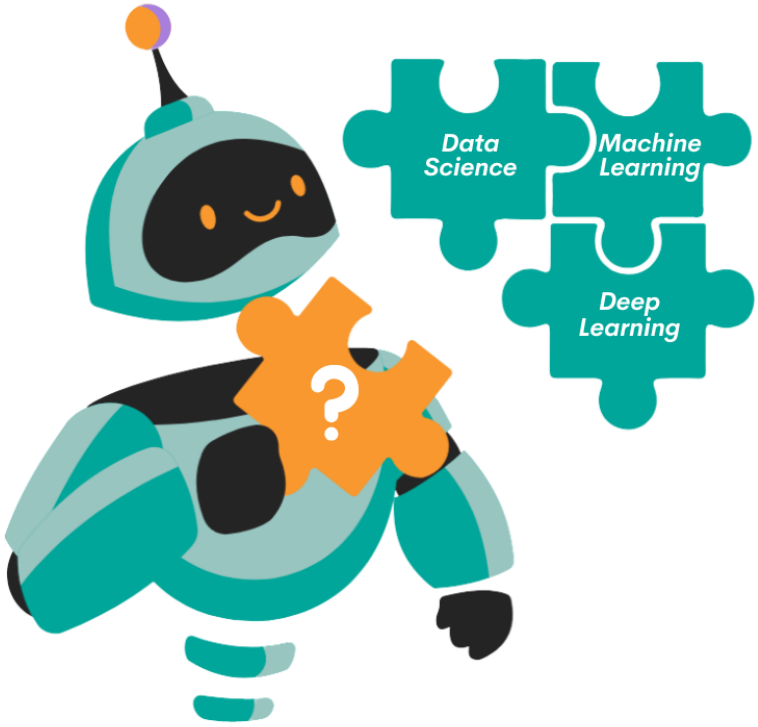
# Data Science vs Machine Learning vs Deep Learning: The Difference

*You never grasped the difference between machine learning and deep learning? This article will help you understand this and how they're used in data science.*




In 2012, the Harvard Business Review posted a viral article describing Data Scientist as the sexiest job in the 21st century. That statement was made not without reason. Nowadays, not only the amount of data grows exponentially bigger, but so does the computational power of a computer, which enables us to perform something with large data that would have been impossible to do in the past.

This, in turn, leads to the introduction of several buzzwords around this subject that you probably have heard so many times until today: *data science*, *artificial intelligence*, *machine learning*, or *deep learning* to mention a few. A large number of buzzwords undoubtedly will lead to confusion, especially for anyone who wants to start their journey in learning data science.

In this article, we are going to discuss the differences between data science vs machine learning vs deep learning. They are three distinct fields of study but closely connected to each other.

Thus, it will make sense to learn about them sequentially, starting from the fundamental one. Here is the structure of what you will learn in this article:

- What data science is
- What artificial intelligence is
- What machine learning is
- What deep learning is
- Machine Learning vs Deep Learning: What are the key differences between Machine Learning and Deep Learning?
- Which one to choose between machine learning and deep learning for your project

So without further ado, let's start with the most fundamental term that you need to know: data science.

## Data Science

On a high level, Data Science is a field that studies how to handle and process data such that we can extract meaningful insight from it. It might sound simple, but there are a lot of things that need to be done to gain insight from the data.

As you might already know, the data that we normally have are not only big but also messy. There are various common sources of mess in our data: missing values, incompatible data types, outliers, or the fact that there are just too many variables in our data.

Data Science helps us to learn about different ways to tackle all those problems such that we can find something meaningful from the data. This means that in data science, we learn the techniques behind *data ingestion*, *data cleaning*, *feature engineering*, *data analysis*, *modeling*, and *data visualization*. Let's dissect the above terms one by one:

## 1. Data Ingestion

This stage describes the process of identifying and assembling relevant data that can be helpful for our analysis from one or several data sources. The data sources can be traditional databases, clusters, or cloud services.

## 2. Data Cleaning

This stage describes the process of correcting inaccurate records within the dataset. This includes removing duplicate records, correcting data types, imputing missing values, or removing records with missing values. Different use cases need different data-cleaning techniques.

## 3. Feature engineering

In some use cases, probably the data that we have are not sufficient enough in order for us to gain insight from them. Thus, what we can do normally is to extract a new feature out of our data such that we get a slightly different perspective from our data. Because of that, this process normally requires domain knowledge.

