Regression Algorithms

Introduction

Regression algorithms are a category of supervised learning techniques used to predict continuous numerical values. These algorithms learn the relationship between input features and a continuous target variable.

What is Regression?

Regression is used when the output variable is continuous, such as price, temperature, salary, or marks. The goal is to find a best-fit relationship between input and output.

```
Input Features  →  Regression Model  →  Continuous Output
```

Types of Regression Algorithms

1. Simple Linear Regression
2. Multiple Linear Regression
3. Polynomial Regression
4. Ridge Regression
5. Lasso Regression

1. Simple Linear Regression

Simple Linear Regression models the relationship between one independent variable and one dependent variable using a straight line.

Equation: $y = mx + c$

Example: Predicting house price based on size.

```
x (Size) → Linear Model → y (Price)
```

2. Multiple Linear Regression

Multiple Linear Regression uses more than one independent variable to predict the dependent variable.

Example: Predicting house price using size, location, and number of rooms.

3. Polynomial Regression

Polynomial Regression models non-linear relationships by transforming features into polynomial terms.

Example: Salary vs experience when growth is not linear.

4. Ridge Regression

Ridge Regression adds a penalty term to reduce overfitting by shrinking coefficients.

5. Lasso Regression

Lasso Regression can shrink some coefficients to zero, effectively performing feature selection.

Regression Model Evaluation Metrics

1. Mean Absolute Error (MAE)
2. Mean Squared Error (MSE)
3. Root Mean Squared Error (RMSE)
4. R-squared (R^2)

Advantages

1. Simple and interpretable models
2. Works well for numeric prediction problems
3. Fast training and prediction

Disadvantages

1. Sensitive to outliers
2. Assumes linear relationship in many cases
3. May underperform on complex data

Real-Life Example

In salary prediction systems, regression algorithms predict employee salary based on experience, education, and skills.

Summary

Regression algorithms are fundamental supervised learning techniques used for predicting continuous values. Choosing the right regression model depends on data complexity and problem requirements.