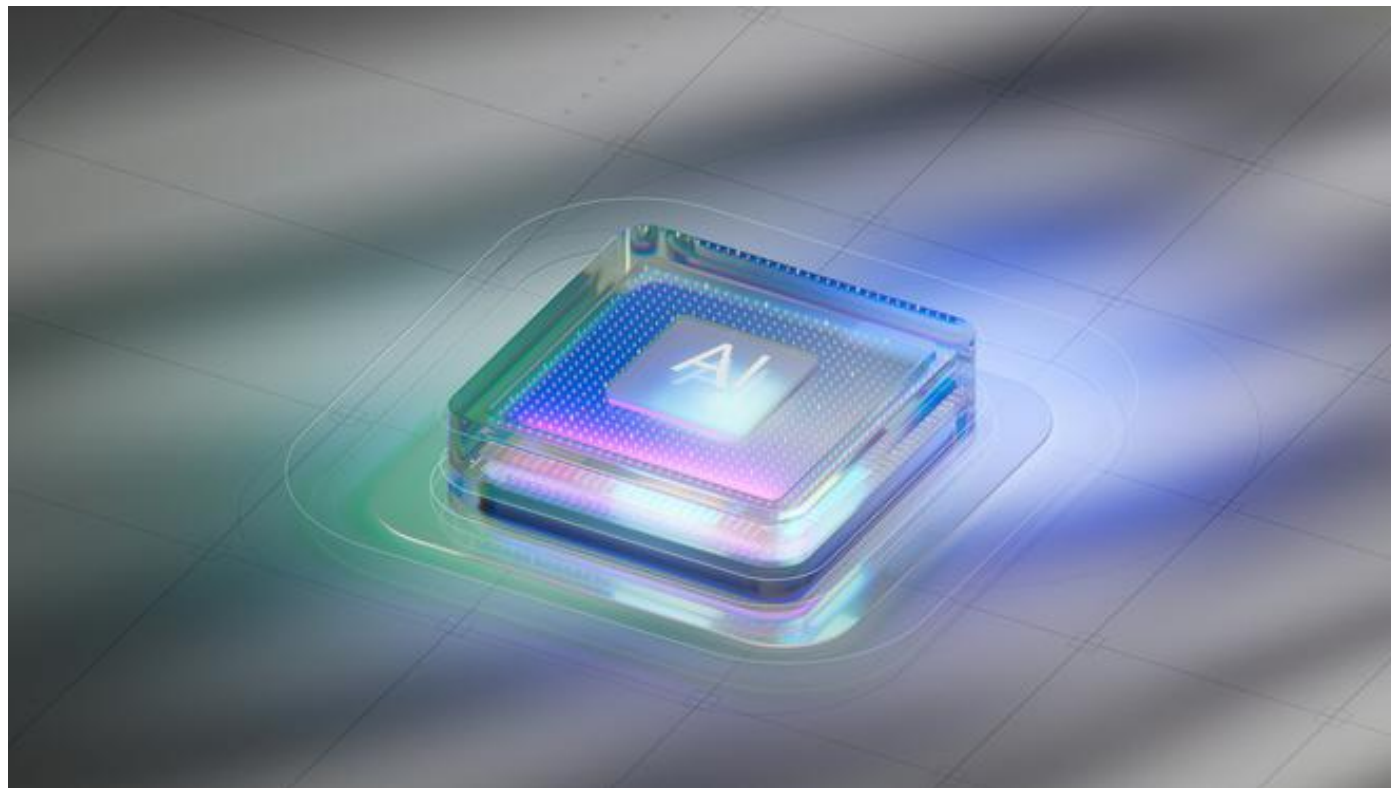


The science of AI and the AI of science

After having evolved from a postwar vision to today's generative and agentic models, powered by neural networks, GPUs, and vast amounts of data, AI is beginning to pose tough questions on ethics, reasoning, and energy

Updated – May 01, 2025 01:16 pm IST

DEBDUTTA PAUL



Representative illustration. The faster and more precise AI models become, the more precise application they will find. | Photo Credit: Milad Fakurian/Unsplash

The fundamental idea behind artificial intelligence (AI) stems from the British mathematician Alan Turing, who in the 1950s defined the **idea of intelligence** in a machine. During World War II, when attempting to break encryption code the Nazis

used to transmit secret messages, he wondered whether machines would be able to find patterns in large amounts of data that humans couldn't. He speculated machines could learn from experience instead of being taught to work from the first principles. Computer scientist John McCarthy coined the term "artificial intelligence" in a summer workshop in 1955 that many leading computer scientists attended.

