An Empirical Research on the Effectiveness of Different LSTM Architectures on Vietnamese Stock Market

Pham Ngoc Hai*
ICT Department, FPT University, Hanoi, Vietnam
haipnhe141371@fpt.edu.vn

Nguyen Manh Tien
ICT Department, FPT University, Hanoi, Vietnam
tiennmhe130912@fpt.edu.vn

Hoang Trung Hieu
ICT Department, FPT University, Hanoi, Vietnam
hieuhthe141638@fpt.edu.vn

Pham Quoc Chung
ICT Department, FPT University, Hanoi, Vietnam
chungpqhe141371@fpt.edu.vn

Nguyen Thanh Son
ICT Department, FPT University, Hanoi, Vietnam
sonnthe140556@fpt.edu.vn

Pham Ngoc Ha
ICT Department, FPT University, Hanoi, Vietnam
hapn10@fe.edu.vn

Ngo Tung Son
ICT Department, FPT University, Hanoi, Vietnam
sonnt69@fe.edu.vn

ABSTRACT

Stock price prediction is a challenging financial time-series forecasting problem. In recent years, on account of the rapid progression of deep learning, researchers have developed highly accurate, state-of-the-art time-series models. Long short-term memory (LSTM) stands out as one of the most reliable architecture at capturing long-time temporal dependences. In Vietnam, there is a lack of research papers that solely focused on the effectiveness of deep-learning in stock price prediction. This paper surveys three different variations of LSTM (Vanilla, Stacked, Bidirectional) when applied to 20 companies' stock prices over a period of 5 years from 2015 to 2020 in the VN-index stock exchange. The results show that Bidirectional LSTM is the most accurate model.

CCS CONCEPTS

• **Information systems->Information retrieval->Retrieval tasks and goals->Information extraction**;

KEYWORDS

LSTM, Stock Price Prediction, Stacked LSTM, Vanilla LSTM, Bidirectional LSTM

ACM Reference Format:
Pham Ngoc Hai, Nguyen Manh Tien, Hoang Trung Hieu, Pham Quoc Chung, Nguyen Thanh Son, Pham Ngoc Ha, and Ngo Tung Son. 2020. An Empirical Research on the Effectiveness of Different LSTM Architectures on Vietnamese Stock Market. In *2020 International Conference on Control, Robotics and Intelligent System (CCRIS 2020), October 27–29, 2020, Xiamen, China.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3437802.3437827

*Corresponding Author.

CCRIS 2020, October 27–29, 2020, Xiamen, China

ACM ISBN 978-1-4503-8805-4/20/10…$15.00
https://doi.org/10.1145/3437802.3437827

1 INTRODUCTION

1.1 Background

At the beginning, conventional statistical models had trouble forecasting them with high accuracy. This is attributed to financial time series' nonstationary, nonlinear and high-noise traits [1]. With the swift development of artificial intelligence, researchers have moved on from traditional methods and started applying deep learning in stock market prediction. In 2010, Nikfarjam et al. [2] surveyed studies which used text mining to extract companies' unquantifiable information then use it to predict future trends of stock values. In 2013, Lin et al. [3] published a stock prediction method using support vector machine to establish a two-part feature selection and prediction model and proved that this method is more generalized than the traditional ones. Wanjawa et al. [4] proposed an artificial network model that is feedforward multi-layer perceptron with error backpropagation to predict stock values in 2014. Later, Zhang et al. [5] proposed a new deep and wide area network (DWNN) that utilize a combination of convolutional neural network (CNN) and recurrent neural network (RNN). Their research proved that the latest DWNN model can reduce the margin of error by 30% in comparison with the general RNN model.

During the process, LSTM model emerges as one of the most accurate architecture to predict stock price, as they do not suffer from the optimization barriers that plague simple recurrent networks (SRNs) [6]. In 2017, Zhao et al. [7] proposed a LSTM model that implemented a time-weighted function and its results surpassed the other models. Jiang et al. [8] proved that LSTM could be better applied to stock forecasting with a combination of LSTM neural network and RNN.

