FIT3080 Assignment 2 Report

Priscilla A. C. Tham

Monash University Malaysia

Priscilla A. C. Tham is now at School of Information Technology, Monash University Malaysia.

ptha0007@student.monash.edu

FIT3080 Assignment 2 Report

I. Implementation of Decision Trees

I.I Which value of depth gave the best test set accuracy? Which value of depth gave the best training set accuracy? Does increasing the depth lead to overfitting on this dataset?

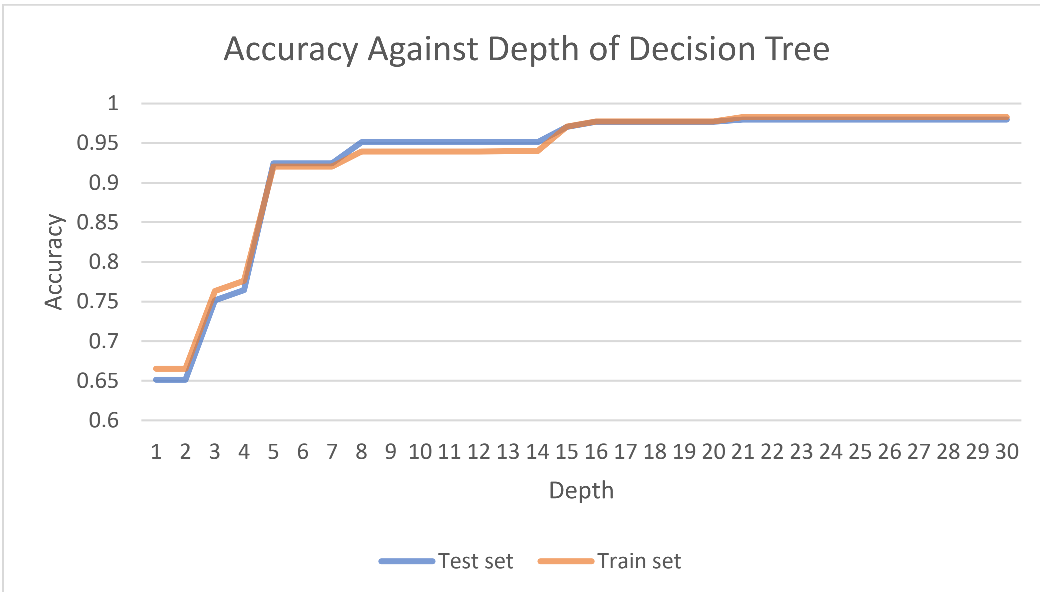

Figure I Accuracy Against Depth of Decision Tree

The accuracy of a decision tree is evaluated with a learning curve. Figure I showed the learning curve for the decision three algorithm on examples obtained from train.txt. Blue refers to the **accuracy** tested **on testing set** whereas orange refers to the **accuracy** tested **on training set**.

First, the tree is trained with the training set consisting of 2109 examples. Its accuracy is then measured with the test set and the train set itself for comparison. The procedure runs by increasing the depth of the tree one at a time from depth = 1 until depth = 30. For each depth, the tree splits by the maximum information gain heuristic.

$$Gain(F) = 1 - \sum_{k=1}^{d} \frac{p_k}{p_k + n_k} B(\frac{p_k}{p_k + n_k})$$

where (1)

F = some attribute F

B(q) = the entropy of a Boolean random variable that is true with probability q

$\frac{p_k}{p_k+n_k}$ = the probability q at the k^{th} example

d = the total number of examples

The minimum sum of entropy of all examples allows us to deduce the maximum information gain following the equation (1).

The curve shows that as the depth of the tree increases, the accuracy also increases. According to the recorded accuracy, testing set has the accuracy of 97% since depth = 21 whereas training set has the accuracy of 98% since depth = 21. This suggests further increase in accuracy should the depth increases further although small.

Now, as both curves are similar and very much aligned, it is safe to assume that the increase in depth did not contribute to overfitting. The curves did not diverge from each other even after reaching the maximum accuracy which indicates that depth is not the controlling variable towards overfitting. Furthermore, the alignment of both curves instead suggests the ideal coverage of the hypothesis space. It means that the hypothesis space constructed upon learning on the training set can accurately classifies the test set examples.

II. (MDP & RL) Soccer

(a) What is $V^{\pi}(1)$ for the policy $\pi(\cdot) = a_S$, the policy that always shoots?

