Computational Drug-target Interaction Prediction based on Graph Embedding and Graph Mining

Maha A. Thafar
King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center (CBRC), Thuwal, Saudi Arabia
Taif University, College of Computers and Information Technology, Taif, Saudi Arabia
maha.thafar@kaust.edu.sa

Somayah Albaradie
King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center (CBRC), Thuwal, Saudi Arabia
King Abdulaziz University, Faculty of Computing and Information Technology, Jeddah, Saudi Arabia
somayah.albaradei@kaust.edu.sa

Rawan S. Olayan
King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center (CBRC), Thuwal, Saudi Arabia
rawan.olayan@kaust.edu.sa

Haitham Ashoor
King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center (CBRC), Thuwal, Saudi Arabia
haitham.ashoor@kaust.edu.sa

Magbubah Essack
King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center (CBRC), Thuwal, Saudi Arabia
magbubah.essack@kaust.edu.sa

Vladimir B. Bajic
King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center (CBRC), Thuwal, Saudi Arabia
vladimir.bajic@kaust.edu.sa

ABSTRACT

Identification of interactions of drugs and proteins is an essential step in the early stages of drug discovery and in finding new drug uses. Traditional experimental identification and validation of these interactions are still time-consuming, expensive, and do not have a high success rate. To improve this identification process, development of computational methods to predict and rank likely drug-target interactions (DTI) with minimum error rate would be of great help. In this work, we propose a computational method for (Drug-Target interaction prediction using Graph Embedding and graph Mining), DTiGEM. DTiGEM models identify novel DTIs as a link prediction problem in a heterogeneous graph constructed by integrating three networks, namely: drug-drug similarity, target-target similarity, and known DTIs. DTiGEM combines different techniques, including graph embeddings (e.g., node2vec), graph mining (e.g., path scores between drugs and targets), and machine learning (e.g., different classifiers). DTiGEM achieves improvement in the prediction performance compared to other state-of-the-art methods for computational prediction of DTIs on four benchmark datasets in terms of area under precision-recall curve (AUPR). Specifically, we demonstrate that based on the average AUPR score across all benchmark datasets, DTiGEM achieves the highest average AUPR value (0.831), thus reducing the prediction error by 22.4% relative to the second-best performing method in the comparison.

ICBBB '20, January 19–22, 2020, Kyoto, Japan

ACM ISBN 978-1-4503-7676-1/20/01…$15.00
DOI:https://doi.org/10.1145/3386052.3386062

CCS Concepts

•Computing Methodologies → Supervised learning by classification

Keywords

Drug discovery; Drug-target interaction prediction; Machine learning; Binary classification; Graph embedding; Graph mining; Bioinformatics; Cheminformatics

1. INTRODUCTION

Drug target interactions (DTIs) prediction is a useful step in drug discovery as well as the prediction of drugs with the same or multiple targets that could cause conflicts between medication or adverse drug reactions [1], [2]. Traditional experimental approaches for identifying DTIs are still time- consuming, expensive, and have low success rates [3]. Therefore, in the last 10 years, research towards developing computational methods for DTIs prediction has received much attention. These computational methods can significantly reduce the time and costs, as well as improve the drug discovery efficiency in comparison with the experimental approaches.

In this work, we propose a computational method DTiGEM (Drug-Target interaction prediction using Graph Embedding and graph Mining) for DTIs prediction. DTiGEM combines similarity-based as well as feature-based techniques. It uses graph embedding, graph-mining, and ML. We evaluate the performance by comparison to six state-of-the-art DTIs predictions methods, using gold-standard benchmark datasets, and show that DTiGEM outperforms these methods.

Here, the structure of this paper is as follow, we discuss the different computational methods that have been developed to predict DTIs in section 2. Datasets, data resources, and data descriptors are discussed in section 3. Section 4 formulates problem and describes the proposed method to predict DTIs. After that, the evaluation metrics and experimental settings are described in section 5. All results and comparison with state-of-

the-art methods are discussed in section 6. Finally, section 7 concludes the work.

download a template from [2], and replace the content with your own material.

2. BACKGROUND

Recent studies to predict DTIs can be grouped into several categories mainly, network analysis-based methods [4]-[7], those based on artificial intelligence (AI) and machine learning (ML) [4]-[9], and those using deep learning (DL) [10]-[12]. Many comprehensive review articles summarized, analyzed, and compared these methods [13]-[18]. Some of these studies integrate two or more approaches to boost the computational prediction.

ML-based methods could implement a feature-based approach [17] or similarity-based approach [19] or both [20]. In a feature-based approach, feature vectors generated by extracting different features from chemical descriptors of drugs and descriptors of targets are representative of known DTIs. One recent example of a feature-based method is cumulative feature subspace boosting for drug-target interaction prediction (CFSBoost) [21], which uses a simple and computationally low cost ensemble learning and boosting classification model for prediction of DTIs based on evolutionary and structurally derived features. Similarity-based ML approaches were built based on the "guilt-by-association" rule that indicates that similar drugs tend to interact with the same target and that a drug can interact with multiple similar targets. These approaches are used to infer DTIs as the link prediction problem in a graph. Many models have been developed based on this assumption, with proven efficacy [16]. For example, one method named self-training bipartite local mode (SELF-BLM) [22] performed k-medoids clustering using drug similarity and protein similarity to classify DTIs as positive, negative, or unknown. Then used a self-training SVM algorithm to identify potential interactions among unknown interactions. Introducing these different types of similarities lead to the development of methods that combine multiple similarity measures.

