2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)
Amity University, Noida, India. Sep 3-4, 2021

2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) | 978-1-6654-1703-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICRITO51393.2021.9596513

Prediction in Stock Price Using of Python and Machine Learning

Anchal Vij [1] , Komal Saxena[2] Ajay Rana[3]
AIIT, Amity University Uttar Pradesh, Noida, India
vijaachal@gmail.com, ksaxena1@amity.edu, ajay_rana@amity.edu

***Abstract*-In the finance industry, the stock price prediction is one of the greatest notable implementations. Price forecast is a demonstration of trying to ascertain the future estimation of stock. This paper clarifies the forecast of a stock closing price utilizing Machine Learning. The dedicated and analytic or the time arrangement inspection is utilized by most of the stockbrokers to creating the stock price forecasting. The programming language is utilized to foresee the stock price prediction utilizing machine learning is Python. Machine learning itself utilizes various techniques to make estimation simpler and genuine. In this report, we recommend a Machine Learning (ML) approach that is destined to be prepared from the accessible stock's information and gain insight and afterward utilizes this gained information for a precise forecast of the stock values. In this set, this investigation utilizes a ML procedure called LSTM to envisage stock values. The objective of the model is to forecast the stock prices and finding out the instabilities in the stock value based on the last sixty days closing prices of the stock.**

Keywords: Finance, Stock price, Prediction, LSTM, Machine learning., Python

I. INTRODUCTION

Stock markets are one of the vital parts of the economy of a nation. As a matter of fact, it's the chief route for establishments to raise money. The financial specialists as and the common citizens are likewise discovering it as a venture apparatus. As the securities exchange impacts individual and public economy vigorously, foreseeing the future predictions of the financial exchange is an essential undertaking while at the same time making the right choice whether to purchase or sell the offer. Be that as it may, it was make an effort to predict the stock value patterns proficiently on the grounds that numerous elements, for example, financial aspects, governmental issues, climate and so on were choosing boundaries [6].

Financial time arrangement changes vigorously and specifically. Such time arrangement is clearly hard to anticipate for the reason that the issue is nonlinear, non-fixed and has a lot of noise [4]. Stock cost is a sort of time arrangement in the economic area. The way to foresee stock pattern, later on, has gotten quite possibly the most import issues by utilizing information on the basis of mining methods. Notwithstanding, the forecast is troublesome from the standard of the productive market speculation that if the marketplace is an effective market, at that point the stock cost will trail an arbitrary work design. Furthermore, a fixed expectation system is likewise unrealistic if the market is effective on the base of that financial specialists will before long find such methodologies and those fruitful estimating rules will prompt fall to pieces [7].

II. LITERATURE REVIEW

Stock price prediction is quite possibly the main subjects to be explored in scholarly and monetary investigates. In any case, a method utilizing machine learning techniques will give a more exact, precise and simpler approach to settle such issues identified with stock prices and market costs.

To make a perdition of stock price, several methods and techniques are being employed. The type of technique usually depends on the user and doesn't much affects the prediction. The most used technique as observed was SVM model. The use of SVM is all-inclusive as it joins pre-preparing of the monetary stock info set, use of different element designing methods, joined with a redid profound deep learning-based framework for stock value pattern forecast. Conducting far-reaching assessments regularly applying machine learning techniques and concluding the stock prices. The technique accomplishes in general high exactness for stock price pattern. With the itemized plan and assessment of forecast term lengths, highlight designing, and information pre-processing strategies, this work adds to the stock examination network mutually in the monetary and specialized fields [2-3].

Stock price prediction is of great value to financial specialists; nonetheless, the stock price is driven by unpredictable factors, for example, microblogs and news that make it challenging to anticipate prices on merely the verifiable information. The gigantic instability underlines the need to viably access the work of outside variables in the stock forecast. Securities exchanges can be anticipated utilizing AI calculations on data confined in online broadcasting and monetary news, as this information can transformation speculators' conduct. In the mentioned method, the calculations are made via online media and monetary news information to find the effect of this information on securities exchange expectation precision for ten resulting days. For refining execution and nature of forecasts, highlight choice and spam tweets decrease is

performed on the informational collections. Additionally, performing analyses to discover such securities exchanges that are hard to anticipate and which are more affected by web-based media and monetary news. We contrast the aftereffects of various calculations to locate a reliable classifier. At long last, for accomplishing greatest prediction exactness, deep learning is utilized and a few classifiers are ensembled [4].

