Supervised Learning: Regression

Elements of AI/ML

Introduction to Supervised Learning

·

Supervised Learning: A type of machine learning where the model is trained using labeled data. The algorithm learns to map input data (features) to the correct output (target).

Goal: Predict outcomes or classify data based on learned patterns.

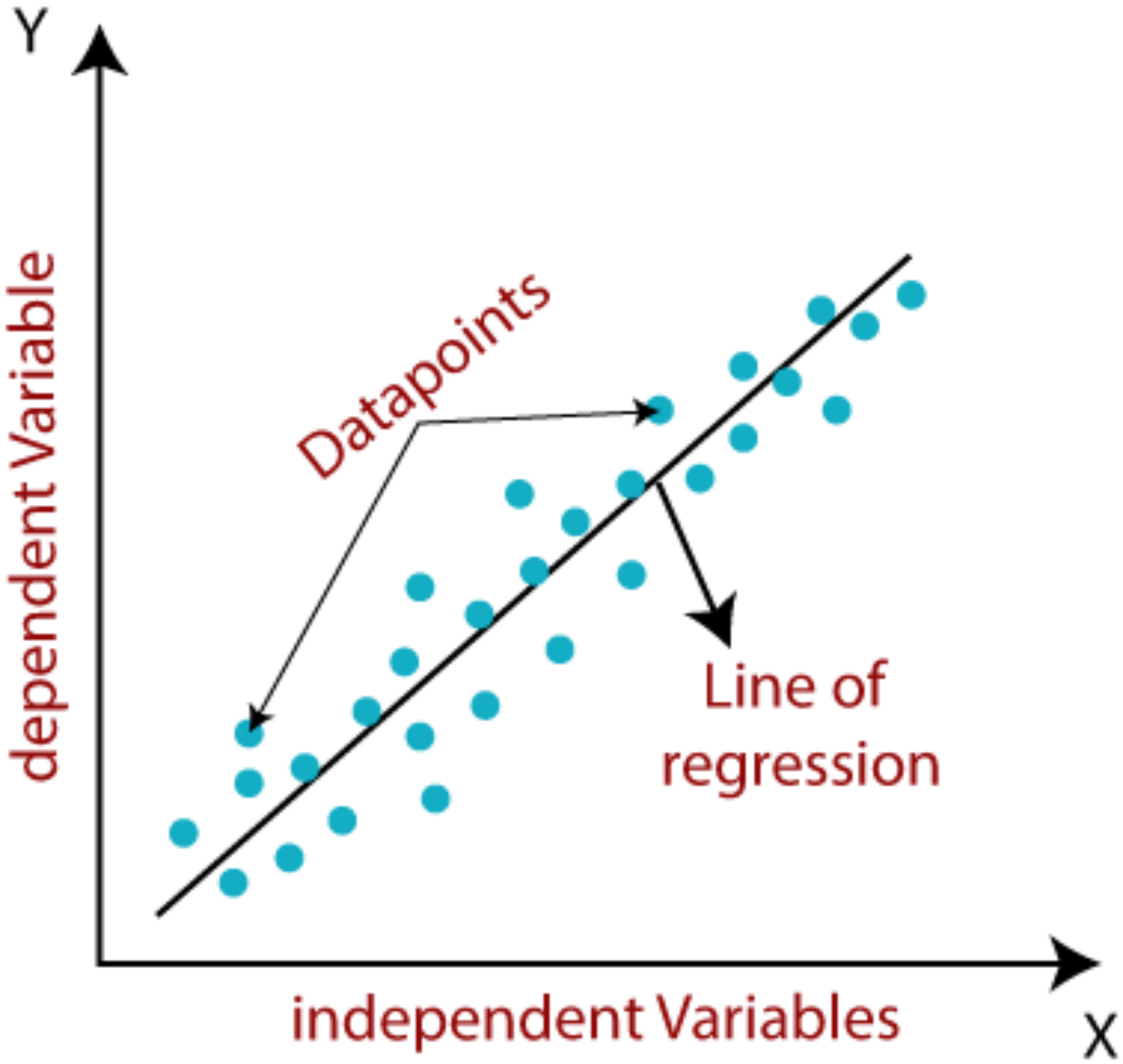
Two types:

1. **Regression**: Predicts continuous output (e.g., house prices).
2. **Classification**: Predicts categorical output (e.g., spam vs non-spam emails)

About Regression

Definition: Regression is a statistical method used in supervised learning to model the relationship between a **dependent variable (target)** and one or more **independent variables (features)**.

