WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Use of Machine Learning Algorithm for Stock Market Analysis

Dr. Uma Pujeri[1], Prof Madhura Phatak[2], Mr Amit Samgir[3], Mr Inderjeet Balotia[4]

[1,2,3,4]School of Computer Engineering and Technology, Dr. Vishwanath Karad MIT World Peace University, Kothrud, Pune, Maharashtra, 411038, India

A B S T R A C T

One of the trickiest tasks in the area of math is anticipating securities exchange. The prediction is affected by a multitude of factors, including physiological as also actual factors, normal and naive behavior, financial backer sentiment, market rumors, etc. As there are so many tangibles and intangibles, stock prices are extremely difficult to anticipate and that too with accuracy. We study information analysis as a distinct edge in this field.

When all information about a firm and financial exchange events are immediately known to all stakeholders and market financial supporters, the effects of those happenings have already been taken into account in the stock price, according to the expert market hypothesis. Accordingly, it is argued that the genuine spot price reflects the influence of any ongoing business sector events and can be used to forecast its future growth. We use a Machine Learning (ML) algorithm with real-time historical stock price data to predict next patter. All influencing factors are represented by their most recent appearance in the stock price, which is how we calculate future patter.

As a result, using the previous stock price as the last appearance of all impacting elements, we use a Machine Learning (ML) algorithm using authentic old stock price data to predict future patterns. Machine learning techniques may unearth examples and bits of knowledge that were not previously explored, and they may be used to generate unerringly close to precise predictions.

This work presents a system utilizing LSTM (Long Short-Term Memory) model and organizations' net development estimation calculation to examine as well as forecast of future development of an organization with regards to the value of stock prices.

Keywords: Machine Learning Algorithm, Long Short-Term Memory model, stock price

Introduction

Information analysis have been utilized in all business sectors for informed decision making. In share market, there are many variables that influence the offer price, and the example of the difference in cost isn't customary. therefore taking a informed decision based on mathematical calculations helps in saving the cost. Artificial Neural Network (ANN) learns from the past information and uses this info to predict the future patterns. Other machine learning methods, such as Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and others, function flawlessly with multivariate time series data. Our model is trained using data from historical stock information, which computes the stock's future cost. This estimated future value is used to estimate how an organization will develop in the future. We also discovered a tendency for future development from different organizations. As a result, we can assess and evaluate the favorable future trajectory of an organization relative to others. Each time a stakeholder request is set for sell or purchase and an exchange is completed, the stock price of a listed firm changes in a stock trade. A trade gathers all sell offers with expected cost per stock (regularly it is more than the cost paid while purchased by the financial backer) and all purchase offers regardless of a cost limit (ordinarily a financial backer expects the future cost of the stock will be more than the ongoing value he is paying now) and a purchase sell exchange is committed when the two offers have a match for example selling bid cost is same with purchasing bid cost of some purchase bid Fama in 1970 [6] proposed proficient market speculation which expresses that in a productive market (where all occasions are referred to all partners as a when it works out) the impact of all market occasions are as of now consolidated in stock costs subsequently foreseeing and analyzing previous occasions or prices is cumbersome. The stock cost of an organization relies upon numerous tangible and intangibles. Full scale financial circumstances also assume a significant part in development or decline of an stock or company in general. A portion of the inherent variables could be organization's net earned profits, liabilities, taxes, dependability, competition in market, actually progressed sequential construction system, excess money for unfavorable circumstances, stakes in unrefined substance provider and completed item wholesalers and so on. Those factors that are past the control of the organization, for example, raw petroleum cost, dollar swapping scale, political soundness, government strategy choice and so on go under extraneous property. Many analysts have taken their stance at involving the authentic stock costs as the reason for time series examination to arrive at future stock costs. A wide range of measurable models were applied since long like moving average (MA), auto regression (AR), weighted moving normal, ARIMA, CARIMA and so on. Later a few non direct models were additionally attempted like GARCH. As of late unique brain network models, developmental calculations are being applied for stock forecast with good progress. Profound neural networks like CNN, RNN are additionally utilized with various parameter settings and highlights. In this paper we will investigate a unique kind of RNN known as LSTM to foresee future organization development in view of past stock costs.

The main Objective of this research is

I. The major goal of this project is to accurately anticipate stock prices and forecast current market trends so that investors may make more knowledgeable and educated investment decisions.

II. The proposed work aims to investigate and enhance supervised learning systems for stock price prediction.

III. The system must have access to a database of previous pricing. Based on previous performance, it must estimate the stock price.Additionally, it must offer a real-time display of the market index.

