Artificial intelligence (**AI**) is the capability of computational systems to perform tasks typically associated with human intelligence, such as learning, reasoning, problem-solving, perception, and decision-making. It is a field of research in computer science that develops and studies methods and software that enable machines to perceive their environment and use learning and intelligence to take actions that maximize their chances of achieving defined goals.[1]

High-profile applications of AI include advanced web search engines (e.g., Google Search); recommendation systems (used by YouTube, Amazon, and Netflix); virtual assistants (e.g., Google Assistant, Siri, and Alexa); autonomous vehicles (e.g., Waymo); generative and creative tools (e.g., language models and AI art); and superhuman play and analysis in strategy games (e.g., chess and Go). However, many AI applications are not perceived as AI: "A lot of cutting edge AI has filtered into general applications, often without being called AI because once something becomes useful enough and common enough it's not labeled AI anymore."[2][3]

Various subfields of AI research are centered around particular goals and the use of particular tools. The traditional goals of AI research include learning, reasoning, knowledge representation, planning, natural language processing, perception, and support for robotics.[a] To reach these goals, AI researchers have adapted and integrated a wide range of techniques, including search and mathematical optimization, formal logic, artificial neural networks, and methods based on statistics, operations research, and economics.[b] AI also draws upon psychology, linguistics, philosophy, neuroscience, and other fields.[4] Some companies, such as OpenAI, Google DeepMind and Meta,[5] aim to create artificial general intelligence (AGI) – AI that can complete virtually any cognitive task at least as well as a human.

Artificial intelligence was founded as an academic discipline in 1956,[6] and the field went through multiple cycles of optimism throughout its history,[7][8] followed by periods of disappointment and loss of funding, known as AI winters.[9][10] Funding and interest vastly increased after 2012 when graphics processing units started being used to accelerate neural networks, and deep learning outperformed previous AI techniques.[11] This growth accelerated further after 2017 with the transformer architecture.[12] In the 2020s, an ongoing period of rapid progress in advanced generative AI became known as the AI boom. Generative AI's ability to create and modify content has led to several unintended consequences and harms. Ethical concerns have been raised about AI's long-term effects and potential existential risks, prompting discussions about regulatory policies to ensure the safety and benefits of the technology.

A **feedforward neural network** is an artificial neural network in which information flows in a single direction – inputs are multiplied by weights to obtain outputs (inputs-to-output).[2] It contrasts with a recurrent neural network, in which loops allow information from later processing stages to feed back to earlier stages.[3] Feedforward multiplication is essential for backpropagation,[4][5][6][7][8] because feedback, where the outputs feed back to the very same inputs and modify them, forms an infinite loop which is not possible to differentiate through backpropagation. This nomenclature appears to be a point of confusion between some computer scientists and scientists in other fields studying brain networks.[9]

Mathematical foundations

Activation function

The two historically common activation functions are both sigmoids, and are described by

The first is a hyperbolic tangent that ranges from -1 to 1, while the other is

the logistic function, which is similar in shape but ranges from 0 to 1. Here is the

output of the -th node (neuron) and is the weighted sum of the input connections. Alternative activation functions have been proposed, including the rectifier and softplus functions. More specialized activation functions include radial basis functions (used in radial basis networks, another class of supervised neural network models).

In recent developments of deep learning, the rectified linear unit (ReLU) is more frequently used as one of the possible ways to overcome the numerical problems related to the sigmoids.

Learning

Vishwas Narayan Nangare Patil, IPS, PPMG is an Indian Police Service officer. He serves as the Additional Director General of Police in the Anti-corruption Bureau, Maharashtra. Previously he was Joint Commissioner of Police (Law & Order), Mumbai. Formerly he was Commissioner of Police, Nashik city.[1][2][3] Patil completed his training in1997. In 2015 he was awarded the President's Police Medal (gallantry) for his role in the counterterrorist operations during the 2008 Mumbai attacks.[4][5][6][7]

Career

See also: 2008 Mumbai attacks

Nangare-Patil was deputy commissioner of police Zone-1 in South Mumbai during the attack.[8] He confronted the terrorists at Taj hotel. He shot one of them in the leg.

During this fight with the terrorists, they shot down one of Nangare's bodyguards and injured one. He had barged in there without any bulletproof vest only with a Glock and if he had not done it people would have died in the marriage hall known as crystal hall.[9][10]

Positions held

- Additional Director General of Police - Anti-Corruption Bureau
- Inspector General of Police (Law and Order) - Mumbai
- Additional Commissioner of Police - Mumbai.[11]
- Superintendent of Police - Thane Rural (2010–11).[12]
- Superintendent of Police - Latur (2002–04)
- Additional Superintendent of Police - Nanded
- Assistant Superintendent of Police – Dhule

Data science is an interdisciplinary academic field[1] that uses statistics, scientific computing, scientific methods, processing, scientific visualization, algorithms and systems to extract or extrapolate knowledge from potentially noisy, structured, or unstructured data.[2]