1.2 Contribution

This study empirically analyzes the accuracy of different LSTM models applied in stock price forecasting. We collected a stock's open, close, high, low price and volume over a time period of 5 years from 2015 to 2020 in the VN-index stock exchange and used it to train and validate the models. Finally, we will rank the effectiveness of the LSTM variants in predict stock price based on the root mean square error of their prediction.

This is one of the earliest studies that apply LSTM models on stock prediction in Vietnam, as there were less than a handful of similar researches done prior to this [9][10][11][12]. Additionally, Ngo et al. [13] conducted empirical research on the predictability and profitability of the candlestick reversal patterns analysis on the Vietnamese stock market and found that it was not effective. Our research provides highly accurate, verified LSTM models capable of predicting a stock price in the third day of the future after accepting data from the past 60 days. The models will be useful for investors looking to maximize profit from the stock market. Moreover, a financial analyst may also implement the model to develop their own trading strategies. The result of this paper is also valuable for deep-learning researchers who are looking to explore the effectiveness of deep-learning models in an economical context.

2 RELATED RESEARCHES

2.1 LSTM

Long Short-Term Memory (LSTM) is a specific recurrent neural network (RNN) architecture that was designed to model temporal sequences and their long-range dependencies more accurately than conventional RNNs [14]. LSTM was first brought up by Sepp Hochreiter and Jurgen Schmidhuber in 1997 [15], it implements a unique set of memory cells, which is built around a central linear unit with a fixed self-connection, instead of the usually hidden layer neurons found in the general RNN. There are three sigmoid layers and one tanh layer in a LSTM memory cell. The LSTM uses its gates architecture (input, forget, and output gates) to filter unnecessary information, maintain and update the memory cells. Firstly, the forget gate decides what information should be discarded or kept from the model. Secondly, the input gate handles the reservation of information on the current cycle input. The input gate is in charge of two tasks. The first task is finding the state of the cell that must be updated; the value to be updated is selected by the sigmoid function. The second task is updating the information to be updated to the cell state. The final gate is the output gate. The output gate controls what the next hidden state should be.

2.2 Vanilla LSTM

A Vanilla LSTM is an LSTM Model that has a single hidden of layer LSTM memory cells and makes predictions with an output layer. [16]. Greff, Srivastava, and Koutn [17] concluded that vanilla LSTM remains a reliable architecture that performs reasonably well on various datasets compare to its extensions even after more than 20 years since its introduction. Nelson et al. [18] is the first researcher to apply Vanilla LSTM to stock price prediction and obtained promising results. This research proved Vanilla LSTM's distinguished ability to capture long-term dependencies.

2.3 Stacked LSTM

Multiple hidden LSTM layers can be stacked one on top of another in what is referred to as a Stacked LSTM model [16]. Stacked LSTM

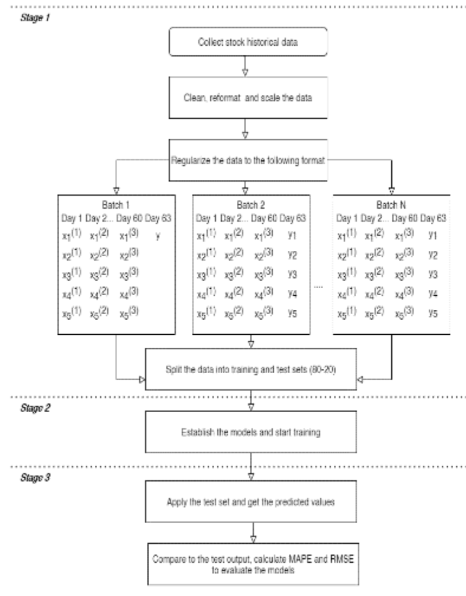

Figure 1: Experimental design flowchart

is an established stock price prediction model. Ojo, Owolawi, and Adisa [19] showed that stacked LSTM is able to predict stock price with certain accuracy. However, according to Zou and Qu [20], interestingly, the stacked-LSTM does not significantly outperform the LSTM in the context of stock price prediction. Instead, the Vanilla LSTM outperformed the stacked-LSTM occasionally. This proves that the more complex representative does not necessarily improve the predictive power.

