



Q.1

a. Define Machine Learning. What are the challenges? (3 Marks)

Machine Learning (ML):

ML is a field of artificial intelligence that enables systems to learn patterns from data and make predictions or decisions without being explicitly programmed.

Challenges:

1. Insufficient/poor quality data
2. Overfitting & underfitting
3. Choosing right algorithms & parameters
4. High computational cost
5. Data privacy & security issues
6. Interpretability of models

b. What is debugging in Machine Learning? Challenges? (7 Marks)

Debugging in ML:

The process of identifying and fixing issues in data, model, features, or training process to improve model performance.

Challenges in ML Debugging:

1. Non-deterministic behavior due to randomness
2. Large and noisy datasets
3. Data leakage
4. Complexity in feature engineering
5. Hyperparameter tuning difficulty
6. Model explainability issues
7. Bias and variance trade-off

c. Briefly explain types of ML. (10 Marks)

1. **Supervised Learning**
 - Labeled data
 - Tasks: Classification, Regression
2. **Unsupervised Learning**
 - Unlabeled data

- Tasks: Clustering, Association
- 3. **Semi-supervised Learning**
 - Mix of labeled + unlabeled
 - Used when labeling is expensive
- 4. **Reinforcement Learning**
 - Learning by interaction & rewards
 - Applications: Robotics, gaming
- 5. **Deep Learning**
 - Multi-layer neural networks
 - Used for images, speech, NLP

Q.2

a. Need of Python in Machine Learning (3 Marks)

- Large ML libraries: **NumPy, Pandas, Scikit-learn, TensorFlow, PyTorch**
- Easy to write & understand
- Strong community support
- Integrates well with data processing tools
- Rich visualization tools: **Matplotlib, Seaborn**

b. What is Regression? Explain two models. (7 Marks)

Regression:

A supervised learning technique that predicts continuous values.

Two models:

1. **Linear Regression**
 - Linearly relates dependent & independent variables.
 - Equation: $y = mx + c$
2. **Polynomial Regression**
 - Fits a nonlinear curve.
 - Example: $y = a + b_1x + b_2x^2 + \dots$

c. Define Perceptron. Explain multilayer perceptron (MLP) with activation function. (10 Marks)

Perceptron:

Basic unit of a neural network performing weighted sum + activation.

MLP:

- Multilayer network: input → hidden layers → output
- Uses **backpropagation**

- Learns complex nonlinear patterns

Activation functions:

- **Sigmoid:** $1/(1+e^{-x})$
- **ReLU:** $\max(0, x)$
- **Tanh**

Q.3

a. Interpret ANN with applications. (3 Marks)

ANN mimics the human brain using neurons and weighted connections. Learns patterns for:

- Image recognition
- Fraud detection
- Speech processing
- Forecasting

b. Support Vector Machine (SVM). Explain Gradient Descent. (7 Marks)

SVM:

- Finds optimal separating hyperplane
- Maximizes margin
- Works well with high-dimensional data

Gradient Descent:

- Optimization method
- Iteratively updates parameters to minimize loss
- Update rule:
 $\theta = \theta - \alpha \times (\text{derivative of loss})$

c. Define Decision Tree. Steps in ID3. (10 Marks)

Decision Tree:

A tree-structured model used for classification/regression.

Steps in ID3:

1. Calculate Entropy for dataset
2. Compute Information Gain for each attribute
3. Choose attribute with highest Gain as root
4. Split dataset
5. Recursively build tree
6. Stop when data is pure or no attributes left

Q.4

a. Types of crossover in Genetic Algorithm (3 Marks)

1. Single-point crossover
2. Two-point crossover
3. Uniform crossover
4. Arithmetic crossover

b. Issues in Decision Tree Learning (any 3) (7 Marks)

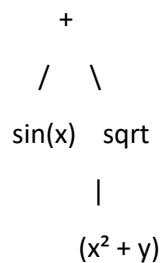
1. Overfitting
2. Bias for attributes with many values
3. High variance
4. Missing values
5. Computational cost for large data

c. Explain GA with steps. Genetic program tree for $\sin(x)+\sqrt{x^2+y}$. (10 Marks)

Steps of GA:

1. Initialization of population
2. Fitness evaluation
3. Selection
4. Crossover
5. Mutation
6. Termination

Genetic Tree:



Q.5

a. Boosting vs Bagging (3 Marks)

Bagging	Boosting
Parallel training	Sequential training
Reduces variance	Reduces bias
Example: Random Forest	Example: AdaBoost, XGBoost

b. Expectation Maximization (EM). Applications. (7 Marks)

EM is an iterative algorithm to estimate parameters of models with hidden variables.

Steps:

- **E-Step:** Estimate missing values
- **M-Step:** Maximize likelihood

Applications:

- Gaussian Mixture Models
- Missing data estimation
- Clustering
- Image restoration

c. Explain Ensemble Techniques. (10 Marks)

1. Bagging
2. Boosting
3. Stacking
4. Random Forest
5. Voting classifier

Q.6**a. Markov Decision Process (3 Marks)**

MDP is a mathematical framework for decision-making with:

- **States**
- **Actions**
- **Transition probabilities**
- **Reward function**

Used in reinforcement learning.

b. Q-Learning algorithm & applications (7 Marks)**Q-Learning:**

Model-free RL algorithm that learns optimal action-value function.

