Learning from Labeled and Unlabeled Vertices in Networks
网络中标记和未标记顶点的学习

KDD，2017

Wei Ye（University of Vienna）

table of Contents

Definition 3.1. **Geometric One-hop Neighborhood.** The geometric one-hop neighbors of a vertex v_i is defined as a set $\mathcal{N}_i^1$ which contains those vertex v_j which can be reached by a random walker from v_i in one step. The geometric one-hop neighborhood is defined as a set $\mathcal{N}^1 = \bigcup_{i=1}^{l} \mathcal{N}_i^1$.

Definition 3.2. **Geometric m-hop Neighborhood.** The geometric m-hop neighbors of a vertex v_i is defined as a set $\mathcal{N}_i^m$ which contains those vertex v_j which can be reached by a random walker from v_i in m steps. The geometric m-hop neighborhood is defined as a set $\mathcal{N}^m = \bigcup_{i=1}^{l} \mathcal{N}_i^m$.

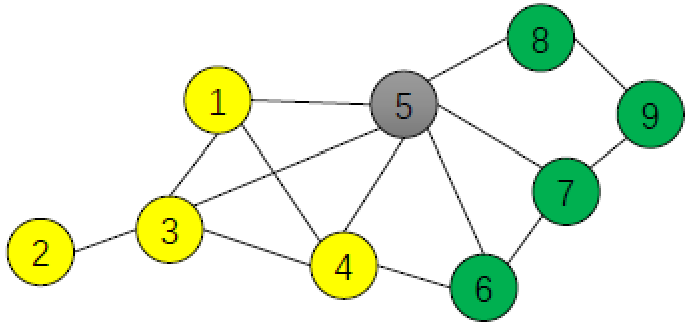
The affinity matrix of vertices is denoted by $\mathrm{A} \in \mathbb{R}^{n \times n}$ with $a_{i,j} = a_{j,i}, a_{i,i} = 0$. The degree matrix D is a diagonal matrix associated with A with $d_{i,i} = \sum_j a_{j,i}$. The random walk transition matrix P is defined as $\mathrm{D}^{-1}\mathrm{A}$.

We denote a vertex v_q in the geometric one-hop neighborhood N^1 by p_q which is the q^{th} row of the transition matrix P

$$f(\mathrm{p}_q) = \mathrm{p}_q \cdot \mathrm{w}^\top + b \tag{1}$$

$$y_q = sign(f(\mathrm{p}_q)) = sign(\mathrm{p}_q \cdot \mathrm{w}^\top + b) \tag{2}$$

Equation (1) does not take the neighborhood relationship into consideration to classify vertices in networks, which would lead to the deterioration of the classifier.



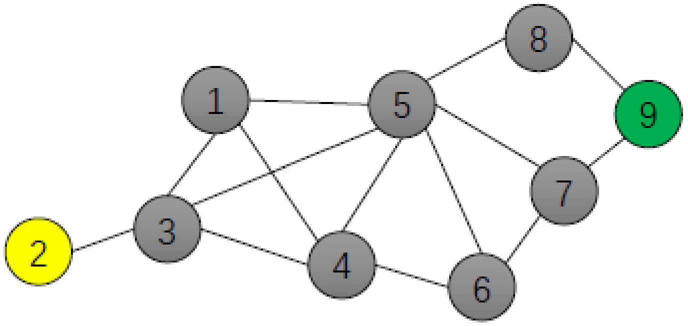
what is the label of the vertex v5?

use a weighted-vote strategy

$$
\begin{aligned}
f(\mathrm{p}_q) &= \mathrm{p}_q \cdot \mathrm{w}^\top + b > 0 \\
y_q &= 1 \\
\text{if } & \frac{\sum_{i \in \mathcal{N}_q^1} a_{i,q} f(\mathrm{p}_i)}{\sum_{i \in \mathcal{N}_q^1} a_{i,q}} \geq +1
\end{aligned} \tag{3}
$$

$$
\begin{aligned}
f(\mathrm{p}_q) &= \mathrm{p}_q \cdot \mathrm{w}^\top + b < 0 \\
y_q &= -1 \\
\text{if } & \frac{\sum_{i \in \mathcal{N}_q^1} a_{i,q} f(\mathrm{p}_i)}{\sum_{i \in \mathcal{N}_q^1} a_{i,q}} \leq -1
\end{aligned} \tag{4}
$$

