IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-4, Apr 2019 E-ISSN: 2347-2693

New Approaches for Stock Index Prediction Using Artifical Neural Networks

Navin S. Patel[1*], Y.T. Krishne Gowda[2]

[1]Dept. of Management Studies, University of Mysore, India
[2]MIT Thandavapura, Mysore, India

[*]*Corresponding Author: navinsp@gmail.com, Tel.: 9731763244*

DOI: https://doi.org/10.26438/ijcse/v7i4.11371141 | Available online at: www.ijcseonline.org

Accepted: 21/Apr/2019, Published: 30/Apr/2019

***Abstract*—** Stock index prediction is a continually evolving field. Prediction is used for technical and fundamental analysis by both short term traders and long terms investors. There are a multitude of theories and methodologies that exist in the area of stock index prediction. These range from older time-series econometric models to newer Artificial Neural Network (ANN) models. The scope and applicability of ANN is widening rapidly with newer and powerful architectures being proposed in the past few years. Significant amount of work has been done using ANN for stock index prediction, but most of it is done on a class of network architecture called Multi Layered Perceptron (MLP). In this work, we make use of two relatively newer ANN architectures – Long-Short Term Memory (LSTM) and Convolutional Neural Network (CNN) for stock index prediction. We find that these two models offer better forecast accuracy compared to the MLP model for the chosen stock index.

***Keywords*—**Stock index prediction, MLP, LSTM, CNN

I. INTRODUCTION

Stock market is a highly sought after place for investment opportunities for both investors and traders. It is a good medium for both short term and long term investors for making profits. Likewise, it also plays an important role for corporate bodies as it is one of their main sources to raise money. Being involved in stock market is risky as the decision making process is not straightforward. With high proliferation of financial and related information of listed companies, made possible by the internet, the job of the investors has been made very difficult, as they will have to gather, analyze, sift all these data to make correct decisions. This information includes historical information, real-time information and macro-economic information. Since long, India has had functioning capital markets and the stock market is an important part of it. The reforms of 1990s have resulted in increased participation of companies in the equity markets which in turn has aroused interest of domestic retail investors, institutions and foreign investors. A major parameter which influences the investment decision is the expectation of future returns and hence forecasting methods play an important role. Traditional time-series techniques and fundamental analysis have been playing a leading role in many of these decisions. The world of computing is witnessing incredible and fast paced changes. This has resulted in newer forecasting methods generally known as 'Soft Computing' methods. Artificial Neural Networks (ANN) belong to this group. Within ANNs, the last few years have seen newer and better architectural paradigms mainly to solve problems in vision, speech and natural language processing. These new developments can also be adopted for time-series forecasting. Much of the earlier work using ANN for time-series forecasting involves a class of architecture called Multi Layered Perceptron (MLP) which is based on feed-forward mechanism.

In this study, we develop a MLP model and two more ANN models – long-short term memory (LSTM) and Convolutional Neural Network (CNN) models and explore if these additional models are an improvement over the MLP based model.
We use National Stock Exchange (NSE) S&P 50 index as the time-series data and:

1. Create a MLP based forecast model and capture forecast accuracy
2. Create LSTM and CNN based models and capture forecast accuracies
3. Compare these three models based on their forecast accuracies

The paper is organized as: Section I contains a brief introduction of stock index forecasting, Section II contains a discussion of previous work done in time-series forecasting using ANN, Section III describes the data used in the study, forecast metrics and the ANN models created, Section IV contains the parameters of the models created and the forecast metrics, Section V describes results and discussion with suggestions for future research

II. RELATED WORK

In [1], the authors proposed a ANN based approach for technical analysis of the TOPIX index to be used for predicting future values. Their unique approach improved the forecast accuracies.

In [2], the authors evaluated the ANN approach for predicting the NSE Nifty 50 Index. Instead of predicting future values, they used the model to predict the direction of the end of day closing value of the time-series. The model was used to predict direction price index value of the stock market. After exploring different features of the ANN model, an optimal model was proposed. The training data set included the trading days from 01-01-2000 to 31-12-2009. 4 years of trading days was used as the validation set. Out-of-sample performance metric was used to measure performance of the neural network The results showed that the network correctly predicted the direction 89.65% times.

In [3], the authors proposed two types of ANN configurations to model the Tehran Stock Exchange index for predicting the future values. MLP based ANN and Elman ANNs were the two models they used. They used MAPE, MAD, and RMSE to capture forecast accuracy. The MLP based model performed better that the Elman based model in terms of lower forecast errors.

In [4], the authors created a MLP based feed forward neural network. This was used to model the one-month daily price of 3 stocks in the Nigerian Stock market. They used a 4-day rolling window-size to predict the price of the next day. MSE was used as the metric to measure the next day forecasting error. The model showed forecasting accuracy rates for the 3 stocks being 50%, 83.3% and 83.3%.

