

Machine learning is one

of the most exciting recent technologies. And in this class, you learn about the state of the art and also gain practice implementing and deploying these algorithms yourself. You've probably use a learning algorithm dozens of times a day without knowing it. Every time you use a web search engine like Google or Bing to search the internet, one of the reasons that works so well is because a learning algorithm, one implemented by Google or Microsoft, has learned how to rank web pages. Every time you use Facebook or Apple's photo typing application and it recognizes your friends' photos, that's also machine learning. Every time you read your email and your spam filter saves you from having to wade through tons of spam email, that's also a learning algorithm. For me one of the reasons I'm excited is the AI dream of someday building machines as intelligent as you or me. We're a long way away from that goal, but many AI researchers believe that the best way to towards that goal is through learning algorithms that try to mimic how the human brain learns. I'll tell you a little bit about that too in this class. In this class you learn about state-of-the-art machine learning algorithms. But it turns out just knowing the algorithms and knowing the math isn't that much good if you don't also know how to actually get this stuff to work on problems that you care about. So, we've also spent a lot of time developing exercises for you to

implement each of these algorithms and see how they work for yourself. So why is machine learning so prevalent today? It turns out that machine learning is a field that had grown out of the field of AI, or artificial intelligence. We wanted to build intelligent machines and it turns out that there are a few basic things that we could program a machine to do such as how to find the shortest path from A to B. But for the most part we just did not know how to write AI programs to do the more interesting things such as web search or photo tagging or email anti-spam. There was a realization that the only way to do these things was to have a machine learn to do it by itself. So, machine learning was developed as a new capability for computers and today it touches many segments of industry and basic science. For me, I work on machine learning and in a typical week I might end up talking to helicopter pilots, biologists, a bunch of computer systems people (so my colleagues here at Stanford) and averaging two or three times a week I get email from people in industry from Silicon Valley contacting me who have an interest in applying learning algorithms to their own problems. This is a sign of the range of problems that machine learning touches. There is autonomous robotics, computational biology, tons of things in Silicon Valley that machine learning is having an impact on. Here are some other examples of machine learning. There's database mining. One of the reasons machine learning has so pervaded is the growth of the web and the growth of automation All this means that we have much larger data sets than ever before. So, for example tons of Silicon Valley companies are today collecting web click data, also called clickstream data, and are trying to use machine learning

algorithms to mine this data to understand the users better and to serve the users better, that's a huge segment of Silicon Valley right now. Medical records. With the advent of automation, we now have electronic medical records, so if we can turn medical records into medical knowledge, then we can start to understand disease better. Computational biology. With automation again, biologists are collecting lots of data about gene sequences, DNA sequences, and so on, and machines running algorithms are giving us a much better understanding of the human genome, and what it means to be human. And in engineering as well, in all fields of engineering, we have larger and larger, and larger and larger data sets, that we're trying to understand using learning algorithms. A second range of machinery applications is ones that we cannot program by hand. So for example, I've worked on autonomous helicopters for many years. We just did not know how to write a computer program to make this helicopter fly by itself. The only thing that worked was having a computer learn by itself how to fly this helicopter. [Helicopter whirling] Handwriting recognition. It turns out one of the reasons it's so inexpensive today to route a piece of mail across the countries, in the US and internationally, is that when you write an envelope like this, it turns out there's a learning algorithm that has learned how to read your handwriting so that it can automatically route this envelope on its way, and so it costs us a few cents to send this thing thousands of miles. And in fact if you've seen the fields of natural language processing or computer vision, these are the fields of AI pertaining to understanding language or understanding images. Most of natural language processing and most of computer vision today is applied machine learning. Learning algorithms are also widely used for self- customizing programs. Every time you go to Amazon or Netflix or iTunes Genius, and it recommends the movies or products and music to you, that's a learning algorithm. If you think about it they have million users; there is no way to write a million different programs for your million users. The only way to have software give these customized recommendations is to become learn by itself to customize itself to your preferences. Finally learning algorithms are being used today to understand human learning and to understand the brain. We'll talk about how researches are using this to make progress towards the big AI dream. A few months ago, a student showed me an article on the top twelve IT skills. The skills that information technology hiring managers cannot say no to. It was a slightly older article, but at the top of this list of the twelve most desirable IT skills was machine learning. Here at Stanford, the number of recruiters that contact me asking if I know any graduating machine learning students is far larger than the machine learning students we graduate each year. So I think there is a vast, unfulfilled demand for this skill set, and this is a great time to be learning about machine learning, and I hope to teach you a lot about machine learning in this class. In the next video, we'll start to give a more formal definition of what is machine learning. And we'll begin to talk about the main types of machine learning problems and algorithms. You'll pick up some of the main machine learning terminology, and

start to get a sense of what are the different algorithms, and when each one might be appropriate.