Data science also integrates domain knowledge from the underlying application domain (e.g., natural sciences, information technology, and medicine).[3] Data science is multifaceted and can be described as a science, a research paradigm, a research method, a discipline, a workflow, and a profession.[4]

Data science is "a concept to unify statistics, data analysis, informatics, and their related methods" to "understand and analyze actual phenomena" with data.[5] It uses techniques and theories drawn from many fields within the context of mathematics, statistics, computer science, information science, and domain knowledge.[6] However, data science is different from computer science and information science. Turing Award winner Jim Gray imagined data science as a "fourth paradigm" of science (empirical, theoretical, computational, and now data-driven) and asserted that "everything about science is changing because of the impact of information technology" and the data deluge.[7][8]

Data science is often described as a multidisciplinary field because it draws on techniques from diverse areas, such as computer science, statistics, information science, and other subject-specific disciplines. Some researchers say that the combination of the different fields is similar to how information science was decades ago (Mayernik, 2023). These similarities help us understand how data science became its own field of study. [9]

Data science is an interdisciplinary field[11] focused on extracting knowledge from typically large data sets and applying the knowledge from that data to solve problems in other application domains. The field encompasses preparing data for analysis, formulating data science problems, analyzing data, and summarizing these findings. As such, it incorporates skills from computer science, mathematics, data visualization, graphic design, communication, and business.[12]

Vasant Dhar writes that statistics emphasizes quantitative data and description. In contrast, data science deals with quantitative and qualitative data (e.g., from images, text, sensors, transactions, customer information, etc.) and emphasizes prediction and action.[13] Andrew Gelman of Columbia University has described statistics as a non-essential part of data science.[14] Stanford professor David Donoho writes that data science is not distinguished from statistics by the size of datasets or use of computing and that many graduate programs misleadingly advertise their analytics and statistics training as the essence of a data-science program. He describes data science as an applied field growing out of traditional statistics.[15]

In 1962, John Tukey described a field he called "data analysis", which resembles modern data science.[15] In 1985, in a lecture given to the Chinese Academy of Sciences in Beijing, C. F. Jeff Wu used the term "data science" for the first time as an alternative name for statistics.[16] Later, attendees at a 1992 statistics symposium at the University of Montpellier II acknowledged the emergence of a new discipline focused on data of various origins and forms, combining established concepts and principles of statistics and data analysis with computing.[17][18]

The term "data science" has been traced back to 1974, when Peter Naur proposed it as an alternative name to computer science. In his 1974 book Concise Survey of Computer Methods, Peter Naur proposed using the term ‘data science’ rather than ‘computer science’ to reflect the growing emphasis on data-driven methods[19][6] In 1996, the International Federation of Classification Societies became the first conference to specifically feature data science as a topic.[6] However, the definition was still in flux. After the 1985 lecture at the Chinese Academy of Sciences in Beijing, in 1997 C. F. Jeff Wu again suggested that statistics should be renamed data science. He reasoned that a new name would help statistics shed inaccurate stereotypes, such as being synonymous with accounting or limited to describing data.[20] In 1998, Hayashi Chikio argued for data science as a new, interdisciplinary concept, with three aspects: data design, collection, and analysis.[18]

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 generalise to unseen data, and thus perform tasks without explicit instructions.[1] Within a subdiscipline in machine learning, advances in the field of deep learning have allowed neural networks, a class of statistical algorithms, to surpass many previous machine learning approaches in performance.

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

Statistics and mathematical optimisation (mathematical programming) methods comprise the foundations of machine learning. Data mining is a related field of study, focusing on exploratory data analysis (EDA) through unsupervised learning.[3][4]

From a theoretical viewpoint, probably approximately correct learning provides a mathematical and statistical framework for describing machine learning. Most traditional machine learning and deep learning algorithms can be described as empirical risk minimisation under this framework.

History

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.[5][6] The synonym *self-teaching computers* was also used during this time period.[7][8]

The earliest machine learning program was introduced in the 1950s when Arthur Samuel invented a computer program that calculated the winning chance in checkers for each side, but the history of machine learning roots back to decades of human desire and effort to study human cognitive processes.[9] 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.[10] 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.[9] 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.[9]

By the early 1960s, an experimental "learning machine" with punched tape memory, called Cybertron, had been developed by Raytheon Company to analyse sonar signals, electrocardiograms, and speech patterns using rudimentary reinforcement learning. It was repetitively "trained" by a human operator/teacher to recognise patterns and equipped with a "goof" button to cause it to reevaluate incorrect decisions.[11] A representative book on research into machine learning during the 1960s was Nils Nilsson's book on Learning Machines, dealing mostly with machine learning for pattern classification.[12] Interest related to pattern recognition continued into the 1970s, as described by Duda and Hart in 1973.[13] In 1981, a report was given on using teaching strategies so that an artificial neural network learns to recognise 40 characters (26 letters, 10 digits, and 4 special symbols) from a computer terminal.[14]

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*."[15] 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?".[16]

Modern-day Machine Learning algorithms are broken into 3 algorithm types: Supervised Learning Algorithms, Unsupervised Learning Algorithms, and Reinforcement Learning Algorithms.[17]

- Current Supervised Learning Algorithms have objectives of classification and regression.
- Current Unsupervised Learning Algorithms have objectives of clustering, dimensionality reduction, and association rule.
- Current Reinforcement Learning Algorithms focus on decisions that must be made with respect to some previous, unknown time and are broken down to either be studies of model-based methods or model-free methods.

