1. Introduction

In general, chemical problems are composed by complex systems. There are several chemical processes that can be described by different mathematical functions (linear, quadratic, exponential, hyperbolic, logarithmic functions, etc.). There are also thousands of calculated and experimental descriptors/molecular properties that are able to describe the chemical behavior of substances. In several experiments, many variables can influence the chemical desired response [1,2]. Usually, chemometrics (scientific area that employs statistical and mathematical methods to understand chemical problems) is largely used as valuable tool to treat chemical data and to solve complex problems [3-8].

Initially, the use of chemometrics was growing along with the computational capacity. In the 80’s, when small computers with relatively high capacity of calculation became popular, the chemometric algorithms and softwares started to be developed and applied [8,9]. Nowadays, there are several softwares and complex algorithms available to commercial and academic use as a result of the technological development. In fact, the interest for robust statistical methodologies for chemical studies also increased. One of the most employed statistical methods is partial least squares (PLS) analysis [10,11]. This technique does not perform a simple regression as multiple linear regression (MLR). PLS method can be employed to a large number of variables because it treats the colinearity of descriptors. Due the complexity of this technique, when compared to other statistical methods, PLS analysis is largely employed to solve chemical problems [10,11].
We can cite some examples of computational packages employed in chemometrics and containing several statistical tools (PLS, MLR, etc.): MATLAB [12], R-Studio [13], Statistica [14] and Pirouette [15]. There are some molecular modeling methodologies as HQSAR [16], CoMFA [17-18], CoMSIA [19] and LTQA-QSAR [20] that also use the PLS analysis to treat their generated descriptors. In general, the PLS method is used to analyse only linear problems. However, when a large number of phenomena and noise are present in the calibration problem, the relationship becomes non-linear [21]. Therefore, artificial neural networks (ANNs) may provide accurate results for very complex and non-linear problems that demand high computational costs [22,23]. One of the most employed learning algorithm is the back-propagation and its main advantage is the use of output information and expected pattern to error corrections [24]. The main advantages of ANN techniques include learning and generalization ability of data, fault tolerance and inherent contextual information processing in addition to fast computation capacity [25]. It is important to mention that since 90’s many studies have related advantages of applying ANN techniques when compared to other statistical methods [23,26-31].

Due to the popularization, there is a large interest in ANN techniques, in special in their applications in various chemical fields such as medicinal chemistry, pharmaceutical, theoretical chemistry, analytical chemistry, biochemistry, food research, etc [32-33]. The theory of some ANN methodologies and their applications will be presented as follows.

2. Artificial Neural Networks (ANNs)

The first studies describing ANNs (also called perceptron network) were performed by McCulloch and Pitts [34,35] and Hebb [36]. The initial idea of neural networks was developed as a model for neurons, their biological counterparts. The first applications of ANNs did not present good results and showed several limitations (such as the treatment of linear correlated data). However, these events stimulated the extension of initial perceptron architecture (a single-layer neural network) to multilayer networks [37,38]. In 1982, Hopfield [39] described a new approach with the introduction of nonlinearity between input and output data and this new architecture of perceptrons yielded a good improvement in the ANN results. In addition to Holpfield’s study, Werbos [40] proposed the back-propagation learning algorithm, which helps the ANN popularization.

In few years (1988), one of the first applications of ANNs in chemistry was performed by Hoskins et al. [41] that reported the employing of a multilayer feed-forward neural network (described in Session 2.1) to study chemical engineering processes. In the same year, two studies employing ANNs were published with the aim to predict the secondary structure of proteins [42,43].

In general, ANN techniques are a family of mathematical models that are based on the human brain functioning. All ANN methodologies share the concept of “neurons” (also called “hidden units”) in their architecture. Each neuron represents a synapse as its biological counterpart. Therefore, each hidden unity is constituted of activation functions that control
the propagation of neuron signal to the next layer (e.g. positive weights simulate the excitatory stimulus and negative weights simulate the inhibitory ones). A hidden unit is composed by a regression equation that processes the input information into a non-linear output data. Therefore, if more than one neuron is used to compose an ANN, non-linear correlations can be treated. Due to the non-linearity between input and output, some authors compare the hidden units of ANNs like a “black box” [44-47]. Figure 1 shows a comparison between a human neuron and an ANN neuron.

