Int. J. Nonlinear Anal. Appl. 12 (2021) No. 2, 133-142
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/IJNAA.2021.5023

Predicting drug-target interaction based on bilateral local models using a decision tree-based hybrid support vector machine

Ali Ghanbari sorkhi [a*], Majid Iranpour mobarakeh [b], Seyed Mohammad Reza Hashemi [c], Maryam Faridpour[d]

[a] *Assistant Professor, Faculty of Electrical and Computer Engineering, University of Science and Technology of Mazandaran, Behshahr, Iran.*
[b] *Assistant Professor, Department of Computer Engineering and IT, Payam Noor University, Tehran, Iran.*
[c] *Young Researchers and Elite Club Qazvin Branch Islamic Azad University, Qazvin, Iran.*
[d] *Department of Electrical and Computer Engineering, Mahdishahr Branch, Islamic Azad University, Mahdishahr, Iran.*

(Communicated by Dr. Ehsan Kozegar)

Abstract

Identifying the interaction between the drug and the target proteins plays a very important role in the drug discovery process. Because prediction experiments of this process are time consuming, costly and tedious, Computational prediction can be a good way to reduce the search space to examine the interaction between drug and target instead of using costly experiments. In this paper, a new solution based on known drug-target interactions based on bilateral local models is introduced. In this method, a hybrid support vector machine based on the decision tree is used to decide and optimize the two-class classification. Using this machine to manage data related to this application has performed well. The proposed method on four criteria datasets including enzymes (Es), ion channels (IC), G protein coupled receptors (GPCRs) and nuclear receptors (NRs), based on AUC, AUPR, ROC and running time has been evaluated. The results show an improvement in the performance of the proposed method.

Keywords: Drug-target interaction, bilateral local model, decision tree, hybrid SVM

1. Introduction

The study of drug-target interaction (DTI) has attracted the attention of many researchers in the field of pharmaceutical science in recent years [16, 2]. In this regard, many efforts have been made to

*Corresponding Author: Ali Ghanbari sorkhi
Email addresses: ali.ghanbari@mazust.ac.ir (Ali Ghanbari sorkhi [a*]), Iranpour@pnu.ac.ir (Majid Iranpour mobarakeh [b]), Smr.hashemi@Shahroodut.ac.ir (Seyed Mohammad Reza Hashemi [c]), maryam.faridpour68@gmail.com (Maryam Faridpour[d])

Received: June 2020 *Accepted:* January 2021

investigate the reuse of existing drugs to discover the interaction between new targets and existing drugs. In fact, the purpose of DTI is to find the interaction between the drug and the targets that leads to a change in drug behavior / use. On the other hand, identifying these interactions minimizes the adverse side effects of the drug [16]. Web-lab experiments have a major challenge in terms of cost and time [6]. In this regard, computational prediction (CP) methods have been used in recent years [3]. Despite the synthesis of a large number of compounds, their target profiles and drug effects are still unknown. In addition, there is no cure for many diseases and many new diseases are reported each year. Due to the importance of this issue, many efforts have been made by researchers to gather information about the various properties of drug compounds, properties, responses and target proteins which has created a large dataset on the interaction between drugs and proteins. Creating large datasets has led to the use of CP with problems such as data complexity and large dimensions, which indicates the need for efficient and powerful algorithms in predicting DTI.

In general, three categories can be introduced for computational methods in DTI [3]. The first category is called ligand-based methods, in which similar drugs tend to share similar properties and usually bind similar proteins [11], which this method predicts interaction using similarities between similar protein ligands. Since these methods do not use proteins sequence information for prediction, it is possible that a new interaction is limited to the relation between known ligands and protein families. On the other hand, the performance of these methods is highly depended on known ligands and if their number is low for a candidate protein, the performance of these methods is greatly reduced [14]. The second category uses three-dimensional structures of drugs and proteins using simulations to predict interaction. These methods are introduced as connection approaches. The main problem of these methods is the lack of known three-dimensional structure of some proteins and their high computational complexity [13, 24]. The third category is chemical-based methods. These methods use information about the drug and the target at the same time. These methods have been considered by many researchers in recent years and can be widely used on biological data that are more accessible [23]. In fact, these methods use data that contains information about processes to predict simultaneously. Process information refers to diagrams of chemical structure and genomic sequences for drugs and targets. This general method can be divided into two categories: feature-based methods and similarity-based methods. In the feature-based method, supervised machine learning methods are used. In fact, in these methods, feature vectors use a set of drug-target pairs with a class label that indicates the presence of interaction (positive sample) or lack of interaction (negative sample) [5, 9]. Of course, it should be noted that negative samples are samples without the drug-target or unknown interaction. In similarity-based methods, two similarity matrices related to drugs and similarity of targets along with the interaction matrix are used respectively, which represents the interaction between drug and targets pairs [20, 1].

In this paper, a bilateral local model based method is used to identify drug-target interaction. In this regard, the drug-drug, target-target matrices and similarity between them have been used. An important phase in local models is decision and classification based on rules extraction. In the proposed method, a combined support vector machine (SVM) based on the decision tree is used. This method focuses on hard data that is close to the decision boundary. In fact, a combined SVM based on the decision tree is used to classify data that is close to the decision boundary. Since the data related to drug-target interaction are unknown in many cases and also these data are difficult data, the use of this method improves performance, on the other hand, since only this type of data by SVM is categorized in the test and for the rest of the data the decision tree is used for classification, the running time of the algorithm is reduced. The details of the proposed method and the steps of the algorithm are described in the following sections.

In the following of this article, in the second section, the proposed method is briefly introduced.

In this section, the bilateral local model and decision based on the combined support vector machine based on the decision tree are described in detail. In the next section, experiments and analysis of the proposed method results which includes the introduction of the database, evaluation criteria and analysis of results in terms of running time is described in detail. Finally, the conclusions and future works are introduced in the fourth section.

