

---

# Few-shot link prediction via graph neural networks for Covid-19 drug-repurposing

---

Vassilis N. Ioannidis<sup>1</sup> Da Zheng<sup>1</sup> George Karypis<sup>1</sup>

## Abstract

Predicting interactions among heterogeneous graph structured data has numerous applications such as knowledge graph completion, recommendation systems and drug discovery. Often times, the links to be predicted belong to rare types such as the case in repurposing drugs for novel diseases. This motivates the task of few-shot link prediction. Typically, GCNs are ill-equipped in learning such rare link types since the relation embedding is not learned in an inductive fashion. This paper proposes an inductive RGCN for learning informative relation embeddings even in the few-shot learning regime. The proposed inductive model significantly outperforms the RGCN and state-of-the-art KGE models in few-shot learning tasks. Furthermore, we apply our method on the drug-repurposing knowledge graph (DRKG) for discovering drugs for Covid-19. We pose the drug discovery task as link prediction and learn embeddings for the biological entities that partake in the DRKG. Our initial results corroborate that several drugs used in clinical trials were identified as possible drug candidates. The method in this paper are implemented using the efficient deep graph learning (DGL) (Wang et al., 2019).

## 1. Introduction

The timeline of the Covid-19 pandemic showcases the dire need for fast development of effective treatments for new diseases. Drug-repurposing is a drug discovery strategy from existing drugs that significantly shortens the time and reduces the cost compared to de novo drug discovery (Sertkaya et al., 2014; Avorn et al., 2015; Setoain et al., 2015). Drug-repurposing leverages the fact that common molecular pathways contribute to different diseases and hence some drugs may be reused (Ashburn & Thor, 2004).

<sup>1</sup>Amazon Web Services AI Lab, Palo Alto, California, USA. Correspondence to: Vassilis N. Ioannidis <ioann006@umn.edu>.

Drug-repurposing relies on identifying novel interactions among biological entities like genes and compounds and can be posed as a link prediction task over a biological network. Several machine learning approaches have been developed for addressing the drug-repurposing task for Covid-19; see e.g. (Gramatica et al., 2014; Zhou et al., 2020; Udrescu et al., 2016; Ioannidis et al., 2020). Towards assisting such machine learning techniques (Ioannidis et al., 2020) created a comprehensive biological knowledge graph relating genes, compounds, diseases, biological processes, side effects and symptoms termed Drug Repurposing Knowledge Graph (DRKG).

However, for novel diseases like Covid-19 only a few interactions are available among viral proteins and possible chemical compounds that may inhibit the related genes. This motivates the framework of few-shot link prediction, where a certain edge type is rare and the model is called to make predictions on the particular edge type.

### 1.1. Related works

Link prediction has been addressed by several works in the context of knowledge-graph (KG) completion. These models rely on embedding the nodes and edges of the KG to a vector space and then train by maximizing the score for existing edges in the KGs; see e.g., (Wang et al., 2017a). An efficient implementation of these models in DGL is presented in (Zheng et al., 2020). Nevertheless, these KGE models do not naturally generalize in the few-shot scenario, where only a few edges are available for a rare edge type, which challenges learning the relation embedding. This was addressed in (Chen et al., 2019), where a meta-learning model is proposed to learn the relation embeddings in an inductive fashion. However, this inductive-relation KGE model require a specialized training scheme, can not learn inductive node embeddings, and can not incorporate node features if available.

Graph convolutional networks learn embeddings for nodes and edges in the graph by applying a sequence of nonlinear operations parametrized by the graph adjacency matrix and utilize node and edge features (Kipf & Welling, 2017; Schlichtkrull et al., 2018). An inductive implementation of these models allows for learning node embeddings in

an inductive fashion (Hamilton et al., 2017). The RGCN model (Schlichtkrull et al., 2018) has been successful in link prediction, where the RGCN is supervised by KGE models (Wang et al., 2017a). However, these GCN models for link prediction inherit the limitation of the KGE models, and are challenged in learning relation embeddings for rare edges types.