Another technique is the use ANN and Rf Methods which we have been applied for foreseeing the following day shutting down figures of stock.

The related and existing work in this area use traditional procedures like linear regression, RWT, MACD and also using some linear models like ARMA, ARIMA, for predicting stock prices [5].

The expanding interest for a precise forecast model brought about broad exploration commitment in models. Therefore, there are different measurable and intelligent strategies existing in writing which are broadly utilized for the prediction that incorporates Fuzzy Logic, ARMA, GARCH, ANN. In the following strategies, ANN is the most fit strategy and discovered to be more proficient in dealing with nonlinearity; then again, the conventional ARMA and GARCH procedures do not have this ability and consequently can't yield huge forecasting exactness in a boisterous climate. As a widespread inexact ANN has been adopted effectively in different examination zones like value of stock will be determining, cash conversion scale gauging, power value anticipating and wind speed estimating, and so on. Further hybrid prediction procedures have been created by analysts with various mixes by using the speculation capacity of ANN that incorporates ANN-ANFIS, ARIMA based GARCH-ANN based ANN and has been applied for NSE Market of stock value determining, wind speed anticipating, gold value estimating, and so on. The FNN is notable in the current writing that has broad applications in various fields. Quite possibly the most famous neural - network called SLFN has extensive applications in relapse and characterization and utilizes notable training calculations like backpropagation (BP) and RLS for tuning its boundaries. The issue lies with the BP calculation where it could be caught in local minima and may knowledge the ill effects of moderate combination. In correlation, the RLS calculation is quicker than BP yet includes more computational multifaceted nature. In these calculations, the obscure boundaries are tuned iteratively in the learning cycle [13][14][15]. These relevant articles have discussed about different Models in different Research Papers Arch Model [16]in this GARCH [17-18] autoregressive Conditional. ARMA Model (Moving Average Model [19,20] Arima Model [21] VAM Model [22]in this they used ANN Model [23].

III. THE PROPOSED SYSTEM

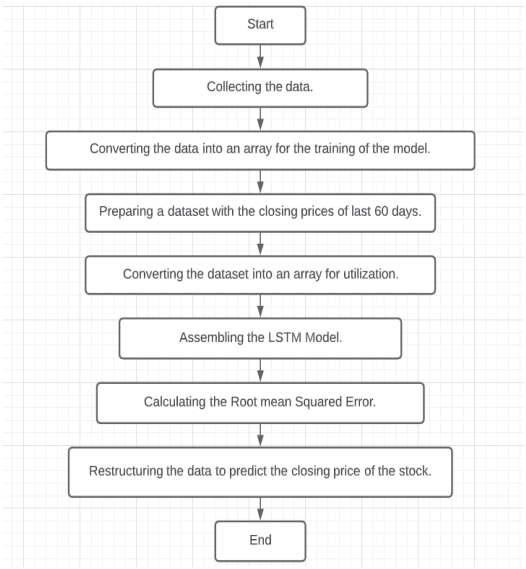

Fig. 1. Proposed system

A. Data set Collected: -

1) First we need to import all the libraries which are essential for the training of the model. Collect the stock statement for the organization 'Alphabet Inc.' utilizing the organizations ticker (GOOGL) from 1st January 2012 to May 10th, 2021 from X Finance-Yahoo.

(url: https//in.finance.yahoo.com/quote/GOOGL?p=GOO GLE &.tsrc=fin-srch).

