No. of Printed Pages : **8**

A9.5-R5 - 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 instructions therein.

(1x10)

1.1 The performance of an agent can be improved by :

(A) Learning

(B) Observing

(C) Perceiving

(D) None of the mentioned

1.2 Select the most appropriate situation for that a blind search can be used.

(A) Real-life situation

(B) Complex game

(C) Small Search Space

(D) All of the above

1.3 Among the given options, which search algorithm requires less memory ?

(A) Depth First Search

(B) Optimal Search

(C) Breadth-First Search

(D) Linear Search

1.4 Which of the following is the branch of Artificial Intelligence ?

(A) Machine Learning

(B) Cyber forensics

(C) Full-Stack Developer

(D) Network Design

1.5 Which term is used for describing the judgmental or commonsense part of problem solving ?

(A) Heuristic

(B) Critical

(C) Value based

(D) Analytical

1.6 Decision Tree algorithm is a type of :

(A) Supervised learning

(B) Unsupervised learning

(C) Active learning

(D) Reinforcement learning

1.7 The output of training process in machine learning is _____.

(A) machine learning algorithm

(B) machine learning model

(C) null

(D) accuracy

1.8 Numbers in R are generally treated as _____ precision real numbers.

(A) single

(B) double

(C) real

(D) imaginary

1.9 If you explicitly want an integer, you need to specify the _____ suffix.

(A) D

(B) R

(C) L

(D) K

1.10 Attributes of an object (if any) can be accessed using the _____ function.

(A) objects()

(B) attrib()

(C) attributes()

(D) obj()

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 instructions therein.

(1x10)

2.1 In supervised learning, given training explain examples of Input and corresponding output, the machine can predict outputs for new inputs.

2.2 Linear regression is a machine learning algorithm based on supervised learning.

2.3 Database Management System is an application of artificial intelligence.

2.4 Artificial Intelligence "Agent" does mapping of precept sequence to an action.

2.5 OLAP stands for Online Available Processing.

2.6 Unsupervised learning is the type of learning in which labeled training data is used.

2.7 Correlation is a statistical measure which determines co-relationship or association of two variable.

2.8 Poisson distribution is a discrete probability distribution.

2.9 library () function is used to watch for all available packages in library.

2.10 Full Form of GUI is Guided User Interface

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 instructions therein.

(1x10)

X		Y	
3.1	Confusion Matrix	**A**	Unsupervised learning
3.2	Robots	**B**	ggplot2
3.3	Classification	**C**	Sensor
3.4	Processing large data sets to identify patterns and relationships	**D**	accuracy
3.5	Clustering	**E**	Data mining
3.6	Data Visualization in R	**F**	Data frame
3.7	It tends to convert a physical attribute to an electrical signal.	**G**	Supervised learning
3.8	it changes an electrical signal to physical action.	**H**	Vectors
3.9	data displayed in a format as a table.	**I**	Reinforcement learning
3.10	a sequence of elements which share the same data type	**J**	Actuator
		K	Application of AI
		L	Data warehousing
		M	Poisson distribution

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

A	regression	**B**	Correlation	**C**	Machine Learning
D	R	**E**	data warehouse	**F**	histogram
G	Supervised learning	**H**	Variance	**I**	Pie chart
J	Arrays	**K**	Root Mean Square Error	**L**	Structured data
M	Unstructured data				

4.1 _____ is the expected value of the squared variation of a random variable from its mean value.

4.2 _____ is a statistical measure that expresses the extent to which two variables are linearly related.

4.3 A _____ is a graph that shows the frequency of numerical data using rectangles.

4.4 A _____ is a statistical technique that relates a dependent variable to one or more independent (explanatory) variables.

4.5 _____ is data that has been organized into a formatted repository, typically a database.

4.6 _____ is a language and environment for statistical computing and graphics.

4.7 _____ is the field of study that gives computers the capability to learn without being explicitly programmed.

4.8 _____ is defined by its use of labeled datasets to train algorithms that to classify data or predict outcomes accurately.

4.9 _____ is a metric used for evaluating the performance of regression algorithm.

4.10 A _____ is a central repository of information that can be analyzed to make more informed decisions.

PART TWO

(Answer any FOUR Questions)

5. (a) Describe the building blocks of artificial intelligence (AI) in detail.

(b) Describe the structure of Agents.

(c) Differentiate between OLAP and OLTP.

(5+5+5)

6. (a) What is classification and regression in machine learning ? Mention different types of regression. Can regression be used for classification ?

(b) Explain the ways/techniques for cleaning the raw data.

(7+8)

7. (a) What are applications of artificial intelligence (AI) ?

(b) What are the metrics to analyze machine learning algorithms ?

(7+8)

8. (a) Explain the neural networks. Explain feed forward and back propagation in detail.

(b) How many data structures R programming uses? Write a R program to convert a given matrix to a 1 dimensional array.

(8+7)

9. Briefly explain the following (**Any three**) :

(a) Confusion Matrix

(b) NAN and INF in R programming

(c) Data mart and Datawarehouse

(d) Different types of agent-based systems.

(5+5+5)

- o O o -

SPACE FOR ROUGH WORK

SPACE FOR ROUGH WORK