   | XC = 0 : 1
|   |   |   |   |   |   |   | XC = 1 : 0
|   |   |   |   |   |   | XM = 1 : 1
|   |   |   |   | XS = 1 : 1
|   |   | XH = 1 :
|   |   |   | XJ = 0 :
|   |   |   |   | XC = 0 :
|   |   |   |   |   | XN = 0 : 1
|   |   |   |   |   | XN = 1 :
|   |   |   |   |   |   | XF = 0 :
|   |   |   |   |   |   |   | XG = 0 : 1
|   |   |   |   |   |   |   | XG = 1 : 0
|   |   |   |   |   |   | XF = 1 : 0
|   |   |   |   | XC = 1 :
|   |   |   |   |   | XM = 0 : 0
|   |   |   |   |   | XM = 1 :
|   |   |   |   |   |   | XF = 0 :
|   |   |   |   |   |   |   | XR = 0 : 1
|   |   |   |   |   |   |   | XR = 1 : 0
|   |   |   |   |   |   | XF = 1 : 1
|   |   |   | XJ = 1 :
|   |   |   |   | XS = 0 : 1
|   |   |   |   | XS = 1 :
|   |   |   |   |   | XG = 0 :
|   |   |   |   |   |   | XB = 0 : 0
|   |   |   |   |   |   | XB = 1 :
|   |   |   |   |   |   |   | XD = 0 : 1
|   |   |   |   |   |   |   | XD = 1 :
|   |   |   |   |   |   |   |   | XE = 0 : 0
|   |   |   |   |   |   |   |   | XE = 1 : 1
|   |   |   |   |   | XG = 1 :
|   |   |   |   |   |   | XC = 0 : 1
|   |   |   |   |   |   | XC = 1 :
|   |   |   |   |   |   |   | XD = 0 : 1
|   |   |   |   |   |   |   | XD = 1 : 0
|   | XT = 1 :
|   |   | XS = 0 :
|   |   |   | XQ = 0 :
|   |   |   |   | XK = 0 :
|   |   |   |   |   | XC = 0 :
|   |   |   |   |   |   | XR = 0 :
|   |   |   |   |   |   |   | XH = 0 :
|   |   |   |   |   |   |   |   | XE = 0 : 0
|   |   |   |   |   |   |   |   | XE = 1 : 1
|   |   |   |   |   |   |   | XH = 1 : 1
|   |   |   |   |   |   | XR = 1 :
|   |   |   |   |   |   |   | XB = 0 :
|   |   |   |   |   |   |   |   | XD = 0 : 0
|   |   |   |   |   |   |   |   | XD = 1 : 1
|   |   |   |   |   |   |   | XB = 1 : 0
|   |   |   |   |   | XC = 1 : 1
|   |   |   |   | XK = 1 :
```

```
|   |   |   |   |   | XD = 0 :
|   |   |   |   |   |   | XF = 0 : 0
|   |   |   |   |   |   | XF = 1 : 1
|   |   |   |   |   | XD = 1 : 0
|   |   |   | XQ = 1 :
|   |   |   |   | XM = 0 :
|   |   |   |   |   | XN = 0 :
|   |   |   |   |   |   | XU = 0 : 1
|   |   |   |   |   |   | XU = 1 : 0
|   |   |   |   |   | XN = 1 :
|   |   |   |   |   |   | XP = 0 :
|   |   |   |   |   |   |   | XB = 0 :
|   |   |   |   |   |   |   |   | XF = 0 : 0
|   |   |   |   |   |   |   |   | XF = 1 : 1
|   |   |   |   |   |   |   | XB = 1 : 1
|   |   |   |   |   |   | XP = 1 : 1
|   |   |   |   | XM = 1 : 0
|   |   |   | XS = 1 :
|   |   |   |   | XL = 0 :
|   |   |   |   |   | XD = 0 :
|   |   |   |   |   |   | XU = 0 : 1
|   |   |   |   |   |   | XU = 1 :
|   |   |   |   |   |   |   | XB = 0 :
|   |   |   |   |   |   |   |   | XE = 0 : 1
|   |   |   |   |   |   |   |   | XE = 1 :
|   |   |   |   |   |   |   |   |   | XC = 0 : 1
|   |   |   |   |   |   |   |   |   | XC = 1 : 0
|   |   |   |   |   |   |   | XB = 1 :
|   |   |   |   |   |   |   |   | XG = 0 :
|   |   |   |   |   |   |   |   |   | XH = 0 : 0
|   |   |   |   |   |   |   |   |   | XH = 1 : 1
|   |   |   |   |   |   |   |   | XG = 1 : 0
|   |   |   |   |   | XD = 1 :
|   |   |   |   |   |   | XG = 0 : 0
|   |   |   |   |   |   | XG = 1 : 1
|   |   |   |   | XL = 1 :
|   |   |   |   |   | XH = 0 :
|   |   |   |   |   |   | XD = 0 :
|   |   |   |   |   |   |   | XQ = 0 : 0
|   |   |   |   |   |   |   | XQ = 1 :
|   |   |   |   |   |   |   |   | XB = 0 : 0
|   |   |   |   |   |   |   |   | XB = 1 : 1
|   |   |   |   |   |   | XD = 1 :
|   |   |   |   |   |   |   | XB = 0 : 1
|   |   |   |   |   |   |   | XB = 1 : 0
|   |   |   |   |   | XH = 1 : 0

