Fundamentals of Artificial Intelligence
Exercise 12: Learning

Laura Lützow

Technical University of Munich

February 9th, 2024

Summary - Learning

- Learning is needed for unknown environments and lazy designers.
- Types of learning: Unsupervised learning, reinforcement learning, supervised learning.
- In supervised learning, one tries to learn a function $y = h(x)$ from input-output pairs.
- It is crucial to find a hypothesis that agrees well with the examples. **Ockham's razor** maximizes a combination of consistency and simplicity.
- **Decision trees** can represent all Boolean functions. The information-gain heuristic efficiently finds simple decision trees.
- The performance of a learning algorithm is measured by the **learning curve**, which shows the prediction accuracy on the **test set** as a function of the **training-set** size.

Recap Decision Tree Learning

- A **decision tree** represents a function that takes a vector of attribute values as input and returns a "decision" – a single Boolean output value.
- We restrict ourselves to discrete inputs.
- A decision tree reaches its decision through a sequence of tests:
  - An internal node represents a test of a property;
  - Edges are annotated with the possible test values;
  - Each leaf node has the Boolean value which should be returned.
- We prefer to find more *compact* decision trees.
- The `Decision-Tree-Learning` algorithm adopts a greedy divide-and-conquer strategy:
  1. Test the most important attribute first.
  2. The test divides the problem into two smaller subproblems that can be solved recursively.

Problem 12.1

I want to create a **decision tree** to tell me whether I am likely to enjoy a particular book.

I have three **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)

Below is the data from 8 books I have read:

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.

Information Gain

$$Gain(A) = B\left(\frac{p}{p+n}\right) - Remainder(A)$$

$$Remainder(A) = \sum_{k=1}^{d} \frac{p_k + n_k}{p+n} B\left(\frac{p_k}{p_k + n_k}\right)$$

- A is an attribute taking d values
- p and n are the number of positive and negative examples
- p_k and n_k are the number positive and negative examples of the k^{th} value of the attribute

Entropy

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

- q is the probability that a Boolean random variable is true

Split 1: Calculate the information gain for each of the attributes.

$$Gain(A) = B\left(\frac{p}{p+n}\right) - Remainder(A)$$

$$Remainder(A) = \sum_{k=1}^{d} \frac{p_k + n_k}{p+n} B\left(\frac{p_k}{p_k + n_k}\right)$$

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

A_1="**Fiction/NF**"

Nr.	F/NF	E?
1	N	0
2	F	0
3	N	0
4	F	0
5	N	1
6	F	1
7	N	1
8	F	1

Split 1: Calculate the information gain for each of the attributes.

A_2="$>$500 Pages?"

$$Gain(A) = B\left(\frac{p}{p+n}\right) - Remainder(A)$$

$$Remainder(A) = \sum_{k=1}^{d} \frac{p_k + n_k}{p+n} B\left(\frac{p_k}{p_k + n_k}\right)$$

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

Nr.	>500?	E?
1	0	0
2	1	0
3	0	0
4	1	0
5	0	1
6	0	1
7	0	1
8	0	1

Split 1: Calculate the information gain for each of the attributes.

$$Gain(A) = B\left(\frac{p}{p+n}\right) - Remainder(A)$$

$$Remainder(A) = \sum_{k=1}^{d} \frac{p_k + n_k}{p+n} B\left(\frac{p_k}{p_k + n_k}\right)$$

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

A_3="Travel?"

Nr.	T?	E?
1	0	0
2	0	0
3	0	0
4	1	0
5	1	1
6	0	1
7	1	1
8	1	1

Split 1: Choose the attribute which leads to the highest information gain.

$$Gain(''Fiction/NF'') = B\left(\frac{4}{8}\right) - \left(\frac{4}{8}B\left(\frac{2}{4}\right) + \frac{4}{8}B\left(\frac{2}{4}\right)\right) = 0$$

$$Gain(''>500\ Pages?'') = B\left(\frac{4}{8}\right) - \left(\frac{6}{8}B\left(\frac{4}{6}\right) + \frac{2}{8}B\left(\frac{0}{2}\right)\right) = 0.3113$$

$$Gain(''Travel?'') = B\left(\frac{4}{8}\right) - \left(\frac{4}{8}B\left(\frac{1}{4}\right) + \frac{4}{8}B\left(\frac{3}{4}\right)\right) = 0.1887$$

Split 1: Split the data set along the attribute ”>**500 Pages?**”.

