No. of Printed Pages : **8**

A9.5-R5.1 : ARTIFICIAL INTELLIGENCE CONCEPTS AND R PROGRAMMING

DURATION : 03 Hours **MAXIMUM MARKS : 100**

OMR Sheet No. :						

Roll No. : | | | | | | |

Answer Sheet No. : | | | | | | |

Name of Candidate : _____ **; Signature of Candidate :** _____

INSTRUCTIONS FOR CANDIDATES :

- Carefully read the instructions given on Question Paper, OMR Sheet and Answer Sheet.
- Question Paper is in English language. Candidate has to answer in English language only.
- There are **TWO PARTS** in this Module/Paper. **PART ONE** contains **FOUR** questions and **PART TWO** contains **FIVE** questions.
- **PART ONE** is Objective type and carries **40** Marks. **PART TWO** is Subjective type and carries **60** Marks.
- **PART ONE** is to be answered in the **OMR ANSWER SHEET** only, supplied with the question paper, as per the instructions contained therein. **PART ONE** is **NOT** to be answered in the answer book for **PART TWO**.
- Maximum time allotted for **PART ONE** is **ONE HOUR**. Answer book for **PART TWO** will be supplied at the table when the Answer Sheet for **PART ONE** is returned. However, Candidates who complete **PART ONE** earlier than one hour, can collect the answer book for **PART TWO** immediately after handing over the Answer Sheet for **PART ONE** to the Invigilator.
- **Candidate cannot leave the examination hall/room without signing on the attendance sheet and handing over his/her Answer Sheet to the invigilator. Failing in doing so, will amount to disqualification of Candidate in this Module/Paper.**
- After receiving the instruction to open the booklet and before answering the questions, the candidate should ensure that the Question Booklet is complete in all respects.

DO NOT OPEN THE QUESTION BOOKLET UNTIL YOU ARE TOLD TO DO SO.

PART - ONE

(Answer all the questions; each question carries ONE mark)

1. Each question below gives a multiple choice of answers. Choose the most appropriate one and enter in the "OMR" answer sheet supplied with the question paper, following the instructions therein. (1x10)

1.1 Which of the following is a key characteristic of intelligent agents in AI ?

(A) Inflexibility

(B) Autonomous decision-making

(C) Manual operation

(D) Centralized control

1.2 What is the output of length(c (1, 3, 5, 7)) in R ?

(A) 3

(B) 5

(C) 4

(D) Error

1.3 Which of the following algorithms is used for grouping unlabelled data into clusters ?

(A) K-Nearest Neighbor

(B) Linear Regression

(C) K-Means

(D) Decision Trees

1.4 OLAP stands for :

(A) Online Analytical Processing

(B) Online Analysis Processing

(C) Online Advanced Processing

(D) Online Analytical Process

1.5 What is the result of "**rep(c(1, 2), times = 3)**" in R ?

(A) c(1, 2, 3, 1, 2, 3, 1, 2, 3)

(B) c(1, 2, 1, 2, 1, 2)

(C) c(1, 1, 1, 2, 2, 2)

(D) c(2, 1, 2, 1, 2, 1)

1.6 In business analytics, which of the following processes is used to extract hidden patterns from large datasets?

(A) Data mining

(B) Data warehousing

(C) OLTP

(D) Data Cleaning

1.7 Which of the following is a performance evaluation metric for classification algorithms?

(A) Root Mean Squared Error (RMSE)

(B) Confusion Matrix

(C) R-squared

(D) Scatter Plot

SPACE FOR ROUGH WORK

1.8 In R programming, which function is used to read CSV files ?

(A) load.csv()

(B) start.csv()

(C) open.csv()

(D) read.csv()

1.9 What is the purpose of the **which()** function in R ?

(A) It determines the mode of a vector

(B) It extracts elements of a list

(C) It returns the index of TRUE values in a logical vector

(D) It calculates the mean of numeric values

1.10 What is the output of the expression "c(1, 2, 3) + c(4, 5)" in R?

(A) Error

(B) c(5, 7, 3)

(C) c(5, 7, 7)

(D) c(1, 2, 3, 4, 5)

2. Each statement below is either TRUE or FALSE. Choose the most appropriate one and enter your choice in the "OMR" answer sheet supplied with the question paper, following the instructions therein. (1x10)

2.1 An intelligent agent always has a fixed program that cannot learn or adapt.

2.2 AI is not required for business data analytics because traditional methods can handle the data effectively.

2.3 Smart cars use AI primarily for entertainment and interior customization, with minimal use in navigation and safety.

2.4 AI-driven automation is transforming industries such as manufacturing and logistics.

2.5 Linear regression can only be applied to tasks where the output is a continuous value.

2.6 In classification tasks, the output is always continuous.

2.7 In R, a list can store different types of data, such as numbers, strings, and vectors, all in the same object.

2.8 A matrix in R can contain both numeric and character data types simultaneously.

2.9 The normal distribution is a discrete probability distribution.

2.10 The variance of a binomial distribution for which n = 30 and p = 0.30 is 7.0.

3. Match words and phrases in column X with the closest related meaning/word(s)/phrase(s) in column Y. Enter your selection in the "OMR" answer sheet supplied with the question paper, following the instructions therein. (1x10)

