ResTran: A GNN Alternative to Learn A Graph with Features

Shota Saito*, Takanori Maehara**, Mark Herbster*
* Dept of CS, UCL **Independent Researcher

paper

TL; DR

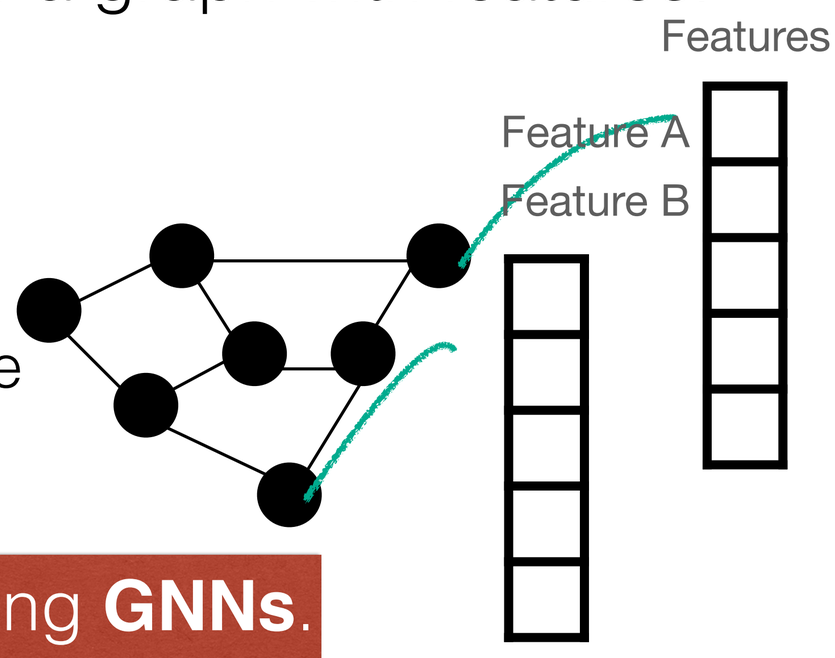
Proposed a GNN alternative learning framework for a graph with features.

Settings

Learn a graph and features associated with vertex

Ex1. **Molecules**:
Vertex as atom, edge as connection, and feature as the kind of atoms

Ex2. **Citation Networks** such as Cora



Modern approach to this setting is using **GNNs**.
Do we have an alternative to GNNs?

$L_b := L + b\mathbf{1}^\top\mathbf{1}$
L: graph Laplacian.
$X := (\mathbf{x}_1, \ldots, \mathbf{x}_n)$
$\mathbf{x}_i$: features at i
$\mathbf{e}_i$: coordinate at i

Proposed Algorithm (**Res**istance **Tran**sformation)

1. Represent vertex i as $\mathbf{v}_i := XL_b^{-1/2}\mathbf{e}_i$
2. Train with **any vector-based model** incl. SVM or NN models

$$f = YourNiceVectorBasedModel(\{\mathbf{v}_i\}_{i \in Train})$$

Typically pseudo-inverse costs $O(n^3)$ but well approximated by $O(m)$

Experiments

Vertex Classification tasks

Heterophilous datasets

	Type	cora	citeseer	pubmed	chameleon	squirrel	actor
GCN	GNN	**79.9 ± 0.9**	67.4 ± 1.1	83.8 ± 0.4	32.7 ± 2.0	23.5 ± 1.1	25.9 ± 0.9
GAT	GNN	74.9 ± 4.2	67.6 ± 0.1	82.8 ± 0.2	32.8 ± 1.8	23.4 ± 1.3	26.4 ± 0.9
SGC	GNN	79.3 ± 1.7	70.2 ± 0.8	67.9 ± 1.8	31.8 ± 1.8	23.5 ± 0.8	26.0 ± 0.8
ResTran + LP	Non-NN	30.6 ± 0.6	20.6 ± 4.6	39.5 ± 1.4	20.3 ± 0.8	20.0 ± 0.3	22.3 ± 2.8
ResTran + SVM	Non-NN	49.1 ± 5.7	45.5 ± 6.7	76.5 ± 2.2	33.6 ± 5.8	31.9 ± 0.9	29.4 ± 0.9
ResTran + VAT	NN	77.6 ± 2.5	68.7 ± 1.1	82.8 ± 0.7	34.0 ± 1.4	27.7 ± 3.5	27.8 ± 1.2
ResTran + AVAE	NN	78.2 ± 1.8	**71.7 ± 1.0**	**83.9 ± 0.7**	**40.7 ± 1.4**	**32.4 ± 0.8**	**29.5 ± 1.3**

Graph Classification tasks

	Type	MUTAG	ENZYMES	NCI1	PROTEINS	BZR
GIN	GNN	89.4±5.3	**46.3±3.6**	**78.5±0.7**	76.8±5.5	88.8±4.1
ResTran + MLP	NN	**89.5±3.7**	45.7±4.0	77.5±2.4	**78.4±3.7**	**89.8±3.2**

Ours is better for heterophilous datasets even with non-NN methods (SVMs!)

Justification via Connections between spectral clustering and k-means

Featureless setting (where $X = I$)

$\{V_\ell\}_{\ell \in [k]}$: a clustering result

k-means using effective resistance = k-means using ResTran

$$\text{argmin}_{\{V_\ell\}_{\ell\in[k]}} \sum_{\ell\in[k]} \sum_{i,j\in V_\ell} \frac{r(i,j)}{|V_\ell|} = \text{argmin}_{\{V_\ell\}_{\ell\in[k]}} \sum_{\ell\in[k]} \sum_{i\in V_\ell} \|\mathbf{v}_i - \mu_\ell\|_2^2$$

$r(i,j)$ is an effective resistance and $r(i,j) = \|\mathbf{v}_i - \mathbf{v}_j\|_2^2$

k-means using ResTran = Spectral Clustering

$$\text{argmin}_{\{V_\ell\}_{\ell\in[k]}} \sum_{\ell\in[k]} \sum_{i\in V_\ell} \|\mathbf{v}_i - \mu_\ell\|_2^2 = \text{argmin}_{\{V_\ell\}_{\ell\in[k]}} SC(\{V_\ell\}_{\ell\in[k]})$$

$\mathbf{v}_i$ is connected to spectral clustering

Graph with Feature Setting: We have more in the paper.