Other recent studies of DTIs prediction that demonstrates their strength are network-based approaches [4], [6], [7], [23]. These works utilize heterogeneous graph and then extract features using different graph-mining techniques. DASPfind [6] finds the simple paths between each source node (drug) and target node (protein) of particular lengths as inferred from a graph. This graph is constructed using known DTIs, drug-drug similarities, and target-target similarities. Then it ranks DTIs based on specific scores and rank predicted DTIs. In finding the top 1% of DTIs, DASPfind outperforms other methods. Although network-based and topology-based DTIs methods proved their strength, these methods are incapable of computing the topological similarities between nodes of the biological graph. They also cannot be scaled to a large graph. Thus, DL methods, which offer a solution for generating features of vertices automatically in a large network were considered for DTIs prediction.

Using a DL based approach is a new trend in the computational prediction of DTIs [10], [24], [25]. The advantage of DL is evident for large-scale data, including data represented as a network. Any heterogeneous network topology has abundant interactions between biomedical entities, and similarity-based methods use this network to predict DTIs based on the diverse array of features for both drugs and their targets. For example, the DL based method, DeepWalk [10], implements short random walks on a heterogeneous network created from biomedically linked datasets. Each of these random walks produces sentences subsequently processed by word2vec. This topology-based DeepWalk report better performance than other methods that use topology-based similarities such as (Jaccard, Simpson, Geometric, Cosine, Pearson Correlation Coefficient (PCC), and SimRank), as well as when using similarity measures derived from chemical structures or genomic sequences.

Feature embedding and graph embedding can be a part of the DL process. Graph embedding and knowledge-graph mining techniques have been used in different studies for drug repositioning and DTI prediction [5], [26]-[28]. Graph embedding technique maps each node to a low-dimensional feature vector, tries to preserve the connection strengths between nodes and learns the distributed representation description for each node [29]. For example, DTINet [5] predicts novel DTIs from a heterogeneous graph and integrates drug-related information from the DrugBank dataset. It learns a low dimensional feature representation that captures the topological properties of each node in the graph and predicts the DTIs based on this feature representation. DTINet is reported to outperform other state-of-the-art methods.

Although these effective computational models for identification of DTIs have achieved significant improvements, there is still much room for improvement by developing different methods. In this study, we propose a computational method DTiGEM (Drug-Target interaction prediction using Graph Embedding and graph Mining) for DTIs prediction. DTiGEM combines similarity-based as well as feature-based techniques. It uses graph embedding, graph-mining, and ML. We evaluate the performance by comparison to six state-of-the-art DTIs predictions methods, using gold-standard benchmark datasets, and show that DTiGEM outperforms these methods.

3. MATERIALS

3.1 Benchmark Datasets

We used four datasets collected and compiled by [30] which are commonly used as a benchmark datasets to evaluate DTI prediction methods. Each one of these four datasets represent one of the four major families of protein targets, namely enzyme (E), ion channel (IC), G-protein-coupled receptor (GPCR), and nuclear receptor (NR). Each dataset includes three types of information: Known DTIs, one drug-drug similarity type, and one target-target similarity type as described in the following subsection. Table 1 provides basic statistics about these four benchmark datasets. The above-mentioned datasets are publicly available at: http://web.kuicr.kyoto-u.ac.jp/supp/yoshi/drugtarget/.

Table 1. Basic statistics of the Yamanishi_08 benchmark datasets

Statistics	*NR*	*GPCR*	*IC*	*Enzyme*
No. Of drugs	54	223	210	445
No. Of targets	26	95	204	664
Known DTIs	90	635	1476	2926
Unknown DTIs	1314	20,550	41,364	292,554

3.2 Data Description

3.2.1 Drugs' chemical data

As mentioned in [30], chemical structures of drugs are collected from the KEGG database, specifically KEGG LIGAND and KEGG DRUG database [31]. Then similarity scores were

calculated for each pair of compounds using SIMCOMP [32]. The drug-drug similarity matrix was constructed and provided with the dataset.

3.2.2 Targets' genomic data

In the DTIs prediction problem, we consider only proteins as drug targets. As described in [30], the amino acid sequence of the proteins was collected from KEGG GENES database [31]. Sequence similarities were computed using the normalized Smith-Waterman scores [33] for each pair of targets. The target-target similarity matrix was constructed and provided with the dataset.

