BUILDING NATURAL TRUST IN ARTIFICIAL INTELLIGENCE

Two new technologies being developed by Fujitsu Laboratories are making AI MORE TRANSPARENT AND ROBUST.

Artificial intelligence (AI) has evolved to become part of our daily lives, providing us with everything from recommended videos on streaming services to automated responses at call centres. Neural networks, known as deep learning, are one of the most powerful technologies driving this advance. Despite this, how and why these networks make decisions remains largely opaque. As one of the world's leading information and communication technology companies, Fujitsu Laboratories is developing AI systems that are centred on people and built on trust, through being transparent and understandable.

Making AI more transparent

Fujitsu Laboratories has been creating AI solutions for more than three decades. Its achievements include the commercialization of Japan's first computer with AI, FACOM Alpha, in 1985. In 2015, Fujitsu launched the Fujitsu AI Zinrai Platform Services, a framework of human-centric AI capabilities that has been used in areas as diverse as cancer genomic medicine to safe navigation for ships. These kinds of applications require a high degree of confidence in AI. As part of Fujitsu Laboratories' philosophy of trusted AI, Fujitsu researchers have developed wide learning and high durability learning, two technologies that create new ways in which AI can empower people.

FUJITSU IS COMMITTED TO THE SAFE AND SECURE DEVELOPMENT OF AI

Deep learning systems are powered by complex neural networks modeled on the human brain, resulting in a black box that is impenetrable to humans. Making AI explainable is critical, not only in industries like manufacturing, but also in fields like healthcare and transport, where lives may be at stake. Meanwhile, deep learning systems also require enormous volumes of training data to make accurate predictions.

In contrast, wide learning works on a simple principle: evaluate every combination of data items in a system to find important combinations, called knowledge chunks. This approach can generate accurate predictions even with a small volume of training data.

In one example, a wide-learning solution was applied to a factory automation scenario to understand why defective products occasionally appeared. Since such defects are rare, the data associated with them was too small for conventional deep learning techniques. A wide learning analysis was performed on a machine that mixes raw materials into a finished product, including sensor data that measured material input, mixing speed, temperature and voltage. By looking at every important pairing of variables, the system identified combinations that can produce defects — a revelation to the factory engineers.

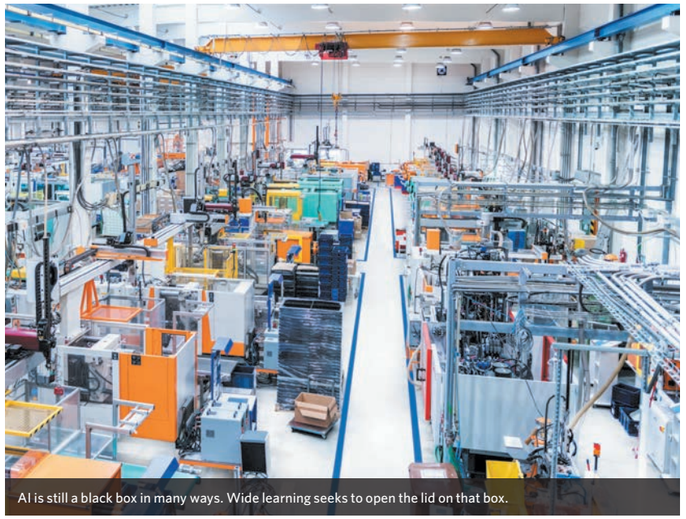
"Using data-mining techniques, wide learning can process 1,000 trillion combinations and select 10,000 to 100,000 important ones in only a few seconds," says Kotaro Ohori, a project manager with Fujitsu Laboratories' AI Core Technologies Project. "Wide learning offers transparency and a low entry barrier when working with small datasets."

Wide learning can also be used in the marketing sector. To effectively target potential customers, marketers use data such as purchase history, gender and other variables that can be mixed into trillions of possible combinations. While marketers could only use a tiny fraction of these, wide learning technology can automatically identify the most important combinations that give the highest number of potential customers likely to make a purchase.

Enhancing the durability of AI

Another problem with AI systems is that their accuracy can deteriorate significantly over time. For example, a financial system built on training data that has become outdated due to new market conditions will not produce accurate results. High durability learning is the world's first technology that can automatically estimate the accuracy of an AI system and, if necessary, update so its predictions remain valid.

High durability learning compares the distribution of data from when an AI system was trained and the distribution changes of data from actual operations. By analyzing differences in those datasets, it can quantitatively estimate the system's accuracy. High durability learning can also automatically adapt an AI system to new input data, maintaining accuracy without expensive retraining.



© yoh4nn/Getty images

AI is still a black box in many ways. Wide learning seeks to open the lid on that box.

© Who is Danny/Shutterstock

High durability learning can update itself when data changes, making it suitable for applications.

This technology was applied to the financial sector using data from 3,800 companies. High durability learning was able to estimate the accuracy of credit risk models within a 3% margin of error. The technology succeeded in maintaining the accuracy of the assessment system at 89%, markedly above the 69% of the conventional approach.

"High durability learning has three merits: it estimates a system's accuracy in real time without manual labeling, it can restore the accuracy after a decline, and it saves on retraining costs," says Katsuhito Nakazawa, a project manager with Fujitsu Laboratories' Trusted AI Project. "As a form of quality-management technology for AI, we aim to apply high durability learning not only to the financial sector, but also to retail and logistics."

Our dependence on artificial learning tools will grow as AI continues to evolve. Fujitsu is committed to the safe and secure development of AI. Through technologies such as wide learning and high durability learning, Fujitsu Laboratories is helping build a human-centric future in which everyone can benefit. ■

FUJITSU

www.fujitsu.com/jp/group/labs/en/

Advertiser retains sole responsibility for content