HORA 2019
www.horacongress.com

International Congress on Human-Computer Interaction, Optimization and Robotic Applications

July 5-7, 2019, Urgup, Nevşehir, Turkey

SETSCI Conference Proceedings
www.set-science.com
8: 158-161, 2019

The Use of a Robust-Adaptive Self Organizing Map to Enhance the Prediction Performance of Clinical Datasets

Naim Ajlouni[1], Alaa Ali Hameed[2], Ali Güneş[3] Adem Özyavaş[4] and Zeynep Orman[5]

[1]*Department of Software Engineering, Istanbul Aydin University, Turkey*
Email: naimajlouni@aydin.edu.tr
[2]*Department of Computer Engineering, Istanbul University-Cerrahpaşa, Turkey*
Email: dr.alaa85@yahoo.com
[3]*Department of Computer Engineering, Istanbul Aydin University, Turkey*
Email: aligunes@aydin.edu.tr
[4]*Department of Computer Engineering, Istanbul Aydin University, Turkey*
Email: ademozyavas@aydin.edu.tr
[5]*Department of Computer Engineering, Istanbul University-Cerrahpaşa, Turkey*
Email: ormanz@istanbul.edu.tr

Abstract – Prediction in the medical field is a challenging problem and as a result many researchers have used different artificial intelligent methods including conventional Self Organizing Map (SOM) to achieve this task. SOM is a specialized clustering technique that has been used in a wide range of applications to solve different problems. Unfortunately, conventional SOM suffers from slow convergence and high steady-state error. The work presented in this paper is based on the recently proposed modified SOM technique introducing a Robust Adaptive learning approach to the SOM (RA-SOM). RA-SOM helps to overcome many of the current drawbacks of the conventional SOM and is able to outperform the SOM in obtaining the winner neuron in a lower learning process time. The efficient and outstanding performance achieved by applying RA-SOM in other research areas is the main driving force behind this work. To verify the improved performance of the RA-SOM, its performance is compared against the performance of other versions of the SOM algorithm, namely GF-SOM, PLSOM, and PLSOM2. The test results proved that the RA-SOM algorithm outperformed the conventional SOM and the other algorithms in terms of prediction time and accuracy. The test results also showed that RA-SOM maintained an efficient performance on the different datasets used, while the case of the other algorithms a more inconsistent performance was recorded, which means that their performance are data type-related.

Keywords – Clinical Data, Prediction, Performance, Quantization Error, Self organizing Map, Robust Adaptive SOM.

I. INTRODUCTION

Efficient and accurate prediction/Diagnosis presents a challenge in many medical applications, as a result, it plays a significant role in clinical data analysis and classifications. Recently data analysis and classifications received increasing attention through numerous applications [1]-[3] in the medical field. With the enormous growth of computational power, data analysis/classification has become more demanding in the field of automated medical prediction/diagnosis.

The Self-Organizing Map (SOM) is an unsupervised learning algorithm introduced by Kohonen [4]. The SOM map consists of a one or two-dimensional (2-D) grid of nodes. These nodes are also called neurons. Each neuron's weight vector has the same dimension as the input vector. The SOM obtains a statistical feature of the input data and is applied to a wide field of data classification [5]-[7]. Prior knowledge of the target output is not required in SOM for the recognition of the process. The algorithm works by finding input features similarities within data objects to define their relation by calculating the distance between them. The nodes output must map to the same weighed vector [8]-[9]. The winner output is defined as the node with the shortest distance between that node and the input vector. The weighted model continues to be updated to obtain the optimal cluster's topology [10]. The training time depends on dataset-size and the ability to find optimal weights within an acceptable time.

Many versions of modified SOM have been proposed to improve the vector quantization and the topology preservation performances, also detecting clusters by applying the different clustering algorithm to SOM [11]-[15].

Berglund & Sitte proposed Parameter-Less SOM (PLSOM) and Parameter-Less SOM2 (PLSOM2) to overcome limitations with Kohonen SOM. PLSOM uses a quadratic function for error fitting in place of the well-known neighborhood size and learning parameters. This method suffers from initial weight distribution overreliance and oversensitivity to outliers. PLSOM2 extended the work of PLSOM by scaling the weights according to input range observed instead of updating them based on the size of error relative to training maximum error [16]-[17].

