International Journal of Engineering and Applied Physics (IJEAP)
Vol. 1, No. 2, May 2021, pp. 162~167
ISSN: 2737-8071

Analysis of Stock Market using Machine Learning

Riteesh Mandi[1], Skanda G[2]
[1] Dept. of Computer Science and Engineering, Dr. Ambedkar Institute of Technology, Bangalore, India.
[2] Dept. of Electronic and Communication Engineering, Dr. Ambedkar Institute of Technology, Bangalore, India.

Article Info

Article history:

Received Nov 13, 2020
Revised Mar 22, 2021
Accepted Mar 25, 2021

Keywords:

Machine Learning
Deep Learning
Stock Market
RNN
LSTM

ABSTRACT

Machine Learning is a prominent area of research that emphasizes on finding patterns in existential data. The field of Machine Learning, can be concisely described as enabling computers to make productive predictions using previous experiences. As there is a large amount of information being available everywhere, it is very important to analyze this data in order to extract some useful information and thus developing algorithms based on this analysis. This can hence be done through data mining and Machine Learning. In addition to many other fields, Machine Learning models have broad applications in the field Bioinformatics. The complexity involved in biological analysis has led to the development of experienced Machine Learning methods. This research paper discusses the importance of a data-driven approach, compared to the formalization of traditional Artificial Intelligence and also primarily focuses on a key approach to forecast company's workflow using Machine learning.

Corresponding Author:

Riteesh M Mandi,
Student,
Dr. Ambedkar Institute of Technology,
Bangalore, Karnataka, India.
Email: riteeshmandi2000@gmail.com

1. INTRODUCTION

Nowadays, Investment firms, hedge funds and also most of the people are using a wide range of financial models and figurine statistics to better perceive market behavior and to create profitable investments and trades. An enormous spectrum information of knowledge is on the market within the sort of previous historical stock costs and company's wise performance data is used with algorithms involving Machine Learning models to and analyze the investments. Investors create educated guesses by analyzing information. Investors usually read the news, study the corporate history, trade trends and different millions of information points to predict future stock market move. Since stock costs square measures are entirely random and unpredictable, it is difficult to predict. So basically, this paper focuses on building a Machine Learning model which can be used to predict and analyze stock market trends, company shares, profits or losses relating to it. In this analysis Neural Network algorithm, Long-Short Term Memory (LSTM), Deep Learning models are used to predict stock prices. Long Short Term Memory (LSTM) architecture is a unique kind of Recurrent Neutral Network (RNN) which was introduced by Hochreiter and Schmidhuber to overcome the weakness of the traditional RNN to learn long term dependencies [1-3].

Long Short-Term Memory (LSTM) networks are peculiar type of Recurrent Neural Network which is efficient enough to learn order dependence in sequence prediction problems. This is type of networks are

usually used in complex problem domain related fields like machine translation, speech recognition, and more. LSTM's are a complex area of deep learning, since it can be quite complex to understand and implement concepts around what LSTMs are, and how labels such as bidirectional and sequence-to-sequence relate to the field [4, 5]. This research paper explains the implementation of Keras to create a LSTM model to predict stock costs mistreatment historical value price terms such as damage and mercantilism volume, and to do complete analysis of data and visualize each the expected price values over time and therefore the best parameters for the model [6-8].

2. RESEARCH METHOD

Data

This research paper uses Tesla Inc. stock historical data, which is imported from investing.com in csv (comma separated file) format wherein the data consists of Date, Price, Open, High, Low, Volume and Percentage change from open price to low.

Metrics

In this research paper performance measure is calculated by the implementation of Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) which can be finally calculated as the difference between predicted and actual merits or values of the target stock which is at adjusted Close price and the delta between the performance of the benchmark model is been implemented using Linear Regression and the Deep Learning models.

Data Exploration

The analysis used in this research paper is in consideration with Tesla Inc. from June 30, 2010 to April 28, 2021 and data is arranged in terms of time series. Goal of this research is to accurately predict the final price (close price) for any given date after training the model. All sample statistics are taken from the Tesla Inc. historical data in csv format as given in Table 1. The prediction was to be made for the final or the Closing (Adjusted close) price of the data. Given that Tesla readily regulates the closing prices, for us forecast is made based on close price.

