National University of Singapore
School of Computing
CS3244 Machine Learning

Tutorial 3: Decision Trees

Issue: February 10, 2022 Due: February 14, 2022

Important Instructions:

- *Your solutions for this tutorial must be TYPE-WRITTEN.*
- *Make TWO copies of your solutions: one for you and one to be SUBMITTED TO THE TUTOR IN CLASS. Your submission in your respective tutorial class will be used to indicate your CLASS ATTENDANCE. Late submission will NOT be entertained.*
- *Indicate your NAME, STUDENT NUMBER, and TUTORIAL GROUP in your submitted solution.*
- *YOUR SOLUTIONS TO QUESTIONS* **TM 3.1** *will be GRADED for this tutorial.*
- *You may discuss the content of the questions with your classmates. But everyone should work out and write up ALL the solutions by yourself.*

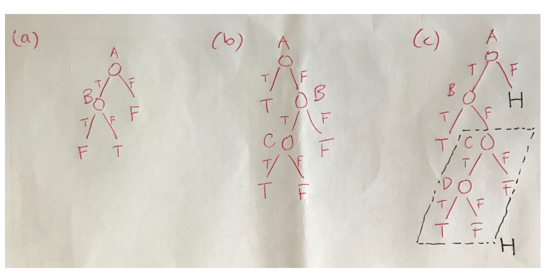
TM 3.1 Give the **smallest** possible decision trees to represent the following boolean functions:

(a) $A \wedge \neg B$

(b) $A \vee (B \wedge C)$

(c) $(A \wedge B) \vee (C \wedge D)$

Solution.



BL 4 This question is recycled from my CS3243 Intro. to AI class because it demonstrates the use of the PLURALITY-VALUE function in the DECISION-TREE-LEARNING algorithm. We will discuss its solution during our tutorial session.

The loans department of a bank has the following past loan processing records, each of which contains an applicant's income, credit history, debt, and the final approval decision. These records can serve as training examples to build a decision tree for a loan advisory system.

Income	CreditHistory	Debt	Decision
$0-5$K	Bad	Low	Reject
$0-5$K	Good	Low	Approve
$0-5$K	Unknown	High	Reject
$0-5$K	Unknown	Low	Approve
$0-5$K	Unknown	Low	Approve
$0-5$K	Unknown	Low	Reject
$5-10$K	Bad	High	Reject
$5-10$K	Good	High	Approve
$5-10$K	Unknown	High	Approve
$5-10$K	Unknown	Low	Approve
Over 10K	Bad	Low	Reject
Over 10K	Good	Low	Approve

(a) Construct a decision tree based on the above training examples. (Note: $\log_2 \frac{x}{y} = \log_2 x - \log_2 y$, $\log_2 1 = 0$, $\log_2 2 = 1$, $\log_2 3 = 1.585$, $\log_2 4 = 2$, $\log_2 5 = 2.322$, $\log_2 6 = 2.585$, $\log_2 7 = 2.807$, $\log_2 8 = 3$, $\log_2 9 = 3.170$, $\log_2 10 = 3.322$, $\log_2 11 = 3.459$, and $\log_2 12 = 3.585$)

Solution.

$$\begin{aligned}
Gain(Decision, Income) &= B(\frac{5}{12}) - H(Decision|Income) \\
H(Decision|Income) &= \frac{6}{12}(-\frac{3}{6}\log_2\frac{3}{6} - \frac{3}{6}\log_2\frac{3}{6}) + \frac{4}{12}(-\frac{1}{4}\log_2\frac{1}{4} \\
&\quad -\frac{3}{4}\log_2\frac{3}{4}) + \frac{2}{12}(-\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2}) \\
&= 0.937 \\
Gain(Decision, Income) &= 0.98 - 0.937 = 0.043
\end{aligned}$$

$$\begin{aligned} H(Decision|CreditHistory) &= \frac{3}{12}(-\frac{3}{3}\log_2\frac{3}{3} - \frac{0}{3}\log_2\frac{0}{3}) + \frac{3}{12}(-\frac{3}{3}\log_2\frac{3}{3} \\ &\quad -\frac{0}{3}\log_2\frac{0}{3}) + \frac{6}{12}(-\frac{2}{6}\log_2\frac{2}{6} - \frac{4}{6}\log_2\frac{4}{6}) \\ &= 0.459 \\ Gain(Decision, CreditHistory) &= 0.98 - 0.459 = 0.521 \end{aligned}$$

$$\begin{aligned} H(Decision|Debt) &= \frac{8}{12}(-\frac{3}{8}\log_2\frac{3}{8} - \frac{5}{8}\log_2\frac{5}{8}) + \frac{4}{12}(-\frac{2}{4}\log_2\frac{2}{4} - \frac{2}{4}\log_2\frac{2}{4}) \\ &= 0.970 \\ Gain(Decision, Debt) &= 0.98 - 0.970 = 0.01 \end{aligned}$$