	X		Y
3.1	In unsupervised learning, the training dataset consists of	**A**	k-nearest neighbors (KNN)
3.2	Unsupervised learning algorithm used for outlier detection	**B**	Scatterplot
3.3	Algorithm used for clustering data points based on their density	**C**	Binomial Distribution
3.4	A distribution that models the number of successes in a fixed number of trials	**D**	Histogram
3.5	A symmetric, bell-shaped distribution where most data points fall near the mean	**E**	DBSCAN
3.6	A plot that shows the relationship between two continuous variables	**F**	PAC
3.7	A plot used to represent the frequency distribution of a dataset	**G**	Input features
3.8	A classification algorithm based on the proximity of data points	**H**	K-Means
3.9	A table showing true positives, false positives, false negatives, and true negatives	**I**	Isolation Forest
3.10	AI-powered route planning, flight recommendations, and traffic predictions	**J**	Travel and Navigation
		K	Normal Distribution
		L	K-Means
		M	Confusion Matrix

4. Each statement below has a blank space to fit one of the word(s) or phrase(s) in the list below. Enter your choice in the "OMR" answer sheet supplied with the question paper, following the instructions therein. (1x10)

A.	Process	**B.**	Significant	**C.**	Data Frame
D.	Environment	**E.**	OLAP, OLTP	**F.**	Clustering
G.	Data mining, Machine learning	**H.**	Correlation	**I.**	Virtual Assistant
J.	par()	**K.**	hist()	**L.**	barplot()
M.	Automobile				

4.1 Artificial intelligence is predicted to play a _____ role in the future.

4.2 Today, AI is used by businesses and governments to automate _____ and make decisions.

4.3 _____ systems are used to optimize decision-making by processing large volumes of historical data, whereas _____ systems are designed for day-to-day transaction processing.

4.4 AI-based systems can provide better insights into data by applying techniques such as _____ and _____.

4.5 In AI, an agent interacts with its _____, which provides the inputs the agent uses to make decisions.

4.6 AI-powered _____ systems, like Siri and Alexa, have become common in smart homes.

4.7 The task of grouping similar data points together is known as _____.

4.8 A _____ in R is a type of data structure used for organizing data in rows and columns, similar to a table.

4.9 To display multiple plots on the same graphic in R, the _____ function can be used.

4.10 _____ is a statistical method used to measure the strength of the relationship between two variables.

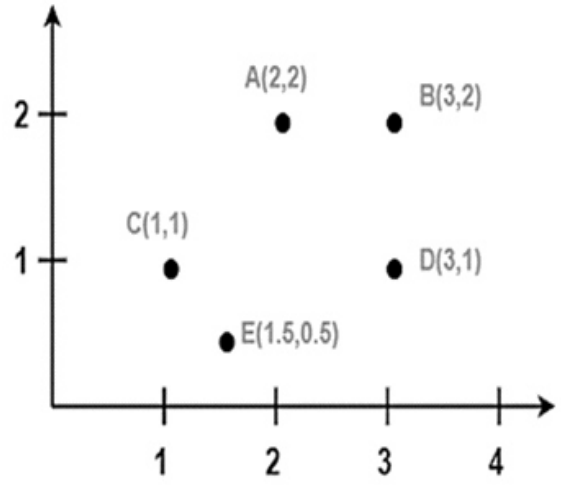
PART - TWO

(Answer any FOUR questions)

5. (a) What are the fundamental elements necessary for designing an AI-based system ? Examine the significance of each element.

(b) How can artificial intelligence enhance security measures in businesses and public areas? Discuss the ethical concerns related to the use of AI in surveillance. **(7+8)**

6. (a) Use k-Means Algorithm to create two clusters.



(b) A data frame in R is two-dimensional, in which each column can contain values of different type.

(i) Describe the steps involved in creating a data frame named StudentData that initially contains three columns: StudentID, Grade, and Score.

(ii) Explain, with reference to appropriate R functions or operators, how each of the following fundamental data-frame operations can be performed on StudentData :

- Adding a new column **Gender**.
- Appending additional rows that supply values for all existing columns.
- Removing the column **Score** from the data frame.
- Deleting the second row of the data frame.

(7+8)

7. (a) Write about the different functions for probability distributions in R that can be used for statistical data analysis.

(b) (i) Explain how the Poisson distribution can be simulated in R. Describe the parameters involved and discuss how a dataset of 1000 values with a mean rate (lambda) of 3 can be generated using the rpois() function.

SPACE FOR ROUGH WORK

(ii) Describe the process of generating a dataset that follows a Binomial distribution using R. What parameters must be defined, and how can 1000 values be generated assuming 10 trials and a probability of success of 0.5 using the rbinom() function ?

(7+8)

8. (a) Implement a Linear Regression model for GDP vs Sales in the below data. The data represents four-wheeler sales with respective GDP year-wise. Also, predict the four-wheeler sales in the year 2024 if the GDP is 7.6. Write R code for implementing Linear Regression and Prediction.

Years	Sales	GDP
2023	30.58	7.94
2022	30.47	7.93
2021	27.9	7.18
2020	26.01	6.54
2019	25.03	5.48
2018	26.65	6.5
2017	26.3	6.2
2016	24.4	5.9

(b) Explain the concept of supervised learning and differentiate it from unsupervised learning.

(8+7)

9. Cluster the following eight points (with (x, y) representing locations) into three clusters A1(2, 10); A2(2, 5); A3(8, 4); A4(5, 8); A5(7, 5); A6(6, 4); A7(1, 2) and A8(4, 9). Initial cluster centers are : A1(2, 10); A4(5, 8) and A7(1, 2).

The distance function between two points a=(x1, y1) and b=(x2, y2) is defined as :

$\rho(a, b) = |x2 - x1| + |y2 - y1|$.

Use k-means algorithm to find the three cluster centers after the second iteration.

15

- o O o -

SPACE FOR ROUGH WORK