Machine learning (**ML**) is a field of study in artificial intelligence concerned with the development and study of statistical algorithms that can learn from data and generalize to unseen data, and thus perform tasks without explicit instructions.[1] Advances in the field of deep learning have allowed neural networks to surpass many previous approaches in performance.[2]

ML finds application in many fields, including natural language processing, computer vision, speech recognition, email filtering, agriculture, and medicine.[3][4] The application of ML to business problems is known as predictive analytics.

Statistics and mathematical optimization (mathematical programming) methods comprise the foundations of machine learning. Data mining is a related field of study, focusing on exploratory data analysis (EDA) via unsupervised learning.[6][7]

From a theoretical viewpoint, probably approximately correct (PAC) learning provides a framework for describing machine learning.

History

[edit]

See also: Timeline of machine learning

The term *machine learning* was coined in 1959 by Arthur Samuel, an IBM employee and pioneer in the field of computer gaming and artificial intelligence.[8][9] The synonym *self-teaching computers* was also used in this time period.[10][11]

Although the earliest machine learning model was introduced in the 1950s when Arthur Samuel invented a program that calculated the winning chance in checkers for each side, the history of machine learning roots back to decades of human desire and effort to study human cognitive processes.[12] In 1949, Canadian psychologist Donald Hebb published the book *The Organization of Behavior*, in which he introduced a theoretical neural structure formed by certain interactions among nerve cells.[13] Hebb's model of neurons interacting with one another set a groundwork for how AIs and machine learning algorithms work under nodes, or artificial neurons used by computers to communicate data.[12] Other researchers who have studied human cognitive systems contributed to the modern machine learning technologies as well, including logician Walter Pitts and Warren McCulloch, who proposed the early mathematical models of neural networks to come up with algorithms that mirror human thought processes.[12]

By the early 1960s, an experimental "learning machine" with punched tape memory, called Cybertron, had been developed by Raytheon Company to analyze sonar signals, electrocardiograms, and speech patterns using

rudimentary reinforcement learning. It was repetitively "trained" by a human operator/teacher to recognize patterns and equipped with a "goof" button to cause it to reevaluate incorrect decisions.[14] A representative book on research into machine learning during the 1960s was Nilsson's book on Learning Machines, dealing mostly with machine learning for pattern classification.[15] Interest related to pattern recognition continued into the 1970s, as described by Duda and Hart in 1973.[16] In 1981 a report was given on using teaching strategies so that an artificial neural network learns to recognize 40 characters (26 letters, 10 digits, and 4 special symbols) from a computer terminal.[17]

Tom M. Mitchell provided a widely quoted, more formal definition of the algorithms studied in the machine learning field: "A computer program is said to learn from experience *E* with respect to some class of tasks *T* and performance measure *P* if its performance at tasks in *T*, as measured by *P*, improves with experience *E*."[18] This definition of the tasks in which machine learning is concerned offers a fundamentally operational definition rather than defining the field in cognitive terms. This follows Alan Turing's proposal in his paper "Computing Machinery and Intelligence", in which the question "Can machines think?" is replaced with the question "Can machines do what we (as thinking entities) can do?".[19]

Modern-day machine learning has two objectives. One is to classify data based on models which have been developed; the other purpose is to make predictions for future outcomes based on these models. A hypothetical algorithm specific to classifying data may use computer vision of moles coupled with supervised learning in order to train it to classify the cancerous moles. A machine learning algorithm for stock trading may inform the trader of future potential predictions.[20]