AI'S OPPENHEIMER

Physics Nobel for Geoffrey Hinton's work on AI acknowledges tech's growing relevance. But it's also a reason to guard against it, as the laureate has made clear

Atanu Biswas

A physics Nobel Prize rarely garners widespread public attention. This makes the 2024 prize awarded to John Hopfield of Princeton University and Geoffrey Hinton of University of Toronto "for foundational discoveries and inventions that enable machine learning with artificial neural networks" an exception. It's a tribute to their scientific contribution as well as a recognition of society's growing reliance on and integration with AI.

What's just as significant, however, both scientists have also spoken out about the dangers it poses to humankind.

Hopfield was recognised for creating "an associative memory that can store and reconstruct images and other types of patterns in data", while Hinton created a technique that can "independently discover properties in data", which is a crucial component of the large artificial neural networks (ANNs) currently in use.

Inspired by the functioning of the brain, the two constructed ANNs that store and retrieve memories and learn from inputs. They did this by employing basic fundamental concepts and techniques from physics. Their groundbreaking work, which began in the 1980s, showed how computer programmes based on neural networks and statistics could serve as the foundation for an entire field. This opened the door for chatbots, facial recognition software, and quick and accurate language translation.

The algorithm that enables machines to learn is called backpropagation, and Hinton is best known for his work on it in the 1980s. However, the true significance of neural networks trained via backpropagation was realised only decades later. Consequently, Hinton, regarded as one of the three "godfathers of AI", jointly received the 2018 Turing Award, also known as the "Nobel Prize of Computing".

It's critical here to evaluate the 2024 physics Nobel as a reminder of how society has dealt with the pervasive anxiety that Hinton and others have engendered. It's no secret that Hinton is scared of the tech he helped create because he believes it has the potential to control or even kill people who aren't ready for it.

But there is a debate on the subject. Several experts and authors have considered the potential for a situation like the one in *Terminator*. For example, Yann LeCun, another 2018 Turing Award recipient and "godfather of AI", believes the AI apocalypse is implausible, while author Yuval Noah Harari has captivatingly depicted an AI-dominated world in which humans are "living inside the hallucinations of nonhuman intelligence".

In May 2023, Hinton made headlines when he abruptly quit his job at Google to "freely speak out" about the potential harms AI could inflict, which ranged from disseminating misinformation to endangering human existence and upending the job market.

Hinton stated in an interview with *BBC Newsnight* that in order to address the impact of AI on inequality, the British govt would have to provide a universal basic income. And in an interview with *NYT*, he said, "It is hard to see how you can prevent the bad actors from using it for bad things."

What Hinton's really done is to highlight the contradictory character of the tech. For, he believes that "in the shorter term", AI would deliver many more benefits than risks, and humans shouldn't stop developing this stuff. He has noted the enormous potential benefits of AI in areas like healthcare, scientific research, and mitigating climate change. The power of AI systems is growing at an unprecedented rate, being very close to human intelligence now, and "they will be much more intelligent than us in the future", he's emphasised.

Significantly, two of this year's three chemistry Nobel laureates – John Michael Jumper and Demis Hassabis – are top AI experts who employed AI to predict protein structures in one of the most fascinating scientific breakthroughs of our time. In fact, many experts are still reluctant to refer to ANNs as the output of physics research.

Thus, the physics Nobel Prize this year may be a measure of the extent to which AI is changing science and everything else. Hinton has gone as far as to say that its "huge influence" on humanity could be comparable to that of the Industrial Revolution in the 18th century.

It's not well-known that machine learning has been applied for a long time in fields that have earned physicists Nobel prizes. Processing and sifting through the enormous volumes of data required to find the Higgs particle was one such application. In contrast, this year's laureates used physics to "train" ANNs.

Just as the Nobel Peace Prize has expanded the domain of who can be honoured and the Literature prize has breached boundaries with awards to Svetlana Alexievich and Bob Dylan in 2015 and 2016, respectively, the Nobel Prize in physics may have broken new ground in 2024.

Is Hinton the modern-day Oppenheimer who helped make the atom bomb and later campaigned against the hydrogen bomb? While there are similarities, there is also a difference: unlike Oppenheimer, Hinton has expressed regret. Astounded by the potential of large language models such as GPT-4, Hinton's felt compelled to warn the public about the potential dangers of the tech he fathered. However, the AI genie is out and it cannot be put back in, only kept on a leash.

The writer is a professor at Indian Statistical Institute, Kolkata