3 EXPERIMENTAL DESIGN

There are three stages in establishing a stock price forecasting model: data collection and preprocessing, model establishment and training, and evaluation of experimental results as in Figure 1

3.1 Data collection

We selected the VN-index as experimental data. The dataset is collected over a 5 years period from 6-12-2015 to 6-12-2020 from the website vndirect.com.vn. There are 20 different companies' stock prices. Each has five different features. The opening price is the price from the first transaction on a trading day. Closing price is the price from the final transaction on a business day. High and low are the highest and lowest price on that day respectively. Volume refers to the number of transactions during that trading day. Figure 2 refers to the candlestick chart of FPT stock from 4/3/2020 to 6/12/2020.

3.2 Data cleaning and preprocessing

We implemented our models in Python using TensorFlow and Keras. The data was collected from a Vietnam website therefore, in order to apply it in Python, we had to reformat the data into their English equivalents. We then split the data into training and test set in 80-20 ratio. We wanted to use the data from the past 60 days to predict

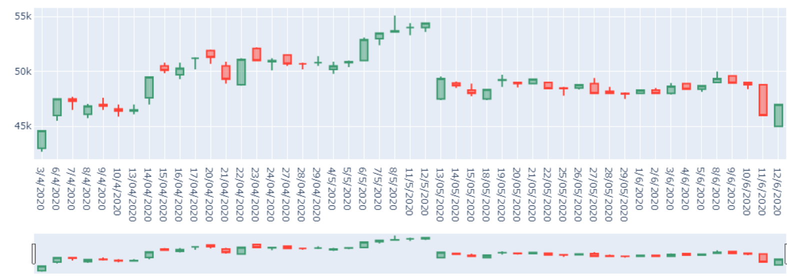

Figure 2: Candlestick chart of FPT Stock

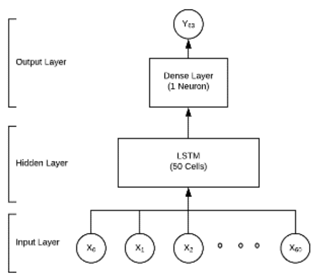

Figure 3: Vanilla LSTM architecture

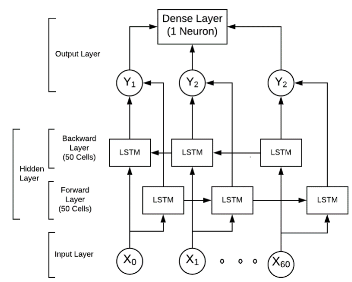

Figure 5: Bidirectional LSTM architecture

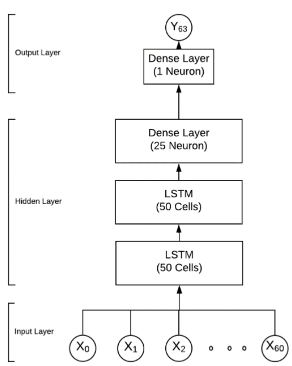

Figure 4: Stacked LSTM architecture

the stock closing price on the third day into the future. Hence, we split both sets into batches. Each batch contains all five features of the data from the past 60 days and the closing price of the third day into the future.

3.3 Model Architectures

The model architectures are illustrated in Figure 3, Figure 4, and Figure 5

3.4 Model evaluation

We evaluated the prediction results and the established prediction model by Root Mean Square Error (RMSE) and Mean Average Percentage Error (MAPE). The smaller the RMSE and, the closer the predicted value to the actual value. The smaller the MAPE, the better the prediction accuracy. The formulas for both error metrics are shown below:

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}}$$

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%$$

where $\hat{y}_i$ stands for predicted values, y_i is the set of actual values and n is the number of observations.

