

“© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.”

Financial Time Series Stock Price Prediction using Deep Learning

Madhu Goyal
Australian Artificial Intelligence Institute
School of Computer Science
University of Technology Sydney
Australia
madhu.goyal-2@uts.edu.au

Abstract— Several research studies have been devoted for the last two decades to make estimates on or to forecast stock prices. Accurate stock prediction movement is still an open question for many companies and financial organizations. This article analyses the stock market prediction using deep learning model. The empirical results reveal the superiority of the LSTM recurrent deep learning model over Feed forward neural network and time series ARIMA model in terms four prediction metrics i.e. mean square error, root mean square error, mean average error and mean average percent error.

Keywords—Stock Prediction, Deep Learning, Recurrent Neural Network

I. INTRODUCTION

Researchers are of the view that prediction of stocks is the most difficult and challenging issue in finance industry. The stock market analysis gives useful information to individuals and institutions about the decorum of the market and thus helping with speculation decisions. The stock markets is highly fluctuating and it is effected by many internal and external unforeseeable factors e.g. economic instability and political situation of the country. The traders use different intelligent systems to make profits or to make decisions about when to trade or how many shares to buy or sell at a particular moment of time.

Many advanced intelligence techniques ranging from using mathematical equations to artificial intelligent models have been proposed by many financial systems for stock price prediction. The prediction will continue to be an interesting area of research making researchers in the domain always desiring to improve existing predictive models. The reason is that institutions and individuals are empowered to make investment decisions and ability to plan and develop an effective strategy about their financial endeavours. The desire of many investors is to lay hold of any forecasting method that could guarantee easy profiting and minimize investment risk from the stock market. Thus it remains a motivating factor for researchers to evolve and develop new predictive models.

In past few years there are several models developed to predict the future value of stocks. The stock forecasting problem is conventionally addressed by time series analysis techniques, i.e., auto regressive integrated moving average (ARIMA)[1, 2] as well as multiple regression models. In real situations, the dynamic of stock index time series is complex and unknown. Using a single classical model cannot produce an accurate forecast for stock price indexes. In Wang's paper

[3] a hybrid method combining linear ESM (exponential smoothing model), ARIMA and nonlinear BPNN (back propagation neural network) techniques was proposed and applied to the two real stock price datasets. The main idea of the hybrid model is to capture different forms of relationships in time series data more effectively. Thus, their study indicated that by combining different models they might develop a compelling hybrid model to generate the more accurate forecast for an incredibly complicated stock price time series.

The rest of this paper is organized as follows. Section 2 reviews machine learning and other techniques used for stock prediction. In Section 3 deep learning and types of deep learning models used in the paper are discussed. Section 4 discusses application of deep learning models on two Australian Energy Companies stock price data of the past five years and the conclusion and future work are presented in section 4 of this paper.

II. RELATED WORK

Many researchers have proposed several advanced soft computing[4] and machine learning techniques for stock price prediction[5]. According to Lai et al [4] determining the best time to buy or sell a stock remains very difficult because there are many factors that may influence the stock prices. This paper establishes a novel financial time series-forecasting model by evolving and clustering fuzzy decision tree for stocks in Taiwan Stock Exchange Corporation (TSEC). This forecasting model integrates a data clustering technique, a fuzzy decision tree (FDT), and genetic algorithms (GA) to construct a decision-making system based on historical data and technical indexes. The hit rate is chosen as a performance measure and the proposed model has shown the best performance above eighty percent of average hit rate when compared with other approaches on various stocks in TSEC.

