Ilkogretim Online - Elementary Education Online, Year; Vol 20 (Issue 5): pp. 1671-1677
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.05.182

A Survey on LSTM-based Stock Market Prediction

Sachin Tiwari*, Department of Computer Science& Engineering, Lakshmi Narain College of Technology, Bhopal, India, sachinmcavds97@gmail.com
Anoop Kumar Chaturvedi, Department of Computer Science& Engineering, Lakshmi Narain College of Technology, Bhopal, India

Abstract- The stock price incorporates variables rate of economic growth, inflation rate, overall economy, trade balance, and monetary system that affect the whole stock market. For investors, the principle of the stock price trend has often been unclear due to numerous significant variables. In developing an investment plan or deciding duration for the purchasing or selling of a stock, the prediction of stock markets provides a crucial function. The stock index's non-linear and dynamic nature estimates the stock market avalue is challenging. Deep learning strategies have emerged as a critical technique in the analysis of dynamic temporal data relations. Several studies of deep learning techniques have been effective in making such a prediction. The Long Short Term Memory (LSTM) has gained popularity for estimating stock market prices. LSTM is a particular form of recurrent neural network (RNN) which implements a gradient descent technique. This paper extensively investigates approaches used for stock market forecasts using LSTM, explains them, and conducts a comparative analysis. The stock market's principal application comprises stock price forecasting, index modeling, risk assessment, and return estimates. We include future directions and summarize the importance of applying LSTM for stock market prediction based on our surveyed papers.

Keyword: Long short-term memory (LSTM), recurrent neural network (RNN), nifty 50, root mean square error (RMSE), prediction, stock prices.

I. INTRODUCTION

For several decades, projections on stock prices have been the area of research [1]. Finance analysts have always experienced significant difficulty in forecasting stocks. Many factors impact the value of stocks, including national politics, market indicators, and trading aspirations. Although stock trading heavily depends on numerous advanced technologies to make buying and selling choices, its effectiveness has so far been restricted. Accurate forecasts are difficult to obtain since financial sector patterns are complex, dynamic, and non-linear [2, 3]. There are typically two primary approaches to forecast the stock value. First one is a fundamental study focused on the strategic and basic knowledge of a business, such as business revenue, investments, and estimated growth figures. The second is the strategy for technical assessment based on past asset performance. There are several pieces of research from different fields of study that aim to tackle this issue.

One popular approach is to use machine learning techniques to learn from past market knowledge to forecast market trends. In the past, financial analysts have usually forecast asset prices. However, data scientists began to address forecasting issues with scientific advancement. Machine learning approaches and, more recently, deep learning have developed new methods for time series forecasting, with deep and layered hierarchy analysis relations between variables. Researchers incorporate machine learning techniques [4-6] to build projection models and improve prediction performance. The next step was to use deep learning to accelerate prediction models' efficiency [7, 8]. Deep learning approaches can detect structure and trends in time-scale predictions such as non-linearity and complexity.The Long Short-Term Memory (LSTM) deep-learning strategies have received tremendous recognition in many fields, including economic analysis, in recent years. LSTM is a form of recurrent network that can differentiate the recent and early instances by providing different weights for each of them while ignoring the memory it identifies to be insignificant.

In contrast to other recurrent neural networks, which can only memorize small sequences, it is suitable for processing long sequences.This survey focuses on the estimation strategies of the stock market based on LSTM. Our survey provides a guide to researchers to the works across the broader range of future LSTM-based work on the stock market. It can be viewed as a detailed overview of how LSTM can be incorporated into a stock-related forecast procedure in different estimations stages.

The remaining article's organization can be summarized as follows: Section 2 explains the deep learning processing approach. Section 3 describes the stock market and elaborates its potential applications; Section 4 presents prediction methods in stock market based on LSTM; Section 5 provides the concluding remarks on the presented survey and opportunities in the field of deep learning stock market prediction.

II. BACKGROUND

In this section, We explain deep learning and an overview of the appearance of a neural network.