In [5], the authors presented a MLP based ANN approach to predict stock indices. The model had customizable parameters. They used a multitude of activation functions along with features for cross-validating different sets. They tested their algorithm on the Nifty stock index and predicted the future values. A best case accuracy of 96% on the dataset was achieved.

In [6], the authors developed feed-forward MLP based ANN to predict future value of 2 indices – NYSE and Nairobi Stock Exchange. They used daily closing price values of five years from 2008 to 20012. RMSE and MAPE were used as prediction performance metrics. In their simulations, they obtained very low MAPE of 0.71%

III. METHODOLOGY

Data:
We decided to use the NIFTY S&P 50 index as the time-series for the ANN models. We selected a period of extreme volatility so as to create robust models to achieve superior prediction performance. We selected the period from 8-Nov-2010 to 19-Dec-2011 when the index value dropped by 26.46%. The series had a total of 261 values which were used as training data for the ANN models. The future 15 values were predicted using the models. The predicted values were then compared with the actual values to derive the forecast metrics.

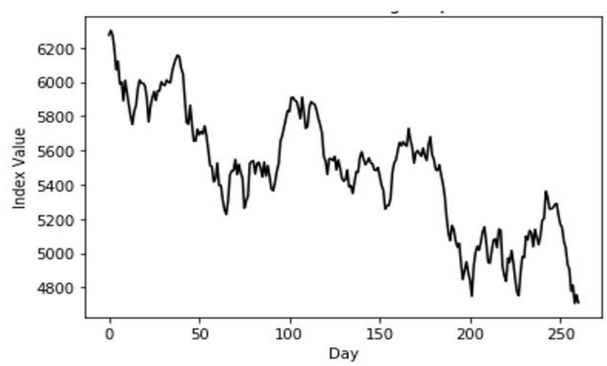

Figure 1: Time-series used in the models

Forecast Metrics:
The main goal of creating robust time-series models is to have good predictability. The models should have low forecasting errors. The central premise of evaluating any forecasting technique is to capture the deviation of the predicted values from the actual value. This is called 'error'. Restricting the evaluation to a single data point of actual values vs. predicted value is neither wise nor useful. Hence, evaluation of the efficacy of forecasting techniques require consideration of more than one data point. In this study, three forecasting horizons were considered: 5-days ahead, 10-days ahead and 15-days ahead.
In this study three metrics were used.

RMSE: The Root Mean Square Error (**RMSE**) measures the difference between values forecasted by a model and the values actually observed. The RMSE is defined as the square root of the mean squared error:

$$\text{RMSE}= \sqrt{\frac{1}{N}\sum_{i=1}^{N}(A_i - F_i)^2}$$

where A_i is actual observed values and F_i is the forecasted value.

MAD: Mean Absolute Deviation tracks the degree of forecast error by summing up the absolute deviations over the period of prediction divided by the prediction period.

$$\text{MAD} = \frac{1}{N}\sum_{i=1}^{N}|A_i - F_i|$$

MAPE: Mean Absolute Percentage Error (MAPE) is the average absolute percent error for each predicted value minus actuals divided by actuals.

$$\text{MAPE} = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{A_i - F_i}{A_i}\right|$$

Metrics were captured for n = 5,10 and 15.

Windowing Approach:

To setup a model for training using ANN, the training data has to be setup in terms of input data and expected output data. As the time-series is auto-correlated in nature, a windowing method is used to setup the training data. In this, a window of *w* consecutive data points is selected. These act as the input data. The *w+1* data point will be the output point for this training sample. For the next training sample, the window moves to the right. Fig. 2 illustrates this for a window size of 4. During the creation of models, the window size is decided by trial and error based on forecasting results

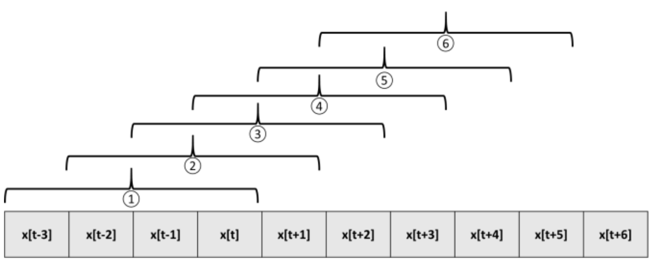

Figure 2: Windowing Approach

MLP-ANN Model

Fig. 3 illustrates a general setup of a MLP based ANN. This shows a window size of four, one hidden layer with five neurons and one output layer. The actual configuration depends on the trial and error.

