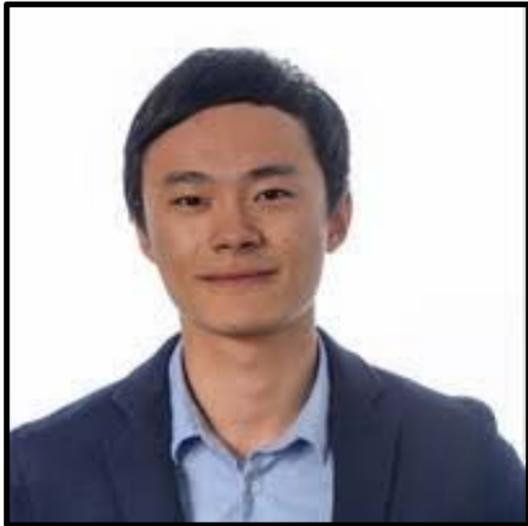




Thesis Proposal

GHC 6501 | Wednesday, October 9, 2019 | 4:00 pm



Learning with Graph Structures and Neural Networks

Yuexin Wu

Abstract

Graph-based learning focuses on the modeling of graphically structured data. Significant applications include the analysis of chemical compounds based on molecular structures, the prediction of solar-energy farm outputs based on radiation sensor network data, the forecasting of epidemiology outbreaks based on geographical relations among cities and social network interactions and so on. Algorithms for graph-based learning have been developed rapidly, addressing the following fundamental challenges:

- To encode the rich information about each individual node and node combinations in a graph, a.k.a. the graph-based representation learning challenge;
- To recover missing edges when graphs are only partially observed, a.k.a. the graph completion challenge;
- To leverage active learning in graphical settings when labeled nodes are highly sparse, a.k.a. the label sparse challenge;
- To enhance the tractability of training and inference over very large graphs, a.k.a. the scaling challenge.

This thesis aims to enhance graph-based machine learning from all the above aspects via the following key contributions:

1. Hierarchical learning of node embedding (H-Emb): Inspired by the great success of BERT and ELMO in recent natural language processing, we propose a new unsupervised learning approach to context-aware node embedding in graphical settings. Different from conventional BERT where the attentions come from the left context or right context of each word, our H-Emb uses all the paths through a specific node to define the rich set of attentions on that node.
2. Graph-enhanced neural learning for node classification: Assuming given graphs are usually incomplete (with missing links) or noisy, directly using them in graph-based classification would be sub-optimal. Thus, we propose to jointly regularize the given graph and optimize the parameters of an NNet during the training of the classifier as our new approach, which combines the strengths of neural classification and spectral transfer of the input graph.
3. Graph convolutional matrix factorization for bipartite edge prediction: For a specific category of graphs, i.e. bipartite graphs, traditional matrix factorization methods could not effectively leverage side information such as similarity measurements within the two groups of nodes. We, therefore, propose to use graph convolutions to enhance the learned factorized representations with the structured side information for better prediction accuracy.
4. Graph-enhanced Active Learning for node classification: Popular AL strategies with successful application in traditional data may not be directly applicable to graphs as they treat all the candidate documents as non-related instances. We propose an approach to active learning over graphs tailored for Graph Neural Networks, which takes both node-internal features and cross-node connections into account for node selection in AL.
5. Successful real-world applications of large-scale graph-based learning: We have investigated the applications of graph-based learning to a variety of real world problems, including multi-graph based collaborative filtering, graph-based transfer learning across languages, graph-based deep learning for epidemiology prediction, graph-enhanced node classification, edge detection and knowledge base completion; we obtained the state-of-the-art results in each of those domains.

https://drive.google.com/file/d/1ukMRRyzkDG4mULoWwA2kuCF5Zig-ik_k/view

COMMITTEE:

Yiming Yang, (Chair)

Aarti Singh

Leman Akoglu

Huan Liu,
(Arizona State University)