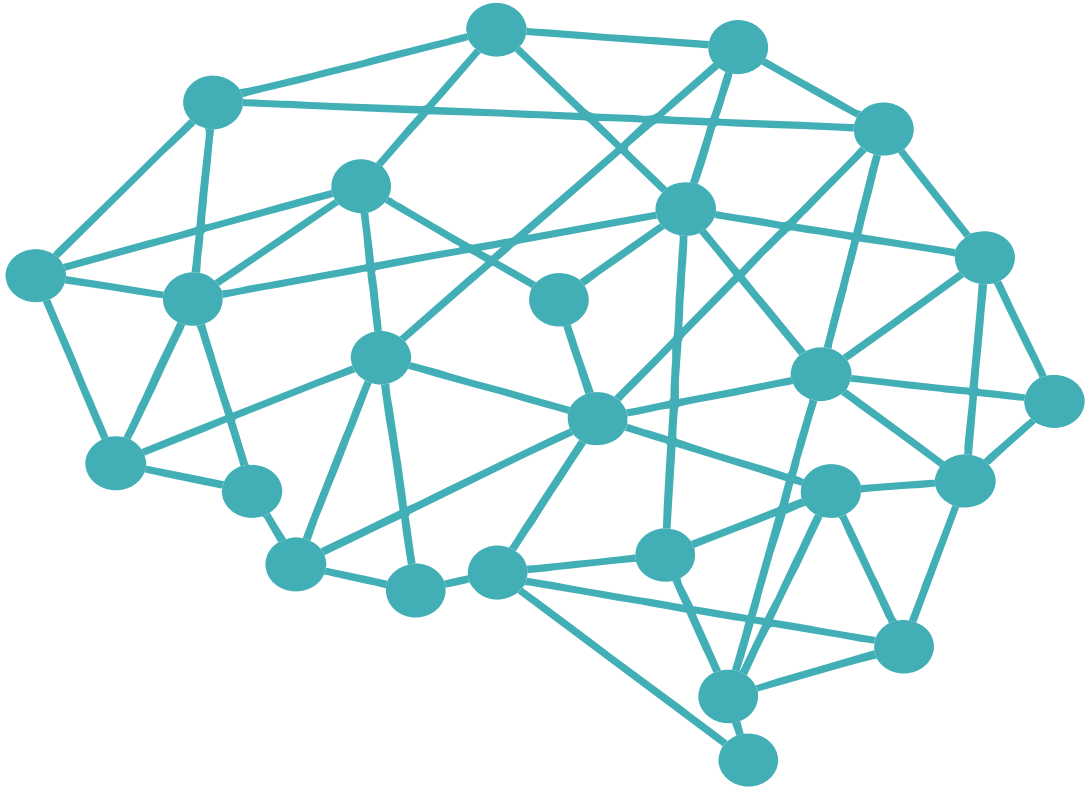
Class Participation Assignment

013- Geometric Graph Convolutional Network

Presented by,
Aiman Younas- MSDS20087
Qazi Danish Ayub -MSDS20075
Muhammad Mubashir Ali -MSDS20085
Hira Saleem - MSDS20102
Mahnoor Imran -MSDS20088



Literature Review

[1]	Introduces Local aggregation framework for GNNs that allows non-local aggregation using convolution and efficient attention-guided sorting.Depending on the disassortative graphs in hand, we can build different non-local GNNs with either MLP or GNNs as the local embedding step. That helps to quickly create non-local GNNs with minimal computational overhead
[2]	Spectral clustering (SC) is implemented into GNNs that can be used to capture long-range graph dependencies.
[3]	Developed a neural network structure known as GIN(Graph Isomorphism Network) The theoretical framework explains that each node in GNN updates its features to capture the structure and features of other nodes. Additionally showed that the graph neural networks are as strong as the Weisfeiler lehman test.
[4]	Proposed modification in the Graph Convolutional Network architecture to utilize edge adjacency information by non-backtracking operator of graph. They named the resulting model as the LGNN(Linear Graph Neural Network). This architecture can also be implemented to the other node wise classification problems. To test this model they used community detection because it has rich theoretical literature.
[5]	To address the issue in generic graphs they proposed the learnable graph convolutional layer(LGCL). In order to convert the graph data into the grid type structured data LGCL is used so that regular convolutional operations can be used on the generic graphs. They also introduced the subgraph training method to reduce the use of excessive memory and resource requirements by methods on graphs.

