

What Are Graph Neural Networks?

GNNs apply the predictive power of deep learning to rich data structures that depict objects and their relationships as points connected by lines in a graph.

Author: Rick Merritt

When two technologies converge, they can create something new and wonderful — like cellphones and browsers were fused to forge smartphones.

Today, developers are applying AI's ability to find patterns to massive graph databases that store information about relationships among data points of all sorts. Together they produce a powerful new tool called graph neural networks.

Graph neural networks apply the predictive power of deep learning to rich data structures that depict objects and their relationships as points connected by lines in a graph.

In GNNs, data points are called nodes, which are linked by lines — called edges — with elements expressed mathematically so machine learning algorithms can make useful predictions at the level of nodes, edges or entire graphs.

An expanding list of companies is applying GNNs to improve drug discovery, fraud detection and recommendation systems. These applications and many more rely on finding patterns in relationships among data points.

Researchers are exploring use cases for GNNs in computer graphics, cybersecurity, genomics and materials science. A recent paper reported how GNNs used transportation maps as graphs to improve predictions of arrival time.

Many branches of science and industry already store valuable data in graph databases. With deep learning, they can train predictive models that unearth fresh insights from their graphs.

"GNNs are one of the hottest areas of deep learning research, and we see an increasing number of applications take advantage of GNNs to improve their performance," said George Karypis, a senior principal scientist at AWS, in a talk earlier this year.

Others agree. GNNs are "catching fire because of their flexibility to model complex relationships, something traditional neural networks cannot do," said Jure Leskovec, an associate professor at Stanford, speaking in a recent talk , where he showed the chart below of AI papers that mention them.

Amazon reported in 2017 on its work using GNNs to detect fraud. In 2020, it rolled out a public GNN service that others could use for fraud detection, recommendation systems and other applications.

To maintain their customers' high level of trust, Amazon Search employs GNNs to detect malicious sellers, buyers and products. Using NVIDIA GPUs, it's able to explore graphs with tens of millions of nodes and hundreds of millions of edges while reducing training time from 24 to five hours.

For its part, biopharma company GSK maintains a knowledge graph with nearly 500 billion nodes that is used in many of its machine-language models, said Kim Branson, the company's global head of AI, speaking on a panel at a GNN workshop .

LinkedIn uses GNNs to make social recommendations and understand the relationships between people's skills and their job titles, said Jaewon Yang, a senior staff software engineer at the company, speaking on another panel at the workshop.

"GNNs are general-purpose tools, and every year we discover a bunch of new apps for them," said Joe Eaton, a distinguished engineer at NVIDIA who is leading a team applying accelerated computing to

GNNs. “We haven’t even scratched the surface of what GNNs can do.”

In yet another sign of the interest in GNNs, videos of a course on them that Leskovec teaches at Stanford have received more than 700,000 views.

To date, deep learning has mainly focused on images and text, types of structured data that can be described as sequences of words or grids of pixels. Graphs, by contrast, are unstructured. They can take any shape or size and contain any kind of data, including images and text.

Using a process called message passing, GNNs organize graphs so machine learning algorithms can use them.

Message passing embeds into each node information about its neighbors. AI models employ the embedded information to find patterns and make predictions.

For example, recommendation systems use a form of node embedding in GNNs to match customers with products. Fraud detection systems use edge embeddings to find suspicious transactions, and drug discovery models compare entire graphs of molecules to find out how they react to each other.

GNNs are unique in two other ways: They use sparse math, and the models typically only have two or three layers. Other AI models generally use dense math and have hundreds of neural-network layers.

A 2009 paper from researchers in Italy was the first to give graph neural networks their name. But it took eight years before two researchers in Amsterdam demonstrated their power with a variant they called a graph convolutional network (GCN), which is one of the most popular GNNs today.

The GCN work inspired Leskovec and two of his Stanford grad students to create GraphSage , a GNN that showed new ways the message-passing function could work. He put it to the test in the summer of 2017 at Pinterest, where he served as chief scientist.

Their implementation, PinSage , was a recommendation system that packed 3 billion nodes and 18 billion edges to outperform other AI models at that time.

Pinterest applies it today on more than 100 use cases across the company. “Without GNNs, Pinterest would not be as engaging as it is today,” said Andrew Zhai, a senior machine learning engineer at the company, speaking on an online panel .

Meanwhile, other variants and hybrids have emerged, including graph recurrent networks and graph attention networks. GATs borrow the attention mechanism defined in transformer models to help GNNs focus on portions of datasets that are of greatest interest.

Looking forward, GNNs need to scale in all dimensions.

Organizations that don’t already maintain graph databases need tools to ease the job of creating these complex data structures.

Those who use graph databases know they’re growing in some cases to have thousands of features embedded on a single node or edge. That presents challenges of efficiently loading the massive datasets from storage subsystems through networks to processors.

“We’re delivering products that maximize the memory and computational bandwidth and throughput of accelerated systems to address these data loading and scaling issues,” said Eaton.

As part of that work, NVIDIA announced at GTC it is now supporting PyTorch Geometric (PyG) in addition to the Deep Graph Library (DGL). These are two of the most popular GNN software frameworks.

NVIDIA-optimized DGL and PyG containers are performance-tuned and tested for NVIDIA GPUs. They provide an easy place to start developing applications using GNNs.

To learn more, watch a talk on accelerating and scaling GNNs with DGL and GPUs by Da Zheng, a senior applied scientist at AWS. In addition, NVIDIA engineers hosted separate talks on accelerating

GNNs with DGL and PyG .

To get started today, sign up for our early access program for DGL and PyG .

Original URL: <https://blogs.nvidia.com/blog/2022/10/24/what-are-graph-neural-networks/>