

RESEARCH ARTICLE

Modified linear regression predicts drug-target interactions accurately

Krisztian Buza^{1,2}*, Ladislav Peška³, Júlia Koller⁴

1 Faculty of Informatics, ELTE – Eötvös Loránd University, Budapest, Hungary, **2** Center for the Study of Complexity, Babes-Bolyai University, Cluj Napoca, Romania, **3** Department of Software Engineering, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic, **4** Institute of Genomic Medicine and Rare Disorders, Semmelweis University, Budapest, Hungary

* These authors contributed equally to this work.

* buza@biointelligence.hu



Abstract

State-of-the-art approaches for the prediction of drug–target interactions (DTI) are based on various techniques, such as matrix factorisation, restricted Boltzmann machines, network-based inference and bipartite local models (BLM). In this paper, we propose the framework of Asymmetric Loss Models (ALM) which is more consistent with the underlying chemical reality compared with conventional regression techniques. Furthermore, we propose to use an asymmetric loss model with BLM to predict drug–target interactions accurately. We evaluate our approach on publicly available real-world drug–target interaction datasets. The results show that our approach outperforms state-of-the-art DTI techniques, including recent versions of BLM.

OPEN ACCESS

Citation: Buza K, Peška L, Koller J (2020) Modified linear regression predicts drug-target interactions accurately. PLoS ONE 15(4): e0230726. <https://doi.org/10.1371/journal.pone.0230726>

Editor: Hans A Kestler, University of Ulm, GERMANY

Received: September 10, 2019

Accepted: March 6, 2020

Published: April 6, 2020

Copyright: © 2020 Buza et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The datasets are available at <http://web.kuicr.kyoto-u.ac.jp/supp/yoshi/drugtarget/>.

Funding: K. Buza was supported by the project ED_18-1-2019-0030. Project no. ED_18-1-2019-0030 (Application domain specific highly reliable IT solutions subprogramme) has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the Thematic Excellence Programme funding scheme. K. Buza received the "Professor Ferencz Rado" Fellowship of the Babes-Bolyai University, Cluj Napoca, Romania. L. Peška

Introduction

When developing new drugs and identifying their side effects [1], pharmaceutical science relies on findings from related branches of science, including statistics and computer science. An essential step in this process is the identification of interactions between drugs and pharmacological targets. Although the existence of interactions can be reliably confirmed by *in vitro* binding assays, see e.g., [2–5], such methods are expensive and time consuming [6]. In order to address this bottleneck, computational approaches have been designed and implemented for the estimation of the probability of interactions. Therefore, most promising candidates for *in vitro* experiments may be selected based on *in silico* approaches.

The importance of drug–target interaction prediction is further emphasised by the costs of drug development. While estimates vary, they agree that it costs hundreds of millions of dollars to bring a new drug to the market, see e.g. [7] for an overview. Furthermore, the process may take more than 10 years in total.

Drug–target interaction prediction (DTI) techniques promise to reduce the aforementioned costs and time, and to support drug repositioning [8], i.e., the use of an existing medicine to treat a disease that has not been treated with that drug yet.

was supported by Charles University grant Progres Q48. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Drug repositioning is especially relevant for the treatment of rare diseases, including neurological disorders. While each of the rare diseases affect only few people, due to the large number of rare diseases, in total 6-8% of the entire population is affected by *one* of those diseases. This results in a paradox situation: although a significant fraction of the population is suffering from one of the rare diseases, it is economically irrational to develop new drugs for many of them. However, drug repositioning may potentially lead to breakthroughs in such cases.

In silico approaches for DTI include techniques based on docking simulations [9], ligand chemistry [10], text mining [11, 12] and machine learning. Text mining is inherently limited to the identification of entities and interactions that have already been documented, although the output of approaches based on text mining, i.e., the identified interactions, may serve as input data for other approaches, such as the ones based on machine learning. A serious limitation of docking simulations is that information about the three-dimensional structure of candidate drugs and targets is required. In many cases, e.g. for G-protein coupled receptors (GPCR) and ion channels, such information may not be available. Moreover, the performance of ligand-based approaches is known to decrease if only few ligands are known.

For the aforementioned reasons, state-of-the-art DTI techniques are based on machine learning [13–17]. Moreover, the increasing interest is also catalysed by the analogies between DTI and the well-studied recommendation tasks [18–20], which resulted in DTI approaches based on matrix factorisation [21–23]. Further recent DTI techniques are based on support vector regression [6], restricted Boltzmann machines [24], network-based inference [25, 26], decision lists [27], positive-unlabelled learning [16] and bipartite local models (BLM) [28]. Extensions of BLM include semi-supervised prediction [29], improved kernels [30], the incorporation of neighbour-based interaction-profiles [31] and hubness-aware regression [19].

Despite all the aforementioned efforts, accurate prediction of drug–target interactions still remained a challenge. In this paper, we propose a new regression technique for accurate DTI predictions. We use a novel loss function that reflects the needs of drug–target interaction better than wide-spread loss functions, such as mean squared error or logistic loss. Our generic framework of asymmetric loss models (ALM) works with various regressors. For simplicity, we instantiate ALM with linear regression which leads to *asymmetric loss linear regression* (ALLR). We propose to use this new regressor in BLM for drug–target interaction prediction. Note that ALM is substantially different from hubness-aware regressors that we used with BLM in our previous work [19]. As ALLR is a modified version of linear regression, we call our approach *Drug–Target Interaction Prediction with Modified Linear Regression*, or MOLIERE for short. We evaluate MOLIERE on publicly available real-world datasets and show that our approach outperforms state-of-the-art DTI techniques, including recent versions of BLM and the cases when conventional loss functions are used. Furthermore, we show that MOLIERE is able to predict medically relevant drug–target interactions that are not contained in the original datasets.

Materials and methods

Data

We used four publicly available real-world drug–target interaction datasets (Table 1), namely Enzyme, Ion Channel (IC), G-protein coupled receptors (GPCR), Nuclear Receptors (NR). The datasets are available at <http://web.kuicr.kyoto-u.ac.jp/supp/yoshi/drugtarget/>. These datasets have been used in various studies, such as [17, 19, 20, 22, 28, 29].

Each dataset contains an interaction matrix \mathcal{M} between drugs and targets, a drug–drug similarity matrix \mathcal{S}^D and a target–target similarity matrix \mathcal{S}^T . Similarities between targets were

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 (IC)	210	204	1476
G-protein coupled receptors (GPCR)	223	95	635
Nuclear Receptors (NR)	54	26	90

<https://doi.org/10.1371/journal.pone.0230726.t001>

determined by the Smith-Waterman algorithm, see [17, 32] for details. Chemical structure similarities between drugs were computed using the SIMCOMP algorithm [33].

Each entry $m_{i,j}$ of the interaction matrix \mathcal{M} indicates whether the interaction between the i -th drug (denoted as d_i) and j -th target (denoted as t_j) is known:

$$m_{i,j} = \begin{cases} +1 & \text{if there is a known interaction between } d_i \text{ and } t_j \\ -1 & \text{otherwise.} \end{cases} \quad (1)$$

Note that in case of these datasets, only the information about the *presence* of interactions is explicit, there is no explicit information about the *absence* of interactions. In particular, the semantics of $m_{i,j} = -1$ is that the corresponding drug d_i and target t_j *may* or *may not* interact. In fact, some of the drug–target pairs denoted as -1 actually interact, however, the interaction was unknown when these datasets were created, roughly 10 years ago. In order to allow for a fair comparison with other works in the literature, in our experiments reported in Tables 2–4, we used the original datasets without these “new” interactions.

In order to illustrate that our approach is indeed able to predict unknown interactions, we show that using the original data, we could predict many of those interactions that have been discovered meanwhile (Tables 5–7).

