FUNDAMENTALS OF ARTIFICIAL INTELLIGENCE
Exercise 12: Learning – Solutions Winter Semester 2023/24
Laura Lützow

Problem 12.1:

I want to create a decision tree to tell me whether I am likely to enjoy a particular book. Below is the data from 8 books I have read, whether I enjoyed them and the attributes:

1 Fiction/Non Fiction
2 Whether the book has $>$ 500 pages (1) or not (0)
3 Whether the book is about travel (1) or not (0)

Book number	Fiction/NF	$>$ 500 pages?	Travel?	Enjoyed?
1	N	0	0	0
2	F	1	0	0
3	N	0	0	0
4	F	1	1	0
5	N	0	1	1
6	F	0	0	1
7	N	0	1	1
8	F	0	1	1

Problem 12.1.1: Use the decision tree learning algorithm with the information gain heuristic to create a decision tree for these data.

We use the Information Gain (Entropy Reduction) heuristic to find the attributes along which to split. The information gain can be computed with:

$$Gain(A) = B\Big(\frac{p}{p+n}\Big) - Remainder(A)$$
$$Remainder(A) = \sum_{k=1}^{d} \frac{p_k + n_k}{p+n} B\Big(\frac{p_k}{p_k + n_k}\Big),$$

where A is an attribute taking d values; p and n are the number of positive and negative examples in the dataset, respectively, and p_k and n_k are the number of positive and negative examples for the k^{th} value of attribute A. For Boolean random variables, which are true with the probability q, the entropy B is defined as:

$$B(q) = -(q\, log_2(q) + (1-q)\, log_2(1-q)).$$

First, we calculate the information gain for each of the attributes "Fiction/NF", "$>$ 500 pages?", and "Travel":

$$Gain(Fiction/NF) = B(\frac{4}{8}) - \Big(\frac{4}{8}B(\frac{2}{4}) + \frac{4}{8}B(\frac{2}{4})\Big) = 0$$
$$Gain(> 500 Pages?\,) = B(\frac{4}{8}) - \Big(\frac{6}{8}B(\frac{4}{6}) + \frac{2}{8}B(\frac{0}{2})\Big) = 0.3113$$
$$Gain(Travel?\,) = B(\frac{4}{8}) - \Big(\frac{4}{8}B(\frac{1}{4}) + \frac{4}{8}B(\frac{3}{4})\Big) = 0.1887$$

Book number	Fiction/NF	$>$ 500 pages?	Travel?	Enjoyed?
1	N	0	0	0
2	F	1	0	0
3	N	0	0	0
4	F	1	1	0
5	N	0	1	1
6	F	0	0	1
7	N	0	1	1
8	F	0	1	1
Entropy Reduction	**0**	**0.3113**	**0.1887**	

Now, we choose the attribute with the highest information gain, which is "$>$ 500 pages". By splitting the original dataset along the attribute "$>$ 500 pages", we obtain two child data sets. For "$>$ 500 pages" $= 1$, we have:

Book number	Fiction/NF	$>$ 500 pages?	Travel?	Enjoyed?
2	F	1	0	0
4	F	1	1	0

Both books have the same value for our goal attribute "Enjoyed", which is 0, which means we are not likely to enjoy books with more than 500 pages.

Now we look at the subdataset where "$>$ 500 pages" $= 0$. Here, the books do not all have the same value for our goal attribute, so we have to perform another test for this subdataset. The remaining attributes along which we can split are "Fiction/NF" and "Travel?". Their information gains for the considered subdataset where "$>$ 500 pages" $= 0$ are computed as follows:

$$Gain(Fiction/NF) = B(\frac{4}{6}) - \left(\frac{4}{6}B(\frac{2}{4}) + \frac{2}{6}B(\frac{2}{2})\right) = 0.2516$$
$$Gain(Travel?) = B(\frac{4}{6}) - \left(\frac{3}{6}B(\frac{1}{3}) + \frac{3}{6}B(\frac{3}{3})\right) = 0.4591$$

Book number	Fiction/NF	$>$ 500 pages?	Travel?	Enjoyed?
1	N	0	0	0
3	N	0	0	0
5	N	0	1	1
6	F	0	0	1
7	N	0	1	1
8	F	0	1	1
Entropy Reduction	**0.2516**	**–**	**0.4591**	

As a split along the attribute "Travel?" would lead to the highest reduction in entropy, we choose this attribute.

