A survey of embedding models of entities and relationships for knowledge graph completion: Supplementary

Dat Quoc Nguyen
VinAI Research, Vietnam
`v.datnq9@vinai.io`

Triple Classification

Task Description

The triple classification task was first introduced by Socher et al. (2013), and since then it has been used to evaluate various embedding models. The aim of this task is to predict whether a triple (h, r, t) is correct or not. For classification, a relation-specific threshold θ_r is set for each relation type r. If the plausibility score of an unseen test triple (h, r, t) is higher than θ_r then the triple will be classified as correct, otherwise incorrect. Following Socher et al. (2013), the relation-specific thresholds are determined by maximizing the micro-averaged accuracy, which is a per-triple average, on the validation set.

Dataset	$\mid \mathcal{E} \mid$	$\mid \mathcal{R} \mid$	#Triples in train/valid/test		
FB13 (Socher et al., 2013)	75,043	13	316,232	5,908	23,733
WN11 (Socher et al., 2013)	38,696	11	112,581	2,609	10,544

Table 1: Statistics of the benchmark datasets for triple classification. In both WN11 and FB13, each validation and test set also contains the same number of incorrect triples as the number of correct triples.

Datasets

Information about benchmark datasets for the triple classification task is given in Table 1. FB13 and WN11 (Socher et al., 2013) are derived from the large real-world KG Freebase (Bollacker et al., 2008) and the large lexical KG WordNet (Miller, 1995), respectively. Note that when creating the FB13 and WN11 datasets, Socher et al. (2013) already filtered out triples from the test set if either or both of their head and tail entities also appear in the training set in a different relation type or order.

Main Results

Table 2 presents the triple classification results of KG completion models on the WN11 and FB13 datasets. The first 9 rows report the performance of models that use TransE/DISTMULT to initialize the entity and relation vectors. The last 12 rows present the accuracy of models with randomly initialized parameters. Note that there are higher triple classification results computed for NTN, Bilinear-COMP and TransE-COMP when entity vectors are initialized by averaging the pre-trained GloVe word vectors (Pennington et al., 2014). It is not surprising because many entity names in WordNet and Freebase are lexically meaningful. However, this is not always the case w.r.t. many domain-specific KGs.

References

Kurt Bollacker, Colin Evans, Praveen Paritosh, Tim Sturge, and Jamie Taylor. 2008. Freebase: A Collaboratively Created Graph Database for Structuring Human Knowledge. In *SIGMOD*, pages 1247–1250.

Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. 2013. Translating Embeddings for Modeling Multi-relational Data. In *NIPS*, pages 2787–2795.

Method	W11	F13	Avg.
CTransR (Lin et al., 2015)	85.7	-	-
TransR (Lin et al., 2015)	85.9	82.5	84.2
TransD (Ji et al., 2015)	86.4	**89.1**	87.8
TEKE_H (Wang and Li, 2016)	84.8	84.2	84.5
TranSparse-S (Ji et al., 2016)	86.4	88.2	87.3
TranSparse-US (Ji et al., 2016)	86.8	87.5	87.2
ConvKB (Nguyen et al., 2018) [*]	<u>87.6</u>	88.8	<u>88.2</u>
TransE-HRS (Zhang et al., 2018)	86.8	88.4	87.6
DISTMULT-HRS (Zhang et al., 2018)	**88.9**	<u>89.0</u>	**89.0**
NTN (Socher et al., 2013)	70.6	87.2	78.9
TransH (Wang et al., 2014)	78.8	83.3	81.1
SLogAn (Liang and Forbus, 2015)	75.3	85.3	80.3
KG2E (He et al., 2015)	85.4	85.3	85.4
Bilinear-COMP (Guu et al., 2015)	77.6	86.1	81.9
TransE-COMP (Guu et al., 2015)	80.3	87.6	84.0
TransR-FT (Feng et al., 2016)	86.6	82.9	84.8
TransG (Xiao et al., 2016)	87.4	87.3	87.4
lppTransD (Yoon et al., 2016)	86.2	88.6	87.4
TransE (Bordes et al., 2013) [*]	86.5	87.5	87.0
TransE-NMM (Nguyen et al., 2016)	86.8	88.6	87.7
TranSparse-DT (Chang et al., 2017)	87.1	87.9	87.5

Table 2: Accuracy results (in %) for triple classification on WN11 (labeled as **W11**) and FB13 (labeled as **F13**) test sets, which are taken from the corresponding papers. "Avg." denotes the averaged accuracy. [*] denotes that scores are taken from Nguyen et al. (2019).

