Learning to Explain Entity Relationships in Knowledge Graphs

extended abstract

Voskarides, N.; Meij, E.; Tsagkias, M.; de Rijke, M.; Weerkamp, W.

DxR 2015 Amsterdam

Proceedings

Proceedings of the 14th Dutch-Belgian Information Retrieval Workshop

Learning to Explain Entity Relationships
in Knowledge Graphs

[Extended Abstract] [*]

Nikos Voskarides[†]
University of Amsterdam
Amsterdam, The Netherlands
n.voskarides@uva.nl

Edgar Meij
Yahoo Labs
London, United Kingdom
emeij@yahoo-inc.com

Manos Tsagkias
904 Labs
Amsterdam, The Netherlands
manos@904labs.com

Maarten de Rijke
University of Amsterdam
Amsterdam, The Netherlands
derijke@uva.nl

Wouter Weerkamp
904 Labs
Amsterdam, The Netherlands
wouter@904labs.com

ABSTRACT

Knowledge graphs are a powerful tool for supporting a large spectrum of search applications including ranking, recommendation, exploratory search, and web search [3]. A knowledge graph aggregates information around entities across multiple content sources and links these entities together, while at the same time providing entity-specific properties (such as age or employer) and types (such as actor or movie). Although there is a growing interest in automatically constructing knowledge graphs, e.g., from unstructured web data [5, 2], the problem of providing evidence on why two entities are related in a knowledge graph remains largely unaddressed. Extracting and presenting evidence for linking two entities, however, is an important aspect of knowledge graphs, as it can enforce trust between the user and a search engine, which in turn can improve long-term user engagement, e.g., in the context of related entity recommendation [1].

We propose a method for explaining the relationship between two entities, which we evaluate on a newly constructed annotated dataset that we make publicly available.[1] In particular, we consider the task of explaining relationships between pairs of Wikipedia entities. We aim to infer a human-readable description for an entity pair given a relationship between the two entities. Since Wikipedia does not explicitly define relationships between entities we use a knowledge graph to obtain these relations. We cast our task as a sentence ranking problem: we automatically extract sentences from a corpus and rank them according to how well they describe a given relationship between a pair of entities. For ranking purposes, we extract a rich set of features and use learning to rank to effectively combine them. Our feature set includes both traditional information retrieval and natural language processing features that we augment with entity-dependent features. These features leverage informa-

tion from the structure of the knowledge graph. On top of this, we use features that capture the presence in a sentence of the relationship of interest. For our evaluation we focus on "people" entities and we use a large, manually annotated dataset of sentences.

Our main contributions are a robust and effective method for explaining entity relationships, detailed insights into the performance of our method and features, and a large manually annotated dataset. Our evaluation shows that our method significantly outperforms state-of-the-art sentence retrieval models for this task. Experimental results also show that using relationship-dependent models is beneficial. Our feature analysis shows that relationship type features are the most important, although entity type features are important as well. This indicates that introducing features based on entities identified in the sentences and the relationship is beneficial for this task. Furthermore, the limited dependency on the source feature type indicates that our method might be able to generalize in other domains. Finally, text type features do contribute to retrieval effectiveness, although not significantly.

REFERENCES

[1] R. Blanco, B. B. Cambazoglu, P. Mika, and N. Torzec. Entity recommendations in web search. In *The Semantic Web–ISWC 2013*.

[2] M. Craven, D. DiPasquo, D. Freitag, A. McCallum, T. Mitchell, K. Nigam, and S. Slattery. Learning to construct knowledge bases from the world wide web. *Artificial Intelligence*, (1–2):69–113.

[3] X. Dong, E. Gabrilovich, G. Heitz, W. Horn, N. Lao, K. Murphy, T. Strohmann, S. Sun, and W. Zhang. Knowledge vault: A web-scale approach to probabilistic knowledge fusion. In *KDD 2014*.

[4] N. Voskarides, E. Meij, M. Tsagkias, M. de Rijke, and W. Weerkamp. Learning to explain entity relationships in knowledge graphs. *ACL-IJCNLP 2015*.

[5] J. Weston, A. Bordes, O. Yakhnenko, and N. Usunier. Connecting language and knowledge bases with embedding models for relation extraction. In *EMNLP 2013*.

---

[*]The full version of this paper was published at ACL-IJNLP 2015 [4].

[†]This work was carried out while this author was visiting Yahoo Labs.

[1]https://github.com/nickvosk/acl2015-dataset-learning-to-explain-entity-relationships