Google AI

USC University of Southern California
Information Sciences Institute

N-GCN: Multi-scale Graph Convolution for Semi-supervised Node Classification

Sami Abu-El-Haija, Amol Kapoor, Bryan Perozzi, ***Joonseok Lee***

Code at: **GitHub**/samihaija/mixhop

Review: Graph ConvNets (GCNs) for` Semi-supervised Node Classification

- Given Graph (Nodes + edges + features); some nodes are labeled.
  - Repeatedly propagate latent features along edges
  - Output label predictions for unlabeled nodes.
- Popularized by [Kipf & Welling, ICLR 2017]

Graph ConvNets (GCNs) for Semi-supervised Node Classification

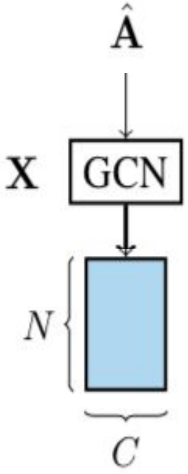
[Kipf & Welling, ICLR 2017]

😃 Inductive

😃 O(#edges) computation

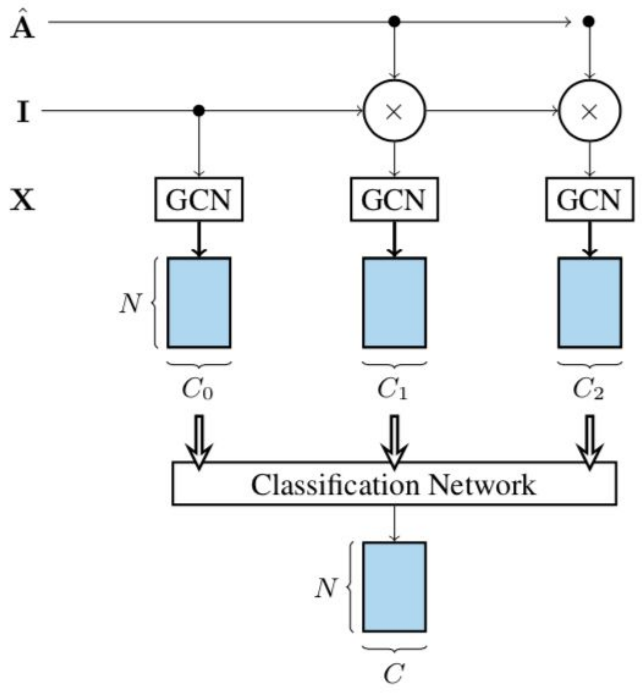
😢 No benefit in more depth (including further nodes)

😢 Affected by label noise



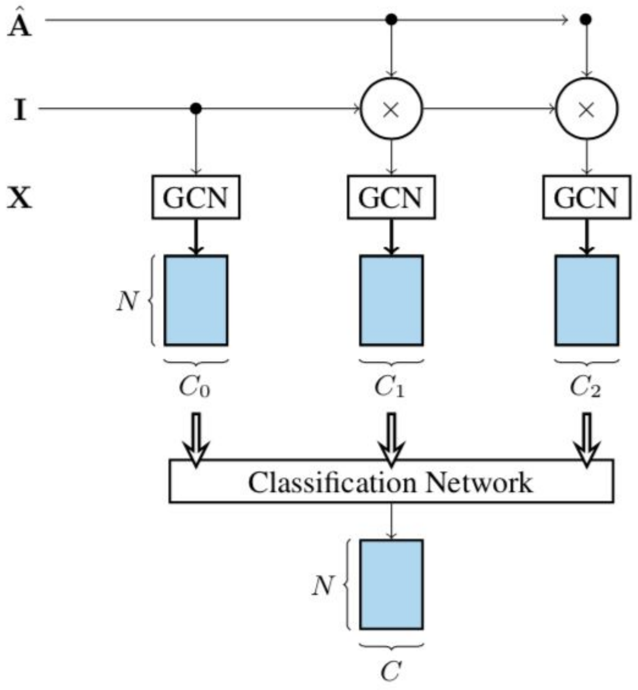
Our Model: Network of GCNs (N-GCN)

- Combines Random Walks and GCNs
- Instantiates multiple GCNs, each on a different power of adjacency matrix
- Combines their output through classification net
- General: can wrap other Graph Nets (GraphSAGE, GAT, …)
- Efficient: no explicit higher power matrix calculation



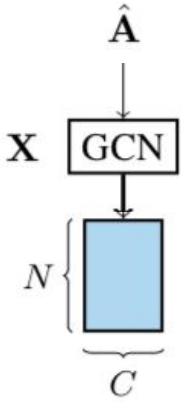
Our Model: Network of GCNs (N-GCN)

