A10.5-R5 : MACHINE LEARNING USING PYTHON

अवधि : 03 घंटे
DURATION : 03 Hours

अधिकतम अंक : 100
MAXIMUM MARKS : 100

ओएमआर शीट सं. :
OMR Sheet No. :

रोल नं. :
Roll No. :

उत्तर-पुस्तिका सं. :
Answer Sheet No. :

परीक्षार्थी का नाम :
Name of Candidate : _____

परीक्षार्थी के हस्ताक्षर :
;**Signature of Candidate :** _____

परीक्षार्थियों के लिए निर्देश : / **Instructions for Candidate :**

कृपया प्रश्न-पुस्तिका, ओएमआर शीट एवं उत्तर-पुस्तिका में दिये गए निर्देशों को ध्यानपूर्वक पढ़ें।	Carefully read the instructions given on Question Paper, OMR Sheet and Answer Sheet.
प्रश्न-पुस्तिका की भाषा अंग्रेजी है। परीक्षार्थी केवल अंग्रेजी भाषा में ही उत्तर दे सकता है।	Question Paper is in English language. Candidate can answer in English language only.
इस मॉड्यूल/पेपर के **दो भाग** हैं। **भाग एक** में **चार** प्रश्न और **भाग दो** में **पाँच** प्रश्न हैं।	There are **TWO PARTS** in this Module/Paper. **PART ONE** contains **FOUR** questions and **PART TWO** contains **FIVE** questions.
भाग एक ''वैकल्पिक'' प्रकार का है जिसके कुल अंक 40 है तथा **भाग दो** ''व्यक्तिपरक'' प्रकार का है और इसके कुल अंक 60 है।	**PART ONE** is Objective type and carries 40 Marks. **PART TWO** is Subjective type and carries 60 Marks.
भाग एक के उत्तर, **ओएमआर उत्तर-पुस्तिका** पर ही दिये जाने हैं। **भाग दो** की उत्तर-पुस्तिका में **भाग एक** के उत्तर **नहीं** दिये जाने चाहिए।	**PART ONE** is to be answered in the **OMR ANSWER SHEET** only. **PART ONE** is **NOT** to be answered in the answer book for **PART TWO**.
भाग एक के लिए अधिकतम समय सीमा **एक घण्टा** निर्धारित की गई है। **भाग दो** की उत्तर-पुस्तिका, **भाग एक** की उत्तर-पुस्तिका जमा कराने के पश्चात् दी जाएगी। तथापि, निर्धारित एक घंटे से पहले **भाग एक** पूरा करने वाले परीक्षार्थी **भाग एक** की उत्तर-पुस्तिका निरीक्षक को सौंपने के तुरंत बाद, **भाग दो** की उत्तर-पुस्तिका ले सकते हैं।	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.
परीक्षार्थी, उपस्थिति-पत्रिका पर हस्ताक्षर किए बिना और अपनी उत्तर-पुस्तिका, निरीक्षक को सौंपे बिना, परीक्षा हॉल/कमरा नहीं छोड़ सकते हैं। ऐसा नहीं करने पर, परीक्षार्थी को इस मॉड्यूल/पेपर में अयोग्य घोषित कर दिया जाएगा।	**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.**
प्रश्न-पुस्तिका को खोलने के निर्देश मिलने के पश्चात् एवं उत्तर लिखना आरम्भ करने से पहले उम्मीदवार जाँच कर यह सुनिश्चित कर लें कि प्रश्न-पुस्तिका प्रत्येक दृष्टि से संपूर्ण है।	After receiving the instruction to open the booklet and before starting to answer the questions, the candidate should ensure that the Question Booklet is complete in all respect.

जब तक आपसे कहा न जाए, तब तक प्रश्न-पुस्तिका न खोलें।
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 supplied with the question paper, following instructions therein.

(1x10)

1.1 Which of the following involves the learning from a training dataset with labelled data using classification and regression models ?
(A) Supervised Learning
(B) Unsupervised Learning
(C) Semi-supervised Learning
(D) Reinforcement Learning

1.2 Which of the following is contained in NumPy library ?
(A) n-dimensional array object
(B) tools for integrating C/C++ and Fortran code
(C) fourier transform
(D) all of the mentioned

1.3 Which of the following will be Euclidean Distance between the two data point A(1,3) and B(2,3) ?
(A) 1
(B) 2
(C) 4
(D) 8

1.4 Which of the following function of Pandas gives the count of the data in the column ?
(A) .count()
(B) .sum()
(C) .mean()
(D) .median()

1.5 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

1.6 What is the objective of backpropagation algorithm ?

(A) To develop learning algorithm for multilayer feedforward neural network

(B) To develop learning algorithm for single layer feedforward neural network

(C) To develop learning algorithm for multilayer feedforward neural network, so that network can be trained to capture the mapping implicitly

(D) none of the mentioned

1.7 Which of the following includes major tasks of NLP ?

(A) Automatic Summarization

(B) Discourse Analysis

(C) Machine Translation

(D) All of the mentioned

1.8 Point out the wrong combination with regards to kind keyword for graph plotting.

(A) 'scatter' for scatter plots

(B) 'kde' for hexagonal bin plots

(C) 'pie' for pie plots

(D) none of the mentioned

1.9 Which of the following is the advantage/s of Decision Trees ?

(A) Possible Scenarios can be added

(B) Use a white box model, If given result is provided by a model

(C) Worst, best and expected values can be determined for different scenarios

(D) All of the mentioned

1.10 What will happen if a learning rate used is too large ?

(A) Network will converge slowly

(B) Network will not converge

(C) Can't Say

(D) Network will converge fast

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 Decision Trees can be used for Classification Tasks.

