Neural Regression with Embeddings for Numeric Attribute Prediction in Knowledge Graphs - Appendix

No Author Given

No Institute Given

1 Experiment Details

All embedding models for link prediction tasks are trained with embedding dimension 32, for 256 epochs, using a learning rate of 0.05 and a mini-batch size of 1024 using the KvsAll scoring technique for standard as well as combined training with LitEm. Similarly, the LitEm is trained with the same embedding dimension as KGE models. We use a dropout of 0.3, and a learning rate of 0.01 and train the LitEm for 200 epochs for literal-prediction tasks using pre-trained embeddings.

2 Literal Embedding Model

Algorithm 1 Literal Embedding Model

Input: Attribute triples $\{(e, a, v)\} \in \mathcal{G}$, Learning rate α, Number of epochs n, Pre-trained entity embeddings $\mathbf{E} \in \mathbb{R}^{|\mathcal{A}| \times d}$,
Initialize $\theta = \{$ attribute embeddings $\mathbf{A} \in \mathbb{R}^{|\mathcal{A}| \times d}$, weight matrices $\mathbf{W}_1, \mathbf{W}_2$, and biases $\mathbf{b}_1, \mathbf{b}_2$ $\}$ ▷ All learnable parameters are part of θ
Output: Trained parameters θ
for epoch = 1 **to** n **do**
  $\hat{y} = \mathbf{W}_2 \cdot (\mathbf{f}(\mathbf{W}_1 \cdot [\mathbf{e}, \mathbf{a}] + \mathbf{b}_1) + [\mathbf{e}, \mathbf{a}]) + \mathbf{b}_2$

  Compute the MAE loss:$\mathcal{L} = \frac{1}{N} \sum(|\hat{y} - v|)$
  Compute the gradient $\nabla_\theta L$ of the loss with respect to θ
  Update the parameters: $\theta \leftarrow \theta - \alpha \nabla_\theta L$
end for
return θ

3 Literal Prediction

Table 1. Numeric Attribute Prediction performance of different embedding models with standard training on FB15k-237, YAGO15K and DB15K evaluated using Mean Absolute Error (MAE) Metric. All the embedding models are trained for 256 epochs, with a learning rate of 0.05 and 32-dimensional real-valued vectors with the KvsAll Scoring technique.

relation	CmplEx	DeCaL	DistMlt	Keci	MuRE	OMult	QMult	TransE
FB15k-237								
release_date	5.317	5.333	5.070	5.022	5.062	5.461	5.140	**4.813**
loc.date_founded	185.434	183.784	182.430	174.937	178.819	182.161	183.216	**169.235**
latitude	9.276	9.351	8.379	8.959	7.573	8.763	9.578	**6.420**
longitude	22.550	22.843	20.609	21.226	20.439	24.194	23.402	**15.935**
loc.area	1.67e6	1.68e6	1.67e6	1.68e6	1.64e6	1.70e6	1.67e6	1.64e6
org.date_founded	58.269	61.449	63.043	64.214	61.223	60.434	63.648	**56.818**
date_of_death	40.127	40.118	44.889	41.223	38.519	37.994	39.873	**35.760**
date_of_birth	22.860	23.716	24.065	22.837	21.428	22.923	22.906	**20.675**
height_meters	**0.066**	0.067	**0.066**	0.067	**0.066**	0.068	0.069	0.074
weight_kg	12.246	11.389	12.980	11.290	12.932	11.713	**10.650**	12.659
pop._number	5.09e6	5.44e6	4.79e6	5.03e6	4.46e6	5.57e6	4.79e6	4.39e6
YAGO15K								
diedOnDate	48.445	47.529	48.797	46.645	46.045	47.261	48.379	**44.932**
happenedOnDate	41.277	42.306	42.882	44.801	36.921	39.471	39.587	**29.847**
Latitude	7.416	7.216	6.972	6.939	6.189	7.245	6.815	**4.652**
Longitude	30.081	27.004	30.644	31.004	26.406	36.895	32.239	**17.490**
BornOnDate	20.690	19.798	20.525	20.776	18.999	20.061	20.516	**17.458**
CreatedOnDate	67.496	67.581	67.695	70.348	62.844	68.904	68.421	**62.785**
DestroyedOnDate	27.670	28.214	29.231	27.093	27.899	27.564	27.649	**23.658**
DB15K								
birthDate	18.318	18.598	18.306	17.619	15.973	18.943	17.691	**14.403**
completionDate	8.275	8.377	8.180	7.259	8.016	8.059	8.426	**5.854**
deathDate	20.115	21.330	21.550	20.036	20.876	21.349	**18.595**	19.432
formationDate	43.400	37.897	39.715	43.236	39.298	**37.228**	38.110	39.945
foundingDate	39.780	40.739	42.937	43.307	43.092	44.545	40.878	**37.933**
height	1.798	1.869	2.280	2.163	**1.663**	2.528	2.189	2.271
releaseDate	11.319	12.818	12.839	12.035	11.728	12.982	12.721	**9.218**
latitude	8.077	8.352	8.928	8.201	8.105	9.474	9.105	**5.944**
longitude	30.947	29.970	30.662	32.365	29.804	34.371	31.704	**21.737**

