SML405

Deep Graph Library: Deep Graph learning at scale

George Karypis

Senior Principal Scientist, AI/ML
AWS

Agenda

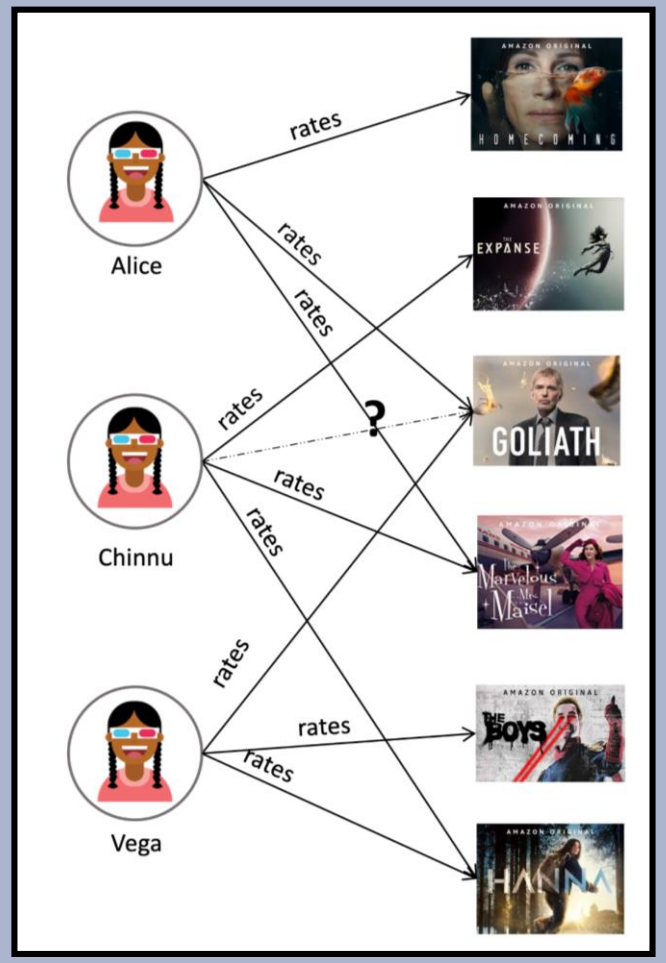
Learning on graphs

Graph neural networks

Deep Graph Library (DGL)

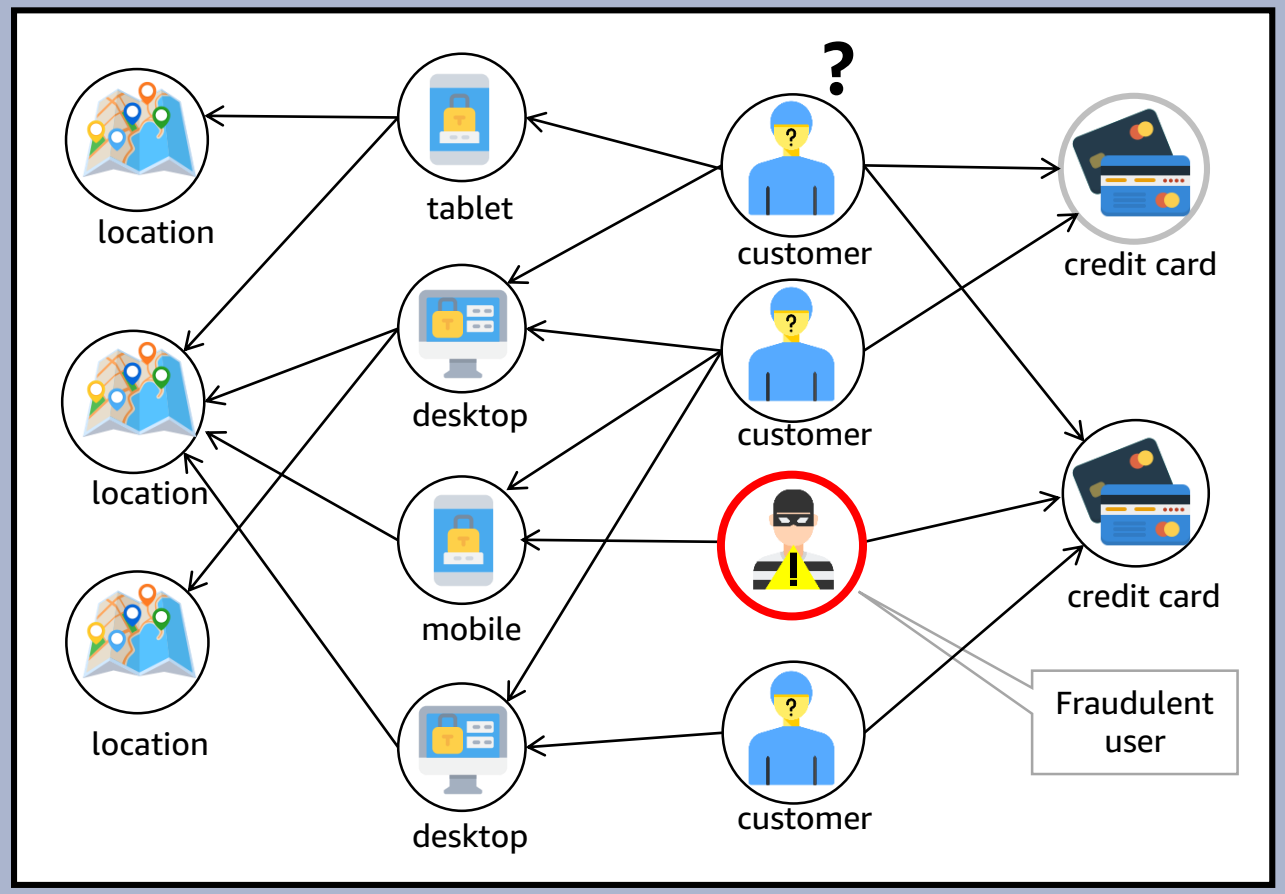
Scaling training to large graphs

Recommending a show



Collaborative (knowledge) graph

Detecting a fraudulent user



Deciding which drug to use for a new disease

BBC Sign in | News | Sport | Reel | Worklife | Travel | Future | More

NEWS

Home | Video | World | US & Canada | UK | Business | Tech | Science | Stories | Entertainment & Arts

Health

Coronavirus: Dexamethasone proves first life-saving drug

By Michelle Roberts
Health editor, BBC News online

16 June 2020 | 2466

https://www.bbc.c

amazon | science — Latest news | Research areas | Conferences & events | Publications

INFORMATION AND KNOWLEDGE MANAGEMENT

Amazon Web Services open-sources biological knowledge graph to fight COVID-19

Knowledge graph combines data from six public databases, includes machine learning tools.

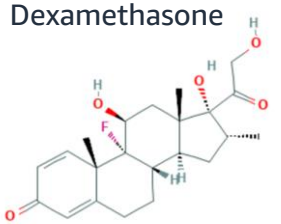
By Cyrus Vahid
June 02, 2020

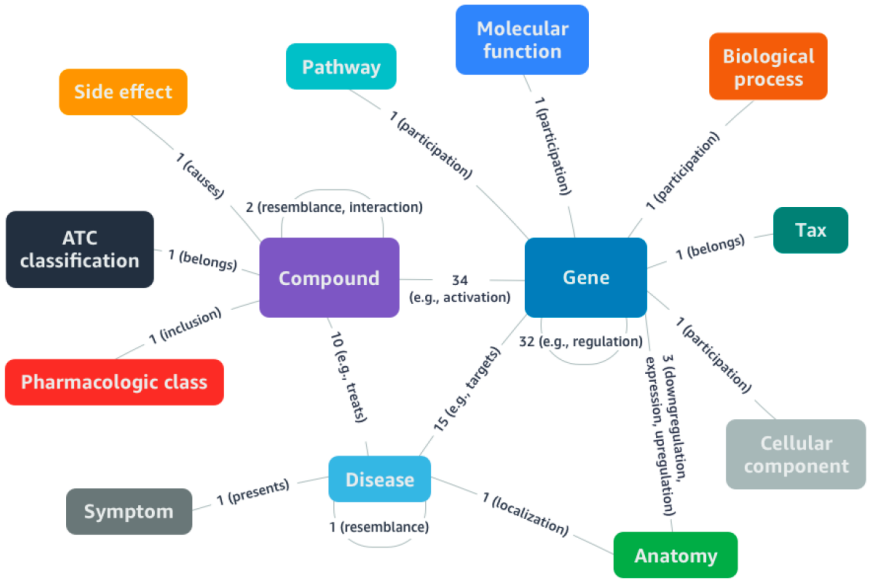
Drug name	Score	Ranking in top-100
Ribavirin	-0.21	0
Dexamethasone	-1.00	4
Colchicine	-1.08	8
Methylprednisolone	-1.16	16
Oseltamivir	-1.39	49
Deferoxamine	-1.51	87

https://github.com/gnn4dr/DRKG
https://arxiv.org/abs/2007.10261

How can I predict which drugs are good candidates for a disease?

