

Recognition of Drug-Target Interaction Patterns using Genetic Algorithm-optimized Bayesian-regularized Neural Networks and Support Vector Machines

Michael Fernandez and Akinori Sarai
Department of Bioscience and Bioinformatics, Kyushu
Institute of Technology (KIT), 680-4 Kawazu, Iizuka,
820-8502 Japan.
michael_llamosa@yahoo.com

Shandar Ahmad
National Institute of Biomedical Innovation, 7-6-8, Saito-
Asagi, Ibaraki-shi, Osaka 5670085, Japan

Abstract— Genetic algorithm (GA) applied to feature selection and model optimization improved the performance of robust mathematical models such as Bayesian-regularized neural networks (BRANNs) and support vector machines (SVMs) on different drug design datasets. The selection of optimum input variables and the adjustment of network weights and biases to optimum values to yield generalizable predictors were optimized by combining Bayesian training and GA based-variable selection. Similarly, kernel and regularization parameters of SVMs were properly set by GA optimization. The predictors were more accurate and robust than previous published models on the same datasets. In addition, feature selection over large pools of molecular descriptors allowed determining the structural and atomic properties of the ligands that are ruling the biological interactions with the target.

Keywords—kernel-based methods, feature selection, enzyme inhibition, structure-activity relationship, *in silico* drug design.

I. INTRODUCTION

One of the main challenges in today's drug design is the discovery of new biologically active compounds on the basis of previously synthesized molecules. Quantitative structure-activity relationship (QSAR) is an indirect ligand-based approach which models the effect of structural features on biological activity. This knowledge is then employed to propose new compounds with enhanced activity and selectivity profile for a specific therapeutic target [1]. QSAR methods are based entirely on experimental structure-activity relationships for enzyme inhibitor or receptor ligands. In comparison to direct receptor-based methods, which include molecular docking and advanced molecular dynamics simulations, QSAR methods do not strictly require the 3D-structure of a target enzyme or even a receptor-effector complex. They are computationally not demanding and allow establishing an *in silico* tool from which biological activity of newly synthesized molecules can be predicted.

3D-QSAR methods, especially comparative molecular field analysis (CoMFA) [2] and Comparative Molecular Similarity Indices Analysis, (CoMSIA) [3] are nowadays used widely in drug design. The main advantages of these methods are that they are applicable to heterogeneous data sets, and they bring a 3D mapped description of favorable and unfavorable interactions, according to physico-chemical properties. In this sense, they provide a solid platform for retrospective hypotheses by means of the interpretation of significant interaction regions. However, some disadvantages of these methods are related to the three-dimensional information and alignment of the molecular structures, since there are uncertainties about different binding modes of ligands, and uncertainties about the bioactive conformations [4].

CoMFA and CoMSIA have emerged as the 3D-QSAR methods most embraced by the scientific community today; however, current reports on QSAR encompass the use of too many forms of the molecular information and statistical correlation methods. The structures can be described by physico-chemical parameters [5], topological descriptors [6] or quantum chemical descriptors [7], etc. The correlation can be obtained by linear methods or non-linear predictors such as artificial neural networks (ANNs) [8] and radial basis function-based support vector machines (RBF-SVM) [9]. Unlike CoMFA, CoMSIA, and linear methods, ANNs and RBF-SVM are able to describe nonlinear relationships, which should bring to a more realistic approximation of the structure-relationship paradigm, since interactions between the ligand and its biological target must be nonlinear.

Besides the complex nature of biological interactions, the enormous variety of molecular descriptors already proposed to correlate with activity arise an undetermined problem where undesirable data overfitting can result. This problem can be handled by implementing feature selection routines that determines relevant descriptors. The present paper describes the application of Bayesian-regularized genetic neural networks (BRGNNs) and genetic algorithm (GA)-optimized SVM (GA-SVM) for feature selection and/or model optimization in drug design area. Firstly, we

describe BRGNN and GA-SVM approaches. Then, we expose their applications to model different own collected drug-target interaction data.

II. MATERIAL AND METHODS

A. DATA SET

In order to study affinities of different series of ligands, biological activities were collected from the literature. Activity measurements were taken as affinity constants (K_i) and ligand concentrations for the 50% (IC_{50}) and 90% (IC_{90}) inhibition of the targets. For modeling, IC_{50} and IC_{90} were converted in their negative logarithmic values (pIC_{50} and pIC_{90}) which are measurement of drug effectiveness, it is the functional strength of the ligand towards the target.

