TDT4171 — Artificial Intelligence Methods
Assignment 6 - Learning from observations

Erik Storås Sommer - 535006

March 2023

Exercise 1

Info

For debugging and learning purposes, I have used graphviz to visualize the decision trees. May therefor need to run 'pip install graphviz' to get the code to run if it is not already installed.

Performance

Running the code and measuring the time it takes iterating 100 times, and then calculating the mean and variance, we get the following results:

```
● eriksommer@Eriks-MBP learning_from_observations % python3 decision_tree.py
  Measure: random
  Number of Trials: 100
  Mean Training Accuracy: 1.0
  Mean Test Accuracy: 0.779642857142857
  Variance Training Accuracy: 0.0
  Variance Test Accuracy: 0.009440178571428572
  Total Time: 0:00:00

  Measure: information_gain
  Number of Trials: 100
  Mean Training Accuracy: 1.0
  Mean Test Accuracy: 0.9971428571428571
  Variance Training Accuracy: 0.0
  Variance Test Accuracy: 0.0004255102040816327
  Total Time: 0:00:02
```

Figure 1: Output from running the code

We see that the time it takes to run the algorithm that allocate a random number as importance to each attribute is under 1 second, while the time it takes to run the algorithm that allocate the expected information gain as importance to each attribute is around 2 seconds. This is due to the exstra calculations that need to be done to calculate the expected information gain. In the other hand we see that the accuracy of the decision tree using the expected information gain as importance to each attribute is much higher than the one using random numbers. This is because the expected information gain is a better measure of how much information an attribute gives us about the class. The variance is also less than for random allocation.

Decision trees

Below we see that the decision tree using random number as importance to each attribute is much more complex and has a greater depth than the one using the expected information gain as importance to each attribute. This makes the decision tree using the expected information gain as importance to each attribute more generalizable and less likely to overfit the data.

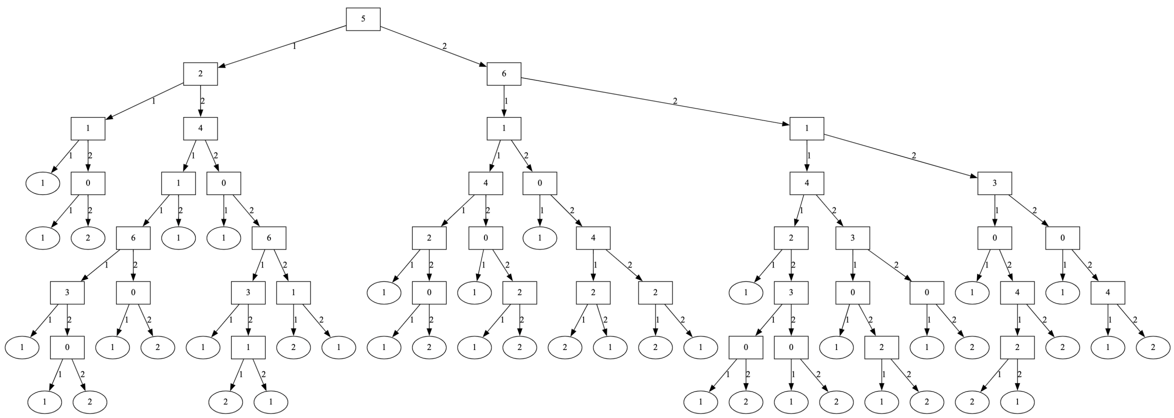

Figure 2: Decision tree using random number as importance to each attribute

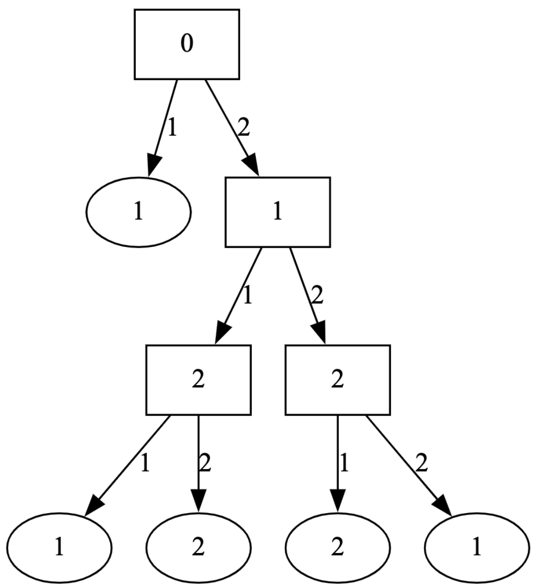

Figure 3: Decision tree using the expected information gain as importance to each attribute

Conclution

We see that the expected information gain is a better measure of how much information an attribute gives us about the class than a random number, though it is more computationally expensive to calculate.