Multidimensional Wavelet Neuron for Pattern Recognition Tasks in the Internet of Things Applications

Olena Vynokurova[1(✉)], Dmytro Peleshko[1], Semen Oskerko[1], Vitalii Lutsan[1], and Marta Peleshko[2]

[1] IT Step University, 83a Zamarstynivs'ka str., Lviv 79019, Ukraine
vynokurova@gmail.com, dpeleshko@gmail.com, semenosker@gmail.com, vl.brony@gmail.com

[2] Lviv State University of Life Safety, 35 Kleparivska, Lviv 79000, Ukraine
marta.peleshko@gmail.com

Abstract. Data mining and processing of Big Data is key problem in developing intelligent Internet of Things (IoT) applications. In this article, the multidimensional wavelet neuron for pattern recognition tasks is proposed. Also, the learning algorithm based on the quadratic error criterion is synthesized. This approach combines the benefits of the neural networks, the neuro-fuzzy system, and the wavelet functions approximation. The proposed multidimensional wavelet neuron can be used to solve a very large class of information processing problems for the Internet of Things applications when signals are fed in online mode from many sensors. The proposed approach is uncomplicated for computational realization and can be implemented in hardware for IoT systems. The proposed learning algorithm is characterized by a high rate of convergence and high approximation properties.

Keywords: Machine learning · Internet of Things
Multidimensional wavelet neuron · Pattern recognition · Classification
Online learning

1 Introduction

IoT produces and accumulates a lot of data of arbitrary natural, which are fed from Internet-connected sensory devices. Therefore, the development of IoT technologies requires new unique solutions for the accumulated data processing in real time, where the method of computational intelligence and machine learning have a lot of advantages as compared to conventional approaches [1–9].

The exponential growth of the stored and processed data caused a rapid transformation of the computational intelligence from a disparate set of heuristic methods into one of the most demanded applied disciplines, transforming all kinds of human activity. While using the intelligent systems becomes universal, there are growing demands for their universality, which means, in particular, stability on any type of data, adaptability to changing conditions, transparency of results interpretation [10–19].

Z. Hu et al. (Eds.): ICCSEEA 2018, AISC 754, pp. 64–73, 2019.
https://doi.org/10.1007/978-3-319-91008-6_7

As the use of intelligent systems becomes widespread, the requirements for their universality grow higher and higher, which means, in particular, a stability over any type of data, an adaptivity to changing conditions, a transparency of the results interpretation [12–23]. Strictly guarantee these properties we can only do by using rigorous mathematical methods based on the computational intelligence theory.

Nowadays, intensive researches are being carried out for the integration of IoT technologies and computational intelligence methods, among them: in [24] authors proposed 4 data mining models for processing IoT data, in [25] authors introduced a systematic manner for reviewing data mining knowledge and techniques in most common applications (classification, clustering, association analysis, time series analysis, and outline detection в IoT applications); in [26] authors ran a survey to respond to some of the challenges in preparing and processing data on the IoT through data mining techniques, in [27] authors attempted to explain the Smart City infrastructure in IoT and discussed the advanced communication to support added-value services for the administration of the city and citizens thereof.

As the analysis shows, in most cases, the existing methods are either not capable of processing the data stream in real time or cannot be implemented based on simple IoT controllers that could allow the development of the cheap IoT applications. Thus, it is important to create new high-speed methods that would have a simplicity of implementation and allow data processing in online mode.

In the paper, the architecture of multidimensional wavelet neuron and its learning algorithm are proposed. The proposed approach is characterized by simplicity of computational implementation and high speed of tuning parameters. Such systems can be used for solving tasks of the classification data, the patterns recognition, the prediction of multidimensional time series, which are generated by Internet-connected sensory devices in IoT applications.

2 The Architecture of Wavelet Neuron and Its Learning Algorithm

The architecture of the wavelet neuron [28] is shown in Fig. 1.

This architecture can be used for prediction or can be modified for binary classification of data in online mode, which are fed from the sensors of smart systems. The architecture of wavelet neuron is simple for implementation in hardware and can be used in IoT controllers and allows processing information in such systems directly.

