CS1110-AIML

Intraday/Swing Stock Predictions

PREPARED BY

Arpit Vyas
(2019BTECHCSE011)
Ankush Choudhary
(2019BTECHCSE007)

FACULTY GUIDE

Dr. Alok Kumar

Department of Computer Science engineering

Institute of Engineering and Technology (IET)

JK Lakshmipat University Jaipur

2021

Abstract

Machine learning algorithms use historical data as input to predict new output values.ML is very powerful tool and is commonly used in finance. But stock Market is a complex structure as apart from technical indicators it is run on human sentiments as well. Machine learning algorithms use historical data as input to predict new output values. Machine learning algorithms use historical data as input to predict new output values. Due to uncertainty in Stock Market, it is difficult to forecast stock price and for intraday or swing it is nearly impossible to get predictions. Still, we tried to predict the trend of stocks with different machine learning algorithm and observed their working and accuracy. We also tried to predict stock price using prophet and regression in live market for day trading and swing trades.

Table of Contents

List of Figures

1. Introduction: -

1.1 Stock Market: -

The stock market broadly refers to the collection of trades with other areas in which the acquisition, sale, and disposal of shares of publicly held companies take place. Such financial transactions are carried out through formal transactions incorporated by the institution (whether physical or electronic) or through over the counter markets operating under a set of defined rules.

We have all heard the word stock in some way. The stock is related to partners and trading companies and must be based in the marketing world. Another word used for stock is the most widely used stock in everyday life. People even time is an investment plan and something that people see as a long-term investment that protects and provides a lot of money during retirement.

Buying a company stock, you buy a small portion of it. People invest the same amount of money in the long run, in the long run, in the long run. Investments that provide long-term and relevant long-term goals. The amount of share you invest today should give you a good harvest tomorrow, but it is not the same.

The market is unpredictable as well as the resources and features taken to complete or stop the set. It has never been the same level and the same pattern still does not predict until then. A certain approach to prediction has been found and estimates and sophisticated calculations are made with the hope of the best but not all resources can be trusted and are not naturally predictable.

1.2 Problem Statement: -

We will be forecasting stock price and trend for day trading as well as swing trade using different machine learning algorithms. We will also test different algorithms to check accuracy after predicting stock price.

1.3 PURPOSE OF THE PROJECT

Stock market forecast is a forecasting system software that reflects the risks involved in investing in stock markets. It predicts stock rates and its exchange rate that allows for basic understanding and statistical analysis in front of users.

Data is considered a digital fuel that offers high aspirations and offers future goals. Information is powerful and is the same as stock. The stock is unpredictable and drastically changes its changing environment. The same ups and downs are not equal

and cannot be easily separated. Dependence on the same deals and flexible resources and agents behind you.

Investment on a financial day determines the opening of the stock market the next day. It has its own dependencies and is fully integrated with the level of finance and profitability. The stock is big and naturally busy. The main theme of the project is to predict dynamic curves and bring the method of predicting and passing the process and algorithms to complete a working resource.

Everything flows pattern. The pattern is a way out, so it is also true in stock. Stock in daily life follows pattern movements. An increase in one material may increase the price in another while lowering the price in others, Source and effect are found based on polarity which may be positive, neutral or negative flow. The correlation of the given polarity is determined and established by the effective and reliable source.

1.4 Machine learning in stock market

Stock and financial markets tend to be unpredictable and even illogical, just like the outcome of the Brexit vote or the last US elections. Due to these characteristics, financial data should be necessarily possessing a rather turbulent structure which often makes it hard to find reliable patterns. Modeling turbulent structures requires machine learning algorithms capable of finding hidden structures within the data and predict how they will affect them in the future. The most efficient methodology to achieve this is Machine Learning and Deep Learning. Deep learning can deal with complex structures easily and extract relationships that further increase the accuracy of the generated results.