Prior to molecular descriptor calculations, 3D structures of the studied compounds were geometrically optimized using the semiempirical quantum-chemical methods implemented in the MOPAC 6.0 computer software by Frank J. Seiler Research Laboratory [10]. Datasets include: cancer therapy targets, HIV target; Alzheimer's disease target, ion channel blockers, antiprotozoan target. ion channel proteins and protein receptor.

Different set of molecular descriptors were computed for encoding the structural information of the targets. Intercorrelation among variables was eliminated and only independent or quasi-independent variables were included in the GA search.

B. BAYESIAN-REGULARIZED GENETIC NEURAL NETWORKS

Back-propagation ANNs are data-driven models the sense that their adjustable parameters are trained to minimize some network performance score F (often equal to MSE):

$$F = MSE = \frac{1}{N} \sum_{i=1}^N (y_i - t_i)^2 \quad (1)$$

In this equations MSE is the mean of the sum of squares of the network errors, N is the number of compounds, y_i is the predicted biological activity of the compound i , t_i is the experimental biological activity of the compound i .

Often predictors can memorize the training examples, but have not learned to generalize to new situations. The Bayesian framework for ANNs is based on a probabilistic interpretation of network training to improve generalization ability of the classical networks. In contrast to conventional network training, where an optimal set of weights is chosen by minimizing an error function, the Bayesian approach involves a probability distribution of network weights. In Bayesian-regularized artificial neural networks (BRANNs), Bayesian approach yields a posterior distribution of network parameters, conditional on the training data and predictions

are expressed in terms of expectations with respect to this posterior distribution [11, 12].

Assuming a set of pairs $D = \{x_i, t_i\}$, where $i = 1 \dots N$ is a label running over the pairs, the data set can be modeled as deviating from this mapping under some additive noise process (v_i):

$$t_i = y_i + v_i \quad (2)$$

If v_i is modeled as zero-mean Gaussian noise with standard deviation σ_v , then the probability of the data given the parameters w is:

$$P(D | w, \beta, M) = \frac{1}{Z_D(\beta)} \exp(-\beta \times MSE) \quad (3)$$

where M is the particular neural network model used, $\beta = 1/\sigma_v^2$ and the normalization constant is given by $Z_D(\beta) = (\pi/\beta)^{N/2}$. $P(D|w, \beta, M)$ is called the likelihood. The maximum likelihood parameters w_{ML} (the what minimises MSE) depends sensitively on the details of the noise in the data [11, 12].

For completing the interpolation model, a prior distribution must be defined which embodies our prior knowledge on the sort of mappings that are "reasonable". Typically this is quite a broad distribution, reflecting the fact that we only have a vague belief in a range of possible parameter values. Once we have observed the data, Bayes' theorem can be used to update our beliefs, and we obtain the density. As a result, the posterior distribution is concentrated on a smaller range of values than the prior distribution. Since a neural network with large weights will usually give rise to a mapping with large curvature, we favor small values for the network weights. At this point, we defined a prior that expresses the sort of smoothness the interpolant is expected to have. The model has a prior of the form:

$$P(w | \alpha, M) = \frac{1}{Z_w(\alpha)} \exp(-\alpha \times MSW) \quad (4)$$

where α represents the inverse of the distribution and the normalization constant is given by $Z_w(\alpha) = (\pi/\alpha)^{N/2}$. MSW is the mean of the sum of the squares of the network weights and is commonly referred to as a regularising function [11, 12].

Considering the first level of inference, if α and β are known, then posterior probability of the parameters w is:

$$P(w | D, \alpha, \beta, M) = \frac{P(D | w, \beta, M) \times P(w | \alpha, M)}{P(D | \alpha, \beta, M)} \quad (5)$$

where $P(w|D, \alpha, \beta, M)$ is the posterior probability, that is the plausibility of a weight distribution considering the information of the data set in the model used, $P(w|\alpha, M)$ is the prior density, which represents our knowledge of the weights before any data is collected, $P(D|w, \beta, M)$ the

likelihood function, which is the probability of the data occurring, given the weights and $P(D|\alpha, \beta, M)$ is a normalization factor, which guarantees that the total probability is 1.

Considering that the noise in the training set data is Gaussian and that the prior distribution for the weights is Gaussian, the posterior probability fulfills the relation:

$$P(w | D, \alpha, \beta, M) = \frac{1}{Z_F} \exp(-F) \quad (6)$$

where Z_F depends of objective function parameters. So under this framework, minimization of F is identical to find the (locally) most probable parameters.

