

Drug–Target Interaction Prediction with Bipartite Local Models and Hubness-aware Regression

Krisztian Buza^{a,b}, Ladislav Peška^c

^a Brain Imaging Center, Research Center for Natural Sciences
Hungarian Academy of Sciences, Budapest, Hungary

^b Knowledge Discovery and Machine Learning,
Rheinische Friedrich-Wilhelms-Universität Bonn, Germany

e-mail: buza@biointelligence.hu

^c Department of Software Engineering,
Faculty of Mathematics and Physics
Charles University, Prague, Czech Republic
e-mail: peska@ksi.mff.cuni.cz

May 25, 2017

Abstract

Computational prediction of drug–target interactions is an essential task with various applications in the pharmaceutical industry, such as adverse effect prediction or drug repositioning. Recently, expert systems based on machine learning have been applied to drug–target interaction prediction. Although hubness-aware machine learning techniques are among the most promising approaches, their potential to enhance drug–target interaction prediction methods has not been exploited yet. In this paper, we extend the Bipartite Local Model (BLM), one of the most prominent interaction prediction methods. In particular, we use BLM with a hubness-aware regression technique, $ECkNN$. We represent drugs and targets in the similarity space with rich set of features (i.e., chemical, genomic and interaction features), and build a projection-based ensemble of BLMs. In order to assist reproducibility of our work as well as comparison to published results, we perform experiments on widely used publicly available drug–target interaction datasets. The results show that our approach outperforms state-of-the-art drug–target prediction techniques. Additionally, we demonstrate the feasibility of predictions from the point of view of applications.

1 Introduction

Due to the large number of drug compounds and pharmacological targets, many of the interactions between these entities are unknown. More complete knowledge about drug–target interactions will not only contribute to better understanding the pharmacology of drugs, but it is also relevant for the prediction of adverse effects and drug repositioning, i.e., use of an existing medicine to treat a disease that has not been treated with that drug yet. The relevance of the later application is also underlined by the fact that only a few dozens of new drugs are approved by FDA each year. Moreover, the average costs related to discovery of a new drug are approximately \$1.8 billion, and the process takes more than 10 years [18].

In addition, the incomplete knowledge about the interactions between drugs and pharmaceutical targets in case of drugs affecting the central nervous system (CNS) further emphasizes the need for computational prediction approaches: while CNS plays an essential role, the costs associated with disorders affecting CNS are enormous: solely in Europe, the total annual costs associated with brain disorders is estimated to be approximately 800 billion EUR [25].

The biochemical validation of hypothesized drug–target interactions is laborious, time-consuming and expensive [31] [49]. Therefore, computational methods have been proposed for the prediction of drug–target interactions [5] [22] [23] [34]. Traditional techniques include approaches based on molecular docking [10] [15] [29], ligand chemistry [21], [26] and text mining [53].

A serious limitation of docking-based approaches is that they require information about the three-dimensional structure of candidate drugs and targets which is often not available, especially for G-protein coupled receptors (GPCRs) and ion channels. Additionally, the performance of ligand-based approaches decrease in case if only few ligands are known.

In response to the above limitations of classic approaches, expert systems based on machine learning techniques have been proposed for the prediction of drug–target interactions [14] [50] [51]. Recent approaches are based on matrix factorization [12] [14] [52], restricted Boltzmann machines [48], network-based inference [9] [11] [35] [43], positive-unlabeled learning [22] and the integration of multiple sources of information [33] [42]. See also [32] and [37] for excellent surveys.

One of the most prominent drug–target interaction prediction techniques is based on Bipartite Local Models (BLMs) [4] in case of which the drug–target interaction prediction task is treated as a link prediction problem in bipartite graphs. Recent works aiming to extend BLMs focused on semi-supervised prediction [50], improved kernels [47] and the incorporation of neighbor-based interaction-profile inferring [24].

However, none of the aforementioned prediction techniques took the presence of hubs into account. With hubs, we mean entities that are connected to surprisingly many other entities of a network. This phenomenon has been observed for various biological, chemical and medical networks, see e.g. [1] [17]. Similar observations can be made for drug–target networks as well, e.g., Fig.1

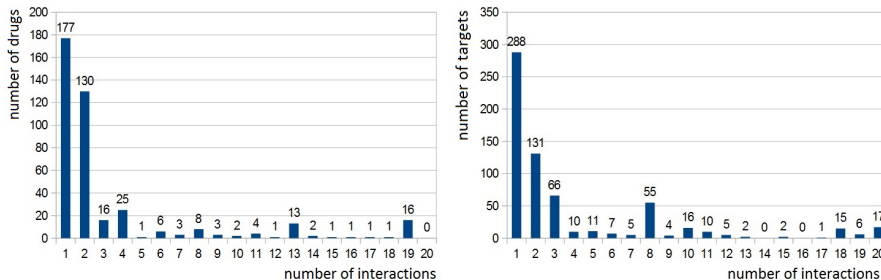


Figure 1: The degree distribution in the *Enzyme* drug–target network. The horizontal axis corresponds to the number of interactions, whereas the vertical axis corresponds to the number of drugs (left) or targets (right). For example, the first column in the left diagram shows that 177 drugs interact with only one (but not necessarily the same) target. In contrast, some drugs (and targets, resp.) participate in surprisingly many interactions, e.g., there are 16 drugs, each of which interacts with 19 targets.

shows the degree distribution for the drugs and targets in the *Enzyme* drug–target interaction network (we will describe the data in Section 2.1). As one can see, the distributions have long tails, i.e., there are drugs (and targets, resp.) that are connected with surprisingly many targets (drugs, resp.) compared to “average” drugs (targets, respectively).

The presence of hubs has been observed in nearest neighbor graphs, see e.g. [8] [28] [46], and hubness-aware classifiers have been developed, see [45] for a survey. More recently, hubness-aware regression techniques, including k -nearest neighbor with error correction ($ECkNN$), were developed that allow for predictions on a continuous scale [7]. Despite the fact that hubness-aware techniques are among the most promising recent machine learning approaches, their potential to enhance drug–target interaction prediction methods has not been exploited yet: to the best of our knowledge, our initial work [6] is the only one aiming to apply hubness-aware models to the drug–target prediction problem.

In this study, we extend Bipartite Local Models and our previous work [6]. We use $ECkNN$ as local model of BLM and propose an enhanced representation of drugs and targets in a multi-modal similarity space (i.e., a representation which incorporates multiple similarity measures). Furthermore, we build a projection-based ensemble and study how the performance depends on the number of base models of the ensemble. As we use hubness-aware local models in the proposed approach, we refer to it as HLM for simplicity. In order to assist reproducibility of our work as well as comparison to published results, we perform experiments on publicly available real-world drug–target interaction datasets. The results show that our approach outperforms other state-of-the-art drug–target prediction techniques.

The rest of this paper is organized as follows: in Section 2 we review the

Table 1: Number of drugs, targets and interactions in the datasets used in our study.

Dataset	# Drugs	# Targets	# Interactions
Enzyme	445	664	2926
Ion Channels	210	204	1476
G-protein coupled receptors (GPCR)	223	95	635
Nuclear Receptors (NR)	54	26	90
Kinase [36]	68	442	1527

background necessary to understand our work. In particular, we focus on BLM and ECkNN. Section 3 presents the proposed approach, followed by its experimental evaluation in Section 4. Finally, conclusions are drawn in Section 5.

