MACHINE LEARNING

Machine learning is defined as the subfield of AI that focuses on the development of the computer programs which have access to data by providing the system with the ability to learn and improve automatically by finding patterns in the database without any human interventions or actions. Based on the data type, i.e., labelled or unlabeled data, the model's training in machine learning has been classified as supervised and unsupervised learning.

Machine Learning Definition

Simply says finds pattern in data and uses those patterns to predict the future. It allows us to discover patterns in existing data and create and make use of a model that identifies those patterns in innovative data. It has gone mainstream.

Why Machine Learning Strategy?

Machine learning is the foundation of countless important applications, including web search, email anti-spam, speech recognition, product recommendations, and more. I assume that you or your team is working on a machine learning application, and that you want to make rapid progress.

Why is Machine Learning so Popular Currently?

- Plenty of data.
- Lots of computer power.
- An effective machine learning algorithm.

Top Machine Learning Companies

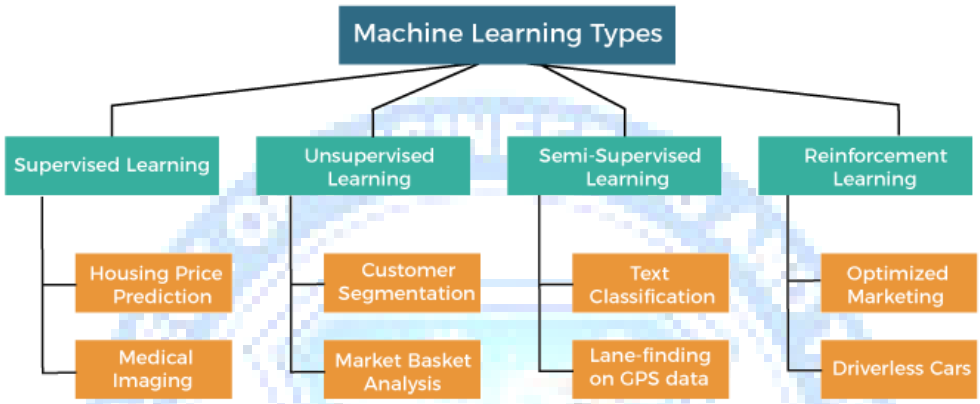
It is becoming an important part of our everyday life. It is really utilized in financial procedures, medical examinations, logistics, posting, and a variety of different fast-rising industries.

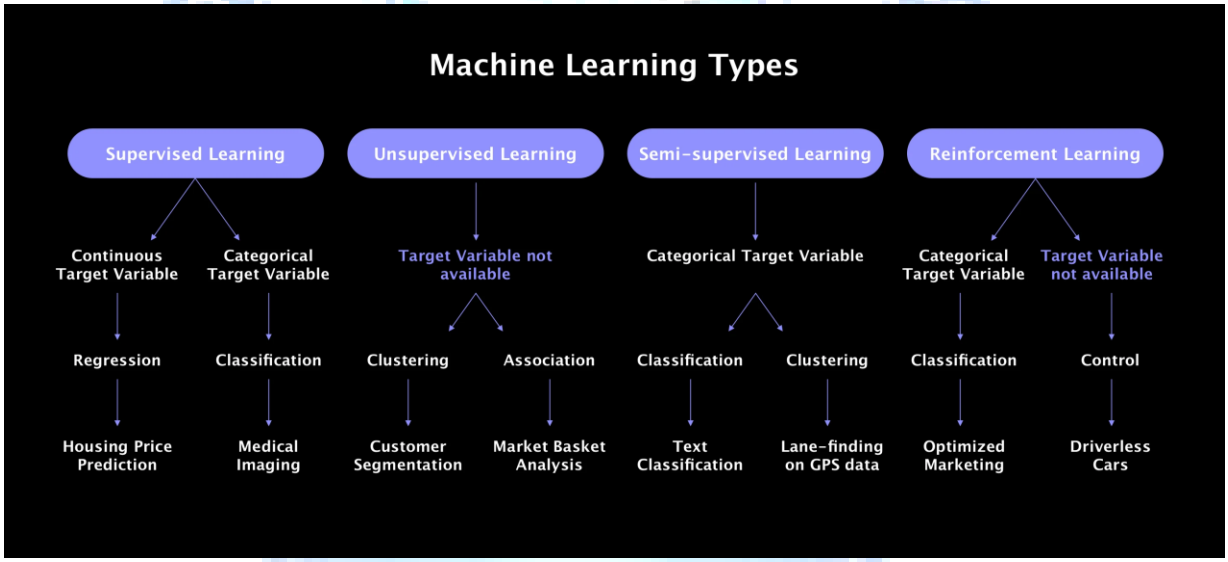
- **Google:** Neural Networks and Machines
- **Tesla:** Autopilot
- **Amazon:** Echo Speaker Alexa
- **Apple:** Personalized Hey Siri

