

Assignment 6 - Jakob Austrheim

søndag 3. mars 2024 23:19

In this assignment I was not able to implement a working decision tree. I was struggling to understand how i could determine what attribute that would give the biggest information gain. I will work with the assignment during the next week as well, as i want to understand it properly.

However, I feel that i have understood a large part of the theory. I would expect the information gain version of importance to perform better than the random one. To determine whether a attribute is good or not one can use entropy. Entropy gives us a measure of uncertainties in the dataset. The information gain method simply subtracts the entropy in the whole dataset by the amount of entropy each different attribute would reduce the dataset with if chosen. By choosing the attributes that reduces the entropy by the most one can easily come to a decision quickly. This would be preferred to the random one, as the random method only choses the attributes to make the decision upon randomly.