http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 6 (5): 718-724, 2007
ISSN 1812-5638

Machine Learning Technique Approaches in Drug Discovery, Design and Development

[1]D. Pugazhenthi and [2]S.P. Rajagopalan
[1]S.R.M. Arts and Science College, Kattankulathur, Kanjipuram District-603203, No. 137/778,
32nd Cross Street, T.P. Chathiram, Chennai-600 030, India
[2]School of Computer Science and Engineering,
MGR University, Chennai-600 095, India

Abstract: Drug discovery refers to the finding of a new drug which could be a completely new compound or a new derivative of existing compounds. Drug discovery is the ultimate goal of drug design which concerned with the design of a chemical compound that exhibits a desired pharmacological activity. Machine learning tools, in particular Support Vector Machines (SVM), Particle Swarm Optimisation (PSO) and Genetic Programming (GP), are increasingly used in pharmaceuticals research and development. They are inherently suitable for use with noisy, high dimensional data, as is commonly used in cheminformatic, bioinformatics and other types of drug research studies. These aspects are demonstrated via review of their current usage and future prospects in context with drug discovery activities.

Key words: Drug, SVM, maccroarray, protein, DNA, QSAR

INTRODUCTION

Pharmaceutical discovery and development is an evolving (Ratti and Trist, 2001) cascade of extremely complex and costly research encompassing many facets. Starting from therapeutic target identification and bioinformatics study through candidate drug discovery and optimisation; to pre-clinical organism-level evaluations and beyond to extensive clinical trials assessing effectiveness and safety of new medicines. In recent years, with products of human genome project helping to reveal many new disease targets to which drug treatments might be aimed, all the major pharmaceutical companies have invested heavily in the routine ultra-High Throughput Screening (uHTS) of vast numbers of drug-like molecules guided by cheminformatic investigations. Due to the enormous expense of failures of candidate drugs late in their development, uHTS *in vitro* assays now cover liabilities such as possible side effects as well as therapeutic properties. In parallel with this, drug design and optimisation increasingly uses computers within *in silico* (virtual) screening (Hou and Xu, 2004). State-of-the-art *in vitro* experiments now employ DNA micro-array chips to simultaneously explore the expression of thousands of genes potentially involved in disease, treatment and toxicity. Similar advancements are now becoming possible in proteomics and metabolomics.

Patient-level genetic and single nuclear polymorphism, SNPs (Roses, 2002) data has become more commonly available supporting conventional observational data in epidemiology, clinical trial treatment response and early safety studies that continue as on-going pharmaco vigilance. The curation and storage of all these individual types of data has become more automated, organized and consistent, providing for greater homogeneity and suitability for exploration. Increasingly, vast integrated research datasets are constructed from larger more in homogeneous combinations of data, from disparate sources and disciplines, to answer novel lines of inquiry and for hypothesis generation, possibly not initially envisaged at the time of planning data collection. However, conventional multivariate statistical methods, i.e., principal components analysis and partial least squares, well established against smaller, lower dimensional datasets, are being stretched. Whilst they remain of great utility and continue to be developed in more scaleable commercial tools, they are inherently linear, tending to render them less suitable toward a plethora of newer, ever more complex problem opportunities. Scientists are thus increasingly using data-mining tools such as recursive partitioning and predictive modelling methods to underpin data exploration, using heavy computation to free up and save scientist time. Consequently, evaluation and early uptake

Corresponding Author: D. Pugazhenthi, S.R.M. Arts and Science College, Kattankulathur, Kanjipuram District-603203, No. 137/778, 32nd Cross Street, T.P. Chathiram, Chennai-600 030, India Tel: 98410 13304

of novel predictive modelling approaches continues within pharmaceuticals research. Whilst uses of artificial neural networks and genetic algorithms are well established in older application areas (Zupan and Gasteiger, 1999), in non-expert hands these may yield suboptimal solutions presenting difficulties in newer areas, including situations when the form of the solution is unclear. More recent machine learning approaches, offer key advantages over these and we here illustrate Support Vector Machines, Genetic Programming and Particle Swarm Optimisation. The current state of their pharmaceuticals R and D application is reviewed and their future prospects assessed.