References

[1] Non-Local Graph Neural Networks, Meng Liu* Zhengyang Wang* Shuiwang Ji
Department of Computer Science & Engineering Texas A&M University College Station, TX

[2] GCN-SL: Graph Convolutional Networks with Structure Learning for Graphs under Heterophily, Mengying Jiang 1 Guizhong Liu 1 Yuanchao Su 2 Xinliang Wu 1-2021

[3] Keyulu Xu, Weihua Hu, Jure Leskovec, and Stefanie Jegelka. How powerful are graph neural networks? In International Conference on Learning Representations (ICLR), 2019

[4] Zhengdao Chen, Lisha Li, and Joan Bruna. Supervised community detection with line graph neural networks. In International Conference on Learning Representations (ICLR), 2019.

[5] Hongyang Gao, Zhengyang Wang, and Shuiwang Ji. Large-scale learnable graph convolutional networks. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data

Main idea of the Project

To address the Fundamental weaknesses of Message Passing Neural Network in which the aggregators limit their ability to represent graph-structured data.

Some useful links for the demos and animations

- https://paperswithcode.com/paper/geom-gcn-geometric-graph-convolutional-1/review/
- https://openreview.net/forum?id=S1e2agrFvS
- https://github.com/graphdml-uiuc-jlu/geom-gcn
- https://colab.research.google.com/drive/1NK5-bI0bReq7foT5neMMglqWlCK4IyKq?usp=sharing
- https://www.researchgate.net/publication/339251875_Geom-GCN_Geometric_Graph_Convolutional_Networks

GEOM-GCN: GEOMETRIC GRAPH CONVOLUTIONAL NETWORKS

ABSTRACT

- Message-passing neural networks (MPNNs) have been effectively employed in a wide range of real-world applications.
- Two fundamental shortcomings of MPNN aggregators limit their capacity to represent graph-structured data.
- The proposed aggregation approach is permutation-invariant and includes three modules.

Introduction

Message-passing neural networks (MPNNs)

Each node sends its feature representation, a “message”, to the nodes in its neighborhood; and then updates its feature representation by aggregating all “messages” received from the neighborhood. By adopting permutation-invariant aggregation functions

Problem Statement

1. The aggregators lose the structural information of nodes in neighborhoods.
2. The aggregators lack the ability to capture long-range dependencies in disassortative graphs

Methodology

GEOMETRIC AGGREGATION SCHEME:

- The proposed aggregation scheme is permutation-invariant and consists of three modules :
- Node embedding,
- Structural Neighborhood,
- Bi-level aggregation.

