

# Supervised Learning and its different types

- Overview of Machine Learning
- What is Supervised Learning?
- Why is it important?
- Types of Supervised Learning
- Applications
- Supervised vs. Unsupervised Learning
- Disadvantages of Supervised Learning

So get ready to dirty your hands with all there is to know about Supervised Learning. Let's get started :)

## Overview of Machine Learning

So for all those of you who do not know what is Machine Learning? Machine Learning, in the simplest of terms, is teaching your machine about something. You collect data, clean the data, create algorithms, teach the algorithm essential patterns from the data and then expect the algorithm to give you a helpful answer. If the algorithm lives up to your expectations, you have successfully taught your algorithm. If not, just scrap everything and start from scratch. That is how it works here. Oh, and if you are looking for a formal definition, Machine Learning is the process of creating models that can perform a certain task without the need for a human explicitly programming it to do something.

There are 3 types of Machine Learning which are based on the way the algorithms are created. They are:

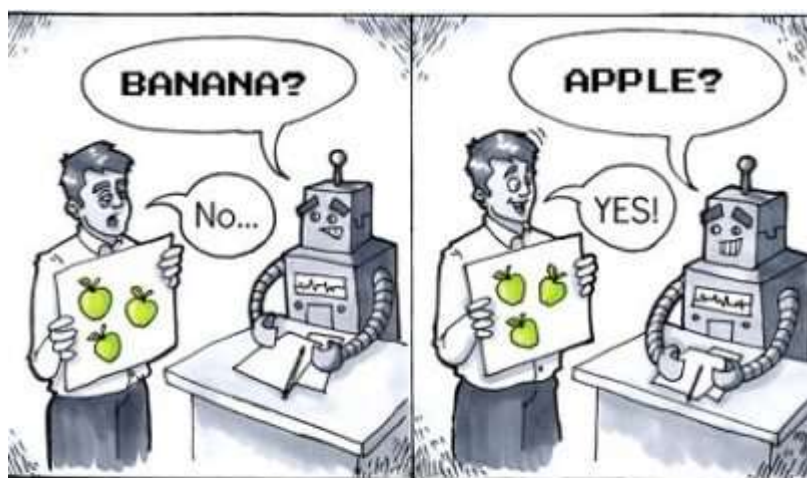
- Supervised Learning – You supervise the learning process, meaning the data that you have collected here is labelled and so you know what input needs to be mapped to what output. This helps you correct your algorithm if it makes a mistake in giving you the answer.
- Unsupervised Learning – The data collected here has no labels and you are unsure about the outputs. So you model your algorithm such that it can understand patterns from the data and output the required answer. You do not interfere when the algorithm learns.
- Reinforcement Learning – There is no data in this kind of learning, nor do you teach the algorithm anything. You model the algorithm such that it interacts with the environment and if the algorithm does a good job, you reward it, else you punish the algorithm. With continuous interactions and learning, it goes from being bad to being the best that it can for the problem assigned to it.

Now that you have a basic idea of what is Machine Learning and the different types of Machine Learning, let us dwell into the actual topic for discussion here and answer What is Supervised Learning? Where is Supervised Learning used? What are the types of Supervised Learning? Supervised Learning Algorithms and much more!

### What is Supervised Learning?

Supervised Learning is the process of making an algorithm to learn to map an input to a particular output. This is achieved using the labelled datasets that you have collected. If the mapping is correct, the algorithm has successfully learned. Else, you make the necessary changes to the algorithm so that it can learn correctly. Supervised Learning algorithms can help make predictions for new unseen data that we obtain later in the future.

This is similar to a teacher-student scenario. There is a teacher who guides the student to learn from books and other materials. The student is then tested and if correct, the student passes. Else, the teacher tunes the student and makes the student learn from the mistakes that he or she had made in the past. That is the basic principle of Supervised Learning.



## Supervised Learning

Let me give another real-life example that can help you understand what exactly is Supervised Learning.

### Example of Supervised Learning

Suppose you have a niece who has just turned 2 years old and is learning to speak. She knows the words, Papa and Mumma, as her parents have taught her how she needs to call them. You want to teach her what a dog and a cat is. So what do you

do? You either show her videos of dogs and cats or you bring a dog and a cat and show them to her in real-life so that she can understand how they are different.



Now there are certain things you tell her so that she understands the differences between the 2 animals.

- Dogs and cats both have 4 legs and a tail.
- Dogs come in small to large sizes. Cats, on the other hand, are always small.
- Dogs have a long mouth while cats have smaller mouths.
- Dogs bark while cats meow.
- Different dogs have different ears while cats have almost the same kind of ears.

Now you take your niece back home and show her pictures of different dogs and cats. If she is able to differentiate between the dog and cat, you have successfully taught her.

So what happened here? You were there to guide her to the goal of differentiating between a dog and a cat. You taught her every difference there is between a dog and a cat. You then tested her if she was able to learn. If she was able to learn, she called the dog as a dog and a cat as a cat. If not, you taught her more and were able to teach her. You acted as the supervisor and your niece acted as the algorithm that had to learn. You even knew what was a dog and what was a cat. Making sure that she was learning the correct thing. That is the principle that Supervised Learning follows.

