CSC7130 Advanced Artificial intelligence
Homework Assignment for Artificial Neural Networks

Irwin King
Dept. Computer Science and Engineering
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong

Handed out on: Friday, 2009/9/25;
Due on: Tuesday, 2009/10/13 at class

Notes

1. Please submit a printed (hardcopy) version of your homework solution to the person at the front desk to be collected and then courier over to CUHK.
2. You may discuss the homework with your classmates in a group. However, you MUST turn in your own work. Do not copy other's work or you will risk failing the class.

You need only to complete the following 8 questions: 1.2.1, 1.2.2, 1.2.4, 1.3, 1.4, 1.6, 1.7, and 1.12.

1 Homework Problems

Exercise 1.1 *This problem requires you to write a mini-proposal to utilize your expertise of neural networks. I expect the answer should be more than a page of write-up (between two to four pages is fine). Think about your workplace and what you do as a professional. Answer the following:*

1. *Select one of the problem areas that neural networks can solve, e.g., classifications, prediction, function approximation, etc., and show examples of instances in your work place that neural networks can be used to solve problems. Explain the problem, the goal, and the possible solutions.*
2. *Take the particular example and suggest the possible input you would need, the type of neural networks used (learning architecture), and the expected output.*

3. *Explain how you would train the neural network, i.e., what is the input and the format of the input.*

Exercise 1.2 *As mentioned, we say that neural networks are at least as powerful as a universal computing machine, i.e., a computer since neural networks can form the basic building blocks of the computer. Draw out neural networks based on the McCulloch-Pitts model for the following logic gates:*

1. *a 1 input, 1 output NOT gate.*
2. *a 3 inputs, 1 output AND gate.*
3. *a 3 inputs, 1 output OR gate.*
4. *a 3 inputs, 1 output NOR gate.*
5. *a 3 inputs, 1 output NAND gate.*
6. *a 2 inputs, 1 output XOR gate.*

Exercise 1.3 *An example of the sigmoid function is defined by*

$$\varphi(v) = \frac{1}{1+\exp(-av)}$$

whose limiting values are 0 and 1.

1. *Show that the derivative of $\varphi(v)$ with respect to v is given by*

$$\begin{aligned} \varphi'(v) &= \frac{d\varphi}{dv} \quad (1) \\ &= a\varphi(v)[1-\varphi(v)] \quad (2) \end{aligned}$$

2. *What is the value of $\varphi'(v)$ at the origin?*

Exercise 1.4 *A neuron j receives inputs from four other neurons whose activity levels are 5, 10, -2, and -10. The respective synaptic weights of neuron j are -0.8, 0.5, 1.0, and -0.9. Calculate the output of neuron j for the following situations:*

1. *The neuron is linear.*
2. *The neuron is represented by a McCulloch-Pitts model with the activation function as*

$$\varphi(v) = \begin{cases} 1, & \text{If } v > 0 \\ 0, & \text{otherwise} \end{cases}.$$

3. *The neuron is model is based on the sigmoid function*

$$\varphi(v) = \frac{1 - \exp(-av)}{1 + \exp(-av)}.$$

Exercise 1.5 *Consider a multilayer feedforward network, all the neurons of which operate in their linear regions. Justify the satement that such a network is equivalent to a single-layer feedforward network.*

Exercise 1.6 *Given the following training patterns for passing and failing of a class:*

$(90, p), (40, f), (30, f), (60, p), (50, f), (45, f), (65, p), (80, p)$

and a simple neural network based on the error correction algorithm. Answer the following questions.

1. *Given that* $\eta = 2, 1, 0.5,$ *and* 0.1 *with the* $w = 1$ *and* $\theta = 100$, *set up a neural network to handle the problem and show what happen when you applied the different* η *values. You may even use a spreadsheet to perform this task if you wish.*
2. *What are some of ways that we can set the initial starting conditions?*
3. *Does the initial starting condition matter? Justify your answer.*
4. *Does it matter what order the training patters are being presented, i.e., sequentially from the beginning to the end or randomly selected? Justify your answer.*

Exercise 1.7 *The delta rule*

$$\Delta w_{kj}(n) = \eta e_k(n) x_j(n)$$

and Hebb's rule

$$\Delta w_{kj}(n) = \eta y_k(n) x_j(n)$$

represent two different methods of learning. List the features that distinguish these two rules from each other.

Exercise 1.8 *Consider the following sets of key patters, applied to a correlation matrix memory:*

$$\mathbf{a}_1 = [1.8, -0.55, 2.5, -10]^T \quad (3)$$
$$\mathbf{a}_2 = [0, 1, 0, 0.1]^T \quad (4)$$
$$\mathbf{a}_3 = [-0.2, 0, 0.9, 0]^T \quad (5)$$

The respective stored patterns are

$$\mathbf{b}_1 = [5, 1, 0]^T \quad (6)$$
$$\mathbf{b}_2 = [-2, 1, 6]^T \quad (7)$$
$$\mathbf{b}_3 = [-2, 4, 3]^T \quad (8)$$

1. *Calculate the memory matrix* $\mathbf{M}$.
2. *Calculate the memory response when* $\mathbf{a}_1$ *is presented.*
3. *Show that the response* $\mathbf{b}$ *is closest to the stored pattern* $\mathbf{b}_1$ *in a Euclidean sense.*

Exercise 1.9 *An autoassociative memory is trained on the following key vectors:*

$$\mathbf{a}_1 = \frac{1}{4}[-2, -3, \sqrt{3}]^T \quad (9)$$

$$\mathbf{a}_2 = \frac{1}{4}[2, -2, -\sqrt{8}]^T \quad (10)$$

$$\mathbf{a}_3 = \frac{1}{4}[3, -1, \sqrt{6}]^T \quad (11)$$

1. *Calculate the angles between these vectors. How close are they to orthogonality with respect to each other?*
2. *Using the generalization of Hebb's rule (i.e., the outer product rule), calculate the memory matrix of the network. Hence, investigate how close to perfect the memory autoassociates.*
3. *A masked version of the key vector* $\mathbf{a}_1$*, namely,*

$$\mathbf{a} = [0, -3, \sqrt{3}]^T$$

is applied to the memory. Calculate the response of the memory, and compare your result with the desired response $\mathbf{a}_1$*.*

Exercise 1.10 *This question concerns about the winner-take-all neural network architecture.*

1. *Design a winner-take-all neural network that uses lateral inhibition to derive the output.*
2. *What are some of the interesting aspects can you observe that is different from other types of neural network architecture?*
3. *Can you say something about the convergence property of your neural network?*

Exercise 1.11 *Can you design a neural network that can output the second winner instead of the first winner? If so, show how. If not, explain why.*

Exercise 1.12 *In the multilayer backpropagation network the momentum constant* α *is normally assigned a positive value in the range* $0 \le \alpha < 1$*. Investigate the difference that would be made in the behavior of the equation*

$$\Delta w_{ji}(n) = -\eta \sum_{t=0}^{n} \alpha^{n-t} \frac{\partial E(t)}{\partial w_{ji}(t)}$$

with respect to time t *if* α *was assigned a negative value in the range* $-1 < \alpha \le 0$.