ISSN: 2581-3404 (Online) IF: 5.445 (SJIF)
IJIRTM, Volume-7, Issue-3, May-2023.

Stock Price Forecast Using Long Short Term Memory (LSTM) Algorithm

Kunal Panthi[1], Dr. Vineet Richhariya[2], Dr. Sadhana K. Mishra[3]
M.Tech. Research Scholar[1], Professor[2], Prof. & Head[3]
Dept. of CSE, LNCT, Bhopal[1,2,3]

Abstract: *One of the most significant activities in the world of finance is stock trading. Trying to anticipate the future value of a stock or other financial instrument traded on a financial exchange is known as stock market prediction. This describes how machine learning was used to predict a stock. The majority of stockbrokers employ technical, fundamental, or time series analysis when making stock predictions. In this article, we suggest a machine learning (ML) strategy that will be trained using the stock market data that is currently available to gather intelligence before using the learned information to make an accurate prediction. We used ML algorithms like LR and LSTM in these, and the summary of model performance parameters shows that LSTM outperforms LR Model for the datasets.*

Keywords: Stock Market, Machine Learning, Predictions, Classification, forecasting, Data mining, Google Stock Forecasting, Logistic Regression (LR), LSTM.

Introduction

Investment firms, hedge funds and even individuals have been using financial models to better understand market behavior and make profitable investments and trades. A wealth of information is available in the form of historical stock prices and company performance data, suitable for machine learning algorithms to process. Can we actually predict stock prices with machine learning? Investors make educated guesses by analyzing data. They'll read the news, study the company history, industry trends and other lots of data points that go into making a prediction. The prevailing theories is that stock prices are totally random and unpredictable but that raises the question why top firms like Morgan Stanley and Citigroup hire quantitative analysts to build predictive models. We have this idea of a trading floor being filled with adrenaline infuse men with loose ties running around yelling something into a phone but these days they're more likely to see rows of machine learning experts quietly sitting in front of computer screens. In fact about 70% of all orders on Wall Street are now placed by software, we're now living in the age of the algorithm.
Stock market prediction is basically defined as trying to determine the stock value and offer a robust idea for the people to know and predict the market and the stock prices. It is generally presented using the quarterly financial ratio using the dataset. Thus, relying on a single dataset may not be sufficient for the prediction and can give a result which is inaccurate. Hence, we are contemplating towards the study of machine learning with various datasets integration to predict the market and the stock trends.

The problem with estimating the stock price will remain a problem if a better stock market prediction algorithm is not proposed. Predicting how the stock market will perform is quite difficult. The movement in the stock market is usually determined by the sentiments of thousands of investors. Stock market prediction, calls for an ability to predict the effect of recent events on the investors. These events can be political events like a statement by a political leader, a piece of news on scam etc. It can also be an international event like sharp movements in currencies and commodity etc. All these events affect the corporate earnings, which in turn affects the sentiment of investors. It is beyond the scope of almost all investors to correctly and consistently predict these hyperparameters. All these factors make stock price prediction very difficult. Once the right data is collected, it then can be used to train a machine and to generate a predictive result.

Machine Learning

When a computer needs to perform a certain task, a programmer' s solution is to write a computer program that performs the task. A computer program is a piece of code that instructs the computer which actions to take in order to perform the task. The field of machine learning is concerned with the higher-level question of how to construct computer programs that automatically learn with experience. A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E. Thus, machine learning algorithms automatically extract knowledge from machine readable information. In machine learning, computer algorithms (learners) attempt to automatically distill knowledge from example data. This knowledge can be used to make predictions about novel data in the future and to provide insight into the nature of the target concepts. Applied to the research at hand, this means that a computer would learn to classify alerts into incidents and non-incidents (task T). A possible performance measure (P) for this task would be the accuracy with which the machine learning program classifies the instances correctly. The training experiences (E) could be labelled instances.

Benefits of Machine Learning

In particular, machine learning plays an essential role in the following three areas of software engineering:

1. Data mining problems where large databases may contain valuable implicit regularities that can be discovered automatically.

2. Difficult-to-program applications, which are too difficult for traditional manual programming.

3. Software applications that customize to the individual user' s preferences, such as personalized advertising.

There are several reasons why machine-learning plays a role in these three domains. First of all, for the classification of security incidents, a vast amount of data has to be analyzed containing historical data. It is difficult for human beings to find a pattern in such an enormous amount of data. Machine-learning, however, seems well-suited to overcome this problem and can therefore be used to discover those patterns. With respect to the difficult-to-program applications, an analyst's knowledge is often implicit, and the environments are dynamic.