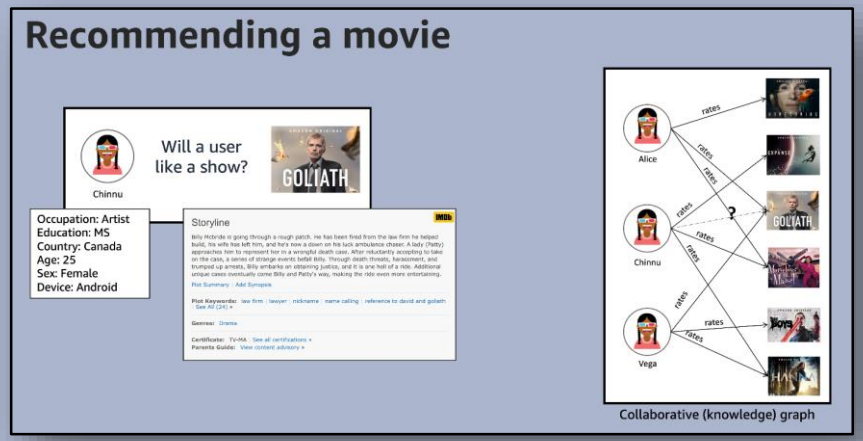
Dexamethasone

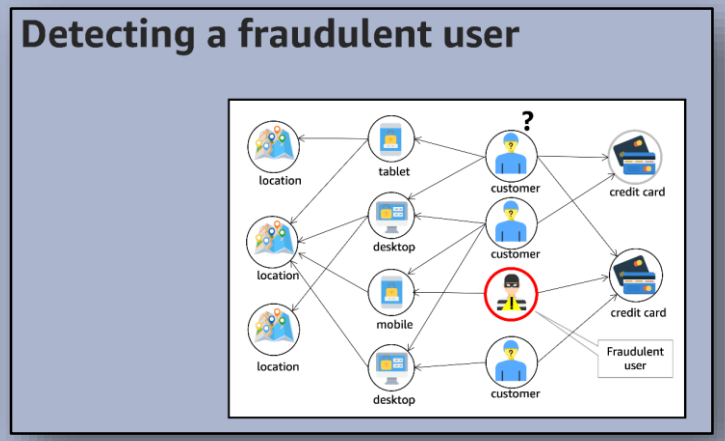


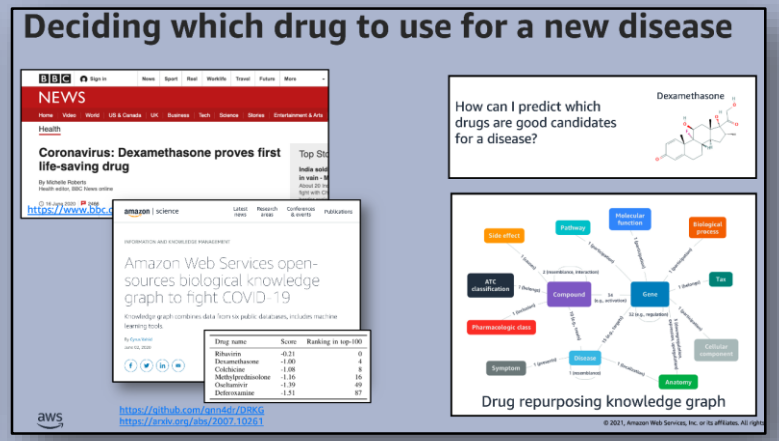


Drug repurposing knowledge graph

The underlying data in all these problems are linked together via graphs.



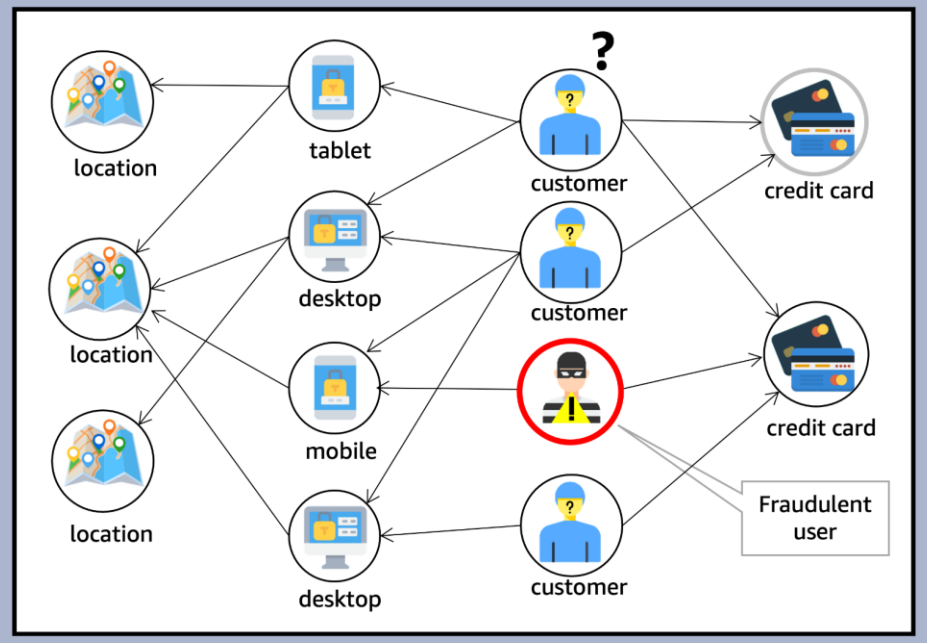




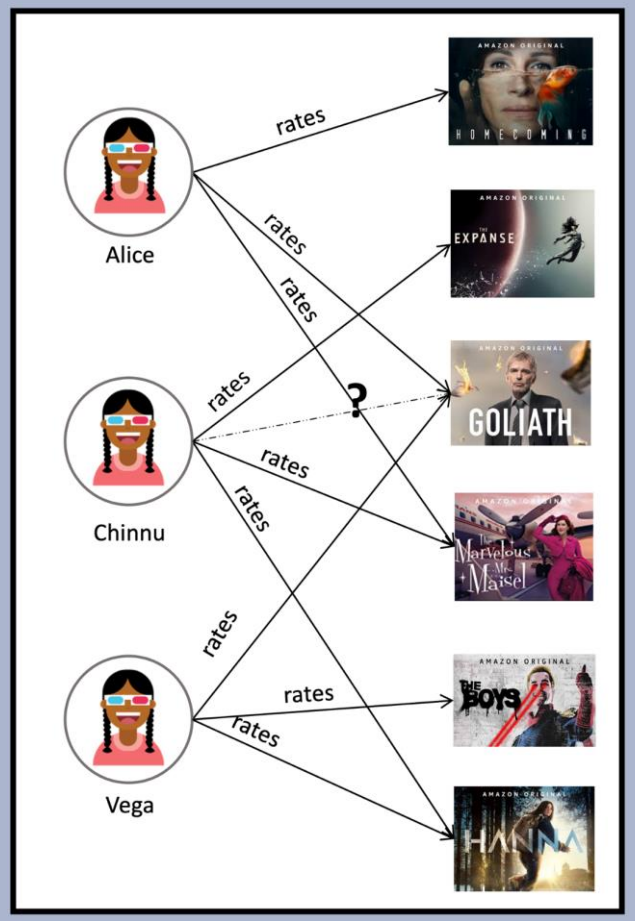
All these problems can be solved by using machine learning approaches based on graph learning.

