

Question Paper Code : 50070

B.E./B.Tech. DEGREE EXAMINATIONS, APRIL/MAY 2024.

Fourth Semester

Computer Science and Engineering (Artificial Intelligence and Machine Learning)

AL 3451 — MACHINE LEARNING

(Common to : Artificial Intelligence and Data Science/Computer Science and Business Systems)

(Regulations 2021)

Time : Three hours

Maximum : 100 marks

Answer ALL questions.

PART A — (10 × 2 = 20 marks)

1. Define Machine Learning.
2. Write some examples for machine learning applications.
3. Mention the merits of Bayesian linear regression.
4. Distinguish between Random Forest and Support Vector Machine.
5. When is supervised learning better than unsupervised learning?
6. Define Expectation Maximization.
7. Differentiate between Single layer and Multilayer Perceptron.
8. List the problems associated with Backpropagation Neural Network.
9. Recall the benefits of the Cross-Validation method.
10. How do you evaluate a Classification Algorithm?

PART B — (5 × 13 = 65 marks)

11. (a) Discuss the following
- (i) Vapnik-Chervonenkis (VC) Dimension. (6)
 - (ii) Probably Approximately Correct (PAC) Learning. (7)

Or

- (b) Write detailed notes on Inductive Bias and Bias variance trade-off.
12. (a) By the method of least squares find the straight line to the data given below.

X	5	10	15	20	25
Y	16	19	23	26	30

Or

- (b) With an example explain the Decision Tree concepts in detail.
13. (a) Compare Bagging, Boosting and Stacking ensemble methods.

Or

- (b) Cluster the following eight points (with (x, y) representing locations) into three clusters using K-means clustering method. A1(2, 10), A2(2, 5), A3(8, 4), A4(5, 8), A5(7, 5), A6(6, 4), A7(1, 2), A8(4, 9).
14. (a) Describe in brief about Multilayer perceptron activation functions.

Or

- (b) Illustrate the following
- (i) Batch Normalization. (7)
 - (ii) Dropout. (6)
15. (a) With neat diagram explain about K-fold Cross Validation technique.

Or

- (b) Elaborate the t test, McNemar's test and K-fold CV paired t test by giving your own example.

PART C — (1 × 15 = 15 marks)

16. (a) Consider the following list that contains name, age, gender and class of sports. In the Gender field males are denoted by the numeric value 0 and females by 1. Using the K-Nearest Neighbor (KNN) algorithm, find class of sports for a girl whose name is Angelina, her k factor is 3, and her age is 5.

Ajay	32	0	Football
Mark	40	0	Neither
Sara	16	1	Cricket
Zaira	34	1	Cricket
Sachin	55	0	Neither
Rahul	40	0	Cricket
Pooja	20	1	Neither
Smith	15	0	Cricket
Laxmi	55	1	Football
Michael	15	0	Football

Or

- (b) The grades of a class of 9 students on a midterm report (X) and on the final examination (Y) are as follows:

X	77	50	71	72	81	94	96	99	67
Y	82	66	78	34	47	85	99	99	68

- (i) Estimate the linear regression line. (12)
- (ii) Estimate the final examination grade of a student who received a grade of 85 on the midterm report. (3)