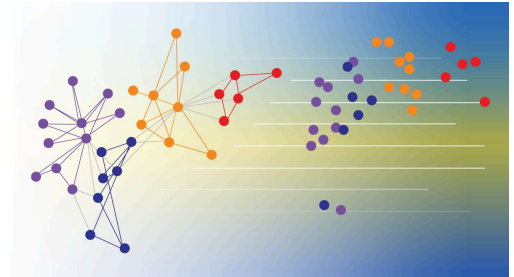# Research

The evolution of graph learning

March 31, 2025

Bryan Perozzi, Research Scientist,
Google Research

We describe how graphs and graph learning have evolved since the advent of PageRank in 1996, highlighting key studies and research.

QUICK LINKS

DeepWalk paper

GCN paper

Research

TensorFlow library

JAX library

Share

[Graphs](#) mathematically represent the connections between things (people, places, objects, etc). They're all around us, in the real world and in our engineered systems — such as social networks, computer networks, communication systems, transportation systems, and biological networks. They provide a powerful and flexible way to model and solve problems involving relationships and connections between objects.

The story of graph theory begins in 1736 with famed mathematician [Leonhard Euler](#), who wondered if one could walk through the city of [Königsberg](#) in [Prussia](#) (now [Kaliningrad](#), [Russia](#)) and cross each of its seven bridges without crossing any of them more than once. In proving that there was no valid solution given this crossing constraint, Euler developed the foundations of modern graph theory. His simple question about bridge crossings is now recognized as one of the famous mathematical problems — the "[Seven Bridges of Königsberg](#)" [problem](#).

Research
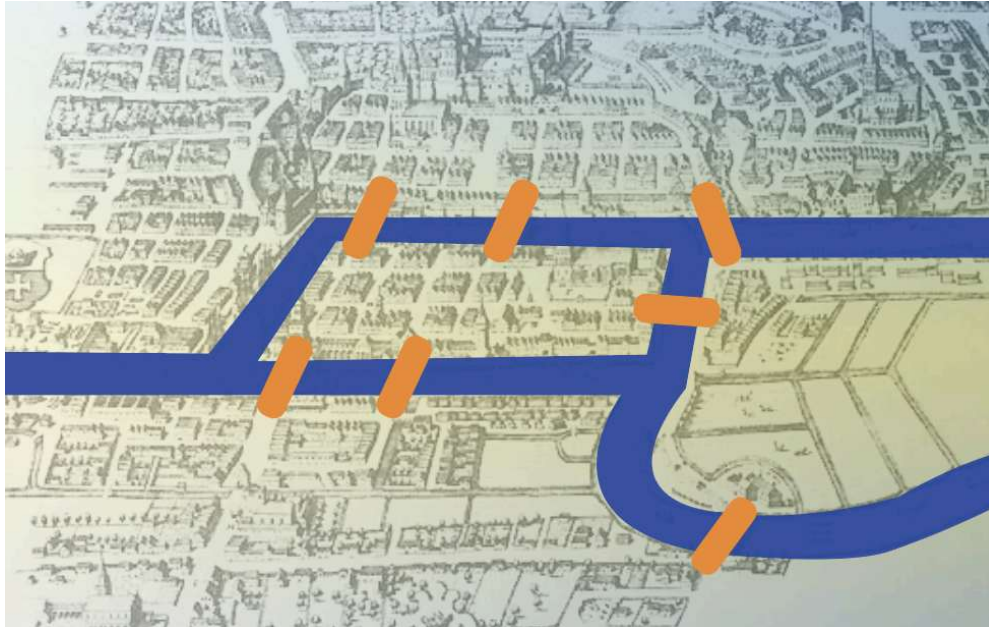
The Seven Bridges of Königsberg connected the city of Königsberg in Prussia, which was split over the [Pregel (now Pregolya) River (image source before editing)](#).

Building on these centuries-old foundations of graph theory, modern applications of graph algorithms have shown particular promise in the realm of computing. Indeed, they have been applied across domains as varied as [traffic prediction](#), [rumor and fake news detection](#), [modeling disease spread](#), [physics simulations](#), and even [understanding why molecules smell](#). Yet the application of graph algorithms to machine learning (ML) was slow to materialize, even though the field had been around for decades.

In this blog post, we trace the recent history of graph-based ML with an emphasis on the role that Google researchers have played in the growth of the field. Beginning with the introduction of [DeepWalk](#) in 2014, which recently won the esteemed [KDD Test of Time Award](#), graph algorithms play an increasingly critical role in modern ML. With new implementations, such as [graph convolutional networks](#) and

Research

Graph algorithms (the pre– deep learning era)

Initial work in graph analysis often focused on developing methods to better understand the structure of graphs. They aimed to uncover hidden patterns, properties, and relationships within graphs (e.g., community structures or centrality within a network) and were concerned with gaining insights into the graph's overall organization and meaning. Meanwhile, parallel efforts focused on designing algorithms to operate over graph structure. These algorithms used the graph as input and performed specific computations or transformations on it (e.g., to calculate shortest paths, maximum flows, etc.). They were concerned with solving well-defined problems based on a graph's existing connections and nodes.

