

# Interweave Neural Networks with Evolutionary Algorithms, Cellular Computing, Bayesian Learning and Ensemble Learning

Jing Zhang

Beijing Union University  
Beijing, The People's Republic of China  
e-mail: zhang1jing4@sina.com

**Abstract**—A neural networks are able to give solutions to complex problems in business intelligence and financial engineering due to their nonlinear processing. This paper consists of a survey of various business intelligence and financial engineering and so on applications based on the neural networks, and also a summary of the recent techniques such as still evolutionary algorithms, cellular computing, Bayesian learning and ensemble learning. The neural networks can be powerful tools in several fields of business intelligence and financial engineering. It can be concluded that, the future of neural networks lies in their use in conjunction with other advanced technologies.

**Keywords**—neural networks; evolutionary algorithms; cellular computing; Bayesian learning; ensemble learning

## I. INTRODUCTION

Neural networks are able to give solutions to complex problems in business intelligence and financial engineering due to their nonlinear processing. Neural networks possess properties which give them advantages over other means. The great majority of applications of neural networks are based on the simple backpropagation algorithm. Recent papers on integration of neural networks with other computing paradigms such as evolutionary algorithms, cellular computing, Bayesian learning and ensemble learning to enhance the performance of fuzzy neural network models are presented.

In recent years, many natural computing models have been successfully applied to the solution of complex problems related to business intelligence and financial engineering. These provide efficient paradigms for fusion of different kinds of information. The aim is to synergistically merge the techniques so that they cooperate with each other in enhancing the overall performance.

The paper gives an overview of the applications of neural networks. Recent papers of neural networks aims at covering the most representative applications of neural networks such as financial diagnosis, money laundering surveillance, business data mining, credit fraud detection, business decision making, etc. The abundant literature that has grown exponentially in the recent years shows that the fields of commerce and financial gives great interest to neural network.

## II. EVOLUTIONARY ALGORITHMS

Martínez-Estudillo, Hervás-Martínez, Gutiérrez and Martínez-Estudillo<sup>[1]</sup> proposed a special assortment method based on a product-unit neural networks. A node is given by

$$y_j = \prod_{i=1}^k x_i^{w_{ji}} \quad (1)$$

where  $k$  is the number of the inputs. If the exponents  $w_{ji}$  in (1) are  $\{0, 1\}$ , then we obtain a higher-order unit, also known as sigma-pi unit. The estimated function  $f_l(\mathbf{x}; \boldsymbol{\theta}_l)$  from each output is given by

$$f_l(\mathbf{x}; \boldsymbol{\theta}_l) = \beta_0^l + \sum_{j=1}^m \beta_j^l B_j(\mathbf{x}, \mathbf{w}_j), l = 1, 2, \dots, J, \quad (2)$$

where  $\mathbf{x} = (x_1, x_2, \dots, x_k)$ ,  $\mathbf{w}_j = (w_{j1}, \dots, w_{jk})$ ,  $\boldsymbol{\theta}_l = (\beta^l, \mathbf{w}_1, \dots, \mathbf{w}_m)$ ,

$B_j(\mathbf{x}, \mathbf{w}_j) = \prod_{i=1}^k x_i^{w_{ji}}$  and  $\beta^l = (\beta_0^l, \beta_1^l, \dots, \beta_m^l)$ . From a statistical point of view, the method can be seen as nonlinear logistic regression where the parameters are estimated using evolutionary computation. It uses softmax transformation and the cross-entropy error function.

Cheng, Bai and Cao<sup>[2]</sup> presented parallel chaos immune evolutionary programming based on clonal selection theory.

$$A_k = \{a_i \mid f(a_i) \geq \bar{f}, i \in N\} \quad (3)$$

where

$$\bar{f} = \frac{1}{N} \sum_{i=1}^N f_i, i = 1, 2, \dots, N. \quad (4)$$

In  $A_k$ , anti-body  $a_i$  will be cloned into  $q_i$  anti-bodies.

$$q_i = \text{Int}(C * P_i), i = 1, 2, \dots, M. \quad (5)$$

$q_i$  is adjusted to be adaptable according to  $C$  and  $p_i$ . The constant  $C$  is a given integer related to the clonal size. Here,  $p_i$ , the probability of antibody  $a_i$ , producing new anti-bodies is

$$P_i = f(i) / \sum_{j=1}^M f(j), i = 1, 2, \dots, M. \quad (6)$$

After population cloning,  $C_k$  replaces population  $A_k$ . Experimental results show that parallel chaos immune evolutionary programming can effectively prevent premature convergence and is of high efficiency.

