**Supervised vs Unsupervised Learning: Difference Between Them**

**Key Difference Between Supervised and Unsupervised Learning**

* In Supervised learning, you train the machine using data which is well “labeled.”
* Unsupervised learning is a machine learning technique, where you do not need to supervise the model.
* Supervised learning allows you to collect data or produce a data output from the previous experience.
* Unsupervised machine learning helps you to finds all kind of unknown patterns in data.
* Regression and Classification are two types of supervised machine learning techniques.
* Clustering and Association are two types of Unsupervised learning.
* In a supervised learning model, input and output variables will be given while with unsupervised learning model, only input data will be given

**In this Difference between tutorial, you will learn:**

**What is Supervised Machine Learning?**

In Supervised learning, you train the machine using data which is well **“labeled**.” It means some data is already tagged with the correct answer. It can be compared to learning which takes place in the presence of a supervisor or a teacher.

A supervised learning algorithm learns from labeled training data, helps you to predict outcomes for unforeseen data. Successfully building, scaling, and deploying accurate supervised machine learning Data science model takes time and technical expertise from a team of highly skilled data scientists. Moreover, Data scientist must rebuild models to make sure the insights given remains true until its data changes.

**What is Unsupervised Learning?**

Unsupervised learning is a machine learning technique, where you do not need to supervise the model. Instead, you need to allow the model to work on its own to discover information. It mainly deals with the unlabelled data.

Unsupervised learning algorithms allow you to perform more complex processing tasks compared to supervised learning. Although, unsupervised learning can be more unpredictable compared with other natural learning deep learning and reinforcement learning methods.

**Why Supervised Learning?**

* Supervised learning allows you to collect data or produce a data output from the previous experience.
* Helps you to optimize performance criteria using experience
* Supervised machine learning helps you to solve various types of real-world computation problems.

**Why Unsupervised Learning?**

Here, are prime reasons for using Unsupervised Learning:

* Unsupervised machine learning finds all kind of unknown patterns in data.
* Unsupervised methods help you to find features which can be useful for categorization.
* It is taken place in real time, so all the input data to be analyzed and labeled in the presence of learners.
* It is easier to get unlabeled data from a computer than labeled data, which needs manual intervention.

**How Supervised Learning works?**

For example, you want to train a machine to help you predict how long it will take you to drive home from your workplace. Here, you start by creating a set of labeled data. This data includes

* Weather conditions
* Time of the day
* Holidays

All these details are your inputs. The output is the amount of time it took to drive back home on that specific day.



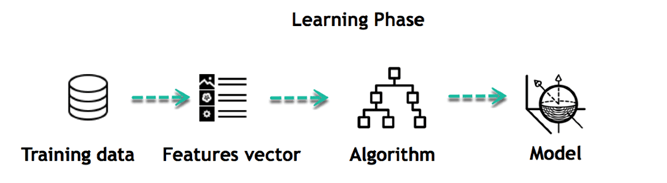
How Supervised Learning works

You instinctively know that if it’s raining outside, then it will take you longer to drive home. But the machine needs data and statistics.

Let’s see now how you can develop a supervised learning model of this example which help the user to determine the commute time. The first thing you requires to create is a training data set. This training set will contain the total commute time and corresponding factors like weather, time, etc. Based on this training set, your machine might see there’s a direct relationship between the amount of rain and time you will take to get home.

So, it ascertains that the more it rains, the longer you will be driving to get back to your home. It might also see the connection between the time you leave work and the time you’ll be on the road.

The closer you’re to 6 p.m. the longer time it takes for you to get home. Your machine may find some of the relationships with your labeled data.



Learning Phase

This is the start of your Data Model. It begins to impact how rain impacts the way people drive. It also starts to see that more people travel during a particular time of day.

**How Unsupervised Learning works?**

Let’s, take the case of a baby and her family dog.



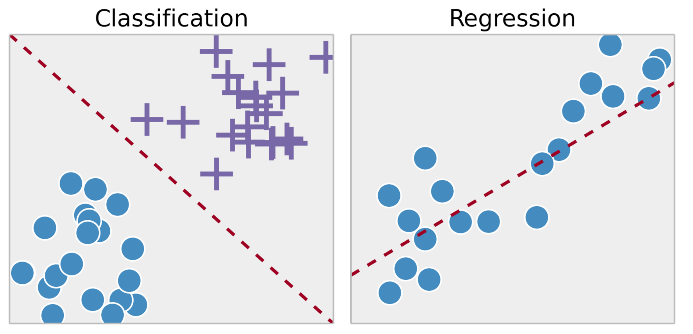
How Unsupervised Learning works

She knows and identifies this dog. A few weeks later a family friend brings along a dog and tries to play with the baby.



Baby has not seen this dog earlier. But it recognizes many features (2 ears, eyes, walking on 4 legs) are like her pet dog. She identifies a new animal like a dog. This is unsupervised learning, where you are not taught but you learn from the data (in this case data about a dog.) Had this been supervised learning, the family friend would have told the baby that it’s a dog.

