Bcl::ChemInfo - Qualitative Analysis Of Machine Learning Models For Activation Of HSD Involved In Alzheimer's Disease

Mariusz Butkiewicz, Edward W. Lowe, Jr., Jens Meiler

Abstract— **In this case study, a ligand-based virtual high throughput screening suite, bcl::ChemInfo, was applied to screen for activation of the protein target 17-beta hydroxysteroid dehydrogenase type 10 (HSD) involved in Alzheimer's Disease. bcl::ChemInfo implements a diverse set of machine learning techniques such as artificial neural networks (ANN), support vector machines (SVM) with the extension for regression, kappa nearest neighbor (KNN), and decision trees (DT). Molecular structures were converted into a distinct collection of descriptor groups involving 2D- and 3D-autocorrelation, and radial distribution functions. A confirmatory high-throughput screening data set contained over 72,000 experimentally validated compounds, available through PubChem. Here, the systematical model development was achieved through optimization of feature sets and algorithmic parameters resulting in a theoretical enrichment of 11 (44% of maximal enrichment), and an area under the ROC curve (AUC) of 0.75 for the best performing machine learning technique on an independent data set. In addition, consensus combinations of all involved predictors were evaluated and achieved the best enrichment of 13 (50%), and AUC of 0.86. All models were computed *in silico* and represent a viable option in guiding the drug discovery process through virtual library screening and compound prioritization *a priori* to synthesis and biological testing. The best consensus predictor will be made accessible for the academic community at www.meilerlab.org**

Keywords – Machine Learning, Quantitative Structure Activity Relation (QSAR), Artificial Neural Network (ANN), Support Vector Machine (SVM), Decision Trees (DT), Kohonen Network, kappa - Nearest Neighbor (KNN), Area under the curve (AUC), Receiver Operator Characteristics (ROC), Enrichment, high-throughput screening (HTS)

Alzheimer's Disease puts a financial burden on society with over $150 billion annually, making it the 3rd most costly disease after heart disease and cancer [1]. In modern drug design, compounds with undesirable biological activity can be eliminated from the available chemical space while optimizing efficacy. The ability to predict active compounds related to cognitive disorders such as Alzheimer's Disease has the potential to reduce the medical cost involved.

The protein target 17-beta hydroxysteroid dehydrogenase type 10 (HSD) has been found in elevated concentrations in the hippocampi of Alzheimer's disease patients. HSD may play a role in the degradation of neuroprotective agents. The inhibition of HSD has been indicated as a possible mean's of treating Alzheimer's disease. Dysfunctions in human 17 beta-hydroxysteroid dehydrogenases result in disorders of biology of reproduction and neuronal diseases, the enzymes are also involved in the pathogenesis of various cancers. HSD has a high affinity for amyloid proteins. Thus, it has been proposed that HSD may contribute to the amyloid plaques found in Alzheimer's patients [2]. Furthermore, HSD degrades neuroprotective agents like allopregnanolone which may lead to memory loss. Therefore, it has been postulated that inhibition of HSD may help lessen the symptoms associated with Alzheimer's.

High-throughput screening (HTS) has become a key technology of pharmaceutical research [3], often more than one million compounds per biological target are screened [2]. At the same time, the number of compounds testable in a HTS experiment remains limited and costs increase linearly with size of the screen [4]. This challenge motivates the development of virtual screening methods which search large compound libraries *in silico* and identify novel chemical entities with a desired biological activity [2].

Machine learning techniques play a crucial role in modeling quantitative structure activity relationships (QSAR) by correlating chemical structure with its biological activity for a specific biological target [3, 5-7]. In recent years the potential of approaches such as Support Vector Machines (SVM) and Artificial Neural Networks (ANN) for establishing highly non-linear relations has become apparent. [18-22]. The algorithms learn to recognize complex patterns and make intelligent decisions based on an established compound library. Imposing such acquired sets of patterns obtained by a learning process, the algorithms are able to recognize not yet tested molecules and categorize them towards a given outcome.

In this study, a cheminformatics software suite named bcl::ChemInfo, incorporates several predictive models using supervised machine learning techniques including artificial neural networks [8], support vector machines with the extension for regression estimation (SVR) [9], decision trees [10], and unsupervised techniques such as kappa nearest neighbors (KNN) [11], and kohonen networks (Kohonen) [12].

