

RESEARCH ARTICLE

Mining Chemical Activity Status from High-Throughput Screening Assays

Othman Soufan¹, Wail Ba-alawi¹, Moataz Afeef¹, Magbubah Essack¹, Valentin Rodionov², Panos Kalnis³, Vladimir B. Bajic^{1*}

1 King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center (CBRC), Thuwal 23955–6900, Saudi Arabia, **2** King Abdullah University of Science and Technology (KAUST), KAUST Catalysis Center (KCC), Thuwal 23955–6900, Saudi Arabia, **3** King Abdullah University of Science and Technology (KAUST), Infocloud Group, Computer, Electrical and Mathematical Sciences and Engineering Division (CEMSE), Thuwal 23955–6900, Saudi Arabia

* vladimir.bajic@kaust.edu.sa



OPEN ACCESS

Citation: Soufan O, Ba-alawi W, Afeef M, Essack M, Rodionov V, Kalnis P, et al. (2015) Mining Chemical Activity Status from High-Throughput Screening Assays. *PLoS ONE* 10(12): e0144426. doi:10.1371/journal.pone.0144426

Editor: A Ganesan, University of East Anglia, UNITED KINGDOM

Received: September 7, 2015

Accepted: November 18, 2015

Published: December 14, 2015

Copyright: © 2015 Soufan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The datasets and implementation of all solutions are available as a MATLAB toolbox online at www.cbrc.kaust.edu.sa/dramote and can be found on Figshare: http://figshare.com/articles/Datasets_Mining_chemical_activity_status_from_high_throughput_screening_assays/1598200 http://figshare.com/articles/Toolbox_Mining_chemical_activity_status_from_high_throughput_screening_assays/1601833.

Funding: This study is supported by the King Abdullah of Science and Technology (KAUST) base research funds of VBB and PK. The funders had no

Abstract

High-throughput screening (HTS) experiments provide a valuable resource that reports biological activity of numerous chemical compounds relative to their molecular targets. Building computational models that accurately predict such activity status (active vs. inactive) in specific assays is a challenging task given the large volume of data and frequently small proportion of active compounds relative to the inactive ones. We developed a method, DRAMOTE, to predict activity status of chemical compounds in HTP activity assays. For a class of HTP assays, our method achieves considerably better results than the current state-of-the-art-solutions. We achieved this by modification of a minority oversampling technique. To demonstrate that DRAMOTE is performing better than the other methods, we performed a comprehensive comparison analysis with several other methods and evaluated them on data from 11 PubChem assays through 1,350 experiments that involved approximately 500,000 interactions between chemicals and their target proteins. As an example of potential use, we applied DRAMOTE to develop robust models for predicting FDA approved drugs that have high probability to interact with the thyroid stimulating hormone receptor (TSHR) in humans. Our findings are further partially and indirectly supported by 3D docking results and literature information. The results based on approximately 500,000 interactions suggest that DRAMOTE has performed the best and that it can be used for developing robust virtual screening models. The datasets and implementation of all solutions are available as a MATLAB toolbox online at www.cbrc.kaust.edu.sa/dramote and can be found on Figshare.

Introduction

Experimental screening of chemical compounds for their biological activity has partial coverage and leaves millions of chemical compounds untested [1]. Such experiments are usually pursued through high-throughput screening (HTS) assays in which chemical molecules (e.g.

role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: VBB is an Academic Editor of PLOS ONE journal. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

drugs) are tested against specific biological targets (e.g. protein) [2]. With existence of emerging and growing public repositories (e.g. PubChem database [3]) that provide access to biological activity information from HTS experiments, there is an opportunity to develop computational methods to predict the biological activities of millions of chemical compounds that remain untested [3, 4]. For example, data mining techniques may help narrow down promising candidate chemicals aimed at interaction with specific molecular targets before they are experimentally evaluated [5–7]. This, in principle, may help in speeding up the drug discovery process. Developing accurate prediction models for *in silico* HTS is however challenging. For datasets such as those obtained from HTS assays, achieving high prediction accuracy may be misleading since this may be accompanied by unacceptable false positive rate [8] as high accuracy does not always imply small proportion of false predictions. The fact that should be considered is that HTS experimental data is usually characterized by a great disproportion of active and inactive chemical compounds out of thousands screened [9]. This class imbalance may affect accuracy and precision of resultant predictors of activity status in individual assays [10]. If the imbalance ratio (IR) between the inactive and active compound classes can be adjusted, the performance may improve [10–12].

In this study we examine robust solutions that can be used for *in silico* screening of compound activity status in individual HTS assays that are characterized by great class imbalance. For such cases, several data mining techniques have been developed to model chemical-target interactions [13–16]. These techniques differ from virtual screening based on ligand-protein docking [17], as they do not require any prior knowledge about the 3D surface representation of the target and its cognate interactor. Also, once trained, data mining models are usually faster than ligand-protein docking models in predicting biological activity status of a given chemical compound [18]. Several web tools for predicting chemical-protein interactions have also been developed [19–22]. Decision trees are used by Han *et al.* [23] to predict the activity of a chemical compound based on the standard set of PubChem features that define chemical fingerprints [24]. The study demonstrated that the great imbalance between data classes limits classification accuracy. Different studies [25, 26] were focused on finding solution to this problem. Cost-sensitive classifiers were explored by Schierz *et al.* [25] to assign a prior importance weight to the minority class for training, whereas an optimization procedure for selecting informative samples, specifically aimed at enhancing performance of support vector machines (SVMs) was also explored [26].

Although good progress has been achieved for building predictive models for HTS data, there are still many issues in current methods that need to be investigated further.

First, many studies have developed prediction models for HTS data without considering precision or other precision relevant scores like F₁Score in optimizing the performance of these models. Recently, some studies [27–29] explored applying random under-sampling or synthetic over-sampling techniques to some assays (BioAssays) from the PubChem database. These studies did not focus on or report the precision of predictions and their impact over the number of false positives, which are highly relevant [9, 11, 12, 30]. In the case of *in silico* screening of chemical activity status, the increased precision will reduce the number of falsely predicted candidate compounds thus reducing the cost of the potential follow up laboratory experiments [8].

Second, generating and selecting a good subset of features is an important step in developing a well-performing prediction model, and may help in the cases of data with large class imbalance [31, 32]. Few efforts, however, have been dedicated for finding strong discriminating features for HTS data [26, 33, 34].

To tackle the above-mentioned problems, in this study we examine robust solutions to be used for *in silico* screening of compound activity status in individual HTS assays. For this

purpose, we run experiments using various state-of-the-art methods and compare their effect on prediction of chemical activity status using different performance metrics. Also, we developed a variant method, DRAMOTE, based on ideas from active learning, which favors selection of precision-informative training samples. We describe the data by a rich set of features that includes PubChem fingerprint features. The set of feature we generated is, to the best of our knowledge, the most comprehensive feature set used for problems of this type. This set of features was further subjected to a feature selection method to propose a set of features that may result in an improved prediction performance in comparison to the PubChem fingerprint features alone. The results of 1,350 *in silico* experiments that involved close to 500,000 interactions, suggest that DRAMOTE is the most efficient variant of data preprocessing in the case of great class imbalance based on the datasets from PubChem we used. DRAMOTE, which favors selection of interactions that enhances the overall precision of a learning model, improves F₁Score on average by over 41%, relative to other methods. Finally, we illustrate the usefulness of our DRAMOTE method through a case study of screening all FDA approved drugs in the DrugBank database [35] against the thyroid stimulating hormone receptor (TSHR) in humans and suggest top 10 candidates that potentially interact with TSHR. Our findings are further partially and indirectly supported by 3D docking results and literature information.

