Hierarchical Graph Representation Learning with Differentiable Pooling

Rex Ying
rexying@stanford.edu
Stanford University

Jiaxuan You
jiaxuan@stanford.edu
Stanford University

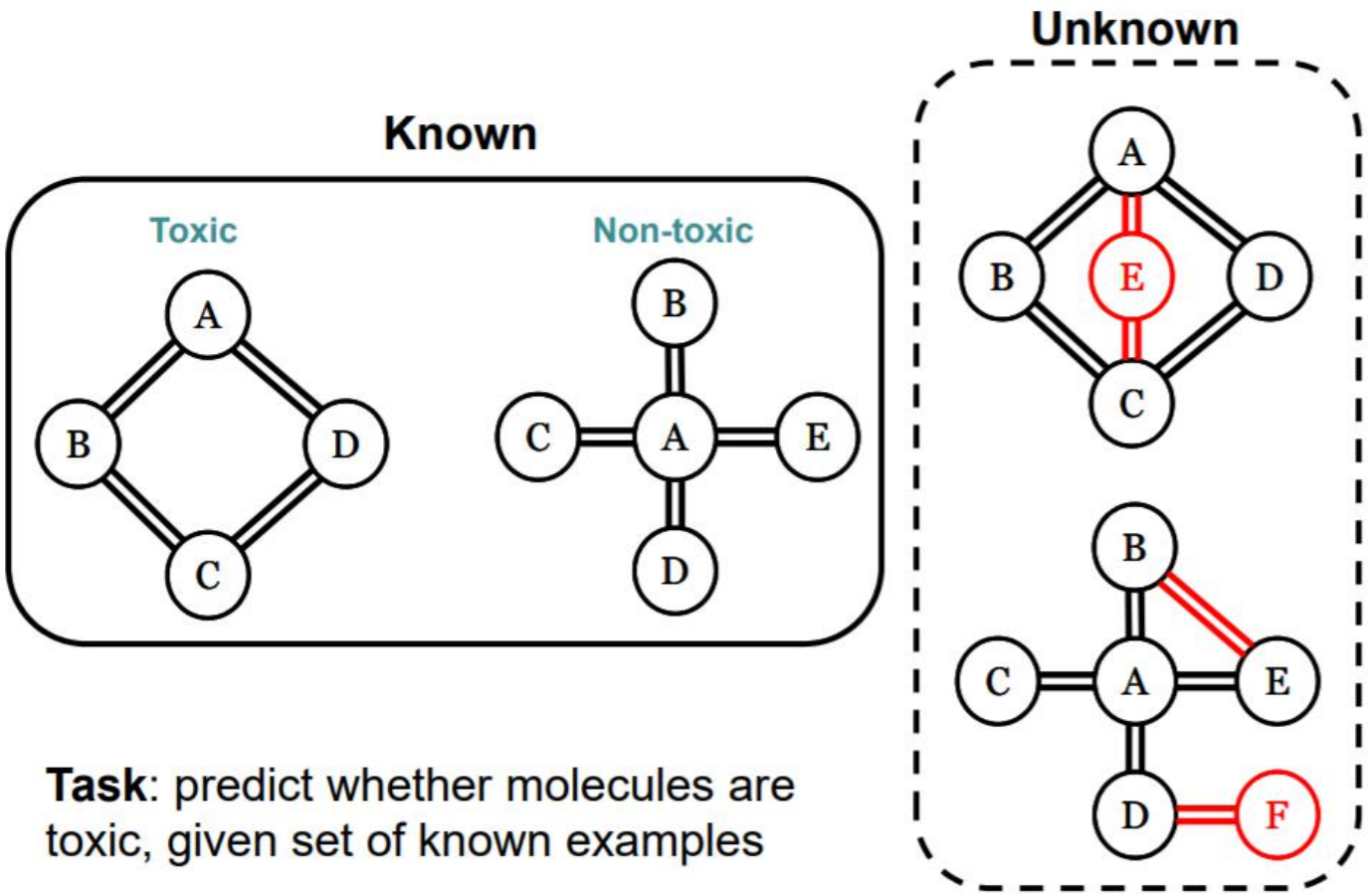
Christopher Morris
christopher.morris@udo.edu
TU Dortmund University