L. Chang, M. Zhu, T. Gu, C. Bin, J. Qian, and J. Zhang. 2017. Knowledge Graph Embedding by Dynamic Translation. *IEEE Access*, 5:20898–20907.

Jun Feng, Minlie Huang, Mingdong Wang, Mantong Zhou, Yu Hao, and Xiaoyan Zhu. 2016. Knowledge graph embedding by flexible translation. In *KR*, pages 557–560.

Kelvin Guu, John Miller, and Percy Liang. 2015. Traversing Knowledge Graphs in Vector Space. In *EMNLP*, pages 318–327.

Shizhu He, Kang Liu, Guoliang Ji, and Jun Zhao. 2015. Learning to Represent Knowledge Graphs with Gaussian Embedding. In *CIKM*, pages 623–632.

Guoliang Ji, Shizhu He, Liheng Xu, Kang Liu, and Jun Zhao. 2015. Knowledge Graph Embedding via Dynamic Mapping Matrix. In *ACL-IJCNLP*, pages 687–696.

Guoliang Ji, Kang Liu, Shizhu He, and Jun Zhao. 2016. Knowledge Graph Completion with Adaptive Sparse Transfer Matrix. In *AAAI*, pages 985–991.

Chen Liang and Kenneth D. Forbus. 2015. Learning Plausible Inferences from Semantic Web Knowledge by Combining Analogical Generalization with Structured Logistic Regression. In *AAAI*, pages 551–557.

Yankai Lin, Zhiyuan Liu, Maosong Sun, Yang Liu, and Xuan Zhu. 2015. Learning Entity and Relation Embeddings for Knowledge Graph Completion. In *AAAI*, pages 2181–2187.

George A. Miller. 1995. WordNet: A Lexical Database for English. *Communications of the ACM*, 38(11):39–41.

Dat Quoc Nguyen, Kairit Sirts, Lizhen Qu, and Mark Johnson. 2016. Neighborhood Mixture Model for Knowledge Base Completion. In *CoNLL*, pages 40–50.

Dai Quoc Nguyen, Tu Dinh Nguyen, Dat Quoc Nguyen, and Dinh Phung. 2018. A Novel Embedding Model for Knowledge Base Completion Based on Convolutional Neural Network. In *NAACL-HLT*, pages 327–333.

Dai Quoc Nguyen, Dat Quoc Nguyen, Tu Dinh Nguyen, and Dinh Phung. 2019. A convolutional neural network-based model for knowledge base completion and its application to search personalization. *Semantic Web*, 10(5).

Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global Vectors for Word Representation. In *EMNLP*, pages 1532–1543.

Richard Socher, Danqi Chen, Christopher D Manning, and Andrew Ng. 2013. Reasoning With Neural Tensor Networks for Knowledge Base Completion. In *NIPS*, pages 926–934.

Zhigang Wang and Juan-Zi Li. 2016. Text-Enhanced Representation Learning for Knowledge Graph. In *IJCAI*, pages 1293–1299.

Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen. 2014. Knowledge Graph Embedding by Translating on Hyperplanes. In *AAAI*, pages 1112–1119.

Han Xiao, Minlie Huang, and Xiaoyan Zhu. 2016. TransG : A Generative Model for Knowledge Graph Embedding. In *ACL*, pages 2316–2325.

Hee-Geun Yoon, Hyun-Je Song, Seong-Bae Park, and Se-Young Park. 2016. A Translation-Based Knowledge Graph Embedding Preserving Logical Property of Relations. In *NAACL-HLT*, pages 907–916.

Zhao Zhang, Fuzhen Zhuang, Meng Qu, Fen Lin, and Qing He. 2018. Knowledge graph embedding with hierarchical relation structure. In *EMNLP*, pages 3198–3207.