Chongqing University of Technology

ATAI
Advanced Technique of Artificial Intelligence

Dual Space Graph Contrastive Learning

Haoran Yang∗
haoran.yang-2@student.uts.edu.au
University of Technology Sydney
Sydney, NSW, Australia

Hongxu Chen∗†
hongxu.chen@uts.edu.au
University of Technology Sydney
Sydney, NSW, Australia

Shirui Pan
shirui.pan@monash.edu
Monash University
Melbourne, VIC, Australia

Lin Li
cathylilin@whut.edu.cn
Wuhan University of Technology
Wuhan, Hubei, China

Philip S. Yu
psyu@cs.uic.edu
University of Illinois at Chicago
Chicago, Illinois, U.S.A

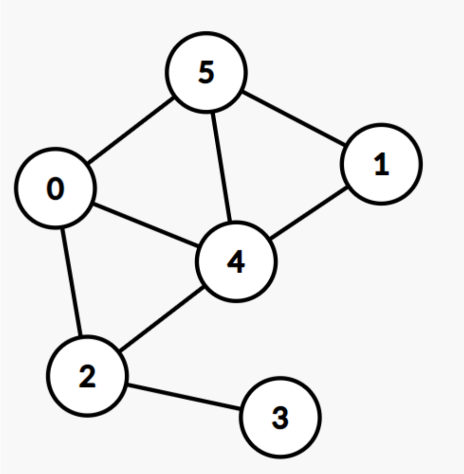
Guandong Xu†
guandong.xu@uts.edu.au
University of Technology Sydney
Sydney, NSW, Australia

—— WWW 2022

2022.12.01 • ChongQing

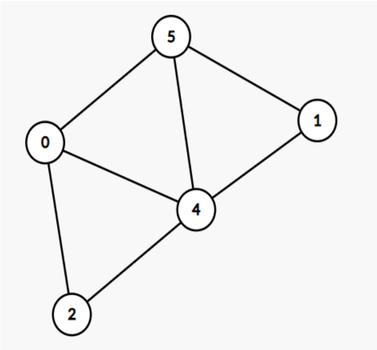
Reported by JiaWei Cheng

Introduction

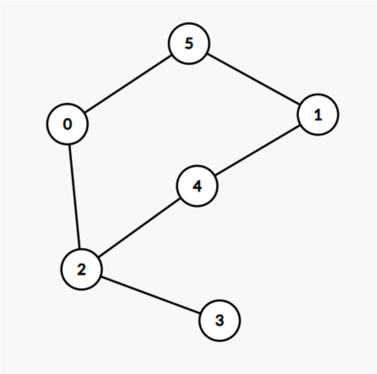


How to generate unique and informative views for graph contrastive learning?