4. METHODS

4.1 Problem Formulation

Here we adopt a network-based method for DTIs prediction. Three subgraphs were used to construct a weighted heterogeneous graph *G (V, E)*. These three subgraphs are Kdti (known DTIs), DDs (drug-drug similarities), and TTs (target-target similarities). The connections between these subgraphs are DDs - Kdti - TTs. This connected graph *G* has two types of nodes: drugs $D = \{d_1, d_2, ..., d_n\}$, and targets $T = \{t_1, t_2, ..., t_m\}$, and three types of edges which are: DDs edges, TTs edges, and Kdti edges between drug and target. Similarity scores represent the edge weights between similar types of nodes. The weights are in the range of (0 1]. The third type of edge is the interaction edges between drugs and targets where the weights are equal to 1. The aim is to find the missing edges between drugs and targets as a link prediction task. All possible drug-target pairs are constructed by generating negative samples between two nodes that have no edges connecting them. We generated features for each pair (drug, target) using different techniques, discussed later. If there is a known interaction for any pair of (drug, target), the class label *y* for this pair is equal to 1; otherwise the class label is equal to zero. The goal is to find novel DTI with high accuracy and low false-positive rate. The problem depiction is shown in Figure 1.

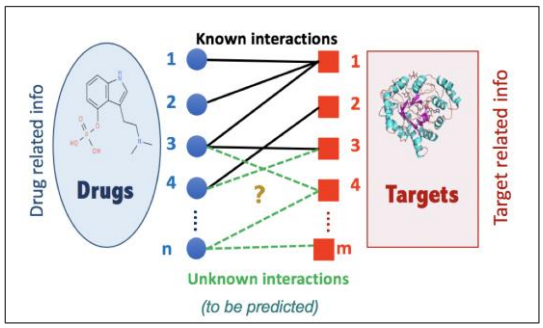

Figure 1. DTIs problem depiction

The proposed method for DTIs prediction is focusing on ML. It combines similarity-based, feature-based, as well as graph-based methods. For similarity-based methods, two types of similarities between each pair of drugs and each pair of targets are calculated in different steps and used for the inference from graph-based on the assumption that similar drugs target similar proteins and similar proteins are targets of the same drug. For feature-based and graph-based methods, features extraction is done based on constructing a heterogeneous DTIs graph and then generating features by calculating different path scores between each (drug, target) pair.

4.2 Graph Embedding Technique

Graph embedding converts the graph data into a low dimensional space in which the graph structural information and graph properties are significantly preserved [34]. Several graph embedding techniques have been applied for random walk in heterogeneous graphs with proven efficacy [29]. One such technique is node2vec [35] which is an algorithmic framework that allows for scalable feature representation learning for heterogeneous graphs. It is a generalized version of DeepWalk [36]. The intuition to use node2vec is to find a mapping of each node to low d-dimensional vector space that preserves the level of node similarity based on neighboring nodes. Two classical search strategies are used to define the neighborhood of a given node for sampling: depth-first search (DFS) and breadth-first search (BFS). Two parameters control the different versions of searching in DeepWalk: return parameter, *p* and in-out parameter, *q*. Parameter *p* controls the likelihood of immediately revisiting a node in the walk, while *q* allows the search to differentiate between "inward" and "outward" nodes. There are other parameters used to control the walk toward different network exploration strategies, as shown in Table 2. More details about node2vec algorithm can be found in [35]. In this work, we applied node2vec technique on the full heterogeneous graph G that consists of the training part of known DTIs, DDs, and TTs. The use of node2vec model applied for this work is shown in Figure 2.

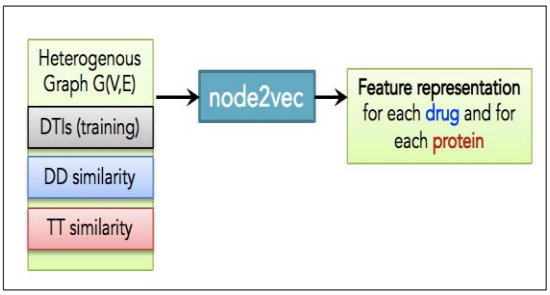

Figure 2. Node2vec model for DTIs network

Grid search are performed to utilize and obtain an optimized set of hyper-parameters. Parameters *p* and *q* were tested with values {0.25, 0.5, 1, 2, 4}, dimension *d* for values {16, 32, 64, 128}, where walk-length took range based on the size of the graph.

Table 2. Node2vec parameters description

Parameter	Description	Default value
dimension	Number of features	128
Walk-length	Length of walk per source	80
num-walk	Number of walks per source	10
p	Return hyperparameter	1
q	In-out hyperparameter	1
worker	Number of parallel workers	8

After applying node2vec on the graph *G* and obtaining a feature representation vector of each node as shown in Figure 2, a cosine similarity is calculated between each pair of drugs and between each pair of targets producing matrices M_d (DDs matrix of size *n x n* where *n* is the number of drugs), and M_t (target-target similarity matrix of size *m x m*, where *m* is the number of targets). The obtained similarity range is [-1 1], because of the existence of negative values of features for some nodes. For this reason, a min-max normalization is applied to both matrices to adjust the range to [0 1]. The benefit of these two steps is the following: First, by applying node2vec on the complete DTIs graph, we obtain feature representation that preserve local neighborhoods of each node (for

drugs or for targets) in a low dimensional space. This means, we capture a meaningful proximity information (e.g., relational and structural) between nodes in the graph. Second, by calculating cosine similarity between feature vectors of each drug pair (or target pair), we improve modeling the similarity between nodes that carry a lot of information. Another advantage of cosine similarity is that even if the two similar nodes are not close based on the Euclidean distance, their feature vectors could still have a small angle between them, indicating their high similarity. Formulating new graph with these new similarities is expected to result in a better representation of the graph instead of using chemical structure similarity of drugs or amino-acid sequence similarity of proteins.