While the idea enjoyed enough attention in the subsequent decade — enough for the first chatbot, ELIZA, to be created in 1966 — funding dipped in the 1970s before rebounding after. By 1997, IBM's Deep Blue defeated chess champion Gary Kasparov, around the same time researchers were starting to build sophisticated artificial neural networks that supercharged machine learning. Soon, the idea emerged that these neural networks, which are computers that process information like networks of neurons in animal brains, could solve most scientific problems.

From ANNs to GPUs

Artificial neural networks, or ANNs, could learn to solve each problem by digesting large amounts of data, mapping the relationships between various problem statements and their respective solutions, and finally recreating these relationships for unsolved problems. This paradigm is called unsupervised learning. In supervised learning, humans label the data with tags the machine picks up on. For example, humans can create a database of images of cats and dogs and label them accordingly. The ANN that 'trains' with the database then 'learns' what 'cat' and 'dog' stand for.

In another type of learning, called reinforcement learning, humans go through the results of what a machine has learnt and provide feedback for the machine to improve.

Every ANN consists of nodes, small computers that accept input signals and provide an output. The nodes are divided into groups called layers. The layers are connected to each other like neurons in the brain: each node in one layer connects to a single other node in the next layer. It's a sandwich: the layers are the two slices of bread and in between are all the connections between the nodes.

Not all connections are equal, some are more important than others. These relationships are adjusted by giving each connection a weight. The greater the weight,

the more important the signal passing along that connection. By adjusting the weights, arrangement of nodes, and number of layers, then, the ANN can be adjusted to learn and process data in different ways.

Machine-learning models that use such ANN architecture are capable of processing in a few hours databases that might take humans several months — as long as they have the requisite computing. This power comes from graphics processing units (GPUs), an advanced version of the central processing units (CPUs) that power home computers. GPUs are specialised to solve multiple mathematical problems simultaneously, speeding up the ANN's learning process.

Machine learning v. artificial intelligence

Recognising patterns in any form of data is in the domain of machine-learning (ML). It has applications in many fields. For example, ML models installed on self-driving cars are trained to check the condition of the cars' various components and if possible perform repairs. In the clinical realm, ML models can learn to find patterns in disorders that lead to new forms of treatment or read test reports to identify the risk of specific diseases.

AI, on the other hand, is broader. It is based on more recent advances in ML that mimic human intelligence in problem-solving — like completing an unfinished sentence the way Arthur Clarke might or create an image in the style of Vincent van Gogh. Such AI models are being rapidly adapted for various applications.

For example, researchers can build ML algorithms that digest the average behaviour of a user's financial accounts, like transaction frequency, spending limits, login times, and device use, according to Jia Zhai, senior associate professor in the Department of Finance at the Xi'an Jiaotong-Liverpool University in Suzhou, China. "If a fraudster gains valid credentials but logs in via an unrecognised device at 3 am and initiates rapid microtransactions, clustering algorithms detect this as an outlier compared to the user's historical behaviour," she said.

Then, more specialised networks called convolutional neural networks look for complex patterns in transactions; recurrent neural networks identify deviations from average

spending behaviour; and graph neural networks examine the connections between accounts, merchants, and IP addresses to uncover hidden money-laundering networks, said Shimeng Shi, assistant professor in the same department and institute.

The capabilities of AI surged from around 2017, when researchers began using ML to process large amounts of data simultaneously using multiple GPUs. A major advance that resulted was the large language model. As private sector enterprises figured out how to apply this and other models to solve different but specific problems, manufacturers and vendors rushed to meet the demand for the underlying hardware. This in turn led to more computing power and faster chips entering the market. Another equally important and roughly simultaneous development was the availability of large datasets on which the new batch of AI/ML models could be trained.

