Node Classification by Message Passing

Yujia Zhu

Node classification is a common semi-supervised learning task in graph. One way to solve this problem is to use node embedding to represent nodes and add an independent classifier to do the classification task. In this article, we will discuss another method: message passing. The intuition of message passing is nodes in a graph have correlations. In the other word, nearby/connected nodes are more likely to have the same label. This is consistent to homophily (Similar nodes are typically close together or directly connected) in the real life.

Relational Classification

The basic idea of relational classification is that the class probability Y_v of node v is a weighted average of class probabilities of its neighbours. Relational Classification is an iterative method. Suppose we have a binary node classification problem. Then we use following steps to realize relational classification:

1. Initialization: For labelled nodes v, initialize label Y_v by its ground truth label $Y_v^* \in \{0, 1\}$. For unlabelled nodes v, initialize $Y_v = 0.5$.
2. Update all nodes **in a random order** until convergence or until maximum number of iterations is reached using following equations

$$P(Y_v = c) = \frac{1}{\sum_{(v,u)\in E} A_{vu}} \sum_{(v,u)\in E} A_{vu} P(Y_u = c)$$

This method is quite easy, thus having at least two shortcomings:

- Covergence is not guaranteed.
- Node feature is not used in the model (We only use relationships between nodes).

Iterative Classification

Iterative classification takes node feature into account. It will classify node v based on its attributes f_v and labels z_v of neighbour set N_v. Iterative classification method will train two classifiers:

1. $\phi_1(f_v)$: predict node label based on node feature vector f_v.
2. $\phi_2(f_v, z_v)$: predict label based on node feature vector f_v and summary z_v of labels of node v's neighbours.

The summary z_v is a vector and is hand-crafted. Possible designs of z_v include:

- Histogram of the number (or fraction) of each label in N_v
- Most common label in N_v
- Number of different labels in N_v

Iterative Classifiers work by two steps:

The first step is to train two classifiers ϕ_1 and ϕ_2. Two classifiers are trained independently on the training set. We can use linear models, neural networks and other classifiers to design classifiers.

The second step is to iterate. The iteration will be operated on the test set until convergence. Concrete steps are as follows.

1. Set label of node v by $\phi_1(f_v)$

2. Update z_v based on Y_u for $u \in N_v$
3. Update Y_v by $\phi_2(f_v, z_v)$
4. Go back to step 2 until convergence

But, convergence of iterative classification is still not guaranteed.

Summary

In this file, we cover how to leverage correlation in graphs to make prediction on nodes. Two crucial methods are relational classification and iterative classification. The first method only uses the correlation between nodes, while the second takes node feature into account.