2.2 Plots may also be adorned with error bars or tables.

2.3 A perceptron adds up all the weighted inputs it receives, and if it exceeds a certain value, it outputs a 1, otherwise it just outputs a 0.

2.4 A graph is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs and utility.

2.5 Training set is used to test the accuracy of the hypotheses generated by the learner.

2.6 Dropout can be applied at visible layer of Neural Network model.

2.7 Machine Translation converts any Human Language to French Language.

2.8 In Python programming language, a Tuple is a collection which is ordered and unchangeable.

2.9 Support Vector machine is a discriminative classifier.

2.10 print (3^4) evaluates to 8.

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	Machine learning	**A.**	free and open-source
3.2	Word Tokenization	**B.**	supervised learning algorithm
3.3	Python	**C.**	python library for data manipulation and analysis
3.4	Cross-validation	**D.**	Step in natural language processing
3.5	Artificial Neural Network	**E.**	focuses on the development of computer programs that can access data and use it learn for themselves
3.6	Sentiment analysis	**F.**	Computer vision
3.7	A tuple can also be created without using parentheses	**G.**	model validation technique
3.8	A field of study focused on the problem of helping computers to see.	**H.**	application of natural language processing
3.9	Face Recognition and Detection	**I.**	Tuple Packing
3.10	Pandas	**J.**	application of computer vision
		K.	{}
		L.	l[4:6]
		M.	Dataframes

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.	Confusion matrix	B.	append	C.	Mutable	D.	False positive
E.	Computer Vision	F.	Underfitting	G.	Deep learning	H.	Seaborn
I.	TensorFlow	J.	Natural language processing	K.	K-NN	L.	Reinforcement learning
M.	Gradient Descent						

4.1 _____ means not to capture enough patterns in the data.

4.2 _____ is the subset of machine learning in artificial intelligence (AI).

4.3 _____ adds its argument as a single element to the end of a list.

4.4 A characteristic of a list data structure is _____.

4.5 An outcome where the model incorrectly predicts the positive class is known as _____.

4.6 _____ deals with how computers can be made to gain high-level understanding from digital images or videos.

4.7 _____ programs the computers to process and analyze large amounts of natural language data.

4.8 Python data visualization library based on matplotlib is _____.

4.9 _____ is Python library for fast numerical computing created and released by Google.

4.10 _____ is a technique for summarizing the performance of a classification algorithm.

PART TWO

(Answer any FOUR questions)

5. (a) What do you mean by the term Python Module ? Mention the different ways to define a module in Python.

   (b) Explain Sentiment Analysis with example.

   (c) State applications of computer vision and explain them in brief. **(5+5+5)**

6. (a) Explain false negative, false positive, true negative and true positive with a simple example.

   (b) What are the different types of Machine Learning ?

   (c) Write a Python program to find the maximum element of each row in a matrix. **(7+5+3)**

7. (a) What is Natural Language Processing (NLP) ? Differentiate between Natural Language and Computer Language. Also, enlist the advantages and disadvantages of NLP.

   (b) Write a Python program to print all integers that are not divisible by either 2 or 3 and lie between 1 and 50.

   (c) Differentiate between unsupervised machine learning and supervised machine learning. **(7+5+3)**

8. (a) Assume we have a set of data from patients who have visited UPMC hospital during the year 2020. A set of features (e.g., temperature, height) have been also extracted for each patient. Our goal is to decide whether a new visiting patient has any of following diseases :

   diabetes, heart disease and Alzheimer (a patient can have one or more of these diseases).

   We have decided to use a neural network to solve this problem. We have two choices : either to train a separate neural network for each of the diseases or to train a single neural network with one output neuron for each disease, but with a shared hidden layer. Which method do you prefer ? Justify your answer.

   (b) Write a Python program to find Armstrong Number in a certain interval. Note: An Armstrong number of three digits is an integer such that the sum of the cubes of its digits is equal to the number itself. For example, $153 = 1*1*1 + 5*5*5 + 3*3*3$ i.e. 153 is an Armstrong number. **(8+7)**

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

   (a) Feed Forward Neural network

   (b) Text Classification with example

   (c) Cross validation in Machine Learning

   (d) Exception handling in Python with example **(5+5+5)**

- o O o -

A10.5-R5 01-22