(a)
$$V^{\pi}(1) = T(1, a_S, S_G)\,[R(1, a_S, S_G) + \gamma V^{*}(S_G)] + T(1, a_S, S_M)\,[R(1, a_S, S_M) + \gamma V^{*}(S_M)]$$
$$= \tfrac{1}{6}(1+0) + \tfrac{5}{6}(0+0)$$
$$= \tfrac{1}{6}$$

(b) What is $Q^{*}(3, a_D)$ in terms of y?

(b)
$$Q^{*}(3, a_D) = T(3, a_D, 4)\,[R(3, a_D, 4) + \gamma V^{*}(4)] + T(3, a_D, S_M)\,[R(3, a_D, S_M) + \gamma V^{*}(S_M)]$$
$$= y\,(0 + V^{*}(4)) + \tfrac{3}{6}(0+0)$$
$$= y\,(V(4))$$
$$= y\,[T(4, a_S, S_G)\{R(4, a_S, S_G) + \gamma V^{*}(S_G)\} + T(4, a_S, S_M)\{R(4, a_S, S_M) + \gamma V^{*}(S_M)\}]$$
$$= y\,[\tfrac{4}{6}(1+0) + \tfrac{2}{6}(0+0)]$$
$$= y\,\tfrac{4}{6}$$

(c) Using $y = 3/4$, complete the first two iterations of value iteration

iteration	Q(1, aD)	Q(2, aD)	Q(3, aD)
0	0	0	0
1	0	0	0
2	1/4	3/8	1/2

iteration	Q(1, aS)	Q(2, aS)	Q(3, aS)	Q(4, aS)
0	0	0	0	0
1	1/6	1/3	1/2	2/3
2	1/6	1/3	1/2	2/3

V(s) = max(Q(s, aD), Q(s, aS))				
iteration	V(1)	V(2)	V(3)	V(4)
0	0	0	0	0
1	1/6	1/3	1/2	2/3
2	1/4	3/8	1/2	2/3

Calculations:

(c)

$$Q_1(1, a_D) = T(1, a_D, 2)[R(1, a_D, 2) + \gamma V_0^*(2)] + T(1, a_D, s_m)[R(1, a_D, s_m) + \gamma V_0^*(s_m)]$$
$$= y(0+0) + \tfrac{5}{6}(0,0)$$
$$= 0$$

$$Q_1(2, a_D) = T(2, a_D, 3)[R(2, a_D, 3) + \gamma V_0^*(3)] + T(2, a_D, s_m)[R(2, a_D, s_m) + \gamma V_0^*(s_m)]$$
$$= y(0+0) + \tfrac{4}{6}(0+0)$$
$$= 0$$

$$Q_1(3, a_D) = T(3, a_D, 4)[R(3, a_D, 4) + \gamma V_0^*(4)] + T(3, a_D, s_m)[R(3, a_D, s_m) + \gamma V_0^*(s_m)]$$
$$= y(0+0) + \tfrac{3}{6}(0+0)$$
$$= 0$$

$$
\begin{aligned}
Q_1(1,a_s) &= T(1,a_s,s_G)\,[R(1,a_s,s_G)+\gamma V_0^*(s_G)] \\
&\quad + T(1,a_s,s_m)\,[R(1,a_s,s_m)+\gamma V_0^*(s_m)] \\
&= \tfrac{1}{6}(1+0)+\tfrac{5}{6}(0+0) \\
&= \tfrac{1}{6}
\end{aligned}
$$

$$
\begin{aligned}
Q_1(2,a_s) &= T(2,a_s,s_G)\,[R(2,a_s,s_G)+\gamma V_0^*(s_G)] \\
&\quad + T(2,a_s,s_m)\,[R(2,a_s,s_m)+\gamma V_0^*(s_m)] \\
&= \tfrac{2}{6}(1+0)+\tfrac{4}{6}(0+0) \\
&= \tfrac{2}{6} = \tfrac{1}{3}
\end{aligned}
$$

$$
\begin{aligned}
Q_1(3,a_s) &= T(3,a_s,s_G)\,[R(3,a_s,s_G)+\gamma V_0^*(s_G)] \\
&\quad + T(3,a_s,s_m)\,[R(3,a_s,s_m)+\gamma V_0^*(s_m)] \\
&= \tfrac{3}{6}(1+0)+\tfrac{3}{6}(0+0) \\
&= \tfrac{3}{6} = \tfrac{1}{2}
\end{aligned}
$$

$$
\begin{aligned}
Q_1(4,a_s) &= T(4,a_s,s_G)\,[R(4,a_s,s_G)+\gamma V_0^*(s_G)] \\
&\quad + T(4,a_s,s_m)\,[R(4,a_s,s_m)+\gamma V_0^*(s_m)] \\
&= \tfrac{4}{6}(1+0)+\tfrac{2}{6}(0+0) \\
&= \tfrac{4}{6} = \tfrac{2}{3}
\end{aligned}
$$

