

Optimization of Ensemble Neural Networks with Type-1 and Interval Type-2 Fuzzy Integration for Forecasting the Taiwan Stock Exchange

Martha Pulido, Patricia Melin and Olivia Mendoza

Abstract This paper describes an optimization method based on particle swarm optimization for ensemble neural networks with type-1 and type-2 fuzzy aggregation for forecasting complex time series. The time series that was considered in this paper to compare the hybrid approach with traditional methods is the Taiwan Stock Exchange (TAIEX), and the results shown are for the optimization of the structure of the ensemble neural network with type-1 and type-2 fuzzy integration. Simulation results show that ensemble approach produces good prediction of the Taiwan Stock Exchange.

Keywords Ensemble neural networks · Time series · Particle swarm · Fuzzy system

1 Introduction

Time series are usually analyzed to understand the past and to predict the future, enabling managers or policy makers to make properly informed decisions. Time series analysis quantifies the main features in data, like the random variation. These facts, combined with improved computing power, have made time series methods widely applicable in government, industry, and commerce. In most branches of science, engineering, and commerce, there are variables measured sequentially in time. Reserve banks record interest rates and exchange rates each day. The government statistics department will compute the country's gross domestic product on a yearly basis. Newspapers publish yesterday's noon temperatures for capital cities from around the world. Meteorological offices record rainfall at many different sites with differing resolutions. When a variable is measured sequentially in time over or at a fixed interval, known as the sampling interval, the resulting data form a time series [1].

M. Pulido · P. Melin (✉) · O. Mendoza
Tijuana Institute of Technology, Tijuana, México, USA
e-mail: pmelin@tectijuana.edu.mx; pmelin@tectijuana.mx

Time series predictions are very important because based on them we can analyze past events to know the possible behavior of futures events and thus can take preventive or corrective decisions to help avoid unwanted circumstances.

The choice and implementation of an appropriate method for prediction has always been a major issue for enterprises that seek to ensure the profitability and survival of business. The predictions give the company the ability to make decisions in the medium and long term, and due to the accuracy or inaccuracy of data this could mean predicted growth or profits and financial losses. It is very important for companies to know the behavior that will be the future development of their business, and thus be able to make decisions that improve the company's activities, and avoid unwanted situations, which in some cases can lead to the company's failure. In this paper we propose a hybrid approach for time series prediction by using an ensemble neural network and its with optimization with particle swarm optimization. In the literature there have been recent produced work of time series [2–10].

2 Preliminaries

In this section we present basic concepts that are used in this proposed method:

2.1 *Time Series and Prediction*

The word “prediction” comes from the Latin prognosticum, which means I know in advance. Prediction is to issue a statement about what is likely to happen in the future, based on analysis and considerations of experiments. Making a forecast is to obtain knowledge about uncertain events that are important in decision-making [6]. Time series prediction tries to predict the future based on past data, it take a series of real data $x_t - n, \dots, x_t - 2, 0, x_t - 1, x_t$ and then obtains the prediction of the data $x_t + 1, x_t + 2, \dots, x_t + n$. The goal of time series prediction or a model is to observe the series of real data, so that future data may be accurately predicted [1, 11].

2.2 *Neural Networks*

Neural networks Neural networks (NNs) are composed of many elements (Artificial Neurons), grouped into layers and are highly interconnected (with the synapses), this structure has several inputs and outputs, which are trained to react (or give values) in a way you want to input stimuli. These systems emulate in some way, the human brain. Neural networks are required to learn to behave (Learning) and

someone should be responsible for the teaching or training (Training), based on prior knowledge of the environment problem [12, 13].

2.3 Ensemble Neural Networks

An Ensemble Neural Network is a learning paradigm where many neural networks are jointly used to solve a problem [14]. A Neural network ensemble is a learning paradigm where a collection of a finite number of neural networks is trained for the same task [15]. It originates from Hansen and Salamon's work [16], which shows that the generalization ability of a neural network system can be significantly improved through ensembling a number of neural networks, i.e. training many neural networks and then combining their predictions. Since this technology behaves remarkably well, recently it has become a very hot topic in both neural networks and machine learning communities [17], and has already been successfully applied to diverse areas such as face recognition [18, 19], optical character recognition [20–22], scientific image analysis [23], medical diagnosis [24, 25], seismic signals classification [26], etc.

