Q&A- AI and ML

1. What is **Machine Learning**?

A subset of Artificial Intelligence (AI) that enables systems to learn from data and improve without explicit programming.

Applications: Used in recommendation systems, fraud detection, speech recognition, self-driving cars, and more.

2. Give one real-life example of **Machine Learning**.

A real-life example of Machine Learning is **email spam filtering**.

3. What is the difference between **Artificial Intelligence (AI)** and **Machine Learning (ML)**?

Artificial Intelligence (AI):

- Broad field of creating intelligent systems.
- Performs tasks like reasoning, decision-making, and problem-solving.
- Includes various approaches: rule-based systems, expert systems, robotics, NLP, and ML.

Machine Learning (ML):

- Subset of AI focused on learning from data.
- Uses algorithms to find patterns and improve over time.
- Types: supervised, unsupervised, reinforcement learning.
- Powers applications like spam filters, recommendations, and fraud detection.

4. What are the **types of Machine Learning**?

a) Supervised Learning
b) Unsupervised Learning
c) Reinforcement Learning

5. What is **supervised learning**? Give an example.

Supervised Learning is a type of ML where models learn from labeled data (inputs with known outputs) to make predictions.

- **Example:** Predicting house prices from features like size, location, and bedrooms.

6. What is **unsupervised learning**? Give an example.

Unsupervised Learning is a type of ML where models work on **unlabeled data** to find hidden patterns or groupings without predefined outputs.

- **Example:** Customer segmentation in marketing, where customers are grouped based on buying behavior.

7. What is **reinforcement learning**? Give an example.

Reinforcement Learning is a type of ML where an agent learns by interacting with an environment, making decisions, and receiving feedback in the form of rewards or penalties.

- **Example:** Training a self-driving car to stay in its lane by rewarding safe driving actions and penalizing mistakes.

8. What is the difference between **training data** and **test data**?

Training Data

- Used to **train** the model.
- Contains input-output pairs so the model can learn patterns and relationships.
- Example: Past emails labeled as *spam* or *not spam*.

Test Data

- Used to **evaluate** the model's performance.
- Contains new, unseen data to check accuracy and generalization.
- Example: New emails the model has never seen before.

9. What are **features** in a dataset?

Features in a dataset are the **individual measurable properties or variables** used as input to a machine learning model. They represent the characteristics that help the model make predictions.

- **Example:** In a house price dataset:
  - Features → size, number of bedrooms, location.
  - Target (label) → house price.

10. What are **labels (targets)** in a dataset?

Labels (or targets) in a dataset are the **outputs or outcomes** that a machine learning model is trained to predict. They are the answers associated with the input features.

- **Example:** In a house price dataset:
  - Features → size, bedrooms, location.
  - **Label (target) → actual house price.**

11. What is the difference between **classification** and **regression**?

Classification

- Predicts **categories or classes**.
- Output is **discrete** (e.g., Yes/No, Spam/Not Spam).
- Example: Predicting if an email is spam.

Regression

- Predicts **continuous numerical values**.
- Output is **continuous** (e.g., price, temperature).
- Example: Predicting house prices.

12. What does the term **model** mean in ML?

In **Machine Learning**, a **model** is the **mathematical representation of the real-world process** that has been trained on data to recognize patterns and make predictions or decisions.

- It is the output of the learning algorithm after training.
- **Example:** A spam detection model that decides whether a new email is *spam* or *not spam* based on learned patterns.

13. What is **overfitting** in ML?

Overfitting is when a model learns the training data too closely (including noise), leading to high training accuracy but poor performance on new, unseen data.

14. What is **underfitting** in ML?

Underfitting occurs when a model is too simple to capture the underlying patterns in the data, resulting in **poor performance on both training and test data**.

15. What is a **confusion matrix** used for?

A **confusion matrix** is used to evaluate a classification model by showing true positives, true negatives, false positives, and false negatives, helping calculate metrics like accuracy and precision.

16. Why do we split data into **training and testing sets**?

We split data into **training and testing sets** to:

- **Train the model** on one portion (training set) so it can learn patterns.
- **Evaluate performance** on unseen data (testing set) to check **generalization**.
- **Prevent overfitting**, ensuring the model works well on new data, not just the training data.

17. What is a **decision tree**?

Definition: Supervised ML model that splits data based on feature values.

Structure: Nodes (features), branches (decisions), leaves (outcomes).

Types: Classification or regression.

Example: Predicting if a student will pass or fail based on study hours and attendance.

18. What is a **linear regression model** used for?

A Linear Regression model is used to predict a continuous numerical outcome based on one or more input features by fitting a straight line to the data.

- **Example: Predicting house prices from size, number of bedrooms, and location.**

19. What does **accuracy** mean in ML?

Accuracy in Machine Learning measures the **proportion of correct predictions** made by a model out of all predictions.

- **Formula:**

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100$$

  ✔ TP = True Positives
  ✔ TNTNTN = True Negatives
  ✔ FPFPFP = False Positives
  ✔ FNFNFN = False Negatives

- **Example:** If a spam filter correctly classifies 90 out of 100 emails, its accuracy = 90%

20. What are some common **applications of Machine Learning** in daily life?

- **Recommendation Systems:** Netflix, YouTube, Amazon suggesting movies, videos, or products.
- **Spam Detection:** Filtering unwanted emails in Gmail or Outlook.
- **Voice Assistants:** Siri, Alexa, Google Assistant understanding and responding to commands.
- **Fraud Detection:** Identifying suspicious transactions in banking.
- **Self-driving Cars:** Autonomous vehicles using ML to navigate safely.
- **Image and Face Recognition:** Unlocking phones or tagging photos on social media.
- **Healthcare:** Predicting diseases or assisting in medical diagnosis.