SUPPORT VECTOR MACHINES

The Support Vector Machine (SVM) arose from Vapniks (1992) concepts of structural risk minimisation and statistical learning theory. An algorithm based upon these ideas was first presented at COLT-92 (Boser *et al*., 1992) and a Support Vector Classifier (SVC) formulation was first presented by Vapnik (1995). Todays' SVC, is a sophisticated synthesis of artificial neural network perception-like hyperplane classifier, backed by a sound theory of learning and convergence. It uses robust linear methods and can apply these within kernel spaces to achieve non-linear classifiers with excellent generalization characteristics. The simplest SVCs are maximal margin binary classifiers, placing the optimal separating hyperplane, centrally giving the largest allowable separation between the nearest data points of opposite classes in the training set. They use uniform-class subsets of these points (known as support vectors) to construct respective bounding hyperplanes defining a margin which models the decision surface. In accordance with statistical learning theory, for bias-variance trade off in learning, this margin-maximisation is tied to a function-limiting to avoid over fitting. In achieving this, SVCs are constrained to minimise an estimated upper bound on expected (not empirical) risk, as derived from statistical learning theory, assuming training data is drawn independently and identically distributed from some unknown distribution p(x,y):

$\{ (x_1, y_1), \ldots, (xl, yl) \}$ xi $\in \Re$, with class yi $\in$. { -1, +1 }

Linear SVCs use the dot product of pairs of input vectors as a distance measure. SVCs can also learn a linear hyperplane after projection of the input to a higher-dimensional kernel-feature space. For efficiency, data mapping to kernel space is not explicitly made, although a sparse new space is effectively created aiding model construction. Kernel spaces allow decision boundaries of apparently arbitrary shape in the input feature space and provide an opportunity to incorporate domain knowledge, enabling solutions to very complex problems of diverse nature (Shawe-Taylor and Cristianini, 2004). Support Vector Regression (SVR) and SVC models achieve a data compression, comprising a linear combination of mapped training examples, the SV subset, using a discovered weighting of input features. Implementations of SVC and SVR are constructed as Linear Programming (LP) or Quadratic Programming (QP) problems using appropriate solver technology. Soft margin SVMs use error terms to handle constraint violations from data-points lying beyond their class margin hyperplane, to enable solutions for noisy, or non-linearly separable data.

Pro's: Sound theory and formalism; use robust linear methods; global optimum for convergence; good accuracy, generalisation and robustness to noise; Few user parameters (regularisation parameter, C; kernel parameters), simplify parameterisation compared to neural nets; implicit feature selection; computationally weakly affected by input dimensionality; sparse solution gives fast prediction; Memory linear in the number of training examples.

Con's: Complex operation and model opaque to end user; optimal parameter configuration is data dependent; cannot handle missing data; computational cost quadratic with number of examples; QP implementations restricted to Mercer kernels; effectively non-parametric density estimators giving point predictions, with no confidences or distributions generated.

SVM APPLICATIONS IN PHARMACEUTICALS RESEARCH

SVM in cheminformatics and Quantitative-Structure Activity Relationship (QSAR) modelling: The role of cheminformatics in drug discovery has been reviewed by Xu and Hagler (2002). An early task is the creation of virtually represented molecules' and assessment of their likely suitability for synthesis and viability for use in the body. The study of drug-likeness and report that SVM predictions were more robust than those from neural networks. Cheminformatics combines chemical properties and high throughput screening measurements, often against novel targets, in large scale structure activity modeling. Trained classifiers enable virtual screening for discovering molecules with specific therapeutic target affinities from potentially millions of virtual representations. Ranking and simple enrichment of actives are key aspects as is the discovery of correlated descriptors. Reducing the scale of subsequent physical screening of synthesised molecules and the number of

synthesis-biotesting cycles for their improvement is an ideal setting for active learning. Finding the bioactive conformations of active molecules is key to understanding their mechanisms of action and thus for improving specificity and selectivity. SVM uses in the wider field of chemistry.

Predicting activity toward therapeutic targets: G-proteins provide such a key interface to intra-cellular signal transduction that G-Protein Coupled Receptors (GPCRs) are the major class of drug targets. Suwa *et al.* (2004) provided physicochemical features of GPCRs and their ligands to a Radial Basis Function-SVC (RBF-SVC) to predict specific G-protein couplings with high degrees of success. RBF-SVR used to predict both antagonist compound metabolism and inhibitory activity toward human glucagons receptor in order to select useful 3-d QSAR features. Heuristis methods are applied SVM to a variety of QSAR problems (Burbidge, 2004) and found good performance can be achieved at the expense of sparsity, i.e., a large number of training points are support vectors.

