2009 World Congress on Nature and Biologically Inspired Computing, NABIC 2009; Coimbatore; 9 December 2009 through 11 December 2009; Category number CFP0995H; Code 79534; Article number 5393728, Pages 312-317; DOI- http://dx.doi.org/10.1109/NABIC.2009.5393728

# Robust Prediction of Stock Indices using PSO based Adaptive Linear Combiner

Ritanjali Majhi, G. Panda, *Senior Member IEEE* and Babita Majhi

***Abstract*— The present paper employs a particle swarm optimization (PSO) based adaptive linear combiner for efficient prediction of various stock indices in presence of strong outliers in the training data. The connecting weights of the model are updated by minimizing the Wilcoxon norm of the error vector by PSO. The short and long term prediction performance of the new model is evaluated with test data and the results obtained are compared with those obtained from the conventional PSO based model. It is in general observed that the proposed model is computationally more efficient, prediction wise more accurate and more robust against outliers in training set compared to those obtained by standard PSO based model.**

## I. Introduction

With the introduction of online trading the stock market has become a forum where the small investors can also earn good profits. In the present era large number of transactions are done in stock market for selling and purchasing shares. Due to large applications in different business transactions, stock market prediction has become a hot topic of research. As the stock market data is non-stationary and volatile the investors feel insecure during investing. In the recent years lot of attention has been devoted to the analysis and prediction of future values and trends of the financial markets. The stock market is affected by many complex events such as business cycles, monetary policies, interest rates and political situations. Due to volatility and non-stationary characteristics of stock indices data it is difficult to build an accurate forecasting model. But even then different financial forecasting methods have been proposed in the literature each of which has its own merits and limitations.

Many practical time series data like the financial data are at times wrongly recorded and hence may be treated as outliers. Outliers are observations which are distinctly different from the remaining data set. It has been observed that the outliers present in the past data used as training set may be as high as 40% [10]. Further while collecting and recording data, the location of and magnitude of corruption by such outliers are also not known. Depending upon the location and magnitude, the outliers may have moderate to severe effects on the performance of an prediction model. In the conventional adaptive forecasting models gradient based learning techniques like the least mean square (LMS) [1] and back propagation (BP) [2] algorithms are used. Such learning algorithms are gradient search type and have been derived using square of the output error as the cost function. It has been reported that update algorithms derived from such simple cost function are not robust against the outliers in the training data. In other words the resulting models developed using corrupted training sets posess poor prediction capability. To improve the forecasting performance in presence of noisy training set choice of suitable norm as cost function is important. In the literature the Wilcoxon norm based regressor has been shown to be insensitive to outliers [3].

In recent years many research papers have appeared in the literature using evolutionary computing tools such as genetic algorithm(GA), particle swarm optimization (PSO), bacterial foraging optimization (BFO) and genetic programming(GP) in developing forecasting models. The stock data are highly time-invariant and are highly influenced by indeterminate dealing. There have been several applications of evolutionary computation tools to portfolio optimization, bankruptcy prediction, financial forecasting, fraud detection and scheduling. GA has been used in developing a forecasting model for portfolio decisions [4]. In a recent paper [5] an efficient investment strategy in portfolio management using GA has been proposed. Bhattacharya et al [6] have developed a GP based trading model for efficient prediction of exchange rates. In another work Chen et al. [7] used GP to derive option pricing formulas using real data from S&P 500 index options for training and testing purpose. The Bacterial Foraging Optimization (BFO) and Particle Swarm Optimization (PSO) techniques are used to develop an efficient forecasting model for prediction of various stock indices in [8] and [9] respectively. The connecting weights of the adaptive linear combiner based model are optimized by the BFO and PSO so that its mean square error(MSE) is minimized.

A new neural network learning machine has been proposed using Wilcoxon norm [10] and has recently been successfully applied for function optimization task in presence of outliers in training samples. Being motivated by this concept, in this paper we propose an adaptive linear combiner model whose weights are updated by minimizing the Wilcoxon norm using PSO based approach. It is expected that the minimization of the Wilcoxon norm of the error vector will yield robust and efficient prediction of stock indices To validate the robustness of the proposed

Ritanjali Majhi is with Centre of Management Studies, National Institute of Technology , Warangal, INDIA (email : ritanjalimajhi@gmail.com)
G. Panda and Babita Majhi are with Department of Electronics & Communication Engineering, National Institute of Technology, Rourkela 769008,India,phone-+91-661-2462455(email: ganapati.panda@gmail.com/babita.majhi@gmail.com)

model, it is employed for short and long range predictions of S&P 500 and DJIA stock indices.

