Artificial intelligence

158 languages

- Article
- Talk
- Read
- View source
- View history

Tools

-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-
-

Appendix

hide

Text

•

Small

Standard

Large

Width

•

Standard

Wide

Color (beta)

•

Automatic

Light

Dark

From Wikipedia, the free encyclopedia

"AI" redirects here. For other uses, see [AI (disambiguation)](), [Artificial intelligence (disambiguation)](), and [Intelligent agent]().

Part of a series on

Artificial intelligence

show

Major goals

show

Approaches

show

Applications

show

Philosophy

show

History

show

Glossary

- v
- t
- e

Artificial intelligence (**AI**), in its broadest sense, is intelligence exhibited by machines, particularly computer systems. 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] Such machines may be called AIs.

Some high-profile applications of AI include advanced web search engines (e.g., Google Search); recommendation systems (used by YouTube, Amazon, and Netflix); interacting via human speech (e.g., Google Assistant, Siri, and Alexa); autonomous vehicles (e.g., Waymo); generative and creative tools (e.g., ChatGPT, 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]

The various subfields of AI research are centered around particular goals and the use of particular tools. The traditional goals of AI research include reasoning, knowledge representation, planning, learning, natural language processing, perception, and support for robotics.[a] General intelligence—the ability to complete any task performable by a human on an at least equal level—is among the field's long-term goals.[4] 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.[5]

Artificial intelligence was founded as an academic discipline in 1956,[6] and the field went through multiple cycles of optimism,[7][8] followed by periods of disappointment and loss of funding, known as AI winter.[9][10] Funding and interest vastly increased after 2012 when deep learning outperformed previous AI techniques.[11] This growth accelerated further after 2017 with the transformer architecture,[12] and by the early 2020s hundreds of billions of dollars were being invested in AI (known as the "AI boom"). The widespread use of AI in the 21st century exposed several unintended consequences and harms in the present and raised concerns about its risks and long-term effects in the future, prompting discussions about regulatory policies to ensure the safety and benefits of the technology.

Goals

The general problem of simulating (or creating) intelligence has been broken into subproblems. These consist of particular traits or capabilities that researchers expect an intelligent system to display. The traits described below have received the most attention and cover the scope of AI research.[a]

Reasoning and problem-solving

Early researchers developed algorithms that imitated step-by-step reasoning that humans use when they solve puzzles or make logical deductions.[13] By the late 1980s and 1990s, methods were developed for dealing with uncertain or incomplete information, employing concepts from probability and economics.[14]

Many of these algorithms are insufficient for solving large reasoning problems because they experience a "combinatorial explosion": They become exponentially slower as the problems grow.[15] Even humans rarely use the step-by-step deduction that early AI research could model. They solve most of their problems using fast, intuitive judgments.[16] Accurate and efficient reasoning is an unsolved problem.

Knowledge representation

An ontology represents knowledge as a set of concepts within a domain and the relationships between those concepts.

Knowledge representation and knowledge engineering[17] allow AI programs to answer questions intelligently and make deductions about real-world facts. Formal knowledge representations are used in content-based indexing and retrieval,[18] scene interpretation,[19] clinical decision support,[20] knowledge discovery (mining "interesting" and actionable inferences from large databases),[21] and other areas.[22]

A knowledge base is a body of knowledge represented in a form that can be used by a program. An ontology is the set of objects, relations, concepts, and properties used by a particular domain of knowledge.[23] Knowledge bases need to represent things such as objects, properties, categories, and relations between objects;[24] situations, events, states, and time;[25] causes and effects;[26] knowledge about knowledge (what we know about what other people know);[27] default reasoning (things that humans assume are true until they are told differently and will remain true even when other facts are changing);[28] and many other aspects and domains of knowledge.