2. The proposed method

In this paper, a new approach to predicting new edges (interactions) by a bilateral network is presented. In this method, information about the target-drug (graph vertices) is used. For more precise expression, $V_d = \{d_1, d_2, \cdots, d_m\} and V_t = \{t_1, t_2, \cdots, t_n\}$ were assumed to be a set of drugs (or potential drugs) and a set of target proteins (or potential targets), respectively. In addition, it is assumed that each drug d_i and target t_j is characterized by a set of related biological properties. Putting the e_{ij} edge between the drug d_i and the target t_j represents the drug's interaction with that target protein which is seen throughout the drug-target interactions set, thus it creates a bilateral graph, a graph in which the edges only allow to pass from one class of vertices (drugs) to another class (targets).

2.1. Bilateral graph inference with local models

This paper presents a solution to the bilateral graph inference problem by training several local models to predict new edges that connect the drug vertices (V_d) to the target vertices (V_t). More precisely, in fact, the presence or absence of the e_{ij} edge between the drug d_i and the target t_j is predicted by algorithm (1).

Algorithm (1) - A bilateral algorithm to check for the presence or absence of an edge between the drug and the target

Input: Drug vertices (V_d), target vertices (V_t), known interaction of (V_d) and (V_t)

1. With the exception of target t_j,
   a. Preparing a list of other known targets d_i in a bilateral network
   b. Preparing a separate list of unknown targets as target by d_i
   c. Known targets are labeled $+1$ and other targets are labeled -1.
2. Using classification to distinguish between data of two classes based on the genomic sequence of targets.
3. Using classification to predict t_j label. In fact, with this process, the presence or absence of the edge between d_i and t_j is checked.
4. The same target is considered constant t_j and then, with the exception of the drug d_i,
   a. Preparing a list of other known drugs by targeting t_j in a bilateral network
   b. Preparing a list of drugs that are not known to targeting t_j.
   c. Similar to the previous steps, known drugs for the t_j target are labeled $+1$ and other drugs are labeled -1.

5. Using classification to distinguish between data of two classes based on the chemical structure of drugs.
6. Using classification to predict d_i label. In fact, in this process, the presence or absence of the edge between d_i and t_j is checked.
Output: Bilateral graph of interaction between drug vertices (V_d), target vertices (V_t)

In Algorithm 1, steps 4 to 6 of this method are very important. The purpose of these steps is to make a second independent prediction of the same edge, if possible. Even if asked to predict exactly one edge in both modes, this is done with a different data set in each mode and using a different classification rule (or a class of rules). This makes two independent predictions for the same edge. In practice, two conditions may occur: in the first mode, the drug may have no known target or in the second mode, the target has no known drug for targeting.

In general, three modes can be considered for the interaction between the drug and the target:

- The drug has no known target and target has at least one known drug for targeting.
- The target has no known drug for targeting and the drug has at least one known target.
- The drug has at least one known target and the target has at least one known drug to target.

The first two modes reflect a situation in which predict unknown interactions involving new combinations of candidate drugs or candidate target proteins out of the training dataset. The third mode is a kind of dual usage of the algorithm that aims at two independent predictions for the same edge. Basically, after two predictions, a function is used to aggregate two predicted values for the same edge. In this paper, the maximum function is used for aggregation.

Since SVM is a local classification, a combination SVM based on the decision tree of the new approach has been used for classification. Based on biological data about the vertices (drugs or targets), the combination SVM based on the decision tree is learned through the labels of these vertices, a function with real values that can assign a score to the drug or the remaining target. In this local model, this process is equivalent to allocating scores to the left edge. Although the prediction -1, +1 is usually obtained by not considering the sign of this score, but the value of the score itself includes a kind of confidence to the prediction. In this paper, all candidate edges are ranked according to the value of the combined SVM prediction based on the decision tree. In cases where there are two scores for the candidate edges, first a rule is chosen to convert these two scores into one score and then this score is ranked.

2.2. Combined SVM based on decision tree

In this paper, a combined SVM method based on the decision tree is used. This fast method is used in the training stage for binary classification. In this method, the focus is on reducing the number of test data points that are classified by SVM, thus the total time consumed in the test is reduced. It is first categorized by the data point tree into data points that are near or far to the SVM decision boundary. Far nodes are classified only by the direct decision tree, while more accurate points require SVM for high accuracy, so the problem involves both univariate and multivariate (SVM) nodes. The combined tree uses SVM only to classify important data points near the decision boundary. The less important points given by the quick univariate nodes are classified without any compromise on the accuracy of the classification. Details of the proposed method are explained below.

The time complexity of SVM can be represented by equation (1):

$$Complexity_{SVM} = d_{numX} O(nN_{sv}) \quad (1)$$

n this regard, d_{numX} is the total number of data points, n is the input dimension, and N_{sv} is the number of support vector hyper planes. Most methods try to reduce the number of SVs which at most can be reduced to some extent; otherwise they will reduce the accuracy. As mentioned above, the combined SVM method focuses on reducing the number of classified data by SVM, and this can achieve high speed without losing accuracy.

The SVM decision function can be considered as equation (2):

$$f(x) = \sum_{i=1}^{N_{sv}} \alpha_i D_i K(x, X_i) + b \quad (2)$$

n this equation,$X_i \in R^n$ is a support vector, D_i is related to the target values (if $D_i = 1$, it is related to class 1 data points, otherwise if $D_i = -1$, it is related to class -1 data points), α_i Lagrangian coefficient, b Bias, N_{sv} the number of support vectors, K the kernel function, and $x \in R^n$ the new data point being tested.

