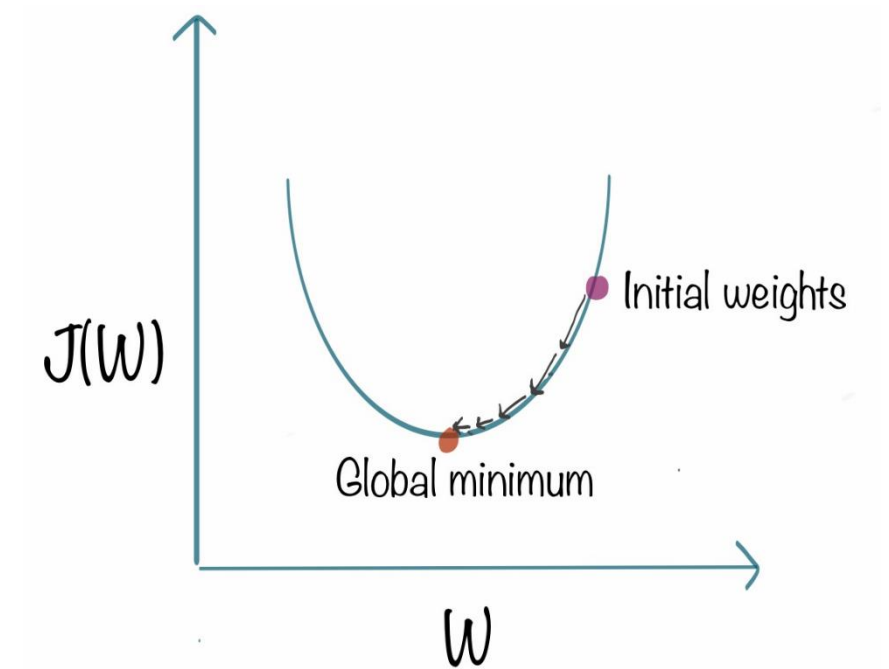


Siddhardhan

Gradient Descent in Machine Learning



Model Parameters

Weights:

Weight decides how much influence the input will have on the output.

$$Y = w * X + b$$

$$Y = w_1 * X_1 + w_2 * X_2 + w_3 * X_3 + b$$

X – feature or input variable

Y – Target or output variable

w – weight

b – bias

Bias:

Bias is the offset value given to the model. Bias is used to shift the model in a particular direction. It is similar to a Y-intercept. 'b' is equal to 'Y' when all the feature values are zero.

Hyperparameters

Learning Rate:

The **Learning Rate** is a tuning parameter in an optimization algorithm that determines the step size at each iteration while moving toward a minimum of a loss function.

Number of Epochs:

Number of Epochs represents the number of times the model iterates over the entire dataset.

Loss Function

Loss function measures how far an estimated value is from its true value.

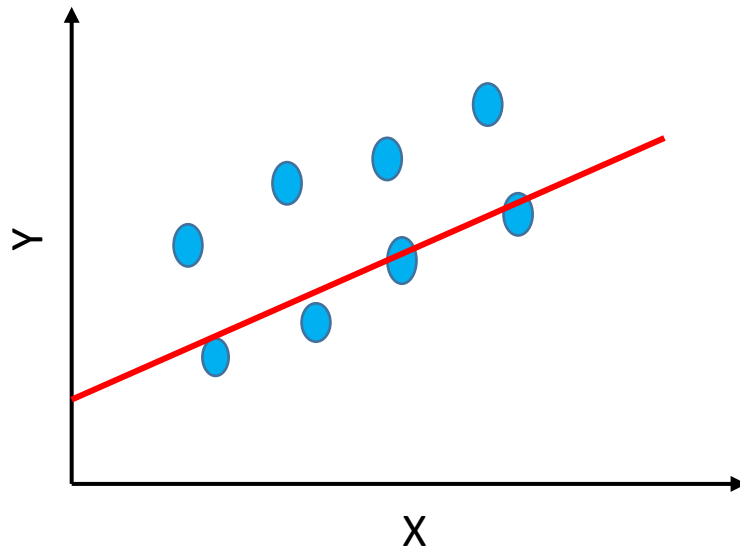
It is helpful to determine which model performs better & which parameters are better.



$$\text{Loss} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Model Optimization

Optimization refers to determining best parameters for a model, such that the loss function of the model decreases, as a result of which the model can predict more accurately.