II. Literature Review

According to [1] In Stock Market is the financial epitome of financial business and trading since it came into existence it has shown the impact of hits low and similarly when it is high. The stock market crash in 2008 showed the world that the business hit the low when the Dow Jones Industrial Average fell 777.68%. Several machine learning algorithms have shown that these stock prices can be predicted and these algorithms can be implemented using the approach of supervised learning. In Supervised Learning, we have test data using this we train the models. Although the results obtained after training the model may differ from the actual but it has been observed that in many cases accuracy is satisfactory. In this paper, the first task is to use web scrapping to collect datasets from stock data. Then we plot the data on the graph, from the graph we can analyze the stock prices going high or low. After this, we will predict stock prices using SVM and Linear Regression, that Linear Regression for stock market analysis is better than the SVM for the same.

According to [2], Stock value examination to a great extent relies upon the capacity to recognize the development of the stock costs and foresee the concealed examples and patterns which the market follows. Information examinations have been developing significance on the financial exchange in the ongoing years. To get the benefit of the contributing, numerous financial specialists need to realize how to examine the significant information from the securities exchange. In a lot of general writing on stock foresee, it a couple of explicit direction show up on the future forecast. Thusly, how to anticipate the stocks from the recovery information, it turns into a significant and impressive issue on market foresees. Share market is one of the most impulsive and spot of high premium on the planet. There are no critical techniques exist to foresee the stock cost. Primarily individuals utilize three different ways, for example, major examination, measurable investigation and Machine Learning to foresee the stock cost of offer market yet none of these strategies are demonstrated as a reliably adequate forecast device. Thus, building up a forecast apparatus is one of the difficult undertakings as stock cost relies upon numerous compelling elements and highlights. we propose a strong technique to anticipate the offer rate utilizing Moving average based model and contrast how it vary and the genuine cost. For that we gather the share market information of most recent a half year of 5years of various classes, diminish their high dimensionality so it will have the option to prepare quicker and productively and make a similar investigation and our strategy for forecast of following day share cost.To legitimize the adequacy of the framework, diverse test information of companies' stock are utilized to confirm the framework These works show that information mining strategies can be applied for assessment of past stock costs and gain significant data by assessing appropriate monetary the request for the best Moving Average model was discovered to be Further, endeavors were made to figure, as exact as could be normal considering the present situation. Data mining systems can be applied on throughout a wide range of time money related data to make models and further calculations.

According to [3], Stock price prediction is one of the most extensively studied and challenging glitches, which is acting so many academicians and industries experts from many fields comprising of economics, and business, arithmetic, and computational science. Predicting the stock market is not a simple task, mainly as a magnitude of the close to random-walk behavior of a stock time series. Millions of people across the globe are investing in stock market daily. A good stock price prediction model will help investors, management and decision makers in making correct and effective Decisions. In this paper, we review studies on supervised machine learning models in stock market predictions. The study discussed how supervised machine learning techniques are applied to improve accuracy of stock market predictions. Support Vector Machine (SVM) was found to be the most frequently used technique for stock price

prediction due to its good performance and accuracy. Other techniques like Artificial Neural Network (ANN), K-Nearest Neighbor (KNN), Naïve Bayes, Random Forest, Linear Regression and Support Vector Regression (SVR) also showed a promising prediction result.

III. Problem Definition

Over the last two decades, humans have grown a lot of dependence on data and information in society and with this advent growth, technologies have evolved for their storage, analysis and processing on a huge scale. The fields of Data Mining and Machine Learning have not only exploited them for knowledge and discovery but also to explore certain hidden patterns and concepts which led to the prediction of future events, not easy to obtain. And one of the difficult things to predict that caught our attention is stock or commonly called as shares.

Stock price prediction is one of the most important topics to be investigated in academic and financial researches. Various Data mining techniques are frequently involved in the studies. To solve this problem. But technique using machine learning/deep learning will give more accurate, precise and simple way to solve such issues related to stock and market prices. There are ample amounts of data about stocks but the most difficult and intriguing thing is to predict the price of these stocks based on old data.

IV. Proposed Work

In the finance world stock trading is one of the most important activities. Stock market prediction is an act of trying to determine the future value of a stock other financial instrument traded on a financial exchange. This explains the prediction of a stock using Machine Learning. The technical and fundamental or the time series analysis is used by the most of the stockbrokers while making the stock predictions. In this paper we propose a Machine Learning (ML) approach that will be trained from the available stocks data and gain intelligence and then uses the acquired knowledge for an accurate prediction.