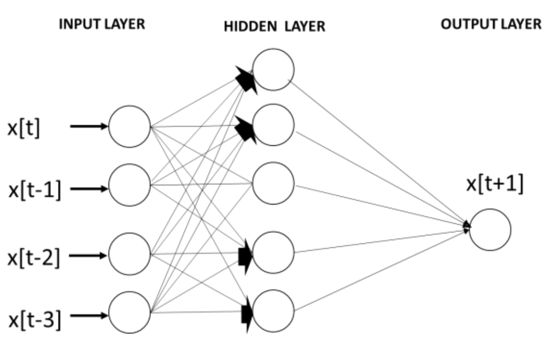

Figure 3: MLP-ANN Model

LSTM-ANN Model

In the conventional MLP-ANN model, the training pattern is considered to be independent, that is, the data at any time instant is not related to each other. This is because it is a feed-forward network where there is no feedback from any layer back into the preceding layers. LSTM architecture is different from these models as they introduce dependency using feedback loops. They can selectively remember or forget values and this is achieved by creating three gates: forget gate, input gate and output gate.

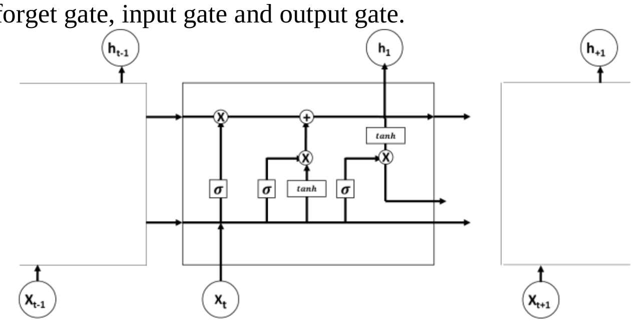

Figure 4: LSTM-ANN Model

Fig.4 illustrates this architecture where x_{t-1} relates to the forget gate, x_t refers to the input gate and x_{t+1} refers to the output gate. The same windowing technique is used for LSTM-ANN model as well.

CNN-ANN Model

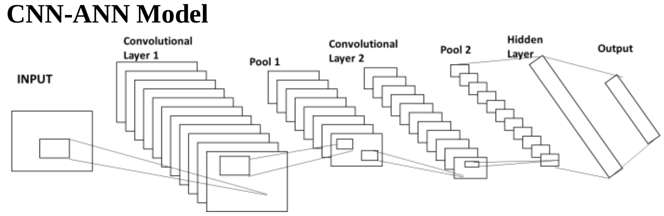

Figure 5: CNN-ANN Model

A CNN is a deep learning model which involves a number of convolutional layers and pooling layers between the input and the output layers. Kernels are used on the convolutional layers to abstract features from them. Though it was designed to work on 3-D images, this architecture can be used for 1-D time-series prediction. The windowing technique is similar to other two models. This model has multiple configurable parameters like the number of convolutional layers, pool size, kernel function, stride size etc. Typically, it is much slower to train than the other two models.

The training of any ANN is based on trial and error. This is due to the fact that many parameters are configurable. In this present instance, following parameters were tuned to get the best forecast metrics

1. The window size for the input which decided the number of input neurons
2. The number of hidden layers, the number of neurons in each layer and the activation function for the neurons.
3. The number of times the test data is passed via the network (epochs) and the batch size
4. The kernel function, the number of convolution layers, pool size (in CNN-ANN).

5. The number of LSTM layers (in LSTM-ANN)

The process of model creation and forecast metrics measurement process is done by trial and error and the best architecture for each type of model is chosen based on lowest forecast accuracy as measured by RMSE.
Fig. 6 shows the high level flowchart of the model training and forecast metrics measurement process.

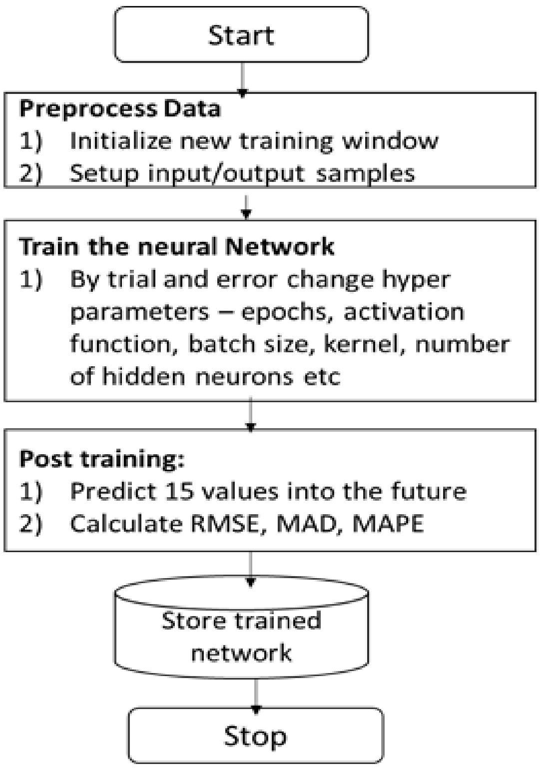

Figure 6: Training and Forecasting Flow

IV. RESULTS AND DISCUSSION

The model parameters and forecast performance metrics are listed below

Table 1: MLP-ANN Model Training Parameters

Parameters	**Value**
Window Size	12
No of Hidden Layers	1
Neurons in Hidden Layer	50
Epochs	50
Batch Size	100