This work is supported by 1R21MH082254 and 1R01MH090192 to Jens Meiler. Edward W. Lowe, Jr. acknowledges NSF support through the CI-TraCS Fellowship (OCI-1122919).

Mariusz Butkiewicz is a graduate student in Chemistry at Vanderbilt University, Nashville TN, 37232;

Edward W. Lowe, Jr., PhD, is a research assistant professor at the Center for Structural Biology at Vanderbilt University, Nashville TN, 37232;

Jens Meiler, PhD, is Associate Professor of Chemistry, Pharmacology, and Biomedical Informatics, Institute for Chemical Biology

Corresponding author: Jens Meiler, PhD, Center for Structural Biology, Vanderbilt University, Nashville, TN 37232.

I. Machine Learning Techniques

A. Unsupervised Learning

The kohonen network represents an unsupervised learning algorithm [12-14]. It is conceptually derived from artificial neural networks consisting of an input layer and a two-dimensional grid of neurons, the kohonen network.

The second unsupervised learning method is the Kappa – Nearest Neighbors [15-17]. This method uses a distance function to calculate pair-wise distances between query points and reference points, where query points are those to be classified.

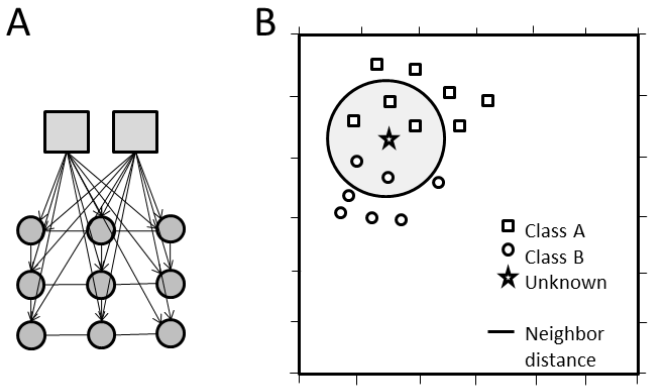

Figure 1: The schematic view of three unsupervised machine learning techniques is presented. A) A Kohonen network is represented by a input layer connected to a grid of nodes, each fully connected with its neighbors. B) The kappa – nearest-neighbors represents the predicted value of a query point as the weighted average of its *kappa* nearest reference points.

B. Supervised Learning

Artificial Neural Networks are successful attempting classification and regression problems in chemistry and biology. The structure of ANNs resembles the neuron structure of a human neural net. Layers of neurons are connected by weighted edges w_{ji}. The input data x_i are summed according to their weights, activation function applied, and output used as the input to neurons of the next layer (Figure 2).

Support Vector Machine learning with extension for regression estimation [18, 19] represents a supervised machine learning approach successfully applied in the past [3, 7]. The core principles lay in linear functions defined in high-dimensional hyperspace [20], risk minimization according to Vapnik's ε - intensive loss function, and structural risk minimization [21] of a risk function consisting of the empirical error and the regularized term.

The Decision Tree learning algorithm [10, 22] determines sets of rules to partition a given training data set. The outcome is a tree diagram or dendrogram (Figure 2) that describes how a given dataset can be classified by assessing a number of predictor variables and a dependent variable.

