

Neuron Model Utilizing Information of Local Samples for Forecasting Management State of Enterprise

Jun Zhai Zhou Yan Liang Lixin Shen

School of Economics and Management, Dalian Maritime University, Dalian 116026, P.R.China

Abstract

In order to conquer the localization of the multi-layered feed forward neural networks, this paper presents a kind of neuron models utilizing information of local samples—the UILS neuron model, including an adaptive neuron model and a self-organization neuron model. Differing from traditional models, it fully employs the experience samples information within the local range and well embodies the association and analogy functions of cerebrum. Through investigating the properties of UILS and learning algorithm, we build a method based on the neuron model for forecasting the management state of enterprise according to the economic and technology indexes. It is verified that using the UILS model, expert experiences can be well expressed in the forecasting results.

Keywords: Neuron model, Neural network, Forecasting, Management state of enterprise

1. Introduction

It is the premise and guarantee for scientific decision to make exact and rational forecasting of the states of management in enterprise [1]-[3]. However, most of already existed forecasting methods, with adopting direct extrapolation, exponential smoothing, and regression analysis and so on, are difficult in dealing with high nonlinear system. Moreover, these methods lay particular stress on qualitative indexes but neglect quantitative indexes, which cause insufficient use of information.

It offers a new approach to solve above questions that utilizing Artificial Neural Networks (ANN) for expressing and getting knowledge [2]-[3]. The basic method is summing up the typical samples according to the experiential knowledge of experts, based on which constructing ANN to get the interior expression of the knowledge by training. Essentially, the neural networks simulate the expert experiences through learning to solve the uncertain questions in the knowledge representation.

Now, common multi-layered feed forward neural networks (FNN) storage the knowledge on the weight values, which' obtain rely on the training of whole samples. In fact, it is a neural network with global property that the corresponding output of the input comes impliedly down to whole samples information. This character possibly weakens the effect of the application of FNN in some complex domains [4].

In order to conquer the localization, literature [4]-[5] put forward a kind of neuron model Utilizing Information of Local Samples. Compared with other neuron models and networks, this neuron model has such characters as follows:

- Taking full advantage of the information of experience samples and using that directly on the output of neuron model;
- Having the local property, which means that the output of neuron model are the weighting sum of experience samples output in local range;
- Having a simple structure and a rapid learning speed.

This neuron model had been applied successfully in lateral prediction for reservoir parameters of the explanation for oil seismic data [4] and in rotary machine fault diagnosis [6]

Based on the literature [4]-[6], this paper generalizes the UILS neuron model, which makes the characters more stand out and the functions more perfect. We study its properties and learning algorithms, then build a method based on the neuron model for forecasting the management state of enterprise and give an example.

2. Adaptive UILS neuron model

The Adaptive UILS neuron model store the experience samples which reflect expert experiences as the standard patterns of input-output-pair. When the neuron receives input information, its inner calculation function collates the information with the experience samples then produce corresponding output information. This process is the embodiment of the cerebral association and analogy functions.

The structure of the model is shown in fig.1. It realizes the mapping from \mathbb{R}^n to \mathbb{R} , which includes an experience samples set $\{(x^k, y^k), x^k \in \mathbb{R}^n, y^k \in \mathbb{R}, k=1, \dots, m\}$, the model parameters $b_{ij} (i, j = 1, \dots, n)$, and four functions which involve the distance function $r_k(\cdot)$, the active function $\varphi(\cdot)$, the weight function $w_k(\cdot)$ and the output function $g(\cdot)$

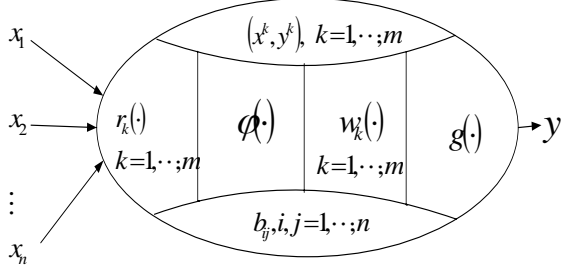


Fig. 1: The structure of the adaptive UILS neuron model.