$$
\begin{aligned}
& y_q \cdot \frac{\sum_{i \in \mathcal{N}_q^1} a_{i,q}(\mathrm{p}_i \cdot \mathrm{w}^\top + b)}{\sum_{i \in \mathcal{N}_q^1} a_{i,q}} \geq 1 \\
\Rightarrow & y_q \cdot \left(\frac{\sum_{i \in \mathcal{N}_q^1} a_{i,q} \cdot \mathrm{p}_i}{d_{q,q}} \cdot \mathrm{w}^\top + b \right) \geq 1 \\
\Rightarrow & y_q \cdot \left(\frac{\mathrm{a}_q \cdot \mathrm{P}}{d_{q,q}} \cdot \mathrm{w}^\top + b \right) \geq 1 \\
\Rightarrow & y_q \cdot (\mathrm{p}_q \cdot \mathrm{P} \cdot \mathrm{w}^\top + b) \geq 1
\end{aligned} \tag{5}
$$



what is the label of the vertex v5?

Geometric m-hop Neighborhood

In the geometric m-hop neighborhood $\mathcal{N}_i^m$, v_q is denoted by p_q^m which is the q-th row of the transition matrix P^m. If we replace p_q in the Inequality (5) with p_q^m, we have:

$$
\begin{aligned}
& y_q \cdot \frac{\sum_{i \in \mathcal{N}_q^m} a_{i,q}(\mathrm{p}_i^m \cdot \mathrm{w}^\top + b)}{\sum_{i \in \mathcal{N}_q^m} a_{i,q}} \geq 1 \\
\Rightarrow & y_q \cdot \left(\frac{\sum_{i \in \mathcal{N}_q^m} a_{i,q} \cdot \mathrm{p}_i^m}{d_{q,q}} \cdot \mathrm{w}^\top + b \right) \geq 1 \\
\Rightarrow & y_q \cdot \left(\frac{\mathrm{a}_q \cdot \mathrm{P}^m}{d_{q,q}} \cdot \mathrm{w}^\top + b \right) \geq 1 \\
\Rightarrow & y_q \cdot (\mathrm{p}_q \cdot \mathrm{P}^m \cdot \mathrm{w}^\top + b) \geq 1 \\
\Rightarrow & y_q \cdot \left(\mathrm{p}_q^m \cdot \mathrm{P} \cdot \mathrm{w}^\top + b \right) \geq 1 \\
\Rightarrow & y_q \cdot \left(\mathrm{p}_q^{m+1} \cdot \mathrm{w}^\top + b \right) \geq 1
\end{aligned}
\tag{6}
$$

where p_q^{m+1} is the representation of v_q in the geometric $(m+1)$-hop neighborhood.

$$
\begin{aligned}
&y_q \cdot \left(\mathbf{p}_q^{m+1} \cdot \mathbf{w}^\top + b\right) \geq 1 - \xi_q \\
&\xi_q \geq 0,\ 1 \leq q \leq l,\ m \geq 1
\end{aligned}
\tag{7}
$$

Similar to SVM,add a slack variable

Unconstrained SVM problem

$$\min_{w,b} \frac{1}{2} w \cdot w^{\top} + \frac{\alpha}{l} \sum_{q=1}^{l} \max(1 - y_q \cdot f(x_q), 0)^2 \qquad (8)$$

represent the vertex v_q by $x_q = p_q^2 + \cdots + p_q^{m+1} = p_q \cdot (P + \cdots + P^m)$.

s

$$\min_{w} \; F(w) = \frac{\lambda}{2}(w \odot d)(w \odot d)^{\top} + \frac{\alpha}{l} \sum_{q=1}^{l} \max(1 - y_q x_q w^{\top}, 0)^2 \quad (9)$$

x_q as $[x_q, 1]$ and w as $[w, b]$.

a dampening factor at each hop

the m^{th} hop is defined as $\rho^m / m!$,

$$x_q = \left[p_q \cdot \left(\frac{\rho^1}{1!} P + \cdots + \frac{\rho^m}{m!} P^m \right), 1 \right],$$

$$d = \left[diag(D^{-\frac{1}{2}})^{\top}, 1 \right]$$

$\odot$ means the Hadamard product.