As it can be seen, the wavelet neuron has similar architecture as neo-fuzzy neuron architecture [29, 30], but it consists of the wavelet synapses WS_i with the wavelet activation functions ($i = 1, 2, \ldots, n$), where the synaptic weights $w_{ji}(k)$ are the adjustable parameters.

When the input of the wavelet neuron gets vector signal

$$x(k) = (x_1(k), x_2(k), \ldots, x_n(k))^T,$$

its output produces a scalar in the form

$$\hat{y}(k) = \sum_{i=1}^{n} f_i(x(k)) = \sum_{i=1}^{n} \sum_{j=1}^{h_i} w_{ji}(k-1)\varphi_{ji}(x_i(k))$$

that is defined both by adjustable weights $w_{ji}(k)$, and by the wavelet functions being used.

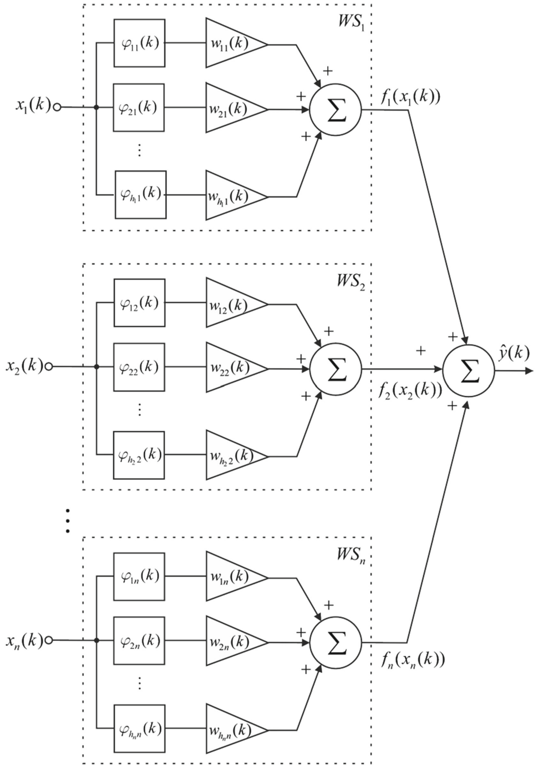

Fig. 1. Wavelet neuron

Here, we will use the one-dimensional wavelet activation function:

$$\varphi_{ji}(x_i(k)) = (1 - t_{ji}^2(k)) \exp\left(-\frac{t_{ji}^2(k)}{2}\right)$$

where $t_{ji}(k) = \left(x_i(k) - c_{ji}(k)\right)\sigma_{ji}^{-1}(k)$; $c_{ji}(k)$ is the parameter that defines a location of the function center; $\sigma_{ji}(k)$ is the parameter that defines the function width.

An additional point to emphasize here is that wavelet function oscillativity does not contradict to a membership function unipolarity since negative values can be treated as low membership levels or non-membership [31].

Learning task is to choose synaptic weights $w_{ji}(k)$ on each iteration k that will optimize the considered quality criterion.

As a learning criterion of the wavelet neuron, a quadratic error function is used that is expressed in wavelet neuron notation as follows:

$$E(k) = \frac{1}{2}(d(k) - \hat{y}(k))^2 = \frac{1}{2}e^2(k) = \frac{1}{2}(d(k) - \sum_{i=1}^{n}\sum_{j=1}^{h_i} w_{ji}(k-1)\varphi_{ji}(x_i(k)))^2 \tag{1}$$

(here $d(k)$ is the external learning signal), and its derivatives subject to adjustable parameters have the form

$$\frac{\partial E(k)}{\partial w_{ji}(k)} = -e(k)\varphi_{ji}(x_i(k)) = -e(k)(1 - t_{ji}^2(k))\exp\left(-\frac{t_{ji}^2(k)}{2}\right) = -e(k)J_{ji}^w(k).$$

Then, by applying gradient procedure to minimize the expression (1), we can express learning method of the wavelet neuron as follows:

$$w_{ji}(k+1) = w_{ji}(k) + \eta^w e(k+1)J_{ji}^w(k+1)$$

where scalar η^w defines the learning step in adjustable parameters space.

