# Introduction to machine learning?

Machine learning is the subset of artificial intelligence which enable the machine to learn automatic from data, improve performance from past experience and make predictions. ML has ability to learn as well as make decisions. It is the type of learning by which a machine has a capability of learning without being explicitly programmed.

For example:

Google maps

Online shopping

Facebook recommended friends list

Google spam filter

What is data and information? How to differentiate both?

input

machine

model

ML

Expected

output

don’t do the things manually all the times in machine learning

## Comparison chart

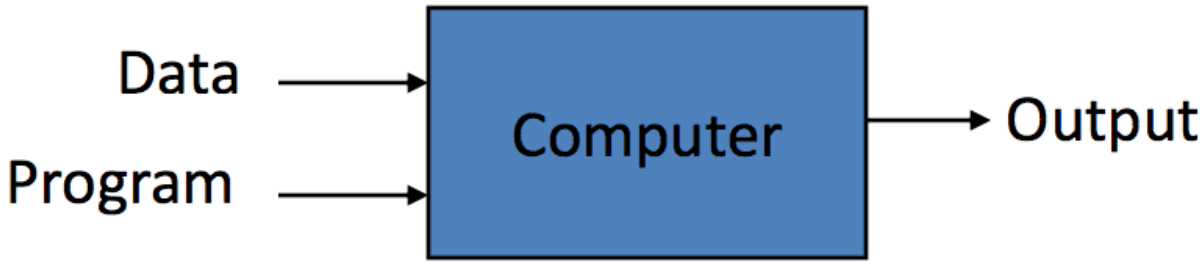
| | Data | Information |
|---|---|---|
| **Meaning** | Data is raw, unorganized facts that need to be processed. Data can be something simple and seemingly random and useless until it is organized. | When data is processed, organized, structured or presented in a given context so as to make it useful, it is called information. |
| **Example** | Each student's test score is one piece of data. | The average score of a class or of the entire school is information that can be derived from the given data. |
| **Etymology** | "Data" comes from a singular Latin word, datum, which originally meant "something given." Its early usage dates back to the 1600s. Over time "data" has become the plural of datum. | "Information" is an older word that dates back to the 1300s and has Old French and Middle English origins. It has always referred to "the act of informing, " usually in regard to education, instruction, or other knowledge communication. |

# Types of data:

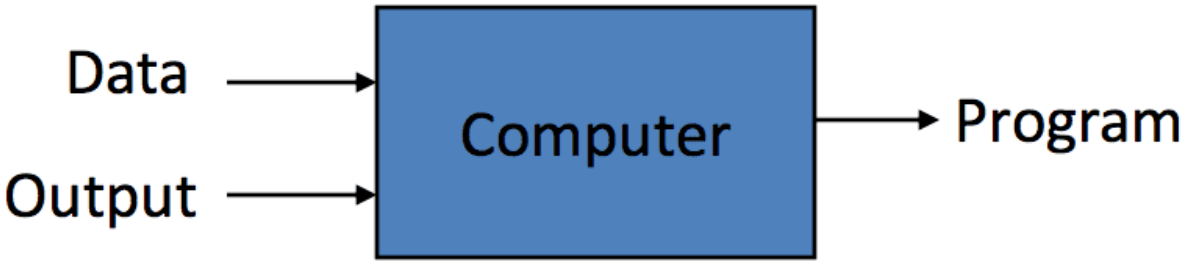
| | Quantitative Data | Qualitative Data |
|---|---|---|
| **Definition** | Quantitative data are the result of **counting** or **measuring** attributes of a population. | Qualitative data are the result of **categorizing** or **describing** attributes of a population. |
| **Data that you will see** | Quantitative data are always numbers. | Qualitative data are generally described by words or letters. |
| **Examples** | Amount of money you have<br>Height<br>Weight<br>Number of people living in your town<br>Number of students who take statistics | Hair color<br>Blood type<br>Ethnic group<br>The car a person drives<br>The street a person lives on |

Difference between traditional programming and machine learning?

## Traditional Programming




## Machine Learning




| Traditional programming | Machine learning |
|---|---|
| is a manual technique in which a programmer instructs the computer by giving inputs (in the form of code) to produce the output. Anyone with a decent programming background can formulate the code using rules manually. | ML is an automatic process we feed in the input and expected output to the machine which there by generate a model |
| Program written once cannot be changed and useless they are updated manually by someone | Model-------→better model<br>When Input are increased |
| Explicitly need of human to generate output | Program generated by algorithm without being explicitly programmed by any human being |

# Application of Machine learning?

Applications of Machine Learning include:

- **Web Search Engine:** One of the reasons why search engines like google, bing etc work so well is because the system has learnt how to rank pages through a complex learning algorithm.