How to jump out local optima？(GCD)

- 1.GD(Gradient Decent）
- 2.CD(Coordinate Decent)
- What is Coordinate Decent?
- For f(x), $\mathrm{x}^{(k-1)} \rightarrow \mathrm{x}^{(k)}$

$$x_1^{(k)} \in \underset{x_1}{\operatorname{argmin}}\; f\left(x_1, x_2^{(k-1)}, x_3^{(k-1)}, \ldots x_n^{(k-1)}\right)$$

$$x_2^{(k)} \in \underset{x_2}{\operatorname{argmin}}\; f\left(x_1^{(k)}, x_2, x_3^{(k-1)}, \ldots x_n^{(k-1)}\right)$$

$$x_3^{(k)} \in \underset{x_2}{\operatorname{argmin}}\; f\left(x_1^{(k)}, x_2^{(k)}, x_3, \ldots x_n^{(k-1)}\right)$$

$$\ldots$$

$$x_n^{(k)} \in \operatorname{argmin}\; f\left(x_1^{(k)}, x_2^{(k)}, x_3^{(k)}, \ldots x_n\right)$$

每一次我们解决了$x_i^{(k)}$，我们都会使用新的值。

GD

$$F'(\mathrm{w}) = \lambda \mathrm{w} \odot \mathrm{d} - \frac{2\alpha}{l} \sum_{j \in \mathbb{I}(\mathrm{w})} y_j \mathrm{x}_j b_j(\mathrm{w}) \tag{10}$$

where $b_j(\mathrm{w}) = 1 - y_j \mathrm{x}_j \mathrm{w}^\intercal$ and $\mathbb{I}(\mathrm{w}) = \{j | b_j(\mathrm{w}) > 0\}$.

We iteratively update w as follows:

$$\mathrm{w}^{t+1} = \mathrm{w}^t - \eta_t F'(\mathrm{w}^t) \tag{11}$$

where η_t is the learning rate at the t^{th} iteration and is chosen from $\{1, \beta, \beta^2, \ldots\}$ by a line search.

The last w in GD is the first w in CD

CD

sequence $\{\mathrm{w}^t\}$ $(t = 0, 1, 2, \ldots)$. At each iteration, w^{t+1} is produced by sequentially updating each entry of w^t with other entries fixed. The process produces a sequence of vectors $\mathrm{w}^{t,i}$ $(i = 1, \ldots, n+1)$, such that $\mathrm{w}^{t,0} = \mathrm{w}^t$, $\mathrm{w}^{t,n+1} = \mathrm{w}^{t+1}$ and

$$\mathrm{w}^{t,i} = [w_1^{t+1}, \ldots, w_i^{t+1}, w_{i+1}^t, \ldots, w_{n+1}^t]$$

Updating $\mathrm{w}^{t,i}$ to $\mathrm{w}^{t,i+1}$ becomes the following one-variable sub-problem:

$$\begin{aligned}
&\min_z \quad F_i(w_1^{t+1}, \ldots, w_i^{t+1}, w_{i+1}^t + z, w_{i+2}^t \ldots, w_{n+1}^t) \\
\equiv &\min_z \quad F_i(\mathrm{w}^{t,i} + z\mathrm{e}_i) \\
= &\min_z \quad \frac{\lambda}{2}\left((\mathrm{w}^{t,i} + z\mathrm{e}_i) \odot \mathrm{d}\right)\left((\mathrm{w}^{t,i} + z\mathrm{e}_i) \odot \mathrm{d}\right)^{\top} \\
&+ \frac{\alpha}{l} \sum_{j \in \mathbb{I}(\mathrm{w}^{t,i} + z\mathrm{e}_i)} (b_j(\mathrm{w}^{t,i} + z\mathrm{e}_i))^2
\end{aligned} \tag{12}$$

where $\mathrm{e}_i \in \mathbb{R}^{1\times(n+1)}$ is a vector with the i^{th} entry 1 and all other entries 0.