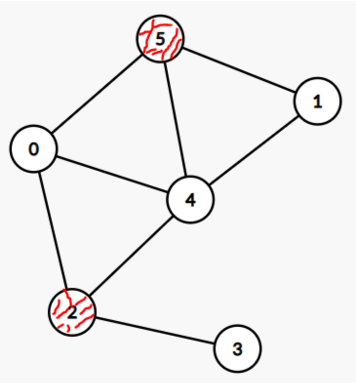
Node dropping



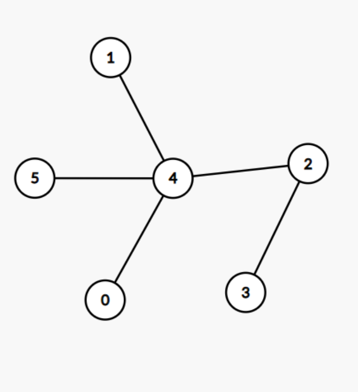
Edge perturbation



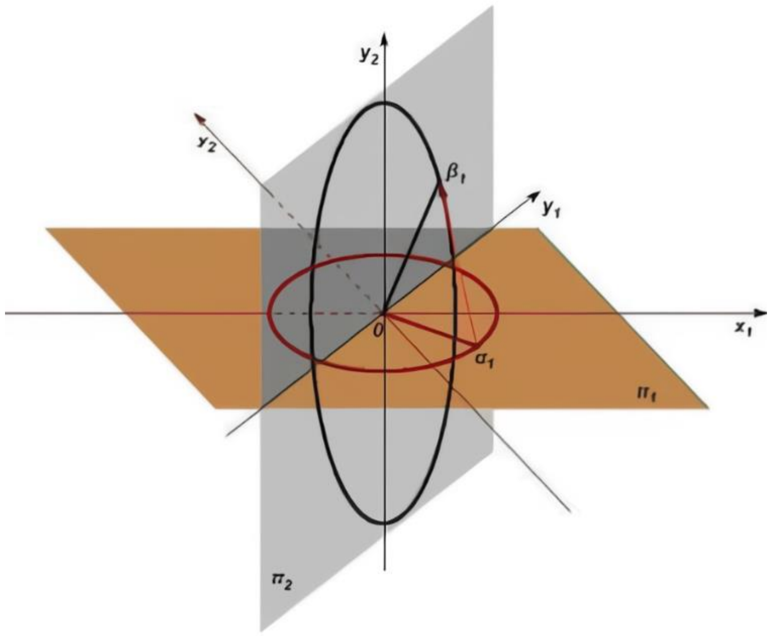
Attribute masking



Subgraph



Introduction

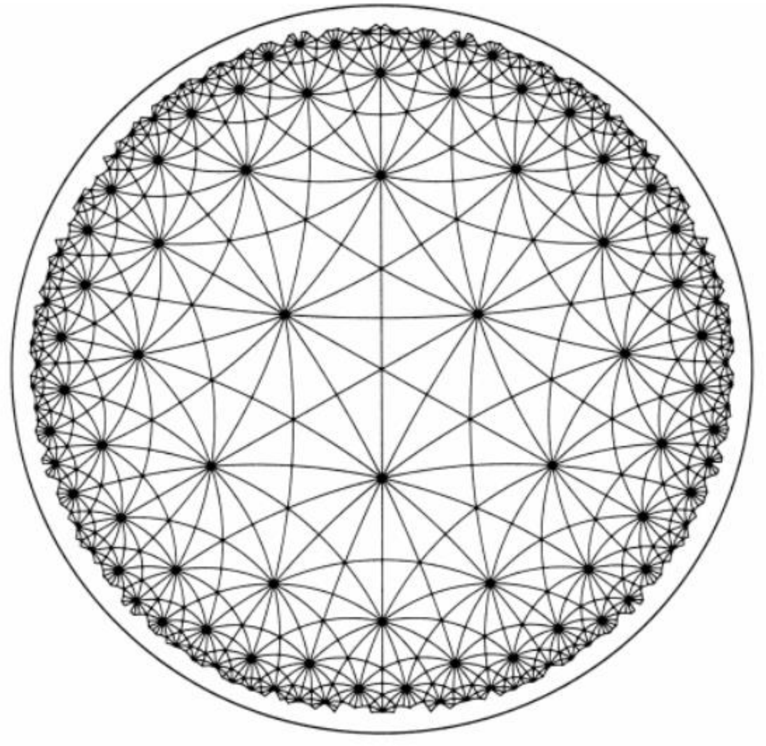


Euclidean Space

Advantages of Euclidean Space:

Compared to the hyperbolic space, vector calculation in Euclidean space is more efficient.