Among the most difficult problems in knowledge representation are the breadth of commonsense knowledge (the set of atomic facts that the average person knows is enormous);[29] and the sub-symbolic form of most commonsense knowledge (much of what people know is not represented as "facts" or "statements" that they could express verbally).[16] There is also the difficulty of knowledge acquisition, the problem of obtaining knowledge for AI applications.[c]

Planning and decision-making

An "agent" is anything that perceives and takes actions in the world. A rational agent has goals or preferences and takes actions to make them happen.[d][32] In automated planning, the agent has a specific goal.[33] In automated decision-making, the agent has preferences—there are some situations it would prefer to be in, and some situations it is trying to avoid. The decision-making agent assigns a number to each situation (called the "utility") that measures how much the agent prefers it. For each possible action, it can calculate the "expected utility": the utility of all possible outcomes of the action, weighted by the probability that the outcome will occur. It can then choose the action with the maximum expected utility.[34]

In classical planning, the agent knows exactly what the effect of any action will be.[35] In most real-world problems, however, the agent may not be certain about the situation they are in (it is "unknown" or "unobservable") and it may not know for certain what will happen after each possible action (it is not "deterministic"). It must choose an action by making a probabilistic guess and then reassess the situation to see if the action worked.[36]

In some problems, the agent's preferences may be uncertain, especially if there are other agents or humans involved. These can be learned (e.g., with inverse reinforcement learning), or the agent can seek information to improve its preferences.[37] Information value theory can be used to weigh the value of exploratory or experimental actions.[38] The space of possible future actions and situations is typically intractably large, so the agents must take actions and evaluate situations while being uncertain of what the outcome will be.

A Markov decision process has a transition model that describes the probability that a particular action will change the state in a particular way and a reward function that supplies the utility of each state and the cost of each action. A policy associates a decision with each possible state. The policy could be calculated (e.g., by iteration), be heuristic, or it can be learned.[39]

Game theory describes the rational behavior of multiple interacting agents and is used in AI programs that make decisions that involve other agents.[40]

Learning

Machine learning is the study of programs that can improve their performance on a given task automatically.[41] It has been a part of AI from the beginning.[e]

There are several kinds of machine learning. Unsupervised learning analyzes a stream of data and finds patterns and makes predictions without any other guidance.[44] Supervised learning requires a human to label the input data first, and comes in two main varieties: classification (where the program must learn to predict what category the input belongs in) and regression (where the program must deduce a numeric function based on numeric input).[45]

In reinforcement learning, the agent is rewarded for good responses and punished for bad ones. The agent learns to choose responses that are classified as "good".[46] Transfer learning is when the knowledge gained from one problem is applied to a new problem.[47] Deep learning is a type of machine learning that runs inputs through biologically inspired artificial neural networks for all of these types of learning.[48]

Computational learning theory can assess learners by computational complexity, by sample complexity (how much data is required), or by other notions of optimization.[49]

Natural language processing

Natural language processing (NLP)[50] allows programs to read, write and communicate in human languages such as English. Specific problems include speech recognition, speech synthesis, machine translation, information extraction, information retrieval and question answering.[51]

Early work, based on Noam Chomsky's generative grammar and semantic networks, had difficulty with word-sense disambiguation[f] unless restricted to small domains called "micro-worlds" (due to the common sense knowledge problem[29]). Margaret Masterman believed that it was meaning and not grammar that was the key to understanding languages, and that thesauri and not dictionaries should be the basis of computational language structure.

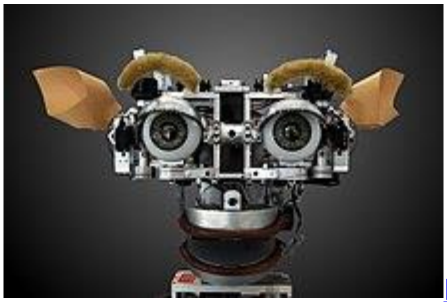
Modern deep learning techniques for NLP include word embedding (representing words, typically as vectors encoding their meaning),[52] transformers (a deep learning architecture using an attention mechanism),[53] and others.[54] In 2019, generative pre-trained transformer (or "GPT") language models began to generate coherent text,[55][56] and by 2023, these models were able to get human-level scores on the bar exam, SAT test, GRE test, and many other real-world applications.[57]

Perception

Machine perception is the ability to use input from sensors (such as cameras, microphones, wireless signals, active lidar, sonar, radar, and tactile sensors) to deduce aspects of the world. Computer vision is the ability to analyze visual input.[58]

The field includes speech recognition,[59] image classification,[60] facial recognition, object recognition,[61]object tracking,[62] and robotic perception.[63]

Social intelligence



Kismet, a robot head which was made in the 1990s; a machine that can recognize and simulate emotions[64]

Affective computing is an interdisciplinary umbrella that comprises systems that recognize, interpret, process, or simulate human feeling, emotion, and mood.[65] For example, some virtual assistants are programmed to speak conversationally or even to banter humorously; it makes them appear more sensitive to the emotional dynamics of human interaction, or to otherwise facilitate human–computer interaction.