4 RESULTS AND DISCUSSIONS

4.1 Overall analysis

The MAPEs and average MAPE of all predicted stock prices are indicated in Table I. Overall, the Bidirectional LSTM model was the most precise model with a MAPE score of 4.3373%, which means it was able to predict stock prices with an accuracy of approximately 95.6627%. Following the bidirectional LSTM model is the Vanilla LSTM model with an accuracy of roughly 95.6045%. The Stacked LSTM was the least accurate model produced the largest MAPE score 4.6397%. The reason is that an overly complex model might overfit the data and perform poorly on the test set. This paper's result is similar with a study conducted by Jia et al [21]. Both papers find that the two-way LSTM is more accurate than the one-way LSTM. However, the margin between the models' errors are not remarkable, which is in line with Greff et al [17]'s research.

Table 1: The MAPEs of the predictions

Ticker	MAPE(%)		
	Vanilla LSTM	Stacked LSTM	Bidirectional LSTM
DAE	2.6359	2.093	2.4078
DBC	6.4411	5.1788	5.5668
EIB	1.7832	2.3836	4.1074
FPT	2.9333	3.4630	3.0518
HAG	6.3326	5.2735	5.4853
HAH	4.0916	3.6789	3.4520
HBC	5.6014	5.2811	5.1547
HPG	3.9972	6.9380	3.8480
HQC	6.5203	7.2600	4.0380
ITD	2.9273	2.5404	6.0313
JVC	6.5768	5.8470	8.5273
MBB	3.1356	3.5524	3.2734
MBS	4.1071	6.2036	4.0111
NAF	2.9747	3.2928	3.3552
NT2	3.5592	6.3567	2.4261
OGC	4.8741	4.3266	3.5801
SHS	6.4462	7.1194	6.0955
STB	3.6628	3.5823	4.1490
STK	4.9134	3.7893	3.8487
VND	2.5302	2.1247	2.3071
Average	4.3022	4.5139	4.2358

4.2 Detailed Analysis

We used the RMSEs for in-depth analysis. Table II specified the RMSEs for all three experimented models in each trading company. The Bidirectional LSTM had the lowest RMSE for 9 stocks. Surprisingly, the overall least accurate model stacked LSTM had the closest predictions for 7 companies whereas Vanilla LSTM only produced the closest predictions for 4 stocks. This is unexpected as LSTM was ranked last in terms of accuracy, trailing the other two models by a fair margin. This is the consequence of the Stacked model performing terribly in a few stocks (HPG, NT2)

The predictions of all three models fit the actual stock values similarly for the majority of the chosen trading firms. All three predicted closing prices of eight stocks are within a close range (100 VND) of each other. The rest of the stocks also have similar predicted prices range aside from a handful of glaring exceptions (EIB, HPG, NT2, ITD, MBS). We delved in the depth of these irregularities in the next few paragraphs.

Figure 6, Figure 7, Figure 8, Figure 9, and Figure 10 showed the graphs of the actual values and predicted values of the 5 mentioned stocks from all three models. Firstly, we split these cases in 2 different categories. The first category was for cases that the RMSE value of one model differs heavily from the rest. The second category are for cases that all three models had a significant variance between their RMSE scores. There were 4 stocks (EIB, HPG, ITD, MBS) belong to the first category whereas only NT2 was classified as a part of the second category.

