ICLR 2019

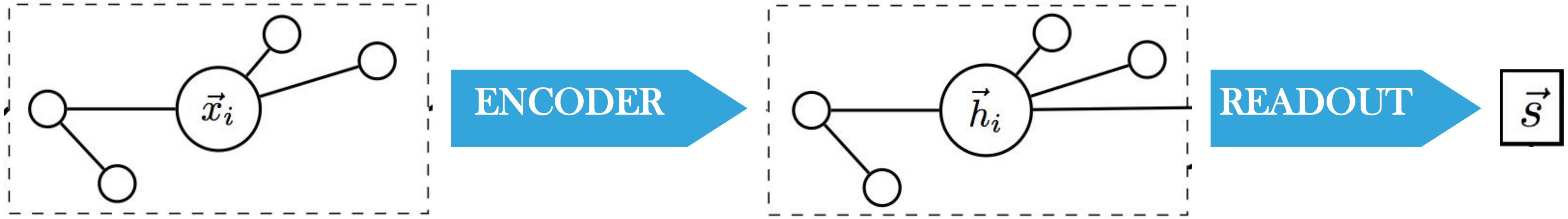
DEEP GRAPH INFOMAX

MOTIVATION

- Most successful methods use supervised learning while most graph data is unlabeled.
- Random walk-based methods over-emphasize proximity information at the expense of structural information.
- High dependence on hyperparameter choice

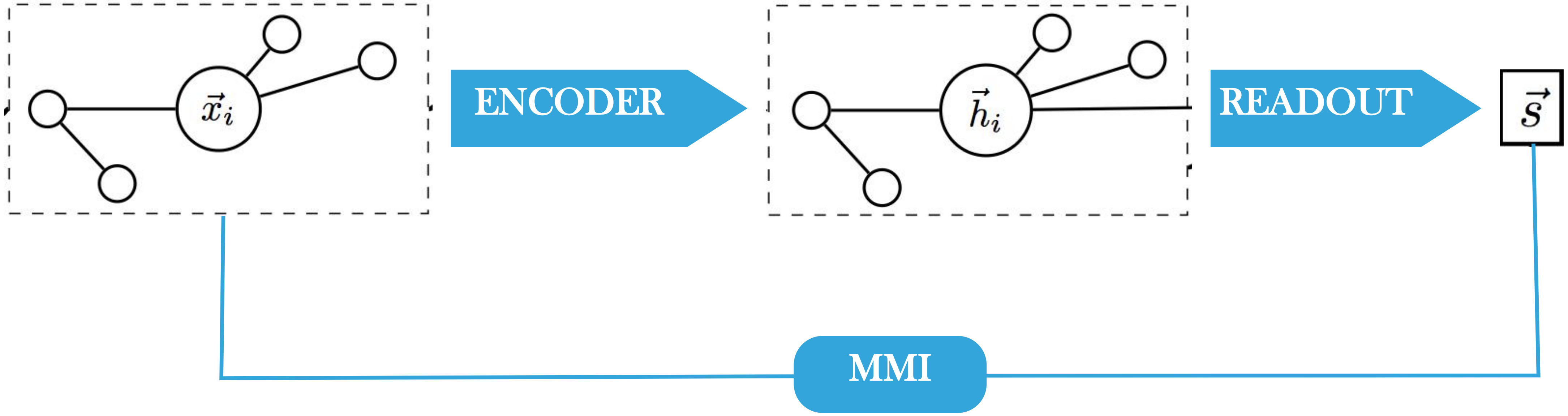
KEY POINT

- A general approach for learning node embeddings in an unsupervised manner.
- Maximize mutual information between a high-level "global" representation and "local" parts of the input.