In general, a neural network ensemble is constructed in two steps, i.e. training a number of component neural networks and then combining the component predictions.

There are also many other approaches for training the component neural networks. Examples are as follows. Hampshire and Waibel [22] utilize different object functions to train distinct component neural networks.

2.4 Fuzzy Systems as Methods of Integration

There exists a diversity of methods of integration or aggregation of information, and we mention some of these methods below.

Fuzzy logic was proposed for the first time in the mid-sixties at the University of California Berkeley by the brilliant engineer Lofty A. Zadeh., who proposed what it's called the principle of incompatibility: "As the complexity of system increases, our ability to be precise instructions and build on their behavior decreases to the threshold beyond which the accuracy and meaning are mutually exclusive characteristics." Then introduced the concept of a fuzzy set, under which lies the idea that the elements on which to build human thinking are not numbers but linguistic labels. Fuzzy logic can represent the common knowledge as a form of language that is mostly qualitative and not necessarily a quantity in a mathematical language that means of fuzzy set theory and function characteristics associated with them [12].

2.5 Optimization

The process of optimization is the process of obtaining the ‘best’, if it is possible to measure and change what is ‘good’ or ‘bad’. In practice, one wishes the ‘most’ or ‘maximum’ (e.g., salary) or the ‘least’ or ‘minimum’ (e.g., expenses). Therefore, the word ‘optimum’ is takes the meaning of ‘maximum’ or ‘minimum’ depending on the circumstances; ‘optimum’ is a technical term which implies quantitative measurement and is a stronger word than ‘best’ which is more appropriate for everyday use. Likewise, the word ‘optimize’, which means to achieve an optimum, is a stronger word than ‘improve’. Optimization theory is the branch of mathematics encompassing the quantitative study of optima and methods for finding them. Optimization practice, on the other hand, is the collection of techniques, methods, procedures, and algorithms that can be used to find the optima [27].

2.6 Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a bio-inspired optimization method proposed by R. Eberhart and J. Kennedy [28] in 1995. PSO is a search algorithm based on the behavior of biological communities that exhibits individual and social behavior [29], and examples of these communities are groups of birds, schools of fish and swarms of bees [29].

A PSO algorithm maintains a swarm of particles, where each particle represents a potential solution. In analogy with the paradigms of evolutionary computation, a swarm is similar to a population, while a particle is similar to an individual. In simple terms, the particles are “flown” through a multidimensional search space, where the position of each particle is adjusted according to its own experience and that of its neighbors. Let x_i denote the position i in the search space at time step t , unless otherwise stated, t denotes discrete time steps. The position of the particle is changed by adding a velocity, $v_i(t)$, to the current position, i.e.

$$\begin{aligned} x_i(t+1) &= x_i(t) + v_i(t+1) \\ \text{with } x_i(0) &\sim U(X_{min}, X_{max}). \end{aligned} \tag{1}$$

3 Problem Statement and Proposed Method

The goal of this work was to implement Particle Swarm Optimization to optimize the ensemble neural network architectures. In this cases the optimization is for each of the modules, and thus to find a neural network architecture that yields optimum results in each of the Time Series to be considered. In Fig. 1 we have the historical data of each time series prediction, then the data is provided to the modules that will

