

Introduction to Graph Neural Networks

Minji Yoon
(CMU)
10707 Introduction to Deep Learning, Spring 2022

Deep Learning
Lecture 10 – Graph Neural Networks

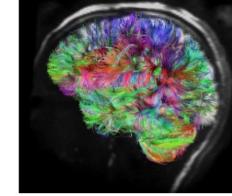
Prof. Dr.-Ing. Andreas Geiger
Autonomous Vision Group,
University of Tübingen / MPIHS

FRIEDRICH-KARL UNIVERSITÄT TÜBINGEN  

(slide: P. Veličković)

Graphs are everywhere

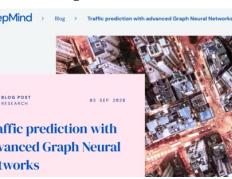



N#Cc1nc(S(=O)(=O)c2ccccc2)sc1S(=O)(=O)c3ccccc3

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

5

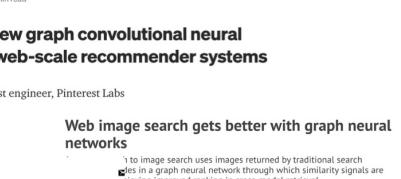
Graph Neural Networks have a large impact on...


Traffic prediction with advanced Graph Neural Networks

Pinterest Engineering | Aug 15, 2018 - 8 min read

PinSage: A new graph convolutional neural network for web-scale recommender systems

Ruining He | Pinterest engineer, Pinterest Labs


Web image search gets better with graph neural networks

amazon | science | PUBLICATION | DECEMBER 4, 2019

P-Companion: A principled framework for diversified complementary product recommendation

By Junghun Han, Tong Zhao, Jin Li, Xin Luo Dong, Christos Faloutsos, Yizhou Sun, Wei Wang


Food Discovery with Uber Eats: Using Graph Learning to Power Recommendations

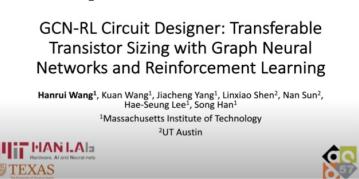
DeepMind | Blog | DECEMBER 19, 2019



Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

6

Graph Neural Networks have a large impact on...


GCN-RL Circuit Designer: Transferable Transistor Sizing with Graph Neural Networks and Reinforcement Learning

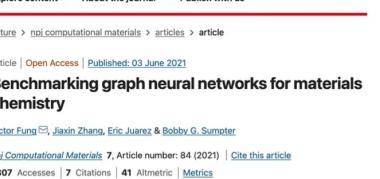
Hanrui Wang¹, Kuan Wang¹, Jiacheng Yang¹, Linxiao Shen², Nan Sun², Hae-Seung Lee¹, Song Han¹
¹Massachusetts Institute of Technology
²UT Austin


The next big thing: the use of graph neural networks to discover particles

September 24, 2020 | Zack Bawitz

Machine learning algorithms can beat the world's hardest video games in minutes and solve complex equations faster than the collective efforts of generations of physicists. But the conventional algorithms still struggle to pick out stop signs on a busy street.

Object identification continues to hammer the field of machine learning — especially when the pictures are incomprehensible and contain noise. And the use of particle detection techniques in high-energy physics experiments. However, a new class of neural networks is helping these models boost their pattern recognition abilities, and the technology may soon be implemented in particle physics experiments to optimize data analysis.


Benchmarking graph neural networks for materials chemistry

Victor Fung¹, Jiaxin Zhang, Eric Juarez & Bobby G. Sumpter

npj Computational Materials 7, Article number: 84 (2021) | Cite this article

7807 Accesses | 7 Citations | 41 Altmetric | Metrics


A graph placement methodology for fast chip design

Asalia Mirosheni¹, Anna Goldie¹, Mustafa Yazar, Joe Wenzle Jian, Ibrahim Sonohor, Shen Wang, Young-Joon Lee, Eric Johnson, Omkar Pathak, Azadeh Naji, Jiwoo Pak, Andy Tong, Kavya Srikrishna, William Heng, Emre Turunc, Quoc V. Le, James Laudon, Richard Ho, Roger Carpenter & Jeff Dean

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

7

Graph Neural Networks have a large impact on...

nature
Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

nature ▾ news > article

NEWS | 01 December 2021
DeepMind's AI helps untangle the mathematics of knots

The machine-learning techniques could sets.
Patterns

Opinion
Neural algorithmic reasoning

Petar Veličković^{1,*} and Charles Blundell¹
¹DeepMind, London, Greater London, UK
Correspondence: petar@google.com
<https://doi.org/10.1038/nature.2021.102375>

We present neural algorithmic reasoning—the art of building neural networks that are able to execute algorithmic computation—and provide our opinion on its transformative potential for running classical algorithms on inputs previously considered inaccessible to them.

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

8

ipam institute for pure & applied mathematics

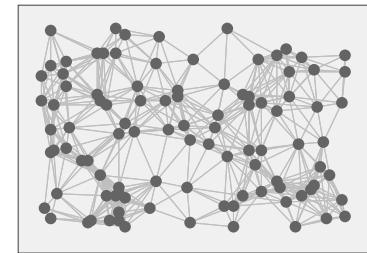
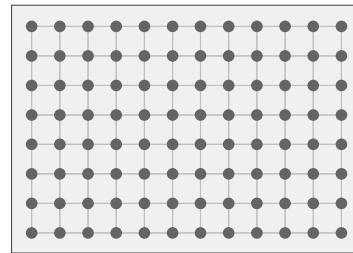
Deep Learning and Combinatorial Optimization

February 22 - 25, 2021



CellPress
OPEN ACCESS

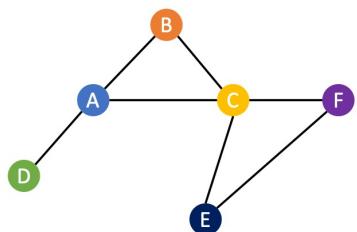
Motivation



- ▶ We know how to build neural nets over larger **grid** structured domains (CNNs)
- ▶ How can we build neural nets for general **graph** structured inputs? (GNNs)
- ▶ Goal: We want to exploit the **local structure** of the graph

5

What is a graph?