2 Materials and Methods

In order to ensure that the paper is self-contained, we begin this section by describing the datasets used in our study and the procedure to obtain drug-drug and target-target similarities. Subsequently, the BLM approach for drug-target interaction prediction is reviewed. This is followed by the description of hubness-aware error correction for nearest neighbor regression.

2.1 Drug-Target Interaction Data

In our study we used five publicly available drug-target interaction datasets from two repositories,¹ namely Enzyme, Ion Channel, G-protein coupled receptors (GPCR), Nuclear Receptors (NR), and Kinase[36]. These datasets have been used in various studies, see e.g. [4] [14] [50], [51] and [39].

Each of the first four datasets contains a binary interaction matrix between drugs and targets, in which each entry indicates whether the interaction between the corresponding drug and target is known or not. In contrast, Kinase contains continuous values of binding affinity for all drug-target pairs of the data. In order to produce a binary interaction matrix, we used the same cutoff threshold as Pahikkala et al. [39]. Tab. 1 shows the number of drugs, targets and interactions in the datasets.

In general, drug-drug and target-target similarities may be computed in many ways. Next, we describe the similarities used in our study. In case of the Enzyme, Ion Channel, GPCR and NR datasets, chemical structure similarities were computed using the SIMCOMP [16] graph-alignment algorithm, in order to obtain drug-drug similarities. For Kinase, we used 2D Tanimoto coefficients as drug-drug similarities.

¹<http://web.kuicr.kyoto-u.ac.jp/supp/yoshi/drugtarget/>
<http://staff.cs.utu.fi/~aatapa/data/DrugTarget/>

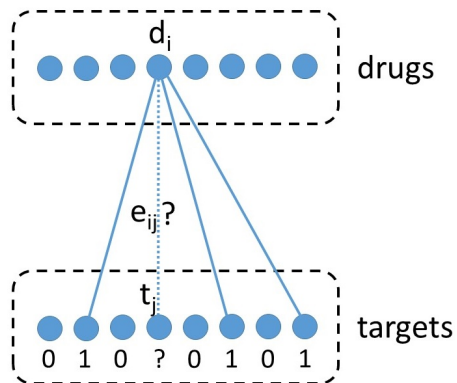


Figure 2: Bipartite Local Models for the prediction of drug–target interactions.

In order to compute the similarity between target proteins of the Enzyme, Ion Channel, GPCR and NR datasets, their amino acid sequences were retrieved from the KEGG GENES [20] database so that similarities between pharmacological targets were determined by sequence alignment methods, such as the Smith-Waterman algorithm. We refer to [51] for more details. For Kinase, we used the normalized Smith-Waterman scores as target–target similarities.

2.2 BLM: Bipartite Local Model

Bipartite Local Models (BLMs) [4] consider the drug–target interaction prediction problem as a link prediction problem in bipartite graphs. As shown in Fig. 2, the vertices in one of the vertex classes of the bipartite graph correspond to drugs, whereas the vertices in the other vertex class correspond to targets. Each edge e_{ij} of the graph corresponds to a known interaction between drug d_i and target t_j .

When predicting the likelihood of an unknown interaction e_{ij} between drug d_i and target t_j , the model computes two independent predictions that are aggregated subsequently.

The first prediction is based on the relations between d_i and the targets. Each target (except t_j) is labeled as “1” or “0” depending on whether or not there is a known interaction between d_i and the target. Then a model is trained to distinguish between “1”-labeled and “0”-labeled targets. Subsequently, this model is applied to predict the likelihood of the unknown interaction e_{ij} . This first prediction for the likelihood of the interaction between drug d_i and target t_j is denoted by $\hat{y}_1(d_i, t_j)$.

The second prediction, denoted by $\hat{y}_2(d_i, t_j)$, is obtained in a similar fashion, but instead of considering the interactions of drug d_i and labeling the targets, the interactions of target t_j are considered and drugs are labeled. The models that make the first and second predictions are called *local models*.

In order to obtain the final prediction of the BLM, we average the predictions of the aforementioned local models:

$$\hat{y}(d_i, t_j) = \frac{\hat{y}_1(d_i, t_j) + \hat{y}_2(d_i, t_j)}{2} \quad (1)$$

Note that instead of averaging, other aggregation functions, such as minimum or maximum are possible as well.

BLM is a generic framework in which various regressors or classifiers can be used as local models. While Bleakley and Yamanishi [4] used support vector machines with a domain-specific kernel, we propose to use a hubness-aware regression technique, ECKNN, which is described in Section 2.4.

2.3 Predictions based on Weighted Profiles

One of the shortcomings of the BLM approach is that it does not handle the case of “new” drugs and “new” targets. With “new” drug, we mean a drug d that does not have any known interaction in the training data. In such cases, all the targets are labeled as “0”, and no reasonable local model can be learned solely from “0”-labeled targets. Similar problem arise in case of “new” targets as well. In order to alleviate this problem, we use the weighted profile [51] approach to obtain predictions for “new” drugs and “new” targets.

Given a “new” drug d_i , and a target t_j , we predict the likelihood of the interaction between d_i and t_j as follows:

$$\hat{y}_1(d_i, t_j) = \frac{\sum_{d' \in \mathcal{D} \setminus \{d_i\}} \text{sim}(d_i, d') i(d', t_j)}{\sum_{d' \in \mathcal{D} \setminus \{d_i\}} \text{sim}(d_i, d')} \quad (2)$$

where \mathcal{D} denotes the set of all drugs, $\text{sim}(d_i, d')$ is the chemical similarity between drugs d_i and d' and $i(d', t_j)$ indicates the presence or absence of interaction between drug d' and target t_j among the known interactions (i.e., interactions in the training set):

$$i(d', t_j) = \begin{cases} 1 & \text{if there is a known interaction between } d' \text{ and } t_j \\ 0 & \text{otherwise.} \end{cases}$$

The intuition behind Eq. (2) is that similar drugs are likely to behave similarly in terms of their interaction with a given target. Therefore, drugs are weighed according to their similarity to the “new” drug d , and we calculate the weighted average of the interaction indicators as prediction.

The case of “new” targets is analogous. Given a “new” target t_j and a drug d_i , the weighted profile approach can be used to calculate the prediction for the likelihood of the interaction between d_i and t_j as follows:

$$\hat{y}_2(d_i, t_j) = \frac{\sum_{t' \in \mathcal{T} \setminus \{t_j\}} \text{sim}(t_j, t') i(d_i, t')}{\sum_{t' \in \mathcal{T} \setminus \{t_j\}} \text{sim}(t_j, t')} \quad (3)$$

where \mathcal{T} denotes the set of all targets and $\text{sim}(t, t')$ is the genomic similarity between targets t and t' .

Although the weighted profile approach is more general than BLM, in the sense that it can be used for “new” drugs and targets as well, the predictions of the weighted profile approach are less accurate than the predictions of BLM. Therefore, we use the weighted profile approach *only* in case of *new* drugs and targets, where the BLM approach does not produce reasonable predictions.

2.4 $\text{EC}k\text{NN}$: k -Nearest Neighbor Regression with Error Correction

In the last decades, various regression schemes have been introduced, such as linear and polynomial regression, support vector regression, neural networks, etc. One of the most popular regression techniques is based on the k -nearest neighbors: when predicting the numeric label on an instance x with k -nearest neighbor regression, the k -nearest neighbors of x (i.e., k most similar instances to x) are determined and the average of their labels is calculated as the predicted label of x . In our case, instances may either correspond to drugs or targets, depending on whether the first or the second prediction of the BLM is calculated.