Pre-Pruned Accuracy
-------------------------------------------
Number of training instances = 600
Number of training attributes = 20
Total number of nodes in the tree = 275
Number of leaf nodes in the tree = 138
Accuracy of the model on the training data = 100.0%
```

```
Number of validation instances = 2000
Number of validation attributes = 21
Accuracy of the model on the validation data = 75.9%

Number of testing instances = 2000
Number of testing attributes = 21
Accuracy of the model on the testing data = 75.85%
Number of Nodes in total 275 and those to be pruned :: 56(e.g 20.0% of 275)

Pruned Tree::
XO = 0 :
 | XM = 0 :
 |  | XF = 0 :
 |  |  | XB = 0 : 0
 |  |  | XB = 1 :
 |  |  |  | XD = 0 : 0
 |  |  |  | XD = 1 :
 |  |  |  |  | XI = 0 : 0
 |  |  |  |  | XI = 1 :
 |  |  |  |  |  | XG = 0 : 1
 |  |  |  |  |  | XG = 1 : 0
 |  | XF = 1 : 0
 | XM = 1 :
 |  | XB = 0 : 1
 |  | XB = 1 :
 |  |  | XI = 0 : 0
 |  |  | XI = 1 :
 |  |  |  | XC = 0 :
 |  |  |  |  | XK = 0 :
 |  |  |  |  |  | XP = 0 : 1
 |  |  |  |  |  | XP = 1 :
 |  |  |  |  |  |  | XS = 0 :
 |  |  |  |  |  |  |  | XG = 0 : 1
 |  |  |  |  |  |  |  | XG = 1 : 1
 |  |  |  |  |  |  | XS = 1 : 0
 |  |  |  |  | XK = 1 : 0
 |  |  |  | XC = 1 : 0
XO = 1 :
 | XI = 0 :
 |  | XM = 0 : 1
 |  | XM = 1 :
 |  |  | XQ = 0 :
 |  |  |  | XF = 0 :
 |  |  |  |  | XL = 0 :
 |  |  |  |  |  | XC = 0 : 1
 |  |  |  |  |  | XC = 1 :
 |  |  |  |  |  |  | XH = 0 : 1
 |  |  |  |  |  |  | XH = 1 :
 |  |  |  |  |  |  |  | XU = 0 :
 |  |  |  |  |  |  |  |  | XB = 0 : 1
 |  |  |  |  |  |  |  |  | XB = 1 : 0
 |  |  |  |  |  |  |  | XU = 1 : 1
 |  |  |  |  | XL = 1 :
 |  |  |  |  |  | XC = 0 :
 |  |  |  |  |  |  | XB = 0 : 0
```

```
|  |  |  |  |  |  | XB = 1 :
|  |  |  |  |  |  |  | XP = 0 : 0
|  |  |  |  |  |  |  | XP = 1 : 1
|  |  |  |  |  | XC = 1 : 1
|  |  |  | XF = 1 : 0
|  |  | XQ = 1 : 0
| XI = 1 : 1


Post-Pruned Accuracy
-------------------------------------
Number of training instances = 600
Number of training attributes = 20
Total number of nodes in the pruned tree = 53
Number of leaf nodes in the pruned tree = 27
Accuracy of the model on the training data = 74.0%

Number of validation instances = 2000
Number of validation attributes = 21
Accuracy of the model on the validation data = 70.1%

Number of testing instances = 2000
Number of testing attributes = 21
Accuracy of the model on the testing data = 70.35%
```