Let $x = (x_1, x_2, \dots, x_n)^T, x^k = (x_1^k, x_2^k, \dots, x_n^k)^T, k=1, \dots, m, X = \{x^1, x^2, \dots, x^m\}$. The calculation process from the input x to the output y of the model is described by such expressions as follows:

$$y = g(x) = \sum_{k=1}^m w_k(x) \cdot y^k \quad (1)$$

$$W_k(x) = \begin{cases} \frac{\varphi_k(x)}{\sum_{j=1}^m \varphi_j(x)} & x \notin X, \sum_{j=1}^m \varphi_j(x) \neq 0 \\ 1 & x \in X, x = x^k \\ 0 & \text{else} \end{cases}, \quad (2)$$

$$k=1, 2, \dots, m$$

$$\varphi(r) = \begin{cases} 1/r & 0 < r \leq 1/3 \\ (27/4)(r-1)^2 & 1/3 < r \leq 1 \\ 0 & 1 < r \end{cases} \quad (3)$$

$$\varphi_k(x) = \varphi(r_k(x)), k=1, 2, \dots, m \quad (4)$$

$$r_k(x) = r_A(x, x^k) = [(x - x^k)A(x - x^k)]^{\frac{1}{2}}$$

$$= [(x - x^k)B^T B(x - x^k)]^{\frac{1}{2}} \quad (5)$$

$$k=1, 2, \dots, m$$

where $A = (a_{ij})_{n \times n}$ is a symmetry positive definite matrix, $a_{ij}, i, j = 1, \dots, n$ are the model parameters. It is hard to keep the symmetry of matrix A in learning process, so we decompose A to $A = B^T B$. $B = (b_{ij})_{n \times n}$ as a nonsingular matrix, which

replaces A to be the model parameter is determined by the learning algorithm.

The actions and characters of such expressions are as follows:

(1) The output function $g(x)$ defined in Eq.(1) realizes the weighting sum of each experience sample as the neuron output y

(2) The weight $W_k(x)$ of each experience sample defined in Eq. (2) is direct ratio to its active value $\varphi_k(x)$ (Eq. (4)). The characters are as follows:

$$i) \sum_{k=1}^m W_k(x) = 1;$$

$$ii) 0 \leq W_k(x) \leq 1, k = 1, 2, \dots, m.$$

(3) The active function $\varphi: \mathbb{R}^+ \rightarrow \mathbb{R}$ defined in Eq. (3) has several characters as follows:

$$i) \varphi(r) \rightarrow +\infty \text{ as } r \rightarrow 0;$$

$$ii) \varphi(r) = 0 \text{ as } r \geq 1;$$

iii) $\varphi(r)$ is monotonically decreasing and continuously differentiable function with respect to the variable r .

(4) The active function φ acts on the distance $r_k(k)$ (Eq. (5)) between various experience samples and the input x , and then get the active value $\varphi_k(x)$ (Eq. (4)) of each sample. The less the distance the bigger the active value and weight, vice versa. When the distance bigger than 1, the active value and weight are 0 at the same time. This cut-off character of φ determines the local character of the neuron model, which is that the experience samples acts on the output y only when the distance between the input of experience samples and x is less than 1 and the weight of that experience samples are not 0.

(5) The distances between each samples and input x are given by Eq. (5). Acted on the distance function $r_A(x, y)$, the experience samples which locate in the super ellipsoid $\{z \in \mathbb{R}^n \mid r_A(z, x) < 1\}$ operate on the output of the input x . The directions of each axis and the length of each radius reflect the influence of the space distribution of the experience samples on the output. Using the learning algorithm, the size and direction of the super ellipsoid can be adjusted adaptively and the local sample information can be taken full advantage of.

If take $A = \text{diag}(1/a_1^2, 1/a_2^2, \dots, 1/a_n^2)$, we get the distance function $r_k(x)$ as

$$r_k(x) = \left(\sum_{i=1}^n (x_i^k - x_i)^2 / a_i^2 \right)^{\frac{1}{2}} \quad (6)$$

Then we get the basic UILS neuron model in literature [4]-[5], which has inferior adaptive capability due to the less adaptive parameters.