4.3 Graph Mining Technique for Drug-target Path Scores

The drug-drug cosine similarity matrix and target-target cosine similarity matrix is used to construct a new heterogeneous graph $\underline{G}$ augmented with the training part of DTIs. Path Score is calculated for each simple path starting from the source node (i.e., drug) and ending with the target node (i.e., target protein) for each (drug, target) pair using *DASPfind* path score as introduced in a previous study [6] and based on the following formula:

$$score\left(d_i, t_j\right) = \sum_{p=1}^{n} \prod \left(P_{weights}\right) \quad (1)$$

where $P = \{p_1, p_2, \ldots, p_n\}$ is the set of paths that connect $drug_i$ to $target_j$. All paths between each drug and target are going through six path structures $C_h = \{C_1, C_2, C_3, C_4, C_5, C_6\}$, and the path scores are calculated for all six path structures. The path length is limited to 3 (i.e., path-length = 2 or 3). These path structures with length = 2 are C_1:D-D-T and C_2:D-T-T, and with length = 3 are C_3:D-D-D-T, C_4:D-T-T-T, C_5:D-D-T-T, and C_6:D-T-D-T, defined in previous works [4]. The set of paths between a pair of $drug_i$ and $target_j$ is denoted by R_{ijh}. We calculated the Path score by multiplying the edge weight score for each path structure, where w is the edge weight as follows:

$$score\left(d_i, t_j, h, q\right) = \prod_{\forall e_x \in P_q} \left(w_x\right) \quad (2)$$

The sum features of path score, as well as the max feature of the path score, are defined in equation 3 and 4 respectively.

$$sumScore\left(d_i, t_j, h\right) = \sum_{\forall Pq \in R_{ijh}} Score\left(d_i, t_j, h, q\right) \quad (3)$$

$$maxScore\left(d_i, t_j, h\right) = MAX_{\forall Pq \in R_{ijh}} \left(score\left(d_i, t_j, h, q\right)\right) \quad (4)$$

Thus, the 12 features are generated for each (drug, target) pair representing maximum path scores as well as sum of the path scores for each path structure described above. The feature vectors constructed using these 12 features are then fed into the DTI prediction model.

4.4 Sampling Imbalanced Datasets

The datasets we use are imbalanced, with the negative samples being much larger than the positive ones. To compensate for this, we applied different resampling techniques [37] and then chose the one that contributes to the best classification performance. In the processing step, we performed resampling on the training data only. It adjusts the data to be balanced. Random oversampling is applied to oversample the minority class (the positive known DTIs in our case) bringing them to the same number as the major class (unknown DTIs) as shown in Figure 3.

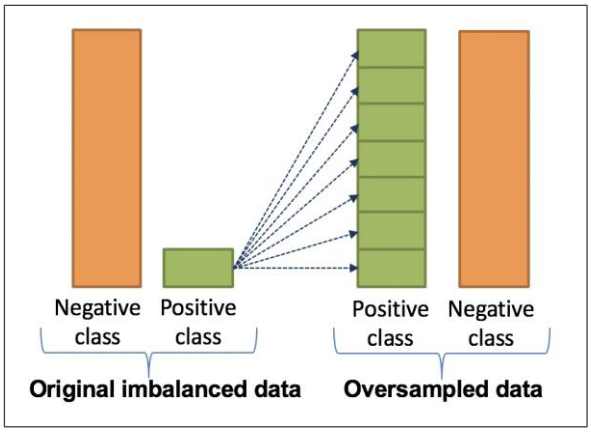

Figure 3. Oversampling the minority class of the training data.

4.5 DTIs Predictive Model

Supervised ML model is used to predict DTIs based on using random forest (RF) classifier. RF classifiers have been shown their efficiency in predictions applied to large datasets. By using an RF classifier, data do not need to be transformed or rescaled. Furthermore, the RF classifier is fast because of parallelism of information processing and it is more robust against the overfitting as well as outliers. We tested different RF parameters to obtain the best performance on the training data. Examples of these parameters are the number of trees, the maximum depth of the trees, the number of features to consider when looking for the best split, the minimum number of samples required to split an internal node, the function to measure the quality of a split, and others. The input to this classifier is the feature vector of several path scores that are explained previously for each drug and target pair (d_i, t_j), and the outputs are the predicted labels showing if there is an interaction or not for each (d_i, t_j) pair.

4.6 The DTiGEM Framework

Figure 4 shows all steps that are applied to obtain the final feature vector (indicated by FV in the Figure 4) for each pair of drug and target (d_i, t_j). These feature vectors are then fed into the RF classifier and output the predicted labels.