Materials and Methods

Datasets

PubChem BioAssay Database. For this study we selected nine datasets from the PubChem BioAssay database where targets are proteins except for one dataset where the target is cell-based. Although we have a special interest in protein targets, we choose a case that is cell-based to illustrate the generality of our method. It is worth noting that all the datasets we chose are based on the confirmatory assays and we avoided selection of primary assays based on recommendation of [25]. The datasets are based on the PubChem's BioAssay protocol, where assays can be referenced by a unique AID identifier. A single BioAssay reports experimental activity results for a set of chemical compounds over a specific biological target, which in most cases is a protein. So, a BioAssay dataset contains a list of chemical compounds with assigned labels, where label '+1' indicates that the compound shows activity with the examined target, while '-1' relates to inactive compounds. Table 1 provides a summary of the datasets used in the study. Eight of these datasets AID: 596, AID: 618, AID: 644, AID: 886, AID: 899, AID: 938, AID: 743042 and AID: 743288, were chosen to demonstrate different imbalance ratios (IR) between the active and inactive compound classes. The ninth one, called BenchSet, is a benchmark dataset that is obtained by merging three BioAssays, AID: 773, AID: 1006 and AID: 1379, as described previously by Li *et al.* [26]. In total, these datasets are composed of 11 BioAssays that represent 487,557 inactive and active interactions and offer a wide variety of class imbalance ratios ranging from 0.26% (i.e. high IR) to 48% (i.e. small IR), where IR is represented as ratio of the number of minority active cases to the number of majority inactive cases. For reporting performance over these datasets, 5-fold cross-validation setup is followed in all computational experiments. Given the large size of our experimental datasets (as shown in Table 1), 5-fold cross-validation for evaluation is a proper choice for computing a representative (i.e. non-biased) estimate [36, 37]. In order to avoid any potential bias, testing data is never used in the training process.

DrugBank database. The DrugBank database data (accessed on August, 2014) was downloaded from the website: <http://www.drugbank.ca/> [35]. The initial database had about 6,800 drug entries including 1,491 FDA-approved drugs. We considered only the FDA-approved drug list to screen the model we developed for thyroid stimulating hormone receptor (TSHR).

Table 1. Summary of experimental datasets including reference IDs in PubChem Database.

Dataset	Target Name (Target)	Type of interacting compounds	Minority Class Size	Majority Class Size	IR Ratio
BenchSet (AID: 773, AID: 1006 and AID: 1379)	Luciferase [<i>Photuris pennsylvanica</i>](Protein)	Inhibitors	487	184,154	1:377
AID 596	Microtubule-associated protein tau [<i>Homo sapiens</i>] (Protein)	Binders	1,391	66,726	1:48
AID 618	Matrix metalloproteinase 1, partial [<i>Homo sapiens</i>] (Protein)	Inhibitors	537	86,197	1:160
AID 644	Rho-associated protein kinase 2 [<i>Homo sapiens</i>] (Protein)	Inhibitors	67	139	1:2
AID 886	Chain B, The Structure Of Wild-Type Human Hadh2 (Protein)	Inhibitors	2,463	64,616	1:26
AID 899	Cytochrome P450 2C19 precursor [<i>Homo sapiens</i>] (Protein)	Inhibitors and Substrates	1,901	6,443	1:3
AID 938	Thyroid stimulating hormone receptor [<i>Homo sapiens</i>] (Protein)	Agonist Activators	1,794	60,806	1:34
AID 743042	Androgen receptor [<i>Homo sapiens</i>] (Protein)	Antagonist Activators	674	6,939	1:10
AID 743288	Hek293 cell line (Cell)	Binders	95	2,128	1:22
Total Interactions			487,557		

doi:10.1371/journal.pone.0144426.t001

Generation of features

Generating and selecting a good subset of features is an important step in developing a well-performing classification model, and may also help in the cases of large class imbalance [31, 32]. A variety of feature sets of varying complexity have been compiled for virtual screening and prediction of biological activity [25, 38]. In this study, we used the combined set of fingerprint features from two major cheminformatics toolkits, RDKit [39] and OpenBabel [40], as well as features from PubChem fingerprints [24]. OpenBabel [40] was specifically used to generate different SMARTS patterns and 3D spectrophore descriptors. In addition, several basic chemical descriptors, such as the molecular weight, number of H-acceptors and donors, and Log-P, were calculated. The finally generated set of features contained 2,940 features. A detailed description of all the features used in the study, as well as the ones we selected, is provided in [S1 Text](#). This, to the best of our knowledge, is the largest set of features compiled for use in prediction of chemical activity status from HTS assays.

Feature selection (FS)

A large set of compiled features, as described in the previous section, leads to generating information of different level of redundancy, as well as introduces features that may not be relevant to the types of biological activity of chemicals as observed in particular assays. A good FS method should be able to remove a lot of such redundant or irrelevant information [41]. FS methods, in general, can be categorized into: the filters, the wrappers, and those based on the embedded FS models [32, 42]. In this study, the wrapper FS model of DWFS tool [43], which selects features so as to maximize the performance of a classifier, is applied. The default setup of DWFS tool was used for FS experiments. As an illustration, an analysis of the effect of FS on the classification performance for one of the datasets can be found in [S2 Text](#).

Classifiers

Six widely used classifiers are applied as a basis for comparing different solutions of the class imbalance problem for activity testing in PubChem assays. These include support vector

machines [44, 45] (SVM) with linear and radial basis function (RBF) kernels, K-nearest neighbors (KNN; K = 3) [46], Linear Discriminant Analysis (LDA) [47], Naïve Bayes Classifier (NBC) [48] and Random Forests (RF) [49]. For SVM, LIBSVM [50] implementation was used for building the different SVM models. The default cost parameter as well as RBF kernel widths were used.

Performance evaluation

Performance of all methods referred to in the results section is obtained from a 5-fold cross-validation. The testing fold was never used in the training phase. Since we performed 5-fold cross-validation, with six classifiers and five class imbalance solutions, we performed 150 (5 folds × 6 classifiers × 5 solutions = 150) experiments for each dataset and 1,350 in total for all nine datasets. We report the average performance over the 5-folds of every dataset, as well as the standard deviation. In addition, we perform significance analysis between the methods using one-way analysis of variance (ANOVA). In cases where there is a significant difference between the methods, we further apply the well-known pair-wise Tukey mean-mean multiple comparison (MCC) to determine which pairs are significantly different, while simultaneously examining all methods [see S1 Table]. Giving the characteristics of this problem and the nature of having highly imbalanced classes, we provide results over many performance metrics to gain a generic view of the performances of different solutions. Let TP be the number of true positives, FP the number of false positives, TN the number of true negatives and FN the number of false negatives. The results in this study are reported based on Eqs (1–7).

$$\text{sensitivity} = TP / (TP + FN) \quad (1)$$

$$\text{specificity} = TN / (TN + FP) \quad (2)$$

$$\text{precision} = TP / (TP + FP) \quad (3)$$

$$G\text{Mean} = \sqrt{\text{sensitivity} \times \text{specificity}} \quad (4)$$

$$F_1\text{Score} = 2 \times \frac{\text{precision} \times \text{sensitivity}}{\text{precision} + \text{sensitivity}} \quad (5)$$

$$F_{0.5}\text{Score} = 1.25 \times \frac{\text{precision} \times \text{sensitivity}}{0.25 \times \text{precision} + \text{sensitivity}} \quad (6)$$

$$\text{ROC AUC} = \text{Area under the receiver operating characteristic curve} \quad (7)$$

The predictions of a classifier for a HTS dataset should result in high precision in order for the set of predicted active compounds to contain as few FP predictions as possible. The number of FPs is a crucial factor in measuring the reliability of predictions as minimizing it leads to increased chances of successful follow up experiments.