2. Methods: -

2.1 Linear Regression

Linear regression may be a basic and commonly used sort of predictive analysis. The overall idea of regression is to look at two things: (1) does a group of predictor variables do an honest job in predicting an outcome (dependent) variable? (2) Which variables especially are significant predictors of the result variable, and in what way do they–indicated by the magnitude and sign of the beta estimates–impact the result variable? These regression estimates are wont to explain the connection between one variable and one or more independent variables. The simplest sort of the regression of y on x with one dependent and one experimental variable is defined by the formula y = c + b*x, where y = estimated dependent variable score, c = constant, b = parametric

statistic , and x = score on the experimental variable .
Naming the Variables. There are many names for a regression's variable . It may be called an outcome variable, criterion variable, endogenous variable, or regress and. The independent variables are often called exogenous variables, predictor variables, or regressors.
Three major uses for multivariate analysis are (1) determining the strength of predictors, (2) forecasting an impact , and (3) trend forecasting.

2.2 K-nearest neighbor-:

K-Nearest Neighbor also referred to as KNN may be a supervised learning algorithm which will be used for regression also as classification problems. Generally, it's used for classification problems in machine learning.
KNN works on a principle assuming every datum falling in almost one another is falling within the same class. In other words, it classifies a replacement datum supported similarity.
Specifically, the KNN algorithm works within the way: find a distance between a question and every one examples (variables) of knowledge , select the actual number of examples (say K) nearest to the query, then decide
• the most frequent label if using for the classification based problems, or
• the averages the label if using for regression-based problems
Therefore, the algorithm hugely depends upon the amount of K, such that
• Value of k – bigger the value of k increases confidence in the prediction.
• Decisions may be skewed if k has a very large value.

2.3 ARIMA

An autoregressive integrated moving average, or ARIMA, may be a statistical analysis model that uses statistic data to either better understand the info set or to predict future trends.
A statistical model is autoregressive if it predicts future values supported past values. For example, an ARIMA model might seek to predict a stock's future prices supported its past performance or forecast a company's earnings supported past periods.
Each component in ARIMA functions as a parameter with a typical notation. For ARIMA models, a typical notation would be ARIMA with p, d, and q, where integer values substitute for the parameters to point the sort of ARIMA model used. The parameters can be defined as:
p: the amount of lag observations within the model; also referred to as the lag order.
d: the amount of times that the raw observations are differenced; also referred to as the degree of differencing.
q: the dimensions of the moving average window; also referred to as the order of the moving average.

2.4 Prophet

Prophet may be a procedure for forecasting statistic data supported an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. It works best with statistic that have strong seasonal effects and a number of other seasons of historical data. Prophet is strong to missing data and shifts within the trend, and typically handles outliers well.

Accurate and fast.

Prophet is employed in many applications across Facebook for producing reliable forecasts for planning and goal setting. We've found it to perform better than the other approach within the majority of cases. We fit models in Stan in order that you get forecasts in only a couple of seconds.

Fully automatic.

Get an inexpensive forecast on messy data with no manual effort. Prophet is strong to outliers, missing data, and dramatic changes in some time series.

Tunable forecasts.

The Prophet procedure includes many possibilities for users to tweak and adjust forecasts. You can use human-interpretable parameters to enhance your forecast by adding your domain knowledge.

