2021 EXAMINATIONS

Lancaster University

Part II

COMPUTING AND COMMUNICATIONS – On-line Assessment

Available Time **[23 Hours]**

Recommended Completion Time **[3 Hours]**

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 Approaches to AI are often categorized into four different schools of thought referring to AI as systems that 'think humanly', 'think rationally', 'act humanly' or 'act rationally'

i. Explain with your understanding, the key idea behind each schools of thought (2 sentences max for each)

[8 marks]

ii. In your opinion, which of these schools of thought best describes the AI of today and why?

[2 marks]

1.b Poole & Markworth, 2017 (Artificial Intelligence: Foundations of Computational Agents) identified **two** types of goals in AI namely: *scientific* and *engineering*. With your understanding, briefly discuss what each of these goals aims to achieve.

[4 marks]

1.c AI affects us in a variety of ways, using real world examples, briefly discuss the following:

i. The social benefits of AI

[3 marks]

ii. Ethical concerns of AI

[3 marks]

1.d Given this definition of machine learning, *"A computer is said to learn from experience **E** with respect to task **T** and some performance measure **P**, if its performance on **T**, as measured by **P**, improved with experience **E**"*.

If my program takes images of handwritten letters (A-Z) as well as the correct letter identification numbers, thereby increasing its ability to correctly identify which letter is in each image. In your opinion,

i. is this a machine-learning problem? If not, why?

[1 mark]

ii. if it is, what type of machine learning problem is it?

[1 mark]

iii. what is the **task** (T), the **experience** (E) and **performance measure** (P) in this scenario? Explain

[3 marks]

[Total 25 marks]

Question 2

2.a A machine learning process involves learning from data.

i. Explain, using real world example problems, how supervised and unsupervised methods work

[4 marks]

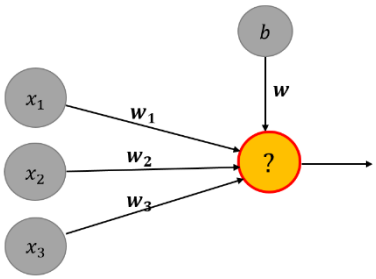
ii. Which of the hierarchical and k-means clustering methods is non-deterministic and why?

[3 marks]

iii. In what ways can the value of ***'k'*** be determined in **'k-means'** clustering?

[3 marks]

2.b Consider this diagram of a simple artificial neural network



i. If the input node values are $x_1 = 2,\ x_2 = 3,\ x_3 = 1$ and the weights are $w_1 = 0.5, w_2 = 0.2, w_3 = 0.08$, what is the value of the **hypothesis function** if the bias value b and its weight w are 1 and 0.25 respectively?

[3 marks]

ii. Explain the essence of the **activation function** using **two** common examples.

[3 marks]

iii. Discuss the role of **gradient descent** in training neural networks

[2 marks]

2.c Looking at the simple neural networks (or **perceptrons**)

i. Explain how they differ from **multi-layer neural networks**

[3 marks]

ii. What is the essence of the **dropout** technique and how does it work?

[4 marks]

[Total 25 marks]

Question 3

3.a Assume five cities {A, B, C, D, E}, which are represented by a fully connected graph. The following table represents the distances between each city (assume the distances are symmetric).

	Distances				
	A	B	C	D	E
A	0				
B	5	0			
C	8	2	0		
D	9	3	7	0	
E	1	4	1	9	0

i. Given below is an initial population consisting of four chromosomes for solving Traveling Salesman Problem (TSP). Calculate the fitness of each chromosome in the initial population

[2 marks]

C_1	ABCDE	C_2	BACED
C_3	ACDBE	C_4	ABCED

ii. In order to select two parents, Roulette Wheel Selection (RWS) algorithm is used. Let's assume that two randomly generated numbers are 0.6 and 0.9. Fill in the following table and specify which two chromosomes are selected as parents.

[3 marks]

Chromosome	C_1	C_2	C_3	C_4
Fitness				
Probability				
Cumulative Probability				

3.b Explain the Naïve Bayes assumption that simplifies the Likelihood expression $P(A_1, A_2, \dots, A_n|C)$ of the Bayes' Rule. Explain your answer in two sentences.

