IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 5 July 2022, accepted 11 August 2022, date of publication 17 August 2022, date of current version 2 September 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3199667

RESEARCH ARTICLE

NNDSVD-GRMF: A Graph Dual Regularization Matrix Factorization Method Using Non-Negative Initialization for Predicting Drug-Target Interactions

JUNJUN ZHANG[1] AND MINZHU XIE[1,2]

[1]Key Laboratory of Computing and Stochastic Mathematics (LCSM) (Ministry of Education), School of Mathematics and Statistics, Hunan Normal University, Changsha 410081, China

[2]College of Information Science and Engineering, Hunan Normal University, Changsha 410081, China

Corresponding author: Minzhu Xie (xieminzhu@hunnu.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 62172028 and Grant 61772197.

ABSTRACT Accurate drug-target interactions (DTIs) prediction can significantly speed up the process of new drug design and development. Recently, many matrix factorization methods have been used to predict DTIs. However, most of them use heuristic and iterative strategies, and their convergence and performance can not be guaranteed. In order to accurately predict DTIs, we propose a new algorithm, NNDSVD-GRMF, our method is based on graph dual regularization non-negative matrix factorization (GDNMF) and non-negative double singular value decomposition (NNDSVD), which considers both the initialization stage of the non-negative matrix factorization and the structural information of the data and features. At the same time, in order to improve the adaptability of the algorithm, the extension of the NNDSVD-GRMF (NNDSVD-WGRMF) is also proposed. Extensive experimental results show that our methods have better performance than other state-of-the-art methods. In the case studies, among the 10 highest-scoring drugs predicted to interact with androgen receptor, 9 drugs have been validated, and among the 10 highest-scoring target proteins predicted to be targeted by the drug nicotine bitartrate, 9 targets have been validated.

INDEX TERMS Drug, graph dual regularization, non-negative double singular value decomposition, target.

I. INTRODUCTION

Drug-target interaction (DTI) identification plays important role in the process of new drug discovery and development. However, all wet experiments to identify DTIs require expensive equipments and costly chemical reagents, and their process is time consuming [1]. With the rapid accumulation of related information of drugs, target proteins and validated DTIs, more and more computational methods have been proposed to predict new DTIs. Based on 3D structures of proteins (targets), Cheng *et al.* [2] used docking simulations to predict compound-protein interactions. Campillos *et al.* [3] introduced a method using drug side-effect similarity to infer the possibility that a drug interacts with a target. The above two methods only apply to drugs with known side-effects or proteins with known 3D structure, and could not perform large scale screening of potential DTIs.

Recently a variety of approaches have been proposed based on the known DTI network, drug (or compound) chemical structures, and target protein sequences. For example, Yamanishi *et al.* [4] used a DTI bipartite graph to represent the known DTI network, and proposed a bipartite graph learning method to predict DTIs. They constructed a similarity matrix according to the shortest distances between all drugs and targets in the the DTI bipartite graph, and used the eigenvalue decomposition of the matrix to embed all drugs and targets into a unified feature space. Two kernel regression models were learned to map drugs and targets to the unified space from their chemical structure space and their sequence space, respectively. The inner product of the

The associate editor coordinating the review of this manuscript and approving it for publication was Shubhajit Roy Chowdhury.

feature vectors in the unified space of a drug and a target was used to measure their interaction possibility. Based on the same data, Bleakley and Yamanishi [5] used bipartite local models (BLM) to predict target proteins for a given drug, and the drugs targeting a given protein. Therefore they obtained two predictions for each drug-protein pair, which were then combined into a prediction score.

Using the interaction profiles of targets (and drugs) from the known DTI network, Laarhoven *et al.* [6] introduced a method called the Gaussian Interaction Profile (GIP) kernel to calculate the similarities between targets (and drugs), and used Regularized Least Squares (RLS) to predict DTIs. Perlman *et al.* [7] integrated multiple types of information including the chemical structures, side effects, ATC codes and related gene expression profiles of drugs, and the sequences and PPI network of targets, then used logistic regression to score DTIs. Mei *et al.* [8] integrated neighbor-based interaction-profile inferring into BLM and proposed a method BLM-NII to make DTI predictions. Based on GIP, Laarhoven and Marchiori [9] proposed a weighted nearest neighbor method to predict DTIs. With the rapid development of machine learning, machine learning based DTI prediction methods have also been proposed. For example, Wang and Zeng [10] used restricted Boltzmann machine to model multiple types of DTIs. Lan *et al.* [11] set unknown interactions as unlabeled samples, and used Random walk with restarts, K nearest neighbor and heat kernel diffusion to divide them into reliable negative samples(RN) and likely negative samples(LN). Based on the positive samples (known DTIs), RN and LN, a weighted support vector machine was chosen as classifier to predict DTIs. Rifaioglu *et al.* proposed DEEPScreen [12], which used a deep convolutional neural network to learn drug features from their 2-D structural images and to predict large-scale DTIs.

In bioinformatics, matrix factorization method has a wide range of applications, such as clustering and feature selection [13], lncRNA-disease associations prediction [14], tumor classification [15] and marker extraction [16]. In DTI prediction, the known DTI network is usually represented by a matrix. In most matrix factorization based methods, the interaction matrix is decomposed into two matrices of low ranks which represent the interaction between a drug and a target as the inner product of their feature vectors. For example, Gönen [17] proposed a kernelized Bayesian matrix factorization method to predict DTIs. Zheng *et al.* [18] proposed a DTI prediction model named Multiple Similarities Collaborative Matrix Factorization (MSCMF), which used an alternating least squares algorithm to decompose the interaction matrix into two low-rank feature factor matrices that are consistent with the similarity matrices of drugs and targets. Liu *et al.* [19] combined logistic matrix decomposition with neighborhood regularization to predict DTIs. Bolgár and Antal [20] proposed a variational Bayesian multiple kernel logistic matrix factorization method using Laplacian regularization, multiple kernel learning, and a variational Bayesian inference process to infer drug-target interaction possibilties. Ezzat *et al.* [21] used graph regularization into matrix factorization to learn drug-target interaction models. Based on $L_{2,1}$ norm graph regularization, Cui *et al.* [22] used matrix factorization to predict DTIs. Liu *et al.* [23] used a graph convolutional network (GCN) followed by a random walk with restart (RWR) to obtain features of the drugs and targets from the related heterozygous data, and a matrix factorization model (DistMult) to predict DTIs. Gao *et al.* [24] proposed a collaborative matrix factorization method with soft regularization to predict DTIs. Although many methods have been proposed for DTIs prediction, the prediction performances are far from satisfactory. The key issue of DTIs prediction is how to efficiently use the existing validated DTIs and exploit the useful information hidden among drugs or targets.

Being simple and practical, the singular value decomposition (SVD) algorithm is usually used to provide an initial solution to matrix factorization, however the negative values in the component matrices make the results hard to explain. To provide better and explainable initial component matrices for matrix factorization, Boutsidis and Gallopoulos [25] proposed an algorithm non-negative double singular value decomposition (NNDSVD), which can enhance the initialization stage of nonnegative matrix factorization. In DTI prediction, the high-dimensional data are in fact sampled from a nonlinear low-dimensional manifold embedded in the high-dimensional space, and according to [26], the model learning performance can be greatly improved if the intrinsic geometrical structure of the manifold have been taken into account. To capture the structural information from both data and the features, Shang *et al.* [27] proposed a graph dual regularization non-negative matrix factorization method (GDMF) for clustering, and their experimental results showed GDMF had better clustering performance and more discriminating power than the general non-negative matrix factorization and graph regularization non-negative matrix factorization. In the paper, we propose a DTI prediction method called NNDSVD-GRMF based on the graph dual regularized non-negative matrix factorization and the non-negative double singular value decomposition. At first a DTI matrix, a drug similarity matrix and a target similarity matrix are constructed based on known DTIs, drug chemical structures and target sequences. Then DTI prediction is transformed into non-negative factorization of the DTI matrix with graph dual regularization terms. The graph dual regularization terms are used to integrate the information from the drug similarity matrix and the target similarity matrix, in order to take the intrinsic geometrical structures of the related manifolds into account. NNDSVD-GRMF can be used to predict novel drug-target interactions for drugs without any known targeted proteins and for proteins without any known drugs targeting them. To improve the adaptability, a weighted extension (i.e. NNDSVD-WGRMF) of the NNDSVD-GRMF is also proposed. Experimental results show that our methods have better performance than other state-of-the-art methods on four datasets under two different experimental scenarios. In case studies involving the target androgen receptor and

TABLE 1. The information of the gold standard datasets.