The present paper is organized as follows. Section II deals with the basic principle of particle swarm optimization (PSO). The Wilcoxon norm based linear combiner (WLC) model using PSO dealt in Section III. The PSO based Wilcoxon norm (W-norm) minimization is discussed in Section IV. The simulation study is carried out in Section V. Finally conclusion is given in Section VI.

## II. BASICS OF PARTICLE SWARM OPTIMIZATION

The Particle swarm optimization(PSO) is an evolutionary computation technique developed by Kennedy and Eberhrt [11]. It is motivated from the simulation of social behavior. In evolutionary computational algorithms evolutionary operators to manipulate the individuals are used. But in PSO these individual are evolved by cooperation and competition among the individuals themselves through generations. Each individual in PSO flies in the search space with a velocity which is dynamically adjusted according to its own as well as its companions' flying experiences. Each individual is treated as a volume-less particle in a D-dimensional space. The $i$th particle is represented as $X_i = [x_{i1}, x_{i2}, ............x_{iD}]^T$. The position giving the best previous value is called the best previous position of any particle and is represented as $P_i = [p_{i1}, p_{i2}, .......p_{iD}]^T$. The rate of change of position for $i$th particle is represented as $V_i = [v_{i1}, v_{i2}, .........v_{iD}]^T$. The global best position at a generation is given by $P_g = [p_{g1}, p_{g2}, .........., p_{gD}]^T$. The velocity and position of the $d$th element of $i$th particle are changed according to (1) and (2) respectively.

$$v_{id} = w * v_{id} + C_1 * rand_1 * (p_{id} - x_{id}) + C_2 * rand_2 * (p_{gd} - x_{id}) \quad (1)$$

$$x_{id} = x_{id} + v_{id} \quad (2)$$

where $C_1$ and $C_2$ are two positive constants where as $rand_1$ and $rand_2$ are two random functions in the range [0, 1]. The second part in (1) is the cognition part which represents the private thinking of the particle itself. The third part is the social part which represents the collaboration among the particles. The new velocity of a particle is calculated using (1) which is according to its previous velocity and the distances of current position from its own best experience(position) and the group's best experience. After that the particle flies towards a new position according to (2). The performance of each particle is evaluated based on a predefined fitness function which is problem dependent. Each particle keeps track of its coordinates in the problem space which are associated with the best solution(fitness) it has achieved so far and is denoted by $P_i$. The overall best location (global) obtained so far by the particles is also tracked by the optimizer and is represented by $P_g = [p_{g1}, p_{g2}, .........., p_{gD}]^T$.

The inertia weight, $w$ has characteristics that are reminiscent of temperature parameter in the simulated annealing. A large inertia weight facilitates a global search while a small inertia weight facilitates a local search. By linearly decreasing the inertia weight from a large value(close to unity) to a small value through the course of PSO run, the PSO tends to have more global search ability at the beginning of the run while possessing more local search ability near the end of the run.

.

## III. WILCOXON NORM BASED LINEAR COMPBINER (WLC) MODEL

Referring to Fig. 1, the adaptive linear combiner is essentially an adaptive finite impulse response (FIR) filter having number of inputs equal to the number of features in the input pattern derived from the stock market series. The weights of the combiner are considered as the particles and initially their values are set to random numbers. A population of such random particles is chosen to represent a set of initial solutions.

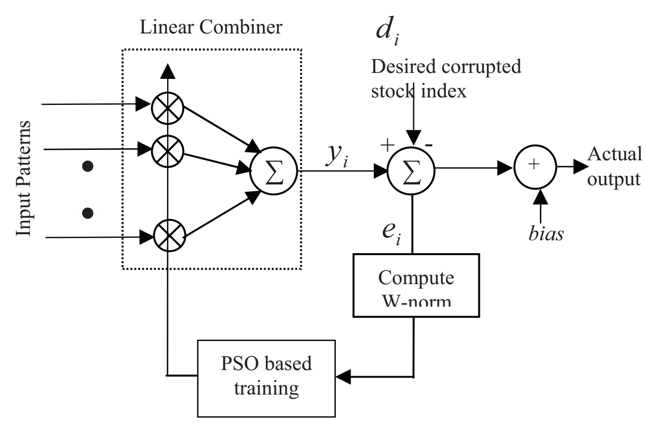


Fig. 1 A stock market forecasting model using adaptive linear combiner with PSO based Wilcoxon norm minimization