Firstly, we examined the first category. In the case of the EIB stock, we observed that the Stacked LSTM and Vanilla LSTM perform similarly with fairly close predicted values (475.0372 and

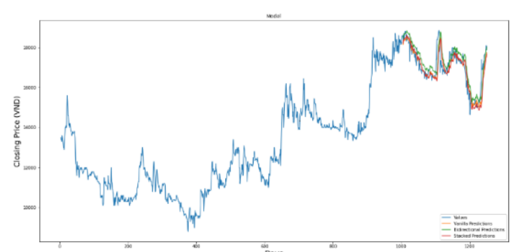

Figure 6: The predicted and actual values of EIB

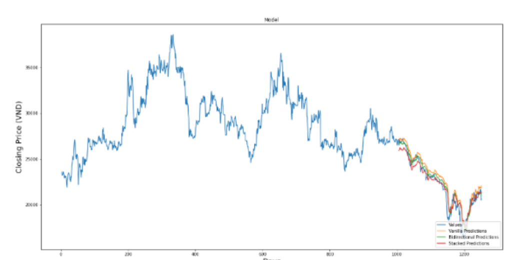

Figure 7: The predicted and actual values of HPG

525.8233 VND respectively) while the Bidirectional LSTM predictions had a huge RMSE score of 818.7812VND. On the other hand, the Stacked LSTM had the worst RMSE value of a staggering 1763.207VND, which is almost twice as much as the other two models as both of their RMSE values are roughly 1000VND for the

Table 2: The RMSEs of the predictions

Ticker	RMSE(VND)		
	Vanilla LSTM	Stacked LSTM	Bidirectional LSTM
DAE	590.7633	516.1296	565.9940
DBC	2272.3792	2104.2446	2155.9364
EIB	475.0372	525.8233	818.7812
FPT	2050.0736	2243.1667	2082.2400
HAG	290.4883	280.0524	260.2646
HAH	628.4380	584.9711	562.3817
HBC	754.6611	721.5700	715.4195
HPG	1073.8445	1763.2075	1054.3759
HQC	117.6520	124.3432	90.6533
ITD	378.8502	352.9765	652.7074
JVC	284.1691	274.6918	384.9212
MBB	806.5758	940.0901	822.9345
MBS	647.4024	893.7921	637.0088
NAF	978.0251	1089.0234	1109.7784
NT2	949.4435	1541.6851	709.1292
OGC	202.3562	184.5116	164.0981
SHS	743.6162	799.2756	730.7302
STB	554.3853	534.8999	9.7284
STK	1149.7765	913.2289	930.5328
VND	441.4045	407.8215	416.3974

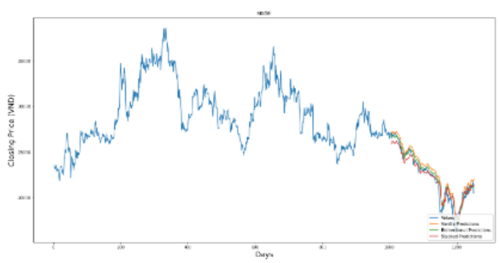

Figure 8: The predicted and actual values of NT2

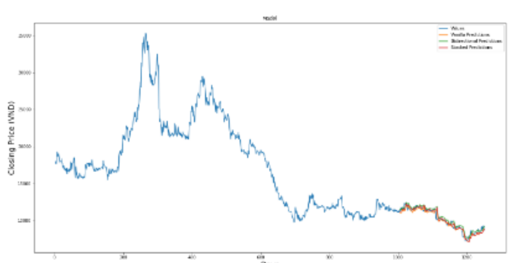

Figure 9: The predicted and actual values of ITD

HPG stock. For the case of the ITD stock, Bidirectional LSTM is the odd one out three among the with a RMSE value of 652.7074 VND, whereas both Stacked LSTM and Vanilla LSTM projected RMSE values of 378.8502 and 352.9762VND respectively. The last case, MBS, featured a comparatively different RMSE score of Stacked LSTM: 893.7921VND while the other models only had RMSE values of

