Group-8 - Graph Machine Learning Assignment - 4

Aayush Sugandh (24AI60R04), Sandipan Majhi (24AI60R21)

31/10/2024

1 Methodology

1.1 Model and Dataset

We implemented the TransE model as described by Bordes et al. (2013) and TransR paper as described in Lin et al. (2015). We tested our model on Nations and Kinships dataset. The dataset has been described in Table-1.

Dataset	Entities	Relationships	Triples
Nations	14 countries	55	1,593
Kinships	104 individuals	26	10,686

Table 1: Summary of Nations and Kinships Knowledge Graph Datasets

1.2 Training

In our case we try to perform an unsupervised training to get embeddings of entities and relationships. We start with random initialization and try to arrive at good embeddings using triplet margin loss as described in Bordes et al. (2013). In order to calculate triplet margin loss, we wrote a Bernoulli Negative Sampler to produce a fixed number of negative samples for a positive sample. We implement the Bernoulli Negative Sampler by calculating corruption rate by calculating average tails per head and average head per tails for each relationship in the dataset. Our training parameters have been explained in Table - 2. For all our experiments we ran our models for **50** Epochs. We employ number of 5 negative samples per positive sample in Kinships dataset because the number of entities and relationships were larger in number.

Dataset	Batch Size	Learning Rate	Optimizer	Negative Samples per Positive
Nations	128	0.001	Adam	1
Kinships	512	0.001	Adam	5

Table 2: Training Parameters for Nations and Kinships Knowledge Graph Datasets

For Nations dataset we use margins of 1,2,3,4 and 5. We also used same margins for Kinships dataset.

1.3 Training and Test Set performance

During training we consistently find that, TransR model was able to achieve lower Training loss than TransE. In terms of Test Set performance, we found that both models were comparable in performance. The training profile has been shown in Table-1.3. The test set perfomance in terms of Mean Rank of correct head or tail and Hit@10 for Nations Dataset are shown in Table - 3 and for Kinships Dataset in Table - 4. In our experiments we report that, for Nations Dataset **TransE** model has best **Mean Rank** of **6.04** and best **Hits@10** of **0.856**. For **TransR** model best **Mean Rank** was **6.14** and best **Hits@10** of **0.836**. For Kinships Dataset, we found, for **TransE** best **Mean Rank** was **48.66** and best **Hits@10** was **0.122**. For **TransR** model, best **Mean Rank** was **50.00** and the best **Hits@10** was **0.113**.

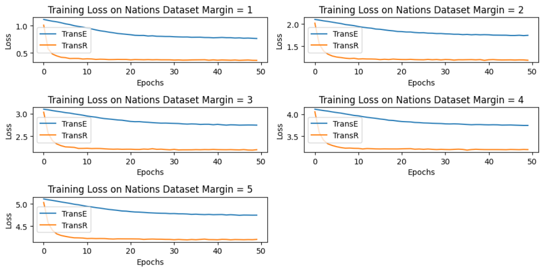

Figure 1: Nations Knowledge Graph Training metrics on TransE vs TransR

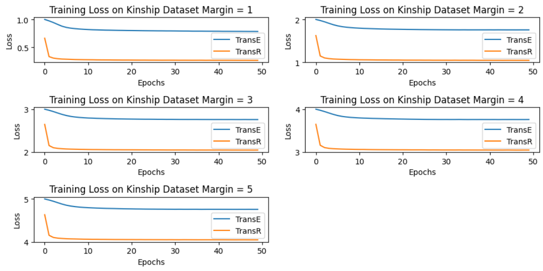

Figure 2: Kinships Knowledge Graph Training metrics on TransE vs TransR

Model	Margin	Mean Rank	Hits@10
TransE	1	6.13	0.828
	2	**6.03**	**0.856**
	3	6.41	0.784
	4	6.43	0.808
	5	6.31	0.816
TransR	1	6.55	0.784
	2	**6.14**	0.821
	3	6.25	**0.836**
	4	6.36	0.821
	5	6.41	0.806

