A STUDY OF EFFECT OF TRAINING PERIOD DURATION AND SHARE SPLIT EVENT ON PREDICTING STOCK PRICE USING NEURAL NETWORKS

[1]G. G. RAJPUT, [2]BHAGWAT H. KAULWAR

[1]Department of Computer Science, Rani Channamma University, Belagavi, India
[2]Department of Computer Science, Solapur University, Solapur, India
E-mail: [1]ggrajput@yahoo.co.in, [2]bhagwat_kaulwar@yahoo.com

Abstract - Artificial Neural Network (ANN) is suitable tool for share market prediction due to its ability to learn the data patterns and generalize their knowledge to recognize the future new patterns. In this paper, we introduce an ANN model using Multilayer Feed-Forward NARX network to predict closing price of a share listed in NSE, in particular for SBI. The experiment under study is expected to observe the effect of training period and duration on prediction by ANN. The experiment also studies the performance of ANN before and after share split event. We designed and tested NARX network for SBI data between 2008 and 2016 for different years, to compare the effect of training period on prediction of closing price. We also observed the effect of events like share split on actual closing price and predicted price of SBI.

Keywords - Artificial Neural Network (ANN), Nonlinear Autoregressive with External Input (NARX), stock price prediction, technical analysis, Gradient descent adaptive back-propagation (GDA)

I. INTRODUCTION

A stock market, equity market or share market is the aggregation of buyers and sellers. People invest and trade in share market to gain premium on their capital. Based on the period for which capital was used, the person is categorised as Trader or Investor. The person is treated as trader, if he/she invests capital for short period of time like a day(s), week(s) or month(s). The trader mostly depends on current situation or technical indicators and targets even small premium. On other hand the person is treated as investor, if he/she invests capital for longer period of time like year(s). Investors mostly depend on fundamental analysis of company, economic situation etc. Investor mostly targets for maximum premium.

Investor/trader can earn premium on their capital through market price of a share. However, there are other means like dividend, bonus or split, through which premium can be earned. These events are very seldom like dividend once or twice in a year and bonus and split after several years. The event like bonus or split has large impact on share price. For example, if a share having face value of Rs. 10/- and price of Rs. 1000/- gets split into 5 shares, then its new face value will be of Rs. 2/- and a price of Rs. 200/-. The quantity of shares will now be 5 instead of 1, to insure total share capital of a company. In other words, investor/trader will have 5 shares of Rs. 200/- instead of 1 share of Rs. 1000/-.

The event like share split is used to increase the liquidity or trading volume of a share. If a share value is higher, may be in thousands, it will be less traded giving less chance for price increase. In such situation, company splits share into multiple shares of lesser amount which leads to attention of large extent of traders and investors. It is observed that share gets considerably high gains in price even within short period. For example, share of SBI got split on 19th Nov 2014 with price from Rs. 2910.50 to Rs. 291.05. But soon on 01st Dec 2014, share recorded a price of Rs. 326.95 as high price, which indicates 12.33 % of returns within 8 trading days. Similarly, ICICI share got split on 03rd Dec 2014 from Rs. 1790.00 to Rs. 358.60. But soon it recorded a high price of 393.40 on 28th Jan 2015 giving 9.70 % returns within 38 trading days.

With advances in technology, there has been considerable increase in research for predicting share market price using data mining techniques. Prediction can be utilized in decision making, whether to buy or sell the particular shares of a given stock. Neural Networks (NN) applied to individual stocks can yield statistically significant predictions of INDEX or stock prices. The NN has two phases, viz. training and testing. The data used in both phases is mutually exclusive. In most of NN applications, larger the data used for training of network, more precise will be the model and lesser will be prediction error. In other words data from multiple years is used to predict single value. In this paper, we analyzed the effect of training period and data on prediction of closing price. Though in most of NN applications training phase has large data, but in share market, value of a share has seasonal effect. Like good or bad monsoon will impact positively or negatively on market. The sentiments of such events can have adverse impact on share. In other words, because of positive sentiments in market, share can record increase in price even if a particular share is in downtrend. After the event is discounted in price, the market will react differently and that particular share will continue its downtrend. This type of behavior is called as trap. If investor/trader buys such stock at high price, they will be trapped for longer time or they have to book heavy losses. In order to observe such patterns, a NN was trained with one year data, two year data and three year data and their performance was measured.

