Embedding Representation for Categorized Knowledge Graphs: The Effectiveness of Hyperparameters and Constraints

Matthew Wai Heng Chung and Hegler Tissot

University Colle London, London, UK
{rmhiwhc,h.tissot}@ucl.ac.uk

Abstract. Learning knowledge representation is an increasingly important technology that supports a variety of machine learning related applications. However, the choice of hyperparameters is seldom justified and relies on exhaustive search. Understanding the effect of hyperparameter combinations on embedding quality is crucial to avoid the inefficient process and enhance practicality of vector representation methods. In this study, we evaluate the effects of distinct hyperparameters and learning constraints focused on translational embedding representation models for multi-relational categorized knowledge graphs. We assess the influence of multiple hyperparameters and constraints regarding the quality of embedding models by contrasting traditional link prediction task accuracy against a classification task. HT: needs review here to generalize hyperparameters The findings provide evidence that lower values of margin are not rigorous enough to help with the learning process, whereas larger values produce much noise pushing the entities beyond to the surface of the hyperspace, thus requiring constant regularization. Finally, the correlation between link prediction and classification accuracy shows traditional validation protocol for embedding models is a weak metric to represent the quality of embedding representation.

Supplementary Material

this is...

Table 1: Examples of embedding scoring functions and their specific definitions.