## 4. Data Analysis

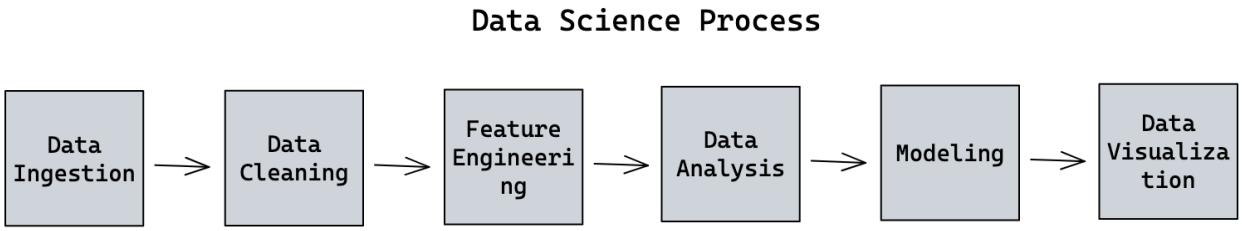
This is the step where we explore and analyze the clean data using Exploratory Data Analysis (EDA) technique. The analysis requires us to know the concept behind statistics, such as hypothesis testing, probability distributions, the measure of spread, etc.

## 5. Modeling

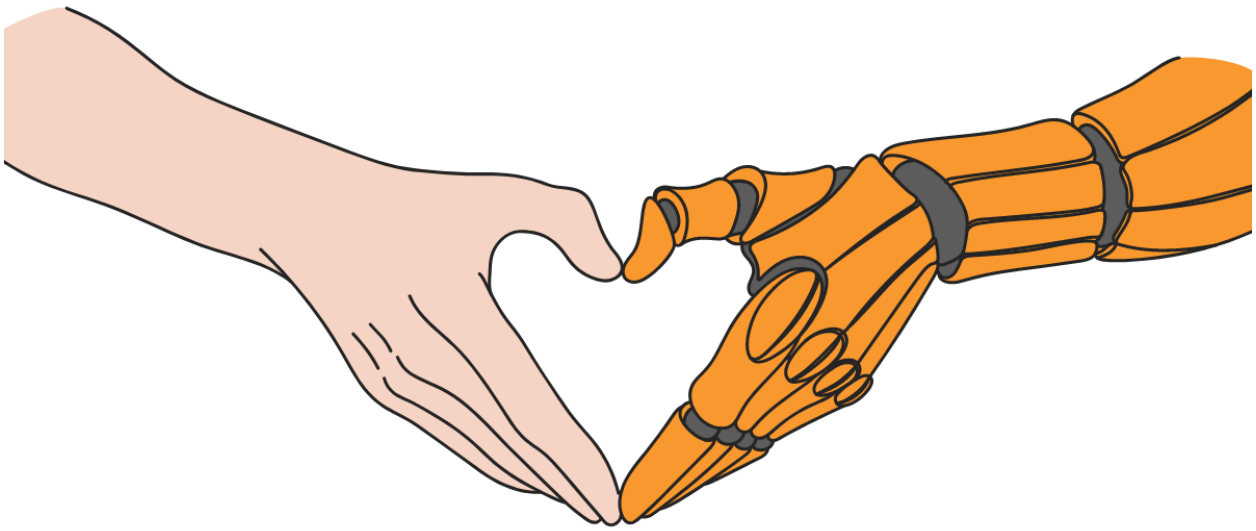
In this step, we want to use the data that we have in order to predict a future event, and this is where the concept of artificial intelligence, machine learning, or even deep learning comes into place. We will discuss all these three terms in more depth in the next section.

### 6. Data visualization

As the name suggests, data visualization aims to visualize the insight that we gain from data. Data visualization helps tremendously to make the insight that we found becomes more presentable and easier to understand by other people.




## Artificial Intelligence



In the previous section, you've seen that one of many parts of the data science process involves dealing with artificial intelligence (AI), but what is AI, exactly? In a nutshell, AI is a field of study that combines the concept of computer science and the use of datasets to solve problems in many use cases with minimal human intervention.

