**Week 1 - Assessment**

**1]What is Machine Learning (ML)?**

**Machine Learning (ML)** is a subset of artificial intelligence (AI) that enables computers to learn from data and make decisions or predictions without being explicitly programmed for specific tasks. Instead of following hard-coded rules, ML algorithms identify patterns in data and improve their performance over time with experience.

**Key Characteristics:**

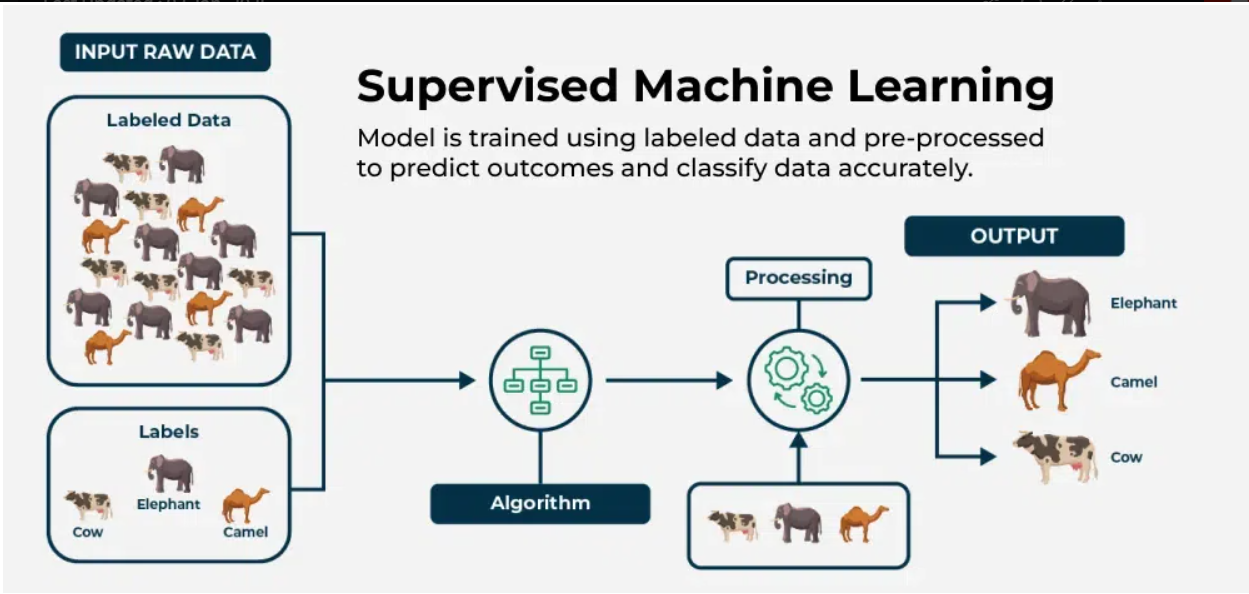
* **Learning from Data:** ML models improve as they are exposed to more data.
* **Pattern Recognition:** They identify patterns and relationships within data.
* **Adaptability:** ML systems can adapt to new, unseen data.

**Real-World Applications:**

* **Email Filtering:** Automatically classifying emails as spam or not spam.
* **Recommendation Systems:** Suggesting products or content based on user behavior.
* **Image Recognition:** Identifying objects or faces in photos.
* **Medical Diagnosis:** Assisting in detecting diseases from medical images.

**2]What is Supervised Machine Learning?**

**Supervised Learning** is a type of machine learning where the model is trained on a labeled dataset. This means each training example is paired with an output label. The model learns to predict the output from the input data.



**How It Works:**

1. **Training Phase:** The model is provided with input-output pairs and learns the mapping between them.
2. **Prediction Phase:** Once trained, the model can predict outputs for new, unseen inputs.

**Analogy:** Think of supervised learning as learning with a teacher. The teacher provides the correct answers during training, and the student (model) learns to generalize from these examples.

**Common Algorithms:**

* **Linear Regression**
* **Logistic Regression**
* **Decision Trees**
* **Support Vector Machines (SVM)**
* **Neural Networks**

**Applications:**

* **Credit Scoring:** Predicting the likelihood of a borrower defaulting on a loan.
* **Speech Recognition:** Transcribing spoken words into text.
* **Medical Diagnosis:** Predicting diseases based on patient data.

***Supervised ML Algorithm***

**How It Works – Step-by-Step**

1. **Input Data**: You provide data with input features (e.g., height, weight, age).
2. **Labels**: Each input also has a correct output label (e.g., disease = "Yes" or "No").
3. **Training**: The algorithm uses the inputs and labels to find patterns or relationships.
4. **Model Building**: It creates a model — like a mathematical equation or decision path.
5. **Prediction**: When you give the model new input data (without labels), it predicts the output.
6. **Evaluation**: You test how accurate the model is using a separate set of labeled data (called the **test set**).

**Real-World Examples**

* **Email Spam Filter**
  + Inputs: Words in the email, sender’s address, etc.
  + Label: “Spam” or “Not Spam”
* **Loan Approval System**
  + Inputs: Income, Credit Score, Loan Amount
  + Label: “Approved” or “Rejected”
* **Medical Diagnosis**
  + Inputs: Symptoms, test results
  + Label: “Disease present” or “Disease absent”

**Common Supervised ML Algorithms**

Here are some of the most-used algorithms in supervised learning:

|  |  |  |
| --- | --- | --- |
| Algorithm | Used For | Example Use Case |
| Linear Regression | Regression (continuous) | Predicting house prices |
| Logistic Regression | Classification | Diagnosing diseases |
| Decision Trees | Both | Credit scoring, Fraud detection |
| Random Forest | Both | Product recommendation |
| Support Vector Machine | Both | Image recognition, Text classification |
| k-Nearest Neighbors | Both | Handwriting recognition |
| Neural Networks | Both | Voice recognition, Self-driving cars |

Note: Some algorithms can be adapted for both regression and classification tasks.

***Regression vs. Classification***

Both regression and classification are types of supervised learning, but they serve different purposes.

**Regression**

**Purpose:** Predict continuous numerical values.

**Examples:**

* Predicting house prices based on features like size and location.
* Forecasting stock prices.
* Estimating temperature for the next week.

**Common Algorithms:**

* Linear Regression
* Ridge Regression
* Lasso Regression
* Support Vector Regression (SVR)

**Classification**

**Purpose:** Predict categorical labels.

**Examples:**

* Determining if an email is spam or not.
* Classifying images as cats or dogs.
* Diagnosing diseases as present or absent.

**Common Algorithms:**

* Logistic Regression
* Decision Trees
* Random Forest
* Support Vector Machines (SVM)
* K-Nearest Neighbors (KNN)

**Comparison Table: Regression vs. Classification**

|  |  |  |
| --- | --- | --- |
| Feature | Regression | Classification |
| **Output Type** | Continuous numerical values | Discrete categories or labels |
| **Example Output** | 250.75, 98.6, -15.0 | "Spam", "Not Spam", "Approved", "Denied" |
| **Goal** | Predict exact values | Assign items to categories |
| **Evaluation Metrics** | Mean Squared Error, R² Score | Accuracy, Precision, Recall, F1 Score |
| **Algorithms** | Linear Regression, SVR | Logistic Regression, SVM, Decision Trees |
| **Use Cases** | Price prediction, Weather forecasting | Email filtering, Disease diagnosis |