2.5 LSTM (Long-short Term Memory Network)

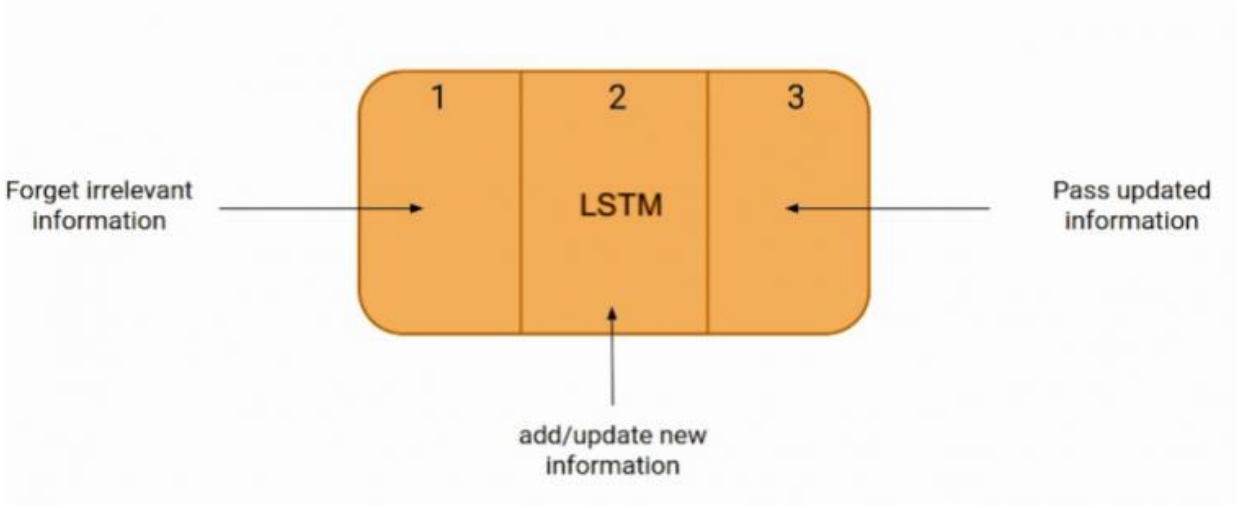
• LSTM may be a special quite recurrent neural network capable of handling long-term dependencies.

• Understand the architecture and working of an LSTM network

Long STM Network is a complicated RNN, a sequential network, that permits information to persist. It is capable of handling the vanishing gradient problem faced by RNN. A recurrent neural network is additionally referred to as RNN is employed for persistent memory.

At a high-level LSTM works considerably like an RNN cell. Here is that the internal functioning of the LSTM network. The LSTM consists of three parts, as shown within

the image below and every part performs a private function.



3. Results: -

3.1 Dataset

	Date	Open	High	Low	Last	Close	Total Trade Quantity	Turnover (Lacs)
0	2018-10-08	208.00	222.25	206.85	216.00	215.15	4642146.0	10062.83
1	2018-10-05	217.00	218.60	205.90	210.25	209.20	3519515.0	7407.06
2	2018-10-04	223.50	227.80	216.15	217.25	218.20	1728786.0	3815.79
3	2018-10-03	230.00	237.50	225.75	226.45	227.60	1708590.0	3960.27
4	2018-10-01	234.55	234.60	221.05	230.30	230.90	1534749.0	3486.05
...	...	...	...	...	...	...	...	...
1230	2013-10-14	160.85	161.45	157.70	159.30	159.45	1281419.0	2039.09
1231	2013-10-11	161.15	163.45	159.00	159.80	160.05	1880046.0	3030.76
1232	2013-10-10	156.00	160.80	155.85	160.30	160.15	3124853.0	4978.80
1233	2013-10-09	155.70	158.20	154.15	155.30	155.55	2049580.0	3204.49
1234	2013-10-08	157.00	157.80	155.20	155.80	155.80	1720413.0	2688.94

1235 rows × 8 columns

Fig.3.1.1 (data set to implement multiple machine leaning algorithms)

Fig3.1.2 Line Plot of above dataset

In Fig.5.1.1 we are taking data set we are taking dataset of 5-year data to use different algorithms to see accuracy and how the model is doing forecast.

