An Approach for Stock Price Forecast using Long Short Term Memory

K.A.Surya Rajeswar[1], Pon Ramalingam[2], Sudalaimuthu.T[3]

[1,2] Department of Management Studies, Hindustan Institute of Technology and Science, Chennai
[3] Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Chennai

Abstract
The Stock price analysis is an increasing concern in a financial time series. The purpose of the study is to analyze the price parameters of date, high, low, and news feed about the stock exchange price. Long short term memory (LSTM) is a cutting-edge technology used for predicting the data based on time series. LSTM performs well in executing large sequence of data. This paper presents the Long Short Term Memory Model has used to analyze the stock price ranges of 10 days and 20 days by exponential moving average. The proposed approach gives better performance using technical indicators of stock price with an accuracy of 82.6% and cross entropy of 71%.

Keywords:
Stock Price, Long short time memory, Cross entropy, Accuracy.

1. Introduction

Stock market price trends raised the financial time series to predict the volatility of stock volume. The use of price trends has been fit into the exchange-traded funds to the trading mentor. The purposes are: understand their trading volume condition using the stock exchange. The raise of stock exchange prediction has brought challenges for social market prices. The feature dimension of price trends was proposed and various approaches was used for forecasting the performance of price predictions with the various role of time, currency, volume using machine learning techniques. The role of the price trends to deliver the liquidity, market capitalization, and mutual fund for the statistical measure of stock moves in the market. The trading mentor emphasizes the real-time information and coming up with pioneering ways to offering the stock exchange. The study originates the price of the stock to fix the initial public offering. A stock market price prediction platform is defined as correlating the feature dimension of market efficiency and stock market price prediction' demands. The field of study deals with the price trends to BID. It has been referred to as the part of the forecasting performance of price. A challenging problem that arises in the stock exchange is the complexity of trade value between the portfolios of price trends. The enormous market order has been generated by the rapid development of the stock exchange, the volume and trade value has been increased in the shareholders, it's become more complicated to the intraday trading. Moreover, few studies have focused on the selection and scoring of the task, capitalization of investment, and performance of investors. LSTM approach to resolving the problem to the price prediction of stocks, analyze the volume, the trade value, the trade type, and investment market in the stock. The exchange of market price has been formed between the trading mentor and investor by the characteristics of stock selection. This study, investigates, to predict the performance of stock market price trends by the supervised learning approach. One of the major aims of the study inferred to the investors in a stock selection and price prediction, this combination to predict the properties in the stock exchange market. In this paper, a study has been undertaken to predict the price trends of stocks.

2. Literature Review

[1] had approached Wavelet Neural Network to optimize the attributes of stock price trends. Using Rough Set theory to reduce the feature dimensions of stock price trends. [2] had exploited Wavelet Neural Network (WNN) to predict stock price trends. By rough set theory, the data has been analyzed by the 5 indices such as composite index, CSI 300 Index, Australian ordinaries index, Japan Nikkei index, and USA jones index. It has determined by the wavelet neural network. Evaluations were based on optimizing the feature dimensions and reduce the computational complexity. The author has discussed the parameter adjustment and weakness of the complexity analysis. [3] had approached the hybrid selection method with the classification model of support vector machine. The model has carried out the prediction of stock price trends. The dataset has analyzed by the Index in Taiwan Economic Journal Database 2008. It has been designed by the performance under the parameter values, supported by the searching techniques of sequential forward search. The limitations of the performance were compared by the machine learning algorithm of the Backpropagation model their computational complexity has reduced 15 % compared with the proposed algorithm. But it gives the structure of the feature selection model [4] the deep learning solutions of a

Manuscript received April 5, 2023
Manuscript revised April 20, 2023
https://doi.org/10.22937/IJCSNS.2023.23.4.22

universal dataset of financial market trends. It includes the sale records of the transactions of the stock exchange. The LSTM units have been used as the rectified linear units with the gradient descent method. This algorithm had optimized by the cost. It's able to cover the stock more than the test data. This method has effectively reduced the computational complexity. [5] had predicted the price trends by Bayes classification and performed by the fractal feature selection. Shanghai Stock Exchange Composite Index (SSECI) dataset has been used to analyze the price trends of the stock exchange with different technical indicators.