Hervás-Martínez, Martínez-Estudillo and Carbonero-Ruz<sup>[3]</sup> proposed a model based on the combination of linear and product-unit models, where the nonlinear basis functions

of the model are given by the product of the inputs raised to arbitrary powers. The general expression of the multilogistic regression model is given by:

$$f_l(\mathbf{x}; \boldsymbol{\theta}_l) = \alpha_0^l + \sum_{i=1}^k \alpha_i^l x_i + \sum_{j=1}^m \beta_j^l \prod_{i=1}^k x_i^{w_{ji}}, l = 1, 2, \dots, J-1, \quad (7)$$

where  $\mathbf{x} = (x_1, x_2, \dots, x_k)$ ,  $\boldsymbol{\theta}_l = (\boldsymbol{\alpha}^l, \boldsymbol{\beta}^l, \mathbf{w}_1, \dots, \mathbf{w}_m)$ ,  $\mathbf{w}_j = (w_{j1}, \dots, w_{jk})$ ,  $\boldsymbol{\alpha}^l = (\alpha_0^l, \alpha_1^l, \dots, \alpha_k^l)$  and  $\boldsymbol{\beta}^l = (\beta_1^l, \dots, \beta_m^l)$ . The model is competitive in terms of the accuracy of the classifier.

Au, Choi and Yu<sup>[4]</sup> have studied the evolutionary neuron network for the forecasting of apparel sales is developed. It is applicable for short term retail forecasting of the fashion retail sales data, which share these features.

Gholizadeh, Salajegheh and Torkzadeh<sup>[5]</sup> proposed a combination of genetic algorithm and neural networks. It found the optimal weight of structures subject to multiple natural frequency constraints. The discrete structural optimization model is given by:

$$\min f(\mathbf{x}) = \sum_{i=1}^{ne} \rho_i x_i l_i \quad (8)$$

$$s.t. g_i(\mathbf{x}) = \frac{\lambda_i}{\lambda_{all}} - 1 \leq 0, i = 1, 2, \dots, m. \quad (9)$$

where  $\rho_i$  and  $l_i$  are weight of unit volume and length of  $i$ th element,  $ne$  is the number of the structural elements,  $\lambda_i$  and  $\lambda_{all}$  are the  $i$ th frequency and allowable frequency. The numerical results demonstrate that the suggested methods provide a more robust tool for structural optimization with frequency constraints.

### III. CELLULAR COMPUTING

Martínez, Toledo, Fernández and Ferrández<sup>[6]</sup> described a new architecture of the hardware implementation of nonlinear multi-layer cellular neural network. The model shown:

$$X_{ij}[n] = \sum_{k,l \in Nr(ij)} A_{kl}[n-1] Y_{kl}[n-1] + \sum_{k,l \in Nr(ij)} B_{kl}[n-1] U_{kl} + I_{ij}, \quad (10)$$

$$Y_{ij}[n] = \frac{1}{2}(|X_{ij}[n] + 1| - |X_{ij}[n] - 1|), \quad (11)$$

where  $I$ ,  $U$ ,  $Y$  and  $X$  denote input bias, input data, output data and state variable of each cell, respectively. The neighbourhood distance  $r$  for cell  $(i, j)$  is given by  $Nr(ij)$  function.  $A$  is the non-linear template for the outputs of the neighbouring cells and  $B$  is the corresponding non-linear weights template for the inputs.