[2 marks]

3.c Take the dataset below that consists of 14 training examples with four attributes (Outlook, Temperature, Humidity, and Wind) and a Binary outcome (Play Tennis) that is class label. Given the below test sample, predict its class label using Naïve Bayes classifier. You DO need to show your calculations.

Question 3 continues on next page…

Test sample: X = (Outlook=*Rain,* Temperature=*Cool,* Humidity=*High,* Wind=*Week*)

Outlook	Temperature	Humidity	Wind	Play Tennis
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Overcast	Hot	High	Weak	Yes
Rain	Mild	High	Weak	Yes
Rain	Cool	Normal	Weak	Yes
Rain	Cool	Normal	Strong	No
Overcast	Cool	Normal	Strong	Yes
Sunny	Mild	High	Weak	No
Sunny	Cool	Normal	Weak	Yes
Rain	Mild	Normal	Weak	Yes
Sunny	Mild	Normal	Strong	Yes
Overcast	Mild	High	Strong	Yes
Overcast	Hot	Normal	Weak	Yes
Rain	Mild	High	Strong	No

[8 marks]

3.d In a medical application domain, suppose we build a classifier for patient screening (True means patient has cancer). Suppose that the confusion matrix is as below. Which of the following situations would you like your classifier to have? Explain your answer in one sentence.

		Predicted	
		True	False
Actual	True	*TP*	*FN*
	False	*FP*	*TN*

A. FP >> FN
B. FN >> FP
C. FN = FP × TP
D. TN >> FP
E. FN × TP >> FP × TN
F. All of the above

[2 marks]

Question 3 continues on next page…

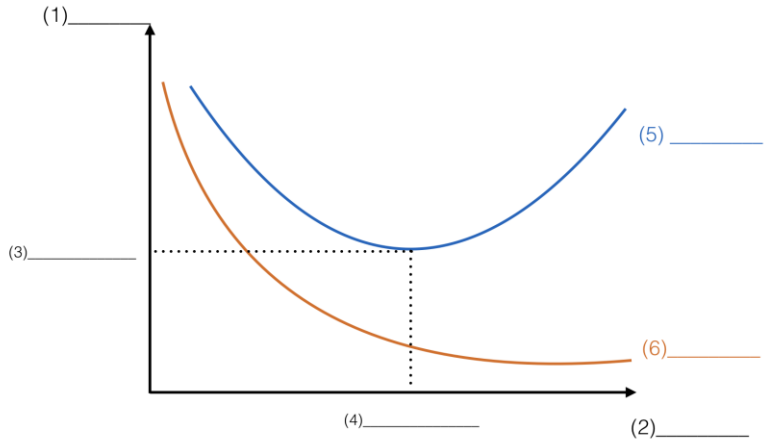
Question 3 continued.

3.e Suppose that there are 50 artificial intelligence related documents in a library of 200 documents. Suppose that a search engine retrieves 10 documents after a user enters 'artificial intelligence' as a query, of which 5 are artificial intelligence related documents. Fill in the following table and calculate the precision and recall values?

TP	FP	TN	FN

[4 marks]

3.f You want to use the figure below to explain the concepts of early stopping to a friend. Fill in the blanks. (1) and (2) describe the axes. (3) and (4) describe the values on the vertical and horizontal axis. (5) and (6) describe the curves.



[4 marks]

[Total 25 marks]

Question 4

4.a Take the dataset below that consists of 14 training examples with four attributes (Outlook, Temperature, Humidity, and Wind) and a Binary outcome (Play Tennis) that is class label.

Outlook	Temperature	Humidity	Wind	Play Tennis
Sunny	Hot	High	Weak	No
Sunny	Hot	High	Strong	No
Overcast	Hot	High	Weak	Yes
Rain	Mild	High	Weak	Yes
Rain	Cool	Normal	Weak	Yes
Rain	Cool	Normal	Strong	No
Overcast	Cool	Normal	Strong	Yes
Sunny	Mild	High	Weak	No
Sunny	Cool	Normal	Weak	Yes
Rain	Mild	Normal	Weak	Yes
Sunny	Mild	Normal	Strong	Yes
Overcast	Mild	High	Strong	Yes
Overcast	Hot	Normal	Weak	Yes
Rain	Mild	High	Strong	No

i. What is the entropy of this dataset?

