UNIVERSITY OF KWAZULU-NATAL
School of Engineering
Discipline of Electrical, Electronic and Computer Engineering
(Howard College Campus)

Examinations: **September 2025**

ENEL4AI H2 : Artificial Intelligence

DURATION :2 HOURS MARKS : 100

Examiners:

Prof. Jules R. Tapamo

Prof. Pieter Van Vuuren

Internal Moderator:

Prof. Tom M. Walingo

Instructions:

1. Answer ALL 5 questions.
2. Start each question on a new page, and number all your answers clearly.
3. Scientific calculators may be used.
4. Ensure you have 6 pages, including this one.
5. **This paper may not be removed from the exam hall.**

Question 1: *Artificial Intelligence and Machine Learning Concepts* **[25 marks]**

(1.1) For each of the following questions choose the best of the available answers.

(a) Which of the following best describes Artificial Intelligence (AI)? [1]

(A) A system that performs only repetitive tasks.
(B) A machine mimicking human intelligence and decision-making.
(C) Any software that processes data.
(D) Only robots that can act like humans.
(E) A database management system.

(b) Which of the following is a limitation of current AI systems? [1]

(A) Ability to learn from big data.
(B) Transfer learning.
(C) Lack of general common sense reasoning.
(D) High accuracy in image classification.
(E) Automation of repetitive tasks.

(c) Which activation function is most commonly used in hidden layers of deep networks due to efficiency and reduced vanishing gradient problems? [1]

(A) Sigmoid
(B) Tanh
(C) ReLU (Rectified Linear Unit)
(D) Softmax
(E) Linear function

(d) Assume a simple MLP model with 4 neurons and inputs= 5,4,3,2,1. The weights to the input neurons are 2, 1,4,5, and 9 respectively. Assume the activation function is a linear, with a constant value of 2. What will be the output ? [2]

(A) 76
(B) 96
(C) 90
(D) 256
(E) 32

(e) Which of the following would have a constant input in each epoch of training a Deep Learning model? [1]

(A) Weight between input and hidden layer
(B) Weight between hidden and output layer
(C) Activation function of output layer
(D) Biases of all hidden layer neurons
(E) None of the above

(f) What is an overfit model? [1]

(A) A model that learns the training dataset too well, performing well on the training dataset but does not perform well on a testing dataset

(B) A model that fails to sufficiently learn the problem and performs poorly on a training dataset and does not perform well on a testing dataset

(C) All of the above

(D) None of the above

(g) Which optimization algorithm is widely used in deep learning for faster convergence compared to standard gradient descent? [1]

(A) AdaBoost

(B) Random Forest

(C) Adam

(D) Naive Bayes

(E) PCA

(h) A convolutional filter of size $f \times f$ applied to an input image of size $n \times n$, with a stride of 1, produces an output image of size: [2]

(A) $n \times n$

(B) $(n - f + 1) \times (n - f + 1)$

(C) $f \times f$

(D) $n \times f$

(E) $\frac{n}{f}$

(i) Why are convolutional layers preferred over fully connected layers for image data? [1]

(A) They use fewer parameters and capture spatial structure

(B) They completely ignore spatial locality

(C) They always run faster on any hardware

(D) They prevent overfitting automatically

(E) They remove the need for backpropagation

(1.2) Answers the following question by **True** or **False**.

(a) Support Vector Machines aims to maximize the margin between classes. [0.5]

(b) With Multilayer Perceptron, Backpropagation computes gradients of the loss with respect to weights. [0.5]

(c) K-Means may converge to a local minimum, not necessarily the global one. [0.5]

(d) Genetic algorithms always guarantee finding the global optimum solution. [0.5]

(1.3) Explain the architecture of a Vision Transformer. You must particularly discuss the following: how images are divided into patches and embedded, the role of positional encoding, and how the Transformer encoder processes the embedded patches. [12]

Question 2: *Search and Control Strategies* **[15 marks]**

(2.1) Considering the TSP-problem for n cities and n-puzzle problems. What are the sizes of their respective search spaces? [2]

(2.2) Given the alphabet $\mathbf{A} = \{0, 1, 2\}$, build a Finite State Machine that recognizes strings that contain the substring **1022**. [6]

(2.3) Suppose you are searching a code to open a door. The code is 4 letters long and is made of the 26 letters of the alphabet a, b,c,. . .,z. You have the letters ordered in alphabetic order and you start writing down possibilities:

$$a, b, c, d, \ldots, y, z, aa, ab, ac, \ldots, zz, \ldots$$

(a) How many strings of 4 or fewer letters can we form with with the letters of the alphabet? [4]

(b) In the above possibilities, in which order are you searching ? [1]

(c) How many possible codes can we create to open the door? [2]

Question 3: *Logic, Knowledge Representation and Reasoning* **[22 marks]**

(3.1) Prove using truth table that

$$[(p \vee q) \wedge (\neg q \vee r)] \Rightarrow (p \vee r)$$

is valid? [2]

(3.2) Translate the following sentence into first order predicate logic:
"**Not all engineering registered students take some humanities electives**" [2]

(c) Given the predicates in Figure 1, translate the following sentences in first-order predicate logic.