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

Split 1: Split the data set along the attribute ">**500 Pages?**".

→ Two child data sets:

1) ">500 Pages?" = 1

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

2) ">500 Pages?" = 0

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

Split 1: Split the data set along the attribute **">500 Pages?"**.

$\rightarrow$ Intermediate decision tree:

Split 2: Calculate the information gain for the remaining two attributes.

$$Gain(A) = B\left(\frac{p}{p+n}\right) - Remainder(A)$$

$$Remainder(A) = \sum_{k=1}^{d} \frac{p_k + n_k}{p+n} B\left(\frac{p_k}{p_k + n_k}\right)$$

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

A_1**=”Fiction/NF”**

Nr.	F/NF	E?
1	N	0
3	N	0
5	N	1
6	F	1
7	N	1
8	F	1

Split 2: Calculate the information gain for the remaining two attributes.

$$Gain(A) = B\left(\frac{p}{p+n}\right) - Remainder(A)$$

$$Remainder(A) = \sum_{k=1}^{d} \frac{p_k + n_k}{p+n} B\left(\frac{p_k}{p_k + n_k}\right)$$

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

A_2=" Travel?"

Nr.	T?	E?
1	0	0
3	0	0
5	1	1
6	0	1
7	1	1
8	1	1

Split 2: Choose the attribute which leads to the highest information gain.

$$Gain(\text{''}\,Fiction/NF\text{''}) = B\left(\frac{4}{6}\right) - \left(\frac{4}{6}B\left(\frac{2}{4}\right) + \frac{2}{6}B\left(\frac{2}{2}\right)\right) = 0.2516$$

$$Gain(\text{''}\,Travel?\text{''}) = B\left(\frac{4}{6}\right) - \left(\frac{3}{6}B\left(\frac{1}{3}\right) + \frac{3}{6}B\left(\frac{3}{3}\right)\right) = 0.4591$$

Split 2: Split the data set along the attribute **"Travel?"**.

$\rightarrow$ Two child data sets:

1) "Travel?" $= 1$

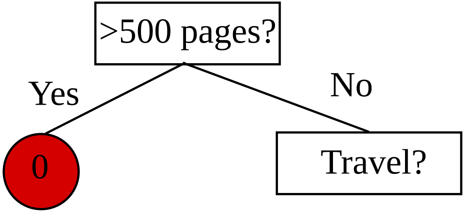
Book number	Fiction/NF	$>$ 500 Pages?	Travel?	Enjoyed?
5	N	0	1	1
7	N	0	1	1
8	F	0	1	1

2) "Travel?" $= 0$

Book number	Fiction/NF	$>$ 500 Pages?	Travel?	Enjoyed?
1	N	0	0	0
3	N	0	0	0
6	F	0	0	1

Split 2: Split the data set along the attribute **"Travel?"**.

→ Intermediate Decision Tree:



Split 3: Calculate the information gain for the remaining attribute.

A_1=**"Fiction/NF"**

$$Gain(A) = B\left(\frac{p}{p+n}\right) - Remainder(A)$$

$$Remainder(A) = \sum_{k=1}^{d} \frac{p_k + n_k}{p+n} B\left(\frac{p_k}{p_k + n_k}\right)$$

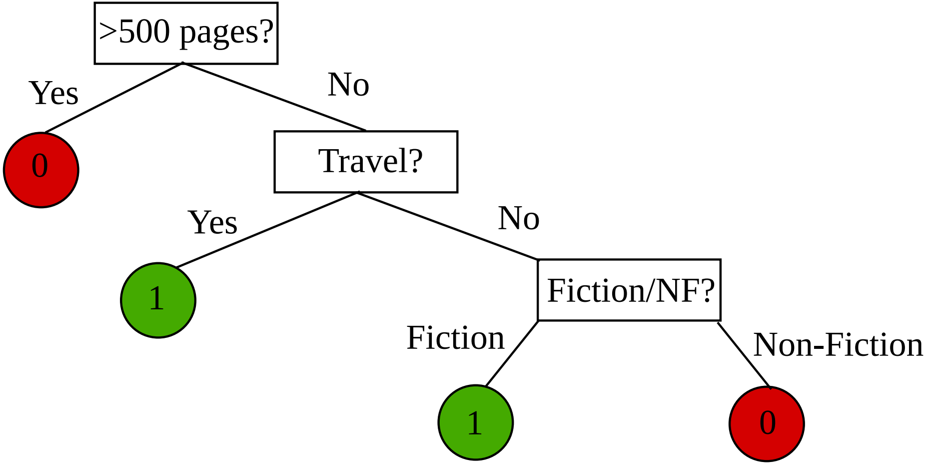
$$B(q) = -(q\, log_2(q) + (1-q)\, log_2(1-q))$$

