IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 12, No. 1, March 2023, pp. 155~161
ISSN: 2252-8938, DOI: 10.11591/ijai.v12.i1.pp155-161

Numerical study of the speed's response of the various intelligent models using the tansig, logsig and purelin activation functions in different layers of artificial neural network

Zineb Laabid[1], Aziz Moumen[1], Khalifa Mansouri[1], Ali Siadat[2]
[1]SSDIA laboratory, ENSET Mohammedia, Hassan II university of Casablanca, Mohammédia, Morocco
[2]LCFC, Arts et Métiers Paris Tech, Metz, France

Article Info

Article history:

Received Sep 30, 2021
Revised Jul 15, 2022
Accepted Aug 13, 2022

Keywords:

Activation function
Artificial neural network
Feed forward back propagation
Innovative materials
Intelligent model

ABSTRACT

Today's world is no longer that of yesterday, the pace with which we live and also the speed is enormous and rapid, that overnight we discover the appearance of new technologies and solutions in all the fields, in particular, that of scientific research. Artificial intelligence plays the main role. Predicting the behavior of new materials using artificial neural networks (ANNs) has become a frequently adopted solution by researchers today. The performance of neural networks depends mainly on the activation functions used. This work was designed to mainly study the impact of these functions on the response speed of an ANN in general, and particularly on the model we are working on to predict the thermomechanical behavior of innovative materials. By using tansig, purelin and logsig in a feed forward back propagation training by Levenberg-Marquardt algorithm, we were able to generate 9 models. For each of these models, we were interested in analyzing the speed's response of the network and studying its regression. Thus, this work was able to show us that choosing the right neuron activation function from one layer to another can clearly influence the performance of the results. Depending on the problem studied, the desired objective and the chosen architecture, the activation function can radically change the result and provide us with the expected efficiency.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Zineb Laabid
Laboratory of Signals, Distributed systems and Artificial Intelligence, Hassan II University of Casablanca
Higher Normal School of Technical Education, P.O. Box 159 28 820 Mohammedia, Morocco
Emaillaabid.zineb@gmail.com

1. INTRODUCTION

Materials science has always been one of the flagship strands of scientific research because the use of matter is found in all areas of human life. It was able to follow and take advantage of the solutions offered by technology through the different eras. At the beginning the goal of the researchers was to discover the elements and materials existing on earth, to study them, to analyze them and to optimize their applications. More recently, their objective is to go beyond the existing and innovate new non-existent materials [1], model them, study them and analyze them to use them for very specific applications; robotics, aviation, automotive, food and others. The combination of two materials whose properties we already know can give us a new one that is more efficient than the other two in terms of characteristics.

At the same time, artificial intelligence (AI) that was designed to mimic human intelligence and go beyond human ability, reasoning and logic represents the future in every way. AI can be used in classification, optimization also in prediction and this in different fields; in medicine for example to predict breast cancer [2], in the field of quality to predict the results of an audit control [3] or in materials science, to predict the properties of innovative materials, bio composites, in our case. AI has already shown its ability to give satisfactory results in predicting the elasticity of a biocomposite to a certain degree of percentage of the mixture and this using artificial neural networks [4], [5]. The performance of an artificial neuron network (ANN) is linked to several parameters, the architecture of the model chooses, number of neurons in each layer, the weight of the bonds, but particularly it is linked the activation function of neurons. An activation function is defined as a mathematical function allowing the passage of information from one neuron to another [6]. In the same neural network, we can find more than an activation function because the latter can be different from one layer to another.

This work aims to study the response of the neural network of choice using three types of activation functions popularly used in AI, namely tansig, logsig and purelin. By using a single architecture with these three activation functions, 9 models will be generated and studied to analyze them in order to allow us to converge towards the fastest and most precise result, and therefore lead us to the most optimum model to predict the behavior of innovative materials.

2. Dataset and model building

2.1. Dataset

AI relies on the results of the latest research using different analytical techniques to arrive at the desired results. So, to train an ANN you need to have a dataset with the data as input and the desired goal as output as information. The objective of training a neural network is to find the relationship between the input and the target provided in the dataset, so that when asked to predict a result, it will be able to provide us with the desired information with a negligible degree of error [7]. Our dataset represents the results from the use of a semi-analytical method called Mori Tanaka [8]. It is made up of two inputs representing the elasticity of the two components of the new material, and the output representing its elasticity.

