Conference Proceeding Issue Published in International Journal of Trend in Research and Development (IJTRD), ISSN: 2394-9333, www.ijtrd.com

Artificial Intelligence

Dr. M. Senthil Murugan,
Assistant Professor in CS, Immaculate College for Women, Viriyur, India

Artificial intelligence (AI), the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings. The term is frequently applied to the project of developing systems endowed with the intellectual processes characteristic of humans, such as the ability to reason, discover meaning, generalize, or learn from past experience. Since their development in the 1940s, digital computers have been programmed to carry out very complex tasks—such as discovering proofs for mathematical theorems or playing chess—with great proficiency. Despite continuing advances in computer processing speed and memory capacity, there are as yet no programs that can match full human flexibility over wider domains or in tasks requiring much everyday knowledge. On the other hand, some programs have attained the performance levels of human experts and professionals in executing certain specific tasks, so that artificial intelligence in this limited sense is found in applications as diverse as medical diagnosis, computer search engines, voice or handwriting recognition, and chatbots.

Learning:

Learning here are a number of different forms of learning as applied to artificial intelligence. The simplest is learning by trial and error. For example, a simple computer program for solving mate-in-one chess problems might try moves at random until mate is found. The program might then store the solution with the position so that, the next time the computer encountered the same position, it would recall the solution. This simple memorizing of individual items and procedures—known as rote learning—is relatively easy to implement on a computer. More challenging is the problem of implementing what is called generalization. Generalization involves applying past experience to analogous new situations. For example, a program that learns the past tense of regular English verbs by rote will not be able to produce the past tense of a word such as jump unless the program was previously presented with jumped, whereas a program that is able to generalize can learn the "add -ed" rule for regular verbs ending in a consonant and so form the past tense of jump on the basis of experience with similar verbs.

Language:

A language is a system of signs having meaning by convention. In this sense, language need not be confined to the spoken word. Traffic signs, for example, form a mini-language, it being a matter of convention that ⚠ means "hazard ahead" in some countries. It is distinctive of languages that linguistic units possess meaning by convention, and linguistic meaning is very different from what is called natural meaning, exemplified in statements such as "Those clouds mean rain" and "The fall in pressure means the valve is malfunctioning."

An important characteristic of full-fledged human languages—in contrast to birdcalls and traffic signs—is their productivity. A productive language can formulate an unlimited variety of sentences.

Large language models like ChatGPT can respond fluently in a human language to questions and statements. Although such models do not actually understand language as humans do but merely select words that are more probable than others, they have reached the point where their command of a language is indistinguishable from that of a normal human. What, then, is involved in genuine understanding, if even a computer that uses language like a native human speaker is not acknowledged to understand? There is no universally agreed upon answer to this difficult question.

AI technology:

In the early 21st century faster processing power and larger datasets ("big data") brought artificial intelligence out of computer science departments and into the wider world. Moore's law, the observation that computing power doubled roughly every 18 months, continued to hold true. The stock responses of the early chatbot Eliza fit comfortably within 50 kilobytes; the language model at the heart of ChatGPT was trained on 45 terabytes of text.

Machine learning:

The ability of neural networks to take on added layers and thus work on more-complex problems increased in 2006 with the invention of the "greedy layer-wise pretraining" technique, in which it was found that it was easier to train each layer of a neural network individually than to train the whole network from input to output. This improvement in neural network training led to a type of machine learning called "deep learning," in which neural networks have four or more layers, including the initial input and the final output. Moreover, such networks are able to learn unsupervised—that is, to discover features in data without initial prompting.

Among the achievements of deep learning have been advances in image classification in which specialized neural networks called convolution neural networks (CNNs) are trained on features found in a set of images of many different types of objects. The CNN is then able to take an input image, compare it with features in images in its training set, and classify the image as being of, for example, a cat or an apple. One such network, Prel-net by Kaiming He and collaborators at Microsoft Research, has classified images even better than a human did.

The achievement of Deep Blue in beating world chess champion Garry Kasparov was surpassed by DeepMind's AlphaGo, which mastered go, a much more complicated game than chess. AlphaGo's neural networks learned to play go from human players and

by playing itself. It defeated top go player Lee Sedol 4–1 in 2016. AlphaGo was in turn surpassed by AlphaGo Zero, which, starting from only the rules of go, was eventually able to defeat AlphaGo 100–0. A more general neural network, Alpha Zero, was able to use the same techniques to quickly master chess and shogi.

