# Logistic Matrix Factorization for Drug-Target Interaction Prediction on STITCH database

**Fen Pei**
Department of Computational & Systems Biology
University of Pittsburgh
fep7@pitt.edu

**Chi Zhang**
Department of Computer Science
University of Pittsburgh
chz54@pitt.edu

## Abstract

As the concept "polypharmacology" becomes popular, one crucial step in drug discovery is to identify drug-target interactions (DTI) for understanding drug Mechanism of Action (MOA) and drug repurposing. However, only a small portion of drug-target interactions have been experimentally validated so far, and experimentally validation process is expensive and time costly. Therefore, to improve the drug discovery efficiency, there' s a great need to develop accurate computational DTI prediction methods. STITCH database ('Search Tool for Interacting Chemicals', http://stitch.embl.de ) is an integrated database for protein-chemical interaction networks. In the 2014 release, it contains 531,716 chemicals (including different stereoisomer), 9,337 proteins, and 1,351,208 experimentally verified chemical-protein interactions for human subset. To date, no publically available model for DTI prediction on STITCH database. In this project, we performed DTI prediction based on STITCH version 4 human subset, employing Logistic Matrix Factorization (LMF) machine learning methods, and compared it with baseline method Probabilistic Matrix Factorization (PMF). The performances of PMF and LMF were evaluated by 10-fold cross-validation. The DTI prediction model helped guide further development of DTI prediction method on even larger database, and the prediction results could be used for drug-repurposing, drug combinations and disease mechanism inferring studies.

## 1 Introduction

Recently the trend of drug discovery and development process has changed from "one-gene, one-target and one-mechanism" strategy to a system level of understanding of drug-target interactions (DTI) and networks [1]. DTIs are quite promiscuous, a drug interacts with multiple targets and a target could have multiple ligands. In addition, multiple targets are often involved in the same disease[2], figuring out DTIs helps better understand drug "polypharmacology" and disease heterogeneity, as well as drug repurposing and side effects analysis. There are estimated about 6000–8000 targets in the human genome that have pharmacological interest [3], but 800-900 of these targets has been involved in approved drugs so far [4]. A huge number of potential drug targets remain to be validated, besides, most of chemicals only annotated with limited number of targets, it is highly possible they interact with other proteins. Therefore, appropriate and powerful computational prediction of DTIs is an important step to guide experimental validation and reduce cost.

Except the traditional docking simulation approaches and ligand-based approaches for DTIs prediction, significant improvements for DTIs prediction have been obtained by many more effective and accurate computational models. These methods are recommendation techniques and mainly fall into two categories: network based methods and machine learning based methods.

## Network based methods

The basic idea of network based methods is calculate the similarity of two drugs or two targets, then the targets of one drug can be predicted by the targets of its similar drugs based on some algorithms. Yang et al. [5] first developed a computational algorithm to infer potential drug targets by systematically analyzing the transformation between the disease state and the desired state in a disease network. Chen et al. [6] developed an effective model of Network-based Random Walk with Restart on the Heterogeneous network (NRWRH) to predict potential drug–target interactions by implementing random walk on the heterogeneous network. Where the random walk is implemented on the heterogeneous networks of drug-target interactions network, drug chemical structure similarity network and target protein sequence similarity network. Cheng et al. [7] developed three supervised inference models to predict drug–target interactions, namely, drug-based similarity inference (DBSI), target-based similarity inference (TBSI) and network-based inference (NBI)). For the DBSI, the score to predict the association between drug ($d_i$) and target ($t_j$) is defined as follows:

$$V_{ij}^{D} = \frac{\sum_{l=1,l\neq i}^{n} S_c(d_i, d_l) a_{lj}}{\sum_{l=1,l\neq i}^{n} S_c(d_i, d_l)}$$

here $S_c(d_i, d_l)$ is the 2D chemical similarity between drug $d_i$ and drug $d_l$, and $a_{lj} = 1$ if drug $d_l$ interacts with target $t_j$. For TBSI, the correspond score is calculated as follows:

$$V_{ij}^{T} = \frac{\sum_{l=1,l\neq i}^{n} S_g(t_j, t_l) a_{il}}{\sum_{l=1,l\neq i}^{n} S_g(t_j, t_l)}$$

here $S_g(t_j, t_l)$ is the genomic sequence similarity between target $t_j$ and target $t_l$. For the NBI, the final score $f_i$ of drug $d_i$ after two-step diffusion is defined as follows:

$$f_i = \sum_{l=1}^{m} \frac{a_{il}}{k(t_1)} \sum_{o=1}^{n} \frac{a_{ol} f_0(o)}{k(d_o)}$$

here $f_0(o) = a_{oj}, o \in \{1, 2, 3, \ldots, n\}, k(d_o) = \sum_{s=1}^{m} a_{os}$ and $k(t_1) = \sum_{s=1}^{n} a_{sl}$ represent the initial resource of drug $d_o$, the number of the targets that interact with $d_o$ and the number of drugs that interact with $t_1$, respectively. Cross validation results demonstrated that NBI, without considering drug chemical structure and target protein sequence information, shown best performance among these three methods.

## Machine learning based methods

Since Yamanishi et al. [8] proposed the first bipartite graph learning method in 2008, a number of machine learning based methods have been developed to predict DTIs. In Yamanishi's method, DTIs are inferred by a kernel regression model, where each drug is denoted by a vector encoding chemical structure information, and each target is denoted by a vector encoding sequence information. Then learn two models $f_c$ and $f_g$ based on a variant of kernel regression model, which were used to map a new compound or a new target on a unified pharmacological feature space. Therefore, this model is able to predict interactions between any new pairs of drugs and targets based on compound structure and protein sequence information. Bleakley and Yamanishi [9] developed a new Bipartite Local Model (BLM) by transforming edge prediction problems into binary classification problems in 2009. Yamanishi further develop a correlation-based model by introducing pharmacological data in 2010 [10]. Wang and Zeng [11] proposed the first learning method to predict not only the binary interactions between drugs and targets but also different types of interactions (i.e. different types of drug actions) in the framework of restricted Boltzmann machines (RBM) based on a multidimensional drug–target network.

Instead of using supervised learning methods, Xia et al. [12] developed a semi-supervised learning method, the manifold regularization method uses labeled and unlabeled information instead of using labeled data alone, which shown better prediction results for DTIs. In the studies mentioned above, these models all treated DTI prediction as binary classification problem, however, in real-life, drug-target interaction is not simple on/off binary relationship. Pahikkala et al. [13] pointed out that more realistic prediction should be obtained by performing regression prediction. They proposed a Kronecker RLS (Kron-RLS) model based on quantitative bioactivity data for kinase inhibitors (such as dissociation constant or inhibition constant) to predict potential drug-target interactions. Recently published multiple kernel learning algorithm (KronRLS-MKL) [14] learned multiple kernel with different weights for drug-drug similarity and target-target similarity, which took full advantage of heterogeneous information sources to identify new interactions.

Cobanoglu et al. [15] developed a Probabilistic Matrix Factorization (PMF)-based active learning methodology to predict drug–target interactions, where the properties of drugs and targets are represented by drug-specific and target-specific latent variables, does not rely on chemical/target similarity or external data collection. Yong Liu et al. [16] further developed a Neighborhood Regularized Logistic Matrix Factorization (NRLMF) model, where the latent variables were learned by logistic matrix factorization, moreover, NRLMF assigned higher levels to positive observations (observed DTIs) than negative observations (unknown DTIs). And the local structure of DTI data has been exploited via neighborhood regularization to achieve better prediction accuracy. Details about PMF and LMF are described in methods part.