The first derivative of (12) with respect to z is:

$$F_i'(z) = \lambda\left(\mathrm{w}_i^{t,i} + z\right)\cdot \mathrm{d}_i - \frac{2\alpha}{l}\sum_{j\in\mathbb{I}(\mathrm{w}^{t,i}+z\mathrm{e}_i))}(y_j x_{j,i}(b_j(\mathrm{w}^{t,i}+z\mathrm{e}_i)) \tag{13}$$

As pointed out in [4], $F_i(z)$ is not twice differentiable at some j, where $b_j(\mathrm{w}^{t,i}+z\mathrm{e}_i) = 0$. Following [4, 19], we define the generalized second derivative of (12) with respect to z as:

$$F_i''(z) = \lambda\mathrm{d}_i + \frac{2\alpha}{l}\sum_{j\in\mathbb{I}(\mathrm{w}^{t,i}+z\mathrm{e}_i))} x_{j,i}^2 \tag{14}$$

The Newton direction at a given z is $\frac{F_i'(z)}{F_i''(z)}$. We start from $z = 0$ and apply a line search $z = z - \eta_i\frac{F_i'(z)}{F_i''(z)}$ until $F_i(z - \eta_i\frac{F_i'(z)}{F_i''(z)}) < F_i(z)$, where η_i is the learning rate for the i^{th} element and is chosen from $\{1, \beta, \beta^2, \ldots\}$.

[4]Kai-Wei Chang, Cho-Jui Hsieh, and Chih-Jen Lin. 2008. Coordinate descent method for large-scale l2-loss linear support vector machines. Journal of Machine Learning Research 9, Jul (2008), 1369–1398.

Complexity analysis.

- Assume that a network has n vertices and r edges, l is the number of the labeled vertices,k is the number of w elements.

$$\tilde{O}(n \cdot (r + k \cdot l)).$$

Experiment

- Network1 has 100 … n are grouped into two classes (3 …). average degree is 21.

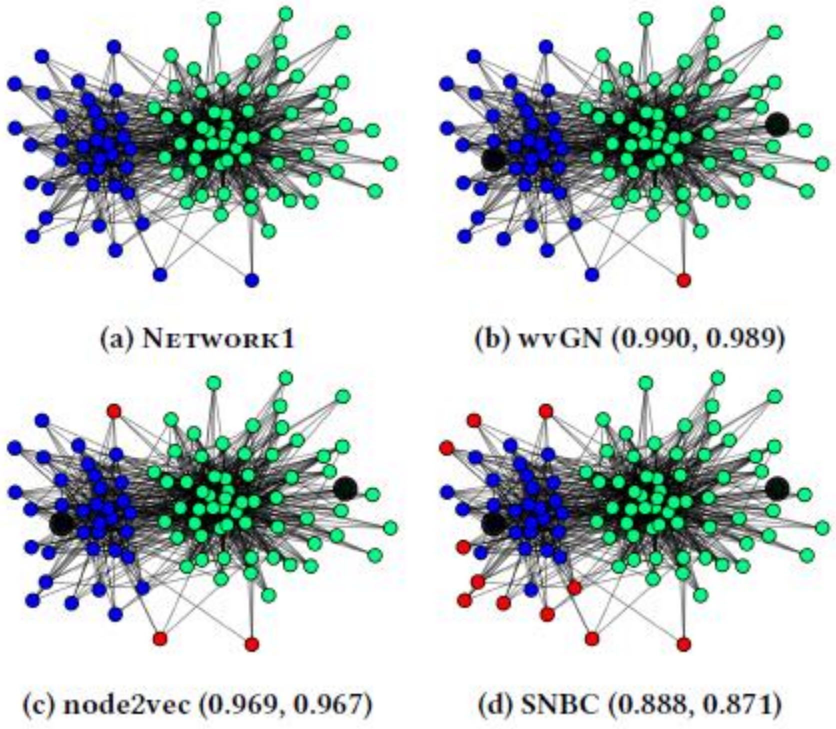
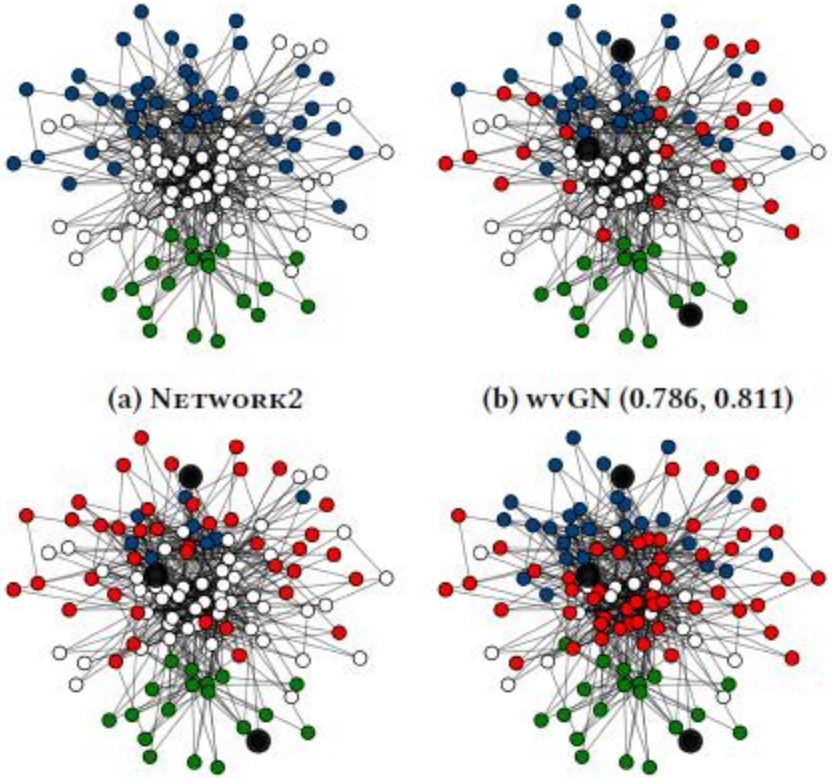


