

Drug-Target Interaction Prediction in Drug Repositioning Based on Deep Semi-Supervised Learning

Meriem Bahi, Mohamed Batouche

► **To cite this version:**

Meriem Bahi, Mohamed Batouche. Drug-Target Interaction Prediction in Drug Repositioning Based on Deep Semi-Supervised Learning. 6th IFIP International Conference on Computational Intelligence and Its Applications (CIIA), May 2018, Oran, Algeria. pp.302-313, 10.1007/978-3-319-89743-1_27. hal-01913886

HAL Id: hal-01913886

<https://hal.inria.fr/hal-01913886>

Submitted on 6 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Drug-Target Interaction Prediction in Drug Repositioning based on Deep Semi-Supervised Learning

Meriem Bahi and Mohamed Batouche

Computer Science Department, Faculty of NTIC
University Constantine 2 - Abdelhamid Mehri
Biotechnology Research Center (CRBt) & CERIST
Constantine, Algeria
{meriem.bahi,mohamed.batouche}@univ-constantine2.dz

Abstract. Drug repositioning or repurposing refers to identifying new indications for existing drugs and clinical candidates. Predicting new drug-target interactions (DTIs) is of great challenge in drug repositioning. This tricky task depends on two aspects. The volume of data available on drugs and proteins is growing in an exponential manner. The known interacting drug-target pairs are very scarce. Besides, it is hard to select the negative samples because there are not experimentally verified negative drug-target interactions. Many computational methods have been proposed to address these problems. However, they suffer from the high rate of false positive predictions leading to biologically interpretable errors. To cope with these limitations, we propose in this paper an efficient computational method based on deep semi-supervised learning (DeepSS-DTIs) which is a combination of a stacked autoencoders and a supervised deep neural network. The objective of this approach is to predict potential drug targets and new drug indications by using a large scale chemogenomics data while improving the performance of DTIs prediction. Experimental results have shown that our approach outperforms state-of-the-art techniques. Indeed, the proposed method has been compared to five machine learning algorithms applied all on the same reference datasets of DrugBank. The overall accuracy performance is more than 98%. In addition, the DeepSS-DTIs has been able to predict new DTIs between approved drugs and targets. The highly ranked candidate DTIs obtained from DeepSS-DTIs are also verified in the DrugBank database and in literature.

Keywords: Drug Repositioning, Drug-Target Interactions, Deep Learning, Semi-Supervised Learning, Stacked Autoencoders, Deep Neural Network

1 Introduction

Over the past decades, de novo drug discovery has become increasingly difficult and risky. This process has grown to be time consuming and expensive. It can

take about 17 years and costs at least one billion dollars. In 2015, Pharmaceutical Research and Manufacturers of America (PhRMA) members had invested more than half a trillion dollars in research and development of a new drug [1], while the number of newly approved drugs and clinical compounds known as New Molecular Entities (NMEs) is steadily declining annually. Therefore, it is beneficial to develop strategies to reduce this time frame, decrease costs and improve success rates [2]. Discovering potential uses for existing drugs, also known as drug repositioning [1], is one strategy which has attracted increasing interests from both the pharmaceutical industry and the research community.

Discovering new indications for existing drugs can be attained through identification of new interactions between drugs and target proteins. The in silico prediction of drug target interaction (DTI) is a challenging task in drug repositioning which lies on two main aspects. First, the volume of chemogenomic data available on drugs and proteins is growing in an exponential manner. Second, the known drug-target interactions pairs are rare. Besides, it is hard to select the negative samples because there are not experimentally verified negative drug-target interactions. To date, a variety of computational methods have been proposed to solve these problems and to accurately predict new interactions between known drugs and targets. They fall into two categories i) Network-based and ii) learning based. However, they suffer from the high rate of false positive predictions leading to biologically interpretable errors.

To overcome these limitations, we propose in this work a novel computational method, namely DeepSS-DTIs, based on deep semi-supervised learning to accurately predict potential new drug-target interactions using large-scale chemical-protein data. This method nicely combines the advantages of the two different methods of feature-based and semi-supervised learning. The rest of the paper is organized as follows. Drug repositioning field is described in section 2. The different computational methods using for drug repurposing are briefly reviewed in Section 3. Section 4 is dedicated to the description of the proposed approach based on the hybrid deep learning architecture. In section 5, the performance of the proposed approach is assessed. In section 6, the list of new predicted interactions is presented. Finally, conclusions and future work are drawn.

