

# Types of Machine Learning

by Taymour | April 12, 2022

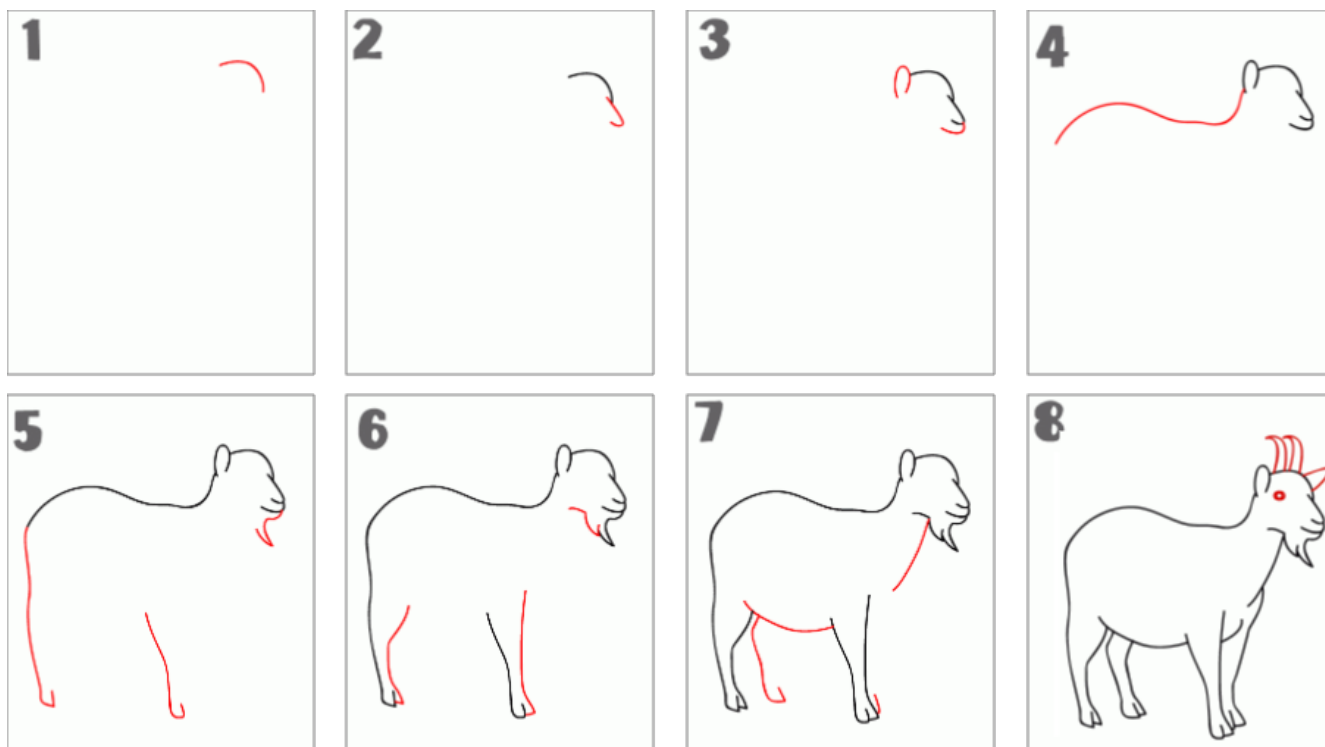
Just when you thought you had machine learning (ML) figured out, you find a new type waiting to be discovered. If you're a marketer or data scientist, staying up-to-date on the different types of machine learning and how they can be used to improve your campaigns and analytics is important. In this blog post, we'll explore the three main types of machine learning:

- Supervised learning
- Semi-supervised learning
- Unsupervised learning

We'll also discuss some of the applications for each type of machine learning. So, what are these different types of machine learning? Let's find out!

## Supervised Learning

*Supervised learning* is when the training data has both the input and output. For example, the training set would include images with a goat and images without a goat, and each image would have features of a goat or something that is not a goat. In this way, you supervise the machine's learning with both the input (e.g., explicitly telling the computer a goat's specific attributes: arched back, size and position of head and ears, size and relative location of legs, presence and shape of backward-arching horns, size and presence of a tail, straightness and length of hair, etc.) and the output ("goat" or "not goat" labels), so the machine can learn to recognize a goat's characteristics. In other words, you're getting the computer to learn a classification system you've created. *Classification algorithms* can sort email into folders (including a spam folder).