There has been a renewed interest in machine learning in last few years. This revival seems to be driven by strong fundamentals – loads of data being emitted by sensors across the globe, with cheap storage and lowest ever computational costs!

Coming to the point, given the amount of confusion on the topic, we thought to create an awesome introductory series of articles on machine learning. The idea is to do away with all the jargons, which might have intimidated you in past and create something which can be read by a 5 year old (ok...sorry, may be a high school pass out)!

## **So what exactly is machine learning? My small experiment...**

Just to make sure I don't over-estimate (or under-estimate) the capability of the target audience, I got hold of 10 people who were completely new to analytics. None of them had heard about machine learning before (yes, there are people like that!). Here is what they said:

- I don't know, may be learning from machines?
- Making machines learn something a.k.a. programming machine software
- Learning with help of computers
- Learning through online courses (!!!)

That was fun! Perfect group to explain machine learning to. Here is how I started explaining to these people:

*Machine Learning refers to the techniques involved in dealing with vast data in the most intelligent fashion (by developing algorithms) to derive actionable insights.*

By this time, they were looking at me as if I have spoken a few things in front of people from Mars! So, I stopped and then asked them a question in return, which they could relate to more:

**KJ:** *What do you think happens when you search for something on Google?*

**Group:** *Google shows up the most relevant web pages related to that search.*

*KJ: That's good! but what really happens so that Google can show these relevant pages to you?*

This time it looked like they were thinking a bit more. Then some one from the group spoke

***Group member:** Google looks at the past clicks from the people to understand which pages are more relevant for those searches and then serves those results on top of search.*

This was a far better attempt. I also had to control my urge to preach that how Google does this is far more smarter way than this simple concept. But, I thought I had a good hook to explain machine learning here. So, I continued:

*KJ: OK, that sounds like a good approach. But, how many searches and what all kind of searches would Google handle regularly?*

***Group:** Must be a real big number – may be a trillion searches every year*

*KJ: So, how do you think Google can serve so many requests with such accuracy? Do you think there are people sitting in Google offices and continuously deciding which search result is relevant and which is not?*

***Group member:** Haven't really thought about it, but no, that sounds humanly impossible to do.*

*KJ: You are right. This is where machine learning comes into play. **Machine learning is a set of techniques, which help in dealing with vast data in the most intelligent fashion (by developing algorithms or set of logical rules) to derive actionable insights (delivering search for users in this case).***

A logical nod from the group, looks like mission accomplished...yay! But wait...

## **Now the common question – How is machine learning different from X?**

The minute you start reading about machine learning, you see various rockets bombarding you with high velocity. These are jargons used loosely in the industry. Here are some of them: Artificial Intelligence, Deep Learning, Data Mining and Statistics.

For your clear understanding, I have explained these terms below in the simple manner. You will also understand the importance of these terms in context of machine learning:

## **X = Artificial Intelligence(AI):**

It refers to the procedure of *programming* a computer (machine) to take *rational*. Ah! what is rational? Rational is the basis of taking a decision.

I mentioned 'rational' instead of intelligence (as expected) because we human beings tend to take decisions which are high on being rational and feasible rather than being explicitly intelligent. This is because *all intelligent decisions needn't be rational and feasible* (my hypothesis). Hence, the central motive behind using AI is to achieve the computer (machine) behave in a dandy fashion in lieu of human guidance instead of being doltish!

AI may include programs to check whether certain parameters within a program are behaving normally. For example, the machine may raise an alarm if a parameter say 'X' crosses a certain threshold which might in turn affect the outcome of the related process.

### **Use of Artificial Intelligence in Machine Learning**

*Machine Learning is a subset of AI* where the machine is trained to learn from its past experience. The past experience is developed through the data collected. Then it combines with algorithms such as Naïve Bayes, Support Vector Machine(SVM) to deliver the final results.

## **X = Statistics:**

At this high level stage, I assume you would know about statistics. If you don't, here's a quick definition, Statistics is that branch of mathematics which utilizes data, either of the entire population or a sample drawn from the population to carry out the analysis and present inferences. Some statistical techniques used are regression, variance, standard deviation, conditional probability and many others. T

### **Use of Statistics in Machine Learning**

Let's understand this. Suppose, I need to separate the mails in my inbox into two categories: 'spam' and 'important'. For identifying the spam mails, I can use a machine learning algorithm known as Naïve Bayes which will check the frequency of the past spam mails to identify the new email as spam. Naïve Bayes uses the statistical technique Baye's theorem( commonly known as conditional probability). Hence, we can say machine learning algorithms uses statistical concepts to execute machine learning.

Additional Information: The main difference between machine learning and statistical models come from the schools where they originated. While machine learning originated from the department of computer science and statistical modelling came down from department of

mathematics. Also any statistical modelling assumes a number of distributions while machine learning algorithms are generally agnostic of the distribution of all attributes.

