Ensemble Neurocomputing Based Oil Price Prediction

Lubna A. Gabralla[1], Hela Mahersia[2], and Ajith Abraham[1,3,4]

[1] Faculty of Computer Science and Information Technology,
Sudan University of Science & Technology, Khartoum, Sudan
lubnagabralla@gmail.com
[2] Image Signal and Information Processing Research Laboratory,
National Engineering School of Tunis (ENIT), University of Tunis El Manar, Tunisia
helamahersia@yahoo.fr
[3] IT4Innovations - Center of excellence, VSB - Technical University of Ostrava,
Czech Republic
[4] Machine Intelligence Research Labs,
Scientific Network for Innovation and Research Excellence, WA, USA
ajith.abraham@ieee.org

Abstract. In this paper, we investigated an ensemble neural network for the prediction of oil prices. Daily data from 1999 to 2012 were used to predict the West Taxes, Intermediate. Data were separated into four phases of training and testing using different percentages and obtained seven sub-datasets after implementing different attribute selection algorithms. We used three types of neural networks: Feed forward, Recurrent and Radial Basis Function networks. Finally a good ensemble neural network model is formulated by the weighted average method. Empirical results illustrated that the ensemble neural network outperformed other models.

Keywords: Oil price prediction, ensemble neural network, computational intelligence.

1 Introduction

Oil is one of the most important topics in the contemporary world and it will remain as a keyword in the world politically, and economically. Oil has unique properties that can be used to control and conquer the world successfully. The history of oil discovery goes back to the 1859, when the first oil was drilled in Pennsylvania, United States [1]. After that it became very useful in the manufacturing engines and cars, planes and machinery. In 1914 during the first World War the head of the French government at the time described that every drop of oil is equal to a drop of blood in an orientation to its importance [2].

At present, oil is the most important source of energy and one of the elements of modern civilization for humans. It is used as a fuel for cars, airplanes, factories and agricultural equipment, trucks, commercial and military ships and electric power generation for homes, workplaces and other places. Oil prices have undergone many changes and instabilities over the years. It was known as oil shocks, and the first

© Springer International Publishing Switzerland 2015

A. Abraham, P. Krömer, & V. Snášel (eds.), *Afro-European Conf. for Ind. Advancement*,
Advances in Intelligent Systems and Computing 334, DOI: 10.1007/978-3-319-13572-4_24

shock was in the October war in 1973, where the price rose from 2.29$ to 10.73$ for a barrel until 1974, and these prices continued to rise, and even achieved strong jump to 32.51$ per barrel in 1981, this what is known as the second oil shock, and the third oil shock was after Iraq's invasion of Kuwait, when the price of oil rose from 17.31$ for the year 1989 to 22.26$, in 1990. Oil prices continued volatility until prices of oil was collapsed in 1998 and the average barrel of oil was around 9.69$ for OPEC. This was as a result of decline in global oil demand after the financial crisis [3]. Several researchers and scientists were interested in studying the factors that influence the oil prices, like climate [4], politics [5], and stock market [6] etc. Owners of the economic sector, such as commercial institutions and companies operating in the field of oil were very much concerned and wanted to know the prices of oil in the coming years in order to determine their economic policies and building plans for the future and make informed decisions, which will help them to avoid the problems of inflation and economic stagnation, losses and financial crises. Recently several artificial intelligence algorithms were used for oil price prediction. Artificial neural networks have many characteristics and does not need any hypotheses (a priori) to be introduced and is able to deal with incomplete information and with the large number of variables and generally it is flexible in modeling [7]. Therefore, this study aims to employ several types of neural network algorithms to develop a computational model that is able to predict oil prices with high accuracy and high performance, which can contribute to the development of the local and global economy. The rest of this paper is organized as follows: After a short literature review in Section 2, Section 3 describes the research methodology in detail. The data used and their divisions are found in Section 4, and experimental results are reported in Section 5 followed by concluding remarks.