2 Drug Repositioning

Drug repositioning or repurposing, rescue or reprofiling (the terms are sometimes used interchangeably) refers to studying drugs that are already approved to treat one disease or condition to see if they are effective for treating other diseases [3]. Finding a new indication of existing drugs is an accelerated route for drug discovery. The process of drug repurposing is generally approved in shorter time frames (3 years). It can reduce about 70% of development cost and decrease the drug safety risk. Because the information about safety, efficacy, and toxicity of an existing drug have been extensively studied and therefore data have already been accumulated toward gaining approval by the U.S. Food and Drug Administration (FDA) for a specific indication.

Most drugs are small compounds that target and interact with therapeutic proteins implicated in a disease of interest to induce perturbation in the protein network [4]. However, approximately 90% of drugs interact not only with the therapeutic target proteins but also with additional proteins resulting in unexpected side effects. The drug side effect may be beneficial for identifying new therapeutic indications [5]. For example, thalidomide is a drug that was developed as a sleeping pill, but it was also found to be useful for easing morning sickness in pregnant women. Unfortunately, it damaged the development of unborn babies. The drug led to the arms or legs of the babies being very short or incompletely formed. More than 10,000 babies were affected around the world. As a result of this disaster, thalidomide was banned [6]. But, thalidomide was redeveloped and repurposing and now it is used as a treatment for leprosy and bone cancer. Many drugs have enormous potential for new therapeutic indications in terms of polypharmacology.

3 In silico Methods for Drug Reprofilng

Identifying drug-target interactions to find new uses of existing and abandoned drugs is a crucial prerequisite and is a major challenge in drug repositioning. Currently, experimental methods of identifying new interactions between drugs and targets are cumbersome. In silico approaches can provide a promising and efficient tool to alleviate this problem, and thus significantly reduce both experimental time and cost of identifying potential DTI. Therefore, so far, there is a strong incentive to seek and develop computational methods to better predict new drug-target interactions. Traditional in silico approaches can be categorized into the ligand-based approach, structure-based approach and text mining approach [7]. The ligand-based approach is based on the concept that similar ligands (or molecules) tend to have similar biological properties. One of these methods is Quantitative Structure-Activity Relationship (QSAR) that predict the bioactivity of a ligand on a target. Given a certain amount of targets, each target builds a predictive model using its known active ligands. Then these built models are used to screen all the drugs to predict the DTIs between drugs and targets. Unfortunately, the problem with this category of a method is that many target proteins have little or no ligand information available. Structure-based methods or molecular docking represent the second category of approaches for drug repositioning. They have been successfully used for predicting drug-target interactions [8]. These methods are based on the same principle of similarity observed for ligands. Proteins with similar structures are likely to have similar functions and to recognize similar ligands. They use the crystallographic structure of target to screen the small molecules and to identify secondary targets of an approved drug. The limitation of these methods is that they require the three-dimensional (3D) structure of a target which is a problem because not all proteins have their 3D structures available [3]. Indeed, for most membrane proteins, like GPCRs, their 3D structure information is still unavailable, as determining their structures is a challenging task. Another approach is the text mining techniques

which are based on keyword searching in the huge number of literature [9], but it suffers from the problem of redundancy in the compound/protein names in the literature.

