

Machine Learning

Machine Learning is the field of study that gives computers the capability to learn without being explicitly programmed. ML is one of the most exciting technologies that one would have ever come across. As it is evident from the name, it gives the computer that makes it more similar to humans: ***The ability to learn.*** Machine learning is actively being used today, perhaps in many more places than one would expect.

Getting started with Machine Learning

Last Updated: 11-05-2020

This article discusses the categories of machine learning problems, and terminologies used in the field of machine learning.

Types of machine learning problems

There are various ways to classify machine learning problems. Here, we discuss the most obvious ones.

1. On basis of the nature of the learning “signal” or “feedback” available to a learning system

- **Supervised learning:** The computer is presented with example inputs and their desired outputs, given by a “teacher”, and the goal is to learn a general rule that maps inputs to outputs. The training process continues until the model achieves the desired level of accuracy on the training data. Some real-life examples are:
 - **Image Classification:** You train with images/labels. Then in the future you give a new image expecting that the computer will recognize the new object.
 - **Market Prediction/Regression:** You train the computer with historical market data and ask the computer to predict the new price in the future.
- **Unsupervised learning:** No labels are given to the learning algorithm, leaving it on its own to find structure in its input. It is used for clustering population in different groups. Unsupervised learning can be a goal in itself (discovering hidden patterns in data).
 - **Clustering:** You ask the computer to separate similar data into clusters, this is essential in research and science.
 - **High Dimension Visualization:** Use the computer to help us visualize high dimension data.
 - **Generative Models:** After a model captures the probability distribution of your input data, it will be able to generate more data. This can be very useful to make your classifier more robust.

A simple diagram which clears the concept of supervised and unsupervised learning is shown below:

As you can see clearly, the data in supervised learning is labelled, where as data in unsupervised learning is unlabelled.

- **Semi-supervised learning:** Problems where you have a large amount of input data and only some of the data is labeled, are called semi-supervised learning problems. These problems sit in between both supervised and

unsupervised learning. For example, a photo archive where only some of the images are labeled, (e.g. dog, cat, person) and the majority are unlabeled.

- **Reinforcement learning:** A computer program interacts with a dynamic environment in which it must perform a certain goal (such as driving a vehicle or playing a game against an opponent). The program is provided feedback in terms of rewards and punishments as it navigates its problem space.