RESEARCH PAPERS OF THE WROCŁAW UNIVERSITY OF ECONOMICS
No. 25 2008
Knowledge Acquisition and Management

Yevgeniy Bodyanskiy, Sergiy Popov
Kharkiv National University of Radio Electronics, Ukraine
bodya@kture.kharkov.ua, serge.popov@gmx.net

MULTILAYER NETWORK OF NEURO-FUZZY UNITS IN FORECASTING APPLICATIONS

Abstract: A new neuron architecture named Neuro-Fuzzy Unit (NFU) is proposed, which is a generalization of elementary Rosenblatt's neuron and Neo-Fuzzy Neuron (NFN) by Yamakawa. NFU eliminates drawbacks of NFN, which complicate its application in multilayer networks. A learning algorithm based on gradient back-propagation procedure is introduced. The obtained results are confirmed by networks comparison in the task of short-term forecasting of electric load time series.

Key words: Neural Networks, Time Series Forecasting, Neuro-Fuzzy Systems.

1. Introduction

A Neo-Fuzzy Neuron (NFN) introduced by T. Yamakawa et al. [Yamakawa et al. 1992] is a simple nonlinear computational structure (Fig. 1) consisting of nonlinear synapses (Fig. 2) and a summation unit that forms its output. The neuron's output is obtained in the following manner

$$y = \sum_{i=1}^{n} \sum_{j=1}^{h} w_{ji} \mu_{ji}(x_i) = \sum_{i=1}^{n} f_i(x_i), \tag{1}$$

where x_i are inputs, μ_{ji} – membership levels, w_{ji} – synaptic weights, h – number of fuzzy intervals, n – number of inputs, y – output.

Membership levels depend on the distance between the input x_i and centers c_{ji}:

$$\mu_{ji}(x_i) = \begin{cases} \dfrac{x_i - c_{j-1,i}}{c_{ji} - c_{j-1,i}}, x_i \in [c_{j-1,i}, c_{ji}]; \\ \dfrac{c_{j+1,i} - x_i}{c_{j+1,i} - c_{ji}}, x_i \in [c_{ji}, c_{j+1,i}]; \\ 0 \text{ otherwise.} \end{cases} \tag{2}$$

$$\sum_{j=1}^{h} \mu_{ji}\left(x_i\right) = 1. \forall i. \tag{3}$$

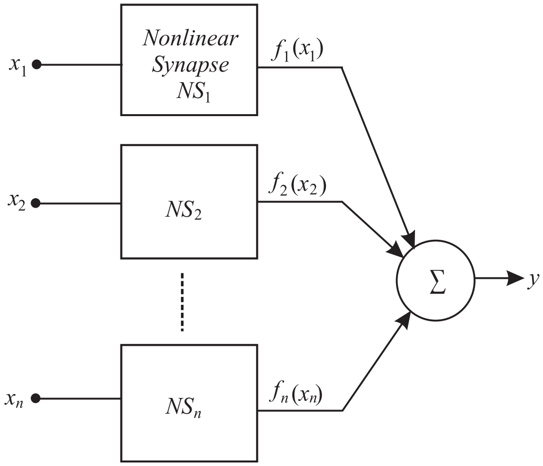

Fig. 1. Neo-Fuzzy Neuron (NFN)

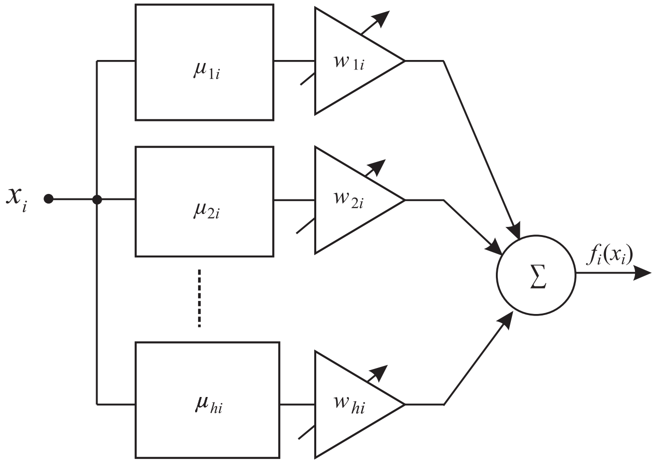

Fig. 2. Nonlinear synapse

Given the current active fuzzy interval p, the output of the nonlinear synapse can be expressed in this way:

$$f_i\left(x_i\right) = \sum_{j=1}^{h} w_{ji}\mu_{ji}\left(x_i\right) = w_{pi}\mu_{pi}\left(x_i\right) + w_{p+1,i}\mu_{p+1,i}\left(x_i\right) = \\ = \frac{c_{p+1,i} - x_i}{c_{p+1,i} - c_{pi}} w_{pi} + \frac{x_i - c_{pi}}{c_{p+1,i} - c_{pi}} w_{p+1,i} = a_i x_i + b_i, \tag{4}$$

where $a_i = \dfrac{w_{p+1,i} - w_{pi}}{c_{p+1,i} - c_{pi}}, b_i = \dfrac{c_{p+1,i} w_{pi} - c_{pi} w_{p+1,i}}{c_{p+1,i} - c_{pi}}$.