In the pre-processing, the data has to be normalized, and the combination was associated by the Association rule mining, By K-cross validation the data has been validated and the performance has been measured by the regression techniques of a polynomial vector regression model. [6] had analyzed by the micro and macro factors of the financial domain. By the cross-validation and combination of an association rule, the accuracy has been analyzed.

The evaluation results were tested by the hyper parameter of combinations. [7] had approached the price trends of Bitcoins. It has been optimized by the bourta algorithm for preprocessing the price distribution. By random forest classifier, LSTM parameters have been optimized. The analyzed dataset has been taken from the year 2013 to 2016. [8] had improved the performance of the financial stock index. It analyzes the movement of Istanbul Stock Exchange price for the financial index. Artificial Neural networks had been used to analyze the movement of stock exchange prices. [9] had developed the random walk-based model to predict the status of the company using the ant colony optimization method to predict the performance of stock trading. [10] had improved the performance using the Support Vector Machine (SVM) method to analyze the HSBC stock price.

It compared the performance of the existing algorithm and proposed algorithm for predicted the stock price. [11] had analyzed the performance of feature selection methodology of PCA and Sequential Forward Selection (SFS) with SVR and concluded that PCA performs better accuracy than SFS. [12] had analyzed the Japan stock price, the study used to analyze the principal component analysis method to monitor the stock price rate.

3. Approaches for Predicting Stock Price

3.1 Stock Price Dimensions

The time series of data has a different dimension to predict the stock price. The data contains stock values of various organizations. The parameters are date, max, min, start, end, and volume.

Date – On which date the price has been taken into account.
Max- Denotes the highest price of the sold-out.
Min – Denotes the lowest price of the sold-out.
End- Close, Price of the stock.
Volume – quantity of the stock.

The dataset has analyzed by machine learning methods [13]. The dataset has split into training and testing data. The train data has 100 records, test data has 50 records. It has been analyzed by the package of caTools.

3.2 LSTM Predictions

By Recurrent Neural Network (RNN) Long Short Term Memory (LSTM) parameters has been trained in a supervised learning method, On a trained data using the optimization algorithms of gradient descent (GD) has paired with backpropagation algorithm (BPA) through the date to compute the GD during the optimization of methods. To change the volume of the stock values of the LSTM is proportional to the derivative error of the price prediction accuracy.

LSTM model is an extensive study of RNN, it mainly handles the price prediction of the daily market values. The inputs of date, the highest price of the stock, lowest price, closing price, and quantity of the stock have considered the feedback of the output and stored in a short period time of memory. The fields of stock price and composition of market exchange values of information have been stored for a long period. A reference of stored information is quite long to predict the time series of price values.

But RNN is exploding the gradients during the training time through the backtracking method. LSTM has handled the feature scaling of the noisy distribution of data and provides a range of parameters. So the complexity has reduced with the O (1). The feature selection part of the DATE parameter has categorized into two different ranges, moving range of 10 days, 20 days of close price values, and the difference between the close price values. It denotes by the A, B, and D.

The main contribution of the categorizing the date parameters has finds the keen accuracy [14] of the forecasting performance of stock price trends in the market. In supervised learning, the first step is labeling the data. Price trends were compared by the Max, and Min of the sold-out stock price range has been identified. The range has been different in all labels, by nominal scaling all the ranges were scaled 0 to 1. Class indicators are denoted by T incorporate different parameters of time and length. It analyzes expected trends and fluctuation in different aspects.

Table 1: Categorize the DATE by moving the range of prices

Class	Meaning
A	10 days close price moving range
B	20 days close price moving range
D	Difference between the stock price ranges of 10.
MFI	Money Flow Index
RSI	Relative Strength Index

4. Results and Discussions

Split the dataset to evaluate the technical indicators with various variables of time, length, number of stock ids, and values during the period. By cross-validation [15, 16], the optimal hyper parameters were identified in the model. The study has used to split the 5 fold cross-validation to train and validate the parameters. An initial train set is said to be 'k's, for each iteration 'k+1" for the validation to split up the dataset. The total number of stock values was 150. The train loss validation results were below the level of the stock exchange, the values were decreased in order of the constituents. The Validation loss was very high in the results. The Sector variables of time, price, currency, and volume were high in that evaluation, so, the stock exchange rate of finance is higher in the companies.