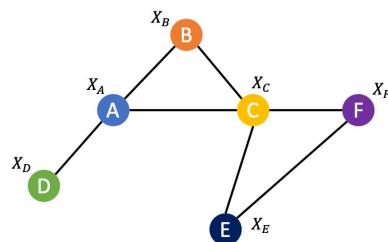
A graph is composed of
 • **Nodes** (also called vertices)
 • **Edges** connecting a pair of nodes presented in an **adjacency matrix**

	A	B	C	D	E	F
A	1	1	1			
B		1				
C	1	1			1	1
D				1		
E					1	
F					1	1

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

3

What is a graph?



A graph is composed of
 • **Nodes** (also called vertices)
 • **Edges** connecting a pair of nodes presented in an **adjacency matrix**

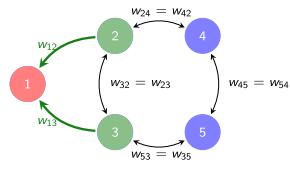
Nodes can have **feature vectors**

A	x_A
B	x_B
C	x_C
D	x_D
E	x_E
F	x_F

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

4

Neighborhoods and Degree



Node 1 neighborhood $\Rightarrow \mathcal{N}(1) = \{2, 3\}$

Node 1 degree $\Rightarrow d_1 = w_{12} + w_{13}$

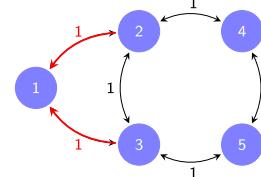
- The **neighborhood** $\mathcal{N}(i)$ of node i is the set of nodes that influence node i :

$$\mathcal{N}(i) = \{j | (i, j) \in \mathcal{E}\}$$

- The **degree** d_i of node i is the sum of weights of its incident edges:

$$d_i = \sum_{j \in \mathcal{N}(i)} w_{ij}$$

Degree Matrix



$$\mathbf{D} = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

- The **degree matrix** \mathbf{D} is a diagonal matrix of degrees: $D_{ii} = d_i$
- We can write \mathbf{D} in terms of the adjacency matrix: $\mathbf{D} = \text{diag}(\mathbf{A}\mathbf{1})$

28

Laplacian Matrix

$$\mathbf{L} = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 3 & -1 & -1 & 0 \\ -1 & -1 & 3 & 0 & -1 \\ 0 & -1 & 0 & 2 & -1 \\ 0 & 0 & -1 & -1 & 2 \end{pmatrix}$$

- The **Laplacian** matrix \mathbf{L} of a graph with adjacency matrix \mathbf{A} is defined as

$$\mathbf{L} = \mathbf{D} - \mathbf{A}$$

- It can also be written explicitly in terms of graph weights

- Off diagonal entries: $L_{ij} = -A_{ij} = -w_{ij}$
- Diagonal entries: $L_{ii} = d_i = \sum_{j \in \mathcal{N}(i)} w_{ij}$

Normalized Adjacency and Laplacian Matrix

- The **normalized adjacency matrix** expresses weights relative to node degrees

$$\bar{\mathbf{A}} = \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}} \Rightarrow \bar{A}_{ij} = \frac{w_{ij}}{\sqrt{d_i d_j}}$$

- The **normalized Laplacian matrix** is similarly defined as

$$\bar{\mathbf{L}} = \mathbf{D}^{-\frac{1}{2}} \mathbf{L} \mathbf{D}^{-\frac{1}{2}} = \mathbf{D}^{-\frac{1}{2}} (\mathbf{D} - \mathbf{A}) \mathbf{D}^{-\frac{1}{2}} = \mathbf{I} - \bar{\mathbf{A}}$$

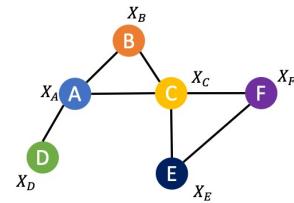
- Normalized operators are more **homogeneous**

30

29

What is Graph Neural Network?

Problem definition

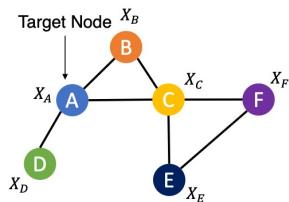


- **Given**
 - A graph
 - Node attributes
 - (part of nodes are labeled)
- **Find**
 - Node embeddings
- **Predict**
 - Labels for the remaining nodes

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

11

Graph Neural Networks

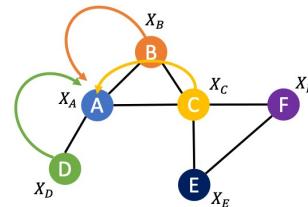


“Homophily: connected nodes are related/informative/similar”

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

12

Graph Neural Networks

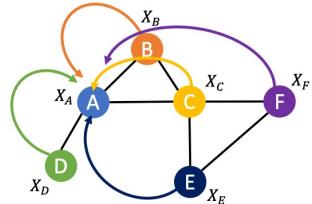


“Homophily: connected nodes are related/informative/similar”

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

13

Graph Neural Networks

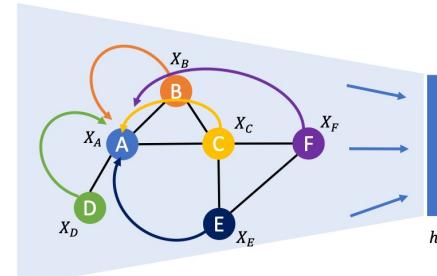


“Homophily: connected nodes are related/informative/similar”