By introducing $(h_i \times 1)$ dimension vectors of variables $\varphi_i(x_i(k)) = (\varphi_{1i}(x_i(k)), \dots, \varphi_{h_i i}(x_i(k)))^T$, $w_i(k) = (w_{1i}(k), \dots, w_{h_i i}(k))^T$, $J_i^w(k) = (J_{1i}^w(k), \dots, J_{h_i i}^w(k))^T$, we can state gradient procedure of the i-th wavelet synapse learning method:

$$w_i(k+1) = w_i(k) + \eta^w e(k+1)J_i^w(k+1).$$

It is possible to improve a convergence of the learning processes by switching from gradient procedures to the Levenberg-Marquardt algorithm for adjusting of neural networks [32].

Using inverse of the sum of matrices lemma and performing obvious transformation [33], we can write the learning algorithms for real time case:

$$\begin{cases} w_i(k+1) = w_i(k) + \dfrac{e(k+1)J_i^w(k+1)}{r_i^w(k+1)}, \\ r_i^w(k+1) = \beta r_i^w(k) + \mathrm{P}J_i^w(k+1)\mathrm{P}^2 \end{cases} \tag{2}$$

where β is the forgetting parameter.

It is clear that given $\beta = 1$, expression (2) acquires stochastic approximation qualities of adaptive Goodwin-Ramadge-Caines algorithm [34], and given $\beta = 0$, it becomes Widrow-Hoff algorithm that is widely used in artificial neural networks theory. Thus, modified learning methods usage does not complicate numerical implementation of the wavelet synapses adjusting procedures virtually and increases the speed of their convergence.

3 The Architecture of Multidimensional Wavelet Neuron and Its Learning Algorithm

In many cases solving the real problems in Internet of Things application is needed the prediction or the classification of multidimensional data, which are fed from some sensors at one time. For this case, we can introduce multidimensional wavelet neuron, which has n inputs, m outputs and h wavelet activation function for each input. For the task of classification and pattern recognition, the sigmoidal functions have to be added to the output layer. The architecture of the multidimensional wavelet neuron for the classification or pattern recognition tasks is shown in Fig. 2.

The input observation vector is fed to the input layer of multidimensional wavelet neuron in the form

$$x(k) = (x_1(k), x_2(k), \dots, x_n(k))^T,$$

than the output of neuron can be written in the from

$$y_j(k) = \frac{1}{1 + \exp(-\gamma u_j)},$$
$$u_j(k) = \sum_{l=1}^{h} \sum_{i=1}^{n} \varphi_{li}(x_i(k)) w_{lij}(k)$$

where k is a number of observation, y_j is j-th output of multidimensional wavelet neuron, γ is a rate of rise of sigmoidal activation function, $\varphi_{li}(x_i(k))$ - l-th wavelet activation function of i-th input of neuron, $w_{lij}(k)$ is j-th synaptic weight of l-th wavelet activation function of i-th input of neuron, $i = 1 \dots n$, $j = 1 \dots m$, $l = 1 \dots h\cdot$

The nodes of the multidimensional wavelet neuron are wavelet neurons, which can be described above.

For the optimization of computational implementation let's rewrite the input of multidimensional wavelet neuron in the form

$$y(k) = W(k)\varphi(x(k))$$

where $\varphi(x(k)) = (\varphi_{11}(x_1), \varphi_{12}(x_2), \dots, \varphi_{1n}(x_n), \varphi_{21}(x_1), \varphi_{22}(x_2), \dots, \varphi_{2n}(x_n), \dots, \varphi_{h1}(x_1), \dots, \varphi_{hn}(x_n))^T$ is $(hn \times 1)$ dimension vector wavelet activation functions,

$$W(k) = \begin{pmatrix} w_{111} & w_{121} & \cdots & w_{1n1} & w_{211} & w_{221} & \cdots & w_{2n1} & \cdots & \cdots & w_{hn1} \\ w_{112} & w_{112} & \cdots & w_{1n2} & w_{212} & w_{222} & \cdots & w_{2n2} & \cdots & \cdots & w_{hn2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \ddots & \vdots \\ w_{11m} & w_{12m} & \cdots & w_{1nm} & w_{21m} & w_{22m} & \cdots & w_{2nm} & \cdots & \cdots & w_{hnm} \end{pmatrix}$$

is $(m \times hn)$ matrix of synaptic weights.

