# Chapter 1
Introduction to Machine Learning

Undoubtedly, Machine Learning is the most in-demand technology in today's market. Its applications range from self-driving cars to predicting deadly diseases such as cancer. The high demand for Machine Learning skills is the motivation behind this Course. In this Course, you will understand all the basic concepts of Machine Learning and its applications in different problem areas of the society.

1.1 Need for Machine Learning

Ever since the technical revolution, we've been generating an immeasurable amount of data. As per research, we generate around 2.5 quintillion bytes of data every single day! It is estimated that by 2020, 1.7MB of data will be created every second for every person on earth.

With the availability of so much data, it is finally possible to build predictive models that can study and analyze complex data to find useful insights and deliver more accurate results. Top Tier companies such as Netflix and Amazon build such Machine Learning models by using tons of data in order to identify profitable opportunities and avoid unwanted risks.

Here's a list of reasons why Machine Learning is so important:

Source:https://www.edureka.co/blog/introduction-to-machine-learning/
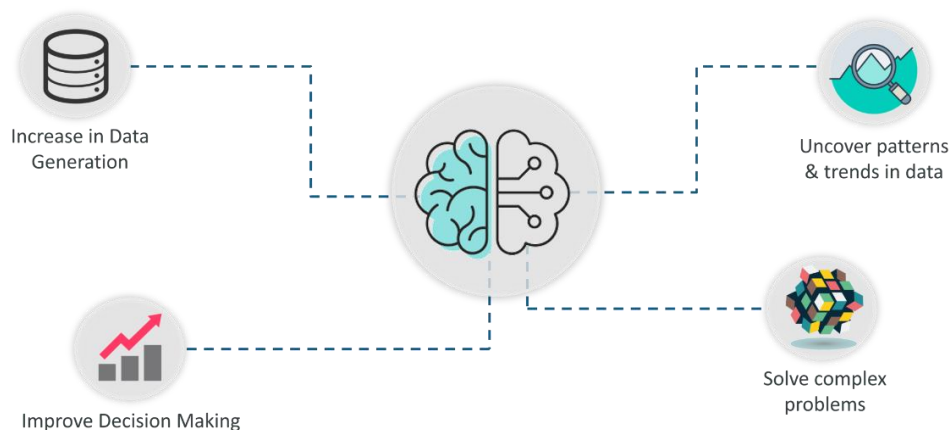
Increase in Data Generation: Due to excessive production of data, we need a method that can be used to structure, analyze and draw useful insights from data. This is where Machine Learning comes in. It uses data to solve problems and find solutions to the most complex tasks faced by organizations.

Improve Decision Making*:* By making use of various algorithms, Machine Learning can be used to make better business decisions. For example,

Machine Learning is used to forecast sales, predict downfalls in the stock market, identify risks and anomalies, etc.

Uncover patterns & trends in data: Finding hidden patterns and extracting key insights from data is the most essential part of Machine Learning. By building predictive models and using statistical techniques, Machine Learning allows you to dig beneath the surface and explore the data at a minute scale. Understanding data and extracting patterns manually will take days, whereas Machine Learning algorithms can perform such computations in less than a second.

Solve complex problems: From detecting the genes linked to the deadly ALS disease to building self-driving cars, Machine Learning can be used to solve the most complex problems.

1.2 Machine Learning Applications

Machine learning is a buzzword for today's technology, and it is growing very rapidly day by day. We are using machine learning in our daily life even without knowing it such as Google Maps, Google assistant, Alexa, etc. Below are some most trending real-world applications of Machine Learning:

Source: https://www.javatpoint.com/applications-of-machine-learning
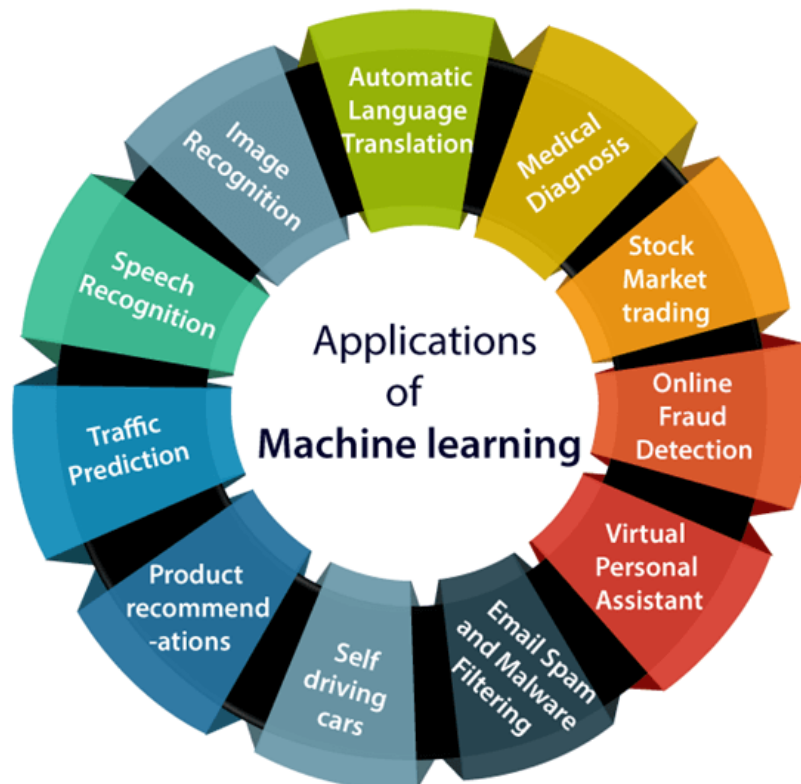
1. Image Recognition
Image recognition is one of the most common applications of machine learning. It is used to identify objects, persons, places, digital images, etc.

Class Notes on Machine Learning And Its Applications By Prof. Anand Singh Jalal

2

The popular use case of image recognition and face detection is, Automatic friend tagging suggestion in Facebook. Facebook provides us a feature of auto friend tagging suggestion. Whenever we upload a photo with our Facebook friends, then we automatically get a tagging suggestion with name, and the technology behind this is machine learning's face detection and recognition algorithm. It is based on the Facebook project named "Deep Face," which is responsible for face recognition and person identification in the picture.

2. Speech Recognition

While using Google, we get an option of "Search by voice," it comes under speech recognition, and it's a popular application of machine learning. Speech recognition is a process of converting voice instructions into text, and it is also known as "Speech to text", or "Computer speech recognition." At present, machine learning algorithms are widely used by various applications of speech recognition. Google assistant, Siri, Cortana, and Alexa are using speech recognition technology to follow the voice instructions.

3. Traffic prediction

If we want to visit a new place, we take help of Google Maps, which shows us the correct path with the shortest route and predicts the traffic conditions. It predicts the traffic conditions such as whether traffic is cleared, slow-moving, or heavily congested with the help of two ways:
- Real Time location of the vehicle form Google Map app and sensors
- Average time has taken on past days at the same time.

Everyone who is using Google Map is helping this app to make it better. It takes information from the user and sends back to its database to improve the performance.

4. Product recommendations

Machine learning is widely used by various e-commerce and entertainment companies such as Amazon, Netflix, etc., for product recommendation to the user. Whenever we search for some product on Amazon, then we started getting an advertisement for the same product while internet surfing on the same browser and this is because of machine learning.

Google understands the user interest using various machine learning algorithms and suggests the product as per customer interest.

As similar, when we use Netflix, we find some recommendations for entertainment series, movies, etc., and this is also done with the help of machine learning.

5. Self-driving cars

One of the most exciting applications of machine learning is self-driving cars. Machine learning plays a significant role in self-driving cars. Tesla, the most popular car manufacturing company is working on self-driving

Class Notes on Machine Learning And Its Applications By Prof. Anand Singh Jalal

3

car. It is using unsupervised learning method to train the car models to detect people and objects while driving.