(i) Every medical doctor is a problem solver [1.5]

(ii) Every problem solver is either intelligent or lazy or gullible. [1.5]

(iii) No medical doctor is gullible. [1.5]

(iv) No problem solver who is competent is lazy. [1.5]

(3.3) Transform the sentences from part (c) into conjunctive normal form. [4]

(3.4) Based on the knowledge base above (statements c-(i) to c-(iv)) and using resolution, prove formally that "**If no problem solver is intelligent, then no medical doctor is competent**". [8]

p1. $MD(x)$: x is Medical Doctor
p2. $PS(x)$: x is Problem Solver
p3. $I(x)$: x is intelligent
p4. $G(x)$: x is Gullible
p5. $L(x)$: x is Lazy
p6. $C(x)$: x is Competent

Figure 1: First order logic predicates

Question 4: *Uncertainty, Fuzzy Logic and Systems* **[18 marks]**

(4.1) Table 1 gives the joint probability distribution of traffic slowdown (S), accident (A), and construction (C).

Table 1: Events (Slowdown, Accident, Construction) Likelihood

S	A	C	p(S,A,C)
T	T	T	0.12
T	T	F	0.24
T	F	T	0.06
T	F	F	0.18
F	T	T	0.14
F	T	F	0.12
F	F	T	0.11
F	F	F	0.03

Compute the probability of the following events:

(a) There is slowdown, accident and construction. [1]

(b) There is an accident [1]

(c) There is slowdown given that there is accident and there is construction. [4]

(4.2) Let the temperature T be a fuzzy variable that can belong to the fuzzy set **hot** and to the fuzzy set **moderately hot**. The concept of **hot** is represented by a fuzzy set **H** whose membership function is :

$$\mu_H(T) = \begin{cases} 0, & \text{if } T < 20^o \\ \frac{T-20}{10}, & \text{if } 20^o \leq T \leq 30^o \\ 1, & \text{if } T > 30^o \end{cases} \quad \text{(eq1)}$$

Let the concept "**moderately hot**" be represented by a fuzzy set **M** whose membership function is :

$$\mu_M(T) = \begin{cases} \frac{T-20}{10}, & \text{if} \quad 20^o \leq T \leq 30^o \\ \frac{35-T}{5}, & \text{if} \quad 30^o < T \leq 35^o \\ 0, & \text{elsewhere} \end{cases} \qquad \text{(eq2)}$$

Let x be the number of customers in the shop. The concepts "**Crowdedness**" and **Quietness** of an ice-cream shop be respectively represented by a fuzzy sets **C** and **Q** whose membership functions are:

$$\mu_C(x) = \begin{cases} 0, & \text{if} \quad x < 25 \\ \frac{x-25}{25} & \text{if} \quad 25 \leq x \leq 50 \\ 1 & \text{if} \quad x > 50 \end{cases} \qquad \text{(eq3)}$$

$$\mu_Q(x) = \begin{cases} 1, & \text{if} \quad x < 10 \\ \frac{50-x}{40}, & \text{if} \quad 10 \leq x \leq 50 \\ 0, & \text{if} \quad x > 50 \end{cases} \qquad \text{(eq4)}$$

(b.1) Draw the graphics for the membership functions of **temperature** and **number of customers**. [3]

(b.2) State the set operations and determine the equations for the membership functions representing the concept "**the temperature is hot OR temperature is moderately hot**" and **temperature is not hot**. [4]

(b.3) Consider the following fuzzy expert system:

RULE 1: **IF** the temperature is hot OR the temperature is moderately hot,
THEN the ice-cream shop is crowded.

RULE 2: **IF** temperature is not hot,
THEN the ice-cream shop is quiet.

Suppose the temperature is 16^o . Determine the expected number of clients in the ice-cream shop. [5]

Question 5:*Genetic algorithms and Naive Bayesian Classifier* **[20 marks]**

(5.1) We want to apply a genetic algorithm for optimization of a simple function of one real variable. The function is defined as

$$f(x) = 2x^3 - 3x + 1$$

and we want to find x from the range $[-1, 2]$ which maximizes the function f.

(a) We will use a binary string as a chromosome to represent real values of the variable x. The required precision for the variable x is two places after the decimal point. How many bits nb are required in a chromosome, as a binary vector of size nb? Show your working. [2]

(b) Give a linear mapping, to map the representation of a chromosome from a binary string $(b_{nb-1}b_{nb-2}\ldots b_0)$ into a real number x from the range $[-1, 2]$. [2]

(c) Given the following chromosome:$v_1 = (10110101)$, Compute the real values x_1 corresponding to the chromosome v_1, and Compute $f(x_1) = eval(v_1)$ [4]

(5.2) A Miner has collected stones and identified them as being either Precious (P) or Normal (N). You are charged with building a Naive Bayesian classifier to predict the Type (P or N) of a given stone as it is found using four features: Roundness (R), color (C), number of holes (H) and surface texture (T). Table 2 gives the responses for collected stones.

Table 2: Responses for collected stones

Instance	**R**	**C**	**H**	**T**	**CL**
x_1	Roundish	grey	3	smooth	P
x_2	Rectangular	grey	3	rough	P
x_3	Roundish	grey	3	rough	P
x_4	Rectangular	grey	3	rough	P
x_5	Rectangular	yellow	2	smooth	P
x_6	Rectangular	grey	2	smooth	N
x_7	Roundish	yellow	2	rough	N
x_8	Rectangular	yellow	3	rough	N
x_9	Roundish	yellow	3	rough	N
x_{10}	Rectangular	yellow	3	rough	N

(a) Calculate the following probabilities:
$P(CL = k)$, for $k \in \{P, N\}$. [2]

(b) Using Naive Bayes classifier, what is the predicted class for a stone that is *Rough*, *Yellow*, *3* and *Rough*. [10]