Albora, Ucan, Ozmen and Ozkan<sup>[7]</sup> have optimised the weight coefficients of the behaviour of the cellular neural network using the recurrent perceptron learning algorithm. The cell dynamical equations may be stated as follows:

$$x_{ij}(n+1) = -x_{ij}(n) + \sum_{kl \in Nr(ij)} A_{(i-k)(j-l)}(n) y_{kl} + \sum_{kl \in Nr(ij)} B_{(i-k)(j-l)}(n) u_{kl} + I, \quad (12)$$

$$y_{ij}(n) = \frac{1}{2}(|x_{ij}(n) + 1| - |x_{ij}(n) - 1|), \quad (13)$$

where  $x$ ,  $y$ ,  $u$ ,  $I$  denote respectively cell state, output, input, bias, and  $j$  and  $k$  are cell indices.

Bertuccio, Coltelli, Nunnari and Occhipinti<sup>[8]</sup> introduce a new method, real-time monitoring of active volcanoes, its foundation is an effective video processing operations implemented by means of cellular neural network architecture. Cellular neural networks are massive parallel analog circuits only place the relationship between the calculation of elements of these programs in a simulation approach to the implementation of almost all of image processing operations.

Gupta and Ogawa<sup>[9]</sup> proposed a novel prejudice method based on the concept of bias coding. The bias-coded is defined by

$$u_{i,j}^* = u_{i,j} + q, \quad (14)$$

$$y_{i,j}^*(n) = y_{i,j}(n) + z, \quad (15)$$

$$a_{k,l}^* = a_{k,l} + s, \quad (16)$$

$$b_{k,l}^* = b_{k,l} + v, \quad (17)$$

where  $q = |\min\{\min_{i,j} u_{i,j}, 0\}|$ ,  $z = |\min\{\min_{i,j} y_{i,j}(0), 0\}|$ ,

$s = |\min\{\min_{i,j} a_{k,l}, 0\}|$ ,  $v = |\min\{\min_{i,j} b_{k,l}, 0\}|$ ,  $y_{i,j}(0)$  represents the output of a cell  $C_{i,j}$  corresponding to initial state  $x_{i,j}(0)$ .

Lent, Tougaw, Brazhnik, Weng, Porod, Liu and Huang<sup>[10]</sup> consider a different approach to the use of coupled quantum-dot cells in the architecture, rather than sound Boolean logic, the use of physical connections to build a near neighbor analog cellular neural network. They have proposed the use of coupling a quantum dot construction figures elements, quantum dot cellular automata.

Fortuna, G. Manganaro, G. Muscato, G. Nunnari<sup>[11]</sup> presented a novel simulator for cellular neural networks. The state transition equation:

$$x_{ij}(k+1) = (1-h)x_{ij}(k) + h \sum_{C(k,l) \in Nr(i,j)} A(i,j,k,l) y_{kl}(k) + \sum_{C(k,l) \in Nr(i,j)} B(i,j,k,l) u_{kl} + hI \quad (18)$$

Breton, Fonvieille, Grenier, Guicheney, Jousset, Roblin and Tamin<sup>[12]</sup> improved greatly in application of neural networks and cellular automata to interpretation of calorimeter data. All of the  $i$  neurons of the network will give an output  $O_i$ , by performing a weighted sum of the  $O_j$  outputs of the  $j$  neurons they are connected with. The output of any neuron is obtained by applying the transition function  $f$  on the weighted sum of its inputs:

$$O_i = f\left(\sum_{j=1}^N W_{ij} O_j - \theta_i\right) \quad (19)$$

where  $\theta_i$  = threshold of neuron  $i$ ,  $W_{ij}$  = connection weights between neurons  $i$  and  $j$ ,  $f(x) = (1 - e^{-kx}) / (1 + e^{-kx})$ .

### IV. BAYESIAN LEARNING

Wu and Baldi<sup>[13]</sup> propose to use Bayesian network which in turn leads to a directed acyclic graph recursive neural network architecture. In the simulations presented, they use a simple scheme

$$T_{ij}(t) = (1 - w)A_{ij}(t) + wA_{ij}(t)(t + k) \quad (20)$$

Propensity is fed into recursive neural networks, derived from a Bayesian network architecture.

