Journal of Harbin Engineering University Vol 44 No. 5
ISSN: 1006-7043 May 2023

Artificial Intelligence Based Stock Price Prediction Using Machine Learning

[1]Dr. J. Vimala Ithayan, [2]Mrs.R.Sujitha, [3] Mrs.V.G.Karthiga, [4]Mrs.S.Vinotha, [5]Mrs.S.Sasikala, [6]J. Saranraj [7]Dr. A. Kingsly Jabakumar

[1]Assistant Professor, Computer Science and Engineering,
Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai
[2] Assistant Professor, Computer Science and Engineering,
Christ the King Engineering College, Coimbatore
[3]Assistant Professor, Computer Science and Engineering,
Christ the King Engineering College, Coimbatore
[4]Assistant Professor, Computer Science and Engineering,
Christ the King Engineering College, Coimbatore
[5]Assistant Professor, Electronics and Communication Engineering,
Sree Sakthi Engineering College, Coimbatore
[6]Assistant Professor, Electronics and Communication Engineering,
Christ the King Engineering College, Coimbatore
[7]Associate Professor, Electronics and Communication Engineering,
Christ the King Engineering College, Coimbatore
[1] drvimalaithayanj@veltech.edu.in, [2]srisuji14@gmail.com, [3]karthiga.mecse@gmail.com,
[4]vinotha2909@gmail.com
[5]sasikala27.ece@gmail.com, [6]er.saranraj89@gmail.com,
[7]kingslyjkumar@gmail.com

Abstract

Stock market has been the Centre of attraction for investors for a long period of time. The investor's goal is to buy the stock, hold it for a period, and then, sell the stock for more investor paid for it. Many people invest to create wealth and to gain a rich reward. By investing in the stock market, it will improve the returns equity. In this project, the main focus is to predict the future stock price movement for more company listed in India. This project used eight months daily basis of historical data to model the relationship using long short term memory (LSTM). By using logistic regression, stock market movement able to predict the stock price movement, either an increasing trend or unchanged or decreasing movement. In this work, LSTM techniques have been utilized for predicting the next day closing price for five companies belonging to different sectors of operation. The financial data: Open, High, Low and Close prices of stock are used for creating new variables which are used as inputs to the model.

Keywords: LSTM , operation, sectors, stock, financial

Introduction

The financial market is a dynamic and composite system where people can buy and sell currencies, stocks, equities and derivatives over virtual platforms supported by brokers. The stock market allows investors to own shares of public companies through trading either by exchange or over the counter markets. This market has given investors the chance of gaining money and having a prosperous life through investing small initial amounts of money, low risk compared to the risk of opening new business or the need of high salary career. Stock markets are affected by many factors causing the uncertainty and high volatility in the market. Although humans can take orders and submit them to the market, automated trading systems (ATS) that are operated by the implementation of computer programs can perform better and with higher momentum in submitting orders than any human. However, to evaluate and control the performance of ATSs, the implementation of risk strategies and safety measures applied based on human judgements are required. Many factors are incorporated and

considered when developing an ATS, for instance, trading strategy to be adopted, complex mathematical functions that reflect the state of a specific stock, machine learning algorithms that enable the prediction of the future stock value, and specific news related to the stock being analysed.

Time-series prediction is a common technique widely used in many real-world applications such as weather forecasting and financial market prediction. It uses the continuous data in a period of time to predict the result in the next time unit. Many time-series prediction algorithms have shown their effectiveness in practice.Prediction of the stock price is one of the most difficult scenarios we have in hand. There are many factors that should be incorporated and taken into account while predicting the stock price. These factors include physical, psychological, rational, irrational behavior etc. All this factor makes the stock price prediction more volatile and very difficult to predict with a high degree of accuracy. The main objective or the motivation behind this comparative analysis is to find out a clear-cut analysis of each of the algorithm which can be further implemented in the prediction of the stock price. This will in turn perform as a preliminary step to incorporate they dynamic changes in the market and also put it in to the algorithm which then gives a real time value.

