

**QUESTION BANK (UNIT WISE)**

**SUBJECT NAME : MACHINE LEARNING**

**COURSE CODE :**

**YEAR : 2023-2024**

**SEM : ODD**

**UNIT-I**

**PART-B**

| **S.No.** | **Coverage** | **Questions** | **Marks** | **Bloom's**  **Level** |
| --- | --- | --- | --- | --- |
| 1 | UNIT-I | **Dataset:**  | **Size** | **Shape** | **Color** | **Fragrance** | **Type** | **Season** | **Flower** | | --- | --- | --- | --- | --- | --- | --- | | Large | Round | Red | Strong | Perennial | Spring | Yes | | Small | Oval | Yellow | Mild | Annual | Summer | No | | Medium | Round | White | Strong | Perennial | Spring | Yes | | Large | Oval | Blue | Mild | Annual | Fall | No | | Medium | Round | Red | Strong | Perennial | Summer | Yes | | Small | Oval | Green | No | Annual | Fall | No | | Medium | Round | Yellow | Mild | Perennial | Spring | Yes | | Large | Oval | Blue | No | Annual | Winter | No | | Medium | Round | Red | Strong | Perennial | Fall | Yes | | Small | Oval | White | Mild | Annual | Winter | No |   given the dataset of plants. The task is to apply the **Find-S algorithm** to identify the most specific hypothesis that predicts whether a plant is a **Flower** or **Non-Flower** based on the given attributes. | **10** | L2 |
| 2 | UNIT-I | You are given the following dataset of **vehicles**, and the task is to apply the **Candidate Elimination algorithm** to determine the most specific and most general hypotheses that classify a vehicle as either a **"Car"** or **"Non-Car"** based on the given attributes.  **Dataset:**   | **Engine** | **Wheels** | **FuelType** | **Doors** | **Color** | **VehicleType** | | --- | --- | --- | --- | --- | --- | | V6 | 4 | Gasoline | 4 | Red | Car | | V8 | 4 | Diesel | 2 | Blue | Non-Car | | V6 | 4 | Gasoline | 4 | Green | Car | | V6 | 6 | Hybrid | 4 | Black | Non-Car | | V8 | 4 | Gasoline | 2 | White | Non-Car | | V6 | 4 | Gasoline | 4 | Yellow | Car | | V6 | 4 | Electric | 4 | Silver | Car | | V8 | 4 | Gasoline | 2 | Red | Non-Car | | V6 | 4 | Gasoline | 4 | Blue | Car | | V6 | 4 | Diesel | 2 | Green | Non-Car | | **10** | L3 |
| 3 | UNIT-I | **Dataset: (Study Hours vs Pass/Fail)**  | **Student** | **Hours Studied (X)** | **Pass/Fail (Y)** | | --- | --- | --- | | 1 | 29 | Fail | | 2 | 15 | Fail | | 3 | 33 | Pass | | 4 | 28 | Pass | | 5 | 39 | Pass |   Dataset of pass or fail in an exam of 5 students given in a table use logistic regression to answer the below 2 questions and Use Z = -64+2\*hours  1.Calculate the probability of passing for students who studied for 35 hours.  2.At least how many hours should a student study to have a probability of passing greater than 93%? | **10** | L2 |
| 4 | UNIT-I | Apply the **Perceptron Learning Rule** to the **AND Gate** problem with the given initial weights=0, bias=0, and learning rate=0.5. |  |  |

