

**Part II**

**COMPUTING AND COMMUNICATIONS – On-line Assessment**

**Available Time** [2.5 Hours]

**Recommended Completion Time** [30 Minutes]

**SCC.361**

---

*Candidates are asked to answer **THREE** questions from **FOUR**; each question is worth a total of 25 marks.*

## Question 1

**1.a** Name one example of artificial intelligence in medicine. Suggest one business and one societal benefit, and discuss one risk of incorporating this example of AI.

[4 marks]

**1.b** You aim to automatically distinguish two types of plant according to the lengths of a randomly-chosen leaf. You introduce a cut-off threshold such that any leaf with a length smaller than 2cm is from a plant of type A and any leaf with a length greater than or equal to 2cm is from a plant of type B.

- i) Give the definition of machine learning according to Mitchell and explain why this is not a machine learning approach, referring to Mitchell's definition. [2 marks]
- ii) Describe the difference between machine learning and traditional programming. [1 mark]
- iii) You are now asked to design a machine learning method for distinguishing these plant types using the leaf measures collected. Describe the relevant experience, task and performance measure of this system. [3 marks]

[6 marks]

**1.c** Consider a convolutional neural network (CNN) consisting of five convolution layers followed by two fully connected (MLP) layers and a softmax layer. The network takes an image as input and generates three outputs. All the weights and biases are set to zero.

- i) Let us assume all activation functions in this network are Sigmoid. If you feed the below image to the network, what is the output? Explain your answer in 1-2 sentences. [3 marks]



- ii) Let us assume all activation functions in this network are ReLU. If you feed the above image to the network, what is the output? Explain your answer in 1-2 sentences. [3 marks]

[6 marks]

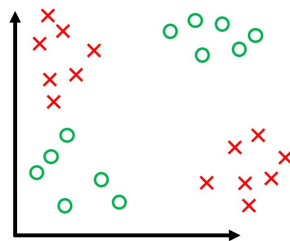
**1.d** Suppose you are given two trained models for a 3-class classification problem as shown below. Each model uses a Softmax layer to generate its output as  $[P_1, P_2, P_3]$ , where  $P_1$ ,  $P_2$  and  $P_3$  denote probability of sample belonging to class A, class B and class C, respectively.

| Sample Number | Actual Class | Model 1 Predicted Output | Model 2 Predicted Output |
|---------------|--------------|--------------------------|--------------------------|
| 1             | C            | [0.1, 0.5, 0.4]          | [0.5, 0.3, 0.2]          |
| 2             | B            | [0.2, 0.6, 0.2]          | [0.4, 0.5, 0.1]          |
| 3             | A            | [0.6, 0.3, 0.1]          | [0.6, 0.2, 0.2]          |
| 4             | A            | [0.1, 0.8, 0.1]          | [0.5, 0.3, 0.2]          |
| 5             | B            | [0.2, 0.1, 0.7]          | [0.3, 0.1, 0.6]          |
| 6             | C            | [0.5, 0.4, 0.1]          | [0.1, 0.5, 0.4]          |

- Compute the Maximum Likelihood of both models. Which model is better? You DO need to show your calculations. [4 marks]
- Compute the Cross-Entropy of both models. Which model is better? You DO need to show your calculations. [4 marks]

[8 marks]

**1.e** Consider the dataset below consisting of samples from two classes (cross and circle). Each sample in this dataset has two features. Is this dataset linearly separable? Explain your answer in one sentence.



[1 mark]

**Total 25 marks**

## Question 2

**2.a** One approach to defining AI is to consider it as “thinking like a human” and an example of this would be cognitive modelling. List the other three approaches to defining AI and give an example of each type.

[3 marks]

**2.b** You are asked to write a method to detect fraudulent transaction on customer accounts. You are given a small list of recent transactions on which to base your method, including (i) distance from the customer’s home to transaction location (dist), (ii) ratio of transaction amount to median transaction amount of the customer (amount-ratio), (iii) whether the transaction was fraudulent or not (fraud). These transactions are given in the below table.

| Record Number | dist | amount-ratio | fraud |
|---------------|------|--------------|-------|
| 1             | 5.8  | 1.5          | Yes   |
| 2             | 2.3  | 1.1          | No    |
| 3             | 6.0  | 0.9          | No    |
| 4             | 3.1  | 0.9          | Yes   |
| 5             | 4.0  | 1.2          | No    |
| 6             | 7.1  | 1.2          | Yes   |