With the rise of web data in the late 1990s and social media in the early 2000s, graph algorithms came into their own. Instead of being mathematical curiosities, they now played a critical role in the rapidly growing Internet. For example, in 1996, Google founders Larry Page and Sergey Brin created PageRank, which would eventually become the backbone of Google Search, and, as such, one of the world's most popular and oft-used graph algorithms. PageRank applied graph theory principles to the web, turning the internet into a giant, interconnected graph of pages (nodes) and hyperlinks (edges). This made it one of the earliest and most influential examples of using graph-based methods to solve real-world problems.

Research

Graph algorithms found success in organizing and operating over graph data, but they had one limitation: it wasn't easy to integrate these discrete algorithms into the world of neural networks. Graph embeddings specifically focus on capturing the relationships between nodes in a graph and preserving the structure of nodes and edges (e.g., adjacency or shortest paths). In contrast, normal neural network embeddings typically represent individual data points (e.g., words or images) based on their features without explicitly modeling relational structure, unless encoded in the data. The key difference lies in the focus on graph-specific relationships in graph embeddings, while regular embeddings focus on feature-based similarity.

In 2014, Bryan Perozzi (Google Research) and colleagues Rami Al-Rfou and Steven Skiena (Stony Brook University) developed DeepWalk, the first practical method for combining graph data with neural networks. DeepWalk uses a neural network encoder to convert graph data into a numeric representation (or *graph embedding*) that captures the similarities between graph objects in a way that's easy for neural networks to understand.

This revolutionized the field of graph analysis and kickstarted the recent flurry of study into graph learning.

Research
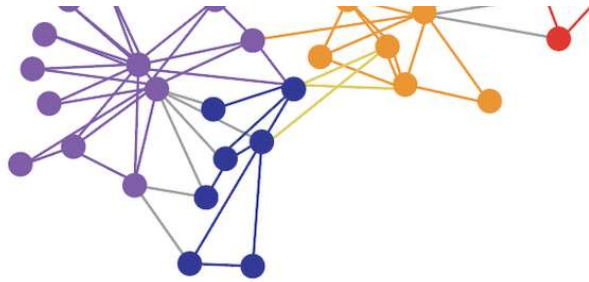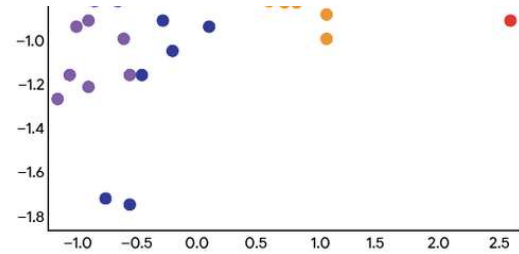
DeepWalk learns a representation that encodes graph information (source: *DeepWalk paper*).

Graph convolutional networks

In 2016, another fundamental leap forward came from Thomas Kipf's (Google DeepMind) thesis work at the University of Amsterdam. Kipf's work introduced graph convolutional networks (GCNs), which can efficiently learn a differentiable mapping over graph structured data, leading to better representations of graph data.

Differential mapping over graph data enables the learning of node and graph representations by applying differentiable operations that aggregate information from neighbors while preserving the graph structure. This approach allows for end-to-end training, capturing both local and global relationships, and scales well to large, dynamic graphs. It has provided a significant advantage in graph learning by improving performance on tasks such as node classification, link prediction, and graph-based recommendations, without the need for manual feature engineering.

Research

(a) Graph Convolutional Network   (b) Hidden layer activations
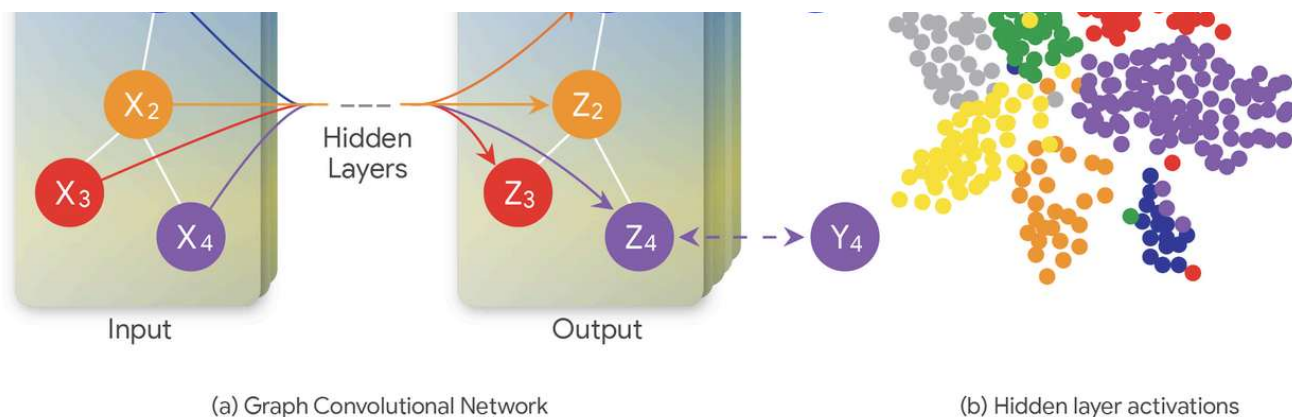
A graph convolution network transforms data over several layers of graph structure.