IV. Using a particular kind of RNN called a Long-Short Term Memory (LSTM) to model and forecast the price of a stock market index in the future.This prediction will be done based on historical data of the Indian Indices(Nifty 50) OR The Stock Prices (ex.SBIN).

V. To minimize the tolerance rate as low as possible.

Motivation behind this research is

The stock market experiences erratic fluctuations, and there are numerous intricate financial indicators. The development of technology has given people the chance to profit steadily from the stock market and has also made it possible for specialists to identify the most useful signs and make more accurate predictions. It is crucial to make accurate market value predictions in order to maximise the return on stock option purchases while minimising risk.

Literature Survey

Stock price prediction may be done utilizing AI and AI models in AI fields. Utilizing the SVM based model for stock price prediction . SVM algorithm is one of the AI algorithms which deals with arrangement calculations. It is utilized to get a new text as a result. Applying Multiple Linear Regression with Interactions to anticipate the pattern in stock costs[3]; Lufuno Ronald Marwala . EMH is not the same as the Random walk speculation however the EMH works fundamentally on Transient examples at foreseeing stock costs. Manh Ha Duong Boris et al 2006 [2] search the reflection between value costs and consolidated funds in Key European countries like UK and Germany. Spending increase in Eu countries ventures is well-suited to results considerably Stronger relationship between the different Eu countries and value costs.The arbitrary walk speculation process is utilized in CIVETS. In an effective and robust securities exchange, the values must follow an irregular walk speculation, with regards to the future value, the qualities are evolving arbitrarily and unpredictable. Regular returns for rising and further developed markets have been tried for irregular strolls. LSTM calculation comprises a Recurrent Neural organization to encode information. The calculation inputs are monetary news headings implantation From Bloomberg and Reuters. Long Momentary Memory with implanted layer and the LSTM with the programmed encoder in the securities exchange for foreseeing stock qualities. The Xiongwen Pang et al [1]. Utilized a programmed encoder and inserted layer to vectorizing the qualities by utilizing LSTM layers. Relationship coefficients in stocks are chosen haphazardly and anticipated utilizing ARIMA and the brain network approach. In this RNN and LSTM calculations are carried out. M. Nabipour et al [7] Utilized different AI and profound learning calculations for anticipating stock values, for example, arbitrary woodland, choice tree and brain organizations. LSTM gives the most exact outcomes and it has the best capacity to fit. LSTM gives the best outcomes while foreseeing stock costs with the least mistake rate (Hyeong Kya Choi,2018 [5]; Huicheng Liu, 2018 [7]; Xiongwen Pang et al, 2020 [1]). Xiao Ding et al. 2020 [4] utilized a simple and powerful connection point to add presence of mind information to the interaction while learning of occasions. The LMS channel is a kind of versatile channel which is utilized for taking care of straight issues.

Equations

Artificial neural networks are computational models that function in the same way as the human nervous system does. These sorts of networks are built using mathematical operations and a set of parameters to decide the result.

There are different Artificial neural networks currently being used:

1. CNN
2. Kohonen Self Organizing Neural Network
3. Radial basis function Neural Network
4. Feedforward Neural Network – Artificial Neuron
5. Long Short Term Memory

Why LSTM?

- Some ANNs produce an output without even thinking about the prior input, assigning a weight matrix to the current input.As a result, information only passes through ANN once, and previous information is not kept. Hence ANN do not perform well where time context is required i.e Time series data.

- The LSTM resolves a major lacuna that recurrent neural networks face i.e. short-memory. The LSTM uses a succession of 'gates,' each with its own RNN, to retain, forget, or discard data points based on a probabilistic model. LSTMs can also be used to handle exploding and vanishing gradient problems.
- With time delay of arbitrary interval , LSTM is more sophisticated and better suited to identify, analyze and predict time series. A benefit of LSTM over competing RNNs, hidden Markov models, and other sequence learning techniques is their relative insensitivity to arbitrary interval length.

LSTM:

The term "Long-Short Term Memory" refers to a particular kind of RNN that can learn long-term dependency (LSTM).RNN can recall long-term inputs with the help of LSTM. The information in its memory can be read, written, and deleted. The memory can be thought of as a closed cell with a closed description that makes the decision whether to store or erase information.

