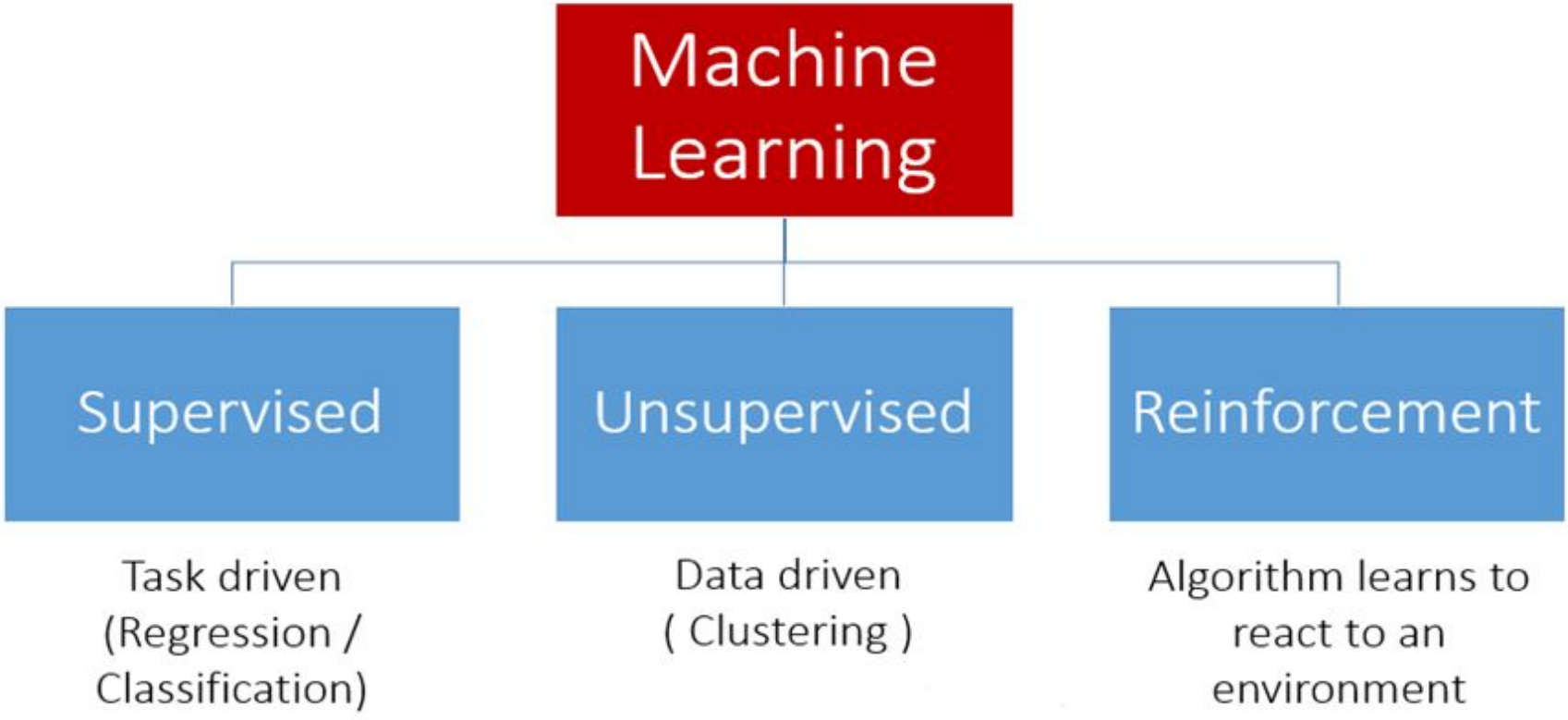
Linear Regression

Shaurya Singhal
Anshu Kumar

Types of Machine Learning



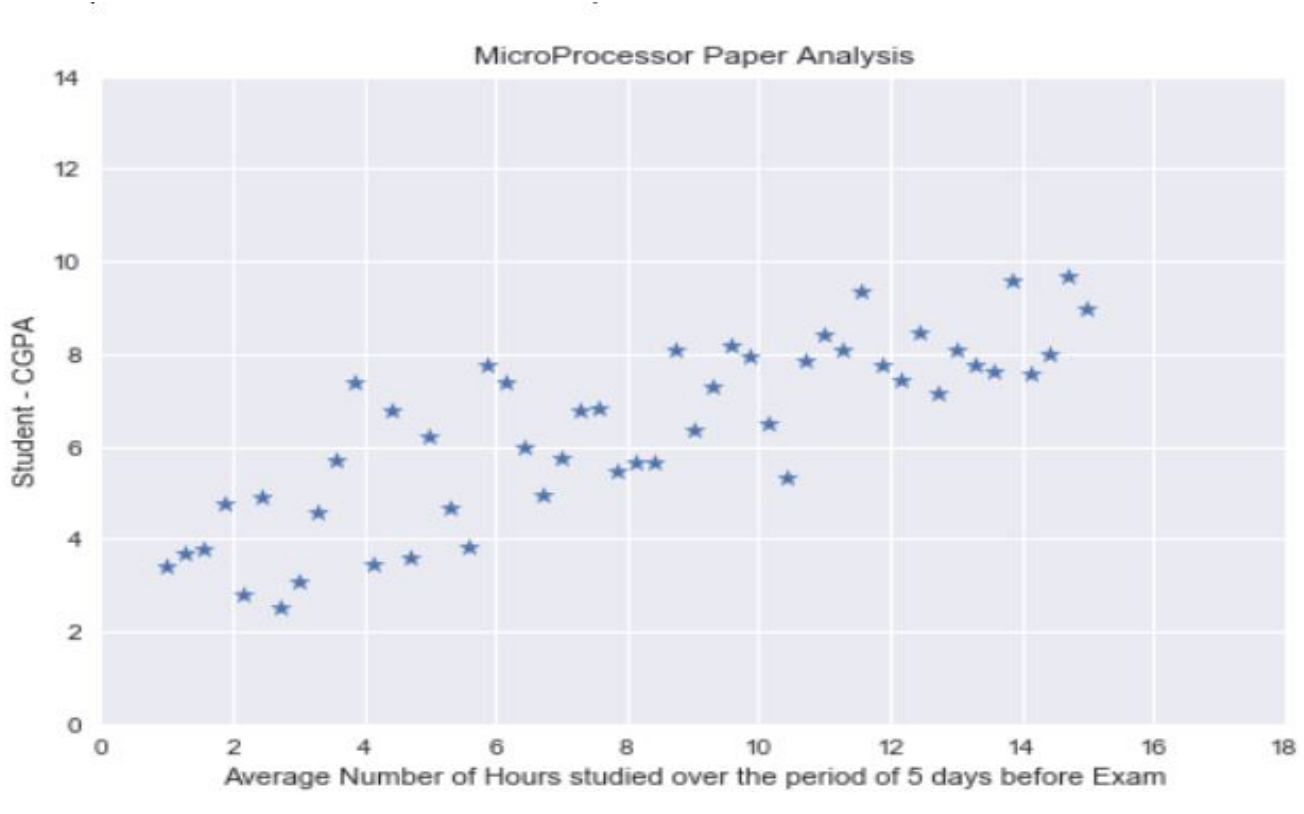
Univariate Linear Regression

Univariate linear regression focuses on determining relationship between one independent variable(X) and one dependent variable(Y).

X(CGPA)	Label(Y) (Number of hrs. study)
7.8	6.6
9.4	12
8.8	10
8.3	9.5
7.3	6
…..	…..

