Graph Markov Neural Nets
Summary

Presenter: Arshdeep Sekhon

https://qdata.github.io/deep2Read

Background: GNN vs SRL for semi-supervised object classification

GNN:

- $p(\boldsymbol{y}_V|\boldsymbol{x}_V) = \prod_{n \in V} p(\boldsymbol{y}_n|\boldsymbol{x}_V)$
- label distribution inferred independently

SRL:

- $p(\boldsymbol{y}_V|\boldsymbol{x}_V) = \frac{1}{Z(\boldsymbol{x}_V)} \prod_{i,j \in E} \psi_{i,j}(\boldsymbol{y}_i, \boldsymbol{y}_j, \boldsymbol{x}_V)$
- label distribution inference (usually intractable Why?)

blue color are points from reference template paper

Method: Graph Markov Neural network

- Intuition: Combine SRL(specifically CRF) + GNN
- Optimization Variational EM because

Figure: Caption

Key Idea

- $q_\theta(\boldsymbol{y_V}|\boldsymbol{x_V})$ (learn labels only from the features, similar to vanilla GCN)
- $p_\phi(\boldsymbol{y_U}|\boldsymbol{y_L}, \boldsymbol{x_V})$ (get label from surrounding labels as well as the node features/ CRF part)

blue color are points from reference template paper

Optimization

- E (Inference) Step: trying to learn $q_\theta(\boldsymbol{y_V}|\boldsymbol{x_V})$(parameterized as a GNN)
  - the labels you don't know use from the p_ϕ
  - also use the known to optimize
- M (Learning) Step: trying to learn $p_\phi(\boldsymbol{y_U}|\boldsymbol{y_L}, \boldsymbol{x_V})$parameterized as a GNN
  - labels you don't know use from q_θ

Optimization

- $O_\theta = O_{\theta,U} + O_{\theta,L}$
- $O_\phi = \sum_{n \in V} log p_\phi(\hat{\boldsymbol{y}}_n | \hat{\boldsymbol{y}}_{nb}, \boldsymbol{x}_V)$
- $Loss = O_\theta + O_\phi$
- First train q_θ, then alternatively both
- Both q_θ and p_ϕ can give final labels for objects

blue color are points from reference template paper

Optimization

- While learning q_θ need true labels for loss function: use the current ones from p_ϕ

optimization.png

blue color are points from reference template paper

Analysis of Amortized Inference

- No Amortized Inference vs with amortized inference