# MACHINE LEARNING
# WEBINAR-2

Vasugi I
Indumathi Prabakeran

BITS Pilani
Pilani Campus

innovate achieve lead

BITS Pilani

Pilani|Dubai|Goa|Hyderabad

# Linear Regression

# Agenda

- Introduction
- Machine Learning approach
- Python Libraries
- Linear Regression
- Demo

# Overview

Machine Learning problem has

- Target/Dependent variable
- Predictor/Independent variables
- Historical data – target and predictor variables

# Overview

Machine learning pipeline consists of:

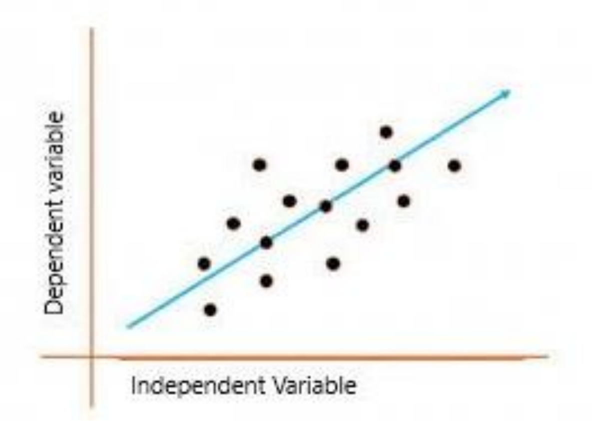
- Importing Data : csv, xls, json etc..
- Exploratory Data Analysis : Univariate and Multivariate
- Data Pre-processing
- Model building: Regression, Classification
- Model Evaluation

# Python Libraries

- Pandas : Data Manipulations
- Numpy : Mathematical operations
- Scikit-learn : Scikit-learn is most popular for classical ML algorithms
- Matplotlib and Seaborn : Visualisations

# Linear Regression

- Linear regression is a the simplest statistical regression method used for predictive analysis
- It shows the linear relationship between the independent(predictor) variable i.e. X-axis and the dependent(target) variable i.e. Y-axis
- Single input variable **X**(independent variable) : **Simple linear regression**



# Linear Regression

To calculate best-fit line linear regression uses a linear equation :