Method	**Scoring function** $f_r(h,t)$	**Specific definitions**
TransE [2]	$\Vert\mathbf{h}+\mathbf{r}-\mathbf{t}\Vert_{l_{1/2}}$	–
HEXTRATO [13]	$\Vert\mathbf{h}_{c_h}+\mathbf{r}-\mathbf{t}_{c_t}\Vert_2$	–
UM [1]	$\Vert\mathbf{h}-\mathbf{t}\Vert_{l_{1/2}}$	–
TransH [14]	$\Vert\mathbf{h}_\perp+\mathbf{r}-\mathbf{t}_\perp\Vert_2^2$	$\mathbf{h}_\perp=\mathbf{h}-\mathbf{w}_r^\top\mathbf{h}\mathbf{w}_r,\quad \mathbf{t}_\perp=\mathbf{t}-\mathbf{w}_r^\top\mathbf{t}\mathbf{w}_r$
TransR [10]	$\Vert\mathbf{h}_\perp+\mathbf{r}-\mathbf{t}_\perp\Vert_2^2$	$\mathbf{h}_\perp=\mathbf{M}_r\mathbf{h},\quad \mathbf{t}_\perp=\mathbf{M}_r\mathbf{t}$
TransD [8]	$\Vert\mathbf{h}_\perp+\mathbf{r}-\mathbf{t}_\perp\Vert_2^2$	$\mathbf{h}_\perp=\mathbf{M}_{r,h}\mathbf{h},\quad \mathbf{t}_\perp=\mathbf{M}_{r,t}\mathbf{t}$ $\mathbf{M}_{r,h}=\mathbf{w}_r\mathbf{w}_h^\top+\mathbf{I}$ $\mathbf{M}_{r,t}=\mathbf{w}_r\mathbf{w}_t^\top+\mathbf{I}$
SE [3]	$\Vert\mathbf{h}_\perp-\mathbf{t}_\perp\Vert_1$	$\mathbf{h}_\perp=\mathbf{M}_{r,h}\mathbf{h},\quad \mathbf{t}_\perp=\mathbf{M}_{r,t}\mathbf{t}$
STransE [12]	$\Vert\mathbf{h}_\perp+\mathbf{r}-\mathbf{t}_\perp\Vert_{l_{1/2}}$	$\mathbf{h}_\perp=\mathbf{M}_{r,h}\mathbf{h},\quad \mathbf{t}_\perp=\mathbf{M}_{r,t}\mathbf{t}$
TransF [4]	$\Vert\mathbf{h}_\perp+\mathbf{r}-\mathbf{t}_\perp\Vert_2^2$	$\mathbf{h}_\perp=\mathbf{M}_{r,h}\mathbf{h},\quad \mathbf{t}_\perp=\mathbf{M}_{r,t}\mathbf{t}$ $\mathbf{M}_{r,h}=\sum_{i=1}^s\alpha_r^{(i)}\mathbf{U}^{(i)}+\mathbf{I}$ $\mathbf{M}_{r,t}=\sum_{i=1}^s\beta_r^{(i)}\mathbf{V}^{(i)}+\mathbf{I}$
TransM [5]	$\theta_r\Vert\mathbf{h}+\mathbf{r}-\mathbf{t}\Vert_{l_{1/2}}$	–
ManifoldE [16]	$\Vert\mathcal{M}(h,r,t)-D_r^2\Vert^2$	linear kernel hyper-spheres: $\mathcal{M}(h,r,t)=\Vert\mathbf{h}+\mathbf{r}-\mathbf{t}\Vert_2^2$ hyperplanes: $\mathcal{M}(h,r,t)=(\mathbf{h}+\mathbf{r}_h)^\top(\mathbf{t}+\mathbf{r}_t)$
TransE-FT [6]	$(\mathbf{h}+\mathbf{r})^\top\mathbf{t}+(\mathbf{t}-\mathbf{r})^\top\mathbf{h}$	–
TransA(Xiao) [15]	$(\vert\mathbf{h}+\mathbf{r}-\mathbf{t}\vert)^\top\mathbf{M}_r(\vert\mathbf{h}+\mathbf{r}-\mathbf{t}\vert)$	–
KG2E [7]	symmetric: $\frac{1}{2}\{\boldsymbol{\mu}^\top\boldsymbol{\Sigma}_r^{-1}\boldsymbol{\mu}+\ln(\det(\boldsymbol{\Sigma}))\}$ asymmetric: $\frac{1}{2}\{\mathrm{tr}(\boldsymbol{\Sigma}_\mathrm{r}^{-1}(\boldsymbol{\Sigma}_\mathrm{h}+\boldsymbol{\Sigma}_\mathrm{t}))+$ $\boldsymbol{\mu}^\top\boldsymbol{\Sigma}_r^{-1}\boldsymbol{\mu}-\ln\frac{\det(\boldsymbol{\Sigma}_h+\boldsymbol{\Sigma}_t)}{\det(\boldsymbol{\Sigma}_r)}\}$	$\mathbf{h}\sim\mathcal{N}(\boldsymbol{\mu}_h,\boldsymbol{\Sigma}_h),\quad \mathbf{t}\sim\mathcal{N}(\boldsymbol{\mu}_t,\boldsymbol{\Sigma}_t),$ $\mathbf{r}\sim\mathcal{N}(\boldsymbol{\mu}_r,\boldsymbol{\Sigma}_r)$ $\mathbf{h}-\mathbf{t}\sim\mathcal{N}(\boldsymbol{\mu}_h-\boldsymbol{\mu}_t,\boldsymbol{\Sigma}_t+\boldsymbol{\Sigma}_h)$ $\boldsymbol{\mu}=\boldsymbol{\mu}_h-\boldsymbol{\mu}_r-\boldsymbol{\mu}_t$ $\boldsymbol{\Sigma}=\boldsymbol{\Sigma}_h+\boldsymbol{\Sigma}_r+\boldsymbol{\Sigma}_t$
TKRL [18]	$\Vert\mathbf{h}_\perp+\mathbf{r}-\mathbf{t}_\perp\Vert$	$\mathbf{h}_\perp=\mathbf{M}_{r,h}\mathbf{h},\quad \mathbf{t}_\perp=\mathbf{M}_{r,t}\mathbf{t}$ $\mathbf{M}_{r,h}=\frac{\sum_{i=1}^n\alpha_i\mathbf{M}_{c_i}}{\sum_{i=1}^n\alpha_i},\quad \alpha_i=\begin{cases}1, c_i\in C_{r_h}\\0, c_i\notin C_{r_h}\end{cases}$ $\mathbf{M}_{r,t}=\frac{\sum_{i=1}^n\alpha_i\mathbf{M}_{c_i}}{\sum_{i=1}^n\alpha_i},\quad \alpha_i=\begin{cases}1, c_i\in C_{r_t}\\0, c_i\notin C_{r_t}\end{cases}$ recursive: $\mathbf{M}_c=\prod_{i=1}^m\mathbf{M}_{c^{(i)}}$ weighted: $\mathbf{M}_c=\sum_{i=1}^m\beta_i\mathbf{M}_{c^{(i)}},$ $\beta_i:\beta_{i+1}=(1-\eta):\eta,\quad \sum_{i=1}^m\beta_i=1$
TransT [11]	$\begin{cases}-\ln p(h\vert r,t,true)+\\ \quad\ln p(h'\vert r,t,true), & h'\neq h\\ -\ln p(t\vert h,r,true)+\\ \quad\ln p(t'\vert h,r,true), & t'\neq t\\ -\ln p(r\vert h,t,true)+\\ \quad\ln p(r'\vert h,t,true), & r'\neq r\end{cases}$	$p(t\vert h,r,true)=$ $\begin{cases}\frac{p(true\vert h,r,t)p(t\vert h,r)}{p(true\vert h,r)}, & p(t\vert h,r)\neq 0\\ 0 & p(t\vert h,r)=0\end{cases}$