Table 1: Historical Stock Price Data of Tesla Inc.

Date	Price (in USD)	Open Price (in USD)	High Price (in USD)	Low Price (in USD)	Volume (in Million)	Change (in %)
22-Apr-21	719.69	742.00	753.35	718.10	35.59	- 3.28%
23-Apr-21	729.40	719.80	737.36	715.46	28.41	1.35%
26-Apr-21	738.20	740.88	748.99	732.79	31.04	1.21%
27-Apr-21	704.74	718.43	724.00	703.46	29.44	- 4.53%
28-Apr-21	694.40	696.43	708.05	693.84	22.27	- 1.47%

Exploratory Data Visualization

This Tesla company stock has been visualized using matplotlib library, the graph is being plotted against closing stock price of the data with the number of days available. So that it can be analyzed how the closing prices of stock varies with number of frames time gap. Figure 1 indicates that there is linear growth in Tesla Inc. stock price for around six years and the main fall of stock prices were between 500 USD to 800 USD in the year 2019 to 2021 due to covid-19 pandemic.

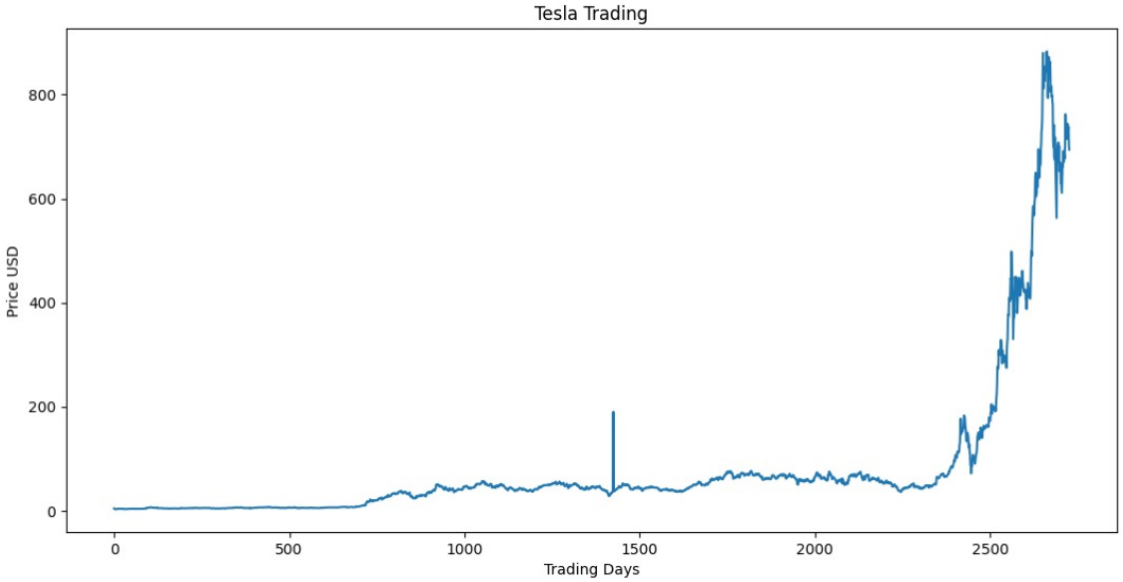

Figure 1: Closing Price of Stock vs Number of Days

In the next section, the implementation of predictive analysis using LSTM method is discussed.

Implementation

There are four steps that clearly describe the predictive analysis using LSTM. Figure 2 explains the implementation of these predictive analysis steps using LSTM.