6. Email Spam and Malware Filtering

Whenever we receive a new email, it is filtered automatically as important, normal, and spam. We always receive an important mail in our inbox with the important symbol and spam emails in our spam box, and the technology behind this is Machine learning. Below are some spam filters used by Gmail:

- Content Filter
- Header filter
- General blacklists filter
- Rules-based filters
- Permission filters

Some machine learning algorithms such as Multi-Layer Perceptron, Decision tree, and Naïve Bayes classifier are used for email spam filtering and malware detection.

7. Virtual Personal Assistant

We have various virtual personal assistants such as Google assistant, Alexa, Cortana, Siri. As the name suggests, they help us in finding the information using our voice instruction. These assistants can help us in various ways just by our voice instructions such as Play music, call someone, open an email, Scheduling an appointment, etc. These virtual assistants use machine learning algorithms as an important part. These assistant record our voice instructions, send it over the server on a cloud, and decode it using ML algorithms and act accordingly.

8. Online Fraud Detection

Machine learning is making our online transaction safe and secure by detecting fraud transaction. Whenever we perform some online transaction, there may be various ways that a fraudulent transaction can take place such as fake accounts, fake ids, and steal money in the middle of a transaction. So to detect this, Feed Forward Neural network helps us by checking whether it is a genuine transaction or a fraud transaction.

For each genuine transaction, the output is converted into some hash values, and these values become the input for the next round. For each genuine transaction, there is a specific pattern which gets change for the fraud transaction hence, it detects it and makes our online transactions more secure.

9. Stock Market trading

Machine learning is widely used in stock market trading. In the stock market, there is always a risk of up and downs in shares, so for this machine learning's long short term memory neural network is used for the prediction of stock market trends.

Class Notes on Machine Learning And Its Applications By Prof. Anand Singh Jalal

4

10. Medical Diagnosis

In medical science, machine learning is used for diseases diagnoses. With this, medical technology is growing very fast and able to build 3D models that can predict the exact position of lesions in the brain. It helps in finding brain tumors and other brain-related diseases easily.

11. Automatic Language Translation

Nowadays, if we visit a new place and we are not aware of the language then it is not a problem at all, as for this also machine learning helps us by converting the text into our known languages. Google's GNMT (Google Neural Machine Translation) provide this feature, which is a Neural Machine Learning that translates the text into our familiar language, and it called as automatic translation.

The technology behind the automatic translation is a sequence to sequence learning algorithm, which is used with image recognition and translates the text from one language to another language.

1.3 Traditional Programming Vs Machine Learning

Traditional Programming refers to any manually created program that uses input data and runs on a computer to produce the output. In Machine Learning, the input data and output are fed to an algorithm to create a program.
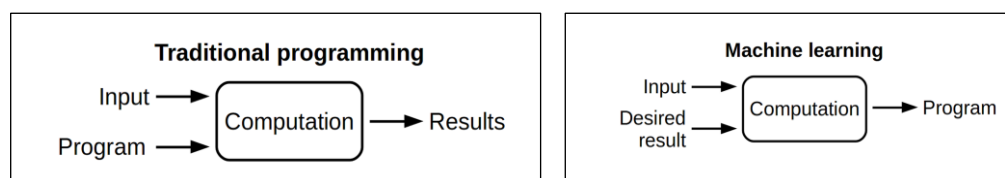

1.4 Machine Learning Definitions

The term Machine Learning was first coined by Arthur Samuel in the year 1959. According to Samuel - "*Machine Learning is the field of study that gives computers the ability to learn without being explicitly programmed.*"

 If you browse through the net about 'what is Machine Learning', you'll get at least 100 different definitions. However, the very first formal definition was given by Tom M. Mitchell (1997) - "*A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E.*"

So, we can say that Machine Learning is a subset of Artificial Intelligence (AI) which provides machines the ability to lean automatically & improve from experience without being explicitly programmed.