However, this tends to give naïve users an unrealistic conception of the intelligence of existing computer agents.[66] Moderate successes related to affective computing include textual sentiment analysis and, more recently, multimodal sentiment analysis, wherein AI classifies the affects displayed by a videotaped subject.[67]

General intelligence

A machine with artificial general intelligence should be able to solve a wide variety of problems with breadth and versatility similar to human intelligence.[4]

Techniques

AI research uses a wide variety of techniques to accomplish the goals above.[b]

Search and optimization

AI can solve many problems by intelligently searching through many possible solutions.[68] There are two very different kinds of search used in AI: state space search and local search.

State space search

State space search searches through a tree of possible states to try to find a goal state.[69] For example, planning algorithms search through trees of goals and subgoals, attempting to find a path to a target goal, a process called means-ends analysis.[70]

Simple exhaustive searches[71] are rarely sufficient for most real-world problems: the search space (the number of places to search) quickly grows to astronomical numbers.

The result is a search that is too slow or never completes.[15] "Heuristics" or "rules of thumb" can help prioritize choices that are more likely to reach a goal.[72]

Adversarial search is used for game-playing programs, such as chess or Go. It searches through a tree of possible moves and counter-moves, looking for a winning position.[73]

Local search

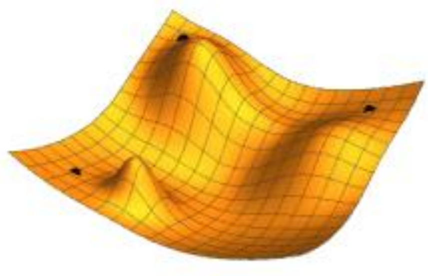


Illustration of gradient descent for 3 different starting points; two parameters (represented by the plan coordinates) are adjusted in order to minimize the loss function (the height)

Local search uses mathematical optimization to find a solution to a problem. It begins with some form of guess and refines it incrementally.[74]

Gradient descent is a type of local search that optimizes a set of numerical parameters by incrementally adjusting them to minimize a loss function. Variants of gradient descent are commonly used to train neural networks.[75]

Another type of local search is evolutionary computation, which aims to iteratively improve a set of candidate solutions by "mutating" and "recombining" them, selecting only the fittest to survive each generation.[76]

Distributed search processes can coordinate via swarm intelligence algorithms. Two popular swarm algorithms used in search are particle swarm optimization (inspired by bird flocking) and ant colony optimization (inspired by ant trails).[77]

Logic

Formal logic is used for reasoning and knowledge representation.[78] Formal logic comes in two main forms: propositional logic (which operates on statements that are true or false and uses logical connectives such as "and", "or", "not" and "implies")[79] and predicate logic (which also operates on objects, predicates and relations and uses quantifiers such as "*Every X* is a *Y*" and "There are *some* *X*s that are *Y*s").[80]

Deductive reasoning in logic is the process of proving a new statement (conclusion) from other statements that are given and assumed to be true (the premises).[81] Proofs can be structured as proof trees, in which nodes are labelled by sentences, and children nodes are connected to parent nodes by inference rules.

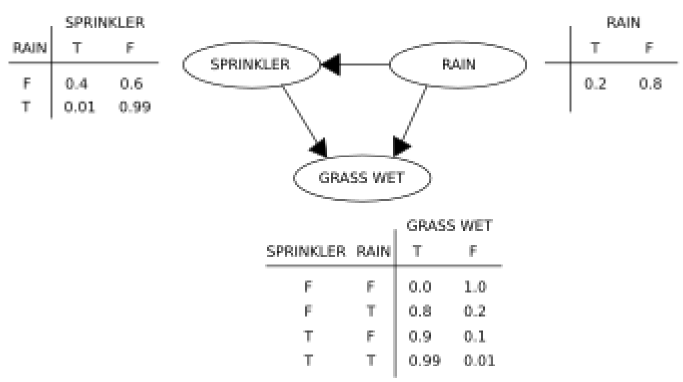
Given a problem and a set of premises, problem-solving reduces to searching for a proof tree whose root node is labelled by a solution of the problem and whose leaf nodes are labelled by premises or axioms. In the case of Horn clauses, problem-solving search can be performed by reasoning forwards from the premises or backwards from the problem.[82] In the more general case of the clausal form of first-order logic, resolution is a single, axiom-free rule of inference, in which a problem is solved by proving a contradiction from premises that include the negation of the problem to be solved.[83]

