Understanding Machine Learning

1 What is Machine Learning?

Machine learning involves learning a function $F: X \to Y$, where:

- X represents the input space (features)
- Y represents the output space (target values)

Depending on the availability of an explicit formula for F:

- If F is known, we can formulate and design an algorithm directly.
- If F is unknown, we rely on machine learning techniques to approximate it.

2 Algorithmic Approach

If we know F, then for a new input X_{new} , we compute the output directly:



where F is a function designed by us.

3 Machine Learning Approach

When F is unknown, we use a learning algorithm to approximate it:

$$(X,Y) \longrightarrow \mathcal{L} \longrightarrow g \approx F$$

where:

- \mathcal{L} represents a learning algorithm that learns from data.
- g is the learned function, which approximates F.

4 When to Apply Machine Learning?

Machine learning is applicable when the following conditions are met:

- 1. Data exists: There must be enough data available to train a model.
- 2. A pattern exists: The data should contain some underlying structure or relationship between input and output.
- 3. The pattern cannot be mathematically defined: If a clear mathematical formulation exists, a traditional algorithm may be a better choice.

5 Where to Apply Machine Learning?

5.1 Movie Recommendation System

Movies can be rated based on:

- Movie Attributes: Genres (Comedy, Action, etc.)
- User Preferences: Whether the user likes a particular genre or not.



5.2 Classification Problem: Logistic Regression

Logistic regression is used when the output is categorical, such as:

• IMDB movie ratings: Good vs. Bad

• Credit loan approval: Based on Age, Years of Experience, Debt, and Salary.



5.3 Regression Problem: Salary Prediction

Regression techniques can be used to predict salary based on performance (e.g., academic grades).



5.4 Spam Filtering Using Nave Bayes

Spam filtering uses Bayesian learning where the probability of an email being spam is determined based on its attributes:

$$P(Y = 1 | x_1, x_2, ..., x_n)$$



6 Why Use Machine Learning?

Machine learning can lead to significant improvements in performance and profitability. For example:

• Netflix reported that a 10% improvement in rating prediction leads to \$1M profit.

7 How to Apply Machine Learning?

When encountering misclassified data points, there are two possible adjustments:

- Shifting the decision boundary
- Rotating the decision boundary

The decision boundary follows the equation:

$$\omega_1 x_1 + \omega_2 x_2 + \omega_3 = 0$$

where:

- Changing ω_3 shifts the boundary.
- Changing ω_1, ω_2 rotates the boundary (affecting slope).

8 Learning Process Diagram

The process of machine learning can be visualized as:



Learned Function

9 Types of Learning

Machine learning can be categorized based on how the model learns from data:

- Supervised Learning: Learning from labeled data.
- Unsupervised Learning: Learning patterns from unlabeled data.
- Semi-supervised Learning: A mix of supervised and unsupervised learning.
- Reinforcement Learning: Learning through rewards and penalties.

Supervised Semi-Supervised Unsupervised

Reinforcement Learning