**Linear Regression**

**Linear Regression** is a fundamental statistical and machine learning algorithm used for predicting continuous outcomes. It establishes a **linear relationship** between the **dependent variable (target)** and one or more **independent variables (features)**. The goal is to find the best-fitting line that minimizes the difference between actual and predicted values.

**Mathematical Equation :**

Simple Linear Regression Model is a model with a single regressor X that has a linear relationship with response variable Y. For a **Simple Linear Regression**, the equation is:

Where :

* Y = Target Variable(Dependent Variable or Response Variable)
* X = Feature Variable (Independent variable)
* = Y- Intercept (constant)
* ​ = Slope (coefficient)
* ε = Random Error Comonent

For **multiple variables (Multiple Linear Regression)**, the equation extends to:

**How It Works**

1. **Training the Model** → The model calculates **coefficients (β-values)** using the **Ordinary Least Squares (OLS)** method, which minimizes the sum of squared errors.
2. **Making Predictions** → Once trained, the model predicts values by substituting input features into the equation.
3. **Evaluation Metrics** :
   * **Mean Squared Error (MSE)** → Measures average squared differences between actual and predicted values.
   * **R-squared (R²)** → Represents the proportion of variance in the target explained by the features.
   * **Adjusted R-squared** → Adjusts R² for the number of predictors to avoid overfitting.

**Advantages :**

✔ Easy to interpret & implement  
✔ Works well for linearly correlated data  
✔ Computationally efficient

**Limitations :**

✖ Assumes a **linear** relationship (not ideal for complex patterns)  
✖ Sensitive to **outliers**  
✖ Prone to **multicollinearity**

**Use Cases :**

* Predicting **house prices** (Boston dataset)
* Estimating **sales revenue** based on marketing spend
* Assessing **credit risk** in finance