In short, Bayesian regularization involves modifying the performance function (F) defined in Equation 5, which is possible improving generalization by adding an additional term that weights by penalizing overly large magnitudes.

$$F = \beta \times MSE + \alpha \times MSW \quad (7)$$

The relative size of the objective function parameters α and β dictates the emphasis for getting a smoother network response. MacKay's Bayesian framework automatically adapts the regularization parameters to maximize the evidence of the training data [11, 12].

The joining of BRANN and GA feature selection (BRGNN) increases the possibilities of BRANNs for modeling pharmaceutical data as reported by Caballero and Fernandez [13]. This method implemented in Matlab environment [14] is relatively fast and considers the whole data set in training process. For other hybrids of ANN and GA the use of the mean square error as fitness function could lead to undesirable well fitted but poor generalized networks as algorithm solutions. In this connection, BRGNN avoids such results by two aspects: (1) keeping network architectures as simple as possible inside the GA framework and (2) implementing Bayesian regulation in the network training function.

Unlike to other GA approaches, the objective of our algorithm is not to obtain a sole optimum model. It yields a sub-population of well fitted models, with MSE lower than threshold value, where the Bayesian's regularization guaranties good generalization abilities (Figure 1). This is due to we used MSE of data training instead of crossvalidation or test set MSE values as cost function and therefore the optimum model cannot be directly derived from the best fitted model yielded by the genetic search. However, from crossvalidation experiments throughout the subpopulation of well fitted models the most generalizable network can be derived with the highest predictive power. This process also assures to avoid chance correlations. This approach has shown to be highly efficient in comparison with crossvalidation-based GA approach since only optimum models, according to the Bayesian regularization, are crossvalidated at the end of the routine and not all the model generated throughout all the search process.

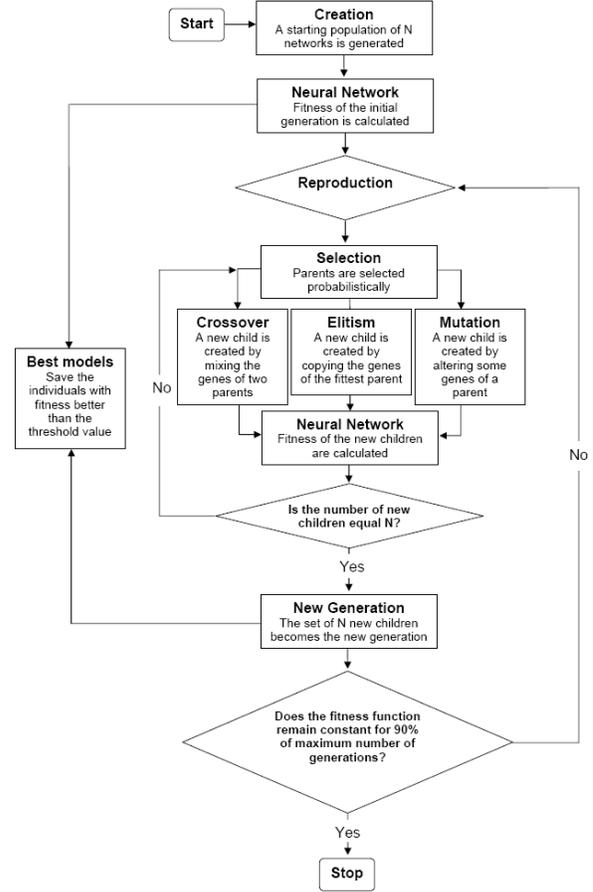


Figure 1. Flow diagram describing the strategy for the genetic algorithm implemented in the BRGNNs.

C. Genetic Algorithm -optimized Support Vector Machine (SVM)

SVM is a machine learning method, which has been used for many kinds of pattern recognition problems [15, 16]. First, the input vectors are mapped onto one feature space (possible with a higher dimension). Secondly, a hyperplane, which can separate two classes, is constructed within this feature space. Only relatively low-dimensional vectors in the input space and dot products in the feature space will involve by a mapping function. SVM was designed to minimize structural risk whereas previous techniques were usually based on minimization of empirical risk. The mapping into the feature space is performed by a kernel function. There are several parameters in the SVM, including the kernel function and regularization parameter. GA-based SVM (GA-SVM) algorithm was implemented for selecting optimum subset of input training vectors and setting the two SVM parameters, regularization parameter and width of the RBF kernel, to optimum values. The toolbox used to implement the SVM with RBF kernel (RBF-

SVM) was LIBSVM for Matlab by Chang and Lin [17] that can be downloaded from: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.