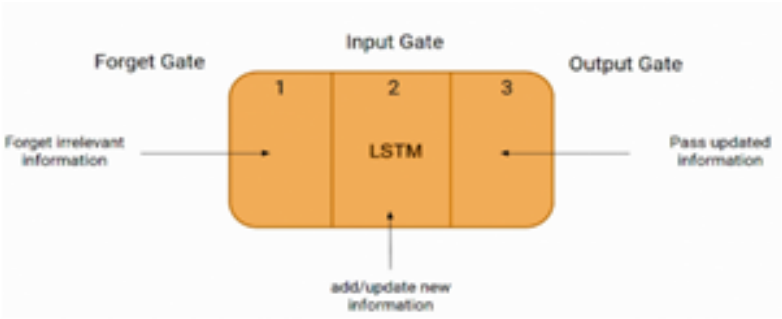

Fig.1: LSTM Architecture

How LSTM Works?

LSTMs algorithm uses a number of gates to streamline how data in a sequence enters, stored , and leaves the network. A typical LSTM has three gates, including:

Forget gate: A neural network is fed with current active input and the prior state. Network creates a vector for each of its component falling between [0, 1]. The network is trained in such a way that when output is 0 then input component is regarded as irrelevant where as when the output is 1 the input is considered to be vital. Based on the previous suppressed state and the current data point in the sequence, the forget gate determines which parts of the long-term memory should be forgotten (have less weight) at this particular instant.

Input gate: With the prior hidden state and current active incoming data, the Input data decides what new information should be added to the network's long term memory(cell state). A new memory update vector is produced by combining the prior concealed state with current incoming active data. Function of the input gate is to determine which elements of the "new memory vector" should kept. Pointwise multiplying is done on the outputs. The concatenated vector that results is subsequently added to the cell state, resulting in long-term memory. The network is being updated.

Output gate: It assess three elements to determine the new hidden state: the recently updated cell state, the previous hidden state, and the new input data. The complete approach for this final phase is as follows:

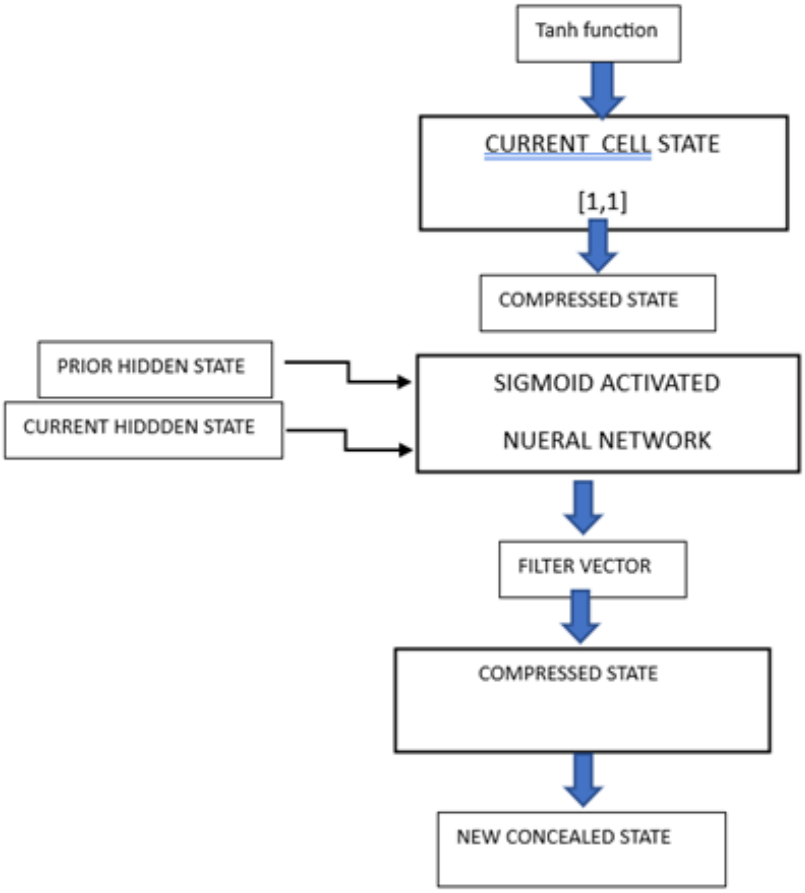

Fig.2: Flowchart of output gate

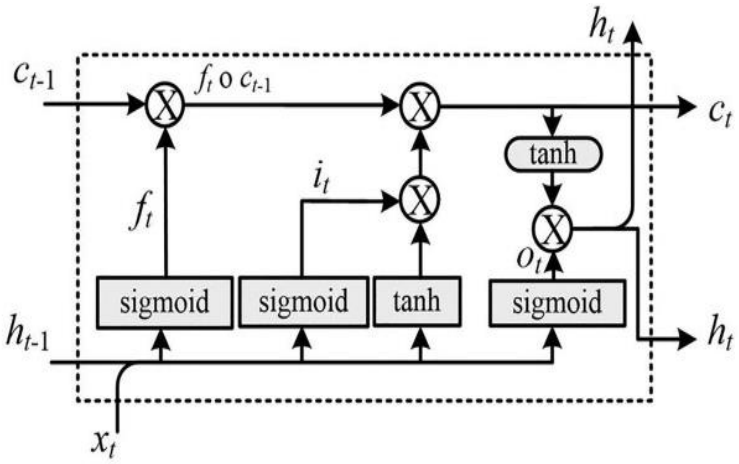

Fig.3 : LSTM Cell

Overall Architecture :

- **Data Selection:**Choosing data for an organisation and dividing it into training and testing is the first stage. We spent 80% of the resources for training and 20% for testing.
- **Pre-processing of data:** In pre-processing,the remaining properties and those needed by the algorithm are not taken into consideration. To acquire numbers inside a certain range, we employ normalisation during pre-processing.
- **Prediction using LSTM:** The LSTM algorithm is employed in this system to forecast stock values. The system first runs the training data through it to train the model. The anticipated values are then contrasted with the actual values during the testing phase.

- **Evaluation:** Involves computing the RMSE, loss, and accuracy.

ALGORITHM :

Step 1: Start

Step 2: Data pre-processing of historical data received

Step 3 : Parse the close price of data set

Step 4: Apply the feature scaling to data, range between 0 to 1

Step 5 : Building data structure with 1 output and 60 timestamps

Step 6 : Using data set received from step 5 and create a RNN, initialize with sequential repressor.

Step 7 : Add first LSTM layer and drop under desired values

Step 8 : Add Output Layer

Step 9: Add adam optimization and loss as a mean squared error during RNN compilation.

Step 10: Use charting tool to visualize result and perform prediction.

Step 11 Show final predicted price of next day

4. Result and Analysis

- Python is used to carry out the LSTM-based model that is proposed.
- Below figures represents the graphs of the inputs given to the lstm model.

Fig.6: Graph of Nifty-50 Close Price History

Fig.7: Graph of SBI Stock Close Price History

- By using 80% -20% data for training and testing respectively we have plotted the graph of predicted and actual values.

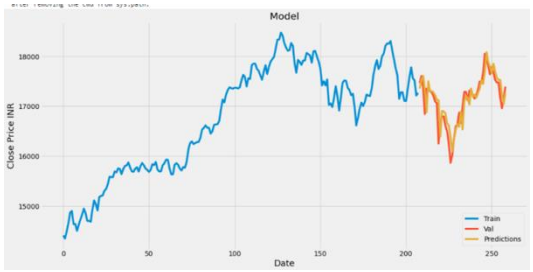

Fig.8: Graph of predicted price and actual price of Nifty-50.

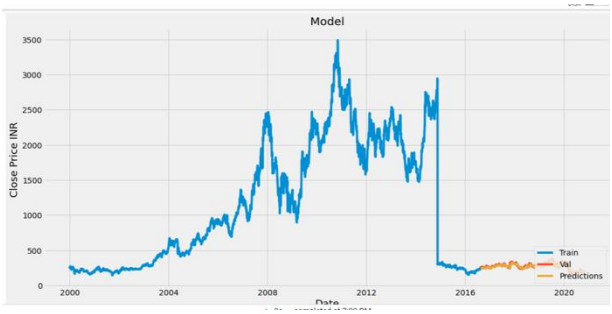

Fig.9: Graph of predicted and actual price of SBI Stock.

- The final output given by the system is the next day predicted closing price of Nifty-50 and SBI.
- **Output:**

```
[[17476.154]]
```

Factors used for determining the hyperparameters of the Model:

The following criteria are taken into account while determining **hyperparameters** such as the number of epochs, nodes, and hidden layers, number of units in a hidden layer, activation function, etc.

Root Mean Squared Error (RMSE): The residuals' standard deviation is what it is (prediction errors). It reveals how dispersed the data is around the line of best fit, or how far apart from the regression line the data points are.

The lower the rmse value, the better the given model is able to fit a dataset. In the proposed model, rmse value depends on how volatile the input prices are. In case of stocks (SBI), the rmse value is relatively much lower than in the case of the Indices (Nifty-50).