Dataset - How univariate data Looks like

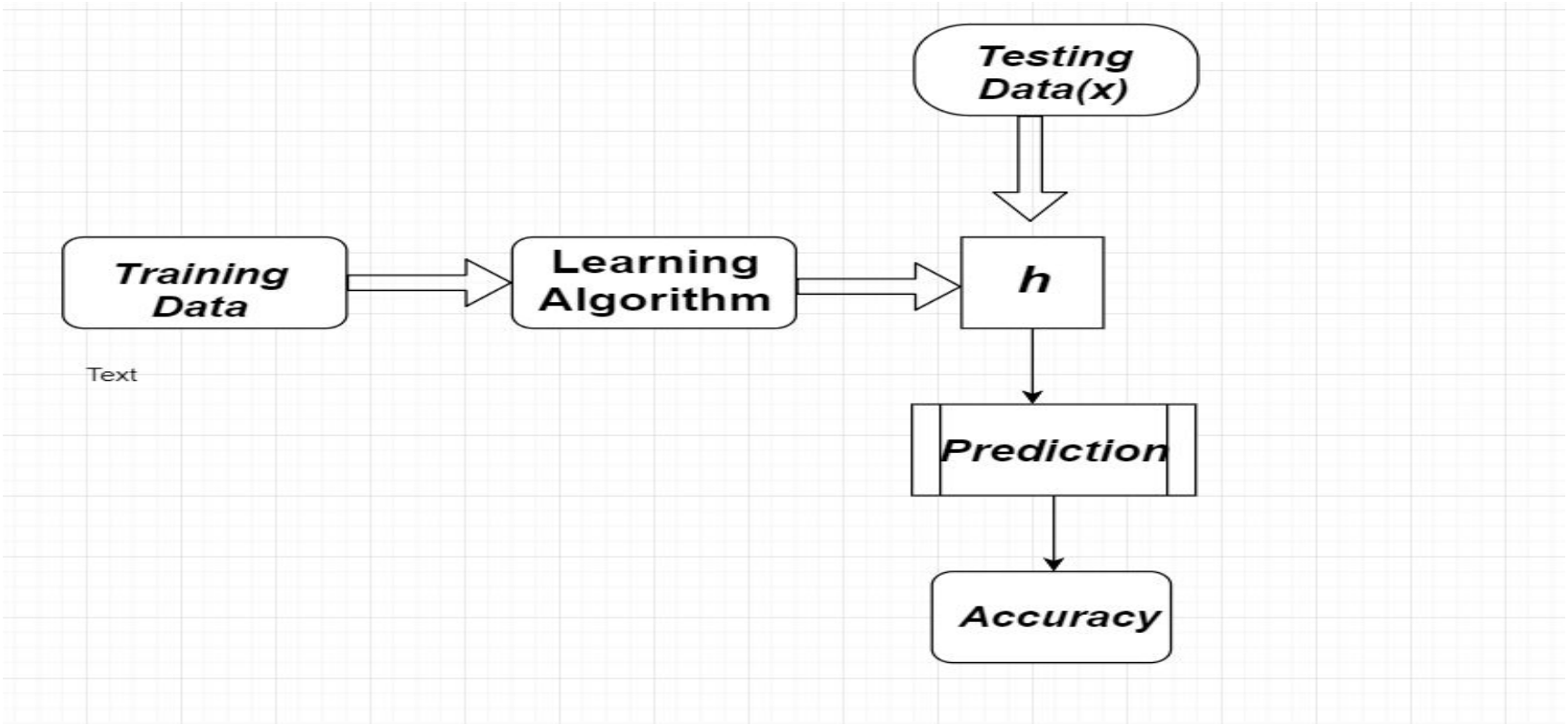
Example : - CGPA Vs Hours of Study (Not Attendence)



