Knowledge Graph Embedding with Transfer Learning on Biterm Topic Model of Entity Description

Qiang Guo,Bingbing Jiang,Liangang Zhang,Huanhuan Chen

Abstract

Knowledge graph embedding aims to encode both entities and relations into a continuous low-dimensional space. Most methods concentrate on the structure information of the knowledge graph whose nodes represent the entities and edges represent the relations between entities. In fact,there are usually concise text for entities,which can be modeled with topic model and well utilized by existing embedding methods. In this paper,we propose a novel method for knowledge graph embedding by uncovering the topics within the short entities descriptions and transferring the topic model to entity embedding. Specifically,we learn the topics by directly modeling the generation of word co-occurrence patterns in the whole entity description and transfer the topic model of the entity description to the knowledge embedding with transfer learning. We evaluate our method on two tasks,including knowledge graph completion and entity classification. Experimental results on Freebase show that,by special modeling the topic of the entity description and transfer learning,our method can significantly outperform state-of-the-art method.

1 Introduction

Hello World.

2 Related Work

Hello World.

2.1 TransE

Given a training set D of triples (h,r,t) composed of two entities h,t in E(the set of entities) and a relationship r in R(the set of relationships),TranE learns the vector embeddings of the entities and the relationships,by minimizing a margin-based ranking criterion over the training set:

2.2 Biterm Topic Model

2.3 Tranfer Learning

Hello World.

3 Our Approach

In this section,we introduce the transfer learning based TransE(TransT) that learns representations of entities and relations utilizing the topic model of short entity description. In TransE and TransT,we have entity set **E** and relation set **R**,and learn to encode both entities and relations in R. Given a KG represented by a set of triples S=(h,r,t) with each triple composed of two entities h,t E and their relation r R,there are also concise text for most entities in the KG. We learn the topics by directly modeling the generation of entity co-occurrence patterns in the entities description corpora by Biterm topic model and transfer the topic model of the entity description to the knowledge embedding with transfer learning method.

3.1 Transferring the Biterm Topic Model to Entities Embedding

From the task of Biterm topic model,we learn the topic distribution of entity description which can represent the semantic of the entities well.Each topic model of the entity can transfer to its embedding by the feature representation transferring method.

$$h^* = h + \lambda \cdot o, t^* = t + \lambda \cdot o \quad (1)$$

$$L = \sum |h^* + r - t^*| \quad (2)$$

3.2 Training the Whole Knowledge Graph

Format manuscripts two columns to a page, in the manner these instructions are formatted. The exact dimensions for a page on A4 paper are:

- Left and right margins: 2.5 cm
- Top margin: 2.5 cm
- Bottom margin: 2.5 cm
- Column width: 7.7 cm
- Column height: 24.7 cm
- Gap between columns: 0.6 cm

Papers should not be submitted on any other paper size. If you cannot meet the above requirements about the production of your electronic submission, please contact the publication chairs above as soon as possible.

4 Experiments

Following ACL 2014 we will also we will attempt to automatically convert your LaTeX source files to publish papers in machine-readable XML with semantic markup in the ACL Anthology, in addition to the traditional PDF format. This will allow us to create, over the next few years, a growing corpus of scientific text for our own future research, and picks up on recent initiatives on converting ACL papers from earlier years to XML.

We encourage you to submit a ZIP file of your LaTeX sources along with the camera-ready version of your paper. We will then convert them to XML automatically, using the LaTeXML tool (`http://dlmf.nist.gov/LaTeXML`). LaTeXML has *bindings* for a number of LaTeX packages, including the ACL 2015 stylefile. These bindings allow LaTeXML to render the commands from these packages correctly in XML. For best results, we encourage you to use the packages that are officially supported by LaTeXML, listed at `http://dlmf.nist.gov/LaTeXML/manual/included.bindings`

5 Conclusion and Future Work

It is also advised to supplement non-English characters and terms with appropriate transliterations and/or translations since not all readers understand all such characters and terms. Inline transliteration or translation can be represented in the order of: original-form transliteration "translation".

Acknowledgments

The acknowledgments should go immediately before the references. Do not number the acknowledgments section. Do not include this section when submitting your paper for review.

References

Alfred V. Aho and Jeffrey D. Ullman. 1972. *The Theory of Parsing, Translation and Compiling*, volume 1. Prentice-Hall, Englewood Cliffs, NJ.

American Psychological Association. 1983. *Publications Manual*. American Psychological Association, Washington, DC.

Association for Computing Machinery. 1983. *Computing Reviews*, 24(11):503–512.

Ashok K. Chandra, Dexter C. Kozen, and Larry J. Stockmeyer. 1981. Alternation. *Journal of the Association for Computing Machinery*, 28(1):114–133.

Dan Gusfield. 1997. *Algorithms on Strings, Trees and Sequences*. Cambridge University Press, Cambridge, UK.