• **TCS:** Machine First Delivery Model with Robotics

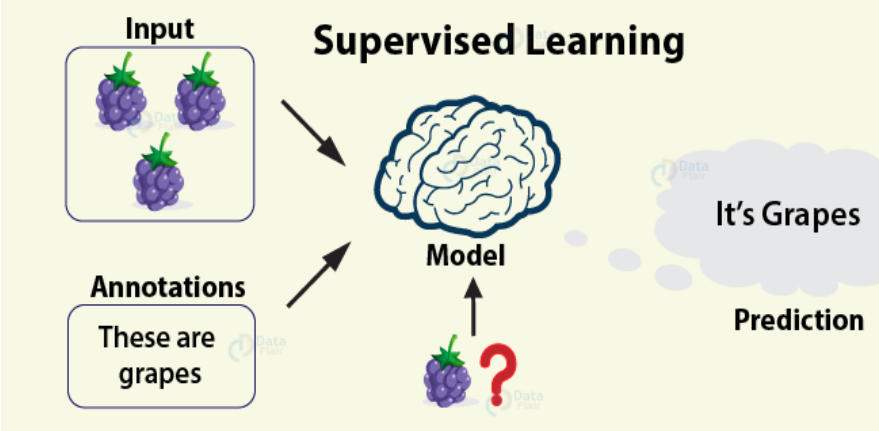
• **Facebook:** Chatbot Army etc.

Types of Machine Learning





1. **Supervised Learning**
   - Supervised Learning is the most popular paradigm for performing machine learning operations. It is widely used for data where there is a precise **mapping** between **input-output data**.
   - The dataset, in this case, is **labeled**, meaning that the algorithm **identifies** the **features explicitly** and **carries out predictions** or **classification** accordingly.
   - As the training period progresses, the **algorithm** is able to **identify** the **relationships** between the **two variables** such that we can **predict** a **new outcome**.



Resulting Supervised learning algorithms are **task-oriented**.

As we provide it with more and more examples, it is able to **learn more properly** so that it can **undertake** the **task** and **yield** us the **output** more accurat

Draft

- Predictive Model
- we have labeled data
- The main types of supervised learning problems include regression and classification problems

Categories of Supervised Machine Learning

Supervised machine learning can be classified into two types of problems, which are given below:

- Classification
- Regression

a) Classification

Classification algorithms are used to solve the classification problems in which the output variable is categorical, such as "Yes" or No, Male or Female, Red or Blue, etc. The classification algorithms predict the categories present in the dataset. Some real-world examples of classification algorithms are Spam Detection, Email filtering, etc.

Some popular classification algorithms are given below:

- Random Forest Algorithm
- Decision Tree Algorithm

- Logistic Regression Algorithm
- Support Vector Machine Algorithm

b) Regression

Regression algorithms are used to solve regression problems in which there is a linear relationship between input and output variables. These are used to predict continuous output variables, such as market trends, weather prediction, etc.

Some popular Regression algorithms are given below:

- Simple Linear Regression Algorithm
- Multivariate Regression Algorithm
- Decision Tree Algorithm
- Lasso Regression

Advantages and Disadvantages of Supervised Learning

Advantages:

- Since supervised learning work with the labelled dataset so we can have an exact idea about the classes of objects.
- These algorithms are helpful in predicting the output on the basis of prior experience.

Disadvantages:

- These algorithms are not able to solve complex tasks.
- It may predict the wrong output if the test data is different from the training data.
- It requires lots of computational time to train the algorithm.