### **X = Deep Learning:**

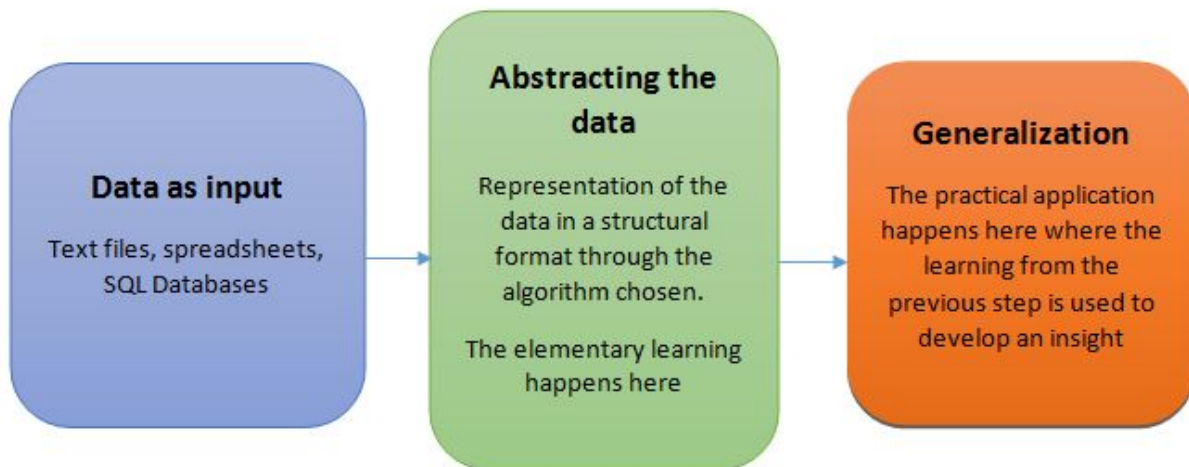
Deep Learning is associated with a machine learning algorithm (Artificial Neural Network, ANN) which uses the concept of human brain to facilitate the modeling of arbitrary functions. ANN requires a vast amount of data and this algorithm is highly flexible when it comes to model multiple outputs simultaneously. ANN is more complex topic and we may do justice to it in an altogether separate article.

### **X = Data Mining:**

During my initial days as an analyst, I always used to muddle the two terms: Machine Learning and Data Mining. But, later I learnt, Data Mining deals with searching specific information. And Machine Learning solely concentrates on performing a given task. Let me cite the example which helped me to remember the difference; Teaching someone how to dance is Machine Learning. And using someone to find best dance centers in the city is Data Mining. Easy!

## **But, How exactly do we teach machines?**

Teaching the machines involve a structural process where every stage builds a better version of the machine. For simplification purpose, the process of teaching machines can broken down into 3 parts:



I shall be covering each of these 3 steps in detail in my subsequent write-ups. As of now, you should understand, these 3 steps ensures the holistic learning of the machine to perform the given task with equal importance. Success of machine depends on two factors:

1. *How well the generalization of abstraction data take place.*
2. *How well the machine is able to put it's learning into practical usage for predicting the future course of action.*

## What are the steps used in Machine Learning?

There are 5 basic steps used to perform a machine learning task:

1. **Collecting data:** Be it the raw data from excel, access, text files etc., this step (gathering past data) forms the foundation of the future learning. The better the variety, density and volume of relevant data, better the learning prospects for the machine becomes.
2. **Preparing the data:** Any analytical process thrives on the quality of the data used. One needs to spend time determining the quality of data and then taking steps for fixing issues such as missing data and treatment of outliers. Exploratory analysis is perhaps one method to study the nuances of the data in details thereby burgeoning the nutritional content of the data.
3. **Training a model:** This step involves choosing the appropriate algorithm and representation of data in the form of the model. The cleaned data is split into two parts –