Datasets	NR	GPCR	IC	E
Interactions	90	635	1476	2926
Drugs	54	223	210	445
Targets	26	95	204	664
Sparseness(%)	93.59	97.00	96.55	99.01

the drug nicotine bitartrate, among the 10 highest-scoring drugs predicted to interact with androgen receptor, 9 drugs have been validated by wet experiments, and among the 10 highest-scoring target proteins predicted to be targeted by the drug nicotine bitartrate, 9 targets have been validated by wet experiments.

II. MATERIALS

We used the same data as in [4]. The known DTI data, the amino acid sequences of protein targets and the chemical structure data of drugs were downloaded from KEGG [28]. Protein targets include the following four classes: nuclear receptor (NR), G protein-coupled receptor (GPCR), ion channel (IC) and enzyme (E). The validated DTIs were considered the gold standard data. We used the same known DTI data as in [4], which have 4 datasets whose names are NR, GPCR, IC and E according to their target classes. The information of the 4 datasets are shown in Table 1. In the NR dataset, there are 90 known interactions between 54 drugs and 26 nuclear receptors; in the GPCR dataset, there are 635 known interactions between 223 drugs and 95 G protein-coupled receptors; in the IC dataset, there are 1476 known interactions between 210 drugs and 204 ion channels; and in the E dataset, there are 2926 known interactions between 445 drugs and 664 enzymes. The known DTIs were denoted as a $n \times m$ matrix X, where n and m are the numbers of drugs and targets, respectively. If the ith drug is validated to interact with the jth target, $X_{ij} = 1$; otherwise, $X_{ij} = 0$.

The structural similarities between drugs were calculated from the chemical structure data using SIMCOMP [29] according to the size of the common substructures between two drugs. The similarities between drugs were represented by a $n \times n$ matrix S^d. The sequence similarities between two target proteins were computed from the sequence data using the normalized Smith-Waterman alignment score [30]. Let $SW(p_1, p_2)$ be the original Smith-Waterman alignment score of p_1 and p_2. The similarity between p_1 and p_2 $s(p_1, p_2) = \frac{SW(p_1,p_2)}{\sqrt{SW(p_1,p_1)}\sqrt{SW(p_2,p_2)}}$. The similarities between targets were represented by a $m \times m$ matrix S^t.

III. METHODS

A. NON-NEGATIVE MATRIX FACTORIZATION

For a $n \times m$ non-negative matrix $X \in \mathbb{R}_+^{n\times m}$, the general form of the non-negative matrix factorization (NMF) of X is to decompose X into two low rank nonnegative component matrices, say $A \in \mathbb{R}_+^{n\times k}$ and $B \in \mathbb{R}_+^{m\times k}$, such that their product well approximates X, i.e. $X \approx AB^T$. The component matrices A and B are also called the latent feature matrices of X. When the square of the Frobenius norm of the difference between X and AB^T is used as the cost function, NMF of X is to solve the following optimization problem:

$$\min \left\| X_{n\times m} - A_{n\times k} B^T_{m\times k} \right\|_F^2 \quad s.t. \quad A, B \geq 0. \tag{1}$$

The most commonly used method to obtain a local optimal solution to the problem is the iterative process proposed by Lee and Seung [31] which uses the following update formulas:

$$A_{ij} \leftarrow A_{ij} \frac{(XB)_{ij}}{(AB^TB)_{ij}},$$
$$B_{ij} \leftarrow B_{ij} \frac{(X^TA)_{ij}}{(BA^TA)_{ij}}. \tag{2}$$

To avoid overfitting, the L2 regular term $(||A||_F^2 + ||B||_F^2)$ is added into the cost function of NMF, and we get the L2 regularized NMF:

$$\min \left\| X - AB^T \right\|_F^2 + \lambda_l(\|A\|_F^2 + \|B\|_F^2)$$
$$s.t. \quad A, B \geq 0. \tag{3}$$

B. GRAPH REGULARIZED NON-NEGATIVE MATRIX FACTORIZATION

Each column of X can be considered as a data point in a n dimension space. In real applications, k is usually much smaller than n and m, and the NMF of X tries to find a low dimension space such that the data points in X could be well represented as linear combinations of k base vectors. Based on the assumption that the two data points close in the latent geometry space will be close in the low dimension spaces, Cai *et. al.* [32] proposed a graph regularized non-negative matrix factorization (GRMF) by adding a graph regularized term. The cost function of GRMF is as follows:

$$O_{gr} = \left\| X - AB^T \right\|_F^2 + \lambda \operatorname{Tr}(B^T(D - W)B)$$
$$s.t. \quad A, B \geq 0, \tag{4}$$

where Tr(.) is the trace of a matrix, W is the weight matrix representing a neighbor graph on the data points, and D is a diagonal matrix that $D_{jj} = \sum_l W_{lj}$. The matrix $D-W$ is called graph Laplacian and denoted by $\mathcal{L}$ in the following. Furthermore, by adding another graph regularizer of the feature space, Shang *et. al.* [27] proposed a graph dual regularization non-negative matrix factorization (GDNMF), whose cost function is:

$$O_{gd} = \left\| X - AB^T \right\|_F^2 + \lambda_b \operatorname{Tr}(B^T\mathcal{L}_bB) + \lambda_a \operatorname{Tr}(A^T\mathcal{L}_aA)$$
$$s.t. \quad A, B \geq 0. \tag{5}$$

C. DRUG-TARGET INTERACTION MATRIX GRAPH REGULARIZED FACTORIZATION

For each drug i, we can select p most similar drugs out from other $n-1$ drugs as its p-nearest neighbors according to the similarity matrix S^d. The p-nearest neighbor set of drug i is denoted by $\mathcal{N}_p^d(i)$. Similarly, we can obtain the p-nearest neighbor set $\mathcal{N}_p^t(j)$ of target j according to the similarity

matrix S^t. We construct a p-nearest neighbor drug graph whose adjacent matrix N^d is defined as follows.

$$N_{ij}^d = \begin{cases} 1, & j \in \mathcal{N}_p^d(i) \text{ and } i \in \mathcal{N}_p^d(j), \\ 0, & j \notin \mathcal{N}_p^d(i) \text{ and } i \notin \mathcal{N}_p^d(j), \\ 0.5, & otherwise. \end{cases} \tag{6}$$

N^d is used to make the similarity matrix S^d sparse by using Equation (7).

$$\hat{S}_{ij}^d = N_{ij}^d S_{ij}^d, \forall i, j. \tag{7}$$

The graph Laplacian of $\hat{S}^d$ is defined as $\mathcal{L}_d = D^d - \hat{S}^d$, where D^d is a diagonal matrix whose diagonal element $D_{ii}^d = \sum_r \hat{S}_{ir}^d$.

Similarly, we construct a p-nearest neighbor target graph whose adjacent matrix is N^t according to Equation (8).

$$N_{ij}^t = \begin{cases} 1, & j \in \mathcal{N}_p^t(i) \text{ and } i \in \mathcal{N}_p^t(j), \\ 0, & j \notin \mathcal{N}_p^t(i) \text{ and } i \notin \mathcal{N}_p^t(j), \\ 0.5, & otherwise. \end{cases} \tag{8}$$

Using N^t, we make a sparse target similarity matrix $\hat{S}^t$ from S^t according to Equation (9).

$$\hat{S}_{ij}^t = N_{ij}^t S_{ij}^t, \forall i, j. \tag{9}$$

The graph Laplacian of $\hat{S}^t$ is $\mathcal{L}_t = D^t - \hat{S}^t$, where D^t is a diagonal matrix whose diagonal element $D_{jj}^t = \sum_q \hat{S}_{jq}^t$.