Applications of Supervised Learning

Some common applications of Supervised Learning are given below:

- **Image Segmentation:**

  Supervised Learning algorithms are used in image segmentation. In this process, image classification is performed on different image data with pre-defined labels.

- **Medical Diagnosis:**

  Supervised algorithms are also used in the medical field for diagnosis purposes. It is done by using medical images and past labelled data with labels for disease conditions. With such a process, the machine can identify a disease for the new patients.

- **Fraud Detection -** Supervised Learning classification algorithms are used for identifying fraud transactions, fraud customers, etc. It is done by using historic data to identify the patterns that can lead to possible fraud.
- **Spam detection -** In spam detection & filtering, classification algorithms are used. These algorithms classify an email as spam or not spam. The spam emails are sent to the spam folder.
- **Speech Recognition -** Supervised learning algorithms are also used in speech recognition. The algorithm is trained with voice data, and various identifications can be done using the same, such as voice-activated passwords, voice commands, etc.

Supervised Learning Use Case

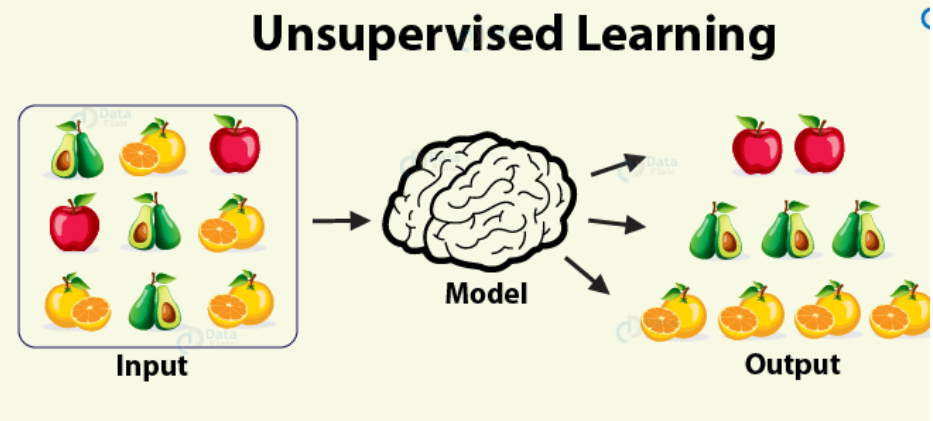
Facial Recognition is one of the most popular applications of Supervised Learning and more specifically –

Artificial Neural Networks.

Convolutional Neural Networks (CNN) is a type of **ANN** used for identifying the faces of people. These models are able to draw features from the image through various **filters**. Finally, if there is a high similarity score between the **input image** and the **image in the database**, a positive match is provided.

Baidu, China's premier search engine company has been investing in **facial recognition**. While it has already installed facial recognition systems in its security systems, it is now **extending** this technology to the major airports of China. Baidu will provide the airports with facial recognition technology that will provide access to the **ground crew** and the **staff**. Therefore, the passengers do not have to wait in long queues for **flight check-in** when they can simply board their flight by scanning their faces.

2. **Unsupervised Learning**
   - In the case of an unsupervised learning algorithm, the data is **not explicitly labeled** into **different classes**, that is, there are **no labels**. The model is able to learn from the data by finding **implicit patterns**.
   - Unsupervised Learning algorithms identify the data based on their **densities**, **structures**, **similar segments**, and other similar features. Unsupervised Learning Algorithms are based on **Hebbian Learning**.
   - **Cluster analysis** is one of the most widely used techniques in supervised learning.



Draft

- Descriptive Model
- The main types of unsupervised learning algorithms include Clustering algorithms and Association rule learning algorithms.

Categories of Unsupervised Machine Learning

Unsupervised Learning can be further classified into two types, which are given below:

- Clustering
- Association

1) Clustering

The clustering technique is used when we want to find the inherent groups from the data. It is a way to group the objects into a cluster such that the objects with the most similarities remain in one group and have fewer or no similarities with the objects of other groups. An example of the clustering algorithm is grouping the customers by their purchasing behaviour.