Machine learning has found applications in many fields beyond gaming and image classification. The pharmaceutical company Pfizer used the technique to quickly search millions of possible compounds in developing the COVID-19 treatment Paxlovid. Google uses machine learning to filter out spam from the inbox of Gmail users. Banks and credit card companies use historical data to train models to detect fraudulent transactions.

The ability of neural networks to take on added layers and thus work on more-complex problems increased in 2006 with the invention of the "greedy layer-wise pretraining" technique, in which it was found that it was easier to train each layer of a neural network individually than to train the whole network from input to output. This improvement in neural network training led to a type of machine learning called "deep learning," in which neural networks have four or more layers, including the initial input and the final output. Moreover, such networks are able to learn unsupervised—that is, to discover features in data without initial prompting.

Among the achievements of deep learning have been advances in image classification in which specialized neural networks called convolution neural networks (CNNs) are trained on features found in a set of images of many different types of objects. The CNN is then able to take an input image, compare it with features in images in its training set, and classify the image as being of, for example, a cat or an apple. One such network, Prel-net by Kaiming He and collaborators at Microsoft Research, has classified images even better than a human did.

Garry Kasparov and Deep Blue

Garry Kasparov and Deep Blue World chess champion Garry Kasparov playing against Deep Blue, the chess-playing computer built by IBM. In 1996 Kasparov won the first match 4−2, but in 1997 he lost to Deep Blue 3 ½−2 ½.

The achievement of Deep Blue in beating world chess champion Garry Kasparov was surpassed by DeepMind's AlphaGo, which mastered go, a much more complicated game than chess. AlphaGo's neural networks learned to play go from human players and by playing itself. It defeated top go player Lee Sedol 4–1 in 2016. AlphaGo was in turn surpassed by AlphaGo Zero, which, starting from only the rules of go, was eventually able to defeat AlphaGo 100–0. A more general neural network, Alpha Zero, was able to use the same techniques to quickly master chess and shogi.

Machine learning has found applications in many fields beyond gaming and image classification. The pharmaceutical company Pfizer used the technique to quickly search millions of possible compounds in developing the COVID-19 treatment Paxlovid. Google uses machine learning to filter out spam from the inbox of Gmail users. Banks and credit card companies use historical data to train models to detect fraudulent transactions.

TikTok account featuring a deepfake of Keanu Reeves

TikTok account featuring a deepfake of Keanu Reeves The "Unreal Keanu Reeves" TikTok account posts include relationship humor and dances.

Deepfakes are AI-generated media produced using two different deep-learning algorithms: one that creates the best possible replica of a real image or video and another that detects whether the replica is fake and, if it is, reports on the differences between it and the original. The first algorithm produces a synthetic image and receives feedback on it from the second algorithm; it then adjusts it to make it appear more real. The process is repeated until the second algorithm does not detect any false imagery. Deepfake media portray images that do not exist in reality or events that have never occurred. Widely circulated deepfakes include an image of Pope Francis in a puffer jacket, an image of former U.S. president Donald Trump in a scuffle with police officers, and a video of Facebook CEO Mark Zuckerberg giving a speech about his company's nefarious power. Such events did not occur in real life.

AI poses certain risks in terms of ethical and socioeconomic consequences. As more tasks become automated, especially in such industries as marketing and health care, many workers are poised to lose their jobs. Although AI may create some new jobs, these may require more technical skills than the jobs AI has replaced.

Moreover, AI has certain biases that are difficult to overcome without proper training. For example, U.S. police departments have begun using predictive policing algorithms to indicate where crimes are most likely to occur. However, such systems are based partly on arrest rates, which are already disproportionately high in Black communities. This may lead to over-policing in such areas, which further affects these algorithms. As humans are inherently biased, algorithms are bound to reflect human biases.

Privacy is another aspect of AI that concerns experts. As AI often involves collecting and processing large amounts of data, there is the risk that this data will be accessed by the wrong people or organizations. With generative AI, it is even possible to manipulate images and create fake profiles. AI can also be used to survey populations and track individuals in public spaces. Experts have implored policymakers to develop practices and policies that maximize the benefits of AI while minimizing the potential risks. In January 2024 singer Taylor Swift was the target of sexually explicit non-consensual deepfakes that were widely circulated on social media. Many individuals had already faced this type of online abuse (made possible by AI), but Swift's status brought the issue to the forefront of public policy.

Google data center

Google data center Aerial view of a Google data center complex in Eemshaven, Netherlands.