For every pattern of input the linear combiner generates the error $e(k)$. Using the error vector, the Wilcoxon norm (W-norm) which is robust against outlier, is computed. Basically the development of the WLC model corresponds to iterative minimization of the cost function, W-norm by changing its weights. This optimization task is accomplished by PSO technique.

## IV. PSO BASED ALGORITHM FOR W-NORM MINIMIZATION OF WLC MODEL

(i) The coefficients or weights of the model are initially chosen from a swarm of $I$ particles (birds). Each particle constitutes $D$ number of parameters and each parameter represents a weight of the linear combiner.

(ii) $K$ number of patterns each containing ten features are obtained from the past stock indices.

(iii) Each of the $K$ patterns are passed through the linear combiner, multiplied with the weights and the partial sums are added together to give $y_i$, where

$$y_i = \sum_{n=1}^{N} w_n x_n \tag{3}$$

The output of the linear combiner, $y_i$ is then compared with corresponding normalized corrupted desired stock index, $d_i$ to produce the error, $e_i$. At the completion of all patterns $K$ errors are produced.

(iv) The Wilcoxon norm (W-norm) of an error vector, $\bar{e}$ is defined using a score function. A score function is a nondecreasing function $\varphi$ which satisfies

$$\int_0^1 \varphi^2(u)du < \infty \tag{4}$$

The score associated with the score function $\varphi$ is given by

$$a(i) = \varphi\left(\frac{i}{i+1}\right), \quad i \in \underline{l} \tag{5}$$

where $l$ is a fixed positive integer.

The W-norm of a given error vector, $\bar{e}$ is defined as

$$\|e\|_w = \sum_{i=1}^{l} a(R(e_i))e_i = \sum_{i=1}^{l} a(i)e_{(i)} \tag{6}$$

where $e = [e_1, \ldots\ldots, e_l]^T \in \Re^l$

and $R(e_i)$ denotes the rank of $e_i$ among $e_1, \ldots\ldots, e_l$. The ordered values of $e_1 \ldots\ldots e_l$ are given by $e_{(1)} \le \ldots\ldots \le e_{(l)}$. The score function used in this paper is defined as

$$\varphi(u) = \sqrt{12}(u - 0.5) \tag{7}$$

(v) The weights of the model are selected in such a way that the W-norm of error vector is minimized. The W-norm minimization operation is carried out using PSO technique.

(vi) In PSO the velocity and position of each bird is updated using (1) and (2) given in Section II.

(vii) The required output is obtained by adding the median of errors to the output of the model.

(viii) In each generation the minimum MSE (MMSE) is obtained and plotted against generation to show the learning characteristics of the model

(ix) The learning process is stopped when MMSE reaches the minimum possible floor level.

(x) The prediction capability of the model so developed is tested with known stock indices.

## V. SIMULATION STUDY

### A. *Experimental Data for training and testing*

The data for the stock market prediction experiments has been collected for Standard's & Poor's 500 (S&P 500), USA and Dow Jones Industrial Average (DJIA), USA. The experimental data used consists of technical indicators and daily close price of the stock indices. The total number of samples for the stock indices is 3228 trading days, from 3rd January 1994 to 23rd October 2006. Each sample consists of the closing price, opening price, lowest price, highest price and the total volume of stocks traded for the day. Ten technical indicators are extracted from the raw data as indicated in Table 1.

The available data is divided into two sets – training and testing sets. The training set consists of 2510 samples and the rest is used for testing. All the inputs are normalized to values between -1 to +1. The normalization is carried out by expressing the data in terms of the maximum and minimum value of the dataset. The training set is corrupted with 10% to 40% outliers having strength between -0.5 and +0.5.

### B. *Training and testing of the forecasting model*

Training of the WLC model is carried out using the PSO algorithm given in Section II and the optimum weights are obtained. Then using the trained model, the forecasting performance is tested using test patterns for one, seven and fifteen days in advance.