II. LITERATURE SURVEY

Because irregular fluctuations occur in stock market, it is very difficult to model its behavior. Stock market can be considered as non-linear deterministic system [1]. ANN has ability to discover nonlinear relationship in input data set without a priori assumption of the knowledge of relation between input and output [10]. Many researchers have proposed different types of models for prediction of an Index or stock. Few of them are described below.

Catalina-Lucia Cocianu [5] et al. conducted a study to compare results of ARIMA model with that of NARX model. They used mean square error (MSE) to evaluate both the models. They used 300 samples of weekly observations of SNP stock, from Bucharest Stock Exchange, between 01/03/2009 and 30/11/2014. The dataset contained opening, closing, highest and lowest price of SNP stock along with seven technical and fundamental analysis indicators. The obtained results were encouraging and entail future work toward extending study in case of using alternative neural models.

Goutam Dutta [6] et al. worked on ANN to predict weekly closing price of SENSEX index of Bombay Stock Exchange, India. They constructed two ANNs, ANN1 and ANN2 with different inputs like weekly closing price of SENSEX, 52-week moving average of weekly closing price, and 5-week moving average of closing price. The other input to ANN1 was 10-week Oscillator for past 200 weeks. The other input to ANN2 was 5-week volatility for past 200 weeks. Both the network was trained using 250 weeks data starting from Jan 1997. The network performance was measured for two-year period beginning Jan 2002. They achieved Mean Absolute Error (MAE) of 3.93 % through ANN1. They also found that Root Mean Square Error (RMSE) dropped from 4.82 % to 4.40 % for ANN1 and from 6.87 % to 6.596 % if last 10 weeks' values were dropped from validation set. They observed that the error towards the end of validation set was higher than the earlier values.

Gitansh Khirbat [7] et al. demonstrated ANN with one input layer, one output layer and two hidden layers. The first hidden layer contained 22 neurons and second hidden layer contained only 2 neurons in it. They used data of 44 days to train the network with 50/500/5000 epochs. They observed that the rate of prediction increases with number of epochs. The ANN forecasted more accurately in 5000 epochs as compared to 50 epochs.

In this paper, we propose a NN design to predict the share price of a share listed in NSE and also consider the effect of split and other factors that influence the Index or share price.

III. ARTIFICIAL NEURAL NETWORK

From the literature review it is clear that, ANN can be constructed to predict a future value of an Index or particular share. During the study, we considered a special class of ANN called Nonlinear Autoregressive with External Input (NARX) model. NARX networks are recurrent neural networks and are well suited for modeling nonlinear systems and specially time series. NARX networks with Gradient descent adaptive back-propagation learning algorithm are considered as better not only because learning is more effective in NARX networks but also they converge much faster and generalize better than other networks [3][4].

Network Architechture

A multilayer feed-forward (MLP) NARX network with back-propagation algorithm is used for the experiments under study. There are basically three layers viz. input layer, output layer and hidden layer. Depending on application under study, there can be more than one hidden layers in the network.

The NARX model can be implemented in many ways, but most commonly used model is shown in figure I which uses a feed-forward neural network with embedded memory.

For architectural model shown in Fig. I, the notation used is NN(d_u,d_y:N) to denote NN with d_u input delays, d_y output delays and N neurons in layer 1. The next level value of time series y(k+1) can be computed by using model in figure I with the past observations u(k), u(k-1),.....,u(k-d_u) and past outputs y(k), y(k-1), ..., y(k-d_y) as inputs [8].

MLP is a feed-forward Neural Network with one or more layers between input and output layer. Feed-forward means that data flows in one direction from input to output layer. From figure I, it is clear that, input data are fed to the neurons in the input layer. After processing within the individual neurons, output values of input layer are given to neurons in hidden layer. Similarly after processing within individual neurons, output of hidden layer is fed to output layer[11].