Tasks in graph learning

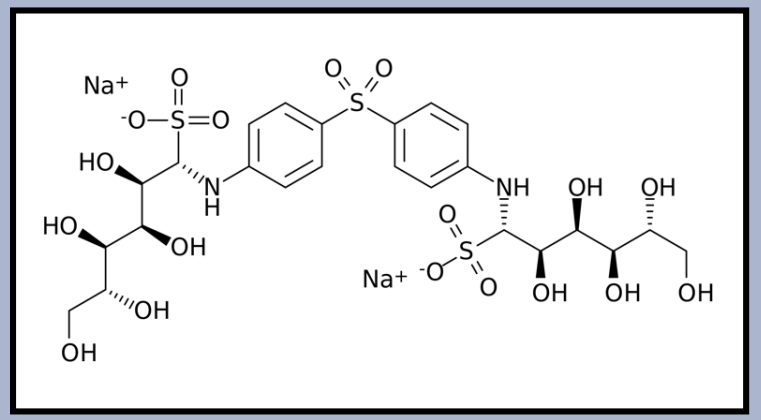
Node-level prediction



Edge-level prediction

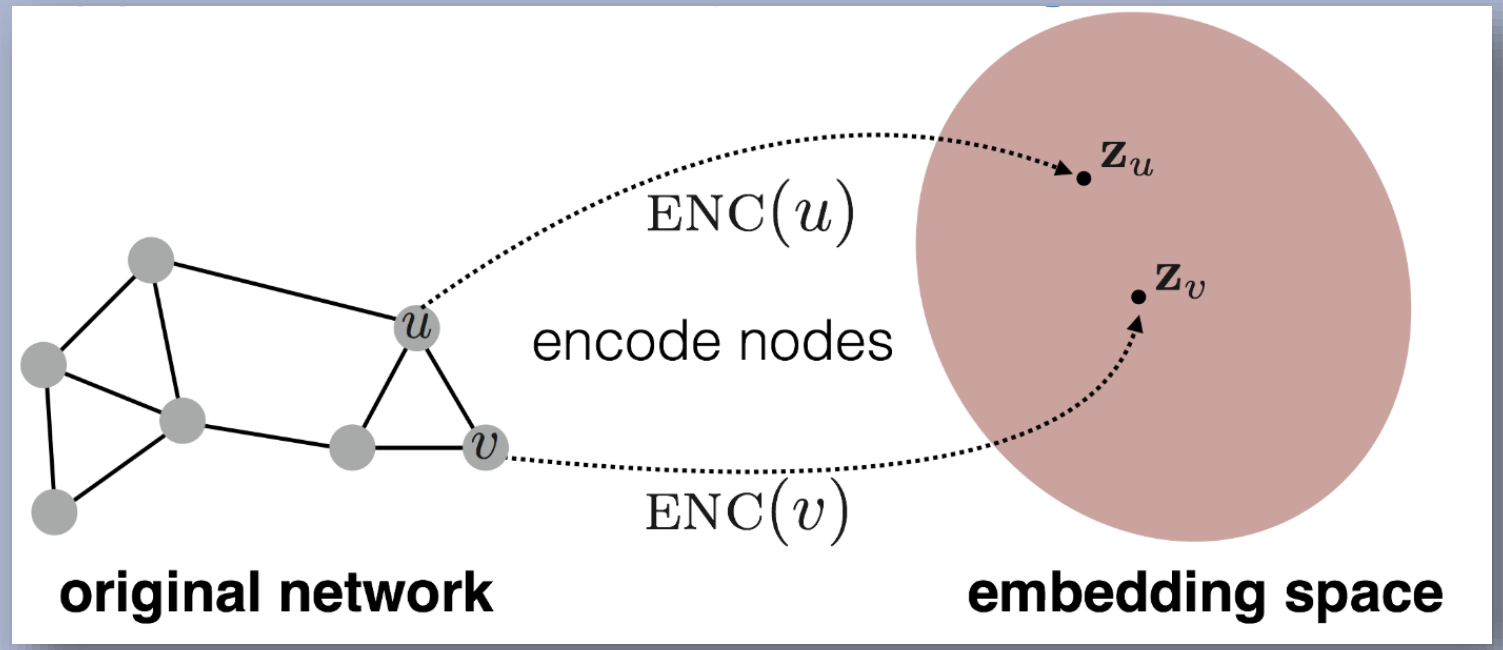


Graph-level prediction



Graph learning and node embeddings

Embed nodes to a low-dimension space by capturing the essential task-specific information and use them to train off-the-self classifiers.



http://snap.stanford.edu/proj/embeddings-www/

Graph embedding approaches

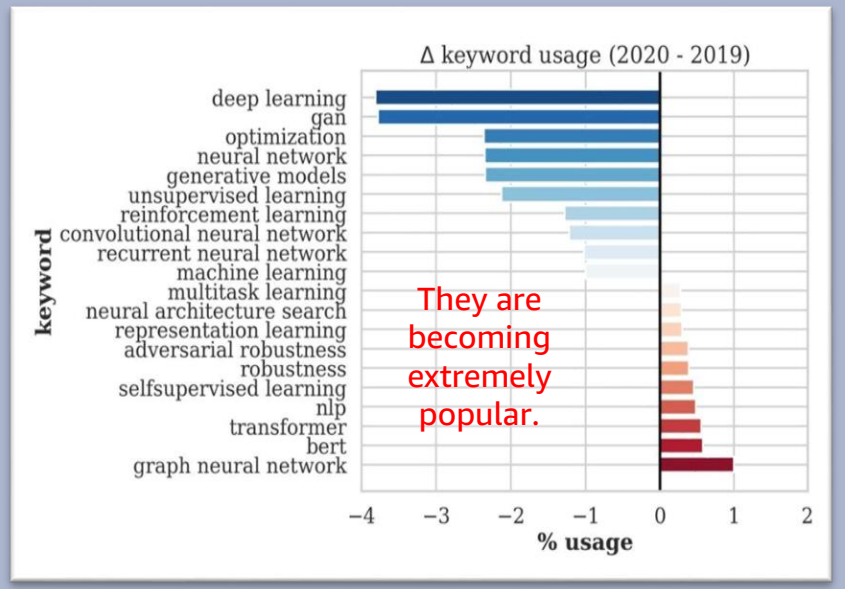
Generate embeddings using manual feature engineering.

Generate embeddings automatically using self-supervised dimensionality reduction approaches.

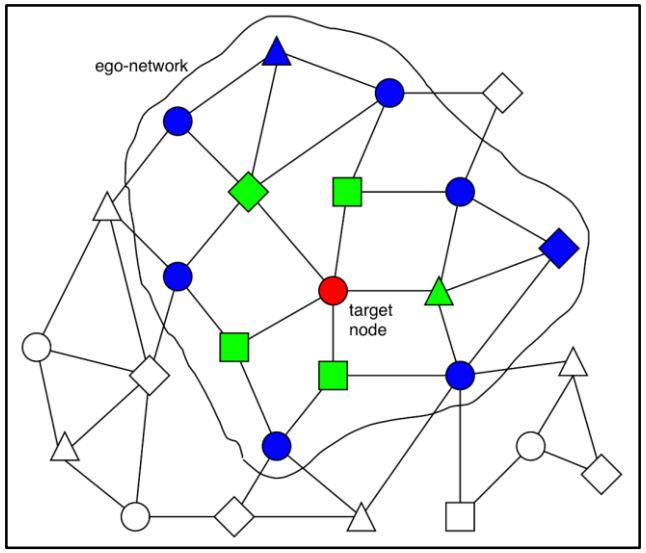
Can we do better?

Graph neural network (GNN)

A family of (deep) neural networks that learn node, edge, and graph embeddings



How do GNNs work?