### STITCH database

STITCH (http://stitch.embl.de) is a database of known and predicted chemical–protein interactions, which integrates the evidence derived from experiments, other databases and literatures. In STITCH fifth release [17], it encompasses more than 9 600 000 proteins from 2031 eukaryotic and prokaryotic genomes, and 430 000 chemicals (not including different stereoisomers). In the human subset we are interested in, there are 559,241 chemicals (including different stereoisomer), 9,920 proteins, and 8,842,952 chemical-protein interactions with experimental confidence scores. Considering the limited computational power of this project, we plan to train a DTI prediction model on STITCH version 4 [18] (released in 2014), which has a smaller human subset: 531,716 chemicals (including different stereoisomer), 9,337 proteins, and 1,351,208 chemical-protein interactions with experimental confidence scores. The experimental confidence score was probabilistic score converted from reported residual kinase activities and kinase affinities, including low (< 0.5), medium (0.5-0.7) and high confidence (> 0.7).

For data like STITCH subset with an order of $10^4$ proteins and $10^5$ chemicals, implying $10^8$ sequence and $10^{10}$ chemical similarity comparisons, which is computationally unfeasible. Therefore, in this project, we plan to learn DTI model using PMF and LMF machine learning method, based on chemical-protein interactions from STITCH version 4, human subset with experimental data. Hyper-parameters selection and performance evaluation are performed through 10-fold cross validation.

## 2 Materials and Methods

### Problem Definition

The DTI network is a bipartite graph with two types of nodes: chemicals and proteins. Each edge represents an interaction between a chemical and a protein, the weight on each edge is determined by the probability of a chemical interacts with a protein. The DTI prediction model is to predict weights on missing edges based on the existing nodes and edges in the network.

## Data Set

The data set we used here was STITCH version 4 (2014), data were downloaded from STITCH website protein chemical links(http://stitch.embl.de), human subset with experimental confidence score greater than 0 was extracted, resulting in N = 531,716 chemicals, M = 9,337 proteins and 1,351,208 chemical-protein interactions among them. The size of interaction space is $N \times M = 4.9646 \times 10^9$, the percent occupancy of the interaction space is 0.027%. On average, there are 2.54 interactions per chemical and 144.72 interactions per protein. The experimental confidence score ranges from 0.063 to 0.900, with mean value of 0.498 and standard deviation of 0.252, histogram of confidence score was shown in **Figure 1**. Confidence score greater than 0.7 were considered as high confidence, between 0.5 and 0.7 were considered as medium confidence, and lower than 0.5 were considered as low confidence.

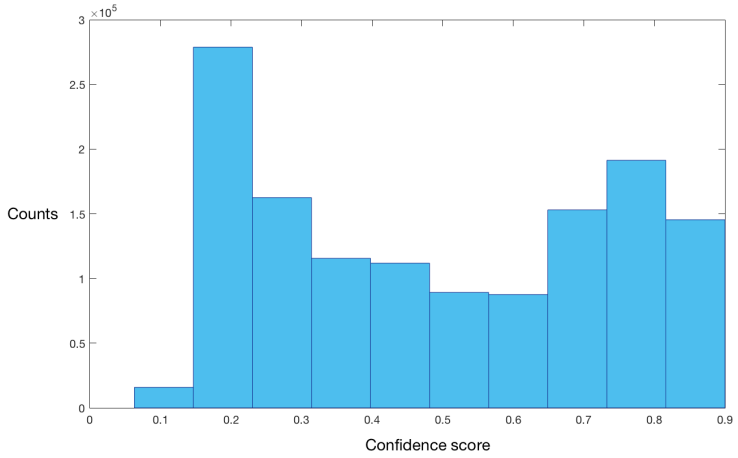


**Figure 1**. Histogram of STITCH version 4 human subset confidence score

## Probabilistic Matrix Factorization (PMF)

The Matrix Factorization technique has been successfully applied to DTI prediction, as descried by Cobanoglu et al. [15], PMF decomposes the connectivity matrix $Y_{N \times M}$, as a bipartite graph of N chemicals $D = \{d_i\}_{i=1}^{N}$, and M proteins $T = \{t_j\}_{j=1}^{M}$. $Y_{ij}$ is defined by the experimental confidence score provided by STITCH, the score is set to be 0 if there's no experimental evidences show interactions. The properties of a chemical $d_i$ and a protein $t_j$ are described by two K-dimensional latent vectors $u_i \in \mathbb{R}^{1 \times K}$ and $v_j \in \mathbb{R}^{1 \times K}$. The interaction probability $p_{ij}$ of a chemical-protein pair is modeled by $p_{ij} = u_i^T v_j$, and the matrix $P_{N \times M}$ is modeled as