The stock market act as a third party who is responsible for successful completion a trade between the buyer and seller. They act as the intermediary to avoid default and cheating in the market, so the exchange will select the shares of trusted firms. Stock which are also called as shares generally represents ownership of a company to retail and institutional investors trading in the stock market. In this work we are trying to predict the future price of a stock and this prediction is expected to be robust, accurate and efficient. The proposed system is studied and planned such that the working is according to the real-life scenarios and should be well suited to the real-world settings. The system or the algorithm which we are using is expected to take all the technical factors and variables that might affect the price of a stock and its performance from the historical data. It has been found that numbers of these techniques have been previously studied up on in the prediction of the stock price

This technique of machine learning also involves the usage of artificial intelligence which have a great deal of impact in this field which in-turn teaches the system to learn and improve from the learning experiences without being programmed in time to time. There are also some traditional methods used in the prediction of stock price using machine learning algorithm such as the Backward propagation, which is also known as the Backward Propagation Errors. We will also be taking in to consideration of machine learning algorithms such as Linear Regression model, Arima model, The Prophet method, Long short-term memory etc. Researchers are now using many techniques in predicting the future stock price, these consist of many ensemble learning techniques. The highlight of these techniques is that they are very less time consuming and low price in predicting the future value price. Stock market prediction for short time window seems to be a random process. The stock price movement over a long time will usually develop a linear curve. The general tendency of the people is to buy the stock as the price may go high or rise in the near future. The main problem with people not investing in stock market is due to the uncertainty associated with the market value.

These refrain people from investing in the stock. So there should be a proper mechanism that will help the people in predicting the price in real life scenarios. The methods used to predict the future price of the stock includes some time-series analysis using regression and other models. Also, some technical analysis and some machine learning algorithms are used for the price prediction

1. Literature Review

"What other people think" has always been an important piece of information for most of us during the decision-making process. The Internet and the Web have now (among other things) made it possible to find out about the opinions and experiences of those in the vast pool of people that are neither our personal acquaintances nor well-known professional critics — that is, people we have never heard of. And conversely, more and

more people are making their opinions available to strangers via the Internet. The interest that individual users show in online opinions about products and services, and the potential influence such opinions wield, is something that is driving force for this area of interest. And there are many challenges involved in this process which needs to be walked all over in order to attain proper outcomes out of them. In this survey we analysed basic methodology that usually happens in this process and measures that are to be taken to overcome the challenges being faced.

The research work done by V Kranthi Sai Reddy Student, ECM, Sreenidhi Institute of Science and Technology, Hyderabad, India. 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 paper explains the prediction of a stock using Machine Learning. The technical and fundamental o the time series analysis is used by the most of the stockbrokers while making the stock predictions. Programming language is used to predict the stock market using machine learning is Python. 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. In this context this study uses a machine learning technique called Support Vector Machine (SVM) to predict stock prices for the large and small capitalizations and in the three different markets, employing prices with both daily and up-to-the-minute frequencies.

The research work done by Lufuno Ronald Marwala 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. The weak form of Efficient Market hypothesis (EMH) states that it is impossible to forecast the future price of an asset based on the information contained in the historical prices of an asset. This means that the market behaves as a random walk and as a result makes forecasting impossible. Furthermore, financial forecasting is a difficult task due to the intrinsic complexity of the financial system. The objective of this work was to use artificial intelligence (AI) techniques to model and predict the future price of a stock market index. Three artificial intelligence techniques, namely, neural networks (NN), support vector machines and neuro-fuzzy systems are implemented in forecasting the future price of a stock market index based on its historical price information. Artificial intelligence techniques have the ability to take into consideration financial system complexities and they are used as financial time series forecasting tools

Two techniques are used to benchmark the AI techniques, namely, Autoregressive Moving Average (ARMA) which is linear modelling technique and random walk (RW) technique. The experimentation was performed on data obtained from the Johannesburg Stock Exchange. The data used was a series of past closing prices of the All Share Index. The results showed that the three techniques have the ability to predict the future price of the Index with an acceptable accuracy. All three artificial intelligence techniques outperformed the linear model. However, the random walk method out performed all the other techniques. These techniques show an ability to predict the future price however, because of the transaction costs of trading in the market, it is not possible to show that the three techniques can disprove the weak form of market efficiency. The results show that the ranking of performances support vector machines, neuro-fuzzy systems, multilayer perceptron neural networks is dependent on the accuracy measure used.

