Week 1

AICTE Internship ID : INTERNSHIP\_174099535967c57b1f336c3

AICTE Internship student Registration ID : STU67b951e7e808e1740198375

**1. What is Machine Learning?**

**Machine Learning (ML)** is a subset of Artificial Intelligence (AI) focused on enabling systems to learn from data and make decisions or predictions without being explicitly programmed.

**Definition:**

Machine Learning is the science of developing algorithms and models that allow computers to learn from and make decisions based on data.

**Why Machine Learning?**

Real-world problems are often too complex to solve with fixed rules. ML provides a data-driven approach where computers learn the patterns from data instead of relying on hardcoded logic.

**Types of Machine Learning:**

1. **Supervised Learning** – Uses labeled data.
2. **Unsupervised Learning** – Finds patterns in unlabeled data.
3. **Reinforcement Learning** – Learns through interaction with an environment using reward-based feedback.

**Real-life Examples:**

* **Spam Detection**: Gmail filters spam emails using ML models.
* **Movie Recommendations**: Netflix suggests movies based on your viewing history.
* **Voice Assistants**: Siri, Alexa understand and respond using speech recognition models.
* **Healthcare**: Predicting disease from medical images or patient history.

**2. What is a Supervised Machine Learning Algorithm?**

**Definition:**

Supervised Learning is a type of ML where the algorithm is trained using labeled input-output pairs. The goal is to learn a function that maps inputs to outputs.

**How It Works:**

* The model is trained on a dataset that includes inputs and correct outputs (labels).
* It learns the relationship between input features and the label.
* Then it can predict outputs for new, unseen data.

**Components:**

* **Features (X)**: Input data
* **Labels (Y)**: Output or target values
* **Model**: The mathematical representation that maps X to Y
* **Loss Function**: Measures prediction error
* **Optimization Algorithm**: Minimizes error by adjusting model parameters

**Example:**

**House Price Prediction**

| **Size (sqft)** | **Bedrooms** | **Price (₹)** |
| --- | --- | --- |
| 1000 | 2 | 30 lakhs |
| 1500 | 3 | 45 lakhs |
| 2000 | 4 | 60 lakhs |

* Input: Size, Bedrooms
* Label: Price
* Task: Predict price for a new house

**Common Supervised Algorithms:**

* **Linear Regression**
* **Logistic Regression**
* **Decision Trees**
* **Support Vector Machines (SVM)**
* **Random Forests**
* **k-Nearest Neighbors (k-NN)**

**3. What is Regression and classification?**

**Regression:**

Regression is a type of supervised learning used to predict continuous numeric values.

**Goal:**

Estimate the relationship between input variables (features) and a continuous output.

**Examples:**

* Predicting **house prices** based on size and location
* Forecasting **sales** or **stock prices**
* Estimating **salary** based on years of experience

**Example:**

| **Experience (Years)** | **Salary (₹ in lakhs)** |
| --- | --- |
| 1 | 3 |
| 2 | 4.5 |
| 3 | 6 |

**Prediction Task**: Estimate salary for 5 years of experience.

**Common Regression Algorithms:**

* Linear Regression
* Polynomial Regression
* Ridge/Lasso Regression
* Support Vector Regression (SVR)
* Decision Trees (for regression)

**Classification:**

Classification is a type of supervised learning used to predict categorical labels or classes.

**Goal:**

Assign a category to each input based on past labeled examples.

**Examples:**

* Spam detection (Spam/Not Spam)
* Disease classification (Positive/Negative)
* Customer segmentation (High/Medium/Low value)
* Image recognition (Cat/Dog/Bird)

**Example:**

| **Age** | **Salary** | **Purchased Product** |
| --- | --- | --- |
| 25 | 40k | No |
| 35 | 70k | Yes |
| 45 | 90k | Yes |

**Prediction Task**: Will a 30-year-old earning 60k buy the product?

**Common Classification Algorithms:**

* Logistic Regression
* k-Nearest Neighbors (k-NN)
* Decision Trees
* Random Forest
* Support Vector Machines (SVM)
* Naïve Bayes