Neural Network Design Parameters for Forecasting Financial Time Series

Assia Lasfer, Hazim El-Baz, and Imran Zualkernan
Engineering Systems Management
American University of Sharjah
Sharjah, UAE

Abstract—**Neural Networks (NN) have been used extensively by researchers and practitioners to forecast financial time series. The forecasting accuracy of NN depends on several design parameters, and fine-tuning them to suit a particular financial time series is essential for attaining lower error levels and minimizing running time. This paper presents the results of a two-level full-factorial Design of Experiment developed to investigate the significant factors that influence the performance of NN in forecasting financial time series. The factors considered in this paper are NN type, number of neurons in the hidden layer, the learning rate of LM algorithm, and the type of output layer transfer function. The methodology is applied to the Morgan Stanley Capital International Index for United Arab Emirates.**

Keywords— Artificial neural networks (NN), Design of experimentsb(DOE), UAE , Financial time series

I. Introduction

The biggest challenge for financial professionals and researchers is the existence of uncertainty and risk. While gambling with investment choices is not an option for most, proper research and planning can greatly reduce risk and guide investors to the right steps to take. Artificial neural networks (NNs) gained great popularity in the financial field for their ability to deal with uncertainty and handle noisy data [1]. Many applications exist that prove the superiority of NN over traditional technical analysis [2]. The features of NN make it convenient to study stock market behaviors; however, although theoretically NN can approximate any function, reaching a good design is a big challenge. NN design is done through calibrating numerous parameters; however, finding the correct combination of parameter values is very challenging given a specific data set. The aim of this study is to identify the important design parameters that may impact NN build for forecasting future moves of the UAE MSCI index. This is done by developing a design of experiment for the NN and performing statistical analysis on the experimental model.

978-1-4673-5814-9/13/$31.00 ©2013 IEEE

Stock market forecasting is a very active field of exploration. Researchers have published several works putting down guidelines to build good NNs. Notably, Kaastra and Boyd [3] discuss a step by step approach for the proper building of NN for forecasting financial and economic time series. They focus on all the important design parameters of back-propagation feed-forward networks and ways of configuring them. Similarly, Zhang et al [4] survey past practices and provide insights on NN modeling issues. Other similar works include Padhiary and Misha [5] who build NNs with an adaptive learning rate to predict long and short term returns. Moreover, among practical published research, most of the surveyed literature works use back-propagation feed-forward networks because of their simplicity and ease of use; however, few others explore other topologies and other learning algorithms.

When considering building a NN, there are many parameters that should be taken into account, and all of them affect the performance of the NN to a certain degree; therefore, finding the most important parameters minimizes building costs and maximizes performance. Some attempt to tackle the problem using simple experimenting of a one-factor-at-a-time fashion as done by Tan and Wittig [6]. Other works like that of Balestrassi et al [7] apply design of experiments with statistical analysis to study NN built for non-linear time series forecasting. It is clearly mentioned that one-factor-at-a-time analysis can lead to unreliable and deceptive results, whereas statistically designed experiments are more efficient. Other works that apply DOE to NN in other fields include Laosiritaworn and Chotchaithanakorn [8] and Behmanesh and Rahimi [9]. By reviewing past literature, it is seen that using design of experiments with NN has not been attempted before for financial forecasting. This work investigates this process in an effort to produce better generalizing NN for the UAE market.

II. Methodology

A. Neural networks

Artificial neural networks are adaptive computational models that are inspired from the biological human brain system. A typical NN consists of multiple neurons organized in a layered fashion and connected to each other forming an inter-dependent network. There are two types of NNs, feed-forward and recurrent (feedback) NNs. Figure 1 shows a typical 3 layered feed-forward network and Figure 2 shows a 3 layered recurrent network.