DOI: 10.36287/setsci.4.5.031

In this paper, the RA-SOM algorithm is used for the prediction/diagnosis using a number of different clinical data types. The performance of the RA-SOM [18]-[19] is compared against well-known algorithms. All algorithms are tested under the same conditions. It is expected that the RA-SOM will outperform the other algorithms as it employs a decreasing adaptive learning rate function. This will cause the algorithm to achieve its desired convergence and accuracy in less run time. The test will be carried out on a number of datasets obtained from UCI and KEEL repository.

The remainder of the paper is organized as follows: Section II reviews the conventional SOM algorithm, Section III Reviews the RA-SOM algorithm, and Section IV presents the simulation results and performance comparison between the RA-SOM and other known algorithms including Kohonen SOM, PLSM, PLSOM2, and GF-SOM. The conclusions and future work are presented in Section V.

II. SOM ALGORITHM

The SOM architecture is composed of input and output layers, connected by link-associated weights. The SOM map uses neuron connections topologies of the hexagonal and rectangular form [1]-[2]. SOM output layers contain $n \times m$ neurons arranged as a two-dimensional grid. The original n-dimensional data are transferred to a two-dimensional map in SOM. In this case the input vector $x_i = \{x_1, x_2, \dots, x_n\}$, $i = 1, 2, \dots, n$, where i is the number of input and n is the input units of the vector. Each i is associated with the map through a weight vector $w = \{w_{n1}, w_{n2}, \dots, w_{nm}\}$.

SOM adopts a number of processes: First step, the $n \times m$ neuron weight vector is initialized randomly, the second step, an input vector x from the dataset is fed into the SOM network. Input vector x is fed to all neurons, at the same time. Third, the distance between the input and output neurons are calculated, then the closest neuron to the input identified (closest-distance) in this case using Euclidean Distance as the Best Matching Unit (BMU). This is denoted by c.

$$c = \arg\min_i (\|w_i(t) - x(t)\|). \quad (1)$$

This is iterated for entire input vectors. In each iteration, the weight vector is updated:

$$w_i(t+1) = w_i(t) + h_{c,i}(t).[x(t) - w_i(t)], \quad (2)$$

where $h_{c,i}$ is the Gaussian neighborhood function given as

$$h_{c,i}(t) = \alpha(t) . \exp\left(-\frac{\|r_c - r_i\|}{2\sigma^2(t)}\right), \quad (3)$$

In this case, $\alpha(t)$ is the learning rate, and $\|r_c - r_i\|$ is the Euclidean distance between the winning neuron c and the neuron i for each updated weight, and $\sigma(t)$ is the width of Gaussian. $\alpha(t) = \delta_\alpha . \alpha(t)$ and $\sigma(t) = \delta_\sigma . \sigma(t)$ which decreases gradually during the learning process by constants factors δ_α and δ_σ, respectively.

III. RA-SOM ALGORITHM

The conventional Kohonen SOM algorithm uses a fixed learning α which is usually between 0-1. The choice of the learning rate affects the speed of the conversion and accuracy of the optimum model. It is known that the higher the learning rate, the faster the convergence. This will not guarantee the accuracy of the data topology (clustering), as data accuracy requires a lower learning rate. Therefore, choosing a high learning rate will provide a high initial convergence, but once this is achieved the algorithm will be forced to diverge to a higher QE due to the inaccuracy in the data topology. On the other hand, choosing a small value for the learning rate will cause a slow convergence which will require many more iterations to achieve the required low QE. This will cause a delay problem and will not be acceptable in the case of big data.

Hence, the RA-SOM introduces an adaptive learning rate $\alpha(t)$ [18]-[19] which is of a decreasing form. It starts by introducing a high learning rate $\alpha(t)$. This is decreased adaptively in subsequent iterations. In this manner the adaptively decreasing learning rate achieves high convergence in the first few iterations, this is then followed by a lower learning rate $\alpha(t)$ to guarantee the continuous high convergence, and achieve the required accuracy of the data clusters.

