CS322 Fall 1999
Module 12 (Neural Network Learning)

Assignment 12

Solution

Question 1

The following is the same data from assignment 11:

Example	*bought*	*edu*	*first*	*visited*	*more_info*
e_1	false	true	false	false	true
e_2	true	false	true	false	false
e_3	false	false	true	true	true
e_4	false	false	true	false	false
e_5	false	false	false	true	false
e_6	true	false	false	true	true
e_7	true	false	false	false	true
e_8	false	true	true	true	false
e_9	false	true	true	false	false
e_{10}	true	true	true	false	true
e_{11}	true	true	false	true	true
e_{12}	false	false	false	false	true

We want to use this data to learn the value of *more_info* as a function of the values of the other variables.

In this assignment we will consider neural network learning for this data. We have a Java applet and a CILog program that can be used to answer this assignment.

(a) Consider neural network learning with no hidden layers. After the network has converged, what are the parameter values? What is the Boolean function that the network represents? Are all the training examples classified correctly (if not, which aren't)? Give two examples, not in the training set, and specify what the predicted values is.

(b) Consider neural network learning with one hidden layer containing two variables. After the network has converged, what are the parameter values? What is the Boolean function that the network represents? Are all the training examples classified correctly (if not, which aren't)? Give two examples, not in the training set, and specify what the predicted values is.

(c) For the network with a hidden layer what is a local minima of the learning rate (within one decimal point)? The value to minimize is the number of steps before the error gets below 1.0. Hint: there is a local minima in the range [0.3, 7.0].

Solution

(a) Consider neural network learning with no hidden layers.

i) After the network has converged, what are the parameter values?

After 200 iterations with a learning rate of 0.5 the parameter values are:

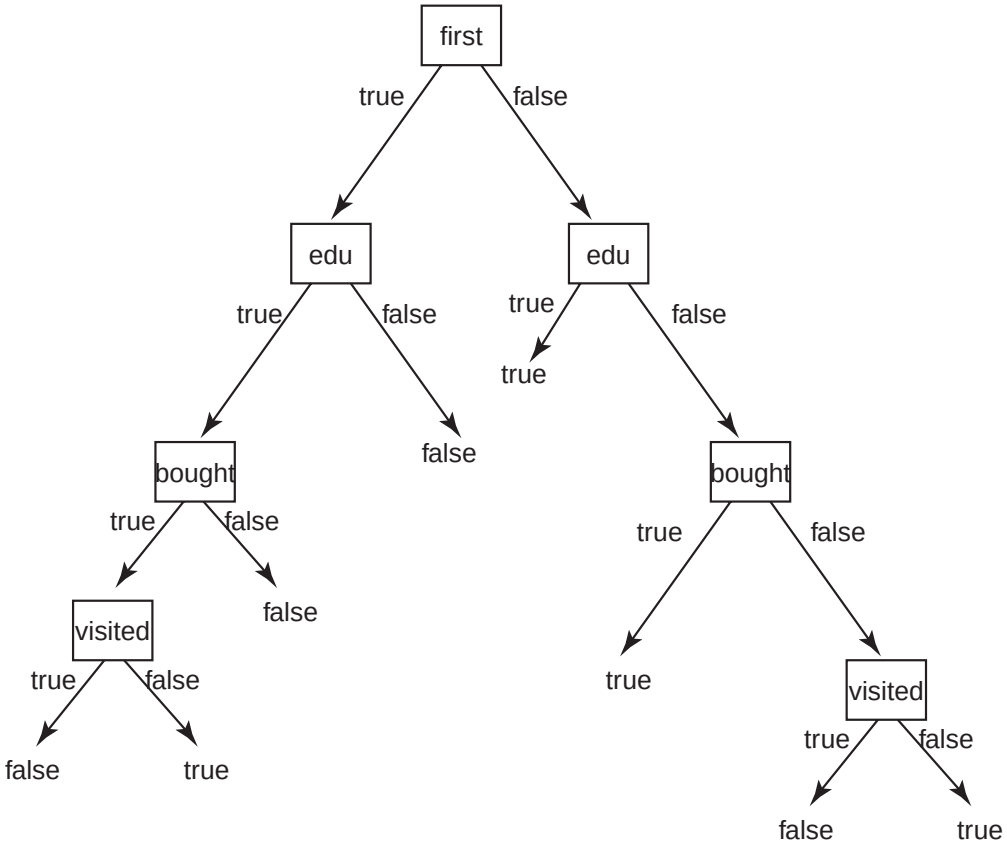
Parameter	Parent	Value
w_0		1.58
w_4	bought	3.96
w_3	edu	3.52
w_2	first	−7.42
w_1	visited	−3.40

ii) What is the Boolean function that the network represents?

When *first* is true, the value of the linear expression is negative unless *bought* and *edu* are true and *visited* is false.

When *first* is false, the value of the linear expression is positive unless *bought* and *edu* are false and *visited* is true.

This can be written as the decision tree:



So the boolean expression is:
(*first* ∧ *bought* ∧ *edu* ∧ ¬*visited*) ∨
(¬*first* ∧ *bought*) ∨
(¬*first* ∧ *edu*) ∨
(¬*first* ∧ ¬*visited*).

iii) Are all the training examples classified correctly (if not, which aren't)?

No. *e*3 is misclassified. The neural network classifies it as false.

iv) Give two examples, not in the training set, and specify what the predicted values is.

The following

bought	*edu*	*first*	*visited*	*more_info*
true	true	true	true	false
true	true	false	false	true
true	false	true	true	false
false	true	false	true	true

(b) Consider neural network learning with one hidden layer containing two variables.

i) After the network has converged, what are the parameter values?
run the applet....

ii) What is the Boolean function that the network represents?
After 200 iterations with learning rate of 0.5, we can have the following table:

bought	*edu*	*first*	*visited*	*more_info*
true	true	true	true	false
true	true	true	false	true
true	true	false	true	true
true	true	false	false	true
true	false	true	true	false
true	false	true	false	false
true	false	false	true	true
true	false	false	false	true
false	true	true	true	false
false	true	true	false	false
false	true	false	true	true
false	true	false	false	true
false	false	true	true	false
false	false	true	false	false
false	false	false	true	false
false	false	false	false	true

This represents the same Boolean function as part (a).

iii) Are all the training examples classified correctly (if not, which aren't)?
Again e3 is misclassified.

iv) Give two examples, not in the training set, and specify what the predicted values is.

(c) For the network with a hidden layer what is a local minima of the learning rate (within one decimal point)? The value to minimize is the number of steps before the error gets below 1.0. Hint: there is a local minima in the range [0.3, 7.0].

There is local minimum at 1.7 or 1.8 (with 42 iterations), another at 2.7 (with 33 iterations) and another at 3.0 (with 34 iterations).