Introduction



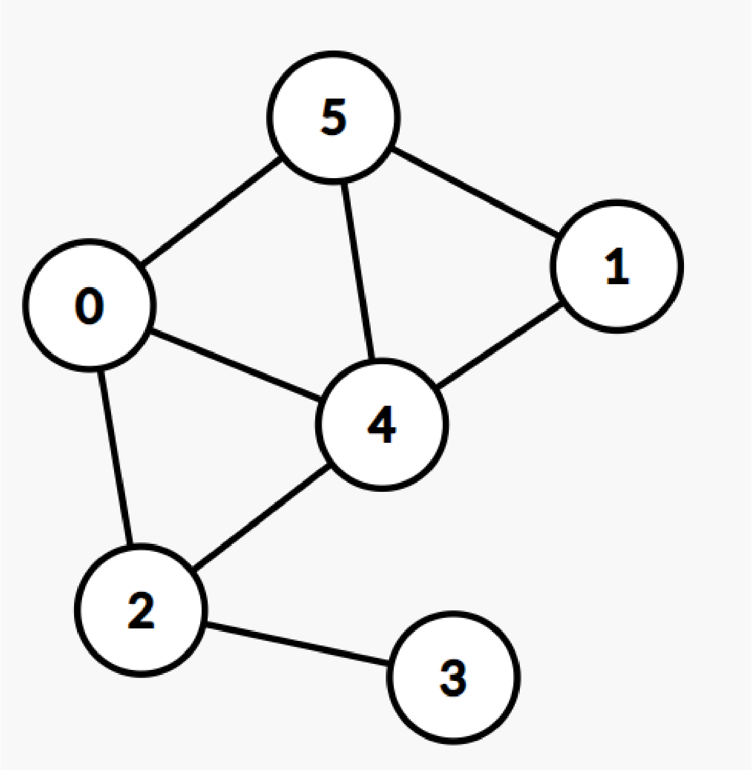
Poincaré ball model

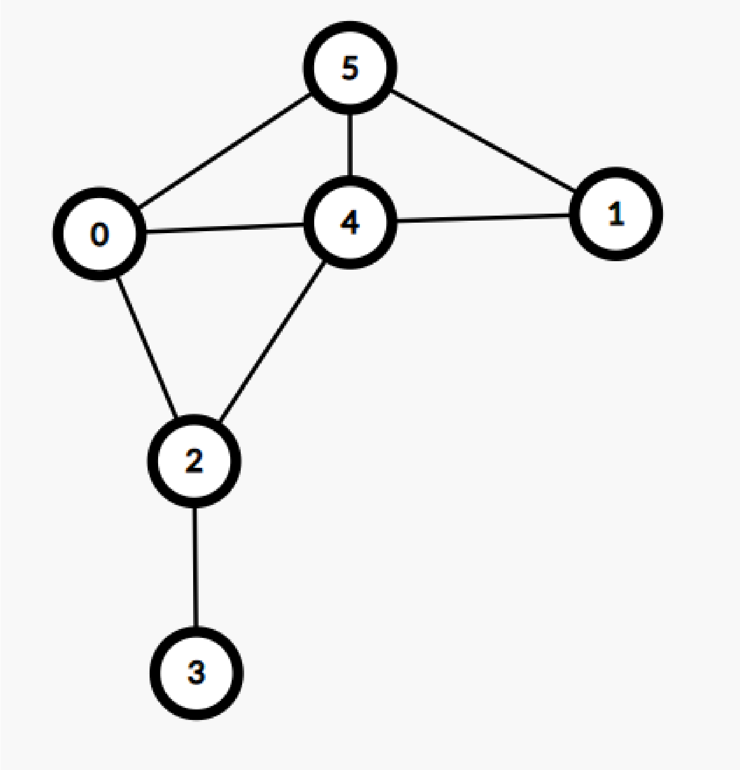
Advantages of Hyperbolic Space:

The first is that its own space size is very different from Euclidean space.

The second advantage is that the hyperbolic space is more capable of capturing hierarchical structures exhibited in graph data.

Introduction





Preliminaries

$$\mathbb{D} = \{(x_1, ..., x_n) : x_1^2 + \cdots + x_n^2 < \frac{1}{c}\}. \quad (1)$$

$$d_{\mathbb{D}}(\mathbf{u}, \mathbf{v}) = \frac{1}{arcosh(1 + \frac{2||\mathbf{u}-\mathbf{v}||^2}{(1-||\mathbf{u}||^2)(1-||\mathbf{v}||^2)})} \quad (2)$$

$$exp_{\mathbf{o}}^c(\mathbf{t}) = tanh(\sqrt{c}||\mathbf{t}||)\frac{\mathbf{t}}{\sqrt{c}||\mathbf{t}||}, \quad (3)$$

the exponential mapping $exp_{\mathbf{o}}^c : \mathcal{T}_{\mathbf{o}}\mathbb{D}_c \rightarrow \mathbb{D}_c$

$$log_{\mathbf{o}}^c(\mathbf{u}) = artanh(\sqrt{c}||\mathbf{u}||)\frac{\mathbf{u}}{\sqrt{c}||\mathbf{u}||}. \quad (4)$$

the logarithmic mapping $log_{\mathbf{o}}^c : \mathbb{D}_c \rightarrow \mathcal{T}_{\mathbf{o}}\mathbb{D}_c$

Preliminaries

$$\mathbf{y} = \sigma(\mathbf{W} \cdot \mathbf{u} + \mathbf{b}). \tag{5}$$

$$\mathbf{W} \otimes \mathbf{u} = exp_{\mathbf{o}}^{c}(\mathbf{W} \cdot log_{\mathbf{o}}^{c}(\mathbf{u})), \tag{6}$$