Most *regression algorithms* are so named for their continuous outputs, which can vary within a range, such as temperature, length, blood pressure, body weight, profit, price, revenue, and so on. The above sales model example would rely on a regression algorithm because revenue—the outcome of interest—is a continuous variable.

## Semi-Supervised Learning

*Semi-supervised learning* occurs when the training data input is incomplete, meaning some portion of it does not have any labels. This approach can be an important check on data bias when humans label it, since the algorithms will also learn from unlabeled data. Gargantuan tasks like web page classification, genetic sequencing, and speech recognition are all tasks that can benefit greatly with a semi-supervised approach to learning.

## Unsupervised Learning

*Unsupervised learning* happens when the training data input contains no output labels at all (most “big data” often lacks output labels), which means you want the machine to learn how to do something without telling it specifically how to do so. This method is useful for identifying data patterns and structures that might otherwise escape human detection. Identified groupings, clusters, and data point categories can result in “ah-ha” insights. In the goat example, the algorithm would evaluate different animal groups and label goats and other animals on its own. Unlike in supervised learning, no mention is made of a goat’s specific characteristics. This is an example of *feature learning*, which automatically discovers the representations needed for feature detection or classification from raw data, thereby allowing the machine to both learn the features and then use them to perform a specific task.

Another example of unsupervised learning is *dimensionality reduction*, which lessens the number of random variables under consideration in a dataset by identifying a group of principal variables. For reasons beyond the scope of this paper, sparsity is always an objective in model building for both parametric and non-parametric models.

Unsupervised machine learning algorithms have made their way into business problem-solving and have proven especially useful in digital marketing, ad tech, and exploring customer information. Software platforms and apps that make use of unsupervised learning include Lotame (real-time data management) and Salesforce (customer relationship management software).

## Other Machine Learning Taxonomies

Supervised, semi-supervised, and unsupervised learning are three rather broad ways of classifying different kinds of machine learning. Other ways to provide a taxonomy of machine learning types are out there, but agreement on the best way to classify them is not widespread. For this reason, many other kinds of machine learning exist, a small sampling of which are briefly described below:

- **Reinforcement Learning**

The algorithm receives positive and negative feedback from a dynamic environment as it performs actions. The observations provide guidance to the algorithm. This kind of learning is vital to autonomous cars or to machines playing games against human opponents. It is also what most people probably have in mind when they think of AI. “Reinforced ML uses the technique called exploration/

exploitation. The mechanics are simple—the action takes place, the consequences are observed, and the next action considers the results and errors of the first action.”

- **Active Learning**

When it is impractical or too expensive for a human to label the training data, the algorithm can be set up to query the “teacher” when it needs a data label. This format makes active learning a variation of supervised learning.

- **Meta Learning**

Meta machine learning is where the algorithm gets smarter with each experience. With every new piece of data, the algorithm creates its own set of rules and predictions based on what it has learned before. The algorithm can keep improving by trying different ways to learn- including testing itself.

- **Transduction Learning**

Transduction in supervised learning is the process of predictions being made on new outputs, based off both training inputs/outputs as well as newly introduced inputs. Unlike inductive models which go from general rules to observed training cases, transduction instead goes from specific observed training cases to even more specific test cases. This usually incorporates Convolutional [Neural Networks](#) (CNNs), Deep Belief Networks (DBN), Deep Boltzman Machine (DBM), or Stacked Auto Coders. Transduction is powerful as it can be used for both supervised and unsupervised learning tasks.

There are many transductive machine learning algorithms available, each with its own strengths and weaknesses. Some of the more popular ones include support vector machines (SVMs), decision trees, and k-nearest neighbors (k-NN). SVMs have been particularly successful in applications such as text classification, image classification, and protein function prediction. Decision trees tend to work well when the data is highly structured (such as financial data). K-

NN is another algorithm that is commonly used in transductive settings. for transductive machine learning, as it is relatively simple to implement and can be used for both regression and classification tasks.

- **Deep Learning**

Deep learning is a subset of machine learning that deals with algorithms inspired by the structure and function of the brain. Deep learning is part of a broader family of machine learning methods based on artificial neural networks with representation learning. It uses networks such as deep neural networks, deep belief networks, and recurrent neural networks. These have been applied to a variety of fields like computer vision, machine hearing, bioinformatics drug design natural language processing speech recognition and robotics.