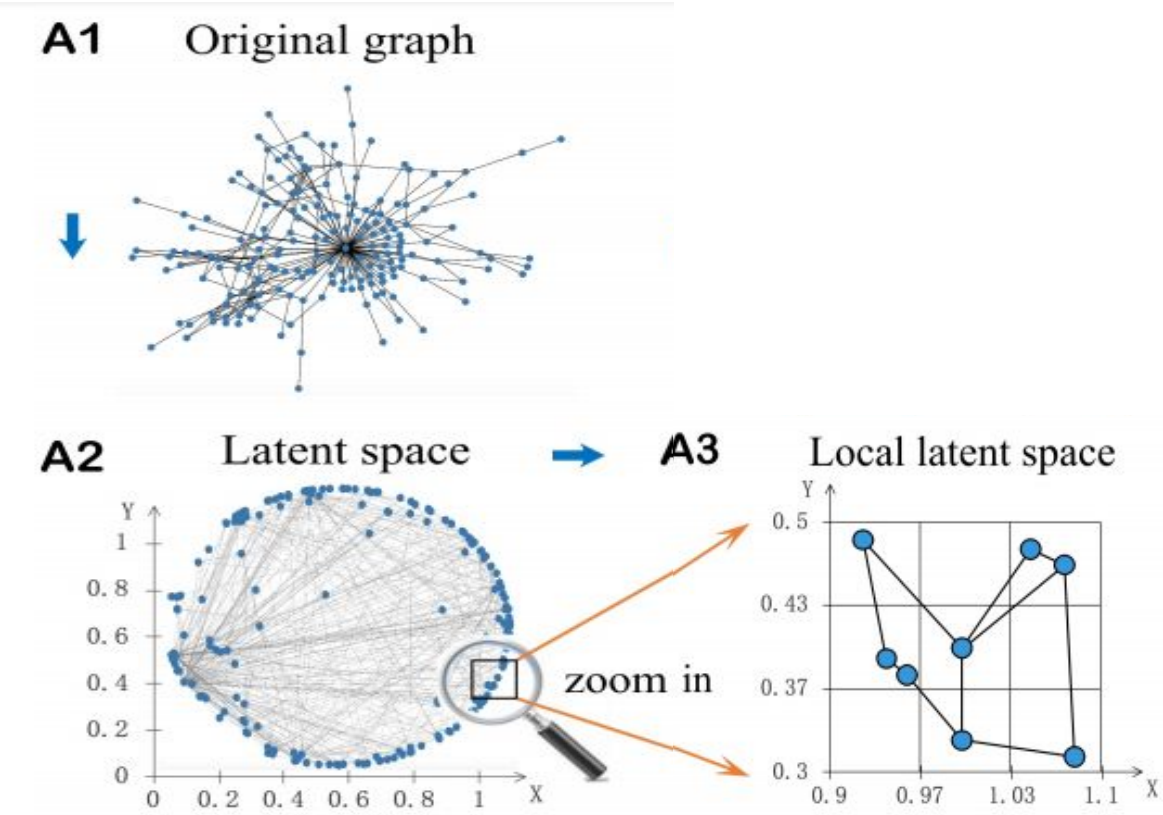
Node Embedding

$G = (V, E)$ $\quad e \in E$

$v \in V$

$f : v \longrightarrow zv$

$\mathbf{zv \in R^d}$



Structural Neighborhood

$$N(v) = (\{Ng(v), Ns(v)\}, \tau)$$

The study uses the concept of structural neighborhood in which it defines the neighbors of a node in a graph Ng(v) and latent space Ns(v) through embedding methods and a relational operator on nodes τ.
In space mapping to create a second kind of neighborhood in latent space: they compare the distance between neighbouring nodes with the parameter ρ , where the parameter's magnitude rises from 0 to the number of neighborhoods in graph space. For each node v, the researchers used a two-stage convolution approach.

Structural Neighborhood Cont'd

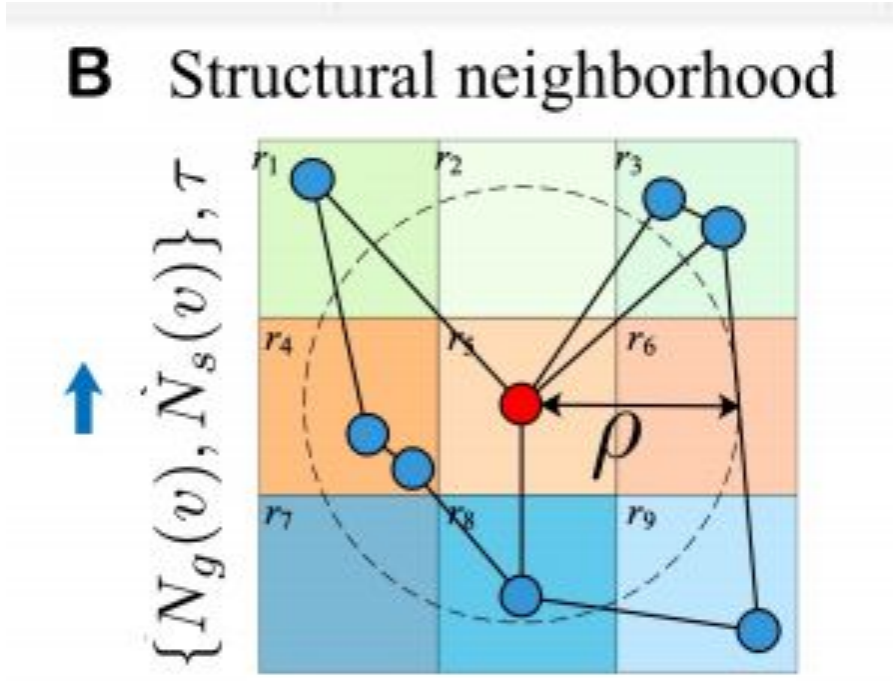
The equation for neighborhood in the graph is given as:

$Ng(v) = \{u | u \in V, (u, v) \in E\}$

The equation for neighborhood in the latent space is given as :

$Ns(v) = \{u | u \in V, d(zu, zv) < \rho\}$

$\tau : (zv, zu) \rightarrow r \in R$



Structural Neighborhood Cont'd

R= {left upper, right upper, left lower, right lower}, and a T (zv, zu,
the relationship "upper" indicates the node nearer to the origin and thus lie in a higher level in a hierarchical graph.