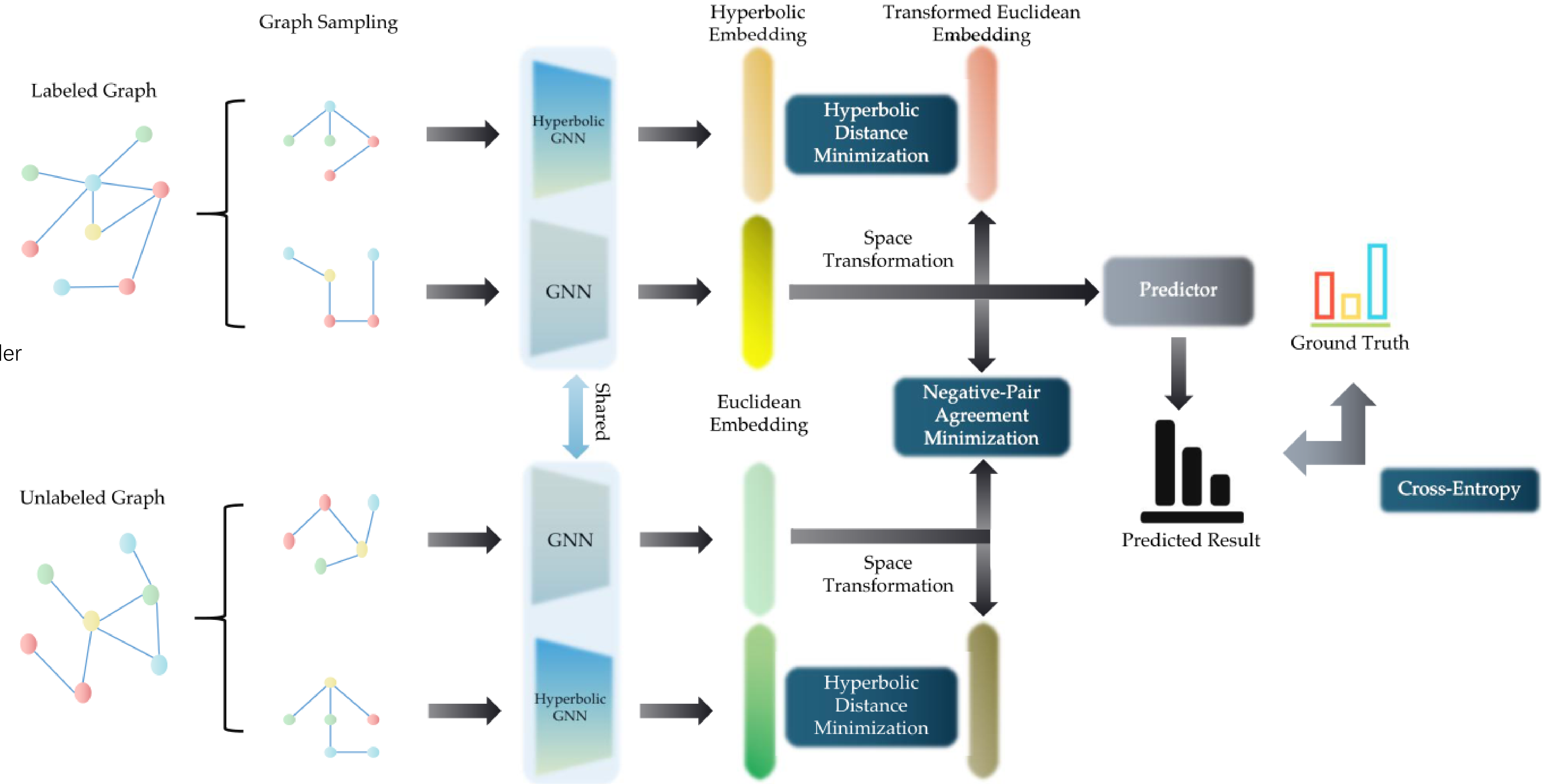
$$\mathbf{u} \oplus \mathbf{b} = exp_{\mathbf{o}}^{c}(log_{\mathbf{o}}^{c}(\mathbf{u}) + \mathbf{b}), \tag{7}$$

$$\mathbf{y} = exp_{\mathbf{o}}^{c}(\sigma(log_{\mathbf{o}}^{c}(\mathbf{W} \otimes \mathbf{u} \oplus \mathbf{b}))). \tag{8}$$