Table 2. Literal Prediction performance of different numeric attribute propagation approaches on FB15k-237 and YAGO15K compared with LitEm on Mean Absolute Error (values for KGA, NAP++ and MrAP) (Values taken from [1])

Relation	KGA	NAP++	MrAP	LitEm
FB15k-237				
date_of_birth	18.900	22.100	**15.000**	18.824
date_of_death	20.600	52.300	**16.300**	32.620
release_date	**4.000**	9.900	6.300	4.258
org.date_founded	49.000	59.300	58.300	**46.741**
loc.date_founded	**76.000**	92.100	98.800	150.873
latitude	2.100	11.800	**1.500**	5.221
longitude	7.100	54.700	**4.000**	11.325
area	**6.10e4**	4.40e5	4.40e5	1.65e6
population_number	**4.00e6**	7.50e6	2.10e7	4.87e6
height_meters	0.077	0.080	0.086	**0.074**
weight_kg	**11.600**	15.300	12.900	12.653
YAGO15K				
BornOnDate	16.300	23.200	19.700	**15.586**
diedOnDate	30.800	45.700	34.000	**40.515**
CreatedOnDate	58.200	83.500	70.400	**57.081**
DestroyedOnDate	23.300	38.200	34.600	**21.436**
happenedOnDate	29.900	73.700	54.100	**26.064**
Latitude	3.400	8.700	**2.800**	3.090
Longitude	7.200	43.100	**5.700**	11.506

4 Link Prediction

Link prediction performances of KGE models with and without combined training with the LitEm model on Mutag and FB15k-237 datasets.

Table 3. Link prediction performance of different embedding models with 32-dimensional real-valued vectors on the FB15k-237 dataset without and with combined training with LitEm, Each model's results on training, test, and validation sets are shown in three rows; bold marks the best per-metric value for each model.

