Module Code: COMP5400M01

Module Title: Bio-Inspired Computing

© **UNIVERSITY OF LEEDS**

School of Computing

Semester 2 2018/2019

Calculator instructions:

- You **are** allowed to use a non-programmable calculator only from the following list of approved models in this examination: **Casio FX-82**, **Casio FX-83**, **Casio FX-85**.

Dictionary instructions:

- You are **not** allowed to use your own dictionary in this examination. A basic English dictionary is available to use: raise your hand and ask an invigilator, if you need it.

Examination Information

- There are **5** pages to this examination.
- There are **2 hours** to complete the examination.
- Answer **all 3** questions.
- The number in brackets **[ ]** indicates the marks available for each question or part question.
- You are reminded of the need for clear presentation in your answers.
- The total number of marks for this examination paper is **60**.
- You are allowed to use annotated materials.

Question 1

(a) Consider the data set of Figure 1. The straight line represents the decision line of a perceptron with two inputs and a bias. Give these weights and bias such that all white squares are classified as one (ignore the black squares for now.)

[4 marks]

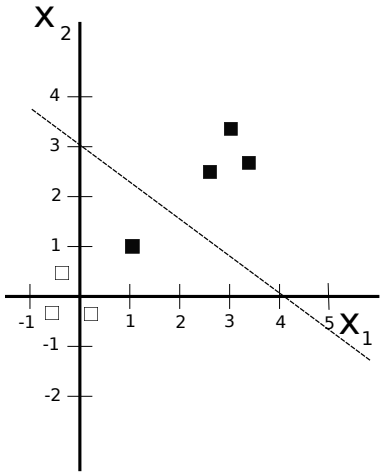

Figure 1: Data set with a linear decision boundary. The decision boundary goes through (4,0) and (0,3).

(b) Now consider the black markers as well. One point, (1, 1), is being misclassified. Carry out one step of the perceptron algorithm and produce new weights and bias. Show your workings

[4 marks]

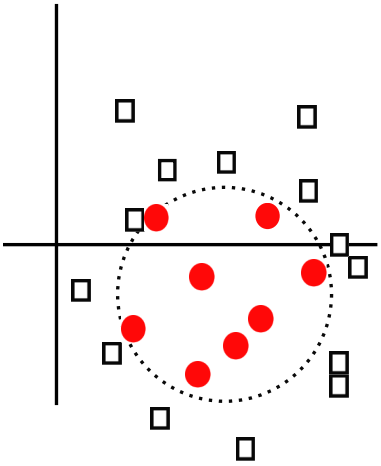

Figure 2: A data set whose classes are approximately divided by a circular decision line.

(c) Consider a data set that can be separated into two classes, and that the decision boundary can be approximated by a circle (Figure 2). Explain why a perceptron of the form $o = f((x - w_1)^2 + (y - w_2)^2 - \theta)$ would be useful as a classifier.

[3 marks]

(d) Derive a learning rule for the perceptron of the previous question, based on steepest gradient descent, for finding values of w_1, w_2 and θ. Assume that the sigmoid function is given by $f(x) = \frac{1}{1+e^{-x}}$.

[6 marks]

(e) Explain why feedforward networks built from artificial neurons with an activation function that is evaluated on the neuron's input need a hidden layer where the neurons have non linear activation functions to be able to solve problems that are not linearly classifiable.

[3 marks]

(f) The iris data set can be learnt by a two-layer network - containing just an input layer and an output layer - if a softmax activation function is used on the output layer. As the data set is not linearly classifiable, this seems a paradox: feedforward networks capable of learning such data sets need at least one hidden layer. Explain why the iris data set is not linearly classifiable.

[3 marks]

(g) Resolve the apparent contradiction that a two-layer network with a softmax activation function can be successfully trained on the iris data set.

[3 marks]

[Question 1 Total: 26 marks]

Question 2

(a) Explain the Ant Colony routing algorithm, and describe the role of stygmergy in this algorithm. It is not necessary to replicate formulae, a qualitative explanation is sufficient.

[4 marks]

(b) You are given a data classification problem. For simplicity, suppose you have decided on a feedforward network architecture, that is: you have determined size and connectivity of all layers in the network, but still need to find appropriate weight values. Describe how you would use a genetic algorithm to solve the data classification problem. You must describe the role of all elements of a standard genetic algorithm in your application.

[6 marks]

(c) Consider a planet far away. Your task is to populate this planet with machines of your design, for example, to survey the planet remotely. Due to load considerations only one machine can be launched. The machine is to land on the planet, mine its surface and to construct copies of itself from the raw materials. You may assume that the machine is equipped with a program and tools that are sufficiently sophisticated to create all the tools and parts that are necessary from the raw material. By what name is this type of machine known?

[1 mark]

(d) There appears to be a paradox in the previous question. The program has to contain, in some way or another, a full description of the machine: not just of the construction unit for a new machine, but also of the program itself. Living creatures face the same problem. Describe a historical approach to the problem which uses 'infinite regress'. Define the term 'infinite regress' and explain why this is a problem in this particular solution.

[3 marks]

(e) Explain a solution to the apparent paradox. Discuss the dual role that the program plays.

[4 marks]

[Question 2 Total: 18 marks]

Question 3

(a) Consider a six node Hopfield network. Give all components of the weight matrix for a network that has learnt the pattern $(1, -1, -1, 1, -1, 1)$.

[3 marks]

(b) Show that a Hopfield network that has been trained using a single pattern will converge to that pattern, or its negative.

[3 marks]

(c) If more than one pattern is stored in a Hopfield network the patterns will interfere. Show this in a simple example. Present a statistical argument that makes plausible that as long as the number of patterns p stored in the network is much smaller than the number of nodes N, the network is still likely to converge on one of the training patterns.

[6 marks]

(d) In a Restricted Boltzmann Machine, full connectivity is replaced by a simpler architecture. Describe this architecture, and give the rationale for making this restriction.

[4 marks]

[Question 3 Total: 16 marks]

[Grand Total: 60 marks]

End