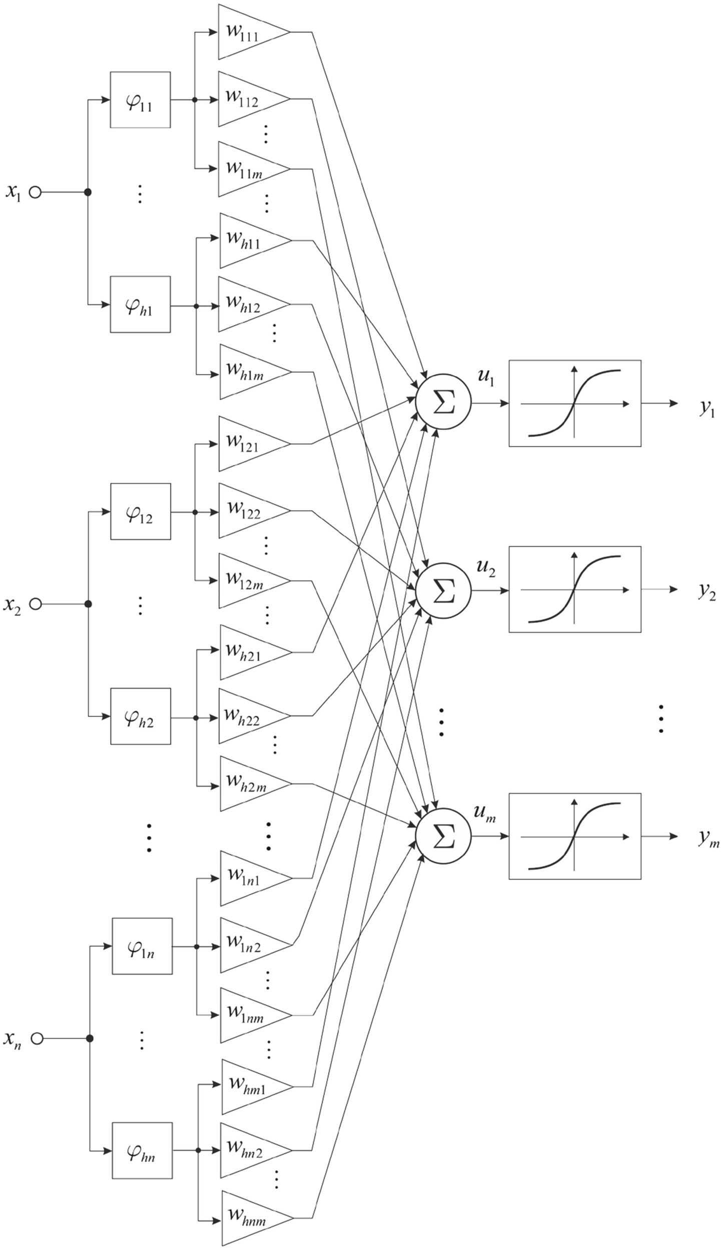

Fig. 2. The architecture of multidimensional wavelet neuron

Due to the synaptic weights of multidimensional wavelet neuron depend on the output systems linearly, we can use the stochastic approximation algorithms, which minimize criterion in the form

$$E_j(k) = \frac{1}{2}e_j^2(k) = \frac{1}{2}(d_j(k) - y_j(k))^2. \quad (3)$$

Minimizing the criterion (3) by synaptic weights $w_{lij}(k)$

$$\frac{\partial E_j}{\partial w_{lij}} = -e_j y_j (1 - y_j)\gamma \varphi_{li}(x_i),$$

we can write learning algorithm in the form

$$w_{lij}(k+1) = w_{lij}(k) - \eta \frac{\partial E_j}{\partial w_{lij}} = w_{lij}(k) + \eta e_j(k) y_j(k)(1 - y_j(k))\gamma \varphi_{li}(x_i(k)) \quad (4)$$

where η is learning rate parameter $(0 < \eta \leq 1)$.