The monolayer feed forward constituted by l AUILS neuron models can realize the mapping from R^n to R^l .

3. Properties of UILS neuron model

When the model parameters confirmed, the relationship that mapping from R^n to R which realized by AUILS can be simplified into $y = g(x)$

Definition 1: The set $F \subset R^n$ is defined as the model action domain, where

$$F = \{x \in R^n \mid \text{there exist some } i, 1 \leq i \leq m, \text{ such that } W_i(x) > 0\}$$

Only when the input $x \in F$, the experience sample set acts on its output y and the input is an effective one.

$$\text{Let } F_k = \{x \in R^n \mid r_k(x) < 1\}, k = 1, 2, \dots, m, \\ \text{then } \varphi_k(x) > 0, \text{ that is } W_k(x) > 0 \text{ if } x \in F_k,$$

hence we get

Theorem 1: The set F_k is convex.

Theorem 2: The model action domain $F = \bigcup_{k=1}^m F_k$.

Theorem 3: The function $y = g(x)$ is continuous in the model action domain F ; it is continuously differentiable in the set $F - X$.

Theorem 4: The function $f(x)$ satisfies the following Lipschitz condition:

$$|f(x) - f(y)| \leq M \|x - y\|^\theta, x, y \in R^n$$

where $M > 0, 0 < \theta \leq 1$

Here the experience sample is chosen from f , namely $y_k = f(x_k), k = 1, 2, \dots, m$, then if $x \in F$, we have $|f(x) - g(x)| \leq c$,

where $c = M \left(\frac{1}{\alpha}\right)^\theta, \alpha = 1 / \max a_i, 1 < i \leq n$. Here we

define the distance function as Eq. (6).

Let $g(p)$ be the function of parameter vector,

where $p = (b_{11}, b_{12}, \dots, b_{1n}, \dots, b_{m1}, b_{m2}, \dots, b_{mn})^T$.

$$U_k(x) = \{p \in R^{n \times n} \mid r_k(x) \geq 1\}$$

$$\text{Let } U(x) = \bigcap_{k=1}^m U_k(x)$$

Theorem 5: The function $y = g(p)$ is continuously differentiable in the set $V(x) = \bar{U}$.

4. Gradient formula for learning parameters

The parameters b_{ij} of above neuron model are confirmed by the learning algorithm.

Let $b = (b_{11}, b_{12}, \dots, b_{1n}, \dots, b_{m1}, b_{m2}, \dots, b_{mn})^T$, the function realized by Eq. (1) ~ Eq. (5) can be simplified as

$$y = G(x, b) \quad (7)$$

Set $\{(x^{(j)}, y^{(j)}), x^{(j)} \in R^n, y^{(j)} \in R, j = 1, 2, \dots, s\}$ as the learning sample set. Choosing a objective function which's parameter vector b

$$E(b) = \frac{1}{2} \sum_{j=1}^s (y^{(j)} - \hat{y}^{(j)})^2 \quad (8)$$

Here $\hat{y}^{(j)} = G(x^{(j)}, b), j = 1, 2, \dots, s$.

The learning algorithm based on steepest decent method is shown as follows:

$$(b_{ij})_l = (b_{ij})_{l-1} - \alpha \frac{\partial E}{\partial (b_{ij})_{l-1}} \quad (9)$$

Here $\alpha \geq 0$ is the step length. We can get $\frac{\partial E}{\partial b_{ij}}$ by the

expressions as follows:

$$\frac{\partial r_k}{\partial b_{ij}} = \frac{[(x_j - x_j^k) \sum_{l=1}^n b_{il} (x_l - x_l^k)]}{r_k},$$

$$k = 1, \dots, m, i, j = 1, \dots, n$$

$$\frac{d\varphi_k}{dr_k} = \begin{cases} -1/r_k^2, 0 < r_k \leq 1/3 \\ (27/2)(r-1), 1/3 < r_k \leq 1 \\ 0, 0 < r_k, k = 1, \dots, m, \end{cases}$$