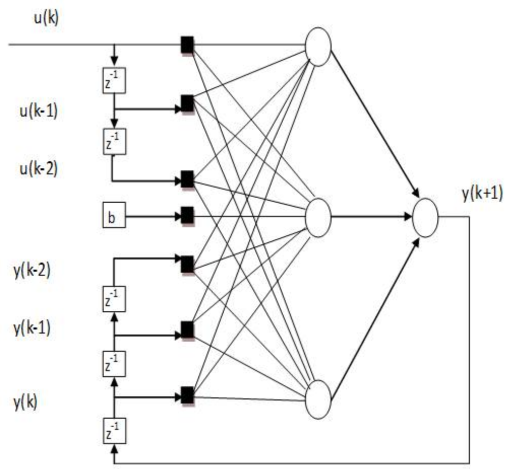


Fig I : NARX model with tapped delay line at input

The connections between neurons are associated with weights. During network training phase, the output of output layer is given back to network as a feedback. If the difference of actual and predicted values is not within acceptable range, the weights are adjusted according to a training algorithm. The back-propagation training algorithm consists of two pass, a forward pass and backward pass. In forward pass weights are fixed and input is applied to input vector. The processed inputs are forwarded from input to hidden layer and from hidden layer to output layer. The output of output layer is considered as network response. After network response is arrived, an error term is calculated by obtaining difference between actual response of the network and expected response specified to network. In backward pass the response of network is given back to network and weights are adjusted so as to minimize the network error i.e. make actual response of network become closer to desired response[12].

IV. DATA AND METHODOLOGY

We have observed that technical parameters are more useful to predict share price. We used daily transaction data like open price, high price, low price, volume, moving averages, deliverable percentage etc. We used daily transaction data of SBI for training and testing NN network. The data collected from National Stock Exchange website for study contains information about technical parameters from the year 2008 to 2016.

A. Data used for network

The company chosen for study is State Bank of India. The data of 2008 to 2016 was used to train and test the ANN. The daily trading data of SBI between 2008 and 2016 is taken from National Stock Exchange website [2]. The National Stock Exchange (NSE) is the leading stock exchange in India and the fourth largest in the World by equity trading volume in 2015, according to World Federation of Exchanges (WFE). It began operations in 1994 and is ranked as the largest stock exchange in India in terms of total and average daily turnover for equity shares every year since 1995, based on annual reports of SEBI. NSE launched electronic screen-based trading in 1994, derivatives trading (in the form of index futures) and internet trading in 2000, which were each the first of its kind in India [9].

Through study we found that NARX network with 12 input variables, 1 hidden layer with 10 neurons, 1 output neuron, feedback delay of 2, Gradient descent adaptive back-propagation (GDA) training algorithm, symmetric sigmoid transfer function in hidden layer and pure linear transfer function in output layer is optimal network. We used daily transaction data that includes close price of SBI on previous trading day, open price of SBI for current trading day, high price of share on trading day, low price of share on trading day, average price of share on trading day, total transactions that trading a share on trading day, total turnover in Rs. for trading share on trading day, and moving averages of shares like 10DMA, 20DMA, 50DMA, 100DMA and 200DMA. Total there are 12 inputs to network. The closing price of the share on trading day is used as output of the network.

B. Network Training

As stated in section IV A, we constructed 19 different structures of NARX network with 12 inputs, one hidden layer with 10 neurons, feedback delay of 2, one output neuron, Gradient descent adaptive back-propagation (GDA) training algorithm, symmetric sigmoid transfer function in hidden layer and pure linear transfer function in output layer. All 19 structures were trained for different period and tested accordingly. Following table I gives details of structures along with training period.

For most of structures, the training period and testing period is same, except for Structure 14 and 19. For Structure 14 training period was of two years and testing period was of one year. This was done to observe effect of share split event on training the NN. For structure 19, training period was of six years and testing period was of four years. This was done to observe networks behavior for longer period of training.