## 1.2. Contributions

This paper addresses the aforementioned limitation of GCN models by introducing a novel inductive-RGCN that learns the relation and the node embeddings in an inductive fashion. The proposed I-RGCN naturally addresses the few-shot link prediction and outperforms competing state-of-the-art models. I-RGCN is also tested in the DRKG for Covid-19 drug-repurposing. The drug discovery task is naturally formulated in a few-shot learning setting. The preliminary results indicate that several drugs used in clinical trials are discovered as possible drug candidates. While this study, by no means recommends specific drugs, it demonstrates a powerful deep learning methodology to prioritize existing drugs for further investigation, which holds the potential of accelerating therapeutic development for COVID-19.

## 2. Few-shot link prediction formulation

Consider the heterogeneous graph with  $T$  node types and  $R$  relation types defined as  $\mathcal{G} := \{\{\mathcal{V}_t\}_{t=1}^T, \{\mathcal{E}_r\}_{r=1}^R\}$ . The  $t$ th node type is defined as  $\mathcal{V}_t := \{v_n^t\}_{n=1}^{N_t}$  and may represent Genes or Chemical compounds in the DRKG. The  $r$ th relation type holds all interactions of a certain type among  $\mathcal{N}^t$  and  $\mathcal{N}^{t'}$   $\mathcal{E}_r := \{(v_n^t, v_{n'}^{t'}) \in \mathcal{N}^t \times \mathcal{N}^{t'}\}$  and may represent that a chemical compound inhibits a gene or that a disease is treated by a chemical compound.

Consider also that each node  $n_t$  is associated with a  $F \times 1$  feature vector  $\mathbf{x}_{n_t}$ . This feature may represent an embedding of the protein sequence associated with a gene (Wang et al., 2017b). In KGs some node types may not have features for these we use an embedding layer to represent their features.

**Few shot link prediction.** Given  $R - 1$  sets of edges  $\{\mathcal{E}_r\}_{r=1}^{R-1}$ , a nodal attribute vector  $\mathbf{x}_{n_t}$  per node  $n_t$ , and a small set of links in the few-shot relation  $\mathcal{E}_R$  with  $|\mathcal{E}_R| \leq K$ , the few-shot link prediction amounts to inferring the missing links of the rare type  $R$ . In the DRKG, this few-shot relation is for example coronavirus treatment.

## 3. Learning inductive embedding for GNNs

The relational GCN (RGCN) (Schlichtkrull et al., 2018) extends the graph convolution operation (Kipf & Welling, 2017) to heterogenous graphs. An RGCN model is com-

prised by a sequence of RGCN layers. The  $l$ th layer computes the  $n$ th node representation  $\mathbf{h}_n^{(l+1)}$  as follows

$$\mathbf{h}_n^{(l+1)} := \sigma \left( \sum_{r=1}^R \sum_{n' \in \mathcal{N}_n^r} \mathbf{h}_{n'}^{(l)} \mathbf{W}_r^{(l)} \right) \quad (1)$$

where  $\mathcal{N}_n^r$  is the neighborhood of node  $n$  under relation  $r$ ,  $\sigma$  the rectified linear unit non linear function, and  $\mathbf{W}_r^{(l)}$  is a learnable matrix associated with the  $r$ th relation. Essentially, the output of the RGCN layer for node  $n$  is a nonlinear combination of the hidden representations of neighboring nodes weighted based on the relation type. The node features are the input of the first layer in the model i.e.  $\mathbf{h}_n^{(0)} = \mathbf{x}_n$ , where  $\mathbf{x}_n$  is the node feature for node  $n$ . For node types without features we use an embedding layer that takes as input an one-hot encoding of the node id.