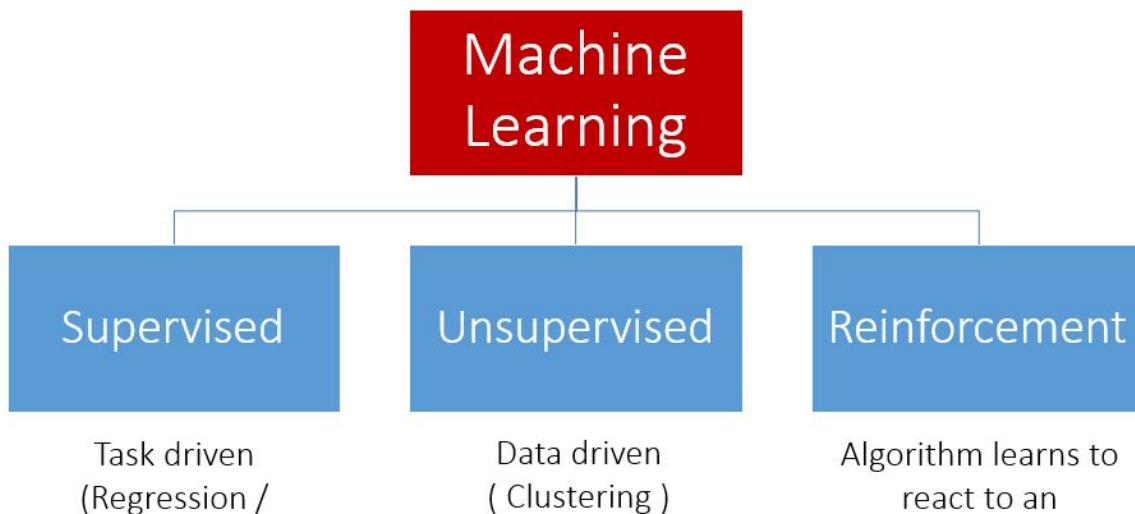
train and test (proportion depending on the prerequisites); the first part (training data) is used for developing the model. The second part (test data), is used as a reference.

4. **Evaluating the model:** To test the accuracy, the second part of the data (holdout / test data) is used. This step determines the precision in the choice of the algorithm based on the outcome. A better test to check accuracy of model is to see its performance on data which was not used at all during model build.
5. **Improving the performance:** This step might involve choosing a different model altogether or introducing more variables to augment the efficiency. That's why significant amount of time needs to be spent **in data collection and preparation**.

Be it any model, these 5 steps can be used to structure the technique and when we discuss the algorithms, you shall then find how these five steps appear in every model!

## What are the types of Machine Learning algorithms?

### Types of Machine Learning



#### Supervised Learning / Predictive models:

Predictive model as the name suggests is used to predict the future outcome based on the historical data. Predictive models are normally given clear instructions right from the beginning as in what needs to be learnt and how it needs to be learnt. These class of learning algorithms are termed as **Supervised Learning**.

For example: Supervised Learning is used when a marketing company is trying to find out which customers are likely to churn. We can also use it to predict the likelihood of occurrence of perils like earthquakes, tornadoes etc. with an aim to determine the Total Insurance Value. Some



examples of algorithms used are: Nearest neighbour, Naïve Bayes, Decision Trees, Regression etc.

### **Unsupervised learning / Descriptive models:**

It is used to train descriptive models where no target is set and no single feature is important than the other. The case of unsupervised learning can be: When a retailer wishes to find out what are the combination of products, customers tends to buy more frequently. Furthermore, in pharmaceutical industry, unsupervised learning may be used to predict which diseases are likely to occur along with diabetes. Example of algorithm used here is: K- means Clustering Algorithm

### **Reinforcement learning (RL):**

It is an example of machine learning where the machine is trained to take specific decisions based on the business requirement with the sole motto to maximize efficiency (performance). The idea involved in reinforcement learning is: The machine/ software agent trains itself on a continual basis based on the environment it is exposed to, and applies it's enriched knowledge to solve business problems. This continual learning process ensures less involvement of human expertise which in turn saves a lot of time!

An example of algorithm used in RL is Markov Decision Process.

Important Note: There is a subtle difference between Supervised Learning and Reinforcement Learning (RL). RL essentially involves learning by interacting with an environment. An RL agent learns from its past experience, rather from its continual trial and error learning process as against supervised learning where an external supervisor provides examples.

A good example to understand the difference is self driving cars. Self driving cars use Reinforcement learning to make decisions continuously – which route to take? what speed to drive on? are some of the questions which are decided after interacting with the environment. A simple manifestation for supervised learning would be to predict fare from a cab going from one place to another.

## **What are the applications of Machine Learning?**

It is very interesting to know the applications of machine learning. Google and Facebook uses ML extensively to push their respective ads to the relevant users. Here are a few applications that you should know:

- **Banking & Financial services:** ML can be used to predict the customers who are likely to default from paying loans or credit card bills. This is of paramount importance as machine learning would help the banks to identify the customers who can be granted loans and credit cards.
- **Healthcare:** It is used to diagnose deadly diseases (e.g. cancer) based on the symptoms of patients and tallying them with the past data of similar kind of patients.
- **Retail:** It is used to identify products which sell more frequently (fast moving) and the slow moving products which help the retailers to decide what kind of products to introduce or remove from the shelf. Also, machine learning algorithms can be used to find which two / three or more products sell together. This is done to design customer loyalty initiatives which in turn helps the retailers to develop and maintain loyal customers.

These examples are just the tip of the iceberg. Machine learning has extensive applications practically in every domain. You can check out a few Kaggle problems to get further flavor. The examples included above are easy to understand and at least give a taste of the omnipotence of machine learning.