Inference in both Horn clause logic and first-order logic is undecidable, and therefore intractable. However, backward reasoning with Horn clauses, which underpins computation in the logic programming language Prolog, is Turing complete. Moreover, its efficiency is competitive with computation in other symbolic programming languages.[84]

Fuzzy logic assigns a "degree of truth" between 0 and 1. It can therefore handle propositions that are vague and partially true.[85]

Non-monotonic logics, including logic programming with negation as failure, are designed to handle default reasoning.[28] Other specialized versions of logic have been developed to describe many complex domains.

Probabilistic methods for uncertain reasoning

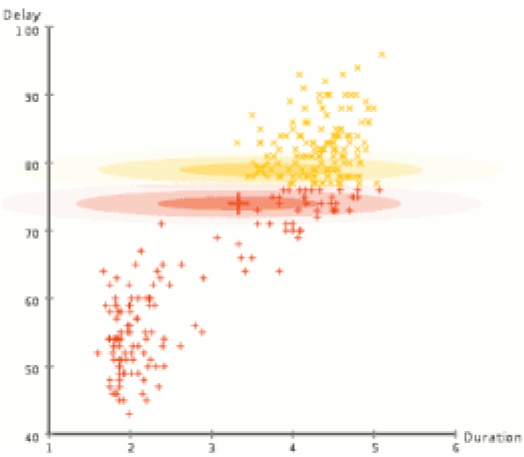


A simple Bayesian network, with the associated conditional probability tables

Many problems in AI (including in reasoning, planning, learning, perception, and robotics) require the agent to operate with incomplete or uncertain information. AI researchers have devised a number of tools to solve these problems using methods from probability theory and economics.[86] Precise mathematical tools have been developed that analyze how an agent can make choices and plan, using decision theory, decision analysis,[87] and information value theory.[88] These tools include models such as Markov decision processes,[89] dynamic decision networks,[90] game theory and mechanism design.[91]

Bayesian networks[92] are a tool that can be used for reasoning (using the Bayesian inference algorithm),[g][94] learning (using the expectation–maximization algorithm),[h][96] planning (using decision networks)[97] and perception (using dynamic Bayesian networks).[90]

Probabilistic algorithms can also be used for filtering, prediction, smoothing, and finding explanations for streams of data, thus helping perception systems analyze processes that occur over time (e.g., hidden Markov models or Kalman filters).[90]



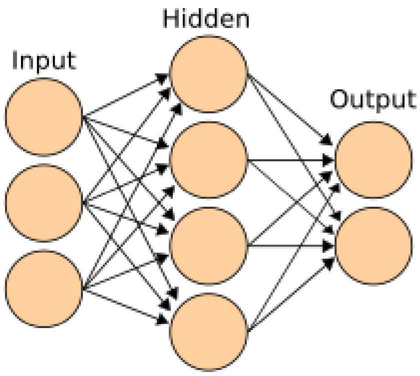
Expectation–maximization clustering of Old Faithful eruption data starts from a random guess but then successfully converges on an accurate clustering of the two physically distinct modes of eruption.

Classifiers and statistical learning methods

The simplest AI applications can be divided into two types: classifiers (e.g., "if shiny then diamond"), on one hand, and controllers (e.g., "if diamond then pick up"), on the other hand. Classifiers[98] are functions that use pattern matching to determine the closest match. They can be fine-tuned based on chosen examples using supervised learning. Each pattern (also called an "observation") is labeled with a certain predefined class. All the observations combined with their class labels are known as a data set. When a new observation is received, that observation is classified based on previous experience.[45]

There are many kinds of classifiers in use.[99] The decision tree is the simplest and most widely used symbolic machine learning algorithm.[100] K-nearest neighbor algorithm was the most widely used analogical AI until the mid-1990s, and Kernel methods such as the support vector machine (SVM) displaced k-nearest neighbor in the 1990s.[101] The naive Bayes classifier is reportedly the "most widely used learner"[102] at Google, due in part to its scalability.[103] Neural networks are also used as classifiers.[104]

Artificial neural networks



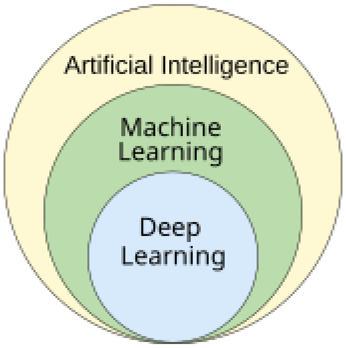
A neural network is an interconnected group of nodes, akin to the vast network of neurons in the human brain.

