DEPARTMENT OF ENGINEERING CYBERNETICS, NORWEGIAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

TDT4171 Exercise 4

Even Ødegaard, Cybernetics and Robotics

April 8, 2016

1 Gain importance

With importance defined as the expected information gain, we obtain great results (28/28), and a really simple tree is generated (Figure 1.1).
The learner based on information gain can be run several times, obtaining the exact same result.

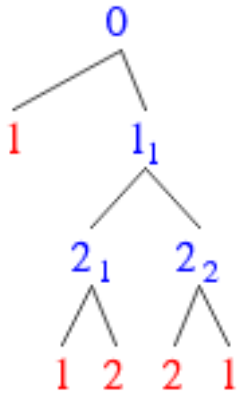

Figure 1.1: Resulting tree from learner based on information gain.

2 Random importance

By allocating a random number as importance to each attribute, we obtain varying results from time to time. Some times we get 28/28 right, and some times below 20. It also generates different trees for every run, as it chooses a random attribute every time. Two of the resulting trees can be seen in Figure 2.1 and 2.2.

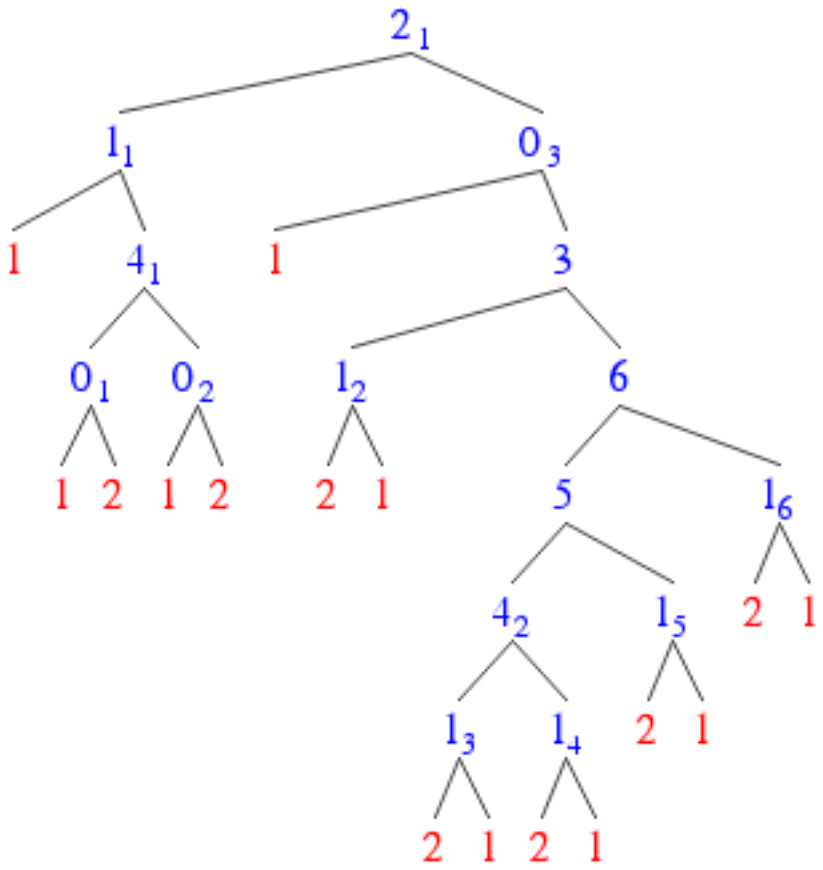

Figure 2.1: Resulting tree from learner based on random importance.

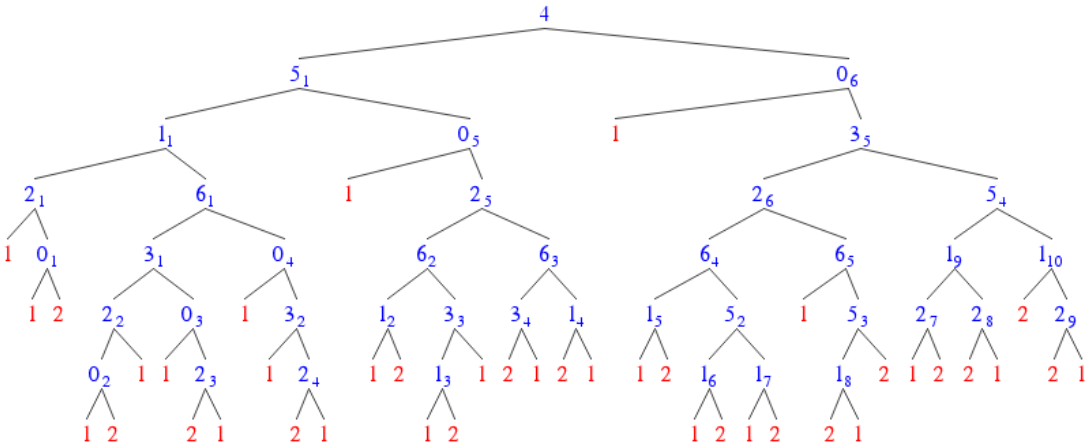

Figure 2.2: Resulting tree from learner based on random importance.

3 Conclusion

Judging by these results it is easy to conclude that defining the importance as the expected information gain is a far more suitable method for this kind of tasks.