The RGCN model in this paper is supervised by a DistMult model (Yang et al., 2014) for link prediction. The loss function

$$\min \sum_{n_t, r, n_{t'} \in \mathbb{D}^+ \cup \mathbb{D}^-} \log(1 + \exp(-y \times \mathbf{h}_{n_t}^\top \text{diag}(\mathbf{h}_r) \mathbf{h}_{n_{t'}})) \quad (2)$$

where  $\mathbf{h}^\top$  denotes the transpose of a matrix,  $\text{diag}(\mathbf{r})$  denotes a diagonal matrix with  $\mathbf{r}$  on its diagonal,  $\mathbf{h}_{n_t}$ ,  $\mathbf{h}_r$ ,  $\mathbf{h}_{n_{t'}}$  are the embedding of the head entity  $n_t$ , relation  $r$  and the tail entity  $n_{t'}$ , respectively and  $\mathbb{D}^+$  and  $\mathbb{D}^-$  are the positive and negative sets of triplets and  $y = 1$  if the triplet corresponds to a positive example and  $-1$  otherwise. The scalar represented by  $\mathbf{h}^\top \text{diag}(\mathbf{h}_r) \mathbf{t}$  denotes the score of triplet  $(h, r, t)$  as given by the DistMult model (Yang et al., 2014). The entity embeddings are obtained by the final layer of the RGCN. The relation type embedding are trained directly from (2).

Such a model (2) is vulnerable when only few training edges are available for a certain relation type. The small number of edges will challenge the learning of the embedding vector  $\mathbf{r}$  for the rare relation.

### 3.1. Inductive RGCN

Certain relation-types may be rare in the training set of links and require a specialized architecture. To address such a few-shot scenario, we introduce a MLP to learn the relation embeddings. Consider the node embeddings  $\{\mathbf{h}_{n_t}\}$  and  $\{\mathbf{h}_{n_{t'}}\}$  extracted from the ultimate layer of the RGCN model where  $n_t \in \mathcal{V}_t$  and  $n_{t'} \in \mathcal{V}_{t'}$ . The proposed MLP learns an embedding for the  $r$ th relation as follows

$$\mathbf{h}_r := \frac{1}{|\mathcal{E}_r|} \sum_{(n_t, n_{t'}) \in \mathcal{E}_r} \sigma(\mathbf{W}_1 \sigma(\mathbf{W}_2 (\mathbf{h}_{n_t} \parallel \mathbf{h}_{n_{t'}}))) \quad (3)$$

Table 1. Statistics of datasets.

	Nodes	Edges	Relation types
IMDb	movie : 4,278	17,106	12
	director : 2,081		
	actor : 5,257		
DBLP	author : 4,057	119,783	12
	paper : 14,328		
	term : 7,723		
	venue : 20		

where  $\parallel$  denotes the vector concatenation. Note that the relation embedding  $\mathbf{h}_r$  is calculated as a nonlinear function of the node embedding for all node pairs  $(n_t, n_{t'})$  participating to a certain relation type  $(n_t, n_{t'}) \in \mathcal{E}_r$ . This allows the I-RGCN to learn relation embeddings in an inductive fashion. This model is supervised by the following loss

$$\mathcal{L}_{\text{FSLP}} := \log(1 + \exp(-y \times \mathbf{h}_{n_t}^\top \text{diag}(\mathbf{h}_r) \mathbf{h}_{n_{t'}})) \quad (4)$$

where  $\mathbf{h}_{n_t}^\top \text{diag}(\mathbf{h}_r) \mathbf{h}_{n_{t'}}$  denotes the triple score and  $y$  is  $-1$  for the negative triples and  $1$  stands for the positive ones. We create negative triples by fixing the head node of a positive triple and randomly selecting a tail node of the same type as the original tail node. Differently from (2), the relation embedding  $\mathbf{h}_r$  is learned in an inductive fashion from the participating node pairs  $(n_t, n_{t'})$ . Hence, upon learning the MLP parameters  $\mathbf{W}_1, \mathbf{W}_2$  the relation embedding will be computed with a forward pass. This obviates the few-shot learning hurdle and enables the model to generalize to rare or even unseen relations.

## 4. Experiments

### 4.1. Few shot link prediction

**Baselines.** We consider the state-of-the-art KGE models RotatE (Sun et al., 2019), ComplEx (Trouillon et al., 2016), and the RGCN model (Schlichtkrull et al., 2018) as baselines for comparison. The parameters of these methods have been optimized via cross validation. We use the IMDB and DBLP datasets (Fu et al., 2020) detailed in Table 1. The total number of edges in the few-shot relation are 1559 for the IMDB and 3534 for the DBLP. In the experiments, we train with only  $K$  links from the few-shot relation and all the links from the other relations and test on the rest edges of the few-shot relation, which are  $|\mathcal{E}_R| - K$ . The nodes in the IMDB and DBLP graphs are associated with feature vectors. Further, information on the datasets is included in the Appendix.