$f(x)$ is a SVM decision function which decision boundary is defined by $f(x) = 0$. This decision boundary divides the feature space into two distinct unique regions, the positive region $f(x) > 0$ (consisting of all predicted points of the training data as class 1) and the negative region $f(x) < 0$ (consisting of all predicted points of the training data as class 2). Based on this decision boundary, closeness measurement $S(x)$ is defined in which calculates small values for data points close to $f(x) = 0$ and large values for data points far from $f(x) = 0$. By selecting a threshold parameter δ, a region is defined around the decision boundary, so that this region includes all training data points by measuring the closeness of $S(x) \leq \delta$. In this section, predictions for all training data points in the δ-region are labeled as class 3. The feature space now has training data points along with its predictions, which represent three different regions in Figure (1).

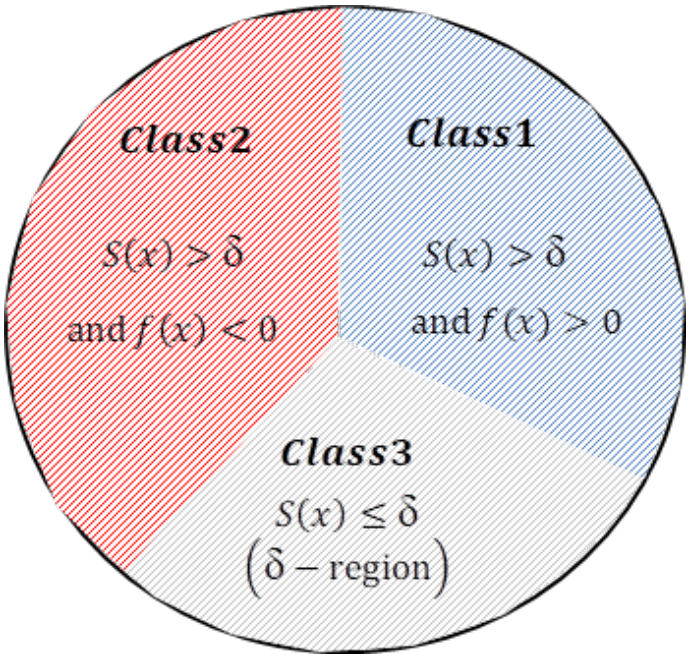

Figure 1: regions related to the three classes based on the decision function f(x) and the closeness and threshold criteria

When training data points with their predictions are available to represent these three regions, the combined SVM decision method based on the decision tree can be used in two steps. First,

a decision tree is trained with this 3-class dataset to identify approximately these three regions. After the decision tree training, the second step is to replace each leaf of class 3 by a sub tree with binary SVM and two leaves. In fact, in this method a once trained binary SVM can be used in several leaves of decision tree (DT). This method consists of both conventional variable decision nodes and multivariate decision nodes (SVM). Univariate decision nodes help make quick decisions for less valuable test data points without the use of multivariate SVM. On the other hand, if the test data points are very important, the univariate decision nodes lead them to the multivariate SVM. Therefore, only a small portion of the test data points are classified with SVM nodes, and the rest are classified much faster using univariate decision nodes, thus the complexity of the total time is reduced.

2.2.1. Measuring closeness and threshold

To identify the δ region, the criterion of closeness between the training data points and the decision boundary $f(x) = 0$ is defined. In this paper, the proposed probable output of the SVM method [12] is used as an adjacent measurement. In general, the SVM decision function $f(x)$ produces uncalibrated values and can be converted to posterior probability estimation by placing a sigmoid function at its output.

$$P(Class1|f(x)) = \frac{1}{(1 + exp(-f(x)))} \quad (3)$$

The above phrase can be modified as follows:

$$\Delta P(x) = P(Class1|f(x)) - 0.5 = \frac{1}{(1 + exp(-f(x)))} - 0.5 \quad (4)$$

Where $\Delta P(x)$ shows the closeness criterion $S(x)$ between the training data points and the decision boundary $f(x) = 0$. It can be noted that, for $f(x) = 0$; $S(x) = 0$, when $f(x) \to \infty$; $S(x) \to 0.5$ and $f(x) \to -\infty$; $S(x) \to -0.5$. So we have a reasonable threshold parameter δ.

The training step of this method is stated in Algorithm (2).

Algorithm (2) - Combined SVM training with decision tree

Input: Binary dataset, δ threshold

1. SVM training with train data and obtaining decision function $f(x)$
2. Classify train data with $f(x)$ in class 1 or class 2, saving predictions in new target
3. Specify data points in the train data with $S(x) \leq \delta$ and changing the predictions of these points in the new target to class 3
4. Decision tree training with train data with new target labels (3 classes) Replacing all leaves of class 3 from the decision tree with a sub tree with SVM and two leaves

The steps of the proposed method are shown in Figure (2).

It should be noted that in this paper the δ threshold value is considered 0.25 according to the validation set.

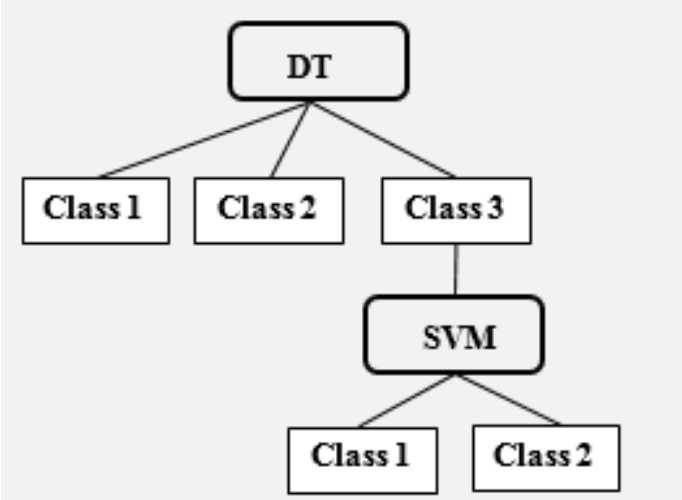

Figure 2: Steps for classification based on the proposed method

3. Experiments and analysis of results

In this section, the experiments and results of the proposed method are analyzed separately. In the first step, the dataset and how to divide it into training and experimental sets are introduced. All experiments and extracted parameters are performed for each dataset under different Cross-Validation Settings (CVS). Then, the proposed method is compared with other new methods based on the Area Under Receiver Operating Characteristic Curve (AUC), the Area Under Precision Recall Curve (AUPR), the Receiver Operating Characteristic (ROC) and the running time.