Class Notes on Machine Learning And Its Applications By Prof. Anand Singh Jalal

5

A Machine Learning process begins by feeding the machine lots of data, by using this data the machine is trained to detect hidden insights and trends. These insights are then used to build a Machine Learning Model by using an algorithm in order to solve a problem.

Source: https://www.edureka.co/blog/introduction-to-machine-learning/
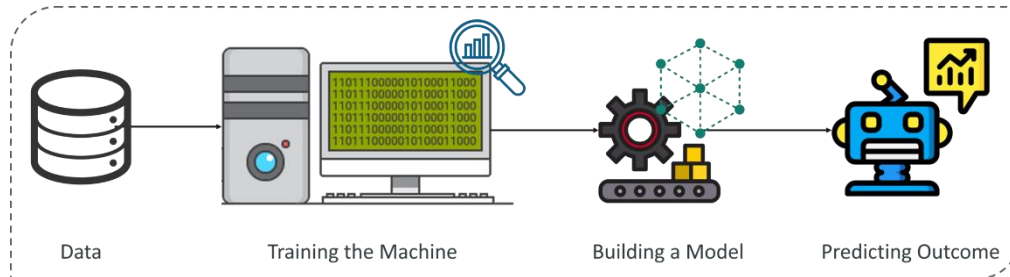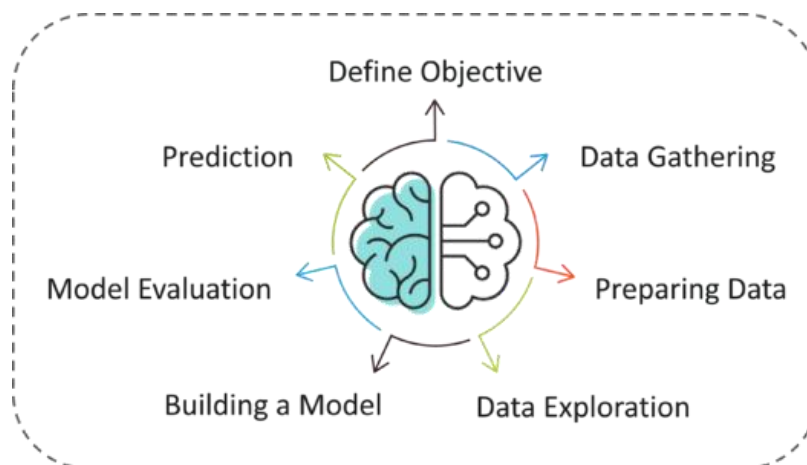
1.5 AI vs ML vs Deep Learning

Artificial Intelligence: Artificial Intelligence is the broader concept of machines being able to carry out tasks in a smarter way. It covers anything which enables the computers to behave like humans.

Machine Learning: Machine Learning is a subset of Artificial Intelligence (AI) which provides machines the ability to lean automatically & improve from experience without being explicitly programmed.

Deep Learning: Deep Learning is a subset of Machine Learning where similar Machine Learning Algorithms are used to train Deep Neural Networks so as to achieve better accuracy in those cases where former was not performing up to the mark.

1.6 Machine Learning Process

The Machine Learning process involves building a Predictive model that can be used to find a solution for a Problem statement.

Source: https://www.edureka.co/blog/introduction-to-machine-learning/

Class Notes on Machine Learning And Its Applications By Prof. Anand Singh Jalal

6

To understand the Machine Learning process let's assume that you have been given a problem that needs to be solved by using Machine Learning. Suppose, the problem is to predict the occurrence of rain in your local area by using Machine Learning. The below steps are followed in a Machine Learning process:

Step 1: Define the objective of the Problem Statement

At this step, we must understand what exactly needs to be predicted. In our case, the objective is to predict the possibility of rain by studying weather conditions. At this stage, it is also essential to take mental notes on what kind of data can be used to solve this problem or the type of approach you must follow to get to the solution.

Step 2: Data Gathering

At this stage, you must be asking questions such as,
- What kind of data is needed to solve this problem?
- Is the data available?
- How can I get the data?

