

Back Propagation in deep learning

What is Bias

Back Propagation (reduce Cost function)
(Batch Gradient Descent, Stochastic
Gradient Descent, Mini-batch
Gradient Descent)

What is Bias

Bias is just like an intercept added in a linear equation. It is an additional parameter in the Neural Network which is used to adjust the output along with the weighted sum of the inputs to the neuron. Moreover, bias value allows you to shift the activation function to either right or left.

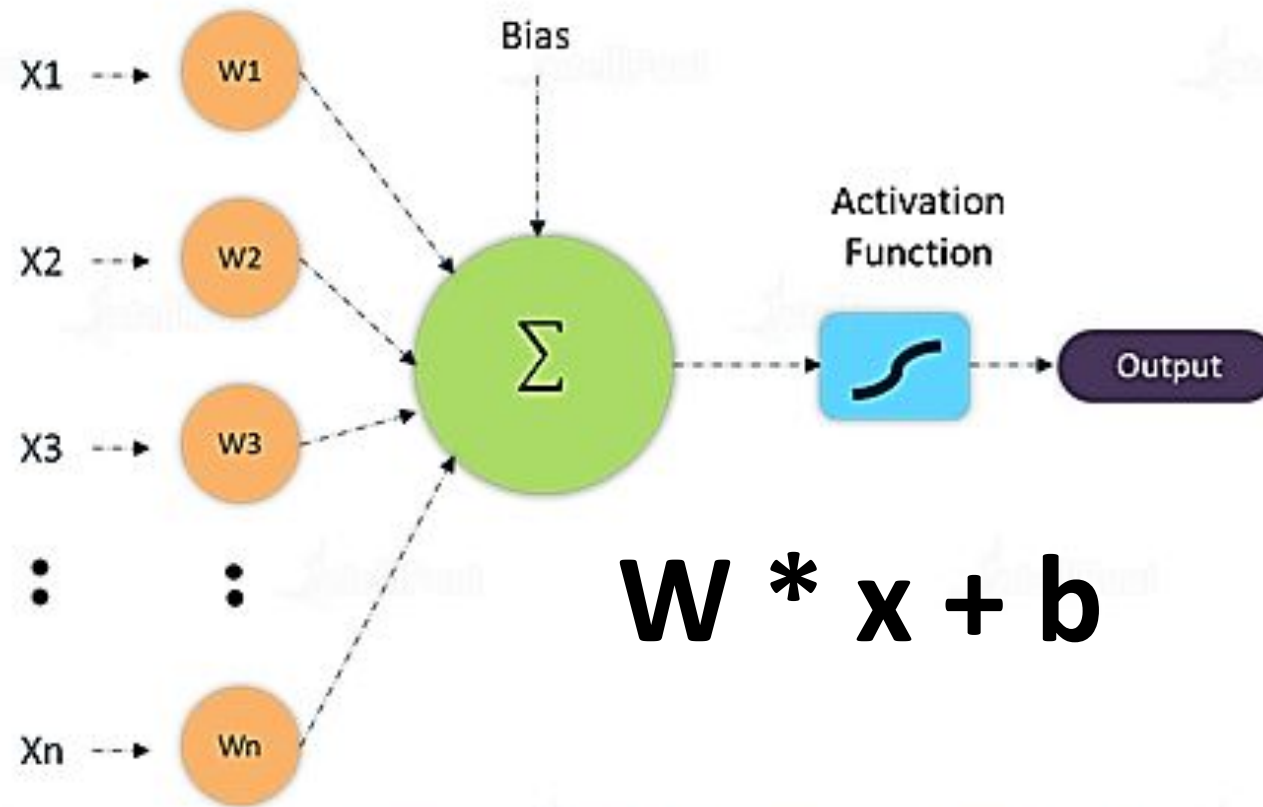
$$\text{output} = \text{sum}(\text{weights} * \text{inputs}) + \text{bias}$$

The output is calculated by multiplying the inputs with their weights and then passing it through an activation function like the Sigmoid function, etc. Here, bias acts like a constant which helps the model to fit the given data. The steepness of the Sigmoid depends on the weight of the inputs.