F₁Score [9] is a summary metric that computes the weighted average of precision and sensitivity. It is also known as balanced F-Score since it balances both precision and sensitivity equally. F_{0.5}Score [11, 51, 52] is another summary metric that weights precision twice as much as sensitivity. Given that intention to use the classifier for computational screening of millions of compounds, sensitivity is of less importance than precision. A conservative sensitivity rate with higher precision will still lead to large number of accurate new findings when screening a large number of candidates in the context of HTS. Thus, we give preference to precision and

$F_{0.5}$ Score as more indicative performance measures in such scenario. We consider also in the results section, discussion over sensitivity and F_1 Score. For other metrics like specificity and ROC AUC scores are reported in details in [S2 Table](#).

Methods

Data preprocessing for class imbalance case. HTS experiments are usually characterized by only a small number of active chemical compounds obtained after screening a big compound set. This nature of imbalanced distributions of the active and inactive compound classes may lead to a degraded classification performance that should be addressed. The class imbalance problem is one of the challenging tasks that received a lot of attention [53–55]. There exists a wide variety of state-of-the-art solutions of the class imbalance problem, which can be categorized abstractly into algorithmic and data-based ones [9]. In our study we consider the following approaches: majority random under-sampling (RU), synthetic minority oversampling technique (SMOTE) [56], granular SVMs for under-sampling (GSVM-RU) [57, 58], majority weighted minority over-sampling technique (MWMOTE) [59] and our precision-aware proposed method DRAMOTE. Further details about the existing methods are provided in [S3 Text](#).

DRAMOTE: our proposed solution. There are certain limitations with the existing solutions for data preprocessing in the case of class imbalance. Methods like RU and SMOTE apply sampling procedures to data without considering the effect of sampling on the classification performance. These methods are independent of any feedback from the classifier and may affect the performance only to a certain limit. In other words, these methods do not provide a mechanism to have a control over precision or other performance metrics. Other algorithms like GSVM-RU, take into account the performance of the classifier, but are limited to a specific classifier, e.g. GSVM-RU is limited to SVM and cannot be applied to other classifiers. MWMOTE needs more parameters for selecting an informative set of minority samples and is limited to optimize the performance over nearest neighbor type of classifiers. We propose here a novel method motivated by ideas from active learning (AL) (for more details about AL see [S3 Text](#)). The method is based on establishing a feedback loop with the classifier to highlight points contributing most to its precision (other performance metrics can be used).

[Fig 1](#) gives a simplified illustration of DRAMOTE, where minority samples are colored based on how informative they are towards minimizing the false positives and this can be compared with SMOTE which does not differentiate between usefulness levels of minority samples. Another major difference between DRAMOTE and SMOTE is choosing the direction for synthetically generating the new samples as illustrated by the blue points that highlight this difference in parts A and B of the [Fig 1](#). Further details about DRAMOTE including mathematical equations and pseudocode details are provided in [S3 Text](#).

Results and Discussion

Performance Comparison

We made a number of experiments to evaluate performance of the methods we used. The results are provided in [Table 2](#) over the analyzed BioAssays. [Table 2](#) shows the 5-fold cross-validation comparison results between the different class imbalance solutions. The summary scores in [Table 2](#) are based on averaging the performance over six types of classifiers for each dataset. Another summary results with statistical significance analysis including p -values can be found in [S1 Table](#) and the detailed results including other performance metrics can be found in [S2 Table](#).