Figure 4: Classification results on NETWORK1. The two labeled vertices are in black and the misclassified vertices are in red. Subcaption: Method (Micro-F1, Macro-F1).

Table 1: Classification results on NETWORK1 with the varying number of labeled vertices (#LV) in each class.

	Micro-F1 (%)					Macro-F1 (%)				
#LV	1	2	3	4	5	1	2	3	4	5
wvGN	**98.7±0.7**	**98.9±0.3**	**99.3±0.5**	**99.0±0.3**	**99.2±0.5**	**98.6±0.7**	**98.8±0.4**	**99.2±0.6**	**98.9±0.3**	**99.2±0.5**
wvGN (full)	**98.7±0.5**	98.8±0.7	**99.3±0.5**	**99.0±0.6**	**99.2±0.5**	**98.6±0.5**	98.7±0.7	**99.2±0.6**	**98.9±0.7**	**99.2±0.6**
node2vec	96.9±0.5	94.5±6.0	94.8±5.1	94.5±5.6	94.2±5.4	96.7±0.5	94.3±5.9	94.6±5.1	94.2±5.5	94.0±5.4
Deepwalk	74.6±26.7	75.6±20.7	77.8±18.9	75.9±20.1	79.6±12.0	74.3±26.9	74.5±22.4	77.0±19.9	75.0±21.2	78.7±12.4
SNBC	85.6±15.6	81.2±17.4	71.4±10.7	71.3±10.5	74.9±10.3	84.0±16.6	80.9±17.8	70.9±10.8	70.5±10.6	73.8±10.0
wvRN	77.9±19.2	66.6±18.5	76.3±21.1	86.7±17.7	92.6±7.3	71.5±25.2	55.1±24.0	70.8±25.7	84.3±20.4	91.7±8.2
SCRN	77.3±22.4	74.2±21.8	81.2±17.2	86.3±13.6	90.2±10.6	66.6±32.3	61.0±31.8	73.5±25.6	80.6±21.5	87.0±16.2
SocDim	55.2±9.9	52.0±8.5	53.7±5.8	57.2±6.2	60.0±6.3	51.8±10.2	49.7±7.5	52.4±4.6	54.9±5.8	58.4±5.4
HeatKernel	82.4±19.8	68.8±20.0	78.6±21.2	88.2±16.3	90.9±11.9	77.7±25.9	57.9±26.1	73.8±25.8	86.6±18.0	90.2±12.3
LGC	60.4±18.6	55.6±18.3	60.5±21.9	62.1±22.1	62.4±17.0	45.1±22.4	39.5±19.7	47.5±26.4	46.9±27.0	45.1±20.1
TSVM	37.8±0	37.5±0	37.2±0	37.0±0	36.7±0	27.4±0	27.3±0	27.1±0	27.0±0	26.8±0
LapSVM	51.9±9.4	52.9±8.8	56.0±8.5	58.2±14.4	65.1±12.8	41.6±8.9	45.6±9.9	49.7±8.2	52.9±16.7	62.4±12.4

- Network2 has 1 … hich are grouped into three classes (1 … vely). average degree is 10.