The Dataset is comprised of total is 2261 rows * 6 columns.

a) ***Trained and Split the Data: -*** Make another information outline with just the end cost and convert it to an array. At this point, we state a variable to accumulation the measurement of the training informational data. We need the preparation informational collection to contain an amount of the information. Make a training informational data collection that encompasses the previous sixty-day closing value that we need to use to anticipate the sixty-one -day closing value of the stock. So, the principal segment in the 'x/_train' informational collection will encompass values from the informational collection from index (0-59) (sixty values complete) and the subsequent section will comprise values from the informational from index 1 to index (sixty values) and more of the same. The 'y/_train' informational collection will contain the sixty-first value situated at index sixty for its the main section and the sixty-second value situated at index 61 of the informational collection for its subsequent value and it goes along in the same manner.

b) ***Convert the Autonomous Data & Prepare the LSTM Model :-*** Then we convert the autonomous train informational collection 'x_train' and subordinate informational index 'y_train' to numpy clusters so they can

```
x_train, y_train = np.array(x_train), np.arr
```

be utilized for preparing the LSTM model. Restructure the information to be three dimensional have taken in the structure [No. of tests, No. of time steps, and number of features]. The LSTM model is expecting a three-dimensional informational set. x_train = np.reshape (x_train, (x_train.shape[0],x_train.shape[1],1))Accumulate the LSTM model to have 2-LSTM layers with Fifty neurons and 2 Dense layers, one with Twenty -five neurons and the other with one neuron. Organize the model exploiting the mean squared error (MSE) Adam optimizer. and the loss function

c) ***Train the Model Utilizing Training Informational Sets***. Batch size is indisputably the quantity of prepared models present in a solitary gathering and epoch is the quantity of redundancies when an entire enlightening set is passed forward and in invert through the neural association.

```
1018/1018 [==============================] - 23s 20ms/step - loss: 0.0018
```

Fig. 2. Showing the loss of data in training.

IV. Results and Discussions

A. LSTM Model should be trained and tested

Testing LSTM Mode Set: Make a test informational set. At this point modification the free test informational set 'x_test' to a numpy cluster thus it tends to be utilized for testing the LSTM model. Restructure the information to be three dimensional in the structure [No. of tests, No. of time steps, and No of features]. This should be done on the grounds that the LSTM model is expecting a 3-dimensional informational collection. Now get the anticipated figures from the model utilizing the test information.

B. RMSE Calculated to extract the Model

Calculate (RMSE) which means, that is a decent degree of how Perfect the model is. An estimation of would demonstrate that this type of models anticipated prices coordinate the real prices from the test informational set correctly. It reduced the figure the improved the model will be performing. Be that as it may, as a rule, it is ideal to utilize different measurements also to really get the knowledge of how well the model is performing.

The formula applied is root mean square error which is equal to the square root of the trained values into the predictions subtracted by the y dependent data multiplied by two.

```
57.96680839371415
```

Fig. 3. Shows the value calculated of root mean square error.

Predicted price is calculated by using the scaler inverse values.

```
[[2281.2083]]
```

Fig. 4. Shows the value projected by the model.

C. Predication of the Model to Provide the estimated Cost

We need to test the model some more and get the expected shutting esteem assessment of Alphabet Inc. for May 10, 20201(10/05/2021). Thus, we should get the assertion, alter the data to a cluster that comprises only the end cost. Now, we will get the latest sixty-day shutting esteems and scale the information to be assessed some place in the scope of 0 and 1 comprehensive. From here on out, we will make an unfilled rundown and add the past sixty-day esteems to it, and subsequently alter it to a NumPy exhibit and rebuild it so we can incorporate the figures into the model. To wrap things up, we will enter the information into the model and get the anticipated price.

Fig. 5. Shows the real stock price

```
Date
2021-05-10    2291.75
Name: Close, dtype: float64
```

The dataset collected shown below in fig.6

Date	High	Low	Open	Close	Volume	Adj Close
2016-01-04	762.200012	747.539978	762.200012	759.440002	3369100	759.440002
2016-01-05	769.200012	755.650024	764.099976	761.530029	2260800	761.530029
2016-01-06	765.729980	748.000000	750.369995	759.330017	2410300	759.330017
2016-01-07	755.309998	735.280029	746.489990	741.000000	3156600	741.000000
2016-01-08	750.119995	728.919983	747.799988	730.909973	2375300	730.909973
...	...	...	...	...	...	...
2021-05-04	2324.989990	2256.679932	2313.550049	2306.830078	2240900	2306.830078
2021-05-05	2335.000000	2308.280029	2328.629883	2314.770020	1331800	2314.770020
2021-05-06	2337.350098	2293.040039	2306.330078	2337.350098	1259500	2337.350098
2021-05-07	2371.250000	2346.159912	2363.889893	2351.929932	1444500	2351.929932
2021-05-10	2331.000000	2284.540039	2328.139893	2291.750000	1544900	2291.750000

1347 rows × 6 columns

Fig. 6. The dataset collected.