An artificial neural network is based on a collection of nodes also known as artificial neurons, which loosely model the neurons in a biological brain. It is trained to recognise patterns; once trained, it can recognise those patterns in fresh data. There is an input, at least one hidden layer of nodes and an output. Each node applies a function and once the weight crosses its specified threshold, the data is transmitted to the next layer. A network is typically called a deep neural network if it has at least 2 hidden layers.[104]

Learning algorithms for neural networks use local search to choose the weights that will get the right output for each input during training. The most common training technique is the backpropagation algorithm.[105] Neural networks learn to model complex relationships between inputs and outputs and find patterns in data. In theory, a neural network can learn any function.[106]

In feedforward neural networks the signal passes in only one direction.[107] Recurrent neural networks feed the output signal back into the input, which allows short-term memories of previous input events. Long short term memory is the most successful network architecture for recurrent networks.[108] Perceptrons[109] use only a single layer of neurons; deep learning[110] uses multiple layers. Convolutional neural networks strengthen the connection between neurons that are "close" to each other—this is especially important in image processing, where a local set of neurons must identify an "edge" before the network can identify an object.[111]

Deep learning



Deep learning[110] uses several layers of neurons between the network's inputs and outputs. The multiple layers can progressively extract higher-level features from the raw input. For example, in image processing, lower layers may identify edges, while higher layers may identify the concepts relevant to a human such as digits, letters, or faces.[112]

Deep learning has profoundly improved the performance of programs in many important subfields of artificial intelligence, including computer vision, speech recognition, natural language processing, image classification,[113] and others. The reason that deep learning performs so well in so many applications is not known as of 2023.[114] The sudden success of deep learning in 2012–2015 did not occur because of some new discovery or theoretical breakthrough (deep neural networks and backpropagation had been described by many people, as far back as the 1950s)[i] but because of two factors: the incredible increase in computer power (including the hundred-fold increase in speed by switching to GPUs) and the availability of vast amounts of training data, especially the giant curated datasets used for benchmark testing, such as ImageNet.[j]

GPT

Generative pre-trained transformers (GPT) are large language models (LLMs) that generate text based on the semantic relationships between words in sentences. Text-based GPT models are pretrained on a large corpus of text that can be from the Internet. The pretraining consists of predicting the next token (a token being usually a word, subword, or punctuation). Throughout this pretraining, GPT models accumulate knowledge about the world and can then generate human-like text by repeatedly predicting the next token. Typically, a subsequent training phase makes the model more truthful, useful, and harmless, usually with a technique called reinforcement learning from human feedback (RLHF). Current GPT models are prone to generating falsehoods called "hallucinations", although this can be reduced with RLHF and quality data. They are used in chatbots, which allow people to ask a question or request a task in simple text.[122][123]

Current models and services include Gemini (formerly Bard), ChatGPT, Grok, Claude, Copilot, and LLaMA.[124] Multimodal GPT models can process different types of data (modalities) such as images, videos, sound, and text.[125]

Specialized hardware and software

Main articles: Programming languages for artificial intelligence and Hardware for artificial intelligence

In the late 2010s, graphics processing units (GPUs) that were increasingly designed with AI-specific enhancements and used with specialized TensorFlow software had replaced previously used central processing unit (CPUs) as the dominant means for large-scale (commercial and academic) machine learning models' training.[126] Specialized programming languages such as Prolog were used in early AI research,[127] but general-purpose programming languages like Python have become predominant.[128]

Applications

Main article: Applications of artificial intelligence

AI and machine learning technology is used in most of the essential applications of the 2020s, including: search engines (such as Google Search), targeting online advertisements, recommendation systems (offered by Netflix, YouTube or Amazon), driving internet traffic, targeted advertising (AdSense, Facebook), virtual assistants (such as Siri or Alexa), autonomous vehicles (including drones, ADAS and self-driving cars), automatic language translation (Microsoft Translator, Google Translate), facial recognition (Apple's Face ID or Microsoft's DeepFace and Google's FaceNet) and image labeling (used by Facebook, Apple's iPhoto and TikTok). The deployment of AI may be overseen by a Chief automation officer (CAO).