3.1. Data set

Yamanishi et al. reviewed information on drugs and target proteins interactions for public databases [23] such as KEGG BRITE [10], RENDA [18], SuperTarget [7] and DrugBank [22]. In this paper, four benchmark datasets are used [14, 23, 15] which are from four different classes of target protein. In fact, these criteria are simulated from public databases. The following is a description of this dataset:

- Enzymes (Es): In this dataset, 445 drugs, 664 targets and 2926 interactions have been extracted.
- Ion channels (IC): In this dataset, 201 drugs, 204 targets and 1476 interactions have been extracted.
- G-protein-coupled receptors (GPCRs): In this dataset, 223 drugs, 95 targets and 635 interactions were extracted.
- Nuclear Receptors (NRs): In this dataset, 54 drugs, 26 targets and 90 interactions have been extracted.

It should be reminded that these datasets are simulated from public databases that are publicly available at: http://web.kuicr.kyoto-u.ac.jp/supp/yoshi/drugtarget/.

In this paper, a similarity matrix between drug-drug and target-target has been used to extract the feature. The similarity of SIMCOMP [8] which is based on the chemical structure of drugs has been used for the drug. Protein sequence information has been used for the target and also the Smith–Waterman scoring method [19] has been used for similarity of targets.

In this paper, Leave-One-Out (LOO) cross-validation setting is used to classify the data. As mentioned in the previous sections, the proposed method is a bilateral way, one side for drug prediction and the other side for target prediction, the combination of these predictions represents the drug-target interaction. In this regard, in the results section, the dataset is considered in three modes. In the article, CVS for drug prediction, CVS for target prediction and interaction prediction are introduced as CVS1, CVS2 and CVS3, respectively [3]. This division is as follows:

- CVS1 / Drug Prediction: All drug profiles are discarded for use as an experiment datasets. It tests the algorithm's ability to predict interactions of new drugs which no cross-information is available.
- CVS2 / Target Prediction: All target profiles are discarded for use as an experiment datasets. It tests the algorithm's ability to predict interactions for new targets.
- CVS3 / Pair Prediction: Random drug-target pairs are discarded as an experiment datasets for prediction. This is a common setting for validation and evaluation.

3.2. Evaluation criteria

To evaluate the performance of the proposed method, the Area Under ROC Curve (AUC) and the Area Under Precision Recall Curve (AUPR) criteria have been used [3]. The criteria are introduced in the following:

- The ROC curve has received more attention in supervised machine learning. One of the ways of checking and evaluating performance is binary classification. The performance of binary classification algorithms is usually measured by indicators called sensitivity or recovery.

But in the ROC graph, both of these indicators are combined and displayed as a curve. ROC curve is often used to evaluate the performance of classification algorithms or to generate classified data. In the case of DTI, this curve is usually used to extract the AUC criterion.

- AUC is a criterion for evaluating the performance of the proposed method in ranking. It uses the ROC curve, which is a graphical graph that shows the ability to detect a positive rate for a method as a function of a false positive rate.

The AUC measures all two-dimensional areas under the ROC curve. Summary of decision thresholds is often used as a measure criterion of classification performance. The AUC interpretation shows a better model in that a random positive example is more than a random negative example.

- AUPR is another criterion used to evaluate the performance of DTI methods in this paper which uses the Precision Recall Curve. This criterion is a ratio of real positive graphs that show positive predictions for each recall rate. AUPR performance evaluation shows that this area under the Precision-Recall Curve punishes more false positives than AUC. AUPR shows a quantitative evaluation of the separation of real interactions from real non-interactions among the predicted scores. Due to the low presence of real drug-target interactions, AUPR is a more important qualitative criterion than AUC that finds real drug-target interactions among predicted scores.