- 😀 Inductive
- 😀 O(#edges) computation
- 😀 Incorporates far nodes! (shallow GCN fed A^i)
- 😀 Resilient to label noise

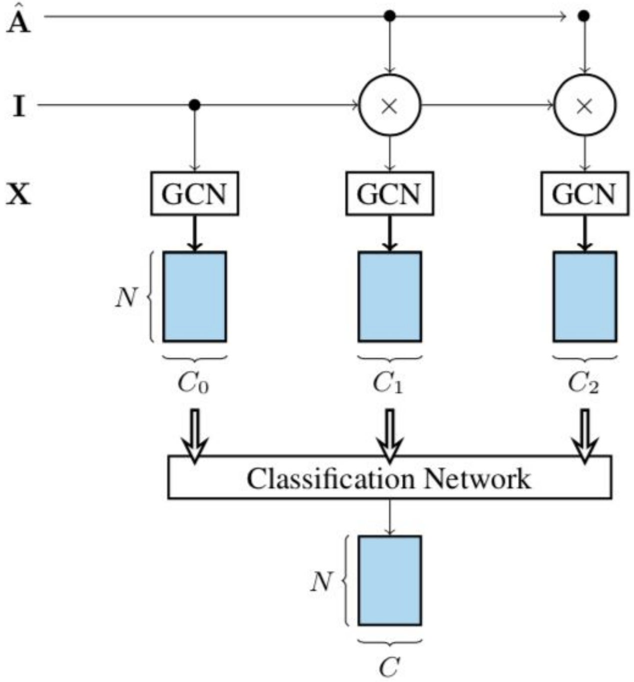


Our Model is General

[Kipf & Welling]



Ours



Can also wrap other Graph Nets (GraphSAGE, GAT, others…)

Transductive Node Classification Tasks

Method (Transductive)	Citeseer	Cora	Pubmed
GCN (Kipf and Welling, 2017)	70.3	81.5	79.0
DCNN (our implementation)	71.1	81.3	79.3
GCN (our implementation)	71.2	81.0	78.8
SAGE (our implementation)	63.5	77.4	77.6
N-GCN (ours)	**72.2**	**83.0**	**79.5**
N-SAGE (ours)	71.0	81.8	79.4

Inductive Node Classification Tasks

Method (Inductive)	PPI
SAGE-LSTM (Hamilton et al., 2017)	61.2
SAGE (Hamilton et al., 2017)	60.0
DCNN (our implementation)	44.0
GCN (our implementation)	46.2
SAGE (our implementation)	59.8
N-GCN (ours)	46.8
N-SAGE (ours)	**65.0**

Handles Noise by shifting attention to further nodes

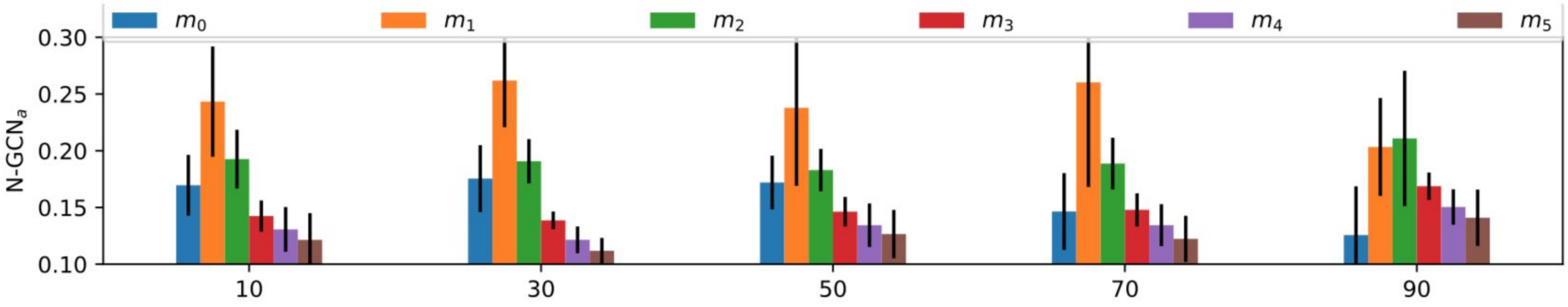


Figure 3: Attention weights (m) for N-GCN$_a$ when trained with feature removal perturbation on the Cora dataset. Removing features shifts the attention weights to the right, suggesting the model is relying more on long range dependencies.

For more results & details please visit our poster **#310** on Thursday (7/25).