Table 2: MLP-ANN Forecast Metrics

Forecasting Window	**RMSE**	**MAD**	**MAPE (%)**
5- Days	198.14	136.82	2.74
10-Days	223.41	172.03	3.44
15-Days	188.92	140.56	2.84

Table 3: LSTM-ANN Model Training Parameters

Parameters	**Value**
Window Size	15
No of LSTM units	50
No of Neurons	100
Epochs	100
Batch Size	100

Table 4: LSTM-ANN Forecast Metrics

Forecasting Window	**RMSE**	**MAD**	**MAPE (%)**
5- Days	208.32	161.85	3.26
10-Days	204.66	167.36	3.36
15-Days	172.71	132.03	2.67

Table 5: CNN-ANN Model Training Parameters

Parameters	**Value**
Window Size	20
No of parallel filters/kernels	32
Pool Size	2
Epochs	100
Batch Size	100

Table 6: CNN-ANN Forecast Metrics

Forecasting Window	**RMSE**	**MAD**	**MAPE (%)**
5- Days	180.391	161.17	3.32
10-Days	142.41	117.92	2.42
15-Days	169.61	145.60	3.04

Three forecast error metrics were captured for three ANN variations for three different forecast horizons. Usually, RMSE is used to rank the models based on forecasting accuracy. Lower RMSE implies a better model. The ability of a model to have a lower error for a longer forecast horizon indicates its robustness. It can be seen that RMSE, MAD and MAPE values of CNN-ANN models for all the three forecast horizons are the lowest compared to MLP-ANN and LSTM-ANN variations. LSTM-ANN metrics are better than MLP-ANN except for the 5-day forecast window. Hence, we can rank the models in following order: CNN-ANN, LSTM-ANN and MLP-ANN. Hence CNN-ANN and LSTM-ANN models offer a better alternative to the often used MLP-ANN model.

V. CONCLUSION AND FUTURE SCOPE

In this paper we presented two powerful alternatives to the hitherto standard MLP-ANN model used for predicting time-series. With the world of ANN undergoing a massive revolution due to the mainstreaming of Artificial Intelligence techniques, newer ANN paradigms are evolving and many these can be evaluated for time-series forecasting. One potential area of future study could be to create hybrid models of LSTM and CNN architectures and explore if they improve the performance of the standalone models.

REFERENCES

[1] Mizono, H., Kosaka, M., Yajma, H., and Komoda, N. (1998). "*Application of Neural Network to Technical Anaysis of Stock Market Prediction*", Studies in Informatics and Control, **Vol.7, No.3, pp.1l1-120**

[2] Majumder, M., and Hussian, M.D.A.(2009). "*Forecasting Of Indian Stock Market Index Using Artificial Neural Network*", National Stock Exchange of India Limited

[3] Naeini, M.P., Taremian, H., and Hashemi, H.B. (2010). "*Stock Market Value Prediction Using Neural Network*", CISIM, **pp: 132 - 136**

[4] Neenwi, S., Asagba, P.O., Kabari, L.G. (2013). "*Predicting the Nigerian Stock Market Using Artificial Neural Network*", European Journal of Computer Science and Information, **Vol I , No.1, pp.30-39**

[5] Kar, A. (2013). "*Stock Prediction using Artificial Neural Network*", Dept. of Computer Science and Engineering, IIT Kanpur

[6] Wanjawa, B.W., and Muchemi, L (2014). "*ANN Model to Predict Stock Prices at Stock Exchange Markets*", Xiv:1502.06434

Authors Profile

Navin S Patel did his Bachelor of Engineering Degree from University of Mysore, and later completed his MBA degree in Finance by topping his university. He is currently pursuing Ph.D. while he is employed in the software industry.. His research work focuses on deep learning and econometric models for time-series prediction.

Dr Y.T Krishne Gowda is the principal of MIT Thandavarapura, Mysore which is part of VTU. He has done his ME from IISc, Bangalore and PhD from IIT, Madras. He has more than 30 years of teaching experience and published more than 20 papers. He has also worked in AICTE as a director.