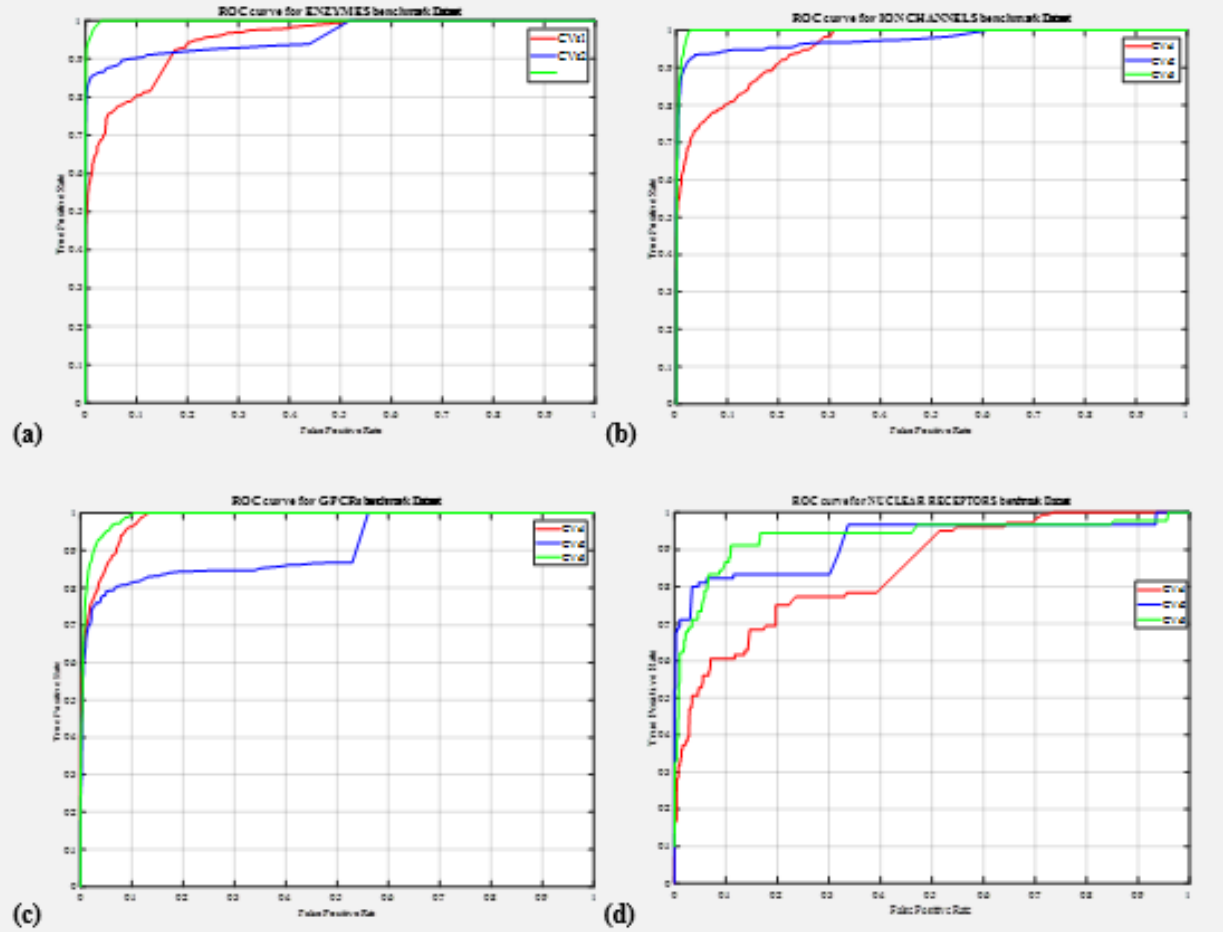

Figure 3: Results obtained based on the ROC curve in four benchmark datasets (a) Es, (b) IC, (c) GPCRs and (d) NRs criteria. In each chart, red is LOO for drug prediction (CVS1), green is LOO for target prediction (CVS2), and blue is LOO for interaction prediction (CVS3).

3.3. Analysis of experiment results

In this section, the results are analyzed based on the ROC curve criterion which indicates the performance of the proposed method in four benchmark datasets.

In the following the ROC curve criterion is used for more detailed analysis.

Figure (3) shows the performance of the proposed method in four benchmark datasets. In this figure, the results obtained in the three CVS are shown separately. According to the results shown, the proposed method has the best performance in all four benchmark datasets in CVS3 (interactions). This point is very valuable because it shows that the combination of two edges related to drug-drug prediction and target-target prediction has had a significant impact on predicting interactions in the proposed method. In general, according to the results, the proposed method shows appropriate performance based on this criterion.

In the following, to evaluate the running time of the algorithm, the proposed method is compared with the standard SVM. In fact, in this method, instead of using a combination SVM based on the decision tree, the standard SVM is used. In the bilateral local model, SVM is used for classification. Table (1) reports the results obtained from running time for four benchmark datasets considering different CVS. The results show an appropriate speed of the proposed method than SVM. The average of running time for each sample is given in seconds.

Table 1: Comparison of the running time of the proposed method based on the combined SVM based on the decision tree and SVM

	CVS1		CVS2		CVS3	
	Proposed Method	SVM	Proposed Method	SVM	Proposed Method	SVM
Es	**0.43**	1.53	**0.65**	1.4	**1.80**	2.9
IC	**0.55**	1.86	**0.87**	1.6	**1.42**	3.5
GPCRs	**0.63**	2.25	**0.88**	2.4	**1.51**	4.70
NRs	**0.96**	3.50	**1.05**	2.8	**2.01**	6.3

Figure (4) provides a comparison of the proposed method based on the AUC and AUPR criteria with the standard SVM. It should be noted that all the parameters used in the SVM of proposed method and the standard SVM are considered similar. As it can be shown from the results obtained from Table (1), Figure (4-a) and Figure (4-b), the proposed method has better performance than the standard SVM method in terms of running time and AUC and AUPR evaluation criteria. This is due to the use of data training based on adjacent to the decision boundary as well as the use of the DT classifier with the SVM classifier.