Tables 2 and 3 report the MRR, Hit-1 and Hit-10 scores of the baseline methods along with the inductive RGCN and the RGCN in the task of few-shot link prediction for the IMDB and DBLP datasets respectively. The I-RGCN significantly outperforms the alternative methods in the task of few-shot link prediction. Specifically, for  $K=10$  the MRR of

the inductive method is one order of magnitude greater. This corroborates the advantage of the inductive relation learning for the few-shot learning task. As the number of training edges increases at  $K=1000$ , it is observed that the RGCN performance approaches the performance of the I-RGCN. This suggests that the I-RGCN method performs well also in non few-shot learning tasks. The worse performance of KGE models is explained since these do not account for node features and do not learn inductive relation embeddings.

To further validate the performance of the I-RGCN we conduct a general link prediction evaluation by splitting the links in training, validation, and testing at random *irrespective* of their relation type. The results for different percentages of training links are reported in Table 4. I-RGCN outperforms even in this training scenario the RGCN and KGEs baselines, which further corroborates the efficiency of the model.

### 4.2. Drug-repurposing via I-RGCN

For this experiment we will utilize the drug-repurposing knowledge graph (DRKG) constructed in (Ioannidis et al., 2020). The DRKG collects interactions from a collection of biological databases such as Drugbank (DS et al., 2017b), GBR (Percha & Altman, 2018), Hetionet (DS et al., 2017a), STRING (D et al., 2019), IntAct (Orchard et al., 2014) and DGIdb (Cotto et al., 2017).

Drug-repurposing aims at discovering the most effective existing drugs to treat a certain disease. Drug-repurposing can be formulated as predicting direct links in the DRKG such as predicting whether a drug treats a disease or as predicting whether a compound inhibits a certain gene which is related to the target disease. Drug-repurposing can be viewed as a few-shot link prediction task since only a few edges are available related to novel diseases in the DRKG.

We use corona-virus related diseases, including SARS, MERS and SARS-COV2, as target diseases representing Covid-19 as their functionality is similar. We aim at predicting links among gene entities associated with the target disease and drug entities. We select FDA-approved drugs in Drugbank as candidates, while we exclude for simplicity drugs with molecule weight less than 250 daltons, as many of certain drugs are actually health drugs. This amounts to 8104 candidate drugs.

We also obtain 442 Covid-19 related genes from the relations extracted from (Gordon et al., 2020; Zhou et al., 2020). Similarly, we obtain the node embeddings for the gene and drugs, and the embeddings for the corresponding relations. Next, we score all triples and rank them per target gene. This way we obtain 442 ranked lists of drugs. Finally, to assess whether our prediction is in par with the drugs used for treatment, we check the overlap among the top 100 pre-

Table 2. Experiment results (%) on the IMDb dataset for k-shot link prediction.

K	MRR				Hit 1				Hit 10			
	ComplEx	RotatE	RGCN	I-RGCN	ComplEx	RotatE	RGCN	I-RGCN	ComplEx	RotatE	RGCN	I-RGCN
10	6.88	11.80	1.32	<b>33.56</b>	1.75	6.74	0.13	<b>25.32</b>	14.25	18.35	1.07	<b>53.55</b>
50	8.48	12.56	15.26	<b>53.24</b>	3.34	7.76	7.42	<b>45.14</b>	15.70	19.12	28.88	<b>69.32</b>
100	8.61	12.57	18.78	<b>53.63</b>	3.44	7.86	9.59	<b>40.27</b>	15.71	18.40	36.84	<b>77.38</b>
1000	68.37	70.09	95.23	<b>96.06</b>	65.48	67.52	91.72	<b>93.56</b>	72.23	73.50	99.72	<b>99.81</b>

Table 3. Experiment results (%) on the DBLP dataset for k-shot link prediction.

