Register Number | | | | | | | | | | |

Code: 20CS51IT

V Semester Diploma Make-Up Examination, July 2024

ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

Duration: 3 Hours **Max.Marks:** 100

INSTRUCTIONS: 1. Answer **any ONE Full** question from each Section.
2. **ONE full** question carries **20** marks.

SECTION-I

1.a) Describe AI and its applications in various fields. 10

.b) Write steps to create repository in GitHub and add file. 10

2.a) Explain how is AI software development life cycle different from traditional software development? 10

2.b) For the following scenarios you are required to build a predictive model. Which machine learning technique/ algorithm can be applied / best suited for stated problems. Justify your recommendation.

i. Predicting the food delivery time

ii. Predicting whether the transaction is fraudulent

iii. Predicting the credit limit of a credit card applicant

iv. To group similar customers of an online grocery store, based on their purchasing patterns, to offer discounts to its customers.

v. Predict the probability of a mechanical system breakdown, based on its system vibration and operating temperature.

SECTION-II

3.a) Handling missing values in a dataset is a crucial data pre-processing step, as missing data can lead to biased or incorrect results in your analysis or machine learning models. Elaborate on how missing values in the data sets can be handled.

3.b) Create two series as shown using pd. series() function. 10

Series A = [20, 30, 40, 50, 60]

Series B = [50, 60, 70, 80, 90]

i. Get the items not common to both.

ii. Identify the smallest and largest element in the Series A.

iii. Find the sum of Series B.

iv. Calculate mean in the Series A.

v. Find median in the given Series B.

4.a) Perform the following operations on Car manufacturing company dataset auto-mpg.csv given below using pandas

i. Read data from an existing file

ii. Statistical details of dataset

iii. Get all cars with 8 cylinders

iv. Get the number of cars manufactured in each year. 10

b) Explain Supervised and Unsupervised learning with examples. 5

c) Compare overfitting with under-fitting. 5

SECTION - III

5.a) A dataset is given to you for creating machine learning model. What are the steps followed before using the data for training the model? Elaborate each step. 10

5.b) Referring to the number of variables or features in a dataset and the focus of analysis. Explain univariate & multivariate data types with examples. 10

6.a) Explain any 2 techniques of cross validation used in Machine Learning. 5

b) Compare "Classification algorithms" with "Clustering algorithm". 5

c) A Machine learning model was built to classify spam emails as "spam"(1) or "not spam"(0). The confusion matrix for the model is as shown below. Evaluate accuracy, precision, recall, specificity and F1-Score.

		Actual	
		1	0
Predicted	1	140	10
	0	5	50

SECTION - IV

7.a) N-grams are a type of linguistic model used in natural language processing (NLP) and computational linguistics. Consider the given sentence: "Artificial Intelligence

is a branch of computer science that focuses on creating intelligent machines capable of performing tasks that typically require human intelligence."

i. Generate bi grams for the above sentence

ii. Generate tri-grams for the above sentence 10

7.b) Summarize any two cloud deployment models 10

8.a) Demonstrate simple linear regression considering a dataset that has two variables: **"Marks"** (dependent variable) 10

"Hours of study" (independent variable)

b) Explain how data exploration, pre-processing of data and splitting of data are performed on datasets. 10

SECTION V

9.a) What are MLOps? Brief explain different stages involved in the Machine Learning Operations (MLOps) lifecycle. 10

b) With examples demonstrate Stemming and Lemmatization normalization techniques. 10

10.a) Discuss any five ethical challenges in AI. 5

b) Explain cloud deployment models. 5

c) With a neat diagram explain components of Docker. 10