The Mean Absolute Percentage Error (MAPE) defined in (8) is computed to compare the performance of various models.

$$MAPE = \frac{1}{K}\sum_{k=1}^{K}\left|\frac{d_k - y_k}{d_k}\right| \times 100 \qquad (8)$$

where $K$ is the number of test patterns.

C. *Results and Discussions*

The proposed robust model of Fig. 1 is simulated to predict the S&P500 and DJIA stock indices closing price one, seven and fifteen days in advance. To compare the performance of the PSO based linear combiner model with error square as cost function is also simulated. The ten technical indicators used for this simulation are EMA10, EMA20, EMA30, ADO, STI, RS19, RSI14, PROC27, CPACC and HPACC. Various parameters used in the simulation study for PSO are :

No. of particle =120, $C_1$=1.042, $C_2$=1.042, no of iterations = 100.

The inertia weight at $k$ th generation is given by

$$w_k = w_0 - \frac{(w_0 - w_1) * k}{itr} \qquad (9)$$

where $k$ = iteration counter (from 1 to $itr$ ), $itr$ = number of iterations, $w_0 = 0.9$ and $w_1 = 0.4$

Figs. 2 and 3 display the actual vs. predicted graphs for S&P500 index for one day ahead during training and testing using WLC model with 20% outlier respectively. Similarly Figs. 4 and 5 show one day ahead actual and predicted DJIA stock value during training and testing with 40% outlier respectively. Table-2 shows the comparison of MAPE obtained by minimizing error square norm and W-norm cost functions in linear combiner model at different outliers (0% to 40%) and for one, seven and fifteen days ahead prediction of S&P 500. The identical results for DJIA prediction is shown in Table 3.

The results of Figs. 2-6 indicate even though outliers are present in the training set the predicted stock indices agree quite well with the actual values. Further, the mean average percentage of error is less in the W-norm minimization based model compared to conventional error square minimization based model. This is true in all cases and against outliers up to 40%.

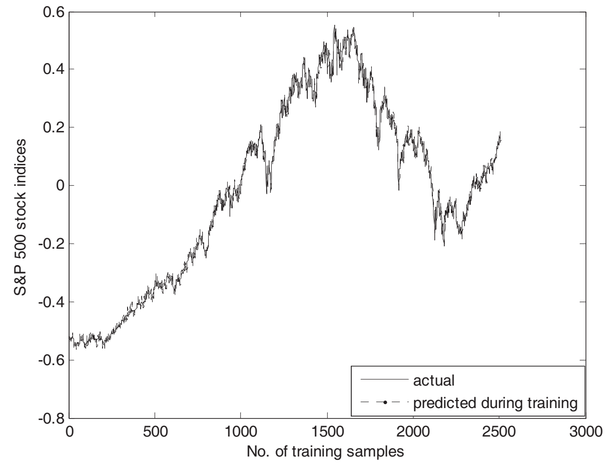


Fig. 2 Comparison of actual and one Day ahead predicted value during training for S&P500 using WLC with 20% outliers

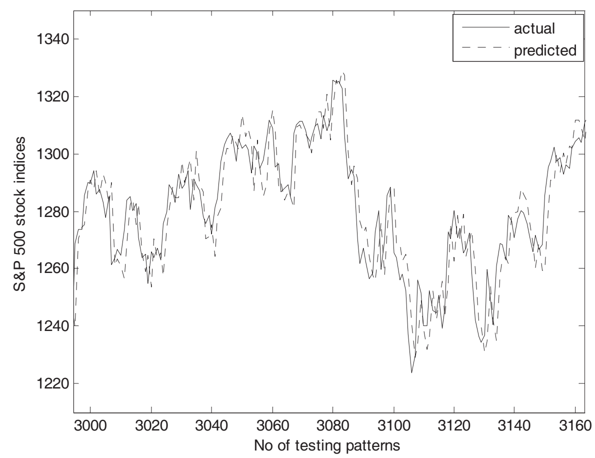


Fig. 3 Comparison of actual and one day ahead predicted value during testing for S&P500 using WLC with 20% outliers