While being intuitive and simple to implement, k -nearest neighbor regression is well-understood from the point of view of theory as well, see e.g. [2], [3], [13], and the references therein for an overview of the most important theoretical results. These theoretical results are also justified by empirical studies: for example, in their recent paper, Stensbo-Smidt et al. found that nearest neighbor regression outperforms model-based prediction of star formation rates [30], while Hu et al. showed that a model based on k -nearest neighbor regression is able to estimate the capacity of lithium-ion batteries [19].

Despite all of the aforementioned advantages of k -nearest neighbor regression, one of its recently explored shortcomings has to be mentioned, namely, the suboptimal performance in the presence of bad hubs. Intuitively, bad hubs are instances that appear as nearest neighbors of many other instances, but have substantially different labels from those instances. The presence of bad hubs has been shown to be related to the intrinsic dimensionality of the data. This means, roughly speaking, that bad hubs are expected in complex data, such as drug–target interaction data. For a more detailed discussion, we refer to [7].

In order to alleviate the detrimental effect of bad hubs, we proposed an error correction technique which is reviewed next. We define the *corrected label* $y_c(x)$ of a training instance x as

$$y_c(x) = \begin{cases} \frac{1}{|\mathcal{R}_x|} \sum_{x_i \in \mathcal{R}_x} y(x_i) & \text{if } |\mathcal{R}_x| \geq 1 \\ y(x), & \text{otherwise} \end{cases}, \quad (4)$$

where $y(x)$ denotes the original (i.e., uncorrected) label of instance x , and \mathcal{R}_x is the set of “reverse neighbors”, i.e. the set of training instances that have x as one of their k -nearest neighbors. In order to make predictions, k -nearest

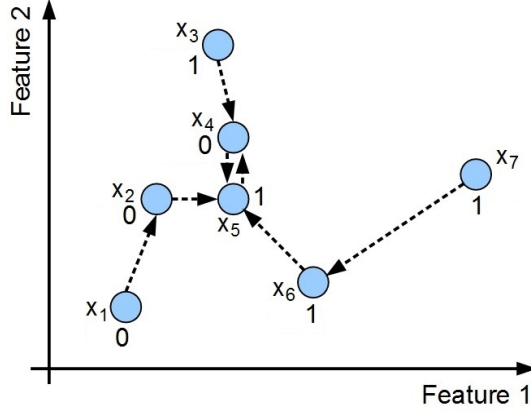


Figure 3: Example used to illustrate error correction. The real values next to each instance x_i denote the labels of the instances.

neighbor regression with error correction (ECkNN) uses the corrected labels. Given a “new” (unlabeled) instance x' , its predicted label $\hat{y}(x')$ is calculated as follows:

$$\hat{y}(x') = \frac{1}{k} \sum_{x_i \in \mathcal{N}(x')} y_c(x_i). \quad (5)$$

where $\mathcal{N}(x')$ denotes the set of nearest neighbors of x' .

Using the example in Fig. 3 we illustrate how the corrected labels are calculated. In Fig. 3 training instances are denoted by circles. They are identified by the symbols $x_1 \dots x_7$. The numeric value (0 or 1) next to each instance shows its label. In order to keep the example simple, we use $k = 1$ nearest neighbor to calculate the corrected labels of training instances. In the figure, directed edges point from each instance to its first nearest neighbor. We only present the calculations for x_4 and x_5 as the procedure is similar in case of the other instances as well. Concretely, the corrected labels of x_4 and x_5 are:

$$y_c(x_4) = \frac{1}{2}(1 + 1) = 1,$$

$$y_c(x_5) = \frac{1}{3}(0 + 0 + 1) = 0.33.$$

Whenever ECkNN predicts the label of a new instance, it uses the corrected labels of the training instances (such as 1 and 0.33 in case of x_4 and x_5) instead of the original labels (i.e., 0 and 1, respectively).

In our study, we used a publicly-available implementation of ECkNN from the PyHubs library.²

²<http://www.biointelligence.hu/pyhubs>

3 HLM: Hubness-aware Local Models

In this section, we develop our approach for drug–target interaction prediction. Following subsections describe components of the method, such as the representation of drugs and targets and the projection-based ensemble.

3.1 Similarity-based Representation

The given drug–drug similarities allow us to represent drugs in terms of their similarities to all the drugs. This means that the first, second, third... features in the representation of drug d correspond to the similarity between d and the first, second, third... drug of the dataset. When using this representation, we say that drugs are represented in the similarity space. Given the target–target similarity matrices, targets may be represented in an analogous way, i.e., using their similarities to all the targets.

While the chemical and genomic similarities between drugs and targets, as described in Section 2.1 are undoubtedly useful, various studies suggest that interaction data (i.e. the known interactions between drugs and targets in our case) may contain highly valuable information as well. For example, in the domain of recommender systems, even a few ratings (i.e., “interactions” between customers and products) may be more informative than metadata about users and products [27]. Therefore, additionally to the given drug–drug and target–target similarities, we propose to compute drug–drug and target–target similarities based on the known interactions (i.e., interaction in the training set). In particular, using the interaction matrix, we calculate the Jaccard-similarity between drugs as well as between targets. Thus the enhanced similarity-based representation of a drug (or target, respectively) consists of its chemical (genomic) similarity to all the drugs (targets) and its interaction-based similarity to all the drugs (targets). This is illustrated in Fig. 4.

3.2 Projection-based Ensemble

We propose to build a projection-based ensemble of BLMs as follows. Given the enhanced similarity-based representation of drugs and targets, we select a random subset of features and use only the selected features when training the local models and making predictions. Denoting the size of the set of selected features by F_D and F_T (for drugs and targets, respectively), the above procedure first projects drugs into F_D -dimensional, and targets into an F_T -dimensional subspace. Subsequently, these lower dimensional representations are used with the prediction models.

The above process of random selection of features and making predictions using the resulting lower-dimensional representation is repeated N -times. This results in an ensemble of N prediction models. As each member of the ensemble is constructed in the same way, their expected prediction accuracies will be similar, therefore, we propose to average the predictions of the members of the

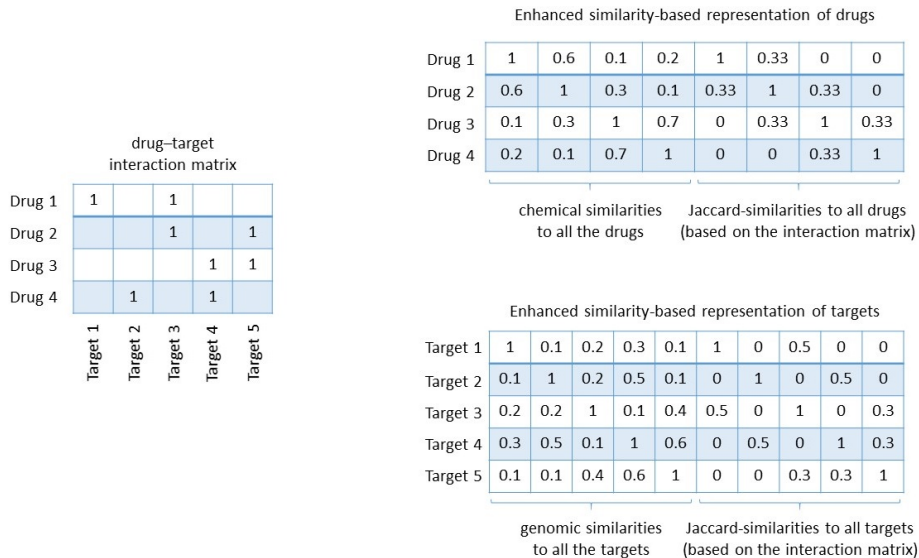


Figure 4: Illustration of enhanced similarity-based representation of drugs and targets

ensemble. Thus the final output of the ensemble is:

$$\hat{y}(d_i, t_j) = \frac{1}{N} \sum_{l=1}^N \hat{y}^{(l)}(d_i, t_j) \quad (6)$$

where $\hat{y}^{(l)}(d_i, t_j)$ is the prediction of the l -th BLM for the interaction between drug d_i and target t_j .