![Figure 1](image)

**Figure 1.** (A) Human neuron; (B) artificial neuron or hidden unity; (C) biological synapse; (D) ANN synapses.

The general purpose of ANN techniques is based on stimulus–response activation functions that accept some input (parameters) and yield some output (response). The difference between the neurons of distinct artificial neural networks consists in the nature of activation function of each neuron. There are several typical activation function used to compose ANNs, as threshold function, linear, sigmoid (e.g. hyperbolic tangent), radial basis function (e.g. gaussian) [25,44-48]. Table 1 illustrates some examples of activation functions.

Different ANN techniques can be classified based on their architecture or neuron connection pattern. The feed-forward networks are composed by unidirectional connections between network layers. In other words, there is a connection flow from the input to output direction. The feedback or recurrent networks are the ANNs where the connections among layers occur in both directions. In this kind of neural network, the connection pattern is characterized by loops due to the feedback behavior. In recurrent networks, when the output signal of a neuron enter in a previous neuron (the feedback connection), the new input data is modified [25,44-47].
### Table 1. Some activation functions used in ANN studies.

<table>
<thead>
<tr>
<th>Function Type</th>
<th>Mathematical Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Threshold</strong></td>
<td>( \phi(r) = \begin{cases} 1 &amp; \text{if } r &gt; \frac{1}{2}; n \geq \frac{1}{2} \ 0 &amp; \text{if } n \leq -\frac{1}{2} \end{cases} )</td>
</tr>
<tr>
<td><strong>Linear</strong></td>
<td>( \phi(r) = \begin{cases} 1 &amp; \text{if } v \geq \frac{1}{2} \ 0 &amp; \text{if } v &lt; 0 \end{cases} )</td>
</tr>
<tr>
<td><strong>Hyperbolic Tangent</strong></td>
<td>( \phi(v) = \text{tanh}(\frac{v}{2}) = \frac{1 - \exp(-v)}{1 + \exp(-v)} )</td>
</tr>
<tr>
<td><strong>Gaussian</strong></td>
<td>( \phi(r) = e^{-r^2} )</td>
</tr>
</tbody>
</table>

Each ANN architecture has an intrinsic behavior. Therefore, the neural networks can be classified according to their connections pattern, the number of hidden unities, the nature of activation functions and the learning algorithm [44-47]. There are an extensive number of ANN types and Figure 2 exemplifies the general classification of neural networks showing the most common ANN techniques employed in chemistry.

**Figure 2.** The most common neural networks employed in chemistry (adapted from Jain & Mao, 1996 [25]).

According to the previous brief explanation, ANN techniques can be classified based on some features. The next topics explain the most common types of ANN employed in chemical problems.
2.1. Multilayer perceptrons

Multilayer perceptrons (MLP) is one of the most employed ANN algorithms in chemistry. The term “multilayer” is used because this methodology is composed by several neurons arranged in different layers. Each connection between the input and hidden layers (or two hidden layers) is similar to a synapse (biological counterpart) and the input data is modified by a determined weight. Therefore, a three layer feed-forward network is composed by an input layer, two hidden layers and the output layer [38,48-50].

MLP is also called feed-forward neural networks because the data information flows only in the forward direction. In other words, the produced output of a layer is only used as input for the next layer. An important characteristic of feed-forward networks is the supervised learning [38,48-50].

The crucial task in the MLP methodology is the training step. The training or learning step is a search process for a set of weight values with the objective of reducing/minimizing the squared errors of prediction (experimental x estimated data). This phase is the slowest one and there is no guarantee of minimum global achievement. There are several learning algorithms for MLP such as conjugate gradient descent, quasi-Newton, Levenberg-Marquardt, etc., but the most employed one is the back-propagation algorithm. This algorithm uses the error values of the output layer (prediction) to adjust the weight of layer connections. Therefore, this algorithm provides a guarantee of minimum (local or global) convergence [38,48-50].

The main challenge of MLP is the choice of the most suitable architecture. The speed and the performance of the MLP learning are strongly affected by the number of layers and the number of hidden unities in each layer [38,48-50]. Figure 3 displays the influence of number of layers on the pattern recognition ability of neural network.