The research work done by Dharmaraja Selvamuthu, Vineet Kumar and Abhishek Mishra Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India. A stock market is a platform for trading of a company’s stocks and derivatives at an agreed price. Supply and demand of shares drive the stock market. In any country stock market is one of the most emerging sectors. Nowadays, many people are indirectly or directly related to this sector. Therefore, it becomes essential to know about market trends. Thus, with the development of the stock market, people are interested in forecasting stock price. But, due to dynamic nature and liable

to quick changes in stock price, prediction of the stock price becomes a challenging task

On the other hand, events extracted from raw texts lacks of common-sense knowledge, such as the intents and emotions of the event participants, which are useful for distinguishing event pairs when there are only subtle differences in their surface realizations. To address this issue, this paper proposes to leverage external common-sense knowledge about the intent and sentiment of the event. Experiments on three event-related tasks, i.e., event similarity, script event prediction and stock market prediction, show that our model obtains much better event embeddings for the tasks, achieving 78% improvements on hard similarity task, yielding more precise inferences on subsequent events under given contexts, and better accuracies in predicting the volatilities of the stock market1.

Markets are mostly a non-parametric, non-linear, noisy and deterministic chaotic system (Ahangar et al. 2010). As the technology is increasing, stock traders are moving towards to use Intelligent Trading Systems rather than fundamental analysis for predicting prices of stocks, which helps them to take immediate investment decisions. One of the main aims of a trader is to predict the stock price such that he can sell it before its value decline, or buy the stock before the price rises. The efficient market hypothesis states that it is not possible to predict stock prices and that stock behaves in the random walk. It seems to be very difficult to replace the professionalism of an experienced trader for predicting the stock price. But because of the availability of a remarkable amount of data and technological advancements we can now formulate an appropriate algorithm for prediction whose results can increase the profits for traders or investment firms. Thus, the accuracy of an algorithm is directly proportional to gains made by using the algorithm

2. Proposed System

There are many conventional techniques available for stock market prediction-based news feed system, but they cannot forecast the prices in the long term as news pertaining to the events taking place in the future cannot be predicted. So, the proposed system predicts the stock market prices using a recurrent neural network and Holt–Winters triple exponential implementation, thus using only historical data to predict the closing price of individual stocks. The system takes input from the user about the amount how much they want to invest, the duration of the investment, and how much loss or profit they can bear. The system uses the information given by the user and applies machine learning algorithms to come up with a solution, suggesting to the user where to invest the money for maximum profit and minimize the risk of loss. The database that is already present in the system is used to analyze the market situation and find an optimal solution. Investing in the stock market is tricky work; the project helps the user and gives them an upper hand in the process. The results are as accurate as they can be. Machine learning algorithms work in real time and manipulate the data in real time, providing a much more efficient way to develop the best solution. With the help of machine learning, the system recognizes the previous patterns and tries to suggest the output of what could be the future price of the stock.

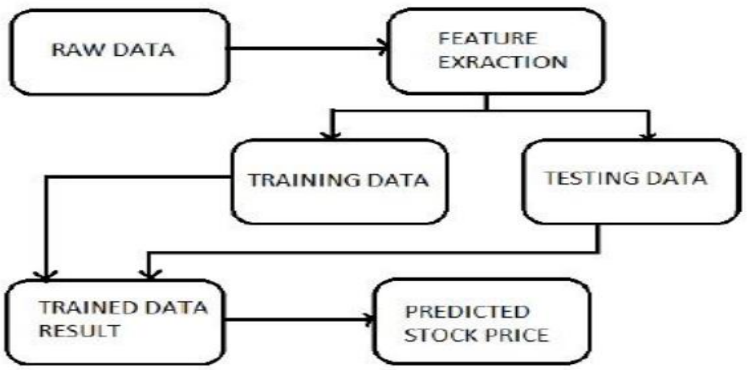

Fig 1: Overall Proposed Architecture

The prediction methods can be roughly divided into two categories, statistical methods and artificial intelligence methods. Statistical methods include logistic regression model, ARCH model, etc. Artificial intelligence methods include multi-layer perceptron, convolutional neural network, naive Bayes network, back propagation network, single-layer LSTM, support vector machine, recurrent neural network, etc. They used Long short-term memory network (LSTM).

The typical LSTM unit consists of a cell, an info door, an entrance door and a door with a view. The cell collects values over discretionary time intervals, and the three inputs manage the progress of data into and out of the cell. The main advantage of the LSTM is its ability to learn context-specific temporal dependence. Each LSTM unit collects information for either a long or short period of time (hence the name) without explicitly using the activation function within the recurrent components. A significant certainty to note is that any cell state is uniquely increased by the output of the overlooked entryway, which changes somewhere in the range of 0 and 1. In other words, the overhead door in the LSTM cell is responsible for both the loads and the capacity to initiate the cell state. Subsequently, data from a past cell state can pass through a cell unaltered rather than expanding or decreasing exponentially at each time-step or layer, and loads can meet their ideal quality in a reasonable measure of time. This allows LSTM's to take care of the evaporating slope issue – as the value put away in the memory cell is not iteratively adjusted, The inclination does not disappear when prepared with back engendering, where markets such as NSE and BSE are considered to be Indian trading entities for our analyzes