Table 2: Link prediction scores for WN18 and FB15k reported in multiple translational models.

Embedding Method	WN18						FB15K					
	MR		MRR		Hits@10		MR		MRR		Hits@10	
	Raw	Filter	Raw	Filter	Raw	Filter	Raw	Filter	Raw	Filter	Raw	Filter
TransE [2]	263	251	–	–	75.4	89.2	243	125	–	–	34.9	47.1
HEXTRATO [13]	–	–	–	–	–	–	116	–	–	–	53.5	–
TransH [14]	401	388	–	–	73.0	82.3	212	87	–	–	45.7	64.4
TransR [10]	238	225	–	–	79.8	79.4	198	77	–	–	48.2	68.7
CTransR [10]	231	218	–	–	79.4	92.3	199	75	–	–	48.4	70.2
TransD [8]	224	212	–	–	79.6	92.2	194	91	–	–	53.4	77.3
TranSparse [9] (shr, S)	237	224	–	–	80.4	93.6	194	88	–	–	53.4	77.7
TranSparse [9] (shr, US)	233	221	–	–	80.5	93.9	191	86	–	–	53.5	78.3
TranSparse [9] (sep, S)	224	221	–	–	79.8	92.8	187	82	–	–	53.3	79.5
TranSparse [9] (sep, US)	223	211	–	–	80.1	93.2	190	82	–	–	53.7	79.9
STransE [12]	217	206	0.469	0.657	80.9	93.4	219	69	0.252	0.543	51.6	79.7
TransF [4]	–	198	–	0.856	–	95.3	–	62	–	0.564	–	82.3
TransM [5]	293	281	–	–	75.7	85.4	197	94	–	–	44.6	55.2
ManifoldE [16] (sphere)	–	–	–	–	80.7	92.8	–	–	–	–	55.7	86.2
ManifoldE [16] (hyperplane)	–	–	–	–	84.2	94.9	–	–	–	–	55.2	88.1
TransA(Xiao) [15]	405	392	–	–	82.3	94.3	155	74	–	–	56.1	80.4
TransA(Jia) [19]	165	153	–	–	–	–	164	58	–	–	–	–
KG2E [7]	342	331	–	–	80.2	92.8	174	59	–	–	48.9	74.0
TransG [17]	357	345	–	–	84.5	94.9	152	50	–	–	55.9	88.2
UM [1]	315	304	–	–	35.3	38.2	1074	979	–	–	4.5	6.3
SE [3]	1011	985	–	–	68.5	80.5	273	162	–	–	28.8	39.8
TKRL [18]	–	–	–	–	–	–	184	68	–	–	49.2	69.4
TransT [11]	137	130	–	–	92.7	97.4	199	46	–	–	53.3	85.4