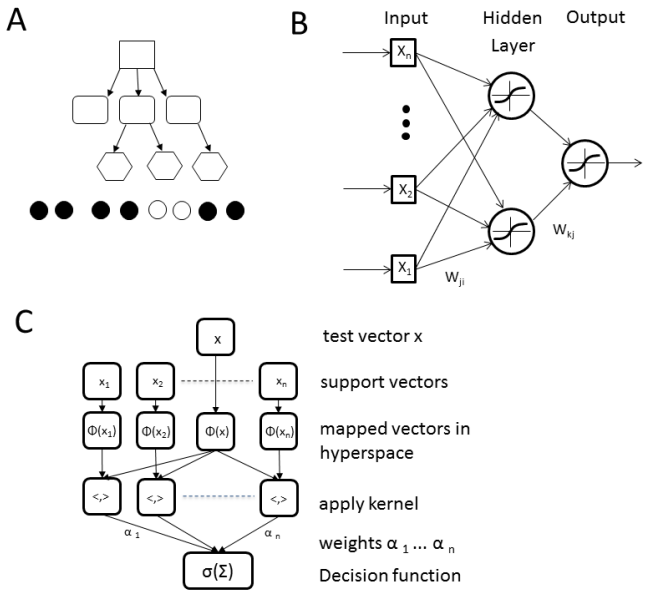

Figure 2: Depictions of Decision Trees (A) , Artificial Neural Networks (B) , and Support Vector Machines (C) are presented. In A), the partitioning algorithm determines each predictor forecast, the value of the dependent variable. The dataset is then successively split into subsets (nodes) by the descriptor that produces the greater purity in the resulting subsets. In B), a three-layer feed forward network is shown using a sigmoid activation function in each neuron. In C), the prediction process of a support vector machine is shown for an unknown vector.

II. Training Data

The protein target 17-beta hydroxysteroid dehydrogenase type 10 (HSD) is part of the family of eleven 17-β hydroxysteroid dehydrogenases that oxidize or reduce steroids at the 17 position [23]. Thus, the biological activity of these steroids is modulated. HSD catalyzes the oxidation of the positive allosteric modulators of GABA, allopregnanolone and allotetrahydrodeoxycorticosterone (3, 5 -THDOC), to 5α-DHP, 5α-DHDOC, respectively [21]. It also inactivates 17β-estradiol [24]. When first identified, HSD was known as endoplasmic reticulum-associated amyloid binding protein (ERAB) [2]. HSD has since been identified as being the only member of the 17β-HSD family to be found in mitochondria [24]. HSD has been found in increased concentrations in the mitochondria of the hippocampi of Alzheimer's disease mice [24] and humans [25]. Several possible relationships between HSD and Alzheimer's disease have been proposed in the literature [23]. HSD has a high affinity for amyloid proteins. Therefore, it has been suggested that HSD may contribute to the amyloid plaques found in Alzheimer's patients [2]. 17β-estradiol is a neuroprotective agent which prevents the degradation of existing neurons via its regulation of the β-amyloid precursor protein metabolism [24]. HSD has been

shown to degrade 17β-estradiol which may lead to neuronal degradation and the accumulation of β-amyloid, forming characteristic plaques [21]. It was also suggested that allopregnanolone reverses memory loss and dementia in the mouse model of Alzheimer's disease [26]. HSD is involved in the degradation of allopregnanolone which may lead to memory loss. Therefore, it has been postulated that inhibition of HSD may help lessen the symptoms associated with Alzheimer's.

The data set for the protein target HSD used in this study was obtained through PubChem [27] (AID 886) and resulted in a final data set of 72,066 molecules. A confirmatory high throughput screen revealed 2,463 molecules which activate the enzyme, as experimentally determined by dose-response curves. Among the actives, 495 compounds with a concentration <= 1 μM were identified. All molecules in the data set were numerically encoded using a series of transformation-invariant descriptors which serve as unique fingerprints. The descriptors (Table I) were calculated using in-house code.

III. Implementation / Method

Our in-house developed C++ class library, the BioChemistryLibrary (BCL), was employed to implement all machine learning algorithms, and descriptor calculations used for this study. A third-party 3D conformation generator, CORINA [28], was used to calculate 3D coordinates for molecules prior to descriptor calculation. The here applied ligand-based virtual high throughput screening suite, bcl::ChemInfo, is part of the BCL library.

A. Dataset Generation

During the training of the models, 10% of the data set was used for monitoring and 10% were used for independent testing of the trained models, leaving 80% for the training data set. The independent data set is put aside and not used during the training process. Each trained model is evaluated by a given quality measure based on this independent data set.

B. Quality Measures

The machine learning approaches are evaluated by means of a receiver operating characteristic (ROC) curve using cross-validated models. ROC curves plot the rate of true positives $TPR = TP/P = TP/(FN + TP)$, or sensitivity versus the rate of false positives $FPR = FP/N = 1 - TN/N = FP/(FP + TN)$, or (1 - specificity). The diagonal represents performance of a random predictor and has an integral or area under the curve (AUC) of 0.5. The QSAR model progressively improves as the AUC-value increases. Often $Precision$ is normalized with the background fraction of active compounds through computation of the enrichment measure:

$$Enrichment = \frac{TP/(TP+FP)}{P/(P+N)} = \frac{Precision}{P/(P+N)} \quad (9)$$

The value represents the factor by which the fraction of active compounds can be theoretically increased above the fraction observed in the original screen through *in silico* ranking. Another introduced measure is the root mean square deviation ($rmsd$).

$$rmsd = \sqrt{\frac{\sum_{i=1}^{n}(exp_i - pred_i)^2}{n}}$$

with exp_i being the experimental value and $pred_i$ the predicted value.

