A Study on Training Criteria for Financial Time Series Forecasting

JingTao YAO
Dept of Information Systems
Massey University
Private Bag 11222
Palmerston North
New Zealand
J.T.Yao@massey.ac.nz

Chew Lim TAN
Dept of Computer Science
National University of Singapore
1 Science Drive 2
Singapore 117543
Singapore
tancl@comp.nus.edu.sg

Abstract

Traditional backpropagation neural networks training criterion is based on goodness-of-fit which is also the most popular criterion forecasting. How ever, in the context of financial time series forecasting, we are not only concerned at how good the forecasts fit their target. In order to increase the forecastability in terms of profit earning, we propose a profit based adjusted weight factor for backpropagation network training. Instead of using the traditional least squares error, we add a factor which contains the profit, direction, and time information to the error function. This article reports the analysis on the performance of several neural network training criteria. The results show that the new approach does improve the forecastability of neural network models, for the financial application domain

1. Introduction

The criteria of forecasting model performance in traditional time series forecasting are error functions. The error functions are generally based on the goodness-of-fit of target and predict time series. Neural networks especially backpropagation networks are similar to conventional regression estimators except for their nonlinearity. Therefore, the error functions are used in the same way as in regression models to judge the goodness of the model fitting. To apply neural network techniques to the financial context, researchers are working on ways to improve the forecastability of neural network based models [2, 3, 6]. Research shows that conventional statistical techniques for forecasting have reached their limitation in applications with nonlinearities in the data set [5]. It is observed that Normalized Mean Square Error (NMSE) and other error functions that are used for financial forecasting models may not make sense in the financial context [1]. Caldwell [1] gives a general review for the performance metrics. Yao *et al.* [9] use the correctness of trend and a paper profit to judge the performance of neural network forecasting model. Caldwell [1] also proposes a new measure, namely, weighted directional symmetry (WDS), which weights errors based on a function of the accuracy of predicted directions. The above works are based on post evaluation of forecasting model performance.

It has been found by using regression models, weighting all data equally (ordinary least squares) are less accurate than discounted least squares, which weight the most recent data more heavily [4]. This is due to the fact that in financial data, especially low frequency data, the structural relationship between an asset price and its determinants changes gradually over time as the economic environment evolves. Thus recent observations should be weighted more heavily than older observations so that the recent information can be emphasized. The Discounted Least Squares neural network model proposed by Refenes *et al.* [6] is to incorporate time factor to neural network forecasting model building.

We are not only concerned at how good the forecasts fit their targets, but we are more interested in profits in the context of financial time series forecasting. In order to increase the forecastability in terms of profit earning, we will propose a profit based adjusted weight factor for backpropagation network training in this paper. Instead of using the traditional least squares error, we add a factor which contains the profit, direction, and time information to the error function. The results show that this new approach does improve the forecastability of neural network models, for the financial application domain. NMSE is one of the most widely used measurements in neural network training. It represents the fit between the neural network predictions and the actual targets. However, a prediction that follows closely the trend of the actual target would also result in a low NMSE. For pattern recognition, it is a very important signal. We argue that although NMSE is a very important signal for pattern recognition, it may not be the case for trading in the context of time series analysis.

The organization of this paper is as follows. In the next section, we will describe the new approach step by step. A section describing several application case studies then follows where we apply the new models to five time series, namely, United States Dow Jones Industrials Index, Hong Kong Hang Seng index, Malaysia Kuala Lumpur Composite Index, Japan Nikkei-225 and Singapore Straits Times Industrials Index forecasting. The results are presented in the same section. The final section contains some discussion and conclusion.

2. New Models Based on Backpropagation Networks

The backpropagation network [7, 8] is one of the most popular and most widely implemented neural network models especially in the financial forecasting domain. It is based on a multi-layered feedforward topology with supervised learning. The network is fully connected with every node in the lower layer linked to every node in next higher layer. These linkages are attached with weight values. The learning of backpropagation neural network is actually an error minimization procedure. The weights are changed according to an error function which compares the neural network outputs with training targets. The error function is a Least Squares function, or Ordinary Least Squares function, which is shown in Equation 1.

$$E_{OLS} = \frac{1}{2N}\sum_{p=1}^{N}(t_p - o_p)^2 \qquad \textbf{(1)}$$

As we have mentioned above, it may not be enough to rely on absolute errors only in financial applications. The following research hypotheses thus were proposed in this study:

[H1] In addition to ordinary least squares error function, a factor representing the profit could be added to the error function in order to improve the forecastability of neural network models.