Artificial neural networks (ANN) models are quite popular, which can learn patterns from the collected data and conclude solution form unknown data. Neural networks are highly predictive and they are found to be technologically versatile, powerful and are ideally suited for market value research [5]. However, some research have shown that ANN it is not appropriate for stock prediction because stock market data has enormous noise and complex dimensionality. Many variants of ANN in terms of hybrid models have been applied

to stock market prediction. Abraham et al. [6] make use of a neural network for one day ahead stock forecasting and a neuro-fuzzy system for analyzing the trend of the predicted stock values. To demonstrate the proposed technique they considered the popular Nasdaq-100 index of Nasdaq Stock and analyzed the 24 months stock data for Nasdaq-100 main index as well as six of the companies listed in the Nasdaq-100 index. Support Vector Machines (SVM) and Support Vector Regression (SVR)[7, 8] have also been used to predict financial time series and gain high predictive accuracy. In the financial time series modeling using SVR the key problem is the high noise. A two-stage modeling approach using independent component analysis (ICA) and support vector regression is proposed in financial time series forecasting to lessen the influence of noise. The proposed approach first uses ICA to the forecasting variables for generating the independent components (ICs). After identifying and removing the ICs containing the noise, the rest of the ICs are then used to reconstruct the forecasting variables which contain less noise and served as the input variables of the SVR forecasting model. In order to gauge the performance of the proposed approach, the Nikkei 225 opening index and TAIEX closing index are used as demonstrative examples. The experimental results show that the proposed model outperforms the SVR model with non-filtered forecasting variables and a random walk model. The researchers [8] have proposed pattern prediction in stock market using least square support vector regression to predict next fifteen days of stock prices.

Recently researchers [9-14] have been using deep learning technique for prediction. This technique employs learning from data with multiple level of abstraction by computational models that are associated with multiple processing layers. There are few deep learning approaches applied to finance [9, 11, 14, 15]. An organized analysis of the use of deep learning networks for stock market analysis and prediction is its ability to extract features from a large set of raw data without relying on prior knowledge of predictors makes deep learning potentially attractive for stock market prediction at high frequencies[14]. This study attempts to provide a comprehensive and objective assessment of both the advantages and drawbacks of deep learning algorithms for stock market analysis and prediction using high-frequency intraday stock returns as input data, we examine the effects of three unsupervised feature extraction methods i.e. Principal component analysis, auto encoder, and the restricted Boltzmann machine on the network's overall ability to predict future market behavior. The empirical results suggest that deep neural networks can extract additional information from the residuals of the autoregressive model and improve prediction performance. In another research [16] four types of deep learning architectures i.e. Multilayer Perceptron (MLP), Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) were used for predicting the stock price of a company based on the historical prices available. It used day-wise closing price of two different stock markets, National Stock Exchange (NSE) of India and New York Stock Exchange (NYSE). It has been observed that CNN outperformed the other deep learning models. The network was able to predict for NYSE even

though it was trained with NSE data. The results obtained were compared with ARIMA model and it has been observed that the neural networks are outperforming the existing linear model (ARIMA).

Some researchers e.g. Ding et al. [11] have combined the neural tensor network and the deep convolutional neural network to predict the short-term and long-term influences of events on stock price movements. In this research market simulation results show that their system is capable of making more profits than previously reported systems trained on S&P 500 stock historical data. Deep learning for stock prediction has been introduced and its performance is evaluated on Google stock price multimedia data (chart) from NASDAQ [15]. The objective of the research is to prove that deep learning can improve stock market forecasting accuracy. In this research two dimensional principal component analysis (PCA) with Deep Neural Network (DNN) method is compared with state of the art method 2-Directional 2-Dimensional Principal Component Analysis (2D) 2PCA and Radial Basis Function Neural Network (RBFNN). It is found that the proposed method is performing better than the existing method RBFNN with an improved accuracy for Hit Rate. Also the results of the proposed model are compared with the Recurrent Neural Network (RNN) and it is found that the accuracy for Hit Rate is improved by 15.6%. But in this paper time series ARIMA method and feedforward neural network and LSTM neural network is used for predicting stock closing price instead of Hit Rate.

III. TIME SERIES

Time-series forecasting is widely used to predict future values based on previously observed values. Some examples of time-series include the prediction of weather, stock market, house price sale etc. In time series the input data is a signal that is defined by observations taken sequentially in time. Basically time-series forecasting is done for non-stationary data. In non-stationary data the mean and standard deviation are not constant over time but instead these metrics vary over time.

