Recurrent Collective Classification

Shuangfei Fan, Bert Huang

Virginia Tech, Blacksburg, VA, USA

The iterative classification algorithm (ICA) is a canonical method for incorporating relational information into node classification [1, 2, 3]. Yet, existing methods for training ICA models rely on the assumption that relational features reflect the true labels of the nodes. Because the assumption is overly optimistic, it causes the learned models to cascade and amplify errors when the assumption is broken in early stages of prediction. In contrast, ICA uses predicted neighbor labels as feedback for each subsequent prediction, which means that if the model was trained expecting these predicted labels to be perfect, it will not be robust to situations where predictions are noisy or inaccurate. A classifier considering these neighbor labels should therefore consider common patterns of misclassification of neighbor labels. We propose the recurrent collective classification (RCC) framework. RCC trains models for iterative classification by treating the intermediate predictions as unknown values, making the models adaptive toward patterns of misclassification.

Figure 1 illustrates the interpretation of iterative collective classification as a recurrent neural network that takes local node features as input and recursively updates its estimates of node classifications. The RCC framework provides a general scheme for how to compute the gradient of the loss function with respect to the classifier parameters, enabling direct minimization of the collective classifier's error. In our experiments, this direct error minimization translates to improved accuracy and robustness on real network data, especially in settings where local classification is very noisy, settings that are particularly challenging for ICA.

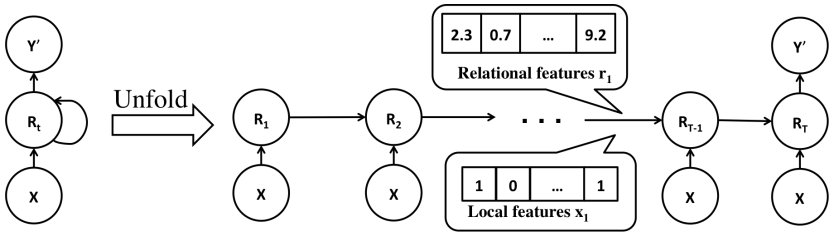

Figure 1: Structure of RCC/ICA prediction. The recurrent form (left) unrolls into a form (right) that explicitly considers each iteration as a separate operation.

Keywords: Collective classification; iterative classification algorithm; recurrent neural networks

[1] Ben London and Lise Getoor. "Collective Classification of Network Data." Data Classification: Algorithms and Applications. CRC Press, 2013.
[2] Jennifer Neville and David Jensen. "Iterative classification in relational data." AAAI Workshop on Learning Statistical Models from Relational Data. 2000.
[3] Prithviraj Sen, Galileo Namata, Mustafa Bilgic, Lise Getoor, Brian Gallagher, and Tina Eliassi-Rad. "Collective classification in network data." AI Magazine. 2008.