Bag of Tricks for Node Classification with Graph Neural Networks

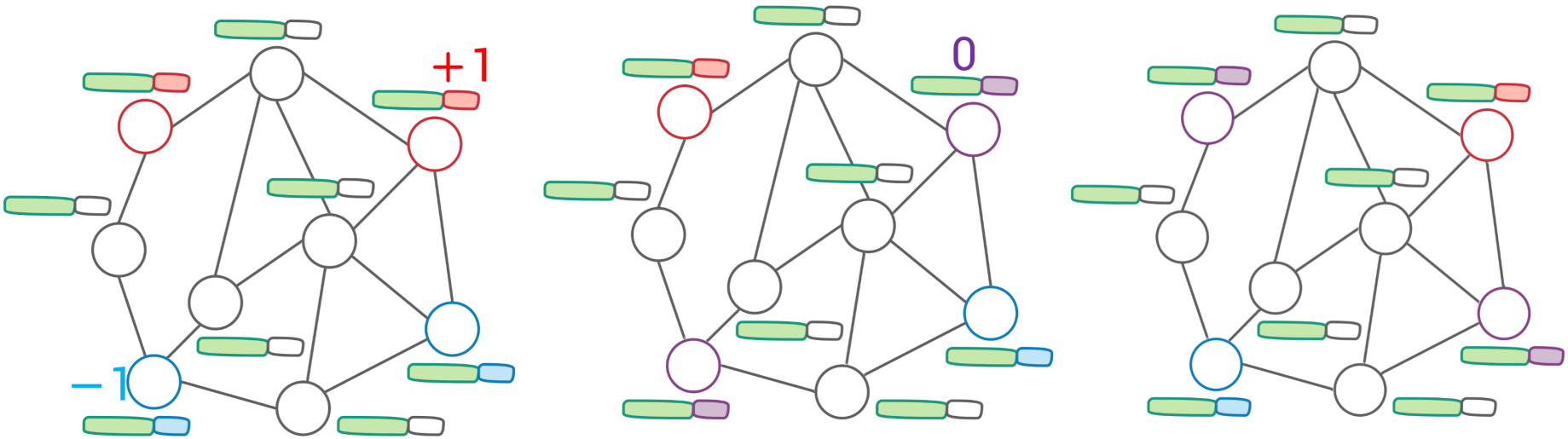
Yangkun Wang[1], Jiarui Jin[1], Weinan Zhang[1], Yong Yu[1], Zheng Zhang[2], David Wipf[2]

