CS 486 - Assignment 2

Bilal Khan bilal2vec@gmail.com

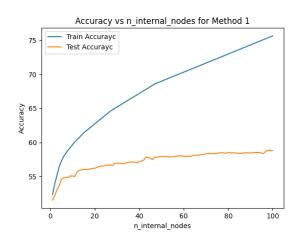
February 24, 2025

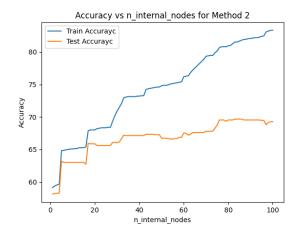
Contents

2 2	3
2.1 a	
1 1	
METHOD 1	
<pre>[leaf] pred = 1 n_doc: 19 node split_order: 5: split on 'atheir [leaf] pred = 1 n_doc: 12</pre>	ature: 4662 i_g = 0.500 n_doc: 1465
<pre>[leaf] pred = 1 n_doc: 10</pre>	a' feature: 1148 i_g = 0.497 n_doc: 1381
<pre>node split_order: 8: split on [leaf] pred = 1 n_doc: 7</pre>	'murder' feature: 2383 i_g = 0.497 n_doc: 1374 n 'proof' feature: 1417 i_g = 0.496 n_doc: 136
<pre>[leaf] pred = 1 n_doc: 6 node split_order: 10: spli- i_g = 0.496 n_doc: 1361</pre>	
[leaf] pred = 1 n_doc: 6 [leaf] pred = 2 n_doc: 1	355

METHOD 2

```
node split_order: 1: split on 'book' feature: 1135 i_g = 0.077 n_doc: 1500
  node split_order: 2: split on 'bible' feature: 5983 i_g = 0.115 n_doc: 155
    [leaf] pred = 1 n_doc: 4
    node split_order: 3: split on 'call' feature: 197 i_g = 0.069 n_doc: 151
      [leaf] pred = 1 \text{ n\_doc}: 2
      node split_order: 4: split on 'sent' feature: 2124 i_g = 0.084 n_doc: 149
        [leaf] pred = 1 n_doc: 2
        node split_order: 9: split on 'controlling' feature: 292 i_g = 0.058 n_doc: 147
          [leaf] pred = 1 n_doc: 1
          [leaf] pred = 2 \text{ n\_doc}: 146
  node split_order: 5: split on 'books' feature: 2437 i_g = 0.059 n_doc: 1345
    node split_order: 6: split on 'sure' feature: 1228 i_g = 0.097 n_doc: 83
      node split_order: 7: split on 'free' feature: 3840 i_g = 0.918 n_doc: 3
        [leaf] pred = 2 n_{doc}: 1
         [leaf] pred = 1 \text{ n\_doc}: 2
      node split_order: 8: split on 'spirit' feature: 5369 i_g = 0.096 n_doc: 80
        [leaf] pred = 1 n_doc: 1
        [leaf] pred = 2 \text{ n\_doc}: 79
    node split_order: 10: split on 'religion' feature: 5240 i_g = 0.035 n_doc: 1262
      [leaf] pred = 1 \text{ n\_doc}: 65
      [leaf] pred = 1 n_doc: 1197
```





2 2

2.1 a

```
P(OC) = 0.6
                                   P(!OC) = 0.4
                                   p(Trav) = 0.05
                                  P(!Trav) = 0.95
                            P(Fraud|Trav) = 0.01
                            P(!Fraud|Trav) = 0.99
                            P(Fraud|!Trav) = 0.004
                           P(!Fraud|!Trav) = 0.996
 P(FP|Trav + Fraud) = P(FP|Trav + !Fraud) = 0.9
P(!FP|Trav + Fraud) = P(!FP|Trav + !Fraud) = 0.1
                      P(FP|!Trav + Fraud) = 0.1
                      P(!FP|!Trav + Fraud) = 0.9
                      P(FP|!Trav+!Fraud) = 0.01
                      P(!FP|!Trav+!Fraud) = 0.99
                        P(IP|OC + Fraud) = 0.02
                        P(!IP|OC + Fraud) = 0.98
                         P(IP|OC+!Fraud) = 0.01
                        P(!IP|OC+!Fraud) = 0.99
                        P(IP|!OC + Fraud) = 0.011
                       P(!IP|!OC + Fraud) = 0.989
                        P(IP|!OC+!Fraud) = 0.001
                        P(!IP|!OC+!Fraud) = 0.999
                               P(CRP|OC) = 0.1
                              P(!CRP|OC) = 0.9
                              P(CRP|!OC) = 0.001
                             P(!CRP|!OC) = 0.999
```

2.2 b

$$\begin{split} P(Fraud) &= \sum_{Trav,FP,IP,OC,CRP} P(Trav)P(OC)P(Fraud|Trav)P(FP|Trav + Fraud) \\ P(IP|OC + Fraud)P(CRP|OC) \\ P(Fraud,FP) &= \sum_{Trav} P(Trav)P(Fraud|Trav)P(FP|Trav + Fraud) \\ &= 0.05 \times 0.01 \times 0.9 + 0.95 \times 0.004 \times 0.1 = 0.00083 \\ P(Fraud,!FP) &= \sum_{Trav} P(Trav)P(Fraud|Trav)P(!FP|Trav + Fraud) \\ &= 0.05 \times 0.01 \times 0.1 + 0.95 \times 0.004 \times 0.9 = 0.00347 \\ P(Fraud) &= P(Fraud,FP) + P(Fraud,!FP) = 0.0043 \end{split}$$

$$P(Fraud|FP+!IP+CRP) = \sum_{Trav,OC} P(Trav)P(OC)P(Fraud|Trav)P(FP|Trav+Fraud) \\ P(IP|OC+Fraud)P(CRP|OC) \\ F(Trav,OC) = 0.05 \times 0.6 \times 0.01 \times 0.9 \times 0.02 \times 0.1 = 5.4e - 7 \\ F(Trav,!OC) = 0.05 \times 0.4 \times 0.01 \times 0.9 \times 0.011 \times 0.001 = 1.98e - 9 \\ F(!Trav,OC) = 0.95 \times 0.6 \times 0.004 \times 0.1 \times 0.02 \times 0.1 = 4.56e - 7 \\ F(!Trav,!OC) = 0.95 \times 0.4 \times 0.004 \times 0.1 \times 0.011 \times 0.001 = 1.672e - 9 \\ P(Fraud|FP+!IP+CRP) = \frac{F(Trav,OC)+F(Trav,!OC)+F(!Trav,OC)+F(!Trav,!OC)}{P(Fraud)} \\ = \frac{5.4e - 7 + 1.98e - 9 + 4.56e - 7 + 1.672e - 9}{0.0043} = 0.00023$$

2.3 c

You want to make a computer purchase so that the system increases its probability of you owning a computer, which in turn reduces the probability of fraud.