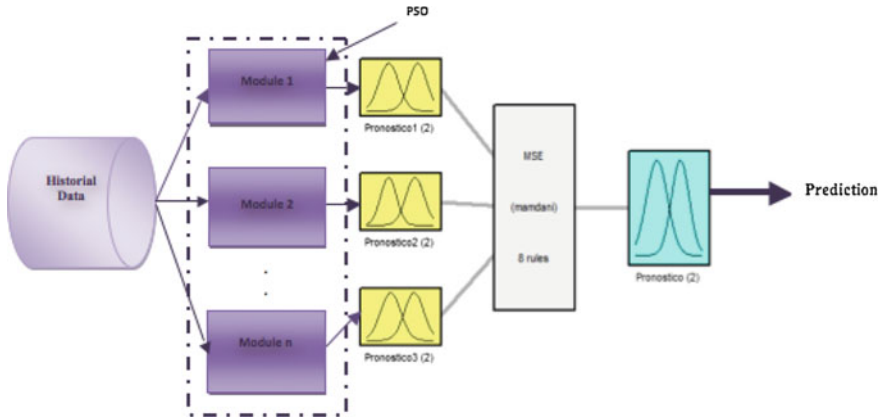


Fig. 1 General architecture of the proposed ensemble model

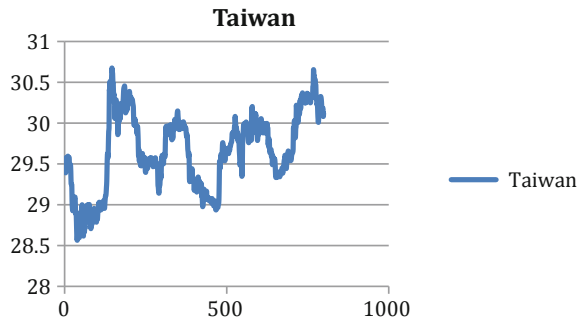
be optimized with the particle swarm optimization for the ensemble network, and then these modules are integrated with integration based on type-1 and type-2 Fuzzy Integration.

Historical data of the Taiwan Stock Exchange time series was used for the ensemble neural network trainings, where each module was fed with the same information, unlike the modular networks, where each module is fed with different data, which leads to architectures that are not uniform.

The Taiwan Stock Exchange (Taiwan Stock Exchange Corporation) is a financial institution that was founded in 1961 in Taipei and began to operate as stock exchange on 9 February 1962. The Financial Supervisory Commission regulates it. The index of the Taiwan Stock Exchange is the TWSE [30].

Data of the Taiwan Stock Exchange time series: We are using 800 points that correspond to a period from 03/04/2011 to 05/07/2014 (as shown in Fig. 2). We used 70% of the data for the ensemble neural network trainings and 30% to test the network [30].

Fig. 2 Taiwan Stock Exchange



Number of Modules	Number of Layers	Neurons 1	...	Neurons n
-------------------	------------------	-----------	-----	-----------

Fig. 3 Particle structure to optimize the ensemble neural network

The objective function is defined to minimize the prediction error as follows:

$$EM = \left(\sum_{i=1}^D |a_i - x_i| / D \right) \quad (2)$$

where \mathbf{a} , corresponds to the predicted data depending on the output of the network modules, \mathbf{X} represents real data, \mathbf{D} the Number of Data points and EM is the total prediction error.

The corresponding particle structure is shown in Fig. 3.

Figure 3 represents the Particle Structure to optimize the ensemble neural network, where the parameters that are optimized are the number of modules, number of layers, and number of neurons of the ensemble neural network. PSO determines the number of modules, number of layers and number of neurons per layer that the neural network ensemble should have, to meet the objective of achieving the better Prediction error.

The parameters for the particle swarm optimization algorithm are: 100 Particles, 100 iterations, Cognitive Component (C1) = 2, Social Component (C2) = 2, Constriction coefficient of linear increase (C) = (0–0.9) and Inertia weight with linear decrease (W) = (0.9–0). We consider a number of 1–5 modules, number of layers of 1–3 and neurons number from 1 to 30.