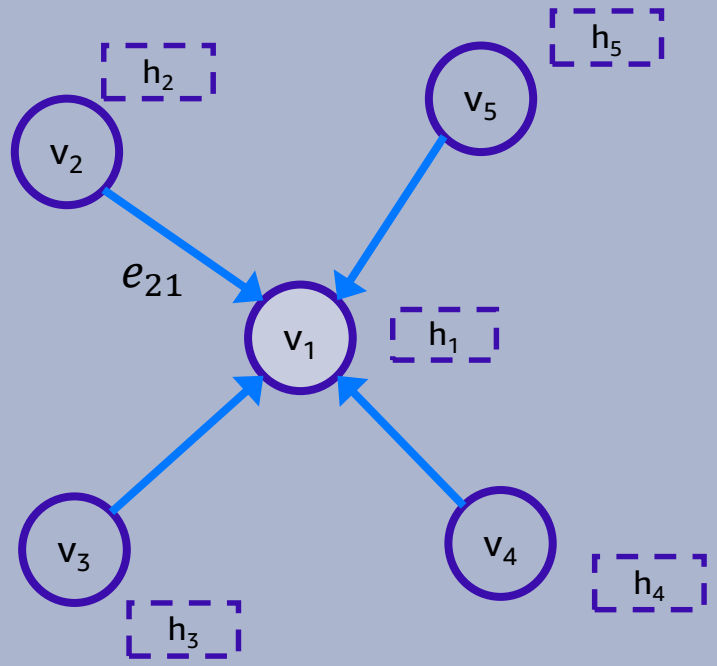
An ego-network around each node is used to learn an embedding that captures task-specific information.

The embeddings use both the structure of the graph and the features of the nodes and edges.

The embeddings are learned in an end-to-end fashion; thus, the predictions are a function of the target node's ego-network.

A general graph neural network formalism

Graph neural networks are based on **message-passing.**



Reduce/aggregate

Message

$$m_v^{(l)} = \sum_{w \in N(v)} M^{(l)}(h_v^{(l-1)}, h_w^{(l-1)}, e_{vw})$$

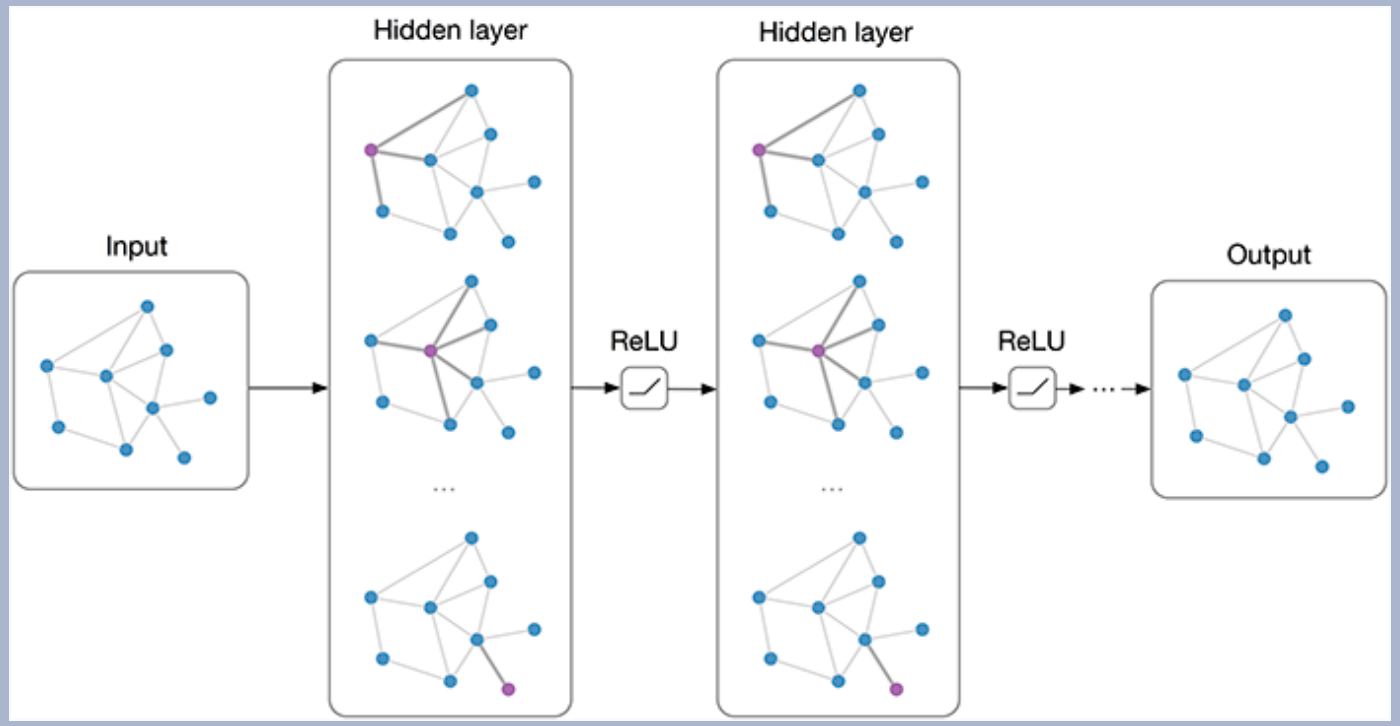
$$h_v^{(l)} = U^{(l)}(h_v^{(l-1)}, m_v^{(l)})$$

Update

GCN: Graph convolution network

$$m_v^{(l)} = \frac{1}{1 + |N(v)|} \sum_{w \in N(v) \cup \{v\}} h_w^{(l-1)}$$

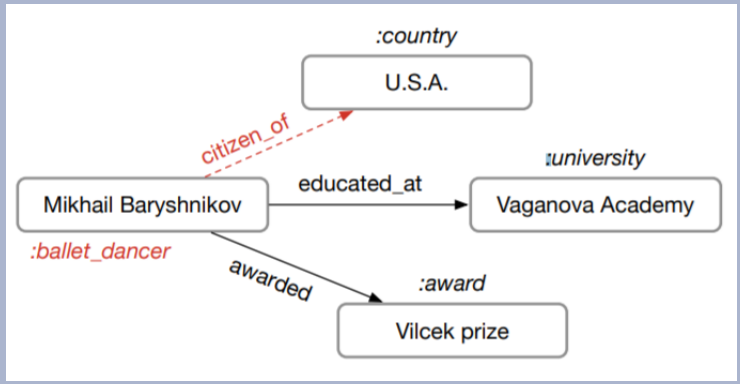
$$h_v^{(l)} = \phi(W^{(l)} m_v^{(l)})$$



https://tkipf.github.io/graph-convolutional-networks/

RGCN: Relational graph convolution networks

Handles graphs whose nodes are connected with different relations

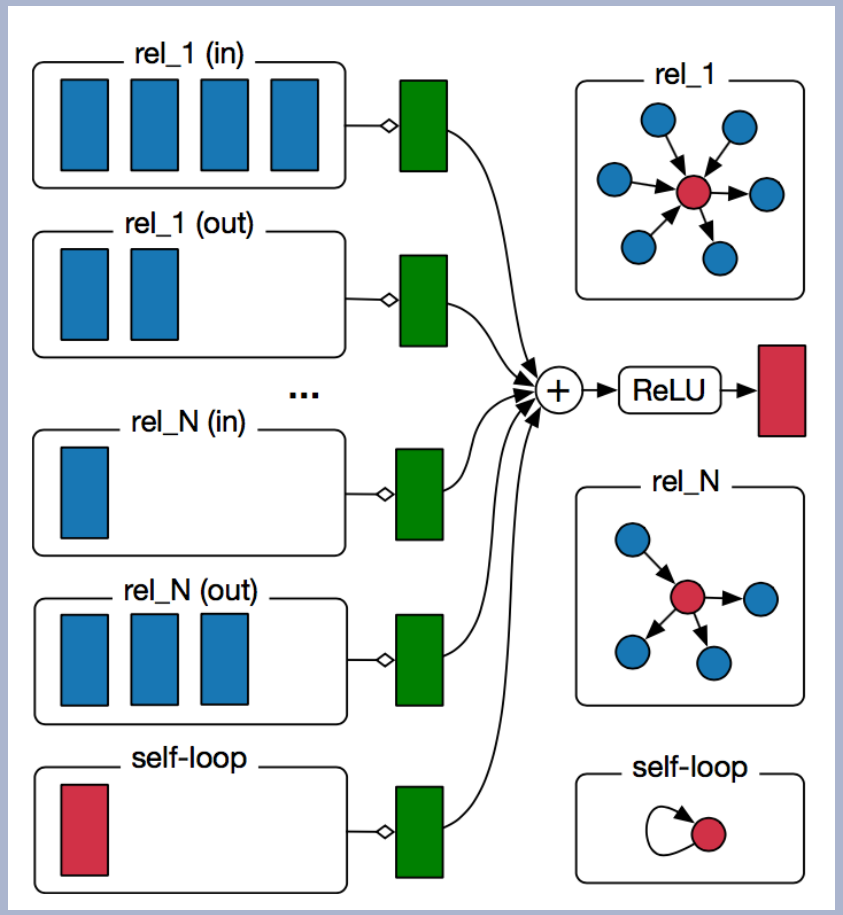


https://arxiv.org/abs/1703.06103

$$M_{vw}^{(l)} = \frac{1}{c_{v,r}} W_r^{(l)} h_w^{(l-1)}$$, r is the relation of e_{vw}