Dataset Common Terminology

1. No. of Samples(m)
2. Input Features(n)
3. Training set
4. Test set
5. Hypothesis

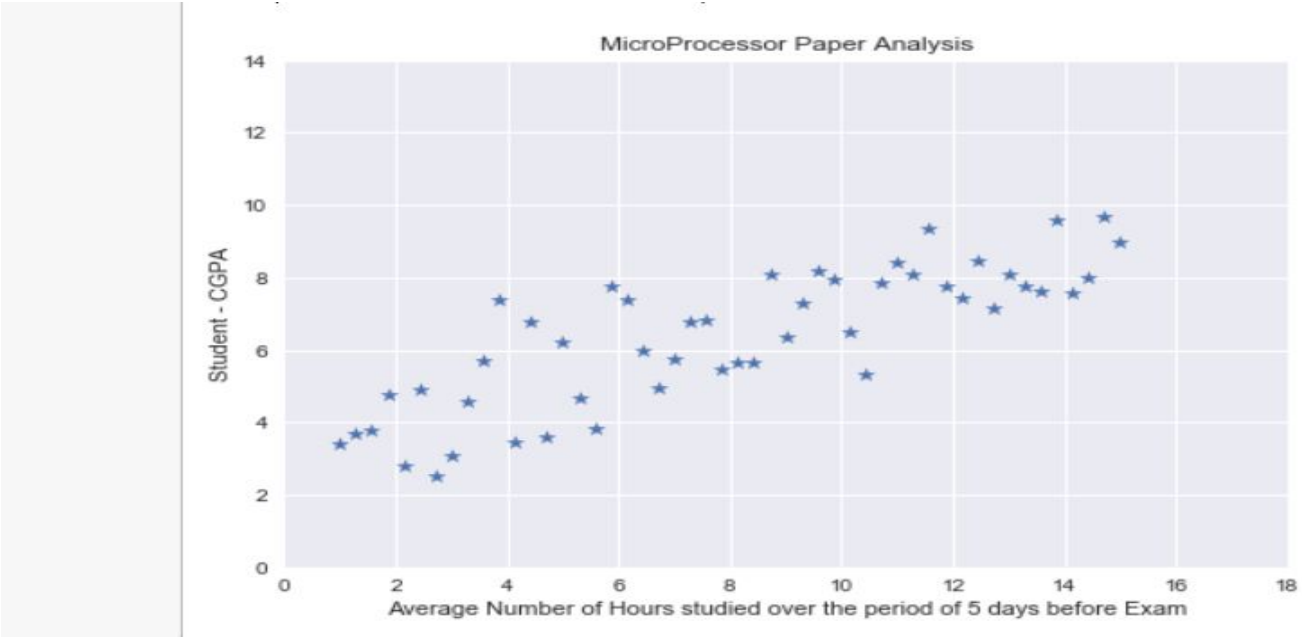
Aim and working (Pipeline)



Text

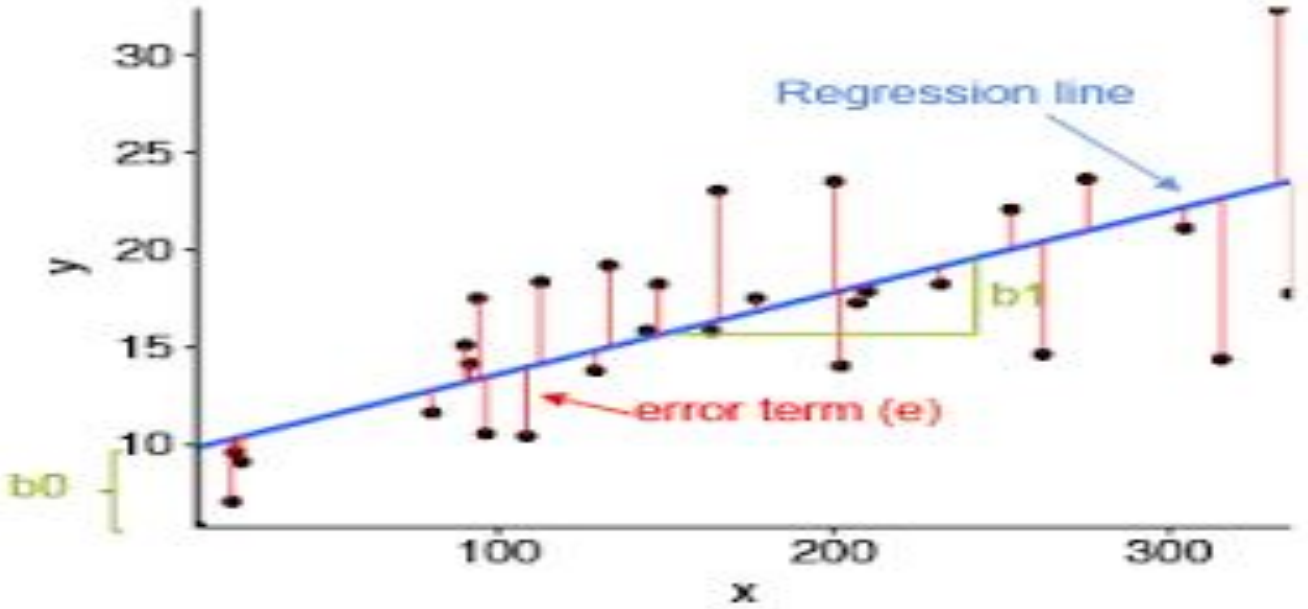
Which line seems right ?