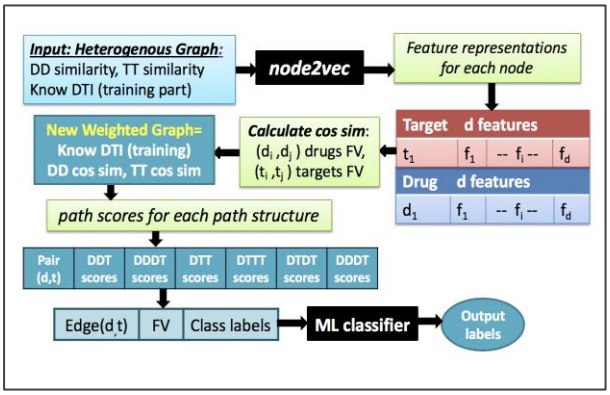

Figure 4. DTiGEM method framework.

5. EVALUATION

5.1 Evaluation Metrics

We calculated recall (also called: true positive rate or sensitivity) and precision (also called positive predictive value) as shown in Equations (5) and (6), respectively, to obtain the area under the precision-recall curve (AUPR) [38], [39]. TP, FN, FP are true-positive predictions, false-negative predictions, and false-positive predictions, respectively.

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

The precision-recall curve is determined based on different precision and recall values at different cut-offs, and then the area under this curve is calculated. AUPR is used to evaluate the performance of the prediction in the case of highly imbalanced data. It provides a proper assessment of how all the predicted scores of true interactions separate from predicted scores of true unknown interactions. Hence, we chose AUPR to be the significant assessment metric in our study and for comparison with the other methods.

5.2 Experiment Settings

In the validation step, we partitioned each dataset into 10 subsets to implement 10-folds cross-validation (CV). We used nine subsets to train the model, and one subset to test the prediction of this model. The process is repeated 10 times using a different subset as the test set. We restricted each fold of the data samples to include both positive and negative samples in training and testing partitions in a stratified fashion. We report performance results for each fold, and the average of the 10 reported results as the average performance.

6. RESULTS

6.1 Performance Comparison with the Existing Methods

For evaluation, we compared the performance of our method with the ML-based and graph-based state-of-art methods such as Multiple kernels learning algorithm (KRONRLS-MKL) [40] (indicated by (MKL) in Table 3 and in Figure 5), Neighborhood Regularized Logistic Matrix Factorization (NRLMF) [41], Dual Network Integrated Logistic Matrix Factorization (DNILMF) [42], Regularized Least Squares with Weighted Nearest Neighbor (RLS-WNN) [13], [43], Identification of Drug Target Interaction using Evolutionary and Structural Features with Boosting (iDTIs-ESBoost) [8], and Advance Local Drug-target Interaction (ALADIN) [44].

The KRONRLS-MKL [40] method used Kronecker regularized least-squares approaches to integrate information from different similarity types. This study demonstrated that utilizing different measures of similarity between drugs and targets improved the performance of DTIs prediction. Subsequently, the regularized least-squares approach was combined with weighted nearest neighbor to develop the RLS-WNN method [13], [43]] that uses bipartite local model (BLM) and compute network similarity in the form of gaussian interaction profile (GIP) kernels [45]. Adding the WNN preprocessing step reinforced the learning process. In the same year, [44] developed the ALADIN method, which extends the bipartite local model (BLM) [46], [47] work by integrating a hubness-aware regression technique coupled with enhanced drug-drug and target-target similarities. It also builds a projection-based ensemble. This method outperformed different versions of BLM. The NRLMF [41] differs from these methods as it models DTI probability using regularized logistic matrix factorization. This method produces two latent vectors, one representing the properties of the drugs and the other representing the properties of the targets. Subsequently, logistic matrix factorization was used to develop the DNILMF [42] method that applied a non-linear similarity fusion technique based on the similarity network fusion method (SNF). This method integrated different similarity measures and then used this final combined measure.

We also compared the performance of our method to iDTIs-ESBoost) [43], a model for DTI prediction that uses evolutionary and structural features and applies a novel technique of data balancing and boosting.

To have a fair comparison of our method with the previously introduced methods, all methods are tested using the same datasets and under the same conditions which are: Random split of drug and target pairs using 10-fold CV. Our method shows high performance and it outperforms other state-of-the-art methods. Table 3 shows the AUPR values for all methods used in comparison, the average AUPR score for each method over the four benchmark datasets, and the average ranking position for each method on each dataset (the lower ranking position, the better is the method). The best results in each row are in bold underlined font, while the second-best results are bold. DTiGEM achieves the best individual AUPR results for each dataset. Moreover, we demonstrate that based on the average AUPR score for all datasets, DTiGEM achieves the highest average AUPR value (0.831). This reduces the error by 22.4% relative to NRLMF, which is the second-best performing method. Moreover, overall, DTiGEM achieves the best ranking position (which is 1) over all datasets. Also, based on the prediction results on the IC and E datasets, the achieved better results could be attributed to the larger sizes of these datasets that help in prediction.