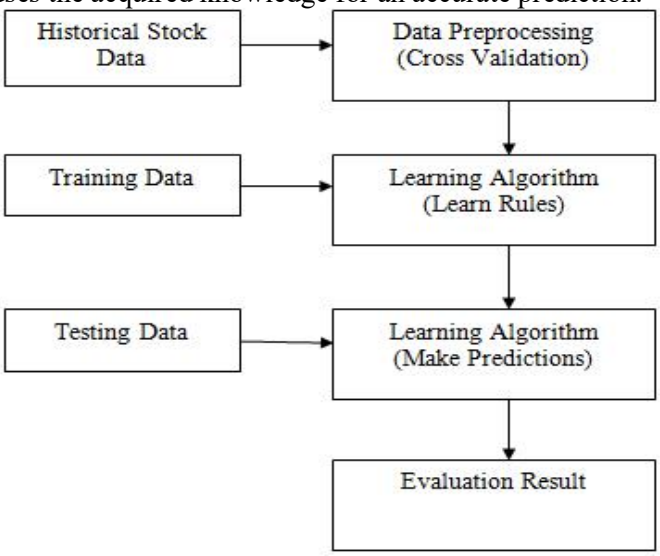


Figure 1: Flow Diagram.

1. Dataset:- we first collect the historical stock data set from publicly available datasets.
2. Data preprocessing: It is the most important phase in prediction models as the data consists of ambiguities, errors, redundancy which needs to be cleaned beforehand. The data gathered from multiple sources first is aggregated and then cleaned as the complete data collected is not suitable for modeling purposes. The records with unique values do not have any significance as they do not contribute much in predictive modeling. Fields with too many null values also need to be discarded.
3. Learn ML Model: After preprocessing we can learn algorithm on training data and make prediction on test dataset.
4. Obtain Result: Based on Prediction we can calculate the performance of the algorithm.

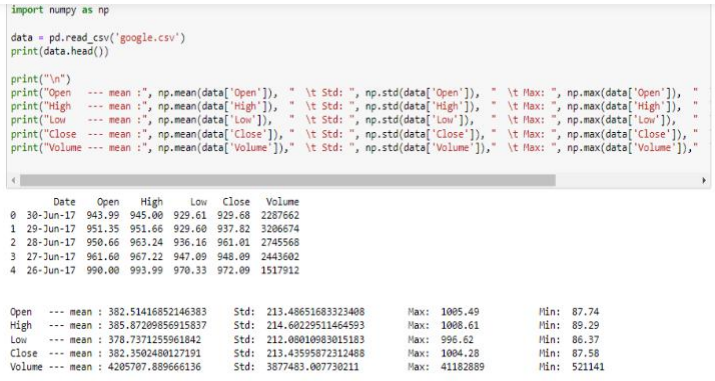
V. Experimental & Result Analysis

For this we proposed the ML algorithm whih use a LSTM Neural Net model able of learning from time series data. This will be programmed in a Jupyter Notebook(iPython) for ease of reproducibility. Using a Keras perpetration of the Tensor Flow library, the result will use a LSTM Neural Net model and will be supported by Pandas DataFrame library for accessible time series dataschema.The measures of performance will be grounded on the prognosticated stock ticker price in comparison to both the factual price and the standard model's prognosticated price. For this we will use a Linear Retrogression model as its primary standard.

This seeks to use Deep literacy models, Long-Short Term Memory(LSTM) Neural Network algorithm, to prognosticate stock prices. For data with timeframes intermittent neural networks(RNNs) come in handy but recent inquiries have shown that LSTM, networks are the most popular and useful variants of RNNs. I'll use Keras to make a LSTM to prognosticate stock prices using literal ending price and trading volume and fantasize both the prognosticated price values over time and the optimal parameters for the model.

