

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

Qiang Guo
UBRI in USTC

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 of entity description and transfer the topic model of the entity description to the knowledge embedding with transfer learning method. 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.

Introduction

Hello world!

Related Work

TransE

Inspired by the word embedding work (Mikolov et al. 2013), TransE (Bordes et al. 2013) model translates the latent feature representations via a relation-specific offset between head and tail entities on the low-dimensional vector space. To learn such embeddings, TransE minimize a margin-based ranking criterion over the training set:

$$L = \sum_{(h,r,t) \in S} \sum_{(h',r,t') \in S'_{(h,r,t)}} [\gamma + d(h+r, t) - d(h'+r, t')]_+ \quad (1)$$

where $[x]_+$ denotes the positive part of x , $\gamma > 0$ is a margin hyperparameter, and

$$S'_{(h,r,t)} = \{(h', r, t) | h' \in E\} \cup \{(h, r, t') | t' \in E\} \quad (2)$$

The optimization is carried out by stochastic gradient descent (in minibatch mode), over the possible h, r and t , with

Copyright © 2016, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

the additional constraints that the L_2 -norm of the embeddings of the entities is 1 (no regularization or norm constraints are given to the label embeddings r).

Topic Model for Short Texts

Topic models have been proposed to uncover the latent semantic structure from text corpus. The conventional topic models, such as LDA and PLSA, implicitly capture the document-level word co-occurrence patterns to reveal topics and thus suffer from the severe data sparsity in short text. BTM (Yan, 2013) can discover more prominent and coherent topics for short text, which explicitly models the word co-occurrence patterns to enhance the topic learning and uses the aggregated patterns in the whole corpus.

Transfer Learning

The study of transfer learning is motivated by the fact that people can intelligently apply knowledge learned previously to solve new problems faster or with better solutions. Given a source domain D_S and learning task T_S , a target domain D_T and learning task T_T , transfer learning aims to help improve the learning of the target predictive function $f_T(\cdot)$ in D_T using the knowledge in D_S and T_S , where $D_S \neq D_T$, or $T_S \neq T_T$.

Our Method

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 \mathbb{R}^k . Given a KG represented by a set of triples $S = (h, r, t)$ with each triple composed of two entities $h, t \in E$ and their relation $r \in R$, 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.

Transferring the Biterm Topic Model to Entities Embedding

From the task of Biterm topic model, we learn the topic distribution θ of entity description which can represent the se-

mantic of the entities well. Each topic model of the entity can transfer to its embedding by the feature representation transferring method.

$$h^* = h + \lambda \cdot \theta_h, t^* = t + \lambda \cdot \theta_t \quad (3)$$

Training the Whole Knowledge Graph

We minimize the following margin-based score function as objective for training:

$$L = \sum_{(h,r,t) \in S} \sum_{(h',r',t') \in S'_{(h,r,t)}} [\gamma + d(h^* + r, t^*) - d(h'^* + r', t'^*)]_+ \quad (4)$$

Experiment

Datasets and Experiment Settings

Datasets In this paper, we adopt FB15K (Bordes et al. 2013), a dataset extracted from a typical large-scale KG Freebase (Bollacker et al. 2008), to evaluate the TranTL model on knowledge graph completion and entity classification.

Table 1: Statistic of data sets

Dataset	#Rel	#Ent	#Train	#Valid	#Test
FB15K	1341	14904	472860	48991	57803

Parameter Settings We implement TransE, TransT for comparison. We train those models with entity/relation dimension n in $\{30, 60, 90\}$. Following (Bordes et al. 2013), we use a fixed learning rate λ among $\{0.0005, 0.001, 0.002\}$, and margin γ among $\{0.5, 1.0, 1.5, 2.0\}$. We also use a fixed transferring rate α among $\{0.4, 0.6, 0.8, 1\}$. The optimal configurations of TransT are: $\lambda = 0.002, \gamma = 1.0, n = 60, \alpha = 1$.

Knowledge Graph Completion

The task of knowledge graph completion aims to complete a triple (h, r, t) when one of h, r, t is missing based on minimizing the score function $S(h, r, t) = ||(h + \alpha\theta_h) + r - (t + \alpha\theta_t)||$

Evaluation We conduct our evaluation on FB15K and consider the knowledge graph completion task as two sub-tasks: entity prediction and relation prediction. Following (Bordes et al. 2013), we use two measures as our evaluation metrics: (1) mean rank of correct entities; (2) proportion of valid entities ranked in top 10 (for entity) or top 1 (for relation). We also follow the two evaluation settings named as "Raw" and "Filter".

Result The results of entity prediction and relation prediction are in Table 2 and Table 3.

Entity Classification

The task of entity classification is a multilabel classification task aiming to predict entity types, which is crucial and widely used in many NLP tasks.

Table 2: Evaluation results on entity prediction

Metric	Mean Rank		Hits@10(%)	
	Raw	Filter	Raw	Filter
TransE	210	119	48.5	66.1
TransT	0	0	0	0

Table 3: Evaluation results on relation prediction

Metric	Mean Rank		Hits@10(%)	
	Raw	Filter	Raw	Filter
TransE	2.91	2.53	69.5	90.2
TransT	0	0	0	0

Evaluation In training, we use entity representations learned by BTM, TransE, TransT trained on FB15K as features. We use Logistic Regression as classifier and one-versus-rest for multilabel classification. For evaluation following (Neelakantan and Chang 2015), we use mean average precision (MAP) which is commonly used in multilabel classification as the evaluation method for entity classification.

Result From Table 4 we observe that TransT outperforms all other models in FB15K.

Table 4: Evaluation results on entity classification

Metric	FB15K
TransE	87.9
BTM	87.9
TransT	87.9

Conclusion and Future Work

In this paper, we introduced a novel method of transferring the topic model of entity description to knowledge embedding. In experiments, we evaluate our method on knowledge graph completion and entity classification. Experimental results show that our method achieves better performances than the baselines on both tasks, which indicates the capability of building representations from entity descriptions.

We will explore the following research directions in future: (1) we will adopt other representation method of entity description, just like *tdidf* or *doc2vec*. (2) there are various of information like textual information of relation or entity types, we will transfer the representation learning of these information to knowledge graph embedding. (3) we will combine different representation of entity description for transfer learning of knowledge graph embedding.

Acknowledgments

AAAI is especially grateful to Peter Patel Schneider for his work in implementing the *aaai.sty* file, liberally using the ideas of other style hackers, including Barbara Beeton. We also acknowledge with thanks the work of George Ferguson for his guide to using the style and BibTeX files — which has been incorporated into this document — and Hans Guesgen, who provided several timely modifications, as well as

the many others who have, from time to time, sent in suggestions on improvements to the AAAI style.