

Large-scale Machine Learning over Graphs

Yiming Yang

Language Technologies Institute, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213
yiming@cs.cmu.edu, <http://www.cs.cmu.edu/~yiming/>

ABSTRACT

Graphs provide powerful representations for statistical modeling of interrelated variables (observed or latent) in a broad range of machine learning applications. Examples include learning and inference based on the dependency structures among words, documents, topics, users, items, web sites, and more. How to best leverage such dependency structures from multiple graphs with massive and heterogeneous types of nodes and relations has posed grand challenges to machine learning theory and algorithms. This talk presents our recent work in this direction focusing on three significant tasks, including 1) a novel framework for fusing multiple heterogeneous graphs into a unified product graph to enable semi-supervised multi-relational learning, 2) the first algorithmic solution for imposing analogical structures in graph-based entity/relation embedding, and 3) a new formulation of neural architecture search as a graph topology optimization problem, with simple yet powerful algorithms that automatically discover high-performing convolutional neural architectures on image recognition benchmarks, and reduce the computational cost over state-of-the-art non-differentiable techniques by several orders of magnitude.

CCS Concepts/ACM Classifiers

- General and reference
- Information systems
- Mathematics of computing

Author Keywords

Learning over graphs, multi-relational learning, analogical structure learning, neural network architecture search, large-scale optimization algorithms

BIOGRAPHY

Yiming Yang is a professor at Carnegie Mellon University with a joint appointment in the Language Technologies Institute and the Machine Learning Department. She received her Ph.D. in Computer Science from Kyoto

University (Japan), and has been a faculty member at CMU since 1996. Her research has centered on machine learning techniques and scalable algorithms and a broad range applications, including extreme-scale text categorization, semi-supervised learning over heterogeneous graphs, analogical structured learning via entity/relation embedding, time series forecasting and anomaly detection, cross-language and cross-domain transfer learning over news media data, deep neural learning for spatiotemporal modeling of social and epidemiological networks, clustering with semi-supervised metric learning for new concept discovery, large-scale optimization for online advertising, and more.



REFERENCES

- [1] Hanxiao Liu, Yiming Yang: Cross-Graph Learning of Multi-Relational Associations. *ICML 2016*: 2235-2243
- [2] Hanxiao Liu, Yuexin Wu, Yiming Yang: Analogical Inference for Multi-relational Embeddings. *ICML 2017*: 2168-2178
- [3] Hanxiao Liu, Karen Simonyan, Yiming Yang: DARTS: Differentiable Architecture Search. *arXiv:1806.09055*.
- [4] Hanxiao Liu, Yiming Yang: Semi-Supervised Learning with Adaptive Spectral Transform. *AISTATS 2016*
- [5] Yiming Yang, Hanxiao Liu, Jaime G. Carbonell, Wanli Ma: Concept Graph Learning from Educational Data. *WSDM 2015*

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

ICTIR'18, September 14-17, 2018, Tianjin, China.

© 2018 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-5656-5/18/09.

DOI: <https://doi.org/10.1145/3234944.3240462>