Figure 5 parts (a, b, c, and d) show the performance for our method and the six state-of-the-art methods applying to NR, GPCR, IC, and Enzyme datasets, respectively. Our method, DTiGEM, outperforms other methods on all datasets but has very close performance to iDTIs-EBoost method on NR dataset. However, iDTIs-EBoost shows low performance on other datasets and the high result on NR dataset may be misleading since the results of NR dataset are not stable (as Figure 5 shows) due to its excessively small size.

Table 3. The AUPR and ranking scores for all comparison methods

Dataset	AUPR of each method						
	MKL	DINLMF	NRLMF	RLS-WNN	iDTIs_EBoost	ALADIN	DTiGEM
NR	0.51	0.66	0.72	0.73	**0.79**	0.59	**0.795**
GPCR	0.67	0.70	0.707	**0.727**	0.50	0.68	**0.733**
IC	0.86	0.87	**0.88**	0.856	0.50	0.87	**0.892**
Enzyme	0.87	**0.89**	0.871	0.849	0.68	0.83	**0.905**
Average AUPR	0.728	0.78	**0.793**	0.791	0.493	0.743	**0.831**
Average Ranking	6	3	**2**	4	7	5	**1**

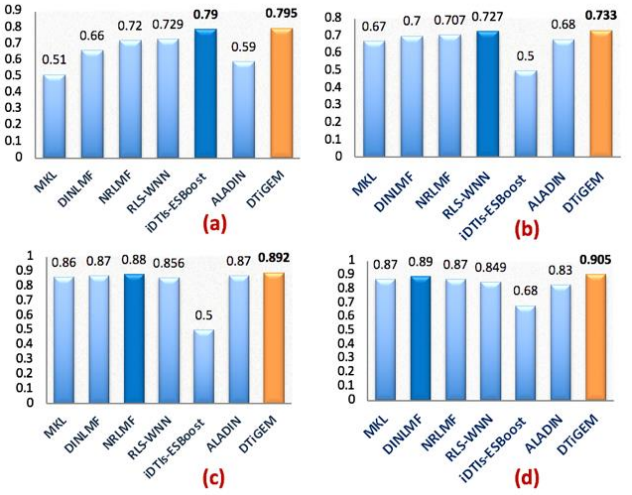

Figure 5. Comparison results of DTiGEM with other methods in terms of AUPR for four datasets.

a) NR, b) GPCR, c) IC, d) Enzyme

7. DISCUSSION AND CONCLUSION

We present a computational method for DTI prediction that integrates different techniques including DL for feature representation to get benefit from the network topology features, graph-mining to extract more features including path scores for different path structure, and ML for classification. The DL is used in node2vec graph embedding technique to generate each node feature vector in a low dimensional space that captures the topology similarity of the neighborhood community. Compared to other state-of-the-art methods used in the comparison, our method achieved the best results.

Our method can further be improved by adding more features, using different embeddings, and filtering the graph edges when computing the path scores using better thresholds. We intend using our method to address a new related problem; predicting the binding affinity between drugs and their target proteins as a regression problem. Furthermore, using DL methods could improve feature extraction and classification.

8. ACKNOWLEDGMENTS

This study is supported by King Abdullah University of Science and Technology (KAUST). M.T., S.A., and V.B.B. were supported by the KAUST Base Research Fund (BAS/1/1606-01-01) to V.B.B. and V.B.B. and M.E. were also supported by KAUST Office of Sponsored Research (OSR) Awards No. FCC/1/1976-24-01.

9. REFERENCES

[1] Y. Nilan, Sellahewa, D., Fernando, S., Gamage, L. and Meedeniya, D., "Analysis of conflicts between medication, adverse drug reactions and diseases.," in IEEE International Conference on Industrial and Information Systems (ICIIS), Peradeniya, Sri Lanka, 2017.

[2] Y. Nilan, Sellahewa, D., Fernando, S., Gamage, L. and Meedeniya, D., "A Clinical Decision Support System for Drug Conflict Identification," in Moratuwa Engineering Research Conference (MERCon), Moratuwa, Sri Lanka, 2018.

[3] S. Morgan, P. Grootendorst, J. Lexchin, C. Cunningham, and D. Greyson, "The cost of drug development: a systematic review," *Health Policy*, vol. 100, no. 1, pp. 4-17, 2011.

[4] R. S. Olayan, H. Ashoor, and V. B. Bajic, "DDR: efficient computational method to predict drug–target interactions using graph mining and machine learning approaches," *Bioinformatics*, vol. 34, no. 7, pp. 1164-1173, 2018.

[5] Y. Luo, X. Zhao, J. Zhou, J. Yang, Y. Zhang, W. Kuang, J. Peng, L. Chen, and J. Zeng, "A network integration approach for drug-target interaction prediction and computational drug repositioning from heterogeneous information," *Nat. Commun.*, vol. 8, no. 1, pp. 573, 2017.

[6] W. Ba-Alawi, O. Soufan, M. Essack, P. Kalnis, and V. B. Bajic, "DASPfind: new efficient method to predict drug-target interactions," *J. Cheminform.*, vol. 8, pp. 15, 2016.