GA was applied for selection of the optimum subset of variables and the optimization of regularization parameter and width of an RBF kernel. We simply concatenated a representation of the parameters to a chromosome encoding the subset of variables used for SVM training [18]. Usually it is not necessary to consider any arbitrary value but only certain discrete values with the form: $n \times 10^k$, where $n=1\dots 9$ and $k=-4\dots 4$. So, these values can be calculated by randomly generating n and k values as integers between $(1\dots 9)$ and $(-4\dots 4)$, respectively. In this way, GA optimized regularization parameter and the width of an RBF kernel.

A three-fold-out crossvalidation assessed model's quality throughout the GA search. Three data subsets were created, two subsets are generated in the crossvalidation process for training the SVM and another subset is then predicted. This process is repeated until all subsets have been predicted. The GA routine minimized the misclassification percent of three-fold-out (MCP_{TFO}) crossvalidation. GA-SVM was implemented in Matlab environment [14].

III. RESULTS AND DISCUSSION

A. Anti-cancer targets

In the context of cancer therapy targets, models were developed to predict the inhibition of farnesyl protein transferase [19], matrix metalloproteinase [20], multiple targets, cyclin-dependent kinase [21], multiple targets antagonists for the luteinizing hormone-releasing hormone (LHRH) receptor [22]. Results from BRGNN modeling of cancer-target datasets appear in Table 1. Numbers of selected features varied accordingly to the size and variability of each data set. The selected features correspond to the molecular descriptors which best describe the affinity of the ligands towards the targets. Models were validated by crossvalidation or/and test set prediction. Validation accuracies were higher than 65% for all datasets.

Cyclin-dependent kinase, LHRH and matrix metalloproteinase inhibitions were modeled using 2D molecular descriptors, which resemble bidimensional distributions of atomic properties on the molecular sketch. Meanwhile, farnesyl protein transferase was modeled by 3D-descriptors encoding distributions of atomic properties on the tridimensional molecular spaces. In the case of the later target, knowledge of the ligand binding mode was used.

It was noteworthy that BRGNN trained with chemical quantum descriptors predicted LHRH antagonist activity with 70% accuracy. Chemical quantum descriptors only encoded information relative to the electronic states of the molecules rather than distribution of chemical groups on the structure. The chemical homogeneity of the macrolides in this dataset suggests a well define and homogenous

electronic pattern that was recognized by the networks after supervised training.

TABLE 1. Details and Accuracies of the Optimum BRGNNs Models.

| Target name | Data size | Numb. Opt. Var. | Validation Accuracy |
|----------------------------------------|--------------------|-----------------|---------------------|
| Farnesyl protein transferase | 78 | 7 | 70% |
| | 128 | 8 | 75% |
| Matrix metalloproteinase | 27-30 ^a | 6 | ~70% ^a |
| | 63-68 ^b | 7 | ~80% ^b |
| LHRH(non-peptide) | 128 | 8 | 75% |
| LHRH (erythromycin A analogs) | 38 | 4 | 70% |
| HIV-1 protease | 55 | 4 | 70% |
| Potassium channel | 29 | 3 | 91% |
| Calcium channel | 60 | 5 | 65% |
| Acetylcholinesterase (tacrine analogs) | 136 | 7 | 74% |
| Acetylcholinesterase (huprine analogs) | 41 | 6 | 84% |
| Cruzain | 46 | 5 | 75% |

a. Average values of five models for MMP-1, MMP-2, MMP-3, MMP-9 and MMP-13 matrix metalloproteinases.

b. Average values of five models for MMP-1, MMP-9 and MMP-13 matrix metalloproteinases.