$$\frac{\partial \phi_k}{\partial b_{ij}} = \frac{d\phi_k}{dr_k} \cdot \frac{\partial r_k}{\partial b_{ij}}, k = 1, \dots, m, i, j = 1, \dots, n$$

$$\frac{\partial W_k}{\partial \varphi_j} = \frac{\sum_{j=1, j \neq k}^m \varphi_j}{\left(\sum_{j=1}^m \varphi_j\right)^2} \text{ as } j = k \quad k, j = 1, \dots, m$$

$$\frac{\partial W_k}{\partial \varphi_j} = \frac{-\varphi_k}{\left(\sum_{j=1}^m \varphi_j\right)^2} \quad \text{as } j \neq k, k, j = 1, \dots, m$$

$$\frac{\partial W_k}{\partial b_{ij}} = \sum_{l=1}^m \frac{\partial W_k}{\partial \varphi_l} \frac{\partial \varphi_l}{\partial b_{ij}} \quad k = 1, \dots, m, i, j = 1, \dots,$$

$$\frac{\partial \hat{y}}{\partial b_{ij}} = \sum_{k=1}^m \frac{\partial \hat{y}}{\partial W_k} \frac{\partial W_k}{\partial a_i} = \sum_{k=1}^m y^k \frac{\partial W_k}{\partial b_{ij}}, i, j = 1, \dots, n$$

$$\frac{\partial E}{\partial b_{ij}} = \sum_{l=1}^s (\hat{y}^{(l)} - y^{(l)}) \frac{\partial \hat{y}^{(l)}}{\partial b_{ij}}, i, j = 1, \dots, n$$

5. Self-organization UILS neural network model

Because the model's parameters, gotten by learning algorithm, reflect system's overall properties that involved in whole experience samples, large amounts of samples are bound to bring out difficulties to neuron learning. Furthermore, excessive samples are not necessary for neurons themselves.

Considering the local property of UILS neuron model in input space, in the case of larger amounts of samples, according to the distribution of samples in the input space, we can use the cluster rules to divide the whole input space into many subsets and each subset has appropriate number of samples. In each subset, we use an UILS neuron model, and then all neurons constitute the network system in the entire input space called self-organization UILS neural network illustrated in Fig. 2.

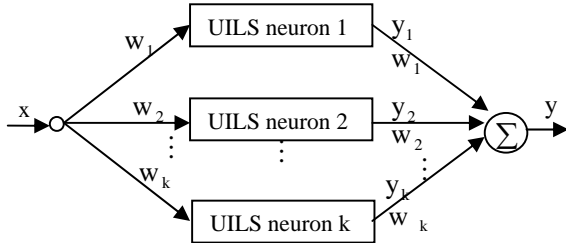


Fig. 2: The structure of self-organization UILS neural network model.

In Fig. 2, k represents the number of subset and neurons. The input of each neuron is described as $w_i x$ and corresponding output is y_i . If x belongs to the k subset, then

$$w_i = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases} \quad (10)$$

And the output of entire network is

$$y = \sum_{i=1}^k w_i y_i = y_k \quad (11)$$

For each input x , there is only one active neuron.

The parameters of neurons are gained by supervised learning, but the weight w_i is gotten by non-supervised learning, so this neural network is self-organizing.

Different cluster algorithm can form different self-organization network model.

6. The method for forecasting management state of enterprise

We apply UILS model to forecasting the states of management in enterprise because of its particular capability in information disposal. The key of forecasting is to build the function relationship between the states of economic technology and management in enterprise. The process is explained as follows:

- (1) Characters extraction. There are many characters represent the states of management in enterprise. We describe it from two aspects: technology usage and economic increase. Technology resource can be analyzed by the index of advanced degree of equipment, index of technology level of equipment, index of input-output of material, and index of educational level of staff. Economic increase is mainly measured by economic benefits and product quality.
- (2) Building sample sets. We quantize the characters and describe their values in numbers between 0 and 1 which are also used to describe the states of management in enterprise. Then, according to business history, we make each value of characters corresponds with each value of states of management to form a sample set.
- (3) Learning neuron model. Use learning algorithm to train ULSI neuron in order to make it represents the relation between the states of management and characters.
- (4) Forecasting. At first, input the new values of characters to trained neuron model, and use the output to describe the states of management. Then qualify the output to get quantitative and qualitative description of states of management.