$$P_{N \times M} = U_{N \times K}^T V_{K \times M}$$

Therefore, the conditional probability over observed interactions is represented as

$$p(Y|U, V, \sigma^2) = \prod_{i=1}^{N} \prod_{j=1}^{M} \left[ f\left(Y_{ij} \middle| u_i^T v_j, \sigma^2\right) \right]^{I_{ij}}$$

where $f(x|\mu, \sigma^2)$ is the Gaussianly distributed probability density function for x, with mean $\mu$ and variance $\sigma$, and $I_{ij}$ is the indicator function equal to 1 if the entry of Y is known and 0 otherwise.

The log likelihood of U and V is

$$\ln\left(p(U,V|Y,\sigma^2,\sigma_u^2,\sigma_v^2)\right) = -\frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{j=1}^{M} I_{ij}\left(Y_{ij} - u_i^T v_j\right)^2 - \frac{1}{2\sigma_u^2}\sum_{i=1}^{N} u_i^T u_i - \frac{1}{2\sigma_v^2}\sum_{j=1}^{M} v_j^T v_j + C$$

To maximize the log likelihood function is to minimize the following function with non-zero square loss and $L_2$ regularization.

$$\min_{U,V} \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M} I_{ij}\left(Y_{ij} - u_i^T v_j\right)^2 + \frac{\lambda}{2}\left(\sum_{i=1}^{N} u_i^T u_i + \sum_{j=1}^{M} v_j^T v_j\right)$$

The PMF model was learned by gradient descent, and yielded the optimal $u_i$ and $v_j$ vector for each chemical and protein. The dot product of $u_i$ and $v_j$ correspond to the probability of chemical $u_i$ interacts with protein $v_j$.

## Logistic Matrix Factorization (LMF)

The Logistic Matrix Factorization (LMF) method share similar idea with PMF, the only difference is that LMF demonstrate the interaction probability $p_{ij}$ between a chemical and a protein by logistic function

$$p_{ij} = \frac{\exp\left(u_i^T v_j\right)}{1 + \exp\left(u_i^T v_j\right)}$$

In this case, the probability $p_{ij}$ is constraint to be within the range 0 to 1. LFM has been demonstrated to be effective for personalized recommendations in previous studies [19]. The correspond log likelihood is represented by

$$\ln\left(p(U,V|P,\sigma^2,\sigma_u^2,\sigma_v^2)\right) = \frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{j=1}^{M} I_{ij}\left(Y_{ij}u_i^T v_j - \ln\left[1 + \exp\left(u_i^T v_j\right)\right]\right) - \frac{1}{2\sigma_u^2}\sum_{i=1}^{N} u_i^T u_i - \frac{1}{2\sigma_v^2}\sum_{j=1}^{M} v_j^T v_j + C$$

To maximize the log likelihood function is to minimize the following function

$$\min_{U,V} \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{M} I_{ij}\left(\ln\left[1 + \exp\left(u_i^T v_j\right)\right] - Y_{ij}u_i^T v_j\right) + \frac{\lambda}{2}\left(\sum_{i=1}^{N} u_i^T u_i + \sum_{j=1}^{M} v_j^T v_j\right)$$

## Stochastic Gradient Descent (SGD) for Matrix Factorization

The gradient for baseline PMF L2 loss

$$\frac{\partial L}{\partial u_i} = -I_{ij}\left(Y_{ij} - u_i^T v_j\right)v_j + \lambda u_i^T$$

$$\frac{\partial L}{\partial v_j} = -I_{ij}\left(Y_{ij} - u_i^T v_j\right)u_i^T + \lambda v_j$$

The gradient for LMF L2 loss

$$\frac{\partial L}{\partial u_i} = -I_{ij}\left(Y_{ij} - \frac{\exp\left(u_i^T v_j\right)}{1 + \exp\left(u_i^T v_j\right)}\right)v_j + \lambda u_i^T$$

$$\frac{\partial L}{\partial v_j} = -I_{ij}\left(Y_{ij} - \frac{\exp\left(u_i^T v_j\right)}{1 + \exp\left(u_i^T v_j\right)}\right)\mathrm{u}_\mathrm{i}^T + \lambda \mathrm{v}_\mathrm{j}$$