At the end of this round, our decision tree looks like this:

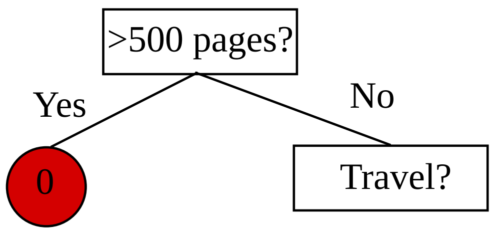

Figure 1: Intermediate Decision Tree

For "Travel?" $= 1$, we have the resulting subdataset:

Book number	Fiction/NF	$>$ 500 pages?	Travel?	Enjoyed?
5	N	0	1	1
7	N	0	1	1
8	F	0	1	1

As all books have the same value for our goal attribute, which is 1, we do not need to perform another test in this subdataset. For "Travel?" $= 0$, we have:

Book number	Fiction/NF	$>$ 500 pages?	Travel?	Enjoyed?
1	N	0	0	0
3	N	0	0	0
6	F	0	0	1
Entropy Reduction	**0.9183**	**–**	**–**	

This subdataset still contains different values for the goal attribute, so we have to perform another test. By computing the information gain for the last attribute ("Fiction/NF"), which we have left, we can check whether a split along this attribute would actually lead to a reduction in entropy:

$$Gain(Fiction/NF) = B(\frac{1}{3}) - \left(\frac{2}{3}B(\frac{0}{2}) + \frac{1}{3}B(\frac{1}{1})\right) = 0.9183$$

The information gain is bigger than zero, which means we can reduce the entropy by performing the split along "Fiction/NF". We obtain the following subdatasets:

Book number	Fiction/NF	$>$ 500 pages?	Travel?	Enjoyed?
1	N	0	0	0
3	N	0	0	0

Book number	Fiction/NF	$>$ 500 pages?	Travel?	Enjoyed?
6	F	0	0	1

For both subdatasets, we can return a clear decision: We are likely to enjoy the fictional books (with less than 500 pages and not about travel), and we are not likely to enjoy the non-fictional books (with less than 500 pages and not about travel).

Our final decision tree looks like this:

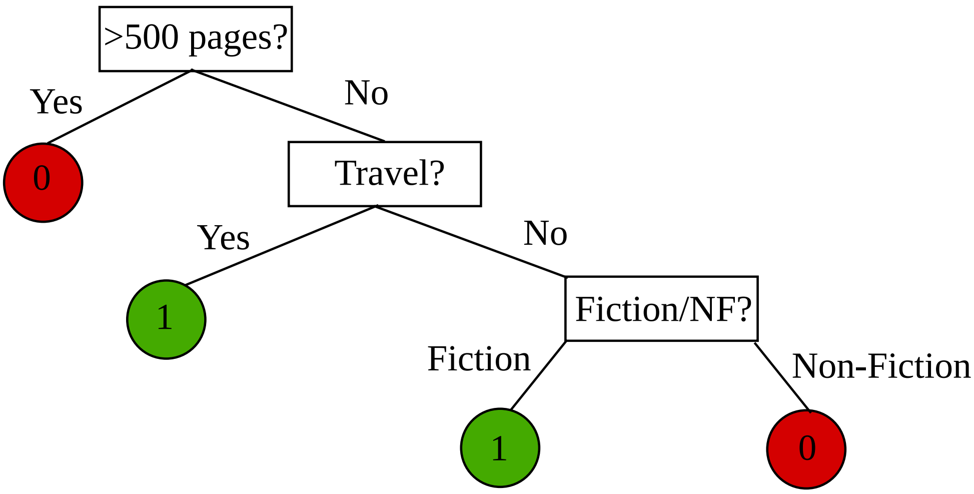

Figure 2: Final Decision Tree

Problem 12.1.2: Philip K. Dick's *Do Androids Dream of Electric Sheep* is fiction, 283 pages and not about travel. Am I likely to enjoy it?

Solution: Yes. And it is an excellent book and highly relevant to ethics in Artificial Intelligence.

Problem 12.2:

Assume we have a robot navigating a maze. The robot is trying to find the goal in the maze. The robot can move in four directions: up, down, left and right. The robot's current state is represented by the coordinates (2,3). The goal is represented by the coordinates (3,3). The robot can take one of the four actions at each state. The reward for taking an action is 1 if the action leads to the goal, 0 otherwise. The Q-table is initialized with all zeros.

Problem 12.2.1: Assume that the robot performs a move to the right. Solve one step of the Q-learning algorithm with learning rate $\alpha = 0.1$.

The Q-learning algorithm is used to find the optimal policy for the robot. The optimal policy is a mapping from states to actions that maximizes the expected reward.

At each step, the robot takes an action and receives a reward. The Q-table is then updated using the following equation:

$$Q(s,a) \leftarrow Q(s,a) + \alpha \cdot (r + \gamma \cdot \max_{a'} Q(s',a') - Q(s,a)),$$

where s is the current state, a is the action taken, r is the reward received, α is the learning rate, γ is the discount factor, s' is the next state, and a' is the action taken in the next state.

In our example, the robot takes the action of moving right. The reward for this action is 1 since it leads to the goal. The Q-table is then updated as follows:

$$\begin{aligned}
&Q(2,3,right) \leftarrow Q(2,3,right) + \alpha \cdot (1 + \gamma \cdot \max_{a'} Q(3,3,a') - Q(2,3,right)) \\
&Q(2,3,right) \leftarrow Q(2,3,right) + \alpha \cdot (1 + \gamma \cdot 0 - 0) \\
&Q(2,3,right) \leftarrow Q(2,3,right) + \alpha \\
&Q(2,3,right) \leftarrow 0 + \alpha \\
&Q(2,3,right) \leftarrow 0.1
\end{aligned}$$