No. of Printed Pages : **8**

A10.5-R5.1 : MACHINE LEARNING USING PYTHON

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)

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

(1x10)

1.1 Is this code valid in Python?

```
>>> m=1,2,23,25
>>> m
```

(A) No, many values will unpack

(B) Yes, (1,2,23,25) will be printed

(C) Yes, 6 will be printed

(D) Yes, [1,2,23,25] will be printed

1.2 What will be the output of this Python code?

```
def find(x, **y):
        print(type(y))
find(" letters",X='1',Y='2')
```

(A) Dictionary

(B) An exception is thrown

(C) String

(D) Tuple

1.3 Which of the following machine learning algorithms can be used with labeled data?

(A) Regression algorithms

(B) Clustering algorithms

(C) Association algorithms

(D) All of the above

1.4 Which of the following is a disadvantage of decision trees?

(A) Decision trees are robust to outliers

(B) Decision trees are prone to be overfit

(C) Both (A) and (B)

(D) None

1.5 Which machine learning models are trained to make a series of decisions based on the rewards and feedback they receive for their actions?

(A) Supervised Learning

(B) Unsupervised Learning

(C) Reinforcement Learning

(D) All of the Above

1.6 Which of the following statement is not correct for neural networks?

(A) Neural networks mimic the human brain.

(B) It can only work for a single input and a single output.

(C) It can be used in image processing.

(D) None

1.7 In a classification problem, which of the following activation function is most widely used in the output layer of neural networks?

(A) Rectifier function

(B) Sigmoid function

(C) Hyperbolic function

(D) All of the above

1.8 Which of following are the components of object recognition system?

(A) Model database, hypothesizer

(B) Feature detector, hypothesis verifier

(C) Both (A) and (B)

(D) None of the above

1.9 Morphological Segmentation

(A) is an extension of propositional logic

(B) Does Discourse Analysis

(C) Separate words into individual morphemes and identify the class of the morphemes

(D) None of the mentioned

1.10 The Bag-of-Words approach

(A) keeps word order, keeps word multiplicity

(B) keeps word order, disregards word multiplicity

(C) disregards word order, disregards word multiplicity

(D) disregards word order, keeps word multiplicity

2. Each statement below is either TRUE or FALSE. Choose the most appropriate one and ENTER in the "OMR" sheet attached to the question paper, following instructions therein.

(1x10)

2.1 Django library is used for game development in python.

2.2 An int can be converted to a string while a string cannot be converted to an int.

2.3 K-means is a clustering algorithm that is guaranteed to converge.

2.4 'A' Search will always expand fewer search nodes than uniform cost search.

2.5 The back-propagation algorithm, when run until a minimum is achieved, always finds the same solution (i.e., weights) no matter what the initial set of weights are.

2.6 A classifier trained on less training data is less likely to overfit.

2.7 Computer Vision Azure services are best for detecting brand logos.

2.8 OCR API works best synchronously.

2.9 Speech- to-text services supports only real-time processing.

2.10 We can use QnA maker and the bot to interact through email channel.

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

(1x10)

X		Y	
3.1	Clustering	**A**	Predict a numeric label based on an item's features.
3.2	Regression	**B**	Group similar items based on their features.
3.3	Classification	**C**	Assign items into a set of predefined categories.
3.4	Inclusiveness	**D**	It is a type of artificial intelligence workload uses sensors to proactively alert users about potential equipment mechanical failures.
3.5	Anomaly detection	**E**	The principle that describes raising awareness of the limitations of responsible AI-based solutions.
3.6	Transparency	**F**	The principle of providing the benefits of responsible AI systems to all parts of society regardless of their gender or ethnicity.
3.7	Named Entity Recognition	**G**	isnumeric()
3.8	You can extract information printed on food product labels by using _____.	**H**	title()
3.9	Check if a string contains only numbers	**I**	Identify words in documents that represent persons, locations, or organizations.
3.10	Capitalize the first character of each word in a string	**J**	Optical Character Recognition
		K	numeric()
		L	titleCase()
		M	Reinforcement learning

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 attached to the question paper, following instructions therein. **(1x10)**

A	Stemming and Lemmatization	**B**	Walk	**C**	Overfitting	**D**	Image Blurring
E	Data	**F**	Method	**G**	Semi-Supervised	**H**	Data Preprocessing
I	Computer Vision	**J**	Density-based	**K**	Image Segmentation	**L**	Machine Learning
M	Chunking						

4.1 When a function is defined inside a class, it is known as _____.

4.2 _____ method is used to recursively list the contents of a directory and all its subdirectories.

4.3 The techniques of keyword normalization are_____.

4.4 _____ is a type of error which results in the failure to predict future observations effectively or fit additional data in the existing model.

4.5 We have a large amount of input data and only some of the data is labeled, is a type of _____ machine Learning.

4.6 A technique that is used to convert the raw data into a clean data set, is known as ____.

4.7 The world of Artificial Intelligence revolves around _____.

4.8 _____ processes would help avoid aliasing while down sampling an image

4.9 _____ clustering algorithms can handle clusters of arbitrary shape.

4.10 The process of extracting phrases from unstructured text is known as _____.

PART TWO

(Answer any FOUR questions)

5. (a) What is k-Nearest Neighbors (KNN) algorithm? Predict the missing values marked with "?" from the given table. Use k = 3 and Manhattan Distance for your calculations.

X1	X2	Y
11	7	Cow
12	8	Cow
13	4	Cow
27	9	Cow
25	5	Buffalo
10	60	Buffalo
4	4	Buffalo
5	11	Buffalo
23	12	?

(b) What are the main steps in a typical Computer Vision pipeline?

(c) Explain semantic and syntactic analysis with example.

(5+5+5)

6. (a) Explain Exception handling in Python with example.

(b) List down Image Noise Filters techniques and explain any two with example..

(c) You want to solve a classification task. You first train your network on 20 samples. Training converges, but the training loss is very high. You then decide to train this network on 10,000 examples. Is your approach to fixing the problem correct? If yes, explain the most likely results of training with 10,000 examples. If not, give a solution to this problem.

(4+6+5)

7. (a) Discuss some pre-processing techniques used to prepare the data in Python.

(b) Explain Back propagation algorithm with example.

(c) Explain difference between Classification and Regression with example.

(5+6+4)

8. (a) Write a Python program to check the validity of a password given by the user. The Password should satisfy the following criteria:

1. Contain at least 1 letter between a and z
2. Contain at least 1 number between 0 and 9
3. Contain at least 1 letter between A and Z
4. Contain at least 1 character from $, #, @
5. Minimum length of password: 6

(b) Explain the principle of the gradient descent algorithm. Accompany your explanation with a diagram.

(c) Provide an intuitive explanation of how the Sliding Window approach works in Object Detection.

(5+5+5)

9. (a) How do Neural Networks distinguish useful features from non-useful features in Computer Vision?

(b) What is Parts-of-speech Tagging? Explain it with example.

(c) Find all frequent itemsets using the Apriori Algorithm having a support of 60% for the following set of transactions.

T1: cereal, butter, cheese, cola
T2: bread, butter, milk, water
T3: cereal, cheese
T4: cereal, bread, butter, water
T5: cola, bread, water
T6: Bread, water, cola, milk, cheese

(5+5+5)

- o O o -

SPACE FOR ROUGH WORK

SPACE FOR ROUGH WORK