There are a number of methods for time series forecasting e.g. Naive approach, moving average, weighted moving average, exponential smoothing and ARIMA. ARIMA model has been used extensively in the field of finance and economics as it is known to be robust, efficient and has a strong potential for short-term share market prediction [1, 2]. Autoregressive Integrated Moving Average (ARIMA) Model converts non-stationary data to stationary data before working on it. ARIMA has three main components i.e. AR (autoregressive term), I (differencing term) and MA (moving average term).

- AR term refers to the past values used for forecasting the next value. The AR term is defined by the parameter 'p' in ARIMA. The value of 'p' is determined using the PACF plot.
- MA term is used to define number of past forecast errors used to predict the future values. The parameter 'q' in ARIMA represents the MA term. ACF plot is used to identify the correct 'q' value.
- Order of differencing specifies the number of times the differencing operation is performed on series to make it stationary. Test like ADF and KPSS can be used to

determine whether the series is stationary and help in identifying the d value.

IV. DEEP LEARNING

Deep learning (DL) is quite popular in machine learning and artificial intelligence. It is a complex set of neural networks wherein learning happens on multiple levels of hidden layers [17]. They have been typically used for complex tasks, such as speech recognition, image recognition, natural language processing and hand writing identification and in bioinformatics [12][17]. Recently researchers have attempted to utilise its potential in time-series data prediction. This technique employs learning from data with multiple level of abstraction by computational models that are associated with multiple processing layers. This method intended to discover complex structure in big data set by using a learning algorithm to predict the result. The machine can learn from source and change its internal parameters by computing the representation in each layer to form the representation in the previous layer. Deep Learning architecture can be illustrated in form of different concept levels that uncover hidden layers in a problem domain. Deep learning models represent dataset in a multi-layer form. The layout consists of three basic layers like input layer, hidden layer and output layer. Input layer is the first layer which can take the input values and pass it to the next layer called hidden layer. The output to the hidden layer is provided based on calculation which is sum of the product of input value and weight. Thus each layer derives from the computation of node and weight of connections among nodes, and each transform represents one level, which will be the input for the next layer. A weight represents the strength of connection between units, which means if the weight of node 1 is greater than node 2 then node 1 has greater influence over node 2 and vice versa. Basically, the weight decides the value of input. If the weight is zero, then the input has no effect and if the weight is negative then increasing the input will decrease the output on hidden layer has neurons (nodes) which apply different conversions to the input data. All the nodes in a hidden layer are connected to each node in the next layer which is output layer where we can predict the desired number of values in a certain range. There are many variations of deep neural networks e.g. deep feed forward network and deep recurrent neural network.

A. Deep Feed Forward Neural Network

Feed forward neural network is the simplest form of neural networks. The data is processed across layers without any loops or cycles. In a feed-forward neural network, the nodes in one layer are connected only to nodes in the next layer.

If the inputs are x_1 , x_2 , and x_3 , then the synaptic weights to be applied to them are denoted as w_1 , w_2 , and w_3 .

Output is

$$y = f(x) = \sum xiwi \quad (1)$$

Where i is the number of inputs.

The network may use types of activation functions other than the sign functions. Examples of other activation functions include linear, sigmoid, and hyperbolic tangent functions. These activations allow the hidden and out-put nodes to produce output values that are nonlinear in their input parameters.

B. Deep Recurrent Neural Network

A recurrent neural network (RNN), unlike feedforward networks propagate data forward and also backward for later processing stages to earlier processing stages: The various types of recurrent neural networks are: Hopfield Networks, Boltzmann machine, Self-organization maps, bidirectional associative memory and Long Short Term Memory (LSTM).

In this paper long short term recurrent neural network (LSTM-RNN) is used for stock prediction. An architecture contained computation units in each memory block in the recurrent hidden layer. The memory block contained memory cells with self-connections storing the temporal state of the network in addition to multiplicative unit that was called 'gate', which controlled the flow of information inputted to unit. An input gate and output gate were included in the original architecture. The input gate controlled the flow of information and activations into the cell that was computed by *sigmoid* and *tanh* function. The output gate controlled the output flow of cell that activation function was computed by using sigmoid and *tanh* function for the rest of the network.

