Exam Neural Networks

October 22, 2007

It is a "closed book" exam: you are not allowed to use any notes, books, etc. You may formulate your answers in Dutch or English. For each problem you get some points; additionally you get 10 points for free. The final grade for this exam is the total number of points you get divided by 10.

1) **Single Perceptron (10+5+2+3 points)**
   a. Describe the working of a single perceptron. Describe in detail the two learning algorithms: Perceptron and Pocket. What are the main properties of these algorithms?

   b. Is it correct that a single unit with a logistic sigmoid activation function may compute a non-linear decision boundary? Justify your answer!

   c. Let us consider a training set in $\mathbf{R}^2$ that consists of one "positive" pattern and one "negative" pattern. Is it true that a single perceptron can always learn to correctly classify these two patterns? Why?

   d. Now let us consider a training set in $\mathbf{R}^2$ that consists of three patterns: one "positive" and two "negatives". Is it true that a single perceptron can always learn to correctly classify these patterns? Why?

2) **Thermometer representation of data (5+15 points)**
   a. Describe the concept of a "thermometer representation" of a continuous input variable x that takes values in [0, 1]. What would be the representation of values: 0.0, 0.23, 0.7, and 0.8, assuming a scale with 5 levels?

   b. Let us assume that a continuous variable x that takes values in [0, 1] is represented with help of a 5-level thermometer scale. Then a network with 5 input nodes and a single linear output node computes a function $f(x)$, f:[0,1]->R. Show that $f(x)$ doesn't have to be a linear function of x. How would you call this function?

3) **Encoder networks and image compression (5 + 10 points)**
   a. Describe the concept of the Encoder Network. What is the architecture of this network, what are the inputs and outputs, and what is this network supposed to learn?

   b. Describe in detail (as much as you can) an application of encoder networks to image compression.

4) Backpropagation (5 + 10 points)

a) Let us consider a feed-forward network shown in figure below. Here we assume that the hidden units have no bias and that the activation function is given by `s(net)=cos(net)`. (Recall that `cos'(x)=-sin(x)`).

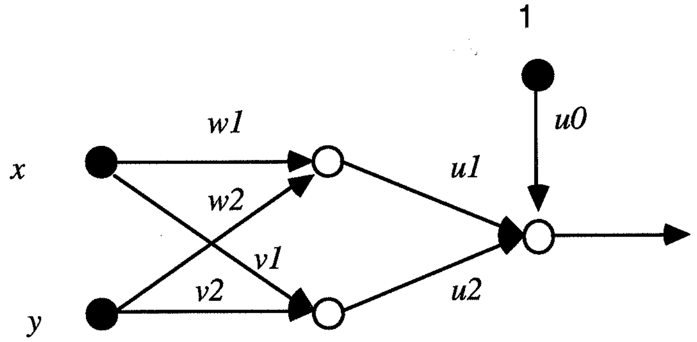

Figure 1. A 2:2:1 feed-forward network with 7 weights.

Express the output of this network as a function of nine variables: x, y, w_1, w_2, v_1, v_2, u_0, u_1, u_2 (write a formula).

b) Derive the update rules for weights u_0, u_1 and w_1. In other words, calculate the partial derivatives of the error function (for a single input vector (x, y) and the target value t) and use them in formulas of the form $u_0^{new} = u_0^{old} + \dots$.

5) Complexity of Backpropagation (2+3+10+5 points)

Let us consider a multi-layer perceptron with 100 inputs, 10 hidden nodes in the first layer and 5 nodes in the output layer.

a) How many weights are used by the network?
b) How many (roughly) operations (additions, multiplications, calling an activation function) should be performed in order to produce an output on a single input vector?
c) How many (roughly) operations are needed for a single step of the backpropagation algorithm (propagating errors backward and updating all weights)?
d) Generalize your findings for a network with K inputs, L nodes in the hidden layer, and M nodes in the output layer.

$Y = 10$ keer $\varphi(K_n \cdot w(n) + \text{bias})$

dan 5 keer $\varphi(L_n \cdot w(n) + \text{bias})$