The adaptive update value evolves during the learning process based on its local sight on the error  $F$ , according to the following rule<sup>[14]</sup>:

$$\Delta_{ij}^{(t)} = \begin{cases} \eta^+ \times \Delta_{ij}^{(t-1)}, & \text{if } \frac{\partial F^{(t-1)}}{\partial w_{ij}} \times \frac{\partial F^{(t)}}{\partial w_{ij}} > 0 \\ \eta^- \times \Delta_{ij}^{(t-1)}, & \text{if } \frac{\partial F^{(t-1)}}{\partial w_{ij}} \times \frac{\partial F^{(t)}}{\partial w_{ij}} < 0 \\ \Delta_{ij}^{(t-1)}, & \text{else} \end{cases} \quad (21)$$

where  $0 < \eta^- < 1 < \eta^+$ . Adaptive backpropagation learning methods, by incorporating the effect of eleven weather parameters and past peak load information.

Kelley and Busemeyer<sup>[15]</sup> contrast three variants of rule-based and associative ‘neural network’ models with two variants of a Bayesian regression forecasting model. They find two primary results: first, there is evidence of the task independence of learning; and the most valid model is a neural network variant. However, if the standard or a large number of different clues relevant to predict the outcome in favor of Bayesian forecasting methods to provide reliable and effective prediction of human response.

Lauret, Fock, Randrianarivony and Manicom-Ramsamy<sup>[16]</sup> proposed the use of Bayesian techniques in order to design an optimal neural network based model for electric load forecasting. A neural network with  $d$  inputs,  $h$  hidden neurons and a linear output unit defines a non-linear parameterized mapping from an input  $x$  to an output  $y$  given by the following relationship:

$$y(x; \mathbf{w}) = \sum_{j=0}^h [w_j f(\sum_{i=0}^d w_{ji} x_i)] \quad (22)$$

The Bayesian approach to the simulation is superior to the traditional neural network learning methods.

Yu and Xu<sup>[17]</sup> proposed four machine learning algorithms which are naïve Bayesian, neural network, support vector machine and relevance vector machine. They develop each algorithm by changing the topologies of the networks to achieve its best possible predicted result. The experiments are performed based on different extracted feature size and training set size. Results show that neural network classifier is unsuitable for using alone as a spam rejection tool.

## V. ENSEMBLE LEARNING

Turning points prediction has long been a tough task in the field of time series analysis due to its strong nonlinearity. In [18], Li, Deng and Luo presented the turning points prediction framework. Further employed to develop a new trading strategy designing approach to financial investment. It is a machine learning-based solution incorporating chaotic dynamic analysis and neural network modeling. They present an ensemble learning based turning points prediction framework. A genetic algorithm based threshold

optimization procedure is described with a novel performance measure. They proposed a trading strategy designing approach based on the turning points prediction framework.

The results show that the integrated use of neural network classification of land cover routine multi-spectral satellite data can lead to significant improvement in classification accuracy. In order to overcome the shortcomings of long-term training time for each set of the network, the network is an effective training, Kalman filter algorithm. According to statistical hypothesis testing, multi-spectral image classification performance compared to the maximum likelihood method and Support Vector Machine classifier. Good generalization accuracy and the calculation by the time the order of 1 hour or less.<sup>[19]</sup>

Minku and Ludermir<sup>[20]</sup> introduced as a way to build clusters and co-evolution of neural network ensembles. This approach has created a neural network ensemble in an innovative manner, through a clear division of the input space through the clustering method. Clustering method allows to reduce the number of nodes constitute a set of neural network, thereby reducing the execution time of the learning process. This is an important feature especially in the use of evolutionary algorithms. The clustering method can also ensure that different neural networks to focus on different parts of the input space, the work of the divide and conquer approach in order to protect and improve the accuracy. In addition, the clustering method will help understand the system so that a simple distributed implementation possible.