Health and medicine

Main article: Artificial intelligence in healthcare

The application of AI in medicine and medical research has the potential to increase patient care and quality of life.[129] Through the lens of the Hippocratic Oath, medical professionals are ethically compelled to use AI, if applications can more accurately diagnose and treat patients.

For medical research, AI is an important tool for processing and integrating big data. This is particularly important for organoid and tissue engineering development which use microscopy imaging as a key technique in fabrication.[130] It has been suggested that AI can overcome discrepancies in funding allocated to different fields of research.[130] New AI tools can deepen the understanding of biomedically relevant pathways. For example, AlphaFold 2 (2021) demonstrated the ability to approximate, in hours rather than months, the 3D structure of a protein.[131] In 2023, it was reported that AI-guided drug discovery helped find a class of antibiotics capable of killing two different types of drug-resistant bacteria.[132] In 2024, researchers used machine learning to accelerate the search for Parkinson's disease drug treatments. Their aim was to identify compounds that block the clumping, or aggregation, of alpha-synuclein (the protein that characterises Parkinson's disease). They were able to speed up the initial screening process ten-fold and reduce the cost by a thousand-fold.[133][134]

Games

Main article: Game artificial intelligence

Game playing programs have been used since the 1950s to demonstrate and test AI's most advanced techniques.[135] Deep Blue became the first computer chess-playing system to beat a reigning world chess champion, Garry Kasparov, on 11 May 1997.[136] In 2011, in a *Jeopardy!* quiz show exhibition match, IBM's question answering system, Watson, defeated the two greatest *Jeopardy!* champions, Brad Rutter and Ken Jennings, by a significant margin.[137] In March 2016, AlphaGo won 4 out of 5 games of Go in a match with Go champion Lee Sedol, becoming the first computer Go-playing system to beat a professional Go player without handicaps. Then, in 2017, it defeated Ke Jie, who was the best Go player in the world.[138] Other programs handle imperfect-information games, such as the poker-playing program Pluribus.[139] DeepMind developed increasingly generalistic reinforcement learning models, such as with MuZero, which could be trained to play chess, Go, or Atari games.[140] In 2019, DeepMind's AlphaStar achieved grandmaster level in StarCraft II, a particularly challenging real-time strategy game that involves incomplete knowledge of what happens on the map.[141] In 2021, an AI agent competed in a PlayStation Gran Turismo competition, winning against four of the world's best Gran Turismo drivers using deep reinforcement learning.[142] In 2024, Google DeepMind introduced SIMA, a type of AI capable of autonomously playing nine previously unseen open-world video games by observing screen output, as well as executing short, specific tasks in response to natural language instructions.[143]

Mathematics

In mathematics, special forms of formal step-by-step reasoning are used. In contrast, LLMs such as *GPT-4* *Turbo*, *Gemini Ultra*, *Claude Opus*, *LLaMa-2* or *Mistral Large* are working with probabilistic models, which can produce wrong answers in the form of hallucinations. Therefore, they need not only a large database of mathematical problems to learn from but also methods such as supervised fine-tuning or trained classifiers with human-annotated data to improve answers for new problems and learn from corrections.[144] A 2024 study showed that the performance of some language models for reasoning capabilities in solving math problems not included in their training data was low, even for problems with only minor deviations from trained data.[145]

Alternatively, dedicated models for mathematic problem solving with higher precision for the outcome including proof of theorems have been developed such as *Alpha Tensor*, *Alpha Geometry* and *Alpha Proof* all from Google DeepMind,[146] *Llemma* from eleuther[147] or *Julius*.[148]

When natural language is used to describe mathematical problems, converters transform such prompts into a formal language such as Lean to define mathematic tasks.