Predicting Absorption Distribution Metabolism Excretion Toxic effects (ADMET): Amongst the first to investigate the utility of SVC in QSAR modeling, (Burbidge, 2004) favorably compared SVC to back-propagation and radial-basis function neural networks and K-nearest-neighbor classifiers against human blood-brain barrier, human oral bioavailability and protein-binding classification problems. P-glycoprotein (P-gp) active molecular transport in bacterial cells may act as effective efflux pump for antibiotics which are substrates, resulting in drug resistance.

SVM IN BIOINFORMATICS

Gene expression micro-array data in the prediction of disease traits: As with SNPs data, dimensionality P of this input can be extremely large (10 Ks of genes) whilst the number of examples N is relatively small (typically a few 10 to 100 s). Whilst it is clear that SVMs are well suited to this kind of situation (Malossini *et al.*, 2004) showed that performance can significantly degrade if some training examples are incorrectly labeled. Furthermore increasing the number of correctly labeled training examples does not counter the presence of incorrectly labeled examples. Large numbers of poorly correlated, correlated and irrelevant genes also diminish performance, making feature selection essential.

Receptor classification and protein function annotation: SVC prediction of the functional classes of proteins from sequence data is now quite common and (Karchin *et al.*, 2002) were first to achieve this for GPCR families and sub-families using efficient hierarchical multi-class SVC tree.

Other bioinformatics applications: Schrattenholz (2004) have reviewed machine learning approaches (including SVMs) to protein sub-cellular localisation for target identification in drug discovery. There is a growing use of SVC prediction of functionally critical sites within proteins.

EVOLUTIONARY COMPUTING

In contrast to the rigorous mathematical approach of SVMs, evolutionary computation (of which genetic programming is the most advanced variant) appeals to metaphor. The basic idea is to use the ideas of Darwinian evolution within the computer. So we have a population of individuals. A fitness function calculates how good each member of the population is. The better ones are selected to be parents for the next generation. Children are created by crossover and/or mutation of the selected individuals from the previous generation. As in natural evolution, the children are not identical to their parents. Some are better, some are worse. So in the next generation, selection will again only allow the better individuals to pass their genes onto the next generation. Hopefully overtime and successive generations the population will improved until an individual with satisfactory performance is found. Such an elegant idea has occurred, apparently independently, to many computer scientists. So who was first, is somewhat controversial. However Turing, Rechenberg, Holland and Fogel all make a claim for primacy. From its diverse starting points several subfields of evolutionary computation (Evolution strategy, genetic algorithms, evolutionary programming, etc.) have thrived. However, because of its simple appeal, it has been successfully applied many times. Examples include: optimisation, particularly of engineering design, scheduling, economic and financial modelling, fraud detection and data mining. Each sub-field lays stress on different aspects of evolution, e.g., crossover versus mutation, large or small populations and should we represent numbers as bits or as floating point numbers. We will concentrate upon a relative new comer, genetic programming.

GENETIC PROGRAMMING

Genetic programming, uses Holland's crossover heavy Genetic Algorithm, to evolve programs. So while

other approaches require the software engineer to design an evolutionary friendly way of representing their problem solution, GP does not force this representation to be fixed up front, instead it too can evolve.

Pro's genetic programming combines a flexible problem representation with a powerful search mechanism. Many computational chemistry problems can be expressed as the problem of finding a computer program e.g., given known properties of a chemical, can we predict some other property (particularly disease binding, toxicology, blood take up). Having recast the problem, the genetic algorithm (GA, used by GP) is a powerful way of searching for a solution which requires minimal assumptions.

Con's genetic programming offers no guarantee that it will find a suitable solution within an acceptable amount of time. In practice GP has solved difficult but economically interesting problems (for which it is known that no guarantee is possible). While many of the new techniques require more computation time, computer power is increasingly available.

