Good Evening all. Hope you are doing well.

Let’s try to understand little more about Machine Learning.

First of all let us know the about the difference between

**Artificial Intelligence vs Machine Learning vs Deep Learning vs Data Science:**

**Artificial Intelligence (AI):**

The main aim of AI is to create an application where it performs all its tasks without human intervention.

Ex: Amazon.in, Chat bots, Google assistance, self-driving cars, Alexa, video streaming recommendations, Netflix etc.

Machine Learning (ML):

It’s part of AI. Here we can explore, analyse, visualize and perform predictions and other tasks of the data by using some statistical tools.

Statistical tools: some algorithms

**Deep Learning (DL): It’s subset of ML.** It is used to mimic the human brain using Multinural networks

Ex: ANN, CNN, Object detect problems

NLP: Natural Language Processing: where we can change our test data into numerical format.

NLP is used in ML, DL, DS and AI.

**Data science:** In short it’s a combination of AI, ML, DL and NLP.

Here we can use all the techniques of AI, ML and DL, NLP using python programming language**.**

**Types of Machine Learning:**

1. **Supervised Machine Learning**
2. **Unsupervised Machine Learning**

**Supervised Machine Learning:** Here we will be working on dependent and independent features.

It’s again classified in to two types:

**Regression and classification:**

**Regression model:** Here we can find continuous variable in output**.**

**Ex: Degree Exp salary**

BE 7 50k

PHD 2 70k

**--- --- ---**

**---- ---- ----**

**--- --- ---**

**--- --- ---**

Here we can find the Degree and Experience are the independent features and Salary would be a dependent feature.

So based on degree and experience the salary of the person is changing continuous. Hence it’s regression problem statement.

**Ex: No of play hours No of study hours Pass/Fail**

9 1 F

7 2 F

3 5 P

**-- -- --**

**-- -- --**

From the above example, we can say that No f ply hours and no of study hours are independent and pass/fail feature is output. And it’s having limited number of variables (pass/fail)

Hence we can say that it’s classification problem.

In this case, if the output is having only 2 variables it’s called BINARY CLASSIFICAION .

If more than 2, then it’s called, MULTICLASS CLASSIFICATION.

**Let us find the below problem statements where regression or classification:**

Flight Price Prediction – REGRESSION

Algerian Fire Forest data – CLASSIFICATION

Air Quality index -- REGRESSION

Tomorrow rain / not --- CLASSIFICAITON

Buy day of the person -- CLASSIFICATION

1. **Unsupervised Machine Learning:** Its depending on grouping/clustering

**Ex: customer segmentaion**

**Age salary spending-score (1-10)**

24 70k 1

26 100k 9

-- -- --

21 20k 9

25 120k 2

**Simple Linear Regression:**

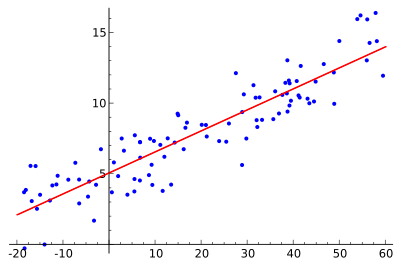
Here our main aim is to create a model. Here we need to train the data and the predict the data.

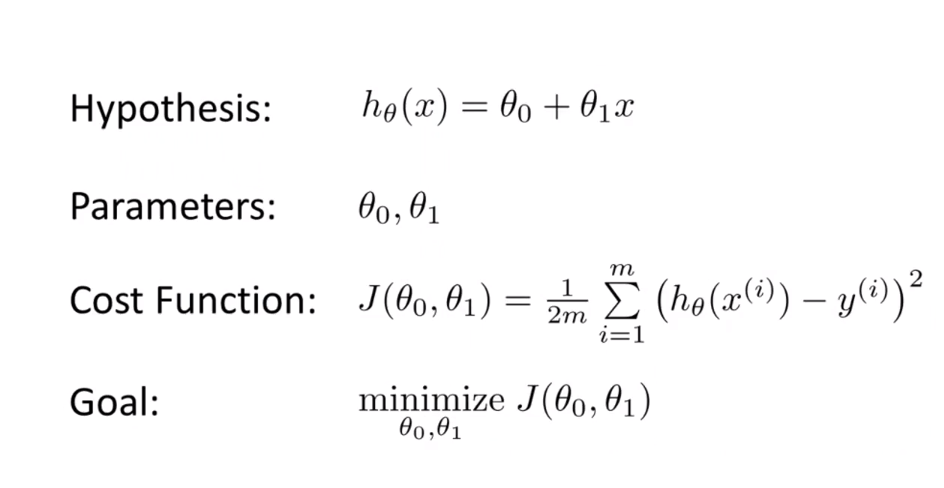
Here we need to find the ‘best fit line’ based on real point and predicted points.

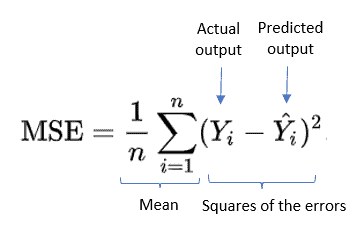
In best fit line the difference between real point and predicted points should be very minimal.