3.2 Moving Average

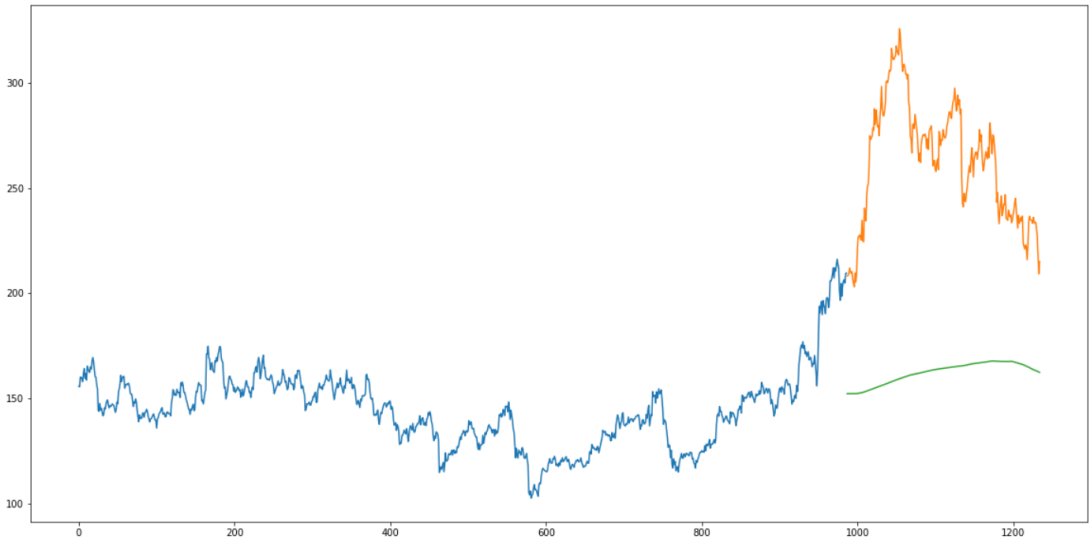


Fig 3.2.1 plot of Moving average

Green is the predicted data set this model is not giving the required result as it is clearly visible in the graph.

The RMSE value is 121.5

3.3 Linear Regression

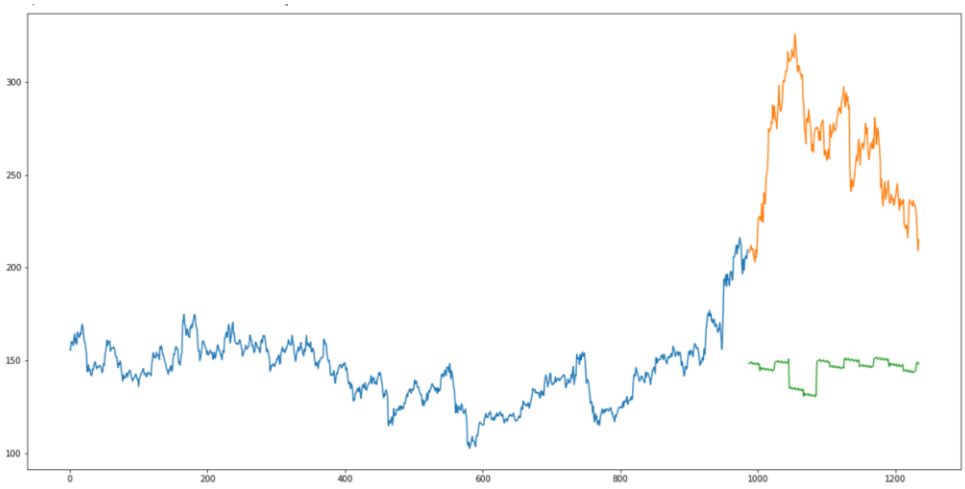


Fig 3.3.1 plot of Moving average

Green is the predicted data set this model is not giving the required result as it is clearly visible in the graph but it shows better results the moving average model.