Some models have been developed to solve challenging problems and reach good results in benchmark tests, others to serve as educational tools in mathematics.[149]

Finance

Finance is one of the fastest growing sectors where applied AI tools are being deployed: from retail online banking to investment advice and insurance, where automated "robot advisers" have been in use for some years. [150]

World Pensions experts like Nicolas Firzli insist it may be too early to see the emergence of highly innovative AI-informed financial products and services: "the deployment of AI tools will simply further automatise things: destroying tens of thousands of jobs in banking, financial planning, and pension advice in the process, but I'm not sure it will unleash a new wave of [e.g., sophisticated] pension innovation."[151]

Military

Main article: Military artificial intelligence

Various countries are deploying AI military applications.[152] The main applications enhance command and control, communications, sensors, integration and interoperability.[153] Research is targeting intelligence collection and analysis, logistics, cyber operations, information operations, and semiautonomous and autonomous vehicles.[152] AI technologies enable coordination of sensors and effectors, threat detection and identification, marking of enemy positions, target acquisition, coordination and deconfliction of distributed Joint Fires between networked combat vehicles involving manned and unmanned teams.[153] AI was incorporated into military operations in Iraq and Syria.[152]

In November 2023, US Vice President Kamala Harris disclosed a declaration signed by 31 nations to set guardrails for the military use of AI. The commitments include using legal reviews to ensure the compliance of military AI with international laws, and being cautious and transparent in the development of this technology.[154]

Generative AI

Main article: Generative artificial intelligence


Vincent van Gogh in watercolour created by generative AI software

In the early 2020s, generative AI gained widespread prominence. GenAI is AI capable of generating text, images, videos, or other data using generative models,[155][156] often in response to prompts.[157][158]

In March 2023, 58% of U.S. adults had heard about ChatGPT and 14% had tried it.[159] The increasing realism and ease-of-use of AI-based text-to-image generators such as Midjourney, DALL-E, and Stable Diffusion sparked a trend of viral AI-generated photos. Widespread attention was gained by a fake photo of Pope Francis wearing a white puffer coat, the fictional arrest of Donald Trump, and a hoax of an attack on the Pentagon, as well as the usage in professional creative arts.[160][161]

Agents

Artificial intelligent (AI) agents are software entities designed to perceive their environment, make decisions, and take actions autonomously to achieve specific goals. These agents can interact with users, their environment, or other agents. AI agents are used in various applications, including virtual assistants, chatbots, autonomous vehicles, game-playing systems, and industrial robotics. AI agents operate within the constraints of their programming, available computational resources, and hardware limitations. This means they are restricted to performing tasks within their defined scope and have finite memory and processing capabilities. In real-world applications, AI agents often face time constraints for decision-making and action execution. Many AI agents incorporate learning algorithms, enabling them to improve their performance over time through experience or training. Using machine learning, AI agents can adapt to new situations and optimise their behaviour for their designated tasks.[162][163][164]

Other industry-specific tasks

There are also thousands of successful AI applications used to solve specific problems for specific industries or institutions. In a 2017 survey, one in five companies reported having incorporated "AI" in some offerings or processes.[165] A few examples are energy storage, medical diagnosis, military logistics, applications that predict the result of judicial decisions, foreign policy, or supply chain management.

AI applications for evacuation and disaster management are growing. AI has been used to investigate if and how people evacuated in large scale and small scale evacuations using historical data from GPS, videos or social media. Further, AI can provide real time information on the real time evacuation conditions.[166][167][168]

In agriculture, AI has helped farmers identify areas that need irrigation, fertilization, pesticide treatments or increasing yield. Agronomists use AI to conduct research and development. AI has been used to predict the ripening time for crops such as tomatoes, monitor soil moisture, operate agricultural robots, conduct predictive analytics, classify livestock pig call emotions, automate greenhouses, detect diseases and pests, and save water.

Artificial intelligence is used in astronomy to analyze increasing amounts of available data and applications, mainly for "classification, regression, clustering, forecasting, generation, discovery, and the development of new scientific insights" for example for discovering exoplanets, forecasting solar activity, and distinguishing between signals and instrumental effects in gravitational wave astronomy. It could also be used for activities in space such as space exploration, including analysis of data from space missions, real-time science decisions of spacecraft, space debris avoidance, and more autonomous operation.

Ethics

Main article: Ethics of artificial intelligence

AI has potential benefits and potential risks. AI may be able to advance science and find solutions for serious problems: Demis Hassabis of Deep Mind hopes to "solve intelligence, and then use that to solve everything else".[169] However, as the use of AI has become widespread, several unintended consequences and risks have been identified.[170] Inproduction systems can sometimes not factor ethics and bias into their AI training processes, especially when the AI algorithms are inherently unexplainable in deep learning.[171]

Risks and harm

Privacy and copyright

Further information: Information privacy and Artificial intelligence and copyright

Machine learning algorithms require large amounts of data. The techniques used to acquire this data have raised concerns about privacy, surveillance and copyright.