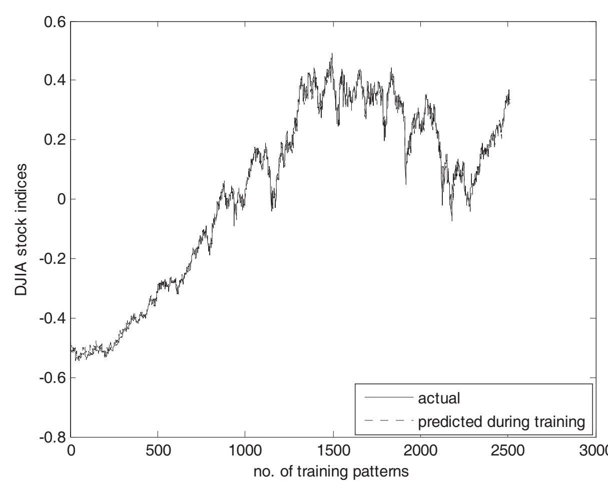


Fig. 4 Comparison of actual and one day ahead predicted value during training for DJIA using WLC with 40% outliers

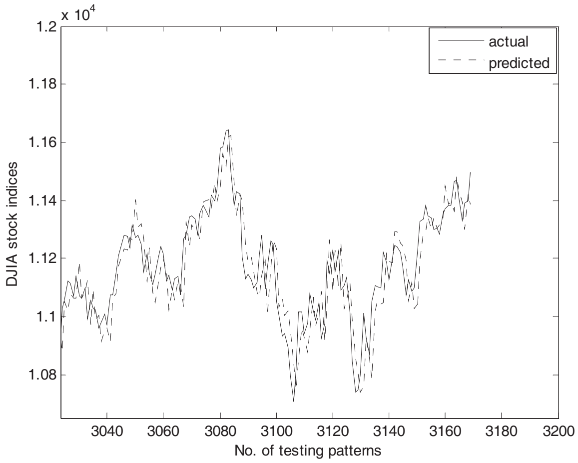


Fig. 5 Comparison of actual and one day ahead predicted value during testing for DJIA using WLC with 40% outliers

TABLE 1
SELECTED TECHNICAL INDICATORS AND THEIR FORMULA

| Technical Indicators used | Formula |
|---|---|
| Exponential Moving Average (EMA) (3 numbers) | $(P \times A) + (\text{Previous EMA} \times (1 - A))$; A=2/(N+1) P – Current Price, A- Smoothing factor, N-Time Period (EMA10, EMA20 and EMA30 are calculated using the given formula) |
| Accumulation / Distribution Oscillator (ADO) | $\frac{(\text{C.P - L.P}) - (\text{H.P - C.P})}{(\text{H.P - L.P}) \times (\text{Period's Volume})}$ C.P – Closing Price, H.P – Highest price, L.P – Lowest price |
| Stochastic Indicator (STI) | $\%K = \frac{(\text{Today's Close - Lowest Low in K period})}{(\text{Highest High in K period - Lowest Low in K period})} \times 100$ $\%D$ = SMA of $\%K$ for the Period. |
| Relative Strength Index (RSI) (2 numbers) | $\text{RSI} = 100 - \frac{100}{1 + (\text{U/D})}$ U= total gain/n, D= total losses/n, n = number of RSI period (RSI9 and RSI14 are calculated using the given formula) |
| Price Rate Of Change (PROC) | $\frac{(\text{Today's Close - Close X-period ago})}{(\text{Close X-period ago})} \times 100$ (PROC27 is calculated using the formula) |
| Closing Price Acceleration (CPACC) | $\frac{(\text{Close Price - Close Price N-period ago})}{(\text{Close Price N-period ago})} \times 100$ |
| High Price Acceleration (HPACC) | $\frac{(\text{High Price - High Price N-period ago})}{(\text{High Price N-period ago})} \times 100$ |

TABLE 2
MAPE COMPARISON FOR S&P 500

| Days Ahead | Outliers(%) | Error square norm | Wilcoxon norm |
|---|---|---|---|
| | 0 | 0.7284 | 0.6494 |
| | 10 | 0.9655 | 0.6060 |
| 1 day | 20 | 1.0211 | 0.6333 |
| | 30 | 0.7785 | 0.6240 |
| | 40 | 0.9508 | 0.6588 |
| | 0 | 1.3974 | 1.3800 |
| | 10 | 1.4810 | 1.3761 |
| 7 days | 20 | 1.5277 | 1.4080 |
| | 30 | 1.4115 | 1.3852 |
| | 40 | 1.5563 | 1.3970 |
| | 0 | 1.8296 | 1.8091 |
| | 10 | 1.9129 | 1.8272 |
| 15 days | 20 | 1.8999 | 1.8396 |
| | 30 | 1.8682 | 1.9762 |
| | 40 | 2.0507 | 1.8333 |

