

2022 EXAMINATIONS



Part II

COMPUTING AND COMMUNICATIONS – On-line Assessment [150 Minutes]

SCC.361 Artificial Intelligence

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

Question 1

1.a Give the definition of an “agent” and describe what it means for an agent to act intelligently. Write how you can define AI in terms of agents.

[4 marks]

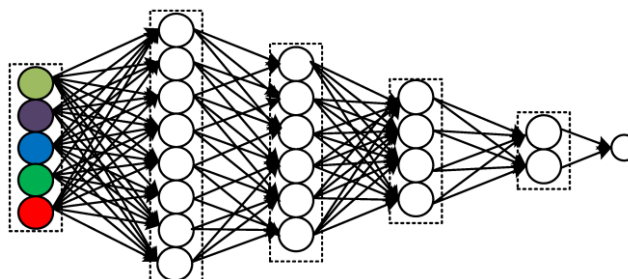
1.b You are asked to write an algorithm to determine whether a person is a child or adult, based on their height. You are given the following table and decide to use k -nearest neighbour algorithm with Euclidean distance. You then want to determine whether a person of height 166.2cm is a child or an adult using your algorithm.

Identifier	1	2	3	3	5	6
Adult/ Child	Adult	Child	Child	Adult	Adult	Child
Height (cm)	170.7	155.6	165.3	175.2	147.3	150.2

- Showing your working, use your KNN algorithm with $k = 1$ to determine if the person of height 166.2cm is an adult or a child. [4 marks]
- Describe the extra step that you would need to take if $k > 1$. [1 mark]
- Use your algorithm to determine whether the same person is an adult or child for $k = 3$ and $k = 5$. [2 marks]
- You now want to use your KNN algorithm to determine the person’s age. What would be the limitation of using your KNN algorithm and what type of approach would be likely to work better? [2 marks]

[9 marks]

1.c You are given the below multi-layer perceptron (MLP) model, which consists of 4 fully connected hidden layers. The network takes an input $x \in \mathcal{R}^5$ and generates a single output. All the weights and biases are initialized to zero.



[Please turn over]

- i) Let us assume all activation functions in this network are Sigmoid. If you feed $x = [1, 2, 3, 4, 5]$ to the network, what is the output? Explain the reason in 2-3 sentences. [2 marks]
- ii) Let us assume all activation functions in this network are ReLU. If you feed $x = [5, 4, 3, 2, 1]$ to the network, what is the output? Explain the reason in 2-3 sentences. [2 marks]

[4 marks]

1.d Suppose you are given two trained models for a binary classification problem (TRUE/FALSE classification) as shown below. The class of interest is the TRUE class.

Sample Number	Actual Class	Model 1 Predicted Output	Model 2 Predicted Output
1	TRUE	0.6	0.8
2	FALSE	0.1	0.2
3	FALSE	0.7	0.7
4	TRUE	0.4	0.3

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

[8 marks]

Total 25 marks

[Please turn over]

Question 2

2.a Using real-world examples, list and discuss two risks and/ or challenges of integrating AI into our daily life.

[4 marks]

2.b I want to write a method to enable my security camera to only start recording when it detects a human and not to record when it detects a cat. I have many labelled examples of cat and person images.

- i) Is this a potential machine learning problem? If so, why? [2 marks]
- ii) If so, what type of machine learning problem is it? Name an approach that you could use to address this problem. [2 marks]
- iii) Give the definition of a computer learning in terms of a task (T), experience (E) and performance measure (M) [1 mark]
- iv) Define the task (T), experience (E) and performance measure (M) for this scenario. [3 marks]

[8 marks]

2.c Two students are given a dataset consisting of 1000 gray-scale images. Each image is 20×20 pixels. The images are divided into 4 different categories/classes including car, cat, apple and mobile. The aim is to design a neural network to perform image classification. The **first** student designs the following network: the input image is first flattened into a 400-dimensional vector and then passed through a fully connected layer with 4 neurons. Then, a softmax layer is added to the end of the network. The **second** student designs the following network: the input image is directly passed through a convolutional layer with four 20×20 filters and a softmax layer is added to the end of the network.

Explain which network you would choose and why.

[5 marks]

2.d Calculate the total number of trainable parameters for a fully connected neural network consisting of 3 hidden layers, each with 10 neurons. The size of the input is 30-dimensional. The network performs binary classification.

[5 marks]

2.e Explain the similarities and differences between the dropout technique and random decision forest.

[3 marks]

[Please turn over]

Total 25 marks

Question 3

3.a Write what is meant by "declarative knowledge" and "imperative knowledge" and give a limitation of each of these.