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

14

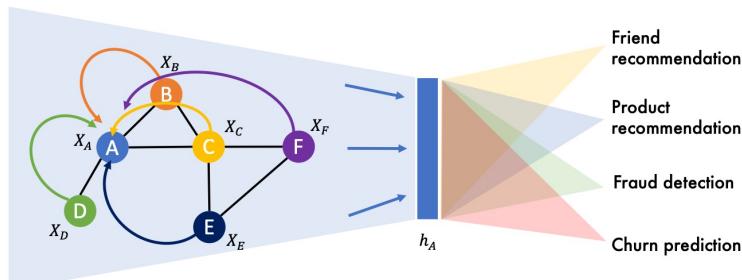
Graph Neural Networks



Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

15

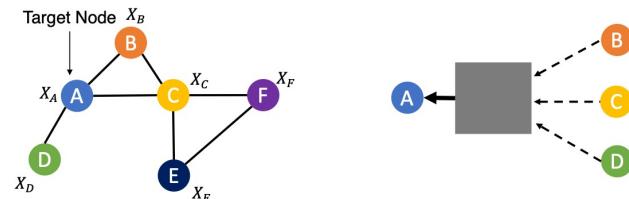
Graph Neural Networks



Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

16

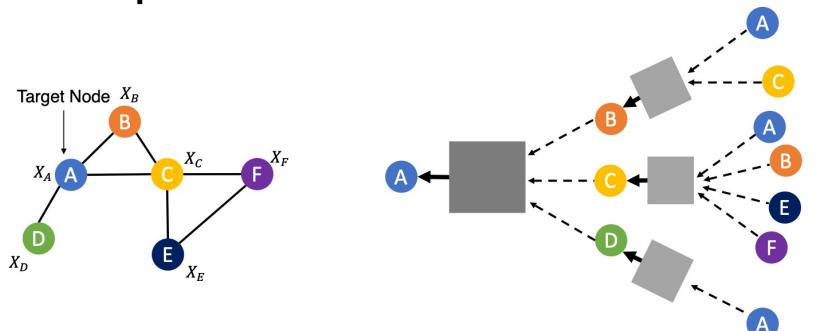
Graph Neural Networks



Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

17

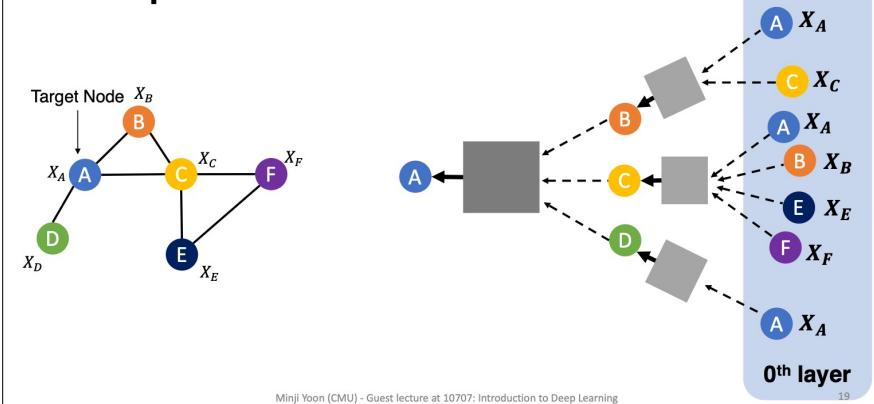
Graph Neural Networks



Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

18

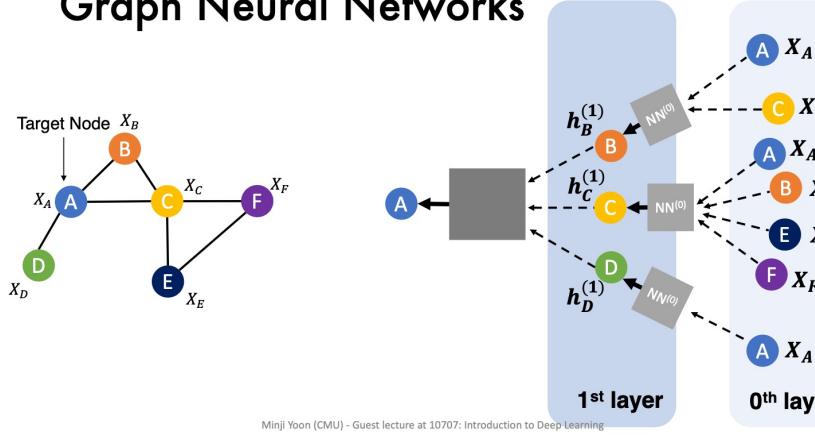
Graph Neural Networks



Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

19

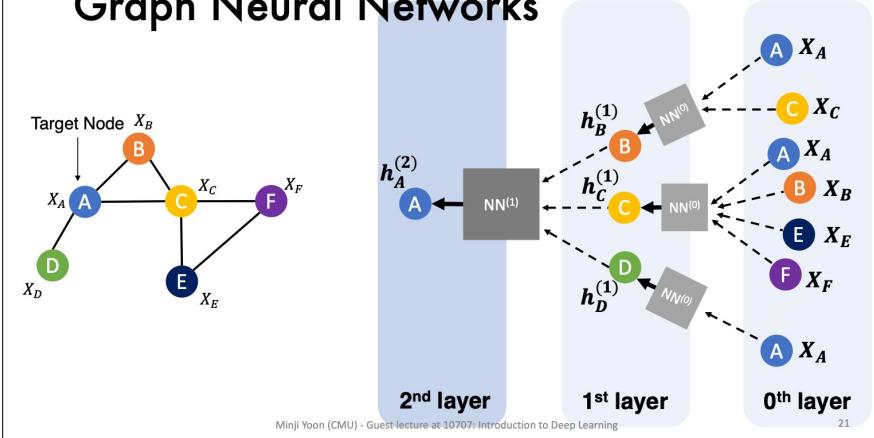
Graph Neural Networks



Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

20

Graph Neural Networks



Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

21

Graph Neural Networks

1. Aggregate messages from neighbors

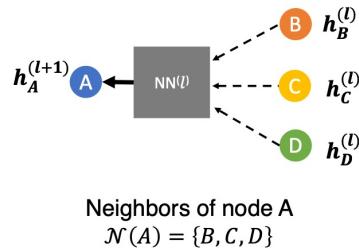
$h_v^{(l)}$: node embedding of v at l -th layer

$\mathcal{N}(v)$: neighboring nodes of v

$f^{(l)}$: aggregation function at l -th layer

$m_v^{(l)}$: message vector of v at l -th layer

$$m_A^{(l)} = f^{(l)} \left(h_A^{(l)}, \{h_u^{(l)} : u \in \mathcal{N}(A)\} \right) \\ = f^{(l)} \left(h_A^{(l)}, h_B^{(l)} h_C^{(l)} h_D^{(l)} \right)$$



Graph Neural Networks

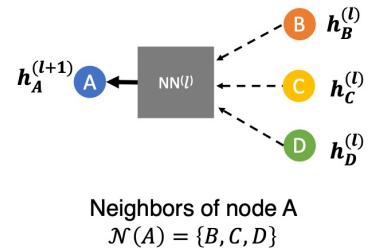
1. Aggregate messages from neighbors

$$m_A^{(l)} = f^{(l)} \left(h_A^{(l)}, \{h_u^{(l)} : u \in \mathcal{N}(A)\} \right) \\ = f^{(l)} \left(h_A^{(l)}, h_B^{(l)} h_C^{(l)} h_D^{(l)} \right)$$

