TYPES OF MACHINE LEARNING:

1. SUPERVISED MACHINE LEARNING: This type of machine learning involves training a model on data that has both input and corresponding output. The goal is to learn the relationship between the input and output so that the model can predict the output for new, unseen inputs

Common tasks

- **Classification** (spam / not spam)
- **Regression** (price prediction)
- **Regression**: Used when the output or target variable is numerical (e.g., predicting house prices based on features like the number of rooms and size)
- **Classification**: Used when the output or target variable is categorical (e.g., predicting if an email is spam or not, or if a student will get placed)

Algorithms

- Linear Regression
- Logistic Regression
- Decision Tree
- KNN
- SVM
- Random Forest

Unsupervised Learning (8:10): In this type, the model learns from data that only has inputs, with no predefined output. The objective is to find patterns, structures, or relationships within the data (8:26-9:01).

Common tasks

- **Clustering**
- **Dimensionality Reduction**
- **Association**

- **Clustering** (9:56): Groups similar data points together into clusters (e.g., segmenting customers based on their behavior) (10:29-10:51).
- **Dimensionality Reduction** (12:32): Reduces the number of input features (columns) in a dataset while retaining most of the important information. This can improve algorithm performance and aid in data visualization (13:00-14:23).
- **Anomaly Detection** (17:07): Identifies unusual data points or outliers that deviate significantly from the norm (e.g., detecting manufacturing defects or fraudulent credit card transactions) (17:10-18:06).
- **Association Rule Learning** (18:11): Discovers interesting relationships or associations between variables in large datasets (e.g., the "beer and diapers" phenomenon in retail, where purchasing diapers often correlates with purchasing beer)

Algorithms

- K-Means
- Hierarchical Clustering
- DBSCAN
- PCA
- Apriori

3. Semi-Supervised Learning

👉 Uses **both labeled and unlabeled data**.

How it works

- Small amount of labeled data
- Large amount of unlabeled data
- Improves accuracy at lower labeling cost

Why it’s useful

- Labeling data is expensive
- Common in real-world datasets

Examples

- Image classification
- Speech recognition
- Web content classification

Techniques

- Self-training
- Co-training
- Label propagation

4. Reinforcement Learning

👉 The model learns by **trial and error** using rewards and penalties.

How it works

- Agent interacts with environment
- Takes actions
- Receives **reward or punishment**
- Goal: maximize total reward

Key terms

- Agent
- Environment

- Action
- Reward
- Policy

Examples

- Game playing (Chess, AlphaGo 🎮)
- Robot navigation 🤖
- Self-driving cars
- Recommendation systems

Algorithms

- Q-Learning
- SARSA
- Deep Q Networks (DQN)