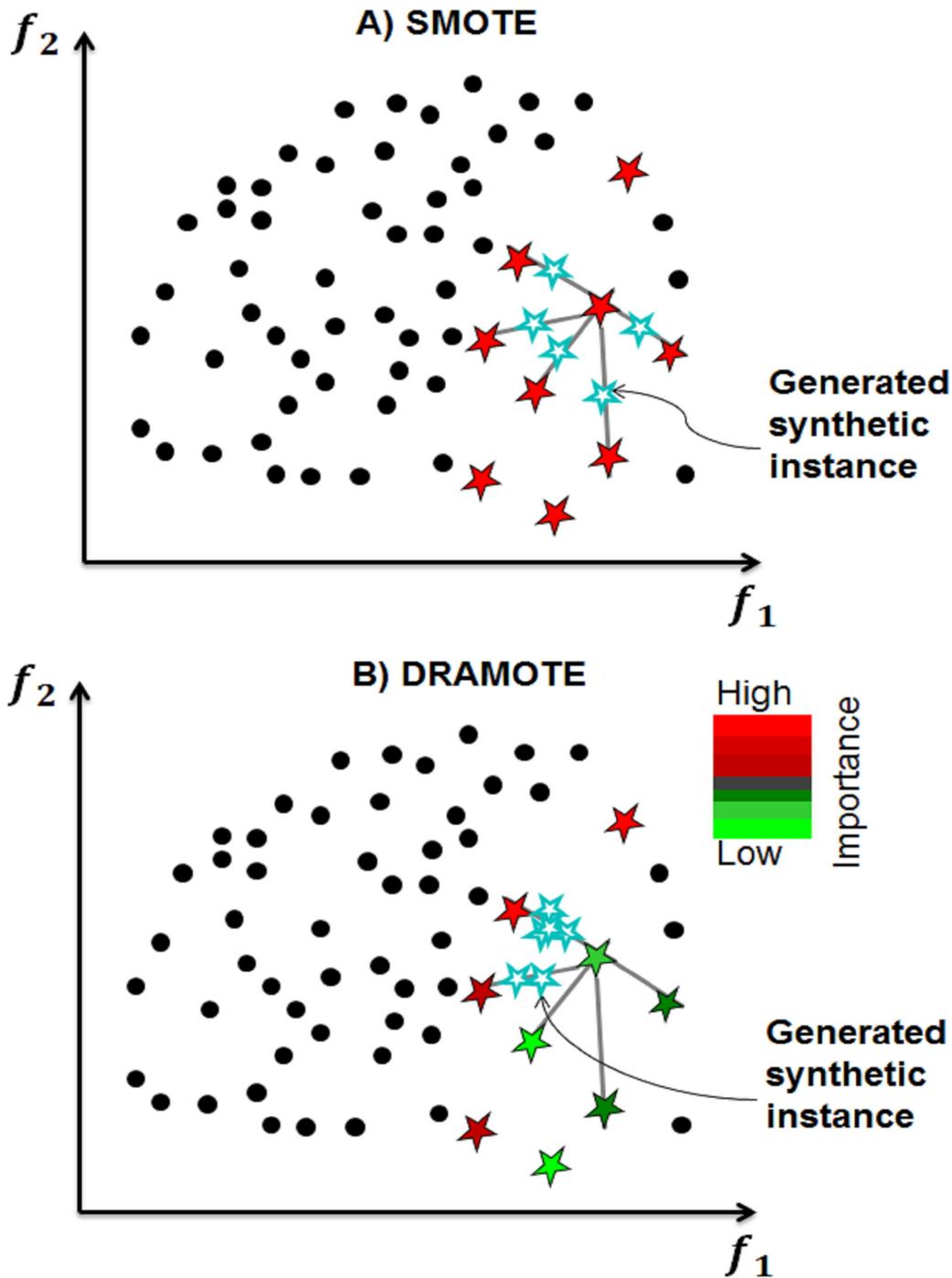


Fig 1. Illustration of generating synthetic instances. A) SMOTE generates the light blue samples by interpolation between a randomly chosen minority sample and k-nearest neighbors. B) DRAMOTE generates the light blue samples by choosing a minority sample based on its importance (i.e. contribution to precision) and the direction towards a safe region. A minority sample (red colored) that is very close to the majority negatives circles will be probably misclassified as a negative one and hence, it should get more support compared to the green colored minority samples. Once a minority sample is chosen, another point needs to be chosen for interpolation. The direction of interpolation can be controlled by choosing a nearest neighbor which is not overlapping with the negative class. This, in turn, helps in providing support for the red colored point while not harming the classifier performance in its surrounding region.

doi:10.1371/journal.pone.0144426.g001

Table 2. Comparison of the data preprocessing methods. Larger standard deviation values are the result of averaging over different types of classifiers in this summary table.

Dataset	Method	Sensitivity %	Precision %	F ₁ Score %	F _{0.5} Score %
BenchSet	RU	85.67 (±2.5)	1.07 (±0.29)	2.93 (±0.56)	1.33 (±0.35)
	GSVM-RU	68.53 (±6)	2.73 (±2.05)	5.13 (±3.7)	3.36 (±2.49)
	SMOTE	62.79 (±15.32)	10.44 (±16.11)	12.44 (±13.64)	10.87 (±15.17)
	MWMOTE	69.49 (±13.18)	4.9 (±6.7)	7.87 (±9.08)	5.75 (±7.52)
	DRAMOTE	58.14 (±19.2)	13.35 (±22.66)	11.62 (±11.42)	11.68 (±16.57)
	[26]	88.46	5	NA ^a	NA ^a
AID 596	RU	75.9 (±3.04)	5.3 (±1.17)	9.89 (±2.07)	6.51 (±1.41)
	GSVM-RU	82.78 (±7.93)	4.56 (±2.78)	8.46 (±4.78)	5.59 (±3.34)
	SMOTE	64.02 (±13.8)	10.9 (±8.95)	16.38 (±9.28)	12.47 (±9.16)
	MWMOTE	62.1 (±14.3)	10.8 (±9.2)	16.11 (±9.34)	12.32 (±9.37)
	DRAMOTE	42.9 (±13.52)	18.41 (±17.81)	19.43 (±9.63)	18 (±13.61)
AID 618	RU	72.54 (±3.41)	1.38 (±0.31)	2.7 (±0.59)	1.71 (±0.38)
	GSVM-RU	52.42 (±11.76)	2.64 (±1.48)	4.89 (±2.59)	3.24 (±1.79)
	SMOTE	43.01 (±17.87)	10.07 (±12.36)	10.93 (±8.36)	10.01 (±10.42)
	MWMOTE	42.34 (±18.53)	10.31 (±12.72)	11.38 (±8.12)	10.24 (±10.49)
	DRAMOTE	29.69 (±15.26)	12.78 (±15.61)	9.73 (±6.09)	10.58 (±10.46)
AID 644	RU	50.29 (±4.46)	35.08 (±2.56)	40.32 (±3.1)	37.3 (±2.49)
	GSVM-RU	71.28 (±12.76)	36.02 (±2.51)	46.62 (±3.23)	39.84 (±2.48)
	SMOTE	47.3 (±14.1)	41.78 (±7.23)	40.95 (±3.21)	41.65 (±3.72)
	MWMOTE	47.37 (±12.37)	42.22 (±6.68)	41.99 (±3.24)	42.4 (±4.42)
	DRAMOTE	40.09 (±8.51)	43.14 (±9.88)	38.84 (±1.64)	41.49 (±5.79)
AID 886	RU	99.54 (±0.31)	67.65 (±2.55)	80.52 (±1.75)	72.27 (±2.31)
	GSVM-RU	99.25 (±0.97)	54.51 (±26.52)	65.87 (±29.63)	58.53 (±27.76)
	SMOTE	96.94 (±4.11)	75.2 (±4.92)	84.43 (±2.51)	78.65 (±3.99)
	MWMOTE	97.03 (±3.27)	74.32 (±4.81)	83.98 (±2.75)	77.9 (±4.06)
	DRAMOTE	94.38 (±8.1)	75.69 (±6.05)	83.55 (±3.72)	78.56 (±4.17)
AID 899	RU	77.65 (±3.43)	45.96 (±7.07)	57.33 (±5.46)	49.89 (±6.7)
	GSVM-RU	97.29 (±3.22)	25.82 (±2.6)	40.69 (±2.84)	30.25 (±2.76)
	SMOTE	70.44 (±8.14)	53.52 (±14.02)	59.07 (±6.9)	55.32 (±11.22)
	MWMOTE	70.5 (±8.48)	52.61 (±13.66)	58.55 (±6.55)	54.55 (±10.83)
	DRAMOTE	64.51 (±8.01)	53.61 (±14.43)	56.73 (±5.38)	54.47 (±10.69)
AID 938	RU	99.42 (±0.41)	66.17 (±2)	79.4 (±1.45)	37.3 (±2.49)
	GSVM-RU	99.16 (±0.5)	45.85 (±17.01)	56.79 (±17.22)	49.64 (±24.09)
	SMOTE	91.86 (±0.9)	80.05 (±1.8)	84 (±1.34)	81.94 (±11.11)
	MWMOTE	94.49 (±8.2)	70.7 (±8)	80.74 (±1.9)	74.41 (±6.24)
	DRAMOTE	91.39 (±4)	81.02 (±2.03)	84.32 (±3)	82.66 (±10.73)
AID 743042	RU	71.34 (±7.44)	17.22 (±2.83)	27.66 (±4)	20.28 (±3.21)
	GSVM-RU	93.21 (±7.7)	11.11 (±0.65)	19.81 (±0.9)	13.47 (±0.74)
	SMOTE	33.38 (±16.32)	36.99 (±21.61)	27.71 (±8.52)	29.84 (±10.97)
	MWMOTE	35.52 (±14.9)	36.54 (±18.4)	30.56 (±7.01)	32.18 (±9.78)
	DRAMOTE	35.38 (±14.13)	38.69 (±20.85)	30.76 (±6.04)	33.03 (±10.23)
AID 743288	RU	68.09 (±5.53)	8.38 (±1.07)	14.89 (±1.77)	10.16 (±1.27)
	GSVM-RU	86.33 (±6.49)	5.76 (±0.4)	10.78 (±0.68)	7.08 (±0.48)
	SMOTE	25.74 (±18.34)	26.99 (±23.95)	24.56 (±6.5)	24.05 (±10.15)
	MWMOTE	23.8 (±17.4)	33.02 (±21.18)	23.75 (±9.67)	25.78 (±10.32)
	DRAMOTE	27.88 (±14.66)	34.13 (±20.58)	27.03 (±6.97)	29.02 (±10.42)

^a NA indicates that a particular measure was not reported in the referenced work

doi:10.1371/journal.pone.0144426.t002

In [Table 2](#), we consider examining more closely sensitivity, precision, F₁Score and F_{0.5}Score [[51](#)] for evaluating classification results on HTS types of data. These performance metrics shall better reflect the impact over the data with imbalanced classes [[11](#)], but other performance metrics like specificity, GMean, and ROC-AUC are also included in [S2 Table](#).