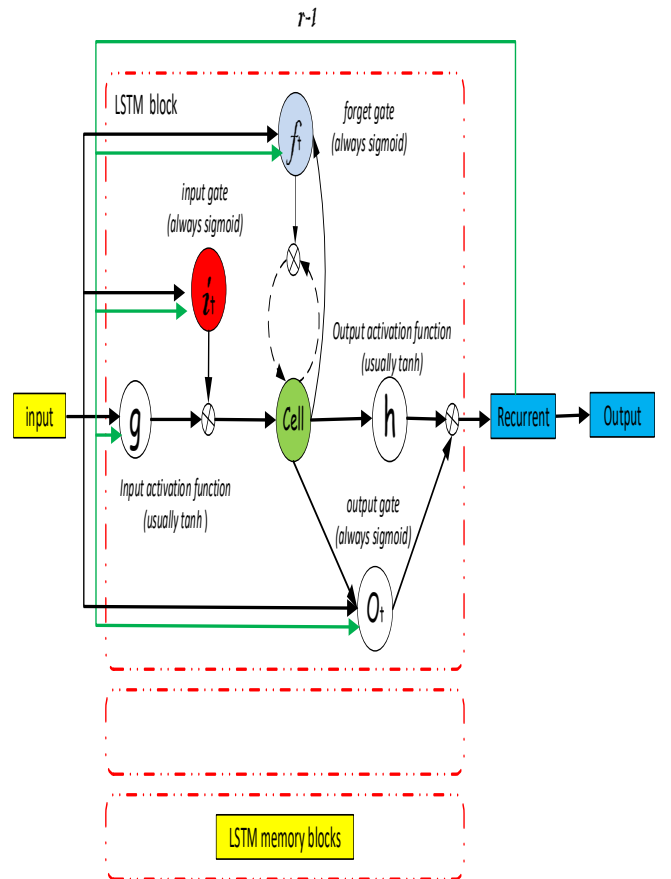


Fig. 1. LSTM-RNN memory cell architecture and memory blocks [18-20].

The forget gate was added to the memory block. This gate prevented a weakness of LSTM models from processing continuous input streams that are not segments into subsequences. The internal state of cell of the forget gate scales run verification before adding an input to the cell through the self-recurrent connection of the cell; therefore, it would forget or reset the cell's memory [7]. This gate used

sigmoid function for computation. Furthermore, in the LSTM architecture peepholes connections (green line) form its internal cells were applied to all gates in the same cell for learning precise timing of the outputs [8] (see Fig 1).

The algorithm model appears as follows:

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t) \quad (2)$$

$$a_i = \sigma(W_a x_i + W_a h_{i-1} + b_a) \quad (3)$$

$$g_i = \sigma(W_g x_i + W_g h_{i-1} + b_g) \quad (4)$$

$$\tilde{h}_i = \tanh(W_h x_i + g_i \circ W_h h_{i-1} + b_h) \quad (5)$$

$$h_i = a_i \circ h_{i-1} + (1 - a_i) \circ \tilde{h}_i \quad (6)$$

Where the W terms are weight matrices value, W_h , W_b , and W_a , are diagonal weight values for next layer to connections. The b terms are bias vectors. The logistic sigmoid function is represented by σ . The input gate, forget gate, and output gate are represented by a , g , and n respectively. All of them are in the same size as the cell output activation vectors h_i , \circ is the element product of the vector \tilde{h}_i is the cell input and cell output activation function, generally and in this research network is \tanh .

V. MODEL EVALUATION

A. Data Source

This research is applied on a historical stock price data of 5 years of the two energy companies' i.e. Origin Energy Limited and Woodside Petroleum Limited (available at Yahoo Finance website). The data has 1267 records and 7 attributes or columns per record. It contains data of Date, Opening price, Highest price, Lowest Price, Closing Price, Adjusted Closing Price and Volume of the stock on a particular day. The histogram of the closing price of the stock of Origin Energy Limited and Woodside Petroleum is shown in Fig 2 and Fig 3.