![Figure 3](http://dx.doi.org/10.5772/51275)

The increase in the number of layers in a MLP algorithm is proportional to the increase of complexity of the problem to be solved. The higher the number of hidden layers, the higher the complexity of the pattern recognition of the neural network.
2.2. Self-organizing map or Kohonen neural network

Self-organizing map (SOM), also called Kohonen neural network (KNN), is an unsupervised neural network designed to perform a non-linear mapping of a high-dimensionality data space transforming it in a low-dimensional space, usually a bidimensional space. The visualization of the output data is performed from the distance/proximity of neurons in the output 2D-layer. In other words, the SOM technique is employed to cluster and extrapolate the data set keeping the original topology. The SOM output neurons are only connected to its nearest neighbors. The neighborhood represents a similar pattern represented by an output neuron. In general, the neighborhood of an output neuron is defined as square or hexagonal and this means that each neuron has 4 or 6 nearest neighbors, respectively [51-53]. Figure 4 exemplifies the output layers of a SOM model using square and hexagonal neurons for a combinatorial design of purinergic receptor antagonists [54] and cannabinoid compounds [30], respectively.

Figure 4. Example of output layers of SOM models using square and hexagonal neurons for the combinatorial design of (a) purinergic receptor antagonists [54] and (b) cannabinoid compounds [30], respectively.

The SOM technique could be considered a competitive neural network due to its learning algorithm. The competitive learning means that only the neuron in the output layer is selected if its weight is the most similar to the input pattern than the other input neurons. Finally, the learning rate for the neighborhood is scaled down proportional to the distance of the winner output neuron [51-53].
2.3. Bayesian regularized artificial neural networks

Different from the usual back-propagation learning algorithm, the Bayesian method considers all possible values of weights of a neural network weighted by the probability of each set of weights. This kind of neural network is called Bayesian regularized artificial neural (BRANN) networks because the probability of distribution of each neural network, which provides the weights, can be determined by Bayes’s theorem [55]. Therefore, the Bayesian method can estimate the number of effective parameters to predict an output data, practically independent from the ANN architecture. As well as the MLP technique, the choice of the network architecture is a very important step for the learning of BRANN. A complete review of the BRANN technique can be found in other studies [56-59].

2.4. Other important neural networks

Adaptive resonance theory (ART) neural networks [60,61] constitute other mathematical models designed to describe the biological brain behavior. One of the most important characteristics of this technique is the capacity of knowledge without disturbing or destroying the stored knowledge. A simple variation of this technique, the ART-2a model, has a simple learning algorithm and it is practically inexpensive compared to other ART models [60-63]. The ART-2a method consists in constructing a weight matrix that describes the centroid nature of a predicted class [62,63]. In the literature, there are several chemical studies that employ the ART-based neural networks [64-73].

The neural network known as radial basis function (RBF) [74] typically has the input layer, a hidden layer with a RBF as the activation function and the output layer. This network was developed to treat irregular topographic contours of geographical data [75-76] but due to its capacity of solving complex problems (non-linear specially), the RBF networks have been successfully employed to chemical problems. There are several studies comparing the robustness of prediction (prediction coefficients, \( r^2 \), pattern recognition rates and errors) of RBF-based networks and other methods [77-80].

The Hopfield neural network [81-82] is a model that uses a binary \( n \times n \) matrix (presented as \( n \times n \) pixel image) as a weight matrix for \( n \) input signals. The activation function treats the activation signal only as 1 or -1. Besides, the algorithm treats black and white pixels as 0 and 1 binary digits, respectively, and there is a transformation of the matrix data to enlarge the interval from 0 – 1 to (-1) – (+1). The complete description of this technique can be found in reference [47]. In chemistry research, we can find some studies employing the Hopfield model to obtain molecular alignments [83], to calculate the intermolecular potential energy function from the second virial coefficient [84] and other purposes [85-86].