The RMSE value is 104.5

3.4 KNN

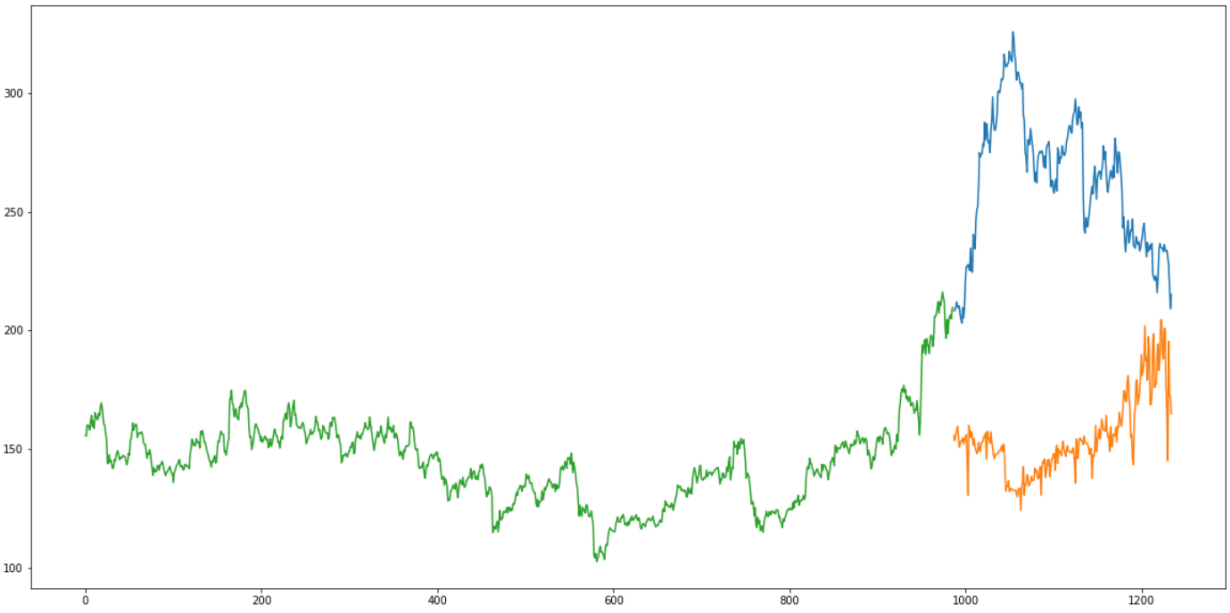


Fig 3.4.1 the plot of KNN

The orange segment shows the prediction of knn model it slightly better than all above mentioned models. The model was able to show the trend of stock moment .

The RMSE value is 115.5

3.5 Auto – ARIMA

Fig 3.5.1 the plot of Auto ARIMA

The green line segment shows the prediction of auto ARIMA which nearly accurate and trend is clearly visible in this model.