$\tau(z_v, z_u)$	$z_v[0] > z_u[0]$	$z_v[0] < z_u[0]$
$z_v[1] < z_u[1]$	left upper	right upper
$z_v[1] > z_u[1]$	left lower	right lower

Table 1: The relationship operator

Bi-level Aggregation

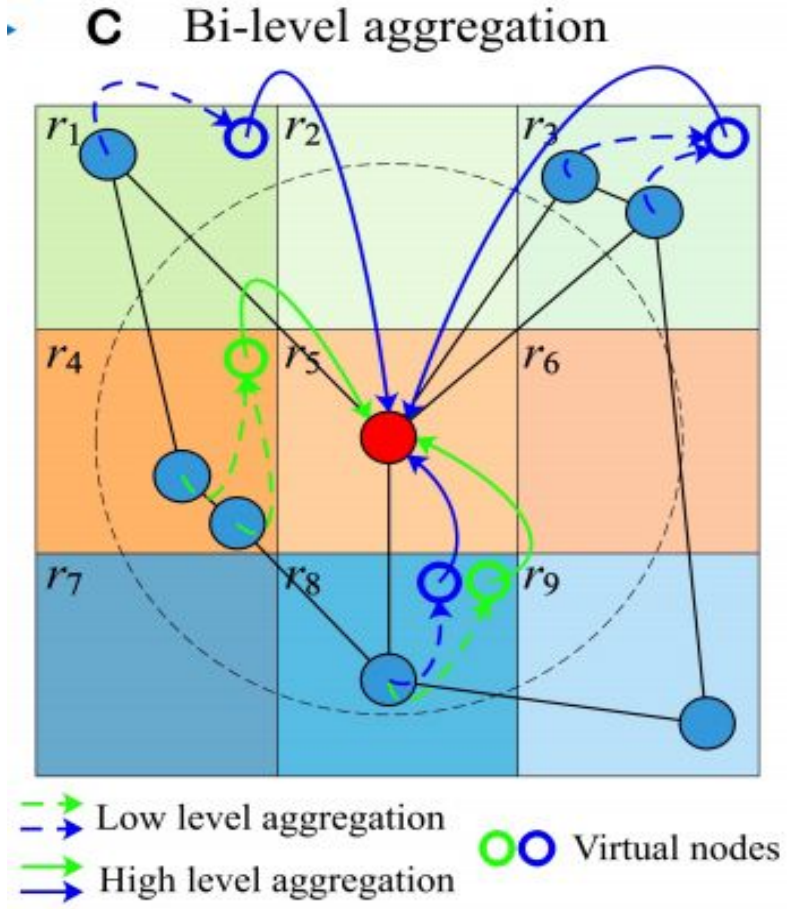
Low level Aggregation: The aggregation function p aggregates the hidden characteristics of nodes that are in the same neighbourhood I and have the same geometric relationship r to a virtual node at the low level.

High level Aggregation: The properties of virtual nodes are further aggregated at the high level by function q. The attributes of virtual nodes and the identification of virtual nodes are both inputs to function q. That is, q can be a function that takes an ordered item as input and returns an ordered object.

$$e_{(i,r)}^{v,l+1} = p(\{h_u^l | u \in N_i(v), \tau(z_v, z_u) = r\}), \forall i \in \{g, s\}, \forall r \in R \quad \text{(Low-level aggregation)}$$

$$m_v^{l+1} = \underset{i \in \{g,s\}, r \in R}{q}((e_{(i,r)}^{v,l+1}, (i,r))) \quad \text{(High-level aggregation)}$$

$$h_v^{l+1} = \sigma(W_l \cdot m_v^{l+1}) \quad \text{(Non-linear transform)}$$



Dataset Statistics

Dataset	Cora	Cite.	Pubm.	Cham.	Squi.	Actor	Corn.	Texa.	Wisc.
# Nodes	2708	3327	19717	2277	5201	7600	183	183	251
# Edges	5429	4732	44338	36101	217073	33544	295	309	499
# Features	1433	3703	500	2325	2089	931	1703	1703	1703
# Classes	7	6	3	5	5	5	5	5	5