Updates after each iteration step

$$u_{i_{new}} \leftarrow u_i - \epsilon_n \frac{\partial L}{\partial u_i}$$

$$v_{j_{new}} \leftarrow v_j - \epsilon_n \frac{\partial L}{\partial v_j}$$

$$\epsilon_n = \frac{\alpha}{n * \gamma + 1}$$

where n is the iteration number of SGD.

**Parameters search**

Hyper parameter K (the dimension of latent vector of each chemical and protein) was selected with singular value decomposition. Top 200 singular values of Y were obtained (**Figure 2**)

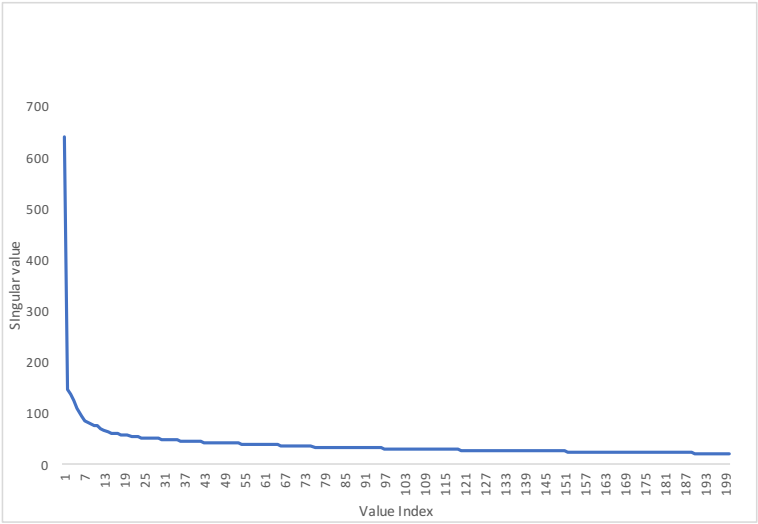


**Figure 2.** Top 200 singular values of chemical-protein matrix Y

The settings of hyper-parameters of each method were as follows

$K \in \{20, 30, 40, 50, 60, 70, 90\}$ the search space of K was set based on the SVD results.

$regularization\ \lambda \in \{5\times10^{-1}, 5\times10^{-2}, 5\times10^{-3}, 5\times10^{-4}, 5\times10^{-5}\}$

step size $\alpha \in \{0.005, 0.01, 0.02, 0.04, 0.1\}$

and $\gamma$ was set to be 0.1.

The model was trained with Spark Scala.

**Performance evaluation**

The performance of PMF and LMF were evaluated under 10-fold cross-validation (CV), mean square error, precision, precision enrichment and recall, recall enrichment were used as the evaluation metrics.

In the 10-fold cross-validation, we randomly separate all known chemical-protein pairs into 10 subsets. For each round, 9 subsets were selected as training data and the remaining 1 subset was validation test set. Mean square error (MSE) was calculate by the sum of each prediction error square divided by the total number of prediction in test set.

$$\mathrm{MSE} = \frac{\sum_{y_{ij} \in test\ set} \left(p_{ij} - y_{ij}\right)^2}{\#\ total\ elements\ in\ test\ set}$$

In our practical recommendation process, chemical-protein interactions with predicted score greater than certain cutoff value were extracted out and considered as potential interactions for research analysis. Therefore we defined our positive pairs as follows, a chemical-protein interaction with confidence score greater than 0.7 (high confidence) is considered as a positive data point, an DTI in test set with both original and predicted scores greater than 0.7 is considered as true positive, an DTI with original score greater than 0.7 but predicted score no larger than 0.7 is considered as false negative, an DTI with original score no greater than 0.7 but predicted score greater than 0.7 is considered as false positive (the performance of cutoff value of 0.5, indicating high and medium confidence was also reported). Then the precision is defined as

$$\mathrm{P(L)} = \frac{\#\ true\ positive}{\#\ true\ positive + \#\ false\ positive}$$

the recall is defined as

$$\mathrm{R(L)} = \frac{\#\ true\ positive}{\#\ true\ positive + \#\ false\ negative}$$