To see where a particular solution stands among all the remaining ones, we also ranked the performance of each of the methods for every classifier based on the F₁Score. We then averaged the rank position for each of the methods. The method with the lowest score is the best performing. We provide in [Table 3](#) the rank position and averaged rank position for each of the methods. [Table 3](#) clearly demonstrates that overall DRAMOTE and SMOTE were the best performing method in terms of F₁Score.

SMOTE, MWMOTE and DRAMOTE are all methods that generate synthetic data with exactly the same number of new over-sampling points. However, DRAMOTE gives preference to generating points contributing more to the precision of a particular classifier. Results of [Table 2](#) confirm this in all nine datasets, based on the fact that DRAMOTE achieves the highest precision with an improvement of a factor of 2.4 relative to precision of every other method, on average. In three out of nine cases SMOTE achieves the best in terms of F₁Score and in four cases the second best. For four out of nine datasets, DRAMOTE (compared to other solutions) achieved the highest F₁Score, while appeared the best in terms of F_{0.5}Score for six out of the nine datasets.

Compared to GSVM-RU that was reported as an effective method for PubChem BioAssays [[26](#)], DRAMOTE shows a significant improvement in precision for five out of nine datasets, while sacrifices sensitivity significantly as compared to GSVM-RU in only three cases [see [S1 Table](#)].

Compounds Interacting with Thyroid Stimulating Hormone Receptor (TSHR)

This section describes a case study for prediction of activity status of FDA drugs with TSHR protein. TSHR is a key protein in the control of thyroid function and belongs to the superfamily of G-protein-coupled receptors (GPCRs) [[60](#)]. Thyroid stimulating hormone (TSH) is the main factor responsible for regulating both differentiated function and growth of thyroid follicular epithelial cells [[61](#)]. Specifically, BioAssay AID 938 in PubChem database is an assay for finding agonists of the TSHR, which is based on stimulation of cAMP production that causes the cyclic nucleotide gated ion channel (CNG) to open to control for compounds signaling through endogenous receptors and other targets of HEK 293 cells.

The biochemical relevance of the 10 top ranked predictions by DRAMOTE was further indirectly supported by *in silico* docking results and literature. For this case study, we use the previous results to select a proper solution to preprocess the data and then, build a system

Table 3. Ranking of methods based on F₁Score for every classifier.

Classifier	RU	GSVM-RU	SMOTE	MWMOTE	DRAMOTE
SVM-L	3	5	1	4	2
SVM-R	4	5	2	3	1
KNN	3	5	2	4	1
LDA	4	5	2	3	1
NBC	1	4	3	5	2
RF	4	5	1	3	2
Average	3.17	4.83	1.83	3.67	1.50

doi:10.1371/journal.pone.0144426.t003

Table 4. Top 10 ranked predictions by DRAMOTE for BioAssay 938 with TSHR protein target.

Rank	DrugBank ID	Drug Name	Description	Ensemble System Score
1	DB00904	Ondansetron	Treatment of nausea and vomiting caused by cytotoxic chemotherapy drugs	0.98
2	DB00962	Zaleplon	Sedative/hypnotic, mainly used for insomnia	0.97
3	DB01349	Tasosartan	Treat patients with essential hypertension	0.966
4	DB00405	Dexbrompheniramine	Treat allergic conditions such as hay fever or urticaria	0.96
5	DB01261	Sitagliptin	Control of type 2 diabetes mellitus	0.958
6	DB06439	Tyloxapol	Non-ionic detergent often used as a surfactant	0.957
7	DB00889	Granisetron	Antiemetic and anti-nauseant for cancer chemotherapy patients	0.954
8	DB01342	Forasartan	Used alone or with other antihypertensive agents to treat hypertension	0.953
9	DB00748	Carbinoxamine	First generation antihistamine that competes with free histamine for binding at HA-receptor sites	0.95
10	DB06267	Udenafil	Treat erectile dysfunction	0.945

doi:10.1371/journal.pone.0144426.t004

based on ensemble of all examined classifiers. Following is the discussion of the results related to this carefully tuned system.

Computational prediction and support. The application of DRAMOTE to the TSHR dataset (AID 938) resulted in precision of 81.02% and sensitivity of 91.39%. After building an ensemble of all six trained classifiers, the performance improved by maintaining similar level of precision (~81%) but with a sensitivity of 98.84% (i.e. more than 7% increase in sensitivity).

We investigated the potential interaction of approved drugs from the DrugBank database [35] over TSHR. The ensemble of classifiers trained using BioAssay (AID 938) as its training set, is used to computationally screen approved drugs extracted from DrugBank. We report the top 10 predictions as candidate drugs with strong potential to be interact with TSHR. Table 4 provides a brief description of each drug and highlights their ranking score based on the ensemble system. The drugs, also, docked to TSHR are shown with their corresponding names and structures in S1 Fig.

Docking simulations can indirectly support the previous top findings in our data-driven approach. While docking simulations are prone to false positives, the presence of consistent levels in binding values between our findings and the top experimentally ranked interactions reported in AID 938 gives more confidence about our suggested candidates having active interaction status with TSHR. Fig 2 illustrates comparison of the docking scores of the top 10 predictions suggested by DRAMOTE, against two other sets of docking experiments we used as references for evaluation. The two sets of docking experiments include the actual top 10 experimental interactions as ranked and reported in PubChem database for AID 938 and another set (Random set) of 10 randomly selected drugs from the approved list of DrugBank [38]. In Fig 2, the listed energies correspond to the lowest predicted binding energy and they are given in kcal/mol as calculated by AutoDock Vina [62]. Part B of Fig 2 provides the root mean squared distance (RMSD) values of the best poses of each drug compound docked to an activation site in TSHR. While difference is not apparent with regard to the free energy values between all the three sets of docking experiments, the RMSD levels achieved by docking predictions using DRAMOTE are very similar to the levels achieved by the experimentally validated ones as compared to the Random set. Detailed docking procedure and scores are provided in S4 Text.

