CSC592CM – Programming Assignment #3

Due 12/20/06 at 3pm in the CS Office

Version 1.0

1 Neural Networks

Neural networks can learn classifications from data. To train a neural network we need to provide it with examples of the classification task. The following is a simple dataset that assigns each point described by $(X1, X2)$ a label $Y \in \{-1, 1\}$,

$X1$	$X2$	Y
1	6	-1
1	4	1
2	1	1
2.5	1.5	1
3	7	-1
3	1	1
4	1	-1
5	3	-1

The goal is to have the neural network learn a classification from this data so that for example given the point $(1, 6)$ it will return a -1. In addition, once we have learned a classification we can predict what the outcome or labels of points not in the dataset might be. For example, we could predict the label for the point $(5, 5)$?

Perhaps the simplest neural network is the *Rosenblatt Perceptron*:

1
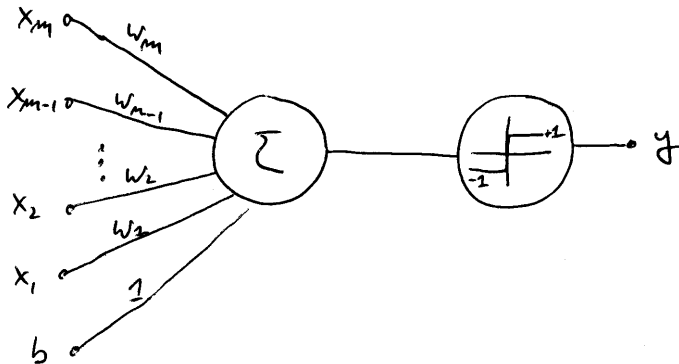
This neural network consists of a single neuron that performs a weighted sum of the network inputs (including a bias term b) and then performs a thresholding step. Notice that there will be as many network inputs as there are independent variables in the example/training dataset plus the bias term.

The training algorithm of this network is given in the following pseudocode,

$\overline{w} \leftarrow \overline{0}$ (*initialize the neural network weights*)
$b \leftarrow 0$ (*initialize the bias*)
$R \leftarrow \max_{1 \leq i \leq n} |\overline{x}_i|$ (*initialize the "radius" of the data*)
$\eta \leftarrow 0 < \eta < 1$ (*initialize the learning rate*)
repeat
    for i in training instances
        if $sign(\overline{w} \bullet \overline{x}_i + b) \neq y_i$ then
            $\overline{w} \leftarrow \overline{w} + \eta y_i \overline{x}_i$
            $b \leftarrow b + \eta y_i R^2$
        end if
    end for
until no mistakes made in the for-loop
return $(\overline{w}, b)$

where $\overline{w} \bullet \overline{x}$ is the dot product between two vectors $\overline{w} = (w_1, w_2, \ldots)$ and $\overline{x} = (x_1, x_2, \ldots)$ defined as $\overline{w} \bullet \overline{x} = w_1 x_1 + w_2 x_2 + \ldots$. The thresholding function $sign$ is defined as

$$sign(x) = \begin{cases} 1 \text{ if } x > 0 \\ -1 \text{ otherwise} \end{cases}$$

Once we have trained the perceptron we can use the weights and the bias term to perform predictions. Given some $\overline{x}$ we can compute the label y for this point using the perceptron,

$$y = sign(\overline{w} \bullet \overline{x} + b)$$

where $\overline{w}$ and b are the weights and the bias, respectively.

Write a Lisp/Scheme program that implements the above perceptron learning algorithm and the prediction function. Train your perceptron on the dataset given above. Validate your perceptron by demonstrating that for each training instance it computes the correct label. Furthermore, predict the label for point $(5, 5)$.

Hand in your source code, the validation runs, and the prediction run.

2 Your Favorite Paper

Write a *one page* essay describing your favorite paper/presentation and why you preferred it to the other papers.