Notes (To Be Deleted)

- Nvidia article has some refs to earliest uses/applications of GNNs at https://blogs.nvidia.com/blog/what-are-graph-neural-networks/; https://ieeexplore.ieee.org/document/4700287 First in 2009, really first application in 2016 https://arxiv.org/abs/1609.02907
- Pinterest in 2017 published GraphSage https://arxiv.org/abs/1706.02216

Other application placeholder

- ETA analysis for travel, 2021
  `https://arxiv.org/pdf/2108.11482.pdf`

Graphical Neural Networks

Dominic DiSanto

Department of Biostatistics, Harvard University

November 27, 2023

Downloadable Slides

Outline

1. Set-Up and Motivation 5ish minutes
2. General Construction 10-20ish minutes
3. Applications and Extensions 20-25ish minutes

Downloadable Slides

Outline (Redux)

Downloadable Slides

Goals

- Outline the architecture of Graphical Neural Networks (GNN)
  - Motivation for necessity of GNN's
  - Provide a general framework
- Describe applications and extensions of the general GNN model
  - List specific instances here

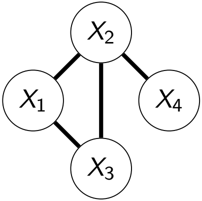
Notation/Set-Up

- Consider the graph $\mathcal{G} = (V, E), E \subseteq V \times V$, where any node v has a related "feature vector" $x_v \in \mathbb{R}^d$

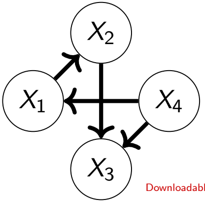
- Let $\mathcal{N}_s(v)$ represent the s-hop neighborhood of any node v (and implicitly $\mathcal{N}(v) \equiv \mathcal{N}_1(v)$)

- Can construct adjacency matrix $A \in \mathbb{R}^{|V| \times |V|}$ describing edge set E
    - $A_{ij} = \mathbb{I}\{(i,j) \in E\}$
    - Adjacency "lists" often used for memory efficiency and permutation invariance

Undirected Graph



Directed Graph



Downloadable Slides

Other application

Image here

Multi-modal Biomedical Data

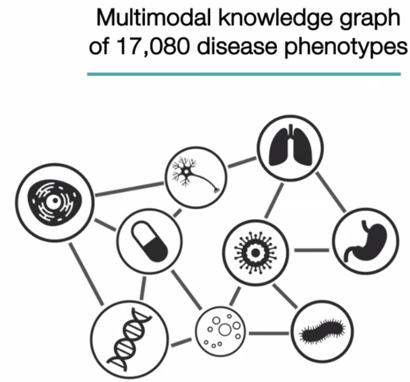
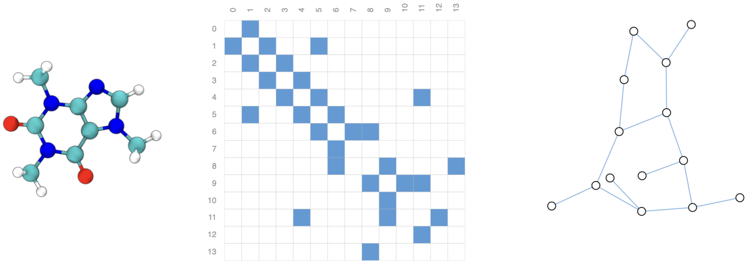


Image courtesy of partial figure from McDermott et al. *Structure-inducing pre-training* [3]

Molecular/Biochemical



(Left) 3d representation of the Caffeine molecule (Center) Adjacency matrix of the bonds in the molecule (Right) Graph representation of the molecule.

Image courtesy of `https://distill.pub/2021/gnn-intro/` [4]

Motivation

- Want to utilize the input structure of the graph
  - Respect/Maintain
  - Update/Estimate
- Why do "regular" NN's/CNN's fail on graphical data? 2009 paper offers pre-GNN data processing led to information loss
- Permutation invariance/Permutation invariant hypotheses
- Applications in chemistry/drug-discovery, large text data(?), multimodal data source analyses

Adjacency Representations

- Adjacency matrix can be prohibitively large but likely sparse, is also permutable but DNN's are not permutation invariant, undesirable
- Can store an adjacency list
- Can use Laplacian matrix $L = \text{diag}(A\mathbf{1}_{|V|}) - A$

What do we estimate about graph structure?

