**Machine learning** (**ML**) is the study of computer algorithms that improve automatically through experience.[[1]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-1) It is seen as a subset of [artificial intelligence](https://en.wikipedia.org/wiki/Artificial_intelligence). Machine learning algorithms build a [mathematical model](https://en.wikipedia.org/wiki/Mathematical_model) based on sample data, known as "[training data](https://en.wikipedia.org/wiki/Training_data)", in order to make predictions or decisions without being explicitly programmed to do so.[[2]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-2)[[3]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-bishop2006-3):2 Machine learning algorithms are used in a wide variety of applications, such as [email filtering](https://en.wikipedia.org/wiki/Email_filtering) and [computer vision](https://en.wikipedia.org/wiki/Computer_vision), where it is difficult or infeasible to develop conventional algorithms to perform the needed tasks.

Machine learning is closely related to [computational statistics](https://en.wikipedia.org/wiki/Computational_statistics), which focuses on making predictions using computers. The study of [mathematical optimization](https://en.wikipedia.org/wiki/Mathematical_optimization) delivers methods, theory and application domains to the field of machine learning. [Data mining](https://en.wikipedia.org/wiki/Data_mining) is a related field of study, focusing on [exploratory data analysis](https://en.wikipedia.org/wiki/Exploratory_data_analysis) through [unsupervised learning](https://en.wikipedia.org/wiki/Unsupervised_learning).[[4]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-4)[[5]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-5) In its application across business problems, machine learning is also referred to as [predictive analytics](https://en.wikipedia.org/wiki/Predictive_analytics).

Machine learning involves computers discovering how they can perform tasks without being explicitly programmed to do so. For simple tasks assigned to computers, it is possible to program algorithms telling the machine how to execute all steps required to solve the problem at hand; on the computer's part, no learning is needed. For more advanced tasks, it can be challenging for a human to manually create the needed algorithms. In practice, it can turn out to be more effective to help the machine develop its own algorithm, rather than have human programmers specify every needed step.[[6]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-Alpaydin2020-6)[[7]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-elements-7)

The discipline of machine learning employs various approaches to help computers learn to accomplish tasks where no fully satisfactory algorithm is available. In cases where vast numbers of potential answers exist, one approach is to label some of the correct answers as valid. This can then be used as training data for the computer to improve the algorithm(s) it uses to determine correct answers. For example, to train a system for the task of digital character recognition, the [MNIST](https://en.wikipedia.org/wiki/MNIST_database) dataset has often been used. [[6]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-Alpaydin2020-6)[[7]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-elements-7)

**Machine learning approaches**[[edit](https://en.wikipedia.org/w/index.php?title=Machine_learning&action=edit&section=2)]

Early classifications for machine learning approaches sometimes divided them into three broad categories, depending on the nature of the "signal" or "feedback" available to the learning system. These were:  
[Supervised learning](https://en.wikipedia.org/wiki/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](https://en.wikipedia.org/wiki/Map_(mathematics)) inputs to outputs.  
[Unsupervised learning](https://en.wikipedia.org/wiki/Unsupervised_learning): No labels are given to the learning algorithm, leaving it on its own to find structure in its input. Unsupervised learning can be a goal in itself (discovering hidden patterns in data) or a means towards an end ([feature learning](https://en.wikipedia.org/wiki/Feature_learning)).  
[Reinforcement learning](https://en.wikipedia.org/wiki/Reinforcement_learning): A computer program interacts with a dynamic environment in which it must perform a certain goal (such as [driving a vehicle](https://en.wikipedia.org/wiki/Autonomous_car) or playing a game against an opponent) As it navigates its problem space, the program is provided feedback that's analogous to rewards, which it tries to maximise. [[3]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-bishop2006-3)

Other approaches or processes have since developed that don't fit neatly into this three-fold categorisation, and sometimes more than one is used by the same machine learning system. For example [topic modeling](https://en.wikipedia.org/wiki/Topic_modeling), [dimensionality reduction](https://en.wikipedia.org/wiki/Dimensionality_reduction) or [meta learning](https://en.wikipedia.org/wiki/Meta_learning_(computer_science)). [[8]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-8) As of 2020, [deep learning](https://en.wikipedia.org/wiki/Deep_learning) had become the dominant approach for much ongoing work in the field of machine learning . [[6]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-Alpaydin2020-6)

History and relationships to other fields[[edit](https://en.wikipedia.org/w/index.php?title=Machine_learning&action=edit&section=3)]

*See also:*[*Timeline of machine learning*](https://en.wikipedia.org/wiki/Timeline_of_machine_learning)