To overcome challenges of traditional methods, chemogenomic approaches have recently attracted increasing attention in drug discovery and repositioning to find new Drug-Protein interactions on a large scale. They simultaneously utilize both the drug and target features (e.g., drug-induced gene expression, chemical structures, side effects, target protein sequences, and biological pathways) and also disease information (e.g., symptomatic state and phenotype) to perform better predictions [10]. Chemogenomic methods can be divided broadly into network-based techniques and learning-based approach. Network-based methods aim at organizing the relationships among drugs and targets in the form of networks to infer unknown drug-target interactions. The drug-target network can be depicted as a connected graph, where each node represents either a drug or a target and the known interactions between drugs and targets corresponding to the lines that link the nodes. These methods have been widely used for computational drug repositioning. For example, Yamanishi et al., [10] integrated the relationship between pharmacological, chemical, and topology spaces of drug-target interaction networks to predict new associations between drugs and targets. Also, Chen et al., [11] developed an effective model of a heterogeneous network, named NRWRH, to predict potential drug-target interactions on a large scale. Liu et al., [12] have developed a network-based inference model for the prediction of potential DTI. A common limitation of these network-based methods is that they mainly look for novel targets which are close to known targets in the network. Learning-based techniques have been extensively used to cope with the drawbacks of the previous methods, under the assumption that similar drugs are likely to interact with similar proteins. The learning-based methods can be divided into supervised and semi-supervised. The supervised-learning approach has been used in two ways including the similarity based-methods and feature-based methods [13]. Similarity-based methods have been developed to predict potential drug-target interactions through the constructed similarity matrices of drug and protein. Nascimento et al., [14] incorporated multiple heterogeneous information sources using multiple kernel learning method for the identification of new DTIs. Furthermore, a key disadvantage of the similarity-based methods is that they cannot be used on large-scale datasets due to the significant computational complexity of measuring similarity matrices. In contrast, the feature-based methods are regarded as more advantageous strategies where drugs and targets are represented by sets of descriptors (i.e., feature vectors). These methods provide meaningful solutions for discovering interest drug-target interactions by identifying features that are highly more discriminative [13]. They can easily be applied to such a dataset and their computational complexity is moderate. The commonly used learning method is to build a supervised binary classification model where the positive class consists of interacting drug-target pairs and the negative class consists of non-interacting drug-target pairs. It takes drug target pairs (DTPs) as input, and the output is whether there is an interaction between

the drug target pair (DTP). However, these models exhibit complicated issues. Since the known DTIs are rare and negative DTIs are difficult or even impossible to achieve because experimentally validated negative samples are not reported and unavailable [15], these methods consider the unknown drug-target interactions as negative samples. This would largely influence the prediction accuracy. Accordingly, the semi-supervised learning approach has been applied to address this problem of imbalanced datasets in drug-target interaction prediction by using the small number of labeled data in conjunction with the numerous unlabeled data. There are only a few studies that have published on semi-supervised learning. That is why researchers are investigating more efforts to develop semi-supervised methods to improve the prediction performance of drug-target interactions. With the increasing of experimental data and increasing complexity of the machine learning algorithms that perform poorly, deep learning methods have been widely applied in many fields of bioinformatics, biology and chemistry [16]. Deep learning methods attract a lot of attention for its better performance and ability to learn representations of data with multiple levels of abstraction. In the drug-repositioning, Wen et al., [17] developed a deep learning method based on deep belief network algorithm to predict new DTIs. They found that deep learning outperforms other state-of-the-art machine learning methods. Wang et al., [18] proposed a stacked autoencoders incorporated with the random forest as the final classifier for predicting interactions between drugs and targets.

4 Materials and Methods

4.1 Data Preparation

The drugs and targets data used in this study were collected from a recent publication [15] which are extracted from DrugBank database. The latter is a unique bioinformatics and cheminformatics resource that combines detailed drug data with comprehensive drug target information. The interactions of drugs and targets were downloaded from drug Target Identifiers category of Protein Identifiers in DrugBank. The Drug target space (DTS) is defined as all possible drug-target pairs (DTPs). In total, there are 5877 drugs and 3348 targets. DTS has 19676196 (that is, $5877 * 3348$) DTPs. Among them, 12674 pairs are positive DTIs (Drug-Target interactions marked as Yes or +1) which have known interaction, and the others are not known (unlabeled data). Because the number of no interaction pairs is much more than the number of interaction pairs, the negative dataset can be randomly selected from the DTS. In this work, we randomly select 12674 drug-target pairs from the DTS as a negative dataset (marked as -1). Therefore, the whole labeled dataset contains 25348 samples, as depicted in Fig. 1.

4.2 Drug-Target Representation

Drugs and targets are represented by sets of descriptors (i.e. feature vectors). These features are classified into two categories: chemical structure of drugs (or

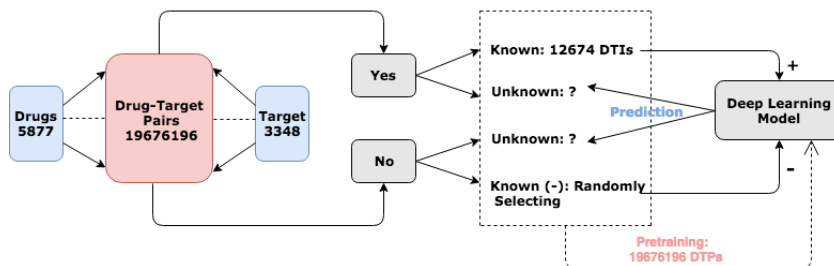


Fig. 1. The Flowchart of Data Flow.