Once you know the types of data that is required, you must understand how you can derive this data. Data collection can be done manually or by web scraping. Coming back to the problem at hand, the data needed for weather forecasting includes measures such as humidity level, temperature, pressure, locality, whether or not you live in a hill station, etc. Such data must be collected and stored for analysis.

Step 3: Data Preparation

The data you collected is almost never in the right format. You will encounter a lot of inconsistencies in the data set such as missing values, redundant variables, duplicate values, etc. Removing such inconsistencies is very essential because they might lead to wrongful computations and predictions. Therefore, at this stage, you scan the data set for any inconsistencies and you fix them then and there.

Step 4: Exploratory Data Analysis

This stage is all about diving deep into data and finding all the hidden data mysteries. EDA or *Exploratory Data Analysis* is the brainstorming stage of Machine Learning. Data Exploration involves understanding the patterns and trends in the data. At this stage, all the useful insights are drawn and correlations between the variables are understood.

For example, in the case of predicting rainfall, we know that there is a strong possibility of rain if the temperature has fallen low. Such correlations must be understood and mapped at this stage.

Step 5: Building a Machine Learning Model

Class Notes on Machine Learning And Its Applications By Prof. Anand Singh Jalal

7

All the insights and patterns derived during Data Exploration are used to build the Machine Learning Model. This stage always begins by splitting the data set into two parts, training data, and testing data. The training data will be used to build and analyze the model. The logic of the model is based on the Machine Learning Algorithm that is being implemented.

In the case of predicting rainfall, since the output will be in the form of True (if it will rain tomorrow) or False (no rain tomorrow), we can use a Classification Algorithm such as Logistic Regression. Choosing the right algorithm depends on the type of problem you're trying to solve, the data set and the level of complexity of the problem.

Step 6: Model Evaluation & Optimization
After building a model by using the training data set, it is finally time to put the model to a test. The testing data set is used to check the efficiency of the model and how accurately it can predict the outcome. Once the accuracy is calculated, any further improvements in the model can be implemented at this stage. Methods like parameter tuning and cross-validation can be used to improve the performance of the model.

Step 7: Predictions
Once the model is evaluated and improved, it is finally used to make predictions. The final output can be a Categorical variable (eg. True or False) or it can be a Continuous Quantity (eg. the predicted value of a stock). In the given problem, for predicting the occurrence of rainfall, the output will be a categorical variable.

1.7 Machine Learning Types

A machine can learn to solve a problem by following any one of the following three approaches. These are the ways in which a machine can learn:
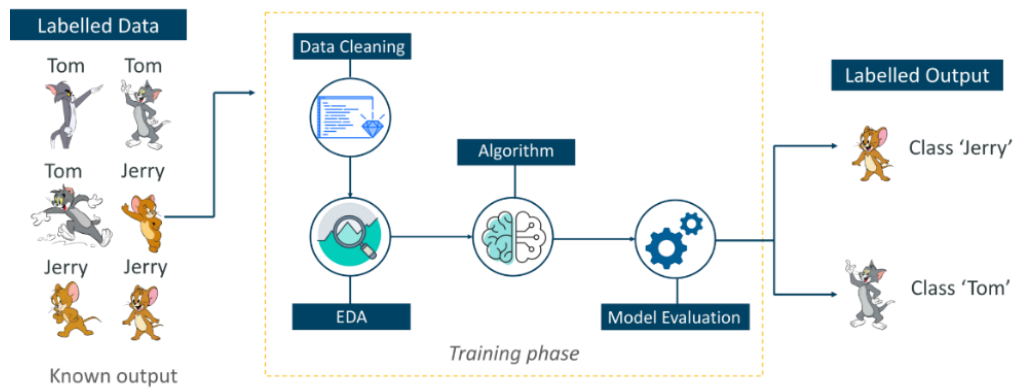1. Supervised Learning
2. Unsupervised Learning
3. Reinforcement Learning

Supervised Learning

"*Supervised learning is a technique in which we teach or train the machine using data which is well labeled.*"
To understand Supervised Learning let's consider an analogy. As kids we all needed guidance to solve math problems. Our teachers helped us understand what addition is and how it is done. Similarly, you can think of supervised learning as a type of Machine Learning that involves a guide. The labeled data set is the teacher that will train you to understand

Class Notes on Machine Learning And Its Applications By Prof. Anand Singh Jalal

8

patterns in the data. The labeled data set is nothing but the training data set.

