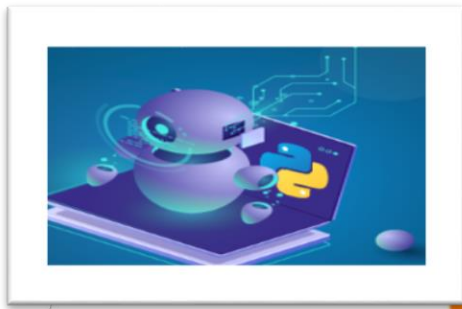
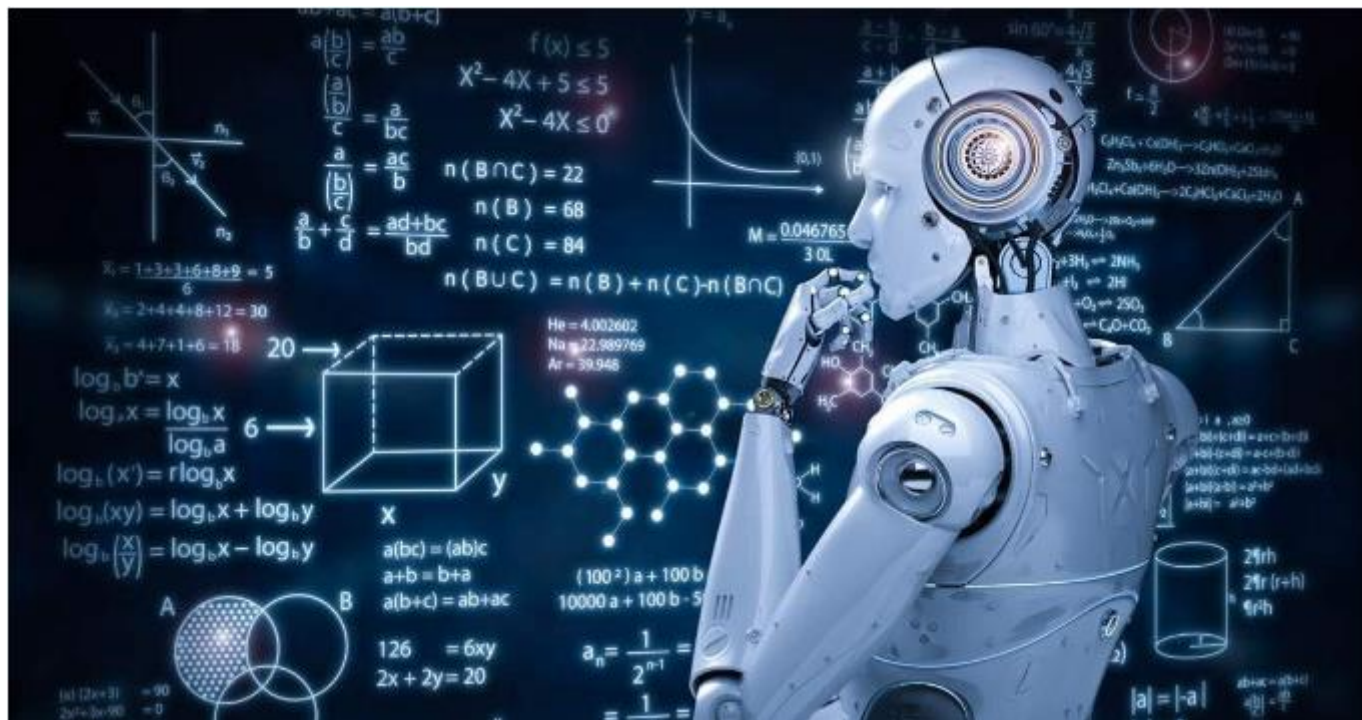


# Data science and Machine learning with Python



Designed by Abdur Rahman Joy - MCSD, MCPD, MCSE, MCTS, OCJP, Sr. Technical Trainer for VFX at IDB BISW (Scholarship program), and C#.net, R, Scala, Kotlin, JAVA, Android/IOS/Windows Mobile Apps, SQL server, Azure, Oracle, SharePoint Development, AWS , CEH, KALI Linux, Python, Data Science, Machine Learning ,Software Testing, Graphics, Multimedia and Game Developer at Joy Infosys and other premises like BITM, SkillsJob, PNTL, Leads Training and New Horizon inc , Cell #: +880-1712587348, email: [jspaonline@gmail.com](mailto:jspaonline@gmail.com). Web URL: <http://www.joyinfosys.com/me>.