The RA-SOM applied in this paper is of the form given by

$$w_i(t+1) = w_i(t) + \alpha(t).[x(t) - w_i(t)], \quad t = 0, 1, \dots \quad (4)$$

where $w_i(t+1)$ is defined as the updated weights, and $\alpha(t)$ is the variable adaptive learning convergence rate which is defined as

$$\alpha(t) = \frac{\lambda}{1 - \beta^t} \quad (5)$$

By substitute (5) in (4) the new format of RA-SOM is defined as

$$w_i(t+1) = w_i(t) + \left(\frac{\lambda}{1 - \beta^t}\right).[x(t) - w_i(t)]. \quad (6)$$

In the RA-SOM, the weight vector w is randomly initialized as a grid of $n \times m$ neurons similar to the conventional SOM algorithm. Then, updating the weights is controlled adaptively through the proposed learning algorithm. RA-SOM optimal weights are obtained in a much shorter time compared to the conventional SOM, PLSOM and PLSOM2 algorithms. Moreover, the optimum weight vectors are also improved and provide lower QE. RA-SOM adaptive learning function (6) starts with high β value while the value of t is small, hence the term ($1-\beta^t$) will be relatively small, therefore $\alpha(t)$ value will be relatively large, which will result in fast convergence of the updated weights in (7). As time t increases, the term ($1-\beta^t$) increases to a value close to unity, and hence $\alpha(t)$ will then be close to or equal to λ, which will result in low error performance in the updated weights of (7).

IV. SIMULATION RESULTS

The proposed algorithm has been tested using two different clinical datasets to assess its performance and ability to carry out prediction efficiently with high accuracy. In this paper, the algorithms were coded using MATLAB R2010b, and the tests were performed using a Core (TM) i7-3612QM CPU (2.10 GHz) PC equipped with 8,00 GB of RAM with Windows 7 Ultimate operating system. The datasets are collected from UCI and KEEL repository. Data were divided into 70% training and 30% testing sets. Datasets used in this test were normalized using Min-Max normalization between 0 and 1. The algorithms tested during the test are conventional Kohonen SOM, GF-SOM, PLSOM, and PLSOM2. The results of these algorithms were then compared against the performance of the RA-SOM under the same tests conditions.

A. Appendicitis dataset

The performance of the RA-SOM algorithm against other algorithms is evaluated using the appendicitis dataset using 2 classes, class label 0 and class label 1. The structure of the K-MAPs consists of 7 input neurons layer with a dimensional gride of 7×2 neurons in the competitive layers.

It can be seen from the results in tables 1 that the RA-SOM outperformed the other algorithms in relation to the minimization of the QE. From Fig. 1 it can also be seen that the lowest QE was obtained by the RA-SOM, followed by the PLSOM2 algorithm, noting that the other algorithms obtained a higher QE compared to the above-mentioned algorithms. However, Fig. 1 and Table 1 also show that the RA-SOM has obtained the lowest QE at 0.123 compared to the PLSOM2 and the other algorithms.

As shown in Fig. 1 the RA-SOM has maintained its high accuracy compared to all other algorithms for all subsequent run times. From the result and Fig. 1 it can be concluded that the RA-SOM obtained high accuracy and preserved the topology of the clustered dataset throughout the run time, with little change compared to other algorithms.

Table 1. QE results of the conventional SOM, GF-SOM, PLSOM, PLSOM2 and RA-SOM algorithms for appendicitis dataset

Algorithms	δ_α	δ_σ	**β**	λ	QE
Conventional SOM	0.57	-	-	-	0.1420
	0.56	-	-	-	0.1420
	0.55	-	-	-	**0.1416**
GF-SOM algorithm	0.7	0.003	-	-	0.1450
	0.6	0.002	-	-	**0.1420**
	0.5	0.001	-	-	0.1440
PLSOM algorithm	1.16	-	-	-	0.1730
	1.15	-	-	-	0.1720
	1.14	-	-	-	**0.1710**
PLSOM2 algorithm	0.9	-	-	-	0.1720
	0.8	-	-	-	**0.1270**
	0.7	-	-	-	0.1340
RA-SOM algorithm	-	-	0.994	0.00064	0.1230
	-	-	0.996	0.00063	0.1250
	-	-	0.994	0.00062	0.1320