2 Related Works

Artificial neural networks (ANN) [8] are designed to represent data by simulating the work of the human brain. ANN's emerged in different areas such as industrial, medical and business, and achieved successful results therefore many researchers also used ANN in the oil industry. Kaboudan [9] selected multilayer perceptron (MLP) and Genetic programming (GP) to forecast crude oil price using monthly data, such as world crude production, OECD consumption, world stocks and lagged crude FOB crude oil price of US imports. Two methods are compared to a random walk and their results proved that GA has an advantage over random walk predictions. Yu et al. [10] constructed an empirical mode decomposition (EMD) based on neural network ensemble learning. They used daily West Texas Intermediate (WTI) data from 1/1/1986 to 30/9/2006 as training and Brent from 20/5/1987 to 30/9/2006 as testing data. Results proved that EMD based neural network ensemble can be used for oil price prediction. Haidar et al. [11] suggested a network to predict the oil prices using two groups of inputs, crude oil futures data, and Dollar index, S&P500, gold price and heating oil price. The authors measured performance by hit rate, root mean square error, correlation coefficient, mean squared error and mean absolute error. The authors concluded that heating oil spot price support forecast crude oil spot price in numerous steps prediction. Alizadeh and Mafinezhad [12] proposed General

Regression Neural network(GRNN) using six factors monthly data to predicting Brent crude oil price. Experiment results show that the model achieved high accuracy in normal ad crisis situations. Mingming and Jinliang [13] collected data covering Brent and West Texas Intermediate (WTI) from 1946 to 2010 and adopted multiple wavelet recurrent neural networks (MWRNNs) to forecast crude oil prices. The study showed that the model has high prediction accuracy. Yu et al. [14] introduced fuzzy ensemble prediction model, support vector machine, radial basis function networks and back-propagation neural networks to predict crude oil prices. They used the data covering a period from January 2000 to December 2007 using West Texas Intermediate and Brent crude oil spot. Results showed that agent-based fuzzy ensemble prediction model outperformed other individual methods in accuracy. Most of the studies in the literature focused on constructing a new model using one percentage of training and testing on other hand, few researchers were interested in using different inputs for testing. So the objective of this paper was to provide a variety of the training and testing percentages with a set of different inputs using several kinds of neural networks to get high accuracy for the model.

3 Research Methodology

3.1 Feed Forward Neural Networks (FFN)

Back propagation [15][16] method is a supervised learning scheme and the most popular technique in multilayer networks when a set of input produces its own actual output and then compare it with the target value by calculating the error, after that error is fed back through the network. The weights of each connection are adjusted to reduce the error by several ways, such as gradient descent etc. until sufficient performance is achieved. To improve the generalization, there are several learning methods such as Levenberg – Marquardt (LM), Bayesian regularization (BR) and BFGS quasi-Newton (BFG-QN) back propagation algorithm [17].

3.2 Recurrent Neural Network (RCN)

RCN is the state of the art in nonlinear time series prediction, system identification, and temporal pattern classification. As the output of the network at time t is used along with a new input to compute the output of the network at time $t+1$, the response of the network is dynamic [8].

3.3 Radial Basis Function (RBF)

Radial basis function network [8] is an artificial neural network that uses radial basis functions as activation functions. The output of the network is a linear combination of radial basis functions of the inputs and neuron parameters. RBF is successful in numerous fields especially for system control, time series and prediction.

3.4 Ensemble Neural Network [18]

The generalized ensemble method find weights for each output that minimizes the MAE of the ensemble. The general ensemble model (GEM) is defined by:

$$F_{GEM} = \sum_{i=1}^{n} \alpha F_i(X) \tag{1}$$

Where $\alpha F_i(x)$ are chosen to minimize the MAE between the outputs and the desired values. Figure 1 illustrates the ensemble neural network approach.

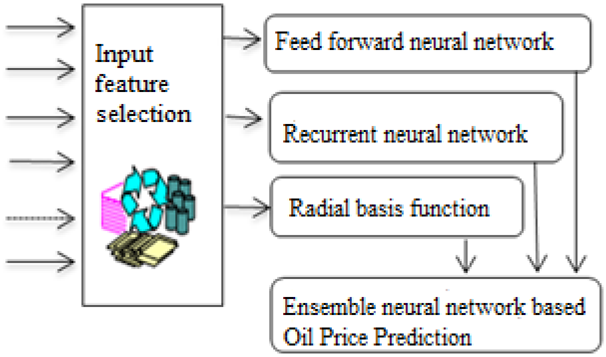

Fig. 1. Ensemble neural network for oil price prediction

4 Experiments

The daily data [19, 20] (from 1999 to 2012) were used to predict the West Taxes Intermediate (Output). The dataset consists of 14 variables as following:

- Date (DT).
- West Texas Intermediate (WTI).
- Federal Fund rate (FFR).
- Volatility Implied Equity Index (VIX).
- The regional Standard & Poor's equity index US, Europe and Asia (SPX).
- Gasoline prices New York & US Gulf Coast (GPNY) & (GPUS).
- Heating oil spot prices (HP).
- Future contracts 1,2,3,4 for WTI (FC1, FC2, FC3, FC4).
- EUR/USD exchange rate (ER).
- Gold prices (GP).