2. Transform messages

$$g^{(l)}$$
: transformation function at l -th layer

$$h_A^{(l+1)} = g^{(l)}(m_A^{(l)})$$



Graph Neural Networks

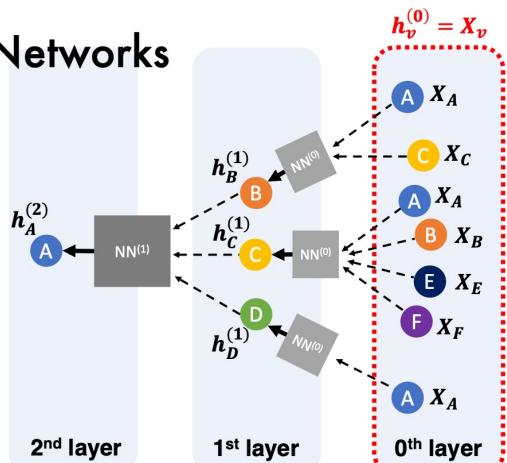
In each layer l ,
for each target node v :

1. Aggregate messages

$$m_v^{(l)} = f^{(l)} \left(h_v^{(l)}, \{h_u^{(l)} : u \in \mathcal{N}(v)\} \right)$$

2. Transform messages

$$h_v^{(l+1)} = g^{(l)}(m_v^{(l)})$$



Graph Neural Networks

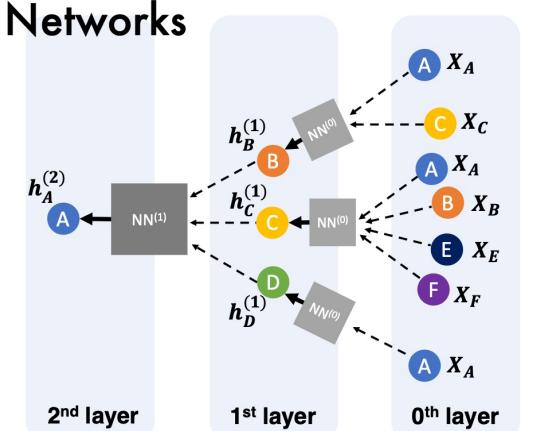
In each layer l ,
for each target node v :

1. Aggregate messages

$$m_v^{(l)} = f^{(l)} \left(h_v^{(l)}, \{h_u^{(l)} : u \in \mathcal{N}(v)\} \right)$$

2. Transform messages

$$h_v^{(l+1)} = g^{(l)}(m_v^{(l)})$$



Graph Neural Networks

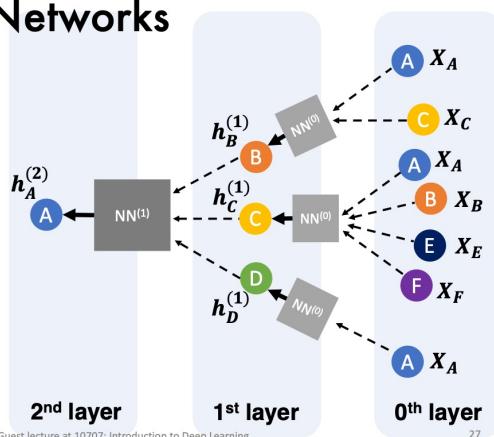
Graph Convolutional Networks^[1]

1. Aggregate messages

$$m_v^{(l)} = \frac{1}{|\mathcal{N}(v) + 1|} \sum_{u \in \mathcal{N}(v) \cup \{v\}} h_u^{(l)}$$

2. Transform messages

$$h_v^{(l+1)} = \sigma(W^{(l)} \circ m_v^{(l)})$$



[1] Kipf, Thomas N., et al. "Semi-supervised classification with graph convolutional networks."

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

Graph Neural Networks

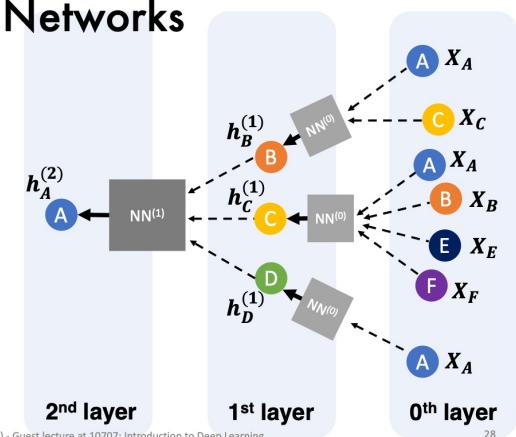
Graph Isomorphism Networks^[2]

1. Aggregate messages

$$m_v^{(l)} = \sum_{u \in \mathcal{N}(v) \cup \{v\}} h_u^{(l)}$$

2. Transform messages

$$h_v^{(l+1)} = \sigma(W^{(l)} \circ m_v^{(l)})$$



[2] Xu, Keyulu, et al. "How powerful are graph neural networks?."

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

Graph Neural Networks

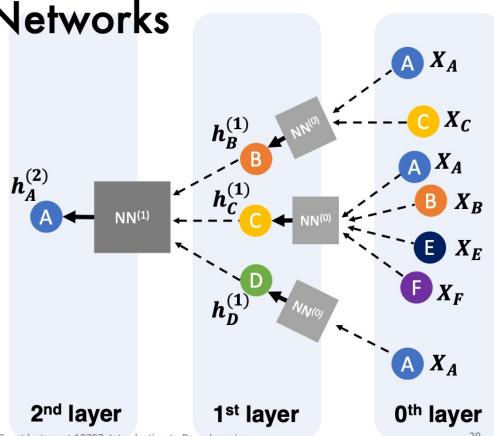
Simplified GCN^[3]

