Performance Evaluation of Knowledge Graph Embedding Approaches under Non-adversarial Attacks

Anonymous

No Institute Given

This pdf contains the results in tables for our study on non aderversarial attacks on Knowledge Graph embedding approaches.

1 Results

1.1 Results on Graph Perturbation

Table 1-5 report average MRR performances of models over no perturbation and different ratios of graph perturbation on the FB15k-237, NELL-995-h100, WN18RR, UMLS, and KINSHIP datasets. Each experiment was conducted 5 times.

Table 1. Link prediction results for graph perturbation on UMLS. Each sequence of three rows for a model report the model performance on the training, validation and test datasets, respectively. Here k denotes the perturbation ratio.

Models	k							
	0%	1%	2%	4%	8%	16%	32%	64%
DistMult	0.906	0.906	0.906	0.901	0.895	0.879	0.845	0.757
	0.700	0.708	0.721	0.734	0.745	0.760	0.757	0.701
	0.693	0.698	0.712	0.724	0.740	0.752	0.746	0.696
ComplEx	0.978	0.973	0.973	0.965	0.953	0.937	0.912	0.798
	0.813	0.818	0.822	0.835	0.847	0.847	0.844	0.753
	0.822	0.823	0.826	0.839	0.842	0.841	0.837	0.748
QMult	0.954	0.948	0.946	0.942	0.927	0.903	0.869	0.765
	0.839	0.840	0.837	0.849	0.839	0.833	0.816	0.729
	0.835	0.834	0.836	0.841	0.832	0.822	0.806	0.725
MuRE	0.981	0.970	0.969	0.960	0.943	0.915	0.874	0.768
	0.916	0.907	0.906	0.895	0.878	0.858	0.825	0.739
	0.909	0.899	0.899	0.887	0.874	0.853	0.823	0.731
Keci	0.974	0.970	0.963	0.958	0.949	0.932	0.893	0.792
	0.791	0.810	0.822	0.836	0.849	0.843	0.828	0.749
	0.806	0.819	0.828	0.837	0.849	0.850	0.828	0.739

Table 2. Link prediction results for graph perturbation on KINSHIP. Each sequence of three rows for a model report the model performance on the training, validation and test datasets. Here k denotes the perturbation ratio.

Models	k							
	0%	1%	2%	4%	8%	16%	32%	64%
DistMult	0.655	0.654	0.646	0.646	0.639	0.618	0.569	0.436
	0.520	0.527	0.519	0.526	0.530	0.523	0.499	0.396
	0.519	0.521	0.517	0.524	0.521	0.516	0.485	0.386
ComplEx	0.866	0.863	0.860	0.852	0.829	0.807	0.736	0.555
	0.731	0.734	0.737	0.739	0.734	0.723	0.671	0.511
	0.738	0.739	0.742	0.741	0.728	0.720	0.668	0.501
QMult	0.793	0.785	0.776	0.764	0.738	0.707	0.630	0.359
	0.690	0.684	0.678	0.673	0.650	0.634	0.578	0.323
	0.684	0.674	0.675	0.657	0.645	0.624	0.563	0.318
MuRE	0.725	0.730	0.721	0.698	0.685	0.637	0.555	0.324
	0.645	0.648	0.645	0.626	0.624	0.585	0.516	0.305
	0.626	0.637	0.626	0.610	0.608	0.568	0.495	0.291
Keci	0.836	0.839	0.841	0.833	0.820	0.794	0.726	0.502
	0.729	0.735	0.733	0.739	0.734	0.717	0.667	0.468
	0.731	0.736	0.733	0.738	0.728	0.715	0.656	0.461

Table 3. Link prediction results for graph perturbation on NELL-995-h100. Each sequence of three rows for a model report the model performance on the training, validation and test datasets. Here k denotes the perturbation ratio.