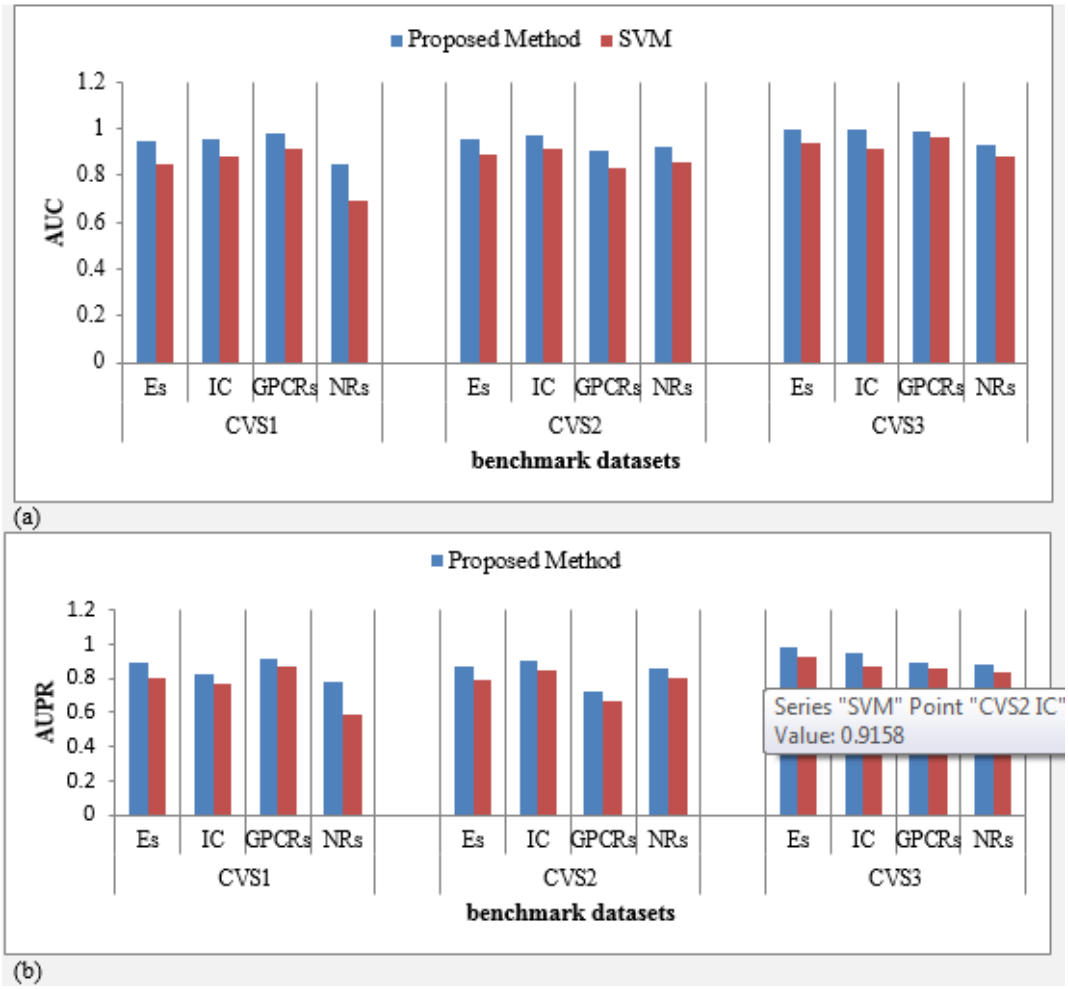

Figure 4: Comparison of the proposed method based on the combined SVM based on the decision tree and the base SVM, a) Comparison based on the AUC criterion and b) Comparison based on the AUPR criterion

Figure (5) shows part of the predicted output for NRs data. In this graph, the edges represent the interaction between the drug and the target. Solid edges represent known interactions, and the dashed line edge is for interactions that are unknown in the dataset and the proposed method is recognized as interaction. Using these prediction edges can be suggesting new potential drug-target interactions. In this graph, the circles represent the drugs and the rectangles represent the target proteins. The values inside the vertex of the rectangle and the vertex of the circle are the target number and the drug number, respectively.