Update rule:

$$Q(s,a)=Q(s,a)+\alpha[R + \gamma \max_{a'} Q(s',a') - Q(s,a)]$$

Applications:

- Robotics
- Game playing
- Traffic control
- Recommendation systems

c. Explain Q-Learning and Q-Function with example (10 Marks)**Q-Function:**

Measures expected future rewards for taking action a in state s .

Example:

Robot navigating grid:

If moving right from state $S1$ gives reward $+5$, $Q(S1, \text{right})$ increases.

Table form:

State	Action	Q-value
S1	Right	5
S1	Up	2

Algorithm updates Q-values until optimal policy is learned.

Q.7 (c) Visualization techniques for AR (10 Marks)

1. 3D Model Visualization

- Displaying virtual 3D objects on real surfaces
- Used in education, medical training

2. Object Tracking Visualization

- Tracks real objects and overlays data (e.g., car part info in automobile AR manuals)

3. Spatial Visualization

- Visualizes digital content anchored to real environment
- Used in interior design apps

4. Gesture-Based Visualization

- User interacts using hand gestures
- Microsoft HoloLens example

5. Interactive Animation Visualization

- Animated digital objects blended with real environment

6. Projection Visualization

- Project digital images on physical objects for AR experience

7. Layer-Based Visualization

- Displays contextual data layers (maps, labels, directions)

Q.8 CASE STUDY – Decision Tree

Dataset: *Play Tennis (14 days)*

Q.8 (a) Explain the process of constructing a decision tree. (5 Marks)

Steps to construct a Decision Tree (ID3 Algorithm):

1. Calculate Entropy of Target Variable (Play Tennis)
2. Calculate Entropy of each attribute
3. Compute Information Gain for each attribute
4. Select attribute with highest Gain as root node
5. Split data according to the chosen attribute
6. Repeat steps recursively for each branch
7. Stop when:
 - All samples belong to same class
 - No attributes left
 - No data left

Q.8 (b) Calculate Gain for Humidity and Wind (10 Marks)

Step 1: Total Entropy (Play Tennis)

Total Yes = 9

Total No = 5

Total = 14

$$\begin{aligned} \text{Entropy}(S) &= -\frac{9}{14} \log_2 \frac{9}{14} - \frac{5}{14} \log_2 \frac{5}{14} \\ \text{Entropy}(S) &= 0.94 \end{aligned}$$

1. INFORMATION GAIN OF HUMIDITY

Split: High and Normal

Humidity = High

Total = 7

Yes = 2

No = 5

$$\text{Entropy} = -\frac{2}{7} \log_2 \frac{2}{7} - \frac{5}{7} \log_2 \frac{5}{7} = 0.863$$

Humidity = Normal

Total = 7

Yes = 7

No = 0

$$\text{Entropy} = 0$$

Weighted Entropy:

$$E(\text{Humidity}) = \frac{7}{14} (0.863) + \frac{7}{14} (0) = 0.4315$$

Information Gain:

$$\text{Gain}(\text{Humidity}) = 0.94 - 0.4315 = 0.5085$$

2. INFORMATION GAIN OF WIND

Split: Weak and Strong

Wind = Weak

Total = 8

Yes = 6

No = 2

$$\text{Entropy} = -\frac{6}{8} \log_2 \frac{6}{8} - \frac{2}{8} \log_2 \frac{2}{8} = 0.811$$

Wind = Strong

Total = 6

Yes = 3

No = 3

$$Entropy = 1$$

Weighted Entropy:

$$E(Wind) = \frac{8}{14}(0.811) + \frac{6}{14}(1) = 0.892$$

Information Gain:

$$Gain(Wind) = 0.94 - 0.892 = 0.048$$

Final Result:

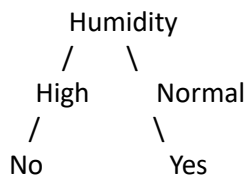
Attribute	Information Gain
Humidity	0.5085
Wind	0.048

Humidity gives the highest gain.

Q.8 (c) Construct the complete decision tree (5 Marks)

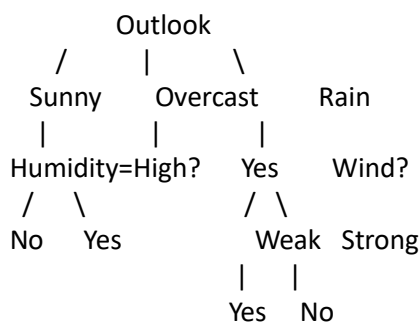
Step 1: Choose the attribute with highest gain → Humidity

Decision Tree:



But we must consider remaining attributes for other splits → Use Outlook for High branch.

Final Tree:



Hence,

- **Overcast → Yes**
- **Sunny → depends on Humidity:**
 - High → No
 - Normal → Yes
- **Rain → depends on Wind:**
 - Weak → Yes
 - Strong → No