Lifted Rule Injection for Relation Embeddings

[Abstract] *

Thomas Demeester
Ghent University - iMinds
Ghent, Belgium
tdmeeste@intec.ugent.be

Tim Rocktäschel
University College London
London, UK
t.rocktaschel@cs.ucl.ac.uk

Sebastian Riedel
University College London
London, UK
s.riedel@cs.ucl.ac.uk

ABSTRACT

Methods based on representation learning currently hold the state-of-the-art in many natural language processing and knowledge base inference tasks. Yet, a major challenge is how to efficiently incorporate commonsense knowledge into such models. A recent approach regularizes relation and entity representations by propositionalization of first-order logic rules. However, propositionalization does not scale beyond domains with only few entities and rules. We present a highly efficient method for incorporating implication rules into distributed representations for automated knowledge base construction. We map entity-tuple embeddings into an approximately Boolean space and encourage a partial ordering over relation embeddings based on implication rules mined from WordNet. Surprisingly, we find that the strong restriction of the entity-tuple embedding space does not hurt the expressiveness of the model and even acts as a regularizer that improves generalization. By incorporating few commonsense rules, we achieve an increase of 2 percentage points mean average precision over a matrix factorization baseline, while observing a negligible increase in runtime.

Keywords

Relation extraction, matrix factorization, implication rules

1. MOTIVATION AND CONTRIBUTIONS

This abstract introduces our recent work [1], where we introduce a highly efficient way to inject first-order implication rules that incorporate commonsense knowledge, while training a relation extraction algorithm for knowledge base population.

The considered relation extraction setting is based on the so-called universal schema [2], in which during training, knowledge base entries can be used hand in hand with surface form representations of the target relations extracted from free-text corpora. Learning to predict whether any of the studied relations hold for a given pair of entities, can be cast into a matrix factorization problem: each relation is represented by a real-valued vector, as well as each pair of entities occurring in the data. During training, these representations or embeddings are learned such that a score obtained by combining them (e.g., by means of the dot product), reflects the validity of the relation between the considered entities.

As recently studied [3], jointly training on the actual training facts and on well-chosen first-order logic rules, leads to improved predictions. For example, injecting the rule `professor_at(X,Y)` $\Rightarrow$ `employee_at(X,Y)`, could lead to improved predictions for the involved relations.

However, the proposed technique of [3] was based on the propositionalization of the formulas in terms of all training facts involved. This means each fact including one of the relations in the rule would contribute a new term to the training objective. As a result, the method does not scale, and in practice only a handful of rules can be injected during training. In our recent work [1], we introduce a method to inject first-order implication rules in a *lifted* way, i.e., independent of the entity pair representations. This turns out to be possible, provided all entity pair embeddings are non-negative, a restriction that appears not to restrict the model's expressiveness, but instead acts as a means of regularization.

The most important advantage of the new method is the fact that it scales very well with the number of rules, such that injecting thousands of rules barely has any effect on the training duration. The model has some other interesting properties, discussed at length in [1].

The relevance of our research in Natural Language Processing in not directly obvious from the general Information Retrieval perspective. However, it could be directly relevant for research in recommendation systems. In fact, the paper could be entirely translated into that domain, where the matrix factorization method underlying our work is a highly popular technique as well. This means external knowledge, for example in the form of logic implication rules, could be injected into recommendation engines. A possible application would be to use such rules to help overcome the cold-start problem.

2. REFERENCES

[1] T. Demeester, T. Rocktäschel, and S. Riedel. Lifted rule injection for relation embeddings. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, November 2016.

[2] S. Riedel, L. Yao, A. McCallum, and B. M. Marlin. Relation extraction with matrix factorization and universal schemas. In *Annual Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL)*, pages 74–84, 2013.

[3] T. Rocktäschel, S. Singh, and S. Riedel. Injecting Logical Background Knowledge into Embeddings for Relation Extraction. In *Annual Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL)*, 2015.

*This abstract appears at *DIR '16, November 25, 2016, Delft, The Netherlands.* It introduces our recent work [1].