Optimality Identification in Epidemiology Using Neural Networks

Samir Talssi
Hassan II University
samirtalssi@gmail.com

Noura Yousfi
Hassan II University
nourayousfi@hotmail.com

Abstract

An optimal control methodology of drug therapy can produce a drug dosing strategy to reducing the cost of treatment, making patient's healing faster or stabilizing his/her case. Some models of the controlled treatment of diseases such as HIV infection were proposed in different works.

In this paper, we use neural network with a single hidden layer to identify the optimality of drug treatment of HIV infection of CD4+ T-cells. The choice of single hidden layer was justified. A training set was generated by a numerical simulation model of treatment through a semi-implicit finite difference method. We used the back-propagation algorithm to perform and demonstrate the effectiveness of this neural approach. The result of this work is a neural network able to mimic and predict the interaction between HIV and the immune system under drug therapy.

Introduction

The human immunodeficiency virus (HIV) is a lentvirus that attacks the immune system and causes the acquired immunodeficiency syndrome (AIDS), a condition in humans in which the immune system begins to fail, leading to life-threatening opportunistic infections.

The CD4+T-cells are a subset of white blood cells that play an important role in the body's immune system. The CD4+T-cells are the key to HIV replication. Since HIV is a retrovirus, it needs cells from a "host" in order to replicate. In HIV cases, CD4+T-cells are the host cells that aid HIV replication. HIV attaches to the CD4+T-cells, allowing the virus to enter and infect the CD4+T-cells, damaging them in the process. The fewer functioning CD4+T-cells, the weaker the immune system and therefore the more vulnerable a person is to infections and illnesses [2].

Some antiretroviral drugs are available nowadays that help the immune system in preventing the infection due to HIV, even though it is not possible to cure it. Use of reverse transcriptase inhibitors is one of the chemotherapies that opposes the conversion of RNA of the virus to DNA (reverse transcription), so that the viral population will be minimal. On the other hand, the CD4+T-cell count remains higher and the host can survive. Another one is the protease inhibitors that prevent the production of viruses from the actively infected CD4+T-cells. This

kind of therapy has prolonged the life of infected individuals; however, the drug cost for this type of treatment is high [3] and side effects can be potentially severe. Therefore, the question is whether there exists a drug treatment schedule that can sustain a low viral load and a healthy immune system while minimizing the amount of drugs used. This suggests an optimal control approach to treatment scheduling.

Significant efforts have been evidenced in the research literature on modeling of physiological and immunological response of HIV in individuals. For excellent reviews of the various types of modeling attempts, see [4, 5]. Some models consider the dynamics of the CD4+T-cells and virus populations as well as the effects of drug therapy [6-8]. There are also some models that include an intracellular delay [9, 10].

Two optimal treatments of HIV infection model were proposed in [1]; in this work, the optimal controls represent the efficiency of drug treatment in inhibiting viral production and preventing new infections. We use a neural network with a single hidden layer to identify the optimality of these drug treatments.

Identification using neural network modeling, permits a nonlinear multivariable processes. The result of this identification seems a good compromise for systems that are difficult to model by conventional methods. A neural network has the ability to learn sophisticated nonlinear relationships, providing an ideal means of modeling complicated nonlinear systems [11-13].

This paper describes the controlled mathematical model of HIV infection with two control terms presented in [1], presents the neural network model for identification, and discusses the identification result.

Mathematical Model of HIV Infection with Two Control Terms

An optimal therapy was presented in [1] in order to minimize the cost of treatment, reduce the viral load, and improve immune response by using the model presented in [14] that incorporates the cure of infected cells. Two controls that measure the efficiency of reverse transcriptase and protease inhibitors were used, respectively, and two types of virus particles were included into the model to examine the effect of protease inhibitors. We use this model and recommend that the reader see those papers for more complete background [1].