3. Applications

Following, we will present a brief description of some studies that apply ANN techniques as important tools to solve chemical problems.
3.1. Medicinal Chemistry and Pharmaceutical Research

The drug design research involves the use of several experimental and computational strategies with different purposes, such as biological affinity, pharmacokinetic and toxicological studies, as well as quantitative structure-activity relationship (QSAR) models [87-95]. Another important approach to design new potential drugs is virtual screening (VS), which can maximize the effectiveness of rational drug development employing computational assays to classify or filter a compound database as potent drug candidates [96-100]. Besides, various ANN methodologies have been largely applied to control the process of the pharmaceutical production [101-104].

Fanny et al. [105] constructed a SOM model to perform VS experiments and tested an external database of 160,000 compounds. The use of SOM methodology accelerated the similarity searches by using several pharmacophore descriptors. The best result indicated a map that retrieves 90% of relevant neighbors (output neurons) in the similarity search for virtual hits.

3.2. Theoretical and Computational Chemistry

In theoretical/computational chemistry, we can obtain some applications of ANN techniques such as the prediction of ionization potential [106], lipophilicity of chemicals [107, 108], chemical/physical/mechanical properties of polymer employing topological indices [109] and relative permittivity and oxygen diffusion of ceramic materials [110].

Stojković et al. [111] also constructed a quantitative structure-property relationship (QSPR) model to predict \( pK_{BH+} \) for 92 amines. To construct the regression model, the authors calculated some topological and quantum chemical descriptors. The counter-propagation neural network was employed as a modeling tool and the Kohonen self-organizing map was employed to graphically visualize the results. The authors could clearly explain how the input descriptors influenced the \( pK_{BH+} \) behavior, in special the presence of halogens atoms in the amines structure.

3.3. Analytical Chemistry

There are several studies in analytical chemistry employing ANN techniques with the aim to obtain multivariate calibration and analysis of spectroscopy data [112-117], as well as to model the HPLC retention behavior [118] and reaction kinetics [119].

Fatemi [120] constructed a QSPR model employing the ANN technique with back-propagation algorithm to predict the ozone tropospheric degradation rate constant of organic compounds. The data set was composed of 137 organic compounds divided into training, test and validation sets. The author also compared the ANN results with those obtained from the MLR method. The correlation coefficients obtained with ANN/MLR were 0.99/0.88, 0.96/0.86 and 0.96/0.74 for the training, test and validation sets, respectively. These results showed the best efficacy of the ANN methodology in this case.
3.4. Biochemistry

Neural networks have been largely employed in biochemistry and correlated research fields such as protein, DNA/RNA and molecular biology sciences [121-127].

Petritis et al. [128] employed a three layer neural network with back-propagation algorithm to predict the reverse-phase liquid chromatography retention time of peptides enzymatically digested from proteomes. In the training set, the authors used 7000 known peptides from D. radiodurans. The constructed ANN model was employed to predict a set with 5200 peptides from S. oneidensis. The used neural network generated some weights for the chromatographic retention time for each aminoacid in agreement to results obtained by other authors. The obtained ANN model could predict a peptide sequence containing 41 aminoacids with an error less than 0.03. Half of the test set was predicted with less than 3% of error and more than 95% of this set was predicted with an error around 10%. These results showed that the ANN methodology is a good tool to predict the peptide retention time from liquid chromatography.

Huang et al. [129] introduced a novel ANN approach combining aspects of QSAR and ANN and they called this approach of physics and chemistry-driven ANN (Phys-Chem ANN). This methodology has the parameters and coefficients clearly based on physicochemical insights. In this study, the authors employed the Phys-Chem ANN methodology to predict the stability of human lysozyme. The data set was composed by 50 types of mutated lysozymes (including the wild type) and the experimental property used in the modeling was the change in the unfolding Gibbs free energy (kJ/mol). This study resulted in significant coefficients of calibration and validation ($r^2=0.95$ and $q^2=0.92$, respectively). The proposed methodology provided good prediction of biological activity, as well as structural information and physical explanations to understand the stability of human lysozyme.

3.5. Food Research

ANNs have also been widely employed in food research. Some examples of application of ANNs in this area include vegetable oil studies [130-138], beers [139], wines [140], honeys [141-142] and water [143-144].

Bos et al. [145] employed several ANN techniques to predict the water percentage in cheese samples. The authors tested several different architecture of neurons (some functions were employed to simulate different learning behaviors) and analyzed the prediction errors to assess the ANN performance. The best result was obtained employing a radial basis function neural network.