The aggregation of the responses of the optimized ensemble neural network is performed with type-1 and type-2 fuzzy systems. In this work the fuzzy system consists of 5 inputs depending on the number of modules of the neural network ensemble and one output is used. Each input and output linguistic variable of the fuzzy system uses 2 Gaussian membership functions. The performance of the type-2 fuzzy aggregators is analyzed under different levels of uncertainty to find out the best design of the membership functions for the 32 rules of the fuzzy system. Previous tests have been performed only with a three input fuzzy system and the fuzzy system changes according to the responses of the neural network to give us better prediction error. In the type-2 fuzzy system we also change the levels of uncertainty to obtain the best prediction error.

Figure 4 shows a fuzzy system consisting of 5 inputs depending on the number of modules of the neural network ensemble and one output. Each input and output linguistic variable of the fuzzy system uses 2 Gaussian membership functions. The performance of the type-2 fuzzy aggregators is analyzed under different levels of uncertainty to find out the best design of the membership functions for the 32 rules of the fuzzy system. Previous experiments were performed with triangular, and Gaussian and the Gaussian produced the best results of the prediction.

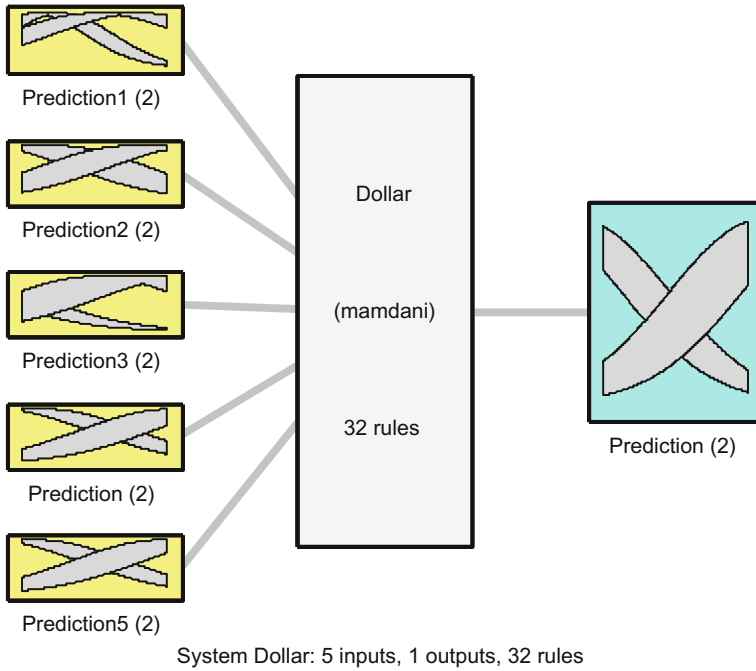


Fig. 4 Fuzzy inference system for integration of the ensemble neural network

Figure 5 represents the 32 possible rules of the fuzzy system; we have 5 inputs in the fuzzy system with 2 membership functions, and the outputs with 2 membership functions. These fuzzy rules are used for both the type-1 and type-2 fuzzy systems. In previous work several tests were performed with 3 inputs, and the prediction error obtained was significant and the number of rules was greater, and this is why we changed to 2 inputs.

4 Simulation Results

In this section we present the simulation results obtained with the genetic algorithm and particle swarm optimization for the Taiwan Stock Exchange.

We consider working with a genetic algorithm to optimize the structure of an ensemble neural network and the best architecture obtained was the following (shown in Fig. 6).

In this architecture we have two layers in each module. In module 1, in the first layer we have 23 neurons and the second layer we have 9 neurons, and In module 2 we used 9 neurons in the first layer and the second layer we have 15 neurons the