Let (x), (y), (v_I) and (v_{NI}) denote, respectively, the concentration in mm^{-o} of uninfected CD4+ T-cells, infected cells, infectious virus and noninfectious virus. The model is given by the nonlinear system of differential equation 1, presented as follows:

$$\begin{cases} \dot{x} = s - dx - (1 - u_1(t))\beta v_I x + ry \\ \dot{y} = (1 - u_1(t))\beta v_I x - (a + r)y \\ \dot{v}_I = (1 - u_2(t))ky + \mu v_I \\ \dot{v}_{NI} = u_2(t)ky - \mu v_{NI} \end{cases} \quad (1)$$

s : Uninfected CD4+ T-cells production rate
d : Uninfected CD4+ T-cells death rate
β : Rate CD4+ T-cells become infected by virus
a : Death rate of infected CD4+ T-cells

Free virus production rate by infected cells at a rate k and cleared at a rate μ. Some infected cells may also revert to the uninfected state by loss of all cccDNA from their nucleus at a rate r.

$u_1(t)$ and $u_2(t)$ are the control functions, and $u_1(t)$ represents the efficiency of drug therapy in blocking new infection, so that infection rate in the presence of drug is $(1-u_1(t))\beta$. $u_2(t)$ represents the efficiency of drug therapy in inhibiting viral production, such that the virion production rate under therapy is $(1-u_2(t))k$.

The existence of an optimal control pair was demonstrated in [1], using a result by Fleming and Rishel [17] and Hattaf and Yousfi [18].

Neural Network for Identification

Several types of neural networks exist; the difference between them is their size, and, in the case of layered neural network architectures, the number of layers in a network, the number of nodes per layer, and the number of connections. A neural network can identify a non-linear correspondence $y = F(x)$ during a learning phase; it can learn how to associate correctly output patterns y to input patterns x. It was proven in [19, 20], based on classic mathematical result of Kolmogorov that for any continuous mapping $f: [0,1]^n \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$, they must exist in a three-layer neural network having an input layer with n processing elements, a hidden layer with (2n + 1) processing elements, and an output layer with m processing elements that implements f exactly. This result gave hope that neural networks would turn out to be able to approximate any function that arise in the real world. George Cybenko [21] proved that a single hidden layer feed-forward neural network can approximate any continuous and multivariate function. He also proved that a failure with such neural nets would be caused by bad weights, learning rate, etc. For more, see [22]. Therefore, a feed-forward neural network with one hidden layer (three layers) is implemented to identify the system (Figure 1). The number of processing elements per layer is fixed, using the Hecht-Nielsen theorem [23].

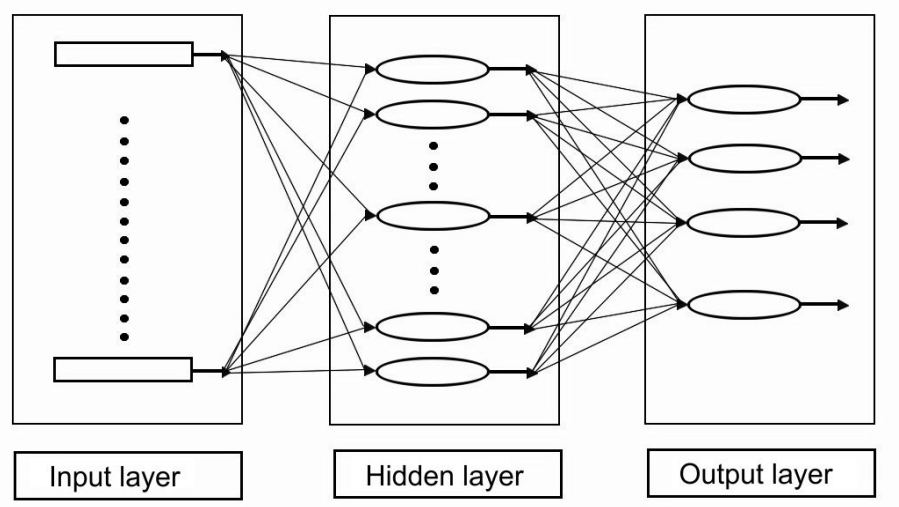

Figure 1. One hidden layer feedforward network architecture with sigmoidal activities

Neural Approach

We have to identify the solutions of controlled system (Figure 1) with variables x, y, v_I , v_{NI} . The semi-implicit finite difference method was used in [1] as a numerical algorithm and interesting results were found; we first reproduced the same results for constructing a learning basis network and then used a neural network for identification according to Figure 2.