Models of ligand-target binding stability of multiple kinase and protease, developed on systematical selection of combined features of target and inhibitors, successfully classified about 80% of more than 1700 and 2200 target-inhibitor pairs for protease and kinase, respectively (Table 2). In general, it was found that hydrophobic and electrostatic natures of the residues along the sequence ruled target-ligand inhibitions, whereas hydrophobicity, resembled as polarizability and van der Waals properties, were the most relevant ligand feature ruling the interaction with targets. Furthermore, polarizability also accounts for the deformability of the ligand for interacting with the active site. This result also suggests that the higher variability of the dataset corresponds to the hydrophobic moieties on the ligands structure. Predictors are available online at [://gibk21.bse.kyutech.ac.jp/llamosa/ProteaseGA-SVM/ProteaseGA-SVM.html](http://gibk21.bse.kyutech.ac.jp/llamosa/ProteaseGA-SVM/ProteaseGA-SVM.html) and <http://gibk21.bse.kyutech.ac.jp/llamosa/KinaseGA-SVM/KinaseGA-SVM.html>.

B. Acetylcholinesterase inhibition

The loss of the basal forebrain cholinergic system is the most significant aspect of neurodegeneration in the brains of neurodegenerative Alzheimer's disease (AD) patients, and it is thought to play a central role in producing the cognitive

impairments [23]. Therefore, enhancement of cholinergic transmission has been regarded as one of the most promising methods for treating (AD) patients. In this regards, we applied BRGNN to model the acetylcholinesterase inhibition by huprine- and tacrine-like inhibitors.

For the huprines [24] and tacrines [25] datasets, the GA allows exploring a wide pool of 3D-descriptors. The predictive capacity of our selected model was evaluated by averaging multiple validation sets generated as members of neural network ensembles (NNEs). The tacrines model showed adequate test accuracy about 71% (Table 1). Meanwhile, the tacrine dataset was also evaluated using NNEs averaging. The ensemble averaging provided reliable statistics. When 40 members were assembled, the NNE provides a reliable high accuracy of 85%. The higher accuracy yielded for the huprine-type dataset depends on the higher chemical variability of tacrine-like inhibitors in comparison to the huprine-like.

TABLE 2. Dataset Details and Accuracies of Optimum GA-SVM Models.

| Target name | Data size | Numb. Opt. Var. | Validation Accuracy |
|---------------------------------------------------|-----------|-----------------|---------------------|
| Voltage-gated K ⁺ channel ^a | ~100 | 3 | ~85% |
| Ghrelin receptor | 23 | 2 | 93% |
| Kinase | >2200 | 10 | 80% |
| Protease | >1700 | 10 | 80% |

a. Average over three physiological variable models

C. HIV-1 protease inhibition

One of the crucial stages in the HIV-1 life cycle is the protease-mediated transformation from the immature, non-dangerous virion, to the mature, infective virus. HIV-1 protease inhibitors have thus become a major target for anti-AIDS drug design, its inhibition has been shown to extend the length and improve the quality of life of AIDS patients. A large number of inhibitors have been designed, synthesized, and assayed, and several HIV-1 protease inhibitors are now utilized in the treatment of AIDS [26]. Cyclic urea derivatives are among the most successful candidates for AIDS targeting and BRGNN was successfully applied to model their activities towards HIV-1 protease [27]. 2D encoding was used in order to avoid tridimensional conformational noise in the dataset and the optimum BRGNN model accurately predicted IC_{50} values with 70% accuracy in validation test for 55 cyclic urea derivatives (Table 1). Inhibitory activity variations due to different substituent groups allocated around the cyclic urea scaffold were learned by the networks and the activity of new compounds were adequately predicted. Here again, despite the information was only 2D relevant, the problem was successfully attained by the nonlinear approach.

D. Potassium-channel and Calcium entry blocker activities

A model of the selective inhibition of the intermediate-conductance Ca²⁺-activated K⁺ (IK_{Ca}) by some clotrimazole analogs was developed BRGNNs [28]. K⁺ channels constitute a remarkably diverse family of membrane-spanning proteins that have a wide range of functions in electrically excitable and unexcitable cells. Several compounds have been shown to block the IK_{Ca} -mediated Ca²⁺-activated K⁺ permeability in red blood cells. Substitutions around triarylmethane scaffold yielded a differential inhibition of the K⁺ channel by triarylmethane analogs that was encoded in 2D descriptors. BRGNN approach yielded a remarkable accurate model describing more than 90% of data variance in validation experiment. Interactions with the ion channel were encoded in topological charge variables and the homogeneity of the dataset again assures very high prediction accuracy.

Similarly, a dataset of analogs of the widely used diltiazem were gathered reporting negative inotropic activities [29]. The experiment measures the myocardial activity of the compounds focusing cardiac failure treatment. However, optimum BRGNN model exhibited poor accuracy about 65%. This low performance reflects the complexity of the cellular cardiac response which is a multifactor event ruled by membrane trespassing and receptor affinities in comparison to single affinity measure such as enzyme inhibition.