$\mathrm{P_{rand}}(L)$ and $\mathrm{R_{rand}(L)}$ are defined as the precision and recall obtained from a random recommendation process, which means each chemical-protein pair is assign an interaction probability of random score within 0 and 1. Then the precision enrichment against random process is defined as

$$\mathrm{E_p(L)} = \frac{\mathrm{P(L)}}{\mathrm{P_{rand}(L)}}$$

the recall enrichment is defined as

$$\mathrm{R_p(L)} = \frac{\mathrm{R(L)}}{\mathrm{R_{rand}(L)}}$$

Mean square error, precision, precision enrichment, recall, recall enrichment values were reported to compare the performance of PMF and LMF through 10-fold cross-validation.

## 3 Results

The optimal parameters setting for PMF is $\mathrm{K} = 30, \alpha = 0.04$ and $\lambda = 0.00005$ and optimal parameter setting for LMF is $\mathrm{K} = 30, \alpha = 0.04$ and $\lambda = 0.00005$ after parameter search. As shown in **Table 1**, the MSE for LMF (0.0227) was much better than that of PMF (0.0436), the precision and recall value got a slight improvement by LMF compared with PMF. To be more specific, for cutoff = 0.7, the precision improved from 0.6228 to 0.6330, recall improved from 0.6003 to 0.6319, 2 to 3 folds better than random recommendation. If we lose the cutoff value to 0.5, higher precision (0.7804) and recall (0.8159) values were obtained.

**Table 1**. The performance of PMF and LMF

| | PMF | LMF |
|---|---|---|
| MSE | 0.0436 ± 0.0015 | 0.0227 ± 0.0002 |
| Precision (score >0.7) | 0.6228 ± 0.0036 | 0.6330 ± 0.0026 |
| Precision Enrichment (score >0.7) | 2.7221 ± 0.0229 | 2.7776 ± 0.0202 |
| Recall (score >0.7) | 0.6003 ± 0.0050 | 0.6319 ± 0.0018 |
| Recall Enrichment (score >0.7) | 1.9954 ± 0.0198 | 2.1181 ± 0.0198 |
| Precision (score >0.5) | 0.7752 ± 0.0037 | 0.7804 ± 0.0028 |
| Precision Enrichment (score >0.5) | 1.8827 ± 0.0065 | 1.8946 ± 0.0074 |
| Recall (score >0.5) | 0.8129 ± 0.0025 | 0.8159 ± 0.0023 |
| Recall Enrichment (score >0.5) | 1.6283 ± 0.0092 | 1.6345± 0.0076 |

# 4 Discussion

In this project, we took the advantage of STITCH experimental confidence score, treated DTI prediction as regression problem instead of simple binary classification. Which is more consistent with the fact that different drug-target pairs have different interaction affinities. The employed LMF method not only helped us constraint the predicted score within 0 and 1, the performance of LMF was also better than simple PMF in our analysis, especially in MSE. However, obtaining precision value of 0.6330 and recall value of 0.6319 for high confidence pairs didn't meet our expectation. It is possible to expect better performance if we employ other regularization method like neighborhood regularization [16] or graph regularization [20] in the future.

There were some other reasons beside the learning method that affect the precision and recall values. One was because of the specific way we define precision and recall. The best performance of LMF got a MSE of 0.0227, which means on average the error of predicted score and original score was 0.1507, indicating the method was applicable to predict potential interactions between drug and target pairs. The average total number of interactions greater than 0.700 in each test data set was 35,629, among which 3,172 (8.9%) between 0.700 and 0.750. Among the DTIs with confidence score within 0.700 and 0.750, only interactions with predicted score greater 0.7 were considered as true positive was a little bit strict, this is one of the reason why we could only get around 0.6 precision and recall values for high confidence setting. The same case when we set a cutoff value of 0.5, since the average total number of interactions with score greater than 0.5 in each test data set was 64,665, among which 5,164 (8.0%) between 0.500 and 0.550. In addition, as indicated by Yamanishi et al. [8], database itself could lead some bias on prediction performance. We need more investigation on better ways to evaluate the model in the future.