The term *machine learning* was coined in 1959 by [Arthur Samuel](https://en.wikipedia.org/wiki/Arthur_Samuel), an American [IBMer](https://en.wikipedia.org/wiki/IBMer" \o "IBMer) and pioneer in the field of [computer gaming](https://en.wikipedia.org/wiki/Computer_gaming) and [artificial intelligence](https://en.wikipedia.org/wiki/Artificial_intelligence). [[9]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-Samuel-9)[[10]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-10) A representative book of the machine learning research during the 1960s was the Nilsson's book on Learning Machines, dealing mostly with machine learning for pattern classification.[[11]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-11) Interest related to pattern recognition continued into the 1970s, as described by Duda and Hart in 1973. [[12]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-12) In 1981 a report was given on using teaching strategies so that a neural network learns to recognize 40 characters (26 letters, 10 digits, and 4 special symbols) from a computer terminal. [[13]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-13)

[Tom M. Mitchell](https://en.wikipedia.org/wiki/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*."[[14]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-Mitchell-1997-14) This definition of the tasks in which machine learning is concerned offers a fundamentally [operational definition](https://en.wikipedia.org/wiki/Operational_definition) rather than defining the field in cognitive terms. This follows [Alan Turing](https://en.wikipedia.org/wiki/Alan_Turing)'s proposal in his paper "[Computing Machinery and Intelligence](https://en.wikipedia.org/wiki/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?".[[15]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-15)

**Relation to artificial intelligence**

As a scientific endeavor, machine learning grew out of the quest for artificial intelligence. In the early days of AI as an [academic discipline](https://en.wikipedia.org/wiki/Discipline_(academia)), some researchers were interested in having machines learn from data. They attempted to approach the problem with various symbolic methods, as well as what were then termed "[neural networks](https://en.wikipedia.org/wiki/Neural_network)"; these were mostly [perceptrons](https://en.wikipedia.org/wiki/Perceptron" \o "Perceptron) and [other models](https://en.wikipedia.org/wiki/ADALINE) that were later found to be reinventions of the [generalized linear models](https://en.wikipedia.org/wiki/Generalized_linear_model) of statistics.[[16]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-16) [Probabilistic](https://en.wikipedia.org/wiki/Probability_theory) reasoning was also employed, especially in automated [medical diagnosis](https://en.wikipedia.org/wiki/Medical_diagnosis).[[17]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-aima-17):488

However, an increasing emphasis on the [logical, knowledge-based approach](https://en.wikipedia.org/wiki/GOFAI) caused a rift between AI and machine learning. Probabilistic systems were plagued by theoretical and practical problems of data acquisition and representation.[[17]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-aima-17):488 By 1980, [expert systems](https://en.wikipedia.org/wiki/Expert_system) had come to dominate AI, and statistics was out of favor.[[18]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-changing-18) Work on symbolic/knowledge-based learning did continue within AI, leading to [inductive logic programming](https://en.wikipedia.org/wiki/Inductive_logic_programming), but the more statistical line of research was now outside the field of AI proper, in [pattern recognition](https://en.wikipedia.org/wiki/Pattern_recognition) and [information retrieval](https://en.wikipedia.org/wiki/Information_retrieval).[[17]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-aima-17):708–710; 755 Neural networks research had been abandoned by AI and [computer science](https://en.wikipedia.org/wiki/Computer_science) around the same time. This line, too, was continued outside the AI/CS field, as "[connectionism](https://en.wikipedia.org/wiki/Connectionism)", by researchers from other disciplines including [Hopfield](https://en.wikipedia.org/wiki/John_Hopfield), [Rumelhart](https://en.wikipedia.org/wiki/David_Rumelhart" \o "David Rumelhart) and [Hinton](https://en.wikipedia.org/wiki/Geoff_Hinton). Their main success came in the mid-1980s with the reinvention of [backpropagation](https://en.wikipedia.org/wiki/Backpropagation).[[17]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-aima-17):25

Machine learning, reorganized as a separate field, started to flourish in the 1990s. The field changed its goal from achieving artificial intelligence to tackling solvable problems of a practical nature. It shifted focus away from the [symbolic approaches](https://en.wikipedia.org/wiki/Symbolic_artificial_intelligence) it had inherited from AI, and toward methods and models borrowed from statistics and [probability theory](https://en.wikipedia.org/wiki/Probability_theory).[[18]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-changing-18) As of 2019, many sources continue to assert that machine learning remains a sub field of AI. Yet some practitioners, for example Dr Daniel Hulme, who both teaches AI and runs a company operating in the field, argues that machine learning and AI are separate. [[7]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-elements-7)[[19]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-19)[[6]](https://en.wikipedia.org/wiki/Machine_learning#cite_note-Alpaydin2020-6)