A simpler way to understand bias is through a constant c of a linear function

$$y = mx + c$$

What is Bias



Schematic Representation of a Neuron in a Neural Network

What is *Gradient Descent* (BGD)

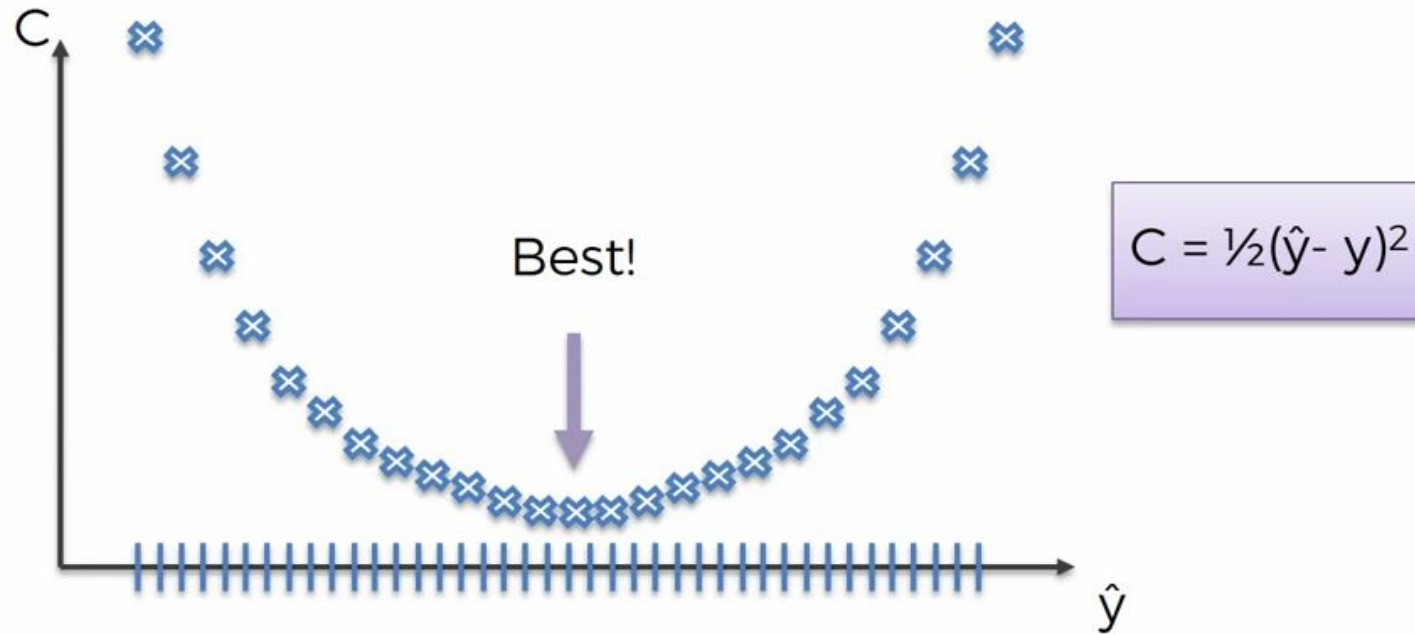
Gradient Descent is an optimization technique that is used to improve deep learning and neural network-based models by minimizing the cost function.

Gradient Descent is a process that occurs in the **backpropagation** phase where the goal is to continuously resample the gradient of the model's parameter in the opposite direction based on the weight w , updating consistently until we reach the **global minimum** of function $J(w)$.

More Precisely,

Gradient descent is an algorithm, which is used to iterate through different combinations of weights in an optimal way....to find the best combination of weights which has a minimum error.

