Ilkogretim Online - Elementary Education Online, 2021; Vol 20 (Issue 3): pp. 3182-3188
http://ilkogretim-online.org
doi: 10.17051/ilkonline.2021.03.329

An Artificial Intelligence Based Mechanism For Stock Market Analysis And Prediction

Keshav Sharma *, Student Member, IEEE

* Department of Computer Science and Engineering, Graphic Era Deemed to be University, Dehradun 248 002, India.

Dr. Prateek Srivastava**, Associate Professor, Department of Computer Science and Engineering, Graphic Era Hill University, Dehradun, *keshav.sharma154201@gmail.com

Abstract

Stock markets are always an attractive investment way to grow capital. With the development of communication technology, the stock markets are getting more popular among individual investors in recent decades. A stock's value can vary depending on various internal and external factors, with this paper the objective is to give a predictive analysis of a company's stock price forthe future from previous parameters and how artificial intelligence (AI) can be used in stock markets by companies/ individualsto get a fair return on their investments. While year by year, the number of shareholders and companies is growing in the stock markets, many try to find a solution to predict a stock market's future trend. The LSTM method is implemented using python in google colab. The long short term memory (LSTM) makes use of recurrent neural network (RNN) and memory storing and reusing. LSTMs are a complex area of deep learning. It can be hard to get your hands around what LSTMs are, and how terms like bidirectional and sequence-to-sequence relate to the field. To test the performance of the proposed mechanism, we have the dataset of BAJAJ FINSERV from the past 10 years is obtained and visualized. The information given in this dataset is used to forecast future stock prices thorough ML algorithms, i.e., SVM, LSTM with RNNs. In the proposed scheme, we have obtained results, like, R2 score was 0.95 on the test dataset and 0.92 on the train dataset. F1 score was 0.15 and accuracy was 96%.

Index Terms Artificial intelligence (AI), Machine learning, stock, market prediction, simulation.

I. INTRODUCTION

Stock market analysis and prediction have always been challenging tasks as the data is non-linear, non-parametric, andnoisy. The markets also vary according to real-time data which are the developments in the real-world scenarios and mayvary according to the particular country and market index. There have been many ways and techniques to implement machine learning in stock value predictions for example support vector machine (SVM), recurrent neural network (RNN) and long short-term memory (LSTM) are the most effective. Machine learning (ML) requires a vast amount of data, on a regular basisto predict effectively [1], [2], [3]. So we have to make a ML model which is trained on small data and makes our machine learning model artificially intelligent. The business requirements are for market prediction are very high as most of the banking and investment

companies looking to invest in technologies for better prediction of their futures and options chain due to their risk factors. The market prospects for individual needs are also high due to the increase in the number of singular investors and people looking to expand their portfolio, especially during pandemic times, indicating rising interest in retail participation. The need for Machine learning in the stock market goes beyond the requirements for individuals or corporate firms [4], [5], [6],[7]. It can be used by a countries government to analyze the current situation and provide measures to improve the economy, manage inflation and market crashes effectively, with prior warning. Data collected from historical records can be utilized to give various visualizations and predictions on how a stock or index will perform in the near future or at least an estimate [8].The target specification and characterization are as follows.