K	MRR				Hit 1				Hit 10			
	ComplEx	RotatE	RGCN	I-RGCN	ComplEx	RotatE	RGCN	I-RGCN	ComplEx	RotatE	RGCN	I-RGCN
10	5.97	7.41	7.42	<b>27.95</b>	1.37	2.13	2.95	<b>13.78</b>	11.81	15.36	12.51	<b>60.51</b>
50	6.32	7.87	17.25	<b>83.42</b>	1.55	2.26	10.95	<b>72.54</b>	12.59	16.64	26.84	<b>96.82</b>
100	7.24	10.66	32.45	<b>90.00</b>	1.99	04.43	23.46	<b>85.27</b>	14.96	20.92	49.85	<b>97.61</b>
1000	36.56	46.51	91.34	<b>96.82</b>	30.62	39.83	86.43	<b>94.41</b>	46.45	59.27	98.59	<b>99.81</b>

Table 4. Experiment results (%) on the IMDb dataset for link prediction.

Metrics	MRR				Hit 1				Hit 10			
	ComplEx	RotatE	RGCN	I-RGCN	ComplEx	RotatE	RGCN	I-RGCN	ComplEx	RotatE	RGCN	I-RGCN
Training links												
95%	94.15	93.97	93.38	<b>95.12</b>	<b>93.75</b>	93.48	89.31	92.75	94.74	94.56	<b>99.29</b>	98.24
90%	88.87	88.89	89.30	<b>93.98</b>	88.25	87.87	83.45	<b>91.52</b>	89.74	90.35	97.66	<b>98.12</b>
80%	78.55	78.90	83.46	<b>90.03</b>	77.45	76.83	76.59	<b>86.70</b>	79.96	81.94	94.72	<b>95.93</b>
70%	69.59	69.56	82.73	<b>87.00</b>	67.98	66.73	76.09	<b>82.95</b>	71.89	73.49	93.76	<b>94.07</b>
60%	60.40	60.90	78.16	<b>81.53</b>	58.49	57.60	70.63	<b>77.01</b>	62.92	65.57	<b>91.14</b>	90.27

dicted drugs and the drugs used in clinical trials per gene. We used 32 clinical trial drugs for Covid-19 to validate our predictions<sup>1</sup>. Table 5 lists the clinical drugs included in the top-100 predicted drugs across all the genes with their corresponding number of hits for the RGCN and I-RGCN. It can be observed, that several of the widely used drugs in clinical trials appear high on the predicted list, and that I-RGCN shows a higher hit rate than RGCN. Hence, the inductive relation prediction module is more appropriate in predicting links when information about the nodes is limited, such as is the case with the novel Covid-19 disease node.

## 5. Conclusion

In this paper we develop a novel I-RGCN that learns inductive relation embeddings and can be applied for few-shot link prediction and drug repurposing. I-RGCN consistently outperforms baseline models in the IMDB and DBLP datasets for few-shot link prediction. We also formulate the Covid-19 drug-repurposing task as a link prediction over the DRKG. I-RGCN successfully identifies a subset of clinical trial drugs for Covid-19 and can be used to assist researchers and prioritize existing drugs for further investigation in the Covid-19 treatment.

<sup>1</sup>The clinical trial drugs were collected from <http://www.covid19-trials.com/>

Table 5. Drug inhibits gene scores for Covid-19. Note that a random classifier will result to approximately 5.3 per drug. This suggests that the reported predictions are significantly better than random.

