

Some Applications of Artificial Intelligence on Biotechnology

Rogério Antonio Strapasson¹, Adenise Lorenci Woiciechowski¹, Luiz Alberto Junior Letti¹, Carlos Ricardo Soccol¹

ABSTRACT

The present work is a revision about neural networks. Initially presents a little introduction to neural networks, fuzzy logic, a brief history, and the applications of Neural Networks on Biotechnology. The chosen sub-areas of the applications of Neural Networks on Biotechnology are, Solid-State Fermentation Optimization, DNA Sequencing, Molecular Sequencing Analysis, Quantitative Structure-Activity Relationship, Soft Sensing, Spectra Interpretation, Data Mining, each one use a special kind of neural network like feedforward, recurrent, siamese, art, among others. Applications of the Neural-Networks in spectra interpretation and Quantitative Structure-activity relationships, is a direct application to Chemistry and consequently also to Biochemistry and Biotechnology. Soft Sensing is a special example for applications on Biotechnology. It is a method to measure variables that normally can't be directly measure. Solid state fermentation was optimized and presenting, as result, a strong increasing of production efficiency.

Keywords: Artificial Neural Networks, Fuzzy Logic, Perceptron, Neural Networks Applications.

Algumas aplicações da Inteligência Artificial em Biotecnologia

RESUMO

O presente trabalho é uma revisão sobre redes neurais. Inicialmente apresenta uma breve introdução a redes neurais, lógica difusa, um breve histórico, e aplicações de Redes Neurais em Biotecnologia. As subáreas escolhidas para aplicação das redes neurais são, Otimização da Fermentação no Estado-Sólido, Sequenciamento de DNA, Análise Molecular Sequencial, Relação Quantitativa Estrutura-Atividade, Sensores inteligentes, Interpretação de espectros, Mineração de Dados, sendo que cada um usa um tipo especial de rede neural, tais como feed forward, recorrente, siamesa, art, entre outros. Aplicações de Redes Neurais em interpretação de espectros e Relação Quantitativa Estrutura-Atividade, como uma aplicação direta à química e consequentemente também para a Bioquímica e Biotecnologia. Os sensores Inteligentes são um exemplo especial de aplicação em Biotecnologia. É um método de medir variáveis que normalmente não podem ser medidas de forma direta. Fermentações no Estado-sólido foram otimizadas e, apresentaram como resultado um forte aumento do rendimento na produção final.

Palavras-chave: Redes Neurais Artificiais, Lógica Difusa, Perceptrons, Aplicações de Redes Neurais.

*Authors for correspondence

¹Bioprocess Engineering and Biotechnology Department, Federal University of Paraná, Curitiba, Brazil, strapasson@ufpr.br*; adenise@ufpr.br; matletti@ufpr.br; soccol@ufpr.br

INTRODUCTION

Artificial Neural Networks (ANN's) or simply Neural Networks (NN's) is an idea inspired in the operation of the human brain. Today it can be defined as a methodology capable to solve with success several relevant problems of the engineering and other areas (Sinvanandam 2006). The neural networks should be capable to accomplish three tasks, common also to the systems based on artificial intelligence (AI) (Haykin 2001):

- a) To Store knowledge;
- b) To apply the stored knowledge to solve problems;
- c) To acquire new knowledge for the experience.

It possesses the natural capacity to make parallel processing, what increases its numeric potentiality.

A neural network is formed by several processing units called artificial neurons. In a neuron model, we can identify four basic elements, as shown on Fig. 1:

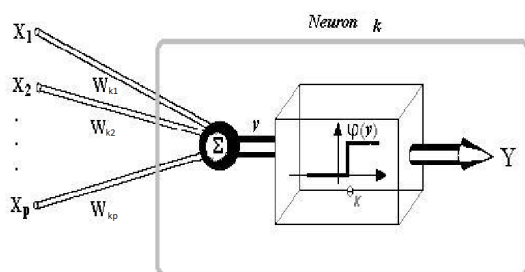


Fig 1. View of a node or neuron. (Almeida, 2002)

- 1) A group of connections called synapses, each one characterized by a synaptic weight k_j , where the index k corresponds to the number of the neuron and j to the entrance incentive. And W is the matrix of synaptic weights. Greater the weight W , higher will be the incentive of the entrance j . X_1 , X_2 and X_p are the input data.
- 2) A sum of entrance signs and their respective weights, resulting in the value γ .
- 3) An activation function (ϕ_γ), also called squashing function that seeks to limit the entrance signs to a certain interval, usually between 0 and 1 or (-1) and 1. The normalization is an important task to optimize

the training time, and the performance of the ANN (Nievola 2004). The activation function is capable to assume several forms, e. g. in step, linear or sigmoidal. The choice of the type of function activation to be used in the neural network will depend on the model type to be built (Strapasson 2003).