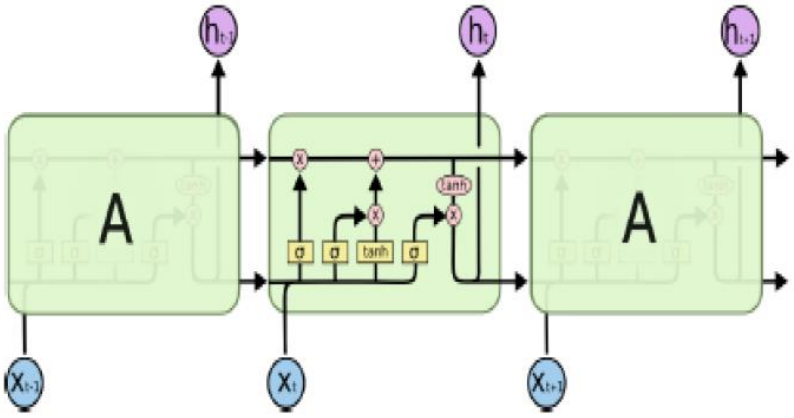

Fig 2: LSTM structure

The typical LSTM unit consists of a cell, an info door, an entrance door and a door with a view. The cell collects values over discretionary time intervals, and the three inputs manage the progress of data into and out of the cell. The main advantage of the LSTM is its ability to learn context-specific temporal dependence. Each LSTM unit collects information for either a long or short period of time (hence the name) without explicitly using the activation function within the recurrent components. A significant certainty to note is that any cell state is uniquely increased by the output of the overlooked entryway, which changes somewhere in the range of 0 and 1. In other words, the overhead door in the LSTM cell is responsible for both the loads and the capacity to initiate the cell state. Subsequently, data from a past cell state can pass through a cell unaltered rather than expanding or decreasing exponentially at each time-step or layer, and loads can meet their ideal quality in a reasonable measure of time. This allows LSTM's to take care of the evaporating slope issue – as the value put away in the memory cell is not iteratively adjusted, The inclination does not disappear when prepared with back engendering, where markets such as NSE and BSE are considered to be Indian trading entities for our analyzes.

The next step is to preprocess the data; in this step the Information Pre-Processing is a significant advance in information mining here the change in crude information into a basic configuration is required. The information which is retrieved from source will be conflicting,

fragmented and it will contain mistakes. The preprocessing step will purify the information; toward the end there is a need to perform highlights scaling which will restrict the factors.

The preparation of the model incorporates cross-approval, which is a very well-founded, projected execution of the model using the preparation information. the purpose of the tuning models is to explicitly tune the calculation training is to add information to the calculation itself. The test sets are immaculate, as a model ought not to be made a decision about dependent on concealed information. Scale up the information to the genuine offer costs. The final step is to draw the data using visualization technique that helps to show the variation of data in the outcome of our algorithm.

The implementation of proposed LSTM model using python which predicts the future price of TATAMOTORS share based on its historical data. The below visualization figure shows the visualization of TATASHARE prediction. In our paper the implementation of an algorithm which predicts the stock price of a share for given period of time, the below graph from our algorithm will show the predicted price of TATAMOTORS share. In the result shown in the below graph is the plotted form our algorithm outcome by applying 96 LSTM units for achieving the accuracy.