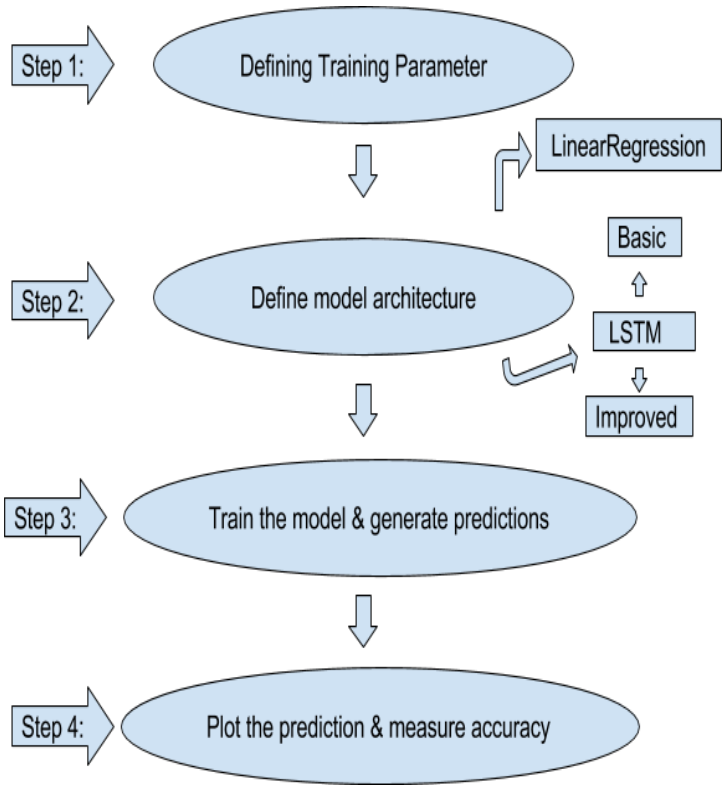

Figure 2: Flowchart of Implementation of LSTM Model

Based on the flow chart from Figure 2, the steps of implementation is as follows:

Step 1. Defining Training Parameters: Here the training parameters and the specifications are defined for which the model needs to be trained. The framework is determined and listed within which the model needs to function in all of its scope.

Step 2. Define Model Architecture**:** Here the set of methods are described functionality, also organization of the mentioned model is defined and implemented. Since Linear Regression and LSTM (Long term short memory) model is used, it is also defined while building the entire model architecture.

Step 3. Train the model and generate predictions: Once all the rules and methods are defined, next step is to train the model based on these methods for various conditions and look out for the expected predictions or the values.

Step 4. Plot the predictions and measure accuracy: After obtaining the results these outcomes have to be plotted graphically and all the coordinates can be analyzed accordingly based on the fed data. Finally, this graph can be used to measure accuracy as well as the precision of the trained model

3. RESULTS AND DISCUSSIONS

As explained in the previous section, the Bench mark model considered is Linear Regression (LR) method and is compared with the basic LSTM and improved LSTM methods.

Benchmark Model

Here the simple regression methodology is employed as primary Benchmark Model. Since the goal is to predict the relative performance and implementation variations of Machine Learning (ML) and deep learning models [17-19]. This simple regression relies on Machine Learning model and was used as error rate comparison between MSE and RMSE utilizing an equivalent dataset, from Figure 3 it is seen that X-axis represents Trading Days, Y-axis represents Closing price in USD, Adjusted Close price is indicated by green line, and Predicted Close price is being indicated by blue line. For the given train score as 0.5117 MSE (0.7153 RMSE), final test score is 0.69205044 MSE (0.83189569 RMSE).