Purpose: Regression helps to predict continuous outcomes, such as predicting house prices, stock market trends, etc.



Types of Regression in Supervised Learning

.

1. Linear Regression: Models the relationship between the dependent and independent variables as a straight line.

2. Polynomial Regression: Extends linear regression by fitting a polynomial equation (non-linear).

3. Ridge and Lasso Regression: Adds regularization to linear regression to prevent overfitting.

4. Logistic Regression: Although used for classification, it’s considered regression because it predicts probabilities for binary outcomes.

Linear Regression in Detail

▪

Definition: Linear Regression is a supervised learning algorithm that models the relationship between a dependent variable and one or more independent variables using a straight line.

- **Simple Linear Regression**: Involves one independent variable.
- **Multiple Linear Regression**: Involves two or more independent variables.

Simple Linear Regression Formula

For **one independent variable** (single feature), the equation simplifies to:

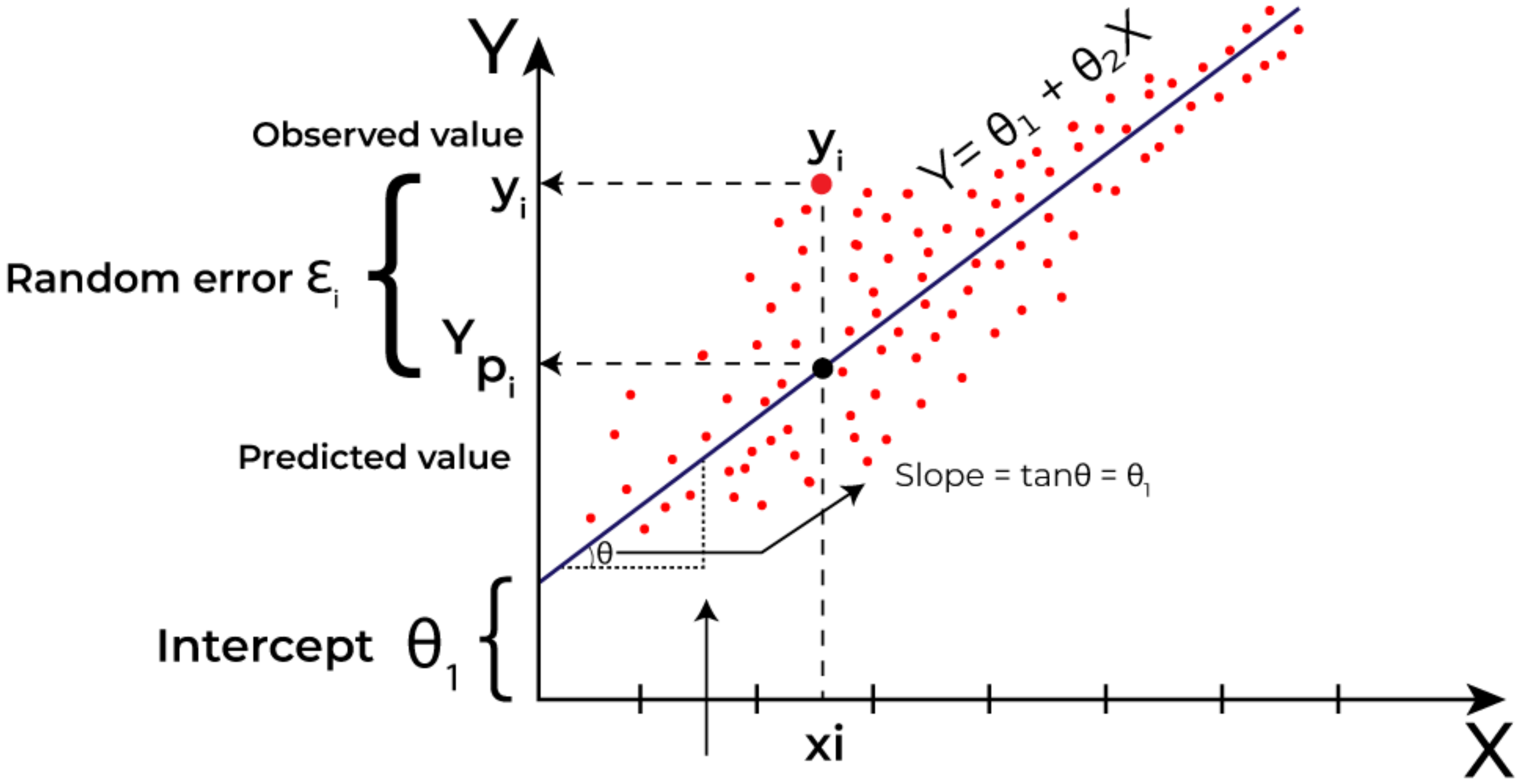
y = w1*x1+ w0+ e

where:

- Y is the dependent variable
- X1, X2, …, Xn are the independent variables
- w0 is the intercept
- W1 are the slopes
- e is error (residual).

Graphical Representation

-



Finding the Best Fit Line

-

Goal: To find the values of β_0 and β_1 that minimize the difference between the predicted and actual values.

Loss Function: Sum of squared errors (SSE)

{SSE} = sum (observed valued - actual value) ^2

$$\text{SSE} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

