

---

# ASSIGNMENT 4

TDT4171 - Methods in Artificial Intelligence

---

*Written by:*

Håkon Ødegård Løvdal

**Spring 2016**



Norwegian University of Science and Technology

# 1 Findings

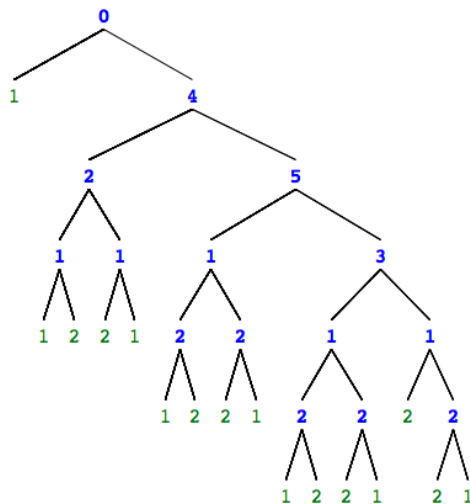
## 1.1 Discussion

After running the algorithm multiple times, with both random and gain IMPORTANCE, the gain importance proved orders of magnitude better than the random importance. The argumentation for this is that the classification of the test-set gives a correctness of 92.86 % every time. On the contrary, the random importance tree gives results ranging from 50% to 100%, but averages around 76% after 1000 runs. I will also mention that the tree produced with gain importance, is drastically smaller than the tree produced from random importance.

As stated above, the use of random importance never gives the same result. This is because of the random importance generating a completely different tree every time.

The learner based on Information Gain gives the same tree for every run. It also gives the same result when running the test-set (92.86% correctness). This is because the training-set and test-set is the same for every run, and therefore the entropy does not change.

## 1.2 Documentation



**Figure 1:** Tree generated from Gain Importance

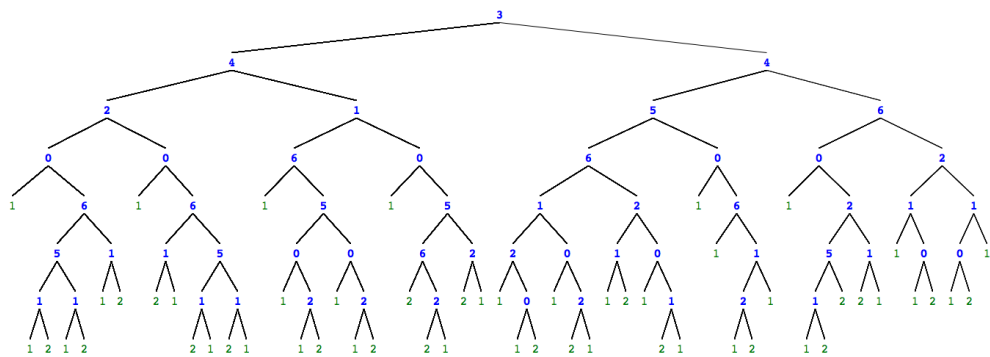
# RESULTS FOR PROVIDED TREES:

Test Result for "gain importance":

26 of 28 (92.86 %) correct classifications

Test Result for "random importance":

20 of 28 (71.43 %) correct classifications



**Figure 2:** Example of tree generated from Random Importance