The projection-based ensemble is illustrated in Fig. 5 for $N = 2$ base prediction models and the selection of $F_D = F_T = 3$ features from the enhanced similarity-based representation.

In order to account for the presence of hubs, we propose to use EC k NN as local model in each of the BLMs. The proposed approach is summarized in Alg. 1.

4 Experimental Evaluation

We begin the experimental evaluation of our approach by the description of the experimental settings. This is followed by comparison to well-known drug-target interaction prediction techniques, and the detailed analysis of HLM. Subsequently, we compare HLM to more recent drug-target interaction prediction techniques and discuss biological feasibility of the predictions.

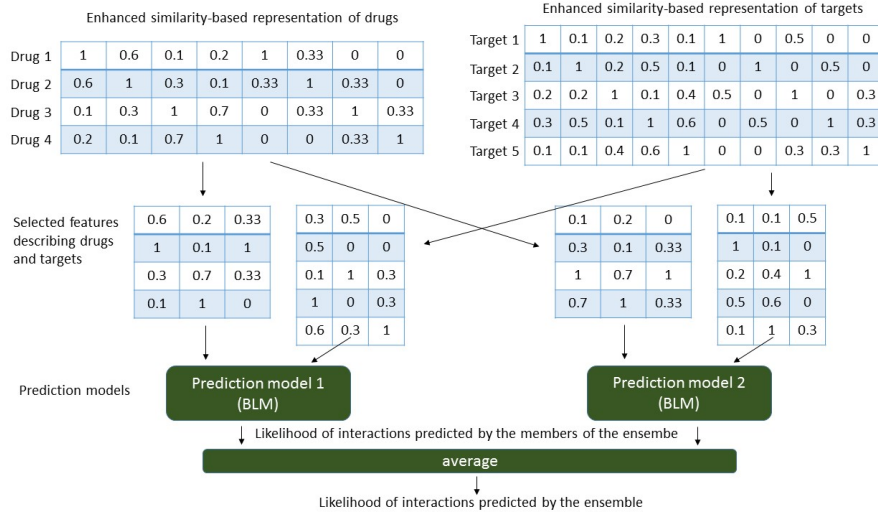


Figure 5: Projection-based ensemble of BLMs using the enhanced similarity-based representation of drugs and targets.

4.1 Experimental Settings

In order to assist reproducibility and comparability with published results, in Sections 4.2–4.4 we used the wide-spread evaluation protocol of Bleakley and Yamanishi [4], i.e., we used leave-one-interaction-out cross-validation³ on the same datasets and evaluated the predictions in terms of Area Under ROC Curve (AUC) and Area Under Precision-Recall Curve (AUPR).

According to the recent study of Pahikkala et al. [39], evaluating drug–target interaction prediction techniques with leave-one-out cross-validation may lead to overoptimistic results. Therefore we used interaction-based⁴ 5×5 -fold cross-validation in the subsequent experiments in Section 4.5 when we compared our approach with more recent drug–target prediction techniques, such as net Laplacian regularized least squares (NetLapRLS) [44], BLM with neighbor-based interaction-profile inferring (BLM-NII) [24] and a combination of weighted nearest neighbor and Gaussian interaction profile kernels (WNN-GIP) [40].

We implemented the proposed approach in Python. We used the EC k NN implementation from the publicly available PyHubs library and methods from the NumPy machine learning library for the calculation of AUC and AUPR. The results presented in in Sections 4.2–4.4 were obtained with $k = 3$ nearest

³Jaccard-similarities between drugs (and targets, resp.) were re-calculated in each iteration of the cross-validation: when making prediction for the pair (d, t) , Jaccard-similarities between drugs were calculated using all the targets except t . Similarly, the Jaccard-similarities between targets were calculated using all drugs except d .

⁴In each round of the cross-validation, the test set contains one fifth of all the drug–target pairs.

Algorithm 1 Hubness-aware local models

Require: Drug–Target interaction matrix I , Drug–drug similarity matrix S_D , Target–target similarity matrix S_T , number of nearest neighbors k , ensemble size N , number of selected features F_D , F_T

Ensure: Likelihood of drug–target interactions

- 1: $D \leftarrow$ enhanced similarity-based representations of drugs
 - 2: $T \leftarrow$ enhanced similarity-based representations of targets
 - 3: **for** $l = 1 \dots N$ **do**
 - 4: $D' \leftarrow$ random sample of D with F_D features
 - 5: $T' \leftarrow$ random sample of T with F_T features
 (Only features of D and T are sampled, all the drugs/targets are kept.)
 - 6: Predict interaction scores with BLM using $EckNN$ as local model and D' and T' as the representation of drugs and targets respectively.
 (Use the weighted profile approach for predictions for “new” drugs/targets.)
 - 7: **end for**
 - 8: Aggregate the predictions that were made in each execution of the loop
-

neighbors. However, we note that we also tried other settings of this parameter, such as $k = 5$, which lead to similar results. We used implementations of NetLapRLS, BLM-NII and WNN-GIP from the publicly available PyDTI software library.⁵

4.2 Comparison with Wide-Spread Drug–Target Interaction Prediction Techniques

In our first experiment, we compared the performance of our approach, HLM, with prominent drug–target interaction prediction techniques from the literature. In particular, we used Kernel-Regression Method (KRM) [51], Bipartite Local Models (BLM) [4] and the weighted profile approach (WP) as baselines.

In this experiment we set the size of the ensemble to $N = 1$ and the number of selected features was set to the number of all the available features, i.e., the enhanced similarity-based representation was used with the local models, without further projection to a lower-dimensional space.

The results are summarized in Tab. 2. As one can see, our approach, HLM outperformed its competitors both in terms of AUC and AUPR.

4.3 The Effect of Hubness-aware Error Correction

In order to examine the effect of hubness-aware error correction, we predicted drug–target interactions without error correction as well. All the experimental settings were the same as in the previous experiment with the only exception that we used simple k -nearest neighbor regressors instead of $EckNN$ as local

⁵<https://github.com/stephenliu0423/PyDTI>

Table 2: The performance of our approach, HLM, and other drug–target interaction prediction techniques from the literature. The ensemble size in HLM is set to $N = 1$.