$$m_v^{(l)} = \sum_{w \in N(v) \cup \{v\}} M_{vw}^{(l)}$$

$$h_v^{(l)} = \phi(W^{(l)} m_v^{(l)})$$



https://arxiv.org/abs/1703.06103

How to train GNN on large graphs

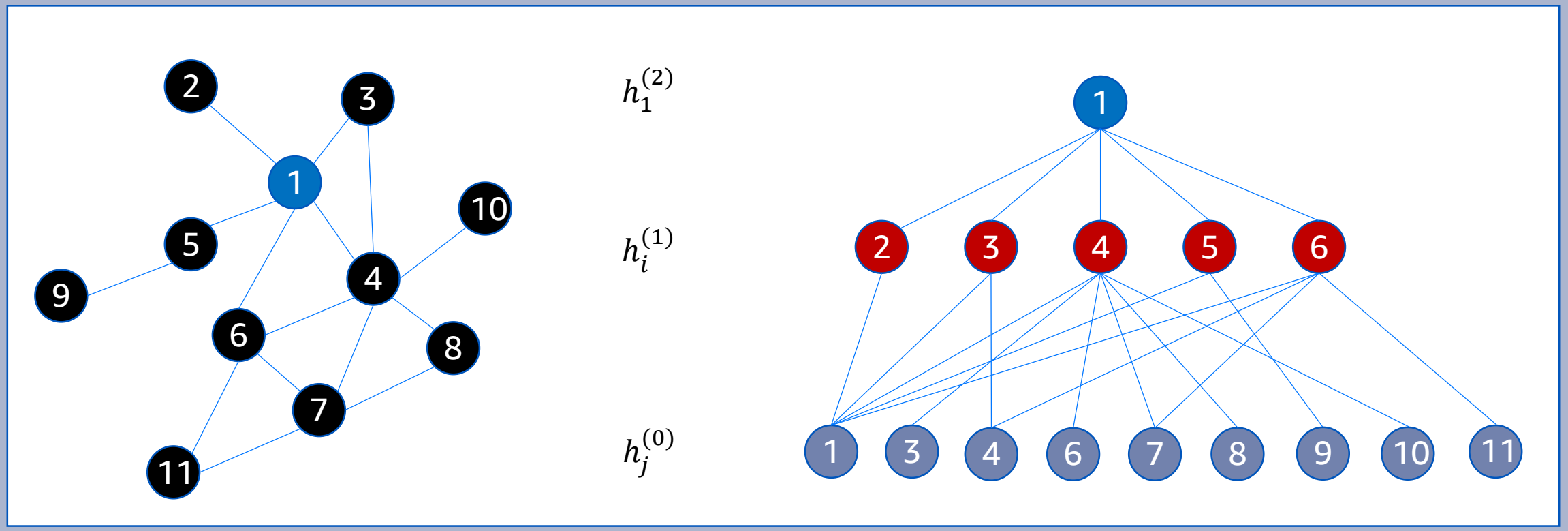
Many applications have extremely large graphs—millions of nodes and billions of edges:

- Social networks
- Recommendation
- Knowledge graph
- …

A typical training method: mini-batch training

A glance of mini-batch training on graphs

A mini batch represents the computation graph for target nodes.

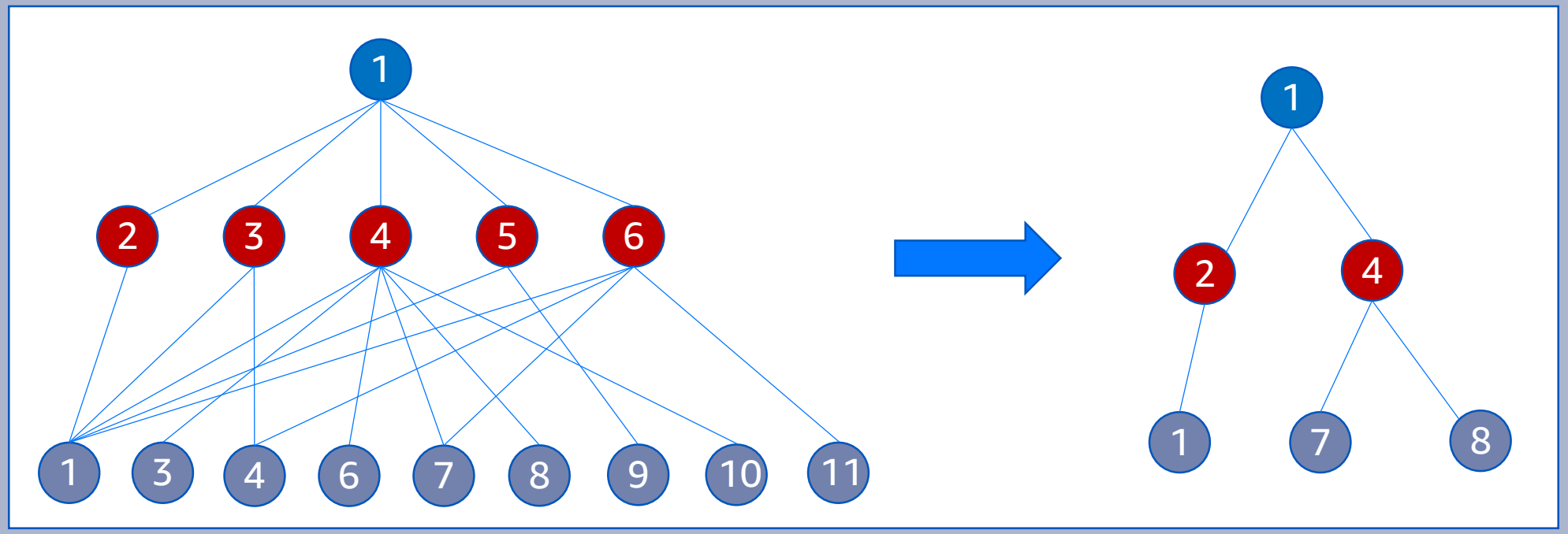


Mini-batch with neighbor sampling

Small-world graphs lead to a huge computation graph.

Prune the computation graph:

- Sample neighbors from a neighbor list of a vertex.