1. Aggregate messages

$$m_v^{(l)} = \frac{1}{|\mathcal{N}(v)|} \sum_{u \in \mathcal{N}(v)} h_u^{(l)}$$

2. Transform messages

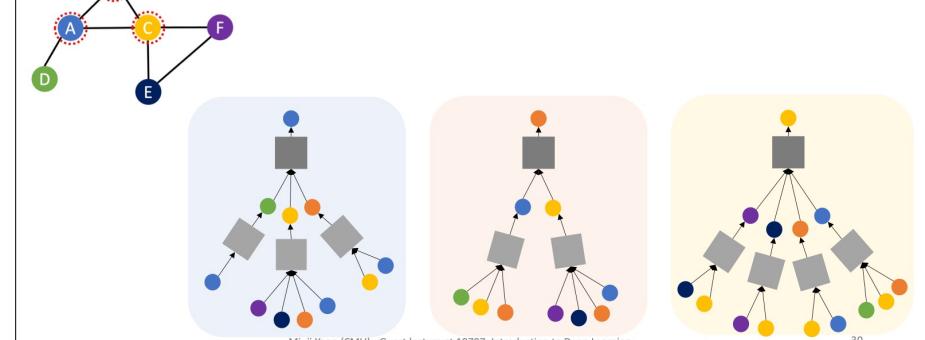
$$h_v^{(l+1)} = W^{(l)} \circ m_v^{(l)}$$



[3] Wu, Felix, et al. "Simplifying graph convolutional networks."

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

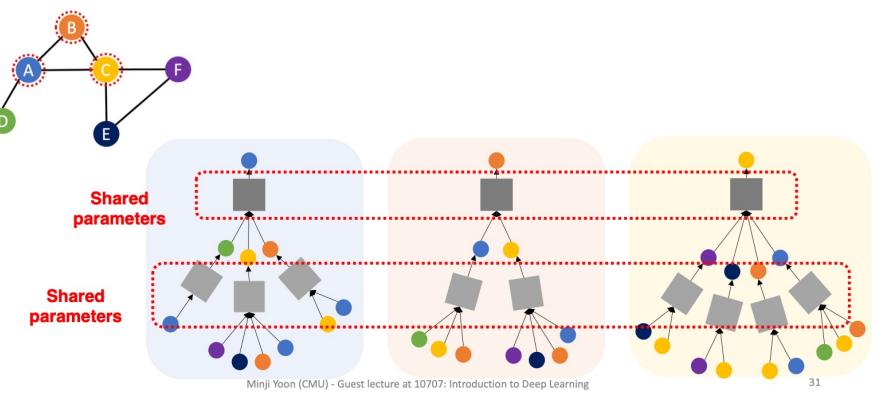
Computation graphs



Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

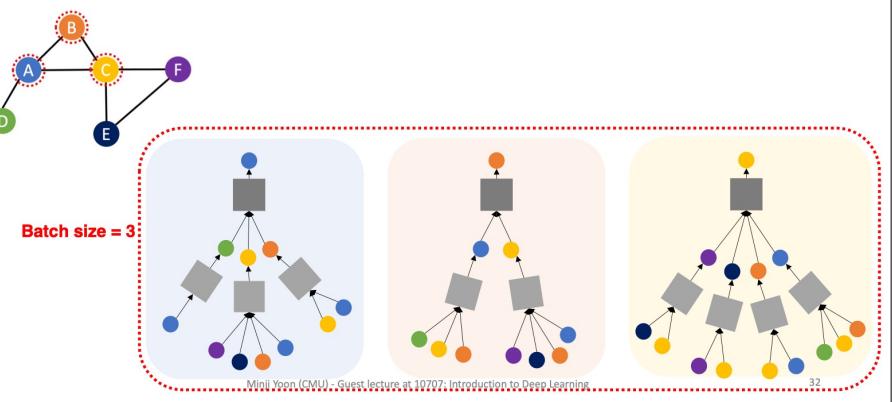
30

Computation graphs

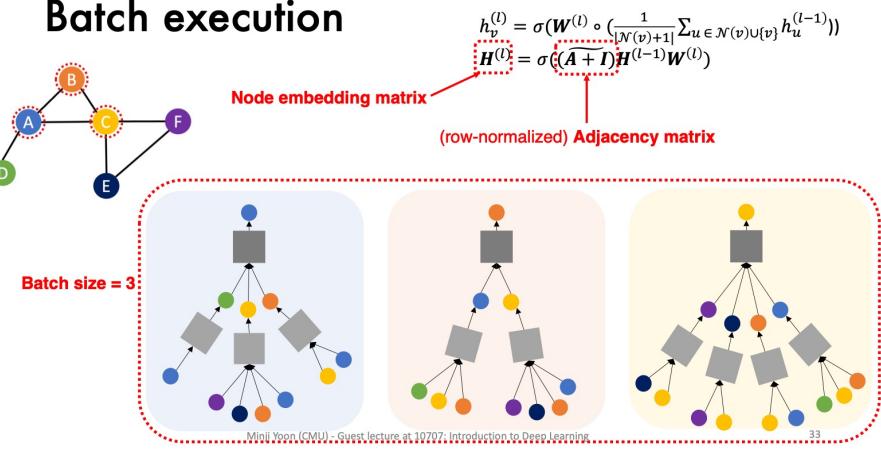


Batch execution

$$h_v^{(l)} = \sigma(\mathbf{W}^{(l)} \circ (\frac{1}{|\mathcal{N}(v)|+1} \sum_{u \in \mathcal{N}(v) \cup \{v\}} h_u^{(l-1)}))$$



Batch execution

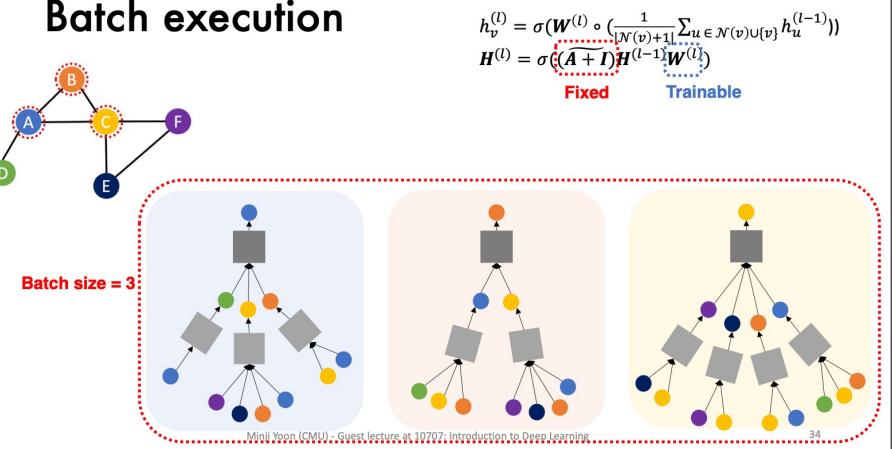


Batch execution

$$h_v^{(l)} = \sigma(\mathbf{W}^{(l)} \circ (\frac{1}{|\mathcal{N}(v)|+1} \sum_{u \in \mathcal{N}(v) \cup \{v\}} h_u^{(l-1)}))$$