4) The bias (θ_k) unit provides a threshold for the activation of the neuron. A typical neuron, k , performs two calculations: a weighted linear combination of its inputs to obtain the activation (θ_k) and a nonlinear mapping of this activation value usually by means of a non-decreasing and differentiable activation function. Chen (1992) used the multilayer feed forward neural network for Fermentation Optimization in the classical paper with the title "Feedback Optimization of Fed-batch Bioreactors via Neural Net Learning" that is ruled by the forward gradient method. The optimal set of weights can be determined iteratively as it is done with the Standard Backpropagation Algorithm:

$$W_{ij}(p+i) = W_{ij}(p) + \eta \Delta W_{ij}(p) \text{ (Equation 1)}$$

$$\Delta W_{ij} = -\partial J(p) / \partial W_{ij}(p) \text{ (Equation 2)}$$

Where: p is the number of learning iterations; η is the Learning velocity tuning parameter; W is the elementary sum of the matrix elements i and j represent neuron location on the matrix. The neural-network result is a simple equation that can be used on a simple electronic calculator, as shown in Equation 3:

$$J(p) = \text{Max } X(t) \cdot V(t) \text{ (Equation 3)}$$

Where: $J(p)$, X , V , and t are the Production of Bioproduct, p (g), cell concentration, fermentation volume, and time in days of fermentation respectively.

In general, a FBNN (Feedforward Back Propagation Neural Network) has N neurons distributed in one input layer, one hidden layer, and one output layer. When an input sequence or vector, x_p , from a pattern p is presented and propagated through the network, a corresponding vector, y_p , is obtained from the output layer. Each connection between neurons has a weight W_{kp} associated with it.

A number of papers have shown that a feedforward network has the potential to approximate any non-linear function like on

medicine in prostate cancer prediction risk (Dombi et al. 2010) and in a neural network that is capable to recognize handwriting, converting handwritten documents into structural text form and recognizing handwritten names (Pradeep et al. 2011).

It has been shown that the two-layered feedforward network can uniformly approximate any continuous function to an arbitrary degree of exactness, providing that the hidden layer(s) contains a sufficient number of nodes.

Unlike the FBNN, Recurrent Neural Networks are more general in the sense that connections are allowed both ways between a pair of neurons and even from a neuron to itself. The RNN (Recurrent Neural Networks) allows the dynamics of the network to be considered, and it is thus useful for temporal association. However, the number of weights (for a fully recurrent network) to be determined may easily become quite large, and for this reason the Cascade Correlation type of network, where connections and nodes are added is required (Fahlmao 1990).

FUZZY LOGIC, ARTIFICIAL INTELLIGENCE AND NEURAL-NETWORKS

For many years, the subjects of artificial intelligence (AI), neural networks, and fuzzy logic were developed by separate intellectual communities. This was due mainly, perhaps, to social and institutional barriers to communication than to inherently different research goals and results. One manifestation of these barriers came in the form of claims, especially from the AI community, that the other approaches could not succeed at solving certain problems, despite progress to the contrary. This divisive period is fortunately substantially behind us. A growing number of models (now performed with computers) synthesize properties of expert production systems, neural networks, and fuzzy logic. Fuzzy ARTMAP is a family of self-organizing neural architectures that are capable of rapidly learning to recognize, test hypotheses about, and predict consequences of analog or binary input patterns occurring in a no stationary time series (Carpenter et al., 1993).

Fuzzy ARTMAP is the classes of Adaptive Resonance Theory (ART) architectures designed for supervise learning. Since the

introduction of ART as a cognitive and neural theory (Grossberg, 1976a, 1976b), an ever-expanding family of ART neural network architectures have been progressively developed, these models include ART1, ART2, ART2-A, ART3, Fuzzy ART, ARTMAP, and Fusion ARTMAP (Chen, 1996).

It can be noticed two main characteristics of the ART computers:

-These self-organized models can operate in an unsupervised, or a supervised, model.

-The technology of fuzzy control has been successfully applied in feed-water treatment in Japan, anaerobic digestion treatment in U.S.A., and Sequencing Batch Reactor (SBR) in Australia (JinLiang et al., 2009).

According to Chaibakhsh and Chaibakhsh (2010), in a biotechnological process that is studied the ANN utilization, it should be tested: Fuzzy Logic (FL), Neural Network (NN) and Genetic Algorithms (GA) and so on. The authors worked with an enzymatic esterification process based on experimental data. They are studying the lipase catalyzed synthesis of adipate ester.