By adding the L2 regular term ($||A||_F^2 + ||B||_F^2$) and two graph regular terms $\mathrm{Tr}(A^T \mathcal{L}_d A)$ and $\mathrm{Tr}(B^T \mathcal{L}_t B)$, the non-negative matrix factorization of the drug-target interaction matrix X is formulated as the optimization problem in Equation (10).

$$\begin{aligned} \min & \left\| X - AB^T \right\|_F^2 + \lambda_l(\|A\|_F^2 + \|B\|_F^2) \\ & + \lambda_d \,\mathrm{Tr}(A^T \mathcal{L}_d A) + \lambda_t \,\mathrm{Tr}(B^T \mathcal{L}_t B), \\ s.t. \quad & A, B \geq 0. \end{aligned} \tag{10}$$

Since the normalized graph Laplacian performs better in many actual applications, we furthermore normalize $\mathcal{L}_d$ and $\mathcal{L}_t$ as Equations (11) and (12), respectively.

$$\widetilde{\mathcal{L}}_d = (D^d)^{-1/2} \mathcal{L}_d (D^d)^{-1/2}, \tag{11}$$

$$\widetilde{\mathcal{L}}_t = (D^t)^{-1/2} \mathcal{L}_d (D^t)^{-1/2}. \tag{12}$$

Finally, the drug-target interaction prediction problem is transformed into the non-negative matrix factorization of X with the L2 regular term and the dual normalized graph regularization terms, which is formulated as Equation (13). The inner product of the ith row of A and the jth row of B is used to predict the interaction between the ith drug and the jth target.

$$\begin{aligned} \min & \left\| X - AB^T \right\|_F^2 + \lambda_l(\|A\|_F^2 + \|B\|_F^2) \\ & + \lambda_d \,\mathrm{Tr}(A^T \widetilde{\mathcal{L}}_d A) + \lambda_t \,\mathrm{Tr}(B^T \widetilde{\mathcal{L}}_t B), \\ s.t. \quad & A, B \geq 0. \end{aligned} \tag{13}$$

TABLE 2. NNDSVD.

Inputs: Matrix $X \in R^{n \times m}$, integer $k < min(m, n)$.
Output: Rank-k nonnegative factors $A \in R^{n \times k}$, $B \in R^{m \times k}$.

1. Compute the largest k singular triplets of X : $[U, S, V] = svd(X, k)$
2. Initialize $A(:,1) = \sqrt{S(1,1)} * U(:,1)$ and $B(:,1) = \sqrt{S(1,1)} * V(:,1)$
3. for $j = 2 : k$
3.1. $u = U(:,j)$; $v = V(:,j)$;
3.2. $up = u^+$; $un = u^-$; $vp = v^+$; $vn = v^-$;
3.3. $upnrm = \|up\|$; $vpnrm = \|vp\|$; $mp = upnrm * vpnrm$;
3.5 $unnrm = \|un\|$;$vnnrm = \|vn\|$; $mn = unnrm * vnnrm$;
3.6. if $mp > mn$:
  $p = up/upnrm$; $q = yp/vpnrm$;$\sigma = mp$;
  else
  $p = un/unnrm$; $q = vn/ynnrm$; $\sigma = mn$;
  endif
3.7. $A(:,j) = \sqrt{S(j,j) * \sigma} * p$ and $B(:,j) = \sqrt{S(j,j) * \sigma} * q$;

$A(:,j)$ denotes the columns of A, $svd()$ denotes Singular Value Decomposition.

D. ALGORITHM

To solve the problem in Equation (13), we design an algorithm NNDSVD-GRMF. NNDSVD-GRMF uses NNDSVD [25] to get the initial values of A and B, and uses an iterative process to update A and B until they converge. The iterative process uses Equation (14) and Equation (15) to update A and B.

$$A = (XB - \lambda_d \widetilde{\mathcal{L}}_d A)(B^T B + \lambda_l I_k)^{-1}, \tag{14}$$

$$B = (X^T A - \lambda_t \widetilde{\mathcal{L}}_t B)(A^T A + \lambda_l I_k)^{-1}. \tag{15}$$

Since the iterative process can only reach a local optimal solution, choosing good initial values for A and B will enhance the possibility that the process reach a global optimal solution. We use NNDSVD to calculate the initial values of A and B, which could lead to rapid convergence of many NMF algorithms [25]. NNDSVD is based on the SVD of X: $X = U\Sigma V^T$, where Σ is a diagonal matrix whose diagonal entry at the ith row and the ith column is the ith largest singular value σ_i of X, and the ith columns u_i and v_i of U and V are the left and right singular vectors corresponding to σ_i, respectively. For a $n \times m$ matrix X with rank k, according to the property of SVD, X equals to the sum of k leading singular factors, *i.e.* $X = \sum_{i=1,..,k} \sigma_i u_i v_i^T$. Let $C_i = \sigma_i u_i v_i^T$, C_i^+ and C_i^- be the nonnegative part and negative part of C_i, respectively. $C_i = C_i^+ - C_i^-$, and $C_i^+ = \sigma_i(u_i^+ v_i^+ + u_i^- v_i^-)$. If $||u_i^+||||v_i^+|| > ||u_i^-||||v_i^-||$, the normalized item $\sqrt{\sigma_i||u_i^+||||v_i^+||}(u_i^+/||u_i^+||)$ is used to initial the i column of A, $\sqrt{\sigma_i||u_i^+||||v_i^+||}(v_i^+/||v_i^+||)$ is used to initial the i column of B. Otherwise $\sqrt{\sigma_i||u_i^-||||v_i^-||}(u_i^-/||u_i^-||)$ and $\sqrt{\sigma_i||u_i^-||||v_i^-||}(v_i^-/||v_i^-||)$ are used. Please see Table 2 for the details.

E. WEIGHTED GRAPH-REGULARIZED MATRIX FACTORIZATION

In order to get some latent features of A and B, weight matrix W is added to the graph regularization matrix

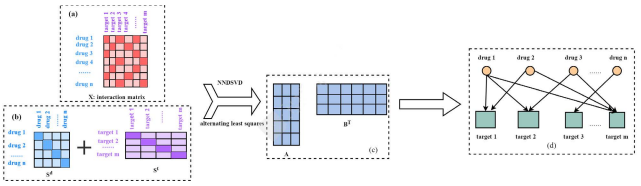


FIGURE 1. A schematic of NNDSVD-GRMF. (a) NNDSVD-GRMF uses known drug-target interaction information to construct adjacency matrix *X*; (b) Graph regularized terms are constructed according to similarity matrices of drugs and targets; (c) Matrix *X* is factorized into matrices *A* and *B* using NNDSVD and alternating least squares; (d) NNDSVD-GRMF obtains interaction scores between drugs and targets.

factorization. The form is

$$\min \left\| W \odot (X - AB^T) \right\|_F^2 + \lambda_l(\|A\|_F^2 + \|B\|_F^2) + \lambda_d Tr(A^T \widetilde{\mathcal{L}}_d A) + \lambda_t Tr(B^T \widetilde{\mathcal{L}}_t B), \tag{16}$$

where $\odot$ represents Hadamard product. Weight matrix $W \in \mathbb{R}_+^{n \times m}$, if X has a known drug-target interaction $W_{ij} = 1$, otherwise, $W_{ij} = 0$, $(i = 1, \cdots, n, j = 1, \cdots, m)$. Using alternating least squares as with NNDSVD-GRMF, the solution is obtained as

$$A_{ik} = \frac{(diag(W)XB + \lambda_d \widetilde{\mathcal{L}}_d A)_{ik}}{(diag(W)B^T B + \lambda_l I_k)_{kk}}, \tag{17}$$

$$B_{jk} = \frac{(diag(W)X^T A + \lambda_t \widetilde{\mathcal{L}}_t B)_{jk}}{(diag(W)A^T A + \lambda_l I_k)_{kk}}, \tag{18}$$

where I is identity matrix and *diag* is diagonal matrix. A schematic of NNDSVD-GRMF is shown in Fig. 1.