Source: https://www.edureka.co/blog/introduction-to-machine-learning/

Consider the above figure. Here we're feeding the machine images of Tom and Jerry and the goal is for the machine to identify and classify the images into two groups (Tom images and Jerry images). The training data set that is fed to the model is labeled, as in, we're telling the machine, 'this is how Tom looks and this is Jerry'. By doing so you're training the machine by using labeled data. In Supervised Learning, there is a well-defined training phase done with the help of labeled data.

There are two kinds of supervise learning:
- Classification
- Regression

Example: Classification Vs Regression
Source: https://medium.com/@ali_88273/regression-vs-classification-87c224350d69 by Ali Reza Kohani

Class Notes on Machine Learning And Its Applications By Prof. Anand Singh Jalal

9

Difference between Classification and Regression

Classification	Regression
• Classification is the task of predicting a discrete class label	• Regression is the task of predicting a continuous quantity
• In a classification problem data is labelled into one of two or more classes	• A regression problem requires the prediction of a quantity
• A classification problem with two classes is called binary, more than two classes is called a multi-class classification	• A regression problem with multiple input variables is called a multivariate regression problem
• Classifying an email as spam or non-spam is an example of a classification problem	• Predicting the price of a stock over a period of time is a regression problem

Source: https://www.edureka.co/blog/interview-questions/machine-learning-interview-questions/

Unsupervised Learning

"Unsupervised learning involves training by using unlabeled data and allowing the model to act on that information without guidance."

Think of unsupervised learning as a smart kid that learns without any guidance. In this type of Machine Learning, the model is not fed with labeled data, as in the model has no clue that 'this image is Tom and this is Jerry', it figures out patterns and the differences between Tom and Jerry on its own by taking in tons of data.

Source: https://www.edureka.co/blog/introduction-to-machine-learning/
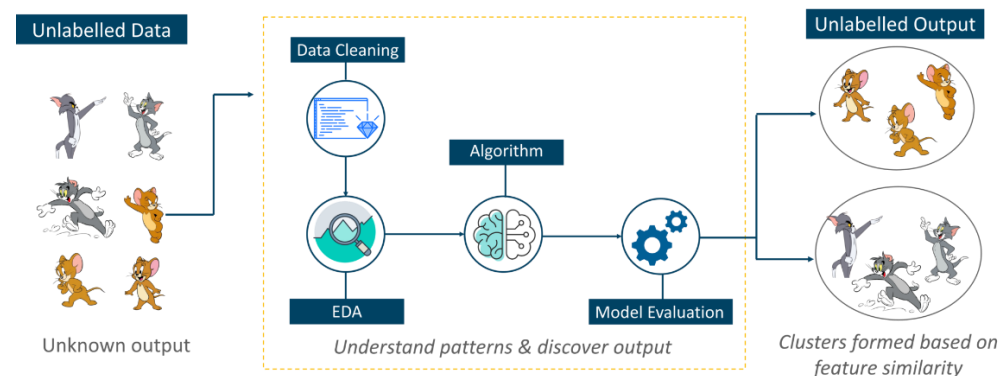
For example, it identifies prominent features of Tom such as pointy ears, bigger size, etc, to understand that this image is of type 1. Similarly, it finds such features in Jerry and knows that this image is of type 2. Therefore, it classifies the images into two different classes without knowing who Tom is or Jerry is.

Google news is another example of unsupervised learning. Google news creates clusters of all the URLs of different e- news portal sharing the same news are mentioned at one place.