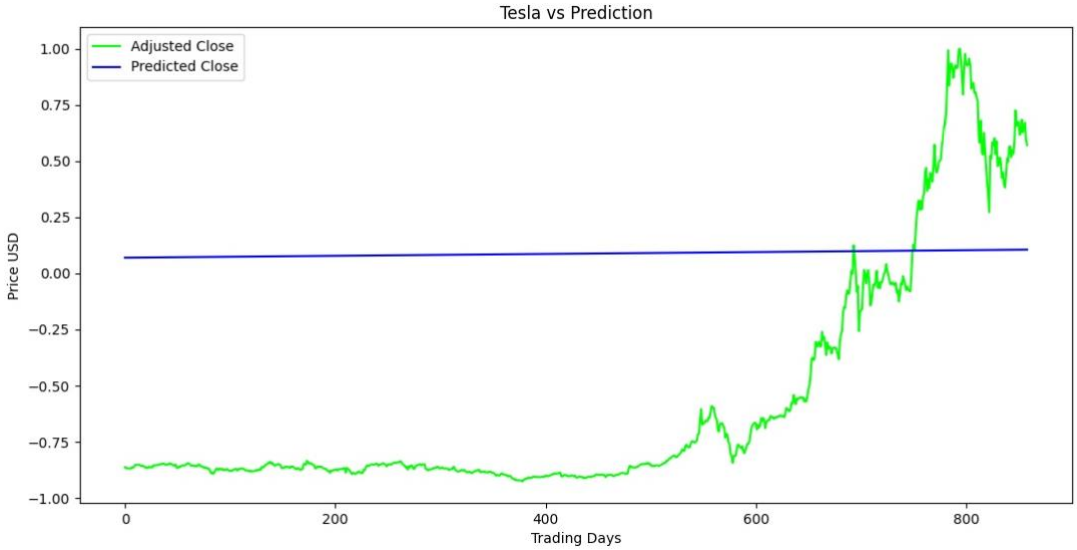

Figure 3: Benchmark Model using Linear Regression

Refinement

Next the fine standardization parameters of LSTM are employed to urge higher predictions. The exact prediction is done by testing and analyzing every parameter so that, the ultimate price for every of them is chosen.

To improve LSTM model the subsequent has been done:

1. The amount of hidden node is exaggerated from a 100 to 128.
2. Dropout of 0.2 is being additional at every layer of LSTM.
3. Batch size is being exaggerated from one to 512.
4. Epochs is exaggerated from one to twenty.
5. Windy is being additional that is capable two.
6. Correct prediction is created supported batch size.

The predicted close price and adjusted close price plot difference for basic LSTM model is shown in Figure 4.

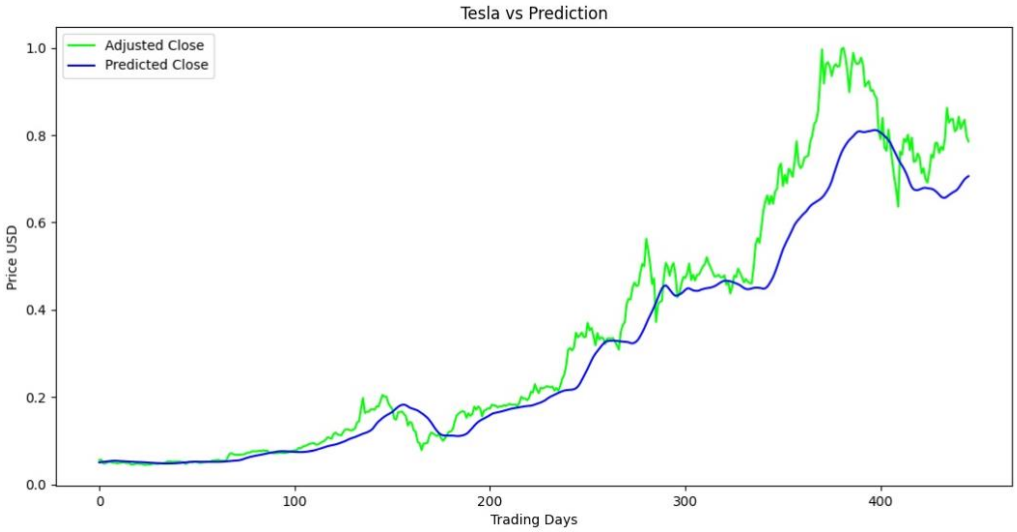

Figure 4: Plot For Adjusted Close and Predicted Close Prices for basic LSTM model

The predicted close price and adjusted close price plot difference for improved LSTM model is shown in Figure 5.

