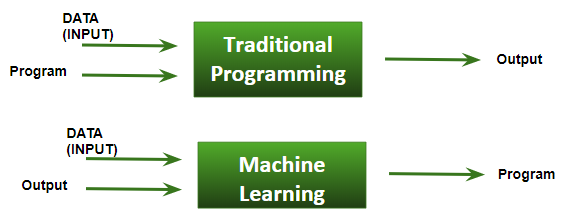
**Machine Learning**

**What Is Machine Learning ?**

* Machine learning is a type of artificial intelligence (AI) that provides computers with the ability to learn without being explicitly programmed.
* Machine learning focuses on the development of Computer Programs that can change when exposed to new data.
* Machine learning involves computer to get trained using a given data set, and use this training to predict the properties of a given new data.



**Machine Learning – Applications**

* **Web Search Engine:** One of the reasons why search engines like google, bing etc work so well is because the system has learnt how to rank pages through a complex learning algorithm.
* **Photo tagging Applications:** Be it facebook or any other photo tagging application, the ability to tag friends makes it even more happening. It is all possible because of a face recognition algorithm that runs behind the application.
* **Spam Detector:** Our mail agent like Gmail or Hotmail does a lot of hard work for us in classifying the mails and moving the spam mails to spam folder. This is again achieved by a spam classifier running in the back end of mail application

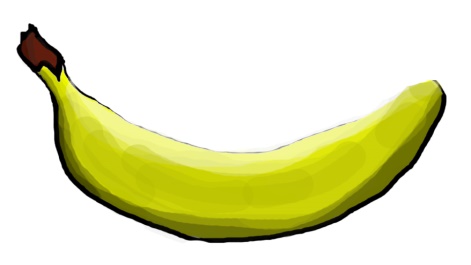
**How ML works?**

* Gathering past data in the form of text file, excel file, images or audio data. The more better the quality of data, the better will be the model learning
* Data Processing – Sometimes, the data collected is in the raw form and it needs to be rectified.  
  Example: if data has some missing values, then it has to be rectified. If data is in the form of text or images then converting it to numerical form will be required, be it list or array or matrix. Simply, Data is to be made relevant and understandable by the machine
* Building up models with suitable algorithms and techniques and then training it.
* Testing our prepared model with data which was not feed in at the time of training and so evaluating the performance – score, accuracy with high level of precision

**Pre-requisites to learn ML:**

* + Linear Algebra
  + Statistics and Probability
  + Calculus
  + Graph theory
  + Programming Skills – Language such as Python, R, MATLAB, C++ or Octave

**Supervised learning**

* Supervised learning as the name indicates a presence of supervisor as teacher.
* Basically supervised learning is a learning in which we teach or train the machine using data which is well labeled that means some data is already tagged with correct answer.
* After that, machine is provided with new set of examples(data) so that supervised learning algorithm analyses the training data(set of training examples) and produces an correct outcome from labeled data.
* **For instance**, suppose you are given an basket filled with different kinds of fruits. Now the first step is to train the machine with all different fruits one by one like this:  
  
* If shape of object is rounded and depression at top having color Red then it will be labelled as –**Apple**.
* If shape of object is long curving cylinder having color Green-Yellow then it will be labelled as –Banana.
* Now suppose after training the data, you have given a new separate fruit say Banana from basket and sked to identify it.  
  
* Since machine has already learnt the things from previous data and this time have to use it wisely.
* It will first classify the fruit with its shape and color, and would confirm the fruit name as BANANA and put it in Banana category.
* Thus machine learns the things from training data(basket containing fruits) and then apply the knowledge to test data(new fruit).

**Supervised learning classified into two categories of algorithms:**

* **Regression**: A regression problem is when the output variable is a real value, such as “dollars” or “weight”.
  + Linear Regression
  + Multiple Regression
  + Polynomial Regression
* **Classification:** A classification problem is when the output variable is a category, such as “Red” or “blue” or “disease” and “no disease”.
  + Logistic Regression
  + K-Nearest Neighbors
  + Support Vector Machines (SVM) & Kernel SVM

**Unsupervised learning**

* Unsupervised learning is the training of machine using information that is neither classified nor labeled and allowing the algorithm to act on that information without guidance.
* Here the task of machine is to group unsorted information according to similarities, patterns and differences without any prior training of data.
* Unlike supervised learning, no teacher is provided that means no training will be given to the machine.
* Therefore machine is restricted to find the hidden structure in unlabeled data by our-self.
* **For instance**, suppose it is given an image having both dogs and cats which have not seen ever.  
  
* Thus machine has no any idea about the features of dogs and cat so we can’t categorize it in dogs and cats.
* But it can categorize them according to their similarities, patterns and differences i.e., we can easily categorize the above picture into two parts.
* First first may contain all pics having **dogs** in it and second part may contain all pics having **cats** in it. Here you didn’t learn anything before, means no training data or examples.

**Unsupervised learning classified into two categories of algorithms:**

* **Clustering**: A clustering problem is where you want to discover the inherent groupings in the data, such as grouping customers by purchasing behavior.
  + K-Means Clustering
  + Hierarchial Clustering
* **Association**: An association rule learning problem is where you want to discover rules that describe large portions of your data, such as people that buy X also tend to buy Y.