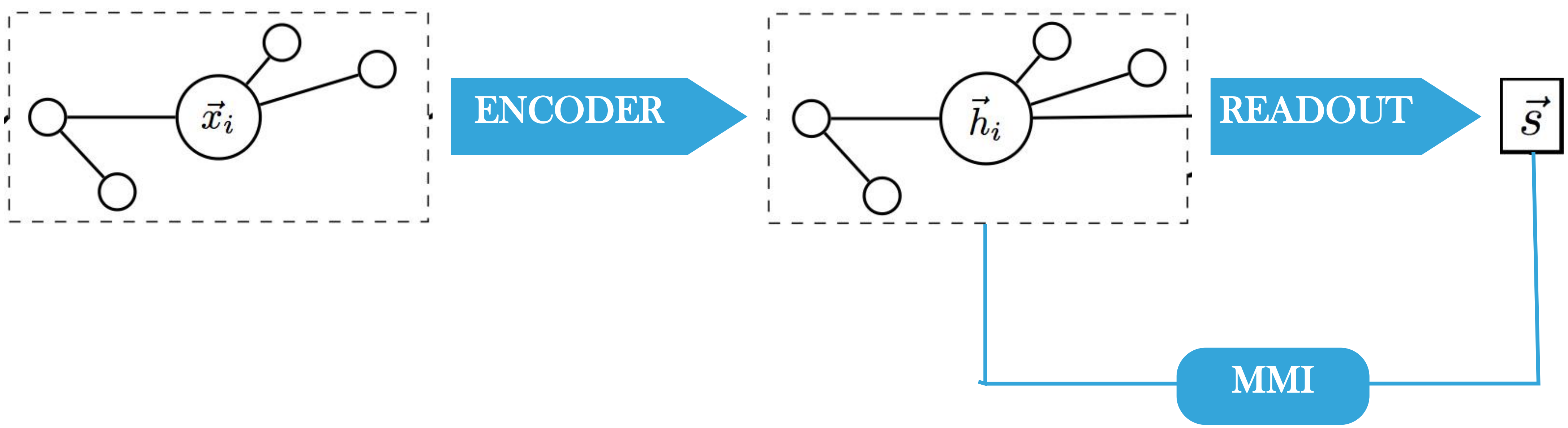
KEY POINT

- A general approach for learning node embeddings in an unsupervised manner.
- Maximize mutual information between a high-level "global" representation and "local" parts of the input.

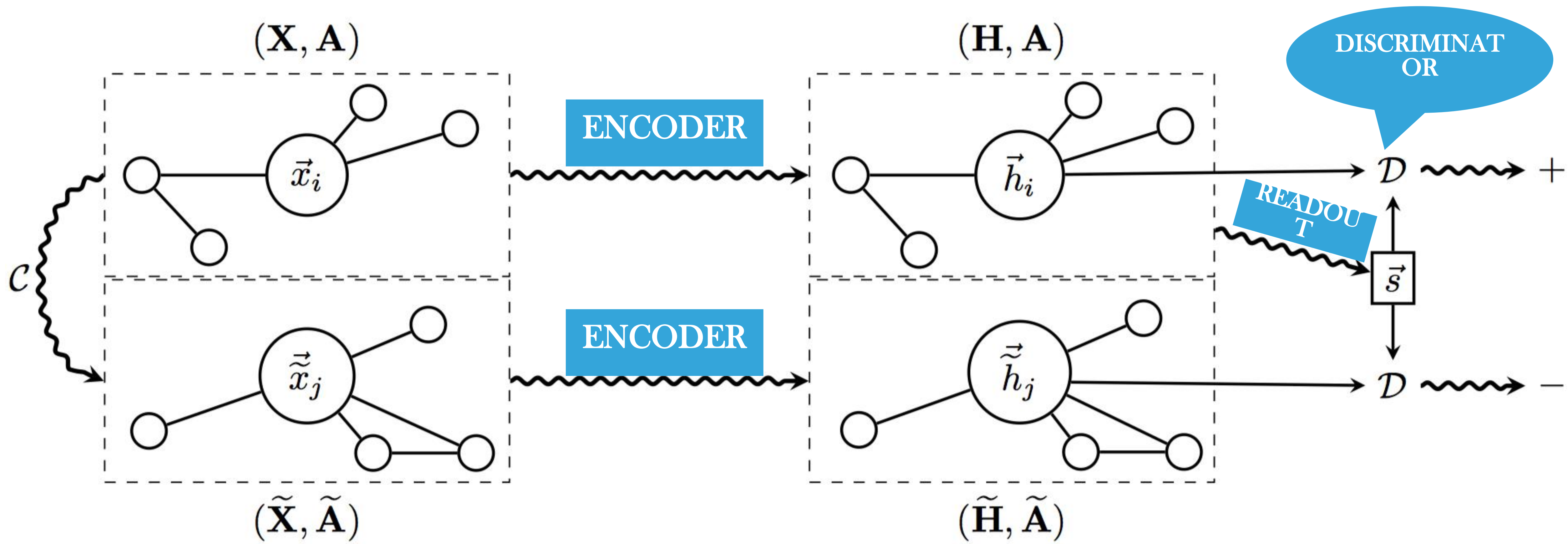


MUTUAL INFORMATION (MI)