We also examined the effect of training and testing data by randomly splitting them as follows:

- 90% - 10% (A)
- 80% - 20% (B)
- 70% - 30% (C)
- 60% - 40% (D)

We used WEKA for pre-processing experiments [21] and formulated 7 different sub datasets, which were derived from the original dataset after implementing several attribute selection algorithms, such as:

- Attributes ranking principal: ranked list of attributes based on evaluated individually each attribute [22].
- Wrapper attributes Selection: It depends on an induction algorithm to estimate the merit of feature subsets [23].
- Relief for regression: Evaluates quality of attributes according to value of the given attribute for the near instance to each other and different predicted (class) value [24].
- Correlation based Feature Selection (CFS): assesses the value of group of attributes by concerning the individual predictive ability of each features as well with the possibility of repetition among the features. Selecting a subset of the original attributes to reduce the dimensionality of the data and then constructing a model from these reduced number of features in some cases could improve the prediction accuracy and performance, and a simpler model that is easier to interpret [19][22]. Table 1 summarizes the results of attribute selection and the 7 sub-data sets (SDS1-SDS7) obtained.

Table 1. Attributes selection methods and their features

Sub dataset	Method	Features
SDS1	Correlation based Feature Selection subset evaluator	WTI; SPX; FG1
SDS2	Correlation based Feature Selection subset evaluator	DT; VIX; WTI; SPX; GPNY; GPUS; HP; ER; FC1; FC2; FC3; FC4
SDS3	Correlation based Feature Selection subset evaluator	VIX; WTI; GPNY; ER; FC1
SDS4	Correlation based Feature Selection subset evaluator	WTI; GPNY; FC1
SDS5	Correlation based Feature Selection subset evaluator	VIX; WTI; GPNY; FC1
SDS6	Wrapper subset evaluator	WTI; FC1
SDS7	Wrapper subset evaluator	WTI; GPUS

5 Experimental Results

5.1 Feed Forward Neural Network

Neural network experiments are accomplished in MATLAB. We used one hidden layer exploring 40-45-50-55-60 neurons and used tan-sigmoidal transfer function for the hidden layer and pure linear function in the output layer. We measured the performance using mean absolute error (MAE).

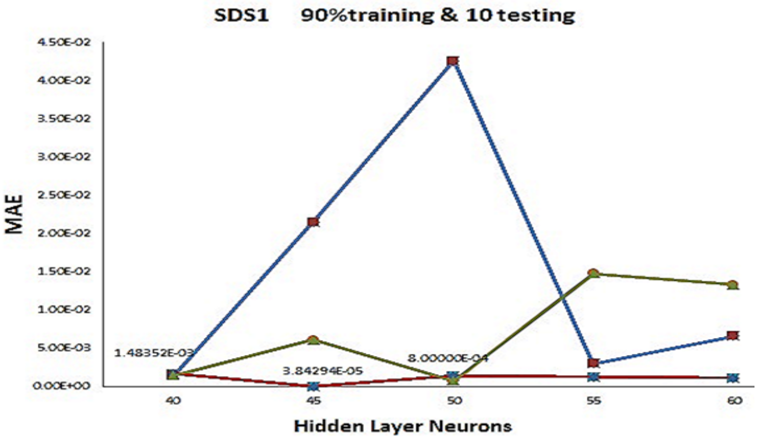

Fig. 2. Comparison of feed-forward networks using three different training algorithms

According to Table 2, in most of the sub-datasets the best results were obtained when using the Bayesian regulation back propagation method with 80% training and 20% testing, and sub-dataset1 (SDS1) achieved MAE= 3.843E-05 with 90% training and 10% testing using 45 neurons. Figure 2 shows the best results in SDS1 using feed-forward networks comparing Levenberg –Marquardt (LM), Bayesian regularization (BR) and BFGS Quasi-Newton (BFG-QN) algorithms.

Recurrent Neural Network

We used a hidden layer with 10 neurons and used three training algorithms Levenberg –Marquardt (LM), Bayesian regularization (BR) and BFGS Quasi-Newton (BFG-QN). Bayesian regularization method outperformed other algorithms by 51.85%. It is noted from Table 3 that for all the sub-datasets in the percentage 80% training and 20% testing is the best (shaded area), except in sub-dataset (SDS5) 90% training and 10% testing is the best. On the other hand the lowest value of the error is 3.941 E-05 when using 90% training and 10% testing with sub-dataset (SDS5).

