Computational Intelligence

A Logical Approach

Problems for Chapter 11

Here are some problems to help you understand the material in Computational Intelligence: *A Logical Approach*. They are designed to help students understand the material and practice for exams.

This file is available in html, or in pdf format, either without solutions or with solutions. (The pdf can be read using the free acrobat reader or with recent versions of Ghostscript).

1 Decision-tree Learning

Consider the data on 4 Boolean attributes a, b, c, and d, where d is the target classification.

	a	b	c	d
e_1	true	true	false	false
e_2	false	true	false	true
e_3	false	true	true	true
e_4	false	false	true	false
e_5	true	false	false	false

In this question we will consider decision-tree learning based on this data.

(a) What is a good attribute to split on first? Explain why.

(b) Draw a decision tree that the top-down myopic decision tree learning algorithm could build. For each node (including the leaves) show which examples are used to determine the classification at that node. (The root note of the tree will be labelled with the list of all of the examples).

(c) Explain how the learning bias inherent in learning decision-trees can be used to classify unseen instances. Give an instance that is not in the training data, show how the above tree classifies that instance. Justify why this is an appropriate classification.

Solution to part (a)

What is a good attribute to split on first? Explain why.

a is a good attribute to split on, as when a is true, all of the examples agree on the value of attribute d, and when a is false the examples are skewed.

b is also a good attribute to split on, as it divides the examples into the same ratios as a.

c is a worse attribute to split on, as knowing the value of c doesn't help us much in determining the value of d.

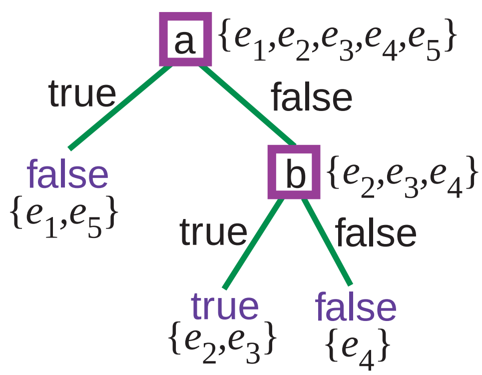
Solution to part(b)

Draw a decision tree that the top-down myopic decision tree learning algorithm could build. For each node (including the leaves) show which examples are used to determine the classification at that node.

Suppose you split on attribute a first. There are 2 examples with attribute a true, namely e_1 and e_5, and these examples agree on the value of d. There are 3 examples with attribute a false, namely e_2, e_3, and e_4, these don't agree on the value of d.

Next you have to choose an attribute to split the examples $\{e_2, e_3, e_4\}$. b is a good choice as the examples with b true all agree on the value of d and the examples with b false all agree on the value of d.

Here is the resultant decision tree showing which examples are used at each node:



Solution to part (c)

Explain how the learning bias inherent in learning decision-trees can be used to classify unseen instances. Give an instance that is not in the training data, show how the above tree classifies that instance. Justify why this is an appropriate classification.

From the 5 assignments of values to a, b and c in the training set, out of the 8 possible assignments, the decision tree makes predictions in all 8 cases. The ability to make predictions in unseen cases is the bias of the learning algorithm. The unseen cases are:

a	b	c	prediction for d
true	true	true	false
true	false	true	false
false	false	false	false

The bias is that c is always irrelevant to the classification, and that when a is true, b is also irrelevant. This is because c always seemed to be irrelevant in the training example. Similarly b was irrelevant when a was true.

2 Decision Tree Evaluation

Write a program that evaluates binary decision trees. A binary decision tree is either a value or of the form $if(Att = Val, T_1, T_2)$, where T_1 and T_2 are decision trees.

You should assume that all of the data on examples is given using the relation:

$$prop(Obj, Att, Val)$$

You need to write a relation:

$$dteval(Obj, DT, Val)$$

That is true if object Obj is classified by decision tree DT as having value Val.

For example, suppose example e_1 defined by

$$prop(e_1, a, true).$$
$$prop(e_1, b, true).$$
$$prop(e_1, c, false).$$

The query

$$?dteval(e_1, if(b = true, if(a = true, true, false), if(c = true, false, true)), Val).$$

has as its answer $Val = true$.

Axiomatize $dteval$. You can assume the predicate $value(V)$ that is true if V is a legal value, as well as the predicates $\neq$ and $prop$.

Solution to decision tree evaluation

$$
\begin{aligned}
& dteval(Ex, V, V) \leftarrow \\
& \quad value(V). \\
& dteval(Ex, if(C, T1, T2), V) \leftarrow \\
& \quad prop(Ex, C, true) \land \\
& \quad dteval(Ex, T1, V). \\
& dteval(Ex, if(C, T1, T2), V) \leftarrow \\
& \quad prop(Ex, C, false) \land \\
& \quad dteval(Ex, T2, V).
\end{aligned}
$$

3 Neural Network

Suppose that a neural network learner uses the network corresponding to the rule:

$$
\begin{aligned}
&\mathit{predicted\_prop}(\mathit{Obj}, d, V) \leftarrow\\
&\quad \mathit{prop}(\mathit{Obj}, a, I_1) \wedge\\
&\quad \mathit{prop}(\mathit{Obj}, b, I_2) \wedge\\
&\quad \mathit{prop}(\mathit{Obj}, c, I_3) \wedge\\
&\quad V \text{ is } f(w_0 + w_1 * I_1 + w_2 * I_2 + w_3 * I_3).
\end{aligned}
$$

where f is the sigmoid function.

$$f(x) = \frac{1}{1 + e^{-x}}$$

(The only property of f you need for this exam is that $f(x) > 0.5$ if and only if $x > 0$.)

Suppose that, after learning, the parameters had the following weights:

w_0	-3
w_1	2
w_2	2
w_3	4

Suppose the neural network classifies as true any example where the predicted value for d is greater than 0.5

(a) How is example e_1 classified, where e_1 is defined by:

$$
\begin{aligned}
&\mathit{prop}(e_1, a, 1).\\
&\mathit{prop}(e_1, b, 1).\\
&\mathit{prop}(e_1, c, 0).
\end{aligned}
$$

(b) Give an example that is classified differently by the neural network and the decision tree

$$\mathit{if}(b = 1, \mathit{if}(a = 1, \mathit{true}, \mathit{false}), \mathit{if}(c = 1, \mathit{false}, \mathit{true})).$$

(which is equivalent to the example decision tree given in the previous problem).

(c) Draw a decision tree that represents the same Boolean function as that represented by the neural network.

Solution to part (a)

How is example e_1 classified, where e_1 is defined by:

$$
\begin{aligned}
&\mathit{prop}(e_1, a, 1).\\
&\mathit{prop}(e_1, b, 1).\\
&\mathit{prop}(e_1, c, 0).
\end{aligned}
$$

The output to the neural network is

$$f(-3+2*1+2*1+4*0)$$

Which is $f(1)$. Thus it classifies e_1 as true.

Solution to part (b)

Give an example that is classified differently by the neural network and the decision tree

$$if(b=1, if(a=1, true, false), if(c=1, false, true)).$$

An example with $b = 1$, $a = 0$, $c = 1$ is classified as false by the decision tree, yet true by the neural network.

Solution to part (c)

Draw a decision tree that represents the same Boolean function as that represented by the neural network.