Dataset	Method	AUC (in %)	AUPR (in %)
Enzyme	HLM	98.4	92.9
	BLM	97.3	84.1
	KRM	96.7	83.1
	WP	95.3	73.2
Ion Channel	HLM	98.5	91.0
	BLM	97.0	77.9
	KRM	96.9	77.8
	WP	96.4	77.0
GPCR	HLM	96.8	74.4
	BLM	95.3	66.7
	KRM	94.7	66.4
	WP	89.0	42.9
NR	HLM	95.5	68.1
	BLM	85.8	60.0
	KRM	86.7	61.0
	WP	80.5	39.4

Table 3: The performance of the proposed approach *with* and *without* hubness-aware error correction (EC).

Dataset		AUC (in %)	AUPR (in %)
Enzyme	HLM (with EC)	98.4	92.9
	predictions without EC	97.4	88.2
Ion Channel	HLM (with EC)	98.5	91.0
	predictions without EC	97.6	89.3
GPCR	HLM (with EC)	96.8	74.4
	predictions without EC	95.0	70.7
NR	HLM (with EC)	95.5	68.1
	predictions without EC	89.9	58.2

models. The results are shown in Tab. 3. The results clearly indicate that error correction indeed contributes to accurate predictions.

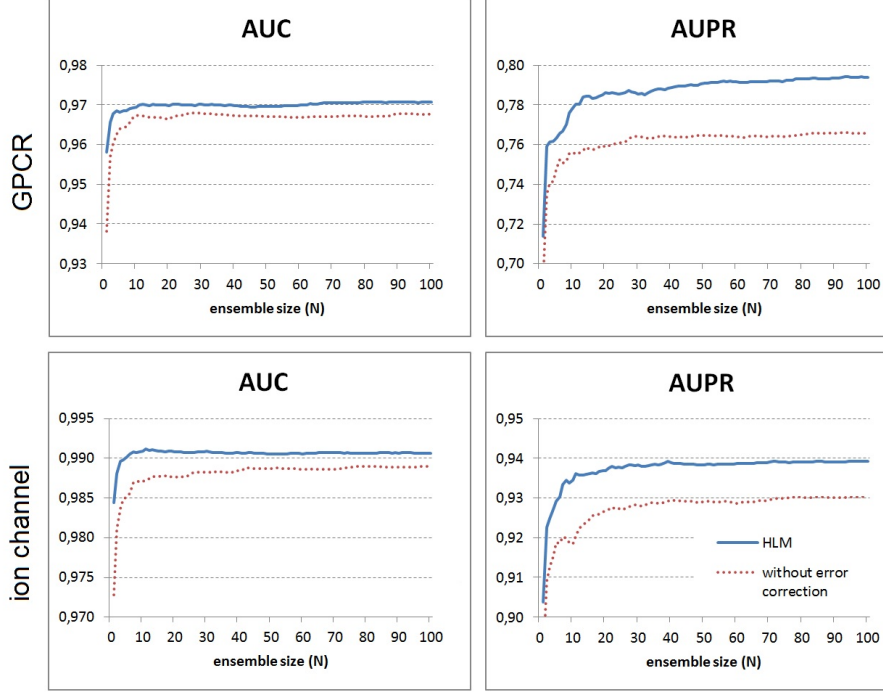


Figure 6: The effect of the ensemble size on the prediction performance in case of the proposed approach (HLM, solid line) and the predictions without error correction (dotted line). Although increasing the ensemble size leads to performance improvements in both cases, we note that in case of HLM, the performance converges to a higher accuracy as measured by both AUC and AUPR.

4.4 The Effect of the Ensemble Size

Next we examined the effect of the ensemble size (N) on the prediction performance. We set the number of selected features $F_T = F_D = 50$ and varied the number of base models (N) in the ensemble. Fig. 6 shows detailed results for two datasets, GPCR and ion channel, however, we note that we observed similar trends in case of the other two datasets as well.

First of all, we could observe that the projection-based ensemble technique described in Section 3.2 systematically increases the performance. For example, on the GPCR dataset in case of $N \approx 100$ base models the AUPR of predictions increased from 74.4 % to 79.4 % compared with the case of using all the available features within one prediction model used in the previous experiments. For the ion channel dataset, the AUPR increased from 91 % to 93.9 %.

In order to examine the interaction between the projection-based ensemble technique and hubness-aware error correction, we also built ensembles using k -

nearest neighbor regression *without* error correction as local models. All the other experimental settings were the same as in case of the proposed approach. In both cases (i.e., with and without error correction) the performance improves with increasing ensemble size. However, in case of the proposed approach the performance converges to a higher accuracy as measured by both AUC and AUPR. This indicates that error correction and the projection based ensemble technique are complementary to each other, i.e., both of them are essential for accurate predictions.

4.5 Comparison with Recent Drug–Target Interaction Prediction Techniques

In order to illustrate that our approach, HLM, is able to outperform recent drug–target interaction prediction techniques as well, we performed experiments with HLM with $N = 25$ base prediction models. Other hyperparameters of HLM were learned via grid-search in internal 5-fold cross-validation on the training data. In particular: the number of nearest neighbors for the local model EC k NN and the number of selected features, were chosen from $\{3, 5, 7\}$ and $\{10, 20, 50\}$.

Hyperparameters of the baselines were learned similarly. In particular: for BLM-NII, the max function was used to generate final predictions and the weight α for the combination of structural and collaborative similarities was chosen from $\{0.0, 0.1, \dots, 1.0\}$. In WNN-GIP, the decay hyperparameter T was chosen from $\{0.1, 0.2, \dots, 1.0\}$ and the weight α for combination of structural and collaborative similarities was chosen from $\{0.0, 0.1, \dots, 1.0\}$. The hyperparameters⁶ of NetLapRLS, were chosen from $\{10^{-6}, 10^{-5}, \dots, 10^2\}$.

As the results in Tab. 4 show, our approach, HLM, outperforms NetLapRLS, BLM-NII and WNN-GIP in all the cases, except the NR dataset, where its performance was similar to that of BLM-NII (the difference statistically insignificant). We note, however, that NR is an exceptionally small dataset, therefore, the results obtained on NR are likely to be less stable compared to other datasets.

We performed paired t-test at significance level $p = 0.01$ in order to check if the differences between the best-performing approach (i.e., HLM in almost all the cases) and the other techniques are statistically significant or not. In Tab. 4 and Tab. 5, asterisk (*) denotes the cases where the difference was found to be statistically significant. As one can see, in the vast majority of the cases our approach, HLM, outperformed its competitors statistically significantly.

We also evaluated the predictions in terms of precision, recall and accuracy. In order to do that, we ranked drug–target pairs according to their predicted interaction scores and considered the top-ranked 5% of the interactions as the ones that are predicted to interact (predicted positives). The rest of the drug–target pairs were considered as the predicted non-interacting drug–target pairs (predicted negatives). As the results in Tab. 5 show, our approach, HLM,

⁶ $\beta = \beta_{drug} = \beta_{target}$ and $\gamma = \gamma_{drug} = \gamma_{target}$

Table 4: The performance (AUC, AUPR) of our approach, HLM, and recent drug-target interaction prediction techniques.