Table 2: Datasets statistics

Experiments

Dataset	Cora	Cite.	Pubm.	Cham.	Squi.	Actor	Corn.	Texa.	Wisc.
GCN	85.77	73.68	88.13	28.18	23.96	26.86	52.70	52.16	45.88
GAT	**86.37**	74.32	87.62	42.93	30.03	28.45	54.32	58.38	49.41
Geom-GCN-I	85.19	**77.99**	**90.05**	60.31	33.32	29.09	56.76	57.58	58.24
Geom-GCN-P	84.93	75.14	88.09	**60.90**	**38.14**	**31.63**	**60.81**	**67.57**	**64.12**
Geom-GCN-S	85.27	74.71	84.75	59.96	36.24	30.30	55.68	59.73	56.67

Table 3: Mean Classification Accuracy (Percent)

Ablation Study On Contributions From Two Neighborhoods

Dataset β	Cora 0.83	Cite. 0.71	Pumb. 0.79	Cham. 0.25	Squi. 0.22	Actor 0.24	Corn. 0.11	Texa. 0.06	Wisc. 0.16
Geom-GCN-I-g	86.26 ↑0.48	**80.64** ↑6.96	**90.72** ↑2.59	**68.00** ↑39.82	**46.01** ↑22.05	31.96 ↑4.04	65.40 ↑12.70	72.51 ↑21.35	68.23 ↑22.35
Geom-GCN-I-s	77.34 ↓8.34	72.22 ↓1.46	85.02 ↓3.11	61.64 ↑33.46	37.98 ↑14.02	30.59 ↑2.67	62.16 ↑9.46	60.54 ↑8.38	64.90 ↑19.01
Geom-GCN-P-g	86.30 ↑0.52	75.45 ↑1.76	88.40 ↑0.27	63.07 ↑34.89	38.41 ↑14.45	31.55 ↑3.63	64.05 ↑11.35	73.05 ↑21.89	69.41 ↑23.53
Geom-GCN-P-s	73.14 ↓12.63	71.65 ↓2.04	86.95 ↓1.18	43.20 ↑15.02	30.47 ↑6.51	**34.59** ↑6.67	**75.40** ↑22.70	**73.51** ↑21.35	**80.39** ↑34.51
Geom-GCN-S-g	**87.00** ↑1.23	75.73 ↑2.04	88.44 ↑0.31	67.04 ↑38.86	44.92 ↑20.96	31.27 ↑3.35	67.02 ↑14.32	71.62 ↑19.46	69.41 ↑23.52
Geom-GCN-S-s	66.92 ↓18.85	66.03 ↓7.65	79.41 ↓8.72	49.21 ↑21.03	31.27 ↑7.31	30.32 ↑2.40	62.43 ↑9.73	63.24 ↑11.08	64.51 ↑18.63

Table 4: Mean Classification Accuracy (Percent)

Analysis Of Embedding Space Combination

Dataset	Cora	Cite.	Pubm.	Cham.	Squi.	Actor	Corn.	Texa.	Wisc.
Geom-GCN-IP	85.13	**79.41**	**90.49**	65.77	**45.49**	**31.94**	**60.00**	66.49	62.75
Geom-GCN-PI	85.09	75.08	85.64	59.19	32.65	29.16	58.11	58.11	58.63
Geom-GCN-IS	84.51	77.83	88.66	58.40	35.29	29.41	54.32	57.57	57.65
Geom-GCN-SI	85.31	75.50	85.52	62.13	32.57	28.97	57.30	60.00	55.10
Geom-GCN-PS	**85.65**	74.84	84.96	56.34	28.27	29.53	58.11	62.43	60.59
Geom-GCN-SP	85.43	75.71	88.00	**65.81**	44.53	31.16	58.38	**67.84**	**65.10**

Table 5: Mean Classification Accuracy

Individual Contribution

Aiman, Danish and Mahnoor

Finding the paper's pertinent aspects, comprehending the code's operation in order to better comprehend the methodology

Mubashir and Hira

To uncover the distinction in our work, we looked for related studies. Assisting in the comprehension of the paper's key terms