Some of the popular clustering algorithms are given below:

- K-Means Clustering algorithm
- Mean-shift algorithm
- DBSCAN Algorithm
- Principal Component Analysis
- Independent Component Analysis

2) Association

Association rule learning is an unsupervised learning technique, which finds interesting relations among variables within a large dataset. The main aim of this learning algorithm is to find the

dependency of one data item on another data item and map those variables accordingly so that it can generate maximum profit. This algorithm is mainly applied in **Market Basket analysis, Web usage mining, continuous production**, etc.

Some popular algorithms of Association rule learning are **Apriori Algorithm, Eclat, FP-growth algorithm.**

Advantages and Disadvantages of Unsupervised Learning Algorithm

Advantages:

- These algorithms can be used for complicated tasks compared to the supervised ones because these algorithms work on the unlabeled dataset.
- Unsupervised algorithms are preferable for various tasks as getting the unlabeled dataset is easier as compared to the labelled dataset.

Disadvantages:

- The output of an unsupervised algorithm can be less accurate as the dataset is not labelled, and algorithms are not trained with the exact output in prior.
- Working with Unsupervised learning is more difficult as it works with the unlabelled dataset that does not map with the output.

Applications of Unsupervised Learning

- **Network Analysis:** Unsupervised learning is used for identifying plagiarism and copyright in document network analysis of text data for scholarly articles.
- **Recommendation Systems:** Recommendation systems widely use unsupervised learning techniques for building recommendation applications for different web applications and e-commerce websites.
- **Anomaly Detection:** Anomaly detection is a popular application of unsupervised learning, which can identify unusual data points within the dataset. It is used to discover fraudulent transactions.
- **Singular Value Decomposition:** Singular Value Decomposition or SVD is used to extract particular information from the database. For example, extracting information of each user located at a particular location.

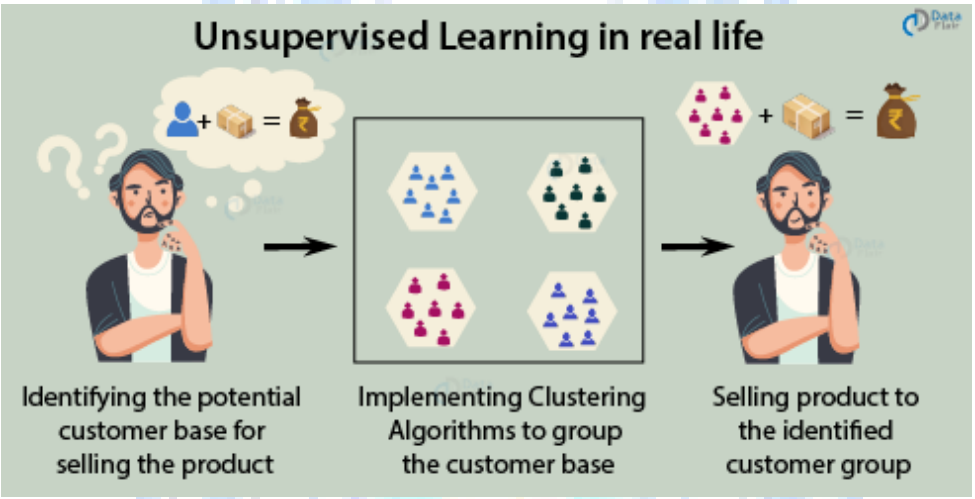
Unsupervised Learning Use Case

One of the most popular unsupervised learning techniques is **clustering**. Using clustering, businesses are able to capture **potential customer segments** for selling their products.

Sales companies are able to identify customer segments that are most likely to use their **services**. Companies can evaluate the customer segments and then decide to sell their product to **maximize the profits**.

One such company that is performing **brand marketing analytics** using Machine Learning is an **Israeli based startup – Optimove**. The goal of this company is to ingest and **process** the **customer data** in order to make it **accessible** to the marketers.

They take it one step further by providing **smart insights** to the marketing team, allowing them to reap the **maximum profit** out of their **product marketing**.