Hypothesis: $h_\theta(x) = \theta_0 + \theta_1 x$



Q - How to find parameters(hypothesis) ?

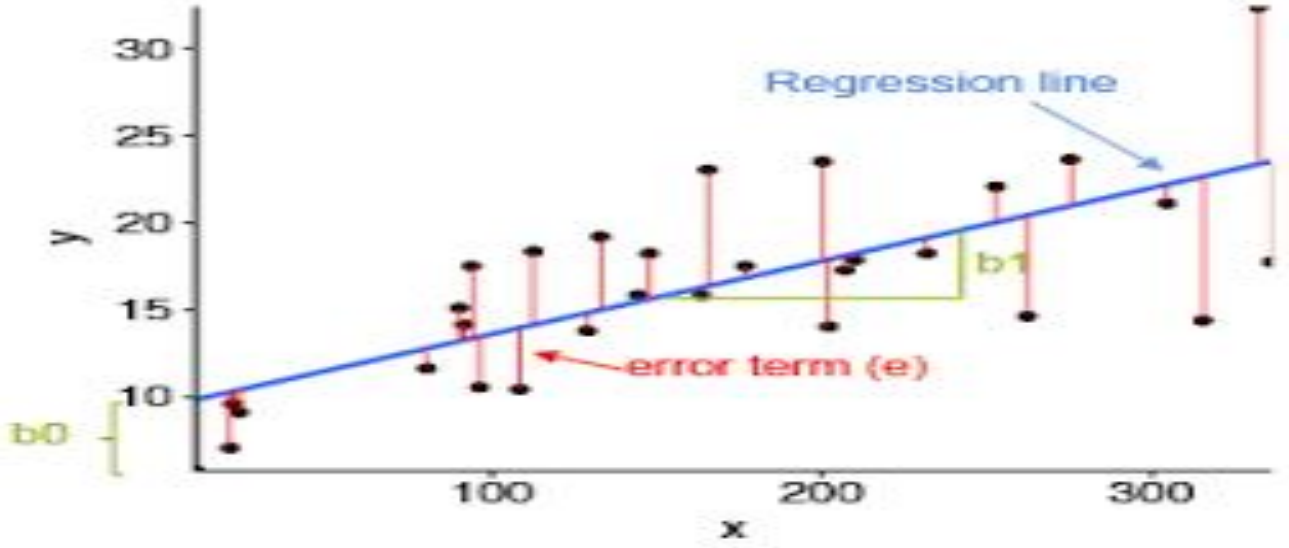
Q - What is the measure of finding Best line ?



Ans - Error Function (How?)

