

USN

--	--	--	--	--	--	--	--	--	--

Course Code

**22AI53**

Fifth Semester B.E. Degree Examinations, February 2025

**MACHINE LEARNING**

(AIML)

Duration: 3 hrs

Max. Marks: 100

- Note:** 1. Answer any FIVE full questions, choosing ONE full question from each module.  
2. Missing data, if any, may be suitably assumed

Q. No Question Marks (RBTL:CO:PI)

**Module-1**

1. a. What is machine learning? Explain different perspectives and issues in machine learning. **07** (1 :1: 1.6.1)  
b. Describe the Find-s algorithm. Explain by taking the Enjoy Sport concept and training instance given below. **08** (2 :3: 2.7.1)

Example	Sky	Air Temp	Humidity	Wind	Water	Forecast	Enjoy Sport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

- c. Discuss about an unbiased learner. **05** (1 :1: 1.6.1)

(OR)

2. a. List and explain the steps to design a learning system in detail. **10** (1 :1: 1.6.1)  
b. Write candidate elimination algorithm. Apply the algorithm to obtain the final version space for the training example. **10** (2 :3: 2.7.1)

Example	Sky	Air Temp	Humidity	Wind	Water	Forecast	Enjoy Sport
1	Sunny	Warm	Normal	Strong	Warm	Same	Yes
2	Sunny	Warm	High	Strong	Warm	Same	Yes
3	Rainy	cold	High	Strong	Warm	Change	No
4	Sunny	Warm	High	Strong	Cool	Change	Yes

**Module-2**

3. a. Illustrate various methods to prepare the data for machine learning. **08** (3 :2: 1.6.1)  
b. Explain the importance of visualizing data before preparing it for a machine-learning model. **06** (2 :2: 1.6.1)  
c. Distinguish between the real data and assumption data in Machine learning. **06** (2 :2: 1.6.1)

(OR)

4. a. Explain how you frame the problem and choose appropriate measures for a dataset in a machine learning Project. **07** (2 :2: 2.7.1)  
b. What is the significance of the MNIST dataset in machine learning? **08** (1 :2: 1.6.1)  
c. Describe error analysis, and why is it crucial in the process of training a machine learning model. **05** (2 :2: 1.6.1)

**Module-3**

5. a. What is the normal equation, and how does it help find the parameters of a linear regression model? **10** (1 :3: 1.6.1)

**Note: (RBTL - Revised Bloom's Taxonomy Level: CO - Course Outcome: PI- Performance Indicator)**

- b. Explain how gradient descent optimizes a model's cost function. Why is the learning rate critical? **10** (2 :3: 1.6.1)

(OR)

6. a. Illustrate polynomial regression limitations of a simple linear model? **08** (3 :3: 1.6.1)  
b. Write a short note on the following: **12** (1 :3: 1.6.1)

- (i) Ridge Regression (ii) Logistic Regression  
(iii) Decision Boundaries in Regressions.

#### Module-4

7. a. What is a decision tree? Explain the training and visualizing a decision tree process using Iris dataset. **10** (1 :3: 1.6.1)  
b. Describe CART training algorithm, and how does it handle classification and regression tasks? **07** (1 :4: 1.6.1)  
c. Explain ensemble learning? Why is it effective? **03** (2 :4: 1.6.1)

(OR)

8. a. Define voting in ensemble learning. List its advantages and disadvantages. **07** (1 :4: 1.6.1)  
b. Illustrate SLR parsing algorithm and explain with example. **06** (3 :4: 2.7.1)  
c. What is the bootstrap technique? How is it used in bagging? **07** (1 :4: 1.6.1)

#### Module-5

9. a. Define Bayesian theorem? What is the relevance and features of Bayesian theorem? Explain the practical difficulties of Bayesian theorem. **08** (2 :4: 1.6.1)  
b. Illustrate medical diagnosis problem in which there are two alternative hypotheses: **12** (3 :5: 2.7.1)

- (i) that the patient has a particular form of cancer (+) and  
(ii) That the patient does not (-).

A patient takes a lab test, and the result comes back positive, The test returns a correct positive result in only 98 % of the cases in which the disease is actually present, and a correct negative result in only 97 % of the cases in which the disease is not present. Furthermore, 0.008 of the entire population have this cancer. Determine whether the patient has Cancer or not using MAP hypothesis.

(OR)

- 10 a. Explain Maximum Posteriori (MAP) and Maximum Likelihood (ML) Hypothesis. Derive the relation for  $h_{MAP}$  and  $h_{ML}$  using Bayesian theorem. **08** (2 :4: 1.6.1)  
b. Apply Naive Bayes classifier for Play Tennis concept learning problem to classify the following novel instance. **12** (3 :5: 2.7.1)

< Outlook: sunny, Temperature: cool, Humidity: high, Wind: strong >

Day	Outlook	Temp.	Humidity	Wind	Play Tennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes
D11	Sunny	Mild	Normal	Strong	Yes
D12	Overcast	Mild	High	Strong	Yes
D13	Overcast	Hot	Normal	Weak	Yes
D14	Rain	Mild	High	Strong	No

\*\* \*\* \*