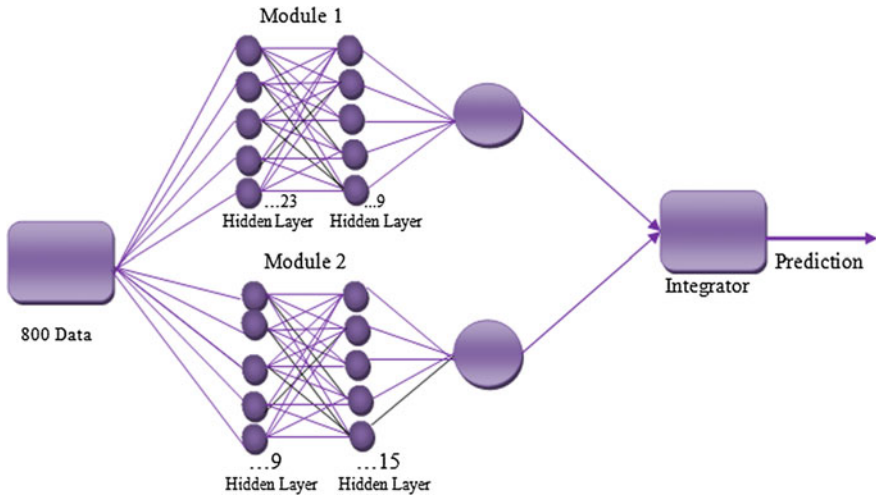


Fig. 6 Prediction with the optimized ensemble neural network with GA of the TAIEX

Table 1 Particle swarm optimization result for the ensemble neural network

No.	Iterations	Particles	Number of modules	Number of layers	Number of neurons	Duration	Prediction error
1	100	100	2	3	13, 16, 2 18, 20, 18	01:48:30	0.002147
2	100	100	2	2	3, 9 14, 19	01:03:09	0.0021653
3	100	100	2	2	20, 4 10, 7	01:21:02	0.0024006
4	100	100	2	2	16, 19 3, 12	01:29:02	0.0019454
5	100	100	2	2	19, 19 24, 17	02:20:22	0.0024575
6	100	100	2	3	21, 14, 23 14, 24, 20	01:21:07	0.0018404
7	100	100	2	2	23, 9 9, 15	01:19:08	0.0013065
8	100	100	2	2	15, 17 9, 22	01:13:20	0.0018956
9	100	100	2	2	20, 16	01:13:35	0.0023377
10	100	100	2	2	23, 8 10, 17	01:04:23	0.0023204

Fuzzy integration is performed initially by implementing a type-1 fuzzy system in which the best result is in experiment of row number 8 of Table 2 with an error of: 0.0235.

Table 2 Result PSO for the type-1 fuzzy integration of the TAIEX

Experiments	Prediction error with fuzzy integration type-1
Experiment 1	0.0473
Experiment 2	0.0422
Experiment 3	0.0442
Experiment 4	0.0981
Experiment 5	0.0253
Experiment 6	0.0253
Experiment 7	0.0253
Experiment 8	0.0235
Experiment 9	0.0253
Experiment 10	0.0253

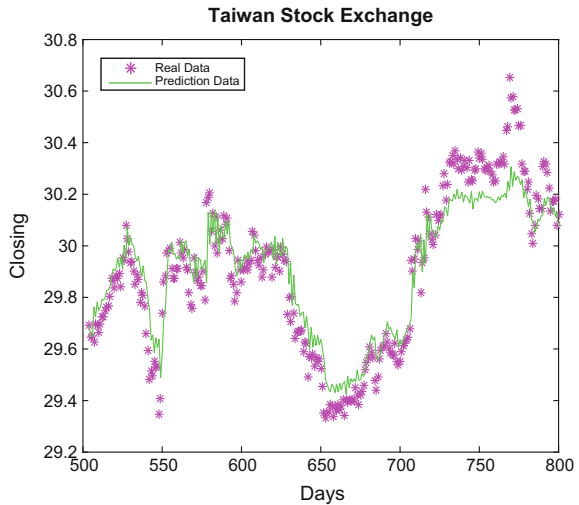
Table 3 Result PSO for the type-2 fuzzy integration of the TAIEX

Experiment	Prediction error 0.3 uncertainty	Prediction error 0.4 uncertainty	Prediction error 0.5 uncertainty
Experiment 1	0.0335	0.033	0.0372
Experiment 2	0.0299	0.5494	0.01968
Experiment 3	0.0382	0.0382	0.0387
Experiment 4	0.0197	0.0222	0.0243
Experiment 5	0.0433	0.0435	0.0488
Experiment 6	0.0121	0.0119	0.0131
Experiment 7	0.01098	0.01122	0.01244
Experiment 8	0.0387	0.0277	0.0368
Experiment 9	0.0435	0.0499	0.0485
Experiment 10	0.0227	0.0229	0.0239