Nr.	F/NF	E?
1	N	0
3	N	0
6	F	1

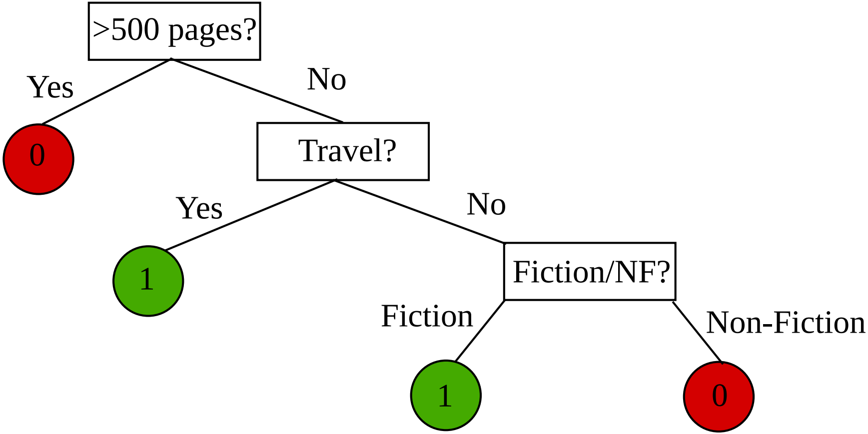
Split 3: Choose the attribute which leads to the highest information gain.

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

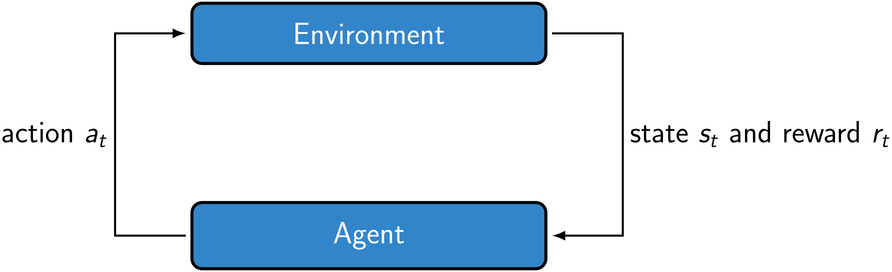
Split 3: Split the data set along the attribute **"Fiction/NF"**.
$\rightarrow$ Final Decision Tree:



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?



Recap Reinforcement Learning



- The agent observes its **state** from the environment.
- The agent acts on the environment through possible **actions**.
- The **reward** indicates how beneficial the last action was with respect to solving the task.

Recap Q-Learning

- Q-Function for the optimal policy

$$Q^*(s,a) = \sum_{s'} P(s'|s,a)[R(s,a,s') + \gamma \max_{a' \in A(s')} Q^*(s',a')],$$

- Update with a learning rate

$$Q_{i+1}(s,a) = Q_i(s,a) + \underset{\substack{\uparrow \\ \text{learning rate}}}{\alpha} \cdot \Big[\underbrace{\sum_{s'} P(s'|s,a)[R(s,a,s') + \gamma \max_{a' \in A(s')} Q_i(s',a')}_{①} - \underbrace{Q_i(s,a)}_{②} \Big]$$

①: recent value
②: prior value
① – ② is called the temporal difference (TD) error.

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 $\alpha = 0.1$.

Initial Q-table:

State	$\uparrow$	$\downarrow$	$\rightarrow$	$\leftarrow$
...	...	...	...	...
(2,3)	0	0	0	0
(3,3)	0	0	0	0

$s = (2, 3) \qquad s' = (3, 3) \qquad a = \rightarrow \qquad R(s, a, (3, 3)) = 1$

Q-learning update equation:

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