MACHINE LEARNING QUESTION

1. What is Machine Learning?

Machine Learning is a branch of Artificial Intelligence that enables a computer system to learn from data and improve its performance without being explicitly programmed.
It focuses on building models that can make predictions or decisions based on data.

---

2. Define Supervised Learning.

Supervised learning is a type of machine learning where the model is trained using labeled data.
Each input data has a corresponding output, and the model learns by comparing its predictions with the correct answers.

---

3. Define Unsupervised Learning.

Unsupervised learning is a type of machine learning where the model is trained using unlabeled data.
The algorithm discovers patterns, structures, or groups in the data without any predefined output.

---

4. What is Reinforcement Learning?

Reinforcement learning is a type of machine learning where an agent learns by interacting with an environment.
The agent receives rewards or penalties based on its actions and learns to maximize the total reward.

---

5. What is a Dataset?

A dataset is a collection of related data used for training and testing a machine learning model.
It consists of input features and may also include output labels depending on the learning type.

---

1. Define Bias and Variance.

Bias is the error caused by incorrect or overly simple assumptions in a learning model, leading to underfitting.
Variance is the error caused by a model being too complex and sensitive to training data, leading to overfitting.

---

2. What is Cross-Validation?

Cross-validation is a technique used to evaluate the performance of a machine learning model.
In this method, the dataset is divided into multiple parts, and the model is trained and tested on different combinations to ensure better accuracy.

---

3. What is an Algorithm?

An algorithm is a step-by-step procedure or set of rules used to solve a problem or perform a task.
In machine learning, algorithms are used to learn patterns from data and make predictions.

1. Explain Machine Learning and its Applications.(5MARKS)

Machine Learning (ML) is a branch of Artificial Intelligence that enables a computer system to learn from data and improve its performance without being explicitly programmed.
It focuses on creating algorithms that identify patterns in data and make predictions or decisions.

In machine learning, models are trained using past data (training data) and then used to predict results on new data. ML reduces human effort and is widely used where manual programming is difficult.

Applications of Machine Learning:

- Spam email detection
- Recommendation systems (Netflix, Amazon)
- Medical diagnosis

- Fraud detection
- Speech and face recognition

---

2. Explain Types of Machine Learning.

Machine Learning is mainly divided into **three types**:

(a) Supervised Learning

In supervised learning, the model is trained using labeled data, where input and output are known.
Examples include **classification** and **regression** problems.

Example: Email spam detection, student marks prediction.

(b) Unsupervised Learning

In unsupervised learning, the model works with unlabeled data and finds hidden patterns or groups in data.

Example: Customer segmentation, clustering.

(c) Reinforcement Learning

In reinforcement learning, an agent learns by interacting with an environment and receiving rewards or penalties.

Example: Game playing, robot navigation.

---

3. What is Overfitting and Underfitting? Explain with Diagrams.

Overfitting

Overfitting occurs when a model learns the training data too well, including noise, and performs poorly on new data.

Underfitting

Underfitting occurs when a model is too simple and cannot capture the underlying pattern of the data.

Diagram Explanation:

- Underfitting → straight/simple line
- Good fit → smooth curve
- Overfitting → complex zig-zag curve

---

4. Explain Performance Evaluation Metrics.

Performance metrics are used to measure how well a machine learning model performs.

Accuracy

Accuracy is the ratio of correctly predicted observations to total observations.

Precision

Precision measures how many predicted positive results are actually positive.

Recall

Recall measures how many actual positive results are correctly predicted.

F1-Score

F1-score is the harmonic mean of precision and recall and gives balanced performance.

---

5. Explain Linear Regression Algorithm.

Linear Regression is a supervised learning algorithm used to predict a continuous output variable based on one or more input variables.
It assumes a linear relationship between input and output.

The equation is:
y = mx + c

It works by minimizing the error between predicted and actual values using techniques like gradient descent.

Example: Predicting house price based on area.

---

6. Explain Logistic Regression Algorithm.

Logistic Regression is a supervised learning algorithm used for classification problems.
It predicts the probability of an event using a sigmoid function.

The output is between **0 and 1**, making it suitable for binary classification.

Example: Disease detection (Yes/No), Spam detection.

---

7. Explain K-Nearest Neighbors (KNN).

KNN is a supervised learning algorithm used for classification and regression.
It classifies a data point based on the majority class of its nearest neighbors.

The value of **K** determines how many neighbors are considered.

Example: Handwriting recognition, recommendation systems.

---

8. Explain Decision Tree Algorithm.

Decision Tree is a supervised learning algorithm used for classification and regression.
It works by splitting data into branches based on conditions until a decision is reached.

Each internal node represents a test, and each leaf node represents an output.

Example: Loan approval system.

9. Explain K-Means Clustering Algorithm.

K-Means is an unsupervised learning algorithm used for clustering data into **K groups**.
It works by assigning data points to the nearest cluster center and updating the centers repeatedly.

The goal is to minimize the distance between data points and cluster centroids.

Example: Customer segmentation.