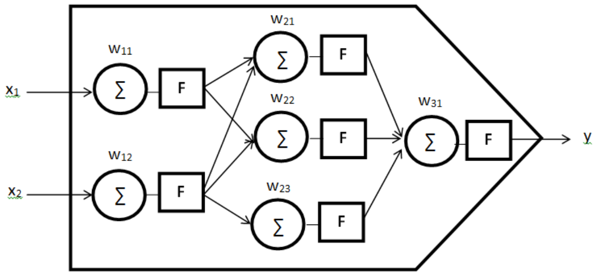

FIGURE 1 FEED-FORWARD NEURAL NETWORK

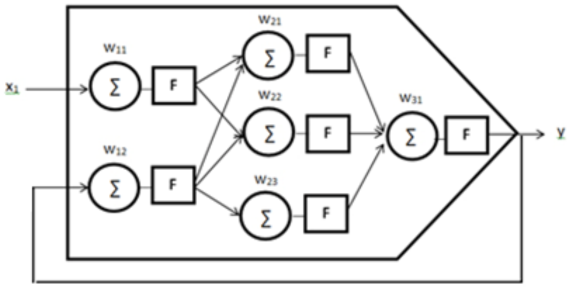

FIGURE 2 RECURRENT NARX NEURAL NETWORK

TABLE I. SUMMARY OF THE FACTOR SETTINGS TO BE USED IN THIS EXPERIMENT AND THEIR VALUES

Input variables	Closing price and EMA	Fixed
Output variable	Closing price	Fixed
Training, testing, validation	70-15-15	Fixed
NN type	Feed-forward, Recurrent(NARX)	Experiment factor
Number of input neurons	2	Fixed
Number of hidden layers	1	Fixed
Number of neurons in hidden layer	2 and (2x3)	Experiment factor
Number of output neurons	1	Fixed
Hidden layer transfer function	Hyperbolic tangent sigmoid	Fixed
Output layer transfer function	Hyperbolic tangent sigmoid and Pure linear	Experiment factor
Training algorithm	Levenberg-Marquart backpropagation	Fixed
Training algorithm (Mu)	0.001 and 1	Experiment factor
Training algorithm (Mu step up and Mu step down)	10 / 0.1	Fixed
Error (performance) function	MSE	Fixed
Max epochs	1000	Fixed
Max training time	10 mins	Fixed

In order to properly design a NN, it is necessary to look at best practices to narrow the choice of variables to use. A thorough survey done by Atsalakis and Valavanis summarizes 100 published articles [10]. The authors note that closing prices and technical indicators are commonly used as input variables. Moreover, 60% of all studies use feed-forward or recurrent NNs with mainly one hidden layer. Kaastra and Boyd [3] also describe necessary steps to obtain the best possible NN; these include rescaling, choice of inputs, number of neurons in hidden layers, choice of transfer function, and learning rate. Table 1 summarizes the parameters used in this study. The closing prices used are for UAE MSCI index from August 2002 till August 2012.

B. Design of experiments

Experiments are used to find significant factors that affect process' outputs, and factor values that give the best output. Properly designing an experiment reduces time and cost and helps focus on the desired information. The objective of conducting experiments is to find which factors are significant and which are not; however, interactions between factors must also be taken into consideration. The presence of significant interactions means that a factor's effect changes at different levels of another factor. In such a scenario, a factorial design is appropriate where factors are varied together [11]. Usually, factorial designs with 2 levels are used. Here, each factor has only a low and a high level. After designing and conducting the experiment, it is necessary to statistically analyze the results. Statistical analysis of the experimental model is important to see how likely is it that a result obtained has occurred by chance alone. Having a level of significance of $p < 0.05$ means that any effect that is likely to happen less than 5% of the time by chance is statistically significant [12]. An ANOVA test compares the variance due to the factor under investigation with the variance due to chance, and this is done by using the F-test. Therefore,

$$F = \frac{\text{Effect Variance}}{\text{Residual Variance}} = \frac{\text{Mean Square of Term}}{\text{Mean Square of Residual}} \tag{1}$$

When conducting the F-test, the null hypothesis is H_o: the tested term in not significant. To test this hypothesis, a p-value associated with a significant level (usually 0.05) is checked. A level is said to be significant if the p-value is less than 0.05 i.e. reject H_o [12].

The model built for this study is a 4 factor full factorial with each factor having two levels. Each treatment is repeated 3 times to calculate the pure error, which gives more accurate results. This results in 48 runs. The response factor is the mean square error (MSE); it is chosen because it is used to train the NN and therefore will better reflect the behavior of the network in the experiment. Table 2 summarizes the model.