For optimizing the learning process, we can rewrite learning algorithm (4) in matrix form

$$W(k+1) = W(k) + \eta\gamma(e(k)\mathrm{e}y(k)\mathrm{e}1 - y(k))) \cdot \varphi^T(x(k))$$

where $e(k) = (e_1(k), e_2(k), \ldots, e_m(k))^T$ is the errors vector, $y(k) = (y_1(k), y_2(k), \ldots, y_m(k))^T$ is the outputs vector, $\odot$ is dot product.

4 Experiments

The effectiveness of proposed recognition system is examined based on both benchmark data and real data. The recognition task of handwritten digits was solved based on MNIST database [35].

The MNIST is handwritten digits database, which was taken from [35]. A training set consists of the 60.000 images, and a test set consists of the 10.000 images. The initial bi-level images from NIST database were normalized (each image was resized to a 20 $\times$ 20 pixel box. After that, the all images were centered in a 28 $\times$ 28 pixel image using the pixels' mass center.

The multidimensional wavelet neuron has 784 inputs, 10 outputs, and 10 wavelet function for each input. The initial synaptic weights values were taken zeros and learning rate parameter was taken $\eta = 0.99\cdot$

The classifiers based on neural networks were taken for the results comparison. Table 1 shows the results of the comparison. As the quality criterion was taken the percentage of the false classified objects based on the testing data image set.

As it may be inferred from the obtained results, the multidimensional wavelet neuron has the best quality of classification among 1-layer neural network classifiers. The 2-layer perceptron has a better quality of classification but has 3.5 times more adjustable parameters than the multidimensional wavelet neuron. This fact influences on the training time of such system and as result – the impossibility of using such system in IoT controllers. The proposed multidimensional wavelet neuron has simple architecture and can be implemented based on Arduino-like controllers for many pattern recognition tasks in IoT application.

Table 1. The classification results of MNIST dataset

Neural network classifier	Preprocessing data	Test error (%)
Multidimensional wavelet neuron (1-layer NN, 10 wavelet activation function for each input)	none	6.5%
Linear 1-layer NN classifier	none	12.0%
Linear 1-layer NN classifier	deskewing	8.4%
Pairwise linear classifier	deskewing	7.6%
2-layer perceptron (300 sigmoidal activation functions in the hidden layer and 10 ones in the output layer)	none	4.7%

5 Conclusion

In this article, the architecture of multidimensional wavelet neuron is proposed. Such system can be used both a classifier and predictor of the multidimensional data sets. The main advantage of the multidimensional wavelet neuron is a simplicity of implementation in the hardware of IoT applications. Also, the learning algorithm of multidimensional wavelet neuron for solving pattern recognition task is proposed. This learning algorithm is the modification of the gradient algorithm and is characterized by high speed of information processing.

The systems based on the proposed multidimensional wavelet neuron can be implemented for solving the problems in IoT applications, Data Stream Mining, Big Data Processing.

The computational experiments are performed based on benchmark and real data sets. The obtained results have confirmed the advantages of the proposed approach in comparison with the existed methods.