C. Feature Selection

A total of 60 descriptor groups resulting in 1,284 descriptor values were generated using the BCL. These 60 categories consisted of scalar descriptors, as well as 2D- and 3D autocorrelation functions, radial distribution functions, and van der Waals surface area weighted variations of each of the non-scalar descriptors (see Table I).

Sequential forward feature selection [29] was used for feature optimization for each machine learning technique individually. It describes a deterministic greedy search algorithm among all features. First, every single feature is selected and five-fold cross-validated models are trained followed by the evaluation of respective objective functions. The top performing feature is elected as the starting subset for the next round. Next, each remaining feature is added to the current subset in an iterative fashion resulting in N-1 feature sets. The best performing feature set is chosen again as the starting set for the next round. This process is repeated until all features are selected and the best descriptor combination is determined.

Additionally, each feature set was trained with 5-fold cross-validation evaluated on an independent data set. The number of models generated during this process for each training method was $\sum_{i=1}^{54} 5(n-1)$. Upon identification of the optimized feature set for each algorithm, any algorithm-specific parameters were optimized using the entire training data set and using 5-fold cross-validation.

Every cross-validation model was evaluated by its quality measure on the independent data set.

IV. Results

Various machine learning methods were evaluated as single predictors, highlighted in Table II. Given the independent data set, a perfect predictor would achieve a theoretical enrichment of 27.

TABLE I
THE MOLECULAR DESCRIPTORS BY NAME AND DESCRIPTION

	Descriptor Name	Description
Scalar descriptors	Weight	Molecular weight of compound
	H-Bond donors	Number of hydrogen bonding acceptors derived from the sum of nitrogen and oxygen atoms in the molecule
	H-Bond acceptors	Number of hydrogen bonding donors derived from the sum of N-H and O-H groups in the molecule
	TPSA	Topological polar surface area in [Å^2] of the molecule derived from polar 2D fragments
Vector descriptors	Identity	weighted by atom identities
2D Autocorrelation (11 descriptors)	Sigma Charge	weighted by σ atom charges
	Pi Charge	weighted by π atom charges
3D Autocorrelation (12 descriptors)	Total Charge	weighted by sum of σ and π charges
	Sigma Electro-negativity	weighted by σ atom electronegativities
Radial Distribution Function (48 descriptors)	Pi Electro-negativity	weighted by π atom electronegativities
	Lone Pair Electro-negativity	weighted by lone pair electronegativities
	Polarizibility	weighted by effective atom polarizabilities

ANNs were trained applying simple propagation as a weight update algorithm each iteration. The $rmsd$ was evaluated as the objective function every step during the feature optimization process. ANNs as a single predictor achieved a theoretical enrichment of 10 (37% of possible maximal enrichment) on an independent dataset. An integral under the ROC curve of 0.83 was obtained.

SVMs were trained using an initial cost parameter *C* of 1.0 and the kernel parameter γ of 0.5 during the feature optimization process. Upon identification of the optimal feature set, the cost and γ parameters were optimized to 2 and 32, respectively. As a single predictor, SVMs achieved a theoretical enrichment of ~12 (44%) and an AUC of 0.75 on an independent dataset.

The KNN algorithm was used to predict the biological activity values of the training, monitoring, and independent data sets. The value of kappa, the number of neighbors to consider, was optimized with the full data set using the optimized feature set determined during the feature selection process. KNNs, as a single predictor, achieved a theoretical enrichment of ~11 (41%), AUC of 0.77 using an optimal kappa of 4.

The Kohonen networks were trained with a network grid dimension of 10 x 10 nodes and a neighbor radius of 4 using the Gaussian neighbor kernel. The best result achieved by Kohonen networks was a theoretical enrichment of ~7 (28%) and an area under the ROC curve of 0.74.