$$
\begin{aligned}
Q_2(1, a_D) &= T(1, a_D, 2)\,[R(1, a_D, 2) + \gamma V_1^*(2)] \\
&\quad + T(1, a_D, s_m)\, R(1, a_0, s_M) + \gamma V_1^*(s_M)] \\
&= y\,(0 + \tfrac{1}{3}) + \tfrac{5}{6}(0+0) \\
&= \tfrac{3}{4}\,(\tfrac{1}{3}) \\
&= \tfrac{1}{4}
\end{aligned}
$$

$$
\begin{aligned}
Q_2(2, a_D) &= T(2, a_0, 3)\,[R(2, a_D, 3) + \gamma V_1^*(3)] \\
&\quad + T(2, a_D, s_m)\,[R(3, a_D, s_m) + \gamma V_1^*(s_m)] \\
&= y\,(0 + \tfrac{1}{2}) + \tfrac{4}{6}(0+0) \\
&= \tfrac{3}{4}\,(\tfrac{1}{2}) \\
&= \tfrac{3}{8}
\end{aligned}
$$

$$
\begin{aligned}
Q_2(3, a_D) &= T(3, a_D, 4)\,[R(3, a_D, 4) + \gamma V_1^*(4)] \\
&\quad + T(3, a_D, s_m)\,[R(3, a_D, s_m) + \gamma V_1^*(s_m)] \\
&= y\,(0 + \tfrac{2}{3}) + \tfrac{3}{6}(0+0) \\
&= \tfrac{3}{4}\,(\tfrac{2}{3}) \\
&= \tfrac{1}{2}
\end{aligned}
$$

$$
\begin{aligned}
Q_2(1, a_s) &= T(1, a_s, s_G)[R(1, a_s, s_G) + \gamma V_1^*(s_G)] \\
&\quad + T(1, a_s, s_m)[R(1, a_s, s_m) + \gamma V_1^*(s_m)] \\
&= \tfrac{1}{6}(1+0) + \tfrac{5}{6}(0+0) \\
&= \tfrac{1}{6}
\end{aligned}
$$

$$
\begin{aligned}
Q_2(2, a_s) &= T(2, a_s, s_G)[R(2, a_s, s_G) + \gamma V_1^*(s_G)] \\
&\quad + T(2, a_s, s_m)[R(2, a_s, s_m) + \gamma V_1^*(s_m)] \\
&= \tfrac{2}{6}(1+0) + \tfrac{4}{6}(0+0) \\
&= \tfrac{2}{6}
\end{aligned}
$$

$$
\begin{aligned}
Q_2(3, a_s) &= T(3, a_s, s_G)[R(3, a_s, s_G) + \gamma V_1^*(s_G)] \\
&\quad + T(3, a_s, s_m)[R(3, a_s, s_m) + \gamma V_1^*(s_m)] \\
&= \tfrac{3}{6}(1+0) + \tfrac{3}{6}(0+0) \\
&= \tfrac{3}{6}
\end{aligned}
$$

$$
\begin{aligned}
Q_2(4, a_s) &= T(4, a_s, s_G)[R(4, a_s, s_G) + \gamma V_1^*(s_G)] \\
&\quad + T(4, a_s, s_m)[R(4, a_s, s_m) + \gamma V_1^*(s_m)] \\
&= \tfrac{4}{6}(1+0) + \tfrac{2}{6}(0+0) \\
&= \tfrac{4}{6}
\end{aligned}
$$

(d) After how many iterations will value iteration compute the optimal values for all states, for this problem?

iteration	Q(1, aD)	Q(2, aD)	Q(3, aD)
0	0	0	0
1	0	0	0
2	1/4	3/8	1/2
3	2/7	3/8	1/2
4	2/7	3/8	1/2

iteration	Q(1, aS)	Q(2, aS)	Q(3, aS)	Q(4, aS)
0	0	0	0	0
1	1/6	1/3	1/2	2/3
2	1/6	1/3	1/2	2/3
3	1/6	1/3	1/2	2/3
4	1/6	1/3	1/2	2/3

V(s) = max(Q(s, aD), Q(s, aS))				
iteration	V(1)	V(2)	V(3)	V(4)
0	0	0	0	0
1	1/6	1/3	1/2	2/3
2	1/4	3/8	1/2	2/3
3	2/7	3/8	1/2	2/3
4	2/7	3/8	1/2	2/3

Calculation:

(c)

$$\begin{aligned} Q_1(1, a_D) &= T(1, a_D, 2)[R(1, a_D, 2) + \gamma V_0^*(2)] \\ &\quad + T(1, a_D, s_m)[R(1, a_D, s_m) + \gamma V_0^*(s_m)] \\ &= y(0+0) + \tfrac{5}{6}(0,0) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Q_1(2, a_D) &= T(2, a_D, 3)[R(2, a_D, 3) + \gamma V_0^*(3)] \\ &\quad + T(2, a_D, s_m)[R(2, a_D, s_m) + \gamma V_0^*(s_m)] \\ &= y(0+0) + \tfrac{4}{6}(0+0) \\ &= 0 \end{aligned}$$

$$\begin{aligned} Q_1(3, a_D) &= T(3, a_D, 4)[R(3, a_D, 4) + \gamma V_0^*(4)] \\ &\quad + T(3, a_D, s_m)[R(3, a_D, s_m) + \gamma V_0^*(s_m)] \\ &= y(0+0) + \tfrac{3}{6}(0+0) \\ &= 0 \end{aligned}$$

$$Q_1(1,a_s) = T(1,a_s,s_G)[R(1,a_s,s_G)+\gamma V_0^*(s_G)] + T(1,a_s,s_m)[R(1,a_s,s_m)+\gamma V_0^*(s_m)]$$
$$= \frac{1}{6}(1+0) + \frac{5}{6}(0+0)$$
$$= \frac{1}{6}$$

$$Q_1(2,a_s) = T(2,a_s,s_G)[R(2,a_s,s_G)+\gamma V_0^*(s_G)] + T(2,a_s,s_m)[R(2,a_s,s_m)+\gamma V_0^*(s_m)]$$
$$= \frac{2}{6}(1+0) + \frac{4}{6}(0+0)$$
$$= \frac{2}{6} = \frac{1}{3}$$

$$Q_1(3,a_s) = T(3,a_s,s_G)[R(3,a_s,s_G)+\gamma V_0^*(s_G)] + T(3,a_s,s_m)[R(3,a_s,s_m)+\gamma V_0^*(s_m)]$$
$$= \frac{3}{6}(1+0) + \frac{3}{6}(0+0)$$
$$= \frac{3}{6} = \frac{1}{2}$$

$$Q_1(4,a_s) = T(4,a_s,s_G)[R(4,a_s,s_G)+\gamma V_0^*(s_G)] + T(4,a_s,s_m)[R(4,a_s,s_m)+\gamma V_0^*(s_m)]$$
$$= \frac{4}{6}(1+0) + \frac{2}{6}(0+0)$$
$$= \frac{4}{6} = \frac{2}{3}$$

$$
\begin{aligned}
Q_2(1, a_D) &= T(1, a_D, 2)\,[R(1, a_D, 2) + \gamma V_1^*(2)] \\
&\quad + T(1, a_D, s_m)\, R(1, a_0, s_M) + \gamma V_1^*(s_M)] \\
&= y\,(0 + \tfrac{1}{3}) + \tfrac{5}{6}(0+0) \\
&= \tfrac{3}{4}(\tfrac{1}{3}) \\
&= \tfrac{1}{4}
\end{aligned}
$$

$$
\begin{aligned}
Q_2(2, a_D) &= T(2, a_0, 3)\,[R(2, a_D, 3) + \gamma V_1^*(3)] \\
&\quad + T(2, a_D, s_m)\,[R(3, a_D, s_m) + \gamma V_1^*(s_m)] \\
&= y\,(0 + \tfrac{1}{2}) + \tfrac{4}{6}(0+0) \\
&= \tfrac{3}{4}(\tfrac{1}{2}) \\
&= \tfrac{3}{8}
\end{aligned}
$$

$$
\begin{aligned}
Q_2(3, a_D) &= T(3, a_D, 4)\,[R(3, a_D, 4) + \gamma V_1^*(4)] \\
&\quad + T(3, a_D, s_m)\,[R(3, a_D, s_m) + \gamma V_1^*(s_m)] \\
&= y\,(0 + \tfrac{2}{3}) + \tfrac{3}{6}(0+0) \\
&= \tfrac{3}{4}(\tfrac{2}{3}) \\
&= \tfrac{1}{2}
\end{aligned}
$$

$$
\begin{aligned}
Q_2(1, a_s) &= T(1, a_s, s_G)\,[R(1, a_s, s_G) + \gamma V_1^*(s_G)] \\
&\quad + T(1, a_s, s_m)\,[R(1, a_s, s_m) + \gamma V_1^*(s_m)] \\
&= \tfrac{1}{6}(1+0) + \tfrac{5}{6}(0+0) \\
&= \tfrac{1}{6}
\end{aligned}
$$