I-RGCN		RGCN	
Drug name	# hits	Drug name	# hits
Dexamethasone	240	Chloroquine	69
Ribavirin	142	Colchicine	41
Colchicine	128	Tetrandrine	40
Chloroquine	115	Oseltamivir	37
Methylprednisolone	86	Azithromycin	36
Tofacitinib	75	Tofacitinib	33
Thalidomide	70	Ribavirin	32
Losartan	64	Methylprednisolone	30
Hydroxychloroquine	48	Deferoxamine	30
Oseltamivir	46	Thalidomide	25
Deferoxamine	34	Dexamethasone	24
Ruxolitinib	23	Bevacizumab	21
Azithromycin	23	Hydroxychloroquine	19
Nivolumab	11	Losartan	19
Tradipitant	11	Ruxolitinib	13
Bevacizumab	10	Ecilizumab	12
Ecilizumab	7	Tocilizumab	11
Baricitinib	6	Anakinra	11
Sarilumab	6	Sarilumab	8
Tetrandrine	6	Nivolumab	6

## References

- Ashburn, T. T. and Thor, K. B. Drug repositioning: identifying and developing new uses for existing drugs. *Nature reviews Drug discovery*, 3(8):673–683, 2004.
- Avorn, J. et al. The \$2.6 billion pill—methodologic and policy considerations. *N Engl J Med*, 372(20):1877–1879, 2015.
- Chen, M., Zhang, W., Zhang, W., Chen, Q., and Chen, H. Meta relational learning for few-shot link prediction in knowledge graphs. *arXiv preprint arXiv:1909.01515*, 2019.
- Cotto, K. C., Wagner, A. H., Feng, Y.-Y., Kiwala, S., Coffman, A. C., Spies, G., Wollam, A., Spies, N. C., Griffith, O. L., and Griffith, M. DGIdb 3.0: a redesign and expansion of the drug–gene interaction database. *Nucleic Acids Research*, 46(D1):D1068–D1073, 11 2017. ISSN 0305-1048. doi: 10.1093/nar/gkx1143. URL <https://doi.org/10.1093/nar/gkx1143>.
- D, S., AL, G., D, L., A, J., S, W., J, H.-C., M, S., NT, D., JH, M., P, B., LJ, J., and von Mering C. String v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *PubMed*, 2019. doi: 10.1093/nar/gky1131.
- DS, H., A, L., C, H., L, B., SL, C., D, H., A, G., P, K., and SE, B. Systematic integration of biomedical knowledge prioritizes drugs for repurposing. *eLife*, 2017a. doi: 10.7554/eLife.26726.
- DS, W., YD, F., AC, G., EJ, L., A, M., JR, G., T, S., D, J., C, L., Z, S., N, A., I, I., Y, L., A, M., N, G., A, W., L, C., R, C., D, L., A, P., C, K., and M, W. Drugbank 5.0: a major update to the drugbank database for 2018 nucleic acids res. *PubMed*, 2017b. doi: 10.1093/nar/gkx1037.
- Fu, X., Zhang, J., Meng, Z., and King, I. Magnn: Metapath aggregated graph neural network for heterogeneous graph embedding. In *Proceedings of The Web Conference 2020*, pp. 2331–2341, 2020.
- Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., O’meara, M. J., Guo, J. Z., Swaney, D. L., Tummino, T. A., Huttenhain, R., et al. A sars-cov-2-human protein-protein interaction map reveals drug targets and potential drug-repurposing. *Nature*, 2020. doi: 10.1038/s41586-020-2286-9.
- Gramatica, R., Di Matteo, T., Giorgetti, S., Barbiani, M., Bevec, D., and Aste, T. Graph theory enables drug repurposing—how a mathematical model can drive the discovery of hidden mechanisms of action. *PLoS one*, 9(1), 2014.
- Hamilton, W., Ying, Z., and Leskovec, J. Inductive representation learning on large graphs. In *Advances in neural information processing systems*, pp. 1024–1034, 2017.
- Ioannidis, V. N., Song, X., Manchanda, S., Li, M., Pan, X., Zheng, D., Ning, X., Zeng, X., and Karypis, G. Drkg - drug repurposing knowledge graph for covid-19. <https://github.com/gnn4dr/DRKG/>, 2020.
- Kipf, T. N. and Welling, M. Semi-supervised classification with graph convolutional networks. In *Proc. Int. Conf. on Learn. Representations*, Toulon, France, Apr. 2017.
- Orchard, S., Ammari, M., Aranda, B., Breuza, L., Briganti, L., Broackes-Carter, F., Campbell, N. H., Chavali, G., Chen, C., Del-Toro, N., et al. The mintact project—intact as a common curation platform for 11 molecular interaction databases. *Nucleic acids research*, 42(D1):D358–D363, 2014.
- Percha, B. and Altman, R. B. A global network of biomedical relationships derived from text. *Bioinformatics*, 34(15):2614–2624, 2018.
- Schlichtkrull, M., Kipf, T. N., Bloem, P., Van Den Berg, R., Titov, I., and Welling, M. Modeling relational data with graph convolutional networks. In *European Semantic Web Conference*, pp. 593–607. Springer, 2018.
- Sertkaya, A., Birkenbach, A., Berlind, A., and Eyraud, J. Examination of clinical trial costs and barriers for drug development: report to the assistant secretary of planning and evaluation (aspe). *Washington, DC: Department of Health and Human Services*, 2014.
- Setoain, J., Franch, M., Martínez, M., Tabas-Madrid, D., Sorzano, C. O., Bakker, A., Gonzalez-Couto, E., Elvira, J., and Pascual-Montano, A. Nfinder: an online bioinformatics tool for searching similar transcriptomics experiments in the context of drug repositioning. *Nucleic acids research*, 43(W1):W193–W199, 2015.
- Sun, Z., Deng, Z.-H., Nie, J.-Y., and Tang, J. Rotate: Knowledge graph embedding by relational rotation in complex space. *arXiv preprint arXiv:1902.10197*, 2019.
- Trouillon, T., Welbl, J., Riedel, S., Gaussier, É., and Bouchard, G. Complex embeddings for simple link prediction. *International Conference on Machine Learning (ICML)*, 2016.
- Udrescu, L., Sbârcea, L., Topîrceanu, A., Iovanovici, A., Kurunczi, L., Bogdan, P., and Udrescu, M. Clustering drug-drug interaction networks with energy model layouts: community analysis and drug repurposing. *Scientific reports*, 6(1):1–10, 2016.