IV. RESULTS

In this section, we analyze the performance of NNDSVD-GRMF and NNDSVD-WGRMF in the following three aspects. First, we analyze the performance of NNDSVD-GRMF and NNDSVD-WGRMF under the two important evaluation indicators through 10-fold cross-validation. Second, we compare the performance of NNDSVD-GRMF and NNDSVD-WGRMF in the case of parameter changes. Third, we evaluate the reliability of NNDSVD-GRMF using case studies.

A. CROSS VALIDATION EXPERIMENTS

To evaluate the performance of DTI prediction methods, 5 repetitions of 10-fold cross-validation is performed in the experiment. In each repetition of 10-fold cross-validation, the dataset is divided into ten parts, 1 part is used for validation, and the remaining 9 parts are used for training, and this procedure is repeated 10 times. The final result is obtained by averaging the 5 repetitions of 10-fold cross-validation results. In order to evaluate the quality of the prediction results, the area under the receiver operating characteristic curve (AUC) and area under the precision-recall curve (AUPR) are used as the main evaluation indicators. In the table, the best AUC and AUPR under each data set are bolded and the standard deviation derived from the experiment is shown in parentheses. At the same time, receiver operating characteristic (ROC) curve and precision recall (PR) curve for each method are plotted.

To fully test the validity of the method, the cross-validation experiments are conducted under the following two scenarios [33].

1) CVd: the complete rows in the DTI matrix X are left out for the testing set. It tests the ability to predict interactions for new drugs,
2) CVt: the complete columns in the DTI matrix X are left out for the testing set. It tests the ability to predict interactions for new targets.

B. COMPARISONS WITH THE STATE-OF-THE-ART METHODS

To demonstrate the effectiveness of NNDSVD-GRMF and NNDSVD-WGRMF in predicting DTIs, we compare NNDSVD-GRMF and NNDSVD-WGRMF with the following the seven methods, namely BLM-NII [8], WKNKN [9], RLS-WNN [6], GRMF [21], WGRMF, CMF [18], SRCMF [24], where WGRMF is the weighted form of GRMF. In fact, BlM-NII, WKNKN and RLS-WNN use the neighborhood interactions for the prediction of DTIs, while the rest of the methods are constructed based on matrix factorization.

1) PARAMETER SETTINGS

In terms of parameter settings, we refer to the original literature of GRMF, WGRMF, CMF and SRCMF, some parameters are automated chosen using grid search [34] based on the AUPR value. The best parameters under each fold are usually different. Based on previous research [21], for GRMF, WGRMF, CMF, SRCMF, NNDSVD-GRMF and NNDSVD-WGRMF, the regularization parameters λ_l was selected from $\{2^{-2}, 2^{-1}, 2^0, 2^1\}$. For GRMF, WGRMF, NNDSVD-GRMF and NNDSVD-WGRMF, λ_d and λ_t were selected from $\{0, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}\}$. For CMF, SRCMF, λ_d and λ_t were selected from $\{2^{-3}, 10^{-2}, 2^{-1}, 2^0, 2^1, 2^1, 2^2, 2^3, 2^4, 2^5\}$. For NNDSVD-GRMF and NNDSVD-WGRMF, the ranks k of matrices were 26 and 49 under the NR and GPCR datasets, respectively. In IC and E datasets, k was selected from $\{50, 100\}$. For GRMF, WGRMF, CMF and SRCMF, the parameter k was selected from $\{50, 100\}$. In BLM-NII, the parameter α was set to 0.5. In the WKNKN, the parameters $K = 5$, $\eta = 0.7$.

2) PREDICTION RESULTS UNDER CVd

Table 3 and Table 4 list the AUC and AUPR values of each algorithm in the CVd scenario, respectively. On the NR dataset, the NNDSVD-WGRMF method has better performance other state-of-the-art methods. On the IC dataset, the AUPR value of NNDSVD-WGRMF has better performance other methods. Importantly, it can be found that the prediction performance is improved with the addition of the weight W. Fig. 2 and Fig. 4 show the histograms of AUC and AUPR with error bars for each algorithm under the CVd scenario, respectively. Fig. 3 and Fig. 5 show ROC curves and

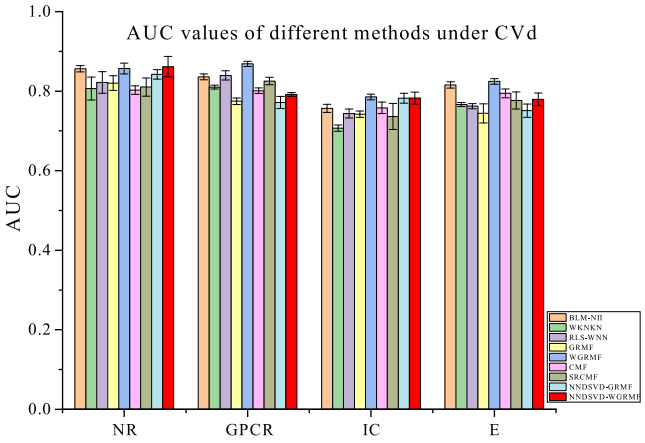


FIGURE 2. In the CVd scenario, the comparison of the AUC values of different methods in the four datasets.

PR curves of different algorithms on the four datasets under CVt, respectively.

3) PREDICTION RESULTS UNDER CVt

The AUC and AUPR values for each algorithm in the CVt scenario are given in Table 5 and Table 6, respectively. From Table 5, it can be found that the AUC of our method has better performance than other state-of-the-art methods on NR, GPCR, IC datasets. From Table 6, we can find that the AUPR values of NNDSVD-WGRMF outperform the other methods on NR and IC datasets. Fig. 6 and Fig. 8 show the histograms of AUC and AUPR with error bars for each algorithm under the CVt scenario, respectively. Fig. 7 and Fig. 9 show ROC curves and PR curves of different algorithms on the four datasets under CVt, respectively.

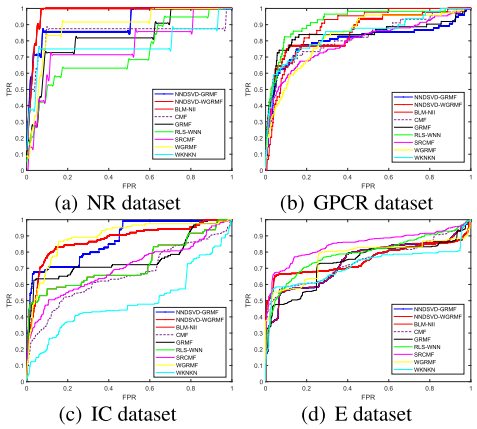


FIGURE 3. ROC curves for different methods are plotted together under CVd on NR dataset, GPCR dataset, IC dataset, E dataset, respectively.

C. SENSITIVITY ANALYSIS

The values of the parameters λ_d, λ_t and λ_l are obtained using grid search. In cross-validation, the training set for each fold is different, which complicates the sensitivity analysis of the parameters.

