Outline

Alex Baier
abaier@uni-koblenz.de

January 18, 2017

1 Introduction

Motivation. Related work. Solution. Evaluation.

2 Foundations

2.1 Wikidata

What is Wikidata? Collaborative, people-driven, etc.. Open World Assumption.
What is an entity? What are statements? References? Qualifiers?
Informal explanation of taxonomy in the use case of Wikidata.
Show problems in Wikidata's taxonomy.
Galárraga [7]

2.2 Taxonomy

- Ontology
  Cimiano et al. [3]
  Galárraga [7]
- Taxonomy
  Cimiano et al. [3]
  Galárraga [7]
- Unconnected taxonomy as in the case of Wikidata
- Root class
- Unlinked class
- Problem statement
  challenge: potentially very high number of target classes

2.3 Similarity

- semantic similarity e.g. distributional, feature-based
  Lin [12]
  Rodríguez and Egenhofer [16]
- geometrical similarity e.g. distance based, cosine

2.4 Similarity-based classification

kNN. Maybe SVM, which is not appropriate for the task at hand, because the number of target classes is too high. Chen et al. [2]
Zhang and Zhou [19]

2.5 Text processing

- N-Gram
  What is an N-Gram? how can it be used? Jurafsky and Martin [10]
- Skip-Gram
  What is a Skip-Gram? Benefits in comparison to N-Grams. Guthrie et al. [9]
- Counting-based word representations
  Short explanation is sufficient. Levy et al. [11]
- Predictive word representations
  Shortly mention models by Mikolov et al.. Maybe mention GloVe too. Levy et al. [11]

3 Analysis of the Wikidata taxonomy

How can classes and unlinked classes be recognized in Wikidata? Analyze whole taxonomy. Identify "relevant" unlinked classes. Repeat analysis for these classes. Summarize important observations.

4 Ontology learning

General concepts. Tasks in ontology learning. Classification of considered problem in the task of ontology learning. Solutions for similar problems.
Cimiano et al. [3]
Wong et al. [18]
d'Amato et al. [4]
Petrucci et al. [14]
Fu et al. [6]

5 Neural networks

Notion of neural networks will be introduced.
Show a schematic for every model.
Purpose of specific NN. How can it be used for the task at hand? What are problems with the NN?

5.1 Recursive neural networks for graph representation

Scarselli et al. [17]

5.2 Deep neural networks for graph representation

Cao et al. [1]
Raghu et al. [15]

5.3 Continuous Bag-of-Words

Mikolov et al. [13]

5.4 Skip-gram with negative sampling

Mikolov et al. [13]
Levy et al. [11]
Goldberg and Levy [8]

5.5 Comparison

Make a decision, which NN will be used for solving the task. Skip-gram with negative sampling.

6 Algorithm

Do I need to explain the actual implementation in detail? It is really not that interesting.

6.1 Challenges

Identify and summarize the challenges of the task.
High number of target classes. Big amount of possibly usable data, what is relevant, what is not?

6.2 Baseline

- Architecture
  Using word embeddings of classes in kNN
  How does this solve the problem definition? How are the challenges addressed?

- Hyper parameters
  configuration of the model, explain decision
  Levy et al. [11]

- Training data
  mapping relevant Wikidata entities to a textual representation

6.3 Variation 1

supplement training data with Wikipedia, what are possible advantages? More data, Wikipedia contains "soft" information? which cannot be represented in Wikidata. what are disadvantages? Wikipedia contains "noise" unlike Wikidata, training will take much longer. How to combine embeddings of Wikidata and Wikipedia? just appending, other options?

7 Evaluation

7.1 Method

Dellschaft and Staab [5]

7.2 Generation of gold standard

pick number (tbd) of random linked classes. the distribution of instances, subclasses, properties per class should similar to the repeated analysis of the observed unlinked classes. remove the subclass properties of the chosen classes. also generate a new set of training data, which reflects the changes, and train the model on this modified data.

7.3 Results

compare baseline algorithm and variation(s). do the results match with my expectations? if not why could this be? mention training and execution time of the algorithms.

8 Conclusion

what was learned? can the developed solution be used in practical application? what about future work?

References

[1] Shaosheng Cao, Wei Lu, and Qiongkai Xu. Deep neural networks for learning graph representations. In Dale Schuurmans and Michael P. Wellman, editors, AAAI, pages 1145–1152. AAAI Press, 2016. URL http://dblp.uni-trier.de/db/conf/aaai/aaai2016.html#CaoLX16.

[2] Yihua Chen, Eric K. Garcia, Maya R. Gupta, Ali Rahimi, and Luca Cazzanti. Similarity-based classification: Concepts and algorithms. J. Mach. Learn. Res., 10: 747–776, June 2009. ISSN 1532-4435. URL http://dl.acm.org/citation.cfm?id=1577069.1577096.