AI-powered devices and services, such as virtual assistants and IoT products, continuously collect personal information, raising concerns about intrusive data gathering and unauthorized access by third parties. The loss of privacy is further exacerbated by AI's ability to process and combine vast amounts of data, potentially leading to a surveillance society where individual activities are constantly monitored and analyzed without adequate safeguards or transparency.

Sensitive user data collected may include online activity records, geolocation data, video or audio.[172] For example, in order to build speech recognition algorithms, Amazon has recorded millions of private conversations and allowed temporary workers to listen to and transcribe some of them.[173] Opinions about this widespread surveillance range from those who see it as a necessary evil to those for whom it is clearly unethical and a violation of the right to privacy.[174]

AI developers argue that this is the only way to deliver valuable applications. and have developed several techniques that attempt to preserve privacy while still obtaining the data, such as data aggregation, de-identification and differential privacy.[175] Since 2016, some privacy experts, such as Cynthia Dwork, have begun to view privacy in terms of fairness. Brian Christian wrote that experts have pivoted "from the question of 'what they know' to the question of 'what they're doing with it'."[176]

Generative AI is often trained on unlicensed copyrighted works, including in domains such as images or computer code; the output is then used under the rationale of "fair use". Experts disagree about how well and under what circumstances this rationale will hold up in courts of law; relevant factors may include "the purpose and character of the use of the copyrighted work" and "the effect upon the potential market for the copyrighted work".[177][178] Website owners who do not wish to have their content scraped can indicate it in a "robots.txt" file.[179] In 2023, leading authors (including John Grisham and Jonathan Franzen) sued AI companies for using their work to train generative AI.[180][181] Another discussed approach is to envision a separate *sui generis* system of protection for creations generated by AI to ensure fair attribution and compensation for human authors.[182]

Dominance by tech giants

The commercial AI scene is dominated by Big Tech companies such as Alphabet Inc., Amazon, Apple Inc., Meta Platforms, and Microsoft.[183][184][185] Some of these players already own the vast majority of existing cloud infrastructure and computing power from data centers, allowing them to entrench further in the marketplace.[186][187]

Substantial power needs and other environmental impacts

See also: Environmental impacts of artificial intelligence

In January 2024, the International Energy Agency (IEA) released *Electricity 2024, Analysis and Forecast to 2026*, forecasting electric power use.[188] This is the first IEA report to make projections for data centers and power consumption for artificial intelligence and cryptocurrency. The report states that power demand for these uses might double by 2026, with additional electric power usage equal to electricity used by the whole Japanese nation.[189]

Prodigious power consumption by AI is responsible for the growth of fossil fuels use, and might delay closings of obsolete, carbon-emitting coal energy facilities. There is a feverish rise in the construction of data centers throughout the US, making large technology firms (e.g., Microsoft, Meta, Google, Amazon) into voracious consumers of electric power. Projected electric consumption is so immense that there is concern that it will be fulfilled no matter the source. A ChatGPT search involves the use of 10 times the electrical energy as a Google search. The large firms are in haste to find power sources – from nuclear energy to geothermal to fusion. The tech firms argue that – in the long view – AI will be eventually kinder to the environment, but they need the energy now. AI makes the power grid more efficient and "intelligent", will assist in the growth of nuclear power, and track overall carbon emissions, according to technology firms.[190]

A 2024 Goldman Sachs Research Paper, *AI Data Centers and the Coming US Power Demand Surge*, found "US power demand (is) likely to experience growth not seen in a generation…." and forecasts that, by 2030, US data centers will consume 8% of US power, as opposed to 3% in 2022, presaging growth for the electrical power generation industry by a variety of means.[191]Data centers' need for more and more electrical power is such that they might max out the electrical grid. The Big Tech companies counter that AI can be used to maximize the utilization of the grid by all.[192]

In 2024, the *Wall Street Journal* reported that big AI companies have begun negotiations with the US nuclear power providers to provide electricity to the data centers. In March 2024 Amazon purchased a Pennsylvania nuclear-powered data center for $650 Million (US).[193]