Table 1: Details of training period & data used to train ANN

Sr. No.	Structure	Training period	Training period in years	Testing period	Testing period in years
1	Structure 1	2008	One year	2009	One year
2	Structure 2	2009	One year	2010	One year
3	Structure 3	2010	One year	2011	One year
4	Structure 4	2011	One year	2012	One year
5	Structure 5	2012	One year	2013	One year
6	Structure 6	2013	One year	2014	One year
7	Structure 7	2014	One year	2015	One year
8	Structure 8	2015	One year	2016	One year
9	Structure 9	2009-10	Two years	2011-12	Two years
10	Structure 10	2010-11	Two years	2012-13	Two years
11	Structure 11	2011-12	Two years	2013-14	Two years
12	Structure 12	2012-13	Two years	2014-15	Two years
13	Structure 13	2013-14	Two years	2015-16	Two years
14	Structure 14	2014-15	Two years	2016	One year
15	Structure 15	2008-10	Three years	2011-13	Three years
16	Structure 16	2009-11	Three years	2012-14	Three years
17	Structure 17	2010-12	Three years	2013-15	Three years
18	Structure 18	2011-13	Three years	2014-16	Three years
19	Structure 19	2000-05	Six years	2006-09	Four years

C. Performance Measurement

To measure the prediction accuracy of the model, the predicted values were compared with actual outputs of sample data. Normalized Mean Square Error (NMSE) is used to evaluate prediction of accuracy of the model. Following formula is used to calculate NMSE.

$$NMSE = \frac{\sum_{t=1}^{N1} (P_t - O_t)^2}{\sum_{t=1}^{N1} (P_t - \overline{P}_t)^2} \quad (1)$$

Where P_t represents actual value of the pre-processed data series i.e. closing price of share, O_t represents observed value or the predicted value i.e. predicted closing price of share for the same day and P_t is the mean of the actual value.

In order to calculate the error percentage, actual closing price and predicted closing price were compared. The formula to calculate error percentage is as follows.

$$Error\% = \frac{|P_t - O_t|}{P_t} * 100 \quad (2)$$

Where P_t represents actual closing price and O_t represents observed or predicted closing price of share. From formula it is clear that, a network with less error % should be considered as best network. The value of 0 for error % indicates that there are no errors in actual and predicted values which indicate perfect prediction.

V. EXPERIMENTAL RESULTS

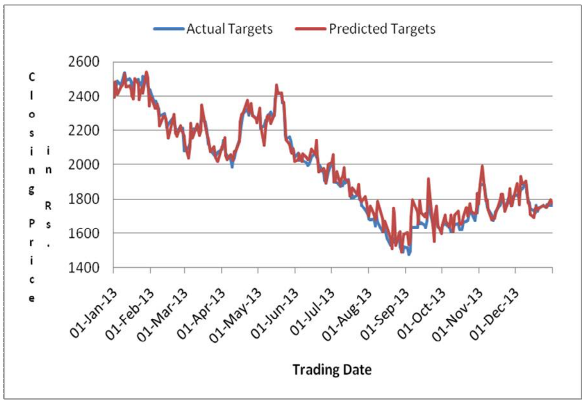
After predicting share prices, performance of all 19 structures were calculated by using Eq (1) and (2). Table II presents the performance details of 19 structures.

Table 2: Performance measure details of ANN structures used in study

Sr. No.	Structure	NMSE	Error %
1	Structure 1	0.06	5.66
2	Structure 2	0.17	5.38
3	Structure 3	0.13	5.13
4	Structure 4	0.42	4.32
5	Structure 5	0.03	1.93
6	Structure 6	0.21	35.45
7	Structure 7	1118.04	302.94
8	Structure 8	0.15	5.46
9	Structure 9	0.39	7.06
10	Structure 10	0.28	5.19
11	Structure 11	0.41	29.37
12	Structure 12	0.26	135.43
13	Structure 13	2201.16	770.97
14	Structure 14	7.56	33.21
15	Structure 15	0.23	5.33
16	Structure 16	0.42	18.12
17	Structure 17	0.08	49.39
18	Structure 18	1.21	232.67
19	Structure 19	0.47	13.69

From table II, minimum NMSE of **0.03** was observed for Structure 5. Also minimum error % was observed for Structure 5 which recorded error of 1.93 %. So Structure 5 with training period of one year and testing of one year is optimal structure of the network. The comparison of closing price of training data and network outputs is shown in plot 1. The comparison between actual closing prices and predicted closing prices of SBI share is shown in Plot 2.