To investigate the effect of changes in parameters λ_d, λ_t and λ_l on the prediction results, we set $\lambda_d = 0$, $\lambda_t = 0$ and $\lambda_l = 0$ under CVd and CVt, respectively. The results are shown in Table 7-10. The results show that NNDSVD-GRMF performs best when λ_d, λ_t, and λ_l are not 0. In the CVd scenario, compared to the change of $\lambda_t = 0$, the change of $\lambda_d = 0$ and $\lambda_l = 0$ has a greater impact on

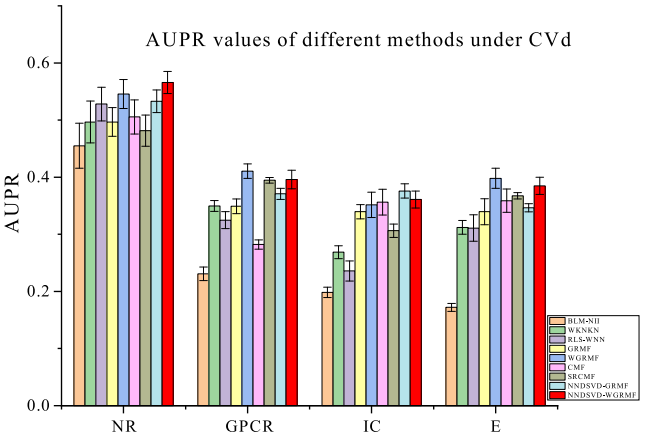


FIGURE 4. In the CVd scenario, the comparison of the AUPR values of different methods in the four datasets.

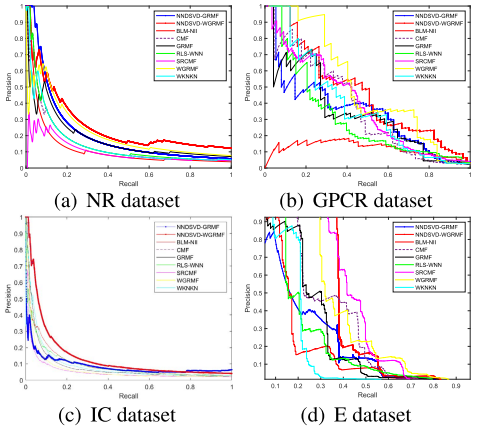


FIGURE 5. PR curves for different methods are plotted together under CVd on NR dataset, GPCR dataset, IC dataset, E dataset, respectively.

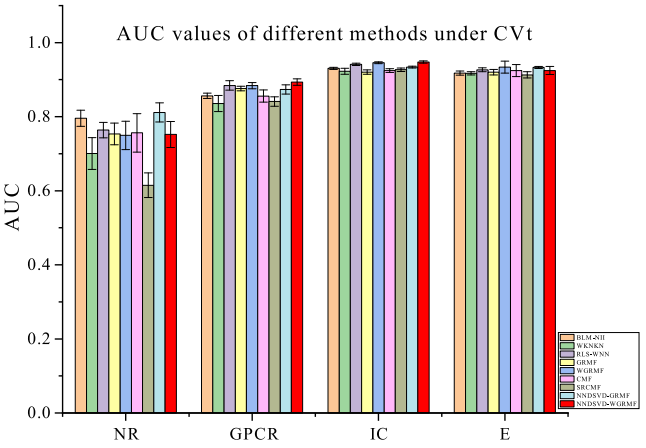


FIGURE 6. In the CVt scenario, the comparison of the AUC values of different methods in the four datasets.

TABLE 3. AUC values of different algorithms in CVd scenario.

Method	NR	GPCR	IC	E
BLM-NII	0.856292(0.0077)	0.836102(0.0073)	0.756714(0.0102)	0.815547(0.0080)
WKNKN	0.806684(0.0289)	0.810142(0.0048)	0.706933(0.0079)	0.766433(0.0050)
RLS-WNN	0.821758(0.0273)	0.839478(0.0116)	0.743888(0.0113)	0.762227(0.0066)
GRMF	0.820413(0.0185)	0.774848(0.0082)	0.742022(0.0080)	0.744108(0.0240)
WGRMF	0.856979(0.0135)	**0.868548**(0.0065)	**0.785357**(0.0070)	**0.824591**(0.0071)
CMF	0.802526(0.0109)	0.801118(0.0069)	0.758156(0.0144)	0.794486(0.0109)
SRCMF	0.810242 (0.0227)	0.825318 (0.0093)	0.736402 (0.0329)	0.776464(0.0214)
NNDSVD-GRMF	0.842158 (0.0121)	0.771546(0.0154)	0.782207(0.0124)	0.750871(0.0166)
NNDSVD-WGRMF	**0.861702** (0.0258)	0.791585(0.0047)	0.782561(0.0152)	0.779424(0.0158)

TABLE 4. AUPR values of different algorithms in CVd scenario.

Method	NR	GPCR	IC	E
BLM-NII	0.455027(0.0395)	0.230746(0.0118)	0.198357(0.0091)	0.172086(0.0068)
WKNKN	0.496622(0.0366)	0.349695(0.0096)	0.268694(0.0113)	0.312078(0.0121)
RLS-WNN	0.528022(0.0294)	0.324815(0.0149)	0.235889(0.0176)	0.310967(0.0232)
GRMF	0.496592(0.0252)	0.349027(0.0129)	0.339622(0.0124)	0.339569(0.0227)
WGRMF	0.545559 (0.0252)	**0.410652**(0.0126)	0.351595(0.0223)	**0.397949**(0.0176)
CMF	0.505449(0.0299)	0.282205(0.0081)	0.356396(0.0227)	0.358833 (0.0205)
SRCMF	0.481308 (0.0273)	0.394653 (0.0049)	0.306309 (0.0116)	0.367386(0.0054)
NNDSVD-GRMF	0.532864(0.0198)	0.370845(0.0098)	**0.375827** (0.0125)	0.346393(0.0072)
NNDSVD-WGRMF	**0.565726** (0.0193)	0.396041(0.0163)	0.360993(0.0148)	0.384971(0.0150)

TABLE 5. AUC values of different algorithms in CVt scenario.

Method	NR	GPCR	IC	E
BLM-NII	0.795604(0.0217)	0.856269(0.0071)	0.930531(0.0029)	0.917814(0.0056)
WKNKN	0.700475(0.0430)	0.835764(0.0217)	0.922583(0.0079)	0.916965(0.0042)
RLS-WNN	0.763799(0.0208)	0.884184(0.0128)	0.941532(0.0031)	0.926638(0.0053)
GRMF	0.753382(0.0293)	0.876011(0.0063)	0.920496(0.0060)	0.920224(0.0074)
WGRMF	0.749512 (0.0384)	0.883883 (0.0083)	0.945641(0.0024)	**0.933971**(0.0161)
CMF	0.75651 (0.0520)	0.855621 (0.0164)	0.924479(0.0051)	0.924598(0.0161)
SRCMF	0.614843 (0.0333)	0.840992 (0.0127)	0.926765 (0.0049)	0.913015 (0.0082)
NNDSVD-GRMF	**0.811495** (0.0257)	0.873582 (0.0124)	0.933894 (0.0026)	0.933011 (0.0028)
NNDSVD-WGRMF	0.751956 (0.0354)	**0.893556**(0.0088)	**0.947729** (0.0030)	0.924824(0.0111)

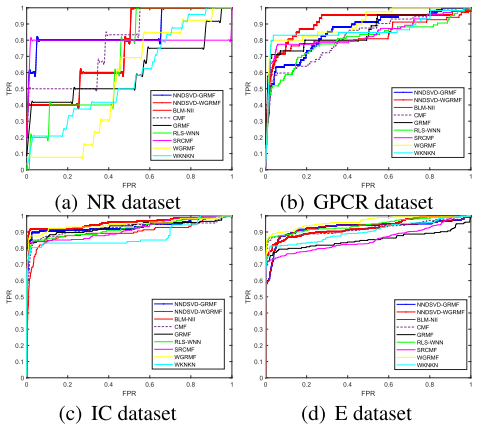


FIGURE 7. ROC curves for different methods are plotted together under CVt on NR dataset, GPCR dataset, IC dataset, E dataset, respectively.