TABLE 3
MAPE COMPARISON FOR DJIA

| Days Ahead | Outliers(%) | Error square norm | Wilcoxon norm |
|---|---|---|---|
| | 0 | 0.7042 | 0.5848 |
| | 10 | 0.6818 | 0.5848 |
| 1 day | 20 | 0.9898 | 0.9841 |
| | 30 | 0.9542 | 0.8399 |
| | 40 | 0.7847 | 0.6224 |
| | 0 | 1.3836 | 1.3229 |
| | 10 | 1.4113 | 1.3910 |
| 7 days | 20 | 1.6418 | 1.3994 |
| | 30 | 1.4277 | 1.4015 |
| | 40 | 1.4760 | 1.4088 |
| | 0 | 1.8692 | 1.8529 |
| | 10 | 1.8975 | 1.8632 |
| 15 days | 20 | 2.3854 | 2.3582 |
| | 30 | 2.1265 | 2.0554 |
| | 40 | 2.2101 | 1.9006 |

## VI. CONCLUSION

The PSO based adaptive linear combiner model for short and long term forecasting of stock indices is developed in this paper using Wilcoxon norm as the robust cost function. To demonstrate the robust performance of the proposed model simulation study is carried out using known stock indices and comparing the same with those obtained from conventional PSO based forecasting model where error square is used as the cost function . The comparison indicates that the proposed model offers superior prediction accuracy and lower MAPE compared to that of the conventional PSO model. Thus the proposed one is a novel promising forecasting model for stock market prediction in presence of strong outliers in the training data.

## REFERENCES


[1] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Second Edition, Pearson.

[2] Clarence N.W. Tan and Gerhard E Wittig, "A Study of the Parameters of a Back propagation Stock Price Prediction Model", Proc. of First New Zealand International Two-Stream Conference on Artificial Neural Networks and Expert Systems, pp. 288-91, 1993.

[3] W. McKean Joseph, "Robust analysis of linear models", Statistical Science, vol. 19, no. 4, pp. 562-570, 2004.

[4] M. Lettau, "Explaining the facts with adaptive agents: the case of Mutual Fund Flows", Journal of Economic Dynamics and Control, vol. 21, 1997.

[5] R. Jiang and K. Y. Szeto, "Discovering investment strategies in Portfolio management" A Genetic Algorithm approach", Proc. Of 9th International conference on Neural Information Processing (ICONP-02), 2002, vol. 3, pp. 1206-1210.

[6] S. Bhattacharya, O. Pictet and G. Zumbach, "Semantics for Genetic Programming based learning in High-frequency financial data", Proc. Of the 3rd Annual Genetic Programming Conference, 1998.

[7] S. H. Chen, C. H. Yeh and W. C. Lee, "Operation pricing with genetic Programming", Pro. of the 3rd Annual Genetic Programming Conference, 1998.

[8] Ritanjali Majhi, G. Panda, G. Sahoo, P. K. Dash and D. P. Das, "Stock Market prediction of S&P 500 and DJIA using Bacterial Foraging Optimization Technique", Proc. of IEEE Congress on Evolutionary Computation (CEC-2007),Singapore,25-28, September, 2007, pp.2569-2575.

[9]Ritanjali Majhi, G. Panda, G. Sahoo and A. Panda, "On the development of Improved Adaptive Models for Efficient Prediction of Stock Indices using Clonal-PSO (CPSO) and PSO Techniques", International Journal of Business Forecasting and Market Intelligence, vol. 1, no. 1, pp.50-67, 2008.

[10] Jer-Guang Hsieh, Yih-Lon Lin and Jyh-Horng Jeng, "Preliminary study on Wilcoxon learning machines", IEEE Trans. on Neural Network, vol. 19, no. 2, pp. 201-211, 2008.

[11] J. Kennedy and R. Eberhart, "Particle Swarm Optimization", in Proc. IEEE Int. Conf. Neural Networks, Perth, Australia, Dec. 1995, pp. 1942-1948.