Models	k							
	0%	1%	2%	4%	8%	16%	32%	64%
DistMult	0.818	0.798	0.777	0.745	0.716	0.653	0.554	0.352
	0.184	0.191	0.197	0.202	0.206	0.204	0.205	0.171
	0.181	0.185	0.191	0.194	0.197	0.196	0.198	0.166
ComplEx	0.877	0.842	0.818	0.768	0.706	0.633	0.517	0.315
	0.152	0.173	0.181	0.186	0.189	0.197	0.201	0.174
	0.150	0.172	0.175	0.181	0.186	0.192	0.196	0.170
QMult	0.811	0.768	0.752	0.717	0.682	0.592	0.447	0.228
	0.148	0.153	0.160	0.157	0.171	0.183	0.184	0.154
	0.147	0.151	0.159	0.159	0.169	0.182	0.184	0.151
MuRE	0.774	0.697	0.743	0.688	0.641	0.615	0.515	0.384
	0.273	0.264	0.270	0.263	0.261	0.265	0.255	0.236
	0.266	0.256	0.262	0.256	0.253	0.258	0.246	0.231
Keci	0.876	0.843	0.811	0.808	0.759	0.676	0.552	0.407
	0.186	0.185	0.193	0.207	0.220	0.233	0.228	0.218
	0.182	0.181	0.190	0.203	0.214	0.227	0.219	0.212

Table 4. Link prediction results for graph perturbation on FB15k-237. Each sequence of three rows for a model report the model performance on the training, validation and test datasets, Here k denotes the perturbation ratio.

Models	k							
	0%	1%	2%	4%	8%	16%	32%	64%
DistMult	0.253	0.256	0.236	0.225	0.208	0.183	0.139	0.069
	0.159	0.170	0.161	0.161	0.152	0.144	0.118	0.061
	0.158	0.169	0.162	0.161	0.151	0.143	0.118	0.060
ComplEx	0.214	0.210	0.219	0.203	0.187	0.161	0.116	0.061
	0.144	0.146	0.152	0.147	0.143	0.127	0.101	0.056
	0.144	0.146	0.151	0.146	0.141	0.128	0.100	0.057
QMult	0.402	0.391	0.389	0.373	0.347	0.310	0.247	0.151
	0.244	0.241	0.244	0.244	0.242	0.231	0.208	0.154
	0.241	0.240	0.242	0.242	0.241	0.230	0.207	0.154
MuRE	0.325	0.299	0.291	0.276	0.256	0.222	0.187	0.143
	0.240	0.229	0.225	0.221	0.212	0.197	0.182	0.158
	0.235	0.226	0.222	0.217	0.209	0.195	0.181	0.156
Keci	0.346	0.327	0.285	0.310	0.260	0.250	0.164	0.123
	0.220	0.212	0.193	0.217	0.196	0.200	0.144	0.120
	0.218	0.210	0.194	0.216	0.195	0.198	0.143	0.120

Table 5. Link prediction results for graph perturbation on WN18RR. Each sequence of three rows for a model report the model performance on the training, validation and test datasets. Here k denotes the perturbation ratio.

Models	k							
	0%	1%	2%	4%	8%	16%	32%	64%
DistMult	0.818	0.837	0.812	0.845	0.816	0.745	0.694	0.434
	0.184	0.318	0.306	0.320	0.320	0.284	0.259	0.141
	0.181	0.314	0.299	0.317	0.314	0.280	0.259	0.140
ComplEx	0.877	0.898	0.897	0.857	0.829	0.783	0.642	0.386
	0.307	0.307	0.308	0.294	0.283	0.260	0.201	0.115
	0.301	0.301	0.303	0.292	0.281	0.262	0.202	0.114
MuRE	0.774	0.982	0.958	0.935	0.916	0.848	0.746	0.483
	0.273	0.383	0.377	0.371	0.359	0.336	0.301	0.186
	0.266	0.380	0.371	0.368	0.354	0.335	0.298	0.185
QMult	0.811	0.954	0.950	0.942	0.919	0.866	0.750	0.500
	0.148	0.192	0.184	0.205	0.200	0.202	0.196	0.126
	0.147	0.186	0.181	0.202	0.197	0.200	0.195	0.124
Keci	0.876	0.920	0.917	0.883	0.881	0.817	0.752	0.497
	0.186	0.333	0.335	0.322	0.330	0.312	0.273	0.160
	0.182	0.329	0.329	0.320	0.329	0.313	0.272	0.162

1.2 Results on Parameter Perturbation

Tables 6-10 report average MRR performances of models over no perturbation and different ratios of parameter perturbation on the FB15k-237, NELL-995-h100, WN18RR, UMLS, and KINSHIP datasets. Each experiment was conducted 5 times.