| 5 | UNIT-I | You are provided with a dataset that tracks the **number of hours spent on social media** and the corresponding **sleep hours** for a group of individuals. The goal is to predict the number of hours a person sleeps based on the time spent on social media using **Simple Linear Regression**.   | **Hours Spent on Social Media (X)** | **Sleep Hours (Y)** | | --- | --- | | 0.5 | 8 | | 1 | 7.8 | | 1.5 | 7.5 | | 2 | 7.2 | | 2.5 | 6.8 | | 3 | 6.5 | | 3.5 | 6.2 | | 4 | 6 | | 4.5 | 5.5 | | 5 | 5 |   predict the sleep hours for an individual who spends **4.75 hours** on social media | **10** | L2 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 6 | UNIT-I | | **Men’s Cloth Sales** | **Women’s Cloth Sales** | **Weekly Sales** | | --- | --- | --- | | **1** | **4** | **1** | | **2** | **5** | **6** | | **3** | **8** | **8** | | **4** | **2** | **2** |       Apply the multiple linear regression on the given dataset and  Predict the weekly sales if men cloth sales are 10 and women cloth sales are 12 | **10** | L3 |
| 7 | UNIT-I | **Describe and differentiate between the three main types of machine learning:**   * **Supervised Learning** * **Unsupervised Learning** * **Reinforcement Learning** | **10** | L3 |
| 8 | UNIT-I | **Perspectives and Issues in Machine Learning:**  a) What are some of the ethical and social issues that arise in the field of machine learning? b) Explain the challenges related to data quality, interpretability, and fairness in machine learning models. | **10** | L3 |
| 9 | UNIT-I | **Supervised Learning:**  a) Define supervised learning. Provide an example where supervised learning can be applied. b) What are the key challenges in supervised learning? Discuss with examples. | **10** | L2 |
| 10 | UNIT-I | Discuss the concept of **Overfitting** and **Underfitting** in machine learning models. How can these issues be addressed and Give Example. | **10** | L2 |
| **PART-A** | | | | |
| **S.No.** | **Coverage** | **Questions** | **Marks** | **Bloom's**  **Level** |
| 1 | UNIT-I | Define supervised learning. | 1 | L2 |
| 2 | UNIT-I | What is a neuron in the context of machine learning? | 1 | L1 |
| 3 | UNIT-I | What is the goal of a learning system in machine learning? | 1 | L1 |
| 4 | UNIT-I | What does 'concept learning' aim to achieve? | 1 | L1 |
| 5 | UNIT-I | What is the 'version space' in concept learning? | 1 | L1 |
| 6 | UNIT-I | What is the perceptron algorithm used for? | 1 | L1 |
| 7 | UNIT-I | What does 'linear separability' mean? | 1 | L1 |
| 8 | UNIT-I | What is the primary difference between regression and classification? | 1 | L1 |
| 9 | UNIT-I | State one limitation of the perceptron. | 1 | L1 |
| 10 | UNIT-I | What is the role of weights in the perceptron model? | 1 | L1 |

**UNIT-II**

**PART-B**

| **S.No.** | **Coverage** | **Questions** | **Marks** | **Bloom's**  **Level** |
| --- | --- | --- | --- | --- |
| 1 | UNIT-II | Assume a simple neural network with one output neuron using the **sigmoid activation function**. Perform a **forward pass** and **backward pass** for the network with the following conditions :   * The actual output (target) y5=0.5 * Learning rate 1η=1 | **10** | L2 |
| 2 | UNIT-II | What is the architecture of a Multi-Layer Perceptron (MLP)? Explain how it differs from a simple perceptron. | **10** | L1 |
| 3 | UNIT-II | Given a 2D dataset with points (x1,y1)=(1,2),(x2,y2)=(3,4),(x3,y3)=(5,6) and a Radial Basis Function (RBF) network with 3 centers at (2,3),(4,5),(6,7) compute the output of the RBF network using Gaussian basis functions. | **10** | L2 |
| 4 | UNIT-II | Explain the "curse of dimensionality" in the context of machine learning. How does the increase in dimensionality affect the performance of models ? Discuss the challenges posed by high-dimensional feature spaces in terms of computational cost, model overfitting. | **10** | L2 |
| 5 | UNIT-II | Explain the concepts of interpolation and extrapolation in machine learning. How do these techniques differ, and in what scenarios would each be applied? Discuss the practical applications of both techniques, | **10** | L2 |
| 6 | UNIT-II | Explain the architecture of a Multi-Layer Perceptron (MLP) and its advantage over a simple perceptron. How does it address the problem of non-linearly separable data? | **10** | L2 |
| 7 | UNIT-II | What is a Support Vector Machine (SVM)? Explain how SVM handles classification and regression tasks. In high-dimensional feature spaces. | **10** | L2 |
| 8 | UNIT-II | You are tasked with building a classification model to distinguish between two types of fruits: **Apples** and **Oranges**, based on two features: **weight (in grams)** and **color intensity** (scale from 0 to 1). You have the following data points:   | **Fruit** | **Weight (grams)** | **Color Intensity** | **Class** | | --- | --- | --- | --- | | Apple | 150 | 0.6 | 1 | | Orange | 200 | 0.2 | -1 | | Apple | 130 | 0.5 | 1 | | Orange | 180 | 0.3 | -1 | | Apple | 160 | 0.7 | 1 | | Orange | 190 | 0.1 | -1 | | Apple | 140 | 0.55 | 1 | | Orange | 210 | 0.15 | -1 |   **Classify the following new data point** using the trained linear SVM:   * Weight: 145 grams * Color Intensity: 0.6 | **10** | L2 |
| 9 | UNIT-II | Differentiate between Linear SVM and Non-Linear SVM | **10** | L1 |
| 10 | UNIT-II | What is the curse of dimensionality? How does it affect machine learning models like Radial Basis Function (RBF) networks and Support Vector Machines (SVMs)? | **10** | L2 |
| **PART-A** | | | | |
| **S.No.** | **Coverage** | **Questions** | **Marks** | **Bloom's**  **Level** |
| 1 | UNIT-II | **** What is the function of the activation function in a multi-layer perceptron (MLP)? | 1 | L2 |
| 2 | UNIT-II | **** What is the purpose of the backpropagation algorithm in training neural networks? | 1 | L3 |
| 3 | UNIT-II | **** How does the forward pass in backpropagation differ from the backward pass? | 1 | L2 |
| 4 | UNIT-II | **** What is the role of the hidden layers in a multi-layer perceptron? | 1 | L2 |
| 5 | UNIT-II | **** What is the key idea behind the "curse of dimensionality"? | 1 | L2 |
| 6 | UNIT-II | **** How does a radial basis function (RBF) network differ from a multi-layer perceptron (MLP)? | 1 | L2 |
| 7 | UNIT-II | **** What is the primary function of the "kernel trick" in support vector machines (SVM)? | 1 | L2 |
| 8 | UNIT-II | **** What is the goal of interpolation in machine learning? | 1 | L2 |
| 9 | UNIT-II | **** What is the main advantage of using radial basis functions in machine learning models? | 1 | L2 |
| 10 | UNIT-II | **** What is the significance of support vectors in support vector machines (SVM)? | 1 | L2 |