As a second phase, to integrate the results of the optimized ensemble neural network a type-2 fuzzy system is implemented, where the best results that are obtained are as follows: with a degree uncertainty of 0.3 a forecast error of 0.01098 is obtained, with a degree of uncertainty of 0.4 the error is of 0.01122 and with a degree of uncertainty of 0.5 the error is of 0.001244, as shown in Table 3.

Figure 7 shows the plot of real data against the predicted data generated by the ensemble neural network optimized with the particle swarm optimization.

Fig. 7 Prediction with the optimized ensemble neural network with PSO of the TAIEX



5 Conclusions

The best result when applying the particle swarm to optimize the ensemble neural network was: 0.0013066 (as shown in Fig. 6 and Table 1). Implemented a type 2 fuzzy system for ensemble neural network, in which the results where for the best evolution as obtained a degree of uncertainty of 0.3 yielded a forecast error of 0.01098, with an 0.4 uncertainty error: 0.01122, and 0.5 uncertainty error of 0.01244, as shown in Table 3. After achieving these results, we have verified efficiency of the algorithms applied to optimize the neural network ensemble architecture. In this case, the method was efficient but it also has certain disadvantages, sometimes the results are not as good, but genetic algorithms can be considered as good technique a for solving search and optimization problems.

Acknowledgements We would like to express our gratitude to the CONACYT, Tijuana Institute of Technology for the facilities and resources granted for the development of this research.

References

1. Cowpertwait, P., Metcalfe, A.: Time Series. In: Introductory Time Series with R, pp. 2–5. Springer, Heidelberg (2009)
2. Castillo, O., Melin, P.: Hybrid intelligent systems for time series prediction using neural networks, fuzzy logic, and fractal theory. *IEEE Trans. Neural Netw.* **13**(6), 1395–1408 (2002)
3. Castillo, O., Melin, P.: Simulation and forecasting complex economic time series using neural networks and fuzzy logic. In: *Proceeding of the International Neural Networks Conference*, vol. 3, pp. 1805–1810 (2001)