Table 6. Link prediction results for parameter perturbation on UMLS. Each sequence of three rows for a model report the model performance on the training, validation and test datasets, Here k denotes the perturbation ratio.

Models	k							
	0%	1%	2%	4%	8%	16%	32%	64%
DistMult	0.906	0.864	0.806	0.703	0.565	0.532	0.537	0.589
	0.700	0.689	0.660	0.593	0.492	0.491	0.512	0.568
	0.693	0.688	0.664	0.592	0.494	0.489	0.502	0.570
ComplEx	0.978	0.940	0.895	0.785	0.695	0.661	0.668	0.695
	0.813	0.796	0.773	0.695	0.622	0.614	0.628	0.663
	0.822	0.794	0.768	0.688	0.615	0.605	0.618	0.653
QMult	0.954	0.921	0.883	0.816	0.726	0.683	0.678	0.682
	0.839	0.808	0.789	0.729	0.669	0.634	0.636	0.639
	0.835	0.803	0.773	0.717	0.659	0.626	0.632	0.632
MuRE	0.981	0.888	0.792	0.675	0.550	0.491	0.462	0.429
	0.916	0.833	0.743	0.647	0.522	0.462	0.429	0.402
	0.909	0.822	0.729	0.630	0.512	0.454	0.431	0.400
Keci	0.974	0.950	0.915	0.826	0.716	0.642	0.578	0.673
	0.791	0.808	0.798	0.732	0.659	0.602	0.544	0.644
	0.806	0.812	0.788	0.733	0.661	0.597	0.537	0.640

Table 7. Link prediction results for parameter perturbation on KINSHIP. Each sequence of three rows for a model report the model performance on the training, validation and test datasets. Here k denotes the perturbation ratio.

Models	k							
	0%	1%	2%	4%	8%	16%	32%	64%
DistMult	0.655	0.570	0.492	0.376	0.179	0.092	0.099	0.093
	0.520	0.484	0.431	0.340	0.164	0.084	0.089	0.085
	0.519	0.479	0.418	0.335	0.166	0.086	0.092	0.085
ComplEx	0.866	0.752	0.604	0.113	0.095	0.106	0.111	0.115
	0.731	0.668	0.553	0.105	0.088	0.098	0.099	0.105
	0.738	0.661	0.544	0.105	0.086	0.097	0.099	0.104
QMult	0.793	0.564	0.114	0.091	0.092	0.091	0.094	0.095
	0.690	0.511	0.107	0.084	0.086	0.086	0.088	0.085
	0.684	0.504	0.106	0.082	0.087	0.084	0.085	0.086
MuRE	0.725	0.591	0.456	0.122	0.057	0.054	0.053	0.054
	0.645	0.556	0.437	0.115	0.054	0.053	0.051	0.053
	0.626	0.540	0.424	0.116	0.053	0.053	0.055	0.054
Keci	0.836	0.755	0.620	0.268	0.098	0.103	0.100	0.104
	0.729	0.677	0.568	0.253	0.091	0.093	0.092	0.093
	0.731	0.670	0.554	0.246	0.090	0.094	0.092	0.094

Table 8. Link prediction results for parameter perturbation on NELL-995-h100. Each sequence of three rows for a model report the model performance on the training, validation and test datasets. Here k denotes the perturbation ratio.