Dataset	Method	AUC (in %)	AUPR (in %)
Enzyme	HLM	96.6	83.2
	BLM-NII	96.6	62.8*
	NetLapRLS	95.9*	78.4*
	WNN-GIP	94.5*	70.8*
Ion Channel	HLM	98.0	86.7
	BLM-NII	96.0*	62.6*
	NetLapRLS	96.6*	82.7*
	WNN-GIP	94.7*	66.3*
GPCR	HLM	94.7	68.6
	BLM-NII	92.9*	38.7*
	NetLapRLS	91.0*	58.0*
	WNN-GIP	92.8*	51.3*
NR	HLM	86.4	57.6
	BLM-NII	87.9	54.3
	NetLapRLS	81.0*	42.8*
	WNN-GIP	86.2	55.0
Kinase	HLM	93.1	61.5
	BLM-NII	70.2*	8.8*
	NetLapRLS	91.3*	57.7*
	WNN-GIP	52.8*	6.2*

outperformed its competitors in terms of precision, recall and accuracy as well.⁷

4.6 Prediction of New Interactions

In this section, by examining some of the highly scored interactions in detail, we illustrate that HLM not only achieves high accuracy in terms of AUC and AUPR, but the predictions are biologically feasible as well.

We begin this discussion by noting that the drug-target interactions contained in the Enzyme, Ion Channel, G-protein coupled receptors (GPCR) and Nuclear Receptors (NR) datasets were extracted from the Kyoto Encyclopedia of Genes and Genomes⁸ (KEGG) *several years ago* and, in order to allow for comparison of prediction techniques, they have been kept unchanged. Meanwhile, additional drug-target interactions have been validated chemically and

⁷On the Enzyme data the difference in accuracy was in the second decimal place (not shown in Tab. 5), however, it was systematic throughout the $5 \times 5 = 25$ rounds of cross-validation, therefore HLM was found to outperform its competitors significantly.

⁸<http://www.kegg.jp/>

Table 5: Precision, recall and accuracy of our approach, HLM, and recent drug-target interaction prediction techniques.

Dataset	Method	Precision (in %)	Recall (in %)	Accuracy (in %)
Enzyme	HLM	17.9	90.4	95.8
	BLM-NII	17.7*	89.7*	95.8*
	NetLapRLS	17.6*	89.1*	95.8*
	WNN-GIP	17.3*	87.8*	95.7*
Ion Channel	HLM	61.8	89.3	97.7
	BLM-NII	51.4*	74.3*	96.7*
	NetLapRLS	57.7*	83.4*	97.3*
	WNN-GIP	59.1*	85.5*	97.5*
GPCR	HLM	45.8	76.7	96.6
	BLM-NII	35.7*	59.9*	95.6*
	NetLapRLS	39.3*	65.7*	95.9*
	WNN-GIP	43.7*	73.1*	96.4*
NR	HLM	64.3	50.8	95.0
	BLM-NII	56.3*	44.2*	94.2*
	NetLapRLS	49.1*	38.3*	93.5*
	WNN-GIP	63.7	50.5	94.9
Kinase	HLM	57.4	56.2	95.6
	BLM-NII	7.7*	7.6*	90.7*
	NetLapRLS	57.3	56.1	95.6
	WNN-GIP	6.4*	6.2*	90.5*

the results have been uploaded to databases, such as KEGG [20], DrugBank [41] or Matador [38].

Therefore, in order to demonstrate that our approach is able to predict new interactions, we trained our approach, HLM, as well as its competitors, BLM-NII, NetLapRLS and WNN-GIP using all the interactions of the original datasets, and ranked the non-interacting drug-target pairs of the original datasets according to their predicted interaction scores. For simplicity, we use the term *predicted new interactions* for the top-ranked 20 drug-target pairs. We say that a predicted new interaction is *validated* if it is included in the current version of KEGG, DrugBank or Matador.

Our approach, HLM, had the overall best performance in terms of the number of validated interactions compared with BLM-NII, NetLapRLS and WNN-GIP. For example, on the Ion Channel and NR datasets, HLM was able to predict 12 and 8 validated interactions, whereas none of its competitors was able to predict more than 6 validated interactions on these datasets. For En-

Table 6: Top 20 new interactions predicted by our approach, HLM, on the Enzyme dataset.

	Drug	Target	Validated	Also predicted by
1	D00542	hsa:1571	yes	BLM-NII, NetLapRLS, WNN-GIP
2	D00437	hsa:1559	yes	BLM-NII, NetLapRLS, WNN-GIP
3	D00139	hsa:1543	yes	BLM-NII, NetLapRLS, WNN-GIP
4	D00574	hsa:1589	yes	NetLapRLS, WNN-GIP
5	D00691	hsa:5152	no	NetLapRLS
6	D00410	hsa:1543	yes	BLM-NII, NetLapRLS, WNN-GIP
7	D00410	hsa:1583	yes	NetLapRLS
8	D00437	hsa:1585	yes	BLM-NII
9	D00691	hsa:8654	no	NetLapRLS
10	D00528	hsa:1549	yes	BLM-NII, NetLapRLS, WNN-GIP
11	D00410	hsa:1585	yes	BLM-NII
12	D00449	hsa:5742	no	—
13	D00126	hsa:246	yes	—
14	D00126	hsa:247	yes	—
15	D03670	hsa:1579	no	BLM-NII, NetLapRLS, WNN-GIP
16	D05458	hsa:4128	yes	—
17	D00691	hsa:5150	yes	—
18	D03670	hsa:51302	no	NetLapRLS
19	D03670	hsa:9420	no	NetLapRLS, WNN-GIP
20	D00691	hsa:50940	no	NetLapRLS

zyme and GPCR datasets, new interactions⁹ are shown in Tab. 6 and Tab. 7. The last columns of these tables show if the same interaction was predicted by other approaches.

Most notably, numerous validated interactions (in **bold**) were only predicted by our approach, HLM, including the interactions between *Ibuprofen* (D00126) and *arachidonate 15-lipoxygenase* (hsa:246) and its second type (hsa:247); as well as the interaction between *Phentermine* (D05458) and *monoamine oxidase A* (hsa:4128); and the interaction between *Dyphylline* (D00691) and *phosphodiesterase 7A* (hsa:5150). On the GPCR dataset, only our approach was able to predict the validated interaction between *Theophylline sodium acetate* (D01712) and *adenosine A2b receptor* (hsa:136), as well as the interaction between *Loxapine* (D02340) and *dopamine receptor D1* (hsa:1812).

⁹For the resolution of drug and target identifiers, see <http://www.kegg.jp/>

Table 7: Top 20 new interactions predicted by our approach, HLM, on the GPCR dataset.

	Drug	Target	Validated	Also predicted by
1	D00283	hsa:1814	yes	BLM-NII, NetLapRLS, WNN-GIP
2	D00437	hsa:152	no	NetLapRLS, WNN-GIP
3	D00095	hsa:155	yes	BLM-NII
4	D00255	hsa:152	yes	NetLapRLS
5	D00227	hsa:136	no	WNN-GIP
6	D00503	hsa:1816	no	—
7	D01712	hsa:136	yes	—
8	D02340	hsa:1812	yes	—
9	D02354	hsa:1814	no	—
10	D04625	hsa:154	yes	BLM-NII
11	D00283	hsa:1132	yes	WNN-GIP
12	D03621	hsa:1132	no	WNN-GIP
13	D02358	hsa:154	yes	BLM-NII
14	D00494	hsa:1815	no	—
15	D02147	hsa:153	yes	BLM-NII
16	D02354	hsa:1816	no	—
17	D00503	hsa:3356	no	BLM-NII
18	D02614	hsa:154	no	BLM-NII
19	D00513	hsa:152	no	NetLapRLS, WNN-GIP
20	D00513	hsa:3363	no	WNN-GIP

5 Conclusion and Outlook

In this study, we considered the drug–target interaction prediction problem which has important applications in understanding the mechanisms of how drugs effect, drug repositioning and prediction of adverse effects. We proposed an extension of BLM, one of the most prominent drug–target prediction models. In particular, we proposed to represent drugs and targets in a multi-modal similarity space, used EC_kNN , a hubness-aware regression approach as local model in BLM and built projection-based ensembles. We called the resulting approach HLM (hubness-aware local models).