[1]Shanghai Jiao Tong University

[2]Amazon

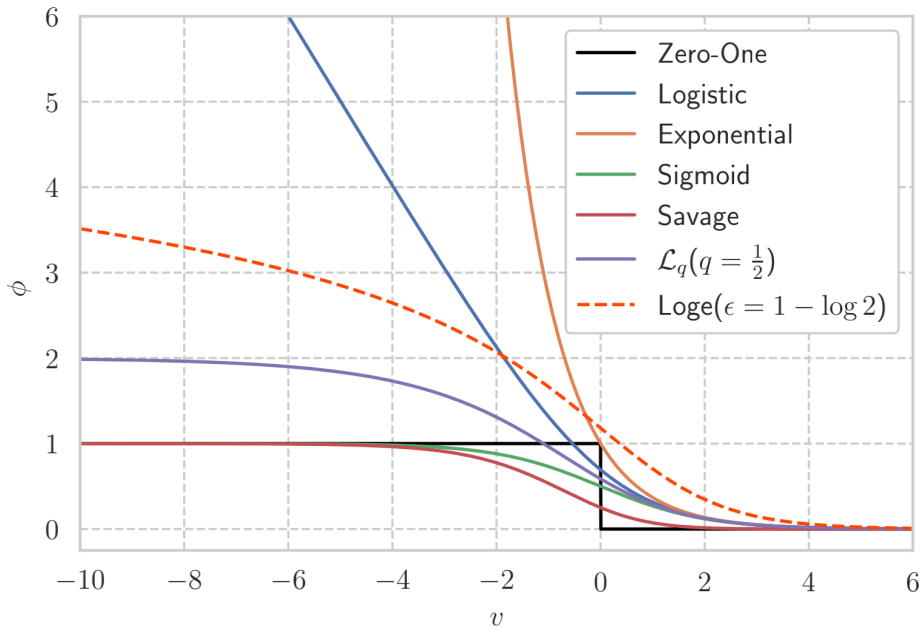
Label Usage

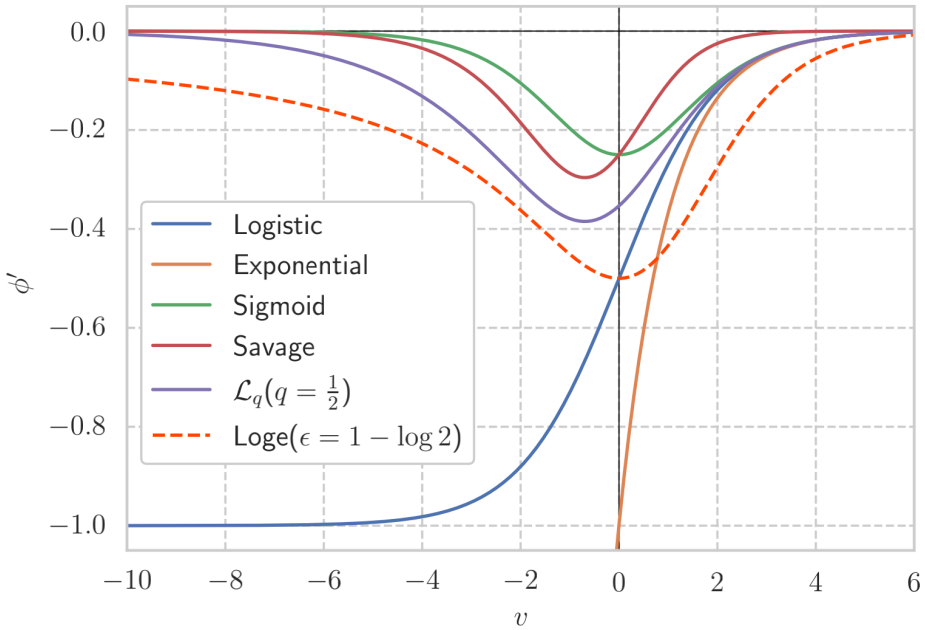
Randomly split the training set into several parts (usually two). Take some training node labels as model input and predict the remaining ones.



Robust Loss Function for Classification

Loss	$\rho(z)$	$\rho(\phi_{logit}(v))$
Logistic	z	$\log(1+\exp(-v))$
Exponential	$\exp(z)-1$	$\exp(-v)$
Sigmoid	$1-\exp(-z)$	$\frac{1}{1+\exp(v)}$
Savage	$(1-\exp(-z))^2$	$\frac{1}{(1+\exp(v))^2}$
$\mathcal{L}_q$	$\frac{1}{q}(1-\exp(-qz))$	$\frac{1}{q}\left(1-\frac{1}{(1+\exp(-v))^q}\right)$
Loge	$\log(\epsilon+z)-\log\epsilon$	$\log(\epsilon+\log(1+\exp(-v)))-\log\epsilon$





Tweaking the GAT Architecture

GAT with symmetric normalized adjacency matrix

Let $\boldsymbol{A}_{att} = \boldsymbol{D}\boldsymbol{\alpha}$, with α being the attention matrix,

$$\boldsymbol{X}^{(l+1)} = \sigma\left(\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\tilde{\boldsymbol{A}}_{att}\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\boldsymbol{X}^{(l)}\boldsymbol{W}_0^{(l)} + \boldsymbol{X}^{(l)}\boldsymbol{W}_1^{(l)}\right),$$

where $\tilde{\boldsymbol{A}}_{att} = \boldsymbol{I} + \boldsymbol{A}_{att}$.

Non-interactive attention

$$\alpha_{ij} = \frac{\exp\left(\mathrm{LeakyReLU}\left(\boldsymbol{a}^T\boldsymbol{x}_j\right)\right)}{\sum_{r\in\mathcal{N}(v_i)}\exp\left(\mathrm{LeakyReLU}\left(\boldsymbol{a}^T\boldsymbol{x}_r\right)\right)}$$

Attention involving edge features

$$\alpha_{ij} = \frac{\exp\left(\mathrm{LeakyReLU}\left(\boldsymbol{a}^T[\boldsymbol{x}_i^V \parallel \boldsymbol{x}_j^V \parallel \boldsymbol{x}_{ij}^E]\right)\right)}{\sum_{r\in\mathcal{N}(v_i)}\exp\left(\mathrm{LeakyReLU}\left(\boldsymbol{a}^T[\boldsymbol{x}_i^V \parallel \boldsymbol{x}_r^V \parallel \boldsymbol{x}_{ij}^E]\right)\right)}$$

Thank you

Full Paper Link: `https://arxiv.org/abs/2103.13355`

Thank you

Full Paper Link: `https://arxiv.org/abs/2103.13355`

Thanks for listening!