Machine Learning: A Complete Overview

1. What is Machine Learning?

Machine Learning is a subset of artificial intelligence (AI) that gives systems the ability to learn and improve from experience without being explicitly programmed.

2. Machine Learning Lifecycle

- Problem Definition

- Data Collection

- Data Cleaning

- EDA (Exploratory Data Analysis)

- Feature Engineering

- Model Selection

- Training

- Evaluation

- Deployment

- Monitoring and Maintenance

3. Types of Machine Learning

Supervised Learning:

Trains on labeled data. Examples include:

- Linear Regression

- Decision Trees

- Random Forest

- Support Vector Machines

Unsupervised Learning:

Trains on unlabeled data to find structure. Examples include:

- K-Means Clustering

- PCA (Principal Component Analysis)

- Hierarchical Clustering

Reinforcement Learning:

An agent learns to act in an environment to maximize reward.

Key elements: Agent, Environment, Reward, Policy.

4. Popular ML Algorithms

- Linear Regression: Predicts continuous values.

- Logistic Regression: Predicts probabilities for classification.

- Decision Trees: Tree-like structure for decision making.

- Random Forest: Multiple decision trees combined (ensemble).

- KNN: Classifies based on nearest neighbors.

- SVM: Finds the optimal separating hyperplane.

- K-Means: Clusters data based on distance.

- Naive Bayes: Based on Bayes' Theorem, assumes feature independence.

5. Model Evaluation Metrics

For classification:

- Accuracy

- Precision

- Recall

- F1 Score

- Confusion Matrix

For regression:

- Mean Absolute Error (MAE)

- Mean Squared Error (MSE)

- R² Score

6. Deployment

After a model is trained and evaluated, it can be deployed using:

- Flask or FastAPI for APIs

- Streamlit for apps

- Docker for containers

- Cloud platforms (AWS, GCP, Azure)