2.2. Architecture and training function

In this step we have to choose the architecture of the neural network to build the 9 models to study. This process consists in defining the topology of the network, therefore the choice of the number of hidden layers and the number of neurons in each layer. Then choose the algorithm to use to train the network.

In our study, the choice fell on a network of three layers as shown in Figure 1. An input layer, an output layer and a hidden layer. The number of neurons in the input layer is identical to the number of input variables of our dataset which is 2 as well as the number of neurons in the output layer is 1 representing the target of our datasast. Regarding the hidden layer, our choice fell on 15 neurons. This choice was not arbitrary, it was after several trials and studies that we made this choice [9]. So, our network topology is 2-15-1. Since this is a regression problem, we opted for the feed forward back propagation trained with Levenberg-Marquardt (ML) algorithm since the latter was able to prove its efficiency by combining it with the feed forward back propagation [10].

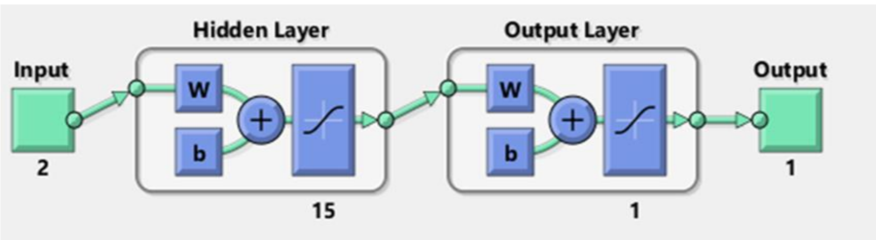

Figure 1. The network architecture to be studied

2.3. Activation functions

At the same time, the activation functions (AF) used in the hidden layer and in the output, layer is tansig, logsig and purelin [11]. These three functions allowed us to build 9 models to study. In every three models, we set the activation function of the hidden layer and we modify that of the output layer, and so on for the other two functions used. In Table 1 we find the representation of the models with the activation functions used in the two hidden layers and the output one.

Table 1. Different activation functions used on the proposel models

Model number	Activation functions hidden layer/output layer
1	tansig/logsig
2	tansig/purelin
3	tansig/tansig
4	logsig/tansig
5	logsig/tansig
6	logsig/purelin
7	purelin/tansig
8	purelin/logsig
9	purelin/purelin

– Tansig: hyperbolic tangent sigmoid transfer function: characterized by its wider interval which can ensure rapid learning. It is defined as,

$$f(x)=\text{Tanh}(x) = \left\{\frac{2}{1+e^{-2x}} - 1\right\} \tag{1}$$

– Logsig: log sigmoid transfer function. It is defined as,

$$f(x)=\text{logsig}(x) = \frac{1}{1+e^{-x}} \tag{2}$$

– Purelin: linear transfer function is typically used for function approximation and regression tasks. It is defined as,

$$f(x)=\text{Purelin}(x) = x \tag{3}$$

3. RESULT AND DISSCUTION

3.1. Intelligent model's number 1, 2 and 3

This is a 2-15-1 topology feed forward back propagation trained with a ML. In these three models we used and fixed tansig as the activation function of the hidden layer [12], and we changed that of the output layer in each model. Table 2 shows us the responses of the network studied in relation to these activation functions. The training of the latter has been done in such a way that at each training epoch we save the trained model by performing a check to avoid overfitting [13]. Among these three models, the most suitable combination giving a faster result is that of the third model, followed by the combination of tansig with tansig.

Table 2. The response time of the three models using the AF tansig in the hidden layer

Model number	AF hidden layer/ AF output layer	Average time (s)	Min time (s)	Max time (s)
1	tansig/logsig	4.32	0.82	7.83
2	tansig/purelin	1.03	0.60	1.47
3	tansig/tansig	0.91	0.52	1.30

A representation of the evolution of the regression coefficient over time for the three models in question clearly shows us that the fastest model is also the one which converges quickly to 1 [14]. The regression coefficient is one of the important indicators proving the performance of an artificial neuron network [15], the closer this coefficient is to 1, the more the network is said to be efficient. Figure 2 illustrates the result clearly. And from Figure 2, we can deduce that when we are looking for regression coefficients close to 1 exhibiting more efficiency, the Tansig/Tansig function quickly leads to this result.