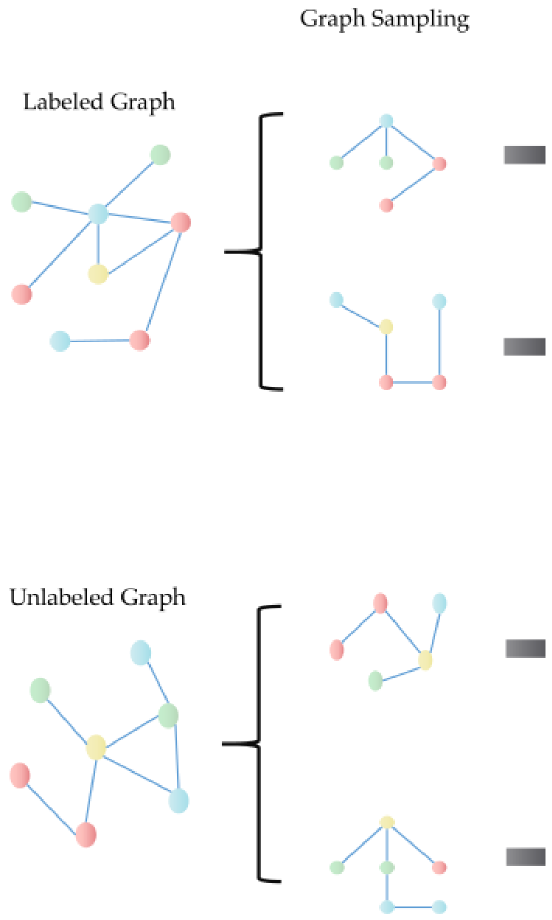
Overview

CommunityStructureExpansionSampler

DiffusionSampler

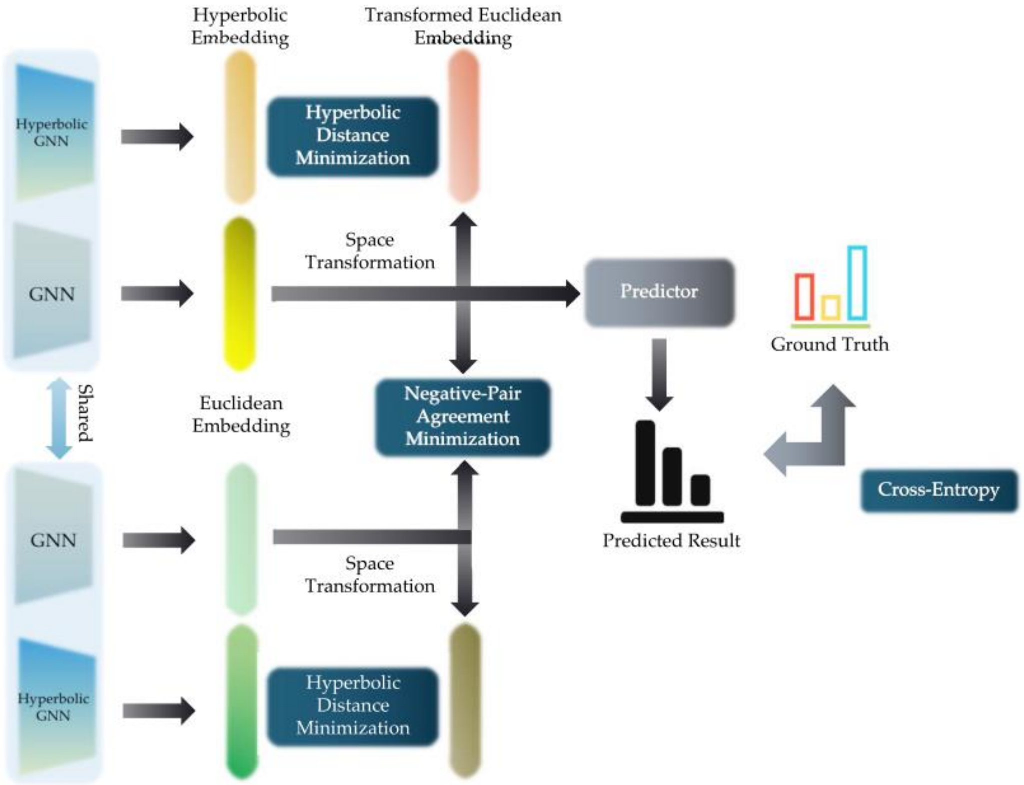


Method



$$SG = S(\mathcal{G}) \tag{9}$$

Method



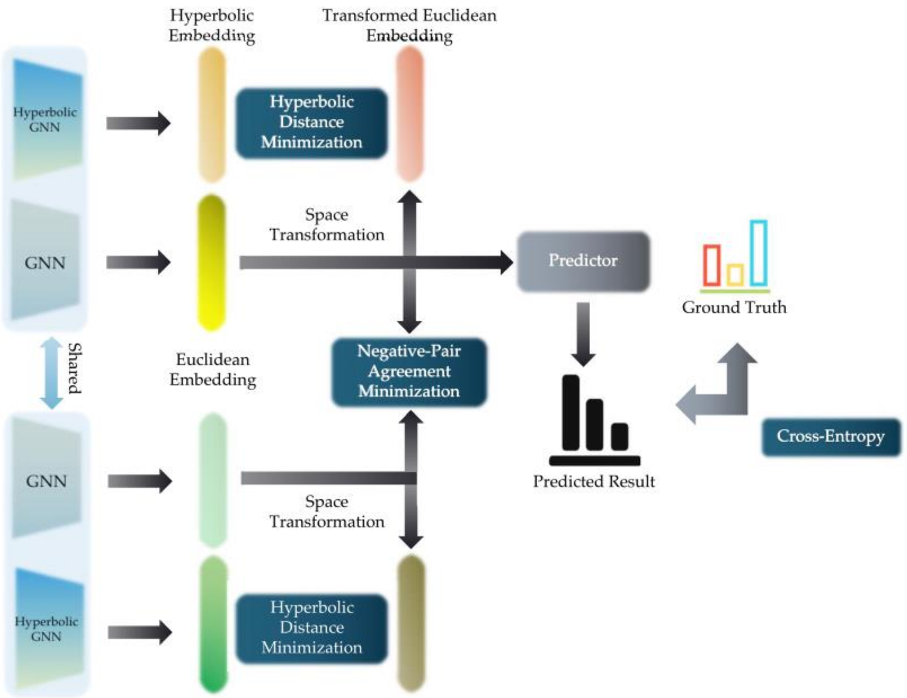
$$\mathcal{H}^{E} = g_{E}(\mathcal{G}). \tag{10}$$