RMSE value of Nifty-50:

```
rmse

241.97255221149098
```

RMSE value of SBI Stock:

```
rmse

16.82912607325965
```

2) **Loss:**A prediction model's ability to accurately forecast the desired outcome is measured by a loss function. According to the epochs in this model, the loss steadily reduces.

3) **Accuracy:** LSTM model has constant accuracy and doesn't variate . The main reason of our constant accuracy is limited number of inputs given to our lstm model. Constant accuracy doesn't affect the output of the model.

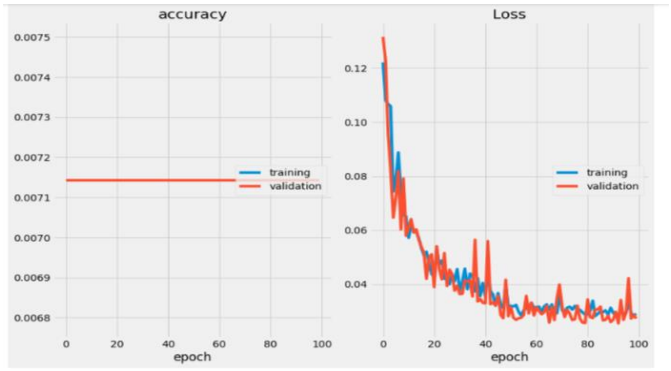

Fig.10: Graphs of Accuracy and Loss VS Epoch of Nifty-50.

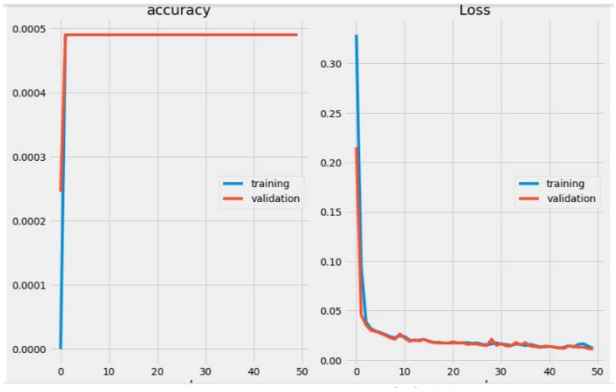

Fig.11: Graphs of Accuracy and Loss VS Epoch of SBI Stock.

5. Conclusion and Future Scope

In pursuing the above work,our aim was to to predict and calculate the future closing price of stock for any given organization. To achieve this we fabricated a web application utilizing the LSTM algorithm. and to optimize the accuracy, we can study and analyze various share price trends across industries and conduct evaluation of graphs over a larger range of time periods.

This framework aims to aid in market analysis and trend evaluation and growth projections for various businesses and enterprises across various time horizons.The accuracy of the model can be further optimized if the model is trained on more data sets.

Future Scope

- the same model may be tested for predicting crypto-currencies and its results can be analyzed and compared.
- Sentiment analysis may be included as one of the input parmeters and the results may be further analyzed for accuracy comparision.
- Same model may also be applied on various foreign markets and its results can be analyzed.

References

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

II. Boriss Siliverstovs & Manh Ha Duong, 2006. "On the Role of Stock Market for Real Economic Activity," Discussion Papers of DIW Berlin 599, DIW Berlin, German Institute for Economic Research.

III. Lufuno Ronald Marwala, " Forecasting the Stock Market Index Using Artificial Intelligence Techniques", A dissertation submitted to the Faculty of Engineering and the Built Environment, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Engineering

IV. Xiao Ding, Kuo Liao, Ting Liu, Zhongyang Li, Junwen Duan Research Centre for Social Computing and Information Retrieval Harbin Institute of Technology, China - Event Representation Learning Enhanced with External Common-sense Knowledge.

V. Huicheng Liu Department of Electrical and Computer Engineering Queen's University, Canada - Leveraging Financial News for Stock Trend Prediction with Attention-Based Recurrent Neural Network.

VI. Hyeong Kyu Choi, B.A Student Dept. of Business Administration Korea University Seoul, Korea = Stock Price Correlation Coefficient Prediction with ARIMA-LSTM Hybrid Model.

VII. M. Nabipour Faculty of Mechanical Engineering, Tarbiat Modares University, 14115-143 Tehran, Iran; Mojtaba.nabipour@modares.ac.ir - Deep Learning for Stock Market prediction.

XI. Uma P, Shamala M, Himangi P,Piyanshu G, " DATA LOSS PREVENTION TECHNIQUES", European Chemical Bulletin 12(S3), 5054 – 5060