The system described below is a "Multi Input/Single Output" type like can be seen in the Fuzzy System Models.

$$\mathbf{X} = [\mathbf{T}, \mathbf{t}, \mathbf{E}, \mathbf{a}]^{\text{Tr}}$$

$$\mathbf{Y} = [\mathbf{y}]^{\text{Tr}}$$

Where: X = Input; Y = Output; T = Temperature; t = Time; E = Enzyme amount; a = alcohol: acid molar ratio; Tr = Transpose; y = Yield.

Fuzzy Logic, from a more general viewpoint, introduces various classes of fuzzy system models. The two basic categories are (1) the Linguistic Models (LMs) of which the Mamdani type Fuzzy Logic Control (FLC) is a prototypical example, and (2) the Takagi - Sugeno - Kang (TSK) Fuzzy Models. The fuzzy models of the first group are essentially a qualitative description of the system behavior, normally cast in terms of natural language; those of the second kind combine linguistic rule antecedents with numerical consequent parts. Besides others, the notions of SISO (Single Input, Single Output) LMs, MISO (Multiple Input, Single Output) LMs, and MIMO (Multiple Input, Multiple Output) LMs

are introduced, and various modes of inference are discussed (Yager, 1994).

Local networks, such as the Radial Basis Function Network (RBFN), are particularly suitable for applications where on-line control and optimization is the main goal (Eikens 1994). In the RBFN, it is necessary to locate centers among the input/output pairs such that the sum of the squares of the distance from the center to the training data set is minimized. These centers are equivalent in concept to the weights in the FBNN. However, the RBFN may fail in predicting values if the prediction space does not contain any center. A global network will perform better in these circumstances (Karim, 1997).

Radial Basis Function (RBF) is a kind of ANN, like ART, that have a different architecture when compared to the classical ANNs, in this case, it is use the Radial Basis Function, that is, a real-valued function whose value depends only on the distance from the origin. The hidden units in Multi Layer Perceptron networks are dependent on the pondered sum of the inputs, transformed by monotonic activation functions. It is different for RBF where the activation functions are determined by a non linear function of the distance of the between the vector of entrance and a reference value. In this second case the hidden units have the activation function that present a radial base defined by its domain (Bishop, 1995).

HISTORY

McCulloch and Pitts (1943) described a simplified, but powerful model, of the real neuron, which still meets in use now in most of the models of Artificial Neural Networks. The model of McCulloch and Pitts (or MCP model) is based on the plausibility neurophysiologic of a neuron carrying out a sum function and limit, where the weights in the connections among the nodes correspond to the synapses inhibiting and activating the job of a real neuron. Although it is not explicit in the original work of McCulloch and Pitts, the nodes MCP are capable to adapt for the adjustment of the weights of each connection. That adaptation process - or learning - it allows the development of procedures for the gradual adjustment of the weights, of the math function that is being constructed.

The decade of 50 was of great success with the implementation of the first simulator of artificial neural network and of the first neuron-computer. Starting from the mathematical model of McCulloch and Pitts (1943) and using the theory of learning of Hebb (1949), and culminating with the creation of perceptron.

Rochester et al (1956) postulated and tested one Neural Network that was run in an IBM 704 computer, which show how the human brain works and that the way as the brain works could be emulated by computer. Their experiments were divided into three sets: The first set of experiments, designed to test parts of the theory advanced in *The Organization of Behavior*, by D. O. Hebb, which simulated a network of 69 neurons with a "Discrete Pulse Model". This set of experiments clearly illustrated the diffuse reverberation that is advanced as an explanation of short term memory. There was, however, no tendency for neurons to group into cell assemblies. The second set of experiments was designed to test P. M. Milner's revision of Hebb's theory with a "Frequency Modulate" Model (FM Model) which kept track of the frequency of firing of 512 neurons but ignored the precise timing of individual firings. All that the machine or the experimenter could know was the frequency at which a neuron was firing, and not the exact instants at which it did fire. The frequency would vary (oscillates) from time to time, so this was called the FM model. Cell assemblies formed and exhibited the "fractionation" and "recruiting" required by the theory. The cell assemblies, however, were not able to arouse one another, so this model was too heavily dominated by environment.

The analytical approach has not been very effective in actually describing the behavior of complicated networks of neurons. However, it has proved effective to simulate such networks and to draw conclusions from the behavior of the simulated network of neurons (Rochester 1956).

Perceptrons

The perceptron (Rosenblatt 1962) is the simplest form of a neural network used for the classification of the special type patterns characterized as linearly separable; a perceptron has only two layers: input and output layers. It computes a linear combination of the network inputs and applies the net input

to produce the output using a threshold output function. An elementary perceptron consists of a single output neuron with adjustable synaptic weights and a threshold. The threshold can be treated as a synaptic weight connected to a fixed input of value 1. Such a fixed input unit is called bias unit. Someone can use the elementary perceptron to solve a pattern classification problem with only two classes. To perform classification with more than two classes requires the use of more output neurons.