Problem formulation

We define the Drug–Target Interaction Prediction problem as follows. We are given a set $\mathcal{D} = \{d_1, \dots, d_n\}$ of n drugs, a set $\mathcal{T} = \{t_1, \dots, t_m\}$ of m pharmaceutical targets, an $n \times n$ drug similarity matrix \mathcal{S}^D , an $m \times m$ target similarity matrix \mathcal{S}^T and an $n \times m$ interaction matrix \mathcal{M} . For

Table 2. The performance of our approach, MOLIERE, compared with BLM and weighted profile (WP) with $k_d = k_t = 5$.

Dataset	Method	AUC	AUPR
Enzyme	MOLIERE	0.990	0.924
	BLM	0.973	0.841
	WP	0.955	0.868
Ion Channel	MOLIERE	0.990	0.954
	BLM	0.970	0.779
	WP	0.974	0.837
GPCR	MOLIERE	0.974	0.837
	BLM	0.953	0.667
	WP	0.943	0.648
NR	MOLIERE	0.921	0.731
	BLM	0.858	0.600
	WP	0.886	0.602

<https://doi.org/10.1371/journal.pone.0230726.t002>

Table 3. The performance of our approach, MOLIERE, compared with the cases when using standard linear regression or logistic regression instead of the proposed regression technique.

Dataset	Method	AUC	AUPR
Enzyme	MOLIERE	0.990	0.924
	linear regression	0.980	0.924*
	logistic regression	0.971	0.870
Ion Channel	MOLIERE	0.990	0.954
	linear regression	0.982	0.952
	logistic regression	0.973	0.891
GPCR	MOLIERE	0.974	0.837
	linear regression	0.963	0.832
	logistic regression	0.917	0.752
NR	MOLIERE	0.921	0.731
	linear regression	0.888	0.724
	logistic regression	0.825	0.676

* In this case, the performance of MOLIERE with ALLR and MOLIERE with linear regression differs in the last digits that are not shown in the table.

<https://doi.org/10.1371/journal.pone.0230726.t003>

Table 4. The performance of our approach, MOLIERE, compared with state-of-the-art DTI techniques. The best results are highlighted using bold font, the symbol •/◦ denotes whether the difference compared with the best approach is statistically significant (•) or not (◦).

Dataset	Method	AUC	AUPR
Enzyme	MOLIERE	0.985	0.897
	BLM-NII	0.966 •	0.628 •
	BRDTI	0.968 •	0.635 •
	HLM	0.966 •	0.832 •
	WNN-GIP	0.945 •	0.708 •
	NetLapRLS	0.959 •	0.784 •
Ion Channel	MOLIERE	0.983	0.912
	BLM-NII	0.960 •	0.626 •
	BRDTI	0.941 •	0.644 •
	HLM	0.980 •	0.867 •
	WNN-GIP	0.947 •	0.663 •
	NetLapRLS	0.966 •	0.827 •
GPCR	MOLIERE	0.952	0.753
	BLM-NII	0.929 •	0.387 •
	BRDTI	0.925 •	0.521 •
	HLM	0.947 ◦	0.686 •
	WNN-GIP	0.928 •	0.513 •
	NetLapRLS	0.910 •	0.580 •
NR	MOLIERE	0.911	0.683
	BLM-NII	0.879 •	0.543 •
	BRDTI	0.868 •	0.397 •
	HLM	0.864 •	0.576 •
	WNN-GIP	0.862 •	0.550 •
	NetLapRLS	0.810 •	0.428 •

<https://doi.org/10.1371/journal.pone.0230726.t004>

Table 5. Top 20 new interactions predicted by our approach, MOLIERE, on the Enzyme dataset. As additional information, we provide whether the interaction is validated (★) or not (–), and if it was predicted by other DTI techniques.

	Drug	Target	Val.	Also predicted by
1	D00542	hsa1571	★	BLM-NII, HLM, NetLapRLS, WNN-GIP, BRDTI
2	D00528	hsa1549	★	BLM-NII, HLM, NetLapRLS, WNN-GIP
3	D00437	hsa1559	★	BLM-NII, HLM, NetLapRLS, WNN-GIP, BRDTI
4	D00097	hsa5743	★	BLM-NII, NetLapRLS, BRDTI
5	D00139	hsa1543	★	BLM-NII, HLM, NetLapRLS, WNN-GIP
6	D03670	hsa1579	–	BLM-NII, HLM, NetLapRLS, WNN-GIP
7	D00574	hsa1589	★	HLM, NetLapRLS, WNN-GIP
8	D03670	hsa9420	–	HLM, NetLapRLS, WNN-GIP
9	D00437	hsa1585	★	BLM-NII, HLM
10	D03670	hsa51004	–	NetLapRLS
11	D03670	hsa51302	–	HLM, NetLapRLS
12	D00410	hsa1543	★	BLM-NII, HLM, NetLapRLS, WNN-GIP
13	D00449	hsa5742	–	HLM
14	D00410	hsa1583	★	HLM, NetLapRLS
15	D00947	hsa4129	–	–
16	D03670	hsa4353	★	–
17	D05458	hsa4128	★	HLM
18	D00043	hsa1990	★	BLM-NII, WNN-GIP
19	D00043	hsa1991	★	BRDTI
20	D00691	hsa5152	–	HLM, NetLapRLS

<https://doi.org/10.1371/journal.pone.0230726.t005>

Table 6. Top 20 new interactions predicted by our approach, MOLIERE, on the GPCR dataset. As additional information, we provide whether the interaction is validated (★) or not (–), and if it was predicted by other DTI techniques.

	Drug	Target	Val.	Also predicted by
1	D00283	hsa:1814	★	BLM-NII, HLM, NetLapRLS, WNN-GIP
2	D04625	hsa:154	★	BLM-NII, HLM, BRDTI
3	D02358	hsa:154	★	BLM-NII, HLM, BRDTI
4	D02614	hsa:154	–	BLM-NII, HLM, BRDTI
5	D00110	hsa:1813	–	–
6	D00765	hsa:1128	★	BRDTI
7	D00760	hsa:1128	–	BRDTI
8	D00604	hsa:147	★	NetLapRLS
9	D02356	hsa:152	–	NetLapRLS
10	D00437	hsa:152	–	HLM, NetLapRLS, WNN-GIP
11	D02147	hsa:153	★	BLM-NII, HLM, BRDTI
12	D00095	hsa:155	★	BLM-NII, HLM
13	D02359	hsa:153	–	BLM-NII, BRDTI
14	D00397	hsa:1131	★	–
15	D00726	hsa:1129	★	–
16	D00255	hsa:152	★	HLM, NetLapRLS
17	D04375	hsa:151	★	–
18	D01103	hsa:1129	★	–
19	D05792	hsa:153	★	BRDTI
20	D01386	hsa:153	★	BRDTI

<https://doi.org/10.1371/journal.pone.0230726.t006>

Table 7. Top 20 new interactions predicted by our approach, MOLIERE, on the IC dataset. As additional information, we provide whether the interaction is validated (★) or not (–), and if it was predicted by other DTI techniques.