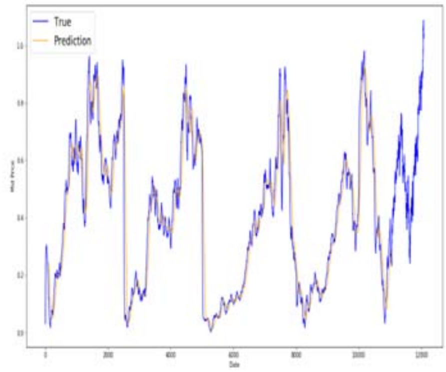

Fig.1. Prediction for 20days

The Average standard deviation of price is 4.56, volume is 3.82, and Sold out price is 3.45. By understanding the statistical significance P= 0.06 the stock price values were evaluated under the better effectiveness of the stock indicators. Figure shows the short prediction range of 0 to 20 days overnight shown in Figure 1, these behaviors are sensible in the features of time, price, and currency for the stock selection task. For better performance, we use the technique of exponential moving average [17]. By using calculated the prediction of the next 10 days by using the stock volume or quantity shown in Figure 2. The Adam optimizer to use to predict the next 10 days price prediction strategy.

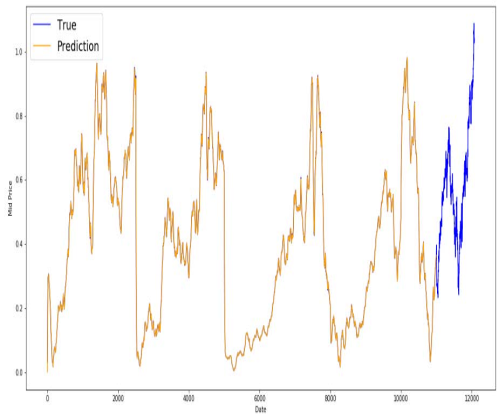

Fig. 2. Prediction for 10 days moving range

(i) Loss Calculation and Optimizer

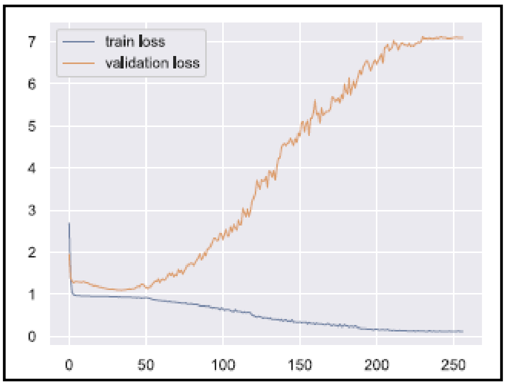

Fig. 3. Comparison of train and validation loss for prices

The unique characteristic when calculating the loss, for each set of predictions and true outputs of train loss and validation loss. From Figure 3 Accuracy has been evaluated and shown in the graph. The training accuracy and validation accuracy were yielded good perf4 for testing accuracy.

Performance between the price variables of 150. The training accuracy yielded a better level of validation using time and volume. The trained accuracy [18] higher in the range of 1.0. It belongs to the range of 0.3 to 1.0 for the 150 Stock variables in the LSTM model shown in Figure 4.

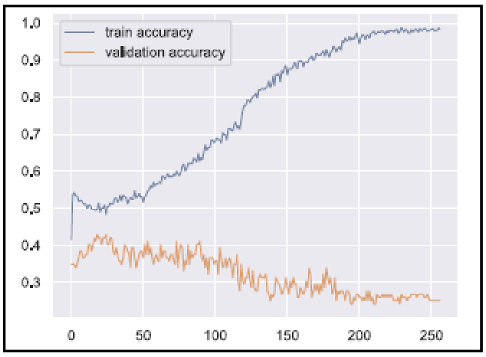

Fig. 4. Comparison of train and validation performance for prices

The training accuracy has linearly raised the level from 0.4 to 1.0 from initial id to stock 150. The training process of the price distribution, the categorical curves in the validation, and the trained accuracy of levels. From 0 to 50 epoch accuracy level has gradually increased up to 100 epoch. Due to the subsequent iterations, 150 epoch onwards.