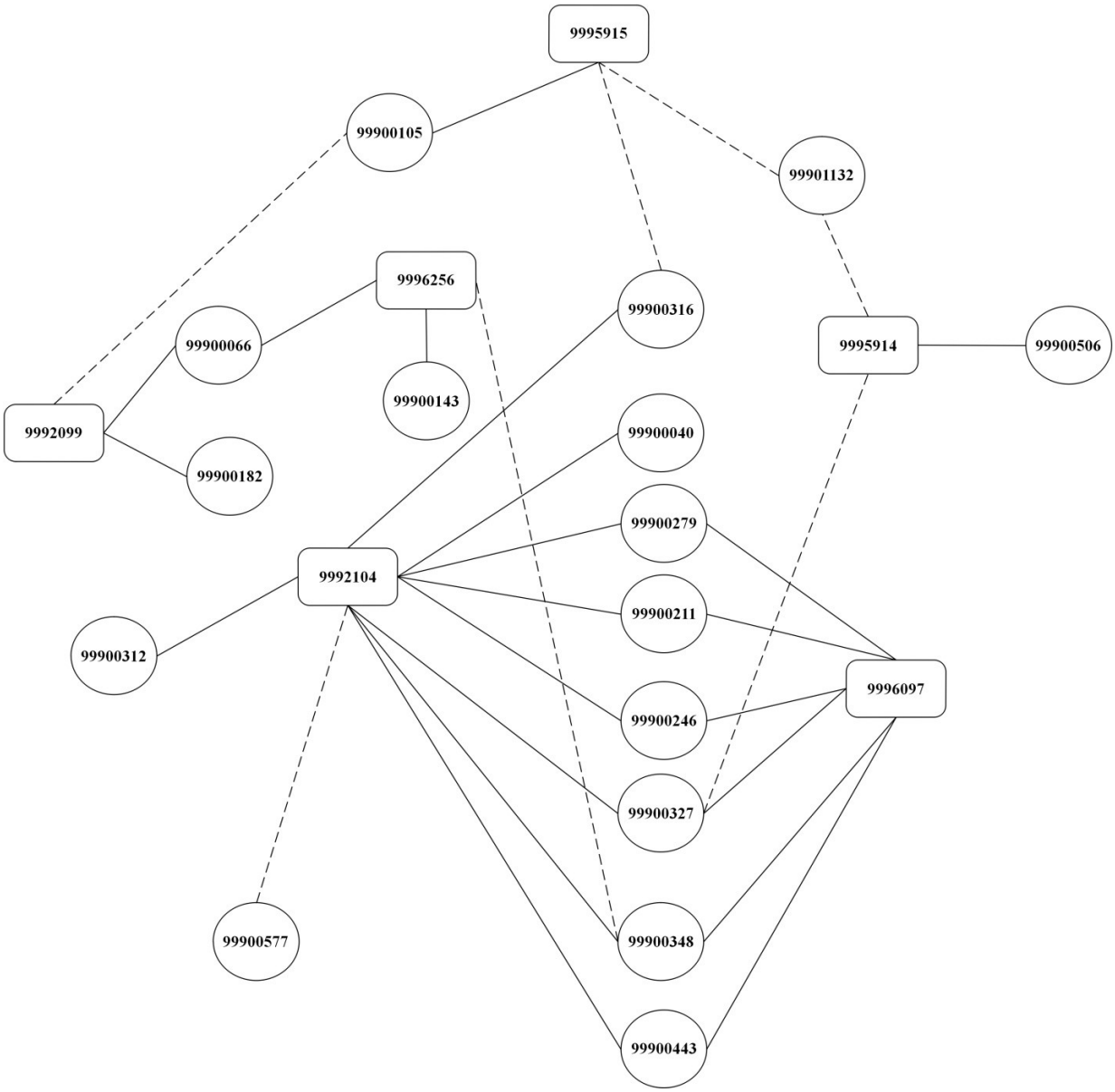

Figure 5: Examples of drug-target interactions in the NRs benchmark datasets, circles represent drugs and rectangles represent target proteins. The solid lines indicate known interactions that the proposed method has correctly identified, dashed lines are unknown interactions detected by the proposed method that may exist between interactive drug and target. The values inside the vertex of the rectangle and the vertex of the circle are the target number and the drug number, respectively

3.4. Comparison with the others methods

In this section, a comparison of the proposed method with the state-of-the-art methods in recent years is presented. In this paper, some methods have been used for comparison [14, 26, 21]. It should be noted that all methods are done from the same dataset and the same CVS. The results of others' study are extracted from articles. The method presented by Mongia et al. [14] has reported the best results based on similarity criteria in four benchmark datasets in 2020. The results based on AUC and AUPR criteria are shown in the following tables. Table (2), Table (3) and Table (4) show the comparison of methods based on AUC criteria. Table (2) shows the results obtained from the methods in four benchmark datasets with CVS1. Similarly, Table (3) and Table (4) show the results based on CVS2 and CVS3, respectively. As can be observed from the results, the proposed method has had an acceptable performance in evaluating DTI in the same dataset compared to other new methods. A very important point in these tables is related to the prediction of the target-drug pair that is shown in CVS3. In Table (4), the proposed method has the best performance in three datasets and the results are very close to Mongia et al. results [14] in the NRs dataset.

Table 2: Comparison of the proposed method with other methods based on AUC criteria in four datasets in CVS1

CVS1	[7]	[25]	[26]	Proposed Method
Es	**0.9683**	0.9272	0.9067	0.9508
IC	**0.9541**	0.9368	0.9286	0.9526
GPCRs	0.8975	0.8966	0.8694	**0.9824**
NRs	0.7502	0.8373	0.8124	**0.8477**

Table 3: Comparison of the proposed method with other methods based on AUC criteria in four datasets in CVS2

CVS2	[7]	[25]	[26]	Proposed Method
Es	0.9460	0.7755	0.7952	**0.9597**
IC	0.9714	0.7669	0.7576	**0.9751**
GPCRs	**0.9567**	0.8800	0.8067	0.9078
NRs	**0.9533**	0.8615	0.8124	0.9201

Table 4: Comparison of the proposed method with other methods based on AUC criteria in four datasets in CVS2

CVS3	[7]	[25]	[26]	Proposed Method
Es	0.9955	0.9705	0.9635	**0.9989**
IC	0.9947	0.9832	0.9786	**0.9969**
GPCRs	0.9785	0.9493	0.9458	**0.9902**
NRs	**0.9660**	0.8679	0.9329	0.9339

Table (5), Table (6) and Table (7) show a comparison of methods based on the AUPR criterion. Table (5) shows the results obtained from the methods in the four benchmark datasets in CVS1. Similarly, Table (6) and Table (7) show the results based on CVS2 and CVS3, respectively. As can be observed from the results, the proposed method has had an acceptable performance in the evaluation of DTI (CVS3) in three benchmark datasets of Es, GPSRs and NRs compared to other new methods. In the IC dataset, the results are very close to Mongia et al. [14].

Table 5: Comparison of the proposed method with other methods based on AUPR criteria in four datasets in CVS1

CVS1	[7]	[25]	[26]	Proposed Method
Es	**0.9041**	0.7808	0.5465	0.8918
IC	**0.9541**	0.7786	0.7437	0.8258
GPCRs	0.8975	0.5989	0.5397	**0.9170**
NRs	0.7502	0.4774	0.4907	**0.7812**

Table 6: Comparison of the proposed method with other methods based on AUPR criteria in four datasets in CVS2

CVS2	[7]	[25]	[26]	Proposed Method
Es	0.8603	0.3848	0.2409	**0.8664**
IC	0.9026	0.3538	0.3090	**0.9050**
GPCRs	**0.8538**	0.4059	0.3463	0.7209
NRs	**0.8773**	0.5203	0.5373	0.8574

Table 7: Comparison of the proposed method with other methods based on AUPR criteria in four datasets in CVS3

CVS3	[7]	[25]	[26]	Proposed Method
Es	0.9660	0.8837	0.8093	**0.9754**
IC	**0.9585**	0.9373	0.8459	0.9484
GPCRs	0.8515	0.7543	0.6933	**0.8922**
NRs	0.8791	0.6383	0.7072	**0.8821**

4. Conclusion

In this paper, a new approach to identifying drug-target interaction is introduced. In the proposed method, a bilateral local model method using combined SVM based on the decision tree is presented. This local method actually considers both drug-drug and target-target effects simultaneously using similarities between these matrices. The use of a combination SVM based on the decision tree improves the reduction of running time in the test phase. Four benchmark datasets have been used to evaluate the proposed method. For accurate evaluation of the method and similar comparison with other methods, different modes of the dataset have been considered in order to predict drugs, predict targets, and predict drug-target interaction. The results obtained in the proposed method show an improvement in the performance of the proposed method based on AUC, AUPR and running time.