Table 3: Comparison of TransE and TransR Models on Nations Test Set

Model	Margin	Mean Rank	Hits@10
TransE	1	50.53	0.094
	2	50.54	0.107
	3	**48.66**	**0.122**
	4	49.38	0.108
	5	49.38	0.122
TransR	1	**50.00**	0.101
	2	50.36	**0.113**
	3	50.34	0.106
	4	50.86	0.112
	5	50.09	0.108

Table 4: Comparison of TransE and TransR Models on Kinships Test Set

1.4 Similarity Performance

In this part of the assignment we take the best margins for TransE model which is 2 and for TransR it is 3 for Nations dataset, and use them in this part of our assignment. Here we used embedding dimension of 30, Adam optimizer and learning rate of 0.001. Then we extract the embeddings for each of samples in our validation set and concatenate the embeddings for head, relation and tail and perform cosine similarity with concatenated embeddings of head, relation and tail of triples in train set. The similarity information has been summarized in Table - 5. We observed that the examples in the validation set were present in the test set also, and moreover there were more triples in the test set too. So we also do a similarity match with the validation set and combined Train and Test set which are summarized in Table - 8. In our comparison tables, we compiled results from both TransE and TransR models.

1.5 Analysis and Discussion

We found that training set losses for **TransR** was consistently lower than **TransE** in Table - 1.3 . But there was no comparable differences on test set performance where **TransE** performance was sometimes better than **TransR**. This could be because **TransR** has higher number of parameters because of relation space specific projection parameters, which could be responsible.

During Test set performance, we see that TransE performs better than TransR when the knowledge graph is bigger and produced a better Mean Rank. On explanation could be, only focussing on relation specific emebeddings could cause TransR models to overfit on specific head or tails. So whenever the model finds out a relation, it kind of only searches in limited space of entities.

During similarity triple matching part, we observed the effect of relation specific projections in actions. If closely check Table - 5, we would find that in TransE model, the similar triples had more variations in suggesting similar relationships. For example, for relation **'reldiplomacy'** we find that, **'exports3'** and **'relexports'** being suggested in the similar triples. On the otherhand, we see that because we used relation specific projections in TransR, we found that all the top 5 similar triple suggested had, same relations. We feel that, TransE helps in choosing diversely similarly matched relations when suggesting similar triples. This could be really helpful if someone tries to design a question answering engine which can be more creative inproducing answers. This is because, **diplomacy** is very closely related to

exports. But one downside of this could be, unreliability in question answering. One might wish to only find out, only specific information for a relation, then TransR succeeds in providing only relation specific information. Hence, TransR produces very reliable and less variability in inferring answers. One area where these could be very useful is answering multihop questions. In these cases, one might use a combination of TransE and TransR embeddings. If too much variability is introduced when triples search is conducted over the space, there could be much larger computation involved in inference and produce vastly different answers each time. This is where one might use a combination of TransE and TransR embeddings.

References

Bordes, A., Usunier, N., García-Durán, A., Weston, J., & Yakhnenko, O. (2013). Translating embeddings for modeling multi-relational data. In *Neural information processing systems.* Retrieved from https://api.semanticscholar.org/CorpusID:14941970

Lin, Y., Liu, Z., Sun, M., Liu, Y., & Zhu, X. (2015, Feb.). Learning entity and relation embeddings for knowledge graph completion. *Proceedings of the AAAI Conference on Artificial Intelligence*, *29*(1). Retrieved from https://ojs.aaai.org/index.php/AAAI/article/view/9491 DOI: 10.1609/aaai.v29i1.9491

Table 5: Top 5 Similar Triples for Validation Data in Nations Dataset (Only Train Set)