3.2. Intelligent model's number 4, 5 and 6

This is a 2-15-1 topology feed forward back propagation trained with an ML. In these three models we used and fixed logsig as the activation function of the hidden layer [16], and we changed that of the output layer in each model. Table 3 shows us the responses of the network studied in relation to these activation functions. The training of the latter has been done in such a way that at each training epoch we save the trained model by performing a check to avoid overfitting. Among these three models, the most suitable combination giving a faster result is that of the sixth model.

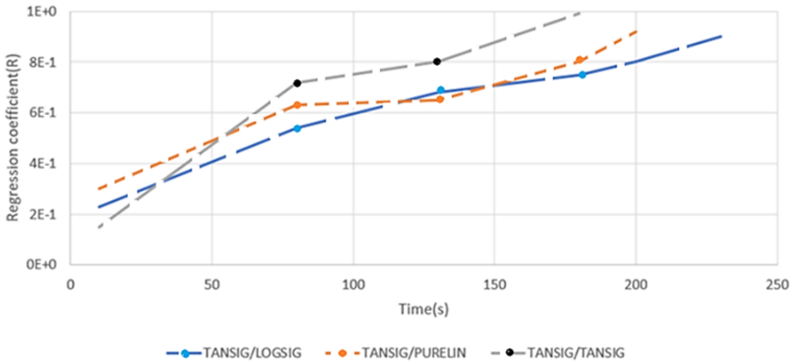

Figure 2. The convergence of the regression coefficient in function of time

Table 3. The response time of the three models using the AF logsig in the hidden layer

Model number	AF Hidden Layer/ AF Output layer	Average time (s)	Min time (s)	Max time (s)
4	logsig/tansig	12.81	3.12	22.5
5	logsig/logsig	19.78	6.9	32.66
6	logsig/purelin	7.24	4.36	10.13

Like wise for these three models using logsig in the hidden layer, it was necessary to follow the evolution of the regression coefficient of each model as a function of time. The representation of the latter in Figure 3 shows us that the three models do not converge. This is due to the use of the logarithmic function in the hidden layer. This activation function was able to provide us with results, but it showed us that its use in the hidden layer will not be able to provide us with effective results.

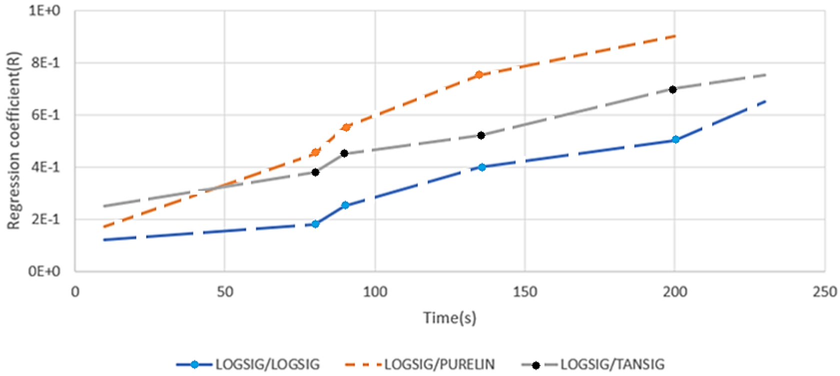

Figure 3. The convergence of the regression coefficient in function of time

3.3. Intelligent model's number 7, 8 and 9

This is a 2-15-1 topology feed forward back propagation trained with an ML. In these three models we used and fixed purelin [17] as the activation function of the hidden layer, and we changed that of the output layer in each model. Table 4 shows us the responses of the network studied in relation to these activation functions. The training of the latter has been done in such a way that at each training epoch we save the trained model by performing a check to avoid overfitting. Among these three models, the most suitable combination giving a faster result is that of the ninth model.

Table 4. The response time of the three models using the AF purelin in the hidden layer

Model number	AF hidden layer/ AF output layer	Average time (s)	Min time (s)	Max time (s)
7	purelin/tansig	0.82	0.52	1.12
8	purelin/logsig	1.9	1.7	2.1
9	purelin/purelin	0.49	0.13	0.85