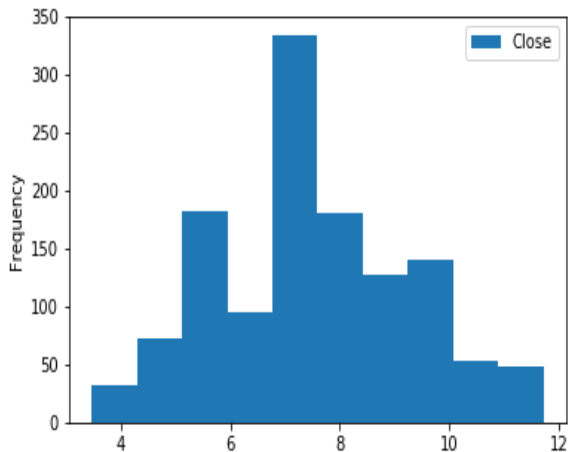


Fig.2. Histogram of Closing Price of Stock (Origin Energy Limited)

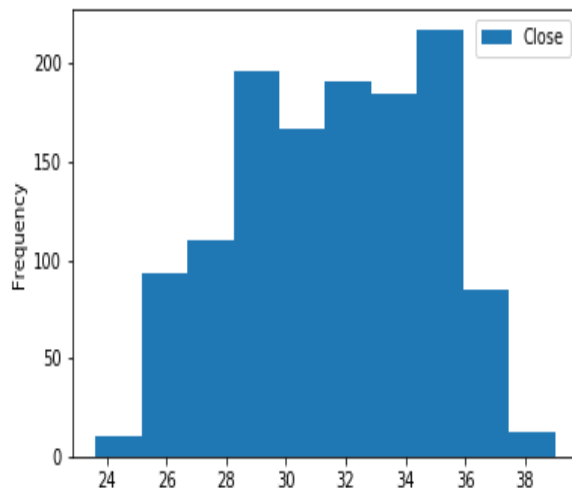


Fig.3. Histogram of Closing Price of Stock (Woodside Petroleum Limited)

B. Evaluation

This research tried to predict the closing price of the stock using time series ARIMA model and Deep learning models i.e. feed forward neural network and LSTM recurrent neural network. The results were generated using the training of 900 records and testing of 367 records of historical stock price data of two Australian energy companies i.e. Origin Energy Limited and Woodside Petroleum Limited. The plots below show predicted stock values of Origin energy and Woodside Petroleum using ARIMA Time series model (Fig 4 and Fig5) and LSTM recurrent neural network (Fig 6 and Fig7).

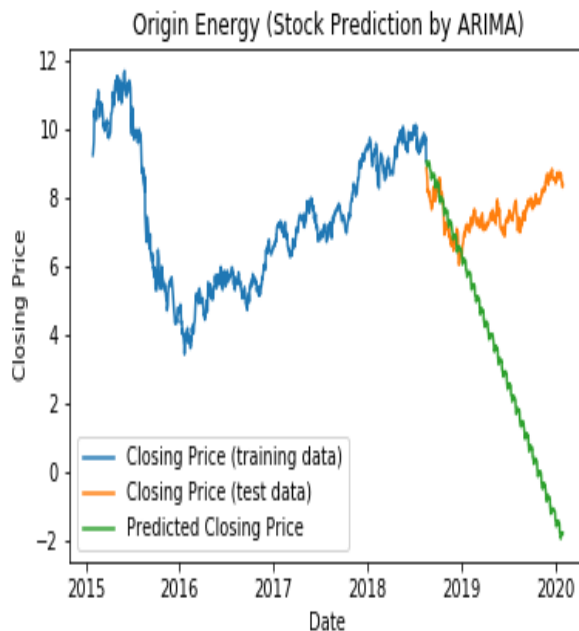


Fig.4. Plot showing Closing Price Prediction of Stock (Origin Energy Limited) using ARIMA