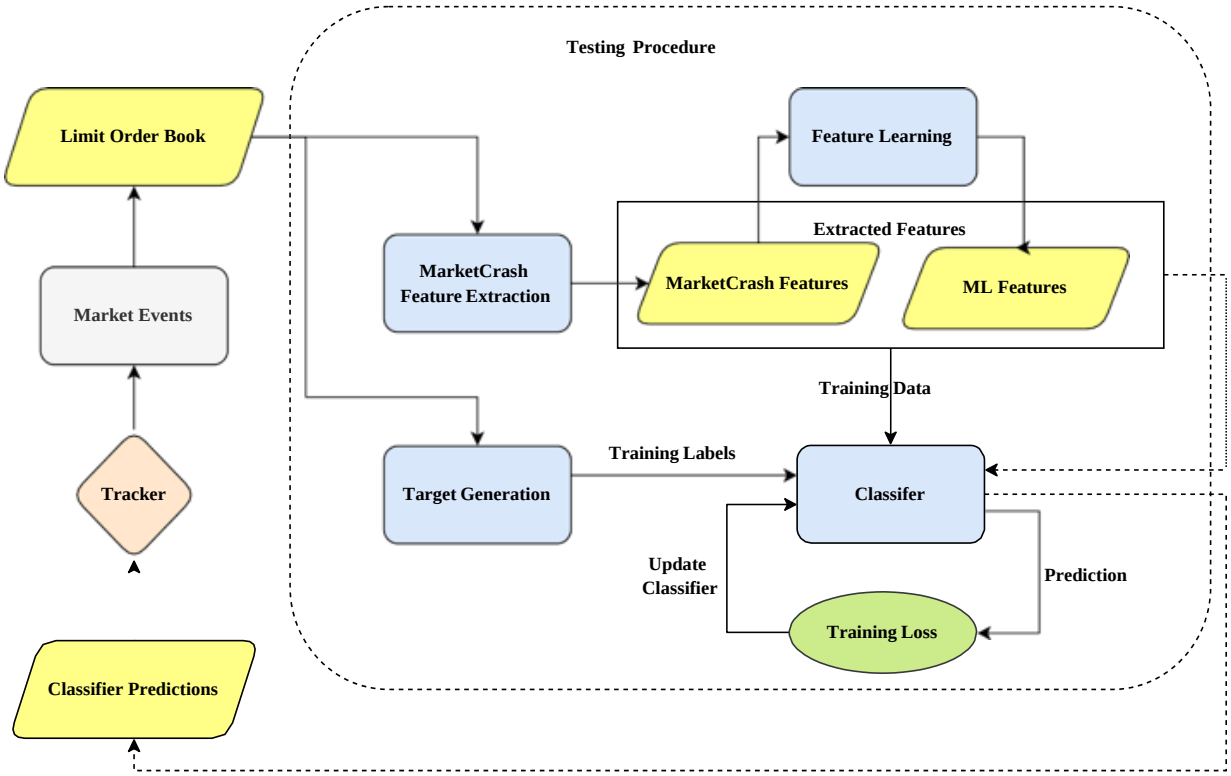
- To enhance the traders view and market predictions.
- Reduce stress and investors/trader's doubt.
- Data visualization and progressive learning.
- Investors' needs i.e., short-term, long-term goals.
- Identifying the current market situations.
- Analyzing the companies' financial, goals, major updates (current/ future).
- Identifying the right time and price point to buy/ sell stock.
- Warning or automatic trigger system to sell at a threshold price point.
- Send news and stock updates automatically.
- Machine learning analysis and predictive data graph of a stock.

II. RELATED WORK

Various researchers have worked for solving the problem of stock market [1], [2], [3]. For such tasks the artificial intelligence(AI) based algorithms seem very helpful. Therefore, these algorithms had been utilized in different papers [4], [5], [6], [7]. Further important results, like, precision, recall, F1 score, computational time, accuracy were estimated and compared [9], [10], [8].

III. PROPOSED SCHEME

The details of the proposed scheme are given below. The process flow diagram of proposed scheme is given in Fig. 1.



Testing data

Fig. 1. Process flow diagram of proposed scheme

A. Business opportunity

Data analysts are good estimators for stock but with the help of AI and data, a bigger and better understanding of the risksand the gains can be obtained and made easier for the consumer of the individual investors and helps a company to reducethe expenses of employees.

B. System designing

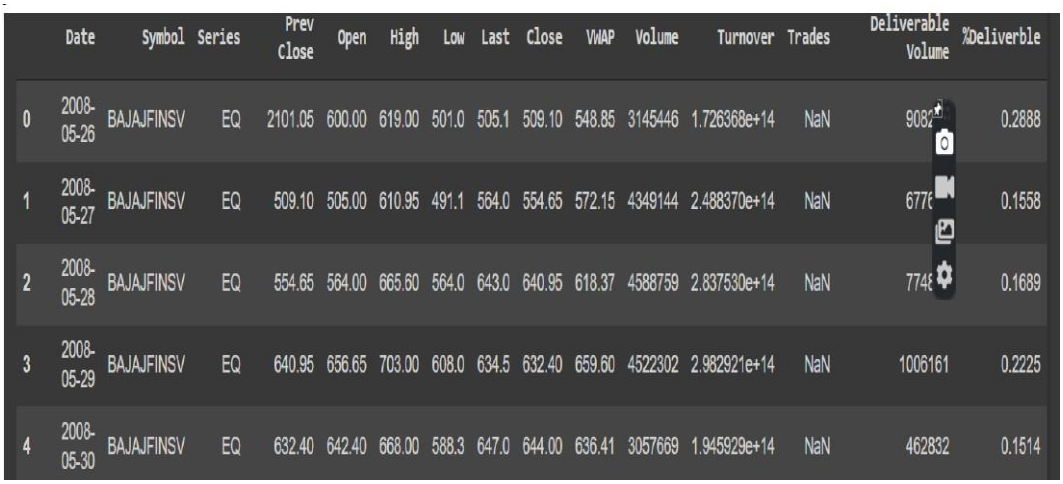
To test the performance of the proposed mechanism, we have the dataset of BAJAJ FINSERV from the past 10 years is obtained and visualized [11]. The information given in this dataset is used to forecast future stock prices using support vector regression and long short-term memory with RNNs.