	Drug	Target	Val.	Also predicted by
1	D00438	hsa:779	★	BLM-NII, HLM, BRDTI
2	D00294	hsa:10060	–	BLM-NII
3	D00552	hsa:6331	★	HLM, NetLapRLS
4	D00542	hsa:3736	–	–
5	D00649	hsa:8911	★	–
6	D00726	hsa:1138	–	BLM-NII, HLM, NetLapRLS, WNN-GIP
7	D03365	hsa:1137	★	HLM
8	D02098	hsa:8645	–	BLM-NII
9	D00538	hsa:6331	★	HLM, NetLapRLS
10	D00136	hsa:116443	–	–
11	D00349	hsa:773	–	BLM-NII, HLM
12	D00640	hsa:6336	★	HLM, NetLapRLS
13	D02356	hsa:6833	–	NetLapRLS
14	D00438	hsa:781	–	HLM
15	D00547	hsa:2570	–	BLM-NII, WNN-GIP
16	D00528	hsa:1080	–	BLM-NII
17	D00477	hsa:6336	★	BLM-NII, HLM, BRDTI
18	D00775	hsa:2898	★	BLM-NII
19	D00799	hsa:3782	–	BRDTI
20	D00638	hsa:8645	–	–

<https://doi.org/10.1371/journal.pone.0230726.t007>

some of the drug–target pairs the presence or absence of interaction is unknown (or simulated to be unknown in order to evaluate our approach). The task is to predict the likelihood of interaction for these unknown pairs.

At the first glance, the above DTI problem seems to be similar to the problems considered in the recommender systems community. Note, however, that most recommender techniques consider only the interactions (“ratings”) because even a few ratings are thought to be more informative than metadata, such as users’ similarity based on their demographic information [34]. In contrast, drug–drug and target–target similarities play an essential role in DTI.

Bipartite local models

BLM considers DTI as a link prediction problem in bipartite graphs [28]. The vertices in one of the vertex classes correspond to drugs, whereas the vertices in the other vertex class correspond to targets. There is an edge $e_{i,j}$ between drug d_i and target t_j if and only if $m_{i,j} = 1$.

The likelihood of unknown interactions is predicted as follows: we consider an unknown pair $u_{i,j} = (d_i, t_j)$ and calculate the likelihood of interaction as the aggregate of two independent predictions.

The first prediction, called *drug-centric prediction* (Fig 1, left panel), is based on the relations between d_i and the targets. Each target t_k (except t_j) is labelled as “+ 1” or “–1” depending on $m_{i,k}$. Then a model is trained to distinguish “+ 1”-labelled and “–1”-labelled targets. Subsequently, this model is applied to predict the likelihood of interaction for the unknown pair $u_{i,j}$. This first prediction is denoted by $\hat{y}'_{i,j}$. (When describing BLM, in accordance with our data, we assumed that only the information about the presence of an interaction is explicit, and therefore we train the model to distinguish known interacting pairs from pairs with unknown

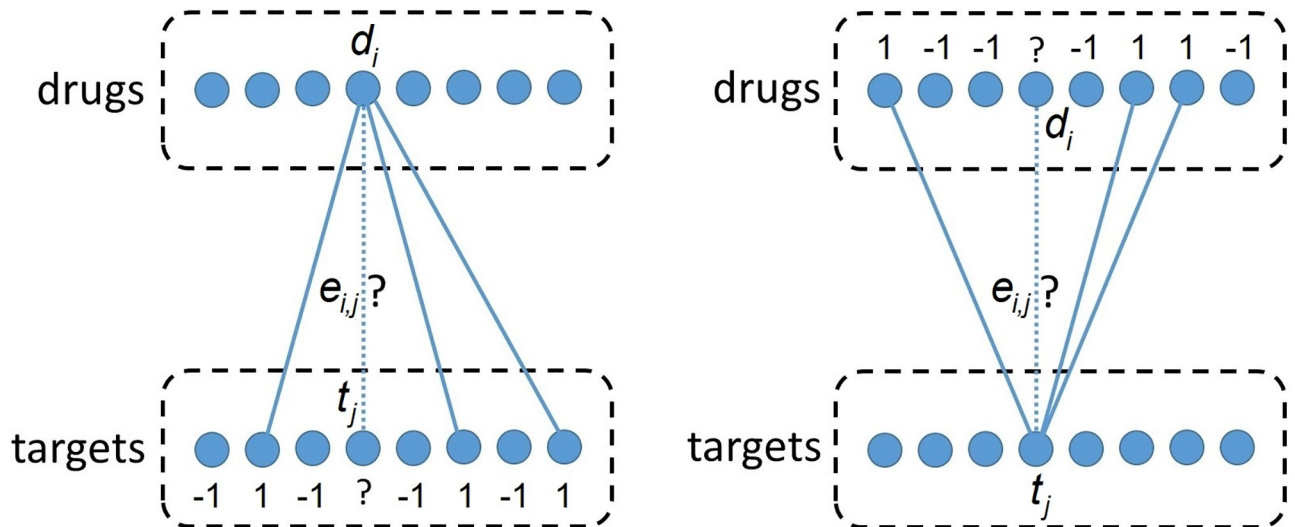


Fig 1. Predictions with BLM. Two predictions are calculated for the likelihood of each unknown interaction, i.e., for the presence of an edge $e_{i,j}$. When calculating the first (second, respectively) prediction, targets (drugs, respectively) are labelled, and a local model is trained using these labels. Subsequently, the local model is used to predict the likelihood of the interaction between d_i and t_j .

<https://doi.org/10.1371/journal.pone.0230726.g001>

status. In contrast, if both *known interacting* and *known non-interacting* drug–target pairs are given, one may train the model using only the known interacting and known non-interacting pairs).

The second prediction, called *target-centric prediction*, $\hat{y}_{i,j}''$, is obtained similarly, but instead of considering the interactions of drug d_i and labelling the targets, the interactions of target t_j are considered and drugs are labelled (Fig 1, right panel). The models that make the first and second predictions are called *drug-centric* and *target-centric local models*.

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

$$\hat{y}_{i,j} = \frac{\hat{y}_{i,j}' + \hat{y}_{i,j}''}{2} \quad (2)$$

Note that instead of averaging, other aggregation functions, such as minimum or maximum are possible as well. According to our observations, our approach achieves most accurate results when the two predictions are averaged. However, the effect of the aggregation function can be considered as minor: when we repeated our experiments reported in Table 4 with *min* and *max* aggregation functions, we observed that our approach consistently outperformed its competitors for all the three aggregation functions. For example, on the GPCR dataset, our approach achieved an AUPR of 0.737 and 0.730 using *min* and *max* respectively, whereas we obtained an AUPR of 0.753 in case of averaging the two predictions.

BLM is a generic framework in which various regressors or classifiers can be used as local models. For example, Bleakley and Yamanishi [28] used support vector machines with a domain-specific kernel, whereas Buza and Peška used a hubness-aware regressor [19]. In our current work, we use BLM with asymmetric loss linear regression which will be described in the next section.

Asymmetric loss models

Local models are the heart of BLM. Next, we propose a new regression technique that we use as a local model.

Given a regression model f_θ where θ is the vector of parameters, f_θ estimates the value of the target y for an instance x as $\hat{y} = f_\theta(x)$. In order to determine the appropriate parameter values θ^* , usually, a loss function $L_D(\theta)$ is minimised:

$$\theta^* = \operatorname{argmin}_\theta L_D(\theta). \quad (3)$$

Note that the actual value of $L_D(\theta)$ depends both on the dataset D and parameters θ . However, once the dataset is fixed, in particular, while the model is being trained using a given training dataset D , the loss can be seen as a function of the parameter vector θ . Therefore, we aim at finding parameters θ^* that minimise the loss. A wide-spread loss function is *mean squared errors*:

$$L_D(\theta) = \frac{1}{|D|} \sum_{(x,y) \in D} (f_\theta(x) - y)^2, \quad (4)$$

where $|D|$ is the number of instances in D .

