GraphSAGE

Inductive Representation Learning on Large Graphs

Thomas Gantz
Alberto Finardi
Tommaso Crippa
Jan Marxen

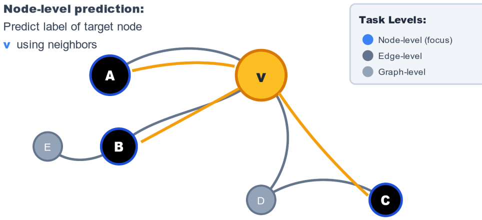
December 2, 2025

Introduction

Graph Prediction Tasks
Introduction

Primary focus: Node-level prediction — predict properties of individual nodes (classification, regression).

- **Node-level (focus):** Predict node attributes or labels using features and neighbor information.
- **Edge-level:** Predict relationships between node pairs (link prediction, edge classification).
- **Graph-level:** Predict properties of whole graphs (e.g., molecule properties).

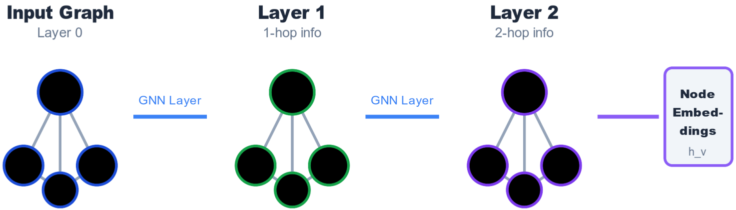


What is a Graph Neural Network?

Introduction

A learnable transformation on graph attributes that:

- Updates node/edge/graph features using **neural networks**
- **Respects graph structure** by aggregating information from neighbors
- Is **permutation invariant** (order of nodes doesn't matter)



Each layer aggregates information from neighbors → deeper layers capture larger neighborhoods

Message Passing: The Core Idea

Three steps repeated at each layer:

1. **Gather:** Collect embeddings from neighboring nodes
2. **Aggregate:** Combine neighbors' info (sum / mean / max)
3. **Update:** Apply learned transform using the aggregated vector

Message Passing Equation

$$h_v^{(k)} = \sigma\left(W \cdot \left[h_v^{(k-1)}, \mathrm{AGG}\left(\{h_u : u \in N(v)\}\right)\right]\right)$$

Analogy: CNNs aggregate over fixed local pixel neighborhoods; GNNs aggregate over variable-size neighbor sets.

Message Passing: Notation

Introduction

- $h_v^{(k)}$ — Representation of node v at layer k
- h_u — Representation of neighbor node u
- $N(v)$ — Set of neighbor nodes of v
- k — Layer index (message-passing hops)
- $\mathrm{AGG}(\cdot)$ — Aggregator function (mean, sum, max, LSTM)
- W — Learnable weight matrix
- σ — Non-linear activation (ReLU, tanh)

Before GraphSAGE: The Problem

Let's audit the limitations of existing approaches...

- **DeepWalk / node2vec:** Transductive: embed every node; new nodes need full retraining.
  - ▷ Not GNNs — random-walk based embeddings, no message passing
- **GCNs:** Often require access to the full graph during training/inference; can be costly to scale.
  - ▷ GNNs — but transductive: fixed node set at training
- **Result:** No compact parametric function to generate embeddings for unseen nodes.

This motivates GraphSAGE's key insight...

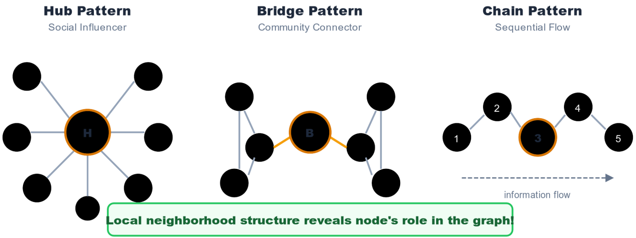
The Key Insight

The Key Insight

Don't learn embeddings for each node...

Learn a FUNCTION that generates embeddings

By sampling & aggregating neighborhood features



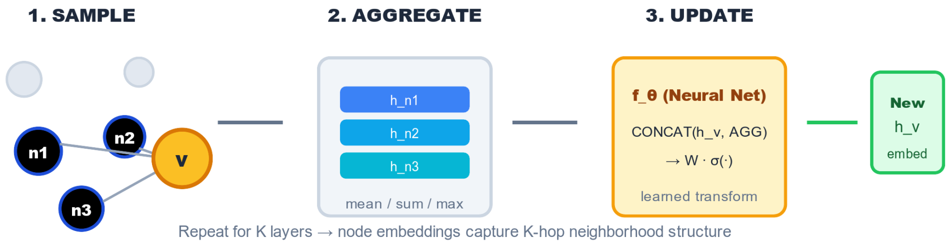
GraphSAGE Framework

GraphSAGE: Inductive Framework

GraphSAGE Framework

Core Principle: Sample + Aggregate

- Learn **aggregator functions** (not node embeddings)
- For any node v: **sample neighbors, aggregate their features**
- Pass through learned neural networks
- **Inference:** Apply same function to unseen nodes



Implementation

Implementation

Implementation

Implementation details coming soon...

Hyperparameters

Hyperparameters

Hyperparameters

Hyperparameter tuning details coming soon...

Results

Results

Experimental results coming soon...

UNIVERSITY OF LUXEMBOURG

Any Question?

UNIVERSITY OF LUXEMBOURG

References

References

[1] Dominique Brodbeck et al. “Domesticating Bead: Adapting an Information Visualization System to a Financial Institution”. In: *Proceedings of the IEEE Symposium on Information Visualization (InfoVis ’97)*. Los Alamitos, CA: IEEE Computer Society, 1997, pp. 73–80.

[2] Aditya Grover and Jure Leskovec. “node2vec: Scalable Feature Learning for Networks”. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016, pp. 855–864.

[3] William L. Hamilton, Rex Ying, and Jure Leskovec. “Inductive Representation Learning on Large Graphs”. In: *Advances in Neural Information Processing Systems (NeurIPS)*. Vol. 30. 2017, pp. 1024–1034. URL: https://arxiv.org/abs/1706.02216.

[4] Susan Havre et al. “ThemeRiver: Visualizing Thematic Changes in Large Document Collections”. In: *IEEE Transactions on Visualization and Computer Graphics* 8.1 (2002), pp. 9–20. DOI: 10.1109/2945.981848.

[5] Thomas N. Kipf and Max Welling. “Semi-Supervised Classification with Graph Convolutional Networks”. In: *arXiv preprint arXiv:1609.02907* (2016).

[6] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. “DeepWalk: Online Learning of Social Representations”. In: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2014, pp. 701–710.

[7] James A. Wise et al. “Visualizing the Non-Visual: Spatial Analysis and Interaction with Information from Text Documents”. In: *Readings in Information Visualization: Using Vision to Think*. Ed. by Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman. San Francisco: Morgan Kaufmann, 1999, pp. 442–445.