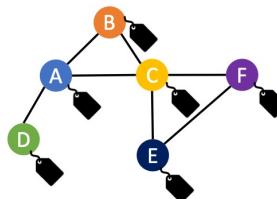
$$\mathbf{H}^{(l)} = \sigma((\mathbf{A} + \mathbf{I}) \mathbf{H}^{(l-1)} \mathbf{W}^{(l)})$$

Fixed Trainable



Downstream tasks

- Node-level prediction

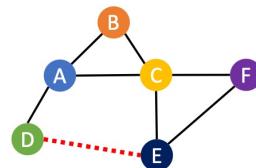


Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

35

Downstream tasks

- Node-level prediction
- Edge-level prediction



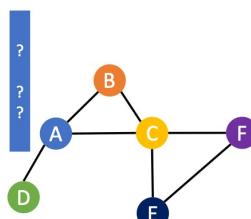
D and E are related enough
to be connected?

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

36

Downstream tasks

- Node-level prediction
- Edge-level prediction
- Attribute-level prediction

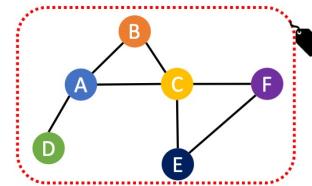


Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

37

Downstream tasks

- Node-level prediction
- Edge-level prediction
- Attribute-level prediction
- Graph-level prediction

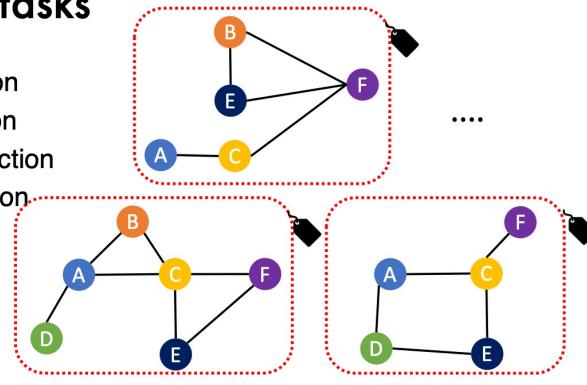


Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

38

Downstream tasks

- Node-level prediction
- Edge-level prediction
- Attribute-level prediction
- Graph-level prediction

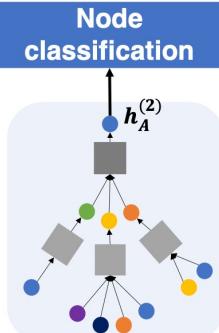


Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

39

Node-level prediction tasks

Node classification

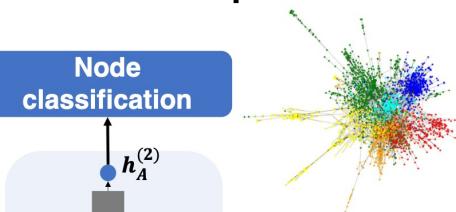


Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

41

Node-level prediction tasks

Node classification



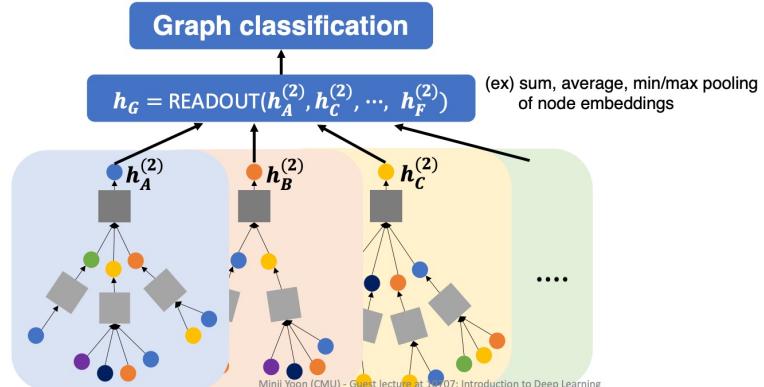
- Classify **papers** into topics on **citation networks**
- Cluster **posts** into subgroups on **Reddit networks**
- Classify **products** into categories on **Amazon co-purchase graphs**

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

42

Graph-level prediction tasks

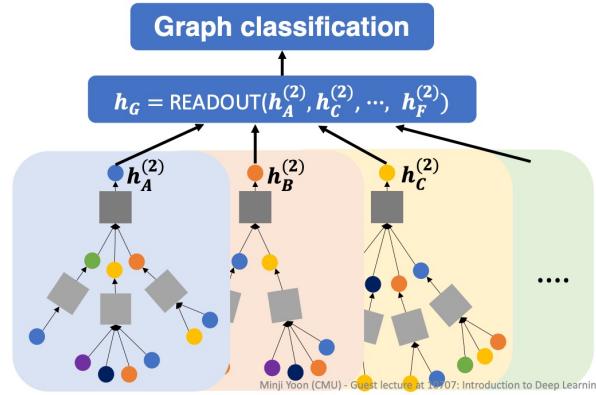
Graph classification



Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

43

Graph-level prediction tasks



- Predict **properties of a molecule (graph)** where nodes are atoms and edges are chemical bonds

44

So far, we have talked about..

