Machine Learning

* New Machine learning are different from the old Machine Learning, new machine do not need programmer to give out specific directions. They learn from previous computations to produce reliable, repeatable decisions and results
* Since now we have cheaper and more powerful computational processing and affordable data storage it’s possible to quickly and automatically produce models that can analyze bigger and more complex data, deliver a better result.
* Requirement for Machine Learning Systems
  + Data Preparation capabilities
  + Algorithms
  + Automation and iterative processes
  + Scalability
  + Ensemble Model
* Facts:
  + In machine learning, a target is called a label.
  + In statistics, a target is called a dependent variable.
  + A variable in statistics is called a feature in machine learning.
  + A transformation in statistics is called feature creation in machine learning.
* Machine Learning Methods
  + Supervised learning: Inferring a function from labeled training data. The data consist a set of training examples. Example is a pair consisting of an input object and a desired output value

(Giving a set call training data that the program can use as a reference, programmer need to be able to predicted the reaction from the program.)

There are couple step for supervised learning:

1. Programmer need to make the decision of what kind of data type it’s going to be put in as the training set.
2. The training sets, data, need to match the fact, the object in real world. Need to make senses
3. Determine the input feature representation of the learned function. The input feature is important that the program would miss understand if programmer don’t represent the input as they want. The input need enough feature for function to work with but too much would confuse the program.
4. Determine what kind of data structure the data would be store in
5. After all the design, programmer test the algorithmic and either do major change or minor change

* Unsupervised Learning: no label are given to the program, so the system do not know what’s the right answer, the algorithmic need to figure it out on its own. The goal is to explore the data and find some structure within.
* Semi-supervised Learning: it is the type of learning fall between supervised learning and unsupervised learning. Scientist find out that unlabeled data, when used in conjunction with a small amount of labeled data. can produce considerable improvement in learning accuracy
  1. Improve supervised learning algorithms
  2. Efficient computation speed
* Reinforcement Learning: inspired by behaviorist psychology, the program interact with a dynamic environment that it has a goal and it will fix itself base on the output and human reactions. The different between supervised learning and reinforcement learning is that there is no simple input given they are focus more on exploration and exploitation. The algorithm discovers through trial and error get the greatest rewards
* Data Mining: process of discovering patterns in large data sets, has been use in different fields
* Machine Learning: goal is to understand the structure of the data a theory behind the model that is mathematically proven. Machine learning model is a validation error on new data, not a theoretical test that proves a null hypothesis. Often uses an iterative approach to learn new data.
* Deep learning: identify objects in images and words in sound are both deep learning, still in development, but a lot of business is ready to apply the technique.
  1. Algorithmic improvements have boosted the performance of deep learning methods
  2. Approach of deep learning increase the accuracy of the models
  3. New class of neural networks has developed that fit well for applications like text translation and image classification
  4. Human to machine interface has been approve by using gesture swipe touch and natural language
  5. Speech Recognition, Image Recognition, Natural Language Processing, Recommendation Systems
  6. Compare to Machine Learning it moves from telling the computer how to solve a problem to training the computer to solve the problem itself
* Artificial Neural Networks: computer systems inspired by biological neural networks, do task by considering examples, without task-specific programming
* Machine learning and statistic are closely related fields
* 3 step of deep learning:
  1. Design the functions and structure of the program
     + Define sets of functions(linear algebra
     + Goodness of functions
     + Pick the best function
  2. Training of Deep learning