The best fit line is obtained by minimizing this error using optimization techniques like **Gradient Descent** or **Ordinary Least Squares (OLS)**

Loss Function

.

Definition: A **loss function** is a mathematical function used to quantify the error between the predicted output and the actual output (true value).

Goal: Minimize the loss function during training to improve the accuracy of predictions.

Types of Loss Functions

Loss functions differ based on the type of machine learning problem you're solving:

1. Mean Squared Error (MSE) (Regression)

2. Mean Absolute Error (MAE) (Regression)

3. Hinge Loss (Classification - SVM)

Mean Squared Error (MSE)

▪

Formula:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

Actual value.

- hat{y_i}: Predicted value.
- n: Number of data points.
- Use Case: Commonly used for regression tasks.
- Explanation: MSE calculates the average squared difference between actual and predicted values.

Mean Absolute Error (MAE)

-

- Formula: $$MAE = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i|$$

Use Case: Also used for **regression tasks**.

Explanation: MAE calculates the average absolute difference between actual and predicted values.

Advantages:

- Less sensitive to outliers compared to MSE.

Disadvantages:

- Does not penalize larger errors as much as MSE.

Gradient Descent

.

Definition: Gradient Descent is an optimization algorithm used to minimize the loss function by adjusting model parameters (weights and biases) iteratively.

Goal: Find the values of parameters that minimize the loss function.