Since $CreditHistory$ has the highest $Gain$, choose it as the root, which has three values, i.e., Bad, $Good$, and $Unknown$. Since all examples for Bad have the same classification (i.e., $Reject$) and all examples for $Good$ have the same classification (i.e., $Approve$), both nodes have no further subtree. For $Unknown$, a subtree for the following subset of examples is to be constructed:

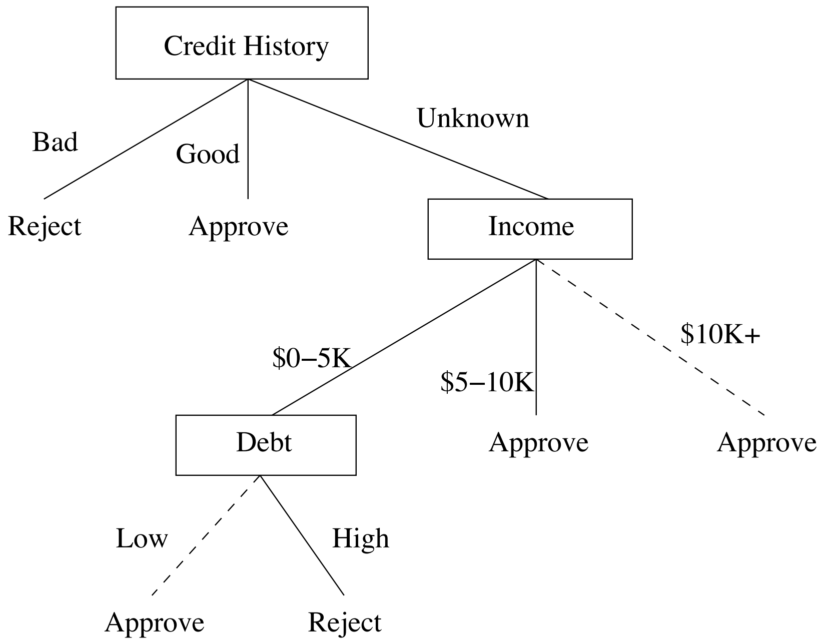
Income	Debt	Decision
$0-5$K	High	Reject
$0-5$K	Low	Approve
$0-5$K	Low	Approve
$0-5$K	Low	Reject
$5-10$K	High	Approve
$5-10$K	Low	Approve

$$B(\frac{2}{6}) = -\frac{2}{6}\log_2\frac{2}{6} - \frac{4}{6}\log_2\frac{4}{6} = 0.918$$

$$\begin{aligned} H(Decision|Income) &= \frac{4}{6}(-\frac{2}{4}\log_2\frac{2}{4} - \frac{2}{4}\log_2\frac{2}{4}) + \frac{2}{6}(-\frac{2}{2}\log_2\frac{2}{2} - \frac{0}{2}\log_2\frac{0}{2}) = 0.667 \\ Gain(Decision, Income) &= 0.918 - 0.667 = 0.251 \end{aligned}$$

$$\begin{aligned} H(Decision|Debt) &= \frac{2}{6}(-\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2}) + \frac{4}{6}(-\frac{1}{4}\log_2\frac{1}{4} - \frac{3}{4}\log_2\frac{3}{4}) = 0.874 \\ Gain(Decision, Debt) &= 0.918 - 0.874 = 0.044 \end{aligned}$$

Since $Income$ has a higher gain than $Debt$, $Income$ is chosen as the root of the subtree under $CreditHistory = Unknown$.



The two decisions on dotted lines are those that invoke the PLURALITY-VALUE function. In the case where $Income = Over10$K, there is **no example** left. So, we do a simple majority vote on the parent examples, which gives the $Approve$ decision. In the other case where $Debt = Low$, there is **no attribute** left. The resulting set of examples is not decisive. There are two approves and one reject. Again, we do a majority vote on this set of examples to give the $Approve$ decision.