While the sum of squared errors is popular, we argue that in case of DTI, it is not fully consistent with the underlying chemical reality. In particular, binding energy may be different for various interactions. Consequently, in case of the presence of an interaction ($y = +1$), we should not penalise a model that predicts a score that is higher than $+1$. Similarly, in case of an unknown interaction ($y = -1$), we do not want to penalise a model that predicts a score that is lower than -1 . Therefore, we propose an asymmetric loss function. First, we define the error of the model f_θ for a single prediction $f_\theta(x)$, for instance x with label y as

$$\operatorname{err}(f_\theta, x, y) = \begin{cases} 0 & \text{if } f_\theta(x) > +1 \text{ and } y = +1 \\ 0 & \text{if } f_\theta(x) < -1 \text{ and } y = -1 \\ (f_\theta(x) - y)^2 & \text{otherwise.} \end{cases} \quad (5)$$

We define *mean asymmetric loss (MAL)* as the mean of the above errors for all instances of the dataset D :

$$MAL_D(\theta) = \frac{1}{|D|} \sum_{(x,y) \in D} \operatorname{err}(f_\theta, x, y). \quad (6)$$

The above loss can be minimised with various optimisation techniques ranging from gradient-based methods to more advanced approaches, see e.g. [35]. For simplicity, we decided to use gradient descent. The partial derivative $\frac{\partial MAL_D(\theta)}{\partial \theta}$ of $MAL_D(\theta)$ is:

$$\frac{\partial MAL_D(\theta)}{\partial \theta} = \frac{1}{|D|} \sum_{(x,y) \in D} \frac{\partial \operatorname{err}(f_\theta, x, y)}{\partial \theta}, \quad (7)$$

where

$$\frac{\partial \text{err}(f_\theta, x, y)}{\partial \theta} = \begin{cases} 0 & \text{if } f_\theta(x) > +1 \text{ and } y = +1 \\ 0 & \text{if } f_\theta(x) < -1 \text{ and } y = -1 \\ 2(f_\theta(x) - y) \frac{\partial f_\theta(x)}{\partial \theta} & \text{otherwise.} \end{cases} \tag{8}$$

In case of linear regression where $x = (x_1, \dots, x_k)$, $\theta = \{w_0, w_1, \dots, w_k\}$, and the model is $f_\theta(x) = w_0 + \sum_{i=1}^k w_i x_i$, the partial derivatives of $\text{err}(f_\theta, x, y)$ according to w_i , $1 \leq i \leq k$, are

$$\frac{\partial \text{err}(f_\theta, x, y)}{\partial w_i} = \begin{cases} 0 & \text{if } f_\theta(x) > +1 \text{ and } y = +1 \\ 0 & \text{if } f_\theta(x) < -1 \text{ and } y = -1 \\ 2(f_\theta(x) - y)x_i & \text{otherwise,} \end{cases} \tag{9}$$

while the partial derivative according to w_0 is

$$\frac{\partial \text{err}(f_\theta, x, y)}{\partial w_0} = \begin{cases} 0 & \text{if } f_\theta(x) > +1 \text{ and } y = +1 \\ 0 & \text{if } f_\theta(x) < -1 \text{ and } y = -1 \\ 2(f_\theta(x) - y) & \text{otherwise.} \end{cases} \tag{10}$$

We propose to use stochastic gradient descent to optimise MAL_D . The pseudocode of the resulting asymmetric loss linear regression (ALLR) is shown in Fig 2.

Weighted profile

One of the shortcomings of the BLM approach is that it does not handle the case of new drugs/targets. With *new drug* (or *new target*, respectively), we mean a drug d (target t) that does not have any known interaction in the (training) data. In such cases, BLM labels all targets (drugs) as “-1”, consequently, no reasonable local model can be learned. In order to alleviate this problem, we use the weighted profiles [17] of the most similar drugs/targets to obtain predictions for new drugs/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:

$$\hat{y}'_{ij} = \frac{\sum_{r \in \mathcal{N}^D(d_i)} m_{r,j} \mathcal{S}_{i,r}^D}{\sum_{r \in \mathcal{N}^D(d_i)} \mathcal{S}_{i,r}^D}, \tag{11}$$

where $\mathcal{N}^D(d_i)$ denotes the set of indices of the k_d most similar drugs to d_i (not including d_i itself) based on the drug–drug similarities \mathcal{S}^D .

The intuition behind Eq (11) 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_i and we calculate the weighted average of the known interactions of other drugs with the same target.

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

Algorithm Asymmetric loss linear regression (ALLR)

Require: Training data D , number of epochs e , learning rate η , standard deviation σ

Ensure: Weights w_0, w_1, \dots, w_k

```

1: Initialise weights  $w_0, w_1, \dots, w_k$  from standard normal distribution with
   zero mean and standard deviation  $\sigma$ 
2: for epoch in  $1 \dots e$  do
3:   for each  $(x, y) \in D$  in random order do
4:      $\hat{y} \leftarrow w_0 + \sum_{i=1}^k w_i x_i$ 
5:     if  $(\hat{y} > 1 \text{ and } y = 1)$  or  $(\hat{y} < -1 \text{ and } y = -1)$  then
6:       continue
7:      $w_0 \leftarrow w_0 - \eta 2(\hat{y} - y)$ 
8:     for  $i$  in  $1 \dots k$  do
9:        $w_i \leftarrow w_i - \eta 2(\hat{y} - y)x_i$ 
10:    end for
11: end for
12: return  $w_0, w_1, \dots, w_k$ 

```

Fig 2. Pseudocode of asymmetric loss linear regression (ALLR).

<https://doi.org/10.1371/journal.pone.0230726.g002>

between d_i and t_j as follows:

$$\hat{y}_{ij}'' = \frac{\sum_{r \in \mathcal{N}^T(t_j)} m_{i,r} \mathcal{S}_{j,r}^T}{\sum_{r \in \mathcal{N}^T(t_j)} \mathcal{S}_{j,r}^T}, \quad (12)$$

where $\mathcal{N}^T(t_j)$ denotes the set of indices of the k_t most similar targets to t_j (not including t_j itself) based on the target-target similarities \mathcal{S}^T .

Although the weighted profile approach is more general than BLM, in the sense that it can be used for new drugs/targets as well, the predictions of the weighted profile approach are usually less accurate than the predictions of BLM. Therefore, we use the weighted profile approach instead of BLM *only* in case of *new* drugs and targets.

Our approach

We summarise our approach as follows. We use BLM for drug–target interaction prediction with the proposed asymmetric loss linear regression as local model in cases when the corresponding drug (target) has at least one known interaction and therefore the local model has at least one positive training instance. When initialising the parameters of ALLR, we use $\sigma = 10^{-8}$. We train each ALLR model with a learning rate $\eta = 10^{-3}$ for $e = 100$ epochs. According to our observations, ALLR is robust in the sense that the aforementioned settings allowed