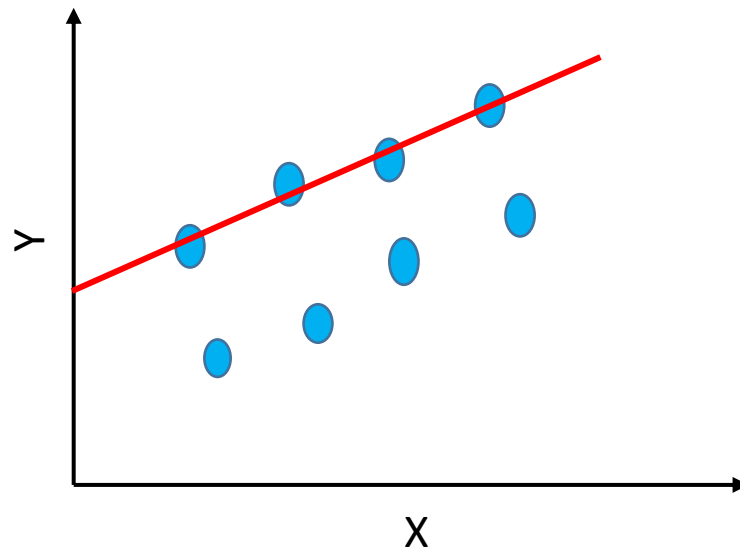


$$Y = w_1X + b_1$$

(w_1 & b_1 are the parameters of the line)

Model Optimization

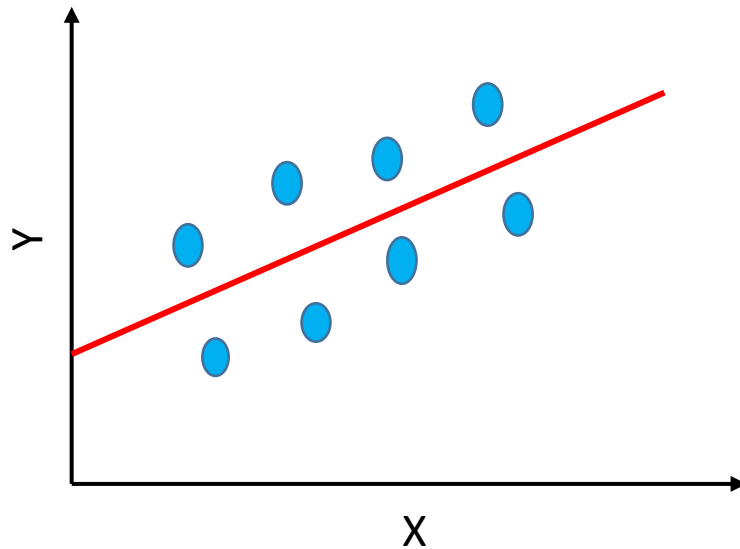
Optimization refers to determining best parameters for a model, such that the loss function of the model decreases, as a result of which the model can predict more accurately.



$$Y = w_2X + b_2$$

Model Optimization

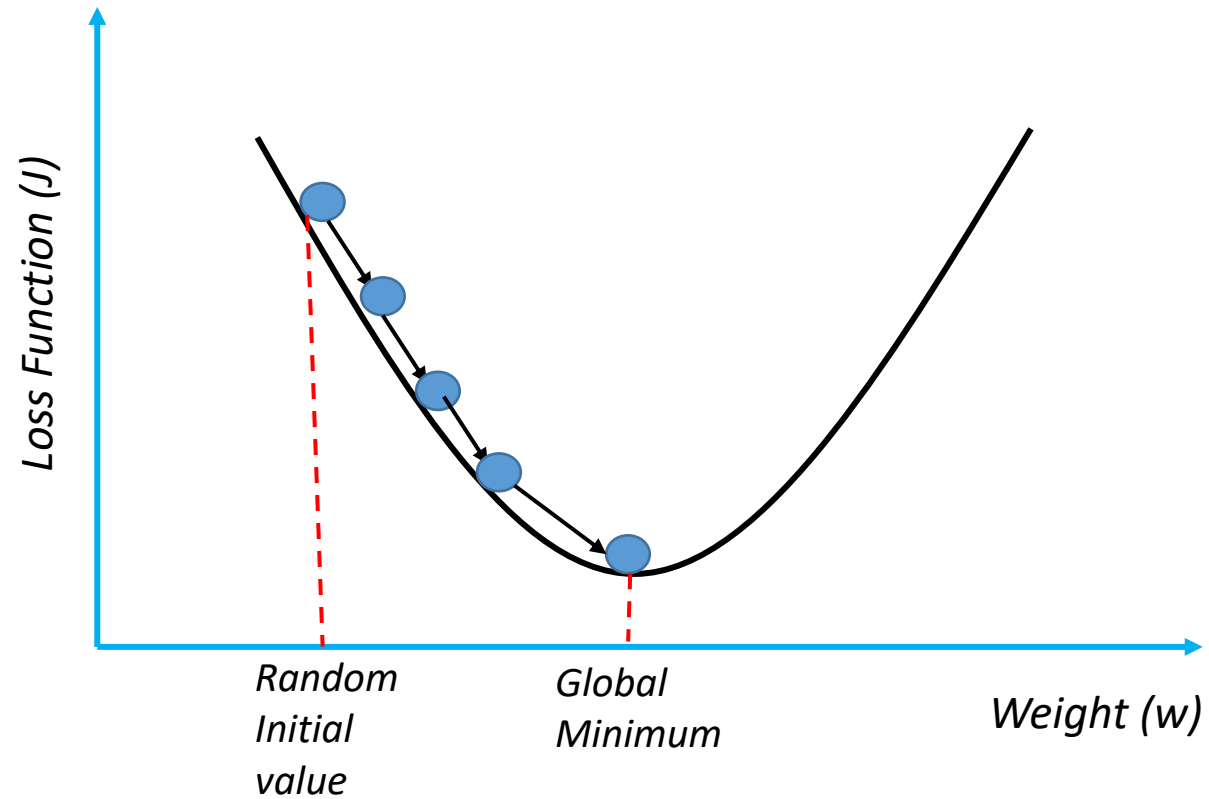
Optimization refers to determining best parameters for a model, such that the loss function of the model decreases, as a result of which the model can predict more accurately.