$$Y = B_0 + B_1 * X$$

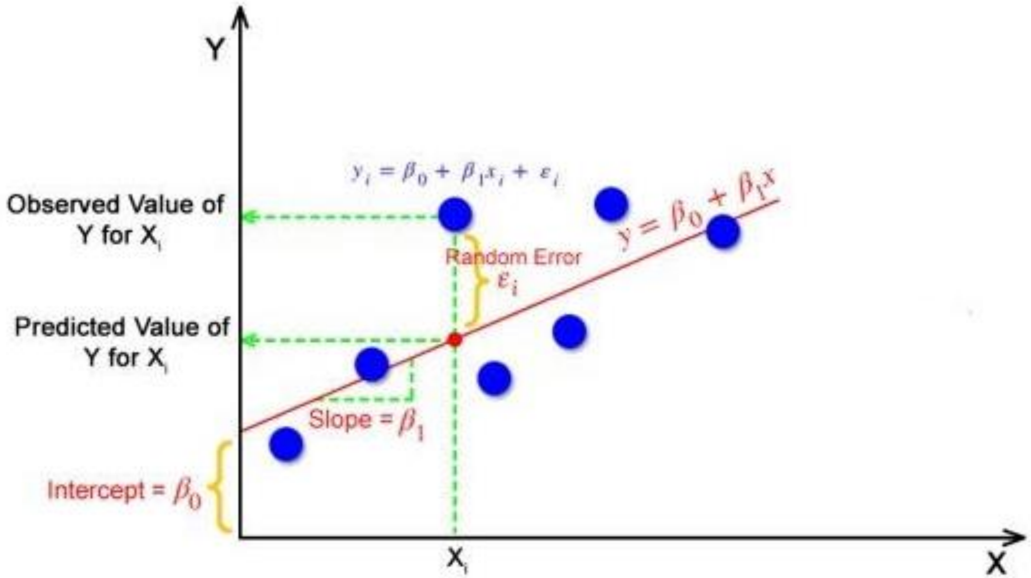
where $Y$ = Dependent Variable

$B_0$ = intercept

$B_1$ = slope

$X$ = Independent variable

**To find best-fit line, get best values for $B_0$ , $B_1$**



# Cost Function : Linear regression

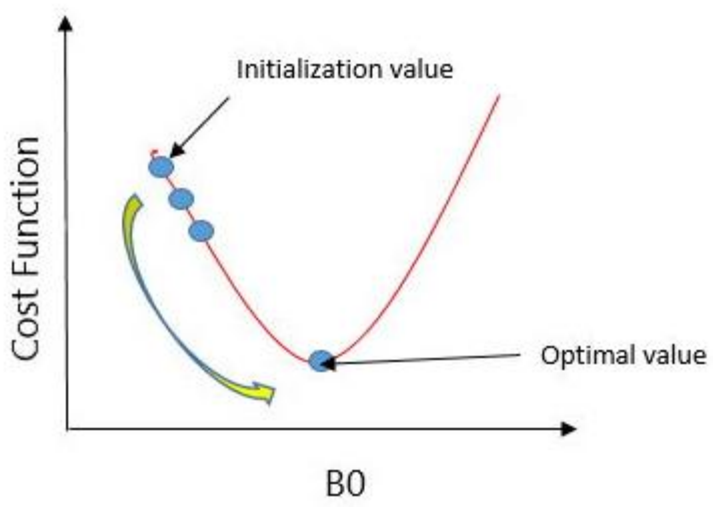
- Cost function helps to get the optimal values for $B_0$ and $B_1$
- Linear Regression, uses **Mean Squared Error (MSE)** cost function
  - average of squared error that occurred between the $\mathbf{y_{predicted}}$ and $\mathbf{y_i}$.
- We calculate MSE :

$$MSE = \frac{1}{N} \sum_{i=1}^{n} (y_i - (B1 x_i + B0))^2$$

- Values of $B_0$ and $B_1$ should be determined such that the MSE value settles at the minima
- This can be done using Gradient Descent method

# Gradient Descent : Linear Regression

- Gradient Descent is an optimization algorithms that optimize the cost function to reach the optimal minimal solution
- Reduce the cost function(MSE) for all data points by updating the values of $B_0$ and $B_1$ iteratively until an optimal solution is obtained



# Gradient Descent : Linear Regression

- To update $B_0$ and $B_1$, we take gradients from the cost function
- To find these gradients, we take partial derivatives for $B_0$ and $B_1$.

$$J = \frac{1}{n} \sum_{i=1}^{n} (B_0 + B_1 \cdot x_i - y_i)^2$$

$$\frac{\partial J}{\partial B_0} = \frac{2}{n} \sum_{i=1}^{n} (B_0 + B_1 \cdot x_i - y_i)$$

$$\frac{\partial J}{\partial B_1} = \frac{2}{n} \sum_{i=1}^{n} (B_0 + B_1 \cdot x_i - y_i) \cdot x_i$$

$$B_0 = B_0 - \alpha \cdot \frac{2}{n} \sum_{i=1}^{n} (pred_i - y_i)$$

$$B_1 = B_1 - \alpha \cdot \frac{2}{n} \sum_{i=1}^{n} (pred_i - y_i) \cdot x_i$$

# Evaluation Metrics : Linear Regression

- Linear regression model can be assessed using various evaluation metrics
  - Root Mean Squared Error (RMSE)

    specifies how close the observed data points are to the predicted values

$$RMSE = \sqrt{\frac{RSS}{n}} = \sqrt{\sum_{i=1}^{n}\left(y_i^{Actual} - y_i^{Predicted}\right)^2 / n}$$

- Residual sum of Squares (RSS) is defined as the sum of squares of the residual for each data point