2.1 Artificial Neural Network (ANN)

ANN [9] is a fundamentally huge, parallel computer models, which mimic human brain activity. A significant number of primary processors connected by weighted links are used in the ANN. The treatment nodes may be called "neurons" by comparison. Each node's performance depends on the available information at the node, either internally stored or accessible via the weighted connections. A different number of neurons can be made up of artificial neural networks (ANNs). At least three layers are used in a neural network - 1) an input layer, 2) hidden layers, and 3) a layer in outputs. Dimensionality or number of nodes in the input layer may be deciding the number of data set characteristics. These nodes are connected to nodes in the hidden layer by the connections called "synapses" (s). For any node in the input layer, the synapses bear some weight. Basically, a neural network learns to change the weight with any synopsis. In hidden layers, the nodes add the weighted sum of inputs to trigger the outputs' values or inputs. The output layer produces a probability vector for the different outputs and chooses one with the lowest error or cost, which minimizes the variations between expected and expected values.

2.2 Recurrent neural networks (RNNs)

A new neural network (RNN) [10] is a particular case of the neural network to predict the next step in the observational process with regard to the previous steps found in the sequence. A significant variant of neural networks is known as RNN, commonly used in numerous issues. Recurrent neural networks (RNNs) are a subset of neural networks that can usually process data from time series and sequential information. Recurring neural networks as a feeding network extension. RNN is called a repeating since, if the output is related to the previously measured values, it does the same job for each variable in the sequence. RNN has a special memory that preserves information that has been computed for a long time. In RNNs, the protected layers serve as internal storage to store the data obtained in previous sequential reading stages. The RNNs receive the data from the hidden layer from the immediate input layer (Input) and an earlier output layer. RNNs are called "recurrent" because they execute the same role with the function of using previously collected data to predict future unknown sequential data for each part of the chain. For a standard generic RNN, the main difficulty consists of recalling only a few earlier steps in the process and not being able to recall longer data sequences. The "memory line" implemented in the recursive Long Short Term Memory (LSTM) network solves this difficult problem. A benefit of the RNNs is that they contribute to long connections in the series between the final output layer and hidden layers of previous data, which creates a problem with the disappearance gradients. Modifications to the RNN units to provide gates and memory units to recall the gauges were made to solve this problem.

2.3 Long Short-Term Memory (LSTM)

LSTM is a specialized category of RNN with a wide variety of applications, including text labeling, time series analysis, voice recognition, and speech recognition. Hochreiter and Schmidhuber [11] suggested the Long-term memory (LSTM) technique. LSTM has the capacity to recall values for future use from the earlier stages of a form of Recurrent Neural Network (RNN).LSTM units are designed to boost RNNs' ability to record previous values and assign priority in older data sequence to learning from newer ones. The earlier data trend can be memorized through some doors along with a memory line in the standard LSTM.

The memory of LSTM units also addresses the issue of disappearing gradients, as the memory allows a gradient to pass from one hidden layer to the next without being decreased. Each LSTM is a series of cells or device modules that collect and store data streams. The cells mimic the transport line in each cell (the upper line), which connects data from the past and gathers it for the current module from one module to the other. LSTMs have been shown to be more potent than typical RNNs. Four neural network layers communicating with a particular approach are made of the LSTM layer. An ordinary LSTM device has three separate components, one cell, one escape door, and a forgotten door. The primary role of cells is to classify the values over random times and to monitor the data flow into and out of the cells. Each LSTM requires three types of doors in order to control each cell's state: Forget Gate produces a number between 0 and 1, where one demonstrates "to keep that completely," while 0 means "to ignore it completely." Memory Gate selects the latest data that must be processed in the cell. Next, a sigmoid layer, called " input layer," selects which values are altered. Then, a tanh layer generates a new candidate value vector that can be added to the state. - The Gate Output determines what output from each cell will be. The returned value and the filtered and newly inserted data are based on the cells.