NFN is appropriate for many applications [Kolodyazhniy et al. 2005; Bodyanskiy et al. 2003], however, it revealed several disadvantages when used to form a multilayer network, known as Fuzzy Kolmogorov's Network (FKN) [Kolodyazhniy, Bodyanskiy 2004; Bodyanskiy et al. 2005; Kolodyazhniy et al. 2006]. The use of linear summation to form the neuron's output leads to unconstrained values of inputs to hidden and output layers. This significantly complicates learning process and degrades network stability.

2. Neuro-Fuzzy Unit

To overcome its disadvantages and make NFN more robust, we propose to add a nonlinear activation function to NFN. This makes a hybrid of a fuzzy system and an elementary neuron of Rosenblatt. We call this new type of neuron a Neuro-Fuzzy Unit (NFU) (Fig. 3). The output of NFU is formed by

$$
\begin{aligned}
& y = \psi\left(\sum_{i=1}^{n}\sum_{j=1}^{h} w_{ji}\mu_{ji}(x_i)\right) = \psi(u) = \\
& = \psi\left(\sum_{i=1}^{n} f_i(x_i)\right) E(k) = \frac{1}{2}\big(d(k) - y(k)\big)^2 = \\
& = \frac{1}{2}e^2(k) = \frac{1}{2}\big(d(k) - \psi(u(k))\big)^2 = \\
& = \frac{1}{2}\left(d(k) - \psi\left(\sum_{i=1}^{n}\sum_{j=1}^{h} w_{ji}\mu_{ji}(x_i(k))\right)\right)^2 = \\
& = \frac{1}{2}\left(d(k) - \psi\left(\sum_{i=1}^{n} w_i^T \mu_i(x_i(k))\right)\right)^2 = \frac{1}{2}\Big(d(k) - \psi\big(w^T \mu(x(k))\big)\Big)^2,
\end{aligned}
\tag{5}
$$

where ψ is a nonlinear activation function, e.g. hyperbolic tangent or sigmoid.

Introduction of additional nonlinearity at the neuron's output provides automatic limiting of output range that is very important for building multilayer networks.

Weight adjustment of NFU is performed with respect to the following quadratic criterion:

$$E(k)=\frac{1}{2}(d(k)-y(k))^2=\frac{1}{2}e^2(k)=\frac{1}{2}(d(k)-\psi(u(k)))^2=$$
$$=\frac{1}{2}\left(d(k)-\psi\left(\sum_{i=1}^{n}\sum_{j=1}^{h}w_{ji}\mu_{ji}(x_i(k))\right)\right)^2= \tag{6}$$
$$=\frac{1}{2}\left(d(k)-\psi\left(\sum_{i=1}^{n}w_i^T\mu_i(x_i(k))\right)\right)^2=\frac{1}{2}\left(d(k)-\psi\left(w^T\mu(x(k))\right)\right)^2,$$

where k is a discrete time, $w_i=(w_{1i},w_{2i},\ldots,w_{hi})^T$, $w=\left(w_1^T,w_2^T,\ldots,w_h^T\right)^T$,
$\mu_i(x_i(k))=(\mu_{1i}(x_i(k)),\mu_{2i}(x_i(k)),\ldots,\mu_{hi}(x_i(k)))^T$,
$\mu(x(k))=\left(\mu_1^T(x_1(k)),\mu_2^T(x_2(k)),\ldots,\mu_n^T(x_n(k))\right)^T$.

