| U.S.N. | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|

# B.M.S. College of Engineering, Bengaluru-560019

**Autonomous Institute Affiliated to VTU**

## June / July 2025 Semester End Main Examinations

**Programme: B.E.** **Semester: V**
**Branch: Information Science and Engineering** **Duration: 3 hrs.**
**Course Code: 23IS5PCMLG** **Max Marks: 100**
**Course: Machine Learning**

**Instructions**: 1. Answer any FIVE full questions, choosing one full question from each unit.
2. Missing data, if any, may be suitably assumed.

**Important Note:** Completing your answers, compulsorily draw diagonal cross lines on the remaining blank pages. Revealing of identification, appeal to evaluator will be treated as malpractice.



| | | UNIT - I | CO | PO | Marks |
|---|---|---|---|---|---|
| 1 | a) | A binary classification model is used to predict whether a patient has a disease (Positive) or not (Negative). The predictions and actual labels for 10 patients are given below:<br><br>**Patient** / **Actual Label** / **Predicted Label**<br>1 / Positive / Positive<br>2 / Negative / Negative<br>3 / Positive / Negative<br>4 / Positive / Positive<br>5 / Negative / Negative<br>6 / Negative / Positive<br>7 / Positive / Negative<br>8 / Positive / Positive<br>9 / Negative / Negative<br>10 / Negative / Positive<br><br>I. Construct the confusion matrix for this classification task.<br>II. Calculate the Accuracy, precision, recall and F1 score. | *CO2* | *PO2* | **10** |
| | b) | Consider you are building a regression model to predict house prices using features such as size, location, number of rooms, and age of the house. During evaluation, the model performs very well on the training data but poorly on the test data.<br><br>I. Identify the problem affecting the model's performance.<br>II. Explain any two solutions to address this issue. | *CO3* | *PO2* | **10** |
| | | **OR** | | | |
| 2 | a) | Imagine You are working on a machine learning project to predict customer churn for a telecom company. You have gathered data on customer demographics, usage patterns, and service feedback. However, after training and testing multiple machine learning | *CO3* | *PO2* | **10** |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | models, you notice the model's performance is not satisfactory due to **"bad data" and "bad algorithm**. explain the challenges posed by "**bad data" and "bad algorithm** in a machine learning project. | | | |
| | | b) | Consider the below dataset which contains information about customers and their subscription plans to a telecom service: Design the steps involved in transforming this raw data into a clean, understandable, and readable format using data preprocessing techniques and choose an appropriate machine learning algorithm based on the preprocessed features to predict Total Spend. | *CO2* | *PO2* | **10** |

| **Cust-ID** | **Age** | **Income** | **Gender** | **Subscription Plan** | **Total Spend ($)** |
|---|---|---|---|---|---|
| 1 | 25 | 50000 | Male | Basic | 150 |
| 2 | 30 | 60000 | Female | Premium | 250 |
| 3 | NaN | 55000 | Male | Basic | 180 |
| 4 | 28 | 48000 | Female | Standard | 200 |
| 5 | 35 | 70000 | Male | Premium | 300 |
| 6 | 40 | 45000 | Female | Standard | 100 |
| 7 | 27 | NaN | Male | Premium | 220 |
| 8 | 50 | 80000 | Female | Basic | 190 |
| 9 | 60 | 95000 | Male | Premium | 350 |
| 10 | 23 | 45000 | Female | Standard | 120 |

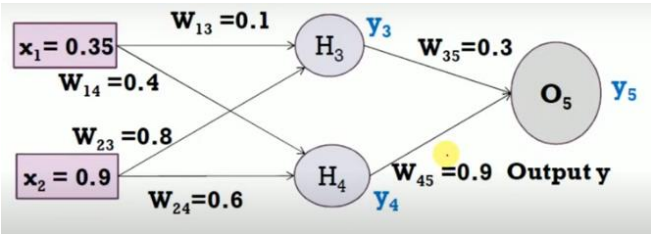
| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | **UNIT - II** | | | |
| 3 | a) | | Imagine you are a data scientist working for a healthcare company. You have a dataset containing patient information and you need to classify whether patients are at risk of developing diabetes. Illustrate the steps used in RIPPER algorithm to build a predictive model & identify the ruleset for the given below dataset. | *CO3* | *PO2* | **10** |