Both feed-forward and back-time dependencies are taught via bi-Directional Neural Networks [12]. Each unit is divided into two units with the same entrance and connected to the same output in a bi-directional LSTM. One unit is used for the sequence of time and the other for the sequence of time. The typical output performs an arithmetic process such as number, average, and concatenation of the two units. Bi-directional LSTM thus makes it useful to benefit from time-series data that span longer.

III. LSTM BASED PREDICTION METHODS

This segment discusses the studies performed by researchers to incorporate and adapt a deep learning methodology to forecast financial stocks. LSTM has demonstrated that it is modeling, visualization, generalization, and self-organizing functionality allow the volatility and non-linear stock market rates to be forecast. The prediction of the stock market behavior is centered on the history of predicted events. The share price forecast of LSTM consists of two levels of deep network preparation modeling and prediction. In the training process, the network creates a collection of interconnecting weights, produces a positive performance outcome, and then matches it with the anticipated value. This method will proceed until the error between acceptable performance and predicted value fulfills the demands such that weights can be reached satisfactorily. The prediction method of the network is to access test data to determine stable weights and thresholds. The key benefit of deep learning is that the non-linear feature can be approximated by a reasonable degree of precision.

Chen et al.[13] have proposed an LSTM forecasting model for the Chinese stock market. The LSTM model proved better than another random prediction process It uses historic market data to forecast whether an equity price will increase, decrease or remain constant one day in addition to stock indices. They have demonstrated that the LSTM is preferable for the financial time series model to the feed-forward neural networks model.

An Artificial Neural Network (ANN) method is proposed by Persio et al.[14] to forecast stock market indices based on the integration of wavelets and Convolutional Neural Networks. Experiments leverage Neural Networks' numerous technologies, including the recurrent neural network, Multi-layer Perceptron (MLP), CNN, and Long Short-Term Memory (LSTM). The efficiency of these approaches was compared to the proposed system, which outperforms all but has very similar LSTM network outcomes. Tests were performed on the historical time series data of the S&P500 and FOREX EUR/USD. The findings show a maximum accuracy of 62 % for S&P500 and 83 % for Forex.

Nelson et al.[15] used LSTMs to incorporate share market and technical research metrics to forecast potential patterns in stock markets. Experimental findings have shown that their proposed LSTM method is more precise than other machine learning models. As a source of information, historical market details (candlesticks) of various stocks from the Brazilian stock exchange (Bovespa) were used. The research findings are promising, with an overall precision of up to 55.9 %.

Bao et al. [16] used the three-stage method of auto-encoding to forecast six potential stock indices. Wavelet transformation was used to minimize stock data of high to low dimensionality. The data were reprocessed with the autoencoder. In order to forecast asset prices, they eventually used an LSTM. In contrast to other models, such as RNN, LSTM, and Wavelet-LSTM, the proposed model's efficiency was higher. The findings reveal that in both statistical precision and viability, the model proposed outperforms other comparable models. Six inventory indices were used to test the forecast capability of the proposed model. These include CSI 300 of China, Tokyo Nikkei index, Nifty 50 on the Indian stock market, Hang Seng Hong Kong index trading, the NSE index of the S & P 500, and DJIA. The OHLC variables, functional indicators, and macroeconomic variables are used in our input variables.

In order to forecast the trends in the Nifty Index, Roondiwala et al.[17] have introduced the LSTM RNN model, which uses features like low, high, close, and open rates. Five years of NIFTY 50 historical records were used for the experiment. This result was a root average RMSE error of 0.0086 after 500 epochs of preparation. However, the document did not explain how over-fitting was prevented. The authors in this study viewed RMSE as a percentage rise in the inventory price every day. The value is almost zero, but RMSE's cumulative value is nearly 1%, marginally more than 100, of the Nifty index.

Fischer et al. [18] employ diverse ways of projection for S&P 500 consisting of deep learning, random forest, and gradients-boosted trees. Surprisingly, deep learning modeling outperformed those methods. This LSTM network outperforms the memory-free procedures with substantial statistical and economic returns of 0.46% per day compared to 0.43% for RAF, 0,32% for normal DNN, and 0.26% for logistic regression. A simpler rules-based trading technique is also built in line with the popular LSTM Portfolio trends.