We performed experiments on widely-used publicly-available datasets, the results of which show that our approach is superior to other drug–target prediction approaches from the literature such as BLM, KRM, BLM-NII, NetLapRLS and WNN-GIP. We examined the effect of hubness-aware error correction and the ensemble size and found that both of these techniques are essential for high accuracy. Subsequently, we demonstrated that our approach is able to predict new drug–target interactions.

Acknowledgment

Krisztian Buza was supported by the National Research, Development and Innovation Office - NKFIH PD 111710 (“BioMining” project) and the János Bolyai Research Scholarship of the Hungarian Academy of Sciences. Ladislav Peska was supported by the Czech grant P46 and Hungarian state stipend no. 119058.

References

- [1] Albert-László Barabási, Natali Gulbahce, and Joseph Loscalzo. Network medicine: a network-based approach to human disease. *Nature Reviews Genetics*, 12(1):56–68, 2011.
- [2] Gérard Biau, Frédéric Céro, and Arnaud Guyader. On the rate of convergence of the bagged nearest neighbor estimate. *The Journal of Machine Learning Research*, 11:687–712, 2010.
- [3] Gérard Biau, Luc Devroye, Vida Dujmović, and Adam Krzyżak. An affine invariant k-nearest neighbor regression estimate. *Journal of Multivariate Analysis*, 112:24–34, 2012.
- [4] Kevin Bleakley and Yoshihiro Yamanishi. Supervised prediction of drug–target interactions using bipartite local models. *Bioinformatics*, 25(18):2397–2403, 2009.
- [5] Bence Bolgár, Adam Arany, Gergely Temesi, Balázs Balogh, Péter Antal, and Peter Matyus. Drug repositioning for treatment of movement disorders: from serendipity to rational discovery strategies. *Current topics in medicinal chemistry*, 13(18):2337–2363, 2013.
- [6] Krisztian Buza. Drug-target interaction prediction with hubness-aware machine learning. In *Proceedings of the 11th IEEE International Symposium on Applied Computational Intelligence and Informatics*, pages 437–440, 2016.
- [7] Krisztian Buza, Alexandros Nanopoulos, and Gábor Nagy. Nearest neighbor regression in the presence of bad hubs. *Knowledge-Based Systems*, 86:250–260, 2015.
- [8] Krisztian Buza, Alexandros Nanopoulos, and Lars Schmidt-Thieme. Insight: efficient and effective instance selection for time-series classification. In *Advances in Knowledge Discovery and Data Mining*, pages 149–160. Springer, 2011.
- [9] Xing Chen, Ming-Xi Liu, and Gui-Ying Yan. Drug–target interaction prediction by random walk on the heterogeneous network. *Molecular BioSystems*, 8(7):1970–1978, 2012.

- [10] Alan C Cheng, Ryan G Coleman, Kathleen T Smyth, Qing Cao, Patricia Soulard, Daniel R Caffrey, Anna C Salzberg, and Enoch S Huang. Structure-based maximal affinity model predicts small-molecule druggability. *Nature biotechnology*, 25(1):71–75, 2007.
- [11] Feixiong Cheng, Chuang Liu, Jing Jiang, Weiqiang Lu, Weihua Li, Guixia Liu, Weixing Zhou, Jin Huang, and Yun Tang. Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS Comput Biol*, 8(5):e1002503, 2012.
- [12] Murat Can Cobanoglu, Chang Liu, Feizhuo Hu, Zoltan N Oltvai, and Ivet Bahar. Predicting drug–target interactions using probabilistic matrix factorization. *Journal of chemical information and modeling*, 53(12):3399–3409, 2013.
- [13] Luc Devroye, Laszlo Györfi, Adam Krzyzak, and Gábor Lugosi. On the strong universal consistency of nearest neighbor regression function estimates. *The Annals of Statistics*, pages 1371–1385, 1994.
- [14] Mehmet Gönen. Predicting drug–target interactions from chemical and genomic kernels using bayesian matrix factorization. *Bioinformatics*, 28(18):2304–2310, 2012.
- [15] Inbal Halperin, Buyong Ma, Haim Wolfson, and Ruth Nussinov. Principles of docking: An overview of search algorithms and a guide to scoring functions. *Proteins: Structure, Function, and Bioinformatics*, 47(4):409–443, 2002.
- [16] Masahiro Hattori, Yasushi Okuno, Susumu Goto, and Minoru Kanehisa. Development of a chemical structure comparison method for integrated analysis of chemical and genomic information in the metabolic pathways. *Journal of the American Chemical Society*, 125(39):11853–11865, 2003.
- [17] Andrew L Hopkins. Network pharmacology: the next paradigm in drug discovery. *Nature chemical biology*, 4(11):682–690, 2008.
- [18] Andrew L Hopkins. Drug discovery: predicting promiscuity. *Nature*, 462(7270):167–168, 2009.
- [19] Chao Hu, Gaurav Jain, Puqiang Zhang, Craig Schmidt, Parthasarathy Gomadani, and Tom Gorka. Data-driven method based on particle swarm optimization and k-nearest neighbor regression for estimating capacity of lithium-ion battery. *Applied Energy*, 129:49–55, 2014.
- [20] Minoru Kanehisa, Susumu Goto, Masahiro Hattori, Kiyoko F Aoki-Kinoshita, Masumi Itoh, Shuichi Kawashima, Toshiaki Katayama, Michihiro Araki, and Mika Hirakawa. From genomics to chemical genomics: new developments in kegg. *Nucleic acids research*, 34(suppl 1):D354–D357, 2006.

