

Makeup Examination – Sept. 2023
V Semester Diploma Examination

ARTIFICIAL INTELLIGENCE & MACHINE LEARNING

(Exam Date / Time: 22nd Sep. 2023 / 2.00 PM)

Time: 3 Hours

Max.Marks: 100

Instructions: (1) Answer 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 Marks

1.b How AI Software Development life cycle differs from traditional software development. Explain. 05 Marks

1.c Summarize the challenges associated with Machine Learning. 05 Marks

2.a Perform the following operations/write code snippet on Car manufacturing company dataset "auto-mpg.csv" given below using pandas. 10 Marks

- i) Read data from a file.
- ii) Calculate mean value of "horsepower".
- iii) Calculate Standard Deviation value of "acceleration".
- iv) Get the number of cars manufactured in each year.

| mpg | cylinders | displacement | horsepower | weight | acceleration | model year | car name |
|-----|-----------|--------------|------------|--------|--------------|------------|------------|
| 18 | 8 | 307 | 130 | 3504 | 12 | 71 | Chevrolet |
| 15 | 8 | 350 | 165 | 3693 | 11.5 | 70 | Skylark |
| 18 | 8 | 318 | 150 | 3436 | 11 | 72 | Plymouth |
| 17 | 8 | 302 | 140 | 3449 | 10.5 | 70 | Ford |
| 14 | 8 | 455 | 225 | 4425 | 10 | 71 | Pontiac |
| 15 | 8 | 390 | 190 | 3850 | 8.5 | 70 | Ambassador |

2.b Explain how is AI software development life cycle different from traditional software development? 10 Marks

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. 10 Marks

3.b 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 Marks

4.a Create two series as shown using pd. series() function. 10 Marks

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.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 Marks

SECTION – III

- 5.a 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. 10 Marks

| | | Actual | |
|-----------|---|--------|----|
| | | 1 | 0 |
| Predicted | 1 | 140 | 10 |
| | 0 | 5 | 50 |

- 5.b Explain Supervised and Unsupervised learning with examples.

5 Marks

- 5.c Compare overfitting with under-fitting.

5 Marks

- 6.a How to Choose the Right Number of Clusters in k-means clustering? Explain any one method.

5 Marks

- 6.b Compare "Classification algorithms" with "Clustering algorithm".

5 Marks

- 6.c Explain with examples: Scalars, Vectors, Matrices, Tensors and Gradients in Linear Algebra.

10 Marks

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 Marks

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

- 8.a With examples demonstrate Stemming and Lemmatization normalization techniques.

10 Marks

- 8b. Explain any 2 techniques of cross validation used in Machine Learning.

5 Marks

- 8c. Brief explain different stages involved in the Machine Learning Operations (MLOps) lifecycle.

5 Marks

SECTION – V

- 9.a Demonstrate simple linear regression considering a dataset that has two variables: "Marks" (dependent variable) and "Hours of study" (independent variable)

10 Marks

- 9.b With a neat diagram explain components of Docker.

- 10.a Summarize any two cloud deployment models .

10 Marks

- 10.b Discuss any five ethical challenges in AI.

10 Marks

10 Marks