Machine Learning Training Session Outline

1. Introduction to Machine Learning

What is Machine Learning?

- Machine Learning (ML) is a subset of **Artificial Intelligence (AI)** that enables systems to **learn from data** and improve their performance **without being explicitly programmed**.
- Instead of writing hard-coded rules, we **train models** on data to make predictions or decisions.

Why Machine Learning?

- **Automation:** ML automates decision-making (e.g., spam detection, recommendation systems).
- **Handling Complex Data:** Finds patterns in large datasets where manual analysis is impossible.
- **Adaptability:** Models improve over time with new data.

Example Applications:

- **Supervised Learning:** Email spam detection, house price prediction.
- **Unsupervised Learning:** Customer segmentation, anomaly detection.
- **Reinforcement Learning:** Self-driving cars, game-playing AI (AlphaGo).

---

2. Types of Machine Learning

1. Supervised Learning (Learning from Labeled Data)

- The model learns from **input-output pairs** (X → y).

- **Goal:** Predict the output for new, unseen data.
- **Types:**

  o **Regression (Continuous Output):** Predict house prices, stock prices.

  o **Classification (Discrete Output):** Spam vs. not spam, image recognition.

2. Unsupervised Learning (Finding Patterns in Unlabeled Data)

- The model learns **hidden structures** from data without labels.
- **Goal:** Discover patterns, groupings, or anomalies.
- **Types:**

  o **Clustering:** Grouping similar customers (e.g., K-Means).

  o **Dimensionality Reduction:** Reducing features (e.g., PCA).

  o **Anomaly Detection:** Fraud detection.

3. Reinforcement Learning (Learning by Trial & Error)

- The model learns by **interacting with an environment** and receiving **rewards/punishments**.
- **Goal:** Maximize cumulative reward (e.g., game AI, robotics).

4. Semi-Supervised & Self-Supervised Learning (Brief Mention)

- Uses a mix of labeled and unlabeled data (common in real-world scenarios).

---

3. How Machine Learning Works

Key Steps in ML:

1. **Data Collection:** Gather relevant data (structured/tabular, images, text).
2. **Data Preprocessing:** Clean, normalize, handle missing values (using Pandas, NumPy).
3. **Feature Engineering:** Select/extract meaningful features (e.g., converting text to numbers).

4. **Model Selection:** Choose an algorithm (Linear Regression, Decision Trees, Neural Networks).
5. **Training:** Fit the model on training data (adjusting weights to minimize error).
6. **Evaluation:** Test on unseen data (using metrics like accuracy, MSE).
7. **Deployment:** Use the model in real-world applications.

Example Workflow (House Price Prediction):

- **Input Features (X):** Size, location, bedrooms.
- **Output (y):** Price.
- **Model:** Linear Regression → Predicts price for new houses.

---

4. Common Machine Learning Algorithms

Supervised:

- **Regression:** Linear Regression, Decision Trees.
- **Classification:** Logistic Regression, SVM, Random Forest.

Unsupervised:

- **Clustering:** K-Means, DBSCAN.
- **Dimensionality Reduction:** PCA, t-SNE.

Deep Learning (Brief Intro):

- Neural Networks for complex tasks (image, speech recognition).