$$\mathcal{H}^{H} = exp_{\mathbf{o}}^{c}(g_{H}(\mathcal{G})). \tag{11}$$

$$\mathcal{H}^{E\to H} = exp_{\mathbf{o}}^{c}(\mathcal{H}^{E}), \tag{12}$$

$$p = \delta(P(\mathcal{H}_{l}^{E})), p \in \mathbf{R}^{K}, \tag{13}$$

Method



$$\mathcal{L}_{contra} = \mathcal{L}_{NCE}^{l} + \mathcal{L}_{NCE}^{u}$$
$$= -log \frac{e^{d_{\mathbb{D}}^{l}(\mathcal{H}_l^H, \mathcal{H}_l^{E\to H})/\tau}}{e^{d_{\mathbb{D}}^{l}(\mathcal{H}_l^H, \mathcal{H}_l^{E\to H})/\tau} + \sum_{i=1}^{N} e^{d_{\mathbb{D}}(\mathcal{H}_l^{E\to H}, \mathcal{H}_{u,i}^H)/\tau}}$$
$$- \frac{\lambda_u}{N} \sum_{i=1}^{N} log \frac{e^{d_{\mathbb{D}}^{u}(\mathcal{H}_{u,i}^H, \mathcal{H}_{u,i}^{E\to H})/\tau}}{e^{d_{\mathbb{D}}^{u}(\mathcal{H}_{u,i}^H, \mathcal{H}_{u,i}^{E\to H})/\tau} + e^{d_{\mathbb{D}}(\mathcal{H}_l^H, \mathcal{H}_{u,i}^{E\to H})/\tau}}, \tag{14}$$

$$\mathcal{L}_{sup} = C(p, p_l), \tag{15}$$

$$\mathcal{L} = \mathcal{L}_{sup} + \omega \cdot \mathcal{L}_{contra}, \tag{16}$$

Experiments

Table 1: Statistics of datasets

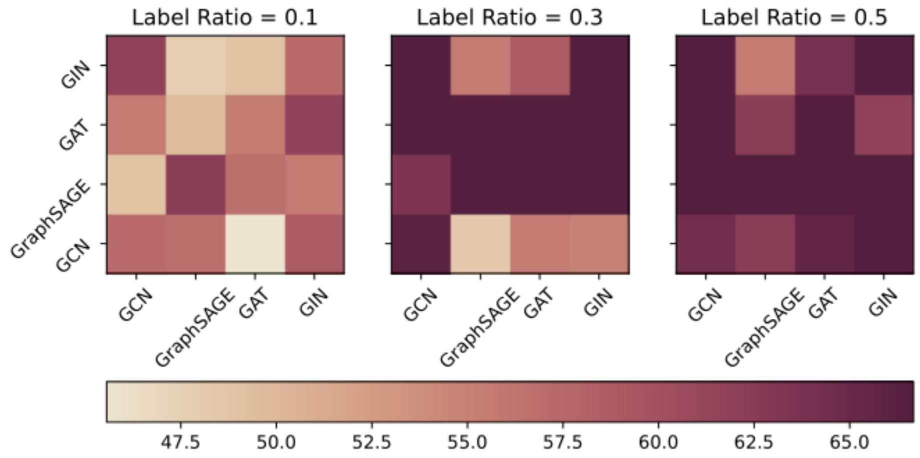
Name \ Statistics	Num. of Graphs	Num. of Classes	Avg. Number of Nodes	Avg. Number of Edges
MUTAG	188	2	17.93	19.79
REDDIT-BINARY	978	2	243.11	288.53
COLLAB	5,000	3	74.49	2457.78