Xiang Ren
xiangren@usc.edu
University of Southern California

William L. Hamilton
wleif@stanford.edu
Stanford University

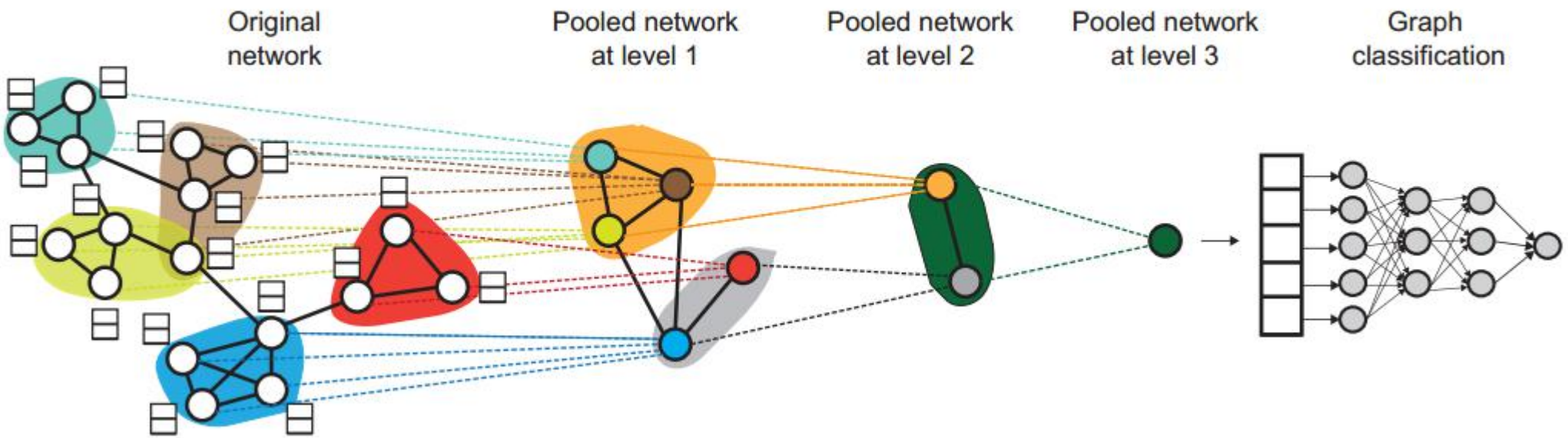
Jure Leskovec
jure@cs.stanford.edu
Stanford University

Graph classification



Task: predict whether molecules are toxic, given set of known examples

DiffPool


Original network
Pooled network at level 1
Pooled network at level 2
Pooled network at level 3
Graph classification

Graph neural networks

- G (A, F)
- Message passing architecture

$$H^{(k)} = M(A, H^{(k-1)}; \theta^{(k)}), \tag{1}$$

$$H^{(0)} = F$$

$$H^{(k)} = M(A, H^{(k-1)}; W^{(k)}) = \mathbf{ReLU}(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^{(k-1)}W^{(k-1)}), \tag{2}$$

Differentiable Pooling

- **Pooling with an assignment matrix:** learning a cluster assignment matrix of layer l+1 over the nodes using the output of at layer l

$$X^{(l+1)} = S^{(l)^T} Z^{(l)} \in \mathbb{R}^{n_{l+1} \times d}, \tag{3}$$

$$A^{(l+1)} = S^{(l)^T} A^{(l)} S^{(l)} \in \mathbb{R}^{n_{l+1} \times n_{l+1}}. \tag{4}$$

Differentiable Pooling

- **Learning the assignment matrix**

- $$Z^{(l)} = \mathrm{GNN}_{l,\mathrm{embed}}(A^{(l)}, X^{(l)}), \tag{5}$$

$$S^{(l)} = \mathrm{softmax}\left(\mathrm{GNN}_{l,\mathrm{pool}}(A^{(l)}, X^{(l)})\right), \tag{6}$$