TABLE II
SINGLE AND CONSENSUS PREDICTOR RESULTS

Method	***rmsd***	**Enrichment (% max)**	**AUC**
ANN/ Kohonen / KNN / SVM	0.91	**13.42 (50)**	**0.86**
ANN/DT/Kohonen/KNN/SVM	1.17	**13.42 (50)**	**0.86**
ANN / KNN / SVM	0.76	13.27 (49)	0.86
ANN / DT / KNN / SVM	1.09	13.27 (49)	0.86
ANN / DT / SVM	1.37	12.87 (48)	0.85
ANN / SVM	0.95	12.86 (48)	0.85
Kohonen / KNN / SVM	0.85	12.81 (47)	0.81
DT / Kohonen / KNN / SVM	1.20	12.81 (47)	0.81
ANN / Kohonen / KNN	0.98	12.74 (47)	0.85
ANN / DT / Kohonen / KNN	1.30	12.74 (47)	0.85
ANN / KNN	0.81	12.55 (47)	0.85
ANN / DT / KNN	1.23	12.55 (47)	0.85
DT / KNN / SVM	1.11	12.52 (46)	0.77
KNN / SVM	0.68	12.44 (46)	0.77
ANN / Kohonen / SVM	1.10	12.35 (46)	0.85
ANN / DT / Kohonen / SVM	1.40	12.35 (46)	0.85
DT / SVM	1.54	12.01 (44)	0.75
SVM	0.84	11.80 (44)	0.75
Kohonen / KNN	0.93	11.78 (44)	0.78
DT / Kohonen / KNN	1.38	11.78 (44)	0.78
Kohonen / SVM	1.10	11.54 (43)	0.80
DT / Kohonen / SVM	1.52	11.52 (43)	0.80
DT / KNN	1.33	11.10 (41)	0.77
KNN	**0.71**	11.08 (41)	0.77
ANN / DT	1.74	10.13 (38)	0.83
ANN	1.25	10.11 (37)	0.83
ANN / Kohonen	1.33	10.03 (37)	0.83
ANN / DT / Kohonen	1.66	10.03 (37)	0.83
Kohonen	1.55	7.43 (28)	0.74
DT / Kohonen	1.97	7.42 (27)	0.74
DT	2.46	4.80 (18)	0.70

The assessed Decision Trees were cross-validated and evaluated resulting in a theoretical enrichment of ~5 (18%) and an AUC of 0.70.

To further evaluate the predictive models, ensemble predictors were also created by averaging the predictions using all possible combinations of models. The best consensus resulted in a $rmsd$ of 0.86 and a theoretical enrichment of 13 (50%) achieved by the ensemble predictor model ANN/Kohonen/KNN/SVM (Table II) (Figure 3).

Additionally, Table II lists all possible combinations introducing consensus predictors. All entries in Table II are sorted by Enrichment in descending order.

V. DISCUSSION

Among single predictors, SVM, KNN, and ANN achieved a comparable enrichment performance (37% to 44%). Kohonen networks (28%) and DTs (18%) underperformed

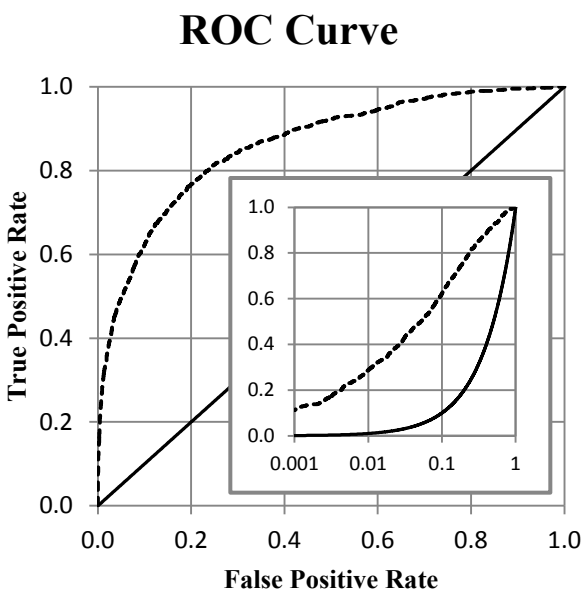

Figure 3: A receiver operating characteristic (ROC) curve for the best consensus model ANN/Kohonen/KNN/SVM is shown with a theoretical enrichment of 13 (50% max enrichment) and an area under the curve (AUC) of 0.86. The sub graph plots the same ROC curve on a logarithmic scale.