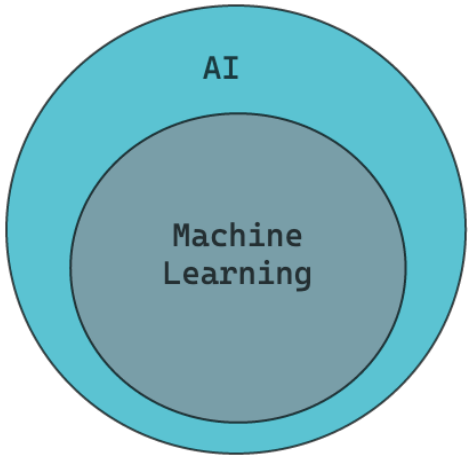
The end product of an AI is a computer that can do a specific task that we intended it to do. Today, AI systems have been applied to many different areas within our life, such as:

- **Customer service**, as an example, chatbots that you usually find in an online shop to assist you with frequently asked questions (FAQ), payments, etc.
- **Recommendation system** that helps you to find the movies that you probably like on movie streaming platforms, or that helps you to find the items that you might also like on an online shop, etc.
- **Fraud detection** that is used in financial services like banks to detect suspicious transactions.

Now that we know the concept of AI, we are ready to deep dive into the next topic, which is machine learning.

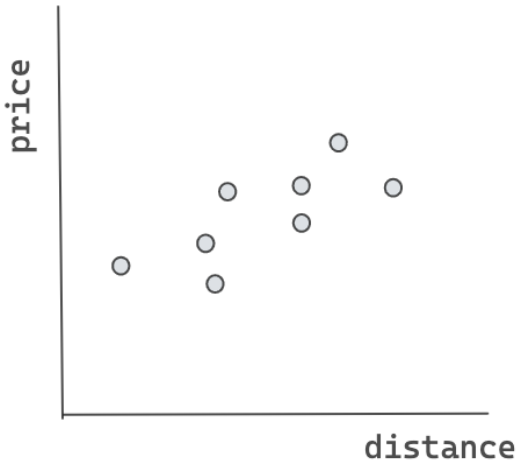
## Machine Learning

Machine learning is a subset of artificial intelligence (AI) in which we use a specific algorithm to process data such that we can use the algorithm to predict the value of unseen data. However, the algorithm that we use in machine learning is designed in such a way that it needs little intervention from us as a programmer.

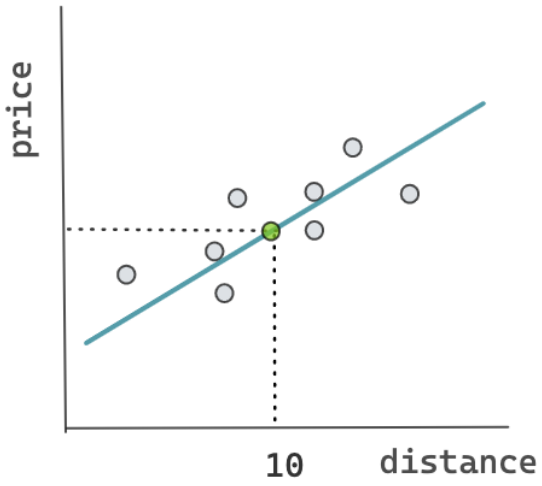



When we create an algorithm for software, normally, what we would do is write a bunch of code with rules and commands that the software should do. In machine learning, however, we let the algorithm learn by itself without the need of us to program it explicitly. This means that we let the computer solve a specific task automatically by looking at the pattern of data.

Let's use an example to make the concept clear. Let's say that you want to take a taxi from your house to the airport, and you know the distance between your house and the airport is roughly 10 km. The thing is, you're not sure how much you need to pay for the taxi to get there. However, what you do have is the data from previous taxi passengers that records the distance traveled and the price that they need to pay for the taxi.




With the data that you have, now the algorithm can draw a line: either a straight line or a non-linear line, depending on the complexity of the data points. Now, since you know that the distance between your house and the airport is 10 km, then you can put this information into the algorithm, and it will give you the prediction of the money that you should spend on a taxi.




In the use case above, we implement a machine learning algorithm called linear regression. You might have noticed that there is nothing fancy about the algorithm above, and we are just implementing the concept of statistics. However, that is the gist of it. The base concept of all machine learning algorithms is statistics, and that's why we need to know about statistics before we delve deep into machine learning.

Common machine learning algorithms that we see in practice include:

## 1. Linear regression

Linear regression is a machine learning algorithm that is used to make a prediction of a value of a variable based on other variables. The variable that we want to predict is called a *dependent variable*, while the variables that we use to help the algorithm to predict the value of our dependent variable are called *independent variables*. The main prerequisite of a linear regression algorithm is that our dependent variable should have a continuous value.