Models	**KGE**				**KGE + LitEm**			
	MRR	**H@1**	**H@3**	**H@10**	**MRR**	**H@1**	**H@3**	**H@10**
ComplEx	0.458	0.352	0.512	0.662	0.415	0.312	0.465	0.616
	0.247	**0.174**	**0.270**	**0.393**	0.234	0.165	0.253	0.372
	0.247	**0.174**	0.269	**0.393**	0.237	0.168	**0.256**	0.372
DeCaL	0.513	0.413	0.565	0.706	0.496	0.393	0.549	0.695
	0.262	**0.186**	**0.286**	**0.415**	0.260	0.183	0.283	0.413
	0.264	**0.187**	**0.288**	**0.418**	0.260	0.184	0.284	0.413
DistMult	0.475	0.372	0.529	0.676	0.454	0.347	0.509	0.660
	0.245	0.170	**0.268**	**0.396**	0.243	0.170	0.266	0.389
	0.247	**0.173**	**0.270**	**0.394**	0.243	0.170	0.265	0.388
Keci	0.543	0.438	0.602	0.744	0.519	0.415	0.576	0.718
	0.260	0.182	0.284	0.417	**0.268**	**0.190**	**0.290**	**0.426**
	0.264	0.186	0.289	0.420	**0.270**	**0.191**	**0.296**	**0.427**
OMult	0.530	0.431	0.584	0.721	0.468	0.366	0.518	0.665
	0.230	0.156	**0.253**	**0.376**	0.230	**0.158**	0.251	0.371
	0.233	**0.160**	**0.255**	**0.378**	0.231	0.159	0.252	**0.374**
QMult	0.544	0.443	0.598	0.736	0.535	0.434	0.589	0.728
	0.255	0.179	0.276	0.410	**0.260**	**0.182**	**0.282**	**0.416**
	0.261	0.184	0.283	0.413	**0.262**	**0.185**	0.283	**0.417**
TransE	0.429	0.306	0.499	0.660	0.409	0.287	0.476	0.638
	0.307	**0.217**	**0.341**	**0.481**	0.301	0.213	0.335	0.473
	0.311	**0.222**	**0.345**	**0.484**	0.306	0.219	0.341	0.480
MuRE	0.493	0.388	0.549	0.693	0.485	0.381	0.539	0.686
	0.267	0.191	0.289	0.420	**0.270**	**0.194**	**0.292**	**0.424**
	0.272	0.195	**0.295**	0.426	**0.274**	**0.198**	0.294	**0.429**

Table 4. Link prediction performance of different embedding models with 32-dimensional real-valued vectors on the Mutagenesis dataset without and with combined training with LitEm, Each model's results on training, test, and validation sets are shown in three rows; bold marks the best per-metric value for each model.

Models	**KGE**				**KGE + LitEm**			
	MRR	**H@1**	**H@3**	**H@10**	**MRR**	**H@1**	**H@3**	**H@10**
ComplEx	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.226	**0.158**	**0.249**	**0.357**	0.174	0.109	0.187	0.315
	0.218	**0.149**	**0.240**	**0.345**	0.172	0.107	0.186	0.314
DeCaL	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.130	0.077	0.131	0.241	**0.154**	**0.093**	**0.160**	**0.280**
	0.128	0.073	0.134	0.244	**0.152**	**0.089**	**0.160**	**0.280**
DistMult	1.000	1.000	1.000	1.000	0.997	0.995	1.000	1.000
	0.177	0.095	0.206	0.322	**0.239**	**0.161**	**0.269**	**0.372**
	0.181	0.098	0.210	0.324	**0.243**	**0.165**	**0.272**	**0.377**
Keci	1.000	1.000	1.000	1.000	0.933	0.875	1.000	1.000
	0.103	0.059	0.105	0.174	**0.158**	**0.109**	**0.162**	**0.246**
	0.095	0.054	0.094	0.166	**0.159**	**0.109**	**0.162**	**0.252**
OMult	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.070	0.030	0.061	0.149	**0.202**	**0.102**	**0.283**	**0.389**
	0.071	0.031	0.065	0.145	**0.199**	**0.102**	**0.278**	**0.379**
QMult	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.224	0.147	0.256	0.359	**0.263**	**0.189**	**0.289**	**0.389**
	0.217	0.138	0.252	0.351	**0.254**	**0.183**	**0.279**	**0.379**
TransE	0.779	0.721	0.817	0.882	0.769	0.710	0.806	0.877
	0.563	0.473	0.620	**0.713**	**0.568**	**0.483**	**0.624**	0.704
	0.548	0.459	0.599	**0.696**	**0.555**	**0.475**	**0.602**	0.687
MuRE	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000
	0.541	0.461	**0.554**	**0.758**	0.532	**0.470**	0.547	0.626
	0.518	0.435	**0.533**	**0.742**	0.511	**0.450**	0.526	0.602

References

1. Wang, J., Ilievski, F., Szekely, P., Yao, K.T.: Augmenting knowledge graphs for better link prediction. arXiv preprint arXiv:2203.13965 (2022)