Cimpoiu et al. [146] used the multi-layer perceptron with the back-propagation algorithm to model the antioxidant activity of some classes of tea such as black, express black and green teas. The authors obtained a correlation of 99.9% between experimental and predicted antioxidant activity. A classification of samples was also performed using an ANN technique with a radial basis layer followed by a competitive layer with a perfect match between real and predicted classes.
4. Conclusions

Artificial Neural Networks (ANNs) were originally developed to mimic the learning process of human brain and the knowledge storage functions. The basic unities of ANNs are called neurons and are designed to transform the input data as well as propagating the signal with the aim to perform a non-linear correlation between experimental and predicted data. As the human brain is not completely understood, there are several different architectures of artificial neural networks presenting different performances. The most common ANNs applied to chemistry are MLP, SOM, BRANN, ART, Hopfield and RBF neural networks. There are several studies in the literature that compare ANN approaches with other chemometric tools (e.g. MLR and PLS), and these studies have shown that ANNs have the best performance in many cases. Due to the robustness and efficacy of ANNs to solve complex problems, these methods have been widely employed in several research fields such as medicinal chemistry, pharmaceutical research, theoretical and computational chemistry, analytical chemistry, biochemistry, food research, etc. Therefore, ANN techniques can be considered valuable tools to understand the main mechanisms involved in chemical problems.

Notes

Techniques related to artificial neural networks (ANNs) have been increasingly used in chemical studies for data analysis in the last decades. Some areas of ANN applications involve pattern identification, modeling of relationships between structure and biological activity, classification of compound classes, identification of drug targets, prediction of several physicochemical properties and others. Actually, the main purpose of ANN techniques in chemical problems is to create models for complex input–output relationships based on learning from examples and, consequently, these models can be used in prediction studies. It is interesting to note that ANN methodologies have shown their power and robustness in the creation of useful models to help chemists in research projects in academy and industry. Nowadays, the evolution of computer science (software and hardware) has allowed the development of many computational methods used to understand and simulate the behavior of complex systems. In this way, the integration of technological and scientific innovation has helped the treatment of large databases of chemical compounds in order to identify possible patterns. However, people that can use computational techniques must be prepared to understand the limits of applicability of any computational method and to distinguish between those opportunities which are appropriate to apply ANN methodologies to solve chemical problems. The evolution of ANN theory has resulted in an increase in the number of successful applications. So, the main contribution of this book chapter will be briefly outline our view on the present scope and future advances of ANNs based on some applications from recent research projects with emphasis in the generation of predictive ANN models.
Author details

Vinicius Gonçalves Maltarollo¹, Káthia Maria Honório¹,² and Albérico Borges Ferreira da Silva³*

*Address all correspondence to: alberico@iqsc.usp.br

1 Centro de Ciências Naturais e Humanas – UFABC – Santo André – SP
2 Escola de Artes, Ciências e Humanidades – USP – São Paulo – SP
3 Departamento de Química e Física Molecular – Instituto de Química de São Carlos – USP – São Carlos – SP

References


http://dx.doi.org/10.5772/51275


[115] Smits, J. R. M., Schoenmakers, P., Stehmann, A., Sijstermans, F., & Chemom, Kate-
man G. (1993). Interpretation of infrared spectra with modular neural-network sys-

Pyrolysis Mass Spectra of Complex Binary and Tertiary Mixtures Using Multivariate


artificial neural network in chromatography-retention behavior prediction and pat-


355-363.

neural network of outer membrane [beta]-strand protein topology. *Protein Sci.*, 7,
2413-2420.


model for the prediction of membrane-spanning amino acid sequences. *Protein Sci.*, 3,
1597-1601.

[124] Dombi, G. W., & Lawrence, J. (1994). Analysis of protein transmembrane helical re-

membrane protein types based on pseudo-amino acid composition. *J. Theor. Biol.*, 242,
941-946.

work for predicting the stability of RNA/DNA hybrid duplexes. *Chemom. Intell. Lab.*, 70,
123-128.


Artificial Neural Networks for the Accurate Prediction of Peptide Liquid Chromatog-