[7] D. Emig, A. Ivliev, O. Pustovalova, L. Lancashire, S. Bureeva, Y. Nikolsky, and M. Bessarabova, "Drug target prediction and repositioning using an integrated network-based approach," *PLoS One*, vol. 8, no. 4, pp. e60618, 2013.

[8] F. Rayhan, S. Ahmed, S. Shatabda, D. M. Farid, Z. Mousavian, A. Dehzangi, and M. S. Rahman, "iDTI-ESBoost: Identification of Drug Target Interaction Using Evolutionary and Structural Features with Boosting," *Sci. Rep.*, vol. 7, no. 1, pp. 17731, 2017.

[9] S. Pathak, and X. Cai, "Ensemble learning algorithm for drug-target interaction prediction," *2017 IEEE 7th International Conference on Computational Advances in Bio and Medical Sciences (ICCABS)*, 2017.

[10] N. Zong, H. Kim, V. Ngo, and O. Harismendy, "Deep mining heterogeneous networks of biomedical linked data to predict novel drug-target associations," *Bioinformatics,* vol. 33, no. 15, pp. 2337-2344, 2017.

[11] M. Tsubaki, K. Tomii, and J. Sese, "Compound–protein interaction prediction with end-to-end learning of neural networks for graphs and sequences," *Bioinformatics,* vol. 35, no. 2, pp. 309-318, 2019.

[12] K. Y. Gao, A. Fokoue, H. Luo, A. Iyengar, S. Dey, and P. Zhang, "Interpretable Drug Target Prediction Using Deep Neural Representation," in IJCAI, 2018, pp. 3371-3377.

[13] A. Ezzat, M. Wu, X.-L. Li, and C.-K. Kwoh, "Computational prediction of drug-target interactions using chemogenomic approaches: an empirical survey," *Brief. Bioinform.,* 2018.

[14] W. Zhang, W. Lin, D. Zhang, S. Wang, J. Shi, and Y. Niu, "Recent Advances in the Machine Learning-Based Drug-Target Interaction Prediction," *Curr. Drug Metab.,* vol. 20, no. 3, pp. 194-202, 2019.

[15] Q. Zhao, H. Yu, M. Ji, Y. Zhao, and X. Chen, "Computational Model Development of Drug-Target Interaction Prediction: A Review," *Curr. Protein Pept. Sci.,* vol. 20, no. 6, pp. 492-494, 2019.

[16] H. Ding, I. Takigawa, H. Mamitsuka, and S. Zhu, "Similarity-based machine learning methods for predicting drug–target interactions: a brief review," *Brief. Bioinform.,* vol. 15, no. 5, pp. 734-747, 2014.

[17] K. Sachdev, and M. K. Gupta, "A comprehensive review of feature based methods for drug target interaction prediction," *J. Biomed. Inform.,* vol. 93, pp. 103159, 2019.

[18] M. Thafar, Raies, A.B., Albradei, S., Essack, M. and Bajic, V.B.,, "Comparison Study of Computational Prediction Tools for Drug-Target Binding Affinities," *Frontiers in Chemistry,* vol. 7, 2019.

[19] L. Kurgan, and C. Wang, "Survey of Similarity-based Prediction of Drug-protein Interactions," *Curr. Med. Chem.,* vol. 26, pp. 1, 2018.

[20] T. He, M. Heidemeyer, F. Ban, A. Cherkasov, and M. Ester, "SimBoost: a read-across approach for predicting drug-target binding affinities using gradient boosting machines," *J Cheminform,* vol. 9, no. 1, pp. 24, 2017.

[21] F. Rayhan, S. Ahmed, D. Md Farid, A. Dehzangi, and S. Shatabda, "CFSBoost: Cumulative feature subspace boosting for drug-target interaction prediction," *J. Theor. Biol.,* vol. 464, pp. 1-8, 2019.

[22] J. Keum, and H. Nam, "SELF-BLM: Prediction of drug-target interactions via self-training SVM," *PLoS One,* vol. 12, no. 2, pp. e0171839, 2017.

[23] X.-Y. Yan, R.-Z. Li, and L. Kang, "Prediction of Drug-Target Interaction with Graph Regularized Non-Negative Matrix Factorization," *Journal of Physics: Conference Series,* vol. 1237, pp. 032017, 2019.

[24] L. Wang, Z.-H. You, X. Chen, S.-X. Xia, F. Liu, X. Yan, and Y. Zhou, "Computational Methods for the Prediction of Drug-Target Interactions from Drug Fingerprints and Protein Sequences by Stacked Auto-Encoder Deep Neural Network," in Bioinformatics Research and Applications, 2017, pp. 46-58.

[25] K. Tian, M. Shao, Y. Wang, J. Guan, and S. Zhou, "Boosting compound-protein interaction prediction by deep learning," *Methods,* vol. 110, pp. 64-72, 2016.