LLMs are located at data centers that require large amounts of electricity. In 2020 Microsoft pledged that it would be carbon neutral by 2030. In 2024 it announced that in the previous fiscal year its carbon emissions had increased by almost 30 percent, mostly from the building materials and hardware required in building more data centers. A ChatGPT query requires about 10 times more electricity than a Google Search. Goldman Sachs has estimated that data centers will use about 8 percent of U.S. electricity in 2030.

As of 2024 there are few laws regulating AI. Existing laws such as the European Union's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) do govern AI models but only insofar as they use personal information. The most wide-reaching regulation is the EU's AI Act, which passed in March 2024. Under the AI Act, models that perform social scoring of citizens' behavior and characteristics and that attempt to manipulate users' behavior are banned. AI models that deal with "high-risk" subjects, such as law enforcement and infrastructure, must be registered in an EU database.

AI has also led to issues concerning copyright law and policy. In 2023 the U.S. government Copyright Office began an initiative to investigate the issue of AI using copyrighted works to generate content. That year almost 15 new cases of copyright-related suits were filed against companies involved in creating generative AI programs. One prominent company, Stability AI, came under fire for using unlicensed images to generate new content. Getty Images, which filed the suit, added its own AI feature to its platform, partially in response to the host of services that offer "stolen imagery." There are also questions of whether work created by AI is worthy of a copyright label. Currently, AI-made content cannot be copyrighted, but there are arguments for and against copyrighting it.

Although many AI companies claim that their content does not require human labor, in many cases, such "groundbreaking" technology is reliant on exploited workers from developing countries. For example, a Time magazine investigation found that OpenAI had used Kenyan workers (who had been paid less than $2 an hour) to sort through text snippets in order to help remove toxic and sexually explicit language from ChatGPT. The project was canceled in February 2022 because of how traumatic the task was for workers. Although Amazon had marketed its Amazon Go cashier-less stores as being fully automated (e.g., its AI could detect the items in a customer's basket), it was revealed that the "Just Walk Out" technology was actually powered by outsourced labor from India, where more than a thousand workers operated as "remote cashiers," leading to the joke that, in this case, AI stood for Actually Indians.

NATURAL LANGUAGE PROCESSING:

natural language processing (NLP), in computer science, the use of operations, systems, and technologies that allow computers to process and respond to written and spoken language in a way that mirrors human ability. To do this, natural language processing (NLP) models must use computational linguistics, statistics, machine learning, and deep-learning models.

Early NLP models were hand-coded and rule-based but did not account for exceptions and nuances in language. For example, sarcasm, idioms, and metaphors are nuances that humans learn through experience. In order for a machine to be successful at parsing language, it must first be programmed to differentiate such concepts. These early developments were followed by statistical NLP, which uses probability to assign the likelihood of certain meanings to different parts of text. Modern NLP systems use deep-learning models and techniques that help them "learn" as they process information. However, such systems cannot be said to "understand" what they are parsing; rather, they use complex programming and probability to generate humanlike responses.

Prominent examples of modern NLP are language models that use artificial intelligence (AI) and statistics to predict the final form of a sentence on the basis of existing portions. One popular language model was GPT-3, from the American AI research laboratory OpenAI, released in June 2020. Among the first large language models, GPT-3 could solve high-school level math problems and create computer programs. GPT-3 was the foundation of ChatGPT software, released in November 2022 by OpenAI. ChatGPT almost immediately disturbed academics, journalists, and others because of concerns that it was impossible to distinguish human writing from ChatGPT-generated writing.

Other examples of machines using NLP are voice-operated GPS systems, customer service chatbots, and language translation programs. In addition, businesses use NLP to enhance understanding of and service to consumers by auto-completing search queries and monitoring social media.

Related Topics: artificial intelligence ChatGPT large language model chatbot semantic network

What do you think?

Is Artificial Intelligence Good for Society?

Explore the Procyon debate

NLP presents certain challenges, especially as machine-learning algorithms and the like often express biases implicit in the content on which models are trained. For example, when asked to describe a doctor, NLP models may be more likely to respond with "He is a doctor" than with "She is a doctor," demonstrating inherent gender bias. Bias in NLP can have real-world consequences. For instance, in 2015 Amazon's NLP program for screening résumés to aid the selection of job candidates was found to discriminate against women, as women were underrepresented in the original training set collected from employees. Moreover, with probability-based NLP models, such as ChatGPT, "hallucinations" may occur, in which a model avoids communicating to the user that it does not know something by responding instead with probable but factually inaccurate text based on the user's prompts.