Multilayer Perceptrons

While simple perceptrons can perform classification only on linearly separable patterns, multilayer perceptrons are general-purpose flexible, non-linear models that, given enough hidden neurons and enough data, can approximate virtually any function to any desired degree of accuracy (White 1992).

USE OF RNs IN BIOTECHNOLOGY

The former Artificial Networks paper that contemplates biotechnology were writing by the 80s and so a great lapse of time is identified. Meistrell (1989) use an ANN to sequence proteins, as well they already know the structure of the protein, so they use it to test the ANNs. They also created the concept of ROC (Receiver Operating Characteristic) that was applied to neural networks. They combine $Y\alpha$ output with $Y\beta$ output. Both variables combined determine important functional groups of the protein, like the amino acid group and the absence of the amino acid residues.

An ROC curve can be constructed by varying the "cutoff value" of the output from a given class detector, and then plotting the true positive rate (customarily on the ordinate) and the corresponding false positive rate (abscissa) for that detector at each cutoff value of its output. The cutoff is the value above which the detector output is interpreted as representing a positive example, and below which the output is considered to represent a negative example (counter-example) of the particular class of interest (Meistrell, 1989).

The success of the protein prediction is denoted so that in 2010 it was just studying sequencing of proteins by neural networks (Mathkour, 2010). It was an integrated approach for protein structure prediction using

artificial neural network. And they concluded that protein prediction is a fundamental problem in bioinformatics, so they used an integrated approach which is very flexible and performs the prediction of protein in set of DNA genetic datasets with a large range success on its executions and results.

In the paper entitled Feedback Optimization of Fed-Batch Bioreactors via Neural Net Learning (Chen and Weigand, 1992); the authors utilized an ANN to optimize a reactor. Fed-batch bioreactor was used in this study.

Kennedy (1993) wrote a review of great importance about Neural Networks and Biotechnology. They accosted in that article three case studies: Controlling Fermentation Processes, Designing Fermentation Media, and Identifying Microorganisms. In 1996 Plamen (1996) propose a GA (General Algorithm) to NN which can be seen in the Appendix.

Karim et al (1997) studied the Global and Local Neural Network Models in Biotechnology. The authors conclude that intracellular components, however, are almost impossible to measure on-line. It was made so and indirect approach by ANNs. The neural network is a viable modeling paradigm that can be used for this purpose.

Leal (1997) used neural networks and "data fusion" to estimate biomass in liquid fermentation. The input parameters were: evolution of carbon dioxide, the dissolved oxygen tension, the fermentation time and pH. Data were normalized so that they were in the range from 0.1 to 0.9. The neural network architecture consisted of an input layer (one neuron for each variable) and an intermediate layer with 6 neurons. The output was the concentration of biomass. The training algorithm was the Levenberg-Marquardt. A data set to training the network was adopted and three others were used for validation. Two forms of data fusion have been used, and the results of average error were 5.30% and 4.58%. The authors also simulated noise on the input variables, and the mean value for the errors was still interesting (5.33% and 4.60%).

A neural network has also been used in hybrid modeling. L. Chen et al (2000) proposed a strategy to model fed-batch fermentation in a stirred tank reactor. First, the authors proposed the set of differential equations to describe biomass production, substrate consumption and product formation. Next using a state transformation that is allows the separation of

the stoichiometric coefficients and the kinetic reaction rates of the model. In the sequence, the stoichiometric coefficients with non linear regression and data fitting were estimated. Finally, neural networks were used to estimate the kinetic reaction rates. The predictions of the fitted model meet the experimental data.

Almeida (2002) wrote an interesting review about biotechnological applications of neural networks (mainly the feed forward propagation model), focusing on the most common mistakes made by some authors. The use of AI has a long tradition in areas such as Environmental Biotechnology, Biochemical Engineering and Food Sciences, but the author also calls attention to the increasing importance of Neural Networks in Cancer Research, Molecular Biology and Genomic and Proteomic Analysis. However, the underlying mechanisms of these tools are not always evaluated properly. Consequently, severe flaws are frequently noticed in some published papers. The main drawback noticed in neural networks software is the over fitting. Another common problem is the incorrect choice of the network topology. To avoid this problem, the author suggests a four point check list to develop neural network software: 1) Scaled Inputs (in order to normalize the entries); 2) Appropriate Learning Procedure (in some cases batch learning is better than backpropagation models); 3) Early Stopping (in order to avoid over fitting); 4) Topology optimization (to determine the optimal number of hidden nodes). He concludes the article with some advices to check the validity of a proposed model and with perspectives for the future.