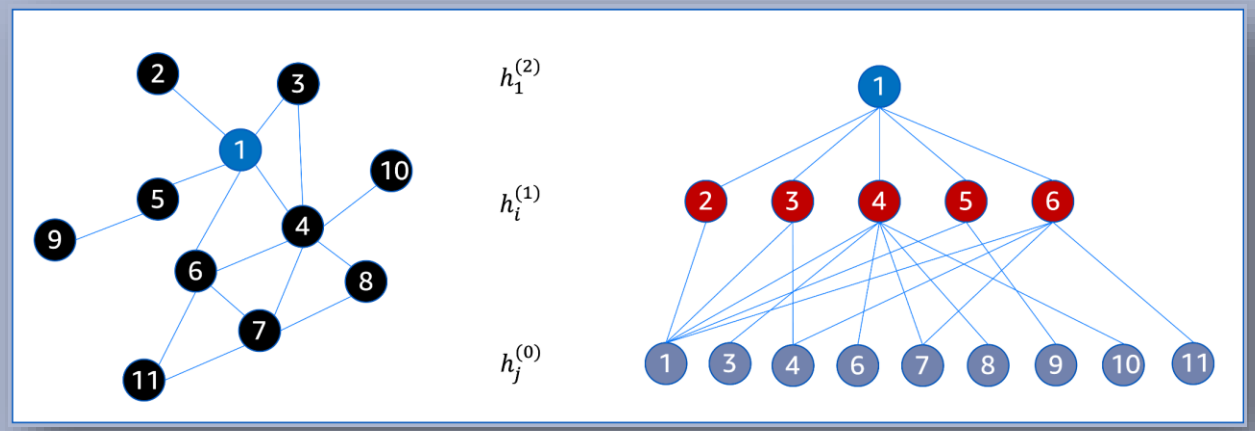


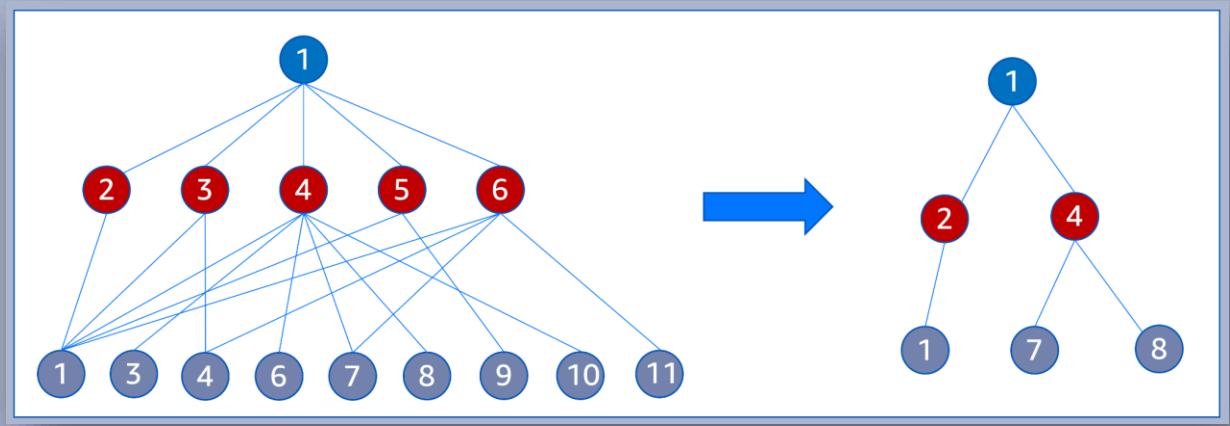
Deep Graph Library

Why do we need something special for GNNs?

Deep learning frameworks are optimized for work-loads whose computational graphs are fixed and for calculations involving dense tensors.

Both do not hold for GNNs.





Development started

First prototype

V0.1
(NeurIPS'18)

V0.2
Sampling APIs

V0.3
Fused message passing
Multi-GPU/-core

V0.3.1
nn.Modules
DGL-LifeSci

V0.4
Heterogeneous graph
DGL-KE

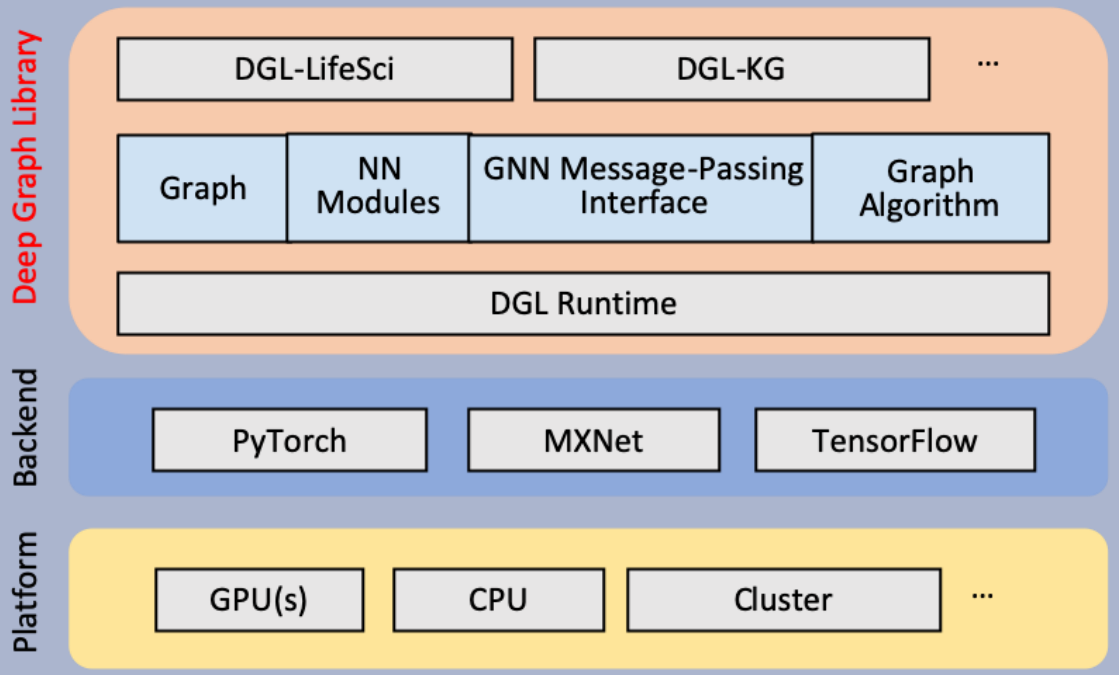
V0.5
TF support
Distributed training
Memory/computation efficiency

V0.6
CPU/GPU improvements
Usability enhancements
Extensive tutorials
Many new examples

2018 2019 2020 2021

aws

DGL's architecture and meta-objective



Forward and backward compatible

- **Forward**: easy to develop new models
- **Backward**: seamless integration with existing frameworks

Fast and scalable

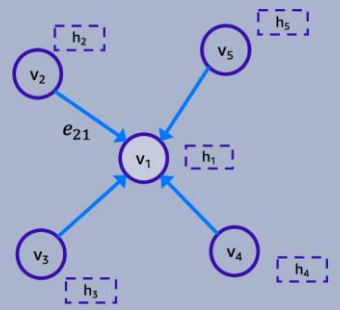
DGL programming interface

Graph as the core abstraction

`DGLGraph`

`g.ndata['h']`

Simple but versatile message passing APIs



$$m_v^{(l)} = \sum_{w \in N(v)} \phi^e(h_v^{(l-1)}, h_w^{(l-1)}, e_{wv})$$

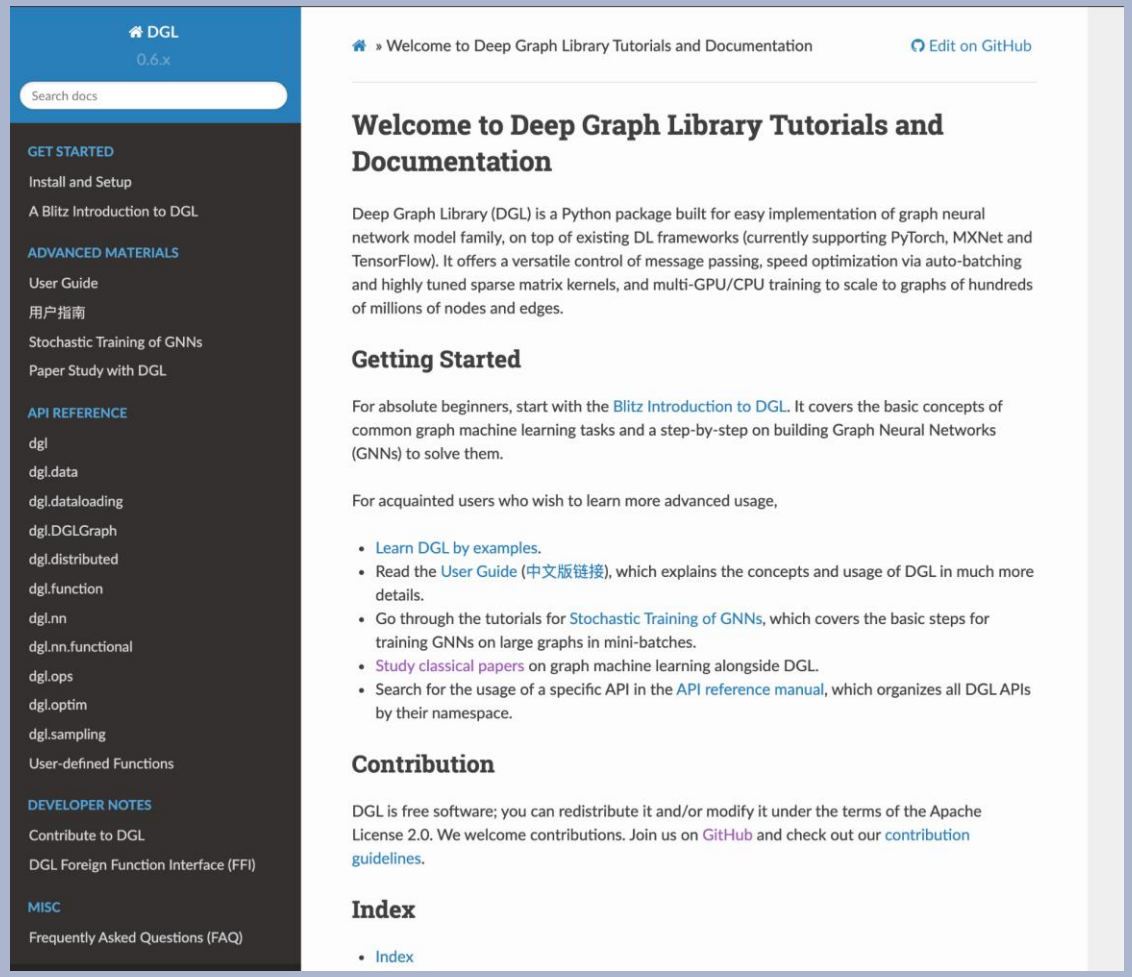
$$h_v^{(l)} = \phi^v(h_v^{(l-1)}, m_v^{(l)})$$