[4 marks]

3.b You are given the following set χ and asked to perform k -means clustering to split the data into 2 clusters, i.e. $k = 2$. The clusters will be labelled χ_1 and χ_2 and you decide to use Manhattan distance.

$$\chi = \{34, 74, 43, 89, 24\}$$

- Set the initialisation of the centre of χ_1 equal to 0 and the initialisation of the centre of χ_2 equal to 100. Perform one iteration of k -means clustering, showing your working. Give the new centre values for the two sets and list which sets each value should belong to. [5 marks]
- Do you need to carry out any more iterations? Justify your answer. [2 marks]
- How would this change if you decide to use Euclidean distance instead? [1 mark]

[8 marks]

3.c Suppose you are asked to write a Genetic algorithm to find an optimal schedule for your coursework submission. Let's assume you have 10 different courseworks; each has a start date, a finish date and number of days required to do the coursework. For example, as shown below, since the start date of the coursework#1 is 10 May 2022, you need to wait until 10 May 2022 to be able to access to the coursework assignment. And since the finish date of the coursework#1 is 20 June 2022, you should submit this coursework by 20 June 2022, otherwise you get a zero mark. The main aim is to minimize the number of missed courseworks.

	Start date	Finish date	Time required
Coursework#1	10 May 2022	20 June 2022	5 days
Coursework#2	15 May 2022	25 May 2022	5 days
Coursework#3	13 May 2022	30 May 2022	4 days
Coursework#4	28 May 2022	15 June 2022	4 days
Coursework#5	20 May 2022	5 June 2022	8 days
Coursework#6	1 June 2022	12 June 2022	4 days
Coursework#7	17 May 2022	16 June 2022	7 days
Coursework#8	19 May 2022	19 June 2022	4 days
Coursework#9	23 May 2022	7 June 2022	5 days
Coursework#10	26 May 2022	17 June 2022	3 days

[Please turn over]

- i) What chromosome encoding is good for this problem? Explain why. [2 marks]
- ii) Show an example chromosome for this problem. [2 marks]
- iii) Design a fitness function which works for this problem. Explain your function in detail. You could use equations to explain your fitness function. [3 marks]
- iv) Calculate the fitness value of the example above in (ii) and show your calculation. [2 marks]

[9 marks]

3.d Suppose we implement a Perceptron to correctly classify all samples in the below dataset (the accuracy of the trained Perceptron is 100%). Each sample has two features (x_1 and x_2) and Y is actual class label. The activation function of the Perceptron is a threshold-based activation function and is defined as below:

$$\varphi(v) = \begin{cases} 0, & v < 0 \\ 1, & v \geq 0 \end{cases}$$

	x_1	x_2	Y
Sample#1	0	0	0
Sample#2	0	1	0
Sample#3	1	0	0
Sample#4	1	1	1

Which of the following combinations would be the weights (W_a and W_b) and bias (b) of the trained Perceptron? Explain your answer in one or two sentences and show your calculations.

- A) $W_1 = 1, W_2 = 1, b = -1.5$
- B) $W_1 = 2, W_2 = 2, b = 1.5$
- C) $W_1 = 1.5, W_2 = 1.5, b = 1$
- D) None of the above

[4 marks]

Total 25 marks

[Please turn over]

Question 4

4.a I have a list of sentences and, in the traditional programming sense, I have a program which gives the class of each word in the sentence (verb, noun etc). The output of my program is a vector of word classes for each sentence. I want to reformulate this as a machine learning problem.

- i) What will be the task of the machine learning problem? [1 mark]
- ii) What inputs will I need? [2 marks]
- iii) What will be the output? [1 mark]

[4 marks]

4.b Feature extraction is a fundamental component in machine learning and artificial intelligence.

- i) State two reasons that Feature Extraction is important in machine learning. [2 marks]
- ii) Feature extraction can be performed manually or automatically. Given one pro and one con of each approach. [4 marks]
- iii) Which of the following matrices (a or b) could be used as an edge detection filter? Justify your answer. [2 marks]

$$a = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

[8 marks]

4.c You are given the sentence "the cat sat on the mat".

- i) Write down the vocabulary for this sentence and, using 1-hot encoding, write down the feature matrix for this sentence. [3 marks]
- ii) How would you construct the Bag of Words from this example? [1 mark]

[4 marks]

4.d Take the dataset below that consists of 4 training examples with three binary features (A, B and C) and a Binary output (Y), which is the class label.

	A	B	C	Y
Sample#1	F	F	T	F
Sample#2	T	T	T	T
Sample#3	F	F	F	T
Sample#4	F	T	F	F

[Please turn over]

- i) Calculate the entropy of this dataset. [1 mark]
- ii) 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. [4 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. If it is not optimal, draw the optimal tree as well. [4 marks]

[9 marks]

Total 25 marks

--- End of paper ---