Two of the main obstacles for a neural network model development are the dimension (number of hidden layers and number of hidden neurons) and the nature of the activation function. Meleiro et al (2003) used a constructive learning algorithm based on the work of Von Zuben and Netto (1995) in order to optimize the results. The model was based in a unit growing learning, which uses an iterative procedure for the incorporation of additional neurons in the hidden layer and parametric models using orthonormal set of basic functions to find the optimal activation functions. The authors related a study case which predicts the behaviour of a bioethanol plant constituted by four tanks reactors arranged in series with biomass recycling. The

objective is to determine a refined control strategy to maximize ethanol production. The predictions with the proposed method (unit growing learning algorithm) provided better results than the batch backpropagation model. The same accuracy was found with 4 neurons in the hidden layer for the former method and 10 neurons in the hidden layer for the last method.

Najafi analyses biofuels with neural networks. The experimental results revealed that using ethanol-gasoline blended fuels increased the power and torque output of the engine marginally. For ethanol blends it was found that the brake specific fuel consumption (BSFC) was decreased while the brake thermal efficiency ($\eta_{b.th.}$) and the volumetric efficiency (η_v) were increased. The concentration of CO and hydrocarbons emissions in the exhaust pipe were measured and found to be decreased when ethanol blends were introduced. This was due to the high oxygen percentage in the ethanol. In contrast, the concentration of CO₂ and NO_x was found to be increased when ethanol is introduced. An ANN model was developed to predict a correlation between brake power, torque, brake specific fuel consumption, brake thermal efficiency, volumetric efficiency and emission components using different gasoline-ethanol blends and speeds as inputs data. About 70% of the total experimental data were used for training purposes, while the 30% were used for testing. A standard back-propagation algorithm for the engine was used in this model. A multilayer perception network (MLP) was used for non-linear mapping between the input and the output parameters. It was observed that the ANN model can predict engine performance and exhaust emissions with correlation coefficient (R) in the range of 0.97–1. Mean relative errors (MRE) values were in the range of 0.46–5.57%, while root mean square errors (RMSE) were found to be very low. This study demonstrates that ANN approach can be used to accurately predict the engine performance and emissions (Najafi et al., 2009).

Solid-State Fermentation

Strapasson (2003) optimized and evaluated the production of fumaric acid using neural networks, as well checked the use and viability of neural networks to process simulation.

A feed-forward network of four neurons in the hidden layer was used to reproduce the

fermentation data obtained in laboratory. Optimization of Fumaric Acid production increases the ratio to 100g/Kg of fumaric acid and dry cassava bagasse.

DNA sequencing

In DNA sequencing ANNs competes with statistical methods approaches of the molecular mechanics, and usually obtain success in predicting the shape of proteins based on their chemical structure. This intensive computational method requires detailed representation of the molecule. And it uses complexes mathematical models.

The gene identification problem is tackled by two complementary approaches: gene search by signal and gene search by content. The search by signal methods identifies signal sequences, such as splice sites, which delimit coding regions. The searched by content methods use various codes measures to determine protein-coding potential of sequences (Wu, 1997).

Neural Networks provide an attractive model in which sequence features for both signals and content can be combined and weighted to improve predictive accuracy (Wu, 1997).

Molecular Sequencing Analysis

Like in the genome project other similar large sequencing projects are gaining a full understanding of its particularities, it is the DNA/RNA and protein Molecular Sequencing Analysis and has become a great challenge in computational molecular biology (Wu, 1997).

The neural network approach has been applied to solve a wide range of molecular sequence analysis problems. With the advancements in both network designs and sequence encoding methods, many neural network systems perform better than other approaches. Several systems have already become useful tools for the molecular biology community.

Neural networks are important tools, especially for the analysis of the large volume of molecular sequence data being generated by the human genome project. Neural decision boundaries between the different classes in a nonparametric fashion, and thereby offer a practical method for solving highly complex pattern classification problems (Wu, 1997).

QSAR

Quantitative structure-activity relationships pose a special problem for fed-forward

networks (FFNs). Unlike the subjects of others applications of FFNs in chemistry, molecules lack an obvious numerical model. "The lack of knowledge on" the physical basis of the phenomenon with which molecular properties are being correlated exacerbates the problem of representation. In many cases, researches have restricted their set of examples to one parent compound and its substituted analogues, allowing then to limit the description of the molecule to simple physical properties of the substituents. For a general representation, molecular graphs, the familiar line drawings used in organic chemistry, can be translated into a list of numbers. In practice, that translation usually includes simple chemical information. The complexity of this field makes FFNs the most flexible generalizing method available (Burns, 1993).

Spectra interpretation

It is a direct application to Chemistry and also to Biotechnology. To a computer, a spectrum is a sequential array of continuous-value data that approximates a function. It typically recognizes as features in a spectrum particular shapes or simple functions, often as simple as a sharp peak at a given position. These recognized patterns can have absolute positions (for example an IR absorbance of 4.35-4.55 μ at 2225 cm^{-1} indicates a nitrile group). Computers and programs do not have the necessity to convert unities to understand values. In a real way computers don't understand it just present the correct answer at the correct time.