Consequently, the trained accuracy level 0.4 onward rose eventually up to the range of 1.0. The stock exchange price, finance, commerce, and properties rate were eventually predicted in the model. From observation sectors mean, standard deviation, and maximization values are yielded good accuracy by the LSTM model.

The Metrics of training and testing data measurement are shown in Table 2. The training metrics of the LSTM model have been analyzed by Cross-Entropy and Accuracy [19,20]. For testing metrics analyzed by the Cross entropy, Accuracy, and F1 measures. It yielded better results compared to the existing methodologies. The Analyzed values of cross-entropy metrics for the training dataset were 0.085. The measured accuracy value of the stock exchange price was 81.23%. The test dataset metrics of cross-entropy value as 0.075. The measured accuracy value was 82.86% and F1 was 8.289.

Table 2: Data Split and metrics of training

Model	Metrics of Training	Metrics of Test
LSTM	Cross entropy, Accuracy	Cross entropy, Accuracy, F1

We adopt 3-fold time split cross-validation to train and validate models on each attributes. The time series split method uses the first k splits for training in each iteration, and the k + 1 split as the validation set.

(ii) Test Predictions Overtime

The LSTM performance of sectors was mid-price of 12 to 14 ranges as in the stock price values of 1100 to 11200. From 11400 the range raises the linearly high in the mid-price of 0.8 value. The highest accuracy raised over 12000. Eventually, the volume of mid-price was 11000, 11200, 11400, 11600, 11800, 12000, 12200, and 12400 in Figure 5 respectively.

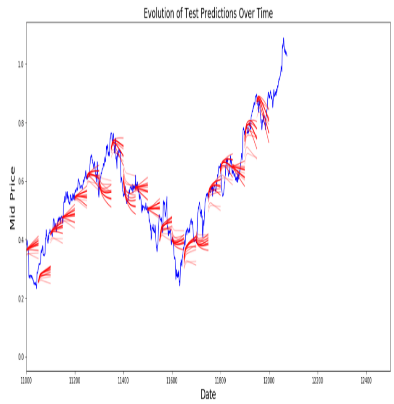

Fig. 5. Evolution of Test price prediction overtime

Based on the analysis of the effectiveness of the information high accuracy values of sectors gives better stock price to predict the outperformance using the price data distribution.
The price prediction steps were followed by the learning rate of the optimizer, the nominal scale of the feature set, by using Adam optimizer to perform the standard LSTM performance for price prediction.

5. Conclusion

The insight of the study has analyzed the Stock Exchange price ranges of 10days and 20 days. The association of the dataset produced the actionable information of train loss and validation loss of the parameters. The price concerning each attribute of date, max price, min price, the volume of everyday stock. The test prediction over time gives the prediction performance of date Vs price. By the analyzed values using cross-entropy metrics for the training, the dataset has measured the accuracy value of the stock exchange price was 81.23%. The test dataset metrics of cross-entropy value as 0.075. The measured accuracy value was 82.86%.

References

[1] Althelaya, K. A., & El-Alfy, E.-S. M, (2018) Evaluation of bidirectional LSTM for short-and long-term stock market prediction. IEEE International Conference on Information and Communication Systems (ICICS), 71-78.

[2] Jiang. (2020). Applications of deep learning in stock market prediction: recent progress. Statistical Finance.

[3] Kang Zhanga, Guoqiang Zhonga, Junyu Donga, ShengkeWangaYongWanga, (2019). Stock Market Prediction Based on Generative Adversarial Network,Procedia Computer Science, 147, 400-406.

[4] EhsanHoseinzade, Saman Haratizadeh. (2019). CNNpred: CNN-based stock market prediction using a diverse set of variables. Expert Systems with Applications, 129(1).

[5] P.Gandhmal, K.Kumar, D. (2019). “Systematic analysis and review of stock market prediction techniques”, Computer Science and Review, 134.

[6] Zhang, Xi, Yunjia Zhang, Senzhang Wang, Yuntao Yao, Bin Xing Fang and Philip S. Yu(2018) “Improving stock market prediction via heterogeneous information fusion.” Knowl. Based Syst. 143, 236-247.