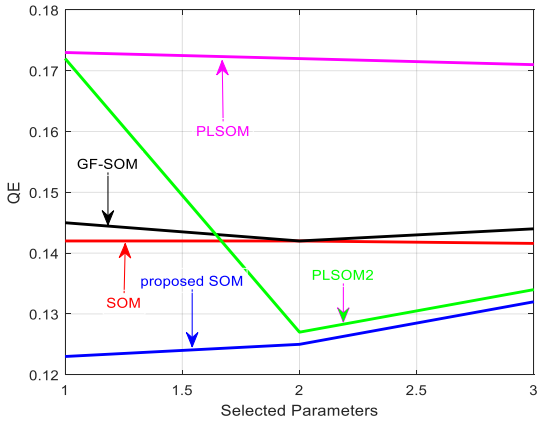

Fig. 1. Comparison of QE measures with various test parameters for appendicitis dataset

B. Wisconsin breast dataset

The dataset consists of 699 instances, 9 attributes, and 2 classes for benign and 4 malignant. In this test, the structure of Kohonen maps consists of 8 neurons in the input layer with a 2D grid of 8×2 neurons in the competitive layer. The basic parameters defining the performance of the algorithms with the relative performance are given in Tables 2, from the table of results it can be seen that the RA-SOM has reached the optimal QE and maintained an improved the QE all subsequent runs from 0.148 down to 0.1479, while the SOM and GF-SOM achieved QE of 0.148 and maintained the QE error to end of test. Both PLSOM2 and PLSOM didn't do very well in reducing the QE as it can be seen in Fig. 2.

From Table 2 it can also be seen that the following parameters $QE = 0.1479, \lambda = 0.0015$ and $\beta = 0.95$ giving lower QE, while the other algorithm showed that once they have reached the optimal $QE = 0.148$ further parameter change had no effect on the optimal QE value, in the case of the PLSOM, it improved relative to higher run times and variation of related parameters namely (algorithm optimal $QE = 0.18$, at $\beta = 0.6$), in the case of the PLSOM2 it can be seen that the algorithm couldn't maintain the dataset cluster topology and as a result the QE increased from $QE = 0.152$ at $\beta = 0.7$ up to $QE = 0.164$ at $\beta = 0.9$.

Table 2. QE results of the conventional SOM, GF-SOM, PLSOM, PLSOM2 and RA-SOM algorithms for wisconsin dataset

Algorithms	δ_α	δ_σ	**β**	λ	QE
Conventional SOM	0.5	-	-	-	0.1485
	0.4	-	-	-	0.1485
	0.3	-	-	-	**0.1484**
GF-SOM algorithm	0.3	1.7	-	-	**0.1484**
	0.2	1.6	-	-	**0.1484**
	0.1	1.5	-	-	**0.1484**
PLSOM algorithm	0.8	-	-	-	0.1880
	0.7	-	-	-	0.1830
	0.6	-	-	-	**0.1800**
PLSOM2 algorithm	0.9	-	-	-	0.1640
	0.8	-	-	-	0.1600
	0.7	-	-	-	**0.1520**
RA-SOM algorithm	-	-	0.95	0.0017	0.1480
	-	-	0.95	0.0016	0.1480
	-	-	0.95	0.0015	**0.1479**

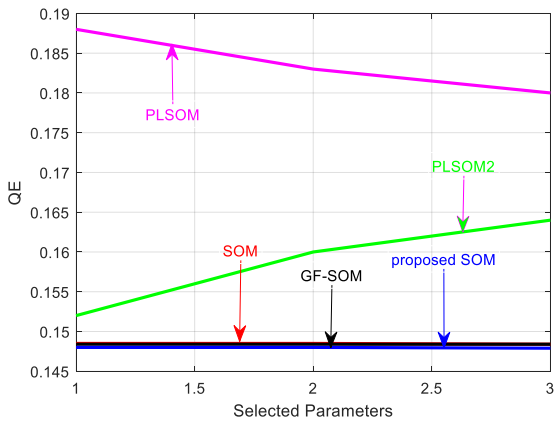

Fig. 2. Comparison of QE measures with various test parameters for wisconsin dataset

V. Conclusion and Future Work

The objective of this work is to implement an intelligent algorithm for the automation of a diagnosis system. In this work, several alternative algorithms including the proposed RA-SOM have been tested under the same conditions. Results show that the RA-SOM performed more efficiently than the other algorithms in the case of both datasets tested. It was also noticed that the RA-SOM not just outperformed the other algorithms, it also maintained the dataset variations. From the tests and results it can be concluded that the increase or reduction of the number of classes, instances and attributes had no effects on the abilities of the RASOM to efficiently converge the QE and to preserve the dataset topology. It is well known that selecting suitable learning parameters is the key to obtaining an optimum model with lower clustering topology error. This is one of the main drawbacks in model estimation and bound to be even a bigger issue in big data contexts, as selecting the optimum parameters one needs to run the program many times and each run may be extremely time-consuming. Therefore, RA-SOM offers more flexibility to obtaining different selection of parameters and thus obtain relevant optimum model quickly and more efficiently.