Take the above example, our dependent variable is the price of the taxi, whilst the independent variable would be the distance. If you notice, the price has a continuous value, and thus linear regression would be a perfect algorithm to use in that case.

You can learn more about the inner workings of linear regression and other regression algorithms → Regression Machine Learning Algorithm.

## 2. Logistic Regression

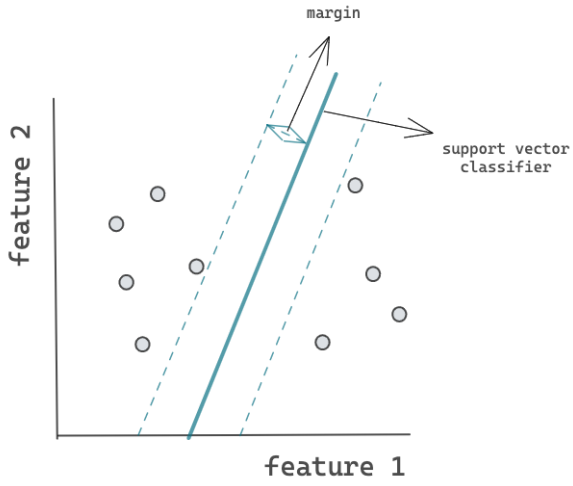
Although it has the word 'regression' on its name, logistic regression is not the same as linear regression. The gist of this algorithm is still the same as linear regression, we have a dependent variable that we want to predict and also independent variables that help the algorithm to make a prediction.

The main difference between linear regression and logistic regression is that logistic regression takes a binary value for the dependent variable instead of a continuous value. For example, whether a customer would churn a credit or not, whether an email is spam or not, whether a student will pass or fail his exam, etc. Thus, logistic regression is one of the algorithms that is commonly used for classification purposes.

You can learn more about the inner workings of logistic regression and other classification algorithms → Classification in Machine Learning.

## 3. Support Vector Machine

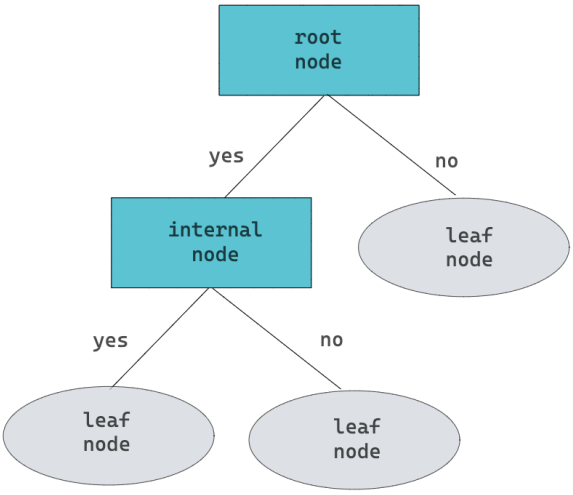
A support vector machine is a machine learning algorithm that can be used for both classification and regression purposes. The main idea behind the support vector machine algorithm is to create a hyperplane that separates different data points with the widest possible margin.




As you can see from the visualization above, the separator on the right-hand side is much more preferable than the one on the left-hand side. If you want to learn more about support vector machines in general, check out this article.

## 4. Decision Tree

Same as support vector machines, a decision tree is a machine learning algorithm that can be used for both classification and regression purposes. As the name suggests, it has a tree-like structure to represent the decision logic of the algorithm.




A decision tree algorithm has a root, which normally is occupied by the purest variable. Next, depending on the value of the data point that we want to predict, it will be passed to either one of the branches into an internal node, where the value of our data point will be assessed again. This cycle will continue until we reach the leaf node, which contains the prediction value of our data point.

You can learn more about the inner workings of a decision tree algorithm → Decision Tree and Random Forest Algorithm Explained.

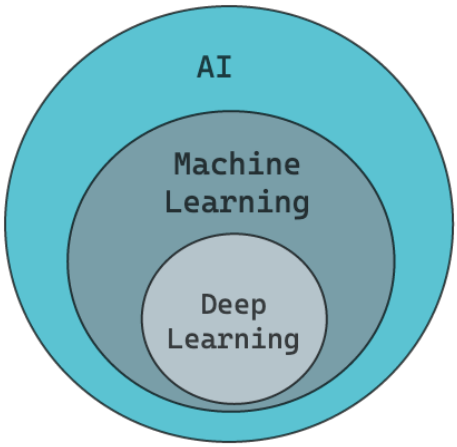
Now that we know about machine learning, let's get into the next concept, which is deep learning.