Message passing and graph neural networks

In 2017, Justin Gilmer's work from Google DeepMind, along with colleagues Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl, enabled graph learning models to compute more complicated functions over graph data in what's known as [messaging passing](). They challenged themselves to build a model that could understand the properties of molecules based on their structures. Through graph learning, their message passing network was able to predict the quantum properties of organic molecules much faster than prior state-of-the-art methods. Since this study, an entire sub-area has emerged using graph learning for [chemical]() and [biological]() applications.

Research
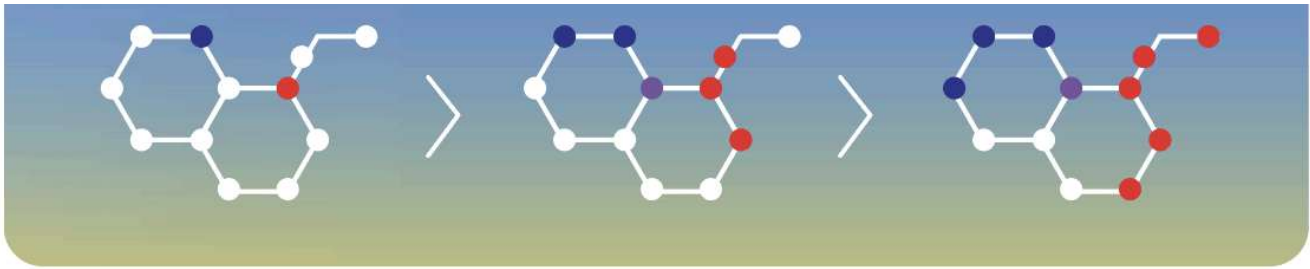
Message Passing Neural Networks use graphs of atom connectivity to infer molecular properties.

Gilmer's body of work (and related research) led to the development of general graph neural network architectures that could represent very flexible models. In 2018, Peter Battaglia (also of Google DeepMind) and colleagues developed the GraphNet model, which was capable of additional feats, such as representing gravity in physics simulations. These graph models have been extended with concepts like attention, which DeepMind's Petar Veličković's added to GNNs in 2018, allowing models to focus on the most important parts of the input data.

The application era: Graph learning in the real world

The entire evolution of graph learning has revolved around enabling better integrations for graph structured data for the sake of improved graph mining. Today, graph learning is rapidly transforming various fields through its ability to model complex relationships. In urban planning, it's being used to predict real-time traffic and analyze mobility patterns. It's also being used in recommendation systems to enhance personalization by leveraging social and knowledge graphs.

Research

The potential for graph learning is unending and a key part of tapping into that potential is knowledge-sharing. In this light, we've released libraries in [TensorFlow](#) and [JAX](#) that provide tools for developers to build models over graph data and to customize their models for processing on TPUs. We recently released [TensorFlow GNN 1.0](#) (TF-GNN), a production-tested library for building GNNs at large scale that supports both modeling and training in TensorFlow as well as the extraction of input graphs from huge data stores.

Looking forward

With the recent remarkable advancements in artificial intelligence one might ask [how can we best integrate graph structured data with artificial intelligence (AI)](#) to allow for the encoding of graphs for large language models (LLMs)? Doing so would yield immense potential for enhanced reasoning, knowledge representation, and contextual understanding, but also presents some challenges, such as representing graphs in a format compatible with LLMs and dealing with large scale graph data. While GNNs offer powerful graph processing capabilities, architectural differences and input modality mismatches have hindered their direct integration into LLM frameworks. However, the promise of richer, more nuanced AI through combined graph and language processing drives ongoing research to bridge this gap.

Ultimately, all of this work will lead to even more powerful understanding of relationships between objects, which will have powerful implications for applications in biology, chemistry, and a wide variety of technology platforms.
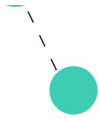
Research

Algorithms & Theory
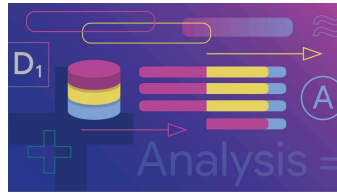
Machine Intelligence

Research

Other posts of interest

MARCH 20, 2025

Load balancing with random job arrivals

Algorithms & Theory · Distributed Systems & Parallel Computing

MARCH 18, 2025

Generating synthetic data with differentially private LLM inference

Machine Intelligence · Natural Language Processing · Security, Privacy and Abuse Prevention

FEBRUARY 13, 2025

Mechanism design for large language models

Algorithms & Theory · Economics & Electronic Commerce · Generative AI

Follow us

Google Research

Google   About Google   Google Products   Privacy   Terms

Help   Submit feedback

Google Research