Lee et al. [19] have applied an RNN approach to forecast share returns. The idea had been to develop portfolios by changing the threatening return levels of the designed RNN layers. In this paper, an ARIMA framework's efficiency compares with that of the LSTM model in financial time series data to evaluate the best attributes of the variables involved in a traditional estimation method.

ModAugNet has been suggested by Baek et al.[20]: an approach to data increase to render stock index over-fitting. The analysis is focused on the current stock market index forecasting platform and consists of two modules; one LSTM avoidance overfit module and the other LSTM prediction module. Moreover, an over-fitting LSTM prevention module is called a prevention module, and an LSTM prediction module is called a prediction module. Empirical experiments were carried out on two representative real-world stock market indexes, namely S&P500 and Korea Composite Stock Price Index 200, to analyze whether our system performs well on various indices (KOSPI200). ModAugNet efficiency is contrasted to the SingleNet model where the Prevention Module lacks in order to test the efficacy of the Prevention Module in our proposed network. The findings show that the current model is outstanding for the prediction. We find that the test success depends on the LSTM prediction module entirely by analyzing the qualified ModAugNet-c.

Chung and Shin [21] suggest a long term network and genetic algorithms (LSTM) hybrid solution (GA). The study's LSTM network comprises two hidden layers that are a profound framework to communicate not-linear and dynamic financial sector functionality more efficiently. GA was used to search in the LSTM network for the optimum or near-optimal time window duration and number of LSTM units. The experimental findings indicate that the suggested solution is lower than MSE, MAE, and MAP, and statistically relevant changes are observed. These overall findings suggest that a GA–LSTM approach may be an efficient way of forecasting stocks to reflect temporal trends. The suggested approach did not take the commercial commission into account in the study and projected only stock exchange valuation and costs. However, trading commissions should be considered for better yields in real-world investing conditions. We chose daily Korea Stock Price Index (KOSPI) data to test the proposed hybrid strategy. The experimental findings reveal that the LSTM and GA hybrid model beats the benchmark model.

The new LSTM model, which integrates the LSTM model with various generalized conditional heteroskedasticity (GARCH) models, is proposed by Kim et al. [22]. The of the work goal is to estimate stock market volatility. To test, the index data from KOSPI 200 was used to find proposed hybrid models integrating the LSTM with one to three GARCH models. Single models, such as GARCH, exponential GARCH, exponentially weighted moveable average, a deep neural feed-forward network(DFN), and LSTM are analyzed in comparative studies with their current methodologies, and the hybrid DFN model integrating a DFN with one of the GARCH models are used. Its efficiency is contrasted with the hybrid LSTM models proposed. The findings show that GEW-LSTM has the lowest prediction errors in terms of mean absolute fallacy (MAE), mean squared error (MSE), modified MAE heteroskedasticity, and adjusted heteroskedasticity (MSE) in a hybrid model combined with three sample forms in the GARCH method (HMSE). The GEW-LSTM MAE is 0.0107 and 37.2 percent smaller than the E-DFN model (0.017), the EGARCH-DFN model that blends the best current model with the best model. Moreover, GEW-LSTM is 57.3% smaller, 24.7%, and MSE, HMAE, and HMSE are 48% smaller.

Kim et al.[23] suggests a model that draws together features acquired from various members of the same data (in particular, the inventory time series and inventory map imagery), which incorporates the long-term functionality memory-convolutions neural network model LSTM-CNN. The proposed model consists of LSTM and CNN, used to derive time features and picture characteristics. The proposed model results are calculated with SPDR S&P 500 ETF data compared to single models (CNN and LSTM). Our LSTM-CNN feature fusion model forecasts market markets by outperforming single ones. Furthermore, a candlestick map is the most fitting picture for stock price forecasts. Using time and image features from the same data, instead of using those features separately, this study shows that prediction errors can efficiently be minimized.