Datasets and Inputs

We will be using the daily prices of the S&P 500 from January 2000 to June 2017, this is a series of data points indexed in time order or a time series. My goal will be to predict the closing price for any given date after training. All of the necessary data for this will come from Google Finance.



```
import numpy as np

data = pd.read_csv('google.csv')
print(data.head())

print("\n")
print("Open   --- mean :", np.mean(data['Open']),  "  \t Std: ", np.std(data['Open']),  "  \t Max: ", np.max(data['Open']),  "
print("High   --- mean :", np.mean(data['High']),  "  \t Std: ", np.std(data['High']),  "  \t Max: ", np.max(data['High']),  "
print("Low    --- mean :", np.mean(data['Low']),   "  \t Std: ", np.std(data['Low']),   "  \t Max: ", np.max(data['Low']),   "
print("Close  --- mean :", np.mean(data['Close']), "  \t Std: ", np.std(data['Close']), "  \t Max: ", np.max(data['Close']), "
print("Volume --- mean :", np.mean(data['Volume']),"  \t Std: ", np.std(data['Volume']),"  \t Max: ", np.max(data['Volume']),"
```

```
        Date    Open    High     Low   Close   Volume
0  30-Jun-17  943.99  945.00  929.61  929.68  2287662
1  29-Jun-17  951.35  951.66  929.60  937.82  3206674
2  28-Jun-17  950.66  963.24  936.16  961.01  2745568
3  27-Jun-17  961.60  967.22  947.09  948.09  2443602
4  26-Jun-17  990.00  993.99  970.33  972.09  1517912


Open   --- mean : 382.51416852146383     Std:  213.48651683323408       Max:  1005.49          Min:  87.74
High   --- mean : 385.87209856915837     Std:  214.60229511464593       Max:  1008.61          Min:  89.29
Low    --- mean : 378.7371255961842      Std:  212.08010983015183       Max:  996.62           Min:  86.37
Close  --- mean : 382.3502480127191      Std:  213.43595872312488       Max:  1004.28          Min:  87.58
Volume --- mean : 4205707.889666136      Std:  3877483.007730211        Max:  41182889         Min:  521141
```

Figure 2: Dataset Loading.

Data Preprocessing

After loading the data we remove unnecessary data, i.e., Date and High value, which is not required in building a machine learning model.

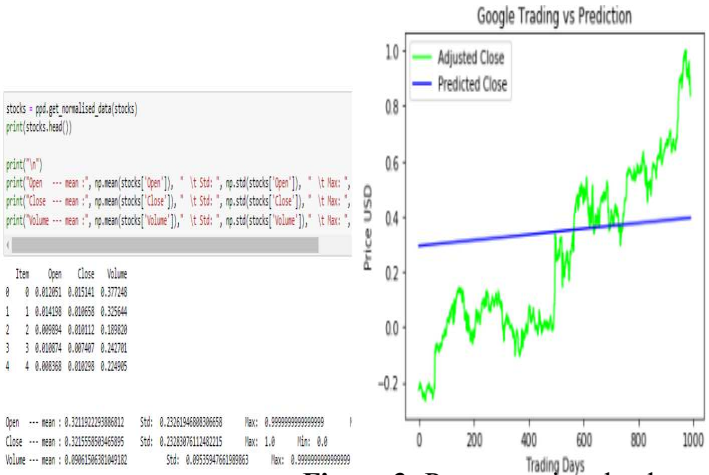


Figure 3: Preprocessing the data

After Preprocessing the data by removing unwanted attributes we Normalise the data using minmaxscaler function and Visualize the data which is shown in figure 4.

Figure 4: Data Visualization.

Model Building

After preprocessing the data we can start building out machine learning model and we will take bench mark model is a simple linear regressor model. For which we follow the following steps.

Step 1: Load the preprocessed data

Step 2: Split data into train and test pair

Step 3: Train a Linear regressor model on training set and get prediction

Step 4: Get prediction on test set

Step 5: Plot the predicted values against actual, which is shown in figure 5.

Figure 5: Predicted values against actual value using LR Model.

Step 6: measure accuracy of the prediction which is shown in figure 6.

```
trainScore = mean_squared_error(X_train, y_train)
print('Train Score: %.4f MSE (%.4f RMSE)' % (trainScore, math.sqrt(trainScore)))

testScore = mean_squared_error(predictions, y_test)
print('Test Score: %.8f MSE (%.8f RMSE)' % (testScore, math.sqrt(testScore)))
```

```
Train Score: 0.1852 MSE (0.4303 RMSE)
Test Score: 0.08133781 MSE (0.28519784 RMSE)
```

Figure 6: Performance Measure of Benchmark LR Model.

Now we can build our proposed Long-Sort Term Memory Model

Step 1 : import keras libraries for smooth implementation of lstm

Step 2 : Split train and test data sets and Unroll train and test data for lstm model

Step 3 : Build a basic Long-Short Term Memory model

Step 4: Train the model

Step 5: make prediction using test data

Step 6: Plot the results which is shown in figure 7.
Step 7: Get the test score.