## ***Introduction to Machine Learning***

### **What is machine learning?**

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

There are multiple definitions exists of machine learning depending upon the understanding. So let's see some of the top definitions about machine learning:

The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

Machine learning is a fundamental subdivision field of artificial intelligence. It is the science of getting computers to learn and act like humans do, and improve their learning over time in autonomous fashion, by feeding them data and information in the form of observations and real-world interactions. When exposed to new data, these computer programs are enabled to learn, change, develop, and grow by themselves. In short; more the data, higher will be the accuracy to learn and predict the results.

Machine learning is a method of data analysis that automates analytical model building. It is a branch of artificial intelligence based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention.

---

Designed by Abdur Rahman Joy - MCSO, MCPD, MCSE, MCTS, OCJP, Sr. Technical Trainer for VFX at IDB BISW (Scholarship program), and C#.net, R, Scala, Kotlin, JAVA, Android/iOS/Windows Mobile Apps, SQL server, Azure, Oracle, SharePoint Development, AWS, CEH, KALI Linux, Python, Data Science, Machine Learning, Software Testing, Graphics, Multimedia and Game Developer at Joy Infosys and other premises like BITM, SkillsJob, PNTL, Leads Training and New Horizon inc, Cell #: +880-1712587348, email: [ispaonline@gmail.com](mailto:ispaonline@gmail.com). Web URL: <http://www.joyinfosys.com/me>.

Machine learning is a form of artificial intelligence that uses algorithms to enable a system to learn from data rather than through explicit programming

Machine Learning is a concept which allows the machine to learn from examples and experience, and that too without being explicitly programmed. So instead of you writing the code, what you do is you feed data to the generic algorithm, and the algorithm/ machine builds the logic based on the given data.

Machine Learning is the science of teaching machines how to learn by themselves. Now, you might be thinking – why on earth would we want machines to learn by themselves? Well – it has a lot of benefits.

Machine learning research is part of research on artificial intelligence, seeking to provide knowledge to computers through data, observations and interacting with the world. That acquired knowledge allows computers to correctly generalize to new setting.

The field of Machine Learning seeks to answer the question “How can we build computer systems that automatically improve with experience, and what are the fundamental laws that govern all learning processes?”

Machine learning is based on algorithms that can learn from data without relying on rules-based programming.

Machine learning at its most basic is the practice of using algorithms to parse data, learn from it, and then make a determination or prediction about something in the world.

Machine learning is the science of getting computers to act without being explicitly programmed.

## Evolution of machine learning

Because of new computing technologies, machine learning today is not like machine learning of the past. It was born from pattern recognition and the theory that computers can learn without being programmed to perform specific tasks; researchers interested in artificial intelligence wanted to see if computers could learn from data. The iterative aspect of machine learning is important because as models are exposed to new data, they are able to independently adapt. They learn from previous computations to produce reliable, repeatable decisions and results. It's a science that's not new – but one that has gained fresh momentum.

While many machine learning algorithms have been around for a long time, the ability to automatically apply complex mathematical calculations to big data – over and over, faster and faster – is a recent development. Here are a few widely publicized examples of machine learning applications you may be familiar with:

- The heavily hyped, self-driving Google car? The essence of machine learning.
- Online recommendation offers such as those from Amazon and Netflix? Machine learning applications for everyday life.
- Knowing what customers are saying about you on Twitter? Machine learning combined with linguistic rule creation.
- Fraud detection? One of the more obvious, important uses in our world today.

# Why is machine learning important?

Resurging interest in machine learning is due to the same factors that have made data mining and Bayesian analysis more popular than ever. Things like growing volumes and varieties of available data, computational processing that is cheaper and more powerful, and affordable data storage.

All of these things mean it's possible to quickly and automatically produce models that can analyze bigger, more complex data and deliver faster, more accurate results – even on a very large scale. And by building precise models, an organization has a better chance of identifying profitable opportunities – or avoiding unknown risks.

## Who's using it?

Most industries working with large amounts of data have recognized the value of machine learning technology. By gleaning insights from this data – often in real time – organizations are able to work more efficiently or gain an advantage over competitors.

## Financial services

Banks and other businesses in the financial industry use machine learning technology for two key purposes: to identify important insights in data, and prevent fraud. The insights can identify investment opportunities, or help investors know when to trade. Data mining can also identify clients with high-risk profiles, or use cyber surveillance to pinpoint warning signs of fraud.

## Government

Government agencies such as public safety and utilities have a particular need for machine learning since they have multiple sources of data that can be mined for insights. Analyzing sensor data, for example, identifies ways to increase efficiency and save money. Machine learning can also help detect fraud and minimize identity theft.

## Health care

Machine learning is a fast-growing trend in the health care industry, thanks to the advent of wearable devices and sensors that can use data to assess a patient's health in real time. The technology can also help medical experts analyze data to identify trends or red flags that may lead to improved diagnoses and treatment.