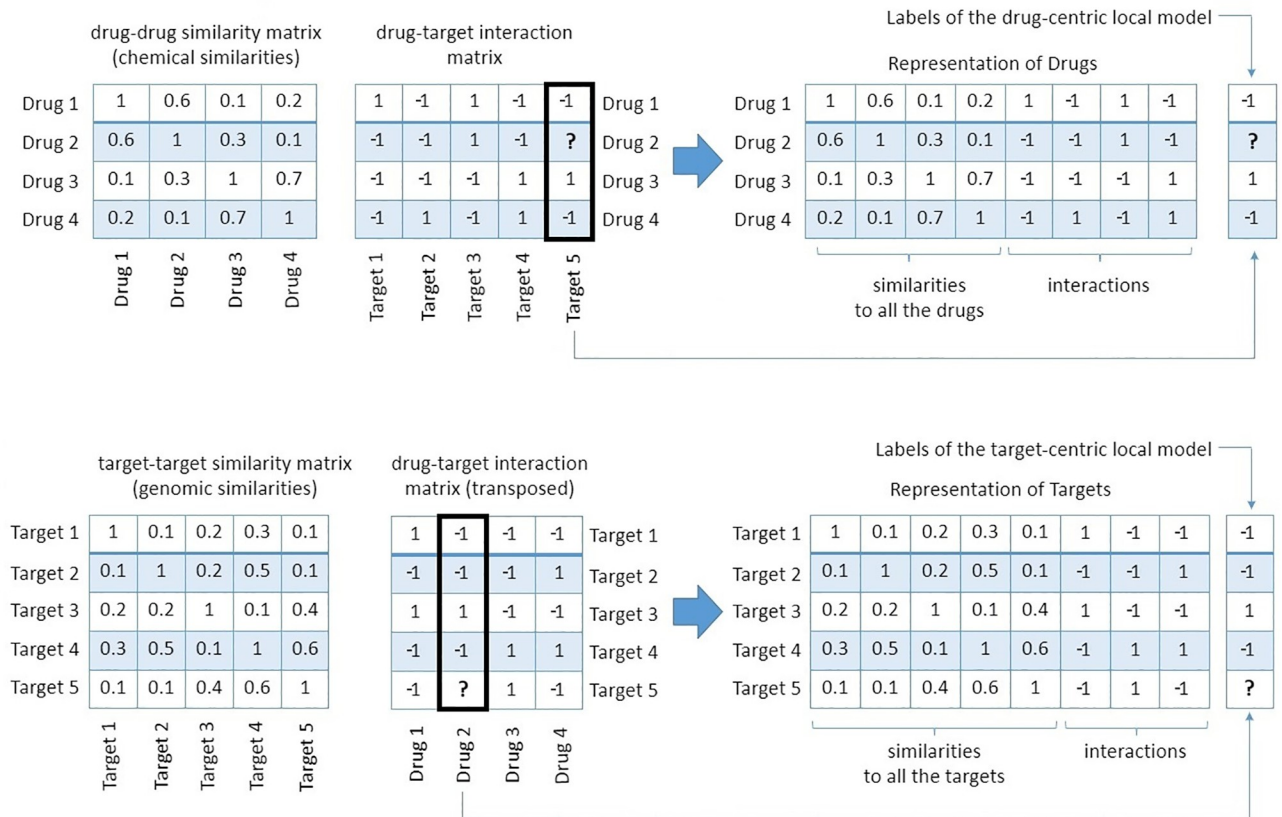


Fig 3. Representation of drugs and targets and labels of local models. In this example, the prediction is made for the interaction denoted by the question mark. Similarities with all drugs (targets, respectively) and interactions with all the targets (drugs), except the interactions with the target (drug) corresponding to the question mark, are used as features. The interactions with the target (drug) corresponding to the question mark are used as labels of the local models. Tables on the right represent the data used by the local model, i.e., ALLR in our case.

<https://doi.org/10.1371/journal.pone.0230726.g003>

ALLR to converge to a model that outperformed other DTI techniques on all the examined datasets (see Table 4).

While predicting the interaction score between drug d and target t with ALLR, we represent each drug (target) as a vector of its similarities to all the drugs (targets) and its interactions, except the interactions with t (or d respectively), because the interactions with d (t) serve as labels for the local models, see Fig 3 for an illustration.

In case of new drugs (targets), we predict the likelihood of interactions using the weighted profile approach with $k_d = k_t = 5$.

Results

Comparison with baselines

As our approach, MOLIERE, is based on BLM, and uses weighted profile (WP) in case of new interactions, first, we compared the performance of MOLIERE to that of the original BLM and WP according to the widely used leave-one-interaction-out cross-validation protocol, see e.g. [28, 30, 31].

The predictions were evaluated both in terms of Area Under ROC Curve (AUC) and Area Under Precision-Recall Curve (AUPR). Table 2 shows that MOLIERE clearly outperforms both BLM and WP, both in terms of AUC and AUPR.

As the proposed asymmetric loss linear regression is a key component of MOLIERE, we examined the performance in case of alternative regression techniques, in particular we examined the cases when we use (a) standard linear regression and (b) logistic regression instead of ALLR. As expected, the proposed asymmetric loss linear regression indeed improves the quality of predictions, see [Table 3](#).

Comparison with recent DTI techniques

We compared MOLIERE with state-of-the-art drug-target interaction prediction techniques: two recent versions of BLM and three further prominent DTI approaches. The former include BLM with neighbour-based interaction-profile inferring (BLM-NII) [31] and hubness-aware regressors as local models (HLM) [19], while the later refer to net Laplacian regularized least squares (NetLapRLS) [29], a combination of weighted nearest neighbour and Gaussian interaction profile kernels (WNN-GIP) [36], and Bayesian Ranking for Drug-Target Interaction Prediction (BRDTI) [20].

Pahikkala et al. [37] pointed out that leave-one-out cross-validation may lead to overoptimistic results. Therefore, in this section, we used the interaction-based 5×5 -fold cross-validation protocol, i.e., 5-fold cross-validation is repeated 5-times with different initial data splits. In each of the 5×5 rounds of the cross-validation, one fifth of the drug-target pairs were in the test data and AUC and AUPR values were calculated. The reported results are averaged values. In order to judge if the observed differences are statistically significant, we used paired t -test with significance threshold of $p = 0.01$.

Essential hyperparameters of BLM-NII, HLM, WNN-GIP, NetLapRLS and BRDTI were learned via grid-search in internal 5-fold cross-validation on the training data. For other hyperparameters that are not expected to have major impact on the results, we used default values according to the publication in which the approach was published.

In particular, for BLM-NII, as proposed by Mei et al. [31], the *max* function was used to generate final predictions and the weight α for the combination of structural and collaborative similarities was learned from $\{0.0, 0.1, \dots, 1.0\}$.

In case of HLM, according to [19], we performed experiments with $N = 25$ base prediction models, while the number of nearest neighbours for the local model ECKNN and the number of selected features, were learned from $\{3, 5, 7\}$ and $\{10, 20, 50\}$ respectively.

The decay hyperparameter of WNN-GIP was learned from $\{0.1, 0.2, \dots, 1.0\}$ and the weight α for combination of structural and collaborative similarities was learned from $\{0.0, 0.1, \dots, 1.0\}$.

The hyperparameters of NetLapRLS ($\beta = \beta_{drug} = \beta_{target}$ and $\gamma = \gamma_{drug} = \gamma_{target}$), were learned from $\{10^{-6}, 10^{-5}, \dots, 10^2\}$.

The content regularisation λ_c of BRDTI was learned from $\{0.1, 0.5, 0.9, 1.5\}$. The number of latent factors f , number of iterations, global regularisation λ_g and initial learning rate η were set to 100, 50, 0.01 and 0.1 respectively.

The results are shown in [Table 4](#) and [Fig 4](#) which show the precision-recall curves for MOLIERE and its competitors. Our approach, MOLIERE, outperforms all the examined competitors in case of Enzyme, Ion Channel, GPCR and NR datasets, both in terms of AUC and AUPR. In the vast majority of the cases, the difference is statistically significant.

The results indicate that our approach, MOLIERE, is the overall best-performing approach out of the examined DTI techniques.