Now with having a basic understanding of what Supervised Learning is, let's also understand what makes this kind of learning important.

Why is it Important?

- Learning gives the algorithm experience which can be used to output the predictions for new unseen data

- Experience also helps in optimizing the performance of the algorithm
- Real-world computations can also be taken care of by the Supervised Learning algorithms

With the importance of Supervised Learning understood, let's take a look at the types of Supervised Learning along with the algorithms!

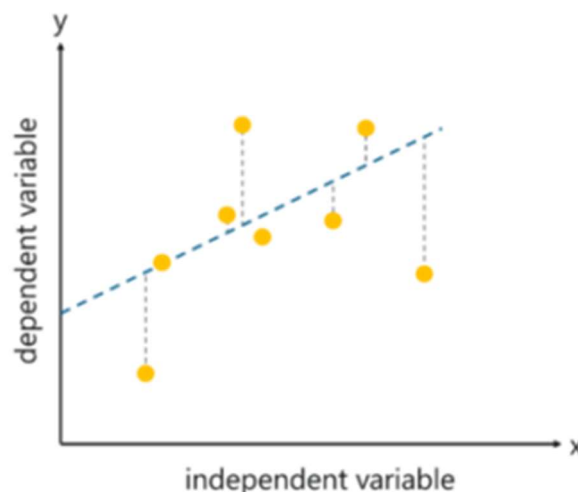
### Types of Supervised Learning

Supervised Learning has been broadly classified into 2 types.

- **Regression**
- **Classification**

Regression is the kind of Supervised Learning that learns from the Labelled Datasets and is then able to **predict a continuous-valued output** for the new data given to the algorithm. It is used whenever the output required is a number such as money or height etc. Some popular Supervised Learning algorithms are discussed below:

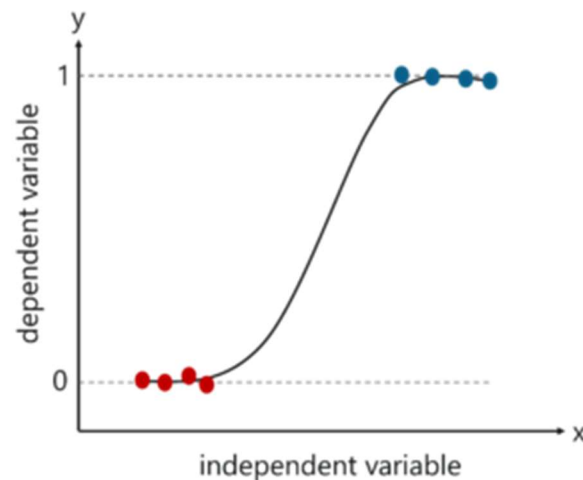
- **Linear Regression** – This algorithm assumes that there is a linear relationship between the 2 variables, Input (X) and Output (Y), of the data it has learnt from. The **Input variable** is called the *Independent Variable* and the **Output variable** is called the *Dependent Variable*. When unseen data is passed to the algorithm, it uses the function, calculates and maps the input to a continuous



value for the output.

- **Logistic Regression** – This algorithm predicts discrete values for the set of Independent variables that have been passed to it. It does the prediction by mapping the unseen data to the **logit function** that has been programmed into

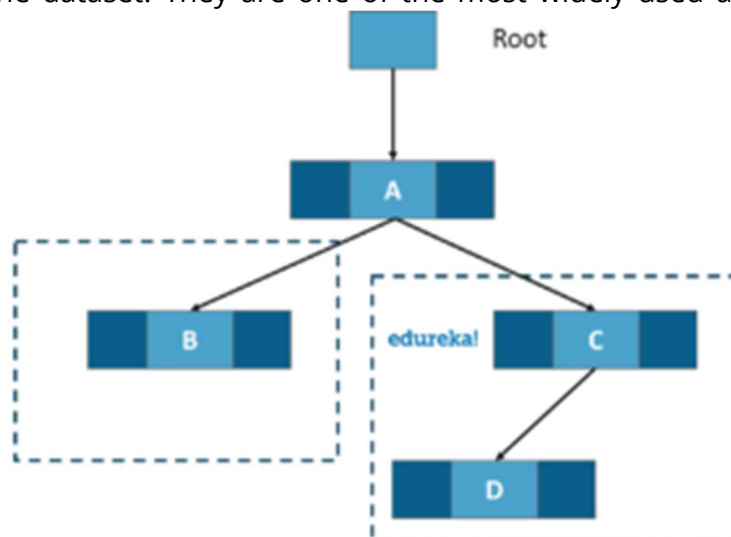
it. The algorithm predicts the probability of the new data and so its output lies



between the range of 0 and 1.