References

[1] F. Cheng, C. Liu, J. Jiang, W. Lu, W. Li, G. Liu, et al., "Prediction of drug-target interactions and drug repositioning via network-based inference," PLoS Comput Biol, vol. 8, p. e1002503, 2012.

[2] Y. Deng, X. Xu, Y. Qiu, J. Xia, W. Zhang, and S. Liu, "A multimodal deep learning framework for predicting drug-drug interaction events," Bioinformatics, 2020.

[3] A. Ezzat, M. Wu, X.-L. Li, and C.-K. Kwoh, "Computational prediction of drug–target interactions using chemogenomic approaches: an empirical survey," Briefings in bioinformatics, vol. 20, pp. 1337-1357, 2019.

[4] A. Ezzat, P. Zhao, M. Wu, X.-L. Li, and C.-K. Kwoh, "Drug-target interaction prediction with graph regularized matrix factorization," IEEE/ACM transactions on computational biology and bioinformatics, vol. 14, pp. 646-656, 2016.

[5] A. Ezzat, M. Wu, X.-L. Li, and C.-K. Kwoh, "Drug-target interaction prediction via class imbalance-aware ensemble learning," BMC bioinformatics, vol. 17, pp. 267-276, 2016.

[6] S. Fakhraei, B. Huang, L. Raschid, and L. Getoor, "Network-based drug-target interaction prediction with probabilistic soft logic," IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 11, pp. 775-787, 2014.
[7] S. Günther, M. Kuhn, M. Dunkel, M. Campillos, C. Senger, E. Petsalaki, et al., "SuperTarget and Matador: resources for exploring drug-target relationships," Nucleic acids research, vol. 36, pp. D919-D922, 2007.
[8] 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," Journal of the American Chemical Society, vol. 125, pp. 11853-11865, 2003.
[9] Z. He, J. Zhang, X.-H. Shi, L.-L. Hu, X. Kong, Y.-D. Cai, et al., "Predicting drug-target interaction networks based on functional groups and biological features," PloS one, vol. 5, p. e9603, 2010.
[10] M. Kanehisa, S. Goto, M. Hattori, K. F. Aoki-Kinoshita, M. Itoh, S. Kawashima, et al., "From genomics to chemical genomics: new developments in KEGG," Nucleic acids research, vol. 34, pp. D354-D357, 2006.
[11] M. J. Keiser, B. L. Roth, B. N. Armbruster, P. Ernsberger, J. J. Irwin, and B. K. Shoichet, "Relating protein pharmacology by ligand chemistry," Nature biotechnology, vol. 25, pp. 197-206, 2007.
[12] M. A. Kumar and M. Gopal, "A hybrid SVM based decision tree," Pattern Recognition, vol. 43, pp. 3977-3987, 2010.
[13] H. Li, Z. Gao, L. Kang, H. Zhang, K. Yang, K. Yu, et al., "TarFisDock: a web server for identifying drug targets with docking approach," Nucleic acids research, vol. 34, pp. W219-W224, 2006.
[14] A. Mongia and A. Majumdar, "Drug-target interaction prediction using Multi Graph Regularized Nuclear Norm Minimization," Plos one, vol. 15, p. e0226484, 2020.
[15] A. Mongia, V. Jain, E. Chouzenoux, and A. Majumdar, "Deep Latent Factor Model for Predicting Drug Target Interactions," in ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 1254-1258.
[16] K. Sachdev and M. K. Gupta, "A comprehensive review of feature based methods for drug target interaction prediction," Journal of biomedical informatics, vol. 93, p. 103159, 2019.
[17] K. Sachdev and M. K. Gupta, "A hybrid ensemble-based technique for predicting drug–target interactions," Chemical Biology & Drug Design, 2020.
[18] I. Schomburg, A. Chang, C. Ebeling, M. Gremse, C. Heldt, G. Huhn, et al., "BRENDA, the enzyme database: updates and major new developments," Nucleic acids research, vol. 32, pp. D431-D433, 2004.
[19] T. F. Smith and M. S. Waterman, "Identification of common molecular subsequences," Journal of molecular biology, vol. 147, pp. 195-197, 1981.
[20] M. A. Thafar, R. S. Olayan, H. Ashoor, S. Albaradei, V. B. Bajic, X. Gao, et al., "DTiGEMS+: drug–target interaction prediction using graph embedding, graph mining, and similarity-based techniques," Journal of Cheminformatics, vol. 12, pp. 1-17, 2020.
[21] T. Van Laarhoven and E. Marchiori, "Predicting drug-target interactions for new drug compounds using a weighted nearest neighbor profile," PloS one, vol. 8, p. e66952, 2013.
[22] D. S. Wishart, C. Knox, A. C. Guo, D. Cheng, S. Shrivastava, D. Tzur, et al., "DrugBank: a knowledgebase for drugs, drug actions and drug targets," Nucleic acids research, vol. 36, pp. D901-D906, 2008.
[23] 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, pp. i232-i240, 2008.
[24] M. A. Yıldırım, K.-I. Goh, M. E. Cusick, A.-L. Barabási, and M. Vidal, "Drug—target network," Nature biotechnology, vol. 25, pp. 1119-1126, 2007.
[25] H. Yu, J. Chen, X. Xu, Y. Li, H. Zhao, Y. Fang, et al., "A systematic prediction of multiple drug-target interactions from chemical, genomic, and pharmacological data," PloS one, vol. 7, p. e37608, 2012.
[26] X. Zheng, H. Ding, H. Mamitsuka, and S. Zhu, "Collaborative matrix factorization with multiple similarities for predicting drug-target interactions," in Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, 2013, pp. 1025-1033.