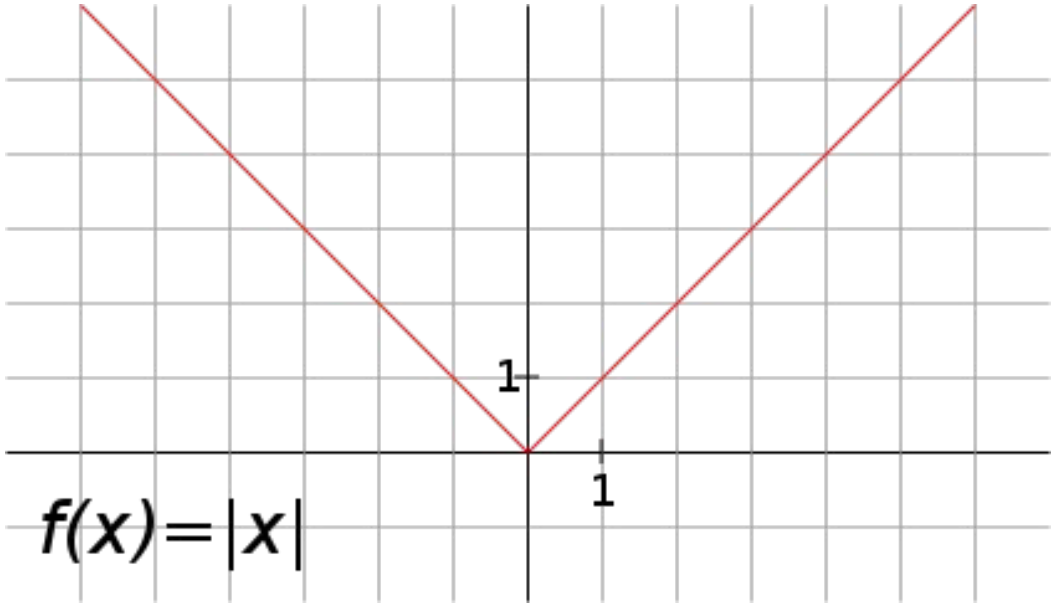
Choice of Error Function

1. $\sum (Y_{pred} - Y_{actual})$ -ve and +ve error added up



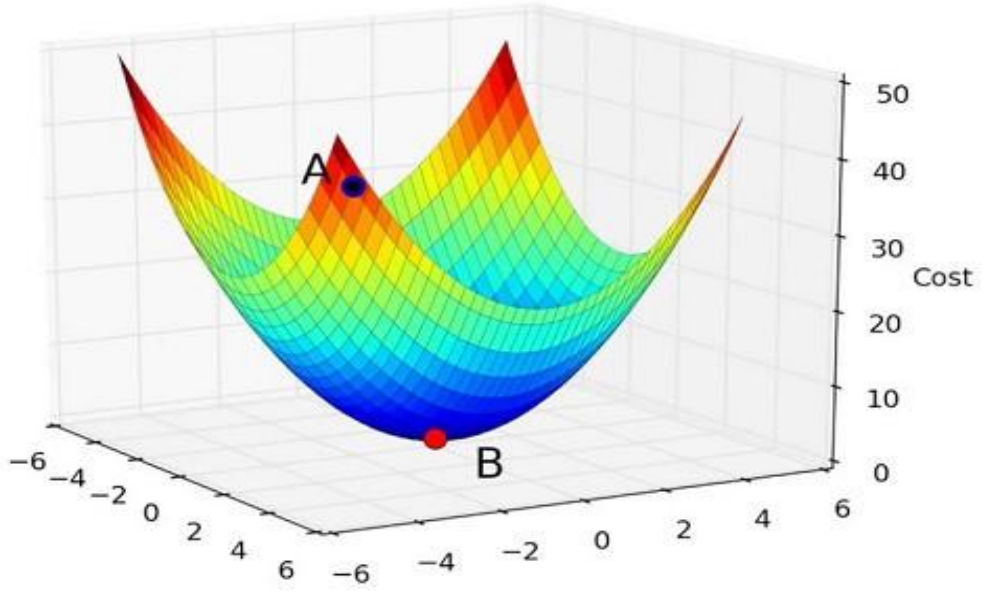
2. $\sum |(Y_{pred} - Y_{actual})|$

Non-differentiable at (Ypred == Yactual)



3. $$\sum (Y_{pred} - Y_{actual})^2$$

Perfect, and this function is Convex in Nature(Helpful)