In a mass spectra a branch peak at 28 atomic mass unit (amu), possibly indicating CO or ethylene. The features that characterize a particular molecule or substructure of interest often spread across the entire spectra lends itself to Fed-forward Networks methodology because of common non linear effects: Some regions distinguish compounds more effectively than others, and the significance of a feature may depend on the context.

Networks have been applied to spectra in four different ways: To recognize them, to quantify then, to classify then, and to transform them into a related function. In the first case, the output of network typically represents the presence of a known substance. In the second case the output measures concentration. In the third the network recognizes patterns associated with particular substructures or

molecular features and identifies these structures or features as output. In the fourth the network generates a function dependent on the entire spectrum (Burns, 1993).

Soft-sensing

Based on the ANN (Artificial Neural Network)-inversion soft-sensing method, the crucial biochemical variables which cannot be directly measured in the penicillin fermentation process were soft-sensed in the work of (Chen, 2010). During the reaction the biomass concentration, the substrate concentration, and the penicillin concentration are very hard to determine directly. But a soft-sensor can determine using the input values like feed flow rate of the substrate, the acid flow rate, and the base flow rate (Chen, 2010). Firstly, the model of the "Assumed Inherent Sensor (AIS)" of the penicillin fermentation process is constructed. The inputs of the AIS are the directly immeasurable variables to be estimated, while the outputs are the directly measurable variables. In this paper, the improved Birol model is used to construct the model of the AIS. Secondly, the soft-sensor is constructed, which is just the inversion of the AIS. The result reveals that only the biomass concentration X and the substrate concentration S can be soft-sensed, but the penicillin concentration P cannot be soft-sensed. Thirdly, the validity of the proposed ANN-inversion soft-sensor is verified by the simulation experiment. Finally, the soft-sensing problem of the penicillin concentration P is discussed. It can be seen that the proposed method is reasonable and reliable compared with other ANN soft-sensing methods.

Data Mining

Current research performed in Molecular Biology involves experiments that consume a great amount of resources and because of their complexity can produce ambiguous results which may require repeating or extensive analysis in order to obtain useful results. In addition, analyzing the DNA, RNA and protein sequence information concerning new experiments involves high demands of current researchers due to the massive volume and complexity of the information involved. Required sequence information is not always available for the organism being studied which means that information from other organisms will have to be correlated and included.

Data mining is defined as the process of discovering patterns in data. The process must be automatic or (more usually) semiautomatic. The patterns discovered must be meaningful in that way lead to some advantage, usually an economic advantage. The data is invariably present in substantial quantities (Witten, 2005).

CONCLUSION

Neural Networks is a versatile tool, in the sense it can be used to solve a wide variety of problems, and a specific tool, in the sense that once a problem has been defined, a certain network can be used to solve the desired problem or situation. In Biotechnology, it is possible to use them perform modeling and simulation of processes and thus, predict results, saving time and money.

The ability of acquire new knowledge implies in the fact that networks can treat with feasibility new problems like old problems, by only a task of training and testing requirements. They can simulate both direct and indirect variables. In the first case it can be cited the variables: fermentation ratio and days of fermentation. Indirect variables cannot be measured directly from the experiments; a very extensively used example is the rate of oxygen micro-adsorbed on the media that can be soft sensed with the help of a neural network.

The different types of neural networks may present better results or not depending on the application. For example recurrent neural networks allow the dynamics of the network to be considered, and it is useful for temporal association (dynamic problems). Each area of its application use a special kind of neural network like feedforward, recurrent, Siamese, art, among others.

In biotechnology the neural networks are used on many areas, a few examples are on biotransformations, to sequence DNA and proteins, design fermentation media, and identifying microorganisms.

Also in biotransformation some interesting applications are in solid-state fermentation, and in liquid fermentation, where are the particular fact that some variables are not directly measured, so it was make one neural network to bypass that problem by measure indirectly that conditions (named soft-sensors), and so utilize the data in a future network just to make the production simulation.

APENDIX

Plamen (1996) propose a GA (Generic Algorithm) to NN (Neural Networks) such as:

begin

$k := 0$

Set the initial population $\Pi(k)$ which constitutes of randomly generated.

Strings of fuzzy controls (agitation rate, RPM). For each RPM(0), RPM(1),..., RPM(N-1) in each population $\Pi(k)$.

Find the resulting concentrations of cell mass - X (k +1), substratum -S (k+1).