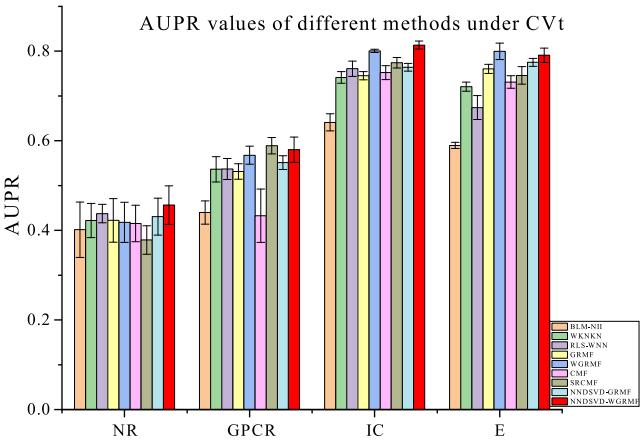


FIGURE 8. In the CVt scenario, the comparison of the AUPR values of different methods in the four datasets.

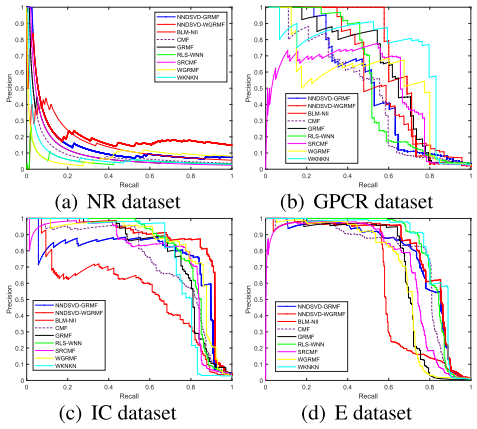
the AUC and AUPR values. In the CVt scenario, compared to the change of $\lambda_d = 0$, the change of $\lambda_t = 0$ and $\lambda_l = 0$ has a greater impact on the AUC and AUPR values. When λ_d, λ_t, λ_l is other parameters, the performance of NNDSVD-GRMF are shown in Fig. 10 and Fig. 11 under CVd and CVt, respectively. Therefore, we can conclude that graph dual regularization and L2 regularization can improve the learning ability of the model, capture more structural information of the graph, and improve the prediction effect.

D. CASE STUDIES

To further demonstrate the effectiveness of NNDSVD-GRMF for predicting DTI, we perform two different types of case studies. In the first case study, we select the target

TABLE 6. AUPR values of different algorithms in CVt scenario.

Method	NR	GPCR	IC	E
BLM-NII	0.40149(0.0618)	0.439848(0.0259)	0.640928(0.0191)	0.589524(0.0069)
WKNKN	0.421919(0.0382)	0.536317(0.0281)	0.741412(0.0131)	0.720789(0.0100)
RLS-WNN	0.437335(0.0206)	0.537046(0.0235)	0.760776(0.0169)	0.674211(0.0266)
GRMF	0.422442(0.0486)	0.531487(0.0175)	0.745256(0.0091)	0.760562(0.0100)
WGRMF	0.417925(0.0447)	0.567606(0.0201)	0.800896(0.0036)	**0.799641**(0.0185)
CMF	0.415443 (0.0407)	0.432831 (0.0596)	0.752132(0.0154)	0.731174(0.0140)
SRCMF	0.378573 (0.0318)	**0.589037**(0.0183)	0.774355 (0.0117)	0.746004 (0.0198)
NNDSVD-GRMF	0.430515 (0.0413)	0.551215 (0.0153)	0.764172(0.0086)	0.775033(0.0090)
NNDSVD-WGRMF	**0.456408** (0.0429)	0.580068 (0.0282)	**0.813607** (0.0088)	0.790897(0.0159)



(a) NR dataset (b) GPCR dataset (c) IC dataset (d) E dataset

FIGURE 9. PR curves for different methods are plotted together under CVt on NR dataset, GPCR dataset, IC dataset, E dataset, respectively.

TABLE 7. AUC results for NNDSVD-GRMF variants under CVd.

Method	NR	GPCR	IC	E
NNDSVD-GRMF	**0.842158**	**0.771546**	**0.782207**	**0.750871**
$\lambda_d = 0$	0.591952	0.671191	0.462721	0.534429
$\lambda_t = 0$	0.823877	0.755510	0.731283	0.738866
$\lambda_l = 0$	0.39899	0.647288	0.620226	0.722817

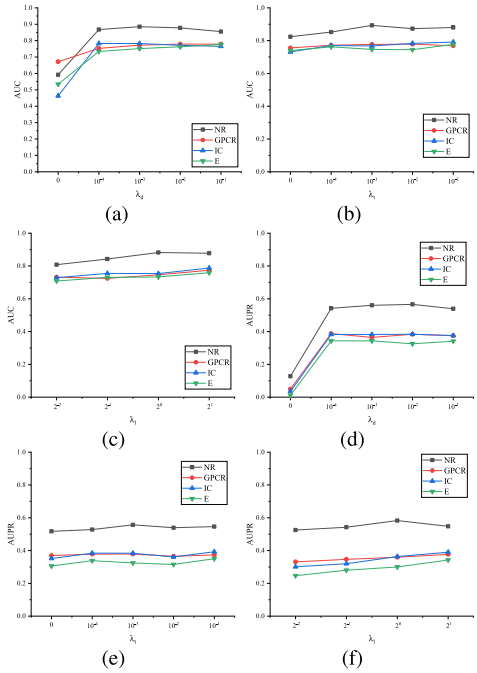
TABLE 8. AUPR results for NNDSVD-GRMF variants under CVd.

Method	NR	GPCR	IC	E
NNDSVD-GRMF	**0.532864**	**0.370845**	**0.375827**	**0.346393**
$\lambda_d = 0$	0.127177	0.0491007	0.0348274	0.011763
$\lambda_t = 0$	0.517326	0.369568	0.351624	0.305765
$\lambda_l = 0$	0.107042	0.295027	0.274093	0.262594

TABLE 9. AUC results for NNDSVD-GRMF variants under CVt.

Method	NR	GPCR	IC	E
NNDSVD-GRMF	**0.811495**	**0.873582**	**0.933894**	**0.933011**
$\lambda_d = 0$	0.713037	0.853216	0.922709	0.905548
$\lambda_t = 0$	0.562456	0.593440	0.499347	0.518962
$\lambda_l = 0$	0.459473	0.782172	0.877774	0.907611

(Androgen receptor) in the NR dataset as the object of the case study. The androgen receptor, as a steroid receptor, not only achieves hormonal regulation of target cells, but it has also been shown to play a role in the proliferation of androgen-refractory cells and in androgen-refractory prostate cancer [35]. For androgen receptors in the NR dataset, androgen receptor interactions with drugs on the NR dataset are



(a) (b) (c) (d) (e) (f)

FIGURE 10. Under CVd, (a), (b) and (c) are the changes of AUC of different parameters λ_d, λ_t and λ_l, respectively; (d), (e) and (f) are the changes of AUPR of different parameter λ_d, λ_t and λ_l, respectively.

TABLE 10. AUPR results for NNDSVD-GRMF variants under CVt.

Method	NR	GPCR	IC	E
NNDSVD-GRMF	**0.430515**	**0.551215**	**0.764172**	**0.775033**
$\lambda_d = 0$	0.429580	0.542078	0.751929	0.760593
$\lambda_t = 0$	0.128248	0.058158	0.037736	0.011504
$\lambda_l = 0$	0.115116	0.273196	0.380932	0.564402

used as the training dataset for NNDSVD-GRMF, and candidate drugs are ranked according to the prediction scores of NNDSVD-GRMF. Then, we select drugs with the top 10 scores and validate them. It can be found that 9 drugs are accurately predicted. The detailed results of the predictions are shown in Table 11.