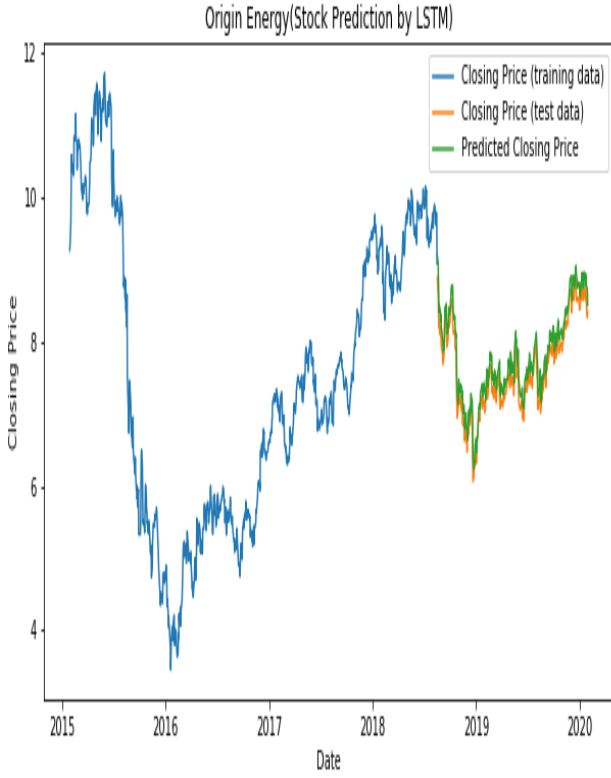


Fig.5. Plot showing Closing Price of Stock (Origin Energy Limited) by LSTM Deep Learning

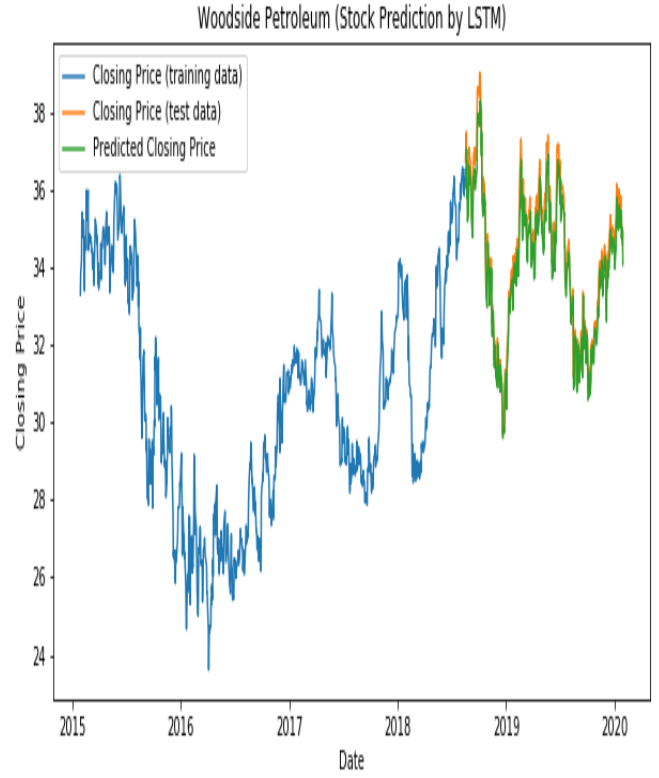


Fig.7. Plot showing Predicted Closing Price of Stock (Woodside Petroleum Limited) using LSTM Recurrent neural Network