E. Antiprotozoan activity

Trypanosoma cruzi, a parasitic protozoan, is the causative agent of the Chagas disease or American trypanosomiasis, one of the most threatening endemics in Central and South America. The primary cysteine protease of T. cruzi, cruzain, is expressed throughout the life cycle and is essential for the survival of the parasite within host cells. Thus, inhibiting cruzain has become interesting for the development of potential therapeutics for the treatment of the Chagas disease. The inhibition constant (K_i) of a set of 46 ketone-based cruzain inhibitors (KCI) against cruzain was successfully modeled by means of data-diverse ensembles of BRGNNs using 2D molecular descriptors with accuracy about 75% [30]. The BRGNNs overcame GA-optimized partial least squares models suggesting that functional dependence between affinity and the chemical features of the inhibitors have a strong nonlinear component.

F. Receptor and ion channel physiological responses

Modeling of target functional properties had been also carried out by genetic algorithm-optimized QSAR. SVM-based function mapping and binary classification were carried out for ghrelin receptor [31] and voltage-gated K⁺ channel functional properties [32], respectively. Target

information was encoded in 2D descriptors calculated over protein sequences. Both regression and classification tasks were properly attained with accuracies about 93% and 85%, respectively (Table 2). Ghrelin receptor model was remarkable accurate and depended on only two descriptors that allowed plotting a functional response surface of the receptor. On the other hand, the accuracy of voltage-gated K^+ channel model was superior to other nine GA-wrapper classifiers [32].

IV. CONCLUSIONS

GA optimization of BRGNNs and SVMs yielded robust and accurate predictors for a variety of ligand-target datasets relevant to drug design nowadays. The methodology also allowed to identification of relevant structural features ruling the chemical-biological interactions in the studied systems.