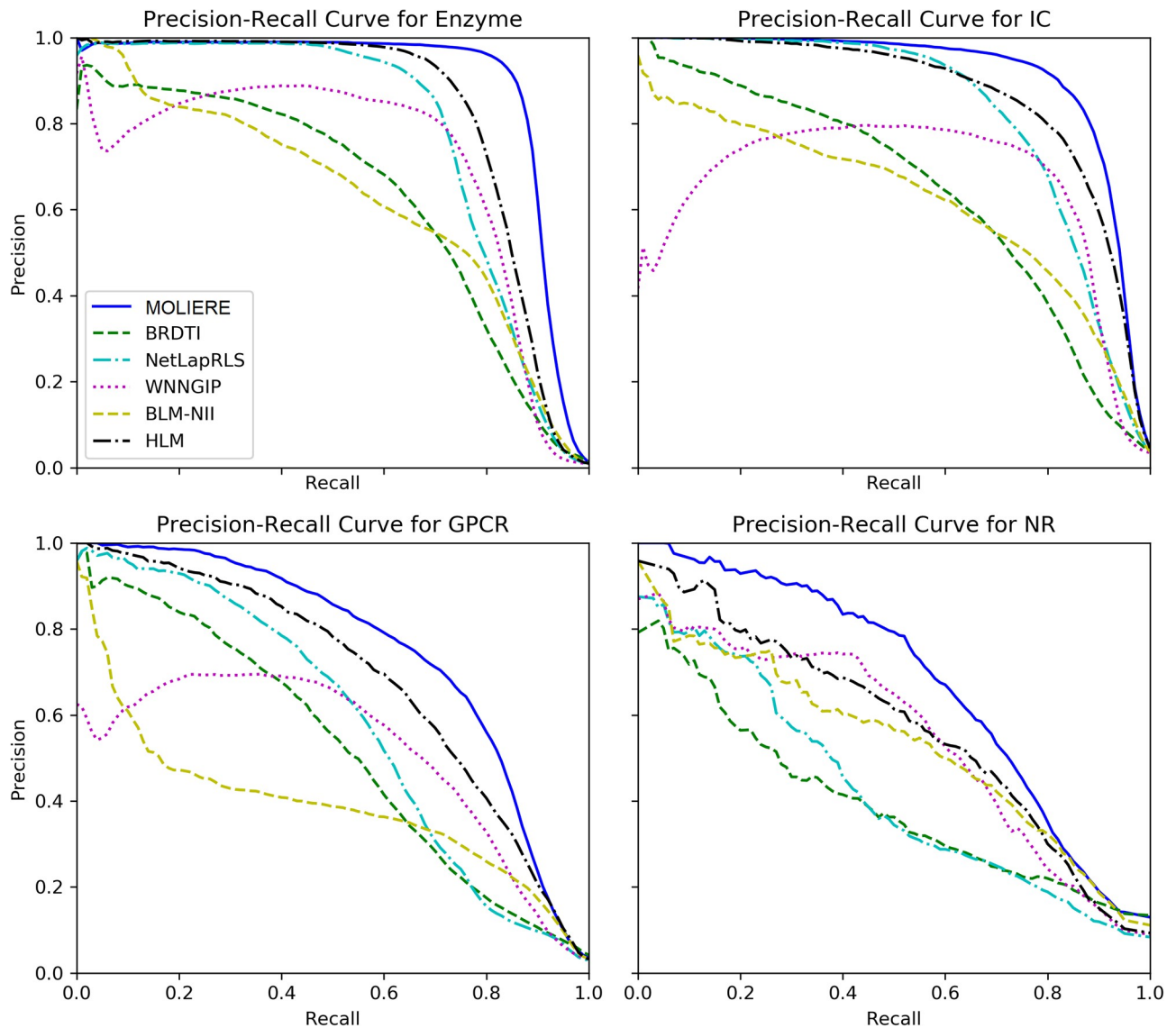


Fig 4. Precision-recall (PR) curves (averaged over the 5×5 folds of the cross-validation). As one can see, our approach, MOLIERE consistently outperforms its competitors: the PR-curve of MOLIERE is consistently above the PR-curves of its competitors.

<https://doi.org/10.1371/journal.pone.0230726.g004>

Prediction of new interactions

In order to demonstrate that our approach is able to predict new interactions, we followed the same protocol as in [19], i.e., we trained our approach, MOLIERE, as well as its competitors, BLM-NII, HLM, NetLapRLS and WNN-GIP using all the interactions of the original datasets. As mentioned before, these datasets have been extracted roughly a decade ago and several additional interactions have been validated meanwhile. Our experiment aims to check whether these recently validated interactions can be predicted based on the original interactions.

In particular, we considered those drug-target pairs that have unknown interaction status in the original datasets. We ranked these pairs according to their predicted interaction scores. For simplicity, we use the term *predicted* interaction for the top-ranked 20 drug-target pairs.

We say that a predicted interaction is *validated* if it is included in the current version of KEGG [38], DrugBank [39] or Matador [40].

The results are shown in Tables 5–7 for Enzyme, GPCR and IC datasets (for drug and target identifiers, see: <http://www.kegg.jp/>). As one can see, many of the predicted interactions are validated. We point out that some of the validated interactions were only predicted by our approach, especially in case of the GPCR dataset.

MOLIERE for drug repurposing

In order to illustrate how our predictions may contribute to drug repurposing, we discuss some of the predicted interactions in more details.

First, we consider Diazoxide (KEGG ID: D00294) and “adenosine triphosphate binding cassette, sub-family C member 9” (ABCC9), also known as “sulfonylurea receptor 2” (SUR2) (KEGG ID: hsa:10060), i.e., the second predicted interaction listed in Table 7. According to KEGG, Diazoxide is an adenosine triphosphate (ATP) sensitive potassium channel opener. It opens potassium channel in beta cells of the pancreas and causes insulin secretion inhibition thus elevating blood sugar level. It is used in insulinoma [41] and congenital hyperinsulinism [42]. Diazoxide treatment can cause pulmonary hypertension and relax smooth muscle [43]. ABCC9 gene encodes a protein that is a subunit of an ATP sensitive potassium channel (ATP-binding cassette transporter) [44]. It is expressed in skeletal and heart muscle and in smooth muscles of the vasculature [44]. Mutation of the gene can cause dilated cardiomyopathy type 10 [45], and reduced cardiac stress adaptation [44]. ABCC9 knock-out mice showed elevated blood pressure and coronary artery vasospasm [46]. Other mutations of the ABCC9 gene (<https://omim.org/entry/601439>) can cause familial atrial fibrillation type 12 and hypertrichotic osteochondrodysplasia (Cantú syndrome) [47]. Diazoxide is not used in the treatment of these diseases. However, as Diazoxide can open the ATP sensitive potassium channel, it would be worth to examine the possible usage of Diazoxide in some ABCC9 gene defects where the transporter still can be activated to some extent.

Next, we consider the predicted interaction between Isradipine (KEGG ID: D00349) and “Calcium Voltage-Gated Channel Subunit Alpha1 A” (CACNA1A), also known as “spinocerebellar ataxia type 6” (SCA6) (KEGG ID: hsa:773), listed in the 11th line of Table 7. According to KEGG, Isradipine is an L type dihydropyridine calcium channel blocker that is used in hypertension. CACNA1A gene encodes the alpha-1A subunit of P/Q type voltage-dependent calcium channel. Mutations of this gene can cause spinocerebellar ataxia, early infantile epileptic encephalopathy, episodic ataxias, hemiplegic migraine and hemiconvulsion-hemiplegia-epilepsy syndrome. Some mutations of the gene increases the density of functional channels and their open probabilities in familial hemiplegic migraine [48]. Since our method takes into consideration the similarity of the two different calcium channels, it may be worth to try Isradipine inhibition of the P/Q type voltage-dependent calcium channel in experimental settings.

Finally, we consider the predicted interaction between Caffeine (KEGG ID: D00528) and Cystic fibrosis transmembrane conductance regulator (CFTR, KEGG ID: hsa:1080), listed in the 16th line of Table 7. Caffeine is a central nervous system stimulant, adenosine receptor antagonist and phosphodiesterase inhibitor (1). CFTR is a chloride channel that conducts chloride ions in lung, pancreas, liver, digestive tract and reproductive tract epithelial cell membranes. According to KEGG, gene mutations can cause cystic fibrosis (CF), hereditary pancreatitis and congenital bilateral absence of vas deferens. In rats caffeine intake increased CFTR chloride secretion in intestine [49]. Although, caffeine consumption is basically not recommended for CF patients, if some of the patients actually drink coffee, it would be interesting to compare their disease status with other CF patients not drinking coffee. Such a survey should

be carefully designed in order to avoid biases. For example, the number of patients involved in the study should be large enough, one should take into account that people who have more severe disease may pay more attention to the health and lifestyle suggestions, while the type of mutations is also important.