The RMSE value is 54.5

3.6 fb-prophet

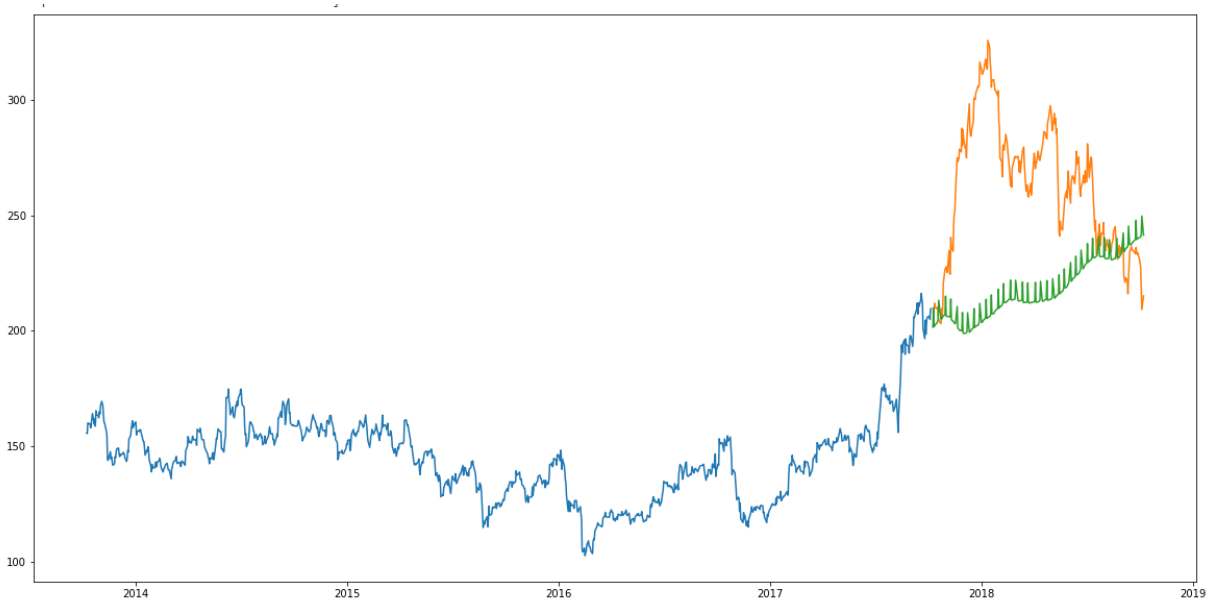


Fig 3.6.1 the plot of fb prophet

The curves and moment of stock price is more clear in prophet model.This is more reliable to use for monitoring stocktrends.

The RMSE value is 48.2

3.7 LSTM

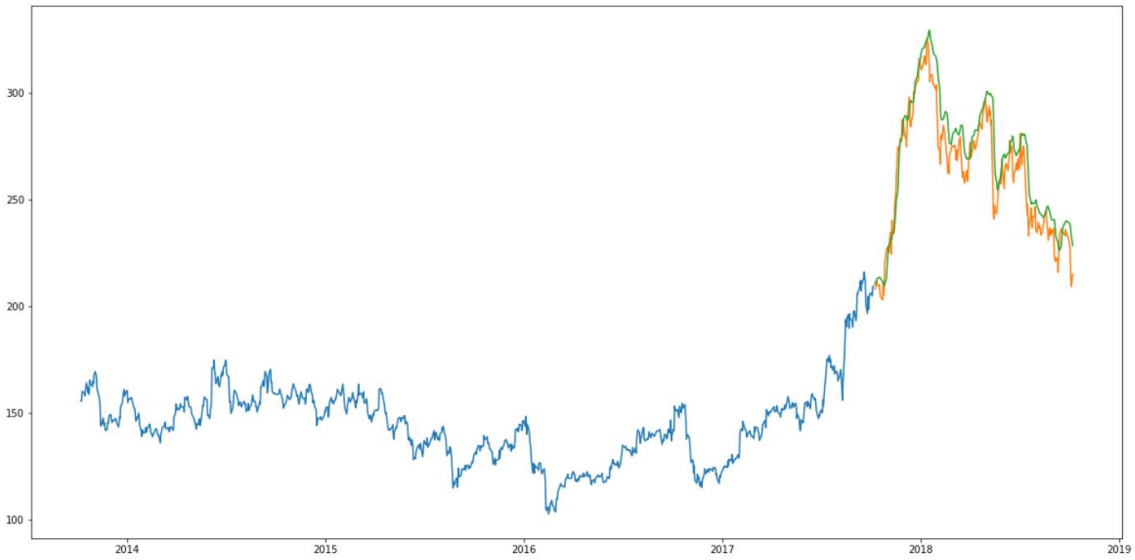


Fig 3.7.1 the plot of LSTM

It is clear that model is successfully predicting stock prices well as the trend of stock moment perfectly.