C. Prototype

"In Google Colab, the LSTM approach is implemented using Python. RNNs, memory storing, and memory reuse are all used by the LSTM. A challenging area of deep learning is LSTMs. Understanding LSTMs and how concepts like bidirectional and sequence-to-sequence relate to the field can be challenging. Few people are better at explaining both the potential of LSTMsand how they operate than the professionals who created them."

D. Algorithmic presentation

1) Support vector machine: "The primary goal of the SVM algorithm is to establish a boundary with the broadest margins, the best line, or a decision boundary that can divide n-dimensional space into classes, allowing us to quickly classify fresh data points in the future. A hyperplane is the name given to this optimal decision boundary. It is simple to have a linear hyper-plane between these two classes in the SVM classifier. The SVM kernel is a function that transforms not separable problems into separable problems by taking a low-dimensional input space and raising its dimension. This is especially helpful in cases involving non-linear separation."



	Date	Symbol	Series	Prev Close	Open	High	Low	Last	Close	VWAP	Volume	Turnover	Trades	Deliverable Volume	%Deliverble
0	2008-05-26	BAJAJFINSV	EQ	2101.05	600.00	619.00	501.0	505.1	509.10	548.85	3145446	1.726368e+14	NaN	9082	0.2888
1	2008-05-27	BAJAJFINSV	EQ	509.10	505.00	610.95	491.1	564.0	554.65	572.15	4349144	2.488370e+14	NaN	6776	0.1558
2	2008-05-28	BAJAJFINSV	EQ	554.65	564.00	665.60	564.0	643.0	640.95	618.37	4588759	2.837530e+14	NaN	7748	0.1689
3	2008-05-29	BAJAJFINSV	EQ	640.95	656.65	703.00	608.0	634.5	632.40	659.60	4522302	2.982921e+14	NaN	1006161	0.2225
4	2008-05-30	BAJAJFINSV	EQ	632.40	642.40	668.00	588.3	647.0	644.00	636.41	3057669	1.945929e+14	NaN	462832	0.1514

Fig. 2. View of used dataset

2) Gradient Boost algorithm: "The method used by XGBoost expands the loss function to the second order using the Taylor expansion, producing enhanced residuals for the first two orders. GBDT is a special case of XGBoost since it turns out thatthe residuals employed by conventional GBDT are the outcome of Taylor's expansion to the first order."

3) XGboost vs. GBDT: "XGboost supports a linear classifier, while GBDT could only use CART as its classifier, First and second derivatives information is used in XGboost compared to only first derivative information is used in GBDT. That is a result of XGboost using the second-order of Taylor expansion. XGboost can better prevent overfitting than GBDT by adding regularization to the cost function, which can reduce the model complexity."