Question: Why do we have such a node where there seems to be ambiguity? There is no clear answer (or at least we don't have the information to determine). One possibility is that we are missing some attribute from the data that will allow us to differentiate at this last node. Another possibility is that there is some non-determinism in the underlying function that we are trying to approximate with this decision tree.

(b) What is decision tree classifier's decision for a person who has 4K yearly income, a good credit history, and a high amount of debt?

Solution. For the individual, the decision tree approves of the application from the root node of the tree.

BL 5 This question is also recycled from my CS3243 Intro. to AI class. We will discuss its solution during our tutorial session.

Given the following training examples about exotic dishes, we want to predict whether or not a dish is appealing based on the input attributes 'Temperature', 'Taste', and 'Size'.

ID	Temperature	Taste	Size	Appealing
1	Hot	Salty	Small	No
2	Cold	Sweet	Large	No
3	Cold	Sweet	Large	No
4	Cold	Sour	Small	Yes
5	Hot	Sour	Small	Yes
6	Hot	Salty	Large	No
7	Hot	Sour	Large	Yes
8	Cold	Sweet	Small	Yes
9	Cold	Sweet	Small	Yes
10	Hot	Salty	Large	No

(a) What is the information gain $Gain(Appealing, Taste)$ associated with choosing the 'Taste' attribute as the root of the decision tree?

Solution.

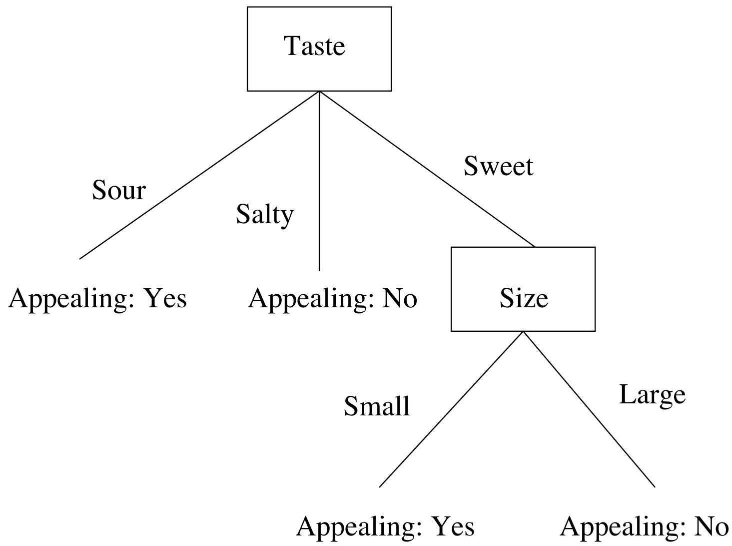
$$\begin{aligned}
Gain(Appealing, Taste) &= B\left(\frac{p}{p+n}\right) - H(Appealing|Taste) \\
H(Appealing|Taste) &= \sum_{i=1}^{d} \frac{p_i+n_i}{p+n} B\left(\frac{p_i}{p_i+n_i}\right) \\
B\left(\frac{5}{10}\right) &= -\frac{5}{10}\log_2\frac{5}{10} - \frac{5}{10}\log_2\frac{5}{10} = 1
\end{aligned}$$

There are three branches 'Salty', 'Sweet', and 'Sour':
Salty - #Yes = 0, #No = 3, Total = 3
Sweet - #Yes = 2, #No = 2, Total = 4
Sour - #Yes = 3, #No = 0, Total = 3

$$\begin{aligned}
H(Appealing|Taste) &= \frac{3}{10}B\left(\frac{0}{3}\right) + \frac{4}{10}B\left(\frac{2}{4}\right) + \frac{3}{10}B\left(\frac{3}{3}\right) \\
&= \frac{4}{10} \times 1 = \frac{2}{5} \\
Gain(Appealing, Taste) &= B\left(\frac{5}{10}\right) - H(Appealing|Taste) \\
&= 1 - \frac{2}{5} = \frac{3}{5}
\end{aligned}$$

(b) Draw a decision tree with 'Taste' as the root.

Solution.



(c) Use the decision tree to predict the class value for the record given by

ID	Temperature	Taste	Size
11	Hot	Salty	Small
12	Cold	Sweet	Large

Solution.

ID 11 - Predicted Appealing = No
ID 12 - Predicted Appealing = No.