1. Graph Neural Network

- Problem definition
- Skeleton
 - Aggregation operation
 - Transformation operation

2. Implementation

- Computation graph
- Batch execution

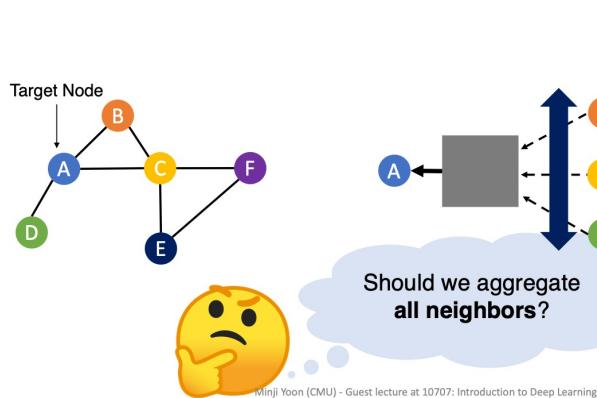
3. Downstream tasks

- Node-level prediction
- Graph-level prediction

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

46

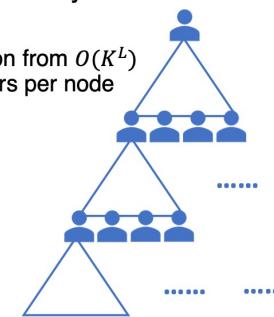
Graph Neural Networks - Width



48

Aggregation Width in GNNs

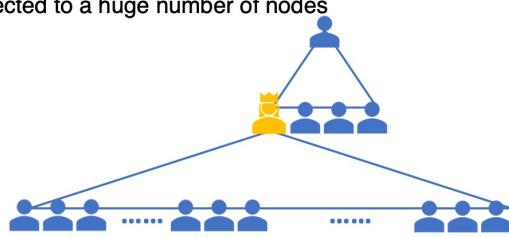
- If we aggregate all neighbors, GNNs have scalability issues
- Neighbor explosion
 - In L -layer GNNs, one node aggregates information from $O(K^L)$ nodes where K is the average number of neighbors per node



53

Aggregation Width in GNNs

- If we aggregate all neighbors, GNNs have scalability issues
- Neighbor explosion
 - Hub nodes who are connected to a huge number of nodes

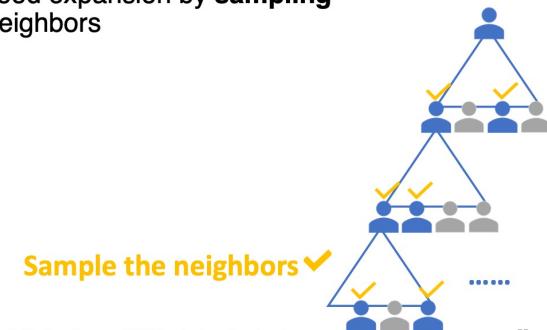


Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

54

Aggregation Width in GNNs

- Limit the neighborhood expansion by **sampling** a fixed number of neighbors



Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

55

Aggregation Width in GNNs

- Random sampling
 - Assign **same** sampling probabilities to all neighbors
 - *GraphSage*^[4]
- Importance sampling
 - Assign **different** sampling probabilities to all neighbors
 - *FastGCN*^[5], *LADIES*^[6], *AS-GCN*^[7], *GCN-BS*^[8], *PASS*^[9]

[4] Will Hamilton, et al. "Inductive representation learning on large graphs"

[5] Jie Chen, et al. "Fastgcn: fast learning with graph convolutional networks via importance sampling"

[6] Difan Zou, et al. "Layer-Dependent Importance Sampling for Training Deep and Large Graph Convolutional Networks"

[7] Wenbing Huang, et al. "Adaptive sampling towards fast graph representation learning"

[8] Ziqi Liu, et al. "Bandit Samplers for Training Graph Neural Networks"

[9] Minji Yoon, et al. "Performance-Adaptive Sampling Strategy Towards Fast and Accurate Graph Neural Networks"

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

56

Aggregation Width in GNNs

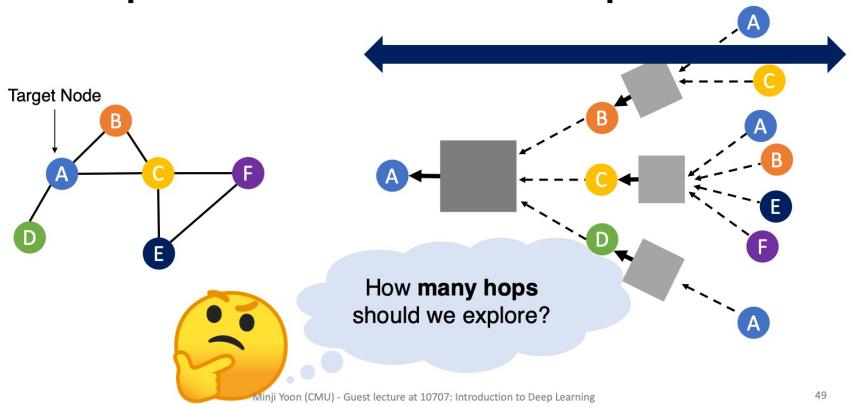
Method	Cora	Citeseer	Pubmed	AmazonC	AmazonP	MsCS	MsPhysics
FastGCN	0.582	0.496	0.569	0.480	0.542	0.520	0.638
AS-GCN	0.462	0.387	0.502	0.419	0.480	0.403	0.516
GraphSage	0.788	0.698	0.792	0.707	0.787	0.766	0.875
GCN-BS	0.788	0.693	0.809	0.736	0.800	0.780	0.887
PASS	0.821	0.715	0.858	0.757	0.855	0.884	0.934

- Node classification task on 7 different real-world graphs
- PASS outperforms all variance-minimizing methods by up to 10.4%

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

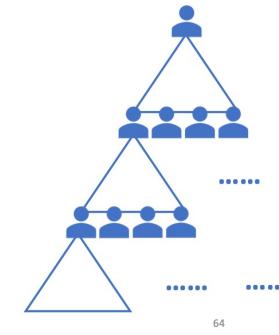
58

Graph Neural Networks - Depth



Aggregation Depth in GNNs

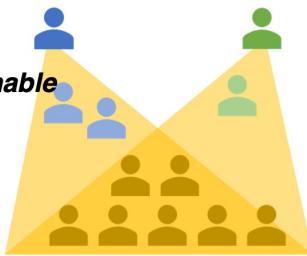
- When we increase the depth L more than this, GNNs face neighbor explosion $O(K^L)$
 - Over-smoothing
 - Over-squashing



Aggregation Depth in GNNs

Over-smoothing^[10]

- When GNNs become deep, nodes share many neighbors
- Node embeddings become *indistinguishable*



[10] Qimai Li, et al. "Deeper Insights into Graph Convolutional Networks for Semi-Supervised Learning"