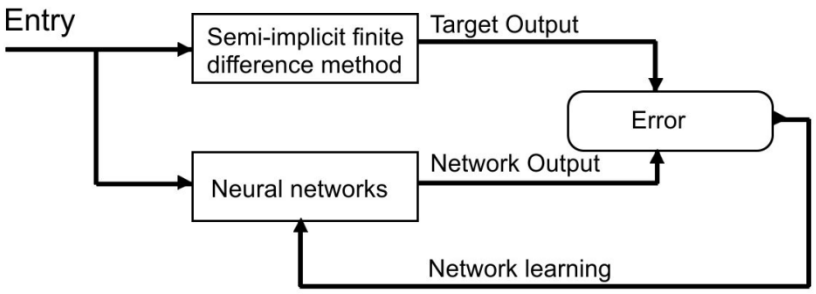

Figure 2. Identification process using neural networks

The following steps describe our method:

a. Generation and normalization of training data
b. Network generation
c. Network learning using back-propagation method and training data
d. Validation of results

Generation and Normalization of Training Data

The network learning using standardized data generates efficient results. Different methods can be used for data normalization in [24], and two effective methods were represented: linear transformation and statistical standards. We use the linear transformation with a slight modification. The principle of this method is simple: first, we look for the minimum and maximum data and then convert the data is converted using

$$x \leftarrow ((x - min)) / ((max - min)) \quad (2)$$

The following changes allow having values between $HI\ (= 0.9)$ and $LO\ (= 0.1)$

$$x \leftarrow ((x - min) / (max - \min)) * (HI - LO) + LO \quad (3)$$

Network Generation

Our neural network is a feed-forward network with one hidden layer of sigmoidal nonlinearities. We used the proven result in [23] to fix the number of neurons in the hidden layer and adjusted this number in simulation. The hidden layer is fully interlayer connected to both input and output layer; each neuron of the hidden layer is connected with all entries, and each network output is connected with all hidden neurons.

The neurons are characterized by their sigmoid activation function, their bias, and the weights of the connections with the entries. We randomly initialize the parameters of the network weights and biases to small values between 0 and 1.

Network Learning Using Back-Propagation Method and Training Data

The network's learning is a stage that poses difficulties most of the time. The back-propagation algorithm has been used in many cases to perform the network's learning [23].

