AutoNet: Automated Network Construction from Massive Text Corpora

Jingbo Shang
University of Illinois at Urbana-Champaign

3/26/2019 at 01:30 pm
CoRE A 301

Abstract

Mining structured knowledge from massive unstructured text data is a key challenge in data science. In this talk, I will discuss my proposed framework, AutoNet, that transforms unstructured text data into structured heterogeneous information networks, on which actionable knowledge can be further uncovered flexibly and effectively. AutoNet is a data-driven approach using distant supervision instead of human curation and labeling. It first constructs networks using phrase mining, entity recognition, and relation extraction methods. It then builds topic taxonomies for further knowledge and insight discovery. Along this line, I have developed a number of state-of-the-art distantly-supervised/unsupervised methods and published them in top conferences and journals. Specifically, I will present my work about phrase/entity mining and taxonomy construction in details, while touching my other research topics slightly. Finally, I will summarize the AutoNet framework with a demo video and conclude by discussing future work collaborating with other disciplines.

Bio

Jingbo Shang is a Ph.D. candidate in Department of Computer Science, University of Illinois at Urbana-Champaign. He received his B.E. from Computer Science Department, Shanghai Jiao Tong University, China. His research focuses on mining and constructing structured knowledge from massive text corpora with minimum human effort. His research has been recognized by many prestigious awards, including Computer Science Excellence Scholarship from CS@Illinois, Grand Prize of Yelp Dataset Challenge in 2015, Google PhD Fellowship in Structured Data and Database Management in 2017, and C.W. Gear Outstanding Graduate Award in 2018. Moreover, his technologies have been transferred and applied in a wide spectrum of industries (e.g.,

high-tech, biomedical, financial) and open-source tools on Github have received thousands of stars and hundreds of forks.

Faculty Host: Yongfeng Zhang