molecular fingerprints) and protein sequence (molecular descriptors). We collected the best features from the recent publication of Ezzat et al., [15]. The latter used the RcpI package to calculate drug features. Examples of drug features include constitutional, topological and geometrical descriptors among other molecular properties. The target features were obtained using the PROFEAT web server. The features that have been used to represent targets are descriptors related to amino acid composition, dipeptide composition, autocorrelation, composition, transition, and distribution, quasi-sequence-order, amphiphilic pseudo-amino acid composition and total amino acid properties. Thus, we obtained 193 and 1290 features for drugs and targets, respectively.

After collecting the features, each drug-target pair is represented by feature vectors that are formed by concatenating the feature vectors of the corresponding drug and target involved. For example, a drug-target pair is represented by the feature vector:

$$[d_1, d_2, d_3, \dots, d_{193}, t_1, t_2, t_3, \dots, t_{1290}]$$

where $[d_1, d_2, \dots, d_{193}]$ is the feature vector corresponding to drug d , and $[t_1, t_2, \dots, t_{1290}]$ is the feature vector corresponding to target t . We refer to these drug-target pairs as instances, and we associate a label (+1 or -1) to each sample.

4.3 DeepSS-DTIs: The proposed Method for Drug Repositioning

The number of known interactions between drugs and targets is limited (less than 0.2% among the DTS) and no negative sample of drug-target interaction is verified experimentally [19]. Thus, it is hard to use only the small part of DTIs to represent the whole sample space and applicability of the model may be biased. In this case, it is necessary to use a semi-supervised learning approach for addressing this problem in drug-target interaction prediction with the small number of labeled data and numerous unlabeled data. In addition, with the sheer size of drug-target pairs available (over twenty million DTPs), it is imperative to use the deep learning method.

The unsupervised pre-training followed by supervised fine-tuning is a way of applying with success the semi-supervised deep learning method. Pre-training is essentially obsolete due to the success of semi-supervised learning which accomplishes the same goals more elegantly by optimizing unsupervised and supervised objectives simultaneously [20]. Unsupervised pre-training is not only still relevant for tasks for which we have small labeled datasets and large unlabeled datasets, but it can also exhibit much better performance in data representation and classification. We can summarize the main advantages of the unsupervised pre-training process as follows:

- A better initialization of the weights in the deep neural network instead of randomly initialized weights which may lead to better convergence and better performing classifiers.
- It acts as some special kind of regularization process which yields a better generalization power.

In this study, the training procedure of our deep learning model DeepSS-DTIs can be divided into two consecutive processes: the layer-wise unsupervised pre-training process using stacked autoencoders, and the supervised fine-tuning process of the deep neural network.

Stacked Autoencoders

Stacked Autoencoders (SAE) is one of popular deep learning model, built with multiple layers of autoencoders, in which the output of each layer is connected to the input of the next layer [21], as depicted in Fig. 2.

An autoencoder (AE) can be considered as a special neural network with one hidden layer. It tries to reconstruct the same features at the output layer using its hidden activations. The AE takes the input and puts it through an encoding function to get the encoding of the input, and then it decodes the encodings through a decoding function to recover (an approximation of) the original input [22]. More formally, let $x \in R^d$ be the input:

$$h = f_e(x) = s_e(W_e x + b_e) \quad (1)$$

$$x_r = f_d(x) = s_d(W_d h + b_d) \quad (2)$$

where $f_e : R^d \rightarrow R^h$ and $f_d : R^h \rightarrow R^d$ are encoding and decoding functions respectively, W_e and W_d are the weights of the encoding and decoding layers, and b_e and b_d are the biases for the two layers. s_e and s_d are elementwise non-linear functions in general, and common choices are sigmoidal functions like tanh or logistic [21].

In general, N-layer stacked autoencoders with parameters $P = \{P^i \mid i \in \{1, 2, \dots, N\}\}$, where $P^i = \{W_e^i, W_d^i, b_e^i, b_d^i\}$ can be formulated as follows:

$$h^i = f_e^i(h^{i-1}) = s_e^i(W_e^i h^{i-1} + b_e^i) \quad (3)$$

$$h_r^i = f_d^i(h_r^{i+1}) = s_d^i(W_d^i h_r^{i+1} + b_d^i) \quad (4)$$

$$h^0 = x \quad (5)$$

SAE plays a fundamental role in unsupervised learning. It is based on a greedy layer-wise training [23]. It can better learn the features of the input information and reduce the original data dimension [24] where the raw data are transformed from layer to layer up to the top layer.