## References


1. Stern, A.M., et al., *A Perspective on Implementing a Quantitative Systems Pharmacology Platform for Drug Discovery and the Advancement of Personalized Medicine.* J Biomol Screen, 2016. **21**(6): p. 521-34.
2. Yamanishi, Y., et al., *Prediction of drug-target interaction networks from the integration of chemical and genomic spaces.* Bioinformatics, 2008. **24**(13): p. i232-40.
3. Overington, J.P., B. Al-Lazikani, and A.L. Hopkins, *How many drug targets are there?* Nat Rev Drug Discov, 2006. **5**(12): p. 993-6.
4. Santos, R., et al., *A comprehensive map of molecular drug targets.* Nat Rev Drug Discov, 2016. **advance online publication**.
5. Yang, K., et al., *Finding multiple target optimal intervention in disease‐related molecular network.* Molecular Systems Biology, 2008. **4**(1).
6. Chen, X., M.-X. Liu, and G.-Y. Yan, *Drug-target interaction prediction by random walk on the heterogeneous network.* Molecular BioSystems, 2012. **8**(7): p. 1970-1978.
7. Cheng, F., et al., *Prediction of Drug-Target Interactions and Drug Repositioning via Network-Based Inference.* PLOS Computational Biology, 2012. **8**(5): p. e1002503.
8. Yamanishi, Y., et al., *Prediction of drug–target interaction networks from the integration of chemical and genomic spaces.* Bioinformatics, 2008. **24**(13): p. i232-i240.
9. Bleakley, K. and Y. Yamanishi, *Supervised prediction of drug–target interactions using bipartite local models.* Bioinformatics, 2009. **25**(18): p. 2397-2403.
10. Yamanishi, Y., et al., *Drug-target interaction prediction from chemical, genomic and pharmacological data in an integrated framework.* Bioinformatics, 2010. **26**(12): p. i246-i254.
11. Wang, Y. and J. Zeng, *Predicting drug-target interactions using restricted Boltzmann machines.* Bioinformatics, 2013. **29**(13): p. i126-i134.
12. Xia, Z., et al., *Semi-supervised drug-protein interaction prediction from heterogeneous biological spaces.* BMC Systems Biology, 2010. **4**(2): p. S6.
13. Pahikkala, T., et al., *Toward more realistic drug-target interaction predictions.* Brief Bioinform, 2015. **16**(2): p. 325-37.
14. Nascimento, A.C., R.B. Prudencio, and I.G. Costa, *A multiple kernel learning algorithm for drug-target interaction prediction.* BMC Bioinformatics, 2016. **17**: p. 46.
15. Cobanoglu, M.C., et al., *Predicting Drug–Target Interactions Using Probabilistic Matrix Factorization.* Journal of Chemical Information and Modeling, 2013. **53**(12): p. 3399-3409.
16. Liu, Y., et al., *Neighborhood Regularized Logistic Matrix Factorization for Drug-Target Interaction Prediction.* PLOS Computational Biology, 2016. **12**(2): p. e1004760.
17. Szklarczyk, D., et al., *STITCH 5: augmenting protein–chemical interaction networks with tissue and affinity data.* Nucleic Acids Research, 2016. **44**(D1): p. D380-D384.
18. Kuhn, M., et al., *STITCH 4: integration of protein–chemical interactions with user data.* Nucleic Acids Research, 2014. **42**(D1): p. D401-D407.
19. Johnson, C.C. *Logistic matrix factorization for implicit feedback data*. in *NIPS 2014 Workshop on Distributed Machine Learning and Matrix Computations*. 2014.
20. Ezzat, A., et al., *Drug-Target Interaction Prediction with Graph Regularized Matrix Factorization.* IEEE/ACM Trans Comput Biol Bioinform, 2016.