If the difference between real point and predicted points are high, we can call that as residuals/errors.

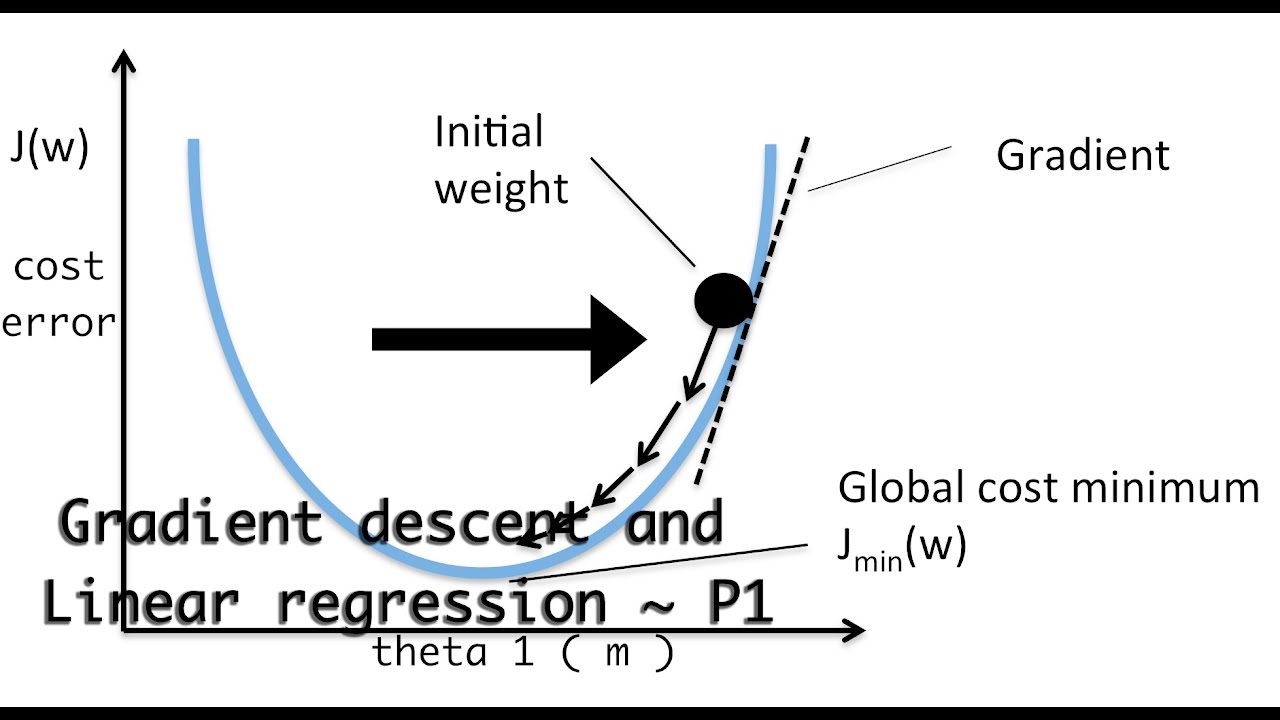
**Equation of straight line: y = mx+c**



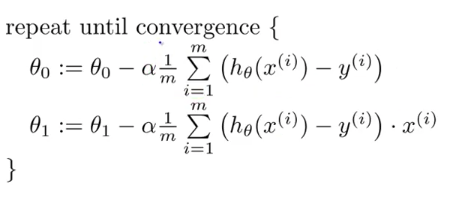
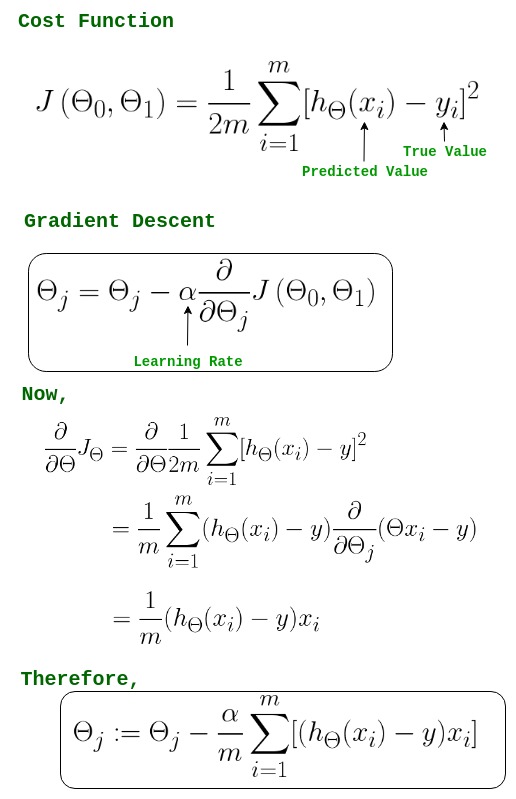




**Convergence Algorithm:**



Here our main aim is to reduce the cost function. With the help of convergence algorithm we can reduce the cost function.



Generally **Alfa** value should be **0.001**.

It takes a lot of time to reach the Global Minima, If alfa value is low,

If alfa value is high, it will never reach the global minima,