Long et al.[24] suggested a profound neural network model that forecasts stock price patterns, using desensitized trade records and public market details. Considering how markets are linked, our approach uses graphical tools to pick appropriate markets and trade details stocks. We used the long-term, short-term memory network to learn the timing details and used BiLSTM and the particular business days' attention mechanism to help the model concentrate more on a promising performance in stock pattern prediction. The model has been tested on the actual details of three Chinese A-shares (CITIC securities, GF Securities, China Pingan) for almost seven years. Deep Market Pattern Neural Network (DSPNN) suggested in this paper. The project's final results showed that the project showed thaSecuritiesway LSTM could predict investors' future and that the technology achieved the best efficiency.

Pang et al. [25] suggested a method based on word vector creation and proved the 'stock vector' principle. The proposed method comprises the long, embedded layer, long-term, neural memory network (LSTM), and the automatic stock encoder, long-term memory neural network. The test results demonstrate that the deep LSTM of the built-in layer is improved. For the Shanghai A-shares composite index, the precision of two versions is 57.2 and 56.9 percent. The Internet of Multimedia of Things (IMMT) for inventory analysis is shown in the visual and analytical findings.

Kelotra et al.[26] developed a method called the deep-convolutionary long-term memory model for Rider-monarch optimization (MBO) for predicting the stock status. The Deep-ConvLSTM model serves as a prediction module that is learned by using the proposed Monarch Butterfly Optimization algorithm (Rider-MBO). The suggested Rider-MBO algorithm consists of the convergence of the rider optimization (ROA) and MBO algorithms. Initially, the stock market data are determined by using the Sparse-Fuzzy C-Means (Sparse-FCM), following with function selection, to represent the features from which the features needed are obtained by clustering. It uses six types of stock market data based on the estimation technique, such as medium squared error (MSE) and root mean squared error (RMSE). The Rider model was used to optimize RMSE and MSE by Rider's Monarch Butterfly process of 26923 and 7 2487.

Yıldırım, Al. [27] provides a hybrid model consisting of a macroeconomic model LSTM and a scientific model LSTM named according to the data forms it utilizes. The results of these data on lateral movement were studied separately. Then we merged the findings to increase prediction accuracy dramatically. The macroeconomic LSTM model uses several financial variables, including interest rates, the Federal Reserve (FED) assets, infancy rates, the Standard and Poor's (S&P) 500, and the German price index (DAX). The pattern in the EUR/USD currency pair is influenced by each element. Two baseline models with macroeconomic or technological predictor data only were introduced. We observed that in terms of both profit accuracies and the number of produced transactions, ME LSTM had marginally improved performance compared to TI LSTM. The gap, however, was rather slight and insignificant. In comparison, it did not improve the accuracy substantially by integrating all features with a single LSTM called ME-TI_ LSTM. In terms of profit accuracy for forecasts over all periods, our proposed hybrid model had the highest possible results (73% on average) (40.37 percent on average).

IV. CONCLUSION

This paper reviews the use of LSTM in stock market forecasting. LSTM plays a significant role in stock market forecasting; it can derive valuable patterns from broad temporal records. The LSTM approach to stock market behavior is a relatively recent, active, and promising area. Although there have been reasonable findings for

LSTM, many researchers are still attempting, by using a hybrid approach and considering more external variables, to improve the precision of the stock market forecast to achieve a more precise prediction. We learned that the stock market has so many external forces as non-linear, unpredictable, and subject to pressures. This article can be used as instructional material for those involved in forecasting stocks using LSTM.

Reference

[1] Murphy, J. J. (1999). *Technical analysis of the financial markets: A comprehensive guide to trading methods and applications*. Penguin.

[2] T. Turner, A Beginner's Guide To Day Trading Online, 2nd ed. New York, NY, USA: Simon and Schuster, 2007.

[3] Duarte, J. B. D., Sarmiento, L. H. T., & Juárez, K. J. S. (2017). Evaluation of the effect of investor psychology on an artificial stock market through its degree of efficiency. *Contaduría y Administración*, *62*(4), 1361-1376.