REFERENCES

- J. Gasteiger. Chemoinformatics: a new field with a long tradition. *Anal Bioanal. Chem.* 2006, 384:57–64.
- R. D. Cramer, D. E. Patterson and J. D. Bunce. Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins. *J Am Chem Soc* 1988, 110:5959–5967.
- G. Klebe, U. Abraham and T. Mietzner. Molecular similarity indices in a comparative analysis (CoMSIA) of drug molecules to correlate and predict their biological activity. *J. Med. Chem.* 1994, 37:4130–4146.
- G. Folkers, A. Merz and D. Rognan. CoMFA: scope and limitations, in: Kubinyi, H. (Ed.), 3D-QSAR in Drug Design. Theory, Methods and Applications, ESCOM Science Publishers BV, Leiden, 1993, pp. 583–618.
- C. Hansch, A. Kurup, R. Garg and H. Gao. Chem-Bioinformatics and QSAR: A Review of QSAR Lacking Positive Hydrophobic Terms. *Chem. Rev.* 2001, 101:6196–72.
- A. Sabljic. Topological indices and environmental chemistry, Karcher, W. and Devillers, J. (Eds.) Practical applications of quantitative structure–activity relationships (QSAR) in environmental chemistry and toxicology, Kluwer Academic Publications, Dordrecht, 1990, pp. 61–82.
- M. Karelson, V. S. Lobanov and A. R. Katritzky. Quantum–Chemical Descriptors in QSAR/QSPR Studies. *Chem. Rev.* 1996, 96: 1027–1043.
- D. J. Livingstone, D. T. Manallack and I. V. Tetko. Data modelling with neural networks: Advantages and limitations. *J. Comput.–Aided Mol. Des.* 1997, 11:135–142.
- R. Burbidge, M. Trotter, B. Buxton and S. Holden. Drug design by machine learning: support vector machines for pharmaceutical data analysis. *Comput Chem*, 2001, 26:5–14.
- MOPAC 6.0. Frank J. Seiler Research Laboratory, Colorado Springs, CO, 1993.
- D. J. C. Mackay. A practical Bayesian Framework for Backprop Networks. *Neural Comput.* 1992, 4:448–472.
- D. J. C. Mackay. Bayesian Interpolation. *Neural Comput.* 1992, 4:415–447.
- J. Caballero and M. Fernández, Linear and non-linear modeling of antifungal activity of some heterocyclic ring derivatives using Multiple Linear Regression and Bayesian-Regularized Neural Networks. *J Mol Model*, 2006, 12:168–181.
- MATLAB 7.0. The Mathworks Inc., Natick, MA, 2004. <http://www.mathworks.com>.
- C. Cortes and V. Vapnik. Support-Vector Networks. *Mach. Learn.* 1995, 20:273–297.
- C. J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Min Knowledge Discovery*, 1998, 2:1–47.
- C. Chih-Chung and L. Chih-Jen. LIBSVM: a library for support vector machines. 2001, <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- H. Fröhlich, O. Chapelle and B. Schölkopf. Feature Selection for Support Vector Machines by Means of Genetic Algorithms. *Proc. 15th IEEE Int. Conf. on Tools with AI*, 2003, 142–148.
- M. P. González, J. Caballero, A. Tundidor-Camba, A. M. Helguera and M. Fernández. Modeling of Farnesyltransferase Inhibition by some Thiol and Non-Thiol Peptidomimetic Inhibitors using Genetic Neural Networks and RDF Approaches. *Bioorg. Med. Chem.* 2006, 14:200–213.
- M. Fernández, J. Caballero and A. Tundidor-Camba. Linear and nonlinear QSAR study of N-hydroxy-2-[(phenylsulfonyl)amino]acetamide derivatives as matrix metalloproteinase inhibitors. *Bioorg Med Chem* 2006, 14: 4137–4150.
- M. Fernández, A. Tundidor-Camba and J. Caballero. Modeling of Cyclin-Dependent Kinase Inhibition by 1H-pyrazolo [3,4-d] pyrimidine Derivatives using Artificial Neural Networks Ensembles. *J Chem Inf Comput Sci* 2005, 45:1884–1895.
- M. Fernández and J. Caballero. Bayesian-regularized genetic neural networks applied to the modeling of non-peptide antagonists for the human luteinizing hormone-releasing hormone receptor. *J Mol Graph Model* 2006, 25:410–422.
- R. T. Bartus, R. L. Dean, B. Beer, and A. S. Lippa. The cholinergic hypothesis of geriatric memory dysfunction. *Science* 1982, 217:408–417.
- M. Fernández and J. Caballero. Ensembles of Bayesian-regularized genetic neural networks for modeling of acetylcholinesterase inhibition by huprines. *Chem Biol Drug Des* 2006, 68:201–12.
- M. Fernandez, M. C. Carreiras, J. L. Marco and J. Caballero. Modeling of acetylcholinesterase inhibition by tacrine analogues using Bayesian-regularized Genetic Neural Networks and ensemble averaging. *J Enz Inh Med Chem* 2006, 21:647–661.
- R. A. Katz and A. M. Skalka. The retroviral enzymes. *Annu. Rev. Biochem.* 1994, 63:133–173.
- M. Fernández and J. Caballero. Modeling of Activity of Cyclic Urea HIV-1 Protease Inhibitors using Regularized-Artificial Neural Networks. *Bioorg Med Chem* 2006, 14:280–294.
- J. Caballero, M. Garriga and M. Fernández. Genetic neural network modeling of the selective inhibition of the intermediate-conductance Ca^{2+} -activated K^+ channel by some triarylmethanes using topological charge indexes descriptors. *J Comput-Aided Mol Des*, 2005, 19:771–789.
- J. Caballero, M. Garriga, and M. Fernández. 2D Autocorrelation modeling of the negative inotropic activity of calcium entry blockers using Bayesian-regularized genetic neural networks. *Bioorg. Med. Chem.* 2006, 14:3330–3340.
- J. Caballero, A. Tundidor-Camba, and M. Fernández, Modeling of the Inhibition Constant (K_i) of Some Cruzain Ketone-Based Inhibitors Using 2D Spatial Autocorrelation Vectors and Data-Diverse Ensembles of Bayesian-Regularized Genetic Neural Networks. *QSAR Comb Sci*, 2006, 26: 27–40.
- J. Caballero, L. Fernández, M. Garriga, J. I. Abreu, Collina, S. and Fernández, M. Proteometric study of ghrelin receptor function variations upon mutations using amino acid sequence autocorrelation vectors and genetic algorithm-based least square support vector machines. *J Mol Graph Model* 2007, 26:166–178.
- M. Fernández, L. Fernández, J. I. Abreu and M. Garriga. of voltage-gated $K(+)$ ion channels from 3D pseudo-folding graph representation of protein sequences using genetic algorithm-optimized support vector machines. *J Mol Graph Model* 2008, 26:1306–1314.