Conclusion

In this paper, we focused on drug–target interaction prediction and proposed a new method, MOLIERE for this task. Despite the fact that MOLIERE is a relatively simple approach, experiments on real-world datasets show that MOLIERE outperforms state-of-the-art DTI methods. By discussing some of the predictions in detail, we showed how our approach may lead to medically relevant hypothesis and support drug repositioning.

As mentioned, the DTI problem shares inherent characteristics with recommender systems tasks, therefore, we expect that MOLIERE will be adapted for recommendation tasks in the future. Furthermore, we point out that the proposed framework of asymmetric loss models is not limited to drug–target interaction prediction, but it may be useful in other cases where the class label is originally continuous (due to the underlying physical, chemical, biological phenomena), but it has been transformed to a binary label.

As for the limitations of our study, we note that our approach is not designed to predict interactions in case of new drugs/targets, i.e., for drugs/targets for which not even one interaction is known. In our current work, we assumed that only *few* new drugs/targets are considered, and we used the simple weighted profile approach in this case. Therefore, further methodical enhancements are required, if predictions are desired for new drugs/targets.

Author Contributions

Methodology: Krisztian Buza.

Software: Krisztian Buza, Ladislav Peška.

Writing – original draft: Krisztian Buza, Ladislav Peška, Júlia Koller.

References

1. Ding Y, Tang J, Guo F. Identification of drug-side effect association via multiple information integration with centered kernel alignment. *Neurocomputing*. 2019; 325:211–224. <https://doi.org/10.1016/j.neucom.2018.10.028>
2. Liu LJ, Lu L, Zhong HJ, He B, Kwong DW, Ma DL, et al. An iridium (III) complex inhibits JMJD2 activities and acts as a potential epigenetic modulator. *Journal of medicinal chemistry*. 2015; 58(16):6697–6703. <https://doi.org/10.1021/acs.jmedchem.5b00375> PMID: 26225543
3. Kang TS, Mao Z, Ng CT, Wang M, Wang W, Wang C, et al. Identification of an iridium (III)-based inhibitor of tumor necrosis factor- α . *Journal of medicinal chemistry*. 2016; 59(8):4026–4031. <https://doi.org/10.1021/acs.jmedchem.6b00112> PMID: 27054262
4. Liu LJ, He B, Miles JA, Wang W, Mao Z, Che WI, et al. Inhibition of the p53/hDM2 protein-protein interaction by cyclometallated iridium (III) compounds. *Oncotarget*. 2016; 7(12):13965. <https://doi.org/10.18632/oncotarget.7369> PMID: 26883110
5. Yang C, Wang W, Li GD, Zhong HJ, Dong ZZ, Wong CY, et al. Anticancer osmium complex inhibitors of the HIF-1 α and p300 protein-protein interaction. *Scientific reports*. 2017; 7:42860. <https://doi.org/10.1038/srep42860> PMID: 28225008
6. Ullrich K, Mack J, Welke P. Ligand Affinity Prediction with Multi-pattern Kernels. In: *International Conference on Discovery Science*. Springer; 2016. p. 474–489.
7. Morgan S, Grootendorst P, Lexchin J, Cunningham C, Greyson D. The cost of drug development: a systematic review. *Health policy*. 2011; 100(1):4–17. <https://doi.org/10.1016/j.healthpol.2010.12.002> PMID: 21256615

8. Zhang J, Li C, Lin Y, Shao Y, Li S. Computational drug repositioning using collaborative filtering via multi-source fusion. *Expert Systems with Applications*. 2017; 84:281–289. <https://doi.org/10.1016/j.eswa.2017.05.004>
9. Cheng AC, Coleman RG, Smyth KT, Cao Q, Souillard P, Caffrey DR, et al. Structure-based maximal affinity model predicts small-molecule druggability. *Nature biotechnology*. 2007; 25(1):71–75. <https://doi.org/10.1038/nbt1273> PMID: 17211405
10. Pérot S, Regad L, Reynès C, Spérandio O, Miteva MA, Villoutreix BO, et al. Insights into an original pocket-ligand pair classification: a promising tool for ligand profile prediction. *PloS one*. 2013; 8(6): e63730. <https://doi.org/10.1371/journal.pone.0063730> PMID: 23840299
11. Cellier P, Charnois T, Plantevit M. Sequential patterns to discover and characterise biological relations. In: *International Conference on Intelligent Text Processing and Computational Linguistics*. Springer; 2010. p. 537–548.
12. Fayruzov T, De Cock M, Cornelis C, Hoste V. Linguistic feature analysis for protein interaction extraction. *BMC Bioinformatics*. 2009; 10(1):374. <https://doi.org/10.1186/1471-2105-10-374> PMID: 19909518
13. Davis J, Santos Costa V, Ray S, Page D. An integrated approach to feature invention and model construction for drug activity prediction. In: *Proceedings of the 24th International Conference on Machine Learning*; 2007. p. 217–224.
14. Fan X, Hong Y, Liu X, Zhang Y, Xie M. Neighborhood Constraint Matrix Completion for Drug-Target Interaction Prediction. In: *Pacific-Asia Conference on Knowledge Discovery and Data Mining*. Springer; 2018. p. 348–360.
15. Jamali AA, Ferdousi R, Razzaghi S, Li J, Safdari R, Ebrahimie E. DrugMiner: comparative analysis of machine learning algorithms for prediction of potential druggable proteins. *Drug discovery today*. 2016; 21(5):718–724. <https://doi.org/10.1016/j.drudis.2016.01.007> PMID: 26821132
16. Lan W, Wang J, Li M, Liu J, Li Y, Wu FX, et al. Predicting drug–target interaction using positive-unlabeled learning. *Neurocomputing*. 2016; 206:50–57. <https://doi.org/10.1016/j.neucom.2016.03.080>
17. Yamanishi Y, Araki M, Gutteridge A, Honda W, Kanehisa M. Prediction of drug–target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics*. 2008; 24(13):i232–i240. <https://doi.org/10.1093/bioinformatics/btn162> PMID: 18586719
18. Buza K, Peška L. ALADIN: A New Approach for Drug–Target Interaction Prediction. *Lecture Notes in Computer Science*. 2017; 10535:322–337. https://doi.org/10.1007/978-3-319-71246-8_20
19. Buza K, Peška L. Drug–target interaction prediction with Bipartite Local Models and hubness-aware regression. *Neurocomputing*. 2017; 260:284–293. <https://doi.org/10.1016/j.neucom.2017.04.055>
20. Peska L, Buza K, Koller J. Drug-target interaction prediction: A Bayesian ranking approach. *Computer methods and programs in biomedicine*. 2017; 152:15–21. <https://doi.org/10.1016/j.cmpb.2017.09.003> PMID: 29054256
21. Bolgar B, Antal P. Bayesian Matrix Factorization with Non-Random Missing Data using Informative Gaussian Process Priors and Soft Evidences. *Journal of Machine Learning Research*. 2016; 52:25–36.
22. Gönen M. Predicting drug–target interactions from chemical and genomic kernels using Bayesian matrix factorization. *Bioinformatics*. 2012; 28(18):2304–2310. <https://doi.org/10.1093/bioinformatics/bts360> PMID: 22730431
23. Zheng X, Ding H, Mamitsuka H, Zhu S. Collaborative matrix factorization with multiple similarities for predicting drug-target interactions. In: *19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; 2013. p. 1025–1033.
24. Wang Y, Zeng J. Predicting drug-target interactions using restricted Boltzmann machines. *Bioinformatics*. 2013; 29(13):i126–i134. <https://doi.org/10.1093/bioinformatics/btt234> PMID: 23812976
25. Chen X, Liu MX, Yan GY. Drug–target interaction prediction by random walk on the heterogeneous network. *Molecular BioSystems*. 2012; 8(7):1970–1978. <https://doi.org/10.1039/c2mb00002d> PMID: 22538619
26. Cheng F, Liu C, Jiang J, Lu W, Li W, Liu G, et al. Prediction of drug-target interactions and drug repositioning via network-based inference. *PLoS Comput Biol*. 2012; 8(5):e1002503. <https://doi.org/10.1371/journal.pcbi.1002503> PMID: 22589709
27. Sönströmd C, Johansson U, Norinder U, Boström H. Comprehensible Models for Predicting Molecular Interaction with Heart-Regulating Genes. In: *7th IEEE International Conference on Machine Learning and Applications*; 2008. p. 559–564.
28. Bleakley K, Yamanishi Y. Supervised prediction of drug–target interactions using bipartite local models. *Bioinformatics*. 2009; 25(18):2397–2403. <https://doi.org/10.1093/bioinformatics/btp433> PMID: 19605421