References

1. Bordes, A., Glorot, X., Weston, J., Bengio, Y.: Joint learning of words and meaning representations for open-text semantic parsing. In: Lawrence, N.D., Girolami, M. (eds.) Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics. Proceedings of Machine Learning Research, vol. 22, pp. 127–135. PMLR, La Palma, Canary Islands (21–23 Apr 2012), http://proceedings.mlr.press/v22/bordes12.html
2. Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., Yakhnenko, O.: Translating embeddings for modeling multi-relational data. In: Burges, C.J.C., Bottou, L., Welling, M., Ghahramani, Z., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems 26, pp. 2787–2795. Curran Associates, Inc. (2013)
3. Bordes, A., Weston, J., Collobert, R., Bengio, Y.: Learning structured embeddings of knowledge bases. In: Conference on Artificial Intelligence (2011)
4. Do, K., Tran, T., Venkatesh, S.: Knowledge graph embedding with multiple relation projections. 2018 24th International Conference on Pattern Recognition (ICPR) pp. 332–337 (2018)
5. Fan, M., Zhou, Q., Chang, E., Zheng, T.F.: Transition-based knowledge graph embedding with relational mapping properties. In: Proceedings of the 28th Pacific Asia Conference on Language, Information and Computing (2014)
6. Feng, J., Huang, M., Wang, M., Zhou, M., Hao, Y., Zhu, X.: Knowlege graph embedding by flexible translation. In: Proceedings of the 15th International Conference on Principles of Knowledge Representation and Reasoning. pp. 557–560 (2015)
7. He, S., Liu, K., Ji, G., Zhao, J.: Learning to represent knowledge graphs with gaussian embedding. In: Proceedings of the 24th ACM International on Conference on Information and Knowledge Management. pp. 623–632. CIKM '15, ACM, New York, NY, USA (2015), http://doi.acm.org/10.1145/2806416.2806502
8. Ji, G., He, S., Xu, L., Liu, K., Zhao, J.: Knowledge graph embedding via dynamic mapping matrix. In: Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing, ACL 2015, July 26-31, 2015, Beijing, China, Volume 1: Long Papers. pp. 687–696 (2015), http://aclweb.org/anthology/P/P15/P15-1067.pdf
9. Ji, G., Liu, K., He, S., Zhao, J.: Knowledge graph completion with adaptive sparse transfer matrix. In: AAAI (2016)
10. Lin, Y., Liu, Z., Sun, M., Liu, Y., Zhu, X.: Learning entity and relation embeddings for knowledge graph completion. In: Proceedings of the 29th AAAI Conference on Artificial Intelligence. pp. 2181–2187. AAAI'15, AAAI Press (2015)
11. Ma, S., Ding, J., Jia, W., Wang, K., Guo, M.: Transt: Type-based multiple embedding representations for knowledge graph completion. In: The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (2017)
12. Nguyen, D.Q., Sirts, K., Qu, L., Johnson, M.: Stranse: a novel embedding model of entities and relationships in knowledge bases. CoRR abs/1606.08140 (2016)
13. Tissot, H.: HEXTRATO: Using ontology-based constraints to improve accuracy on learning domain-specific entity and relationship embedding representation for knowledge resolution. In: Proceedings of the 10th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2018, Volume 1: KDIR, Seville, Spain, September 18-20, 2018. pp. 70–79 (2018), https://doi.org/10.5220/0006923700700079

14. Wang, Z., Zhang, J., Feng, J., Chen, Z.: Knowledge graph embedding by translating on hyperplanes. In: Brodley, C.E., Stone, P. (eds.) Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence. pp. 1112–1119. AAAI Press (2014)
15. Xiao, H., Huang, M., Hao, Y., Zhu, X.: Transa: An adaptive approach for knowledge graph embedding. CoRR abs/1509.05490 (2015), http://arxiv.org/abs/1509.05490
16. Xiao, H., Huang, M., Zhu, X.: From one point to a manifold: Knowledge graph embedding for precise link prediction. In: Proceedings of the 25th International Joint Conference on Artificial Intelligence. pp. 1315–1321. IJCAI'16, AAAI Press (2016), http://dl.acm.org/citation.cfm?id=3060621.3060804
17. Xiao, H., Huang, M., Zhu, X.: Transg : A generative model for knowledge graph embedding. In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). pp. 2316–2325. Association for Computational Linguistics (2016), http://aclweb.org/anthology/P16-1219
18. Xie, R., Liu, Z., Sun, M.: Representation learning of knowledge graphs with hierarchical types. In: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence. pp. 2965–2971. IJCAI'16, AAAI Press (2016), http://dl.acm.org/citation.cfm?id=3060832.3061036
19. Yantao Jia, Yuanzhuo Wang, H.L.X.J.X.C.: Locally adaptive translation for knowledge graph embedding. In: Proceedings of the 30th AAAI Conference on Artificial Intelligence. pp. 992–998. AAAI Press (2016)