[H2] In addition to ordinary least squares error function, a factor representing the time could be added to the error function in order to improve the forecastability of neural network models.

[H3] If profit and time are useful factors of error function to improve the forecastability, an even better result could be achieved by using the combination of both of them.

2.1 Implementation of H1: Directional Profit Model

In financial applications, profit gain is the major goal. To reflect this point in evaluating a forecasting models performance, Yao *et al.* [9] use the correctness of trend and a paper profit. Based on the directional symmetry, Caldwell [1] proposes a Weighted Directional Symmetry (WDS) function. Comparing with the targets, incorrectly predicted directions are penalized more heavily than the correct predictions. The definition of WDS is as follows,

$$WDS = \frac{100}{n}\sum_{p=1}^{N} d_p \,|\, t_p - o_p \,| \qquad \textbf{(2)}$$

where

$$d_p = \begin{cases} g & if\,(t_p - t_{p-1})(o_p - o_{p-1}) \le 0 \\ h & other \end{cases}$$

and *g* and *h* are constants or some function of t_p, *g=1.5* and *h=0.5* as suggested by Caldwell for instance.

WDS is only a post measure of the neural network performance in the financial context. For financial trading, the correctness of trend or direction is sometime more important than the goodness-of-fit. If the correctness is only based on how the forecasts fit the targets, we might not earn what we expected according to our forecasts. This is the main idea of WDS.

To further consider the profit driven procedure of financial trading, we understand that not only the direction but also the amount of change is important. The WDS weights should be adjusted more if a wrong direction is forecasted for a big change. Weights will be adjusted less if a right direction is forecasted for a big change. Based on Hypothesis 1, we then propose another new profit adjust factor which is a function of changes and direction as follows,

$$f_{DP}(p) = F(|\, t_p - t_{p-1} \,|, sign(\Delta t_p, \Delta o_p)) \qquad \textbf{(3)}$$

The Directional Profit adjustment factor is defined as

$$f_{DP}(p) = \begin{cases} a_1 & if\ \Delta t_p * \Delta o_p > 0\ and\ |\Delta t_p| \le \sigma \\ a_2 & if\ \Delta t_p * \Delta o_p > 0\ and\ |\Delta t_p| > \sigma \\ a_3 & if\ \Delta t_p * \Delta o_p < 0\ and\ |\Delta t_p| \le \sigma \\ a_4 & if\ \Delta t_p * \Delta o_p < 0\ and\ |\Delta t_p| > \sigma \end{cases} \quad \textbf{(4)}$$

where σ is a threshold of the changes of sample data. We use the standard deviation of the training data set (including validation set).

$$\sigma = \frac{1}{n}\sum_{p=1}^{N}(t_p - \mu) \quad \textbf{(5)}$$

where μ is the mean of the target series. In the case of our first experiment[1], *a_1=0.5, a_2=0.8, a_3=1.2* and *a_4=1.5*. The new error function will be,

$$E_{DP} = \frac{1}{2N}\sum_{p=1}^{N} f_{DP}(p)(t_P - o_p)^2 \quad \textbf{(6)}$$

2.2 Implementation of H2: Discounted Least Squares Model

The Discounted Least Squares function, which is based on Hypothesis 2, proposed by Refenes *et al.* [6] is given by

$$E_{DLS} = \frac{1}{2N}\sum_{p=1}^{N} w(p)(t_P - o_p)^2 \quad \textbf{(7)}$$

where N, t_p and o_p are as same as in Equation 1 and *w(p)* is an adjustment of the contribution of observation *I* to the overall error

$$w(p) = \frac{1}{1+e^{(a - \frac{2ap}{N})}} \quad \textbf{(8)}$$

The discount rate *a* will decide the function which actually is a function of the recency of the observation as it was used in conventional regression model [4]. We use a=6 as suggested by Refenes *et al*. [6] in this study.