29. Xia Z, Wu LY, Zhou X, Wong ST. Semi-supervised drug-protein interaction prediction from heterogeneous biological spaces. *BMC Systems Biology*. 2010; 4(Suppl 2):S6. <https://doi.org/10.1186/1752-0509-4-S2-S6> PMID: 20840733
30. van Laarhoven T, Nabuurs SB, Marchiori E. Gaussian interaction profile kernels for predicting drug-target interaction. *Bioinformatics*. 2011; 27(21):3036–3043. <https://doi.org/10.1093/bioinformatics/btr500> PMID: 21893517
31. Mei JP, Kwok CK, Yang P, Li XL, Zheng J. Drug-target interaction prediction by learning from local information and neighbors. *Bioinformatics*. 2013; 29(2):238–245. <https://doi.org/10.1093/bioinformatics/bts670> PMID: 23162055
32. Davis MI, Hunt JP, Herrgard S, Ciceri P, Wodicka LM, Pallares G, et al. Comprehensive analysis of kinase inhibitor selectivity. *Nature Biotechnology*. 2011; 29(11):1046–1051. <https://doi.org/10.1038/nbt.1990> PMID: 22037378
33. Hattori M, Okuno Y, Goto S, Kanehisa M. 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*. 2003; 125(39):11853–11865. <https://doi.org/10.1021/ja036030u> PMID: 14505407
34. Pilászy I, Tikk D. Recommending new movies: even a few ratings are more valuable than metadata. In: 3rd ACM Conf. on Recommender Systems; 2009. p. 93–100.
35. Suciú M, Lung RI, Gaskó N. Noisy extremal optimization. *Soft Computing*. 2017; 21(5):1253–1270. <https://doi.org/10.1007/s00500-015-1858-3>
36. van Laarhoven Twan and Marchiori Elena. Predicting drug-target interactions for new drug compounds using a weighted nearest neighbor profile. *PloS one*. 2013; 8(6):e66952. <https://doi.org/10.1371/journal.pone.0066952> PMID: 23840562
37. Pahikkala T, Airola A, Pietilä S, Shakyawar S, Sz wajda A, Tang J, et al. Toward more realistic drug-target interaction predictions. *Briefings in Bioinformatics*. 2015; 16(2):325–337. <https://doi.org/10.1093/bib/bbu010> PMID: 24723570
38. Kanehisa M, Goto S, Hattori M, Aoki-Kinoshita KF, Itoh M, Kawashima S, et al. From genomics to chemical genomics: new developments in KEGG. *Nucleic acids research*. 2006; 34(Suppl 1):D354–D357. <https://doi.org/10.1093/nar/gkj102> PMID: 16381885
39. Wishart David S and Knox Craig and Guo An Chi and Shrivastava Savita and Hassanali Murtaza and Stothard et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic acids research*. 2006; 34(Suppl 1):D668–D672. <https://doi.org/10.1093/nar/gkj067> PMID: 16381955
40. Günther Stefan and Kuhn Michael and Dunkel Mathias and Campillos Monica and Senger Christian and Petsalaki et al. SuperTarget and Matador: resources for exploring drug-target relationships. *Nucleic acids research*. 2008; 36(Suppl 1):D919–D922. <https://doi.org/10.1093/nar/gkm862> PMID: 17942422
41. Ferrer-Garcia J, Gonzalez-Cruz VI, Navas-DeSolis S, Civera-Andres M, Morillas-Arino C, Merchante-Alfaro A, et al. Management of malignant insulinoma. *Clinical and Translational Oncology*. 2013; 15(9):725–731. <https://doi.org/10.1007/s12094-012-0996-7> PMID: 23460559
42. Banerjee I, Salomon-Estebanez M, Shah P, Nicholson J, Cosgrove K, Dunne M. Therapies and outcomes of congenital hyperinsulinism-induced hypoglycaemia. *Diabetic Medicine*. 2018;.
43. Timlin MR, Black AB, Delaney HM, Matos RI, Percival CS. Development of Pulmonary Hypertension During Treatment with Diazoxide: A Case Series and Literature Review. *Pediatric cardiology*. 2017; 38(6):1247–1250. <https://doi.org/10.1007/s00246-017-1652-3> PMID: 28642988
44. Schumacher T, Benndorf RA. ABC transport proteins in cardiovascular disease—A brief summary. *Molecules*. 2017; 22(4):589. <https://doi.org/10.3390/molecules22040589>
45. Bienengraeber M, Olson TM, Selivanov VA, Kathmann EC, O’Cochlain F, Gao F, et al. ABCC9 mutations identified in human dilated cardiomyopathy disrupt catalytic K ATP channel gating. *Nature genetics*. 2004; 36(4):382. <https://doi.org/10.1038/ng1329> PMID: 15034580
46. Chutkow WA, Pu J, Wheeler MT, Wada T, Makielski JC, Burant CF, et al. Episodic coronary artery vasospasm and hypertension develop in the absence of Sur2 K ATP channels. *The Journal of clinical investigation*. 2002; 110(2):203–208. <https://doi.org/10.1172/JCI15672> PMID: 12122112
47. Harakalova M, van Harssel JJ, Terhal PA, van Lieshout S, Duran K, Renkens I, et al. Dominant missense mutations in ABCC9 cause Cantu syndrome. *Nature genetics*. 2012; 44(7):793. <https://doi.org/10.1038/ng.2324> PMID: 22610116
48. Hans M, Luvisetto S, Williams ME, Spagnolo M, Urrutia A, Tottene A, et al. Functional consequences of mutations in the human $\alpha 1A$ calcium channel subunit linked to familial hemiplegic migraine. *Journal of*

Neuroscience. 1999; 19(5):1610–1619. <https://doi.org/10.1523/JNEUROSCI.19-05-01610.1999> PMID: [10024348](https://pubmed.ncbi.nlm.nih.gov/10024348/)

49. Wei X, Lu Z, Yang T, Gao P, Chen S, Liu D, et al. Stimulation of Intestinal Cl-Secretion Through CFTR by Caffeine Intake in Salt-Sensitive Hypertensive Rats. *Kidney and Blood Pressure Research*. 2018; 43(2):439–448. <https://doi.org/10.1159/000488256> PMID: [29558753](https://pubmed.ncbi.nlm.nih.gov/29558753/)