Supervised vs Unsupervised Learning: Notes

1. Definition

- Supervised Learning:
  - Model learns from labeled data (input-output pairs).
  - Goal: Predict outcomes for new, unseen data.

- Unsupervised Learning:
  - Model learns from unlabeled data (no predefined outputs).
  - Goal: Discover hidden patterns or structures.

2. Key Differences

Feature	Supervised Learning	Unsupervised Learning
Data	Labeled	Unlabeled
Objective	Predict outcomes	Discover patterns/clusters
Algorithms	Regression, Classification	Clustering, Dimensionality
Examples	Spam detection, Fraud det.	Customer segmentation

3. Algorithms

- Supervised Learning:
  - Regression: Linear Regression, Logistic Regression
  - Classification: Decision Trees, SVM, Random Forests

- Unsupervised Learning:
  - Clustering: K-Means, DBSCAN

  - Dimensionality Reduction: PCA, t-SNE

4. Output

- Supervised:

  - Specific predictions (e.g., classify as spam/not spam).

- Unsupervised:

  - Groups data into clusters or uncovers hidden patterns.

5. Applications

- Supervised Learning:

  - Medical diagnosis

  - Sentiment analysis

  - Stock price prediction

- Unsupervised Learning:

  - Market segmentation

  - Anomaly detection

  - Recommendation systems

6. Advantages and Challenges

- Supervised Learning:

  - Advantages: Accurate predictions, measurable performance.

  - Challenges: Requires large labeled datasets.

- Unsupervised Learning:

  - Advantages: Works with unlabeled data, useful for exploration.

  - Challenges: Hard to validate and interpret results.