DRUG RESEARCH APPLICATIONS OF GENETIC PROGRAMMING

In most Pharmaceutical applications, the evolved programs are models. That is, while we can view them as programs which we run and which produce answers, mostly GP is restricted to producing functions. These take known facts or measurements (e.g., number of positively charged ions, presence of aromatic rings, acidity) and produce a single number. Then we treat the number as a prediction. For example, a positive number might indicate that the evolved model predicts that the molecule inhibits normal enzyme activity. There is an increasing body of work using evolutionary computation in Biology. For example there are now at least two annual workshops. BioGEC (2002-05) is held in conjunction with the GECCO conference and EvoBIO (2003-05) which is co-sited with EuroGP. Genetic programming figures heavily in both. The June, 2004 special issue of the GP journal featured biological applications.

GP IN CHEMINFORMATICS AND QSAR

Genetic programming has been used for combinatorial design (Nicolotti *et al*., 2002) modelling drug bioavailability and GP ensembles of ANNs have been developed to predict p450 inhibition

GP IN BIOINFORMATICS

Hot topics include sequence alignment (typically of either DNA or proteins); protein localisation (Heddad *et al*., 2004) using genetic algorithms etc., to infer phylogenetics trees for classification and prediction recognizing parts of proteins (e.g., transmembrane regions) or in the case of DNA, creating algorithms to find promoters and other gene regulatory sites. Infrared spectroscopy (wave number), DNA chip and Single Nucleotide Polymorphisms (SNPs) (Reif *et al*., 2004) datasets are noted for having huge numbers of input features. In these cases, while a predictive model might be of use, the immediate problem is to discover which of the thousands of data actually relate to the underlying biology. GP based prediction has also been used with DNA chip data in a mode in which, although it generates predictive models, the principle interest is to use GP to sift hundreds or thousands of inputs in order to discover which genes are important to a metabolic process or to reduce the number of inputs required so a diagnostic test is practicable. While GAs can achieve high multi-class accuracy, they are also commonly combined with other classifiers, e.g., linear SVM, naive Bayes and k-nearest neighbor, where the bit string GA selects which genes can be used by the second classifier. It is no wonder that GP is increasingly being used in Bioinformatics data mining and increasingly this includes: modelling genetic interactions and organisms; inferring metabolic pathways and gene regulatory networks.

PARTICLE SWARM OPTIMISATION

Particle Swarm Optimization (PSO) is a population based stochastic optimisation method inspired by observation of swarms of insects, shoals of fish, etc., For example, millions of insects can build complex cathedral termite mounds, apparently without central or hierarchical control. Instead each individual acts by themselves in response its environment. Chemical signals provide simple distributed communication between nearby (in space and time) agents. PSO simplifies still further swarms for use in computers for optimisation. The agents are abstracted to particles (like electrons, protons etc., from Physics). These have position and speed. They interact with each other via spring like forces. The particles fly over the problem space. Each time step they sample where they are to determine how good it is. If it is better than any place they have visited, an attractive force is set up which attracts them back to it. There is a similar social cognitive force which attracts the particle to the best place found by the particle's neighbors. A binary extension of PSO (BPSO) is made by replacing the continuous search space by a probability space, i.e., 0.1 in each dimension. At each time step the particle's location is probabilistically converted to a binary string. e.g., a particle at 0.94 along a particular dimension of the problem, has a 94% chance of sampling binary value 1 (true) and only a 6% chance of sampling false.

Pro's: PSO and BPSO are capable of solving a wide range of very different applications without expensive human up front design.

Con's: Like every blind (i.e., problem independent) search technique, PSO do not have a guarantee of success. Nevertheless, as we shall see, despite being originally designed for classic optimisation benchmarks, PSO have been successfully transferred to biological applications.

BIOLOGICAL APPLICATIONS OF PARTICLE SWARM OPTIMISATION

Unlike genetic programming, at present, the use of Particle Swarm Optimisation (PSO) in pharmaceutical research is relatively unexplored. However it is common to use PSO in conjunction with other approaches. This hybrid approach comes from the fact that PSOs naturally search extensively, making them suitable for finding good regions. Often, currently, a more exploitive local method is needed to refine the good starting points found by PSOs into excellent solutions. However as PSOs and their features such as friction (constriction) and momentum become better understood, we anticipate PSOs will tend to be used in a more dominant role.