TABLE II. SIGNIFICANT EFFECTS FOR MODEL 1

Factor	Name	Low	High
A	Type	FFNN	RNN
B	H- neurons	2	6
C	Output TF	Tan	Linear
D	Mu	0.001	1

III. Results

After conducting a 2x2x2x2 ANOVA test with 95% confidence interval, the list of significant effects is retrieved. There is a main effect of number of hidden layer neurons (B) with networks made of 6 neurons having a lower average MSE than networks with 2 neurons, $F(1, 41) = 7.93$, $p<0.05$. There is also a main effect of output layer transfer function (C) where networks with a pure linear function in the output layer resulting in lower average MSE than networks with a hyperbolic tangent function, $F(1, 41) = 26.68$, $p<0.05$. Moreover, there is a main effect of learning rate Mu (D) where initializing it to 1 results in a lower average MSE than initializing it to 0.001, $F(1, 41) = 8.93$, $p<0.05$. Furthermore, there is an interaction between the type of network and the number of hidden layer neurons (AB) with a much larger difference in MSE between having 2 or 6 hidden neurons when the type of network is recurrent NARX. When the NN is feed-forward, changing the number of hidden neurons does not cause a significant difference in mean MSE. Only when the NN is recurrent (NARX), the size of the network makes a difference. On average, having a NARX network with 6 neurons in the hidden layer gives the lowest average MSE, $F(1, 41) = 5.42$, $p<0.05$. Finally, there is also an interaction between the number of hidden layer neurons and output layer transfer function (BC); when number of neurons is 2, changing the output transfer function from hyperbolic tan to pure linear is more significant than when there are 6 neurons in the hidden layer. This being said, a network with 6 hidden neurons gives lower average MSE, $F(1, 41) = 7.15$, $p<0.05$. Figure 3 and Figure 4 show the significant effects and interactions. The regression equation is:

$$
\begin{aligned}
\text{MSE UAE} = & \\
&+4.769\text{E-}004 \\
&-2.457\text{E-}006 \quad * A \\
&-3.721\text{E-}005 \quad * B \\
&-6.825\text{E-}005 \quad * C \\
&-3.949\text{E-}005 \quad * D \\
&-3.077\text{E-}005 \quad * A * B \\
&+3.534\text{E-}005 \quad * B * C
\end{aligned} \tag{2}
$$

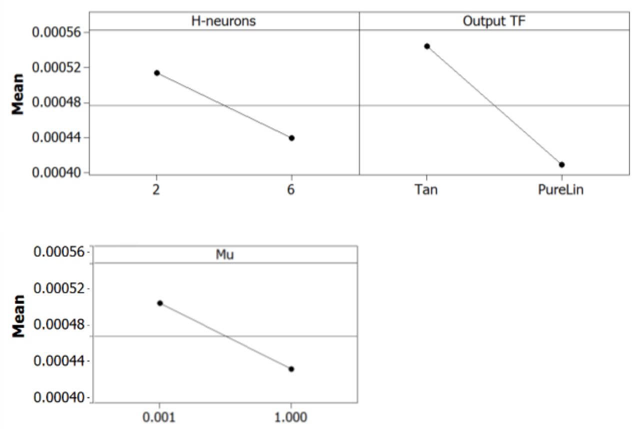

Figure 3 UAE MSE significant main effects

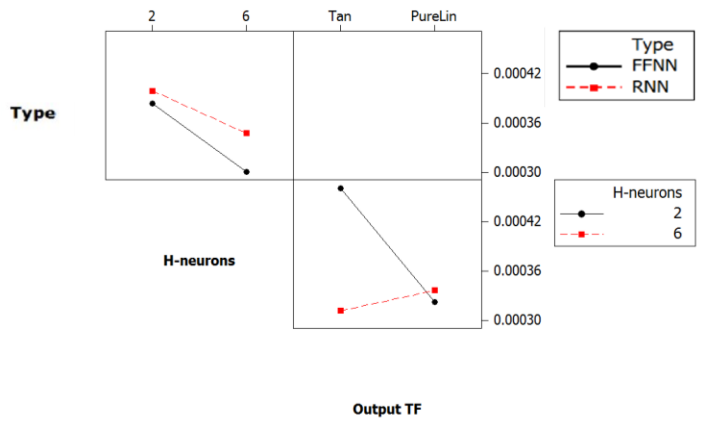

Figure 4 UAE MSE significant interactions

After conducting the ANOVA test, it is necessary to diagnose the model and make sure that the assumptions of the ANOVA are met. These two assumptions are that residuals are normally distributed and that their variance is constant.