# Deep Learning

Machine learning algorithms can do a variety of tasks and work very well when we have structured data, such as tabular data for example. But what if we have unstructured data, such as images, texts, or voices? Well, we can use machine learning algorithms to do that, but there are two major drawbacks:

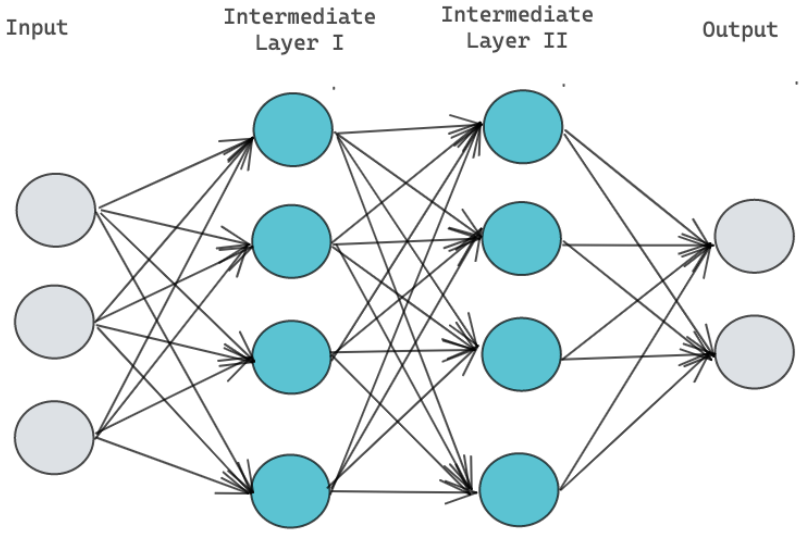
- First, they require a lot of human interventions as we need to pre-process the data into an organized format such that we can feed these data into our machine learning algorithms
- Second, machine learning algorithms' performance can be plateaued even if we supply more data to them.

And this is where we normally shift our attention to deep learning.




Deep learning can be described as a subset of machine learning. This is because deep learning models have more complexity than conventional machine learning models in terms of their architecture.

Deep learning consists of a layered structure called neural networks, as you can see in the visualization below:



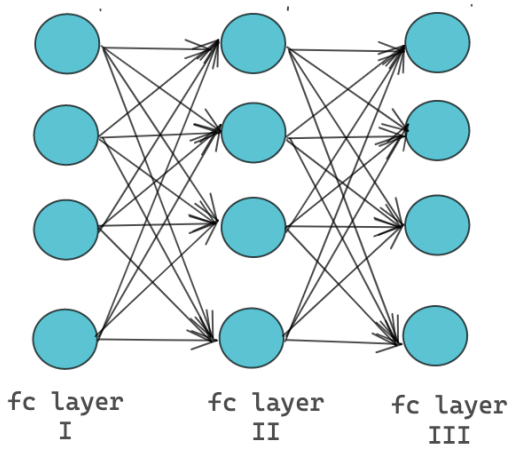

From the visualization above, the leftmost layer is our input features, the rightmost layer is our output, and the layers in between are intermediate layers. In the visualization above, we only have two intermediate layers.

The more intermediate layers our neural network has, the deeper the network will be, the more parameters it will have, and the higher the probability that we will get better performance from our network. Neural networks with two or more intermediate layers can be considered deep, hence the name deep learning.

These deep stacks of intermediate layers in a neural network often result in better performance than conventional machine learning algorithms, especially when we have a lot of data. The intermediate layers of a neural network can differ depending on the application and the goal of our deep learning model. Commonly used layers in a deep learning model are:

## 1. Fully connected layer

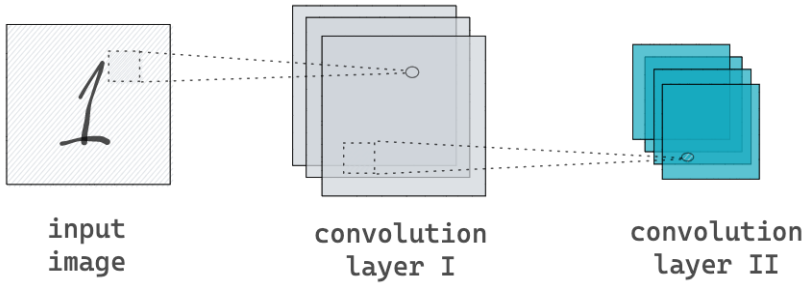
A fully connected layer or commonly also referred to as a linear layer, is one of the commonly used layers in neural networks. This layer connects every input to every neuron, as you can see in the visualization below:




Normally, fully connected layers are only applied in the last few layers of a neural network because it's computationally expensive. In the last few layers of a common deep learning model, the important part of the input features has been extracted by other layers. Thus, the dimension of the inputs has also been reduced, and this is where we can apply fully connected layers.

## 2. Convolutional layer

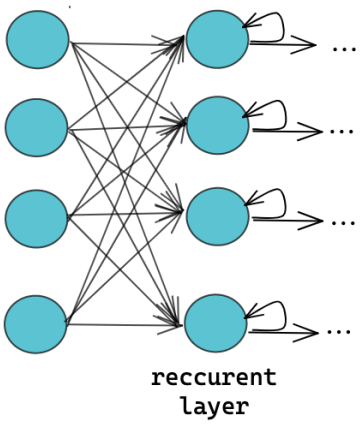
Convolutional layers, as the name suggests, are the main building block of Convolutional Neural Networks (CNN). These layers are commonly used when we're dealing with 2D image data, although they can also deal with 1D and 3D data.




The convolutional layers' main purpose is feature extraction. They extract important features (normally from image data) such that the dimension of the original data is smaller, but the important information from that feature is still preserved. In a common CNN model, the extracted features from several convolutional layers will normally be passed into a fully connected layer to determine the prediction result.

## 3. Recurrent layer

Recurrent layers are the main building block of Recurrent Neural Networks (RNN) and their variants, such as LSTM and GRU. We normally use this type of neural network when we're dealing with time series data or sequential data, such as stock price prediction, speech recognition, or image captioning.




The main difference between recurrent neural networks and any other type of neural network is their ability to memorize the previous inputs to influence the result of current inputs and outputs. Another difference is that the neurons in each recurrent layer have the same weight, whereas the neurons of any other layers normally have different weights.

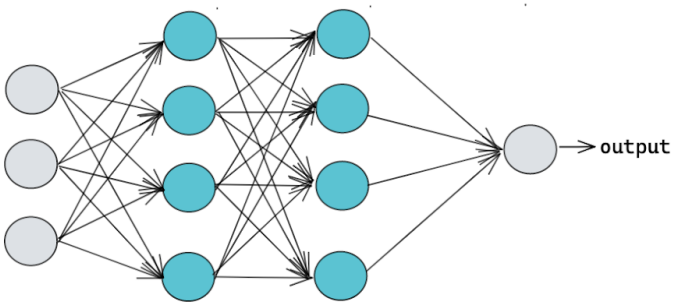
# Machine Learning vs Deep Learning: What are the key differences between ML and Deep Learning?

Now that we know what machine learning and deep learning actually are, you might be wondering: what are the differences between machine learning and deep learning?

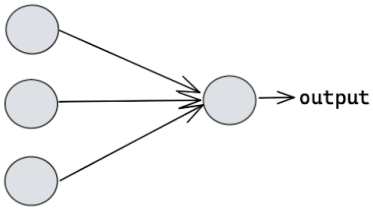
There are several differences between machine learning and deep learning, such as:

## 1. Machine Learning vs Deep Learning: The architecture

Machine learning models tend to have simpler architecture and decision logic than deep learning models. Take logistic regression as an example. It can be described as a one-layer neural network, while deep learning models normally have a deep stack of layers.




neural networks




logistic regression

There are pros and cons that come out of this. With such a deep stack of layers, deep learning models usually perform better than machine learning models, but it comes with the cost that we are ‘forced’ to treat deep learning models as a black box, meaning it’s challenging to explain why our deep learning model behaves the way it does.

Meanwhile, the machine learning model’s decision logic is easier to explain. Thus, in a business setting where the thought process of a model’s prediction is very important, then machine learning models would be preferable instead of deep learning models.