$$
\begin{aligned}
Q_2(2, a_s) &= T(2, a_s, s_G)\,[R(2, a_s, s_G) + \gamma V_1^*(s_G)] \\
&\quad + T(2, a_s, s_m)\,[R(2, a_s, s_m) + \gamma V_1^*(s_m)] \\
&= \tfrac{2}{6}(1+0) + \tfrac{4}{6}(0+0) \\
&= \tfrac{2}{6}
\end{aligned}
$$

$$
\begin{aligned}
Q_2(3, a_s) &= T(3, a_s, s_G)\,[R(3, a_s, s_G) + \gamma V_1^*(s_G)] \\
&\quad + T(3, a_s, s_m)\,[R(3, a_s, s_m) + \gamma V_1^*(s_m)] \\
&= \tfrac{3}{6}(1+0) + \tfrac{3}{6}(0+0) \\
&= \tfrac{3}{6}
\end{aligned}
$$

$$
\begin{aligned}
Q_2(4, a_s) &= T(4, a_s, s_G)\,[R(4, a_s, s_G) + \gamma V_1^*(s_G)] \\
&\quad + T(4, a_s, s_m)\,[R(4, a_s, s_m) + \gamma V_1^*(s_m)] \\
&= \tfrac{4}{6}(1+0) + \tfrac{2}{6}(0+0) \\
&= \tfrac{4}{6}
\end{aligned}
$$

$$
\begin{aligned}
Q_3(1,a_D) &= T(1,a_D,2)\,[R(1,a_D,2) + \gamma V_2^*(2)] \\
&\quad + T(1,a_D,S_m)\,[R(1,a_D,S_m) + \gamma V_2^*(S_m)] \\
&= \gamma\left(0 + \tfrac{3}{8}\right) + \tfrac{5}{6}(0+0) \\
&= \tfrac{3}{4}\left(\tfrac{3}{8}\right) \\
&= \tfrac{9}{32} \approx \tfrac{2}{7}
\end{aligned}
$$

$$
\begin{aligned}
Q_3(2,a_D) &= T(2,a_D,3)\,[R(2,a_D,3) + \gamma V_2^*(3)] \\
&\quad + T(2,a_D,S_m)\,[R(2,a_D,S_m) + \gamma V_2^*(S_m)] \\
&= \gamma\left(0 + \tfrac{1}{2}\right) + \tfrac{4}{6}(0+0) \\
&= \tfrac{3}{4}\left(\tfrac{1}{2}\right) \\
&= \tfrac{3}{8}
\end{aligned}
$$

$$
\begin{aligned}
Q_3(3,a_D) &= T(3,a_D,4)\,[R(3,a_D,4) + \gamma V_2^*(4)] \\
&\quad + T(3,a_D,S_m)\,[R(3,a_D,S_m) + \gamma V_2^*(S_m)] \\
&= \gamma\left(0 + \tfrac{2}{3}\right) + \tfrac{3}{6}(0+0) \\
&= \tfrac{3}{4}\left(\tfrac{2}{3}\right) \\
&= \tfrac{1}{2}
\end{aligned}
$$

$$Q_3(1,a_s) = T(1,a_s,S_G)[R(1,a_s,S_G) + \gamma V_2^*(S_G)] + T(1,a_s,S_m)[R(1,a_s,S_m) + \gamma V_2^*(S_m)]$$
$$= \frac{1}{6}(1+0) + \frac{5}{6}(0+0)$$
$$= \frac{1}{6}$$

$$Q_3(2,a_s) = T(2,a_s,S_G)[R(2,a_s,S_G) + \gamma V_2^*(S_G)] + T(2,a_s,S_m)[R(2,a_s,S_m) + \gamma V_2^*(S_m)]$$
$$= \frac{2}{6}(1+0) + \frac{4}{6}(0+0)$$
$$= \frac{1}{3}$$

$$Q_3(3,a_s) = T(3,a_s,S_G)[R(3,a_s,S_G) + \gamma V_2^*(S_G)] + T(3,a_s,S_m)[R(3,a_s,S_m) + \gamma V_2^*(S_m)]$$
$$= \frac{3}{6}(1+0) + \frac{3}{6}(0+0)$$
$$= \frac{1}{2}$$

$$Q_3(4,a_s) = T(4,a_s,S_G)[R(4,a_s,S_G) + \gamma V_2^*(S_G)] + T(4,a_s,S_m)[R(4,a_s,S_m) + \gamma V_2^*(S_m)]$$
$$= \frac{4}{6}(1+0) + \frac{2}{6}(0+0)$$
$$= \frac{2}{3}$$

