3BA2
Tutorial
Neural Networks

1. In this question we consider the use of a feedforward neural network in the design of an E-mail filtering system. The objective of the system is to rank incoming messages as Important or Normal depending on the attributes of some messages.

   An analysis of the problem has shown that three key predictive features of the importance of a message are:

   Domain: The origin of the message; can be Internal, Ireland or the rest if the World.

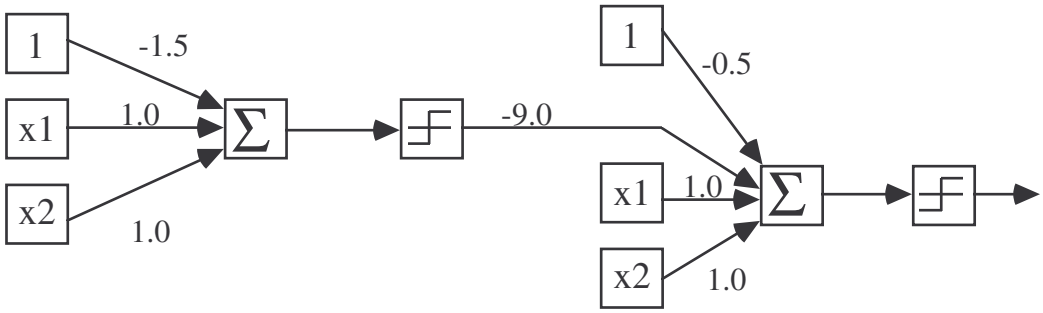
   Size: Can be Small, Medium or Large.

   Type: Whether the message was addressed to a Mailing List or an Individual.

   In the Table below there are 9 examples of email messages described according to these features. Describe the architecture of a feedforward neural network that could learn the relationships described in this table. Explain in outline how this network would be trained and how it would operate in practice.

Message	Domain	Size	Type	Rank
No.1	Internal	Medium	Personal	**Important**
No.2	Internal	Medium	Mailing List	**Normal**
No.3	Internal	Small	Personal	**Important**
No.4	World	Small	Personal	**Important**
No.5	World	Large	Mailing List	**Normal**
No.6	IE	Small	Mailing List	**Normal**
No.7	IE	Small	Personal	**Important**
No.8	World	Large	Mailing List	**Normal**
No.9	IE	Large	Personal	**Normal**

2. (i) Even though the XOR Problem is linearly unseparable, the following arrangement of two Perceptrons is able to classify the XOR inputs correctly. Show that this is true. Assume that the transfer function for each neuron is a step function that outputs 1 for positive and 0 for negative input.



   The -9.0 weight linking the output of the first neuron to the input of the second is much larger than is necessary. What is the minimum value that this can have and still produce the correct results?

   (12 marks)

   (ii) Explain the difference between supervised and unsupervised learning. An ART network is an unsupervised learning mechanism that requires a vigilance parameter; why is the vigilance parameter required.

   (8 marks)

3. (a) The following table illustrates a three-parity relationship, i.e. the **p** bit is set so as to give the 4 bits (**x1,x2,x3,p)** odd parity. Would you expect a single perceptron to be able to learn this relationship between **p** and the three inputs **x1,x2,x3**. If not why not?

x1	x2	x3	p
0	0	0	1
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	0

(8 marks)

(b) Discrimination networks are a popular retrieval mechanism in case-based reasoning. Explain using examples how discrimination networks work.

(12 marks)

4. (a) A manual personal loan approval system is based on a form that records important pieces of information. This information is used in determining whether a loan should be approved or not. The decision is recorded as Accept or Reject. A typical example of the information recorded on the loan application form is as follows:

Amount required:		(Numeric)
Period of Repayment in Months		(Integer)
Loan Purpose		(Discrete Classes)
	e.g.	Home Improvement
		Car Purchase
		Personal
		House Loan
Life Assurance?		(y/n)
Number of Dep. Children		(Integer)
Age		(Integer)
Marital Status		(Discrete Classes)
Residential Status		(Discrete Classes)
Occupation		(Discrete Classes) ;general categories
Have Cheque Card		(y/n)
Have Current a/c		(y/n)
Decision		**(Accept/Refer)**

What type of Neural Network (NN) architecture would be appropriate for this application? Describe in outline how an NN solution to this problem would be implemented.

5. In a Hopfield Network, to store a set of binary patterns:

$\mathbf{s}(p) = (s_1(p),...,s_i(p),...s_n(p))$, the weight matrix $\mathbf{W} = \{w_{ij}\}$ is given by:

$$w_{ij} = \sum_p [2s_i(p) - 1][2s_j(p) - 1] \quad \text{for} \quad i \neq j$$

and $w_{ii} = 0$.

What would be the weight matrix after storing the following patterns?