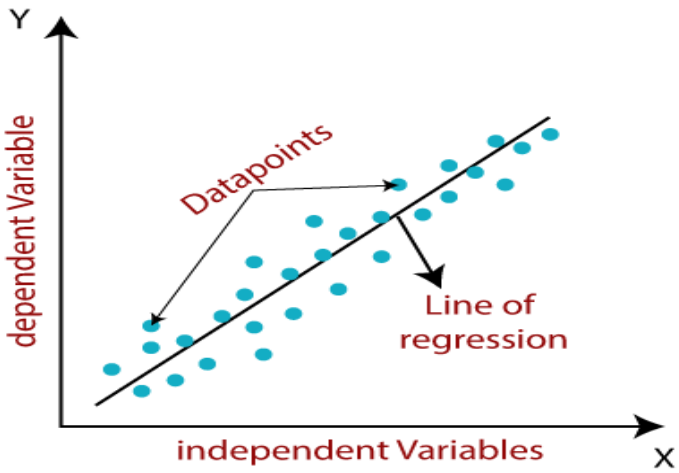
Linear Regression

Linear regression is the process of modeling the relationship between two variables by fitting a linear equation to observed data. When the linear equation is known, we can predict the value of a dependent variable by putting the value of the independent variable to the linear equation. Linear regression is used to predict the value of a continuous variable.

We assume that there is a linear relationship between the variables before applying linear regression. Once a regression model has been fit to a group of data, examination of the residuals (the deviations from the fitted line to the observed values) allows us to investigate the validity of our assumption that a linear relationship exists.

A simple linear equation look like: ***y = b0 + b1*x***



Where b0 and b1 are coefficients, y is output variable and x is input variable. In higher dimensions when we have more than one input (x), the line is called a plane or a hyper-plane.

We need to find the value of coefficients so that the residue and cost function is minimized. The cost function is the sum of squares of difference between predicted output and actual output divided by no of observations. We start with some value for b0 and b1, then we keep on changing b0 and b1 to reduce the cost function until we end up at minimum.

Logistic Regression:

Logistic regression is used for classification problems. It is used to predict a binary outcome based on a set of independent variables. We don't require a linear relationship between input and output variables to model the data using logistic regression but we require that the independent variable is binary or dichotomous. It can also be used to calculate the probability that an individual belongs to one of the groups. Once the model is derived, the probability is calculated using the following equation:

$$p = exp(\beta o + \beta 1(x)) = e^{(\beta o + \beta 1(x))}$$

If the probability of an observation belonging to a class A is equal to or more than 0.5 then the observation is classified class A.

While deriving the model, we start at some value of βo and $\beta 1,$ then we keep on changing βo and $\beta 1$ to reduce the loss function until we end up at minimum. A cost function is a measure of fit between a mathematical model of data and the actual data. The cost function is defined as:

$$\text{Cost}(h\theta(x),y) = \begin{cases} -\log(h\theta(x)) & \text{if } y = 1 \\ -\log(1-h\theta(x)) & \text{if } y = 0 \end{cases}$$