References

1. Mahdavinejad, M.S., Rezvan, M., Barekatain, M., Adibi, P., Barnaghi, P., Sheth, A.P.: Machine learning for Internet of Things data analysis: a survey. Dig. Commun. Netw. (in Press, 2017)
2. Atzori, L., Iera, A., Morabito, G.: The Internet of Things: a survey. Comput. Netw. **54**(15), 2787–2805 (2010)
3. Cecchinel, C., Jimenez, M., Mosser, S., Riveill, M.: An architecture to support the collection of big data in the Internet of Things. In: Proceedings of 2014 IEEE World Congress on Services, pp. 442–449 (2014)
4. Joakar, A.: A methodology for solving problems with data science for Internet of Things. Data Science for Internet of Things. https://www.datasciencecentral.com/profiles/blogs/a-methodology-for-solving-problems-with-datascience-for-internet. Accessed 27 Nov 2017
5. Ni, P., Zhang, C., Ji, Y.: A hybrid method for short-term sensor data forecasting in Internet of Things. In: Proceedings of 2014 11th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), pp. 369–373 (2014)

6. Alam, F., Mehmood, R., Katib, I., Albeshri, A.: Analysis of eight data mining algorithms for smarter Internet of Things (IoT). Procedia Comput. Sci. **98**, 437–442 (2016)
7. Bodyanskiy, Y., Vynokurova, O., Pliss, I., Peleshko, D.: Hybrid adaptive systems of computational intelligence and their on-line learning for green it in energy management tasks In: Kharchenko, V., Kondratenko, Y., Kacprzyk, J. (eds.) Green IT Engineering: Concepts, Models, Complex Systems Architectures. Studies in Systems, Decision and Control, vol. 74, pp. 229–244. Springer, Cham (2017)
8. Peleshko, D., Peleshko, M., Kustra, N., Izonin, I.: Analysis of invariant moments in tasks image processing. In: Proceedings of 2011 11th International Conference the Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), Polyana-Svalyava, pp. 263–264 (2011)
9. Ivanov, Y., Peleshko, D., Makoveychuk, O., Izonin I., Malets I., Lotoshunska, N, Batyuk, D. Adaptive moving object segmentation algorithms in cluttered environments. In.: Proceedings of 2015 15th International Conference the Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), Lviv, pp. 97–99 (2015)
10. Rutkowski, L.: Computational Intelligence: Methods and Techniques. Springer, Berlin (2008)
11. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, Berlin (2006)
12. Murphy K.P.: Machine Learning: a Probabilistic Perspective. MIT Press (2012)
13. Bodyanskiy, Y.V., Vynokurova, O.A., Dolotov, A.I.: Self-learning cascade spiking neural network for fuzzy clustering based on group method of data handling. J. Autom. Inf. Sci. **45**, 23–33 (2013)
14. Hu, Z., Bodyanskiy, Y.V., Tyshchenko, O.K., Tkachov, V.M.: Fuzzy clustering data arrays with omitted observations. Int. J. Intell. Syst. Appl. (IJISA) **9**(6), 24–32 (2017). https://doi.org/10.5815/ijisa.2017.06.03
15. Bodyanskiy, Y., Vynokurova, O., Setlak, G., Peleshko, D., Mulesa, P.: Adaptive multivariate hybrid neuro-fuzzy system and its on-board fast learning. Neurocomputing **230**, 409–416 (2017)
16. Hu, Z., Bodyanskiy, Y.V., Tyshchenko, O.K., Samitova, V.O.: Possibilistic fuzzy clustering for categorical data arrays based on frequency prototypes and dissimilarity measures. Int. J. Intell. Syst. Appl. (IJISA) **9**(5), 55–61 (2017). https://doi.org/10.5815/ijisa.2017.05.07
17. Hu, Z., Bodyanskiy, Y.V., Tyshchenko, O.K., Samitova, V.O.: Fuzzy clustering data given on the ordinal scale based on membership and likelihood functions sharing. Int. J. Intell. Syst. Appl. (IJISA) **9**(2), 1–9 (2017). https://doi.org/10.5815/ijisa.2017.02.01
18. Babichev, S., Taif, M.A., Lytvynenko, V.: Inductive model of data clustering based on the agglomerative hierarchical algorithm. In: Proceedings of 2016 IEEE 1st International Conference on Data Stream Mining and Processing, DSMP 2016, pp. 19–22 (2016)
19. Babichev, S., Taif, M.A., Lytvynenko, V., Osypenko, V.: Criterial analysis of gene expression sequences to create the objective clustering inductive technology. In: Proceedings of 2017 IEEE 37th International Conference on Electronics and Nanotechnology, ELNANO 2017, pp. 244–248 (2017)
20. Bodyanskiy, Y., Vynokurova, O., Pliss, I., Setlak, G., Mulesa, P.: Fast learning algorithm for deep evolving GMDH-SVM neural network in data stream mining tasks. In.: Proceedings of 2016 IEEE 1st International Conference on Data Stream Mining and Processing, DSMP 2016, pp. 257–262 (2016)
21. Hu, Z., Bodyanskiy, Y.V., Tyshchenko, O.K., Samitova, V.O.: Fuzzy clustering data given in the ordinal scale. Int. J. Intell. Syst. Appl. (IJISA) **9**(1), 67–74 (2017). https://doi.org/10.5815/ijisa.2017.01.07