- i) Showing your working, use  $k$ -nearest neighbours with  $k = 1$  and Manhattan distance as the distance measure to determine whether the two following transactions are fraudulent. [5 marks]

| Record Number | dist | amount-ratio |
|---------------|------|--------------|
| N1            | 7.5  | 1.4          |
| N2            | 2.0  | 1.1          |

- ii) Describe how you would extend this to  $k > 1$  and determine whether records N1 and N2 represent fraudulent transactions for  $k = 3$ . [2 marks]
- iii) Do you consider this to be a machine learning solution? Explain your answer. [2 marks]

[9 marks]

**2.c** Consider a system of nonlinear equations shown below:

$$\begin{aligned}f_1(x) &= x_1 - 0.26378153 - 0.1874589x_4x_3x_9 \\f_2(x) &= x_2 - 0.37784532 - 0.16232323x_1x_{10}x_6 \\&\dots \\f_{10}(x) &= x_{10} - 0.42989891 - 0.211234554x_4x_8x_1\end{aligned}$$

The objective is:

$$\min(f_1(x), f_2(x), \dots, f_{10}(x))$$

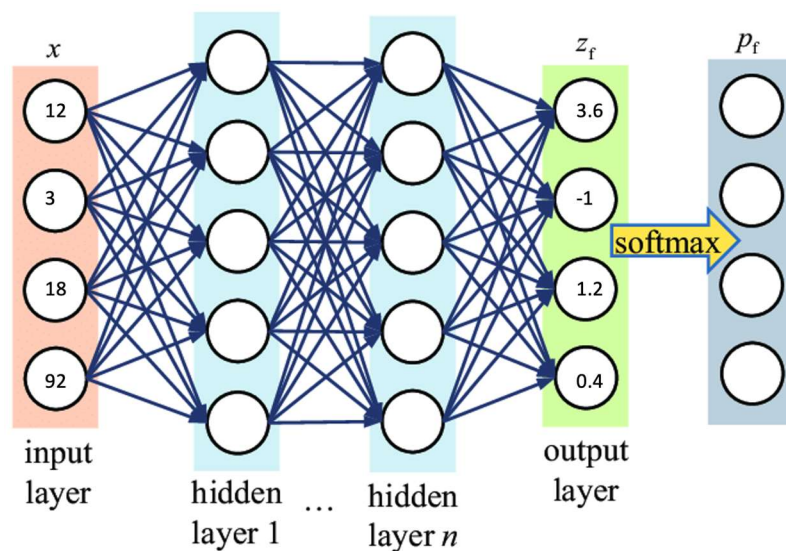
- i) Suppose you are asked to write a Genetic algorithm to solve this problem. What chromosome encoding method is most appropriate for this problem? Explain your answer in one sentence. [2 marks]

- ii) Show an example chromosome representing a candidate solution for this problem. [2 marks]
- iii) Consider a population with 5 chromosomes and a fitness function  $F$  with values as shown in the table below. Our aim is to minimise  $F$ . Suppose RWS algorithm is used to select a parent chromosome. The generated random number,  $r$ , where  $0 < r \leq 1$ , is  $r = 0.585$ . Which chromosome is selected? You DO need to show your calculations. [5 marks]

| Chromosome | $c_1$ | $c_2$ | $c_3$ | $c_4$ | $c_5$ |
|------------|-------|-------|-------|-------|-------|
| F          | 25    | 61    | 13    | -4    | 2     |

[9 marks]

**2.d** Consider a 4-class multi-layer perceptron (MLP) model shown below. Calculate the output of the softmax layer. You DO need to show your calculations.



[4 marks]

**Total 25 marks**

### Question 3

**3.a** We often define AI in terms of “agents.”

- Give the definitions of an agent and a computational agent. [2 marks]
- Give one example of a computational agent and one example of a non-computational agent. Justify your choices. [3 marks]
- Give the definition of AI in terms of agents. [1 marks]

[6 marks]

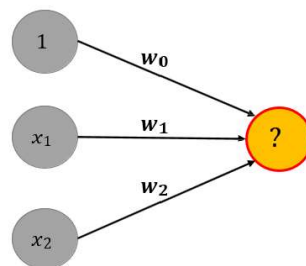
**3.b** Name and explain three reasons that classic machine learning algorithms, such as  $k$ -nearest neighbour might struggle to learn from data.

