

Accepted Manuscript

Knowledge triple mining via multi-task learning

Zhao Zhang, Fuzhen Zhuang, Xuebing Li, Zhengyu Niu, Jia He,
Qing He, Hui Xiong

PII: S0306-4379(17)30680-4
DOI: <https://doi.org/10.1016/j.is.2018.09.003>
Reference: IS 1336

To appear in: *Information Systems*

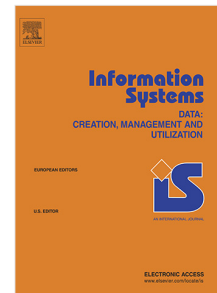
Received date: 20 November 2017

Revised date: 10 July 2018

Accepted date: 15 September 2018

Please cite this article as: Z. Zhang, et al., Knowledge triple mining via multi-task learning, *Information Systems* (2018), <https://doi.org/10.1016/j.is.2018.09.003>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



Knowledge Triple Mining via Multi-Task Learning

Zhao Zhang^{†,‡}, Fuzhen Zhuang^{*†,‡}, Xuebing Li[†],
Zhengyu Niu[¶], Jia He^{†,‡}, Qing He^{†,‡}, Hui Xiong[§]

[†]Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS),
Institute of Computing Technology, CAS, Beijing 100190, China

[‡]University of Chinese Academy of Sciences, Beijing 100049, China

[¶]Baidu Inc., Beijing, China

[§]Management Science and Information Systems Department
Rutgers Business School, Rutgers University, Newark, NJ, USA

*Corresponding Author, zhuangfuzhen@ict.ac.cn

Abstract

Recent years have witnessed the rapid development of knowledge bases (KBs) such as WordNet, Yago and DBpedia, which are useful resources in AI-related applications. However, most of the existing KBs are suffering from incompleteness and manually adding knowledge into KBs is inefficient. Therefore, automatically mining knowledge becomes a critical issue. To this end, in this paper, we propose to develop a model (S²AMT) to extract knowledge triples, such as <Barack Obama, wife, Michelle Obama>, from the Internet and add them to KBs to support many downstream applications. Particularly, because the seed instances¹ for every relation is difficult to obtain, our model is capable of mining knowledge triples with limited available seed instances. To be more specific, we treat the knowledge triple mining task for each relation as a single task and use multi-task learning (MTL) algorithms to solve the problem, because MTL algorithms can often get better results than single-task learning (STL) ones with limited training data. Moreover, since finding proper task groups is a fatal problem in MTL which can directly influence the final results, we adopt a clustering algorithm to find proper task groups to further improve the performance. Finally, we conduct extensive experiments on real-world datasets and the experimental results clearly validate the performance of our MTL algorithms against STL ones.

Keywords: multi-task learning, knowledge mining, relation extraction, knowledge graph construction

1. Introduction

Nowadays, knowledge bases (KBs) are extremely useful resources for query expansion, coreference resolution, question answering and information retrieval.

¹In this paper, seed instances refer to labeled positive instances.

Download English Version:

<https://daneshyari.com/en/article/11263013>

Download Persian Version:

<https://daneshyari.com/article/11263013>

[Daneshyari.com](https://daneshyari.com)