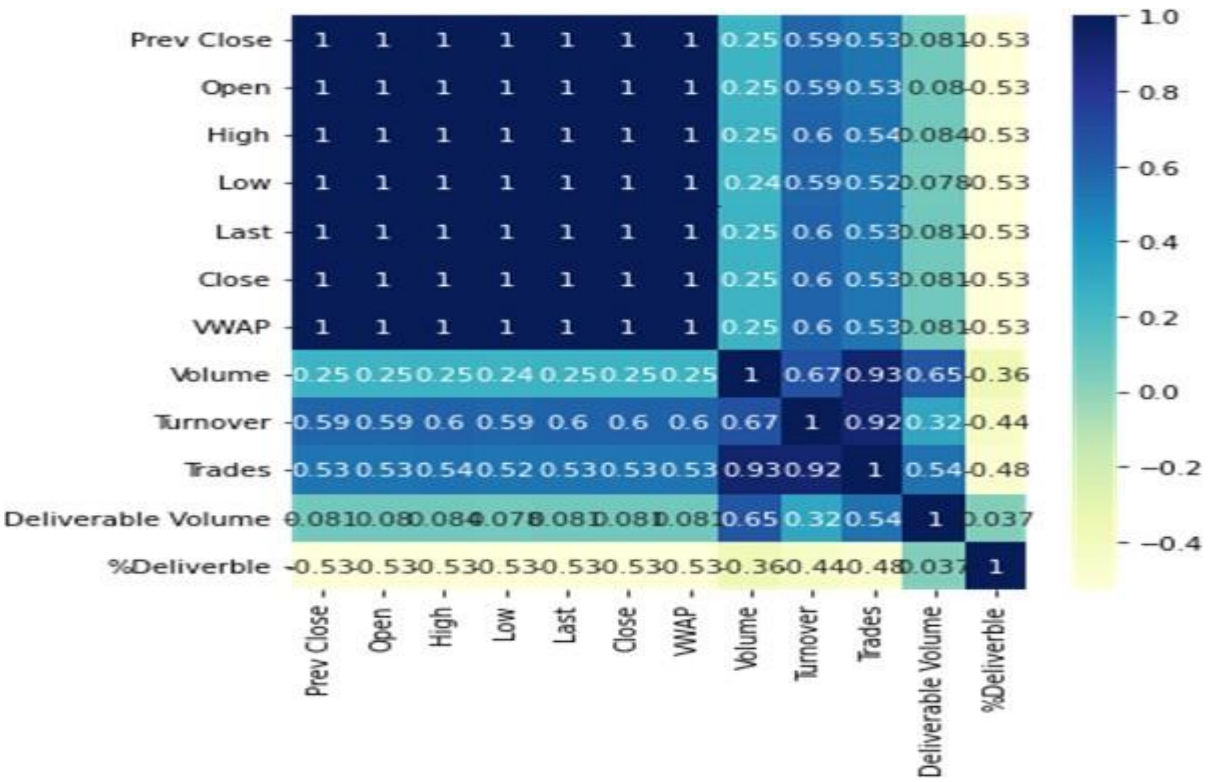


Fig. 3. Performed processing on the dataset

4) Long Short Term Memory (LSTM): "The model is very popular in time series forecasting, and this is the reason why this model is chosen for this task. The historical prices of BAJAJ FINSERV are collected automatically using or downloaded from open source. After pre-processing the columns have main significance in the dataset. The stock prices for this new durationwill be predicted by the already trained LSTM model and the predicted prices will be plotted against the original prices to visualize the model's accuracy."

E. Practical implementation

To test the performance of the proposed mechanism, we have the dataset of BAJAJ FINSERV from the past 10 yearsis obtained and visualized [11]. The information given in this dataset is used to forecast future stock prices thorough ML algorithms, i.e., SVM, LSTM with RNNs. The view of dataset is given in Fig. 2.

The processing, which has been done on the dataset is given in Fig. 3.