$$\text{update\_all}(\phi^e, \Sigma, \phi^v)$$

ϕ^e, ϕ^v, Σ can be user-defined functions (**UDF**s) or **built-in** symbolic functions.

DGL's API functionalities

- Graph queries
- Message passing API
- Mini-batch generation
- Distributed training (distributed graph, distributed tensors, etc.)
- nn.Modules for popular GNN models
- Graph transformations
- Multi-graph operations (e.g., batching and readout)
- Graph datasets

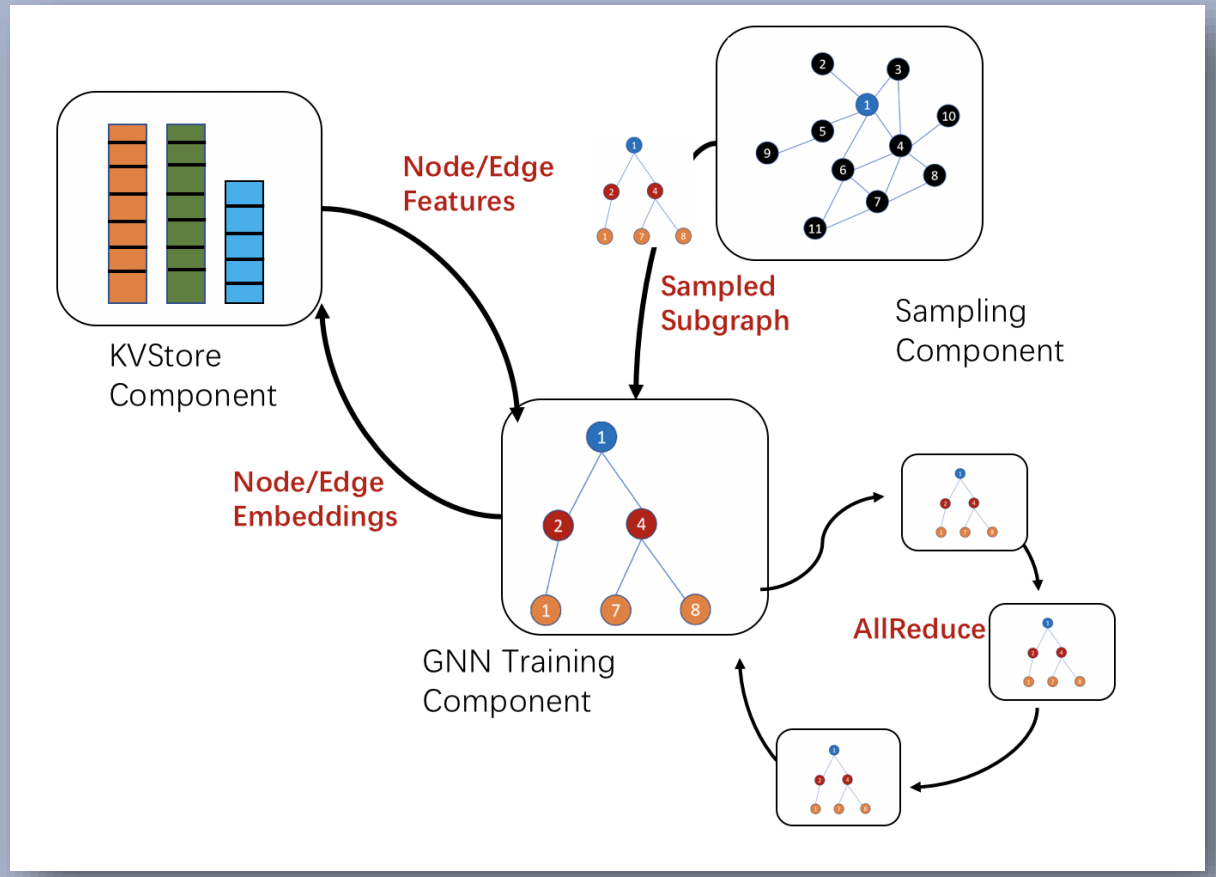


Scaling training to large graphs

Distributed training

Design

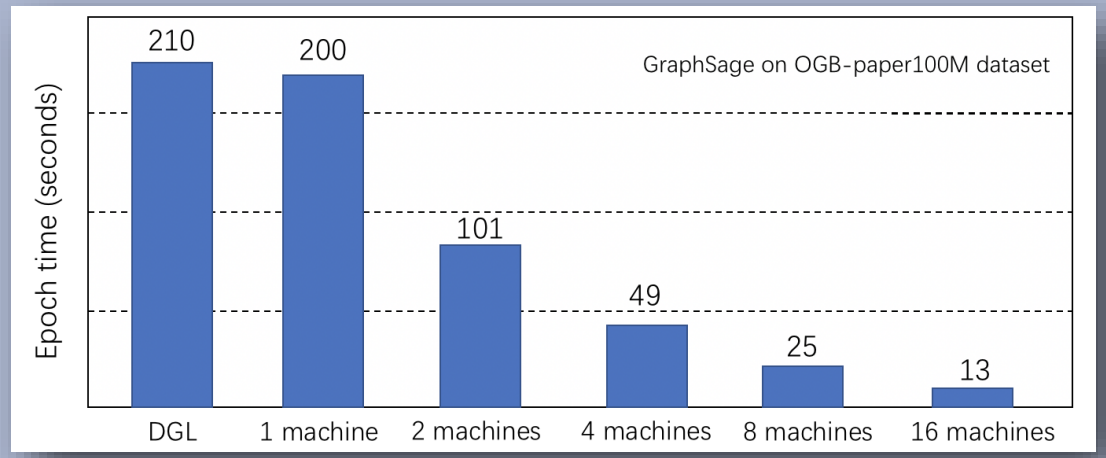
- Partitions the graph across the machines.
- Co-locates data with computations.
- Data parallel mini-batch training.
- Synchronous GNN model parameter updates.
- Asynchronous sparse embedding updates.