Validation Triple	Top 5 Similar Triples (Score)
['brazil', 'commonbloc1', 'india']	['netherlands', 'commonbloc1', 'india'] (0.9854) ['uk', 'commonbloc1', 'india'] (0.9770) ['india', 'commonbloc1', 'netherlands'] (0.9746) ['india', 'commonbloc1', 'uk'] (0.9709) ['india', 'commonbloc1', 'usa'] (0.9688)
['burma', 'intergovorgs3', 'indonesia']	['burma', 'relngo', 'indonesia'] (0.9182) ['burma', 'exports3', 'indonesia'] (0.8374) ['burma', 'relngo', 'cuba'] (0.8137) ['burma', 'intergovorgs3', 'usa'] (0.8006) ['burma', 'relintergovorgs', 'cuba'] (0.7946)
['china', 'accusation', 'uk']	['china', 'accusation', 'india'] (0.9926) ['china', 'accusation', 'usa'] (0.9740) ['china', 'accusation', 'ussr'] (0.9134) ['indonesia', 'accusation', 'usa'] (0.8985) ['cuba', 'accusation', 'usa'] (0.8788)
['cuba', 'reldiplomacy', 'china']	['cuba', 'exports3', 'china'] (0.9447) ['cuba', 'reldiplomacy', 'indonesia'] (0.9245) ['indonesia', 'reldiplomacy', 'china'] (0.8955) ['cuba', 'relexports', 'china'] (0.8749) ['cuba', 'relintergovorgs', 'indonesia'] (0.8618)
['egypt', 'embassy', 'uk']	['egypt', 'embassy', 'india'] (0.9926) ['egypt', 'embassy', 'netherlands'] (0.9894) ['egypt', 'embassy', 'brazil'] (0.9770) ['egypt', 'embassy', 'usa'] (0.9740) ['egypt', 'embassy', 'poland'] (0.9526)

Table 6: TransE Model

Validation Triple	Top 5 Similar Triples (Score)
['brazil', 'commonbloc1', 'india']	['jordan', 'commonbloc1', 'netherlands'] (0.9286) ['brazil', 'commonbloc1', 'jordan'] (0.8210) ['india', 'commonbloc1', 'netherlands'] (0.8160) ['jordan', 'commonbloc1', 'uk'] (0.8071) ['netherlands', 'commonbloc1', 'india'] (0.8059)
['burma', 'intergovorgs3', 'indonesia']	['burma', 'intergovorgs3', 'egypt'] (0.9997) ['cuba', 'intergovorgs3', 'egypt'] (0.9445) ['indonesia', 'intergovorgs3', 'egypt'] (0.9416) ['poland', 'intergovorgs3', 'egypt'] (0.9414) ['burma', 'intergovorgs3', 'brazil'] (0.8569)
['china', 'accusation', 'uk']	['ussr', 'accusation', 'israel'] (0.8475) ['china', 'accusation', 'india'] (0.6684) ['india', 'accusation', 'uk'] (0.6047) ['israel', 'accusation', 'jordan'] (0.4896) ['china', 'accusation', 'indonesia'] (0.4283)
['cuba', 'reldiplomacy', 'china']	['egypt', 'reldiplomacy', 'china'] (0.9984) ['brazil', 'reldiplomacy', 'egypt'] (0.9933) ['egypt', 'reldiplomacy', 'brazil'] (0.9866) ['egypt', 'reldiplomacy', 'cuba'] (0.9845) ['cuba', 'reldiplomacy', 'ussr'] (0.9809)
['egypt', 'embassy', 'uk']	['egypt', 'embassy', 'poland'] (0.9981) ['jordan', 'embassy', 'uk'] (0.9938) ['indonesia', 'embassy', 'poland'] (0.9827) ['usa', 'embassy', 'poland'] (0.9588) ['usa', 'embassy', 'burma'] (0.9586)

Table 7: TransR Model

Table 8: Comparison of Top 5 similar for Validation data in Nations Dataset(Train + Test Set)

