```
■ README.md
decisionTree
An implementation of the ID3 algorithm for decision trees.
Answers to questions (by parts)
Part A
Plots of error vs tree depth
Monks-1 data
                     Training and Testing Error Curves for Monk1
     50
                                                                            Training Error
                                                                            Test Error
     40
     30
     20
     10
       0
                                                                                        10
                                                                        8
                     2
                                                Depth
Monks-2 data
                     Training and Testing Error Curves for Monk2
                                                                           Training Error
                                                                           Test Error
     35
     30
     25
  Errors
00
     15
     10
       0
                                                                                        10
                     2
                                                                        8
                                                Depth
Monks-3 data
                     Training and Testing Error Curves for Monk3
                                                                           Training Error
     50
                                                                            Test Error
     40
     30
     20
     10
       0 -
                                                                                        10
                     2
                                                                        8
                                                Depth
Part B
Learned Decision Tree and Confusion Matrices for depths 1 and 2 on Monks1 data.
Decision Tree of Depth 1
  Node value = None
    Feature f1{1, 2, 3}
              Node value = 1
              Label value = 0
              Node value = 2
              Label value = 1
              Node value = 3
              Label value = 1
The Confusion Matrix on the Training Set for Depth of 1:
048 | 014
031 | 031
The Confusion Matrix on the Test Set for Depth of 1:
144 | 072
144 | 072
Decision Tree of Depth 2
  Node value = None
    Feature f1{1, 2, 3}
              Node value = 1
               Feature f2{1, 2, 3}
                         Node value = 1
                         Label value = 1
                         Node value = 2
                         Label value = 0
                         Node value = 3
                         Label value = 0
              Node value = 2
               Feature f2{1, 2, 3}
                         Node value = 1
                         Label value = 0
                         Node value = 2
                         Label value = 1
                         Node value = 3
                         Label value = 0
              Node value = 3
               Feature f2{1, 2, 3}
                         Node value = 1
                         Label value = 0
                         Node value = 2
                         Label value = 0
                         Node value = 3
                         Label value = 1
The Confusion Matrix on the Training Set for Depth of 2:
041 | 021
000 | 062
The Confusion Matrix on the Test Set for Depth of 2:
144 | 072
000 | 216
Part C SciKit Learn's Decision tree for Monks1 data
Tree Generated by SciKit Learn on Monks data
                                        X[4] <= 1.5
gini = 0.5
samples = 124
value = [62, 62]
                                                X[0] \le 1.5
                                    gini = 0.0
                                                gini = 0.453
                                   samples = 29
                                                samples = 95
                                  value = [0, 29]
                                               value = [62, 33]
                                         gini = 0.326
                                                      gini = 0.494
                                         samples = 39
value = [31, 8]
                                                      samples = 56
                                                     value = [31, 25]
                                                                    X[3] \le 1.5
                              gini = 0.0
                                          gini = 0.0
                                                       gini = 0.0
                                                                   gini = 0.424
                             samples = 8 value = [0, 8]
                                                      samples = 20
value = [20, 0]
                                         samples = 31
                                                                   samples = 36
value = [11, 25]
                                         value = [31, 0]
                                                                                           X[5] <= 1.5
gini = 0.463
                                                       X[5] \le 1.5
                                                       gini = 0.337
                                                       value = [3, 11]
                                                                                           value = [8, 14]
                                                       X[4] <= 2.5
gini = 0.444
                                                                                                                   X[4] <= 2.5
gini = 0.337
                                                                                           X[4] \le 2.5
                                            gini = 0.0
                                                                                           gini = 0.469
                                           samples = 5
                                                       samples = 9
value = [3, 6]
                                                                                           samples = 8
                                                                                                                   samples = 14
                                           value = [0, 5]
                                                                                           value = [5, 3]
                                                                                                                  value = [3, 11]