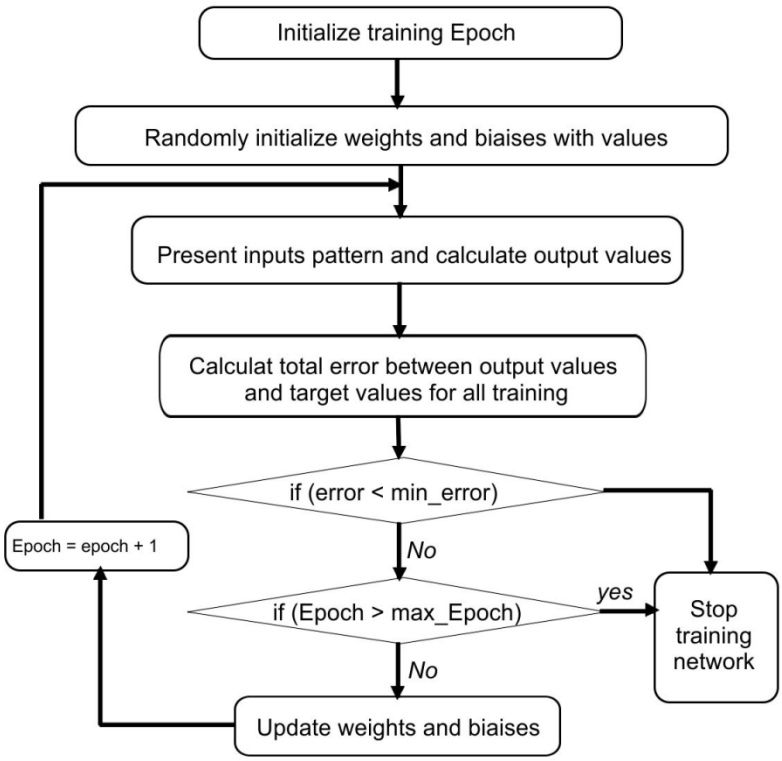

Figure 3. Training process

A particular case of training data is fed through the network in a forward direction, producing results at the output layer:

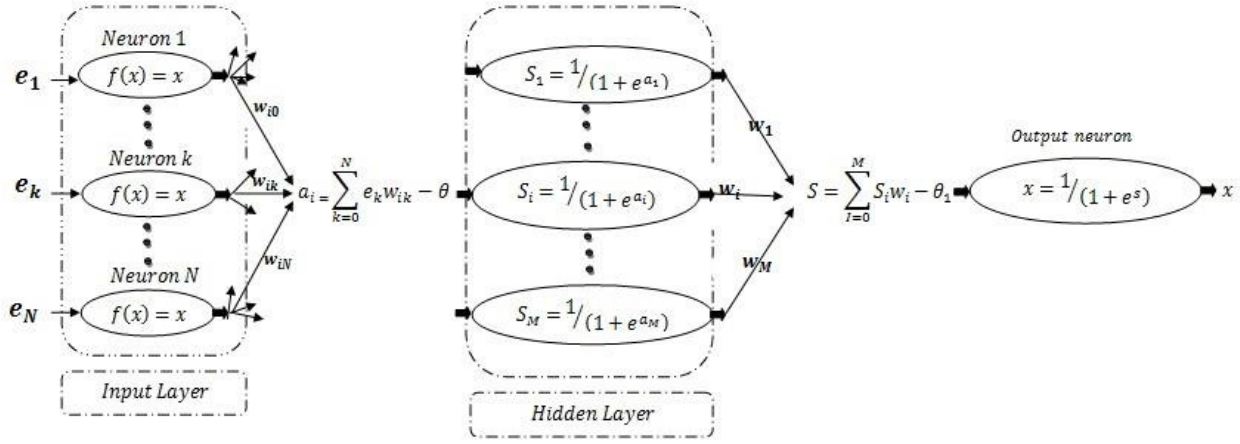

Figure 4. Propogation process in neural network

The modified back-propagation method presented in [25] is employed in our training process to ensure and accelerate the algorithm convergence. In this version of back-propagation, the

network weights are not updated after each pattern is presented. Rather, the weights are modified only after all training input patterns have been presented.

Changes in weight are calculated as follows:

$$\Delta w_{ij}(t+1) = \eta \sum_p (\delta_{pj} a_{pi}) + \alpha . \Delta w_{ij}(t) \quad (4)$$

Where

$$\delta_{pj} = a_{pj}(1 - a_{pj}) \sum_x \delta_x w_{px} \quad \text{if intermediate node} \quad (5)$$

$$\delta_{jp} = a_{jp}(1 - a_{jp})(t_{jp} - a_{jp}) \quad \text{if output node} \quad (6)$$

i is the "emitting" or "preceding" layer of nodes, j is the "receiving" or "subsequent" layer of nodes, ij is the layer of weights between node layers i and j, pj is the layer of weights between node layers p and j, weights are specified by w, node activations are specified by a, delta values for nodes are specified by δ, subscripts refer to particular layers of nodes (i, j, p) or weights (ij, jp), "sub-subscripts" refer to individual weights and nodes in their respective layers, learning rate η, and α is the momentum factor.