22. Bodyanskiy, Y., Setlak, G., Peleshko, D., Vynokurova, O.: Hybrid generalized additive neuro-fuzzy system and its adaptive learning algorithms. In: Proceedings of 2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS, pp. 328–333 (2015)
23. Perova, I., Pliss, I.: Deep hybrid system of computational intelligence with architecture adaptation for medical fuzzy diagnostics. Int. J. Intell. Syst. Appl. (IJISA) **9**(7), 12–21 (2017). https://doi.org/10.5815/ijisa.2017.07.02
24. Bin, S., Yuan, L., Xiaoyi, W.: Research on data mining models for the internet of things. In: Proceedings of 2010 IEEE International Conference on Image Analysis and Signal Processing, pp. 127–132 (2010)
25. Chen, F., Deng, P., Wan, J., Zhang, D., Vasilakos, A.V., Rong, X.: Data mining for the internet of things: literature review and challenges. Int. J. Distrib. Sens. Netw. **11**(8), 1–14 (2015)
26. Tsai, C.-W., Lai, C.-F., Chiang, M.-C., Yang, L.T.: Data mining for internet of things: a survey. IEEE Commun. Surv. Tutor. **16**(1), 77–97 (2014)
27. Zanella, A., Bui, N., Castellani, A., Vangelista, L., Zorzi, M.: Internet of 905 Things for smart cities. IEEE Internet of Things J. **1**(1), 22–32 (2014)
28. Bodyanskiy, Y., Lamonova, N., Pliss, I., Vynokurova, O.: An adaptive learning algorithm for a wavelet neural network. Exp. Syst. **22**(5), 235–240 (2005)
29. Yamakawa, T., Uchino, E., Miki, T., Kusanagi, H.: A neo-fuzzy neuron and its application to system identification and prediction of the system behavior. In: Proceedings of 2nd International Conference on Fuzzy Logic and Neural Networks, pp. 477–483, Iizuka, Japan (1992)
30. Miki, I., Yamakawa, I.: Analog implementation of neo-fuzzy neuron and its on-board learning. In: Computational Intelligence and Applications. pp. 144–149. WSES Press, Piraeus (1999)
31. Mitaim, S., Kosko, B.: What is the best shape for a fuzzy set in function approximation? In: Proceedings of 5th IEEE International Conference on Fuzzy Systems, Fuzz 1996, vol. 2, pp. 1237–1213 (1996)
32. Shepherd, A.J.: Second-Order Methods for Neural Networks. Springer, London (1997)
33. Bodyanskiy, Y., Kolodyazhniy, V., Stephan, A.: An adaptive learning algorithm for a neuro-fuzzy network. In: Reusch, B. (eds.) Computational Intelligence. Theory and Applications. Fuzzy Days 2001. LNCS, vol. 2206, pp. 68–75. Springer, Heidelberg (2001)
34. Goodwin, G.C., Ramadge, P.J., Caines, P.E.: A globally convergent adaptive predictor. Automatica **17**(1), 135–140 (1981)
35. MNIST Homepage. http://yann.lecun.com/exdb/mnist/. Accessed 19 Nov 2017