And enzyme activity, A (k+1) and evaluate each string in $\Pi(k)$.

while $k < N$ do:

begin

$k := k+1$

Assign the probabilities to each string in $\Pi(k - 1)$ which are proportional to the value of the evaluation function for each string and randomly (applying this probabilities) generate the new population $\Pi(k)$ perform crossovering and mutation on the strings in $\Pi(k)$ calculate the evaluation function for each string of $\Pi(k)$

end

end.

Nomenclature Section in Alphabetical Definitions Order:

$\varphi(\gamma)$ – Activation Function

AI – Artificial intelligence

AIS – Assumed Inherent Sensor

ANN – Artificial Neural Network

amu. – Atomic Mass Unit

ART – Adaptive Resonance Theory

A – Enzyme Activity

ART1 – Is the variety of ART Networks, accepting only binary inputs

ART2 – Extends network capabilities to support continues inputs

ART2-A – Is a Stream line from ART-2 with a drastically accelerated runtime, and with qualitative results being only rarely inferior to the full ART-2 implementation

ART3 – Builds on ART2 by simulating rudimentary neurotransmitter regulation of synaptic activity

ARTMAP – Predictive ART

X – Cell concentration

cm – Centimeter (10^{-2} m)

cm^{-1} – Number of wavelengths that that fit in a cm

BSFC – Brake Specific Fuel Consumption

DNA – Deoxyribonucleic acid

IBM 704 – First Mass-Produced Computer

FBNN – Feedforward Back Propagation Neural Network

V – Fermentation Volume

FL – Fuzzy Logic

FLC – Fuzzy Logic Control

FM – Frequency Modulate

Fuzzy ART – Implements fuzzy logic into ART's pattern recognition

Fuzzy ARTMAP – Is the classes of ART architectures designed for supervise learning

$\Pi(k)$ – Initial population

GA – Generic Algorithms

Hz - Hertz

IBM – International Business Machines

γ – Input Function

X_p – Input Sequence

θ_k – Input at neuron k

κ – Knot, node, neuron

p – Learning interaction number

η – Learning velocity

k (i,j) – neuron k and its coordinates

LM – Linguistic model

MCP – McCulloch & PITTS Model

$\mu - 10^{-6}$ m

$w_{(i,j)}$ – matrix element (i,j)

MLP – Multiple Layer Perceptron

MIMO – Multiple Impute / Multiple Output

MIS – Multiple Input / Single Output

MISO – Multiple Impute / Single Output

NN – Neural Network

γ – Output

γ_p – Output sequence

QSAR – Quantitative Structure-Activity Relationships

RBFN – Radial Basis function Networks

ROC – Receiver Operating Character

RNA – Ribonucleic Acid

RNN – Recurrent Neural Network

RPM – Revolutions per Minute 1/60 Hz

RMSE – Root Mean Square Errors

SBR – Sequencing Batch Reactor

w – Synaptic Weight

W – Sum of matrix w elements

SISO – Single Input / Single Output

SPK – Sum of each p interaction for the value at the k neuron

Y_α – Sum of component α output

Y_β – Sum of component β output

T – Transpose

TSK – Takagi-Sugen-Kang Fuzzy Model

t – Time (days)