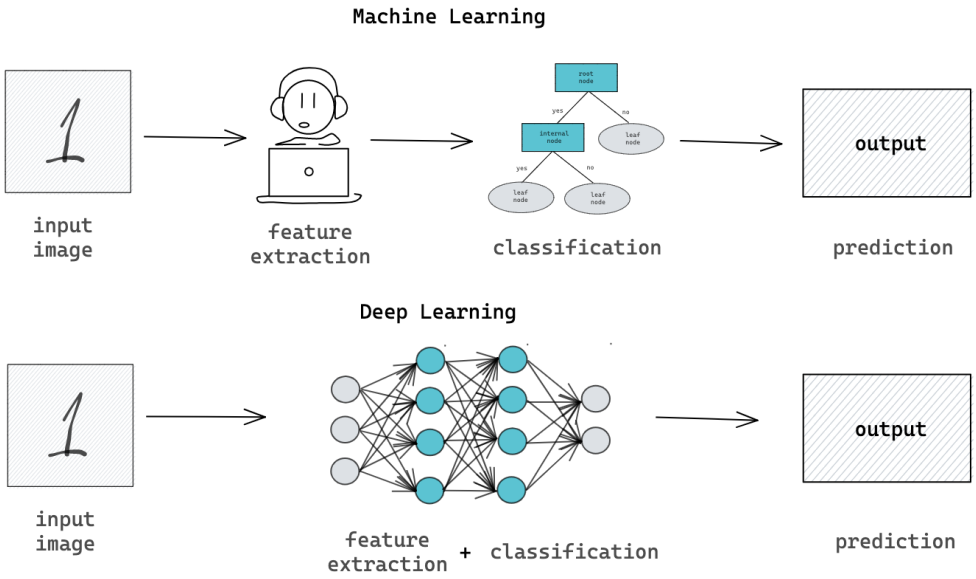
## 2. Machine Learning vs Deep Learning: Human Intervention

Deep learning models require a lot less human intervention in terms of their training process in comparison with machine learning models.

As previously mentioned, machine learning models work best with structured data, such as tabular data. However, this doesn’t mean that they can’t work with unstructured data like images or texts. It means that we need to manually extract the features from images or texts before we are able to train a machine learning model. With a deep learning model, we can just use the

images or texts as our input, and it will do feature extraction automatically during the training process.

As an example, let's say we want to train a model to classify images. Before we can train a machine learning model, we need to manually rearrange the pixel in the image such that it becomes flat and can be fed into a machine learning model of our choice. The consequence of this is that the number of features becomes huge, depending on the size of the image. If we use a deep learning model, we can just feed the model with the image, and the convolutional layers within the model will automatically extract the feature.




## 3. Machine Learning vs Deep Learning: Data, Training Time, and Computational Power

Deep learning models require much more data compared to machine learning models due to their deep stack of layers. The more data we have, the higher the chance our deep learning model will perform properly. If we have sufficient amounts of data, then both machine learning and deep learning models would perform similarly well, but deep learning models will have a higher plateau in terms of their peak performance.

Machine learning models work with thousands of data, while a deep learning model can work with millions of data. This factor, alongside with the architecture of the model, is the reason why machine learning models are much faster to train in comparison with deep learning models.

To sum up, here are some key differences between machine learning and deep learning:

| Aspect | Machine Learning | Deep Learning |
|---|---|---|
| Data Volume | Hundreds or thousands of data | Millions of data |
| Computational Cost | ML model doesn't need a lot of computational power (Using CPU is often sufficient) | DL model needs a lot of computational power (The use of GPU is often necessary) |
| Training Time | ML model takes less time to train | DL model takes more time to train |
| Feature Engineering | Needs to be done explicitly by human | DL model can automatically learn important feature during training |
| Interpretability | ML models' behavior are easier to interpret | DL models' behavior are more difficult to interpret |
| Application Examples | Customer segmentation, Recommendation system, Fraud detection | Natural Language Processing, Computer Vision |

## Machine Learning or Deep Learning: Which One to Choose

As you can see from the previous section, both machine learning and deep learning have their own pros and cons. Depending on the project's requirement, it might be better to use one approach over another, such as:

### 1. Data Types and Availability

If you have structured data such as tabular data and it consists of thousands of entries, then using one of the machine learning algorithms would be enough. Using deep learning for such conditions would be overkill.

Meanwhile, if you have millions of data, then you can shift your attention to deep learning. Deep learning models need plenty of data for them to function properly. Also, if your data are unstructured, such as images, text, or voices, then deep learning is definitely the one to go.

However, this doesn’t mean that machine learning can’t handle unstructured data. Using machine learning when you have unstructured data means that you need to do a lot of work in advance to transform the data into the desired format before you can train a particular machine learning model.