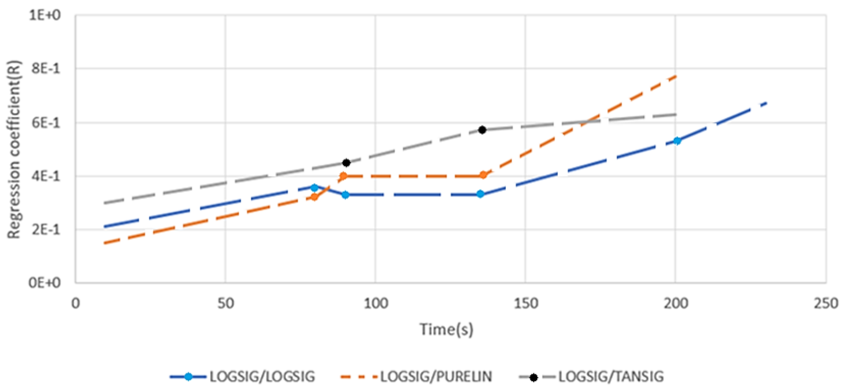

Figure 4. The convergence of the regression coefficient in function of time

3.4. Fastest model validation

In this part, we will proceed to the validation of the fastest model, and this by asking it to predict the behavior of a new material. Polypropylene reinforced with 15% biofiller [18]. The ANN was trained by testing and error checking in such a way as to see the minimum squared error along with the maximum regression value. Using 1,000 repeated epochs of training, the regression produced by our model is shown in the Figure 5. The regression coefficient of polypropylene [19] reinforced at 15% with a choosen bio load is 0.99 and that converges to 1 and signified that our model is efficient and performant [20]–[25].

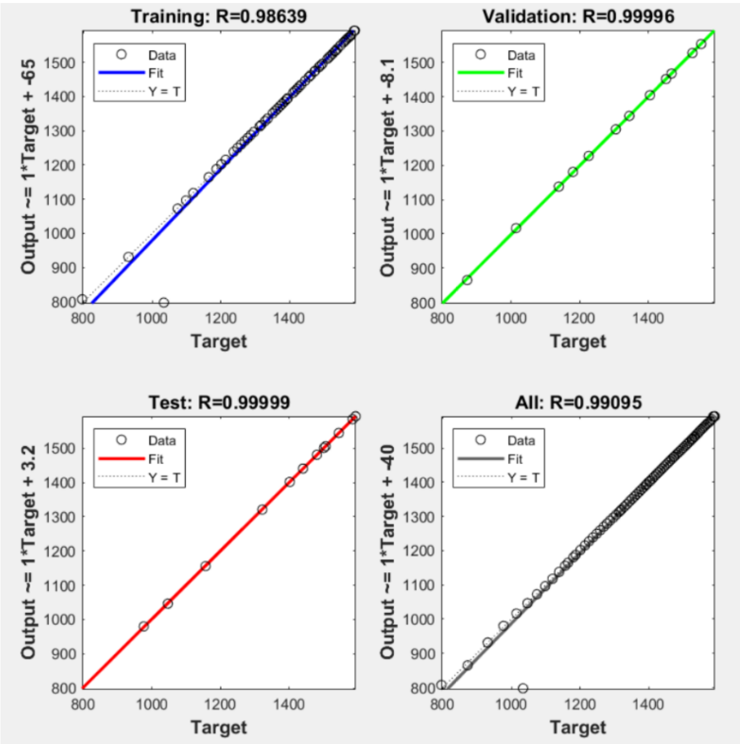

Figure 5. Performance validation by regression analysis

4. CONCLUSION

In this article we worked on an artificial neural network that meets the need to predict the properties of a new material, to study its response by playing on the activation functions of neurons. The activation function of a neuron is a very important parameter that influences the results. Choosing the right activation function clearly leads to effective results. The aim was to study the behavior of a feed forward back propagation which consists of three layers. Using tansig, purelin and logsig in activating neurons from both hidden and output layers, we were able to build 9 models that we studied. The speed of response as well as the regression of each of this model was our goal. We were able to show that in the prediction of the characteristics of biocomposites, the use of the tansig/tansig function allowed us to have a good regression coefficient close to 1 more quickly than using the other combinations of the functions of activation. The activation function remains a mechanism by which neurons process and transmit information through the different layers. For the tansig, the processing and the transmission are done in a wide interval which allows a fast learning to converge towards a better regression.

REFERENCES

[1] Z. Laabid, A. Moumen, A. Lakhdar, and K. Mansouri, "Towards the numerical implementation of neural network to predict the mechanical characteristics of bio composites," *IOP Conference Series: Materials Science and Engineering*, vol. 1126, no. 1, p. 012010, 2021, doi: 10.1088/1757-899x/1126/1/012010.