- **Photo tagging Applications:** Be it facebook or any other photo tagging application, the ability to tag friends makes it even more happening. It is all possible because of a face recognition algorithm that runs behind the application.
- **Spam Detector:** Our mail agent like Gmail or Hotmail does a lot of hard work for us in classifying the mails and moving the spam mails to spam folder. This is again achieved by a spam classifier running in the back end of mail application.
- **Augmentation**:Machine learning, which assists humans with their day-to-day tasks, personally or commercially without having complete control of the output. Such machine learning is used in different ways such as Virtual Assistant, Data analysis, software solutions. The primary user is to reduce errors due to human bias.
- **Automation**:Machine learning, which works entirely autonomously in any field without the need for any human intervention. For example, robots performing the essential process steps in manufacturing plants.
- **Finance Industry:**Machine learning is growing in popularity in the finance industry. Banks are mainly using ML to find patterns inside the data but also to prevent fraud.
- **Government organization:**The government makes use of ML to manage public safety and utilities. Take the example of China with the massive face recognition. The government uses Artificial intelligence to prevent jaywalker.
- **Healthcare industry:** Healthcare was one of the first industry to use machine learning with image detection.
- **Marketing:**Broad use of AI is done in marketing thanks to abundant access to data. Before the age of mass data, researchers develop advanced mathematical tools like Bayesian analysis to estimate the value of a customer. With the boom of data, marketing department relies on AI to optimize the customer relationship and marketing campaign.

Goals of machine learning

# Goals of Machine Learning:

The goal of ML, in simpler words is to understand the nature of (human & other forms of) learning, and to build learning capability in computers. To be more specific there are three aspects of goals of ML.

1. To make the computers smarter , more intelligent.
2. To develop computational models of human learning process and perform computer simulations.
3. To explore new learning methods and develop general learning algorithms independent of applications.

These are goals that we attain using Machine Learning.

Types of machine learning/methods of machine learning

1.supervised learning: it is sub-category of ML and AI which is defined by use of label datasets to train algorithm that is use to classify the data or predict outcomes accurately

"Supervised learning algorithms are used when the output is classified or labeled. These algorithms learn from the past data that is inputted, called training data, runs its analysis and uses this analysis to predict future events of any new data within the known classifications. The accurate prediction of test data requires large data to have a sufficient understanding of the patterns. The algorithm can be trained further by comparing the training outputs to actual ones and using the errors for modification of the algorithms."

Working: it use a training set to teach the training model to yield the desire output this training dataset includes input and output which allow the model to learn overtime the algorithm measures its accuracy through the loss function adjusting until the error has been sufficiently minimized it can be separated into two types :

1.classification 2. Regression

Classification: uses the algorithm to accurately test data into specific categories. It recorgnizes specific entities within the dataset and attempts to draw some conclusion on how those entities should label or defined

Regression(dependent and independent variable): it is used to understand the relationship between dependent variable and independent variable. It is commonly used to make projections such for sales revenue for a given business

input

machine model

ML

Expected

output

2.Unsupervised learning: model itself find the patterns and insides from given data.
The goal of unsupervised learning is that to find the underline structure of dataset , group that data according to similarities and represent that data in a compressed format.

Interpretation algorithm processing circle

triangl

square

3.semi-supervised learning: it is the type of learning coming under between supervised and unsupervised where the unlabelled data amount is large as compared to labeled data to train model. The goal is to learn a function that can accurately predict the output variable based on the input variable, similar to supervised learning, however unlike supervised learning algo is trained on a dataset that contains both label and unlabel data

Example: student is under supervision of instructors at home and college further if that student is self analysing the same concept without any help from instructors It come under unsupervised under semi-supervised student has to care itself ,after analyzing the same concept under guidance of an Instructor of college

An image archive can contain only some of its data labeled, e.g. Dog, cat, mouse, and a large chunk of images remain unlabelled.

4.reinforcement learning: it aims to maximize the rewards by their hit and trail methods/ actions. It is a feedback based ML technique in which an agent learn to behave in an environment by performing the actions and seeing thre result of actions for each good action, the agent gets positive feedback and for each bad action the agent gets feedback or penalty

The agent learn automatically using feedbacks without any labelled data unlike supervised since there is no label data, so the agent is bound to learn by experience only

The agent interact with the environment and explores it by itself the primary goal of an agent is to improve the performance by getting the maximum +ve rewards