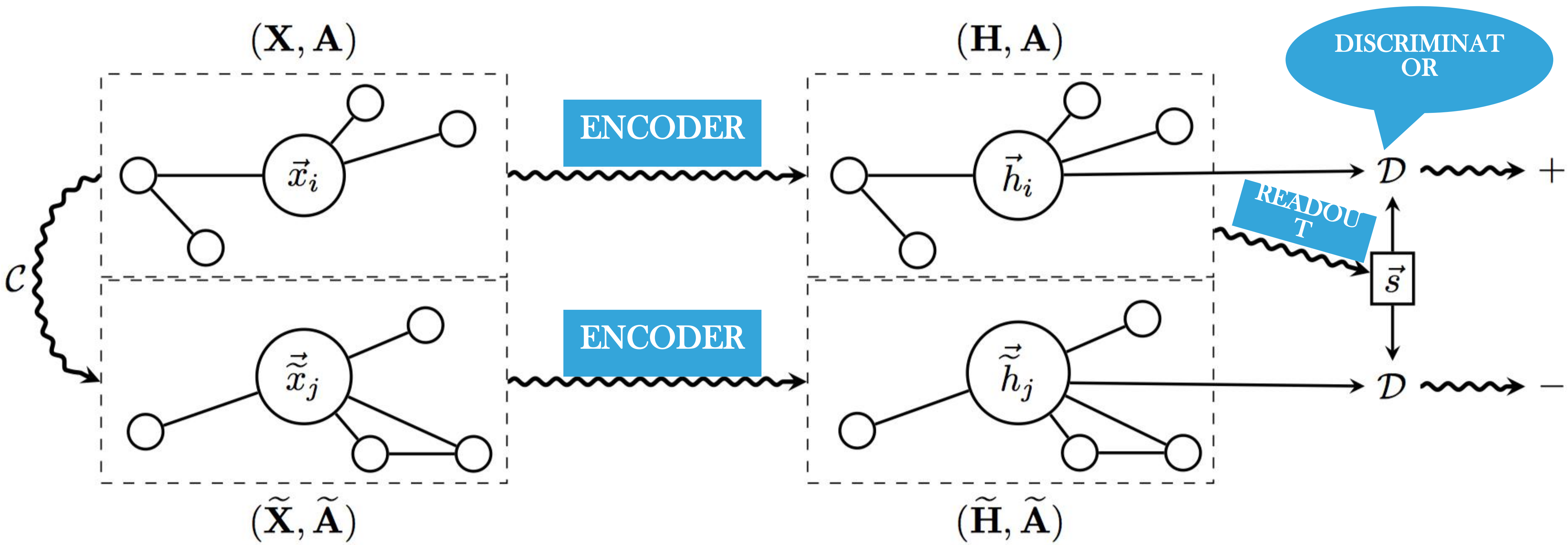
- *互信息*表示两个变量X与Y是否有关系，以及关系的强弱
- 通过考察其联合概率分布$p(X,Y)$与边缘概率分布乘积$p(X)p(Y)$之间的 KL散度来判断独立程度 $I(X,Y) = KL(p(X,Y)||p(X)p(Y))$



OVERVIEW OF DEEP GRAPH INFOMAX



OVERVIEW OF DEEP GRAPH INFOMAX



$$\mathcal{L} = \frac{1}{N+M}\left(\sum_{i=1}^{N}\mathbb{E}_{(\mathbf{X},\mathbf{A})}\left[\log \mathcal{D}\left(\vec{h}_i, \vec{s}\right)\right] + \sum_{j=1}^{M}\mathbb{E}_{(\widetilde{\mathbf{X}},\widetilde{\mathbf{A}})}\left[\log\left(1 - \mathcal{D}\left(\vec{\tilde{h}}_j, \vec{s}\right)\right)\right]\right)$$

EXPERIMENT

- Encoder
  - Transductive learning (GCN)
  - Inductive learning on large graphs(GraphSAGE-GCN)
- Readout $$\mathcal{R}(\mathbf{H}) = \sigma\left(\frac{1}{N}\sum_{i=1}^{N}\vec{h}_i\right)$$
- Discriminato $$\mathcal{D}(\vec{h}_i, \vec{s}) = \sigma\left(\vec{h}_i^T \mathbf{W} \vec{s}\right)$$