[4] Hassan, M. R., Nath, B., & Kirley, M. (2007). A fusion model of HMM, ANN and GA for stock market forecasting. *Expert systems with Applications*, *33*(1), 171-180.

[5] W. Huang, Y. Nakamori, and S.-Y. Wang, ''Forecasting stock market movement direction with support vector machine,'' Comput. Oper. Res., vol. 32, no. 10, pp. 2513–2522, Oct. 2005.

[6] Sun, J., & Li, H. (2012). Financial distress prediction using support vector machines: Ensemble vs. individual. *Applied Soft Computing*, *12*(8), 2254-2265.

[7] Maqsood, H., Mehmood, I., Maqsood, M., Yasir, M., Afzal, S., Aadil, F., ... & Muhammad, K. (2020). A local and global event sentiment based efficient stock exchange forecasting using deep learning. *International Journal of Information Management*, *50*, 432-451.

[8] Long, W., Lu, Z., & Cui, L. (2019). Deep learning-based feature engineering for stock price movement prediction. *Knowledge-Based Systems*, *164*, 163-173.

[9] Dongare, A. D., Kharde, R. R., & Kachare, A. D. (2012). Introduction to artificial neural network. *International Journal of Engineering and Innovative Technology (IJEIT)*, *2*(1), 189-194.

[10] DiPietro, R., & Hager, G. D. (2020). Deep learning: RNNs and LSTM. In *Handbook of medical image computing and computer assisted intervention* (pp. 503-519). Academic Press.

[11] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.

[12] Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, *45*(11), 2673-2681.

[13] Chen, K., Zhou, Y., & Dai, F. (2015, October). A LSTM-based method for stock returns prediction: A case study of China stock market. In *2015 IEEE international conference on big data (big data)* (pp. 2823-2824). IEEE.

[14] Di Persio, L., & Honchar, O. (2016). Artificial neural networks approach to the forecast of stock market price movements. *International Journal of Economics and Management Systems*, *1*.

[15] Nelson, D. M., Pereira, A. C., & de Oliveira, R. A. (2017, May). Stock market's price movement prediction with LSTM neural networks. In *2017 International joint conference on neural networks (IJCNN)* (pp. 1419-1426). IEEE.

[16] Bao, W., Yue, J., & Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PloS one*, *12*(7), e0180944.

[17] Roondiwala, M., Patel, H., & Varma, S. (2017). Predicting stock prices using LSTM. *International Journal of Science and Research (IJSR)*, *6*(4), 1754-1756.

[18] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, *270*(2), 654-669.

[19] Lee, S. I., & Yoo, S. J. (2017). A deep efficient frontier method for optimal investments. *Department of Computer Engineering, Sejong University, Seoul*, *5006*.

[20] Baek, Y., & Kim, H. Y. (2018). ModAugNet: A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. *Expert Systems with Applications*, *113*, 457-480.

[21] Chung, H., & Shin, K. S. (2018). Genetic algorithm-optimized long short-term memory network for stock market prediction. *Sustainability*, *10*(10), 3765.

[22] Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications*, *103*, 25-37.

[23] Kim, T., & Kim, H. Y. (2019). Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. *PloS one*, *14*(2), e0212320.

[24] Long, J., Chen, Z., He, W., Wu, T., & Ren, J. (2020). An integrated framework of deep learning and knowledge graph for prediction of stock price trend: An application in Chinese stock exchange market. *Applied Soft Computing*, *91*, 106205.

[25] Pang, X., Zhou, Y., Wang, P., Lin, W., & Chang, V. (2020). An innovative neural network approach for stock market prediction. *The Journal of Supercomputing*, *76*(3), 2098-2118.

[26] Kelotra, A., & Pandey, P. (2020). Stock market prediction using optimized deep-convlstm model. *Big Data*, *8*(1), 5-24.

[27] Yıldırım, D. C., Toroslu, I. H., & Fiore, U. (2021). Forecasting directional movement of Forex data using LSTM with technical and macroeconomic indicators. *Financial Innovation*, *7*(1), 1-36.