Note that t represents the iteration number rather than the presentation number, since the weights are updated only once per iteration through all training patterns.

The total error for all training patterns is calculated at any iteration. If the total error is reduced, then the learning rate η is multiplied by a factor $\varphi > 1$ for the next iteration. Else, if the error is more than a few percent above the previous value, all changes to the weights are rejected, and η is multiplied by a factor $\phi < 1$ and α is reduced to 0. For more, see [25].

Simulation Study

The therapy is described by four variables: x, y, v_I, v_{NI} denoting respectively the concentration in mm^{-3} of uninfected CD4+ T-Cells, infected cells, infectious virus, and noninfectious virus. We first generate and normalize a workforce of 2500 data (x, y, v_I, v_{NI}) using the semi-implicit finite difference method given in [1]. 2000 data are used for learning the network and 500 to test network performance training. And finally, we generate a new 1000 data using our network to validate and compare the identification results. The back-propagation algorithm is applied several times (in the best case, 25 times). At each time, we randomly reinitialize the network settings to give a new stimulus to the calculations.

Figures 5-8 show the evolution of the concentration in mm^3 under controlled therapy: the uninfected CD4+T-cells (Figure 5), infected cells (Figure 6), infectious virus (Figure 7), and noninfectious virus (Figure 8).

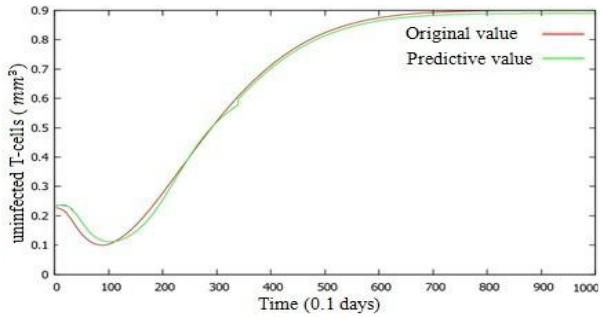

Figure 5. Uninfected CD4+T-cells

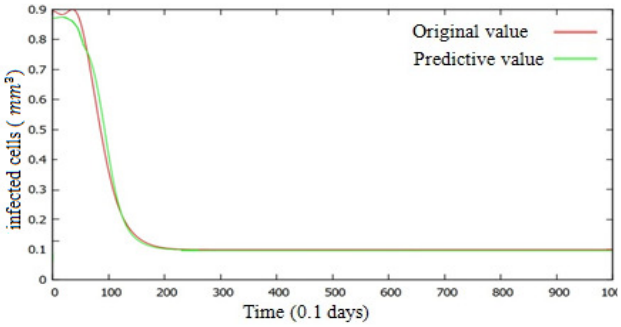

Figure 6. Infected cells

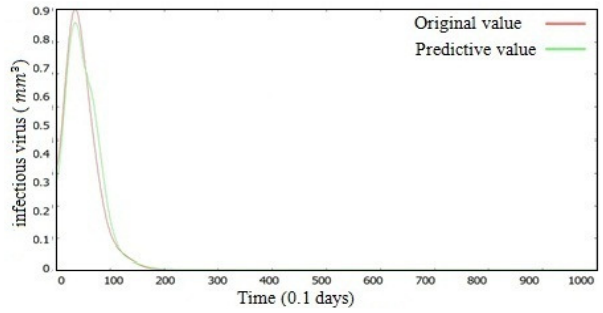

Figure 7. Infectious virus

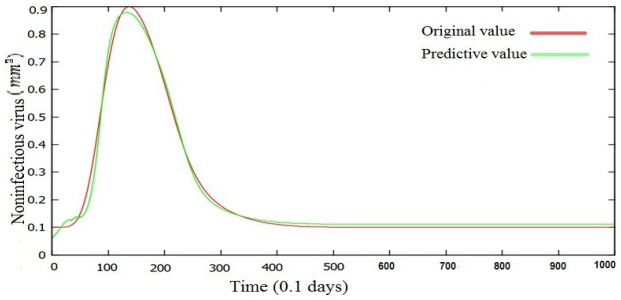

Figure 8. Noninfectious virus

The graphs present the predictive and the originals values of x, y, v_I and v_{NI} obtained with our feed-forward neural network. Graphically, it is clear that the predictive values have the same behavior as the original; the small perturbation noticed is generated at a random initialization of the network parameters.

