

STOCK MARKET PREDICTION USING MACHINE LEARNING

Odaiah Rachapally¹ and Appala Raju Uppala²

^{1,2}Department of ECE, Geethanjali College Engineering and Technology, Telangana, India

¹Odaiahr.ece@gcet.edu.in ²uar77.ece@gcet.edu.in (Corresponding Author)

Abstract

Our objective is to implement machine learning model to predict the stock trend. The majority of machine learning algorithms are most powerful technique to implement in predicting stock prices or trend. The objective is to predict the stock trend in order to make more efficient and accurate investment decisions. We have implemented a model for machine learning that the mathematical calculations, functions, and other external factors to achieve better stock trend accuracy and providing the profitable trades. We implemented Moving average indicator which is also a powerful indicator in technical analysis of the selected stock. Generally speaking, there are two of trading, the first one is day trading which is also called Intraday which usually exists through the day's conclusion and the second one is swing trading, short term investment or long-term investment which usually prolongs to weeks to months and sometimes years as well. The LSTM algorithm which is used in our project is considered one of the most trending and powerful algorithms which will provide promising results and it will be a hope for profitable trade. This algorithm which will provide promising results and it helps for a profitable trade. Historical data is employed in this model. The greatest asset for any retail investors or stock analysers is having LSTM based prediction model. Because it provides a great accuracy and efficient prediction for a stock. The main objective, to be precise is to find the stock trend whether it is bullish or bearish (uptrend or down trend).

Keywords — lstm, moving average, machine learning, intraday.

I. Introduction

The stock market is a vigorous and complex system where people buy and sell equities, commodities, and derivatives over online trading platforms which is supported by stockbroker. Any person can own the shares of a company being a retail investor by buying or selling the equities on markets. This stock market is a boon to the people because it will provide you a financial independence and risk-free life compared to investing in a new business or the need of high salary career.

Stock markets are dependent on many external factors like a news from the company, government policies, inflation rate around the globe, and also dependant on other countries financial markets. This factor causes instability and high volatility in the market for the better interface while trading the equities automated trading systems (ATS) are used [1]. This is implemented by the computer programs for better performance in receiving and submitting the orders.

However, to gauge and manipulate the automated trading systems (ats), the safety measures and risk-free strategies are made based upon the human judgements are required. various factors a are acquired and adopted to implement a trading strategy with ats. there will be a machine learning algorithm powerful tool to predict the stock trend of stock under analysis. time series prediction is a most popular technique which is used for predictionprocess such as weather prediction, and stock market prediction. itusually uses the historical data to forecast the outcome of a future event.

There are several algorithms for predicting time series which show very accurate and efficient results. the most common algorithm is recurrent neural network (rnn) and a algorithm named long short-term memory (lstm) which is current most powerful machine learning algorithm used in the prediction models.many researchers and financial market investors are proposedvarious models based on this algorithm. in our algorithm we used lstm to implement our model.

II. LITERATURE STUDY

The research work done by Karthikeyan C, Sahaya Anselin Nisha A, Anandan P, Prabha R, Mohan D, Vijendra Babu D of ECEand CSE departments of various colleges, India. Stock value analysis is a great examination process to analyse a particular stockthat how it reacts based on the movement of candles and historicaldata. Technical analysis has been developing significance on the financial exchange in the upcoming years. To set up a risk-free trade and strategy every investor must learn the stock examinationand analysis. In a lot of thesis papers on stock predictions, it a fewof directions show up on the future forecast. Share market is amongthe most passionate and spot of high surcharge on the planet. Primarily utilize three different ways i.e., major examination, measurable analysis and machine learning to predict the stock costof offer market but still these three yet to be insufficient data.Thus building up a prediction model is one of the difficult tasks as stockcost depends upon the various factors.

III. METHODOLOGY

The model which we selected i.e.; Stock price prediction will be a composite problem. There are numerous criteria that are to be considered which the initial data is not so analytical. But using Datapre-processing and machine learning algorithm i.e., LSTM one can develop model from previous historical data, and it can train the LSTM model based on that data to make appropriate and suitable assumptions. We need data for data pre-processing, it's called as dataset [2]. This dataset is acquired from yahoo finance website in this project. The dataset consists of our selected stock data. Data will the stock price of the selected stock between the selected timelines. Date,open, close, low, high, volume, and adjacent closed are some of the columns that make up this table. All the required data was converted into excel sheets i.e., csv format, which was initially read, pre-processed, pandas and other libraries are utilized, data frames are produced. Then the required column data is taken for further processi.e., to train the LSTM model. Following this, scaling of data needs tobe processed. It involves making roundoff for the trained data. This whole process is processed using sklearn library which helps in pre- processing.

i) Data pre-processing:

Data pre-processing is employed to sort out the large datasets to identify patterns and relationships, which involves converting raw data into a clean data which is easy for data pre-processing. This process involves checking out for the null values, plotting the basic regression models. Pre-processing refers to the transformations applied to our data before feeding it to the algorithm [3].

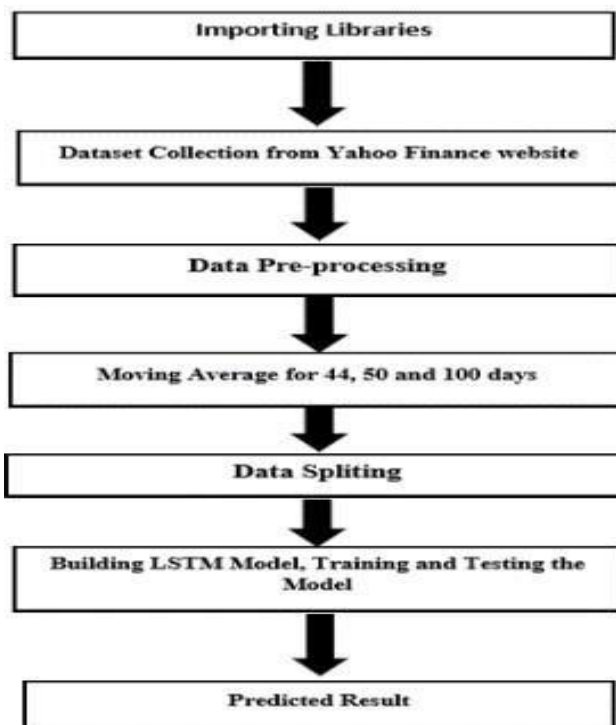


Figure 1: Flow Chart of the Model

Data Pre-processing is a technique that is used to convert the raw data into a clean and error-free data set. In other words, whenever the data is gathered from different sources it is collected in raw format which is not suitable for the analysis, that's why pre-processing is mandatory.

ii) Moving Average:

Moving average is the summation of the past period of time stock price i.e., open or close is taken and ought to be divided by period of time. The period of time is flexible, most common moving average indicator used is 44 days moving average and we also added 50 days, 100 days, and 150 days. Average of the selected from the selected timeline based on historical data which can show movements in the candles accordingly and may be applied for technical analysis. Sample Regression model is represented in figure 1.

iii) Data Splitting:

Data splitting is when data is split into two or more groups. subsets. Typically, with a two-part split, one part is employed to assess or verify the data and the other to train the model. Data splitting is significant part of data science, particularly for creating models based on data[4]. This technique helps ensure the creation of data models and processes that use data models such as machinelearning are accurate.

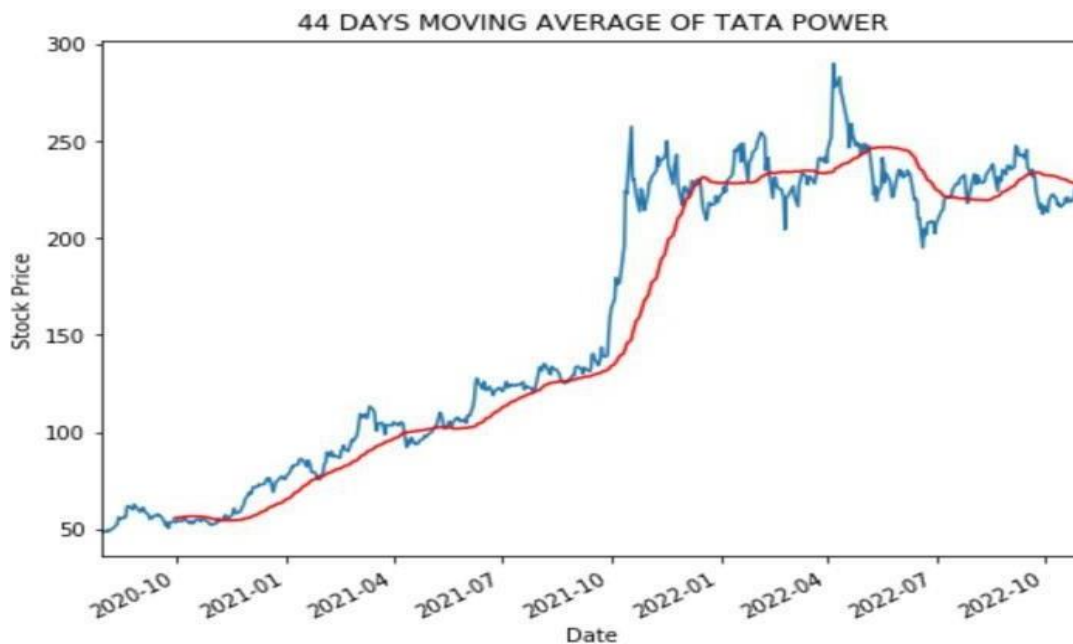


Figure 2: Regression Model of 44 days Moving average

iv) LSTM, Training and Testing: Creating a Simple LSTM with Keras:

Using Keras and Tensorflow makes recurrent type neural networks much easier to build. It's easier to build neural networks with these libraries than from bottom. The best reason to build a neural network from bottom is to examine and understand how neural networks work. In practical situations, using a library like Tensorflow is the best suitable approach. It's straightforward and simple to build a neural network with Tensorflow and Keras, let's take a look at howto use Keras that can build our LSTM.

Importing the Right Modules:

Importing the proper modules is the first thing we must do. For this example, we're going to be working with tensorflow. The bottom two imports are optional, but they speed up writing so that we don't have to enter tf.keras.layers every time we add a layer.

Adding Layers to Your Keras LSTM Model It's quite easy and comprehensible to build an LSTM in Keras. All that's really required for an LSTM neural network is that it should have LSTM cells or at least one LSTM dense layer. If we add many forms of layers and cells, we can still call our neural network an LSTM, but it would be more accurate. To build an LSTM, the first thing we're going to do is start a Sequential model. Afterwards, we'll add an LSTM layer. This is what makes this an LSTM network. Then we'll add a batch normalization layer or scaled data and a dense (fully

Train and Fit the Keras LSTM Model:

After creating the LSTM model, we need to train the model and then test the trained model. Historical data is fed as input to the training data. Training the data includes scaling of data. Scaling of data means we will be making approximate round-off of the data values. The scaled data is between 0-1. There are a few factors which are very essential in controlling the training process. Those parameters are called Hyperparameters. Hyperparameters are used to tune the training model. We also find the scaled data for this individual sets of Hyperparameters. Our primary aim to train the LSTM model is to acquire predicted data from our trained model from the inputs of test dataset. The cycle will continue until we get proper estimated values. At last, we take predicted output from trained data and inputs of test data.

In general, training and testing process of any model is initially divided into 80% - 20%. We can take variable data such as 75% - 25%. After training the data the model is compiled. The model is trained in keras, we have initialized the fit function. For using the fit function, we'll need to pass in the training data for x and y, the validation, and the

```
In [34]: # Create the training data set
# Create the scaled training data set
train_data = scaled_data[0:int(training_data_len), :]
# Split the data into x_train and y_train data sets
x_train = []
y_train = []

for i in range(60, len(train_data)):
    x_train.append(train_data[i-60:i, 0])
    y_train.append(train_data[i, 0])
    if i <= 61:
        print(x_train)
        print(y_train)
        print()

# Convert the x_train and y_train to numpy arrays
x_train, y_train = np.array(x_train), np.array(y_train)
# Reshape the data
x_train = np.reshape(x_train, (x_train.shape[0], x_train.shape[1], 1))
x_train.shape
```

epochs. For this example, we have trained the data for just 1 epoch.

IV. MODEL RESULT

The data set gone through train and test sequence and below are the results:

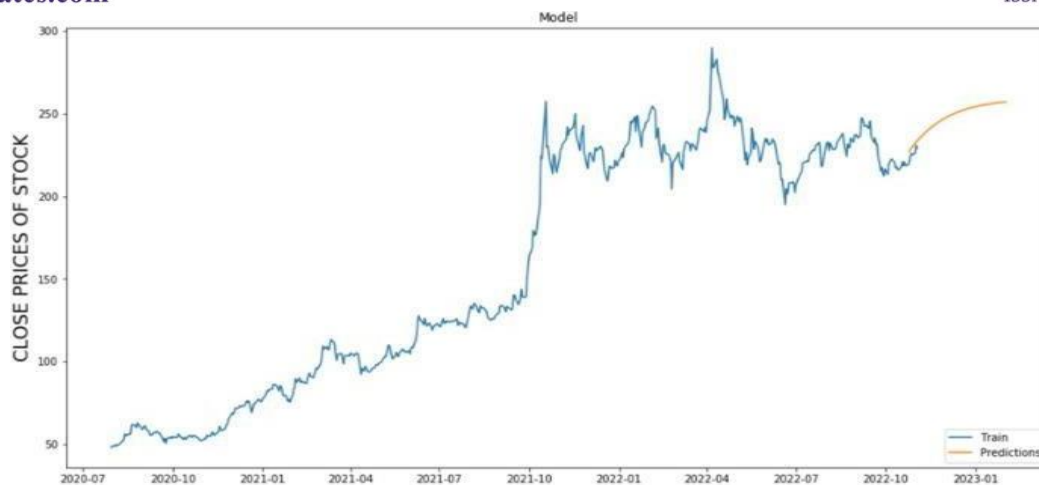


Figure3: Regression model of prediction of the stock price.

V. CONCLUSION

Two techniques are considered in this project LSTM and moving average. Moving average will be a great indicator in stock analysis[1]. The LSTM algorithm which is used in our project is considered one of the most trending and significant methods for machine learning that will provide promising results and it helps for a profitable trade. Historical data is used in this model. The greatest asset for any retail investors or stock analyzers is having LSTM based prediction model. Because it provides a great accuracy and efficient prediction for a stock. This prediction model can be utilised in the future to improve by taking large data from stock datasets [2].

REFERENCES

- [1]. S. Kavitha, S. Varuna and R. Ramya, "A Comparative Analysis on Linear Regression and Support Vector Regression," of Green Engineering and Technologies(ICGET), 2016.
- [2]. Dr.Subbarao Polamuri, "A Survey on Stock Market Prediction Using Machine Learning Techniques," DOI:[10.1007/978-981-15-1420-3_101](https://doi.org/10.1007/978-981-15-1420-3_101)), May, 2020.
- [3]. Dr. Subbarao Polamuri, "Stock Market Prediction Using Machine Learning <https://doi.org/10.1016/j.procs.2021.10.071>), 2021.
- [4]. Jagruti Hota 1 , Sujata Chakravarty 2 , Bijay K. Paikaray 3 and Harshvardhan Bhojar 4, "A Survey on Stock Market Prediction Using Machine Learning Techniques, 2022.