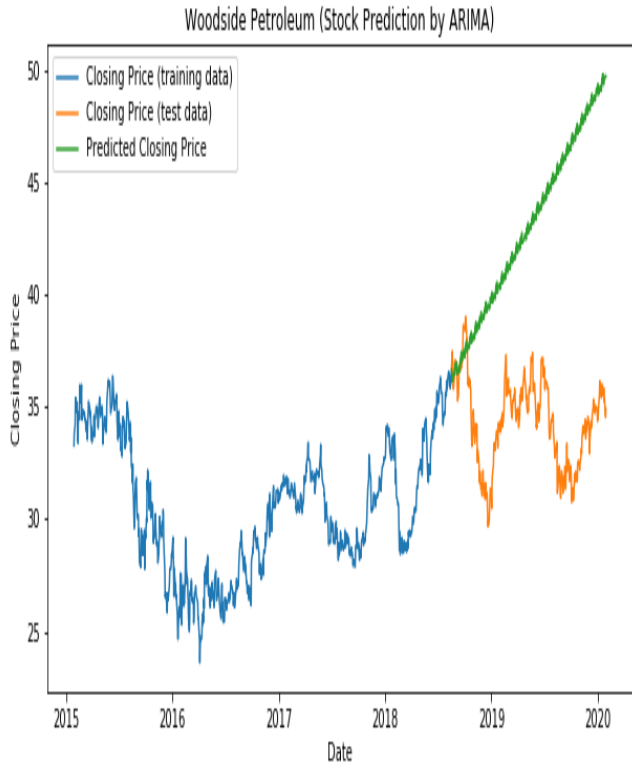


Fig.6. Plot showing Predicted Closing Price of Stock (Woodside Petroleum Limited) using ARIMA

The prediction was measured using four metrics i.e. mean square error (MSE), root mean square error (RMSE), mean average error (MAE) and mean average percent error (MAPE)(Table1). The results show that prediction by LSTM recurrent neural network has lesser all the four error metrics as compared the feedforward neural network for both the energy companies.

$$MSE = \frac{\sum_{i=1}^N (Predicted_i - Actual_i)^2}{N} \quad (7)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Predicted_i - Actual_i)^2}{N}} \quad (8)$$

$$MAE = \frac{\sum_{i=1}^N |Predicted_i - Actual_i|}{N} \quad (9)$$

$$MAPE = \frac{100}{N} \sum_{i=1}^N \left| \frac{Predicted_i - Actual_i}{Predicted_i} \right| \quad (10)$$

The error of prediction of the stock Origin energy and Woodside petroleum is least (Origin Energy: MSE=0.0173, RMSE=0.13, MAE=0.0156, MAPE=1.5) (Woodside Petroleum: MSE=0.213, RMSE=0.4585, MAE=0.6665, MAPE=6) using LSTM recurrent neural network as compared to the feedforward neural network and the time series ARIMA method.

TABLE 1: PREDICTION METRICS OF TWO ENERGY COMPANIES

NAME OF COMPANY	MODEL	MSE	RMSE	MAE	MAPE
ORIGIN ENERGY LIMITED	ARIMA	28.03	5.29	4.13	18.09
	DEEP FEED FORWARD NEURAL NETWORK	.0562	.2372	.1904	19
	LSTM RECURRENT NEURAL NETWORK	.0173	.1315	.0156	1.5
WOODSIDE PETROLEUM LIMITED	ARIMA	97.32	9.86	8.71	19.52
	DEEP FEED FORWARD NEURAL NETWORK	.3123	.5588	.4460	44
	LSTM RECURRENT NEURAL NETWORK	.2103	.4585	.6665	6

VI. CONCLUSION

Prediction of time-series data in complex, dynamic systems is among the most difficult of problems to address. No model can offer perfect predictions as there is no means to capture all the relevant data and, even if there was, computational limitations would severely limit the model's performance. In this paper, we have presented stock prediction with time series and deep learning model. The results show that deep learning model LSTM recurrent neural network offers a very promising and encouraging solution to sequence and time series related problems. All the four prediction metrics i.e. mean square error, root mean square error, mean average error and mean averaged percent error is least for the LSTM recurrent neural network as compared to the feedforward neural network and time series ARIMA method. However, the disadvantage is that more time and system resources are required for training a model using deep learning. In future, work will be done to experiment with larger datasets and with more examples to determine whether other deep RNNs can better exploit the time series data.