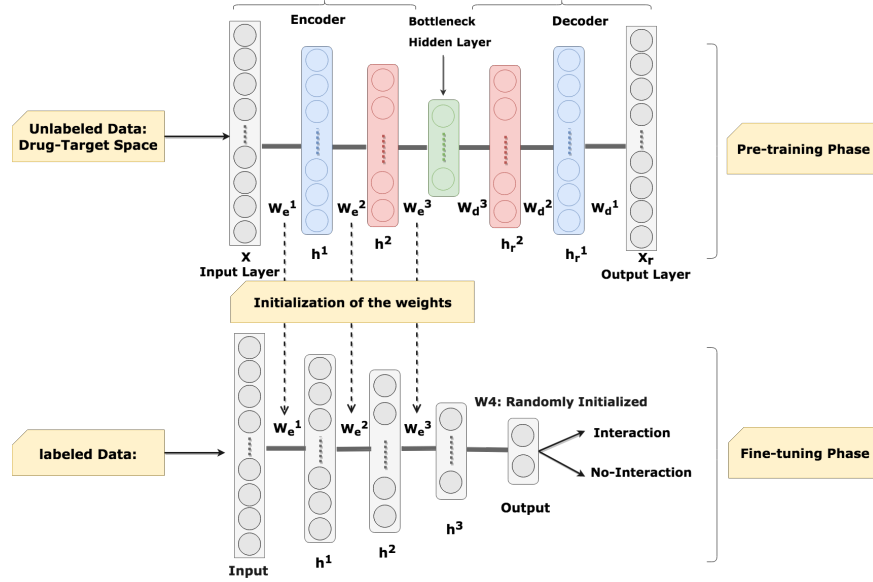


Fig. 2. The proposed Drug Semi-Supervised Model. Top: The Stacked Autoencoders after training. Down: The pre-trained Deep Neural Network initialized with Stacked Autoencoders' weights. For the simplicity, all biases are excluded from the figure.

The layer-wise unsupervised pre-training of stacked autoencoders process is as follows:

1. Train the bottom most autoencoder using the unlabeled data.
2. After training, we remove the decoder layer, we construct a new autoencoder by taking the latent representation of the previous auto-encoder as input.
3. Train the new autoencoder. Note the parameters (weights and bias) of the encoder from the previously trained autoencoder are fixed when training the newly constructed autoencoder.
4. Repeat step 2 and 3 until all encoding layers are trained. The activation function is usually the sigmoid function or tanh function.

The supervised fine-tuning process is as follows :

1. After training, we use the weights of the unsupervised stacked autoencoders model to initialize the weights of the supervised deep neural networks model (DNN).
2. Fine-tune all the parameters of all deep neural networks with stochastic gradient descent using back-propagation. As shown in Fig. 2.
3. Initialize randomly the output layer parameters of deep neural networks.

5 Experimental Results

5.1 Measurement of prediction quality

To assess the performance of the proposed method based on deep semi-supervised learning for prediction drug-target interactions in drug repositioning, we used four measures namely the area under the receiver operator characteristic curve (AUC), the accuracy rate (AR), the sensitivity (SE) and the specificity (SP) with 5-fold cross-validation. The statistical measures are defined as follows:

Accuracy Rate (AR): It measures the percentage of samples that are correctly classified.

$$(AR) = \frac{TP+TN}{TP+TN+FP+FN} * 100$$

Sensitivity (SE): It measures the accuracy on positive samples.

$$(SE) = \frac{TP}{TP+FN} * 100$$

Specificity (SP): It measures the accuracy on negative samples.

$$(SP) = \frac{TN}{TN+FP} * 100$$

With TP , FP , TN and FN the numbers of true-positive, false-positive, true-negative and false-negative predictions, respectively. In a two-class prediction problem, the outcomes are labeled either as positive (p) or negative (n). If the prediction and actual value are all p , it is called a TP ; if the prediction value is p while the actual value is n , it is called a FP . Conversely, if the prediction and actual value are all (n), it is called a TN ; if the prediction value is n while the actual value is p , it is called a FN .