(a) Network2 (b) wvGN (0.786, 0.811)

(c) node2vec (0.650, 0.622) (d) TSVM (0.539, 0.537)

Figure 5: Classification results on Network2. The three labeled vertices are in black and the misclassified vertices are in red. Subcaption: Method (Micro-F1, Macro-F1).

Table 2: Classification results on Network2 with the varying number of labeled vertices (#LV) in each class.

#LV	Micro-F1 (%)					Macro-F1 (%)				
	1	2	3	4	5	1	2	3	4	5
wvGN	66.7±11.4	71.8±11.0	**77.3±14.5**	**89.0±2.7**	**88.9±2.5**	67.7±12.2	73.3±11.6	**79.1±15.0**	**90.9±2.5**	**90.9±2.1**
wvGN (full)	**67.4±10.1**	**79.9±7.6**	69.6±15.8	86.3±3.5	84.1±5.8	**68.1±10.4**	**82.3±7.1**	71.4±15.1	87.7±4.2	85.3±6.0
node2vec	60.4±9.7	70.5±7.3	61.8±11.5	67.9±10.9	67.9±10.9	58.9±12.3	71.8±8.2	61.2±12.6	67.8±12.3	67.8±12.3
Deepwalk	42.4±7.4	46.5±7.1	42.9±12.6	51.2±5.9	51.2±5.9	40.7±7.2	45.2±6.8	41.7±11.7	48.5±5.8	48.5±5.8
SNBC	35.8±6.7	45.5±12.8	48.7±12.3	56.0±5.0	57.2±4.8	35.3±12.0	43.8±13.1	46.0±13.9	56.7±5.8	57.9±4.9
wvRN	45.0±12.1	56.0±15.7	56.9±11.3	68.1±7.5	68.1±7.5	36.2±13.9	49.4±17.8	53.9±14.9	64.8±11.6	64.8±11.6
SCRN	45.9±14.6	50.3±14.8	57.3±10.3	57.2±4.8	65.2±10.3	36.3±16.4	39.7±18.3	50.1±15.2	57.9±4.9	58.7±17.4
SocDim	31.0±7.9	37.0±7.6	42.9±7.2	44.5±4.7	44.5±4.7	29.1±6.7	35.4±6.5	41.3±6.7	42.1±5.2	42.1±5.2
HeatKernel	48.3±10.1	57.3±16.4	55.3±12.6	65.1±8.0	65.1±8.0	40.2±14.5	52.2±17.7	52.5±14.7	59.6±12.0	59.6±12.0
LGC	31.2±5.1	31.3±4.7	31.5±8.0	35.2±4.8	35.2±4.8	29.6±3.5	29.0±3.6	28.4±6.7	32.1±3.8	32.1±3.8
TSVM	54.2±5.0	64.6±7.5	51.7±4.1	52.0±4.9	52.0±4.9	53.5±5.0	64.7±6.7	51.4±4.1	51.7±4.9	51.7±4.9
LapSVM	37.6±13.2	49.4±10.7	58.3±10.5	63.1±8.5	63.1±8.5	23.4±8.6	39.1±11.2	56.9±12.3	66.0±7.9	66.0±7.9

- For the real-world data, we use four popular relational datasets
- CoRA, PubMed, IMDb， Wikipedia

Table 3: Classification results on CoRA (#vertices: 24,519, #edges: 92,207, #classes: 10) with the varying percent of labeled vertices (%LV). N/A means the results are not available because the algorithm is not finished in one week.