$$Q_4(1,a_D) = T(1,a_D,2)\,[R(1,a_D,2) + \gamma V_3^*(2)]$$
$$+\, T(1,a_D,S_m)\,[R(1,a_D,S_m) + \gamma V_3^*(S_m)]$$
$$= y\,(0 + \tfrac{3}{8}) + \tfrac{5}{6}\,(0+0)$$
$$= \tfrac{3}{4}\,(\tfrac{3}{8})$$
$$= \tfrac{9}{32} \approx \tfrac{2}{7}$$

$$Q_4(2,a_D) = T(2,a_D,3)\,[R(2,a_D,3) + \gamma V_3^*(3)]$$
$$+\, T(2,a_D,S_m)\,[R(2,a_D,S_m) + \gamma V_3^*(S_m)]$$
$$= y\,(0 + \tfrac{1}{2}) + \tfrac{4}{6}\,(0+0)$$
$$= \tfrac{3}{4}\,(\tfrac{1}{2})$$
$$= \tfrac{3}{8}$$

$$Q_4(3,a_D) = T(3,a_D,4)\,[R(3,a_D,4) + \gamma V_3^*(4)]$$
$$+\, T(3,a_D,S_m)\,[R(3,a_D,S_m) + \gamma V_3^*(S_m)]$$
$$= y\,(0 + \tfrac{2}{3}) + \tfrac{3}{6}\,(0+0)$$
$$= \tfrac{3}{4}\,(\tfrac{2}{3})$$
$$= \tfrac{1}{2}$$

$$Q_4(1,a_s) = T(1,a_s,S_G)[R(1,a_s,S_G)+\gamma V_3^*(S_G)] + T(1,a_s,S_m)[R(1,a_s,S_m)+\gamma V_3^*(S_m)]$$
$$= \frac{1}{6}(1+0) + \frac{5}{6}(0+0)$$
$$= \frac{1}{6}$$

$$Q_4(2,a_s) = T(2,a_s,S_G)[R(2,a_s,S_G)+\gamma V_3^*(S_G)] + T(2,a_s,S_m)[R(2,a_s,S_m)+\gamma V_3^*(S_m)]$$
$$= \frac{2}{6}(1+0) + \frac{4}{6}(0+0)$$
$$= \frac{1}{3}$$

$$Q_4(3,a_s) = T(3,a_s,S_G)[R(3,a_s,S_G)+\gamma V^*(S_G)] + T(3,a_s,S_m)[R(3,a_s,S_m)+\gamma V^*(S_m)]$$
$$= \frac{3}{6}(1+0) + \frac{3}{6}(0+0)$$
$$= \frac{1}{2}$$

$$Q_4(4,a_s) = T(4,a_s,S_G)[R(4,a_s,S_G)+\gamma V_3^*(S_G)] + T(4,a_s,S_m)[R(4,a_s,S_m)+\gamma V_3^*(S_m)]$$
$$= \frac{4}{6}(1+0) + \frac{2}{6}(0+0)$$
$$= \frac{2}{3}$$

Therefore, iteration = 3 computed the optimal values for all states. The purpose of iteration = 4 is to indicate the computation of the optimal values in the previous iteration when no further updates are made, as the values no longer change.

(e) For what range of values of y is $Q^*(3,a_S) \geq Q^*(3,a_D)$?

(e)

$$Q(3, a_s) \geq Q(3, a_D)$$

$$T(3, a_s, s_G)[R(3, a_s, s_G) + \gamma V^*(s_G)] + T(3, a_s, s_m)[R(3, a_s, s_m) + \gamma V^*(s_m)] \geq T(3, a_D, 4)[R(3, a_D, 4) + \gamma V^*(4)] \quad T(3, a_D, s_m)[R(3, a_D, s_m) + \gamma V^*(s_m)]$$

$$\frac{3}{6}(1+0) + \frac{3}{6}(0+0) \geq y(0 + V^*(4)) + \frac{3}{6}(0+0)$$

$$\frac{3}{6} \geq y V^*(4)$$

$$\geq y[T(4, a_s, s_G)[R(4, a_s, s_G) + \gamma V^*(s_G)] + T(4), a_s, s_m)[R(4), a_D, s_m) + \gamma V^*(s_m)]]$$

$$\geq y \frac{4}{6}$$

$$3 \geq 4y$$

$$\frac{3}{4} \geq y$$

$$y \leq \frac{3}{4}$$

Therefore, $0 < y \leq \frac{3}{4}$ because y cannot be negative as y is a probability.