                                                                                                                                  X[0] \le 2.5
gini = 0.397
                                                       X[4] \le 3.5
                                                                                                 X[2] \le 1.5
                                           gini = 0.0
samples = 1
                                                                                                                    gini = 0.0
                                                       gini = 0.469
                                                                                      gini = 0.5
                                                                                                                   samples = 3 value = [0, 3]
                                                       samples = 8
value = [3, 5]
                                                                                     samples = 4
value = [2, 2]
                                                                                                samples = 4
value = [3, 1]
                                                                                                                                 samples = 11
value = [3, 8]
                                           value = [0, 1]
                                                       X[1] <= 2.5
gini = 0.375
                                                                                                                                  X[1] \le 2.5
                                                                                     X[3] \le 2.5
                                                                                                             X[0] \le 2.5
                                                                          gini = 0.0
                                                                                                 gini = 0.0
                                gini = 0.5
                                                                                     gini = 0.444
                                                                                                              gini = 0.5
                                                                                                                                  gini = 0.48
                                                                                                                                                 gini = 0.278
                                                       samples = 4
                                                                          samples = 1
                                                                                                  samples = 2
                                                                                                             samples = 2
value = [1, 1]
                               samples = 4 value = [2, 2]
                                                                                                                                 samples = 5
value = [2, 3]
                                                                                     samples = 3
                                                                                                                                                 samples = 6
                                                                         value = [0, 1]
                                                                                                 value = [2, 0]
                                                       value = [1, 3]
                                                                                     value = [2, 1]
                                                                                                                                                 value = [1, 5]
              X[0] \le 2.5
gini = 0.5
                                                 X[2] \le 1.5
gini = 0.5
                                X[0] \le 2.5
                                                                         X[1] \le 2.5
                                                                                                                                                             gini = 0.0
                                                                                      gini = 0.0
                                                                                                  gini = 0.0
                                                                                                              gini = 0.0
                                                                                                                         gini = 0.0
                                                                                                                                      gini = 0.0
                                                                                                                                                 gini = 0.0
                                gini = 0.5
                                                                          gini = 0.5
                                                                                                                                                            samples = 5
value = [0, 5]
                                                             samples = 2
                                                                                      samples = 1
                                                                                                 samples = 1
                                                                                                             samples = 1
                                                                                                                         samples = 3
                                                                                                                                     samples = 2
                                                                                                                                                 samples = 1
                                                  samples = 2
              samples = 2
                                samples = 2
                                                                          samples = 2
                                                             value = [0, 2]