- Wang, M., Yu, L., Zheng, D., Gan, Q., Gai, Y., Ye, Z., Li, M., Zhou, J., Huang, Q., Ma, C., et al. Deep graph library: Towards efficient and scalable deep learning on graphs. *arXiv preprint arXiv:1909.01315*, 2019.
- Wang, Q., Mao, Z., Wang, B., and Guo, L. Knowledge graph embedding: A survey of approaches and applications. *IEEE Transactions on Knowledge and Data Engineering*, 29(12):2724–2743, 2017a.
- Wang, S., Sun, S., Li, Z., Zhang, R., and Xu, J. Accurate de novo prediction of protein contact map by ultra-deep learning model. *PLoS computational biology*, 13(1): e1005324, 2017b.
- Yang, B., Yih, W.-t., He, X., Gao, J., and Deng, L. Embedding entities and relations for learning and inference in knowledge bases. *arXiv preprint arXiv:1412.6575*, 2014.
- Zheng, D., Song, X., Ma, C., Tan, Z., Ye, Z., Dong, J., Xiong, H., Zhang, Z., and Karypis, G. Dgl-ke: Training knowledge graph embeddings at scale. *arXiv preprint arXiv:2004.08532*, 2020.
- Zhou, Y., Hou, Y., Shen, J., Huang, Y., Martin, W., and Cheng, F. Network-based drug repurposing for novel coronavirus 2019-ncov/sars-cov-2. *Cell discovery*, 6(1): 1–18, 2020.

## A. Datasets

We use the IMDB and DBLP datasets (Fu et al., 2020) detailed in Table 1, where the third column denotes the total size of edges in the few-shot relation that is  $|\mathcal{E}_R|$ . The nodes in the IMDB and DBLP graphs are associated with feature vectors. The original datasets in (Fu et al., 2020) are used for node classification. We adapt the datasets and create new edge types, where the edges are parametrized by the label of the associated nodes. For example, the edge type (director, directed, movie) becomes (director, directed\_drama, movie) if the associated movie is in the drama genre, and the same transformation undergoes the (actor, played, movie) relation. Since, there are 3 labels for movies, this way the original 4 edge types become 12. The same transformation happens in the DBLP dataset.