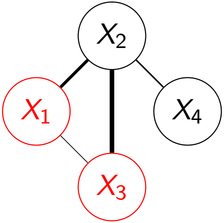
See supp note 1 on Multimodal learning with graphs

$$\text{GNN's can be } \begin{cases} \text{Node-wise } \Phi(\mathcal{G}, x) : (x \in V) \to \mathbb{R}^m \\ \\ \text{Edge-wise } \Phi(\mathcal{G}, e) : (e \in E) \to \mathbb{R}^m \\ \\ \text{Graph-level level } \Phi(\mathcal{G}) : \mathcal{G} \end{cases}$$

Downloadable Slides

Outline (Redux)

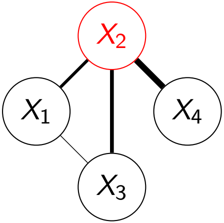
An overly simple representation:



$$V^{(0)} \longrightarrow \Phi_V(V) \longrightarrow V^{(K)}$$

$$E^{(0)} \longrightarrow \Phi_E(E) \longrightarrow E^{(K)}$$

$$\mathcal{G}^{(0)} \longrightarrow \Phi_{\mathcal{G}}(\mathcal{G}) \longrightarrow \mathcal{G}^{(K)}$$



General Framework[1]

General structure:

Initialize $h^{(0)} \leftarrow x_v, \forall v \in V$

for $\kappa = 0, \ldots, K$ **do**:

  for $v \in \mathcal{G}$ **do**:

    $h_{agg}^{\kappa+1} \leftarrow \text{Aggregate}(\{h_u^{(\kappa)} \mid u \in \mathcal{N}_v\})$

    $h^{(\kappa+1)} \leftarrow \text{Update}(h^{(\kappa)}, h_{agg}^{(\kappa)})$

$h_{\mathcal{G}} \leftarrow \text{Readout}(h_v^K | v \in \mathcal{G})$

Can succinctly represent the κth layer as:

$$\mathbf{h}_v^{(\kappa+1)} = \text{Update}\left(x_v^{(\kappa)}, \text{Aggregate}(h_v^{(\kappa)}, x_u^{(\kappa)}, e_{u,v}^{(\kappa)})\right)$$

Choices of (differentiable) functions for Aggregate, Update, and Readout determine the architecture of your GNN

[1] See [1, 5]

Graph Convolutional Network

Proposed in 2017 by Thomas Kipf, Max Welling [2]

$$\mathbf{h}_v^{(\kappa+1)} = \sigma\left(\mathbf{\Omega}\mathbf{H}^{(\kappa)}\mathbf{\Theta}\right)$$

- $\sigma(\cdot) = \mathrm{ReLu}(\cdot)$ activation function
- Learned weight/parameter matrix $\mathbf{\Theta}$
- Including $\mathbf{\Omega}$ normalizing matrix (with known closed form) for computational stability

Review paper, comment on applications briefly (KG setting)

Aggregate/Update

Readout

- Permutation invariant function
- Simple function vs pooling

Outline (Redux)

Downloadable Slides

References I

- Some diagrams generated in conjunction with ChatGPT 3.5

[1] Yasha Ektefaie et al. "Multimodal learning with graphs". en. In: *Nature Machine Intelligence* 5.4 (Apr. 2023). Number: 4 Publisher: Nature Publishing Group, pp. 340–350.

[2] Thomas N. Kipf and Max Welling. *Semi-Supervised Classification with Graph Convolutional Networks*. arXiv:1609.02907 [cs, stat]. Feb. 2017.

[3] Matthew B. A. McDermott et al. "Structure-inducing pre-training". en. In: *Nature Machine Intelligence* 5.6 (June 2023). Number: 6 Publisher: Nature Publishing Group, pp. 612–621.

[4] Benjamin Sanchez-Lengeling et al. "A Gentle Introduction to Graph Neural Networks". en. In: *Distill* 6.9 (Sept. 2021), e33.

References II

[5] Keyulu Xu et al. *How Powerful are Graph Neural Networks?* arXiv:1810.00826 [cs, stat]. Feb. 2019.

Appendix Slides