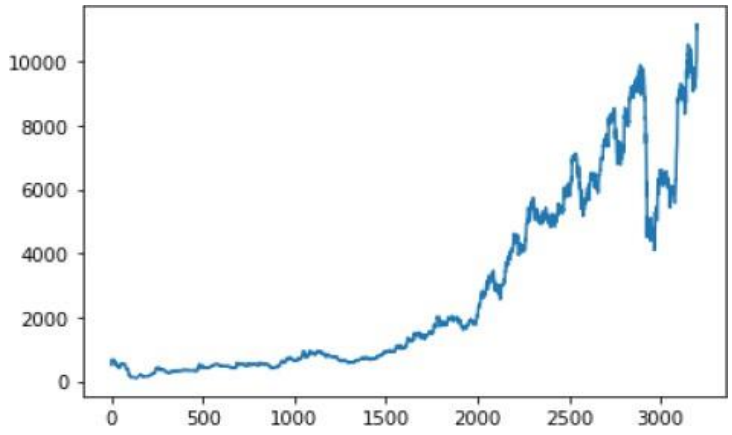


Fig. 4. Graphical view of obtained results

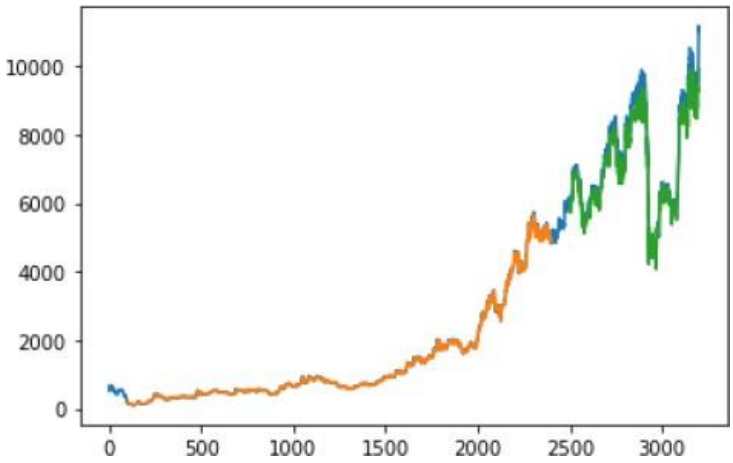


Fig. 5. Graphical view of obtained results

F. Obtained results

We have obtained following results. The R2 score was 0.95 on the test dataset and 0.92 on the train dataset. F1 score was
0.15 and accuracy was 96%. The graphical views of obtained results are given in Fig. 4 and Fig. 5.

IV. CONCLUSION

AI is being used by various departments of the world and is progressively increasing in the way people are learning to implement it. Most hedge fund managers and individuals have already used or are implementing and improving the Machine learning algorithms to give betterment predictions of the stock prices. Although prediction of stock price varies on multiple (micro and macro) factors a rough estimation can be made with the above-mentioned techniques. In the proposed scheme, we have obtained results, like, R2 score was 0.95 on the test dataset and 0.92 on the train dataset. F1 score was 0.15 and accuracy was 96%.

REFERENCES

[1] W. Chen, X. Li, and D. D. Zeng, "Simple is beautiful: Toward light prediction markets," IEEE Intelligent Systems, vol. 30, no. 3, pp. 76–80, 2015.

[2] D. Yan, G. Zhou, X. Zhao, Y. Tian, and F. Yang, "Predicting stock using microblog moods," China Communications, vol. 13, no. 8, pp. 244–257, 2016.

[3] W. Zhang, C. Li, Y. Ye, W. Li, and E. W. Ngai, "Dynamic business network analysis for correlated stock price movement prediction," IEEE IntelligentSystems, vol. 30, no. 2, pp. 26–33, 2015.

[4] X. Li, H. Xie, Y. Song, S. Zhu, Q. Li, and F. L. Wang, "Does summarization help stock prediction? a news impact analysis," IEEE Intelligent Systems,vol. 30, no. 3, pp. 26–34, 2015.

[5] J. Bollen and H. Mao, "Twitter mood as a stock market predictor," Computer, vol. 44, no. 10, pp. 91–94, 2011.

[6] F. Z. Xing, E. Cambria, and R. E. Welsch, "Intelligent asset allocation via market sentiment views," IEEE Computational Intelligence Magazine, vol. 13,no. 4, pp. 25–34, 2018.

[7] J. Yin, R. Wang, S. Ju, Y. Bai, and J. Z. Huang, "An asymptotic statistical learning algorithm for prediction of key trading events," IEEE IntelligentSystems, vol. 35, no. 2, pp. 25–35, 2020.

[8] P. Somani, S. Talele, and S. Sawant, "Stock market prediction using hidden markov model," in 2014 IEEE 7th Joint International Information Technologyand Artificial Intelligence Conference, 2014, pp. 89–92.

[9] L. Zhao and L. Wang, "Price trend prediction of stock market using outlier data mining algorithm," in 2015 IEEE Fifth International Conference onBig Data and Cloud Computing, 2015, pp. 93–98.

[10] A. Sharma, D. Bhuriya, and U. Singh, "Survey of stock market prediction using machine learning approach," in 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), vol. 2, 2017, pp. 506–509.

[11] Satnam, "Bajaj Finance Dataset Since 2002 to July 2020," 2020. [Online]. Available: {https://www.kaggle.com/general/177027}