2.3 Implementation of H3: Time Dependent Directional Profit Model

DP Model emphasizes on the profit while DLS model emphasizes on time. Basically, a profit factor, $f_{DP}(p)$, or a time factor, *wp(p),* is incorporated. To include emphasis on the time concept, we incorporate Refenes's approach to our DP model based on Hypothesis 3. We name it Time dependent Directional Profit model (TDP) which is proposed as follows,

$$E_{TDP} = \frac{1}{2N}\sum_{p=1}^{N} f_{TDP}(p)(t_P - o_p)^2 \quad \textbf{(9)}$$

where $f_{TDP} = f_{DP}(p) * w(p)$. The weight changing is the same as the approach used in the traditional OLS backpropagation network except that an adjustment factor is introduced for both DP and TDP models.

3. Experiment Results of Hypotheses

Three models, namely Refenes's Discounted Least Squares (DLS) model, Directional Profit (DP) model and Time dependent Directional Profit (TDP) model are implemented to test on the three hypotheses. They are benchmarked with traditional Ordinary Least Squares (OLS) model. Four major Asian stock market indices, Hong Kong Hang Seng index (HS), Malaysia Kuala Lumpur Composite Index (KL), Japan Nikkei-225 (NK) and Singapore Straits Times Industrials Index (ST) together with United States Dow Jones Industrials Index (DJ) are applied to these four models in our study.

We partition each stock index into seven time series in this study. There are 1960 data entries for each stock under study. Each time series contains 280 trading days of stock indices. For example, the seventh time period of HS indices, chosen for our experiment, ranges from 3 August 1995 to 17 September 1996. The first 260 days are used as training data which include 80 days for validation purpose. The remaining 20 days are used for testing the forecastability of the neural network model. Training period: Monday 3 August 1995 to Thursday 25 April 1996. Validation period: Friday 26 April 1996 to Monday 19 August 1996. Testing period: Tuesday 20 August 1996 to Tuesday 17 September 1996. The time period and the statistic results of the data under study are shown in Table 1.

A one hidden layer network is selected as the neural network model in this study. To simplify our experiment, the learning rate, η, is set to 0.25 and the momentum rate, α, is set to 0.9. The structure of the neural network is 30-10-1, i.e. 30 inputs, 10 nodes in the hidden layer and one output. Thus 30 consecutive days of data are fed to the network to forecast the index of the following day.

---

[1] We use *a_1=0.5, a_2=1, a_3=1* and *a_4=1.5* at first (*a_1* and *a_4* following Caldwell). We then use *h=0.1* as the step and find out that *a_2=0.8* and *a_3=1.2* perform the best among our experiments. Further optimization may be conducted for the *a_i* in future.

As we are focusing on the profit, instead of using NMSE only, we introduce a paper profit [9] measure for the model performance. Table 2 compares the average paper profit achieved by different models for the whole data set of five markets under study. The preference of DLS over OLS is of 60%. The preference of DP over OLS is of 80%. The improvement of these two models in Japan and Singapore markets are significant and consistent. While in the other markets, there is an improvement in Malaysia for DLS model and the United States and Malaysia for DP model. The performance of both models in Hong Kong is worse, although it is only a slightly worse for DLS, than the traditional OLS model. With combination of profit and time factors, the preference of TDP over OLS is 100%. The improvement in Malaysia, Japan, and Singapore is significant. The results shown here have moderately supported H1 and H2 but a full support for H3.

Table 1: Statistic Results for Stock Index Time Series in Five Markets

	Dow Jones	Hang Seng	Kuala Lumpur	Nikkei	Straits Times
Minimum	2243.040	2093.000	327.340	14309.409	993.160
Mean	3524.442	6260.082	738.201	23011.814	1735.970
Maximum	5889.200	12201.089	1314.460	38915.859	2493.710
St. Derivation	902.099	2996.524	253.020	6113.686	417.006
Variance('000)	813.782	8979.157	64.019	37377.965	173.893
Skew	0.955	0.255	0.343	0.990	0.181
First Day	06 Feb 89	08 Nov 88	19 Oct 88	25 Jan 89	06 Dec 88
Last Day	16 Sept 96	17 Sept 96	17 Sept 96	17 Sept 96	17 Sept 96

Table 2: Comparison of Paper Profits Achieved. (Excess: Excess Return over benchmark OLS).

