

Sl. No.

**A10.5-R5 : 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 respect.

**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** Which of the following command is used to shuffle a list 'xyz' ?

- (A) random.shufflelist(xyz)
- (B) shuffle(xyz)
- (C) L.shuffle(xyz)
- (D) random.Shuffle(xyz)

**1.2** Which of the following is true about Naive Bayes ?

- (A) Assumes that all the features in a dataset are equally important
- (B) Assumes that all the features in a dataset are independent
- (C) Both (A) and (B)
- (D) None of the above

**1.3** For what purpose Feedback neural networks are primarily used ?

- (A) Classification
- (B) Feature mapping
- (C) Pattern mapping
- (D) None of the above

**1.4** The SVM's are less effective when :

- (A) The data is linearly separable
- (B) The data is clean and ready to use
- (C) The data is noisy and contains overlapping points
- (D) None of the above

**1.5** Which of the following is a widely used and effective machine learning algorithm based on the idea of bagging ?

- (A) Decision Tree
- (B) Regression
- (C) Random Forest
- (D) Classification

**1.6** What is the result of round (0.5) - round (-0.5) ?

- (A) 0.0
- (B) 2.0
- (C) 1.0
- (D) 0.25

**1.7** Which of the following is example of low level feature in an image ?

- (A) HOG
- (B) SIFT
- (C) HAAR features
- (D) All of the above

- 1.8 What will be the output of the following Python code ?
- ```

i=0
while i<3:
print(i)
i+=1
else:
print(0)

```
- (A) 0 1 2 0  
(B) 0 1 2 3  
(C) 2 1 0  
(D) 1 2 0
- 1.9 Which of the following is a challenge when dealing with computer vision problems ?
- (A) Variations due to geometric changes (like pose, scale etc.)  
(B) Variations due to photometric factors (like illumination, appearance etc.)  
(C) Image occlusion  
(D) All of the above
- 1.10 What is the output of this expression,  $3*2**3$  ?
- (A) 26  
(B) 24  
(C) 216  
(D) 214
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 If we're achieving 100% training accuracy very easily, we need to check to verify if we are overfitting our data.
- 2.2 To blur an image, you can use a linear filter.
- 2.3 A Pearson correlation between two variables is zero but, still their values can still be related to each other.
- 2.4 Quantizing an image will reduce the amount of memory required for storage.
- 2.5 In 1954 Robert Fraz developed the first learning machine in which connection strengths could be adapted automatically & efficiently.
- 2.6 Each point which will always be misclassified in 1-NN which means that you will get the 0% accuracy.
- 2.7 Neural networks have high computational rates than conventional computers.
- 2.8 Datasets which have a clear classification boundary will function best with SVM's.
- 2.9 LogLoss evaluation metric can have negative values.
- 2.10 Median filter technique helps reduce noise to a good enough extent.

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  | Bagging                               | A. | keyword normalization        |
| 3.2  | High-bias machine learning algorithms | B. | Maximize goodness-of-fit     |
| 3.3  | Levenshtein and Soundex Techniques    | C. | Generative model             |
| 3.4  | Conditional Random Field (CRF)        | D. | delta rule                   |
| 3.5  | Lemmatization and Stemming Techniques | E. | Discriminative model         |
| 3.6  | Hidden Markov Model (HMM)             | F. | String matching              |
| 3.7  | Activation function                   | G. | Linear Discriminant Analysis |
| 3.8  | backpropagation law                   | H. | Emotion                      |
| 3.9  | Minimize error                        | I. | Bootstrap Aggregating        |
| 3.10 | Sentiment                             | J. | Sigmoid                      |
|      |                                       | K. | Classes                      |
|      |                                       | L. | Generalization               |
|      |                                       | M. | Regression                   |

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 | Classes           | B | Over fitting           | C | Capital letter                                                         | D | Stop word         |
| E | Tokenization      | F | Support Vector Machine | G | True false and None are capitalized while the others are in lower case | H | Objects           |
| I | Systematically    | J | Boltzmann Machine      | K | Summarization                                                          | L | Increase variance |
| M | decrease variance |   |                        |   |                                                                        |   |                   |

- 4.1 All key words in Python are in \_\_\_\_\_.
- 4.2 The process of converting a sentence or paragraph into tokens is referred to as \_\_\_\_\_.
- 4.3 Detecting \_\_\_\_\_ from an image is a computer vision use case.
- 4.4 low-bias machine learning algorithms include: Decision Trees, k-NN and \_\_\_\_\_.
- 4.5 One of the most basic Deep Learning models is a \_\_\_\_\_, resembling a simplified version of the Multi-Layer Perceptron.
- 4.6 Averaging the output of multiple decision trees helps \_\_\_\_\_.
- 4.7 You say your model is biased if it \_\_\_\_\_ under or over predicts the target variable.
- 4.8 Text \_\_\_\_\_ is an NLP use case.
- 4.9 \_\_\_\_\_ is one of the disadvantage of decision trees.
- 4.10 The process of removing words like "and", "is", "a", "an", "the" from a sentence is called as \_\_\_\_\_.

**PART TWO**

**(Answer any FOUR questions)**

5. (a) Explain the Values, Types, Keywords, Statements, Expressions, and Variables in Python programming language with examples.
- (b) What is CSV ? Describe the characteristics of the CSV format. When use CSV ?
- (c) Discuss the relation between tuples and lists, tuples and dictionaries in detail.
- (5+5+5)**
6. (a) What are the different types of Machine Learning ? Differentiate between them.
- (b) What is the procedure of building Decision Tree using ID3 algorithm with Gain and Entropy.
- (c) What is SVM ? What is the goal of the support vector machine (SVM) ?
- (5+5+5)**
7. (a) Explain Navies Bayes Classifier with example.
- (b) Explain steps of face recognition process using PCA.
- (8+7)**
8. (a) What is Artificial Neural Network ? Explain appropriate problem for Neural Network Learning with its characteristics.
- (b) Explain the concept of a Perceptron and represent the Boolean functions of AND or OR using perceptron.
- (c) Write the differences between Single Layer Perceptron (SLP) and Multi Layer Perceptron (MLP).
- (5+5+5)**
9. (a) Explain with a diagram organization of natural language processing understanding system.
- (b) Explain lexical resource wordnet used in natural language processing
- (8+7)**

- o O o -

---

**SPACE FOR ROUGH WORK**

---

SPACE FOR ROUGH WORK