**UNIT-III**

**PART-B**

| **S.No.** | **Coverage** | **Questions** | **Marks** | **Bloom's**  **Level** |
| --- | --- | --- | --- | --- |
| 1 | UNIT-III | Explain the process of constructing a decision tree. | **10** | L2 |
| 2 | UNIT-III | Compare Bagging and Boosting in ensemble learning. | **10** | L3 |
| 3 | UNIT-III | **Construct a Decision Tree:** A dataset contains the following features for loan approval:   * + Age: Young, Middle-aged, Old   + Income: High, Medium, Low   + Credit Score: Good, Bad   + Class (Target): Approved, Not Approved   Using the following dataset, construct a decision tree using CART algorithm for predicting loan approval. Use the **Gini Index** as the splitting criterion.   | **Age** | **Income** | **Credit Score** | **Class** | | --- | --- | --- | --- | | Young | High | Good | Approved | | Young | Medium | Bad | Not Approved | | Middle-aged | Medium | Good | Approved | | Old | Low | Good | Not Approved | | **10** | L3 |
| 4 | UNIT-III | **Nearest Neighbor Problem:** A dataset contains the following points:   * (2, 3), Class A * (5, 4), Class B * (9, 6), Class A * (4, 7), Class B * (8, 1), Class A   Predict the class for the point (6, 5) using the **k-Nearest Neighbors** algorithm with k=3. Compute distances and show the majority vote. | **10** | L2 |
| 5 | UNIT-III | **K-Means Clustering Problem:** You are tasked with clustering the following points into 2 clusters using the k-means algorithm: Points: (2, 2), (4, 4), (6, 6), (8, 8)   * Initialize centroids at (2, 2) and (8, 8). | **10** | L2 |
| 6 | UNIT-III | Discuss the key differences between supervised learning and unsupervised learning. Provide examples where each approach is suitable. | **10** | L2 |
| 7 | UNIT-III | What are the advantages and limitations of using the k-means algorithm for clustering? | **10** | L2 |
| 8 | UNIT-III | What are the advantages and limitations of using the k-nn algorithm? | **10** | L2 |
| 9 | UNIT-III | What are the different methods to combine multiple classifiers? | **10** | L2 |
| 10 | UNIT-III | What are the challenges of combining multiple classifiers? | **10** | L2 |
| **PART-A** | | | | |
| **S.No.** | **Coverage** | **Questions** | **Marks** | **Bloom's**  **Level** |
| 1 | UNIT-III | **** What is a decision tree used for in machine learning? | 1 | L2 |
| 2 | UNIT-III | **** How does the ID3 algorithm help in constructing decision trees? | 1 | L3 |
| 3 | UNIT-III | **** What is the difference between classification trees and regression trees? | 1 | L2 |
| 4 | UNIT-III | **** What is ensemble learning in machine learning? | 1 | L1 |
| 5 | UNIT-III | **** How does boosting improve the performance of weak classifiers? | 1 | L2 |
| 6 | UNIT-III | **** What is bagging, and how does it help reduce variance in model predictions? | 1 | L2 |
| 7 | UNIT-III | **** What is the purpose of combining classifiers in machine learning? | 1 | L2 |
| 8 | UNIT-III | **** What does a Gaussian Mixture Model (GMM) represent in statistics? | 1 | L1 |
| 9 | UNIT-III | **** What is the k-nearest neighbor (K-NN) algorithm used for? | 1 | L1 |
| 10 | UNIT-III | **** How does the K-means algorithm work in unsupervised learning? | 1 | L2 |