A literature review of our top predictions points out that **Tasosartan** (third ranked prediction) and **Forasartan** (eighth ranked prediction) are both angiotensin II receptor antagonist. These drugs are used to treat hypertension [63] and are known to block the renin-angiotensin system thereby protecting the kidney from damage caused by increased kidney blood pressure

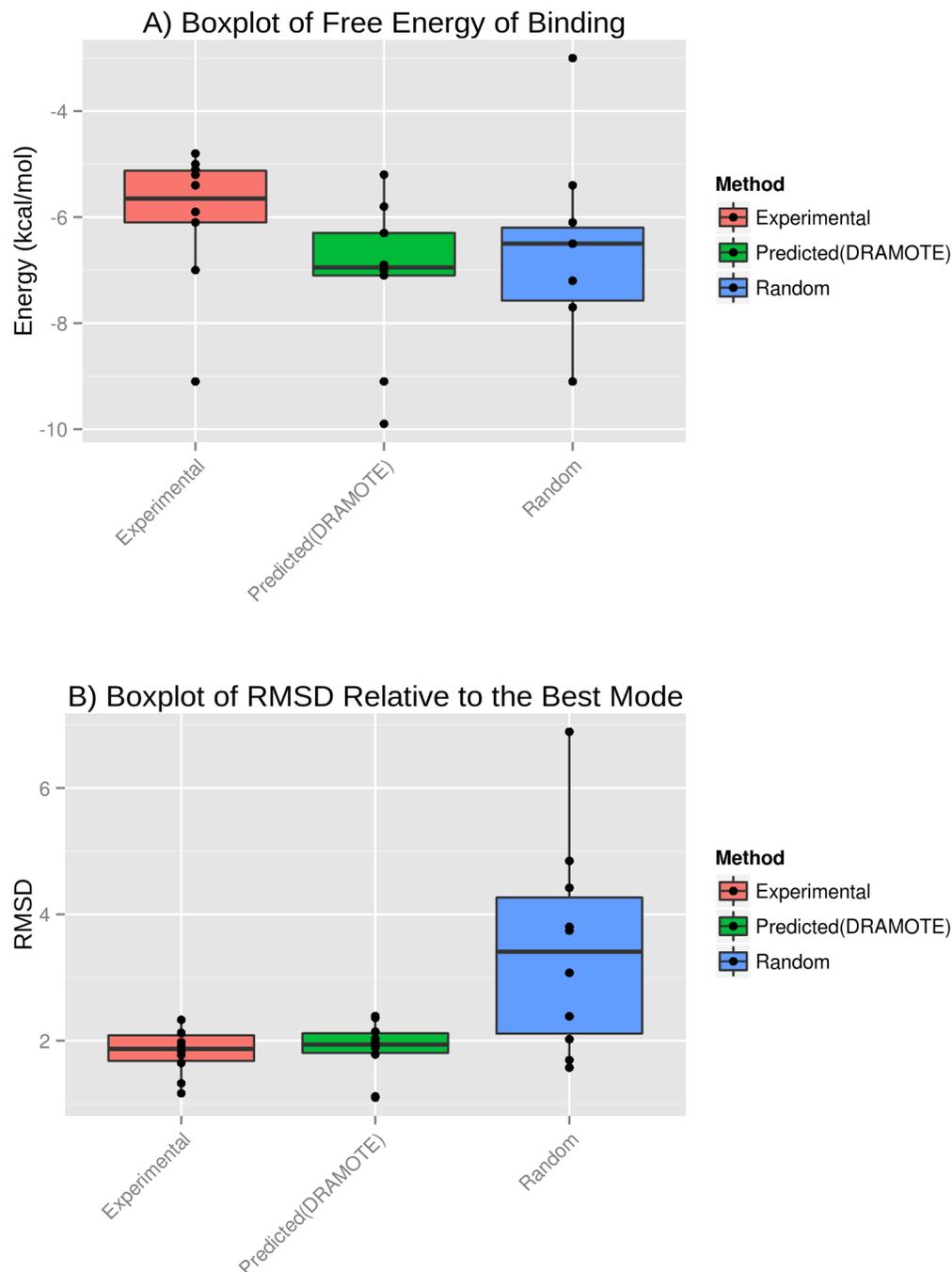


Fig 2. Boxplot over free energy of binding and RMSD values for experimental, random and DRAMOTE docking results. The random set is based on choosing 10 random drugs from approved drugs list in DrugBank database. The experimental set includes the top 10 drugs as listed in the original BioAssay AID 938 of PubChem database.

doi:10.1371/journal.pone.0144426.g002

[64]. Several studies have demonstrated a positive correlation between high blood pressure and the concentration of thyroid stimulating hormone [65–67]. Literature review for remaining top ranked drugs can be found in [S5 Text](#). These findings strengthen our proposition that the proposed top 10 predictions could be candidate drugs for interacting with TSHR. In order to show

that DRAMOTE can be used for drugs for different diseases other than those related to TSHR, an additional top ranked list is included in [S6 Text](#) for 17beta-Hydroxysteroid Dehydrogenase Type 10 (17β-HSD10) as the protein target in AID 886 assay. The expression level of this protein is elevated in the brains of Alzheimer's disease patients [68]. Thus, the predicted/suggested drugs ([S6 Text](#)) could serve as potential drugs for Alzheimer's disease aimed to inhibit expression of 17β-HSD10 since AID 886 assay is testing inhibition of 17β-HSD10.

Conclusions

In this study, we extensively compare several state-of-the-art methods that handle class imbalance problem based on advanced sampling techniques. The results based on approximately 500,000 interactions suggest that DRAMOTE can be used for developing robust virtual screening models to recognize candidate chemical compounds for potential activity with specific molecular targets in specific assays. Moreover, we applied DRAMOTE to screen for drugs likely to interact with the TSHR as a case study and we presented the top 10 drugs that potentially interact with TSHR along with indirect supporting evidence of their validity from literature and simulated 3D docking.

Supporting Information

S1 Fig. Docking output results for Carbinoxamine, Granisetron, Ondansetron, Zalepon, Sitagliptin, Forasartan, Tasosartan, Udenafil, Tyloxapal with TSHR. The orange color highlights the top docking results of a drug binding to the chosen activation site.

(TIFF)

S1 Table. Extended comparison of existing and proposed methods including an analysis of significance of difference between the reported performance metrics.

(DOCX)

S2 Table. Detailed comparison results for each dataset. Mean and variance of 5-fold cross-validation performance scores are displayed for each method and for each used classifiers.

(DOCX)

S1 Text. Summary description of features generated for chemical compounds. The file also includes most of the features we selected after applying variable selection over the originals set of generated features.

(DOCX)

S2 Text. Effect of feature selection results on classification performance.

(DOCX)

S3 Text. Details about the existing state-of-the-art solutions used in the study and their input parameters. The file includes also all information about DRAMOTE and its procedure.

(DOCX)

S4 Text. Detailed docking scores including the set of random selected drugs and description of the docking procedure.

(DOCX)

S5 Text. Extended literature review of the top predicted FDA drugs for the TSHR in humans.

(DOCX)

S6 Text. A list of the top ranked prediction by DRAMOTE for potential drugs interacting with 17 β -HSD10 in humans.

(DOCX)

Acknowledgments

The authors thank Dr. Hammad Naveed, Ahmed Elshewy, Loqmane Seridi, Dr. Salim Bougouffa, Haitham Ashoor and Dr. Mahmut Uludag for multiple insightful and valuable discussions about experimental design and results presentation. The computational analysis for this study was performed on Dragon and SnapDragon compute clusters of Computational Bioscience Research Center at King Abdullah University of Science and Technology (KAUST). This study is supported by the KAUST base research funds of VBB and PK.

Author Contributions

Conceived and designed the experiments: OS WB VB. Performed the experiments: OS. Analyzed the data: OS WB MA ME PK VB. Contributed reagents/materials/analysis tools: OS WB VR VB. Wrote the paper: OS WB MA ME VR PK VB.