Models	k							
	0%	1%	2%	4%	8%	16%	32%	64%
DistMult	0.818	0.806	0.810	0.691	0.518	0.393	0.359	0.268
	0.184	0.186	0.180	0.173	0.165	0.155	0.134	0.106
	0.181	0.180	0.174	0.163	0.157	0.148	0.126	0.099
ComplEx	0.877	0.844	0.779	0.633	0.444	0.287	0.230	0.185
	0.152	0.155	0.164	0.162	0.150	0.118	0.098	0.096
	0.150	0.153	0.158	0.155	0.143	0.112	0.092	0.085
QMult	0.811	0.786	0.735	0.630	0.519	0.388	0.287	0.185
	0.148	0.145	0.138	0.124	0.109	0.101	0.105	0.093
	0.147	0.142	0.135	0.119	0.103	0.095	0.097	0.087
MuRE	0.774	0.750	0.791	0.766	0.728	0.626	0.419	0.298
	0.273	0.272	0.269	0.274	0.257	0.232	0.181	0.145
	0.266	0.262	0.262	0.264	0.246	0.227	0.175	0.143
Keci	0.876	0.865	0.845	0.790	0.584	0.418	0.303	0.184
	0.186	0.196	0.194	0.206	0.194	0.176	0.139	0.105
	0.182	0.192	0.191	0.199	0.185	0.170	0.131	0.097

Table 9. Link prediction results for parameter perturbation on FB15k-237. Each sequence of three rows for a model report the model performance on the training, validation and test datasets, Here k denotes the perturbation ratio.

Models	k							
	0%	1%	2%	4%	8%	16%	32%	64%
DistMult	0.253	0.243	0.258	0.245	0.197	0.135	0.122	0.122
	0.159	0.156	0.167	0.159	0.131	0.099	0.097	0.103
	0.158	0.156	0.167	0.158	0.129	0.097	0.094	0.101
ComplEx	0.214	0.213	0.223	0.207	0.178	0.119	0.122	0.126
	0.144	0.144	0.151	0.143	0.122	0.088	0.106	0.122
	0.144	0.145	0.150	0.143	0.120	0.087	0.105	0.120
QMult	0.402	0.403	0.403	0.380	0.279	0.173	0.154	0.148
	0.244	0.244	0.245	0.237	0.193	0.153	0.148	0.148
	0.241	0.243	0.243	0.235	0.192	0.150	0.146	0.146
MuRE	0.325	0.313	0.316	0.311	0.283	0.236	0.182	0.149
	0.240	0.230	0.234	0.230	0.220	0.203	0.172	0.150
	0.235	0.227	0.231	0.226	0.216	0.200	0.169	0.147
Keci	0.346	0.282	0.279	0.295	0.150	0.111	0.102	0.108
	0.220	0.189	0.185	0.202	0.132	0.102	0.099	0.108
	0.218	0.186	0.182	0.202	0.131	0.100	0.097	0.107

Table 10. Link prediction results for parameter perturbation on WN18RR. Each sequence of three rows for a model report the model performance on the training, validation and test datasets. Here k denotes the perturbation ratio.

Models	k							
	0%	1%	2%	4%	8%	16%	32%	64%
DistMult	0.906	0.842	0.866	0.757	0.767	0.611	0.438	0.289
	0.700	0.307	0.331	0.279	0.293	0.239	0.124	0.071
	0.693	0.304	0.326	0.275	0.289	0.240	0.126	0.067
ComplEx	0.978	0.831	0.819	0.880	0.780	0.587	0.171	0.090
	0.813	0.276	0.274	0.300	0.271	0.194	0.027	0.006
	0.822	0.272	0.271	0.299	0.268	0.194	0.026	0.005
QMult	0.954	0.945	0.956	0.937	0.910	0.662	0.415	0.135
	0.839	0.169	0.175	0.156	0.173	0.101	0.026	0.005
	0.835	0.162	0.170	0.155	0.170	0.102	0.027	0.006
MuRE	0.983	0.998	0.979	0.975	0.973	0.938	0.836	0.715
	0.399	0.403	0.405	0.389	0.395	0.377	0.332	0.285
	0.397	0.400	0.399	0.382	0.379	0.366	0.327	0.279
Keci	0.974	0.941	0.858	0.913	0.899	0.430	0.386	0.174
	0.791	0.339	0.302	0.325	0.323	0.144	0.117	0.051
	0.806	0.336	0.303	0.325	0.323	0.145	0.119	0.049