The RMSE value is 11.7

3.8 SWING trade predictions using fb-prophet

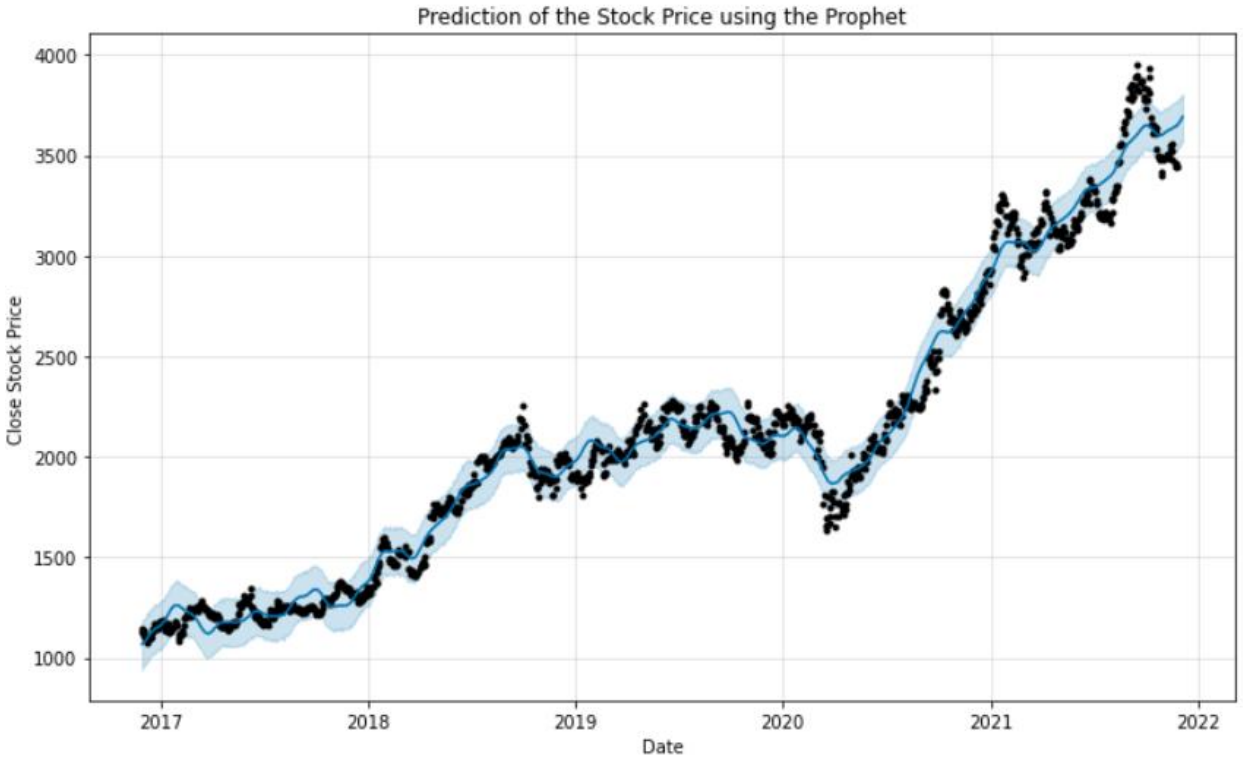


Fig.3.8.1 forecasting stock data for 10 days to estimate stock momentum for swing trades

We Successfully build the fb-prophet ml model to forecast stock price for swing trades and also checked this in live market and completed 2 successful swing trades using this model in stocks like TCS and Jai Corp.