References

1. Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. *Nature reviews Drug discovery*. 2004; 3(8):673–83. PMID: [15286734](#)
2. Dudley JT, Deshpande T, Butte AJ. Exploiting drug–disease relationships for computational drug repositioning. *Briefings in bioinformatics*. 2011: bbr013.
3. Wang Y, Xiao J, Suzek TO, Zhang J, Wang J, Zhou Z, et al. PubChem's BioAssay database. *Nucleic acids research*. 2012; 40(D1):D400–D12.
4. Wang Y, Xiao J, Suzek TO, Zhang J, Wang J, Bryant SH. PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic acids research*. 2009; 37(suppl 2):W623–W33.
5. He Z, Zhang J, Shi X-H, Hu L-L, Kong X, Cai Y-D, et al. Predicting drug-target interaction networks based on functional groups and biological features. *PloS one*. 2010; 5(3):e9603. doi: [10.1371/journal.pone.0009603](#) PMID: [20300175](#)
6. Bresso E, Grisoni R, Marchetti G, Karaboga AS, Souchet M, Devignes M-D, et al. Integrative relational machine-learning for understanding drug side-effect profiles. *BMC bioinformatics*. 2013; 14(1):207.
7. Kim J, Shin M. An integrative model of multi-organ drug-induced toxicity prediction using gene-expression data. *BMC bioinformatics*. 2014; 15(Suppl 16):S2. doi: [10.1186/1471-2105-15-S16-S2](#) PMID: [25522097](#)
8. Nagamine N, Shirakawa T, Minato Y, Torii K, Kobayashi H, Imoto M, et al. Integrating statistical predictions and experimental verifications for enhancing protein-chemical interaction predictions in virtual screening. *PLoS computational biology*. 2009; 5(6):e1000397. doi: [10.1371/journal.pcbi.1000397](#) PMID: [19503826](#)
9. He H, Garcia EA. Learning from imbalanced data. *Knowledge and Data Engineering, IEEE Transactions on*. 2009; 21(9):1263–84.
10. Novianti PW, Jong VL, Roes KC, Eijkemans MJ. Factors affecting the accuracy of a class prediction model in gene expression data. *BMC bioinformatics*. 2015; 16(1):199.
11. Davis J, Goadrich M. The relationship between Precision-Recall and ROC curves. *Proceedings of the 23rd international conference on Machine learning: ACM; 2006*. p. 233–40.
12. Chen P, Huang JZ, Gao X. LigandRFs: random forest ensemble to identify ligand-binding residues from sequence information alone. *BMC bioinformatics*. 2014; 15(Suppl 15):S4. doi: [10.1186/1471-2105-15-S15-S4](#) PMID: [25474163](#)
13. Webb SJ, Hanser T, Howlin B, Krause P, Vessey JD. Feature combination networks for the interpretation of statistical machine learning models: application to Ames mutagenicity. *Journal of cheminformatics*. 2014; 6(1):8. doi: [10.1186/1758-2946-6-8](#) PMID: [24661325](#)
14. Liu X, Xu Y, Li S, Wang Y, Peng J, Luo C, et al. In Silico target fishing: addressing a “Big Data” problem by ligand-based similarity rankings with data fusion. *Journal of cheminformatics*. 2014; 6(1):33.

15. Munkhdalai T, Li M, Batsuren K, Park HA, Choi NH, Ryu KH. Incorporating domain knowledge in chemical and biomedical named entity recognition with word representations. *Journal of cheminformatics*. 2015;(Suppl 1):S9.
16. Akhondi SA, Hettne KM, van der Horst E, van Mulligen EM, Kors JA. Recognition of chemical entities: combining dictionary-based and grammar-based approaches. *J Cheminform*. 2015; 7(Suppl 1):S10.
17. Schneidman-Duhovny D, Nussinov R, Wolfson HJ. Predicting molecular interactions in silico: II. Protein-protein and protein-drug docking. *Current medicinal chemistry*. 2004; 11(1):91–107. PMID: [14754428](#)
18. Xie X-Q, Chen J-Z. Data mining a small molecule drug screening representative subset from NIH PubChem. *Journal of chemical information and modeling*. 2008; 48(3):465–75. doi: [10.1021/ci700193u](#) PMID: [18302356](#)
19. Kuhn M, von Mering C, Campillos M, Jensen LJ, Bork P. STITCH: interaction networks of chemicals and proteins. *Nucleic acids research*. 2008; 36(suppl 1):D684–D8.
20. Sakakibara Y, Hachiya T, Uchida M, Nagamine N, Sugawara Y, Yokota M, et al. COPICAT: a software system for predicting interactions between proteins and chemical compounds. *Bioinformatics*. 2012; 28(5):745–6. doi: [10.1093/bioinformatics/bts031](#) PMID: [22257668](#)
21. Liu X, Vogt I, Haque T, Campillos M. HitPick: a web server for hit identification and target prediction of chemical screenings. *Bioinformatics*. 2013.
22. Wang X, Chen H, Yang F, Gong J, Li S, Pei J, et al. iDrug: a web-accessible and interactive drug discovery and design platform. *Journal of cheminformatics*. 2014; 6(1):1–8.
23. Han L, Wang Y, Bryant SH. Developing and validating predictive decision tree models from mining chemical structural fingerprints and high-throughput screening data in PubChem. *BMC bioinformatics*. 2008; 9(1):401.
24. PubChem. PubChem Substructure Fingerprint 2009 [cited 2013 2/25/2013]. Available from: ftp://ftp.ncbi.nlm.nih.gov/pubchem/specifications/pubchem_fingerprints.txt.
25. Schierz AC. Virtual screening of bioassay data. *Journal of cheminformatics*. 2009; 1:21. doi: [10.1186/1758-2946-1-21](#) PMID: [20150999](#)
26. Li Q, Wang Y, Bryant SH. A novel method for mining highly imbalanced high-throughput screening data in PubChem. *Bioinformatics*. 2009; 25(24):3310–6. doi: [10.1093/bioinformatics/btp589](#) PMID: [19825798](#)
27. Rafati-Afshar AA, Bouchachia A, editors. *An Empirical Investigation of Virtual Screening. Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on*; 2013: IEEE.
28. Zakharov AV, Peach ML, Sitzmann M, Nicklaus MC. QSAR Modeling of Imbalanced High-Throughput Screening Data in PubChem. *Journal of chemical information and modeling*. 2014; 54(3):705–12. doi: [10.1021/ci400737s](#) PMID: [24524735](#)
29. Hao M, Wang Y, Bryant SH. An efficient algorithm coupled with synthetic minority over-sampling technique to classify imbalanced PubChem BioAssay data. *Analytica chimica acta*. 2014; 806:117–27. doi: [10.1016/j.aca.2013.10.050](#) PMID: [24331047](#)
30. Bunkhumpornpat C, Sinapiromsaran K, Lursinsap C. Safe-level-smote: Safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem. *Advances in Knowledge Discovery and Data Mining*: Springer; 2009. p. 475–82.
31. Forman G. An extensive empirical study of feature selection metrics for text classification. *The Journal of machine learning research*. 2003; 3:1289–305.
32. Guyon I, Elisseeff A. An introduction to variable and feature selection. *The Journal of Machine Learning Research*. 2003; 3:1157–82.
33. Cheng T, Li Q, Wang Y, Bryant SH. Binary classification of aqueous solubility using support vector machines with reduction and recombination feature selection. *Journal of chemical information and modeling*. 2011; 51(2):229–36. doi: [10.1021/ci100364a](#) PMID: [21214224](#)
34. Rao H, Li Z, Li X, Ma X, Ung C, Li H, et al. Identification of small molecule aggregators from large compound libraries by support vector machines. *Journal of computational chemistry*. 2010; 31(4):752–63. doi: [10.1002/jcc.21347](#) PMID: [19569201](#)
35. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic acids research*. 2006; 34(suppl 1): D668–D72.
36. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Ijcai1995*. p. 1137–45.
37. Braga-Neto UM, Dougherty ER. Is cross-validation valid for small-sample microarray classification? *Bioinformatics*. 2004; 20(3):374–80. PMID: [14960464](#)