3. Semi-Supervised Learning

Semi-Supervised learning is a type of Machine Learning algorithm that lies between Supervised and Unsupervised machine learning. It represents the intermediate ground between Supervised (With Labelled training data) and Unsupervised learning (with no labelled training data) algorithms and uses the combination of labelled and unlabeled datasets during the training period.

Although Semi-supervised learning is the middle ground between supervised and unsupervised learning and operates on the data that consists of a few labels, it mostly consists of unlabeled data. As labels are costly, but for corporate purposes, they may have few labels. It is completely different from supervised and unsupervised learning as they are based on the presence & absence of labels.

To overcome the drawbacks of supervised learning and unsupervised learning algorithms, the concept of Semi-supervised learning is introduced. The main aim of semi-supervised learning is to effectively use all the available data, rather than only labelled data like in

supervised learning. Initially, similar data is clustered along with an unsupervised learning algorithm, and further, it helps to label the unlabeled data into labelled data. It is because labelled data is a comparatively more expensive acquisition than unlabeled data.

We can imagine these algorithms with an example. Supervised learning is where a student is under the supervision of an instructor at home and college. Further, if that student is self-analysing the same concept without any help from the instructor, it comes under unsupervised learning. Under semi-supervised learning, the student has to revise himself after analyzing the same concept under the guidance of an instructor at college.

Advantages and disadvantages of Semi-supervised Learning

Advantages:

- It is simple and easy to understand the algorithm.
- It is highly efficient.
- It is used to solve drawbacks of Supervised and Unsupervised Learning algorithms.

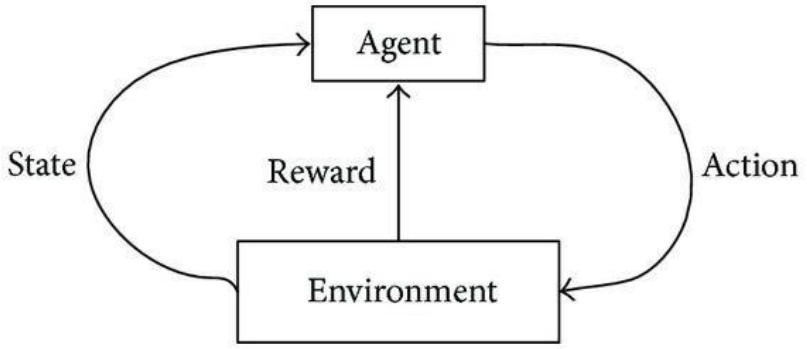
Disadvantages:

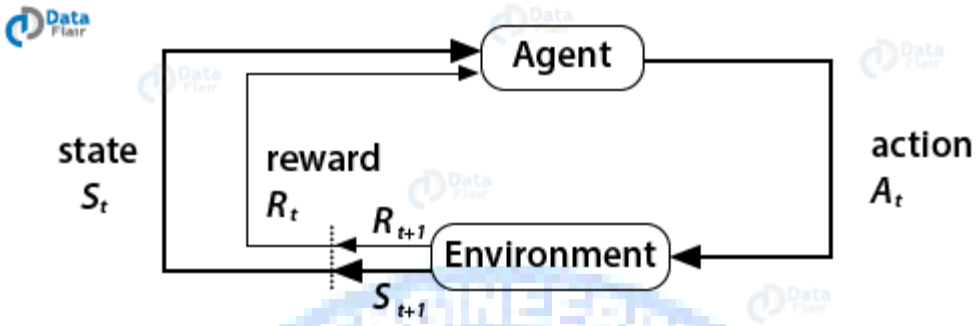
- Iterations results may not be stable.
- We cannot apply these algorithms to network-level data.
- Accuracy is low.

4. Reinforcement Learning

Reinforcement Learning covers more area of Artificial Intelligence which allows machines to interact with their dynamic environment in order to reach their goals. With this, machines and software agents are able to evaluate the ideal behavior in a specific context.

With the help of this reward feedback, agents are able to learn the behavior and improve it in the longer run. This simple feedback reward is known as a reinforcement signal.





The agent in the environment is required to take actions that are based on the **current state**. This type of learning is different from Supervised Learning in the sense that the training data in the former has **output mapping** provided such that the model is capable of learning the **correct answer**.