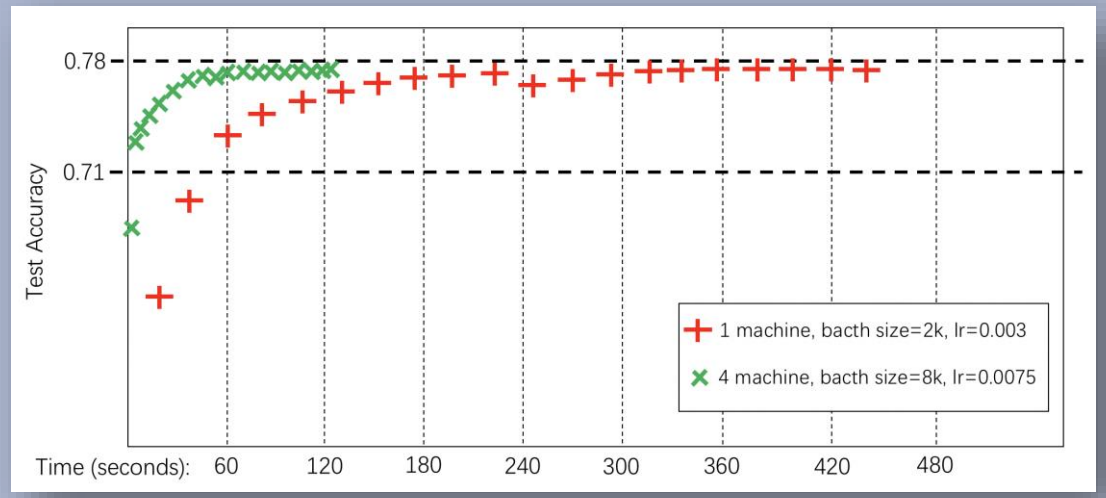


Distributed training

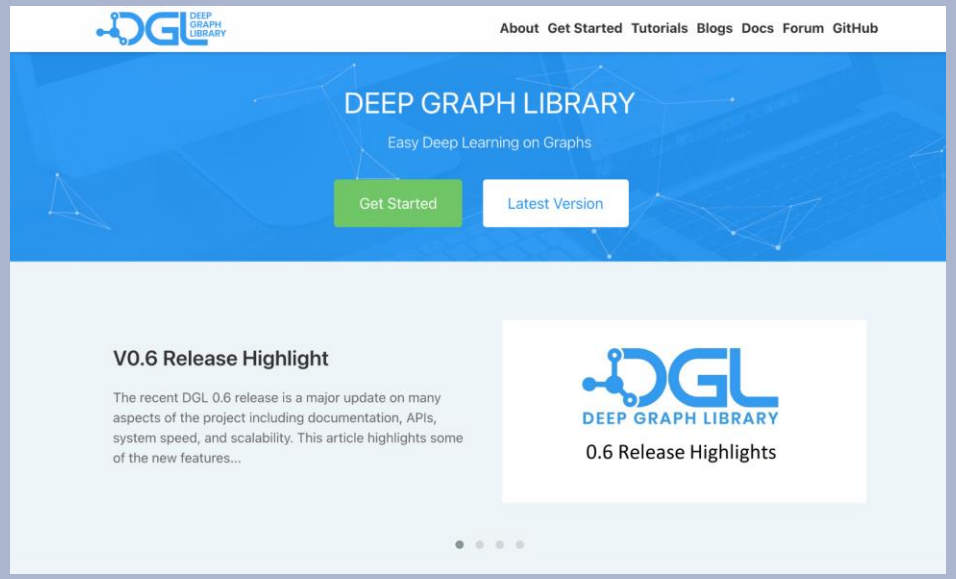
Scale to a graph with 110 million nodes, 3.2 billion edges

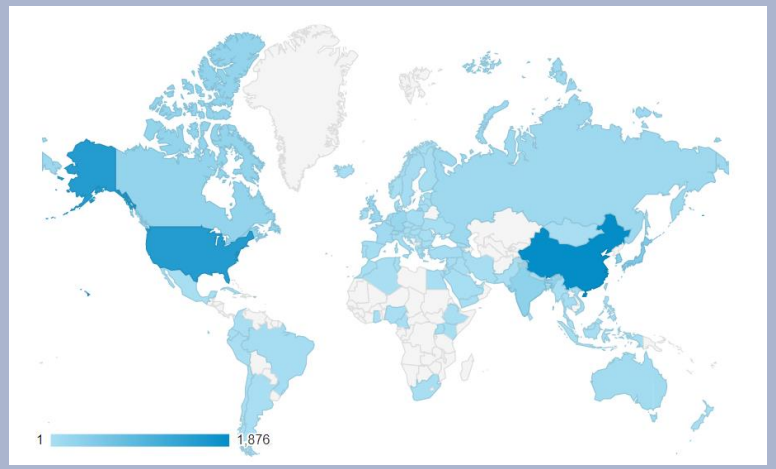
Linear speedup with no compromise in model accuracy

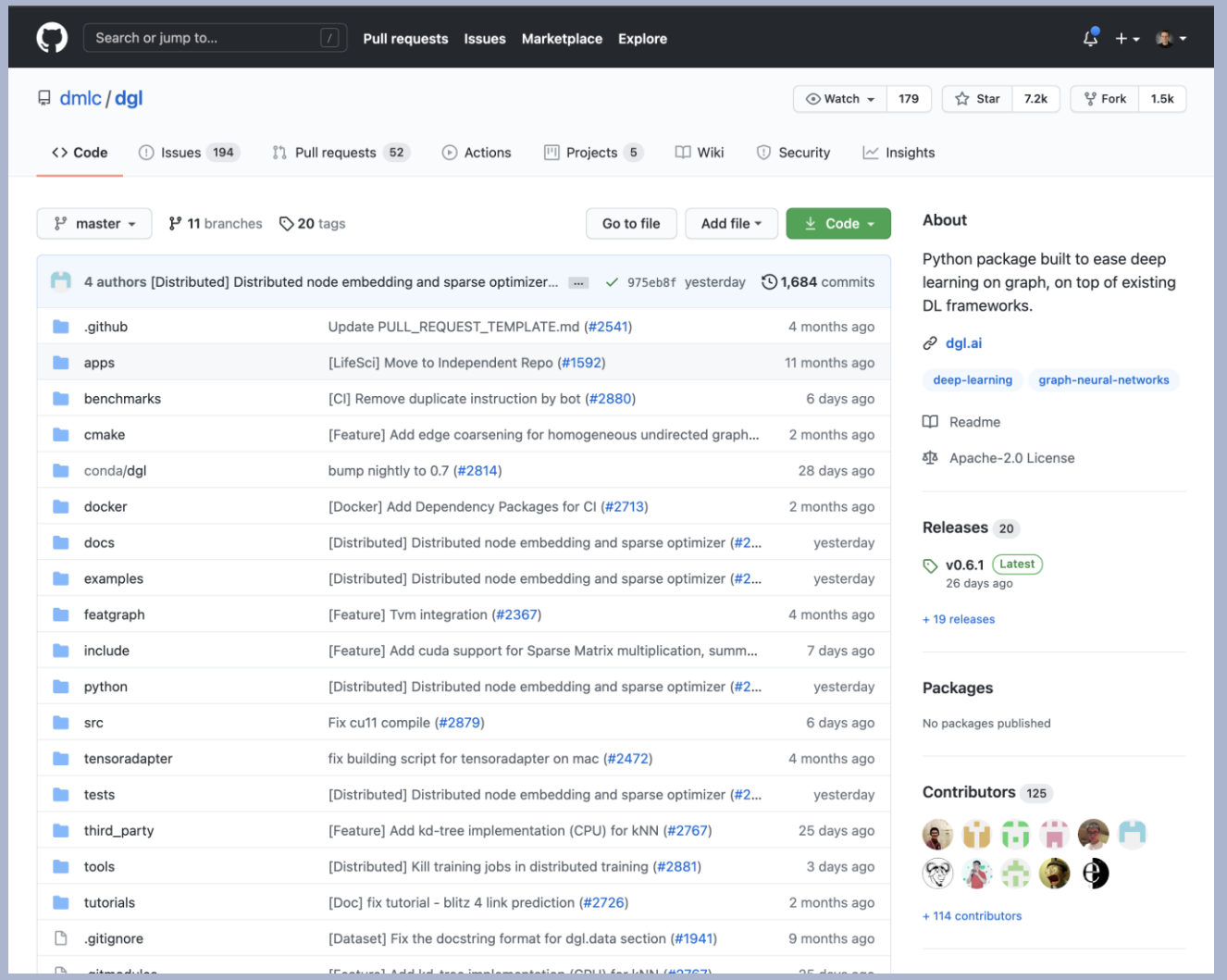




DGL is a vibrant open-source effort







https://www.dgl.ai

Learn ML with AWS Training and Certification

LEARN TO APPLY MACHINE LEARNING TO YOUR BUSINESS, UNLOCKING NEW INSIGHTS AND VALUE

Courses appropriate for your role, including developers, data scientists, data platform engineers, and business decision-makers

65+ free digital machine learning courses from AWS experts let you learn from real-world challenges solved by Amazon

Build credibility and confidence with AWS Certification, including *AWS Certified Machine Learning – Specialty*

Go deeper with labs, white papers, tech talks, and more by accessing the AWS Ramp-Up Guide

Visit aws.training/MachineLearning

aws

Thank you!