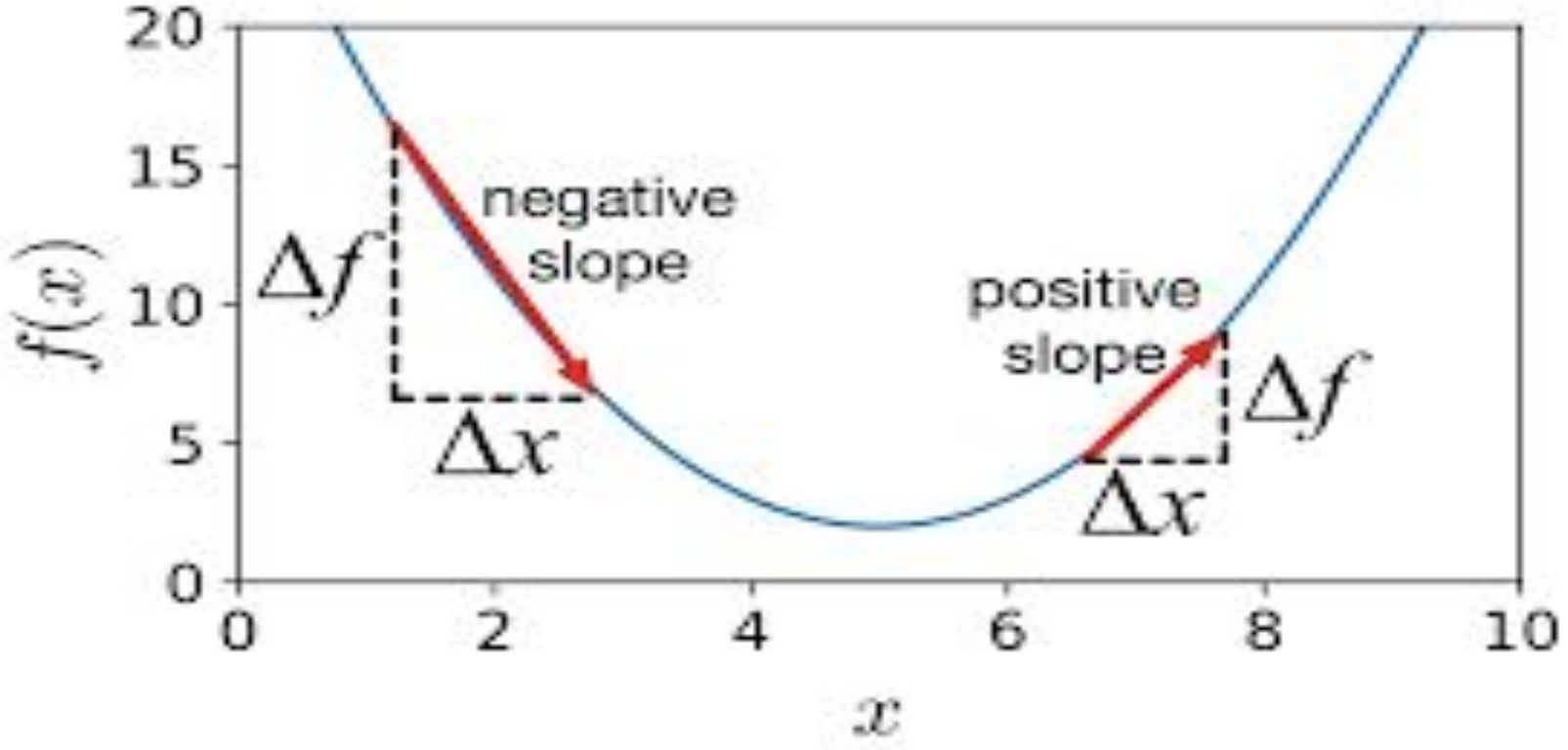
Goal - Minimizing the Error Function

Gradient Descent

1. *An Iterative method to move toward the minimum point(using Gradient).*

2. **Gradient descent** is an optimization algorithm used to minimize some function by iteratively moving in the direction of **steepest descent** as defined by the negative of the **gradient**. In machine learning, we use **gradient descent** to update the parameters of our model.

How we got the formulas of Gradient Descent ?



Updation Rule

Hypothesis: $h_\theta(x) = \theta_0 + \theta_1 x$

repeat until convergence {

$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta_0, \theta_1)$$

(for $j = 1$ and $j = 0$)

}

Derivative Derivation

repeat until convergence {

$$\theta_0 := \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)$$

$$\theta_1 := \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right) \cdot x^{(i)}$$

}

Convergence Criteria

- Number of Iteration

- Change in Error

Let's See the code (Working!)

- Visualizing Line(Hypothesis)
- Behaviour of Error Function using no. of iteration, change in error.

- Variation of Learning rate
- Visualizing Convex Function