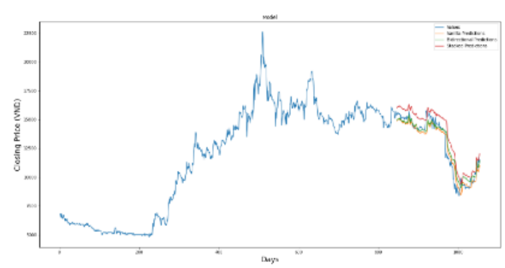

Figure 10: The predicted and actual values of MBS

roughly 650.000VND. Notably, the Vanilla LSTM never significantly underperformed compare to the stacked LSTM and bidirectional LSTM. Thus, we concluded that the issue lied in the complex architecture of the Stacked and Bidirectional LSTM models. The stock market is influenced by many factors other than the stock prices such as the overall economy, investors' reading habit and even rumors [22][23]. Both models tried to put too much emphasis on the available numerical features while ignoring other unquantifiable factors that were not included in the features.

Lastly, we inspected the second category. NT2 is the only case in this category. The RMSE values of all models differ from each other by a huge margin: 949.4436, 1541.6851 and 709.1292VND for Vanilla, Stacked and Bidirectional respectively. Our hypothesized reason is that the data from the test set and training set for this stock follow completely different patterns. We noticed that the training set constantly fluctuate in a wide range, whereas the test set generally followed a downward trend with a sudden recovery in the end. It is likely that the models fit the training set very well but couldn't do the same on a training set that follow a completely different trend.

5 CONCLUSION

This paper empirically studies three different variants of LSTM (Vanilla, Stacked and Bidirectional LSTM) in stock price prediction. Overall, the models had similar performance with Bidirectional LSTM being the most accurate. However, there were irregular cases and we were able to analyze in depth and narrow down the causes.

Our future work has several directions. Our proposed models were able to predict future stock closing prices with a notable accuracy. Nonetheless, simply considering the impact of numerical features is somewhat lacking and may not be able fully make use of the LSTM architecture. Therefore, we can directly process or quantify qualitative features such as stock-related news and company information to enhance and the accuracy of the models. Moreover, we can implement an attention model to improve the performance of the LSTM models.