Fig. 7. Graph showing the information collected.

```
[array([0.04574613, 0.04696721, 0.04568187, 0.03497271, 0.02907769,
       0.03033967, 0.03750833, 0.02245241, 0.02935815, 0.01714748,
       0.02216614, 0.02186232, 0.02660051, 0.03757844, 0.030661  ,
       0.03076031, 0.02128977, 0.03923767, 0.04686201, 0.0523656 ,
       0.05828978, 0.03986866, 0.02856359, 0.01321555, 0.01344922,
       0.01161473, 0.01502083, 0.01473456, 0.01504423, 0.02132483,
       0.02969699, 0.02124887, 0.02393637, 0.02799101, 0.02112032,
       0.02322946, 0.02803191, 0.02554304, 0.02107942, 0.03565626,
       0.03408465, 0.02947501, 0.02867456, 0.01849708, 0.0189236 ,
       0.02586436, 0.02981384, 0.03723372, 0.0403711 , 0.04056391,
       0.0445309 , 0.04518524, 0.04339162, 0.04733525, 0.04610251,
       0.04464775, 0.04305863, 0.04214721, 0.0495145 , 0.05094589])]
[0.04776761781170197]
```

Fig. 8. Training Dataset in form of array.

Fig. 9. The actual data as compared to the values predicted by the model.

Date	Close	Predictions
2020-04-16	1257.430054	1200.463379
2020-04-17	1279.000000	1214.906494
2020-04-20	1261.150024	1229.016602
2020-04-21	1212.160034	1238.780029
2020-04-22	1258.410034	1238.742065
...	...	...
2021-05-04	2306.830078	2294.892090
2021-05-05	2314.770020	2293.347168
2021-05-06	2337.350098	2288.900879
2021-05-07	2351.929932	2286.801025
2021-05-10	2291.750000	2287.978027

269 rows × 2 columns

Fig. 10. Shows all the valid predicted prices.

V. CONCLUSION

We have used LSTM method to Prepare the model and forecast the stock value for google Inc in this research article. The proposed model produces the highest level of performance with relation to the root Least squared error. To verify the effectiveness of the model we have used more than thousand Alphabet stock.

REFERENCES

[1] Kranthi.V Stock Market Prediction Using Machine Learning. 10.13140/RG.2.2.12300.77448. (2018).

[2] Chouhan, Lokesh & Agarwal, Navanshu & Parmar, Ishita & Saxena, Sheirsh & Arora, Ridam & Gupta, Shikhin & Dhiman, Himanshu. Stock Market Prediction Using Machine Learning. 10.1109/ICSCCC.2018.8703332. (2018).

[3] Shen.J. Shafiq, M.O. Short-term stock market price trend prediction using a comprehensive deep learning system. J Big Data 7, 66 (2020). https://doi.org/10.1186/s40537-020-00333-6.

[4] Khan, W., Ghazanfar, M.A., Azam, M.A. et al. "Stock market prediction using machine learning classifiers and social media, news. J Ambient Intell Human Compute". (2020). https://doi.org/10.1007/s12652-020- 01839-w.

[5] Vijh .M, Chandola. D, Tikkiwal V.Anand, Kumar. A, Stock Closing Price Prediction using Machine Learning Techniques, Procedia Computer Science, Volume 167, 2020, Pages 599-sixty6, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2020.03.326.

[6] Billah M., Waheed S. & Hanifa A. "Stock market prediction using an improved training algorithm of neural network," 2016 2nd International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), Rajshahi, 2016, pp. 1-4, doi: 10.1109/ICECTE.2016.7879611.