Radial Basis Function Network

We constructed the network until it reached a maximum number of neurons or the sum-squared error falls beneath an error goal. Table 4 shows the results obtained using the seven sub-datasets and different number of neurons: 40, 45, 50, 55 and 60. The shaded area indicates the best results when using 60 neurons in few sub-datasets then followed by 55 neurons. According to the percentage of training and testing sub-dataset (SDS4 - SDS5-SDS6 - SDS7) achieved the best results with 80% training &20% testing. The best results over all sub-datasets is an MAE of 2.206 E-05 in SDS6 with 80% training & 20 % testing using 45 neurons.

Table 2. Performance of FFN

Sub-datasets	Data	Mean Absolute Error			Hidden layer neurons
		LM	BR	BFG-QN	
SDS1	A	1.48352E-03	**3.84294E-05**	8.00000E-04	45
	B	4.81400E-03	**1.35000E-04**	4.00000E-04	45
	C	5.77900E-03	**3.55000E-04**	1.50000E-03	45
	D	8.15200E-03	**4.38000E-04**	6.30000E-03	40
SDS2	A	9.17000E-04	2.83700E-03	**9.00000E-04**	40
	B	4.48000E-04	2.18100E-03	**4.00000E-04**	40
	C	2.74700E-03	**1.02200E-03**	1.90000E-03	45
	D	3.15700E-03	**7.69000E-04**	1.00000E-03	60
SDS3	A	1.83600E-03	**1.26594E-04**	1.10000E-03	50
	B	3.00400E-03	**1.24897E-04**	9.00000E-04	50
	C	1.21500E-02	5.31100E-03	**3.10000E-03**	45
	D	1.69940E-02	4.06200E-03	**1.80000E-03**	45
SDS4	A	5.58850E-02	**5.74155E-05**	7.80000E-03	50
	B	2.86700E-03	**6.45000E-05**	1.60000E-03	50
	C	2.48740E-02	**3.04484E-04**	7.70000E-03	50
	D	1.67079E-02	**1.94000E-04**	2.40000E-03	50
SDS5	A	3.50300E-03	**9.65000E-04**	2.00000E-03	55
	B	1.40900E-03	**6.80000E-05**	8.00000E-04	60
	C	2.19900E-02	**3.73327E-04**	2.80000E-03	40
	D	4.28700E-02	**2.10900E-03**	2.60000E-03	40
SDS6	A	1.44560E-02	**6.23000E-05**	5.70000E-03	40
	B	1.94854E-02	**6.04640E-05**	3.10000E-03	60
	C	3.23723E-01	**3.49000E-04**	4.95000E-02	55
	D	1.07220E-01	**1.62000E-04**	2.25000E-02	55
SDS7	A	5.69780E-02	**1.90200E-03**	1.92300E-01	40
	B	1.39500E-02	**1.97000E-04**	9.70000E-03	55
	C	9.65800E-02	**2.25700E-03**	3.45000E-02	40
	D	2.83030E-01	**4.50000E-04**	3.32000E-02	55

Experiments Using Ensemble Method

We compared the results of three different types of neural networks and observed that the RBF network outperformed other methods in obtaining the lowest error (MAE= 2.206 E-05). Also the data set using training 80% and testing 20% accomplished the best results in all the neural network methods. In the feed-forward and radial basis networks, the best results were obtained when using 45 neurons. RBF networks out performed again in the time factor, as it was faster than feed-forward and recurrent neural network.

To further improve the results, we used the ensemble neural network. We selected the best results using the category of 80% training and 20% testing for three neural networks. Experimental results illustrate that the proposed ensemble neural network is superior to other methods by achieving the lowest MAE = 2.186E-05 as shown in Table 5.