Classification, on the other hand, is the kind of learning where the algorithm needs to map the new data that is obtained to any one of the 2 classes that we have in our dataset. The classes need to be mapped to either 1 or 0 which in real-life translated to 'Yes' or 'No', 'Rains' or 'Does Not Rain' and so forth. The output will be either one of the classes and not a number as it was in Regression. Some of the most well-known algorithms are discussed below:

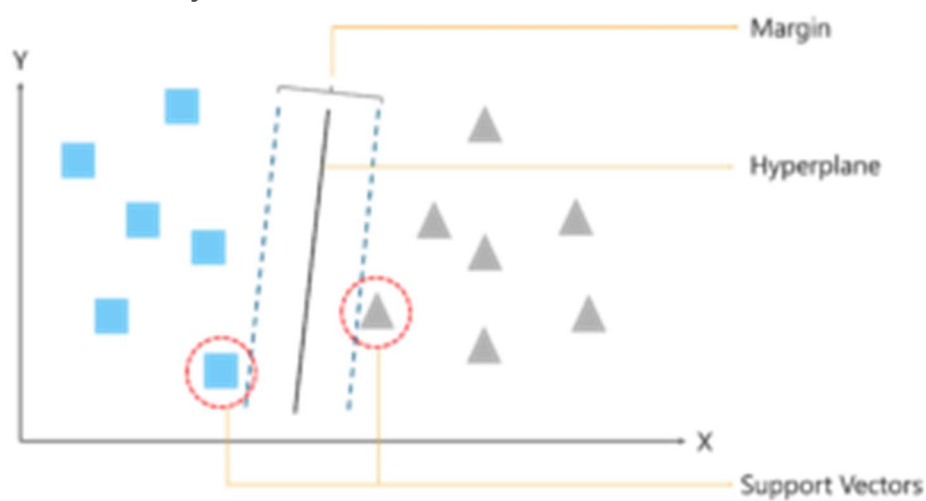
- **Decision Tree** – Decision Trees classify based on the feature values. They use the method of **Information Gain** and find out which feature of the dataset gives the best of information, make that as the root node and so on till they are able to classify each instance of the dataset. Every branch in the Decision Tree represents a feature of the dataset. They are one of the most widely used algorithms for



classification.

- **Naive Bayes Classifier** – Naive Bayes algorithms assume that the features of the dataset are all independent of each other. They work great on large datasets. Directed Acyclic Graphs (DAG) is used for the purpose of classification.
- **Support Vector Machines (SVM)** – SVM algorithms are based on the statistical learning theory of Vap Nik. They use Kernel functions which are a central concept

for most of the learning tasks. These algorithms create a hyper-plane that is used to classify the two classes from each other.



So, I hope you have a clear understanding of the 2 types of Supervised Learning and a few of the most popular algorithms in them. Let's move over to its applications.

### Applications of Supervised Learning

Supervised Learning Algorithms are used in a variety of applications. Let's go through some of the most well-known applications.

- **BioInformatics** – This is one of the most well-known applications of Supervised Learning because most of us use it in our day-to-day lives. BioInformatics is the storage of Biological Information of us humans such as fingerprints, iris texture, earlobe and so on. Cellphones of today are capable of learning our biological information and are then able to authenticate us bringing up the security of the system. Smartphones such as iPhones, Google Pixel are capable of facial recognition while OnePlus, Samsung is capable of In-display finger recognition.
- **Speech Recognition** – This is the kind of application where you teach the algorithm about your voice and it will be able to recognize you. The most well-known real-world applications are virtual assistants such as Google Assistant and Siri, which will wake up to the keyword with your voice only.
- **Spam Detection** – This application is used where the unreal or computer-based messages and E-Mails are to be blocked. G-Mail has an algorithm that learns the different keywords which could be fake such as “You are the winner of something” and so forth and blocks those messages directly. OnePlus Messages App gives the user the task of making the application learn which keywords need to be blocked and the app will block those messages with the keyword.
- **Object-Recognition for Vision** – This kind of application is used when you need to identify something. You have a huge dataset which you use to teach your algorithm and this can be used to recognize a new instance. Raspberry Pi algorithms which detect objects are the most well-known example.

Those were some of the places where Supervised Learning has shined and shown its grit in the real world of today. With that, let us move over to the differences between Supervised and Unsupervised learning.

### Supervised vs. Unsupervised Learning

Parameter	Supervised Learning	Unsupervised Learning
Dataset	Labelled	Unlabelled
Method of Learning	Guided learning	The algorithm learns by itself using dataset
Complexity	Simpler method	Computationally complex
Accuracy	More Accurate	Less Accurate

### Disadvantages of Supervised Learning

Supervised Learning has a lot of challenges and disadvantages that you could face while working with these algorithms. Let's take a look at these.

- You could overfit your algorithm easily
- Good examples need to be used to train the data
- Computation time is very large for Supervised Learning
- Unwanted data could reduce the accuracy
- Pre-Processing of data is always a challenge
- If the dataset is incorrect, you make your algorithm learn incorrectly which can bring losses

So now that we have finished all the disadvantages, let's retrace back and summarize what we have learnt today.

We had an overview of what Machine Learning is and its various types. We then understood in depth of what supervised learning is, why is it so important. Later, we went through the various types of supervised Learning which are regression and classification. After that, we discussed the various algorithms, the applications of supervised Learning, differences between Supervised and Unsupervised Learning and the disadvantages that you may face when you work with supervised Learning Algorithms.