Exam Neural Networks

January 6, 2009

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+5+5 points)**
   a. Describe the working of a single perceptron. Describe in detail the two learning algorithms: Perceptron and Pocket. What can be said about the convergence of both algorithms?

   b. Define the concept of linear separability. Give an example of 2 sets of points in 3-dimensional space, R^3, A and B, such that A and B are not linearly separable.

   c. There are 16 possible boolean functions of two variables (every function can be defined by a table with 4 rows that represent all possible combinations of 2 logical variables and a column of four of 0's and 1's that determine the corresponding values of the function). How many of these functions are not linearly separable? (For example, we know that at least one of these functions, XOR, is not linearly separable.)

   d. Let us consider a training set that consists of 10 patterns that represent digits 0, 1, .., 9, each pattern being a vector of 256 of 0's and 1's that represents a 16x16 bit map image of a digit. Is it possible to train a single perceptron to separate the digit 7 from the remining nine digits? Extensively justify your answer.

2) **Backpropagation (10+5+15+10 points)**
   a. Describe the back-propagation algorithm (including formulas).

   b. What can be said about the training process when the learning rate parameter is too big? And what happens when the learning parameter is too small?

   c. Briefly describe 3 techniques of speeding up the backpropagation algorithm: Adaptive Gradient Descent, Resilient Backpropagation and Quick Propagation.

d. Let us consider a simple multilayer perceptron with two input units, two hidden units and one output unit. Additionally, let us assume that each unit uses the logistic sigmoid activation function. It is well-known that such a network can be trained to learn the exclusive OR problem, XOR. Is it true that this network can be trained to learn any of the 16 boolean functions of two arguments that are mentioned in question 1.c? Justify your answer.

3) **Self-organizing Maps (15 +10 points)**

a. Describe in detail the architecture, working, and training a Self-Organizing Map (SOM) network.

b. Describe in detail the WEBSOM application.