Job completion time prediction is a critical task. Chen<sup>[21]</sup> presented the following model:

$$\min S \quad (23)$$

$$s.t. J_m = \sum_{k=1}^K \sum_{i=1}^n \mu_{i(k)}^m e_{i(k)}^2 \quad (24)$$

$$e_{\min}^2 = \min_{p \neq q} (\sum_{all j} (\bar{x}_{(p)j} - \bar{x}_{(q)j})^2), \quad (25)$$

$$S = \frac{J_m}{n \times e_{\min}^2}, \quad (26)$$

$$K \in Z^+. \quad (27)$$

where  $K$  is the required number of categories,  $n$  is the number of examples,  $\mu_{i(k)}$  represents the membership of example  $i$  belonging to category  $k$ ,  $e_{i(k)}$  measures the distance from example  $i$  to the centroid of category  $k$ ,  $m > 1$  is a parameter to increase or decrease the fuzziness.

In [22], a multistage non-linear radial basis function neural network ensemble forecasting model, the rate of foreign exchange forecasts. In the process of integration of modeling, the first phase will have a large number of single radial basis function neural network model. In the second phase, a conditional generalized minimum variance method is used to select the appropriate members of the orchestra. In the final phase, a radial basis function networks for the purpose of neural network ensemble forecasting. The experimental results show that forecasts the use of the method is still superior to other methods to use in this study in the same measurement.

In [23], an empirical mode decomposition of a neural network ensemble learning paradigm to the world crude oil spot price forecast. To this end, the spot price of crude oil was first broken down into a series of limited, and often is a small country, a number of intrinsic mode functions. And then the 3-layer feed forward neural network model, with model features inherent in each extract so that the inherent tendency of these models can be accurately predicted. Finally, all models predict the outcome of the functions inherent in the combination of adaptive linear neural networks, the development of a band of the original output of crude oil price series. Verification and testing, the two main crude oil price series, west Texas Intermediate spot price of crude oil and Brent crude spot price, is used to test the validity of empirical mode decomposition of the neural network ensemble learning methods. Empirical results show that the attractiveness of the proposed empirical mode decomposition of the neural network ensemble learning paradigm.

## VI. CONCLUSION

In this paper, we have tried to review on the applications of neural networks to different business intelligence, financial engineering, etc. In the near future, we believe that the researchers will benefit from the use of the recent advances in neural networks, evolutionary algorithms, cellular computing, Bayesian learning and ensemble learning, and their combinations. Neural networks will certainly be one of the key technologies for business intelligence and financial engineering. It can be concluded that, the future of neural networks lies in their use in conjunction with other advanced technologies.

