

COMP3411 Tutorial - Week 8

Reinforcement Learning and Decision Trees

Question 1 - Q-Learning

Consider a world with two states $S = \{S_1, S_2\}$ and two actions $A = \{a_1, a_2\}$, where the transitions δ and reward r for each state and action are as follows:

$$\begin{aligned}\delta(S_1, a_1) &= S_1 & r(S_1, a_1) &= 0 \\ \delta(S_1, a_2) &= S_2 & r(S_1, a_2) &= -1 \\ \delta(S_2, a_1) &= S_2 & r(S_2, a_1) &= +1 \\ \delta(S_2, a_2) &= S_1 & r(S_2, a_2) &= +5\end{aligned}$$

- (i) Draw a picture of this world, using circles for the states and arrows for the transitions.
- (ii) Assuming a discount factor of $\gamma = 0.9$, determine:
- (a) the optimal policy $\pi^* : S \rightarrow A$
 - (b) the value function $V^* : S \rightarrow R$
 - (c) the Q function $Q : S \times A \rightarrow R$
- (iii) Write the Q values in a table.

Q	a_1	a_2
S_1		
S_2		

- (iv) Trace through the first few steps of the Q-learning algorithm, with all Q values initially set to zero. Explain why it is necessary to force exploration through probabilistic choice of actions in order to ensure convergence to the true Q values.

Question 2 - Decision Trees

Consider the task of predicting whether children are likely to be hired to play members of the Von Trapp Family in a production of The Sound of Music, based on these data:

height	hair	eyes	hired
short	blond	blue	+
tall	red	blue	+
tall	blond	blue	+
tall	blond	brown	-
short	dark	blue	-
tall	dark	blue	-
tall	dark	brown	-
short	blond	brown	-

1. Compute the information gain (entropy) for each of the three attributes (height, hair, eyes) in terms of classifying objects as belonging to the class, + or - .
2. Construct a decision tree based on the minimum entropy principle.

Question 3 - Pruning

The Laplace error estimate for pruning a node in a Decision Tree is given by:

$$E = 1 - \frac{n + 1}{N + k}$$

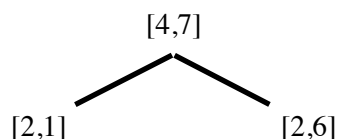
where

N is the total number of items

n is the number of items in the majority class

k is the number of classes

Given the following subtrees should the children be pruned or not? Show your calculations



Question 4

Construct a Decision Tree for the following set of examples.

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

What class is assigned to the instance {D15, Sunny, Hot, High, Weak}?