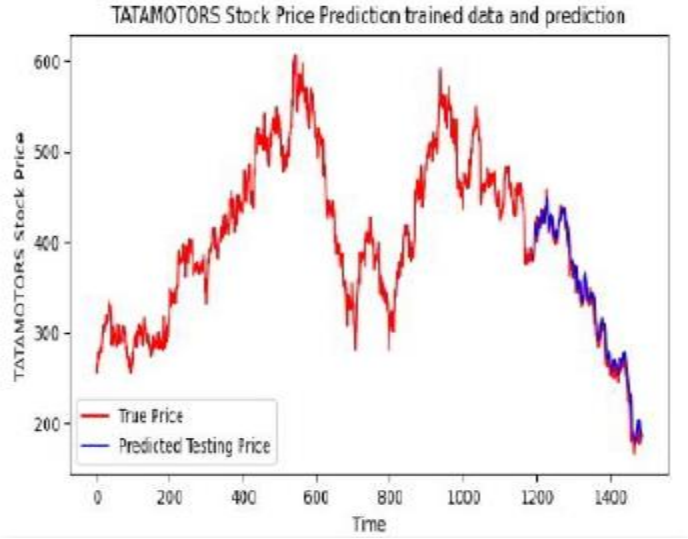

Fig 3: Predicted Testing Stock Price

The graph has been plot for whole data set along with some part of trained data. the graph is showing the open price of TATAMOTORS share for 1484th day's opening price with very minimal loss. the algorithm has plotted the graph successfully along with the predicted price testing price (blue) and true price (red), there is a slight difference in predicting the price between the predicted price testing price (blue) and true price (red), which proves that our algorithm is able to predict the with minimum loss rate for the given complete data set of a particular share.

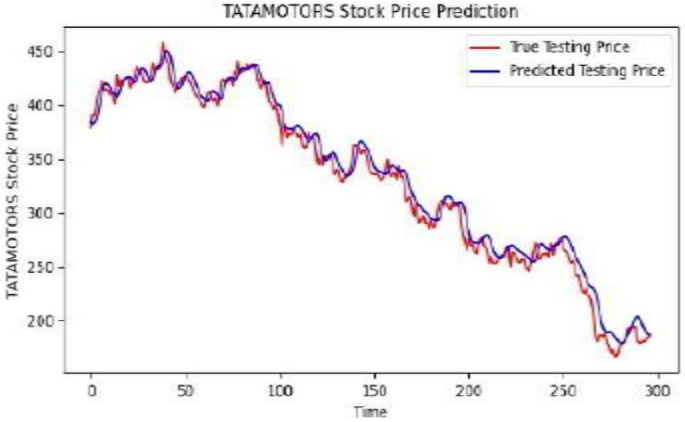

Fig 4: Predicted Stock price

3. Conclusion

In this work, we are predicting closing stock price of any given organization, we developed a web application for predicting close stock price using LSTM algorithms for prediction. We have applied datasets belonging to Google, Nifty50, TCS, Infosys and Reliance Stocks and achieved above 95% accuracy for these datasets. The study of the share is carried out in this paper and it can be carried out for several shares in the future. Prediction could be more reliable if the model trains a greater number of data sets using higher computing capacities, an increased number of layers, and LSTM modules

References

[1] Stock Price Prediction Using LSTM on Indian Share Market by Achyut Ghosh, Soumik Bose1, Giridhar Maji, Narayan C. Debnath, Soumya Sen

[2] S. Selvin, R. Vinayakumar, E. A. Gopalkrishnan, V. K. Menon and K. P. Soman, "Stock price prediction using LSTM, RNN and CNN-sliding window model," in International Conference on Advances in Computing, Communications and Informatics, 2017.

[3] Murtaza Roondiwala, Harshal Patel, Shraddha Varma, “Predicting Stock Prices Using LSTM” in Undergraduate Engineering Students, Department of Information Technology, Mumbai University, 2015.

[4] Xiongwen Pang, Yanqiang Zhou, Pan Wang, Weiwei Lin, “An innovative neural network approach for stock market prediction”, 2018

[5] Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan Department of Computer Science and Engineering National Institute of Technology, Hamirpur – 177005, INDIA - Stock Market Prediction Using Machine Learning.

[6] Pranav Bhat Electronics and Telecommunication Department, Maharashtra Institute of Technology, Pune. Savitribai Phule Pune University - A Machine Learning Model for Stock Market Prediction.

[7] Anurag Sinha Department of computer science, Student, Amity University Jharkhand Ranchi, Jharkhand (India), 834001 - Stock Market Prediction Using Machine Learning.

[8] V Kranthi Sai Reddy Student, ECM, Sreenidhi Institute of Science and Technology, Hyderabad, India - Stock Market Prediction Using Machine Learning.

[9] Asset Durmagambetov currently works at the mathematics, CNTFI. Asset does research in Theory of Computation and Computing in Mathematics, Natural Science, Engineering and Medicine. Their current project is 'The Riemann Hypothesis-Millennium Prize Problems' - stock market predictions.

[10] Mariam Moukalled Wassim El-Hajj Mohamad Jaber Computer Science Department American University of Beirut - Automated Stock Price Prediction Using Machine Learning.

[11] Manh Ha Duong Boriss Siliverstovs June 2006 - The Stock Market and Investment.

[12] Dharmaraja Selvamuthu , Vineet Kumar and Abhishek Mishra Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India - Indian stock market prediction using artificial neural networks on tick data.