	Micro-F1 (%)					Macro-F1 (%)				
%LV	1%	3%	5%	7%	9%	1%	3%	5%	7%	9%
wvGN	62.8±2.8	**71.6±0.9**	**74.2±0.6**	**75.4±0.5**	**75.8±0.4**	**52.8±2.0**	**63.0±1.0**	**66.1±1.0**	**67.7±0.6**	**68.3±0.8**
node2vec	49.7±15.4	50.3±7.3	50.4±7.7	54.2±7.5	54.8±8.0	42.2±10.0	43.3±6.5	46.4±5.9	49.2±5.3	50.7±7.0
Deepwalk	23.2±1.6	16.5±1.9	18.8±1.4	24.1±0.2	26.7±0.9	15.0±0.5	14.5±1.2	16.9±1.0	20.0±0.1	21.6±0.7
SNBC	50.1±2.6	63.0±1.1	66.5±0.9	68.0±0.9	68.2±0.9	27.2±2.4	49.2±2.1	54.4±0.1	57.3±1.5	57.8±1.0
wvRN	**65.2±1.4**	70.0±0.7	72.1±0.5	73.2±0.6	74.3±0.3	52.5±2.2	59.3±1.0	61.9±0.8	63.7±0.6	64.9±0.6
SCRN	64.3±1.9	70.3±0.7	72.6±0.4	73.6±0.4	74.6±0.3	50.8±3.0	59.2±0.9	62.2±0.7	63.9±0.5	65.2±0.5
SocDim	49.3±0.9	55.6±0.5	59.6±0.6	62.5±0.8	63.6±0.4	27.8±1.3	44.8±0.8	49.9±0.9	53.4±1.2	54.8±0.8
HeatKernel	64.3±2.0	69.6±0.7	72.0±0.5	73.1±0.6	74.2±0.2	51.7±3.1	59.1±1.0	62.0±0.9	63.7±0.6	64.9±0.4
LGC	47.4±2.6	48.7±2.5	48.6±1.8	48.7±2.0	48.5±1.7	23.1±0.3	24.5±2.9	23.1±2.7	23.3±1.6	22.3±2.3
TSVM	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
LapSVM	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 4: Classification results on PubMed (#vertices: 19,717, #edges: 44,324, #classes: 3) with the varying percent of labeled vertices (%LV).

	Micro-F1 (%)					Macro-F1 (%)				
%LV	1%	3%	5%	7%	9%	1%	3%	5%	7%	9%
wvGN	63.0±5.7	**75.8±1.2**	**78.4±0.8**	**79.0±0.6**	**79.8±0.5**	59.5±6.4	**73.9±1.1**	**76.7±1.0**	**77.3±0.8**	**78.2±0.5**
node2vec	64.0±4.0	63.2±6.1	64.5±5.5	65.6±5.7	61.1±6.2	61.8±4.5	69.5±8.1	61.6±6.8	62.7±6.8	57.4±7.4
Deepwalk	33.2±1.4	35.0±1.0	35.2±0.8	35.5±0.8	36.1±0.9	32.7±1.2	34.5±0.9	34.7±0.9	35.0±0.6	35.7±0.9
SNBC	52.1±7.2	71.3±1.8	72.3±0.6	78.4±0.6	79.5±0.5	44.2±10.0	68.0±1.5	73.7±1.0	76.3±1.0	77.7±0.6
wvRN	35.8±0.8	35.9±0.5	35.9±0.4	35.9±0.3	36.0±0.3	33.2±0.3	33.5±0.3	33.3±0.4	33.4±0.3	33.3±0.2
SCRN	36.4±0.9	36.2±0.3	36.3±0.5	36.2±0.3	36.2±0.3	33.1±0.3	33.6±0.2	33.4±0.4	33.4±0.4	33.3±0.3
SocDim	42.6±1.6	47.6±1.9	51.5±1.2	55.2±2.1	57.3±1.7	37.7±3.2	43.3±4.1	47.7±3.1	53.1±3.3	55.4±2.2
HeatKernel	**67.9±2.4**	73.3±1.1	76.5±0.6	77.7±0.6	78.8±0.4	**65.9±2.4**	71.8±1.0	75.0±0.6	76.4±0.7	77.5±0.4
LGC	62.2±10.0	71.5±6.3	75.6±2.3	76.6±1.6	76.6±1.7	56.8±11.9	68.9±0	72.7±3.1	74.1±1.8	74.1±1.8
TSVM	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
LapSVM	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 5: Classification results on IMDb (#vertices: 19,359, #edges: 362,079, #classes: 21) with the varying percent of labeled vertices (%LV).