In 2014 Ian Goodfellow and others introduced generative adversarial networks (GANs) with realistic data synthesis.[18] By 2016 AlphaGo obtained victory against top human players using reinforcement learning techniques.[19]

In machine learning, a **neural network** or **neural net** (**NN**), also called **artificial neural network** (**ANN**), is a computational model inspired by the structure and functions of biological neural networks.[1][2]

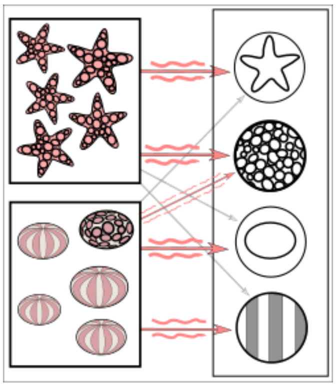
A neural network consists of connected units or nodes called *artificial neurons*, which loosely model the neurons in the brain. Artificial neuron models that mimic biological neurons more closely have also been recently investigated and shown to significantly improve performance. These are connected by *edges*, which model the synapses in the brain. Each artificial neuron receives signals from connected neurons, then processes them and sends a signal to other connected neurons. The "signal" is a real number, and the output of each neuron is computed by some non-linear function of the totality of its inputs, called the *activation function*. The strength of the signal at each connection is determined by a *weight*, which adjusts during the learning process.

Typically, neurons are aggregated into layers. Different layers may perform different transformations on their inputs. Signals travel from the first layer (the *input layer*) to the last layer (the *output layer*), possibly passing through multiple intermediate layers (*hidden layers*). A network is typically called a deep neural network if it has at least two hidden layers.[3]

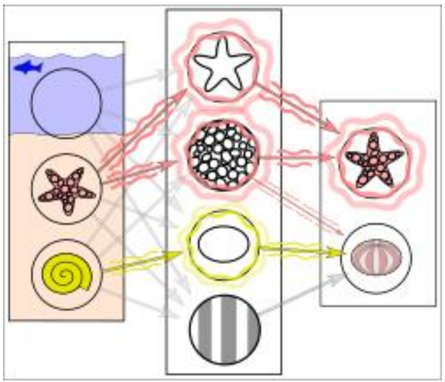
Artificial neural networks are used for various tasks, including predictive modeling, adaptive control, and solving problems in artificial intelligence. They can learn from experience, and can derive conclusions from a complex and seemingly unrelated set of information.

Training

Neural networks are typically trained through empirical risk minimization, which is based on the idea of optimizing the network's parameters to minimize the difference, or empirical risk, between the predicted output and the actual target values in a given dataset.[4] Gradient-based methods such as backpropagation are usually used to estimate the parameters of the network.[4] During the training phase, ANNs learn from labeled training data by iteratively updating their parameters to minimize a defined loss function.[5] This method allows the network to generalize to unseen data.



Simplified example of training a neural network in object detection: The network is trained by multiple images that are known to depict starfish and sea urchins, which are correlated with "nodes" that represent visual features. The starfish match with a ringed texture and a star outline, whereas most sea urchins match with a striped texture and oval shape. However, the instance of a ring textured sea urchin creates a weakly weighted association between them.



Subsequent run of the network on an input image (left):[6] The network correctly detects the starfish. However, the weakly weighted association between ringed texture and sea urchin also confers a weak signal to the latter from one of two intermediate nodes. In addition, a shell that was not included in the training gives a weak signal for the oval shape, also resulting in a weak signal for the sea urchin output. These weak signals may result in a false positive result for sea urchin.

In reality, textures and outlines would not be represented by single nodes, but rather by associated weight patterns of multiple nodes.

History

Main article: History of artificial neural networks

Early work

Today's deep neural networks are based on early work in statistics over 200 years ago. The simplest kind of feedforward neural network (FNN) is a linear network, which consists of a single layer of output nodes with linear activation functions; the inputs are fed directly to the outputs via a series of weights. The sum of the products of the weights and the inputs is calculated at each node. The mean squared errors between these calculated outputs and the given target values are minimized by creating an adjustment to the weights. This technique has been known for over two centuries as the method of least squares or linear regression. It was used as a means of finding a good rough linear fit to a set of points by Legendre (1805) and Gauss (1795) for the prediction of planetary movement.[7][8][9][10][11]