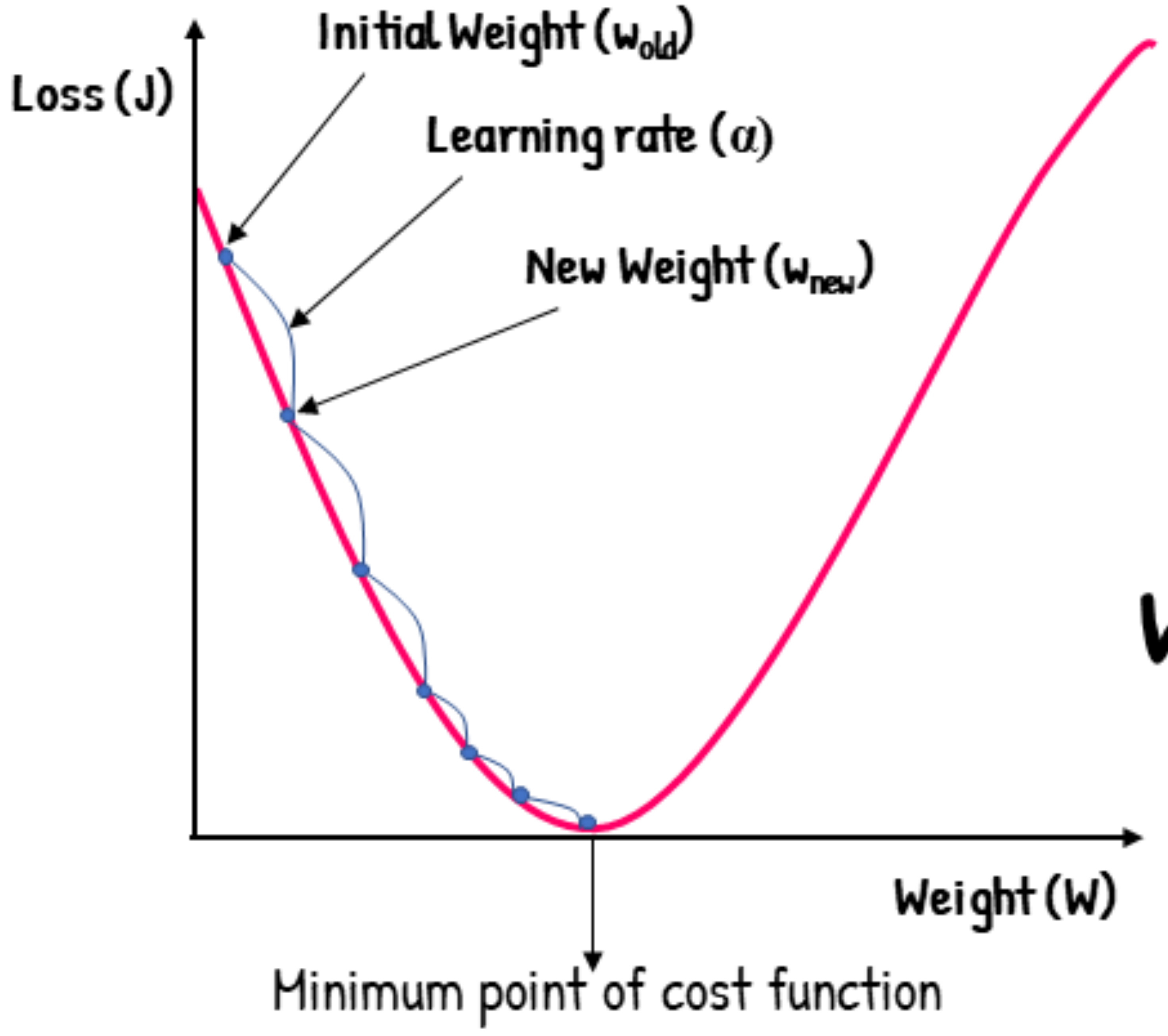
Variants of Gradient Descent

1. Batch Gradient Descent
2. Stochastic Gradient Descent (SGD
3. Mini-Batch Gradient Descent:

Diagram

-
- .

Gradient Descent



$$w_{new} = w_{old} - \alpha \frac{\delta J}{\delta w}$$

Advantages of Linear Regression

- **1. Simple to Understand**: Linear regression provides a clear relationship between dependent and independent variables.

  2. Easy to Implement: Easy to interpret and implement using common programming languages (Python, R, etc.).

  3. Computationally Efficient: Fast training and testing compared to more complex models.

Disadvantages of Linear Regression

1. Assumes Linearity: Only captures linear relationships between features and target.

2. Sensitive to Outliers: Outliers can distort the model and lead to inaccurate predictions.

3. Limited Expressiveness: Cannot model complex patterns in data (e.g., non-linear relationships

Summary

▪

Key Points:

- Supervised learning: Regression is used for predicting continuous outputs.
- Linear regression provides a simple model for estimating relationships between variables.
- There are many variations of regression techniques, each suited to specific problem types.