**Linear Regression**

Linear regression predicts continuous/real or numeric values such as sales, salary, product price, etc.

Linear regression shows a linear relationship between the independent variable or variables and a dependent variable.

**Main definition:**

Linear regression is a supervised machine learning model that predicts the continuous/real or numeric values using the concept of linear relationship.

It is based on the line equation i.e., y = mx + c

Where, m= linear regression coefficient

x= independent variable

y= dependent variable

c= intercept of line

**Two types of Linear regression:**

1. **Simple linear regression:**

When the dependent variable depends on single independent variables then it is known as Simple linear regression.

1. **Multiple linear regression:**

When the dependent variable depends on multiple independent variables then it is known as multiple linear regression.

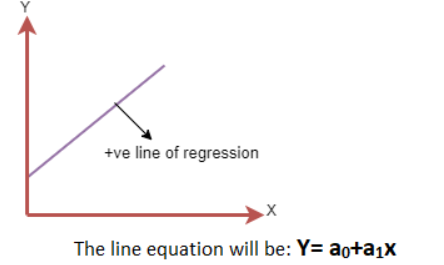
**Linear regression line:**

The linear line showing the relationship between the dependent variable and the independent variable is called the Linear regression line

Linear regression lines can show two types of relationship:

**1] Positive Linear Relationship:**

If the dependent variable increases on the Y-axis and the independent variable increases on the X-axis, then such a relationship is called a Positive linear relationship.



**2] Negative Linear Relationship:**

If the dependent variable decreases on the Y-axis and the independent variable increases on X-axis, then such a relationship is called as a Negative Linear relationship.

A graph of a line of equation

Description automatically generated

**Cost Function:**

To determine the best-fit line, the cost function is used.

We can use the cost function to find the accuracy of the mapping function, which maps the input variable to output variable. The mapping function is also known as the **Hypothesis function**.

For linear regression, the **Mean Squared Error (MSE)** cost function is used, which is the average of squared error between the predicted values and the actual values.

**MSE formula:**

A math equations and formulas

Description automatically generated with medium confidence

**Residual:**

The distance between the actual value and the predicted value is called residual. If the distance between the actual value and predicted value is high then the residual will be high, and so the cost function will be high and vice versa.

**Gradient Descent:**

Gradient descent is used to minimize the MSE by calculating the gradient descent of the cost function.

It is done by selecting a random value of coefficient and the iteratively updating the values to reach the minimum cost function.