PSO IN CHEMINFORMATICS AND QSAR

In QSAR a few teams have used a two stage approach. In the first stage a binary PSO is used to select a few (typically 3-7) features as inputs to supervised learning method. In Liu *et al.* (2004) the BPSO selects 7 of 85 features. Then linear models of drug activity (IC50) with two enzymes, COX-1 and COX-2, are constructed. Some existing drugs (e.g., Aspirin) bind to both COX enzymes, leading to potentially fatal side-effects. Liu *et al.* (2004) produce models which can potentially differentiate between binding to the two enzymes by virtual chemicals, i.e., as an aid to *in silico* design of drugs before the decision is made to manufacture and test the physical chemicals. Both Wang *et al.* (2004) and Shen *et al.* (2004) use feed-forward artificial neural networks to classify the bio-activity of chemicals using a few features selected by a BPSO. Wang *et al.* (2004) investigates two ways of using PSO to train the ANN. Either the network is trained in a conventional way or by using another PSO. Shen *et al.* (2004) also consider replacing the ANN by a k-nearest neighbor classifier in combination with kernel regression. While they note some differences, many approaches turn out to have similar performance at predicting which chemicals will be carcinogenic. The datasets cover typically only cover a few chemicals but a large number of features are computed for each from the chemical's formula. One can reasonably argue that some form of feature selection, i.e., choosing which attributes can be used by the ANN, is essential.

PSO IN BIOINFORMATICS

The problem of small but wide datasets becomes even more apparent when dealing with DNAchip datasets. Xiao *et al.* (2003) suggests a novel combination of PSO and Self Organising Maps (SOM). Instead of finding the few relevant genes, they use SOM to pick clusters of similarly behaved genes from datasets with thousands of gene measurements. The PSO swarm is seeded with crude results produced by the SOM and then used to refine the clusters.

CONCLUSIONS

Whilst the above survey clearly demonstrates a wide coverage of relevant problem areas, it remains unclear as to the underlying extent to which these reported machine learning approaches are actually deployed within pharma R and D, so their importance here is difficult to ascertain. Although becoming less sporadic, it seems that the use of machine learning is still largely driven by individuals either with their own expertise and/or external expert resources. Conventional statistical methods are currently better known and understood by scientists. They benefit from their traditional supporting design of experiments, data capture and preparation making them difficult to displace on a wider scale. Statisticians continue to dominate pharmaceutical company quantitative analysis groups. However statistics is becoming increasingly computational and recognising alternative approaches (Breiman, 2001) as existing (usually hypothesis testing) methods are found lacking. This is generally due to the increasing need for data exploration and hypothesis generation in the face of growing data, problem complexities and ad hoc experimental design inadequacies and from compromises due to cost and lack of prior knowledge. An important recent problem is the integrated analysis of combined omics-type data in surrogate biomarker and systems biology research. Here the numerical dominance of variables from genomics, currently swamps those from other types of data in existing methods where all variables (as opposed to the fundamentally different types of information) are treated equally. As individual methods and accompanying validation procedures may only partly cope with problems, multiple methods are often used for comparative analyses in the hope that inappropriate model biases, costly false negatives or effort-producing false positives, are minimised. SVMs have, however, proved their worth in many areas and for this technology

to make further applications advances there is a need for easier derivation of problem-specific kernel representations, i.e., using structured (ontological) data, or kernel-based data-fusion (Lanckriet *et al*., 2004) adequate ways of handling missing data; more widespread generation of confidence measures of prediction and attention to statistical power of datasets in model selection, which itself continues to present problems especially for SVR end-users. Similar kinds of difficulties hamper the uptake of evolutionary methods by non-expert users, although model transparency (as well as performance) here is a strongly recognized benefit and worthy commercially available tools are now appearing. Encouragingly, the machine learning research community keeps aware and responds to publicized needs. Deficiencies in individual methods are being countered by customizations, ensemble and hybrid approaches. For example, in QSAR, individual classifiers can be inadequate in the face of vast molecular spaces and multi-mechanism problems. GP classifier fusion was developed to form good ensembles of weak or niche classifiers using Receiver Operating Characteristics (ROC) curve area as fitness. Whilst GAs are commonly used as feature selectors for SVM they are becoming integrated (Li, 2005) and sophisticated hybrids of complementary evolutionary and SVM technique are appearing for kernel development, parameter tuning, alternative QP solvers and model selection. An ease of blending of these and other techniques incorporating multi-objective capabilities is awaited with anticipation for challenges in areas like gene regulatory mechanisms discovery (Burckin *et al*., 2005), selectively non-selective drug design (Roth *et al*., 2004), clinical trials simulation and personalized of medicines.