in comparison. In contrast, consensus models clearly outperform single technique predictive models. The majority of ensemble models rank above the average (Table II). Among the best ensemble predictors, ANNs, SVMs, and KNN in combination yield the best enrichment result (50%). Adding each one of the mentioned techniques to an ensemble increased the predictive accuracy, respectively. This implies that each single technique supplies a distinct contribution to the predictive accuracy of the final model (ANN/Kohonen/KNN/SVM). The inclusion of decision trees into the final ensemble did not change the overall performance (see ANN/DT/Kohonen/KNN/SVM). Both predictors achieved the same enrichment (50%). $Rmsd$, as a quality measure, discriminates distinguished predictive models by precision rather than accuracy of the predictions. Sorting Table II by $rmsd$ indicates that KNNs contribute among all top ranking predictors.

VI. Conclusions

In this research, we present a case study application of the ligand-based virtual high throughput screening suite, bcl::ChemInfo, on the target protein 17-beta hydroxysteroid dehydrogenase type 10 (HSD). HSD is involved in Alzheimer's Disease which puts a tremendous financial burden on today's society. QSARs models were developed to identify biologically active compounds for the activation of HSD. We have shown that the best consensus predictor achieved 50% of the maximal enrichment on an independent dataset. The best consensus predictor will be made accessible for the academic community at www.meilerlab.org.

Acknowledgements

The authors thank the Advanced Computing Center for Research & Education (ACCRE) at Vanderbilt University for hardware support.

References

[1] D. L. Leslie and S. K. Inouye, "The Importance of Delirium: Economic and Societal Costs," *Journal of the American Geriatrics Society,* vol. 59, pp. S241-S243, 2011.

[2] S. D. Yan, J. Fu, C. Soto, X. Chen, H. Zhu, F. Al-Mohanna, K. Collison, A. Zhu, E. Stern, T. Saido, M. Tohyama, S. Ogawa, A. Roher, and D. Stern, "An intracellular protein that binds amyloid-beta peptide and mediates neurotoxicity in Alzheimer's disease," *Nature,* vol. 389, pp. 689-95, Oct 16 1997.

[3] R. Mueller, A. L. Rodriguez, E. S. Dawson, M. Butkiewicz, T. T. Nguyen, S. Oleszkiewicz, A. Bleckmann, C. D. Weaver, C. W. Lindsley, P. J. Conn, and J. Meiler, "Identification of Metabotropic Glutamate Receptor Subtype 5 Potentiators Using Virtual High-Throughput Screening," *ACS Chemical Neuroscience*.

[4] R. D. Cramer, 3rd, D. E. Patterson, and J. D. Bunce, "Recent advances in comparative molecular field analysis (CoMFA)," *Prog Clin Biol Res,* vol. 291, pp. 161-5, 1989.

[5] V. V. Zernov, K. V. Balakin, A. A. Ivaschenko, N. P. Savchuk, and I. V. Pletnev, "Drug discovery using support vector machines. The case studies of drug-likeness, agrochemical-likeness, and enzyme inhibition predictions," *J Chem Inf Comput Sci,* vol. 43, pp. 2048-56, Nov-Dec 2003.

[6] A. Bleckmann and J. Meiler, "Epothilones: Quantitative Structure Activity Relations Studied by Support Vector Machines and Artificial Neural Networks," *QSAR Comb. Sci.,* vol. 22, pp. 719-721, 2003.

[7] M. Butkiewicz, R. Mueller, D. Selic, E. Dawson, and J. Meiler, "Application of Machine Learning Approaches on Quantitative Structure Activity Relationships," presented at the IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology, Nashville, 2009.

[8] D. Winkler, "Neural networks as robust tools in drug lead discovery and development," *Molecular Biotechnology,* vol. 27, pp. 139-167, 2004.

[9] B. Schoelkopf, "SVM and Kernel Methods," *www,* 2001.

[10] J. R. Quinlan, "Induction of Decision Trees," *Machine Learning,* vol. 1, pp. 81-106, 1986.

[11] M. Shen, Y. Xiao, A. Golbraikh, V. K. Gombar, and A. Tropsha, "Development and Validation of k-Nearest-Neighbor QSPR Models of Metabolic Stability of Drug Candidates," *Journal of Medicinal Chemistry,* vol. 46, pp. 3013-3020, 2003.

[12] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological cybernetics,* vol. 43, pp. 59-69, 1982.

[13] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE,* vol. 78, pp. 1464-1480, 1990.