Class Notes on Machine Learning And Its Applications By Prof. Anand Singh Jalal

10

Reinforcement Learning

"Reinforcement Learning is a part of Machine learning where an agent is put in an environment and he learns to behave in this environment by performing certain actions and observing the rewards which it gets from those actions."
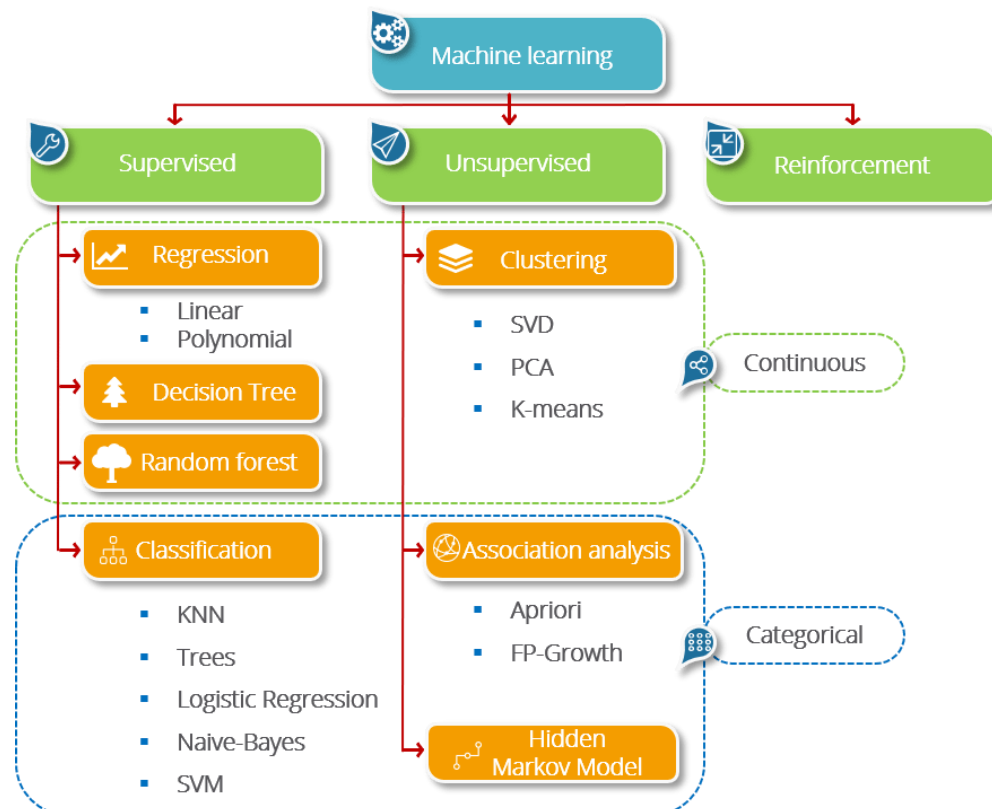
Reinforcement Learning is mainly used in advanced Machine Learning areas such as self-driving cars, AplhaGo, etc.

1.8 Difference between Types of Learning

	Supervised Learning	Unsupervised Learning	Reinforcement Learning
Definition	The machine learns by using labelled data	The machine is trained on unlabelled data without any guidance	An agent interacts with its environment by producing actions & discovers errors or rewards
Type of problems	Regression & Classification	Association & Clustering	Reward based
Type of data	Labelled data	Unlabelled data	No pre-defined data
Training	External supervision	No supervision	No supervision
Approach	Map labelled input to known output	Understand patterns and discover output	Follow trail and error method
Popular algorithms	Linear regression, Logistic regression, Support Vector Machine, KNN, etc	K-means, C-means, etc	Q-Learning, SARSA, etc

Source: https://www.edureka.co/blog/interview-questions/machine-learning-interview-questions/

Summary

Source: https://www.edureka.co/blog/what-is-machine-learning/

Class Notes on Machine Learning And Its Applications By Prof. Anand Singh Jalal

11