References

[1] T. Kohonen, "The self-organizing maps", *Proceedings of the IEEE* , vol. 78, issue 9, pp. 1464-1480, 1990.

[2] T. Kohonen, E. Oja, O. Simula, A. Visa and J. Kangas, 'Engineering applications of the self-organizing map", *Proceedings of the IEEE*, vol. 84, issue 9, pp.1358-1384 1996.

[3] S. L. Shieh and I. E. Liao, "A new approach for data clustering and visualization using self-organizing maps", Expert Systems with Applications, vol. 39, pp. 11924–11933 2012.

[4] T. Kohonen, Self-Organizing Maps, Springer, Berlin, 2001.

[5] J. Vesanto and E. Alhoniemi, "Clustering of the self-organizing map", IEEE Transactions on Neural Networks, vo. 11, issue 3, pp. 586–600, 2002.

[6] I. Lapidot, H. Guterman and A. Cohen, "Unsupervised speaker recognition based on competition between self-organizing maps", IEEE Transactions on Neural Networks, vol. 13, issue 4, pp. 877–887, 2002.

[7] T. Kohonen, Self-Organization and Associative Memory Process, Springer-Verlag, Berlin, 1989.

[8] M. Hoffmann, "Numerical control of Kohonen neural network for scattered data approximation", Numerical Algorithms, vol. 39, issue 1, pp. 175-186, 2005.

[9] F. Marini, J. Zupan and A. L. Magri, "Class-modeling using Kohonen artificial neural networks", Analytica Chimica Acta, vol. 544, issue 1, pp. 306-314, 2005.

[10] A. Astel, S. Tsakovski, P. Barbieri and V. Simeonov, "Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets", Water Research, vol. 41, issue 19, pp. 4566-4578, 2007.

[11] K. Appiah, A. Hunter, P. Dickinson and H. Meng, "Implementation and applications of tri-state self-organizing maps on FPGA", IEEE Transactions on Circuits and Systems for Video Technology, vol. 22, issue 8, pp. 1150–1160, 2012.

[12] D. Brugger, M. Bogdan and W. Rosenstiel, "Automatic cluster detection in Kohonen's SOM", IEEE Transactions on Neural Networks, vol. 19, issue 3, pp. 442-459, 2008.

[13] S. C. Chi and C. C. Yang, "A two-stage clustering method combining ant colony SOM and k-means", Journal of Information Science and Engineering, vol. 24, issue 5, pp. 1445-1460, 2008.

[14] M. Cottrell, P. Gaubert, C. Eloy, D. Francois, G. Hallaux, J. Lacaille and M. Verleysen, "Fault prediction in aircraft engines using self-organizing maps", International Workshop on Self-Organizing Maps, Springer Berlin Heidelberg, vol. 5629, pp. 37-44, 2009.

[15] K. Tasdemir, P. Milenov and B. Tapsall, "Topology-based hierarchical clustering of self-organizing maps", IEEE Transactions on Neural Networks, vol. 22, issue 3, pp. 474-485, 2011.

[16] E. Berglund and J. Sitte, "The parameterless self-organizing map algorithm", IEEE Transactions on Neural Networks, vol. 17, issue 2, pp. 305-316, 2006.

[17] E. Berglund, "Improved PLSOM algorithm", Applied Intelligence, vol. 32, issue 1, pp. 122-130, 2010.

[18] A. A. Hameed, Karlik, B., Salman, M. S., Eleyan, G. "Robust adaptive learning approach to self-organizing maps", Knowledge-Based Systems, vol. 171, pp. 25-36, 2019.

[19] A. A. Hameed, N. Ajlouni and B. Karlik, "Robust adaptive SOMs challenges in a varied datasets analytics", In International Workshop on Self-Organizing Maps, pp. 110-119, 2019.