Neural Networks Homework #1.

Prof: Cheng-Yuan Liou, TA: Jiun-Wei Liou

1   Power of Neural Networks

1. It's well known that 3-2-1 network cannot implement all 2^{2^3} functions while 3-3-1 network can implement all possible functions. Is it possible for 3-2-2-1 network to implement all 2^{2^3} functions? (Hint: what function 3-2-1 network cannot implement?)

2. Show that for N inputs and three layer perceptron structure, N-N-1 network is able to accomplish the N input xor function.

2   Hidden Tree in the Multilayer Perceptron

1. A dataset is represented in Fig. 1(a). The multilayer perceptron used to separate o and x can be shown as Fig. 1(b). The four lines in Fig. 1(a) denote the four percaptrons in the first layer. Please draw the hidden tree for this network to correctly classify the dataset.
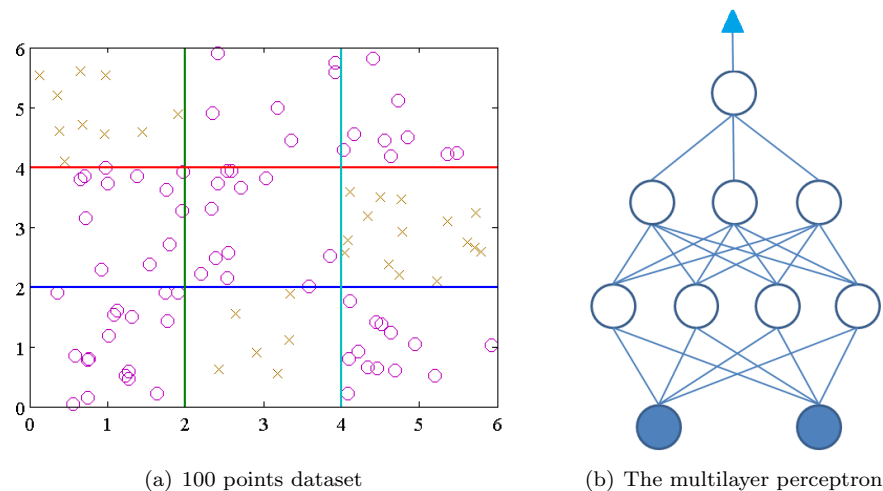

(a) 100 points dataset

(b) The multilayer perceptron

Figure 1: Dataset and multilayer perceptron for classification

2. If an x is put in $(1, 3)$, which parts of this tree contain mixed set of data with o and x?

3   Backpropagation Algorithm

A dataset with 1000 2-dimensional data with labels can be obtained from the course website. The dataset consist 1000 lines of x, y, and the classified value. The NetTalk algorithm can refer to `http://papers.cnl.salk.edu/PDFs/Parallel%20Networks%20That%20Learn%20to%20Pronounce%20English%20Text%201987-3351.pdf`. The smoothing parameter α can be set to 0 and ignore the smoothing effect.

1. Implement a multilayer perceptron according to the update equations mentioned in the NetTalk paper to classify the dataset. The structure used in the multilayer perceptron is 2-4-3-1. The output error curve during training should be drawn.

2. After training, we can construct the hidden tree for the obtained network. Show the hidden tree for all layers except input to the first hidden layer and idnetify if there exists redundent neurons or mixed regions.

3. Redo the previous works for using 2-8-6-1 structure on the same dataset and briefly stats your findings.

Hint: The 2-dimensional dataset are distributed within $x \in [-1, 1]$ and $y \in [-1, 1]$. The initial weight range, the learning rate, and number of iterations should be considered.

Notes: Suggested length of your homework report is no more than 8 pages.