All together, the next major advance took shape: generative AI, where an AI model didn't just analyse what was in front of it but also put existing information together in new ways, e.g. creating an image based on a user's text instructions. Perhaps the most well-known products that make such capabilities available to users are ChatGPT and DALL-E, both made by US-based company OpenAI. Shimeng Shi also said financial firms have been trying to "help their clients to generate real-time trade ideas" using "AI-empowered tools" that are out of view.

The technology isn't a silver bullet, of course. Completely autonomous AI agents are not yet a reality because of their tendency to "hallucinate", i.e. invent information that doesn't exist in the real world. This happens when an AI model is confronted with a particular kind of data it hasn't been trained on, causing it to mix them up with data it is familiar with.

Precision, speed, structure

"Your model is as good as your data," Aditi Shanmugam, a research associate of analytics and databases at Bengaluru-based startup Ultrahuman who uses AI models to draw inferences in health data, said. "For any good model, you need lots of data with good diversity," Debnath Pal, professor in the Department of Computational and Data Sciences at the Indian Institute of Science (IISc), Bengaluru, added.

The next thing a good model needs after training data is hardware resources. “Each data centre — especially a large one with AI GPUs — can consume as much power as a whole nuclear power plant will produce,” Akash Pradhan, a member of the technical staff at chip-maker AMD, said. The machines also generate a large amount of heat of their own, which means they need to be cooled, which requires even more power.

If the machines are performing a particularly complex task, the data they are manipulating need to be stored in high-speed hard drives.

Given all these requirements, most of the better AI research today — especially of the cutting edge variety — is led by big corporations with deep pockets.

But it may not always be this way. Many computer scientists are working on techniques to lower the power and hardware requirements for specific models without also compromising the latter’s problem-solving ability.

For example, Rakesh Sengupta, director of the Centre for Creative Cognition at S.R. University in Warangal is working on a technique called pruning. In a **recent paper**, he proposed a method in which some connections in a neural network are cut while the most important ones are preserved, then retraining the model to work with the smaller set of connections. He expressed belief that we can “trim” existing models without sacrificing their reliability. “I feel customising small language models for specialised tasks in healthcare or robotics will be most” improved, he added.

The faster and more precise AI models become, the more precise application they will find — “whether it’s neural prosthetics or brain-computer interfaces or some [other] technologies that can interface seamlessly with the brain,” Sengupta said.

Most AI researchers use the most accessible model and data to achieve specific goals. In their absence, researchers draw up datasets from first principles and mix them with available ones to create more complete, yet also more reliable, datasets.

For example, Pal said, material science researchers integrate experimental data of the properties of materials with synthetic data of the presence of other materials to create

synthetic datasets that are complete and contain more information for the models to search for. “After doing all these experiments, you may be able to figure out that, ‘oh, if I dope with this material, then I would get that property’. Such experiments are being done and then it is kind of reducing the time to realise those compositions,” Pal said.

But defining the problems and arriving at solutions is not always straightforward, and often depends on factors that require researchers to dig deep into the specific peculiarities in the data and the models.

For example, Adway Mitra, an assistant professor in the Centre of Excellence in Artificial Intelligence at IIT-Kharagpur, expressed belief that there is considerable scope to use AI models to improve weather and seasonal predictions, especially of Indian monsoons. This is what he does. Often, weather data exists as a combination of textual, visual, and numerical data “We first condense the space of all weather patterns to a small number (about 10) of ‘typical’ patterns, and our claim is that every day’s weather pattern is an approximate or noisy version of any one of these ‘typical’ patterns,” Mitra explained. Generative AI models train on these datasets and create new data from them that are easier to analyse and represent as mathematical structures.

But real-world weather data is often noisy and difficult to interpret, and weather is a complex system with lots of parameters across various locations and times. “The key technical challenge is the availability of weather data,” Mitra said.

Weather data has structures that an ML model must be able to work with. Mitra’s research focuses on what kind of algorithms or models scientists can use to best utilise that structure. Thus, researchers like Mitra are turning the idea of AI back to where it started from: while machines are good at understanding patterns, at the end of the day, the patterns must be supported by physics because weather patterns are created by physical processes. The question researchers **are thus asking** is: “How can we constrain machine learning so that it provides us values which are consistent with the different laws of physics?” This exercise, Mitra said, will bring down the number of computations the AI models will need to perform to make accurate weather predictions and thus demand less power and data storage infrastructure.