3.9 Day trading model

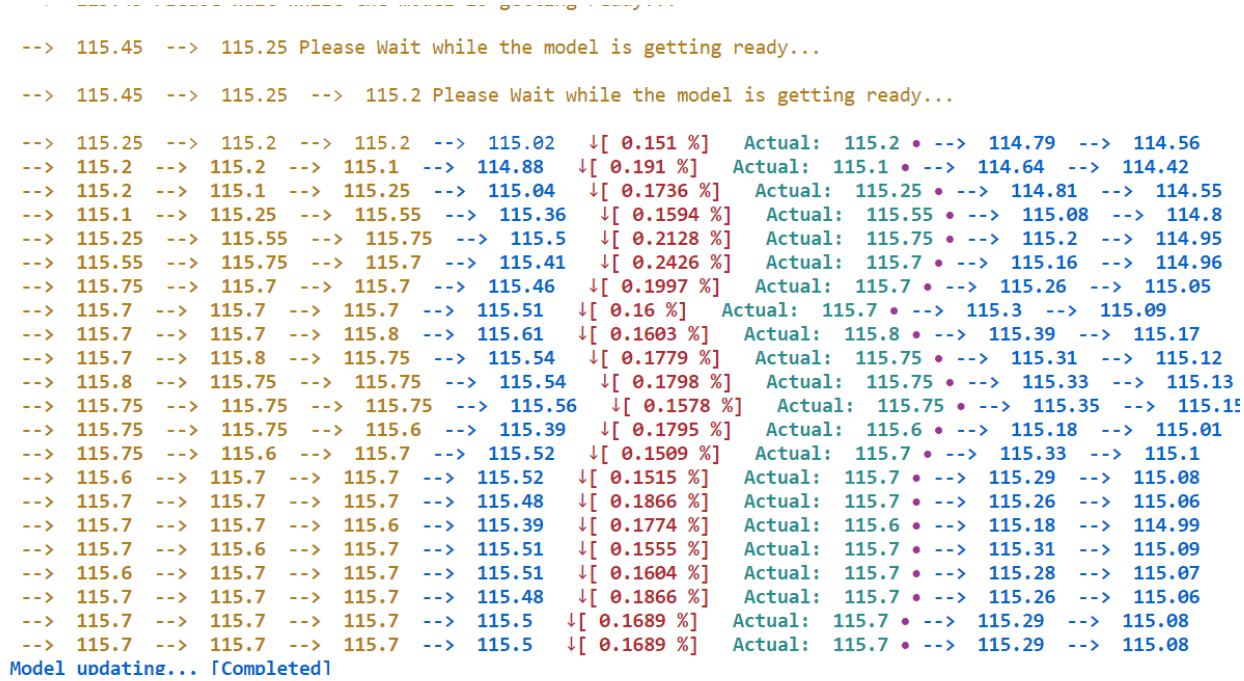


Fig 3.9.1 day trading result of machine learning model

It shows the result of the model that help predict the stock price after given time frame int this figure we time frame as 1min so the model is predicting price of stock after every minute and also with the predicted it get better and better the more we run it

4. Discussions: -

To conclude stock is an unpredictable mechanism which follows the segments of chain and the dependencies of the same are unpredictable. It is defined to be an curve which keeps on changing and turning the price from low to high and vice-versa.

It is not Possible to predict such complex financial segments with machine learning as they work on news and fundamentals of company more than the technical analysis and human sentiment cannot be predicted by machine leaning algorithms.

But still these models worked pretty good and apart from price these models help us see the trend of particular stocks.

I observed that they worked quite good on nifty-50 stocks valued more than 1400 but I don't have a statistical proof to show it but this was my 1 month observation

We checked along with many algorithms and found that LSTM fb-prophet and ARIMA were good fit for these problems.

4.1 Future scope: -

- We can use news headline to check for sentiments of the market and use those results to build a better model to forecast stock price.
- As it is difficult t predict stock price, predicting for day trading is not worth it as well as for long it is risky as it lacks fundament knowledge of stock price. So it is best to predict stock price for short term like 2 - 4month

5. References: -

[1] Md. Rafiul Hassan and Baikunth Nath, "Stock Market forecasting using Hidden Markov Model: A New Approach", Proceeding of the 2005 5th International conference on intelligent Systems Design and Application 0-7695-2286-06/05, IEEE 2005.

[2] P. Hajek, Forecasting Stock Market Trend using Prototype Generation Classifiers, WSEAS Transactions on Systems, Vol.11, No. 12, pp. 671-80, 2012.

[3] Kyoung-jae Kim, Ingoo Han. "Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index". Expert Systems with Applications, Volume 19, Issue 2, 2000, Pages 125-132, ISSN 0957-4174.

https://facebook.github.io/prophet/

https://www.analyticsvidhya.com/blog/2021/05/knn-the-distance-based-machine-learning-algorithm/

6.Appendix

Code link: -

https://github.com/arpitgits/Stock-Market-predictions-using-different-ML-algorithms-for-swing-and-day-trading