## 2. Hardware availability

Before conducting the modeling step in a data science process, we need to also look at the resources that we have in hand. Deep learning models require huge computational power such that the usage of GPU becomes necessary. This is understandable since a deep learning model has more parameters and, thus, requires a lot more internal computations and a longer training time.

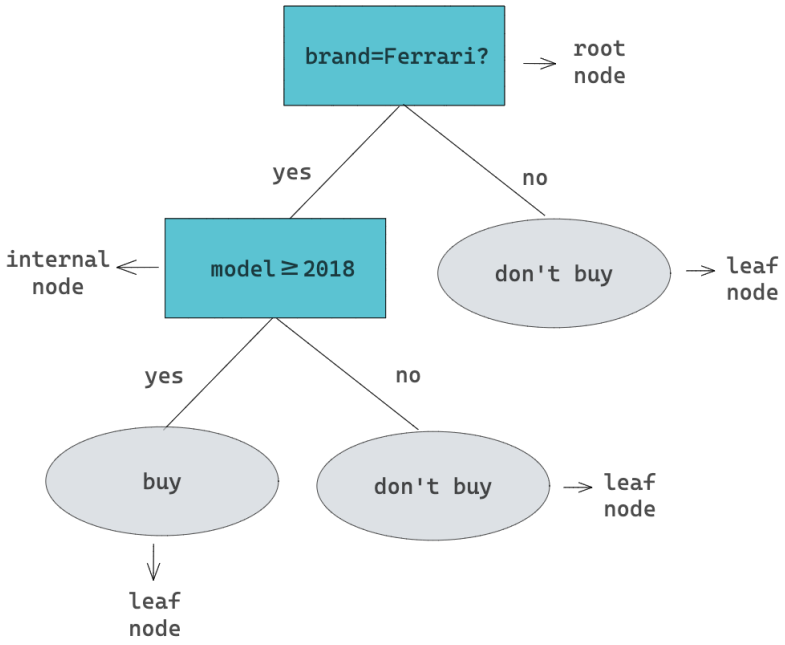
If you only have access to CPUs, then it will be better to use conventional machine learning models.

## 3. Model Interpretability

Depending on the goal of your project, model interpretability would be a crucial factor in the end. Let's say you're developing an AI model to detect malware. During the training process, the model has shown a promising accuracy result. However, the model's performance after it has been deployed varies from time to time: sometimes, it predicts malware correctly, but sometimes it predicts the malware wrongly.

The source of uncertainty of your model becomes difficult to find if you're using deep learning models. As you already know, deep learning models consist of a deep stack of intermediate layers, which makes it difficult for us to understand the decision logic of a model.

Meanwhile, conventional machine learning algorithms tend to have simpler architecture, and thus, the decision logic can be easily tracked and understood. As an example, let's say that we have built a decision tree classifier to predict whether a person would buy a car or not. We can visualize the decision logic of the model as follows:




From the visualization above, we can follow the decision logic of our decision tree in a simple way. If a car's brand is not Ferrari, then a person wouldn't buy the car. If yes, then he/she will look at the model of the car. If the car's model is 2018 or newer, then he/she will buy the car.

However, if you really need to use deep learning models but also want to gain insight into how your model predicts an outcome, there are now different methods that you can use to at least interpret the model's prediction, such as using SHAP value, LIME, or Integrated Gradients.

## Conclusion

In this article, we have learned the distinction between data science vs machine learning vs deep learning. Data science is a field of study to handle and process data such that, in the end, we can gain insight from it. Meanwhile, machine learning and deep learning are two fields of study that play an important part in one of many data science life cycles. Machine learning is a subset of AI, whilst deep learning is a subset of machine learning.

Machine learning and deep learning differ in terms of their architecture, human intervention, data volume, training time, and the computational power that they need. Thus, it's important for us to assess the data and hardware that we have in the beginning before deciding whether to use machine learning and deep learning.

If you want to learn about in-depth machine learning algorithm concepts, you can easily check out other StrataScratch articles, such as:

- If you want to learn about different kinds of machine learning algorithms → Machine Learning Algorithms

- If you want to learn more about the differences between supervised vs unsupervised paradigms in machine learning → Supervised vs Unsupervised Learning
- If you want to learn more about unsupervised paradigm in machine learning → Unsupervised Machine Learning Algorithm