η_v – Volumetric Efficiency

REFERENCES

- Almeida, J. S. Predictive Non-Linear Modeling Of Complex Data By Artificial Neural Networks. *Current Opinion in Biotechnology*. 13:72–76. Elsevier Science Ltd (2002).
- Angelov, P. Intelligent Optimal Control Of Biotechnological Processes. In: 4th European Congress on Intelligent Techniques and Soft Computing. Aachen, Germany (1996).
- Antsaklis, P. J. Neural Networks for Control Systems Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, IN 46556. IEEE Log Number 9034405 (1990).
- Bishop, C.M. Neural Network for Pattern Recognition. Oxford University Press. New York, U.S.A. (1995).
- Burns, J. A. and Whitesides, G. M. Feed-Forward Neural-Networks in Chemistry: Mathematical Systems for classification and Patern Recognition. *Chemical Reviews*. 8, 93 (1993).
- Carpenter, G. A. Grossberg, S. Learning, Categorization, Rule Formation, and Prediction by Fuzzy Neural Networks. Center of adaptive Systems and Department of Cognitive Neural Systems. Bostom: Boston University (1993).
- Chaibakhsh, A. Chaibakhsh, N. Fuzzy Modeling and Optimization of Biochemical Processes: A Case Study. International Conference on Chemistry and Chemical Engineering (ICCCE) (2010).
- Chastrette, M. El Aidi, C. Structure-Bell-Pepper Odor Relationship for Pyrazines and Pyridines Using Neural Networks. *Lab. De Chim. Org. Physiq. et Synthét* 83-96 (1996).
- Chen, C. H. Fuzzy Logic and Neural Network Handbook Ed. McGraw-Hill p 1.4; R006.3F997 (1996).
- Chen, Q. Weigand, W.A. Control Applications Maryland Univ., College Park, MD;, First IEEE Conference on ANN. Dept. of Chem. Eng., p 72 – 77; Dayton, OH (1992).
- Chen Weiliang, Zhang Kaifeng, Lu Chao, Dai Xianzhong, DING Yuhan, Beijing, China, P. R. China (2010).
- Dombi, G.W; Rosbolt, J.P; Severson, R. K; Neural network analysis of employment history as a risk factor for prostate cancer. *Comp. in Biol. and Med.* 40 751–757 (2010).
- Eikens, B. and Karim, M. N. Real-time control of a waste neutralization process using radial basis functions. *Proc. Adchem' 94 Conf.*, Kyoto, Japan, p. 125-130 (1994).
- F. J. Von Zuben and M. L. A. Netto. Unit-Growing Learning Optimizing The Solvability Condition For Model-Free Regression. *Proc. of the IEEE Intern. Conf. on Neural Networks*. Vol. 2:795-800 (1995).
- Fahlmao, S. E. and Lebeire, C.: The cascade correlation learning architecture. In Touretzky, D. S. (ed.), *Adv. in Neural Inform. Proc. Syst.* Morgan Kaufmann, Germany (1990).
- Haykin, S. *Redes Neurais princípios e práticas*, 2. ed. Bookman (2001).
- Hebb, D. O., *The Organization of Behavior*, New York: John Wiley and Sons Inc. (1949).
- JinLiang; FeiLuo; Ren-hui Yu; Yu-ge Xu; Waste water Effluent Prediction Based on Fuzzy-Rough Sets RBF Neural Networks College of Automation Science and engineering South China University of Technology at Guangzhou in China, Key Laboratory of Measurement and Control of CSE (School of Automation, Southeast University), Ministry of Education, Nanjing (2009).
- Lashkarbolooki, M. Hezave, A. Z. Artificial Neural Network As An Applicable Tool To Predict The Binary Heat Capacity Of Mixtures Containing Ionic Liquids. *Fluid Phase. Equil.* 324 p102– 107 (2012).
- M. Hassan, M. Ahmad, An integrated approach for protein structure prediction using artificial neural network. *Sec. Intern. Conf. on Computer Engineering and Applications* (2010).

- Mcculloch, W.S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*. n.5 p115-133 (1943).
- Meleiro, L. A; Von Zuben, F; Maciel Filho, R; Constructive Neural Network in Model-Based Control of a Biotechnological Process (2003).
- Michael, L. Meistrell, M.D. Evaluation of Neural Network Performance by Receiver Operating Characteristic Analysis: Examples from the Biotechnology Domain (1989).
- Montague G. and Morris J. Neural-Network Contributions in Biotec. Dept of Chemical process Engineering, University of Newcastle. U.K. (1980).
- Nievola, J. C. Redes Neurais Artificiais Editora da PUC Paraná (2004).
- Najafi, G; Ghobadian, B. and Tavakoli, T. and Buttsworth, D. R. and Yusaf, T. F. and Faizollahnejad, M; Performance and exhaust emissions of a gasoline engine with ethanol blended gasoline fuels using artificial neural network *Applied Energy*, 86, (5), 630–639. http://www.sciencedirect.com/science/ob=MIimg&_imagekey=B6V39-.pdf (2009).
- Pradeep, J; Srinivasan, E; Himavathi, S; Diagonal Based Feature Extraction for Handwritten Alphabets Recognition System using Neural Networks. *International Journ. of Comput. Scienc. & Informat. Technol. (IJCSIT)* 2011 Feb; vol 3 n. 1.
- Rochester, N., Holland, J.H., Haibt L.H. and Duda, W.L. (1956). Tests on a Cell Assembly Theory of the Action of the Brain Using a Large Digital Computer, *IRE Transaction of Information Theory* IT-2:80-93
- Ronald R.Yager and Dimitar P. Essentials of Fuzzy Modeling and Control. Filev New York, N Y: John Wiley (1994).
- Sinvanandam, S. N. Smathi, S. Deepa, S. N. Introduction to Neural Networks using MatLab 6.0 New Delhi Tata Mc Graw-Hill Publishing Company Limited (2006).
- Strapasson, R. A. Otimização das Variáveis do Bioprocesso de Fermentação no Estado Sólido para Produção de Ácido Fumárico a Partir de Resíduos Agroindustriais, Utilizando Redes Neurais. Curitiba: Univers. Fed. do Paraná (2003).
- Witten, I. H.; Frank, E. *Data Mining Practical Machine Learning Tools and Techniques*. Elsevier Inc. (2005).

Recebido: 10/10/2013
Received: 10/10/2013

Aprovado: 15/01/2014
Approved: 01/15/2014