In the second type of case study, we aim to assess whether NNDSVD-GRMF could be applied to drugs without known interaction targets. We select drug (Nicotine bitartrate) from the IC dataset as the subject of the case study. Nicotine bitartrate, a drug that acts on nicotinic receptors, has been shown to reduce falls in Parkinson's patients [36]. In the case study, the interactions of nicotine bitartrate with the target

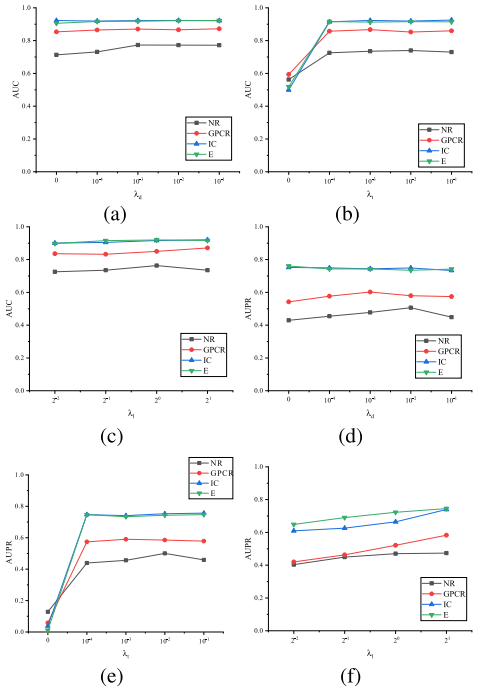


FIGURE 11. Under CVt, (a), (b) and (c) are the changes of AUC of different parameters λ_d, λ_t and λ_l, respectively; (d), (e) and (f) are the changes of AUPR of different parameter λ_d, λ_t and λ_l, respectively.

TABLE 11. Predicted drugs for Androgen receptor in the NR dataset.

Rank	Drug	Drug KEGG id	Evidence
1	Mifepristone	D00585	Unknown
2	**Fluoxymesterone**	**D00327**	KEGG& Drugbank
3	**Flutamide**	**D00586**	KEGG& Drugbank
4	**Nilutamide**	**D00965**	KEGG& Drugbank
5	**Bicalutamide**	**D00961**	KEGG& Drugbank
6	**Oxandrolone**	**D00462**	KEGG& Drugbank
7	**Nandrolone**	**D00956**	KEGG& Drugbank
8	**Testosterone**	**D00075**	KEGG& Drugbank
9	**Progesterone**	**D00066**	Drugbank
10	**Levonorgestrel**	**D00950**	Drugbank

Drugs that interact with Androgen receptor are in bold.

TABLE 12. Predicted targets for Nicotine bitartrate in the IC dataset.

Rank	Target	Target KEGG id	Evidence
1	**CHRNA2**	**hsa1135**	KEGG& Drugbank
2	**CHRNA6**	**hsa8973**	KEGG& Drugbank
3	**CHRNA3**	**hsa1136**	KEGG& Drugbank
4	**CHRNA9**	**hsa55584**	KEGG& Drugbank
5	**CHRNA10**	**hsa57053**	KEGG& Drugbank
6	**BFNC**	**hsa1137**	KEGG
7	**CHRNA7**	**hsa1139**	KEGG& Drugbank
8	**CHRNA1**	**hsa1134**	KEGG
9	**CHRNA5**	**hsa1138**	KEGG& Drugbank
10	BRGDA9	hsa3752	Unknown

Targets that interact with Nicotine bitartrate are in bold.

obtained from the IC dataset are used as a training dataset for NNDSVD-GRMF, and candidate targets are ranked according to prediction scores. Then, we select targets with the top 10 scores and validate them. It can be found that 9 targets are accurately predicted. The detailed results of the predictions are shown in Table 12.

V. CONCLUSION

In order to accurately predict DTIs, we proposed new methods NNDSVD-GRMF based on graph dual regularized matrix factorization. This method can accurately obtain the structural information on the drug and target manifold. In the process of solving the objective function, we use non-negative double singular value decomposition to enhance the initial stage of non-negative matrix factorization. In order to improve adaptability, the weighted form of NNDSVD-GRMF (i.e. NNDSVD-WGRMF) has also been proposed. In the experiment, our method has better performance compared with other state-of-the-art methods. At the same time, our method also has a better performance for certain specific targets and drug predictions.

The better performance of our method may be attributed to the following factors. First of all, NNDSVD-GRMF not only adds graph dual regularization, but also considers the initial stage of matrix factorization. The graph dual regularization enables the model to obtain more information about the data structure. The use of non-negative double singular value decomposition in the initial stage of matrix factorization not only affects the convergence of matrix factorization, but also improves the quality of the solution. Second, as a weighted extension of NNDSVD-GRMF, NNDSVD-WGRMF has better adaptability to datasets. Third, our method may be more suitable for solving the problem of DTIs prediction.

As future work, other multi-manifold regularization methods and network embedding technologies may be used to further improve the prediction performance. For example, in multi-manifold regularization [37], the manifold and coefficient matrix can be combined with multi-manifold regularization to maintain the local geometry of drug and target. When we reduce the dimensionality, the network embedding can preserve the structure of the network [38].

REFERENCES

[1] M. Bagherian, E. Sabeti, K. Wang, M. A. Sartor, Z. Nikolovska-Coleska, and K. Najarian, "Machine learning approaches and databases for prediction of drug-target interaction: A survey paper," *Briefings Bioinf.*, vol. 22, no. 1, pp. 247–269, Jan. 2020, doi: 10.1093/bib/bbz157.

[2] A. C. Cheng, R. G. Coleman, K. T. Smyth, Q. Cao, P. Soulard, D. R. Caffrey, A. C. Salzberg, and E. S. Huang, "Structure-based maximal affinity model predicts small-molecule druggability," *Nature Biotechnol.*, vol. 25, no. 1, pp. 5–71, 2007. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/17211405

[3] M. Campillos, M. Kuhn, A. C. Gavin, L. J. Jensen, and P. Bork, "Drug target identification using side-effect similarity," *Science*, vol. 321, no. 5886, pp. 263–266, 2008.

[4] Y. Yamanishi, M. Araki, A. Gutteridge, W. Honda, and M. Kanehisa, "Prediction of drug-target interaction networks from the integration of chemical and genomic spaces," *Bioinformatics*, vol. 24, no. 13, pp. i232–i240, Jul. 2008.

[5] K. Bleakley and Y. Yamanishi, "Supervised prediction of drug-target interactions using bipartite local models," *Bioinformatics*, vol. 25, no. 18, pp. 2397–2403, 2009, doi: 10.1093/bioinformatics/btp433.

[6] T. Laarhoven, S. B. Nabuurs, and E. Marchiori, "Gaussian interaction profile kernels for predicting drug-target interaction," *Bioinformatics*, vol. 27, no. 21, p. 3036, Nov. 2011.

[7] L. Perlman, A. Gottlieb, N. Atias, E. Ruppin, and R. Sharan, "Combining drug and gene similarity measures for drug-target elucidation," *J. Comput. Biol.*, vol. 18, no. 2, pp. 45–133, 2011.