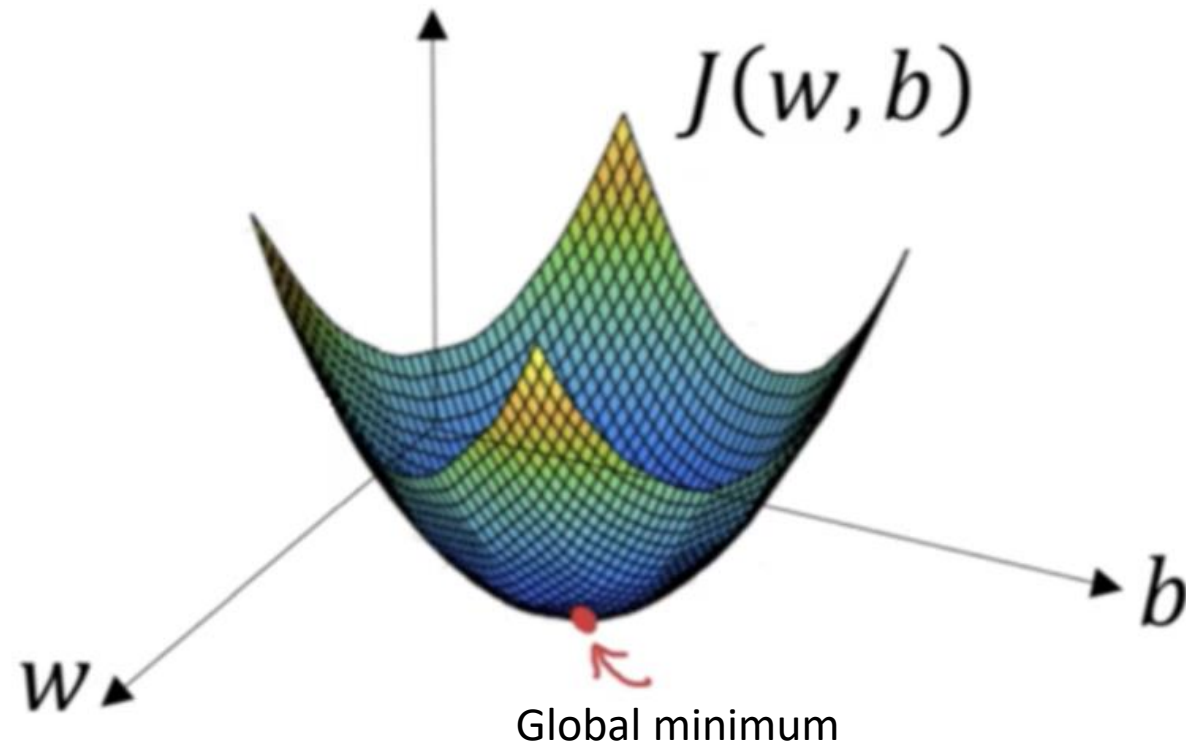
$$Y = w_3X + b_3$$

Hence, w_3 & b_3 are the best parameters

Gradient Descent



Gradient Descent in 3 Dimension



Gradient Descent

Gradient Descent is an optimization algorithm used for minimizing the loss function in various machine learning algorithms. It is used for updating the parameters of the learning model.

$$w_2 = w_1 - L * dw$$

$$b_2 = b_1 - L * db$$

$$\frac{dy}{dx}$$

w --> weight

b --> bias

L --> Learning Rate

dw --> Partial Derivative of loss function with respect to ~~m~~ w

db --> Partial Derivative of loss function with respect to ~~c~~ b