## Retail

Websites recommending items you might like based on previous purchases are using machine learning to analyze your buying history. Retailers rely on machine learning to capture data, analyze it and use it to personalize a shopping experience, implement a marketing campaign, price optimization, merchandise supply planning, and for customer insights.



## Oil and gas

Finding new energy sources. Analyzing minerals in the ground. Predicting refinery sensor failure. Streamlining oil distribution to make it more efficient and cost-effective. The number of machine learning use cases for this industry is vast – and still expanding.

## Transportation

Analyzing data to identify patterns and trends is key to the transportation industry, which relies on making routes more efficient and predicting potential problems to increase profitability. The data analysis and modeling aspects of machine learning are important tools to delivery companies, public transportation and other transportation organizations.

## Applications of Machine Learning

### 01- Machine Learning in Healthcare



Machine learning in healthcare is one such area which is seeing gradual acceptance in the healthcare industry. Google recently developed a machine-learning algorithm to identify cancerous tumors in mammograms, and researchers in Stanford University are using deep learning to identify skin cancer. Machine Learning (ML) is already lending a hand in diverse situations in healthcare. ML in healthcare helps to analyze thousands of different data points and suggest outcomes, provide timely risk scores, precise resource allocation, and has many other applications.

## 02- Machine Learning in Videos Surveillance



The video surveillance system nowadays are powered by AI that makes it possible to detect crime before they happen. They track unusual behavior of people like standing motionless for a long time, stumbling, or napping on benches etc. The system can thus give an alert to human attendants, which can ultimately help to avoid mishaps. And when such activities are reported and counted to be true, they help to improve the surveillance services. This happens with machine learning doing its job at the backend.

### 03- Machine Learning in Fraud Detection



Experts predict online credit card fraud to soar to a whopping **\$32 billion** in 2020. That's more than the profit made by Coca Cola and JP Morgan Chase combined. That's something to worry about. Fraud Detection is one of the most necessary Applications of Machine Learning. The number of transactions has increased due to a plethora of payment channels — credit/debit cards, smartphones, numerous wallets, UPI and much more. At the same time, the amount of criminals have become adept at finding loopholes. Whenever a customer carries out a transaction — the Machine Learning model thoroughly x-rays their profile searching for suspicious patterns. In Machine Learning, problems like fraud detection are usually framed as classification problems

### 04- Machine Learning in Social Media



Social media services utilizes machine learning for their own and user benefits. You have seen one of the most common applications of Machine Learning is **Automatic Friend Tagging Suggestions** in Facebook or any other social media platform. You have also noticed **People You May Know Suggestion** in Facebook where it shows some of those person profiles that you can become friends with. It is done by continuously notices the friends that you connect with, the profiles that you visit very often, your interests, workplace, or a group that you share with someone etc. Both these suggestion features are using machine learning behind.



## 05- Machine Learning in Retail Industry



Businesses organizations that are in the retail industry or e-commerce companies have been using advanced machine learning applications including Recommendation systems, Chat-bot applications, Predictive Analytics system, etc. to innovate and enhance their business processes. A number of big Retail and E-commerce industries like Walmart, Amazon, Alibaba, Flipkart have successfully incorporated AI and Machine Learning technologies across their entire sales cycles from logistics to sales to post-sales services, thus improve results as well as business processes.