	Dow Jones	Hang Seng	Kuala Lumpur	Nikkei	Straits Times
OLS	16.812	17.812	4.274	16.658	19.389
DLS	7.255	16.322	16.373	22.799	37.870
DP	20.213	1.120	8.359	25.323	29.517
TDP	19.618	24.672	31.897	32.642	37.951
Best Model	DP/TDP	TDP	TDP	TDP	TDP
Excess(TDP)	2.806	6.861	27.623	15.984	18.562
Excess(DP)	3.401	-16.692	4.085	8.665	10.128

4. Concluding Remarks

Most applications of financial time series forecasting rely on the goodness-of-fit as their performance criterion. However, a well fitted prediction times series with the target series does not necessarily mean a good profit achieved by the forecasts. As the ultimate goal of using financial forecasts is to make profit, the financial profitability becomes the major concern of selecting a forecasting model. We propose a profit model to meet the real goal of financial forecasting. We understand that success in one instance does not necessarily mean the same for other cases. Therefore five markets and seven data sets are experimented in our study. The results show that this proposed new criterion does improve the forecastability of neural network forecasting models. It provides another way to improve neural network model in order to suit the application domain requirement.

To emphasize time or profit alone would not be good for financial forecasting. A time dependent profit model combining these two factors has been shown to achieve about 50% improvement with about 3% excess annual return above the benchmark OLS model even for the worst case. We have gained insights from this study in that emphasizing time or profit alone appears not sufficient. The combination of these two factors in our model has been shown to be promising model with up to 27.6% excess annual return above benchmark OLS model or at least 2.8% excess annual return in the worst case.

There are still rooms for further improvement. Instead of using just time delayed indices, other input parameters [6, 9] could be introduced for better forecasting in the future. While the proposed model will be further fine tuned, the present research has been worthwhile in demonstrating the need to consider time dependency and direction of profit change in formulating a new criterion for financial forecasting with neural networks.

References

[1] Caldwell, R.B., "Performances Metrics for Neural Network-based Trading System Development", *NeuroVe$t Journal*, Vol. 3, No. 2. 1995, pp22-26.

[2] Hutchinson, J.M., A.W. Lo and T. Poggio, "A Nonparametric Approach to Pricing and Hedging Derivative Securities via Learning Networks", *The Journal of Finance,* Vol. 49, No. 3, July 1994, pp851-889.

[3] LeBaron, B. and A.S. Weigend, "A Bootstrap Evaluation of the Effect of Data Splitting on Financial Time Series", *IEEE Transactions on Neural Networks,* Vol. 9, No. 1, 1998, pp213-220.

[4] Makridakis, S., "The Accuracy of Extrapolation (Time Series) Methods: Results of a Forecasting Competition", *Journal of Forecasting*, Vol. 1, 1982, pp111-153.

[5] Refenes, A. N., A. Zapranis, and G. Francis, "Stock Performance Modeling Using Neural Networks: A Comparative Study with Regression Models", *Neural Network*, Vol 5, 1994, pp961-970.

[6] Refenes, A.N., Y. Bentz, D.W. Bunn, A.N. Burgess, A.D. Zapranis, "Financial Time Series Modeling with Discounted Least Squares Backpropagation", *Neurocomputing,* Vol. 14, No. 2, 1997, pp123-138.

[7] Rumelhart, D.E., G.E. Hinton, R.J. Williams, "Learning Internal Representations by Error propagation", in: *Parallel Distributed Processing*, Volume 1, D.E. Rumelhart, J.L. McClelland (Eds.), MIT Press, Cambridge, MA, 1986, pp318-362.

[8] Werbos, P.J., *Beyond Regression: New Tools for prediction and analysis in the behavioral Science,* PhD Thesis, Harvard University, 1974.

[9] Yao, J.T., H.L. Poh, "Equity Forecasting: a Case Study on the KLSE Index", *Neural Networks in Financial Engineering,Proceedings of 3rd International Conference On Neural Networks in the Capital Markets*, Oct 1995, London, A.-P N. Refenes, Y. Abu-Mostafa, J. Moody and A. Weigend (Eds.), World Scientific , 1996, pp341-353.