What Is Machine Learning?

Machine learning as a field grew out of the need to get computers to solve problems that are difficult to program as a sequence of instructions. Take, for example, that we want a computer to perform the task of recognizing faces in an image. One will realize that it is incredibly complicated, if not impossible to develop a precise instruction set that will satisfactorily perform this task. However, by drawing from the observation that humans improve on performing complex functions from past experiences, we can then attempt to develop algorithms and methods that enable the computer to establish a system for solving complex tasks based off prior experiences without being explicitly programmed. The set of methods and algorithms for discovering patterns in data is what is known as machine learning.

Two classical definitions of machine learning are that of Arthur Samuel in 1956 who described machine learning as "the ability for computers to learn without being explicitly programmed" and Tom Mitchell in 1997 who defined machine learning as "the process of teaching a computer to perform a particular task by improving its measure of performance with experience."

Machine learning is an interdisciplinary field of study that brings together techniques from the fields of computer science, statistics, mathematics, and the cognitive sciences which include biology, psychology, and linguistics, to mention just a few. While the idea of learning from data has been around the academic community for several decades, its entry into the mainstream technology industry began in the early 2000s. This growth coincided with the rise of humongous data as a result of the web explosion as people started sharing data over the Internet.

The Role of Data

Data is at the core of machine learning. It is central to the current evolution and further advancement of this field. Just as it is for humans, it is the same way for machines. Learning is not possible without data.

Humans learn how to perform tasks by collecting information from the Environment. This information is the data the brain uses to construct patterns and gain an understanding of the Environment. For a human being, data is captured through the sense organs. For example, the eyes capture visual data, the ears capture auditory data, the skin receives tactile data, while the nose and tongue detect olfactory and taste data, respectively.

As with humans, this same process of learning from data is replicated with machines. Let's take, for example, the task of identifying spam emails. In this example, the computer is provided email examples as data. It then uses an algorithm to learn to distinguish spam emails from regular emails.

The Cost of Data

Data is expensive to collect, and high-quality data is even more costly to capture due to the associated costs in storing and cleaning the data. Over the years, the paucity of data had limited the performance of machine learning methods. However, in the early 1990s, the Internet was born, and by the dawn of the century, it became a super highway for data distribution. As a result, large and diverse data became readily available for the research and development of machine learning products across various domains.

In this chapter, we covered the definition and history of machine learning and the importance of data. Next, we will take it further by discussing the principles of machine learning in Chapter 14.