## REFERENCES

- [1] F.J. Martínez-Estudillo, C. Hervás-Martínez, P.A. Gutiérrez, A.C. Martínez-Estudillo, "Evolutionary product-unit neural networks classifiers," *Neurocomputing*, Vol. 72, Dec. 2008, pp. 548-561.
- [2] Cheng Bo, Bai Zhifeng, Cao Binggang, "State of charge estimation based on evolutionary neural network," *Energy Conversion and Management*, Vol. 49, Oct. 2008, pp. 2788-2794.
- [3] C. Hervás-Martínez, F.J. Martínez-Estudillo, M. Carbonero-Ruz, "Multilogistic regression by means of evolutionary product-unit neural networks," *Neural Networks*, Vol. 21, Sep. 2008, pp. 951-961.
- [4] Kin-Fan Au, Tsan-Ming Choi, Yong Yu, "Fashion retail forecasting by evolutionary neural networks," *International Journal of Production Economics*, vol. 114, Aug. 2008, pp. 615-630.
- [5] S. Gholizadeh, E. Salajegheh, P. Torkzadeh, "Structural optimization with frequency constraints by genetic algorithm using wavelet radial basis function neural network," *Journal of Sound and Vibration*, Vol. 312, Apr. 2008, pp. 316-331.
- [6] J. Javier Martínez, F. Javier Toledo, Eduardo Fernández, José M. Ferrández, "A retinomorph architecture based on discrete-time cellular neural networks using reconfigurable computing," *Neurocomputing*, Vol. 71, Jan. 2008, pp. 766-775.
- [7] A. Muhittin Albora, Osman N. Ucan, Atilla Ozmen, Tulay Ozkan, "Separation of Bouguer anomaly map using cellular neural network," *Journal of Applied Geophysics*, Vol. 46, Feb. 2001, pp. 129-142.
- [8] L. Bertucco, M. Coltelli, G. Nunnari, L. Occhipinti, "Cellular neural networks for real-time monitoring of volcanic activity," *Computers & Geosciences*, Vol. 25, Mar. 1999, pp. 101-117.
- [9] N. K. Gupta, Hidemitsu Ogawa, "Optoelectronic architecture for cellular neural networks based on a new bias method and spatial coding scheme," *Optics Communications*, Vol. 138, May 1997, pp. 11-15.
- [10] Geza Toth, Craig S. Lent, P. Douglas Tougaw, Yuriy Brazhnik, Weiwen Weng, Wolfgang Porod, Ruey-Wen Liu, Yih-Fang Huang, "Quantum cellular neural networks," *Superlattices and Microstructures*, Vol. 20, Dec. 1996, pp. 473-478.
- [11] L. Fortuna, G. Manganaro, G. Muscato, G. Nunnari, "Parallel simulation of cellular neural networks," *Computers & Electrical Engineering*, Vol. 22, Jan. 1996, pp. 61-84.
- [12] V. Breton, H. Fonvieille, P. Grenier, C. Guicheney, J. Jousset, Y. Roblin, F. Tamin, "Application of neural networks and cellular automata to interpretation of calorimeter data," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, Vol. 362, Aug. 1995, pp. 478-486.
- [13] Lin Wu, Pierre Baldi, "Learning to play Go using recursive neural networks," *Neural Networks*, Vol. 21, Nov. 2008, pp. 1392-1400.
- [14] Lalit Mohan Saini, "Peak load forecasting using Bayesian regularization, Resilient and adaptive backpropagation learning based artificial neural networks," *Electric Power Systems Research*, Vol. 78, Jul. 2008, pp. 1302-1310.
- [15] Hugh Kelley, Jerome Busemeyer, "A comparison of models for learning how to dynamically integrate multiple cues in order to forecast continuous criteria," *Journal of Mathematical Psychology*, Vol. 52, Jun. 2008, pp. 218-240.
- [16] Philippe Lauret, Eric Fock, Rija N. Randrianarivony, Jean-François Manicom-Ramsamy, "Bayesian neural network approach to short time load forecasting," *Energy Conversion and Management*, Vol. 49, May 2008, pp. 1156-1166.
- [17] Bo Yu, Zong-ben Xu, "A comparative study for content-based dynamic spam classification using four machine learning algorithms," *Knowledge-Based Systems*, Vol. 21, May 2008, pp. 355-362.
- [18] Xiuquan Li, Zhidong Deng, Jing Luo, "Trading strategy design in financial investment through a turning points prediction scheme," *Expert Systems with Applications*, Vol. 36, May 2009, pp. 7818-7826.
- [19] Morton J. Canty, "Boosting a fast neural network for supervised land cover classification," *Computers & Geosciences*, December 2008, in press.
- [20] Fernanda L. Minku, Teresa B. Ludermir, "Clustering and co-evolution to construct neural network ensembles: An experimental study," *Neural Networks*, Vol. 21, Nov. 2008, pp. 1363-1379.
- [21] Toly Chen, "A hybrid fuzzy-neural approach to job completion time prediction in a semiconductor fabrication factory," *Neurocomputing*, Vol. 71, Oct. 2008, pp. 3193-3201.
- [22] Lean Yu, Kin Keung Lai, Shouyang Wang, "Multistage RBF neural network ensemble learning for exchange rates forecasting," *Neurocomputing*, Vol. 71, Oct. 2008, pp. 3295-3302.
- [23] Lean Yu, Shouyang Wang, Kin Keung Lai, "Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm," *Energy Economics*, Vol. 30, Sep. 2008, pp. 2623-2635.