| **PatientID** | **Age** | **BMI** | **Glucose Level** | **Blood Pressure** | **Family History** | **Physical Activity** | **Diabetes Risk** |
|---|---|---|---|---|---|---|---|
| P001 | 45 | 28 | 130 | 85 | Yes | Low | Yes |
| P002 | 30 | 22.5 | 90 | 75 | No | High | No |
| P003 | 60 | 32 | 145 | 88 | Yes | Low | Yes |
| P004 | 25 | 20 | 85 | 70 | No | High | No |
| P005 | 50 | 35 | 155 | 95 | Yes | Low | Yes |
| P006 | 40 | 25 | 100 | 80 | No | Moderate | No |

| | | | | | | |
|---|---|---|---|---|---|---|
| | b) | Consider given a dataset about students and their likelihood of passing a final exam. Construct a decision tree using the CART algorithm to classify whether a student is likely to pass. Perform one iteration of the algorithm, showing the calculations for the Gini Index, selecting the best splitting attribute, and providing the resulting decision tree. | *CO2* | *PO2* | **10** |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | **StudentID** | **Attendance Rate** | **Previous Grades** | **Extracurricular Activities** | **Pass Exam** | | | |
| | | | S101 | High | Good | Yes | Yes | | | |
| | | | S102 | Low | Poor | No | No | | | |
| | | | S103 | Moderate | Average | Yes | Yes | | | |
| | | | S104 | Low | Poor | No | No | | | |
| | | | S105 | Moderate | Average | Yes | Yes | | | |
| | | | S106 | Low | Poor | No | No | | | |
| | | | **OR** | | | | | | | |
| | 4 | a) | Develop a Decision tree model on Wine dataset consists of 178 samples of wine, with 13 features and 3 classes. Demonstrate both pre-pruning and post-pruning techniques for a Decision Tree classifier and print confusion matrix and classification report. | | | | | ***CO4*** | ***PO3*** | **10** |
| | | | Alcohol \| Malic_Acid \| Ash \| Ash_Alcanity \| Magnesium \| Total_Phenols \| Flavanoids \| Nonflavanoid_Phenols \| Proanthocyanins \| Color_Intensity \| Hue \| OD280 \| Proline \| Customer_Segment<br>14.23 \| 1.71 \| 2.43 \| 15.6 \| 127 \| 2.8 \| 3.06 \| 0.28 \| 2.29 \| 5.64 \| 1.04 \| 3.92 \| 1065 \| 1<br>13.2 \| 1.78 \| 2.14 \| 11.2 \| 100 \| 2.65 \| 2.76 \| 0.26 \| 1.28 \| 4.38 \| 1.05 \| 3.4 \| 1050 \| 1<br>13.16 \| 2.36 \| 2.67 \| 18.6 \| 101 \| 2.8 \| 3.24 \| 0.3 \| 2.81 \| 5.68 \| 1.03 \| 3.17 \| 1185 \| 1 | | | |
| | | b) | Given the dataset below, generate the decision tree for 1 iteration<br>Outlook \| Temperature \| Humidity \| Wind \| Played football(yes/no)<br>Sunny \| Hot \| High \| Weak \| No<br>Sunny \| Hot \| High \| Strong \| No<br>Overcast \| Hot \| High \| Weak \| Yes<br>Rain \| Mild \| High \| Weak \| Yes<br>Rain \| Cool \| Normal \| Weak \| Yes<br>Rain \| Cool \| Normal \| Strong \| No<br>Overcast \| Cool \| Normal \| Strong \| Yes<br>Sunny \| Mild \| High \| Weak \| No<br>Sunny \| Cool \| Normal \| Weak \| Yes<br>Rain \| Mild \| Normal \| Weak \| Yes<br>Sunny \| Mild \| Normal \| Strong \| Yes<br>Overcast \| Mild \| High \| Strong \| Yes<br>Overcast \| Hot \| Normal \| Weak \| Yes<br>Rain \| Mild \| High \| Strong \| No | | | ***CO2*** | ***PO2*** | **10** |
| | | | **UNIT - III** | | | | | | | |
| | 5 | a) | Discuss the key features of Bayesian learning methods. Explain how Bayesian learning differs from traditional machine learning methods, and outline the advantages and challenges associated with using a Bayesian approach. | | | | | ***CO2*** | ***PO2*** | **10** |
| | | b) | Consider the following hypothetical data concerning student characteristics and whether or not each student should be hired.<br>Student ID \| Attendance \| GPA \| Effort \| Hired (Class)<br>1 \| Excellent \| Good \| High \| Yes | | | | | ***CO2*** | ***PO2*** | **10** |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | 2 | Poor | Poor | Low | No | | |
| | | | 3 | Excellent | Poor | High | Yes | | |
| | | | 4 | Good | Good | Medium | Yes | | |
| | | | 5 | Poor | Poor | High | No | | |
| | | | 6 | Excellent | Good | High | Yes | | |