(f) Now consider Q-learning, in which we do not have a model (T, y, and k above), and instead learn from a series of experienced transitions. Using a learning rate of α=1⁄2 execute Q-learning on these episodes:

iteration	Q(1, aD)	Q(1, aS)	Q(2, aD)	Q(2, aS)	Q(3, aD)	Q(3, aS)	Q(4, aS)
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
4	0	0	0	0	0	0	1/2
5	0	0	0	0	0	0	1/2
6	0	0	0	0	0	0	1/2
7	0	0	0	0	1/4	0	1/2
8	0	0	0	0	1/4	0	1/4
9	0	0	0	0	1/4	0	1/4
10	0	0	1/8	0	1/4	0	1/4
11	0	0	1/8	0	1/4	1/2	1/4
12	0.06	0	1/8	0	1/4	1/2	1/4
13	0.06	0	1/8	0	1/4	1/2	1/4
14	0.09	0	1/8	0	1/4	1/2	1/4
15	0.09	0	1/3	0	1/4	1/2	1/4
16	0.06	0	1/3	0	1/8	1/2	1/4

Calculations:

(f) $Q_1(1, a_D) \leftarrow (1-\alpha)\, Q(1, a_D) + \alpha[R(1, a_D, 2) + \gamma V_0^*(2)]$

$\leftarrow \frac{1}{2} Q_0(1, a_D) + \frac{1}{2}(0+0)$

$\leftarrow \frac{1}{2}(0) + \frac{1}{2}(0+0)$

$\leftarrow 0$

$$Q_2(2, a_D) \leftarrow (1-\alpha)\, Q_0(2, a_D) + \alpha\, [R(2, a_D, 3) + \gamma V^*(3)]$$
$$\leftarrow \tfrac{1}{2}(0) + \tfrac{1}{2}(0+0)$$
$$\leftarrow 0$$

$$Q_3(3, a_D) \leftarrow (1-\alpha)\, Q_0(3, a_D) + \alpha\, [R(3, a_D, 4) + \gamma_2 V^*(4)]$$
$$\leftarrow \tfrac{1}{2}(0) + \tfrac{1}{2}(0+0)$$
$$\leftarrow 0$$

$$Q_4(4, a_S) \leftarrow (1-\alpha)\, Q_0(4, a_S) + \alpha\, [R(4, a_S, S_G) + \gamma V_3^*(S_G)]$$
$$\leftarrow \tfrac{1}{2}(0) + \tfrac{1}{2}(1+0)$$
$$\leftarrow \tfrac{1}{2}$$

$$Q_5(1, a_D) \leftarrow (1-\alpha)\, Q_4(1, a_D) + \alpha\, [R(1, a_D, 2) + \gamma V_4^*(2)]$$
$$\leftarrow \tfrac{1}{2}(0) + \tfrac{1}{2}(0+0)$$
$$\leftarrow 0$$

$$Q_6(2, a_D) \leftarrow (1-\alpha)\, Q_5(2, a_D) + \alpha\, [R(2, a_D, 3) + \gamma V_5^*(3)]$$
$$\leftarrow \tfrac{1}{2}(0) + \tfrac{1}{2}(0+0)$$
$$\leftarrow 0$$

$$Q_7(3, a_D) \leftarrow (1-\alpha)\, Q_6(3, a_D) + \alpha\, [R(3, a_D, 4) + \gamma V_6^*(4)]$$
$$\leftarrow \tfrac{1}{2}(0) + \tfrac{1}{2}(0+\tfrac{1}{2})$$
$$\leftarrow \tfrac{1}{4}$$

$$Q_8(4, a_S) \leftarrow (1-\alpha)\, Q_7(4, a_S) + \alpha\, [R(4, a_S, S_m) + \gamma V_7^*(S_m)]$$
$$\leftarrow \tfrac{1}{2}(\tfrac{1}{2}) + \tfrac{1}{2}(0+0)$$
$$\leftarrow \tfrac{1}{4}$$

$$Q_9(1, a_D) \leftarrow (1-\alpha)\, Q_8(1, a_D) + \alpha\, [R(1, a_D, 2) + \gamma V_8^*(2)]$$
$$\leftarrow \tfrac{1}{2}(0) + \tfrac{1}{2}(0+0)$$
$$\leftarrow 0$$

$$Q_{10}(2, a_D) \leftarrow (1-\alpha)\, Q_9(2, a_D) + \alpha\, [R(2, a_D, 3) + \gamma V_9^*(3)]$$
$$\leftarrow \tfrac{1}{2}(0) + \tfrac{1}{2}(0 + \tfrac{1}{4})$$
$$\leftarrow \tfrac{1}{8}$$