4. Castillo, O., Melin, P.: Simulation and forecasting complex financial time series using neural networks and fuzzy logic. In: Proceedings the IEEE the International Conference on Systems, Man and Cybernetics, vol. 4, pp. 2664–2669 (2001)
5. Karnik, N., Mendel, M.: Applications of type-2 fuzzy logic systems to forecasting of time-series. *Inf. Sci.* **120**(1–4), 89–111 (1999)
6. Kehagias, A., Petridis, V.: Predictive modular neural networks for time series classification. *Neural Netw.* **10**(1), 31–49 (2000), 245–250 (1997)
7. Maguire, L.P., Roche, B., McGinnity, T.M., McDaid, L.J.: Predicting a chaotic time series using a fuzzy neural network. *Inf. Sci.* **112**(1–4), 125–136 (1998)
8. Melin, P., Castillo, O., Gonzalez, S., Cota, J., Trujillo, W., Osuna, P.: Design of Modular Neural Networks with Fuzzy Integration Applied to Time Series Prediction, vol. 41/2007, pp. 265–273. Springer, Heidelberg (2007)
9. Yadav, R.N., Kalra, P.K., John, J.: Time series prediction with single multiplicative neuron model, soft computing for time series prediction. *Appl. Soft Comput.* **7**(4), 1157–1163 (2007)
10. Zhao, L., Yang, Y.: PSO-based single multiplicative neuron model for time series prediction. *Expert Syst. Appl. Part 2* **36**(2), 2805–2812 (2009)
11. Brockwell, P.T., Davis, R.A.: Introduction to Time Series and Forecasting. Springer, New York, pp 1–219 (2002)
12. Jang, J.S.R., Sun, C.T., Mizutani, E.: Neuro-Fuzzy and Soft Computing. Prentice Hall (1996)
13. Multaba, I.M., Hussain, M.A.: Application of neural networks and other learning. In: Technologies in Process Engineering. Imperial Collage Press (2001)
14. Sharkey, A.: Combining Artificial Neural Nets: Ensemble And Modular Multi-net Systems. Springer, London (1999)
15. Sollich, P., Krogh, A.: Learning with ensembles: how over-fitting can be useful. In: Touretzky, D.S., Mozer, M.C., Hasselmo, M.E. (eds.) Advances in Neural Information Processing Systems, Denver, CO, vol. 8, pp. 190–196. MIT Press, Cambridge, MA (1996)
16. Hansen, L.K., Salomon, P.: Neural network ensembles. *IEEE Trans. Pattern Anal. Mach. Intell.* **12**(10), 993–1001 (1990)
17. Sharkey, A.: One combining Artificial of Neural Nets. Department of Computer Science University of Sheffield, U.K. (1996)
18. Gutta, S., Wechsler, H.: Face recognition using hybrid classifier systems. In: Proceedings of the ICNN-96, Washington, DC, pp. 1017–1022. IEEE Computer Society Press, Los Alamitos, CA (1996)
19. Huang, F.J., Huang, Z., Zhang, H.-J., Chen, T.H.: Pose invariant face recognition. In: Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition, Grenoble, France. IEEE Computer Society Press, Los Alamitos, CA (2000)
20. Drucker, H., Schapire, R., Simard, P.: Improving performance in neural networks using a boosting algorithm. In: Hanson, S.J., Cowan Giles, J.D. (eds.) Advances in Neural Information Processing Systems, Denver, CO, vol. 5, pp. 42–49. Morgan Kaufmann, San Mateo, CA (1993)
21. Hampshire, J., Waibel, A.: A novel objective function for improved phoneme recognition using time- delay neural networks. *IEEE Trans. Neural Netw.* **1**(2), 216–228 (1990)
22. Mao, J.: A case study on bagging, boosting and basic ensembles of neural networks for OCR. In: Proceedings of the IJCNN-98, Anchorage, AK, vol. 3, pp. 1828–1833. IEEE Computer Society Press, Los Alamitos, CA (1998)
23. Cherkauer, K.J.: Human expert level performance on a scientific image analysis task by a system using combined artificial neural networks. In: Chan, P., Stolfo, S., Wolpert, D. (eds.) Proceedings of the AAAI-96 Workshop on Integrating Multiple Learned Models for Improving and Scaling Machine Learning Algorithms, Portland, OR, AAAI, pp. 15–21. Press, Menlo Park, CA (1996)
24. Cunningham, P., Carney, J., Jacob, S.: Stability problems with artificial neural networks and the ensemble solution. *Artif. Intell. Med.* **20**(3), 217–225 (2000)
25. Zhou, Z.-H., Jiang, Y., Yang, Y.-B., Chen, S.-F.: Lung cancer cell identification based on artificial neural network ensembles. *Artif. Intell. Med.* **24**(1), 25–36 (2002)

26. Shimshon, Y.N.: Intrator classification of seismic signal by integrating ensemble of neural networks. *IEEE Trans. Signal Process.* **461**(5), 1194–1201 (1998)
27. Antoniou, A., Sheng, W. (eds.): Practical optimization algorithms and engineering applications. In: *Introduction Optimization*. Springer, pp. 1–4 (2007)
28. Eberhart, R., Kennedy, J.: A new optimizer using swarm theory. In: *Proceedings of the 6th International Symposium Micro Machine and Human Science (MHS)*, pp. 39–43, October 1995
29. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: *Proceedings of the IEEE International Conference Neural Network (ICNN)*, Nov. 1995, vol. 4, pp. 1942–1948
30. Taiwan Bank Database: www.twse.com.tw/en (April 03, 2011)