**Types of Supervised Machine Learning Techniques**



Types of Supervised Machine Learning Techniques

**Regression:**

Regression technique predicts a single output value using training data.

Example: You can use regression to predict the house price from training data. The input variables will be locality, size of a house, etc.

**Classification:**

Classification means to group the output inside a class. If the algorithm tries to label input into two distinct classes, it is called binary classification. Selecting between more than two classes is referred to as multiclass classification.

**Example**: Determining whether or not someone will be a defaulter of the loan.

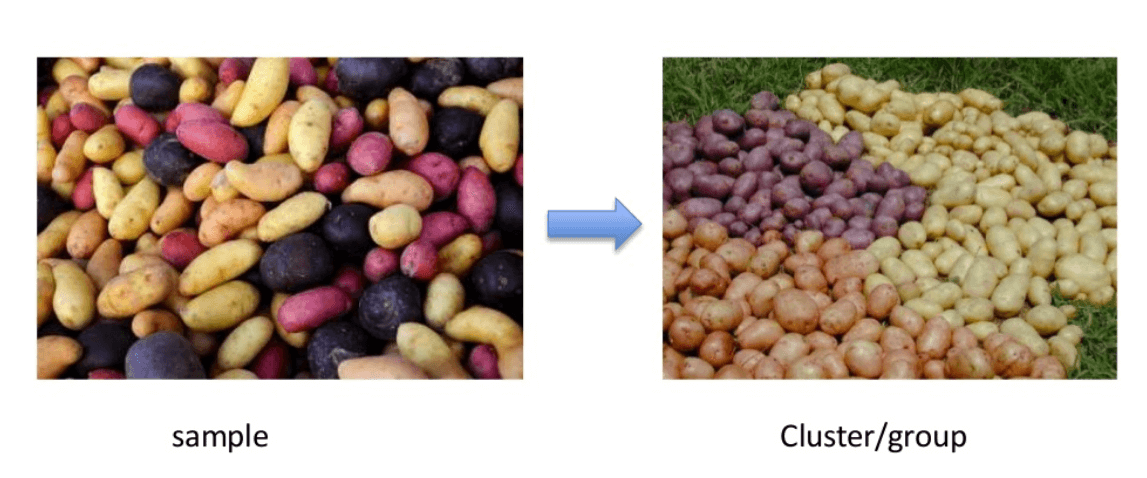
**Strengths**: Outputs always have a probabilistic interpretation, and the algorithm can be regularized to avoid overfitting.

**Weaknesses**: Logistic regression may underperform when there are multiple or non-linear decision boundaries. This method is not flexible, so it does not capture more complex relationships.

**Types of Unsupervised Machine Learning Techniques**

Unsupervised learning problems further grouped into clustering and association problems.

**Clustering**



Clustering is an important concept when it comes to unsupervised learning. It mainly deals with finding a structure or pattern in a collection of uncategorized data. Clustering algorithms will process your data and find natural clusters(groups) if they exist in the data. You can also modify how many clusters your algorithms should identify. It allows you to adjust the granularity of these groups.

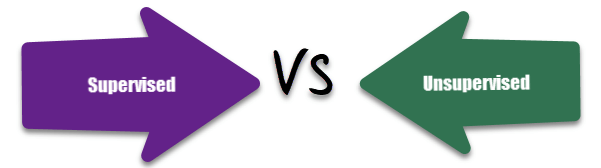
**Association**

Association rules allow you to establish associations amongst data objects inside large databases. This unsupervised technique is about discovering exciting relationships between variables in large databases. For example, people that buy a new home most likely to buy new furniture.

Other Examples:

* A subgroup of cancer patients grouped by their gene expression measurements
* Groups of shopper based on their browsing and purchasing histories
* Movie group by the rating given by movies viewers

**Difference Between Supervised and Unsupervised Learning**



Supervised vs. Unsupervised Learning

| **Parameters** | **Supervised machine learning technique** | **Unsupervised machine learning technique** |
| --- | --- | --- |
| Process | In a supervised learning model, input and output variables will be given. | In unsupervised learning model, only input data will be given |
| Input Data | Algorithms are trained using labeled data. | Algorithms are used against data which is not labeled |
| Algorithms Used | Support vector machine, Neural network, Linear and logistics regression, random forest, and Classification trees. | Unsupervised algorithms can be divided into different categories: like Cluster algorithms, K-means, Hierarchical clustering, etc. |
| Computational Complexity | Supervised learning is a simpler method. | Unsupervised learning is computationally complex |
| Use of Data | Supervised learning model uses training data to learn a link between the input and the outputs. | Unsupervised learning does not use output data. |
| Accuracy of Results | Highly accurate and trustworthy method. | Less accurate and trustworthy method. |
| Real Time Learning | Learning method takes place offline. | Learning method takes place in real time. |
| Number of Classes | Number of classes is known. | Number of classes is not known. |
| Main Drawback | Classifying big data can be a real challenge in Supervised Learning. | You cannot get precise information regarding data sorting, and the output as data used in unsupervised learning is labeled and not known. |