Table 3. Performance of RCN

Sub-datasets	Data	Mean Absolute Error		
		LM	**BR**	**BFG-QN**
SDS1	**A**	**1.14102E-03**	1.27500E-03	2.52400E-03
	B	1.17400E-03	2.48800E-03	**5.79000E-04**
	C	1.03650E-02	**6.98300E-03**	1.25780E-02
	D	1.29940E-02	**7.69800E-03**	1.29150E-02
SDS2	A	4.60000E-04	**3.77000E-04**	2.18328E-02
	B	2.22000E-04	**1.74000E-04**	1.18390E-02
	C	**1.36700E-03**	1.44800E-03	7.03600E-02
	D	8.61000E-04	**3.32000E-04**	3.45920E-02
SDS3	A	**5.22000E-04**	4.43900E-03	2.55500E-03
	B	3.57762E-04	**1.05000E-04**	6.40100E-03
	C	4.21100E-03	**2.82000E-04**	4.10660E-02
	D	6.82000E-04	**1.68000E-04**	1.67200E-02
SDS4	**A**	7.68000E-04	**4.80285E-05**	1.58640E-02
	B	5.20102E-05	**3.94799E-05**	7.16800E-03
	C	4.73000E-04	2.17658E-04	**2.00530E-02**
	D	2.08400E-03	**1.81824E-04**	6.17200E-03
SDS5	A	**3.94114E-05**	9.52860E-05	4.38700E-03
	B	1.16000E-04	**1.12000E-04**	4.13600E-03
	C	**4.41680E-04**	3.77708E-01	1.42909E-01
	D	4.95000E-04	**3.62000E-04**	9.70200E-03
SDS6	A	**1.83283E-04**	1.51900E-03	3.33800E-03
	B	**1.72000E-04**	1.89800E-03	1.58400E-03
	C	**1.82716E-03**	6.89000E-03	6.86900E-03
	D	**9.69000E-04**	5.59500E-03	4.03800E-03
SDS7	A	5.80000E-04	**2.48000E-04**	1.01690E-02
	B	**1.47000E-04**	2.19000E-04	2.20500E-03
	C	1.30400E-03	**1.01200E-03**	4.03600E-03
	D	**4.30824E-04**	6.10500E-03	1.03750E-02

Table 4. Performance of Ensemble neural network

Prediction models	MAE
Feed forward	6.0464E-05
Recurrent	3.9479E-05
Radial Basis Function	2.2191E-05
Ensemble	**2.1862E-05**

Table 5. Performance of RBF

Sub-datasets	Data	Mean Absolute Error based on the number of neurons				
		40	45	50	55	60
SDS1	A	1.03146E-04	9.34926E-05	1.00340E-04	5.70070E-05	5.08729E-05
	B	1.53000E-04	1.48000E-04	1.09000E-04	2.58000E-04	**2.40010E-05**
	C	4.84160E-05	5.01626E-05	5.02828E-05	5.02436E-05	4.98570E-05
	D	**3.88153E-05**	**3.81595E-05**	**3.81548E-05**	**3.81543E-05**	3.81548E-05
SDS2	A	5.05680E-03	**2.70885E-03**	**1.53678E-03**	**1.33255E-03**	**1.36009E-03**
	B	**4.09600E-03**	3.23000E-03	3.02800E-03	1.85800E-03	1.68000E-03
	C	6.13000E-03	6.04800E-03	5.16800E-03	4.47700E-03	2.85700E-03
	D	8.42200E-03	8.58200E-03	7.28400E-03	5.07200E-03	4.30200E-03
SDS3	A	**2.97000E-04**	**2.93000E-04**	**2.91000E-04**	**2.86000E-04**	2.75000E-04
	B	7.31000E-04	7.51000E-04	7.21000E-04	3.81000E-04	**2.24000E-04**
	C	1.50700E-03	6.52000E-04	5.11000E-04	9.74410E-04	6.96864E-04
	D	1.60300E-03	1.56300E-03	1.51400E-03	1.52400E-03	1.54100E-03
SDS4	A	2.11138E-04	2.09055E-04	2.09022E-04	2.09002E-04	2.08946E-04
	B	**6.08910E-05**	**6.23224E-05**	**6.23330E-05**	**6.23224E-05**	**6.35603E-05**
	C	1.25481E-04	1.10437E-04	1.13763E-04	1.16843E-04	1.16840E-04
	D	4.06000E-04	3.84000E-04	3.81000E-04	3.80000E-04	3.80000E-04
SDS5	A	1.00800E-03	8.17000E-04	3.54000E-04	1.60000E-04	4.34000E-04
	B	**1.53000E-04**	**1.03000E-04**	**1.26000E-04**	**1.04000E-04**	**1.21483E-04**
	C	4.14000E-04	4.01671E-04	4.01344E-04	4.28000E-04	2.34880E-04
	D	2.54000E-04	2.16000E-04	1.97000E-04	1.85000E-04	8.74530E-05
SDS6	A	9.04430E-02	9.04440E-02	9.04440E-02	9.04440E-02	9.04440E-02
	B	**2.21914E-05**	**2.20646E-05**	**2.21172E-05**	**2.21087E-05**	**2.21087E-05**
	C	4.51944E-05	4.51615E-05	4.49053E-05	4.48180E-05	4.47060E-05
	D	3.51792E-05	3.41134E-05	3.40980E-05	3.35330E-05	3.35330E-05
SDS7	A	2.34000E-04	2.85000E-04	2.76000E-04	2.72000E-04	2.76000E-04
	B	**4.03840E-05**	**4.03361E-05**	**3.98230E-05**	**3.98158E-05**	**3.98230E-05**
	C	5.37473E-05	5.34101E-05	5.34009E-05	5.36361E-05	5.36107E-05
	D	1.21000E-04	1.20000E-04	1.20000E-04	1.16000E-04	1.16000E-04