5.2 Cross-validation results

We compared our approach to five state-of-the-art machine learning algorithms reported in the literature [15] which are Random Forest, SVM, Decision Tree, Nearest Neighbor and ensemble learning. The obtained results are summarized

in table 1 and show that our method outperforms other methods in all measurements.

As shown in table 1, the results obtained by our method DeepSS-DTIs using H2O platform are more than 0.98 (98%) in almost all measurements. The *AUC*, accuracy, sensitivity, and specificity of test set are 0.9980, 0.9853, 1 (100%) and 1 (100%) respectively. Because the number of positive DTIs is much fewer than that of negative in Drug-Target space and the purpose of the model is to predict the true positive DTI, the sensitivity (*SE*) is a more important evaluation metric among the four evaluation metrics. The obtained results by our approach are clearly better than the ones reported in [15]. Our method achieved an *AUC* of 0.998, which is 9.8% higher than the ensemble classifier learning (or class imbalance method) with an *AUC* of 0.900. This method is well suited for the prediction of new drug-target interactions. The other methods such as Decision Trees, SVM, Nearest Neighbor and Random Forest, yield to heterogeneous results. This supports our claim that using both a semi-supervised and deep learning technique is important for improving the prediction performance. Overall, the cross-validation between the results of our approach (DeepSS-DTIs) and those of five different machine learning algorithms applied all on the same datasets, clearly demonstrates that the DeepSS-DTIs method gained the best performance in *AUC*, *AR*, *SP* and *SE*. This indicates that the built DeepSS-DTIs model is reliable and can be further applied for novel DTIs prediction.

Table 1. Performance assessment of the proposed method

Prediction model	AUC (%)	AR (%)	SE (%)	SP (%)
Our method (DeepSS-DTIs)	99.80	98.53	100	100
SVM	80.40	81.08	76.31	85.85
Nearest Neighbor	81.40	82.26	78.50	85.37
Ensemble classifier	90.00	89.65	88.80	88.54
Random Forest	85.5	81.84	79.58	84.11
Decision Tree	76.00	75.05	51.67	66.80

6 Predicting new drug-target interactions

After confirming the performance of our method (DeepSS-DTIs) in comparison against other state-of-the-art methods, we tested the ability of our built model to correctly predict interactions on the remaining of the drug-target space (DTS) and ranked them by their probability. The table 2 shows the list of the top 10 probability predicted DTIs by DeepSS-DTIs with H2O platform.

In order to evaluate the reliability of new predicted interactions, we consulted the literature and the DrugBank database with the predicted relationships be-

tween drugs and targets. We found that some of the drug predicted by our method are validated by relevant literature and show potentiality for further study.

Table 2. Top10 probability scoring DTIs predicted by our model.

Rank	Drug-ID	Target-ID	Description
1	DB00246	P08909	Ziprasidone,5-hydroxytryptamine receptor
2	DB00126	P40238	Vitamin C, Thrombopoietin receptor
3	DB00363	P08909	Clozapine, 5-hydroxytryptamine receptor
4	DB00463	Q16478	Metharbital, Glutamate receptor ionotropic
5	DB00898	P15587	Ethanol, Xylose isomerase
6	DB00312	Q13002	Pentobarbital, kainate 2
7	DB01045	Q16850	Rifampicin, Lanosterol 14-alpha demethylase
8	DB00933	P21728	Mesoridazine, D(1A) dopamine receptor
9	DB00114	P19235	Pyridoxal Phosphate, Erythropoietin receptor
10	DB00128	Q13332	L-Aspartic Acid, Receptor-type tyrosine-protein phosphatase S

7 Conclusion and future work

Identifying drug-target interactions (DTIs) is a key area in drug repositioning. In this paper, we have presented an effective method for predicting both new drugs and detecting new targets for drug repositioning based on deep semi-supervised learning dealing with unbalanced data using a small number of known interactions in conjunction with the many unknown interactions. The cross-validation experiments demonstrated that the proposed approach (DeepSS-DTIs) outperforms the previous methods for drug-target interaction prediction. As future work, we expect to scale up the proposed approach by using sparkling water (Spark+H2O) to handle big data and improve performances.