The deep learning algorithm has been shown to outperform other methods in many tasks. However, deep learning is also a complex field, and requires significant expertise to design and train effective models. In addition, deep learning algorithms are often resource intensive, requiring large amounts of data and computational power. As a result, deep learning is typically used in combination with other machine learning methods, such as shallow neural networks or support vector machines.

- **Robot Learning**

Algorithms create their own “curricula,” or learning experience sequences, to acquire new skills through self-guided exploration and/or interacting with humans.

One such algorithm, known as Q-learning, is often used in reinforcement learning. Q-learning involves an agent that interacts with its environment by taking actions and receiving rewards or punishments in return. The goal of the agent is to learn a policy that will allow it to maximize its reward.

Q-learning can be used to teach a robot how to perform a task such as opening a door. The robot would first need to be able to identify the door and then figure out how to open it. In order to do this, the robot would need to trial and error different actions until it found the one that worked best.

Once the robot has learned how to open the door, it could then be taught how to close it. This could be done by rewarding the robot for taking the correct action and punishing it for taking the wrong action.

Through Q-learning, robots can learn a wide variety of skills. This type of learning is becoming increasingly important as robots are being used in more complex tasks such as search and rescue, manufacturing, and healthcare.

## • **Ensemble Learning**

This approach combines multiple models to improve predictive capabilities compared to using a single model. "Ensemble methods are meta-algorithms that combine several machine learning techniques into one predictive model in order to *decrease variance* (bagging), *decrease bias* (boosting), or *improve predictions* (stacking)."

This approach is often used in data science competitions, where the goal is to achieve the highest accuracy possible. Ensemble machine learning has been used to win many well-known competitions, such as the Netflix Prize and the KDD Cup.

While ensemble machine learning can be very effective, it is important to remember that it comes with some trade-offs. This approach can be more complex and time-consuming than using a single model. In addition, the results of ensemble machine learning can be difficult to interpret.

Today's powerful computer processing and colossal data storage capabilities enable ML algorithms to charge through data using brute force. At a micro-level, the evaluation typically consists of answering a "yes" or "no" question. Most ML algorithms employ a *gradient descent* approach (please see a related post [Neural Networks \(NN Models\)](#) as an example of supervised learning .known as *back propagation*. To do this efficiently, they follow a sequential iteration process.

In lieu of multivariate differential calculus, an analogy may help the reader better understand this class of algorithms. Let's say a car traverses through a figurative mountain range with peaks and valleys (loss function). The algorithm helps the car navigate through different locations (parameters) to find the mountain range's lowest point by finding the path with the steepest descent. To overcome the problem of the car getting stuck in a valley (local minimum) that is not the lowest valley (global minimum), the algorithm sends multiple cars, each with the same mission, to different locations. After

every turn, the algorithm learns whether the cars' pathways help or hinder the overarching goal of getting to the lowest point.

More technically, the gradient descent algorithm's goal is to minimize the *loss function*. To improve the prediction's accuracy, the parameters are *hyper-tuned*, a process in which different model families are used to find a more efficient path to the global minimum. Sticking with the mountain range analogy, an alternative route may fit the mountain topology better than others, thereby enabling a car to find the lowest point faster. Most of these ML models make relatively few assumptions about the data (e.g., continuity and differentiability are assumed for gradient descent algorithms) or virtually no assumptions (e.g., most clustering algorithms, such as K-means or random forests); the essential characteristic that makes it "machine learning" is the consecutive processing of data using trial and error. Stochastic (can be analyzed statistically) models make assumptions about how the data is distributed: "I've seen that mountain range before, and I don't need to send a whole fleet of cars to accomplish the mission." Related blog post: [Statistical Modeling vs. Machine Learning](#)

How iteration can lead to more precise predictions than stochastic methods should make sense to you: The exact routes are calculated rather than estimated with statistics. However, because the loss (a.k.a. cost) function will usually be applied to a slightly different mountain range, the statistical models, which are more generalizable, may ultimately lead to a superior prediction (or interpolation). However, while statistics and non-parametric ML are clearly different, non-parametric ML algorithms' iterative elements may be used in a statistical model, and conversely, non-parametric ML algorithms may contain stochastic elements.

A fuller understanding of the differences between these machine learning variations and their highly-specific, continuously-evolving algorithms requires a level of technical knowledge beyond the scope of this paper. A more fruitful avenue of inquiry here is to explore the relationship of ML and statistics as they relate to AI.