                                                                                     value = [1, 0]
                                                                                                 value = [0, 1]
                                                                                                             value = [1, 0]
                                                                                                                         value = [0, 3]
                                                                                                                                    value = [2, 0]
                                                                                                                                                 value = [1, 0]
                                value = [1, 1]
                                                  value = [1, 1]
                                                                         value = [1, 1]
              value = [1, 1]
                                                  gini = 0.0
  gini = 0.0
              gini = 0.0
                          gini = 0.0
                                      gini = 0.0
                                                              gini = 0.0
                                                                                      gini = 0.0
                                                                          gini = 0.0
                                                 samples = 1
value = [1, 0]
                                                                                     samples = 1
value = [1, 0]
                                                             samples = 1
             value = [1, 0]
                                     value = [0, 1]
                                                             value = [0, 1]
                                                                         value = [0, 1]
Confusion Matrix for Monks-1 Using Scikit:
200 | 016
036 | 180
Part D Custom decisionTree and SciKit Learn's Decision tree for SPECT data
Plot of error vs tree depth (decisionTree class)
                 Training and Testing Error Curves for SPECT data
     60
                                                                            Training Error
                                                                           Test Error
     55
     50
     45
  Errors
05
     35
     30
     25
                                                                                        10
                     2
                                                                        8
                                                Depth
Decision Tree of Depth 1
  Node value = None
    Feature f1{0, 1}
              Node value = 0
              Label value = 0
              Node value = 1
              Label value = 1
The Confusion Matrix on the Training Set for Depth of 1:
018 | 022
011 | 029
The Confusion Matrix on the Test Set for Depth of 1:
087 | 085
003 | 012
Decision Tree of Depth 2
  Node value = None
    Feature f1{0, 1}
              Node value = 0
               Feature f2{0, 1}
                         Node value = 0
                         Label value = 0
                         Node value = 1
                         Label value = 1
              Node value = 1
               Feature f2{0, 1}
```

```
Node value = 0
                        Label value = 1
                       Node value = 1
                        Label value = 1
The Confusion Matrix on the Training Set for Depth of 2:
022 | 018
013 | 027
The Confusion Matrix on the Test Set for Depth of 2:
105 | 067
003 | 012
Tree Generated by SciKit Learn on SPECT data
                                                                                                       X[12] <= 0.5
gini = 0.5
samples = 80
value = [40, 40]
                                                                                                   True
                                                                                                                  False
                                                                                                                   X[7] <= 0.5
gini = 0.293
                                                                                              X[10] \le 0.5
```

gini = 0.44

samples = 52 value = [35, 17] samples = 28 value = [5, 23] X[15] <= 0.5 gini = 0.343 X[0] <= 0.5 gini = 0.463 X[9] <= 0.5 gini = 0.111  $X[2] \le 0.5$ gini = 0.397samples = 17 value = [1, 16] samples = 41 value = [32, 9] samples = 11 value = [3, 8] samples = 11 value = [4, 7] X[4] <= 0.5 gini = 0.245 samples = 7 value = [1, 6]  $X[21] \le 0.5$  $X[6] \le 0.5$  $X[18] \le 0.5$ gini = 0.0gini = 0.0gini = 0.0gini = 0.444gini = 0.32samples = 10gini = 0.295samples = 2samples = 1samples = 5samples = 10samples = 39 value = [32, 7] samples = 6value = [0, 2]value = [1, 0]value = [0, 5]value = [0, 10]value = [2, 8]value = [4, 2]X[8] <= 0.5 gini = 0.251 X[9] <= 0.5 gini = 0.48  $X[5] \le 0.5$  $X[3] \le 0.5$  $X[2] \le 0.5$ gini = 0.0gini = 0.0gini = 0.444gini = 0.48gini = 0.444samples = 5value = [0, 5]samples = 4 value = [0, 4]samples = 3samples = 34 value = [29, 5] samples = 5value = [3, 2]samples = 5 value = [2, 3]samples = 3value = [1, 2]samples = 3value = [1, 2]value = [3, 0]X[18] <= 0.5 gini = 0.271  $X[1] \le 0.5$  $X[4] \le 0.5$ gini = 0.0gini = 0.444gini = 0.375samples = 2value = [2, 0]samples = 1value = [1, 0]samples = 2value = [0, 2]samples = 3samples = 2samples = 1samples = 1samples = 31samples = 3samples = 4value = [3, 0]value = [0, 2]value = [1, 0]value = [1, 0] value = [1, 2]value = [1, 3]value = [26, 5]X[19] <= 0.5  $X[19] \le 0.5$ gini = 0.0gini = 0.0gini = 0.0gini = 0.444gini = 0.245samples = 2value = [0, 2]samples = 1value = [1, 0]samples = 3value = [0, 3]samples = 1value = [1, 0]samples = 28 value = [24, 4] samples = 3value = [2, 1]gini = 0.0 samples = 1 value = [1, 0] gini = 0.5gini = 0.5gini = 0.204samples = 2 value = [1, 1] samples = 2value = [1, 1]samples = 26value = [23, 3]gini = 0.0 samples = 4 value = [4, 0] gini = 0.236samples = 22 value = [19, 3] X[7] <= 0.5 gini = 0.198 gini = 0.375 samples = 4value = [3, 1]samples = 18 value = [16, 2] X[3] <= 0.5 gini = 0.219 samples = 2samples = 16 value = [14, 2] value = [2, 0] $X[1] \le 0.5$ gini = 0.0gini = 0.231samples = 1value = [1, 0]samples = 15value = [13, 2] gini = 0.245samples = 14 value = [12, 2] samples = 1value = [1, 0]Scikit Learn on SPECT Data, Confusion Matrix: 118 | 054 003 | 012