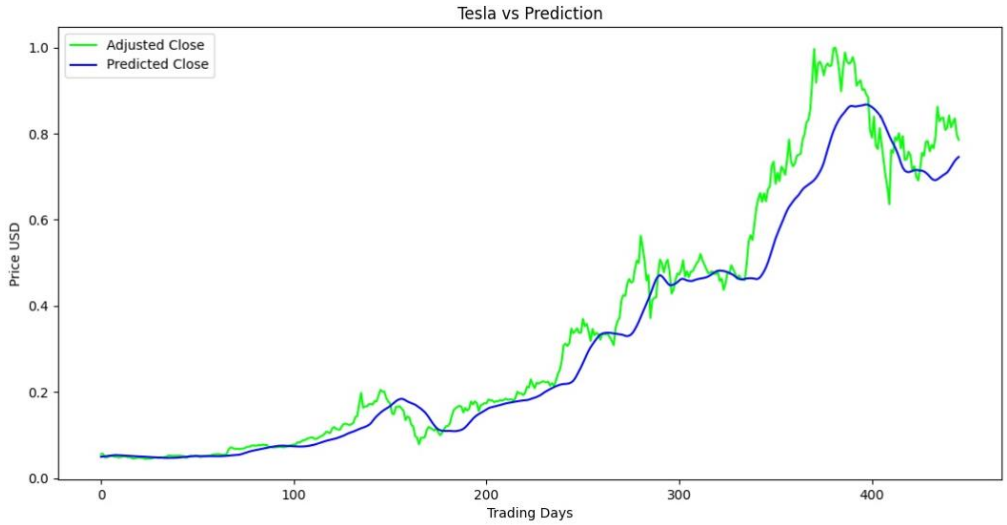

Figure 5: Plot for Adjusted Close and Predicted Close Prices for improved LSTM model

As seen from Figure 4 to Figure 5 the LSTM model is improved my Mean Squared Error (MSE) for testing set, from 0.00690415 MSE to 0.005017 MSE.

4. CONCLUSION

Comparing the Benchmark model or the regression technique to the ultimate improved LSTM model, the Mean Square Error (MSE) is improved from 0.00690415 MSE to 0.005017 MSE. This important decrease in error rate clearly shows that final model have surpassed the fundamental and benchmark model.

REFERENCES

[1] S. O. Ojo, P. A. Owolawi, M. Mphahlele and J. A. Adisa, "Stock Market Behaviour Prediction using Stacked LSTM Networks*," 2019 International Multidisciplinary Information Technology and Engineering Conference (IMITEC), 2019, pp. 1-5, doi: 10.1109/IMITEC45504.2019.9015840.

[2] S. Liu, G. Liao and Y. Ding, "Stock transaction prediction modeling and analysis based on LSTM," 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), 2018, pp. 2787-2790, doi: 10.1109/ICIEA.2018.8398183.

[3] K. Chen, Y. Zhou and F. Dai, "A LSTM-based method for stock returns prediction: A case study of China stock market," 2015 IEEE International Conference on Big Data (Big Data), 2015, pp. 2823-2824, doi: 10.1109/BigData.2015.7364089.

[4] W. Jhang, S. Gao, C. Wang and M. Hsieh, "Share Price Trend Prediction Using Attention with LSTM Structure," 2019 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD), 2019, pp. 208-211, doi: 10.1109/SNPD.2019.8935806.

[5] M. Faraz, H. Khaloozadeh and M. Abbasi, "Stock Market Prediction-by-Prediction Based on Autoencoder Long Short-Term Memory Networks," 2020 28th Iranian Conference on Electrical Engineering (ICEE), 2020, pp. 1-5, doi: 10.1109/ICEE50131.2020.9261055.

[6] Y. Zeng and X. Liu, "A-Stock Price Fluctuation Forecast Model Based on LSTM," 2018 14th International Conference on Semantics, Knowledge and Grids (SKG), 2018, pp. 261-264, doi: 10.1109/SKG.2018.00044.

[7] K. A. Althelaya, E. M. El-Alfy and S. Mohammed, "Evaluation of bidirectional LSTM for short-and long-term stock market prediction," 2018 9th International Conference on Information and Communication Systems (ICICS), 2018, pp. 151-156, doi: 10.1109/IACS.2018.8355458.

[8] S. Kumar and D. Ningombam, "Short-Term Forecasting of Stock Prices Using Long Short Term Memory," 2018 International Conference on Information Technology (ICIT), 2018, pp. 182-186, doi: 10.1109/ICIT.2018.00046.