**Artificial Intelligence**

Artificial intelligence (AI) is wide-ranging branch of computer science concerned with building smart machines capable of performing tasks that typically require human intelligence. AI is an interdisciplinary science with multiple approaches, but advancements in [machine learning](https://builtin.com/machine-learning) and deep learning are creating a paradigm shift in virtually every sector of the tech industry.

**Machine Learning**

Machine learning (ML) is a type of artificial intelligence ([AI](https://searchenterpriseai.techtarget.com/definition/AI-Artificial-Intelligence)) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning [algorithms](https://whatis.techtarget.com/definition/algorithm) use historical data as input to predict new output values.

**Types of machine learning**

Classical machine learning is often categorized by how an algorithm learns to become more accurate in its predictions. There are four basic approaches: [supervised](https://searchenterpriseai.techtarget.com/definition/supervised-learning) learning, [unsupervised](https://whatis.techtarget.com/definition/unsupervised-learning) learning, semi-supervised learning and reinforcement learning. The type of algorithm a data scientist chooses to use depends on what type of data they want to predict.

**Supervised learning:**

 In this type of machine learning, [data scientists](https://searchenterpriseai.techtarget.com/definition/data-scientist) supply algorithms with labeled training data and define the variables they want the algorithm to assess for correlations. Both the input and the output of the algorithm is specified.

**Unsupervised** **learning**:

This type of machine learning involves algorithms that train on unlabeled data. The algorithm scans through data sets looking for any meaningful connection. Both the data algorithms train on and the predictions or recommendations they output are predetermined.

**Semi-supervised learning:**

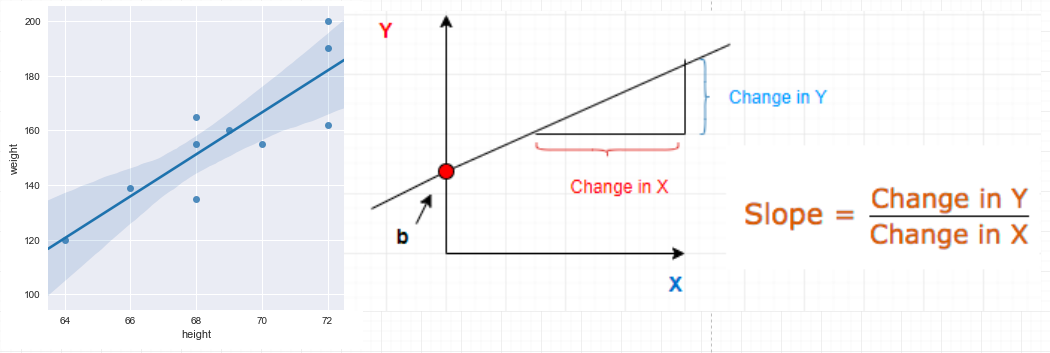
 This approach to machine learning involves a mix of the two preceding types. Data scientists may feed an algorithm mostly labeled [training data](https://searchenterpriseai.techtarget.com/feature/Using-small-data-sets-for-machine-learning-models-sees-growth), but the model is free to explore the data on its own and develop its own understanding of the data set.

**Reinforcement learning:**

[Reinforcement learning](https://searchenterpriseai.techtarget.com/definition/reinforcement-learning) is typically used to teach a machine to complete a multi-step process for which there are clearly defined rules. Data scientists program an algorithm to complete a task and give it positive or negative cues as it works out how to complete a task. But for the most part, the algorithm decides on its own what steps to take along the way.

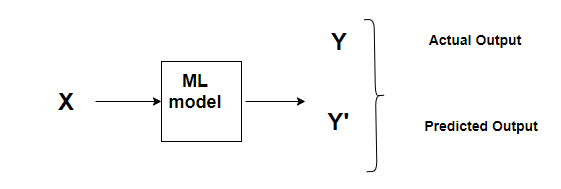
**Machine Learning Model**

* Let us draw an arbitrary line in space that passes through some of the data points. The equation of this straight line would be **Y = mX + b** where **m** is the slope and **b** is its intercept on the Y-axis.



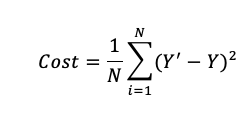
**Predictions**

Given a known set of inputs and their corresponding outputs, A machine learning model tries to make some predictions for a new set of inputs.



# Cost Function

A **Cost Function/Loss Function** evaluates the performance of our Machine Learning Algorithm. The**Loss function** computes the error for a single training example while the **Cost functi**on is the average of the loss functions for all the training examples. Henceforth, I shall be using both the terms interchangeably.



**Learning Rate:**

This size of steps taken to reach the minimum or bottom is called **Learning Rate**. We can cover more area with larger steps/higher learning rate but are at the risk of overshooting the minimal. On the other hand, small steps/smaller learning rates will consume a lot of time to reach the lowest point.

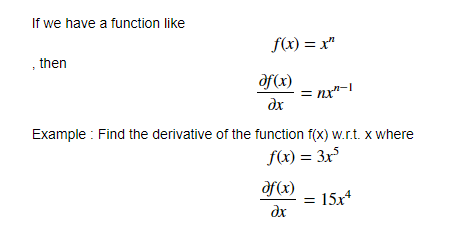
## Step-by-step:

There are two parameters in our cost function we can control: mm (weight) and bb (bias). Since we need to consider the impact each one has on the final prediction, we need to use partial derivatives. We calculate the partial derivatives of the cost function with respect to each parameter and store the results in a gradient.

## Derivatives

Machine learning uses derivatives in optimization problems. Optimization algorithms like gradient descent use derivates to actually decide whether to increase or decrease the weights in order to increase or decrease any objective function.

Say for instance we have a function:



## Gradient Descent:

At a theoretical level, gradient descent is an algorithm that minimizes functions. Given a function defined by a set of parameters, gradient descent starts with an initial set of parameter values and iteratively moves toward a set of parameter values that minimize the function. This iterative minimization is achieved using calculus, taking steps in the negative direction of the function [gradient](http://en.wikipedia.org/wiki/Gradient).It involves the following steps:

1. Initialize parameters (Weights and bias) at random position or simply as zero.
2. Calculate cost function (J).
3. Take the partial derivative of the cost function with respect to Weights and bias (dW and db).
4. Change parameters values as:
   * Wnew = W – learning rate \* dW
   * Bnew = b – learning rate \* db
5. Again, start from step 2 with new values of W and b and repeat the same for ‘n’ no. of iterations. With each iteration, the value of cost will progressively decrease and eventually end up with flat value.