Whereas, in the case of reinforcement learning, there is **no answer key** provided to the agent when they have to **perform** a **particular task**. When there is no training dataset, it **learns** from its **own experience**.

Categories of Reinforcement Learning

Reinforcement learning is categorized mainly into two types of methods/algorithms:

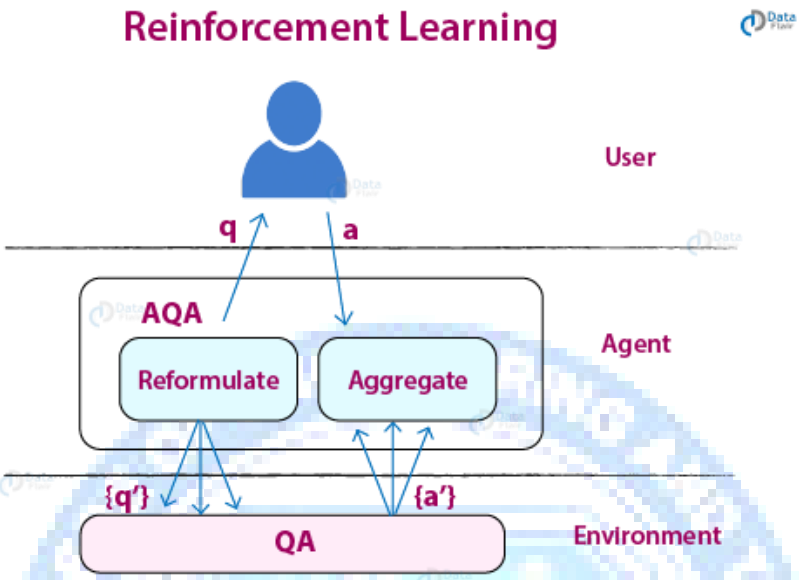
- **Positive Reinforcement Learning:** Positive reinforcement learning specifies increasing the tendency that the required behaviour would occur again by adding something. It enhances the strength of the behaviour of the agent and positively impacts it.
- **Negative Reinforcement Learning:** Negative reinforcement learning works exactly opposite to the positive RL. It increases the tendency that the specific behaviour would occur again by avoiding the negative condition.

Reinforcement Learning Use Case

Google's Active Query Answering (AQA) system makes use of reinforcement learning.It reformulates the questions asked by the user.

For example, if you ask the AQA bot the question – "What is the birth date of Nikola Tesla" then the bot would reformulate it into different questions like "What is the birth year of Nikola Tesla", "When was Tesla born?" and "When is Tesla's birthday".

This process of reformulation utilized the traditional sequence2sequence model, but Google has integrated reinforcement Learning into its system to better interact with the query based environment system.



This is a deviation from the **traditional seq2seq** model such that all the tasks are carried out using **reinforcement learning** and **policy gradient** methods. That is, for a given **question q0**, we want to obtain the best possible **answer a***.

The goal is to **maximize** the award **a* = argmaxa R(ajq0)**.

Real-world Use cases of Reinforcement Learning

- **Video Games:**

  RL algorithms are much popular in gaming applications. It is used to gain super-human performance. Some popular games that use RL algorithms are **AlphaGO** and **AlphaGO Zero**.

- **Resource Management:**

  The "Resource Management with Deep Reinforcement Learning" paper showed that how to use RL in computer to automatically learn and schedule resources to wait for different jobs in order to minimize average job slowdown.

- **Robotics:**

  RL is widely being used in Robotics applications. Robots are used in the industrial and manufacturing area, and these robots are made more powerful with reinforcement learning. There are different industries that have their vision of building intelligent robots using AI and Machine learning technology.

- **Text Mining**

  Text-mining, one of the great applications of NLP, is now being implemented with the help of Reinforcement Learning by Salesforce company.

Advantages and Disadvantages of Reinforcement Learning

Advantages

- It helps in solving complex real-world problems which are difficult to be solved by general techniques.
- The learning model of RL is similar to the learning of human beings; hence most accurate results can be found.
- Helps in achieving long term results.

Disadvantage

- RL algorithms are not preferred for simple problems.
- RL algorithms require huge data and computations.
- Too much reinforcement learning can lead to an overload of states which can weaken the results.