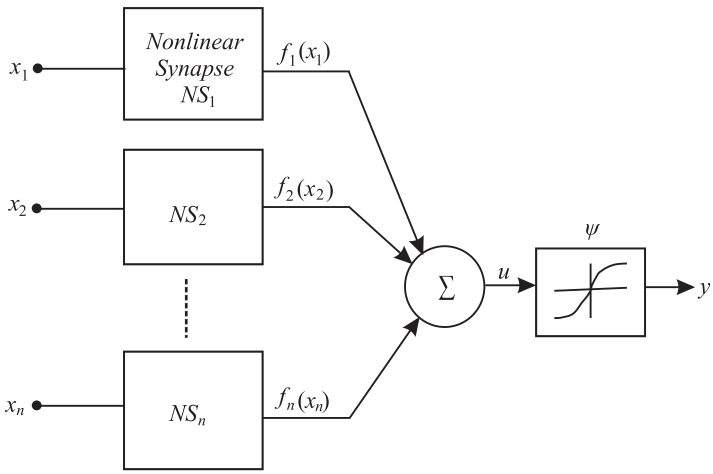

Fig. 3. Neuro-Fuzzy Unit (NFU)

To minimize (6), a gradient descent learning algorithm can be applied

$$w(k+1)=w(k)-\eta(k)\nabla_w E(k)=$$
$$=w(k)+\eta(k)e(k)\frac{\partial e(k)}{\partial u(k)}\nabla_w u(k)= \tag{7}$$
$$=w(k)+\eta(k)e(k)\psi'(u(k))\mu(x(k)),$$

where $\eta(k)$ is a learning rate.

3. Simulation results

To validate our theoretical results we compared the performance of the proposed NFU network against FKN network. Short-term electric load forecasting problem is used as a benchmark. Power generation and transmission companies need to know future electric load to balance it with power supply. Both low and high power supply incur additional costs or penalties, that is why accurate electric load forecasting is crucial: improvement of forecasting accuracy by as low as 0.1% may save thousands of dollars.

We used electric load time series from Puget Sound Power and Light Company, obtained from a well-know forecasting competition [Engle et al. 1992]. We did not fully reproduce the setup of the original competition but instead used the period from January 16th, 1985 to January 15th, 1987 for networks training, and period from January 16th, 1987 to January 15th, 1988 for testing. The networks forecasted electric load on an hourly basis for 24 hours ahead using information about past load and temperature, and future temperature and day of week (total of 8 inputs).

The aim was to compare two types of networks and not to obtain the best possible forecast for a given problem, so we did not get deeply into the inputs and model structure selection processes. Both networks had 10 neurons in one hidden layer and a single output neuron.

The forecasting error (MAPE) on the test set was 2.78% for FKN and 2.63% for NFU network. Besides better accuracy, as expected, NFU network was much more stable during training because outputs of the hidden layer always lie between –1 and 1.

4. Conclusions

The proposed Neuro-Fuzzy Unit solves several problems inherent to Neo-Fuzzy Neuron which stem from its linear activation mechanism. Adding nonlinear activation function significantly improves network learning stability and approximating power, as shown in simulations on electric load time series.

References

Bodyanskiy Ye., Gorshkov Ye., Kolodyazhniy V. (2005). Neuro-fuzzy Kolmogorov's network with a hybrid learning algorithm. In: *Proc. XI Int. Conf. "Knowledge – Dialogue – Solution"*, 2. FOI Commerce, Sofia, pp. 622-627.

Bodyanskiy Ye., Kokshenev I., Kolodyazhniy V. (2003). An adaptive learning algorithm for a neo-fuzzy neuron. In: *Proc. 3rd Int. Conf. of European Union Society for Fuzzy Logic and Technology (EUSFLAT 2003)*. EUSFLAT, Zittau, pp. 375-379.

Engle R. et al. (1992). *Construction of the Puget Sound forecasting model*. *EPRI Project RP919*. Quantitative Economic Res. Inst., San Diego, CA.

Kolodyazhniy V., Bodyanskiy Ye. (2004). Fuzzy Kolmogorov's Network. In: *Proc. 8th Int. Conf. Knowledge-Based Intelligent Information and Engineering Systems International (KES 2004)*. Eds. M. Negoita, R.J. Howlett, L.C. Jain. Springer, Berlin, pp. 764-771.

Kolodyazhniy V., Bodyanskiy Ye., Otto P. (2005). Universal approximator employing neo-fuzzy neurons. In: *Computational Intelligence: Theory and Applications*. Ed. B. Reusch. Springer, Berlin, pp. 631-640.

Kolodyazhniy V. et al. (2006). Neuro-fuzzy Kolmogorov's network with a modified perceptron learning rule for classification problems. In: *Computational Intelligence: Theory and Applications*. Ed. B. Reusch. Springer, Berlin, pp. 41-49.

Yamakawa T. et al. (1992). A neo fuzzy neuron and its applications to system identification and prediction of the system behavior. In: *Proc. 2nd Int. Conf. Fuzzy Logic and Neural Networks*. Iizuka, Japan, pp. 477-483.