Plot 1: Plot of actual closing price of training data and models outputs for Structure 5.



Plot 2: Plot of actual and predicted closing prices of SBI for Structure 5.

From table I, it is clear that, for structures 1 to 8 training period is of one year. Also, for structures 1, 2, 3, 4, 5, and 8 recorded NMSE and error % are substantially good as compared to all other structures. Structure 6 and 7 gave irrelevant results because split period was part of either training or testing phase. From this observation it clear that one year training period gives better results as compared to other compositions like two years, three years, etc.

From table II, it is clear that when year 2014 is considered in either training or testing phase, network fails to predict proper values which lead to higher error percentage and NMSE. Structures 6, 7, 11, 12, 13, 14, 16, 17, 18 are related to year 2014 either in training or testing phase. And for all these networks error % is very high. Structures 1, 2, 3, 4, 5, 8, 9, 10, 15 and 19 are not related to year 2014 and all of them have given lesser NMSE and good or average error %. In Nov 2014, one share of SBI was got split into ten shares which lead to drastic change in share price. Share price of SBI on 19th Nov 2014 was 2910.50/-. So on 20th Nov 2014, one share of SBI was converted to 10 shares with a new price value of 291.05/-. It is observed that, for such an event, network identified drastic change in predicted values, but they were not close to actual closing price. Plot 3 is for actual and predicted prices of shares for Structure 6, which was trained for year 2013 and tested for year 2014. From plot it is clear that, predicted values also decreased in same direction but after that, they were not close to actual values.

Plot 3: Plot of actual and predicted closing prices of SBI during year 2014 for Structure 6.

If we ignore split data from analysis, the performance of network is improved substantially. For example, for structure 6, NMSE was recorded as 0.21 and error % of 35.45. But after ignoring split data NMSE was improved to 0.19 and error % was improved substantially to 7.29. Plot 4 shows comparison of actual close values and predicted values after excluding data after split event.

Plot 4: Plot of actual and predicted closing prices of SBI during year 2014 for Structure 6 after ignoring split data.

CONCLUSION

ANN based NARX model has been studied in this paper for prediction of closing price of SBI. A feed-forward back-propagation NARX network with three layers viz. input, hidden and output, with 12 input variables, 10 hidden neurons, 1 output neuron, feedback delay of 2, Gradient descent adaptive back-propagation (GDA) training algorithm, Symmetric sigmoid transfer function in hidden layer and pure linear transfer function in output layer is considered for all structures of network. The network 5 with training period of one year (2012) and prediction year 2013 gave optimal solution with NMSE 0.03 and error percentage of 1.93. The experimental results showed that NARX neural network can be designed and used for predicting the share prices of a particular company, even for highly volatile Indian stock market.

We also observed that, for a short term period share prices are affected more by current sentiments and seasonal parameter than fundamental parameters. The optimal network for one year training data is Structure 5 with NMSE of 0.03 and error of 1.93 %. The optimal network for two year training data is Structure 10 with NMSE of 0.28 and error of 5.19 %. The optimal network for three year training data is Structure 15 with NMSE of 0.23 and error of 5.33 %. From this data and observation it is clear that data should be specific. Hence unlike most of NN applications, the data used for training NN shouldn't be more. Even a data of year is enough to train the NN. In other words networks with training data of shorter period gave good results as compared to NN with training data of larger period.