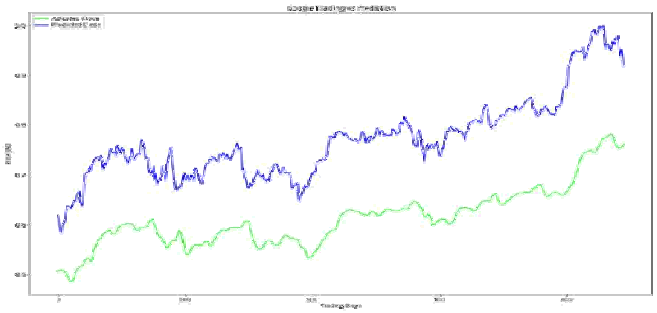


Figure 7: Predicted values against actual value using LSTM Model.

```
trainScore = model.evaluate(X_train, y_train, verbose=0)
print('Train Score: %.8f MSE (%.8f RMSE)' % (trainScore, math.sqrt(trainScore)))

testScore = model.evaluate(X_test, y_test, verbose=0)
print('Test Score: %.8f MSE (%.8f RMSE)' % (testScore, math.sqrt(testScore)))
```

```
Train Score: 0.00268990 MSE (0.05186425 RMSE)
Test Score: 0.02585540 MSE (0.16079614 RMSE)
```

Figure 8: Performance Measure of LSTM Model.

Comparison of models

For this we will measure performance using the mean squared difference between predicted and actual values of the target stock at adjusted close price and the delta between the performance of the benchmark model (Linear Regression) and our primary model (LSTM). We now analyze the models with parameters such as RMSE (Root Mean Square Error) and MSE (Mean Square Error) and the comparison summary.

Parameter	Existing LR Model	Proposed LSTM
RMSE	28.51	16.07
MSE	8.13	2.58

Table 1: Summary of Models.

VI. Conclusion

A forecasting algorithm is a process that seeks to predict future values based on the past and present data. This historical data points are extracted and prepared trying to predict future values for a selected variable of the dataset. During market history there have been a continuous interest trying to analyze its tendencies, behavior and random reactions. From the summary of model performance parameters, we can see that LSTM Model performs better than the benchmark LR model for the datasets.

References

[1] 1. Bhawna Panwar; Gaurav Dhuriya; Prashant Johri; Sudeept Singh Yadav; Nitin Gaur" Stock Market Prediction Using Linear Regression and SVM" in 2021, IEEE .

[2] Karthikeyan C, 2Sahaya Anselin Nisha A, 3Anandan P, 4Prabha R, 5Mohan D, 6Vijendra Babu D" Predicting Stock Prices Using Machine Learning Techniques" in IEEE 2021.

[3] Zaharaddeen Karami Lawa, Hayati Yassin, Rufai Yusuf Zakari " Stock Market Prediction using Supervised Machine Learning Techniques: An Overview" , in IEEE 2021.

[4] Pawee Werawithayaset, Suratose Tritilanunt, " Stock Closing Price Prediction Using Machine Learning" , in IEEE, 2019.

[5] H. L. Siew and M. J. Nordin, "Regression techniques for the prediction of stock price trend," 2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE), Langkawi, 2012, pp. 1-5.

[6] K. V. Sujatha and S. M. Sundaram, "Stock index prediction using regression and neural network models under non normal conditions," INTERACT-2010, Chennai, 2010, pp. 59-63.

[7] 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), Wuhan, 2018, pp. 2787-2790.

[8] T. Gao, Y. Chai and Y. Liu, "Applying long short term momory neural networks for predicting stock closing price," 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, 2017, pp. 575-578.

[9] K. A. Althelaya, E. M. El-Alfy and S. Mohammed, "Evaluation of bidirectional LSTM for short-and longterm stock market prediction," 2018 9th International Conference on Information and Communication Systems (ICICS), Irbid, 2018, pp. 151-156.

[10] M. Usmani, S. H. Adil, K. Raza and S. S. A. Ali, "Stock market prediction using machine learning techniques," 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, 2016, pp. 322-327.

[11] K. Raza, "Prediction of Stock Market performance by using machine learning techniques," 2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT), Karachi, 2017, pp. 1-1.

[12] R Dubey, D Rathore, D Kushwaha, JP Maurya, "An empirical study of intrusion detection system using feature reduction based on evolutionary algorithms and swarm intelligence methods", International Journal of Applied Engineering Research 12 (19), 2017. pp. 8884-8889.