**UNIT-IV**

**PART-B**

| **S.No.** | **Coverage** | **Questions** | **Marks** | **Bloom's**  **Level** |
| --- | --- | --- | --- | --- |
| 1 | UNIT-IV | Describe the working of Locally Linear Embedding (LLE) and Isomap . | **10** | L2 |
| 2 | UNIT-IV | Describe dimensionality reduction techniques and explain their importance in machine learning | **10** | L3 |
| 3 | UNIT-IV | Describe the working of Linear Discriminant Analysis and Principal Component Analysis. | **10** | L2 |
| 4 | UNIT-IV | A Genetic Algorithm is used to optimize a function. The initial population has chromosomes represented as binary strings:   * Parent 1: 10101 * Parent 2: 11010   Apply the following genetic operations:   * **Crossover** at position 3 * **Mutation** at position 4 of the first offspring   Write the resulting offspring. | **10** | L1 |
| 5 | UNIT-IV | Explain the working of a Genetic Algorithm (GA). Discuss its key components | **10** | L2 |

| 6 | UNIT-IV | **Perform Linear Discriminant Analysis (LDA) on the given dataset** You have two classes of data:   * Class A: (2, 3), (3, 3), (2, 4) * Class B: (6, 8), (7, 9), (6, 7) | **10** | L3 |
| --- | --- | --- | --- | --- |
| 7 | UNIT-IV | **Compute the principal component of following data-**  **CLASS 1**   * **X = 2 , 3 , 4** * **Y = 1 , 5 , 3**   **CLASS 2**   * **X = 5 , 6 , 7** * **Y = 6 , 7 , 8** | **10** | L3 |
| 8 | UNIT-IV | Discuss the advantages and disadvantages of using Genetic Algorithms (GAs). | **10** | L2 |
| 9 | UNIT-IV | Explain the concept of Factor Analysis. How does it help in dimensionality reduction | **10** | L2 |
| 10 | UNIT-IV | Discuss various techniques used to reduce the curse of dimensionality in machine learning. | **10** | L2 |
| **PART-A** | | | | |
| **S.No.** | **Coverage** | **Questions** | **Marks** | **Bloom's**  **Level** |
| 1 | UNIT-IV | **** What is dimensionality reduction in machine learning? | 1 | L2 |
| 2 | UNIT-IV | **** What is the purpose of Linear Discriminant Analysis (LDA)? | 1 | L3 |
| 3 | UNIT-IV | **** How does Principal Component Analysis (PCA) help in reducing the dimensionality of data? | 1 | L2 |

| 4 | UNIT-IV | **** What is the main difference between Factor Analysis and PCA? | 1 | L2 |
| --- | --- | --- | --- | --- |
| 5 | UNIT-IV | **** What is Independent Component Analysis (ICA) used for in data processing? | 1 | L2 |
| 6 | UNIT-IV | **** How does Locally Linear Embedding (LLE) work for dimensionality reduction? | 1 | L2 |
| 7 | UNIT-IV | **** What is Isomap, and how does it differ from traditional methods like PCA? | 1 | L1 |
| 8 | UNIT-IV | **** What is the least squares optimization method used for? | 1 | L1 |
| 9 | UNIT-IV | **** What is the role of genetic algorithms in evolutionary learning? | 1 | L1 |
| 10 | UNIT-IV | **** What are genetic operators, and how are they used in genetic algorithms? | 1 | L1 |