From this study, we also conclude that the events like share split can't be trained to NN. NN can predict the change but the gap between actual and predicted values will be more after split event. It indicates that, a NN network has to be retrained for specific period after events like share spilt or bonus. For example, the error % for Structure 11 with training period of 2011-12 and testing period of 2013-14 was 29.37 and NMSE was 0.41. But if we ignore data after share spilt event, error % decreased to 5.54 and NMSE record was 0.12. Similar observations were found in other structures where year 2014 is included in either phase. E.g. Structure 6 recorded NMSE of 0.21 and error % of 35.45. But if we ignore split data from analysis, NMSE was dropped to 0.19 and error % was dropped to 7.29. From these observations it is clear that, same NN works properly till share spilt event but behaved unexpectedly after event. Even we used two years of data (2014-15) to train network and predicted 2016 values with Structure 14, but both performance measures NMSE of 7.56 and Error % of 33.21 were high as split event data was a part of training phase. So it is recommended that, the NN trained before share spilt event shouldn't be used after event. After specific period is passed with event, the NN has to be retrained with new data. Structure 8 indicates the same. NN in structure 8 was trained for 2015 data, data of entire year after share split event, and tested for data of 2016. The NMSE of structure 8 was recorded as 0.15 with error of 5.46 %, which is better if compared with structure including year 2014 in either training or testing.

ACKNOWLEDGMENT

This research paper is made possible through the help and support from parents, teachers, family, friends and colleagues of Ph. D.
I would like to thank Prof. S. D. Lokhande, Principal, Sinhgad College of Engineering, Pune for his continuous guidance and suggestions for developing

the NN model. Despite busy schedule and positional responsibilities, they have always given me time for discussion which leads to improve the experimental model.

I thank to my fellow Ph. D. colleagues for their support and continuously being there to help me whenever I needed.

I sincerely thank to my parents, family and friends for their consistent support and encouragement for study. My parents are always my source of inspiration. They have always supported me during my educational journey. It is because of them, I have achieved this much of success.

I thank everyone who have directly or indirectly supported and helped me during this process of research.

REFERENCES

[1] Abhyankar, A. Copeland, L.S., & Wong, W., Uncovering nonlinear structure in real-time stock-market indexes: The S&P 500, the DAX, the Nikkei 225, and the FTSE-100, Journal of Business & Economic Statistics, 15, 1–14., 1997

[2] www.nseindia.com/products/content/equities/eq_security.htm

[3] Tsungnan Lin, Bill G. Horne, Peter Tino, C. Lee Giles, Learning long-term dependencies in NARX recurrent neural networks, IEEE Transactions on Neural Networks, Vol. 7, No. 6, 1996, pp. 1329-1351

[4] Yang Gao, Meng Joo Er, NARMAX time series model prediction: feed-forward and recurrent fuzzy neural network approaches, Fuzzy Sets and Systems, Vol. 150, No. 2, 2005, pp.331-350

[5] Cătălina-Lucia COCIANU, Hakob GRIGORYAN, An Artificial Neural Network for Data Forecasting Purposes, Informatica Economică vol. 19, no. 2/2015, DOI: 10.12948/issn14531305/19.2.2015.04

[6] Goutam Dutta, Pankaj Jha, Arnab Kumar Laha & Neeraj Mohan, Artificial Neural Network Models for Forecasting Stock Price Index in the Bombay Stock Exchange, JOURNAL OF EMERGING MARKET FINANCE, 5:3 (2006)

[7] Gitansh Khirbat, Rahul Gupta and Sanjay Singh, Optimal Neural Network Architecture for Stock Market Forecasting, Communication Systems and Network Technologies (CSNT), 2013 International Conference on , pp.557-561, 6-8 April 2013

[8] Eugen Diaconescu, The use of NARX Neural Networks to predict Chaotic Time Series, ISSN:1991-8755, Issue 3, Volume 3, March 2008.

[9] https://www.nseindia.com/global/content/about_us/about_us.htm

[10] M.T. Hagan, H.B. Demuth and., M. Beale, "Neural network design". PWS Publishing Company, Boston, 1996.

[11] A. Victor Devadoss and T. Antony Alphonnse Ligori, "Forecasting of Stock Prices Using Multi Layer Perceptron", International Journal of Computing Algorithm, Volume 2, December 2013, Pages 440-449

[12] Simon Haykin, "Neural Network A comprehensive Foundation", Second edition, Prentice Hall, pp. 161-173, 1998.

★★★