Looking at the observed results, it is evident that the MSE decreases as the output transfer function changes from hyperbolic tangent to pure linear. Unlike the hyperbolic tangent function, the pure linear function does not transform the input it gets from the hidden layer; instead, the results of the hidden layer are summed up in the output layer and are passed as is. Changing the factor of number of neurons in the hidden layer deals with the size of the NN, and as the number of neurons increases the mean MSE decreases. This could be due to the increased ability of the NN to better learn patterns with more neurons in the hidden layer. A larger NN, to a certain extent, has a larger vector of weights to modify and is thus more capable of generalizing. The results found are specific to this model and to the size of network chosen.

The interaction between number of neurons in the hidden layer and the output transfer function is a significant interaction. When number of neurons is low (2), the change from tan function to pure linear is more significant. However, when the number of neurons is high (6), the change in output transfer function is not as significant. This could be because when the NN is larger in size, the burden of learning becomes more dependent on the size of the network, and the effect of the hidden neurons increases, so the effect of the output transfer function becomes less. However, when the size is small, the effect of the output transfer function becomes more apparent. A larger NN is also a better generalizer than a small one given the studied levels. Strictly for this model, this means that for the UAE market it is generally better to build NNs with 6 hidden layer neurons.

Furthermore, the interaction between the number of hidden layer neurons and the network type is significant. When the NN is feed-forward, changing the number of hidden neurons does not cause a significant difference in mean MSE. Only when the NN is recurrent (NARX) then the size of the network makes a difference. This could be due to the different way NARX networks handle data and learn.

IV. CONCLUSION

Using NNs to build predictive models for financial time series is an active area of research; moreover, design of experiments with statistical analysis is a common practice for finding significant factors, but it has not been used before with NN in the financial field. This research combines DOE and NN to build the best performing NN given a selected list of factors. The results show that it is possible to find the most significant NN parameters using factorial designs. Furthermore, the effect of changing each parameter and interaction on the mean performance can be studied, whether increasing or decreasing. This proposed method leads to a reduction in costs and a reduction in trial and error when building NNs.

REFERENCES

[1] M. Zekic, "Neural network applications in stock market predictions - a methodology analysis," in Proceedings of the 9th International Conference on Information and Intelligent Systems, pp. 255-263, 1998.

[2] K. Assaleh, H. El-Baz, S. Al-Salkhadi, "Predicting Stock Prices Using Polynomial Classifiers: The Case of Dubai Financial Market," Journal of Intelligent Learning Systems and Applications, vol. 3, no. 2, pp. 82-89, May 2011.

[3] I. Kaastra and M. Boyd, "Designing a neural network for forecasting financial and economic time series," Neurocomputing, vol. 10, no. 3, pp. 215-236, 1996.

[4] G. Zhang, B., Patuwo, B., and M. Y Hu, "Forecasting with artificial neural networks: The state of the art," International Journal of Forecasting, vol. 14, no. 1, pp. 35-62, March 1998.

[5] P. K. Padhiary and A. P. Mishra, "Development of improved artificial neural network model for stock market prediction," International Journal of Engineering Science and Technology, vol. 3, no. 2, pp. 1576-1581, March 2011.

[6] C. N. W. Tan and G. E. Witting, "A study of the parameters of a backpropagation stock price prediction model," in Artificial Neural Networks and Expert Systems, Dunedin, pp. 288 – 291, 1993.

[7] P. P. Balestrassi, E. Popova, A. P. Paiva, J. W. MarangonLima, " Design of experiments on neural network's training for nonlinear time series forecasting " Neurocomputing, vol. 72, pp. 1160 – 1178, 2009.

[8] W. Laosiritaworn and N. Chotchaithanakorn, "Artificial neural networks parameters optimization with design of experiments: An application in ferromagnetic materials modeling," Chiang Mai J. Sci., vol. 36, no. 1, pp. 83 - 91, 2009.

[9] R. Behmanesh and I. Rahimi, "Using combination of optimized recurrent neural network with design of experiments and regression for control chart forecasting," in Business Engineering and Industrial Applications Colloquium, Kuala Lumpur, pp. 435 – 439, 2012.

[10] G. S. Atsalakis and K. P. Valananis, "Surveying stock market forecasting techniques - Part II: Soft computing methods," Expert Systems with Applications, vol. 36, no. 3, pp. 5932 - 5941, April 2009.

[11] D. C. Montgomery, Design and Analysis of Experiments, 4th ed. Arizona, United States of America: John Wiley & Sons, 1997.

[12] J. R. Turner and J. F. Thayer, Introduction to Analysis of Variance. Thousand Oaks, California, United States of America: Sage Publicatons, 2001.