[7] Zhao. L. & L. Wang, "Price Trend Prediction of Stock Market Using Outlier Data Mining Algorithm," 2015 IEEE Fifth International Conference on Big Data and Cloud Computing, Dalian, 2015, pp. 93-98, doi: 10.1109/BDCloud.2015.19.

[8] Jaeheon Chun, Jaejoon Ahn, Youngmin Kim & Sukjun Lee (2020) Using Deep Learning to Develop a Stock Price Prediction Model Based on Individual Investor Emotions, Journal ofBehavioral Finance, DOI: 10.1080/154275sixty.2020.1821686.

[9] Sharma.N and Juneja A. "Combining of random forest estimates using LSboost for stock market index prediction," 2017 2nd International Conference for Convergence in Technology (I2CT), Mumbai, 2017, pp. 1199-1202, doi: 10.1109/I2CT.2017.8226316.

[10] Wang. Zhaoxia, Seng-Beng Ho, Zhiping Lin, "Stock Market Prediction Analysis by Incorporating Social and News Opinion and Sentiment", Data Mining Workshops (ICDMW) 2018 IEEE International Conference on, pp. 1375-1380, 2018.

[11] Engle. R. F "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation," Econometrical, vol. 50, no. 4, pp. 987–1008, 1982.

[12] Bollerslev.T "Generalized autoregressive conditional heteroskedasticity," Journal of Econometrics, vol. 31, no. 3, pp. 307–327, 1986.

[13] Girish G. P, "Spot electricity price forecasting in Indian electricity market using autoregressive-GARCH models", Energy Strategy Reviews, vol. 11-12, pp. 52–57, 2016.

[14] P. Burlando, R. Rosso, L. G. Cadavid, and J. D. Salas, "Forecasting of short-term rainfall using ARMA models," Journal of Hydrology, vol. 144, no. 1-4, pp. 193– 211, 1993.

[15] Contreras.J, Espínola. R, Nogales.V, Conejo. A.J, "ARIMA models to predict next-day electricity prices", IEEE Transactions on Power Systems, vol. 18, no. 3, pp. 1014–1020, 2003.

[16] Geurts. M, G. E. P. Box,G. M. Jenkins, "Book review: time series analysis: forecasting and control", Journal of Marketing Research, vol. 14, no. 2, p. 269, 1977.

[17] P.F. Pai, Lin C.S, "A hybrid ARIMA and support vector machines model in stock price forecasting", Omega, vol. 33, no. 6, pp. 497–505, 2005.

[18] Wang. J.Z. Wang J. J, Z Zhang G, Guo S.P. "Forecasting stock indices with back propagation neural network", Expert Systems with Applications, vol. 8, no. 11, pp. 14346–14355, 2011.

[19] C.T Lin. Yeh H.Y, "Empirical of the Taiwan stock index option price forecasting model - applied artificial neural network", Applied Economics, vol. 41, no. 15, pp. 1965–1972, 2009.

[20] K.j, Kim, I. Han, "Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index", Expert Systems with Applications, vol. 19, no. 2, pp. 125–132, 2000.

[21] S. Manu, Y, D Chen, Sohn, K. Shimada, K. Hirasawa, "Stock price prediction using neural networks with RasID- GA," IEEJ Transactions on Electrical and Electronic Engineering", vol. 4, no. 3, pp. 392–403, 2009.

[22] H. Ince and T. B. Trafalis, "Short term forecasting with support vector machines and application to stock price prediction", International Journal of General Systems, vol. 37, no. 6, pp. 677–687, 2008.

[23] Sai Y, Zhang F. T, Zhang T. "Research of Chinese stock index futures regression prediction based on support vector machines", Chinese Journal of Management Science, vol. 3, pp. 35–39, 2013.

[24] Liu.Y "Novel volatility forecasting using deep learning-long short-term memory recurrent neural networks", Expert Systems with Applications, vol. 132, pp. 99–109, 2019.

[25] Y. Baek & H. Y. Kim, Mod Aug Net: "A new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module", Expert Systems with Applications, vol. 113, pp. 457–480, 2018.