[14] T. Kohonen, "Self-organization and associative memory," *Self-Organization and Associative Memory, 100 figs. XV, 312 pages.. Springer-Verlag Berlin Heidelberg New York. Also Springer Series in Information Sciences, volume 8,* vol. 1, 1988.

[15] M. Shen, Y. Xiao, A. Golbraikh, V. K. Gombar, and A. Tropsha, "Development and validation of k-nearest-neighbor QSPR models of metabolic stability of drug candidates," *J. Med. Chem,* vol. 46, pp. 3013-3020, 2003.

[16] F. Nigsch, A. Bender, B. van Buuren, J. Tissen, E. Nigsch, and J. B. O. Mitchell, "Melting point prediction employing k-nearest neighbor algorithms and genetic parameter optimization," *J. Chem. Inf. Model,* vol. 46, pp. 2412-2422, 2006.

[17] S. Ajmani, K. Jadhav, and S. A. Kulkarni, "Three-dimensional QSAR using the k-nearest neighbor method and its interpretation," *J. Chem. Inf. Model,* vol. 46, pp. 24-31, 2006.

[18] L. M. Berezhkovskiy, "Determination of volume of distribution at steady state with complete consideration of the kinetics of protein and tissue binding in linear pharmacokinetics," *Journal of pharmaceutical sciences,* vol. 93, pp. 364-74, Feb 2004.

[19] J. C. Kalvass, D. A. Tess, C. Giragossian, M. C. Linhares, and T. S. Maurer, "Influence of microsomal concentration on apparent intrinsic clearance: implications for scaling in vitro data," *Drug metabolism and disposition: the biological fate of chemicals,* vol. 29, pp. 1332-6, Oct 2001.

[20] B. Schoelkopf and A. J. Smola, *Learning with Kernels*. Cambridge, Massachusetts: The MIT Press, 2002.

[21] X. Y. He, J. Wegiel, Y. Z. Yang, R. Pullarkat, H. Schulz, and S. Y. Yang, "Type 10 17beta-hydroxysteroid dehydrogenase catalyzing the oxidation of steroid modulators of gamma-aminobutyric acid type A receptors," *Mol Cell Endocrinol,* vol. 229, pp. 111-7, Jan 14 2005.

[22] M. R. H. M. Maruf Hossain, James Bailey, "ROC-tree: A Novel Decision Tree Induction Algorithm Based on Receiver Operating Characteristics to Classify Gene Expression Data," 2006.

[23] J. Adamski and F. J. Jakob, "A guide to 17beta-hydroxysteroid dehydrogenases," *Mol Cell Endocrinol,* vol. 171, pp. 1-4, Jan 22 2001.

[24] X. Y. He, G. Y. Wen, G. Merz, D. Lin, Y. Z. Yang, P. Mehta, H. Schulz, and S. Y. Yang, "Abundant type 10 17 beta-hydroxysteroid dehydrogenase in the hippocampus of mouse Alzheimer's disease model," *Brain Res Mol Brain Res,* vol. 99, pp. 46-53, Feb 28 2002.

[25] P. Hovorkova, Z. Kristofikova, A. Horinek, D. Ripova, E. Majer, P. Zach, P. Sellinger, and J. Ricny, "Lateralization of 17beta-hydroxysteroid dehydrogenase type 10 in hippocampi of demented and psychotic people," *Dement Geriatr Cogn Disord,* vol. 26, pp. 193-8, 2008.

[26] J. M. Wang, C. Singh, L. Liu, R. W. Irwin, S. Chen, E. J. Chung, R. F. Thompson, and R. D. Brinton, "Allopregnanolone reverses neurogenic and cognitive deficits in mouse model of Alzheimer's disease," *Proc Natl Acad Sci U S A,* vol. 107, pp. 6498-503, Apr 6.

[27] "PubChem http://pubchem.ncbi.nlm.nih.gov," ed.

[28] J. Gasteiger, C. Rudolph, and J. Sadowski, "Automatic Generation of 3D-Atomic Coordinates for Organic Molecules," *Tetrahedron Comput. Method.,* vol. 3, pp. 537-547, 1992.

[29] K. Z. Mao, "Orthogonal forward selection and backward elimination algorithms for feature subset selection," *IEEE Trans Syst Man Cybern B Cybern,* vol. 34, pp. 629-34, Feb 2004.