Towards AI agents

Sandeep Juneja, a professor of computer science and director of Safexpress Centre for Data, Learning and Decision Sciences at Ashoka University, said corporations like Google have **large data-driven AI models** that are already doing this at scale — but that they may be running out of data to train with. On the other hand, he added, academicians in India and even worldwide don't have the computational capacity to develop such large models to develop nuanced weather predictions. He said models like DeepSeek provide hope as they have been able to use “clever” tricks to use small amounts of data to train the models efficiently.

But Chiranjib Bhattacharyya, a professor in the Department of Computer Science and Automation at IISc, said that even DeepSeek's model is large compared to what academics can presently access.

Lixian Qian, an associate dean for research and professor in the Department of Intelligent Operations and Marketing at Xi'an Jiaotong-Liverpool University, works on autonomous vehicles that use AI algorithms to model their complex environment, predict the movement of objects on the road, and decide how the vehicle moves to avoid accidents. While there has been significant integration of AI into autonomous vehicles, he said practical challenges remain — and AI has the ability to address them. “AI algorithms can increase the number of tests on autonomous driving systems in diverse driving environments, so that the potential problems could be uncovered and diagnosed in advance.”

In a sense, then, we are slowly transitioning from a world of generative AI to agentic AI. AI agents are more powerful than the present versions of AI which still specify on particular tasks. They integrate the power of different functionalities into an ecosystem that can be empowered to make particular decisions.

For example, AI assistants may one day be able to parse data about a person's life, including their hobbies, expenses, health conditions, work, and life priorities, and help them with tasks like booking appointments or filling out forms. However, how much of such a technology will be accessible and usable to people at large will depend on data

privacy protections and technological literacy. Bhattacharya said social scientists and law scholars will play an important role in shaping how such systems fit into our lives.

Sohini Majumdar, a software engineering manager at Salesforce, agreed the time for agentic AI was near. Many business platforms are increasingly using agentic AI instead of simple chatbots to integrate their business and increase their impact. However, she added, fundamental challenges remain in using generative AI models too. The principal challenge is understanding why an AI model outputs one specific business decision rather than another — especially if the output deviates from a human understanding of the business. So she and her colleagues use yet other AI models to validate the decisions suggested by generative AI. Their aim is to understand what a model is doing and how to tweak various inputs so that the model does what she wants it to. In this way, her team will be able to make automated decisions and trust them as well.

According to Bhattacharyya, the fundamental problem boils down to AI models currently lacking the ability to reason. Pal agreed: “What is the path that [a model] follows? Is it following the same path that as a human I would want it to follow to do this inference? That we don’t know.” Mathematicians, computer scientists, and physicists are currently trying to untangle this Gordian knot.

Pradhan of AMD said these challenges are fundamental: despite neural networks being based on the human brain, the way the machines learn and the way the human brain functions are different. A fundamental difference is how the computational blocks in an artificial intelligence model — the GPUs — are different sites from where the parameters of the model are stored. In the brain, these two are stored at the same location. Second, chemical reactions run the brain whereas electricity runs digital machines. The challenges, Pradhan said, can be mitigated in neuromorphic computing, where the hardware more closely mimics how neural networks in our brain operate.

“Instead of you writing code to emulate a neural network, your hardware is the neural network,” he said. Functional neuromorphic computers of the future are expected to require less power and have the ability to update its model automatically when it encounters new data — just like our brain. But there are multiple hardware and software challenges to be surmounted before it can be realised, Pradhan said.

Sengupta is sceptical of how much AI will truly mimic us. While each generation of humans has been more comfortable with the increasing presence of smarter gadgets and software, and the tools have changed us too, there might be a natural barrier to how much AI might affect us. But it has also made us think deeply about our technologies. Just like how we constantly grapple with understanding our own existence, we might have to do the same when we invite AI into every aspect of our lives.

Debdutta Paul is a freelance science journalist.

Published – May 01, 2025 12:00 pm IST