[3 marks]

**3.c** Describe the Support Vector Machine (SVM) approach to dividing points and explain how this can lead to a classification result. Name two different kernel tricks, describe why kernel tricks might be needed and how they can improve classification results.

[5 marks]

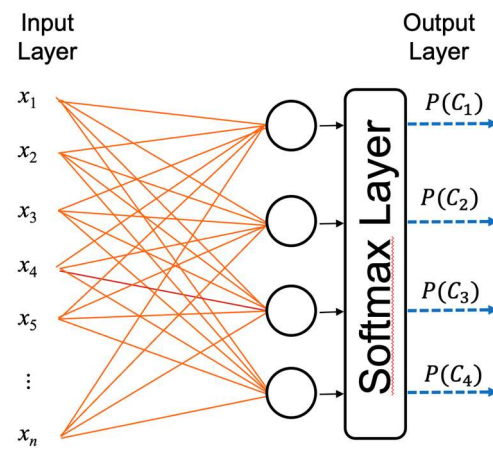
**3.d** You aim to build a simple perceptron system, shown below, to determine whether or not it will rain given two variables  $x_1$  and  $x_2$ .



- Write the corresponding hypothesis function, given the determined weights  $w_0 = 1.5$ ,  $w_1 = 7$  and  $w_2 = -3$ . Evaluate this equation for  $(x_1, x_2) = (2.4, 1.3)$  and write the answer. [2 marks]
- Name the type of function that should be used to transform the result of the hypothesis function into a classification result and name two examples of such functions. [3 marks]

[5 marks]

**3.e** Consider a dataset consisting of 2000 RGB images. Each image is  $10 \times 10$  pixels. The images are divided into 4 classes. The aim is to design a neural network to perform image classification. Suppose the input image is flattened and fed to the below neural network. What is the number of parameters of this network? You DO need to show and explain your calculations.



[6 marks]

Total 25 marks

#### Question 4

**4.a** Name the key scientific and engineering goals in artificial intelligence and describe the difference between them.

[4 marks]

**4.b** You are asked to help to split a company's customers into two groups ( $\Gamma_1$  and  $\Gamma_2$ ) to create a more effective marketing strategy. You have information of the customers' age and income, given in the table below. To achieve this, you will use  $k$ -means clustering with maximum distance as the distance measure, given by

$$|a - b|_\infty = \max_i |a_i - b_i|$$

where  $|x|$  denotes the absolute value of  $x$ .

| Customer Number  | 1  | 2  | 3  | 4  | 5  | 6  |
|------------------|----|----|----|----|----|----|
| Age (years)      | 36 | 37 | 48 | 25 | 32 | 59 |
| Income (x£1,000) | 41 | 22 | 42 | 21 | 14 | 28 |

- You are given the initial centre (0,0) for the age and income of group  $\Gamma_1$  and (50,50) for the age and income of group  $\Gamma_2$ . Perform one iteration of  $k$ -means clustering to determine group membership. You must show your working and present the membership of the two groups using the customer numbers. [5 marks]
- Perform a second iteration of  $k$ -means clustering, showing your working and presenting the membership of the two groups. [5 marks]
- State whether you need to carry out any more iterations and justify your answer. [2 marks]

[12 marks]

**4.c** For this question, you are going to answer a couple of sub-questions regarding the dataset shown in the table below. You will be trying to determine whether a file is infected by a virus based on its three features including writable, updated and size.

|         |   | WRITABLE | UPDATED | SIZE  | CLASS    |
|---------|---|----------|---------|-------|----------|
| DATASET | 1 | yes      | no      | small | infected |
|         | 2 | yes      | yes     | large | infected |
|         | 3 | no       | yes     | med   | infected |
|         | 4 | no       | no      | med   | clean    |
|         | 5 | yes      | no      | large | clean    |
|         | 6 | no       | no      | large | clean    |

- Calculate the entropy of this dataset. You DO need to show your calculations. [2 marks]
- Use the ID3 algorithm (as discussed in the lecture) and draw the tree that would be learned by the algorithm. You DO need to show your calculation for each step of the algorithm. [5 marks]



iii) Is the tree in (ii) built by ID3 algorithm optimal? More precisely, does the tree in (ii) get zero training error with minimal depth? Explain your answer in 2-3 lines. [2 marks]

**[9 marks]**

**Total 25 marks**