38. Kong X, Yu PS, editors. Semi-supervised feature selection for graph classification. Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining; 2010: ACM.
39. Landrum G. RDKit. Q2; 2010.
40. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. *Journal of cheminformatics*. 2011; 3(1):1–14.
41. Zhu L, Yang J, Song JN, Chou KC, Shen HB. Improving the accuracy of predicting disulfide connectivity by feature selection. *Journal of computational chemistry*. 2010; 31(7):1478–85. doi: [10.1002/jcc.21433](https://doi.org/10.1002/jcc.21433) PMID: [20127740](https://pubmed.ncbi.nlm.nih.gov/20127740/)
42. Saeys Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. *bioinformatics*. 2007; 23(19):2507–17. PMID: [17720704](https://pubmed.ncbi.nlm.nih.gov/17720704/)
43. Soufan O, Klefogiannis D, Kalnis P, Bajic VB. DWFS: A Wrapper Feature Selection Tool Based on a Parallel Genetic Algorithm. *PloS one*. 2015; 10(2):e0117988. Epub 2015/02/27. doi: [10.1371/journal.pone.0117988](https://doi.org/10.1371/journal.pone.0117988) PMID: [25719748](https://pubmed.ncbi.nlm.nih.gov/25719748/); PubMed Central PMCID: PMC4342225.
44. Boser BE, Guyon IM, Vapnik VN, editors. A training algorithm for optimal margin classifiers. The Fifth Annual Workshop on Computational Learning Theory 1992: ACM.
45. Cortes C, Vapnik V. Support-vector networks. *Machine learning*. 1995; 20(3):273–97.
46. Cover T, Hart P. Nearest neighbor pattern classification. *Information Theory, IEEE Transactions on*. 1967; 13(1):21–7.
47. Bishop CM. *Pattern Recognition and Machine Learning*. Jordan M, Kleinberg J, Scholkopf B, editors: Springer Science + Business Media; 2006.
48. Mitchell TM. *Machine learning*. WCB. McGraw-Hill Boston, MA; 1997.
49. Breiman L. Random forests. *Machine learning*. 2001; 45(1):5–32.
50. Chang C-C, Lin C-J. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*. 2011; 2(3):27.
51. Santoni FA, Hartley O, Luban J. Deciphering the code for retroviral integration target site selection. *PLoS computational biology*. 2010; 6(11):e1001008. doi: [10.1371/journal.pcbi.1001008](https://doi.org/10.1371/journal.pcbi.1001008) PMID: [21124862](https://pubmed.ncbi.nlm.nih.gov/21124862/)
52. Maitin-Shepard J, Cusumano-Towner M, Lei J, Abbeel P, editors. Cloth grasp point detection based on multiple-view geometric cues with application to robotic towel folding. *Robotics and Automation (ICRA), 2010 IEEE International Conference on*; 2010: IEEE.
53. Van Hulse J, Khoshgoftaar TM, Napolitano A, editors. Experimental perspectives on learning from imbalanced data. Proceedings of the 24th international conference on Machine learning; 2007: ACM.
54. Japkowicz N, editor *Learning from imbalanced data sets: a comparison of various strategies*. AAAI workshop on learning from imbalanced data sets; 2000.
55. Chawla NV, Japkowicz N, Kotcz A. Editorial: special issue on learning from imbalanced data sets. *ACM SIGKDD Explorations Newsletter*. 2004; 6(1):1–6.
56. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. *arXiv preprint arXiv:11061813*. 2011.
57. Tang Y, Zhang Y-Q, Chawla NV, Krasser S. SVMs modeling for highly imbalanced classification. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*. 2009; 39(1):281–8.
58. Tang Y, Zhang Y-Q, editors. Granular SVM with repetitive undersampling for highly imbalanced protein homology prediction. *Granular Computing, 2006 IEEE International Conference on*; 2006: IEEE.
59. Barua S, Islam M, Yao X, Murase K. MWMOTE—Majority Weighted Minority Oversampling Technique for Imbalanced Data Set Learning. *Knowledge and Data Engineering, IEEE Transactions on*. 2014; 26(2):405–25.
60. Szkudlinski MW, Fremont V, Ronin C, Weintraub BD. Thyroid-stimulating hormone and thyroid-stimulating hormone receptor structure-function relationships. *Physiological Reviews*. 2002; 82(2):473–502. PMID: [11917095](https://pubmed.ncbi.nlm.nih.gov/11917095/)
61. Vassart G, DUMONT JE. The Thyrotropin Receptor and the Regulation of Thyrocyte Function and Growth*. *Endocrine Reviews*. 1992; 13(3):596–611. PMID: [1425489](https://pubmed.ncbi.nlm.nih.gov/1425489/)
62. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry*. 2010; 31(2):455–61. doi: [10.1002/jcc.21334](https://doi.org/10.1002/jcc.21334) PMID: [19499576](https://pubmed.ncbi.nlm.nih.gov/19499576/)
63. Hagmann M, Nussberger J, Naudin R, Burns T, Karim A, Waeber B, et al. SC-52458, an orally active angiotensin II-receptor antagonist: inhibition of blood pressure response to angiotensin II challenges and pharmacokinetics in normal volunteers. *Journal of cardiovascular pharmacology*. 1997; 29(4):444–50. PMID: [9156352](https://pubmed.ncbi.nlm.nih.gov/9156352/)

64. Naik P, Murumkar P, Giridhar R, Yadav MR. Angiotensin II receptor type 1 (AT 1) selective nonpeptidic antagonists—A perspective. *Bioorganic & medicinal chemistry*. 2010; 18(24):8418–56.
65. Åsvold BO, Bjørø T, Nilsen TI, Vatten LJ. Association between blood pressure and serum thyroid-stimulating hormone concentration within the reference range: a population-based study. *The Journal of Clinical Endocrinology & Metabolism*. 2007; 92(3):841–5.
66. Turchi F, Ronconi V, di Tizio V, Boscaro M, Giacchetti G. Blood pressure, thyroid-stimulating hormone, and thyroid disease prevalence in primary aldosteronism and essential hypertension. *American journal of hypertension*. 2011; 24(12):1274–9. doi: [10.1038/ajh.2011.144](https://doi.org/10.1038/ajh.2011.144) PMID: [21850059](https://pubmed.ncbi.nlm.nih.gov/21850059/)
67. Jian W-X, Jin J, Qin L, Fang W, Chen X, Chen H, et al. Relationship between thyroid-stimulating hormone and blood pressure in the middle-aged and elderly population. *Singapore medical journal*. 2013; 54(7):401–5. PMID: [23900471](https://pubmed.ncbi.nlm.nih.gov/23900471/)
68. Yang S-Y, He X-Y, Isaacs C, Dobkin C, Miller D, Philipp M. Roles of 17 β -hydroxysteroid dehydrogenase type 10 in neurodegenerative disorders. *The Journal of steroid biochemistry and molecular biology*. 2014; 143:460–72. doi: [10.1016/j.jsbmb.2014.07.001](https://doi.org/10.1016/j.jsbmb.2014.07.001) PMID: [25007702](https://pubmed.ncbi.nlm.nih.gov/25007702/)