| | | | 7 | Good | Poor | Low | No | | |
| | | | Use a Naive Bayes classifier to determine whether or not someone with excellent attendance, poor GPA, and High effort should be hired. | | | | | | |
| | | | **OR** | | | | | | |
| | 6 | a) | Define Bayes' Theorem and explain its components. In a factory, two machines produce parts. Machine 1 produces 70% of the parts, and Machine 2 produces 30% of the parts. The probability that a part produced by Machine 1 is defective is 0.02, and the probability that a part produced by Machine 2 is defective is 0.05. A part is selected at random, and it is found to be defective. Using **Bayes' Theorem**, calculate the probability that the defective part was produced by Machine 1. | | | | | ***CO2*** | ***PO2*** | **10** |
| | | b) | Discuss how Maximum Likelihood and Least Square Error Hypothesis techniques fit into the Bayesian paradigm for parameter estimation with suitable derivations. | | | | | ***CO1*** | | **10** |
| | | | **UNIT - IV** | | | | | | | |
| | 7 | a) | i) Explain the concept of the curse of dimensionality and how it affects machine learning algorithms.<br>ii) Assume you are working with the IRIS dataset and wish to reduce its dimensionality to 2 using Principal Component Analysis (PCA). Write a Python script to:<br>I. Load the IRIS dataset.<br>II. Apply PCA to reduce its dimensions to 2.<br>III. Use an ensemble learning algorithm<br>IV. Evaluate the model's performance on the transformed data by outputting its accuracy and a classification report. | | | | | ***CO4*** | ***PO3*** | **4+6** |
| | | b) | In ensemble learning, various techniques are used to improve model performance by combining multiple models. In this context, the following concepts are important:<br>I. Differentiate between Hard Voting and Soft Voting. For each, provide a code example that demonstrates how they work in a classification problem.<br>II. Differentiate between Bagging and Boosting. Explain the key differences between these two techniques and illustrate each with an example using a classification dataset. | | | | | ***CO1*** | | **10** |
| | | | **OR** | | | | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 8 | a) | Apply the K-Means algorithm on the given dataset, use centroid as (1.0,1.0) & (5.0,7.0) for two clusters.<br>Individual \| Variable 1 \| Variable 2<br>1 \| 1.0 \| 1.0<br>2 \| 1.5 \| 2.0<br>3 \| 3.0 \| 4.0<br>4 \| 5.0 \| 7.0<br>5 \| 3.5 \| 5.0<br>6 \| 4.5 \| 5.0<br>7 \| 3.5 \| 4.5 | *CO2* | *PO2* | **10** |
| | | b) | An ensemble technique in machine learning combines the predictions of multiple models to create a stronger model. One such technique focuses on combining several weak classifiers to produce a strong classifier. Identify and explain the ensemble technique that combines multiple weak classifiers to create a strong classifier. | *CO2* | *PO2* | **10** |
| | | | **UNIT - V** | | | |
| | 9 | a) | Explain the concept of a Perceptron with a neat diagram. Discuss the Perceptron training rule and the conditions under which the Perceptron rule fails. Why does it become necessary to apply the Delta rule in such cases? | *C01* | | **10** |
| | | b) | Explain the Backpropagation Algorithm. How does it learn the weights for a multilayer network? | *CO1* | | **10** |
| | | | **OR** | | | |
| | 10 | a) | i) Derive the Gradient Descent rule and explain the conditions under which it is applied.<br>ii) Differentiate between Gradient Descent and Stochastic Gradient Descent. | *C01* | | **10** |
| | | b) | Consider a simple neural network with the following configuration:<br>• **Input Layer:** 2 neurons (inputs X1 and X2)<br>• **Hidden Layer:** 2 neurons with sigmoid activation.<br>• **Output Layer:** 1 neuron Y=05<br>• **Learning Rate (η)** = 0.1<br>Perform 1 forward pass and 1 backward pass<br>  | *CO2* | *PO2* | **10** |



******