Validation Triple	**Top 5 Similar Triples (Score)**
['brazil', 'commonbloc1', 'india']	['brazil', 'commonbloc1', 'india'] (1.0000) ['netherlands', 'commonbloc1', 'india'] (0.9854) ['uk', 'commonbloc1', 'india'] (0.9770) ['india', 'commonbloc1', 'netherlands'] (0.9746) ['india', 'commonbloc1', 'uk'] (0.9709)
['burma', 'intergovorgs3', 'indonesia']	['burma', 'intergovorgs3', 'indonesia'] (1.0000) ['burma', 'relngo', 'indonesia'] (0.9182) ['burma', 'exports3', 'indonesia'] (0.8374) ['burma', 'relngo', 'cuba'] (0.8137) ['burma', 'reldiplomacy', 'indonesia'] (0.8106)
['china', 'accusation', 'uk']	['china', 'accusation', 'uk'] (1.0000) ['china', 'accusation', 'india'] (0.9926) ['china', 'accusation', 'usa'] (0.9740) ['indonesia', 'accusation', 'uk'] (0.9245) ['china', 'accusation', 'ussr'] (0.9134)
['cuba', 'reldiplomacy', 'china']	['cuba', 'reldiplomacy', 'china'] (1.0000) ['cuba', 'exports3', 'china'] (0.9447) ['cuba', 'reltreaties', 'china'] (0.9322) ['cuba', 'reldiplomacy', 'indonesia'] (0.9245) ['indonesia', 'reldiplomacy', 'china'] (0.8955)
['egypt', 'embassy', 'uk']	['egypt', 'embassy', 'uk'] (1.0000) ['egypt', 'embassy', 'india'] (0.9926) ['egypt', 'embassy', 'netherlands'] (0.9894) ['egypt', 'embassy', 'brazil'] (0.9770) ['egypt', 'embassy', 'usa'] (0.9740)

Table 9: TransE Model

Validation Triple	**Top 5 Similar Triples (Score)**
['brazil', 'commonbloc1', 'india']	['brazil', 'commonbloc1', 'india'] (1.0000) ['jordan', 'commonbloc1', 'netherlands'] (0.9286) ['brazil', 'commonbloc1', 'jordan'] (0.8210) ['india', 'commonbloc1', 'netherlands'] (0.8160) ['jordan', 'commonbloc1', 'uk'] (0.8071)
['burma', 'intergovorgs3', 'indonesia']	['burma', 'intergovorgs3', 'indonesia'] (1.0000) ['burma', 'intergovorgs3', 'egypt'] (0.9997) ['cuba', 'intergovorgs3', 'egypt'] (0.9445) ['indonesia', 'intergovorgs3', 'egypt'] (0.9416) ['poland', 'intergovorgs3', 'egypt'] (0.9414)
['china', 'accusation', 'uk']	['china', 'accusation', 'uk'] (1.0000) ['indonesia', 'accusation', 'uk'] (0.8595) ['ussr', 'accusation', 'israel'] (0.8475) ['china', 'accusation', 'india'] (0.6684) ['india', 'accusation', 'uk'] (0.6047)
['cuba', 'reldiplomacy', 'china']	['cuba', 'reldiplomacy', 'china'] (1.0000) ['egypt', 'reldiplomacy', 'china'] (0.9984) ['cuba', 'reldiplomacy', 'egypt'] (0.9935) ['brazil', 'reldiplomacy', 'egypt'] (0.9933) ['egypt', 'reldiplomacy', 'brazil'] (0.9866)
['egypt', 'embassy', 'uk']	['egypt', 'embassy', 'uk'] (1.0000) ['egypt', 'embassy', 'poland'] (0.9981) ['egypt', 'embassy', 'burma'] (0.9980) ['jordan', 'embassy', 'uk'] (0.9938) ['indonesia', 'embassy', 'poland'] (0.9827)

Table 10: TransR Model