Tab. 1 shows the partial data of the states of economic technology and management of one enterprise to train neuron model. The coefficient of correlation between values of estimated by the neuron model and actual is 0.99. Tab. 2 shows the partial forecasting results of management state which can make us confirm that the forecasting results can reflect the trend of history data.

7. Conclusions

State of manage	Index of advanced degree of equipment	Index of input-output of material	Index of educational level of staff	Ratio of pre tax profits to assets	Lose rate for quality
excellent (1.0)	0.85	0.20	0.75	0.45	0.10
	0.87	0.22	0.70	0.42	0.11
	0.90	0.19	0.72	0.44	0.16
good (0.8)	0.70	0.15	0.65	0.35	0.18
	0.73	0.13	0.63	0.32	0.17
	0.68	0.14	0.60	0.33	0.15
middle (0.6)	0.66	0.12	0.57	0.30	0.25
	0.65	0.13	0.51	0.29	0.24
	0.63	0.11	0.56	0.28	0.24
bad (0.4)	0.50	0.07	0.45	0.20	0.30
	0.52	0.05	0.42	0.25	0.31
	0.54	0.04	0.41	0.24	0.32

Table 1: The partial data of states of economic technology and management

Forecasting result	Index of advanced degree of equipment	Index of input-output of material	Index of educational level of staff	Ratio of pre tax profits to assets	Lose rate for quality
0.99 (excellent)	0.83	0.21	0.73	0.44	0.12
0.98 (excellent)	0.86	0.23	0.67	0.39	0.11
0.89 (good)	0.78	0.17	0.69	0.40	0.14
0.83 (good)	0.72	0.16	0.68	0.38	0.17
0.62 (middle)	0.68	0.14	0.55	0.26	0.31
0.54 (middle)	0.57	0.09	0.50	0.32	0.23
0.44 (bad)	0.60	0.06	0.50	0.22	0.33
0.43 (bad)	0.57	0.09	0.51	0.26	0.31

Table 2: The partial forecasting results.

The method for forecasting management state of enterprise based on UILS model possesses several characters as follows:

- (1) Set the expert experience as knowledge which are utilized and stored directly.

- (2) Efficiently combine the expert experiences to form new knowledge.
- (3) The forecasting process simulates the associative and analogical function of the human brain, which is intelligent.

The examples indicate that the UILS neuron model, with fast learning and strong adaptability, can extract rules from history data and refresh the rules with samples' increasing and changing, which provides a new method and tool for forecasting the states of management in enterprise.

References

- [1] C.G. She and Y.M. Xi, The theoretics review on our country's corporation early-warning studies. *Forecasting*, 22(2):23-29, 2003. (In Chinese)
- [2] X.Y. Huang and S.H. Xiao, Neural network warning system and its application in enterprise motion. *Systems Engineering and Electronics*, 17(10):50-58, 1995. (In Chinese)
- [3] H. Ren and X.S. Xu., A warning system of enterprise's crisis management. *Journal of Wuhan Automotive Polytechnic University*, 25(6): 153-157, 2003.(In Chinese)
- [4] J. Zhai and Y.J. Feng, A kind of neuron model of local property. *Acta Electronica Sinica*, 25(11):113-115, 1997. (In Chinese)
- [5] Y.J. Feng, Q.W. Ran and W.G. Mao, A new kind of artificial neuron models. *Journal of Harbin Institute of Technology*, Vol.E-1:13-15, 1994.
- [6] J. Zhai, X.J. Yan and Y. Chen, A method of adaptive neuron model and its application. In: *Proceedings of Fifth IEEE International Conference on Cognitive Informatics*, Beijing (ICCI 2006, Vol.1), pp.47-52, 2006.