[3] P. Cimiano, A. Mädche, S. Staab, and J. Völker. Ontology learning. In S. Staab and R. Studer, editors, Handbook on Ontologies, International Handbooks on Information Systems, pages 245–267. Springer, 2nd revised edition edition, 2009. URL http://www.uni-koblenz.de/~staab/Research/Publications/2009/handbookEdition2/ontology-learning-handbook2.pdf.

[4] Claudia d'Amato, Steffen Staab, Andrea G. B. Tettamanzi, Tran Duc Minh, and Fabien L. Gandon. Ontology enrichment by discovering multi-relational association rules from ontological knowledge bases. In Sascha Ossowski, editor, SAC, pages 333–338. ACM, 2016. ISBN 978-1-4503-3739-7. URL http://dblp.uni-trier.de/db/conf/sac/sac2016.html#dAmatoSTMG16.

[5] Klaas Dellschaft and Steffen Staab. On how to perform a gold standard based evaluation of ontology learning. In Proceedings of the 5th International Conference on The Semantic Web, ISWC'06, pages 228–241, Berlin, Heidelberg, 2006. Springer-Verlag. ISBN 3-540-49029-9, 978-3-540-49029-6. doi: 10.1007/11926078_17. URL http://dx.doi.org/10.1007/11926078_17.

[6] Ruiji Fu, Jiang Guo, Bing Qin, Wanxiang Che, Haifeng Wang, and Ting Liu. Learning Semantic Hierarchies via Word Embeddings. Acl, pages 1199–1209, 2014.

[7] Luis Galárraga. Rule Mining in Knowledge Bases. PhD thesis, Telecom ParisTech, 2016.

[8] Yoav Goldberg and Omer Levy. word2vec explained: deriving mikolov et al.'s negative-sampling word-embedding method. CoRR, abs/1402.3722, 2014. URL http://arxiv.org/abs/1402.3722.

[9] David Guthrie, Ben Allison, W. Liu, Louise Guthrie, and Yorick Wilks. A closer look at skip-gram modelling. In Proceedings of the Fifth international Conference on Language Resources and Evaluation (LREC-2006), Genoa, Italy, 2006.

[10] Daniel Jurafsky and James H. Martin. N-Grams. In Speech and Language Processing, chapter N-Grams. 2014.

[11] Omer Levy, Yoav Goldberg, and Ido Dagan. Improving distributional similarity with lessons learned from word embeddings. Transactions of the Association for Computational Linguistics, 3:211–225, 2015. ISSN 2307-387X. URL https://transacl.org/ojs/index.php/tacl/article/view/570.

[12] Dekang Lin. An information-theoretic definition of similarity. In Proceedings of the Fifteenth International Conference on Machine Learning, ICML '98, pages 296–304, San Francisco, CA, USA, 1998. Morgan Kaufmann Publishers Inc. ISBN 1-55860-556-8. URL http://dl.acm.org/citation.cfm?id=645527.657297.

[13] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. Efficient estimation of word representations in vector space. CoRR, abs/1301.3781, 2013. URL http://arxiv.org/abs/1301.3781.

[14] Giulio Petrucci, Chiara Ghidini, and Marco Rospocher. Using recurrent neural network for learning expressive ontologies. CoRR, abs/1607.04110, 2016. URL http://arxiv.org/abs/1607.04110.

[15] M. Raghu, B. Poole, J. Kleinberg, S. Ganguli, and J. Sohl-Dickstein. On the expressive power of deep neural networks. ArXiv e-prints, June 2016.

[16] M. Andrea Rodríguez and Max J. Egenhofer. Determining semantic similarity among entity classes from different ontologies. IEEE Trans. on Knowl. and Data Eng., 15(2):442–456, February 2003. ISSN 1041-4347. doi: 10.1109/TKDE.2003.1185844. URL http://dx.doi.org/10.1109/TKDE.2003.1185844.

[17] F. Scarselli, M. Gori, Ah Chung Tsoi, M. Hagenbuchner, and G. Monfardini. The Graph Neural Network Model. IEEE Transactions on Neural Networks, 20(1):61–80, jan 2009. ISSN 1045-9227. doi: 10.1109/TNN.2008.2005605. URL http://ieeexplore.ieee.org/document/4700287/.

[18] Wilson Wong, Wei Liu, and Mohammed Bennamoun. Ontology learning from text: A look back and into the future. ACM Comput. Surv., 44(4):20:1–20:36, September 2012. ISSN 0360-0300. doi: 10.1145/2333112.2333115. URL http://doi.acm.org/10.1145/2333112.2333115.

[19] Min-Ling Zhang and Zhi-Hua Zhou. A k-Nearest Neighbor Based Algorithm for Multi-label Classification. volume 2, pages 718–721 Vol. 2. The IEEE Computational Intelligence Society, 2005. URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1547385.