

Supervised Learning

Introduction

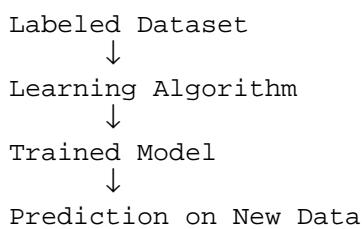
Supervised Learning is a type of machine learning where the model is trained using labeled data. Each training example consists of input features along with the correct output label. The model learns a mapping from inputs to outputs.

Labeled Data

Labeled data means that for every input data point, the correct output is already known. This helps the model learn by comparing its predictions with the actual answers.

Input Data → Correct Output
Email Text → Spam / Not Spam

How Supervised Learning Works



Types of Supervised Learning

Supervised learning problems are mainly divided into two categories:

- 1 **Regression:** Used when the output is a continuous value.
- 2 **Classification:** Used when the output is a class or category.

Regression

Regression models predict numerical values based on input features.

Example: Predicting house price, temperature forecasting.

Input Features → Regression Model → Numeric Output

Classification

Classification models assign inputs into predefined categories or classes.

Example: Email spam detection, disease diagnosis.

Input Features → Classification Model → Class Label

Common Supervised Learning Algorithms

- 1 Linear Regression
- 2 Logistic Regression
- 3 Decision Tree
- 4 K-Nearest Neighbors (KNN)

5 Support Vector Machine (SVM)

6 Random Forest

Advantages

- 1 High accuracy with labeled data
- 2 Easy to evaluate performance
- 3 Well-suited for real-world prediction tasks

Disadvantages

- 1 Requires large labeled datasets
- 2 Labeling data can be expensive
- 3 May not generalize well with poor data

Real-Life Example

In a student performance prediction system, historical student data with known results is used to train a model that predicts future academic performance.

Summary

Supervised learning is one of the most widely used machine learning approaches. It works best when labeled data is available and clear prediction goals are defined.