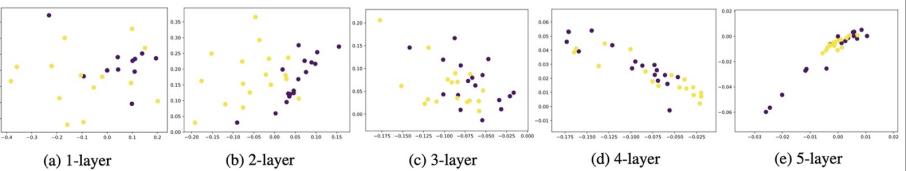
Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

65

Aggregation Depth in GNNs

Over-smoothing^[10]

- Node embeddings of Zachary's karate club network with GNNs



[10] Qimai Li, et al. "Deeper Insights into Graph Convolutional Networks for Semi-Supervised Learning"

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

66

Aggregation Depth in GNNs

Mitigate over-smoothing

PairNorm^[11]

- Keep total pairwise squared distance (TPSD) **constant** across layers
- Push away pairs that are not connected

$$\text{TPSD}(\vec{x}) = \sum_{(i,j) \in \mathcal{E}} \|\vec{x}_i - \vec{x}_j\|_2^2 + \sum_{(i,j) \notin \mathcal{E}} \|\vec{x}_i - \vec{x}_j\|_2^2 = C$$

Connected pairs Disconnected pairs

[11] Lingxiao Zhao, et al. "PAIRNORM: TACKLING OVERSMOOTHING IN GNNs"

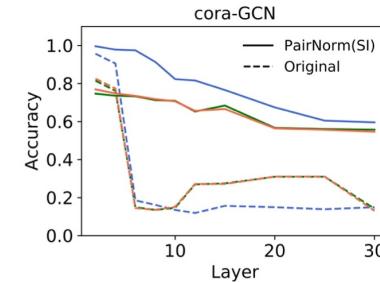
Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

67

Aggregation Depth in GNNs

Mitigate over-smoothing

PairNorm^[11]

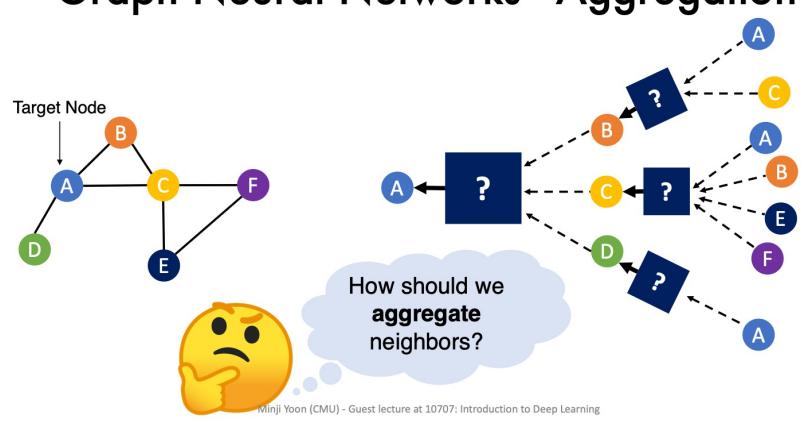


[11] Lingxiao Zhao, et al. "PAIRNORM: TACKLING OVERSMOOTHING IN GNNs"

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

68

Graph Neural Networks - Aggregation



Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

50

Aggregation strategy in GNNs

GCN^[1]

- Average embeddings of neighboring nodes

[1] Kipf, Thomas N., et al. "Semi-supervised classification with graph convolutional networks."

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

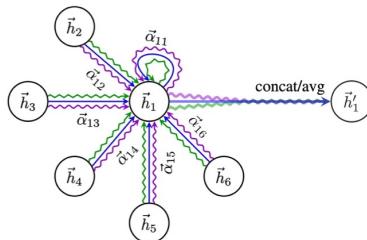
76

Aggregation strategy in GNNs

- GAT^[14]

- Different weights to different nodes in a neighborhood
- Multi-head attention

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(\vec{a}^T [\mathbf{W}\vec{h}_i || \mathbf{W}\vec{h}_j]))}{\sum_{k \in \mathcal{N}_i} \exp(\text{LeakyReLU}(\vec{a}^T [\mathbf{W}\vec{h}_i || \mathbf{W}\vec{h}_k]))}$$



[14] Petar Veličković, et al. "GRAPH ATTENTION NETWORKS."

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

77

Aggregation strategy in GNNs

In each layer l :

Aggregate over neighbors

$$m_v^{(l-1)} = f^{(l)}(h_v^{(l-1)}, \{h_u^{(l-1)} : u \in \mathcal{N}(v)\})$$

Core part of GNNs

Transform messages

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

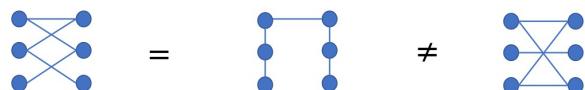
Any neural network module can fit in
1-layer MLP is commonly used

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

78

Aggregation strategy in GNNs

- How powerful are Graph Neural Networks?^[2]
- Metric
 - Graph-level prediction task
 - Can a GNN model distinguish two non-isomorphic graphs?



[2] Keyulu Xu, et al. "HOW POWERFUL ARE GRAPH NEURAL NETWORKS?"

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

84

Aggregation strategy in GNNs

- Can we make more powerful GNNs?
 - Very active area, with many open problems

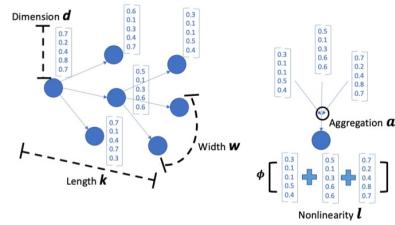
Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

87

Neural Architecture Search for GNNs

- Which *width, depth, and aggregation strategy* are proper for a given graph and task?

Width?
Depth?
Aggregation?



Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

96

Neural Architecture Search for GNNs

- Finding proper *width, depth, and aggregation strategy* for a given graph and task **automatically**^[1,2,3]

Here is the GNN you requested



[23] Minji Yoon., et al. "Autonomous Graph Mining Algorithm Search with Best Speed/Accuracy Trade-off"
[24] Kaixiong Zhou, et al. "Auto-GNN: Neural Architecture Search of Graph Neural Networks"
[25] Yang Gao, et al. "GraphNAS: Graph Neural Architecture Search with Reinforcement Learning"

Minji Yoon (CMU) - Guest lecture at 10707: Introduction to Deep Learning

97