References

1. Barratt, M.J., Frail, D.E.: Drug repositioning: Bringing new life to shelved assets and existing drugs. John Wiley & Sons (2012)
2. Zhang, R.: An ensemble learning approach for improving drug–target interactions prediction. In: Proceedings of the 4th International Conference on Computer Engineering and Networks, Springer (2015) 433–442
3. Li, J., Zheng, S., Chen, B., Butte, A.J., Swamidass, S.J., Lu, Z.: A survey of current trends in computational drug repositioning. *Briefings in bioinformatics* **17**(1) (2015) 2–12
4. Li, Y.Y., Jones, S.J.: Drug repositioning for personalized medicine. *Genome medicine* **4**(3) (2012) 27
5. Napolitano, F., Zhao, Y., Moreira, V.M., Tagliaferri, R., Kere, J., DAmato, M., Greco, D.: Drug repositioning: a machine-learning approach through data integration. *Journal of cheminformatics* **5**(1) (2013) 30

6. Shim, J.S., Liu, J.O.: Recent advances in drug repositioning for the discovery of new anticancer drugs. *International journal of biological sciences* **10**(7) (2014) 654
7. March-Vila, E., Pinzi, L., Sturm, N., Tinivella, A., Engkvist, O., Chen, H., Rastelli, G.: On the integration of in silico drug design methods for drug repurposing. *Frontiers in pharmacology* **8** (2017)
8. Phatak, S.S., Zhang, S.: A novel multi-modal drug repurposing approach for identification of potent ack1 inhibitors. In: Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing, NIH Public Access (2013) 29
9. Peng, L., Liao, B., Zhu, W., Li, K.: Predicting drug-target interactions with multi-information fusion. *IEEE journal of biomedical and health informatics* (2015)
10. Yamanishi, Y.: Chemogenomic approaches to infer drug–target interaction networks. *Data Mining for Systems Biology: Methods and Protocols* (2013) 97–113
11. Chen, X., Liu, M.X., Yan, G.Y.: Drug-target interaction prediction by random walk on the heterogeneous network. *Molecular BioSystems* **8**(7) (2012) 1970–1978
12. Cheng, F., Liu, C., Jiang, J., Lu, W., Li, W., Liu, G., Zhou, W., Huang, J., Tang, Y.: Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS computational biology* **8**(5) (2012) e1002503
13. Mousavian, Z., Masoudi-Nejad, A.: Drug–target interaction prediction via chemogenomic space: learning-based methods. *Expert opinion on drug metabolism & toxicology* **10**(9) (2014) 1273–1287
14. Nascimento, A.C., Prudêncio, R.B., Costa, I.G.: A multiple kernel learning algorithm for drug-target interaction prediction. *BMC bioinformatics* **17**(1) (2016) 46
15. Ezzat, A., Wu, M., Li, X.L., Kwoh, C.K.: Drug-target interaction prediction via class imbalance-aware ensemble learning. *BMC bioinformatics* **17**(19) (2016) 509
16. Chen, H., Zhang, Z.: A semi-supervised method for drug-target interaction prediction with consistency in networks. *PLoS One* **8**(5) (2013) e62975
17. Wen, M., Zhang, Z., Niu, S., Sha, H., Yang, R., Yun, Y., Lu, H.: Deep-learning-based drug–target interaction prediction. *Journal of Proteome Research* **16**(4) (2017) 1401–1409
18. Wang, L., You, Z.H., Chen, X., Xia, S.X., Liu, F., Yan, X., Zhou, Y., Song, K.J.: A computational-based method for predicting drug–target interactions by using stacked autoencoder deep neural network. *Journal of Computational Biology* (2017)
19. Zhao, J., Cao, Z.: A label extended semi-supervised learning method for drug-target interaction prediction. *studies* **13** (2015) 21
20. Erhan, D., Bengio, Y., Courville, A., Manzagol, P.A., Vincent, P., Bengio, S.: Why does unsupervised pre-training help deep learning? *Journal of Machine Learning Research* **11**(Feb) (2010) 625–660
21. Zhou, Y., Arpit, D., Nwogu, I., Govindaraju, V.: Is joint training better for deep auto-encoders? *arXiv preprint arXiv:1405.1380* (2014)
22. Oliveira, T.P., Barbar, J.S., Soares, A.S.: Multilayer perceptron and stacked autoencoder for internet traffic prediction. In: *IFIP International Conference on Network and Parallel Computing*, Springer (2014) 61–71
23. Candel, A., Parmar, V., LeDell, E., Arora, A.: Deep learning with h2o. *H2O. ai Inc* (2016)
24. Zabalza, J., Ren, J., Zheng, J., Zhao, H., Qing, C., Yang, Z., Du, P., Marshall, S.: Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging. *Neurocomputing* **185** (2016) 1–10