REFERENCES

Boser *et al*., 1992. A training algorithm for optimal margin classifiers. 5th Annual ACM Workshop, COLT, 1992.

Breiman, 2001. Random forests. Machine Learning, 45: 5-32.

Burckin *et al*., 2005. Exploring functional relationships between components of gene expression machinery. Nat. Struct. Mol. Biol., 12: 175-182.

Burbidge, 2004. Heuristic methods for support vector machines with applications to drug discovery. Ph.D Thesis, University College London, London University, UK.

Heddad *et al*., 2004. Evolving Regular Expression-based Sequence Classifiers for Protein Nuclear Localization. In: Applications of Evolutionary Computing. Raidl *et al*. (Eds.), LNCS 3005, pp: 31-40.

Hou and Xu, 2004. Recent development and application of virtual screening in drug discovery. An overview. Curr. Pharma. Design, 10: 1011-1033.

Karchin *et al*., 2002. Classifying G-protein coupled receptors with support vector machines. Bioinformatics, 18: 147-159.

Lanckriet *et al*., 2004. A statistical framework for genomic data fusion. Bioinformatics, 20: 2626-2635.

Li, 2005. Preclinical *in vitro* screening assays for drug-like properties: Drug discovery today. Technologies, 2: 179-185.

Liu *et al*., 2004. QSAR and classification models of a novel series of COX-2 selective inhibitors: 1, 5-diary limidazoles based on support vector machines. J. Comput. Aided Mol. Design, 18: 389-399.

Malossini *et al*., 2004. Assessment of SVM reliability for microarrays data analysis. In: Proc. 2nd European Workshop on Data Mining and Text Mining for Bioinformatics. Pisa, Italy, 2004.

Nicolotti *et al*., 2002. Multiob jective optimization in quantitative structure-activity relationships: Deriving accurate and interpretable QSARs. J. Med. Chem., 45: 5069-5080.

Ratti and Trist, 2001. Continuing evolution of the drug discovery process in the pharmaceutical industry. Pure Applied Chem., 73: 67-75.

Reif *et al*., 2004. Integrated analysis of genetic, genomic and proteomic data. Expert Rev. Proteomics, 1: 67-75.

Roses, 2002. Genome-based pharmacogenetics and the pharmaceutical industry. Nat. Rev. Drug Discov., 1: 541-549.

Roth *et al*., 2004. Magic shotguns versus magic bullets: selectively non-selective drugs for mood disorders and schizophrenia. Nat. Rev. Drug Discovery, 3: 353-359.

Shawe-Taylor and Cristiamini, 2004. Kernel Methods for Pattern Analysis. Cambridge University Press. ISBN: 0 521 81397 2.

Schrattenholz, 2004. Proteomics: How to control highly dynamic patterns of millions of molecules and interpret changes correctly? Drug discovery today. Technologies, 1: 1-8.

Shen *et al*., 2004. Hybridized particle swarm algorithm for adaptive structure training of multilayer feed-forward neural network. QSAR studies of bioactivity of organic compounds. J. Computational Chem., 25: 1726-1735.

Suwa *et al*., 2004. GPCR and G-protein Coupling Selectivity Prediction Based on SVM with Physico-Chemical Parameters. GIW 2004. Poster Abstract: P056. http://www.jsbi.org/journal/GIW04/GIW04Poster.html.

Vapnik, V., 1992. Principles of Risk Minimization for Learning Theory. In: Advances in Neural Information Processing Systems. Moody, J.E., S.J. Hanson and R.P. Lippmann (Eds.), Vol. 4; Morgan Kaufmann Publishers, Inc.

Vapnik, V.N., 1995. The Nature of Statistical Learning Theory. Springer: New York.

Wang *et al*., 2004. Particle swarm optimization and neural network application for QSAR. In HiCOMB.

Xiao *et al*., 2003. Gene clustering using self-organizing maps and particle swarm optimization. In HiCOMB.

Xu and Hagler, 2002. Chemoinformatics and drug discovery. Molecules, 7: 566-600.

Zupan and Gasteiger, 1999. Neural Networks in Chemistry and Drug Design. An Introduction. 2nd Edn., John Wiley, ISBN: 3-527-29778-2, pp: 400.