Conclusion

The life expectancy of patients infected with HIV has increased dramatically with the advent of antiviral treatment. In this paper, we described the development of a new approach for the identification and prediction of optimal treatment of HIV drugs in numerical models of the environment using neural network techniques. Training data were generated using the data of [1] and were normalized to converge back propagation.

This approach provides us with an independent perspective mechanism of ideas specific to the therapy and allows rapid understanding of the behavior of HIV during treatment. This speed will help us to make good decisions with the aim of reducing the cost of treatment and avoiding critical phases of the disease. There is no effective treatment for HIV infection to cure this infection; existing treatments can only block the evolution of the virus in the body and maintain the balance between virus and defense system.

Among the identified risk factors, tobacco consumption is of paramount importance because it can be modified at the individual level. It turns out, however, that its control is difficult because of the existence of these psychosocial factors in persons living with HIV. This tobacco constraint will be added to the constraints of our approach in our next work, on the control treatment viral of HIV stress of smoking.

References

[1] Hattaf, K., & Yousfi, N. (2012). Two Optimal Treatments of HIV Infection Model. *World Journal of Modeling and Simulation*, *8*(1), 27-35.

[2] Bushman, F. D., Nabel, G. L., & Swanstrom, R. (2011). *HIV: From Biology to Prevention and Treatment*. New York: Cold Spring Harbor Laboratory.

[3] Gilks, G. F., Crowley, S., Ekpini, R., Gove, S., Perriens, J., Souteyrand, Y., & Sutherlan, D. (2006). *The WHO Public-Health Approach to Antiretroviral Treatment against HIV in Resource-Limited Settings*. Geneva, Switzerland: Department of HIV/AIDS, World Health Organization.

[4] Culshaw, R. V. (2004). Review of HIV Models: The Role of the Natural Immune Response and Implication for Treatment. *Journal of Biological Systems, 12*, 123-135.

[5] Callaway, D. S., Perelson, A. S. (2002). HIV-Infection and Low Steady State Viral Loads. *Bulletin of Mathematical Biology*, *64*, 29-64.

[6] Perelson, A. S., & Nelson, P. W. (1999): Mathematical Analysis of HIV-1 Dynamics in Vivo. *Society for Industrial and Applied Mathematics*, *41*(1), 3-44.

[7] Anderson, R. M., & May, R. M. (1988). Epidemiology Parameters of HIV Transmission. *Nature*, *333*, 514-519.

[8] Bonhoeffer, R., May, M., & Shaw, G. M. (1997). Network, Virus Dynamics and Drug Therapy. *Proceedings of the National Academy of Science, USA*, *94*, 6971-6976.

[9] Dumrongpokaphan, T., Lenbury, Y., Ouncharoen, R., & Xu, Y. (2007). An Intracellular Delay-Differential Equation Model of the HIV Infection and Immune Control. *Research India Publications*, *2*(1) 75-99.

[10] Nelson, P. W., Murray, J. D., & Perelson, A. S. (1999). A Model of HIV-1 Pathogenesis That Includes an Intracellular Delay. *Mathematical Biosciences*, *163*, 201-215.

[10] Chu, S. R., Shoureshi, R., & Tenorio, M. (1992). Neural Networks for System Identification. *Control Systems Magazine*, *10*(3), 31-35.

[11] Barron, A. (1991). Universal Approximation Bounds for Superposition of a Sigmoidal Function. *IEEE Tranactions on Information Theory,* IT-39, 930-945.