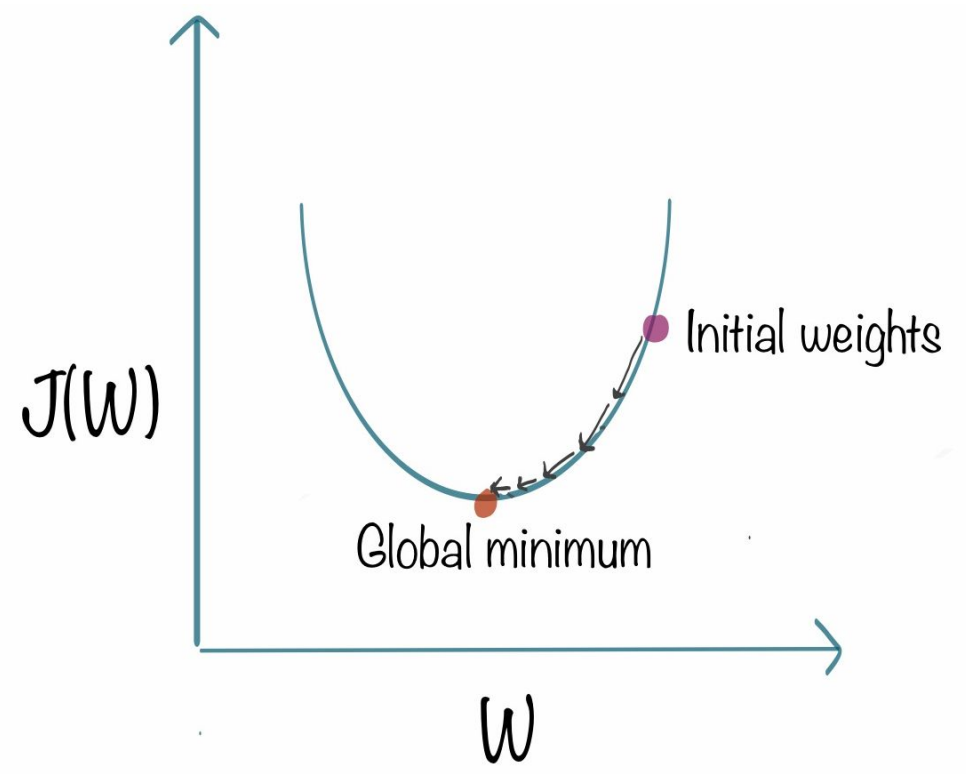
Brute force algorithm



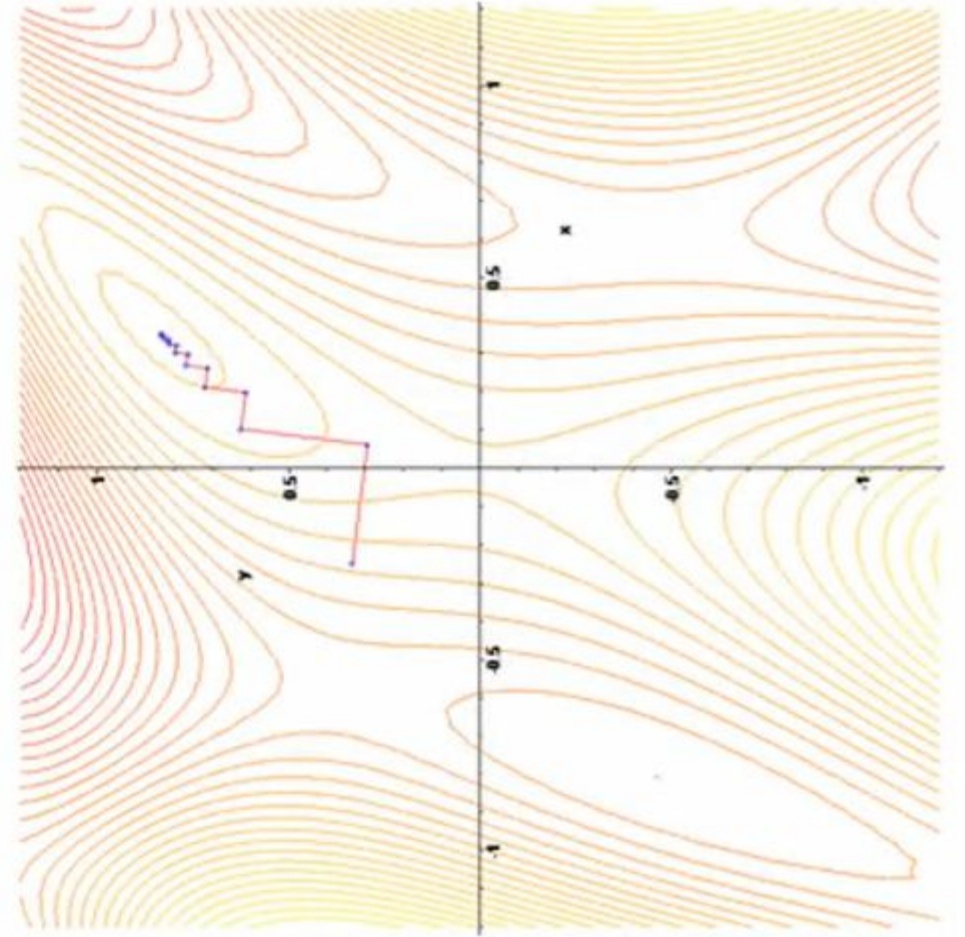
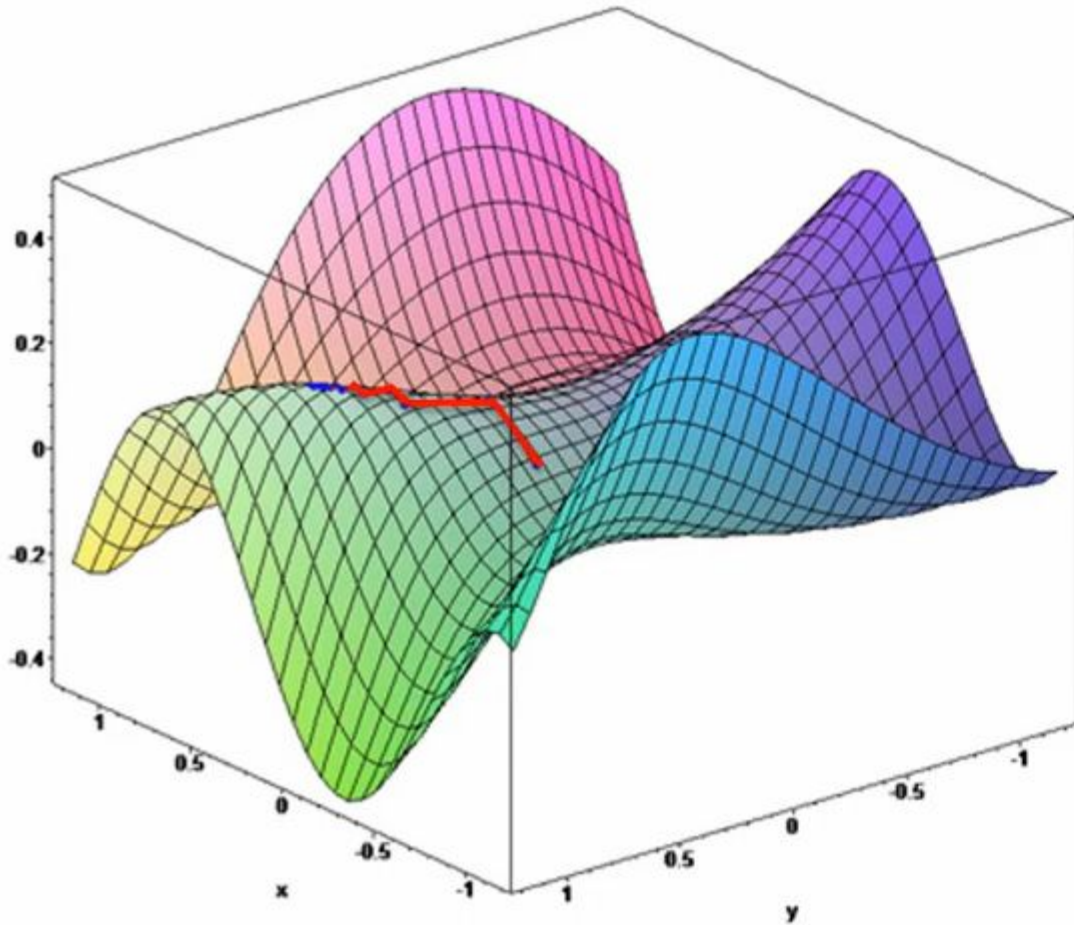
Curse of dimensionality

Brute Force Algorithms refers to a programming style that does not include any shortcuts to improve performance, but instead relies on sheer computing power to try all possibilities until the solution to a problem is found. A classic example is the traveling salesman problem (TSP).

What is *Gradient Descent*



What is *Gradient Descent*



Useful link

<https://towardsdatascience.com/understanding-the-mathematics-behind-gradient-descent-dde5dc9be06e>