[2] T. A. Assegie, R. L. Tulasi, and N. K. Kumar, "Breast cancer prediction model with decision tree and adaptive boosting," *IAES International Journal of Artificial Intelligence*, vol. 10, no. 1, pp. 184–190, 2021, doi: 10.11591/ijai.v10.i1.pp184-190.

[3] S. Goto *et al.*, "New artificial intelligence prediction model using serial prothrombin time international normalized ratio measurements in atrial fibrillation patients on Vitamin K antagonists: GARFIELD-AF," *European Heart Journal - Cardiovascular Pharmacotherapy*, vol. 6, no. 5, pp. 301–309, 2020, doi: 10.1093/ehjcvp/pvz076.

[4] C. T. Chen and G. X. Gu, "Machine learning for composite materials," *MRS Communications*, vol. 9, no. 2, pp. 556–566, 2019, doi: 10.1557/mrc.2019.32.

[5] M. Markus, "Artificial neural networks," *Hydroinformatics: Data Integrative Approaches in Computation, Analysis, and Modeling*, pp. 409–435, 2005, doi: 10.1533/9780857099440.275.

[6] A. Apicella, F. Donnarumma, F. Isgrò, and R. Prevete, "A survey on modern trainable activation functions," *Neural Networks*, vol. 138, pp. 14–32, 2021, doi: 10.1016/j.neunet.2021.01.026.

[7] S. Ye, B. Li, Q. Li, H. P. Zhao, and X. Q. Feng, "Deep neural network method for predicting the mechanical properties of composites," *Applied Physics Letters*, vol. 115, no. 16, 2019, doi: 10.1063/1.5124529.

[8] A. Moumen, A. Lakhdar, M. Jammoukh, L. Zaheri, and K. Mansouri, "Optimization of the mechanical and morphological properties of polypropylene bio-loaded by argan nut shell particles with different theoretical and numerical models," *2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science, ICECOCS 2020*, 2020, doi: 10.1109/ICECOCS50124.2020.9314595.

[9] X. Li, Z. Liu, S. Cui, C. Luo, C. Li, and Z. Zhuang, "Predicting the effective mechanical property of heterogeneous materials by image based modeling and deep learning," *Computer Methods in Applied Mechanics and Engineering*, vol. 347, pp. 735–753, 2019, doi: 10.1016/j.cma.2019.01.005.

[10] S. Sapna, "Backpropagation learning algorithm based on levenberg marquardt algorithm," pp. 393–398, 2012, doi: 10.5121/csit.2012.2438.

[11] J. Feng and S. Lu, "Performance analysis of various activation functions in artificial neural networks," *Journal of Physics: Conference Series*, vol. 1237, no. 2, pp. 111–122, 2019, doi: 10.1088/1742-6596/1237/2/022030.

[12] J. Adnan, N. G. N. Daud, M. T. Ishak, Z. I. Rizman, and M. I. A. Rahman, "Tansig activation function (of MLP network) for cardiac abnormality detection," *AIP Conference Proceedings*, vol. 1930, 2018, doi: 10.1063/1.5022900.

[13] M. S. Santos, J. P. Soares, P. H. Abreu, H. Araujo, and J. Santos, "Cross-validation for imbalanced datasets: Avoiding overoptimistic and overfitting approaches [Research Frontier]," *IEEE Computational Intelligence Magazine*, vol. 13, no. 4, pp. 59–76, 2018, doi: 10.1109/MCI.2018.2866730.

[14] A. A. Nagaraja *et al.*, "Flux prediction using artificial neural network (ANN) for the upper part of glycolysis," *PLoS ONE*, vol. 14, no. 5, 2019, doi: 10.1371/journal.pone.0216178.

[15] D. W. Abueidda, M. Almasri, R. Ammourah, U. Ravaioli, I. M. Jasiuk, and N. A. Sobh, "Prediction and optimization of mechanical properties of composites using convolutional neural networks," *Composite Structures*, vol. 227, 2019, doi: 10.1016/j.compstruct.2019.111264.

[16] F. R. Hashim *et al.*, "Solar Location Estimation using Logsig Based Activation Function using Artificial Neural Network Approach," *Defence Science , Engineering & Technology*, vol. 4, no. 1, pp. 1–7, 2021.

[17] L. R. Chandran, N. Jayagopal, L. S. Lal, C. Narayanan, S. Deepak, and V. Harikrishnan, "Residential load time series forecasting using ANN and classical methods," *Proceedings of the 6th International Conference on Communication and Electronics Systems, ICCES 2021*, no. 1508–1515, pp. 1508–1515, 2021, doi: 10.1109/ICCES51350.2021.9488969.