[26] M. Alshahrani, M. A. Khan, O. Maddouri, A. R. Kinjo, N. Queralt-Rosinach, and R. Hoehndorf, "Neuro-symbolic representation learning on biological knowledge graphs," *Bioinformatics,* vol. 33, no. 17, pp. 2723-2730, 2017.

[27] G. Crichton, Y. Guo, S. Pyysalo, and A. Korhonen, "Neural networks for link prediction in realistic biomedical graphs: a multi-dimensional evaluation of graph embedding-based approaches," *BMC Bioinformatics,* vol. 19, no. 1, pp. 176, 2018.

[28] I. Abdelaziz, A. Fokoue, O. Hassanzadeh, P. Zhang, and M. Sadoghi, "Large-scale structural and textual similarity-based mining of knowledge graph to predict drug–drug interactions," *Journal of Web Semantics,* vol. 44, pp. 104-117, 2017.

[29] P. Goyal, and E. Ferrara, "Graph embedding techniques, applications, and performance: A survey," *Knowledge-Based Systems,* vol. 151, pp. 78-94, 2018.

[30] Y. Yamanishi, M. Araki, A. Gutteridge, W. Honda, and others, "Prediction of drug–target interaction networks from the integration of chemical and genomic spaces," 2008.

[31] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima, "KEGG: new perspectives on genomes, pathways, diseases and drugs," *Nucleic Acids Res.,* vol. 45, no. D1, pp. D353-D361, 2017.

[32] M. Hattori, N. Tanaka, M. Kanehisa, and S. Goto, "SIMCOMP/SUBCOMP: chemical structure search servers for network analyses," *Nucleic Acids Res.,* vol. 38, no. Web Server issue, pp. W652-6, 2010.

[33] S. B. Smith, W. Dampier, A. Tozeren, J. R. Brown, and M. Magid-Slav, "Identification of common biological pathways and drug targets across multiple respiratory viruses based on human host gene expression analysis," *PLoS One,* vol. 7, no. 3, pp. e33174, 2012.

[34] H. Cai, V. W. Zheng, and K. C. Chang, "A Comprehensive Survey of Graph Embedding: Problems, Techniques, and Applications," *IEEE Trans. Knowl. Data Eng.,* vol. 30, no. 9, pp. 1616-1637, 2018.

[35] A. Grover, and J. Leskovec, "node2vec: Scalable Feature Learning for Networks," *KDD,* vol. 2016, pp. 855-864, 2016.

[36] B. Perozzi, R. Al-Rfou, and S. Skiena, "DeepWalk: Online Learning of Social Representations," in Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, 2014, pp. 701-710.

[37] G. Lemaître, F. Nogueira, and C. K. Aridas, "Imbalanced-learn: A python toolbox to tackle the curse of imbalanced datasets in machine learning," *J. Mach. Learn. Res.,* vol. 18, no. 1, pp. 559-563, 2017.

[38] J. Davis, and M. Goadrich, "The Relationship Between Precision-Recall and ROC Curves," in Proceedings of the 23rd International Conference on Machine Learning, 2006, pp. 233-240.

[39] J. Grau, I. Grosse, and J. Keilwagen, "PRROC: computing and visualizing precision-recall and receiver operating

characteristic curves in R," *Bioinformatics,* vol. 31, no. 15, pp. 2595-2597, 2015.

[40] A. C. A. Nascimento, R. B. C. Prudêncio, and I. G. Costa, "A multiple kernel learning algorithm for drug-target interaction prediction," *BMC Bioinformatics,* vol. 17, pp. 46, 2016.

[41] Y. Liu, M. Wu, C. Miao, P. Zhao, and X.-L. Li, "Neighborhood Regularized Logistic Matrix Factorization for Drug-Target Interaction Prediction," *PLoS Comput. Biol.,* vol. 12, no. 2, pp. e1004760, 2016.

[42] M. Hao, S. H. Bryant, and Y. Wang, "Predicting drug-target interactions by dual-network integrated logistic matrix factorization," *Sci. Rep.,* vol. 7, pp. 40376, 2017.

[43] A. Ezzat, P. Zhao, M. Wu, X.-L. Li, and C.-K. Kwoh, "Drug-Target Interaction Prediction with Graph Regularized Matrix Factorization," *IEEE/ACM Trans. Comput. Biol. Bioinform.,* vol. 14, no. 3, pp. 646-656, 2017.

[44] K. Buza, and L. Peska, "ALADIN: a new approach for drug–target interaction prediction," *Joint European Conference on Machine Learning and,* 2017.

[45] T. van Laarhoven, S. B. Nabuurs, and E. Marchiori, "Gaussian interaction profile kernels for predicting drug–target interaction," *Bioinformatics,* vol. 27, no. 21, pp. 3036-3043, 2011.

[46] K. Bleakley, and Y. Yamanishi, "Supervised prediction of drug–target interactions using bipartite local models," *Bioinformatics,* vol. 25, no. 18, pp. 2397-2403, 2009.

[47] K. Buza, and L. Peška, "Drug–target interaction prediction with Bipartite Local Models and hubness-aware regression," *Neurocomputing,* vol. 260, pp. 284-293, 2017.