[12] Chen, S., Billings, S. A. & Grant, P. M. (1990) Nonlinear System Identification Using Neural Networks. *International Journal of Control*, *51*(6), 1191-1214.

[13] Hattaf, K., & Yousfi, N. (2011). Dynamics of HIV Infection Model with Therapy and Cure Rate. *International Journal of Tomography and Statistics*, *16*(11), 74-80.

[14] Schuitemaker, H., Koot, M., Kootstra, N. A., Dercksen, M. W., de Miedema, R. E., Tersmette Goede, M., van Steenwijk, R. P., Lange, J. M., & Schattenkerk, F. (1992). Biological Phenotype of Human Immunodeficiency Virus Type 1 Clones at Different Stages of Infection: Progression of Disease Is Associated with a Shift from Monocytotropic to T-Cell-Tropic. *Virus Populations*, *66*(3),1354.

[15] Pantaleo, G., Graziosi, C., Demarest, J. F., Butini, L., Montroni, M., Fox, C. H., Orenstein, J. M., Kotler, D. P., & Fauci, A. S. (1993). HIV Infection Is Active and Progressive in Lymphoid Tissue during the Clinically Latent Stage of Disease. *Nature*, *362*(6418), 355-8.
[16] Fleming, W., & Rishel, R. (1975). *Deterministic and Stochastic Optimal Control*. New York: Springer Verlag.
[17] Hattaf, K., & Yousfi, N. (2011). Dynamics of HIV Infection Model with Therapy and Cure Rate. *International Journal of Tomography and Statistics*, *16*(11), 74–80.
[18] Kurkova, V. (1992). Kolmogorov's Theorem and Multilayer Neural Networks. *Neural Networks*, *5(*3),501-506.
[19] Brattka, V. (2003). A Computable Kolmogorov Superposition Theorem. Computability and Complexity in Analysis. *Informatik Berichte*, *272*, 7-22.
[20] Cybenko, G. (1989). Approximation by Superpositions of a Sigmoidal Function. *Mathematics of Control, Signals and Systems*, *2*(4) 303-314.
[21] Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer Feedforward Networks Are Universal Approximators. *Journal of Neural Networks*, *2*,(5), 359-366.
[22] Hecht-Nielsen, R. (1989). Theory of the Backpropagation Neural Network. *The International Joint Conference on Neural Networks*. San Diego, CA.
[23] Shanker, M., Hu,, M. Y., & Hung, M. S. (1996). Effect of Data Standardization on Neural Network Training. *Omega, 24*(4), 385-397.
[24] Vogl, T. P., Mangis, J. K., Rigler, A. K., Zink, W. T., & Alkon, D. L. (1988). Accelerating the Convergence of the Back-Propagation Method. *Biological Cybernetics*, *59*, 257-263.
[25] Taghi Ameli, M., Shivaie, M., & Moslehpour, S. (2011). Transmission Network Expansion Planning Based on Hybridization Model of Probabilistic Neural Networks and Harmony Search Algorithm. *IAJC-ASEE International Conference.* Paper #ENG 108.
[26] Jin, Y., & Eydgahi, A. (2008). Monitoring of Distributed Pipeline Systems by Wireless Sensor Networks. *IAJC-IJME International Conference*. Paper #IT 304.

Biographies

SAMIR TALSSI is a professor at higher institutes of applied engineering in Casablanca. His interest area is artificial intelligence and its application in epidemiology. Talssi was a member the High Council Education in Morocco and has over 10 years of experience as manager and educator. Professor Talssi may be reached at samirtalssi@gmail.com.

NOURA YOUSFI is currently a professor in the Department of Mathematics and Computer Science in Mohammedia Hassan II University. She is director of several doctoral student research projects. Yousfi has over 20 years of experience as a scientist, manager, and teacher. She is a biomathematics specialist. Dr. Yousfi may be reached at nourayousfi@hotmail.com.