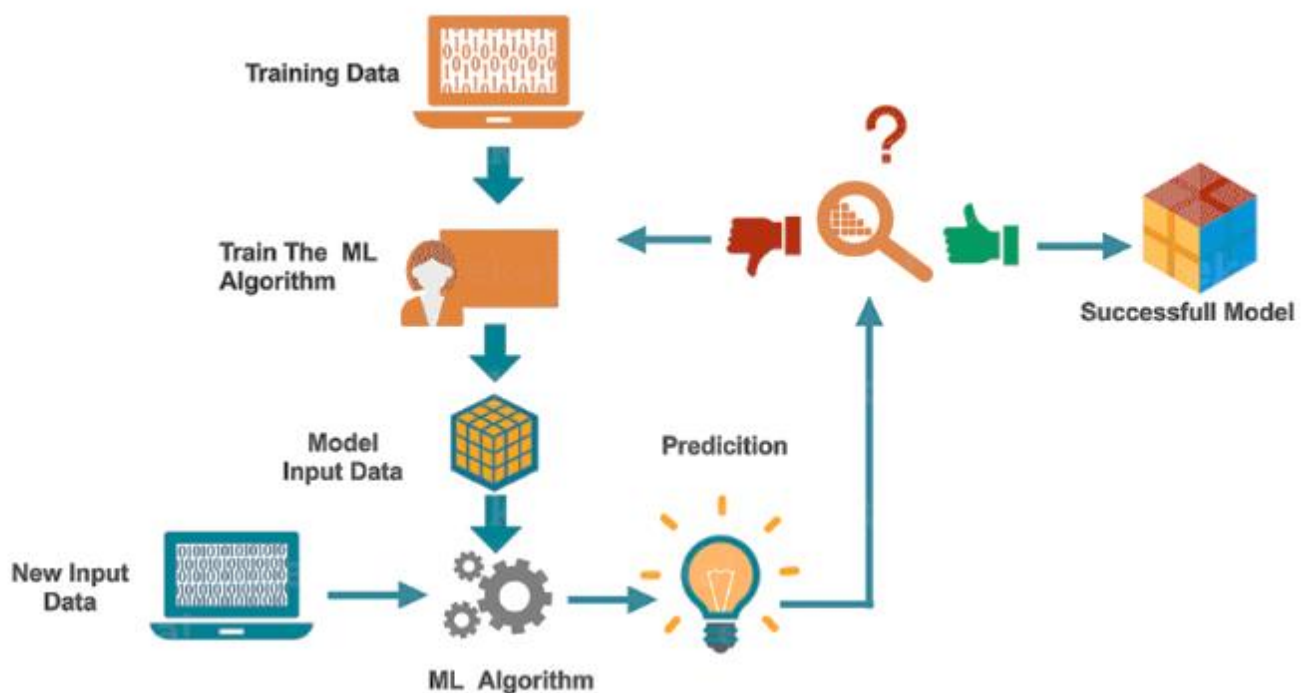


## How does Machine Learning Work?

Machine Learning algorithm is trained using a training data set to create a model. When new input data is introduced to the ML algorithm, it makes a prediction on the basis of the model.

The prediction is evaluated for accuracy and if the accuracy is acceptable, the Machine Learning algorithm is deployed. If the accuracy is not acceptable, the Machine Learning algorithm is trained again and again with an augmented training data set.

This is just a very high-level example as there are many factors and other steps involved.



# Steps of Machine Learning

## 1 - Data Collection

- The quantity & quality of your data dictate how accurate our model is
- The outcome of this step is generally a representation of data (Guo simplifies to specifying a table) which we will use for training
- Using pre-collected data, by way of datasets from Kaggle, UCI, etc., still fits into this step

## 2 - Data Preparation

- Wrangle data and prepare it for training
- Clean that which may require it (remove duplicates, correct errors, deal with missing values, normalization, data type conversions, etc.)
- Randomize data, which erases the effects of the particular order in which we collected and/or otherwise prepared our data
- Visualize data to help detect relevant relationships between variables or class imbalances (bias alert!), or perform other exploratory analysis
- Split into training and evaluation sets

## 3 - Choose a Model

- Different algorithms are for different tasks; choose the right one

## 4 - Train the Model

- The goal of training is to answer a question or make a prediction correctly as often as possible
- Linear regression example: algorithm would need to learn values for  $m$  (or  $W$ ) and  $b$  ( $x$  is input,  $y$  is output)
- Each iteration of process is a training step

## 5 - Evaluate the Model

- Uses some metric or combination of metrics to "measure" objective performance of model
- Test the model against previously unseen data
- This unseen data is meant to be somewhat representative of model performance in the real world, but still helps tune the model (as opposed to test data, which does not)
- Good train / eval split? 80/20, 70/30, or similar, depending on domain, data availability, dataset particulars, etc.

## 6 - Parameter Tuning

- This step refers to *hyper parameter* tuning, which is an "art form" as opposed to a science
- Tune model parameters for improved performance
- Simple model hyper parameters may include: number of training steps, learning rate, initialization values and distribution, etc.

## 7 - Make Predictions

- Using further (test set) data which have, until this point, been withheld from the model (and for which class labels are known), are used to test the model; a better approximation of how the model will perform in the real world

## Machine Learning: The technology leaders

In addition to Microsoft, Google, Facebook, IBM and Amazon, Apple also spends enormous financial resources on the use and further development of Machine Learning. IBM's Watson supercomputer is still the best-known appliance for Machine Learning. Watson is mainly used in the medical and financial sectors. As already mentioned, Facebook uses Machine Learning for image recognition, Microsoft for the speech recognition system Cortana, Apple for Siri. Of course, Machine Learning is also used at Google, both in the area of image services and search engine ranking.

Cloud providers such as Google, Microsoft, Amazon Web service and IBM have now created services for Machine Learning. With their help it is also possible for developers who do not have specific Machine Learning knowledge to develop applications. These applications are able to learn from a freely definable set of data. Depending on the provider, these platforms have different names:

- IBM: Watson
- Amazon: Amazon Machine Learning
- Microsoft: Azure ML Studio
- Google: Tensor flow



# Types of Machine Learning

There some variations of how to define the types of Machine Learning Algorithms but commonly they can be divided into categories according to their purpose and the main categories are the following:

