Journal of Population Therapeutics & Clinical Pharmacology

RESEARCH ARTICLE
DOI: 10.47750/jptcp.2023.30.12.015

Stock Market Price Prediction Using Machine Learning

Saranya K[1]*, Vijayashaarathi S[2], Sasirekha N[3], Koushikrajaa M[4], Lohith Raksha S[5]

[1,2,3]Assistant Professor, Department of Electronics and Communication Engineering, Sona College of Technology, Salem, India

[4,5]Department of Electronics and Communication Engineering, Sona College of Technology, Salem, India

***Corresponding author:** Saranya K, Assistant Professor, Department of Electronics and Communication Engineering, Sona College of Technology, Salem, India,
Email: saranya.k@sonatech.ac.in

Submitted: 06 March 2023; Accepted: 19 April 2023; Published: 15 May 2023

ABSTRACT

Stock Market value prediction is changing into a well-liked and vital topic in financial and educational studies. From previous couple of decades, there's associate explosive increase in the average person's interest for stock exchange. The goal of stock market prediction is to foresee the stock price of a specific company over the long term. Accurate prediction of exchange value may be a terribly difficult task due to its unstable and non-linear nature of the stock markets. With the introduction of computer science Artificial Intelligence (AI) and Advanced Machine Learning strategies of prediction have proved to be efficient in predicting the stock costs. In this Paper, Linear Regression and LSTM techniques are used for predicting the longer-term stock value of two corporations of totally different fields of operation. In order to forecast the stock prices over the long run, these machine learning algorithms employ a dataset of stock prices from previous years. These datasets were obtained from Yahoo Finance with values of Date, Open, High, Low, Close, Volume value. Though share market will never be foretold with hundred Percent accuracy, because of its unstable and imprecise nature, this paper aims at proving the potency of LSTM model for predicting the longer-term stock costs with most possible accuracy.

Keywords: *Machine Learning, Linear regression, Long short term memory (LSTM), Neural Networks, ANN, RNN*

INTRODUCTION

Stock market price prediction is a difficult task, even for experienced investors and analysts. However, with the advent of machine learning, it has become possible to use algorithms and computational models to analyze and make predictions about the stock market. In this introduction, In this article, we'll talk about some of the difficulties and restrictions associated with utilizing machine learning to forecast stock values.

A kind of artificial intelligence called "machine learning" uses models and algorithms to anticipate the future and make decisions based on facts. It can be applied to a wide range of tasks, including stock market price prediction. In the context of stock market prediction, machine learning algorithms can be trained on historical

data about various stocks, including their prices, volumes, and other factors. These algorithms can then be used to make predictions about the future prices of these stocks.

One of the main challenges of using machine learning for stock market prediction is the high level of noise and unpredictability in the stock market. Numerous variables, such as the state of the economy, current politics, and even natural calamities, can affect prices. This makes it difficult for even the most sophisticated algorithms to make accurate predictions. Additionally, the stock market is constantly changing, and the algorithms used for prediction must be regularly updated and retrained to keep up with these changes.

Despite these challenges, many investors and analysts are using machine learning to make informed decisions about the stock market. By leveraging the power of computation and data, these techniques can provide valuable insights and help investors make better predictions about the future direction of stock prices. While machine learning is not a magic bullet for stock market prediction, it is an important tool that can help investors navigate the complex and ever-changing world of finance.

LITERATURE SURVEY

Predicting stock prices has been a long-standing challenge for researchers in the field of finance and machine learning. Accurate prediction of stock prices can provide valuable insights for investors and traders, allowing them to make informed decisions about buying and selling stocks[1]. In recent years, machine learning techniques have been widely used to predict stock prices.

One of the earliest works on stock price prediction using machine learning was done using a neural network to predict stock returns. Since then, several studies have used machine learning strategies to forecast stock prices, including decision trees, support vector machines(SVM)[2].

Other machine learning techniques that have been used for stock price prediction include extreme learning machines, linear regression, and deep learning. Deep learning, in particular, has gained popularity in the field due to its ability to learn complex patterns in data and make accurate predictions[3]. In addition to traditional machine learning techniques, some researchers have also explored the use of hybrid approaches that combine machine learning with other techniques such as natural language processing (NLP) and sentiment analysis. For example, An ML model used NLP to extract sentiment from financial news articles and combined it with machine learning algorithms to predict stock prices[4]. Overall, the use of machine learning for stock price prediction has shown promising results, and there is ongoing research in this area to improve the accuracy and robustness of these predictions.

Existing System

The objective of this research is to create a machine learning model that can forecast stock market prices in the future. The following steps outline the proposed work for this prediction model: Data collection: The first step is to collect a large dataset of historical stock prices, as well as any relevant financial and economic indicators that may influence stock prices. This ML model will be tested and trained using this data.

Data pre-processing: Next, collected data will need to be cleaned and pre-processed to ensure that it is in a suitable format for the machine learning model. This may involve filling in missing values, normalizing numerical data, and encoding categorical data.

Feature engineering: Once the data is clean and pre-processed, the next step is to identify and create relevant features that will be used as input to the machine learning model. This may involve using technical indicators, such as moving averages and relative strength index, as well as fundamental data, such as earnings per share and price-to-earnings ratio.

Model selection and training: The features have been found and retrieved, and the subsequent step is to choose a suitable algorithm for machine learning and train it using the data. This may involve experimenting with different model architectures and hyperparameters, and using techniques such as cross-validation to ensure that the model is robust and generalizes well to new data.

J Popul Ther Clin Pharmacol Vol 30(12):e130–e136; 15 May 2023.

Model evaluation: Once the model is trained, it will need to be evaluated to determine its accuracy and reliability for predicting stock prices. This may involve comparing the model's predictions to the actual stock prices and calculating various performance metrics.

ML Model deployment: Deploying the model in a production setting is the last step if it performs well on the evaluation data., where it can be used to make real-time predictions of stock prices. This may involve integrating the model into a trading platform or building a web application that allows users to input their own stock ticker and receive predictions.

METHODOLOGY

Step 1: Accumulate historic inventory market statistics for the inventory or index you desire to expect. This records should consist of the date, last rate, and any other applicable records consisting of buying and selling extent. All this data are obtained from Yahoo finance website. It is one of most trusted and updated website of all time.

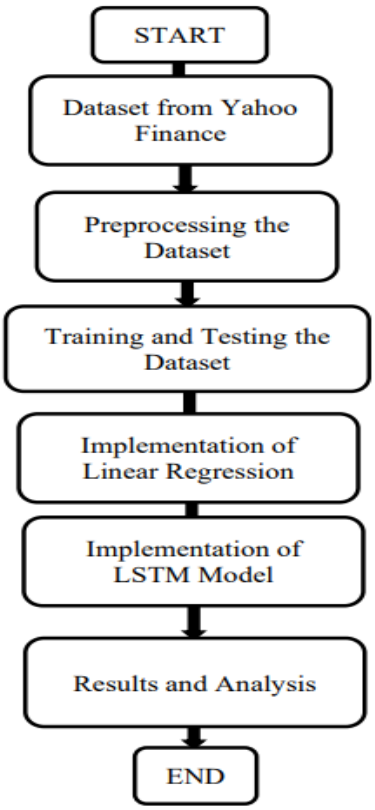

FIG 1: Methodology

Step 2: Easy and pre-process the information. This could include managing missing values, normalizing the statistics, and growing a new dataset with relevant features for the version.

Step 3: Split the dataset into schooling and checking out sets. The training set may be used to train the LSTM version, even though its system productivity may be assessed using the tested set.

Step 4: Construct the LSTM version. This can be finished the use of a library along with Keras or TensorFlow. The version ought to have several layers, such as an input layer, LSTM layer, and output layer.

Step 5: Educate the LSTM version the usage of the education set. This may require adjusting the model's hyperparameters which include the quantity of neurons within LSTM layer, batch size, and number of epochs.

Step 6: Compare the LSTM version the usage of the trying out set. This could be finished by means of comparing the model's anticipated expenses to the actual fees in the testing set.

Step 7: Use the educated LSTM version to make predictions on new, unseen records. That is a excessive stage overview of the manner, each step may also require greater distinctive consideration and extra steps relying on the complexity of the hassle, records and assets available to you.

Proposed Work

Collection Of Dataset

The dataset is the starting point in this model. It contains the information about the stock price for the recent years. For example: Date, Open value, Close value, Adjacent close, Volume of stock prices. By using the dataset only we can train the machine to give the better output for the prediction.

Dataset Filtering

Basically, the dataset is all about the Stock price details of particular company for the past recent years. Sometimes the Datasets may contain null values and also may contain duplicate values. It will decrease the accuracy of the algorithm. So,

the Process of Dataset Filtering is the process of removing the null values and duplicate rows and values in the dataset [5]. By this, accuracy of the model will be increased.

Splitting The Dataset

The complete dataset will be divided into the Training set and the Testing set in this instance. The training set is utilized to teach the machine how to get the desired results. Additionally, the testing set is utilized to determine whether the machine is producing the required or accurate results. The ideal splitting ratio for any prediction model is 30:70.

Prediction And Classification

Linear Regression

Linear regression is one of the most basic and extensively used Machine Learning algorithms. It is a method of performing predictive analysis using statistics. Linear regression generates forecasts for continuous, real, and numeric variables including sales, pay, age, and product price, among others [6]. The linear regression procedure, also referred to as linear regression, shows a linear relationship between one or more independent variables and a dependent variable. Given that regression analysis shows a linear relationship, it can be used to ascertain how the value of the dependent variable evolves in relation to the value of the independent variable [7].

A machine learning algorithm based on supervised learning is linear regression. It executes a regression operation. Regression uses independent variables to model a goal prediction value. It is primarily employed to determine how variables and forecasting relate to one another. The type of link between dependent and independent variables that multiple regression models consider, as well as the quantity of independent variables utilized, are what make them different [8]. Here we have used Linear regression for two particular companies.

Long Short-Term Memory

Long Short-Term Memory networks are referred to as the LSTM Network model. These neural networks are capable of understanding hard relationships in general. Long-term dependency issues were generally intended to be prevented by using LSTM models, and they are generally successfully prevented [9]. I will walk you through how to apply LSTMs in time-series forecasting in this article. The LSTM Network models typically include the ability to selectively remove or add data, which is controlled by a unique structure known as gates [10].

LSTM processing starts with determining what data we need to toss out of the cell. Choosing what additional data to enter in the cell is the subsequent step. Finally, we choose the product we desire. Typically, the output is influenced by the state of the cells. Let's use time series forecasting to anticipate stock prices as an example to better understand how LSTM works [11]. I firmly believe that it's essential to comprehend the mathematical operations involved in the cell's construction if one is to comprehend why LSTMs function and to gain an intuitive understanding of the numerical uncertainty that underlies the model and enables it to fit to a variety of data samples [12].

As everyone is aware, CNN accurately interprets the data using a picture. Therefore, we must first transform our auditory data into images. So, CNN is not suitable for forecasting future stock. How and what kind of data are used to train the LSTM depends on what we want it to perform; the weights will adjust themselves to best approximate the answer that we seek [13]. Characterization is a vague concept that only serves to highlight how the concealed state is more interested in the most frequent time. We did not use CNN because it is best suited for character recognition only not for time series data [14]. So, we have used Long Short-Term memory algorithm for two particular companies.

Standardization And Data Visualization

When continuous independent variables are measured at several scales, the idea of standardization becomes relevant. To standardize the range of properties of a particular input data set of stock prices, standardizing is a famous technique that is frequently used as a pre-processing step in most machine learning and

deep learning models. For optimization methods like LSTM that use stock price inputs in ML algorithms, this strategy of rescaling value with a distribution value between 0 and 1 is helpful.

Data highlights, test data analytics, and model output evaluation are only a few of the numerous analytical activities for which data visualization is crucial. These visuals are way better and easier to understand the dataset. Data visualization is significant because it makes data easier to see, interact with, and comprehend. Machine learning is being used by businesses more and more to collect vast volumes of data that can be slow and difficult to filter through, understand, and explain. This process may be sped up, and information can be presented to stakeholders and business owners in ways they can understand. No of their degree of skill, the correct visualization can put everyone on the same page, whether it is straightforward or complex. Here we have used two different company's stock price datasets for data visualization.

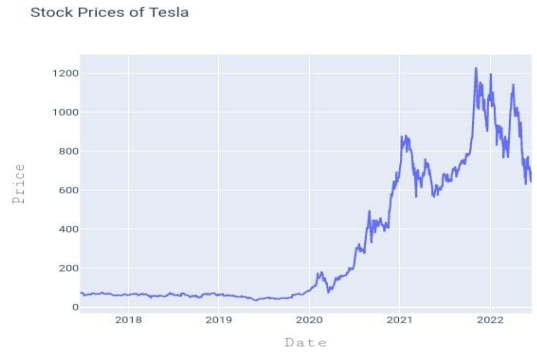

FIG 2: Stock value of tesla for the past 5 years

FIG 3: Stock value of Amazon for the past 5 years

RESULTS AND DISCUSSION

Linear Regression Output

FIG 4: Stock price prediction for Tesla by linear regression

FIG 5: Stock price prediction for Amazon by linear regression

By observing the prediction made by Linear regression, we clearly came to know that the prediction are not accurate. And we found that linear regression are suitable for only linear kind of datasets. For the dataset like stock price, weather changes dataset this linear regression algorithm is not suitable for the precise prediction.

Lstm Model Output

J Popul Ther Clin Pharmacol Vol 30(12):e130–e136; 15 May 2023.

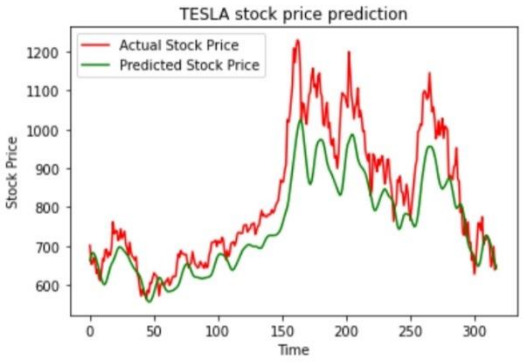

FIG 6: Stock price prediction for Amazon by LSTM model

FIG 7: Stock price prediction for Amazon by LSTM model

By observing the prediction made by LSTM model, we clearly that the predictions are mostly accurate and we obtained the desired output. With the ability of the LSTM model to store the dataset in its memory and able to train and predict with its own memory, this kind of precise prediction is possible.

CONCLUSION

In conclusion, ML machines using LSTM algorithm for stock value forecast can be a promising approach due to its ability to effectively handle time series data and capture long-term dependencies. However, it is important to note that LSTM models, like any other models, are not immune to overfitting and are sensitive to the quality and quantity of the data used for training. Additionally, evaluation of the LSTM algorithm's performance is essential, against other models and benchmarks to ensure its effectiveness in a specific market or industry. Despite these challenges, LSTM networks can provide valuable insights and predictions that can aid in making informed investment decisions. By this prediction we came to know that LSTM model is the best model to predict the future stock price with the most possible accuracy.

REFERENCES

1. Tuarob, S., Wettayakorn, P., Phetchai, P. et al. DAViS, “a unified solution for data collection, analyzation, and visualization in real-time stock market prediction.” Financ Innov 7, 56 – 2021
2. Pang, X., Zhou, Y., Wang, P. et al. “An innovative neural network approach for stock market prediction”. J Supercomput 76, - 2020
3. R. Anand, T. Shanthi, R.S. Sabeenian and S. Veni on a title "Real time noisy dataset implementation of optical character identification using CNN" in International Journal of Intelligent Enterprise – 2020
4. Gopinathan R, Durai S, “Stock market and macroeconomic variables: new evidence from India.” Financ Innov 5:12 – 2019
5. Fischer T, Krauss C, “Deep learning with long short-term memory networks for financial market predictions”. Eur J Oper Res. S0377221717310652 – 2017
6. M Qiu, Y Song and F. Akagi, “Application of artificial neural network for the prediction of stock market returns: The case of the Japanese stock market[J]”, Chaos Solitons & Fractals, vol. 85, pp. 1-7,- 2016.
7. Khaidem, L., Saha, S., & Dey, S. R. “Predicting the direction of stock market prices using random forest”. arXiv preprint arXiv:1605.00003 – 2016
8. A H Moghaddam, M H Moghaddam and M. Esfandyari, “Stock market index prediction using artificial neural network:[J]”, Journal of Economics Finance & Administrative Science, - 2016.
9. Cavalcante RC, Brasileiro RC, Souza VLF, Nobrega JP, Oliveira ALI, “Computational intelligence and financial markets: a survey and future directions.” Expert Syst Appl 55:194–211- 2016
10. M Qiu and Y. Song, “Predicting the Direction of Stock Market Index Movement Using an Optimized Artificial Neural Network Model[J]”, PLoS ONE, vol. 11, no. 5,- 2016.
11. Ding X, Zhang Y, Liu T, Duan J (eds) “Deep learning for event-driven stock prediction”. In: International conference on artificial intelligence – 2015

12. X Y Lv, S L Sun and H Liu, “Stock Price Prediction Model Based on BA Neural Network and its Applications[J]”, Advanced Materials Research, vol. 989-994, no. 9, pp. 1646-1651, - 2014.
13. R Anand, R.S Sabeenian, Deepika Gurang, R Kirthika and Shaik Rubeena on a title "AI based Music Recommendation system using Deep Learning Algorithms" in IOP Conference Series: Earth and Environmental Science in 2013.

J Popul Ther Clin Pharmacol Vol 30(12):e130–e136; 15 May 2023.