Experiments

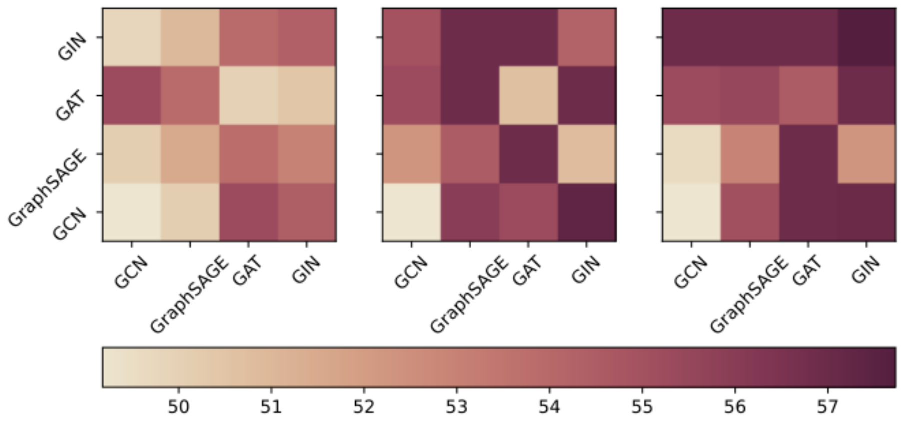
Table 2: Comparison experiment results of classification accuracies and standad error of all the comparing methods (the best results are in bold-face).

Dataset	Label Ratio \ Methods	GCN	GraphSAGE	GAT	GIN	GCC	GraphCL	DSGC
MUTAG	0.1	56.11(std 19.02)	52.78(std 19.14)	58.89(std 18.23)	50.56(std 20.98)	61.67(std 16.15)	57.78(std 13.44)	**62.22(std 15.50)**
	0.3	62.78(std 15.54)	57.22(std 18.05)	62.78(std 14.22)	54.44(std 18.01)	63.89(std 14.06)	62.78(std 12.99)	**66.11(std 12.41)**
	0.5	61.11(std 14.60)	63.33(std 15.57)	58.89(std 17.60	60.56(std 19.88)	65.56(std 13.57)	59.44(std 16.76)	**66.67(std 12.37)**
REDDIT-BINARY	0.1	52.27(std 7.54)	51.55(std 13.07)	53.71(std 12.66)	53.51(std 7.05)	51.65(std 7.36)	54.54(std 7.44)	**55.26(std 6.99)**
	0.3	56.60(std 6.08)	55.88(std 11.35)	54.43(std 7.72)	54.33(std 10.12)	53.40(std 10.68)	56.19(std 5.68)	**57.32(std 5.67)**
	0.5	55.67(std 6.96)	53.61(std 8.33)	57.63(std 7.99)	53.40(std 9.00)	52.37(std 8.81)	**58.14(std 5.73)**	57.73(std 4.35)
COLLAB	0.1	38.98(std 13.78)	38.58(std 14.20)	38.74(std 11.77)	38.48(std 10.88)	37.68(std 13.38)	46.72(std 7.78)	**50.08(std 5.79)**
	0.3	38.54(std 9.07)	42.90(std 13.44)	42.24(std 11.38)	38.56(std 4.62)	37.78(std 13.26)	48.12(std 7.51)	**50.48(std 5.14)**
	0.5	35.14(std 10.13)	36.96(std 12.19)	42.64(std 9.51)	40.24(std 6.41)	38.74(std 6.81)	46.76(std 7.20)	**52.00(std 1.39)**

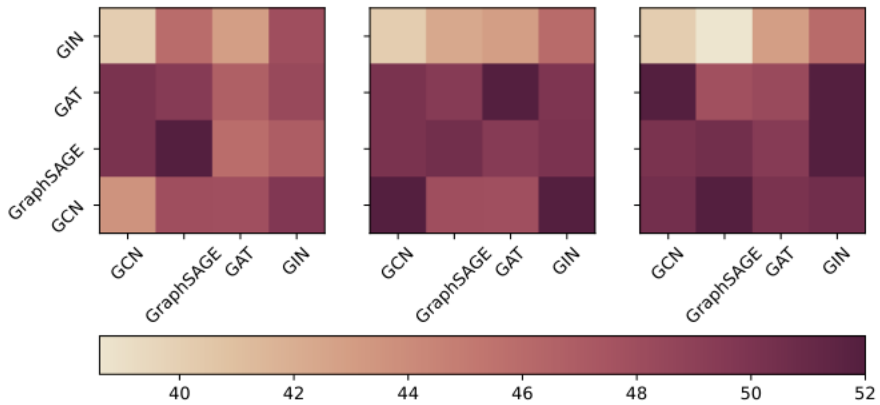
Experiments



(a) Performances of DSGC with different pairs of graph encoders for the Euclidean and Hyperbolic spaces on dataset MUTAG.