REFERENCES

[1] Bontempi, G., Ben Taieb, S. and Le Borgne,Y. 2013. Machine Learning Strategies for Time Series Forecasting. Business Intelligence, 62-77. DOI= 10.1007/978-3-642-36318-4_3.
[2] Nikfarjam, A., Emadzadeh, E. and Muthaiyah, S. 2010. Text mining approaches for stock market prediction. 2010 The 2nd International Conference on Computer and Automation Engineering (Singapore, 2010). 256-260. DOI= 10.1109/ICCAE.2010.5451705
[3] Lin, Y., Guo, H. and Hu J. 2013. An SVM-based approach for stock market trend prediction. The 2013 International Joint Conference on Neural Networks. (Dallas, TX, 2013).1-7. DOI= 10.1109/IJCNN.2013.6706743
[4] Wanjawa, B. W. and Muchemi, L. 2014. ANN Model to Predict Stock Prices at Stock Exchange Markets. arXiv:1502.06434
[5] Zhang, R., Yuan, Z. and Shao, X. 2018.A New Combined CNN-RNN Model for Sector Stock Price Analysis. 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC). (Tokyo, 2018). 546-551. DOI= 10.1109/COMPSAC.2018.10292
[6] Kolen, J. and Kremer,S. 2001. Gradient Flow in Recurrent Nets: The Difficulty of Learning Long-Term Dependencies. A Field Guide to Dynamical Recurrent Networks. 237-243. DOI= 10.1109/9780470544037.ch14
[7] Zhao, Z., Rao, R., Tu, S. and Shi, J. 2017. Time-Weighted LSTM Model with Redefined Labeling for Stock Trend Prediction. 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (Boston, MA, 2017). 1210-1217. DOI= 10.1109/ICTAI.2017.00184
[8] Jiang, Q., Tang, C., Chen, C., Wang, X. and Huang, Q. 2018. Stock Price Forecast Based on LSTM Neural Network. Proceedings of the Twelfth International Conference on Management Science and Engineering Management (2018). 393-408. DOI= 10.1007/978-3-319-93351-1_32
[9] Do, Q. and Trang, T. 2020. Forecasting Vietnamese stock index: A comparison of hierarchical ANFIS and LSTM. Decision Science Letters. 9, (2020). 193-206. DOI: 10.5267/j.dsl.2019.11.002
[10] Lien Minh, D., Sadeghi-Niaraki, A., Huy, H. D., Min, K. and Moon, H. Deep Learning Approach for Short-Term Stock Trends Prediction Based on Two-Stream Gated Recurrent Unit Network. IEEE Access. 6, (2018). 55392-55404. DOI= 10.1109/ACCESS.2018.2868970
[11] Luu, Q., Nguyen, S. and Pham, U. 2020. Time series prediction: A combination of Long Short-Term Memory and structural time series models. Science & Technology Development Journal – Economics – Law and Management. 4, 1 (2020). 500-515. DOI= 10.32508/stdjelm.v4i1.593
[12] Huynh, H., Dang, L. and Duong, D. 2017. A New Model for Stock Price Movements Prediction Using Deep Neural Network. Proceedings of the Eighth International Symposium on Information and Communication Technology. (2017). 57-62. DOI= 10.1145/3155133.3155202
[13] Ngo Tung Son, Le Van Thanh, Tran Quy Ban, Duong Xuan Hoa, and Bui Ngoc Anh. 2018. An analyze on effectiveness of candlestick reversal patterns for Vietnamese stock market. In Proceedings of the 2018 International Conference on Information Management & Management Science (IMMS '18). ACM, New York, NY, USA, 89-93. DOI: https://doi.org/10.1145/3277139.3277161
[14] Sak, H., Senior. A and Beaufays, F. 2014. Long short-term memory recurrent neural network architectures for large scale acoustic modeling. INTERSPEECH. (2014). 338-342
[15] Hochreiter, S. and Schmidhuber, J. 1997. Long Short-Term Memory. Neural Computation. 9, 8 (15 Nov. 1997). 1735-1780. DOI= 10.1162/neco.1997.9.8.1735
[16] Brownlee, J. Deep Learning with Python: Develop Deep Learning Models on Theano and TensorFlow Using Keras; Machine Learning Mastery: Melbourne, Australia, 2017
[17] Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R. and Schmidhuber, J. 2017. LSTM: A Search Space Odyssey. IEEE Transactions on Neural Networks and Learning Systems. 28, 10 (Oct. 2017). 2222-2232. DOI= 10.1109/TNNLS.2016.2582924.
[18] Nelson, D. M. Q., Pereira, A. C. M. and de Oliveira, R. A. 2017. Stock market's price movement prediction with LSTM neural networks. 2017 International Joint Conference on Neural Network (Anchorage, AK, 2017). 1419-1426, DOI= 10.1109/IJCNN.2017.7966019.
[19] Ojo, S. O., Owolawi, P. A., Mphahlele, M. and Adisa, J. A. 2019. Stock Market Behaviour Prediction using Stacked LSTM Networks. 2019 International Multidisciplinary Information Technology and Engineering Conference (Vanderbijlpark, South Africa, 2019). 1-5. DOI= 10.1109/IMITEC45504.2019.9015840.
[20] Zou, Zhichao and Zihao Qu. "Using LSTM in Stock prediction and Quantitative Trading." (2020).
[21] Jia, M., Huang, J., Pang, L. and Zhao, Q. 2019. Analysis and Research on Stock Price of LSTM and Bidirectional LSTM Neural Network. Proceedings of the 3rd International Conference on Computer Engineering, Information Science & Application Technology (ICCIA 2019). DOI= 10.2991/iccia-19.2019.72
[22] Cruz, F. and Gomes, M.. The influence of rumors in the stock market: a case study with Petrobras. Transinformação, 25, 3 (2013). 187-193. DOI= 10.1590/s0103-37862013000300001
[23] Fedyk, A. 2016. Research: How Investors' Reading Habits Influence Stock Prices: 2016. Harvard Business Review, 2016