	Micro-F1 (%)					Macro-F1 (%)				
%LV	1%	3%	5%	7%	9%	1%	3%	5%	7%	9%
wvGN	24.5±7.7	**40.8±1.9**	**41.5±0.9**	**43.4±0.9**	**44.1±0.8**	8.7±1.2	11.6±0.5	12.5±0.5	13.4±0.7	13.6±0.5
node2vec	11.5±2.7	14.7±1.1	14.1±0.5	15.3±1.0	18.0±1.1	8.6±1.9	11.9±0.4	11.7±0.2	12.1±0.3	12.8±0.4
Deepwalk	15.2±1.7	13.3±0.7	15.9±0.2	24.8±1.9	32.8±1.2	11.2±0.6	10.7±0.3	11.1±0.5	11.2±0.5	11.0±0.2
SNBC	20.3±8.3	35.8±2.4	35.3±1.0	34.8±0.7	34.8±0.8	7.6±1.5	12.3±0.7	14.4±0.4	**15.6±0.4**	**16.5±0.3**
wvRN	33.3±5.0	36.0±0.5	36.4±0.4	37.0±0.3	37.4±0.3	10.0±0.7	10.4±0.3	10.3±0.2	10.4±0.2	10.3±0.2
SCRN	33.6±6.2	36.7±0.9	37.1±0.3	37.8±0.4	38.0±0.4	9.3±0.9	9.5±0.4	9.5±0.2	10.0±0.3	10.0±0.2
SocDim	**37.2±1.8**	38.6±1.2	40.3±0.1	40.8±0.4	41.1±0.3	7.6±0.5	8.0±0.5	8.8±0	9.2±0.4	0.095±0.003
HeatKernel	30.8±6.8	35.6±1.2	37.2±0.6	39.3±0.7	41.2±0.8	**11.7±1.4**	**13.9±0.6**	**14.7±0.5**	15.5±0.6	16.4±0.6
LGC	37.1±4.0	39.4±0.1	39.7±0.1	39.9±0.1	39.9±0.1	8.3±0.4	9.0±0.2	9.4±0.2	9.4±0.1	9.5±0.1
TSVM	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
LapSVM	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 6: Classification on WIKIPEDIA (#vertices: 4,777, #edges: 184,812, #classes: 40) with the varying percent of labeled vertices (%LV).

	Micro-F1 (%)					Macro-F1 (%)				
%LV	1%	3%	5%	7%	9%	1%	3%	5%	7%	9%
wvGN	41.6±1.0	**45.3±0.6**	**44.9±1.0**	**45.4±1.0**	**45.4±0.7**	**6.5±0.4**	6.8±0.4	6.3±0.4	6.8±0.5	6.6±0.3
node2vec	29.3±5.0	31.2±1.8	31.1±2.4	31.7±2.6	34.3±4.2	6.4±0.4	**7.1±0.2**	**7.6±0.5**	**8.0±0.6**	**8.3±0.5**
deepwalk	17.7±1.5	14.0±1.2	13.5±1.4	13.8±1.8	15.7±1.8	4.2±0.4	4.1±0.2	4.1±0.3	4.0±0.3	4.0±0.2
SNBC	**41.9±0.5**	42.5±0.4	42.4±0.6	42.5±0.6	42.6±0.5	4.4±0.2	4.4±0.3	4.4±0.5	4.5±0.3	4.6±0.4
wvRN	1.6±1.1	4.2±2.5	7.8±4.0	10.4±3.5	13.0±6.6	0.7±0.4	1.1±0.6	1.6±0.5	2.0±0.7	2.2±0.6
SCRN	1.7±1.3	4.2±2.1	8.3±4.1	11.8±3.1	15.0±6.2	0.7±0.4	1.1±0.5	1.7±0.5	2.1±0.5	2.4±0.5
SocDim	33.9±1.5	32.4±1.6	32.6±0.8	33.3±0.9	33.5±1.0	5.6±0.2	6.4±0.3	6.5±0.4	6.8±0.5	7.2±0.4
HeatKernel	1.3±1.0	4.3±3.3	7.7±4.3	10.5±3.8	12.8±8.7	0.6±0.4	1.0±0.7	1.5±0.6	1.7±0.6	2.0±0.7
LGC	36.7±6.5	38.9±0.1	39.0±0.1	39.0±0.1	39.1±0.1	3.0±0.3	3.0±0	3.0±0	3.0±0	2.9±0
TSVM	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
LapSVM	27.4±13.9	41.9±0.6	42.0±0.6	42.2±0.3	42.2±0.6	2.4±0.7	4.1±0.4	4.2±0.3	4.3±0.3	4.2±0.4

Thank you !