[7] Nabipour, M., P. Nayyeri, H. Jabani, A. Mosavi, E. Salwana, and Shahab S. 2020. "Deep Learning for Stock Market Prediction" Entropy 22, no. 8: 840.

[8] Chung, Hyejung, and Kyung-shik Shin. 2018. "Genetic Algorithm-Optimized Long Short-Term Memory Network for Stock Market Prediction" Sustainability 10, no. 10: 3765.

[9] Li, Xiaodong, Pangjing Wu and Wenpeng Wang,(2020) “Incorporating stock prices and news sentiments for stock market prediction: A case of Hong Kong.” Inf. Process. Manag. 57: 102212.5).

[10] Xingyu Zhou, Zhisong Pan, Guyu Hu, Siqi Tang, and Cheng Zhao (2018). “Stock Market Prediction on High-Frequency Data Using Generative Adversarial Nets. Computational Intelligence in Data-Driven Modelling and Its Engineering Applications.

[11] Pang, Xiong Wen, Yanqiang Zhou, Pan Wang, Weiwei Lin and Victor I. C. Chang(2018). “An innovative neural network approach for stock market prediction.” The Journal of Supercomputing 76 : 2098-2118.

[12] Moghar, Adil, Hamiche, Mhamed (2020). Stock Market Prediction Using LSTM Recurrent Neural Network. Procedia Computer Science, 170, 1168-1173.

[13] J. Lee, R. Kim, Y. Koh and J. Kang (2019), "Global Stock Market Prediction Based on Stock Chart Images Using Deep Q-Network," in IEEE Access, vol. 7, pp. 167260-167277.

[14] Nam, Ki-Eun and Nohyoon Seong (2019), “Financial news-based stock movement prediction using causality analysis of influence in the Korean stock market.” Decis. Support Syst. 117: 100-112.

[15] Ananthi, M. & Vijayakumar, K.. (2021), “Stock market analysis using candlestick regression and market trend prediction (CKRM)”. Journal of Ambient Intelligence and Humanized Computing. 12.

[16] Wei Huang, Yoshiteru Nakamori, and Shou-Yang Wang. (2015), “Forecasting stock market movement direction with support vector machine”. Comput. Oper. Res. 32, 10, 2513–2522.

[17] Mohamed M. Mostafa, (2016) Forecasting stock exchange movements using neural networks: Empirical evidence from Kuwait. Expert Syst. Appl. 37, 9, 6302–6309.

[18] Hyun Joon Jung, Aggarwal J.K (2017), "A Binary Stock Event Model for stock trends forecasting: Forecasting stock trends via a simple and accurate approach with machine learning", 11th International Conference on Intelligent Systems Design and Applications (ISDA), Pages 714- 719, 2011.

[19] Jigar Patel, Sahil Shah, Priyank Thakkar , K Kotecha (2015),“Predicting stock and stock price index movement using Trend Deterministic Data Preparation and machine learning techniques”, Expert Systems with Applications, Volume 42, Issue 1, Pages 259-268.

[20] Shin-Fu Wu, Shie-Jue Lee(2015), "Employing local modeling in machine learning-based methods for time-series prediction", Expert Systems with Applications, Volume 42, Issue 1, Pages 341-354.

K.A.Surya Rajeswar Completed B.E in Computer Science and Engineering at Hindustan Institute of Technology and Science. He has completed Masters of Business administration at Hindustan Institute of Technology and Science. He is pursuing PhD in the department of management studies and his area of research includes stock price prediction using machine learning techniques. He has published papers in international conferences.

Pon Ramalingam is working as a Registrar and professor in the Department of Management studies at Hindustan Institute of Technology and Science, Chennai, India. He obtained more than 15 publications in the reputed International Journals. His research areas includes management studies. He is a life time member of CSI, ISTE, he is a member of ACM and IEEE.

T. Sudalaimuthu is working as a Professor in the Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Chennai, India. He completed his Ph.D degree from Hindustan University, Chennai. India. He is a Certified Ethical Hacker. He obtained more than 45 publications in the reputed International Journals. He obtained many awards in his career includes Best Teacher Award, Best innovator Award. His research areas are Network Security, Advanced Computing. He is a life time member of CSI, ISTE, he is a member of ACM and IEEE. He received both Indian and Australian Innovation patent.