- [21] Michael J Keiser, Bryan L Roth, Blaine N Armbruster, Paul Ernsberger, John J Irwin, and Brian K Shoichet. Relating protein pharmacology by ligand chemistry. *Nature biotechnology*, 25(2):197–206, 2007.
- [22] Wei Lan, Jianxin Wang, Min Li, Jin Liu, Yaohang Li, Fang-Xiang Wu, and Yi Pan. Predicting drug–target interaction using positive-unlabeled learning. *Neurocomputing*, 2016.
- [23] Charles J Manly, Shirley Louise-May, and Jack D Hammer. The impact of informatics and computational chemistry on synthesis and screening. *Drug discovery today*, 6(21):1101–1110, 2001.
- [24] Jian-Ping Mei, Chee-Keong Kwoh, Peng Yang, Xiao-Li Li, and Jie Zheng. Drug–target interaction prediction by learning from local information and neighbors. *Bioinformatics*, 29(2):238–245, 2013.
- [25] J Olesen, A Gustavsson, Mikael Svensson, H-U Wittchen, and B Jönsson. The economic cost of brain disorders in europe. *European Journal of Neurology*, 19(1):155–162, 2012.
- [26] Stéphanie Pérot, Leslie Regad, Christelle Reynès, Olivier Spérandio, Maria A Miteva, Bruno O Villoutreix, and Anne-Claude Camproux. Insights into an original pocket-ligand pair classification: a promising tool for ligand profile prediction. *PloS one*, 8(6):e63730, 2013.
- [27] István Pilászy and Domonkos Tikk. Recommending new movies: even a few ratings are more valuable than metadata. In *Proceedings of the third ACM conference on Recommender systems*, pages 93–100. ACM, 2009.
- [28] Miloš Radovanović, Alexandros Nanopoulos, and Mirjana Ivanović. Hubs in space: Popular nearest neighbors in high-dimensional data. *The Journal of Machine Learning Research*, 11:2487–2531, 2010.
- [29] Matthias Rarey, Bernd Kramer, Thomas Lengauer, and Gerhard Klebe. A fast flexible docking method using an incremental construction algorithm. *Journal of molecular biology*, 261(3):470–489, 1996.
- [30] Kristoffer Stensbo-Smidt, Christian Igel, Andrew Zirm, and Kim Steenstrup Pedersen. Nearest neighbour regression outperforms model-based prediction of specific star formation rate. In *Big Data, 2013 IEEE International Conference on*, pages 141–144. IEEE, 2013.
- [31] S Joshua Swamidass. Mining small-molecule screens to repurpose drugs. *Briefings in bioinformatics*, 12(4):327–335, 2011.
- [32] [Chen, Xing and Yan, Chenggang Clarence and Zhang, Xiaotian and Zhang, Xu and Dai, Feng and Yin, Jian and Zhang, Yongdong. Drug–target interaction prediction: databases, web servers and computational models. *Briefings in bioinformatics*, 17\(4\):696–712, 2016.](#)

- [33] Cheng, Feixiong and Li, Weihua and Wu, Zengrui and Wang, Xichuan and Zhang, Chen and Li, Jie and Liu, Guixia and Tang, Yun. Prediction of polypharmacological profiles of drugs by the integration of chemical, side effect, and therapeutic space. *Journal of chemical information and modeling*, 53(4):753–762, 2013.
- [34] Cheng, Feixiong and Zhou, Yadi and Li, Jie and Li, Weihua and Liu, Guixia and Tang, Yun. Prediction of chemical–protein interactions: multitarget-QSAR versus computational chemogenomic methods. *Molecular BioSystems*, 8(9):2373–2384, 2012.
- [35] Cheng, Feixiong and Zhou, Yadi and Li, Weihua and Liu, Guixia and Tang, Yun. Prediction of chemical-protein interactions network with weighted network-based inference method. *PLoS One*, 7(7):e41064, 2012.
- [36] Davis, Mindy I and Hunt, Jeremy P and Herrgard, Sanna and Ciceri, Pietro and Wodicka, Lisa M and Pallares, Gabriel and Hocker, Michael and Treiber, Daniel K and Zarrinkar, Patrick P. Comprehensive analysis of kinase inhibitor selectivity. *Nature biotechnology*, 29(11):1046–1051, 2011.
- [37] Ding, Hao and Takigawa, Ichigaku and Mamitsuka, Hiroshi and Zhu, Shan-feng. Similarity-based machine learning methods for predicting drug–target interactions: a brief review. *Briefings in Bioinformatics*, 15(5):734–747, 2014.
- [38] Günther, Stefan and Kuhn, Michael and Dunkel, Mathias and Campillos, Monica and Senger, Christian and Petsalaki, Evangelia and Ahmed, Jessica and Urdiales, Eduardo Garcia and Gewiss, Andreas and Jensen, Lars Juhl and others. SuperTarget and Matador: resources for exploring drug-target relationships. *Nucleic acids research*, 36(suppl 1):D919–D922, 2008.
- [39] Pahikkala, Tapio and Airola, Antti and Pietilä, Sami and Shakyawar, Sushil and Sz wajda, Agnieszka and Tang, Jing and Aittokallio, Tero. Toward more realistic drug-target interaction predictions. *Briefings in Bioinformatics*, 16(2):325–337, 2015.
- [40] van Laarhoven, Twan and Marchiori, Elena. Predicting drug-target interactions for new drug compounds using a weighted nearest neighbor profile. *PloS one*, 8(6):e66952, 2013.
- [41] Wishart, David S and Knox, Craig and Guo, An Chi and Shrivastava, Savita and Hassanali, Murtaza and Stothard, Paul and Chang, Zhan and Woolsey, Jennifer. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic acids research*, 34(suppl 1):D668–D672, 2006.

- [42] Wu, Zengrui and Cheng, Feixiong and Li, Jie and Li, Weihua and Liu, Guixia and Tang, Yun. SDTNBI: an integrated network and chemoinformatics tool for systematic prediction of drug–target interactions and drug repositioning. *Briefings in bioinformatics*, page bbw012, 2016.
- [43] Wu, Zengrui and Lu, Weiqiang and Wu, Dang and Luo, Anqi and Bian, Hanping and Li, Jie and Li, Weihua and Liu, Guixia and Huang, Jin and Cheng, Feixiong and others. In silico prediction of chemical mechanism of action via an improved network-based inference method. *British journal of pharmacology*, 173(23):3372–3385, 2016.
- [44] Xia, Zheng and Wu, Ling-Yun and Zhou, Xiaobo and Wong, Stephen TC. Semi-supervised drug-protein interaction prediction from heterogeneous biological spaces. In *BMC systems biology*, volume 4, page S6. BioMed Central Ltd, 2010.
- [45] Nenad Tomašev, Krisztian Buza, Kristóf Marussy, and Piroška B Kis. Hubness-aware classification, instance selection and feature construction: Survey and extensions to time-series. In *Feature selection for data and pattern recognition*. Springer-Verlag, 2015.
- [46] Nenad Tomašev, Miloš Radovanovic, Dunja Mladenec, and Mirjana Ivanovic. A probabilistic approach to nearest-neighbor classification: Naive hubness bayesian knn. In *Proc. CIKM*, 2011.
- [47] Twan van Laarhoven, Sander B Nabuurs, and Elena Marchiori. Gaussian interaction profile kernels for predicting drug–target interaction. *Bioinformatics*, 27(21):3036–3043, 2011.
- [48] Yuhao Wang and Jianyang Zeng. Predicting drug-target interactions using restricted boltzmann machines. *Bioinformatics*, 29(13):i126–i134, 2013.
- [49] Steven Whitebread, Jacques Hamon, Dejan Bojanic, and Laszlo Urban. Keynote review: in vitro safety pharmacology profiling: an essential tool for successful drug development. *Drug discovery today*, 10(21):1421–1433, 2005.
- [50] Zheng Xia, Ling-Yun Wu, Xiaobo Zhou, and Stephen TC Wong. Semi-supervised drug-protein interaction prediction from heterogeneous biological spaces. *BMC systems biology*, 4(Suppl 2):S6, 2010.
- [51] Yoshihiro Yamanishi, Michihiro Araki, Alex Gutteridge, Wataru Honda, and Minoru Kanehisa. Prediction of drug–target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics*, 24(13):i232–i240, 2008.
- [52] Xiaodong Zheng, Hao Ding, Hiroshi Mamitsuka, and Shanfeng Zhu. Collaborative matrix factorization with multiple similarities for predicting drug–target interactions. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1025–1033. ACM, 2013.

- [53] Shanfeng Zhu, Yasushi Okuno, Gozoh Tsujimoto, and Hiroshi Mamitsuka. A probabilistic model for mining implicit chemical compound–generelations from literature. *Bioinformatics*, 21(suppl 2):ii245–ii251, 2005.