[18] P. Sahu and M. K. Gupta, "A review on the properties of natural fibres and its bio-composites: effect of alkali treatment," *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications*, vol. 234, no. 1, pp. 198–217, 2020, doi: 10.1177/1464420719875163.

[19] P. N. Tri, A. Guinault, C. Sollogoub, P. N. Tri, A. Guinault, and C. S. Élaboration, "Élaboration et propriétés des composites polypropylène recyclé / fibres de bambou To cite this version : HAL Id : hal-01200819 Science Arts & Métiers (SAM)," 2015.

[20] P. Tubaro and A. A. Casilli, "Micro-work, artificial intelligence and the automotive industry," *Journal of Industrial and Business Economics*, vol. 46, no. 3, pp. 333–345, 2019, doi: 10.1007/s40812-019-00121-1.

[21] C. Krittanawong, H. J. Zhang, Z. Wang, M. Aydar, and T. Kitai, "Artificial intelligence in precision cardiovascular medicine," *Journal of the American College of Cardiology*, vol. 69, no. 21, pp. 2657–2664, 2017, doi: 10.1016/j.jacc.2017.03.571.

[22] R. Cerutti *et al.*, "NAD+-dependent activation of Sirt1 corrects the phenotype in a mouse model of mitochondrial disease," *Cell Metabolism*, vol. 19, no. 6, pp. 1042–1049, 2014, doi: 10.1016/j.cmet.2014.04.001.

[23] Z. Laabid, A. Moumen, A. Lakhdar, and K. Mansouri, "Toward the prediction of the elasticity of bio loaded polypropylene using artificial neural networks," pp. 49–57, 2022, doi: 10.1007/978-981-16-5559-3_5.

[24] J. Bilski, B. Kowalczyk, A. Marchlewska, and J. M. Zurada, "Local levenberg-marquardt algorithm for learning feedforwad neural networks," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 10, no. 4, pp. 299–316, 2020, doi: 10.2478/jaiscr-2020-0020.

[25] A. Moumen, Z. Laabid, A. Lakhdar, and K. Mansouri, "Prediction of the mechanical properties of Polypropylene reinforced with Snail Shell Powder with a Deep Neural Network Model and the Finite Element Method," *IOP Conference Series: Materials Science and Engineering*, vol. 1126, no. 1, p. 012009, 2021, doi: 10.1088/1757-899x/1126/1/012009.

BIOGRAPHIES OF AUTHORS

Zineb Laabid was born in Mohammedia in 1990. She is a PhD student at the University Hassan II Casablanca, ENSET Institute since 2020. Here research is focused on Contribution to the numerical characterization of bio-based polymers by intelligent models. Also, she has contributed to the experimental and numerical characterization of bio-based polymers subjected to aging. She can be contacted at email: laabid.zineb@gmail.com.

Aziz Moumen was born in Inezgane in 1987.He is a PhD student at the University Hassan II Casablanca, ENSET Institutesince 2018.He has the expertise in modeling the mechanical properties of the biomaterials classified as eco composites and obtained by natural bio loading. His focus is based on contributing to solve the demands of production companies in terms of continuous improvement of industrial materials. He can be contacted at email: moumenaziz1@gmail.com.

Khalifa Mansouri was born in 1968 at AZILAL, Morocco. He is now a teacher of computer science and researcher at the University Hassan II Casablanca, ENSET Institute. His research is focused on e-learning systems, industrial systems (modeling, optimization, and numerical computing). Diploma ENSET Mohammedia in 1991, CEA in 1992 and Ph.D (Calculation and optimization of structures) in 1994 to Mohammed V University in Rabat, HDR in 2010and Ph.D(Computer Science) in 2016 in Hassan II University of Casablanca. He can be contacted at email: khmansouri@hotmail.com.

Ali Siadat received his PhD in robotics and digital signal processing from Institut National Polytechnique de Lorraine in France. He started his carrier as associate professor in computer and industrial engineering at Arts et Metiers ParisTech and became full professor in 2016 at the same engineering school. Since 2011, he is head of applied mathematics and computer Engineering department. He served also as Guest Professor in several universities in China, Iran, Morocco and Brazil. His research interests include artificial intelligence, information system, knowledge formalization and operations research applied to manufacturing fields. He can be contacted at email: ali.siadat@ensam.eu.