6 Conclusions

In this paper, we presented and ensemble neural network model for prediction of oil price. The model is based on three different types of neural networks: feed-forward, recurrent and radial basis function networks. The network structure was selected after many experiments including a number of hidden neurons and several learning methods. In addition, four different groups of training and testing were experimented and many attribute selection algorithms were implemented, which leaded to 7 different sub-datasets. The results illustrate that the radial basis function achieved the best MAE and less time to run when compared it with other individual methods. Ensemble methods were found to be superior when compared to the individual neural networks and learning methods.

References

1. http://history1800s.about.com/od/oil/a/first-oil-well.htm
2. http://www.theforbiddenknowledge.com/../the_rothschild_bloodline.htm
3. http://www.opec.org/opec_web/en/publications/457.htm
4. Pindyck, R.S.: The dynamics of commodity spot and futures markets: a primer. The Energy Journal, 1–29 (2001)
5. Griffin, J.M.: OPEC behavior: a test of alternative hypotheses. The American Economic Review, 954–963 (1985)
6. Soytas, U., et al.: World oil prices, precious metal prices and macroeconomy in Turkey. Energy Policy 37(12), 5557–5566 (2009)
7. Maimon, O.Z., Rokach, L.: Data mining and knowledge discovery handbook, vol. 1. Springer (2005)
8. Abraham, A.: Artificial neural networks. In: Handbook of Measuring System Design (2005)
9. Kaboudan, M.: Compumetric forecasting of crude oil prices. In: Proceedings of the 2001 Congress on Evolutionary Computation. IEEE (2001)
10. Yu, L., Wang, S., Lai, K.K.: Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. Energy Economics 30(5), 2623–2635 (2008)
11. Haidar, I., Kulkarni, S., Pan, H.: Forecasting model for crude oil prices based on artificial neural networks. In: International Conference on Intelligent Sensors, Sensor Networks and Information Processing, ISSNIP 2008. IEEE (2008)
12. Alizadeh, A., Mafinezhad, K.: Monthly Brent oil price forecasting using artificial neural networks and a crisis index. In: 2010 International Conference on Electronics and Information Engineering (ICEIE). IEEE (2010)
13. Mingming, T., Jinliang, Z.: A multiple adaptive wavelet recurrent neural network model to analyze crude oil prices. Journal of Economics and Business 64(4), 275–286 (2012)
14. Yu, L., Wang, S., Lai, K.K.: A generalized Intelligent-agent-based fuzzy group forecasting model for oil price prediction. In: IEEE International Conference on Systems, Man and Cybernetics, SMC 2008, IEEE (2008)
15. Reed, R.D., Marks, R.J.: Neural smithing: supervised learning in feedforward artificial neural networks. Mit Press (1998)
16. Chauvin, Y., Rumelhart, D.E.: Backpropagation: theory, architectures, and applications. Psychology Press (1995)
17. Demuth, H., Beale, M., Hagan, M.: Neural network toolbox™ 6. User's guide (2008)
18. Maqsood, I., Khan, M.R., Abraham, A.: An ensemble of neural networks for weather forecasting. Neural Computing & Applications 13(2), 112–122 (2004)
19. http://www.eia.gov
20. http://www.usgold.com
21. Garner, S.R.: Weka: The waikato environment for knowledge analysis. In: Proceedings of the New Zealand Computer Science Research Students Conference. Citeseer (1995)
22. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. The Journal of Machine Learning Research 3, 1157–1182 (2003)
23. Hall, M.A.: Correlation-based feature selection for machine learning. The University of Waikato (1999)
24. Robnik-Šikonja, M., Kononenko, I.: An adaptation of Relief for attribute estimation in regression. In: Machine Learning: Proceedings of the Fourteenth International Conference, ICML1997 (1997)