REFERENCES

- [1] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159-175, 2003/01/01/ 2003.
- [2] A. A. Ariyo, A. O. Adewumi, and C. K. Ayo, "Stock Price Prediction Using the ARIMA Model," in *2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation*, 2014, pp. 106-112.
- [3] J.-J. Wang, J.-Z. Wang, Z.-G. Zhang, and S.-P. Guo, "Stock index forecasting based on a hybrid model," *Omega*, vol. 40, no. 6, pp. 758-766, 2012.
- [4] R. K. Lai, C.-Y. Fan, W.-H. Huang, and P.-C. Chang, "Evolving and clustering fuzzy decision tree for financial time series data forecasting," *Expert Systems with Applications*, vol. 36, no. 2, Part 2, pp. 3761-3773, 2009/03/01/ 2009.
- [5] G. S. Atsalakis and K. P. Valavanis, "Surveying stock market forecasting techniques – Part II: Soft computing methods," *Expert Systems with Applications*, vol. 36, no. 3, Part 2, pp. 5932-5941, 2009/04/01/ 2009.
- [6] A. Abraham, B. Nath, and P. K. Mahanti, "Hybrid Intelligent Systems for Stock Market Analysis," Berlin, Heidelberg, 2001, pp. 337-345: Springer Berlin Heidelberg.
- [7] C.-J. Lu, T.-S. Lee, and C.-C. Chiu, "Financial time series forecasting using independent component analysis and support vector regression," *Decision Support Systems*, vol. 47, no. 2, pp. 115-125, 2009/05/01/ 2009.
- [8] S. Kaushik and N. Singhal, "Pattern Prediction in Stock Market," presented at the Proceedings of the 22nd Australasian Joint Conference on Advances in Artificial Intelligence, Melbourne, Australia, 2009.
- [9] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A Fast Learning Algorithm for Deep Belief Nets," *Neural Computation*, vol. 18, no. 7, pp. 1527-1554, 2006.
- [10] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to Forget: Continual Prediction with LSTM," *Neural Computation*, vol. 12, no. 10, pp. 2451-2471, 2000.
- [11] X. Ding, Y. Zhang, T. Liu, and J. Duan, "Deep learning for event-driven stock prediction," in *Twenty-fourth international joint conference on artificial intelligence*, 2015.
- [12] P. Chantamit-o-pas and M. Goyal, "Prediction of Stroke Using Deep Learning Model," Cham, 2017, pp. 774-781: Springer International Publishing.
- [13] P. Chantamit-o-pas and M. Goyal, "Long Short-Term Memory Recurrent Neural Network for Stroke Prediction," Cham, 2018, pp. 312-323: Springer International Publishing.
- [14] E. Chong, C. Han, and F. C. Park, "Deep learning networks for stock market analysis and prediction: Methodology, data representations, and case studies," *Expert Systems with Applications*, vol. 83, pp. 187-205, 2017/10/15/ 2017.
- [15] R. Singh and S. Srivastava, "Stock prediction using deep learning," *Multimedia Tools and Applications*, journal article vol. 76, no. 18, pp. 18569-18584, September 01 2017.
- [16] H. M. G. E.A, V. K. Menon, and S. K.P, "NSE Stock Market Prediction Using Deep-Learning Models," *Procedia Computer Science*, vol. 132, pp. 1351-1362, 2018/01/01/ 2018.
- [17] C. C. Aggarwal, "Neural Networks and Deep Learning," *Springer International Publishing*, 2018.
- [18] F. A. Gers, N. N. Schraudolph, and J. Schmidhuber, "Learning precise timing with LSTM recurrent networks," *Journal of machine learning research*, vol. 3, no. Aug, pp. 115-143, 2002.
- [19] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory recurrent neural network architectures for large scale acoustic modeling," in *Fifteenth Annual Conference of the International Speech Communication Association*, Singapore, 2014, pp. 338-342, Singapore, 2014.
- [20] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE transactions on neural networks and learning systems*, 2017.