$$Q_{11}(3, a_S) \leftarrow (1-\alpha)\, Q_{10}(3, a_S) + \alpha\, [R(3, a_S, S_G) + \gamma V_{10}(S_G)]$$
$$\leftarrow \tfrac{1}{2}(0) + \tfrac{1}{2}(1 + 0)$$
$$\leftarrow \tfrac{1}{2}$$

$$Q_{12}(1, a_D) \leftarrow (1-\alpha)\, Q_{11}(1, a_D) + \alpha\, [R(1, a_D, 2) + \gamma V_{11}(2)]$$
$$\leftarrow \tfrac{1}{2}(0) + \tfrac{1}{2}(0 + \tfrac{1}{8})$$
$$\leftarrow \tfrac{1}{16}$$

$$Q_{13}(2, a_S) \leftarrow (1-\alpha)\, Q_{12}(2, a_S) + \alpha\, [R(2, a_S, S_m) + \gamma V^*(S_m)]$$
$$\leftarrow \tfrac{1}{2}(0) + \tfrac{1}{2}(0 + 0)$$
$$\leftarrow 0$$

$$Q_{14}(1, a_D) \leftarrow (1-\alpha)\, Q_{13}(1, a_D) + \alpha\, [R(1, a_D, 2) + \gamma V^*(2)]$$
$$\leftarrow \tfrac{1}{2}(\tfrac{1}{16}) + \tfrac{1}{2}(0 + \tfrac{1}{8})$$
$$\leftarrow \tfrac{1}{32} + \tfrac{1}{16}$$
$$\leftarrow \tfrac{3}{32}$$

$$Q_{15}(2, a_D) \leftarrow (1-\alpha)\, Q_{14}(2, a_D) + \alpha\, [R(2, a_D, 3) + \gamma V^*(3)]$$
$$\leftarrow \tfrac{1}{2}(\tfrac{1}{8}) + \tfrac{1}{2}(0 + \tfrac{1}{2})$$
$$\leftarrow \tfrac{1}{16} + \tfrac{1}{4}$$
$$\leftarrow \tfrac{5}{16}$$

$$Q_{16}(3, a_D) \leftarrow (1-\alpha)\, Q_{15}(3, a_D) + \alpha\, [R(3, a_D, S_m) + \gamma V^*(S_m)]$$
$$\leftarrow \tfrac{1}{2}(\tfrac{1}{4}) + \tfrac{1}{2}(0 + 0)$$
$$\leftarrow \tfrac{1}{8}$$

Appendix

I.I Code Structure

small_test.txt is a file with categorical data used to debug the implemented decision tree. It is assumed that the first few values on each line are binary features and the last values on each line is the binary label while debugging. On the other hand, train.txt and test.txt are files with categorical data with the first 38 values on each line as the binary features and the last value on each line is the binary label.

The files are opened and read into the system using the function read_datafile. The function returns two list, 1) list of all 38 features of each line and 2) list of all label of each line. This is later converted to numpy 2-dimensional array in the main function.

compute_accuracy

compute_accuracy is a function to evaluate the accuracy of the decision tree predicting a case correctly. Evaluation is based on the It returns accuracy in the form of probability; hence, the probability format will be used throughout this report.

DecisionTree

DecisionTree is the class handling training and testing the decision tree. The class initializes the depth limit (depth_limit) which limits the height of the tree during training.

train

train is a method inside the DecisionTree class implementing the decision tree algorithm. It “creates” a decision tree by splitting the best feature (parent node) at each iteration until it ends with a label as its leaf nodes. The best feature is chosen by evaluating each possible feature splitting among 38 × number of lines in data maximum information gain heuristic.

$$Gain(F) = 1 - P(positive)H(positive) + P(negative)H(negative)$$

where (1)

F = some feature F

positive = left child node

negative = right child node

Simplified, the method ***calculateEntropy*** calculates H(positive) and H(negative) whereas the method ***calculateOverallEntropy*** calculates P(positive)H(positive) + P(negative)H(negative) portion. The method ***chooseBestSplit*** evaluates all possible splits obtained from ***getPotentialSplits*** according to their overall entropy. The minimum overall entropy will result in the maximum information gain following (1).

predict

predict is a method to classify a case passed using the decision tree "created" from train method. By running through the decision tree from the parent node (a condition), going left if the case value of the feature in the parent node satisfies the condition or right otherwise, we would reach a leaf node containing a label. This label is the predicted classification of the case.