[1 mark]

ii. If you build an optimal decision tree according to the Information Gain (IG) score, what should be the Information Gain (IG) scores for "Temperature" and "Humidity" attributes. Fill in the below table. You DO need to show your calculations.

Attribute	"Outlook"	"Temperature"	"Humidity"	"Wind"
Information Gain	0.246	?	?	0.0478

[8 marks]

iii. Which attribute would be chosen as the root node of the tree? Why?

[1 mark]

Question 4 continues on next page…

Question 4 continued.

4.b Consider the following confusion matrix of a classification system with three possible classes on 600 samples.

		Predicted Class		
		Class="C1"	Class="C2"	Class="C3"
Actual Class	Class="C1"	150	40	10
	Class="C2"	20	160	20
	Class="C3"	20	40	140

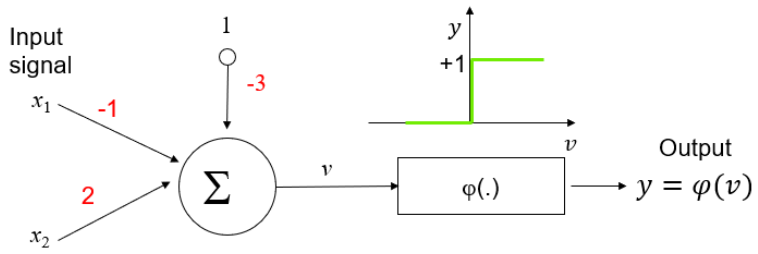
i. Calculate accuracy of this classification system.

[1 mark]

ii. Calculate precision, recall, and F1-Score for class "C1".

[3 marks]

4.c Consider the following perceptron (with weights given on the connections).



i. Fill in the table below by computing the output of this perceptron on the following training data.

[1.5 marks]

x_1	x_2	target	y
1	1	1	
-2	1	0	
0	1	1	

ii. What is the accuracy of this perceptron on this training set?

[0.5 mark]

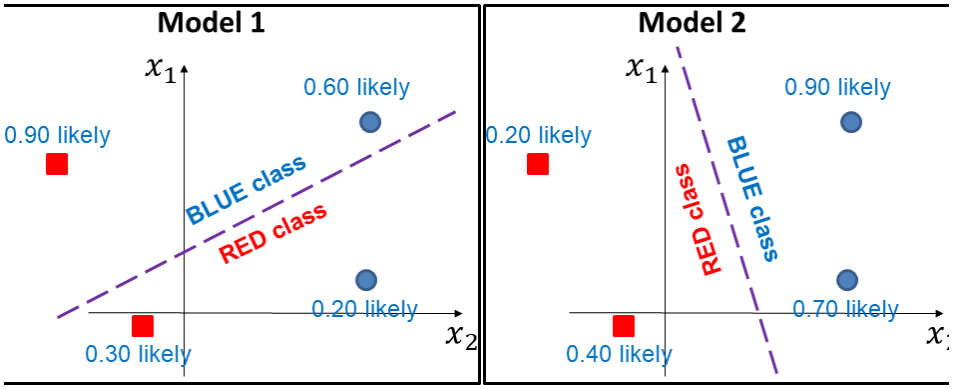
iii. Write the equation of the separation line defined by this perceptron.

[3 marks]

Question 4 continues on next page…

Question 4 continued.

4.d Suppose you are given two learned models for a binary classification problem (BLUE/RED classification) as shown below. The class of interest is the BLUE class (i.e. the probability next to each sample is the probability of a sample be BLUE). BLUE samples are shown as circles and RED samples are shown as squares.



i. Compute the Maximum Likelihood of both models. Which model is better? <u>You DO need to show your calculations</u>.

[3 marks]

ii. Compute the Cross-Entropy of both models. Which model is better? <u>You DO need to show your calculations.</u>

[3 marks]

[Total 25 marks]

--- End of Paper ---