BRAINWARE UNIVERSITY

Term End Examination 2025-2026
Programme – B.Tech.(CSE)-AIML-2023/B.Tech.(CSE)-DS-2023/B.Tech.(CSE)-2023
Course Name – Machine Learning and its Applications
Course Code - PCC-CSM601/PCC-CSD601/PCC-CSG601
(Semester VI)

Full Marks : 60 **Time : 2:30 Hours**

[The figure in the margin indicates full marks. Candidates are required to give their answers in their own words as far as practicable.]

Group-A
(Multiple Choice Type Question) 1 x 15=15

1. *Choose the correct alternative from the following :*

(i) Identify the mathematical concept fundamental for understanding optimization algorithms in machine learning.

a) Calculus b) Linear Algebra
c) Probability Theory d) Number Theory

(ii) Name the mathematical technique that is commonly used for dimensionality reduction in machine learning.

a) Singular Value Decomposition b) Fourier Transform
c) Taylor Series Expansion d) Wavelet Transform

(iii) Identify the condition in a symmetric distribution concerning mean, median, and mode.

a) Mean < Median b) Mean > Median
c) Mean = Median = Mode d) Mean = Mode only

(iv) Identify the relationship among mean and median in a normal distribution.

a) Mean < Median b) Mean > Median
c) Mean = Median d) Mean and Median differ

(v) Select from the given options that is not a type of machine learning.

a) Supervised learning b) Unsupervised learning
c) Transferred learning d) Reinforcement learning

(vi) Identify example of an unsupervised learning task.

a) Predicting house prices based on square footage and neighbourhood b) Grouping customers into different segments based on their purchase history
c) Detecting fraudulent transactions in financial data d) Identifying sentiment (positive, negative) from customer reviews

(vii) Identify the type of learning, that is used for labeled training data.

a) Supervised learning b) Unsupervised learning
c) Reinforcement learning d) Semi-supervised learning

(viii) Identify the phase of the CRISP-DM methodology would you explore and pre-process the data to make it suitable for modeling

a) Data Understanding. b) Data Preparation.
c) Modeling. d) Evaluation.

(ix) Identify the phase includes transforming raw data into a structured format suitable for analysis in CRISP-DM.

a) Data Understanding b) Data Preparation
c) Evaluation d) Deploymen

(x) Identify the primary objective of EDA

a) Cleaning the dataset b) Summarizing the dataset and finding patterns
c) Training the model d) Evaluating model performance

(xi) Identify a box plot is useful for detecting:

a) Missing values b) Outliers
c) Data types d) Feature scaling

(xii) Select the factor that affects model selection in machine learning.

a) Training accuracy only b) Model interpretability and generalization
c) Number of hidden layers only d) Number of input features only

(xiii) Select the metric suitable for evaluating regression models.

a) Accuracy b) F1-score
c) Mean Squared Error d) Confusion Matrix

(xiv) Select the evaluation metric commonly used in logistic regression.

a) R-Squared b) Mean Squared Error
c) Adjusted R-Squared d) F1-Score

(xv) Identify the type of function used in logistic regression.

a) Sigmoid b) Linear
c) Polynomial d) Exponential

Group-B

(Short Answer Type Questions) 3 x 5=15

2. Calculate the probability density function if mean and variance are 60 and 7 respectively, for a dataset and x = 65. (3)
3. Define Machine Learning and give examples. (3)
4. Contrast the purpose of the Data Understanding phase in CRISP-DM. (3)
5. Explain the importance of model selection in machine learning. (3)
6. Correlate the applications of linear algebra with machine learning. (3)

OR

Summarize with examples the use of linear algebra concepts, such as vectors, matrices, and matrix operations, in machine learning algorithms. (3)

Group-C

(Long Answer Type Questions) 5 x 6=30

7. Define sampling distribution and Binomial distribution and its importance in machine learning. (5)
8. Illustrate the types of unsupervised learning. (5)
9. Illustrate the seven major steps involved in the machine learning life cycle. (5)
10. Construct the techniques for handling missing data. Explain when you should use imputation vs. removal of missing values. (5)
11. Compare decision trees and logistic regression for classification tasks. (5)
12. Explain the working of linear regression using a real-time dataset and evaluate its predictive performance with an example. (5)

OR

Compare simple linear regression and multiple linear regression with suitable examples. (5)

****