[8] J. P. Mei, C. K. Kwoh, P. Yang, X. L. Li, and J. Zheng, "Drug-target interaction prediction by learning from local information and neighbors," *Bioinformatics*, vol. 29, no. 2, pp. 238–245, 2013.
[9] T. van Laarhoven and E. Marchiori, "Predicting drug-target interactions for new drug compounds using a weighted nearest neighbor profile," *PLoS One*, vol. 8, no. 6, Jun. 2013, Art. no. e66952.
[10] Y. H. Wang and J. Y. Zeng, "Predicting drug-target interactions using restricted Boltzmann machines," *Bioinformatics*, vol. 29, no. 13, pp. 126–134, 2013.
[11] W. Lan, J. Wang, M. Li, J. Liu, Y. Li, F.-X. Wu, and Y. Pan, "Predicting drug-target interaction using positive-unlabeled learning," *Neurocomputing*, vol. 206, pp. 50–57, Sep. 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0925231216304337
[12] A. S. Rifaioglu, E. Nalbat, V. Atalay, M. J. Martin, R. Cetin-Atalay, and T. Doğan, "DEEPScreen: High performance drug-target interaction prediction with convolutional neural networks using 2-D structural compound representations," *Chem. Sci.*, vol. 11, no. 9, pp. 2531–2557, 2020.
[13] N. Yu, M.-J. Wu, J.-X. Liu, C.-H. Zheng, and Y. Xu, "Correntropy-based hypergraph regularized NMF for clustering and feature selection on multi-cancer integrated data," *IEEE Trans. Cybern.*, vol. 51, no. 8, pp. 3952–3963, Aug. 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9130073/
[14] J.-X. Liu, Z. Cui, Y.-L. Gao, and X.-Z. Kong, "WGRCMF: A weighted graph regularized collaborative matrix factorization method for predicting novel LncRNA-disease associations," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 1, pp. 257–265, Jan. 2021.
[15] C.-N. Jiao, Y.-L. Gao, N. Yu, J.-X. Liu, and L.-Y. Qi, "Hyper-graph regularized constrained NMF for selecting differentially expressed genes and tumor classification," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 10, pp. 3002–3011, Oct. 2020.
[16] C.-Y. Wang, Y.-L. Gao, X.-Z. Kong, J.-X. Liu, and C.-H. Zheng, "Unsupervised cluster analysis and gene marker extraction of scRNA-seq data based on non-negative matrix factorization," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 1, pp. 458–467, Jan. 2022.
[17] M. Gönen, "Predicting drug-target interactions from chemical and genomic kernels using Bayesian matrix factorization," *Bioinformatics*, vol. 28, no. 18, pp. 2304–2310, 2012, doi: 10.1093/bioinformatics/bts360.
[18] X. Zheng, H. Ding, H. Mamitsuka, and S. Zhu, "Collaborative matrix factorization with multiple similarities for predicting drug-target interactions," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*. New York, NY, USA: Association for Computing Machinery, Aug. 2013, pp. 1025–1033, doi: 10.1145/2487575.2487670.
[19] Y. Liu, M. Wu, C. Miao, P. Zhao, and X.-L. Li, "Neighborhood regularized logistic matrix factorization for drug-target interaction prediction," *PLOS Comput. Biol.*, vol. 12, no. 2, Feb. 2016, Art. no. e1004760, doi: 10.1371/journal.pcbi.1004760.
[20] B. Bolgár and P. Antal, "VB-MK-LMF: Fusion of drugs, targets and interactions using variational Bayesian multiple kernel logistic matrix factorization," *BMC Bioinf.*, vol. 18, no. 1, p. 440, Dec. 2017, doi: 10.1186/s12859-017-1845-z.
[21] A. Ezzat, P. Zhao, M. Wu, X.-L. Li, and C.-K. Kwoh, "Drug-target interaction prediction with graph regularized matrix factorization," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 14, no. 3, pp. 646–656, May 2017.
[22] Z. Cui, Y.-L. Gao, J.-X. Liu, L.-Y. Dai, and S.-S. Yuan, "$L_{2,1}$-GRMF: An improved graph regularized matrix factorization method to predict drug-target interactions," *BMC Bioinf.*, vol. 20, no. S8, p. 287, Jun. 2019.
[23] Z. Liu, Q. Chen, W. Lan, H. Pan, X. Hao, and S. Pan, "GADTI: Graph autoencoder approach for DTI prediction from heterogeneous network," *Frontiers Genet.*, vol. 12, Apr. 2021, Art. no. 650821.
[24] L.-G. Gao, M.-Y. Yang, and J.-X. Wang, "Collaborative matrix factorization with soft regularization for drug-target interaction prediction," *J. Comput. Sci. Technol.*, vol. 36, no. 2, pp. 310–322, Apr. 2021, doi: 10.1007/s11390-021-0844-8.
[25] C. Boutsidis and E. Gallopoulos, "SVD based initialization: A head start for nonnegative matrix factorization," *Pattern Recognit.*, vol. 41, no. 4, pp. 1350–1362, Apr. 2008, doi: 10.1016/j.patcog.2007.09.010.
[26] Y.-X. Wang and Y.-J. Zhang, "Nonnegative matrix factorization: A comprehensive review," *IEEE Trans. Knowl. Data Eng.*, vol. 25, no. 6, pp. 1336–1353, Jun. 2013. [Online]. Available: https://ieeexplore.ieee.org/document/6165290/
[27] F. Shang, L. C. Jiao, and F. Wang, "Graph dual regularization non-negative matrix factorization for co-clustering," *Pattern Recognit.*, vol. 45, no. 6, pp. 2237–2250, 2012.
[28] M. Kanehisa, S. Goto, M. Hattori, K. F. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki, and M. Hirakawa, "From genomics to chemical genomics: New developments in KEGG," *Nucleic Acids Res.*, vol. 34, pp. D354–D357, Jan. 2006.
[29] M. Hattori, Y. Okuno, S. Goto, and M. Kanehisa, "Development of a chemical structure comparison method for integrated analysis of chemical and genomic information in the metabolic pathways," *J. Amer. Chem. Soc.*, vol. 125, no. 39, pp. 11853–11865, 2003.
[30] T. Smith and M. Waterman, "Identification of common molecular subsequences," *J. Mol. Biol.*, vol. 147, no. 1, pp. 195–197, Mar. 1981.
[31] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Int. Conf. Neural Inf. Process. Syst.*, 2000, pp. 1–7.
[32] D. Cai, X. He, J. Han, and T. S. Huang, "Graph regularized nonnegative matrix factorization for data representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1548–1560, Aug. 2011.
[33] T. Pahikkala, A. Airola, S. Pietila, S. Shakyawar, A. Szwajda, J. Tang, and T. Aittokallio, "Toward more realistic drug-target interaction predictions," *Briefings Bioinf.*, vol. 16, no. 2, pp. 325–337, Mar. 2015.
[34] J. Bergstra and Y. Bengio, "Random search for hyper-parameter optimization," *J. Mach. Learn. Res.*, vol. 13, no. 2, 2012.
[35] O. L. Zegarra-Moro, L. J. Schmidt, H. Huang, and D. J. Tindall, "Disruption of androgen receptor function inhibits proliferation of androgen-refractory prostate cancer cells," *Cancer Res.*, vol. 62, no. 4, pp. 1008–1013, 2002. [Online]. Available: https://cancerres.aacrjournals.org/content/canres/62/4/1008.full.pdf
[36] A. Lieberman, T. E. Lockhart, M. C. Olson, V. A. S. Hussain, C. W. Frames, A. Sadreddin, M. McCauley, and E. Ludington, "Nicotine bitartrate reduces falls and freezing of gait in Parkinson disease: A reanalysis," *Frontiers Neurol.*, vol. 10, no. 424, 2019. [Online]. Available: https://www.frontiersin.org/article/10.3389/fneur.2019.00424
[37] L. Zong, X. Zhang, L. Zhao, H. Yu, and Q. Zho, "Multi-view clustering via multi-manifold regularized non-negative matrix factorization," *Neural Netw.*, vol. 88, pp. 74–89, Apr. 2017.
[38] Z. Liu, Q. Chen, W. Lan, J. Liang, Y. P. Chen, and B. Chen, "A survey of network embedding for drug analysis and prediction," *Current Protein Peptide Sci.*, vol. 22, no. 3, pp. 237–250, 2021. [Online]. Available: https://www.eurekaselect.com/article/107859

JUNJUN ZHANG received the M.S. degree in mathematics and computational mathematics from the Guilin University of Electronic Science and Technology, Guilin, China, in 2019. He is currently pursuing the Ph.D. degree with the School of Mathematics and Statistics, Hunan Normal University. His research interest includes the application of optimization algorithms to drug association prediction.

MINZHU XIE received the M.E. and Ph.D. degrees in computer science from Central South University, China, in 2003 and 2008, respectively. He is currently a Professor with the College of Information Science and Engineering, Hunan Normal University, China. His current research interest includes bioinformatics.