Quantifying Biases in Attribute Inference on Social Networks

Espín-Noboa L., Karimi F., Ribeiro B., Lerman K., Wagner C.

MOTIVATION

Relational classification (RC) infers the class label of a node based on the class label of its neighbors. For example:

It can predict **your** political leaning based on the political leaning of your **friends**.

Our connections carry important information:
We tend to connect to people who are **similar** to us, but also **different** from us, e.g., because they are **popular**. All these connections shape the structure of the network.

DOES THE STRUCTURE OF THE NETWORK AFFECT PERFORMANCE AND BIAS OF RELATIONAL CLASSIFICATION?

METHODS

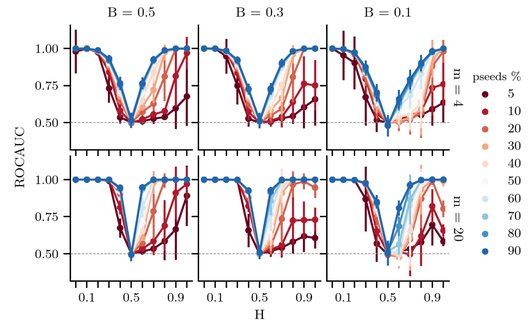
1. **Different types of social networks:**
   We generate synthetic *scale-free undirected* networks with different levels of: *homophily, fraction of minorities (group size), edge density, and number of nodes.*

2. **Modeling and inference:**
   We run RC using the *network-only Bayes classifier* (nBC) and *relaxation labeling* for modeling and inference, respectively.

3. **Evaluation and correlations:**
   We quantify correlations between the input and output in order to ***verify whether certain characteristics of the network can explain good/bad performance and biases against majorities/minorities.***

Network structure (*homophily, preferential attachment, group size, and edge density*) **explains performance** and **bias** in relational classification.

Click or **take a picture** to **see more details.**



RESULTS

- **Heterophilic networks, H<0.5**
  (when different nodes connect to each other)
  Achieve perfect classification (ROCAUC=1.0) and unbiased results with small training samples (e.g., 5% of nodes).

- **Homophilic networks, H>0.5**
  (when similar nodes connect to each other)
  Require at least 30% of all nodes in the training sample to achieve high performance (ROCAUC>0.7). Results are mostly biased towards majority nodes.

DISCUSSION

Why is heterophily easier to predict?
When training samples are small, the probability of correctly observing each type of edge is low. This probability is even lower in homophilic networks because two types of edges are required, whereas heterophilic networks require only one type. Additionally, when networks are unbalanced (B<0.5), heterophilic networks benefit from popular nodes (i.e., preferential attachment).

Lisette Espín-Noboa
@lespin
/lisette-espin
www.lisetteespin.info

gesis Leibniz-Institut für Sozialwissenschaften
USC University of Southern California Information Sciences Institute
PURDUE UNIVERSITY
UNIVERSITÄT KOBLENZ · LANDAU