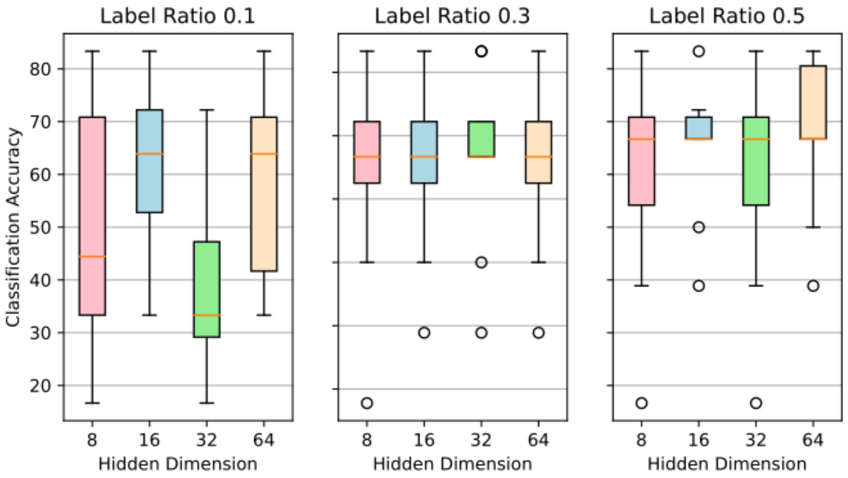


(b) Performances of DSGC with different pairs of graph encoders for the Euclidean and Hyperbolic spaces on dataset REDDIT-BINARY.

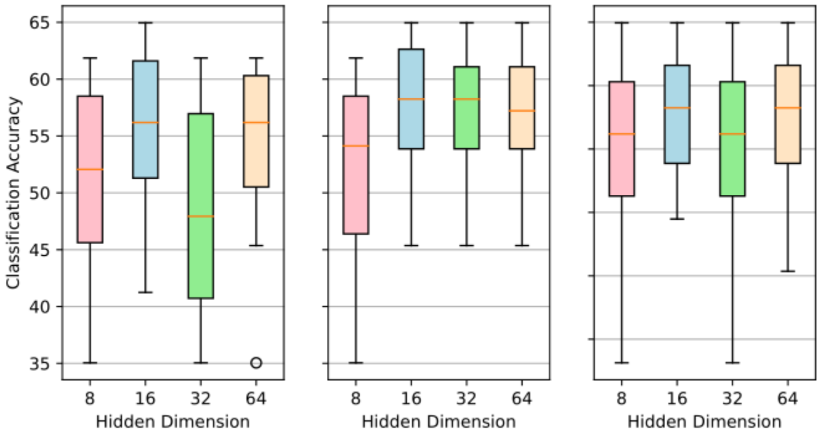


(c) Performances of DSGC with different pairs of graph encoders for the Euclidean and Hyperbolic spaces on dataset COLLAB.

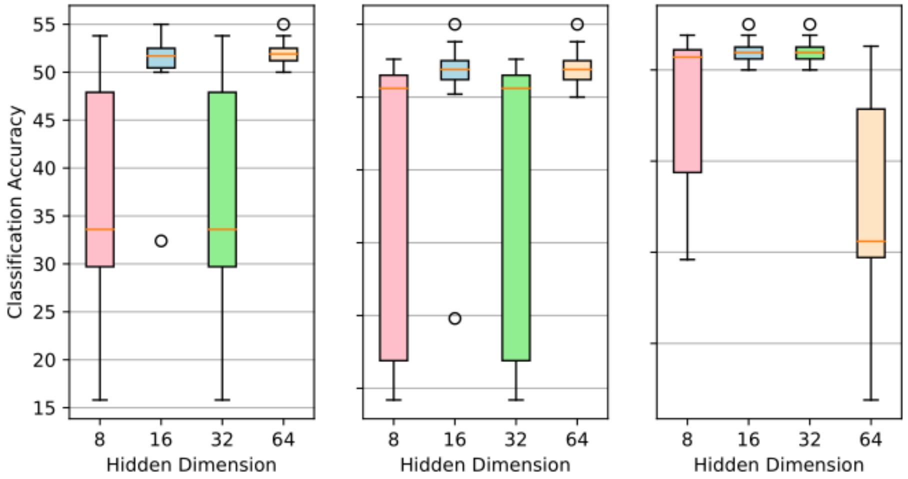
Experiments



(a) Performances of DSGC with different hidden dimension on dataset MUTAG.



(b) Performances of DSGC with different hidden dimension on dataset REDDIT-BINARY.



(c) Performances of DSGC with different hidden dimension on dataset COLLAB.

Thanks!