

## Training Day-7 Report:

### What is Linear Regression?

Linear regression is a type of supervised machine learning algorithm that computes the linear relationship between the dependent variable and one or more independent features by fitting a linear equation to observed data.

When there is only one independent feature, it is known as **Simple Linear Regression**, and when there are more than one feature, it is known as **Multiple Linear Regression**.

Similarly, when there is only one dependent variable, it is considered **Univariate Linear Regression**, while when there are more than one dependent variables, it is known as **Multivariate Regression**.

### Why Linear Regression is Important?

The interpretability of linear regression is a notable strength. The model's equation provides clear coefficients that elucidate the impact of each independent variable on the dependent variable, facilitating a deeper understanding of the underlying dynamics. Its simplicity is a virtue, as linear regression is transparent, easy to implement, and serves as a foundational concept for more complex algorithms.

Linear regression is not merely a predictive tool; it forms the basis for various advanced models. Techniques like regularization and support vector machines draw inspiration from linear regression, expanding its utility. Additionally, linear regression is a cornerstone in assumption testing, enabling researchers to validate key assumptions about the data.

### Types of Linear Regression:-

There are two main types of linear regression:

#### Simple Linear Regression

This is the simplest form of linear regression, and it involves only one independent variable and one dependent variable. The equation for simple linear regression is:

$$y = \beta_0 + \beta_1 X$$

where:

- Y is the dependent variable
- X is the independent variable
- $\beta_0$  is the intercept
- $\beta_1$  is the slope

#### Multiple Linear Regression

This involves more than one independent variable and one dependent variable.

The equation for multiple linear regression is:

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

where:

- Y is the dependent variable
- $X_1, X_2, \dots, X_p$  are the independent variables
- $\beta_0$  is the intercept
- $\beta_1, \beta_2, \dots, \beta_n$  are the slopes

*The goal of the algorithm is to find the best Fit Line equation that can predict the values based on the independent variables.*

In regression set of records are present with X and Y values and these values are used to learn a function so if you want to predict Y from an unknown X this learned function can be used. In regression we have to find the value of Y, So, a function is required that predicts continuous Y in the case of regression given X as independent features.