**UNIT-V**

**PART-B**

| **S.No.** | **Coverage** | **Questions** | **Marks** | **Bloom's**  **Level** |
| --- | --- | --- | --- | --- |
| 1 | UNIT-V | Explain the concept of Reinforcement Learning (RL). How does it differ from supervised and unsupervised learning? | **10** | L2 |
| 2 | UNIT-V | Explain the steps involved in training a Reinforcement Learning (RL) agent. Discuss the components | **10** | L2 |
| 3 | UNIT-V | Describe how the agent’s decisions are influenced by exploration and exploitation strategies. | **10** | L3 |
| 4 | UNIT-V | Define a Markov Chain. How is it used in MCMC methods for sampling from complex distributions? | **10** | L2 |
| 5 | UNIT-V | Describe the concept of "Markov Chain Monte Carlo" (MCMC). What is its significance in machine learning? | **10** | L2 |
| 6 | UNIT-V | What is the role of "sampling" in MCMC methods? Why is it necessary for approximating in complex models? | **10** | L2 |
| 7 | UNIT-V | Consider a Bayesian Belief Network with the following structure:   * **Node A**: Weather (Rain or No Rain) * **Node B**: Person carrying an umbrella (Yes or No) * **Node C**: Person gets wet (Yes or No)   The conditional probability tables (CPTs) are as follows:   * P(A=Rain)=0.6 and P(A=No Rain)=0.4 * P(B=Yes∣A=Rain)=0.8, P(B=Yes∣A=No Rain)=0.3 * P(C=Yes∣B=Yes)=0.1, P(C=Yes∣B=No)=0.7 * **Tasks:**  1. Calculate the probability that the person gets wet, i.e., P(C=Yes) 2. Calculate the joint probability P(A=Rain, B=Yes, C=Yes). | **10** | L2 |
| 8 | UNIT-V | Consider a Bayesian Belief Network with three variables: **X** (Weather), **Y** (Person goes out), and **Z** (Person carries an umbrella). The following conditional probability tables (CPTs) are provided:   * P(X=Sunny)=0.5, P(X=Cloudy)=0.3, P(X=Rainy)=0. 2 * P(Y=Yes∣X=Sunny)=0.8, P(Y=Yes∣X=Cloudy)=0.6, P(Y=Yes∣X=Rainy)=0.4 * P(Z=Yes∣Y=Yes,X=Rainy)=0.9, P(Z=Yes∣Y=Yes,X=Cloudy)=0.4 , P(Z=Yes∣Y=No)=0.   **Tasks:**   1. Calculate the probability that the person carries an umbrella, P(Z=Yes) 2. Calculate the joint probability P(X=Rainy,Y=Yes,Z=Yes) | **10** | L2 |
| 9 | UNIT-V | Explain the concept of **Reinforcement Learning (RL)**. How does it differ from supervised and unsupervised learning? Provide a real-world example to illustrate the working of RL | **10** | L2 |
| 10 | UNIT-V | Discuss the working principles of **Markov Chain Monte Carlo (MCMC) Methods**. Highlight their significance in RL with an example. | **10** | L3 |
| **PART-A** | | | | |
| **S.No.** | **Coverage** | **Questions** | **Marks** | **Bloom's**  **Level** |
| 1 | UNIT-V | **** What is reinforcement learning in the context of machine learning? | 1 | L2 |
| 2 | UNIT-V | **** What does the "Getting Lost" example demonstrate in reinforcement learning? | 1 | L2 |
| 3 | UNIT-V | **** What is the role of exploration and exploitation in reinforcement learning? | 1 | L2 |
| 4 | UNIT-V | **** What is a Markov Chain Monte Carlo (MCMC) method used for in statistics? | 1 | L2 |
| 5 | UNIT-V | **** What is the purpose of a proposal distribution in Markov Chain Monte Carlo methods? | 1 | L2 |
| 6 | UNIT-V | **** How does Markov Chain Monte Carlo sampling work in practice? | 1 | L2 |
| 7 | UNIT-V | **** What are graphical models in machine learning? | 1 | L2 |
| 8 | UNIT-V | **** What is the primary purpose of Bayesian networks in probabilistic modeling? | 1 | L2 |
| 9 | UNIT-V | **** How do Markov Chains help in MCMC sampling methods? | 1 | L2 |
| 10 | UNIT-V | **** What is the difference between a Bayesian network and a general graphical model? | 1 | L2 |

**\*\*\*\*\*\***