STOCK PRICE PREDICTION USING TIME SERIES ANALYSIS

[1st] Viral I Shah, [2nd] Prof. Hetal Gaudani
[1] Computer Department of G. H. Patel College of Engineering & Technology in Anand, Gujarat Technological University, India
[2] Associate Professor, Department of G. H. Patel College of Engineering & Technology in Anand, Gujarat Technological University, India

Abstract:

Stock price prediction is an important investment topic in the finance and economics. It's applied in different fields like 'medicine', 'environment', 'education', 'crime' etc. This performs vast process of stock price prediction using the ARIMA model and time series techniques. Prediction plays a very important thing in stock market business. In literature survey ARIMA model is extensively used for brief term prediction of stock value. ARIMA model has an intense potential algorithm for short-term prediction. It would be used for time series modelling data and forecasting trends in the future. First, we are going to extract area unit from daily without relying on previous information of prediction. We are going to explore varying deep learning algorithms to improve the accuracy of prediction and comparing our result with LSTM, ARIMA and Prophet Algorithms.

Keywords— Standard average and exponential moving average, Lstm, Arima, Prophet

Introduction:

Time series analysis and forecasting in statistic information finds it significance in several applications like 'business', 'stock market exchange', 'weather', 'electricity demand', 'cost and usage of product like fuels electricity'. The statement of the statistic knowledge provides the organization with necessary data that's needed for creating necessary decision. The prediction created helps the organization to create helpful selections for the expansion and development of the corporate or organization. A stock exchange, equity or the share market is the assembly of traders and sellers of the stocks that represent possession claim on businesses; these might include securities listed on a public stock market, still as stock that's solely listed in private. There are many algorithms has been work about the stock market prediction. Stock market is a critical part of economy of any country. This is a lovely feature of investment in stocks, compared to different less liquid investments like property and different immobile assets. This allows the businesses to be traded publicly, and elevate additional financial wealth for expansion by selling shares of ownership of the organization in the public sector. To issue shares for the investors to invest in the stock market, a company has to get listed to a stocks exchange through the primary market of the stock exchange they will issue the shares and find the funds for business requirements.

The Data:

In this study data was collected from NASDAQ link. Real time data of American airlines was taken in consideration for present study. NASDAQ is give the free API for the stock forecasting. There are many manually data are available like Reliance, Wipro etc. Consisting of applicants all the stock history attributes pertaining like 'date', 'time', 'open', 'close', 'high' & 'low'.

Table 1: Attributes of Stock

1	Date
2	Time
3	Open

4	Close
5	High
6	Low

Table: 2 First 10 Dataset of American airlines

Date	Open	High	Low	Close	Volume	Name
04-01-2016	41.29	41.3501	40.3	40.91	12037151	AAL
05-01-2016	41.18	41.49	39.95	40.52	10514164	AAL
06-01-2016	40.05	41.7	40	41.23	12393780	AAL
07-01-2016	40.97	41.52	39.68	40.45	11312876	AAL
08-01-2016	40.94	41.36	40.26	40.37	8853096	AAL
11-01-2016	40.56	41.2	39.9	41.08	15877511	AAL
12-01-2016	41.15	42.36	40.48	42	12635992	AAL
13-01-2016	42.17	42.74	39.66	40.1	12718577	AAL
14-01-2016	40.21	40.78	38.86	40.55	11756473	AAL
15-01-2016	39.04	39.89	38.37	38.76	18924825	AAL

Related work:

Ayodele A. Adebiyi, Aderemi O. Adewumi and Charles K. Ayo in their paper had predict the stock price with the Bayesian classification and ARIMA model for the Nokia stock market index [2]. In this paper the ARIMA model is better than compare of Bayesian classification. The standard error for the Nokia stock Bayesian [2.3736] and ARIMA model [0.7872].

San-hong Liu and Fang Zhou have performed stock price using different machine learning techniques like KNN, KNN-ANN, BP Neural network [3]. It will decrease the precision of forecast where we use the KNN. KNN-ANN algorithm use for the optimize weight. In this paper KNN-ANN is better.

Udit Agarwal and A Sai Sabitha in their paper fetch the 2500 records of stock archive data with the time series forecasting and data mining techniques. In this research paper, data reduction and pre-processing of stock data by data mining techniques.[4]

'G. Mahalakshmi', 'Dr.S.Sridevi' and 'Dr.S.Rajaram' survey on the time series data. In this paper show that many affected factors are effect on stock market sector like 'election', 'politics', and 'Economics condition'. This research paper states about all the different types of time series data that are present; how these techniques could be analysed and predicted using the various forecasting models [5]. The aggregate invariant fuzzy time series forecasting techniques RMSE 51.14.

'Poonam Somani', 'Shreyas Talele' and 'Suraj Sawant' have performed the many techniques like HMM model, neural network and SVM model [6]. This techniques predict on ICICI, SBI and IDBI bank. Compare to other model HMM model is more effectively in extracting information from the data sets. This model use for the small amount of data because the training data are increasing then their performance are decreasing.

Implementation:

"One step ahead prediction via averaging":

Averaging mechanism permits you to predict by acting future stock value as a median of the previously observed stock costs. Doing this just for one occasion using one time step can generate quite unhealthy results. You will look at two explore techniques such as; standard averaging and exponential

moving average. you may calculated by taking the square Error between verity worth at one step ahead and therefore the expected worth and averaging it over all the predictions.

Standard Average method:

You can understand the trouble of this issue by first attempting to model this is an average calculation problem. First you will see how normal averaging works.

$$x_{t+1} = 1/N \sum_{i=t-N}^{t} x_i$$

In other words, you say the prediction at t+1 is the average value of all the stock prices you observed within a window of t to t−N.

Exponential Moving Average

In the exponential moving average method, calculate xt+1xt+1 as,

- xt+1 = EMAt = γ × EMAt-1 + (1-γ) xt where EMA0 = 0 and EMA is the exponential moving average value you maintain over time.

This formula is generally used to calculate the exponential moving average from t+1t+1 time step and on the basis of that it provides one step ahead prediction. γγ shows the contribution of the latest prediction is to the Exponential Moving Average. For example, a γ=0.2 γ=0.2 gets only 20% of the present value into the EMA.

LSTM - Long Short-Term Memory

Long short term memory models are terribly powerful time-series models. This model will predict a random range of steps into the longer term. A LSTM cell (or module) has 5 essential parts useful that allows it to model for the prediction.

• "Cell state" (ctct) - This describes that the internal memory of the cell will store each short-term memories and long-term memories.

• "Hidden state" (htht) - This State is output state data/information that calculates current input, previous hidden state and current cell input that we tend to finally use to forecast the long term stock exchange rates. Additionally to the above, the hidden state can consider whether to only retrieve the short and long-term or both types of memory to be stored in the cell state for forecasting the next prediction.

• "Input gate" (itit) - It decides how much information flows from this input to the cell state.

• "Forget gate" (ftft) - It decides what proportion information flows from the present input and also the previous cell state into the present cell state.

• "Output gate" (otot) - Decides what proportion information flows from this cell state into the hidden state.

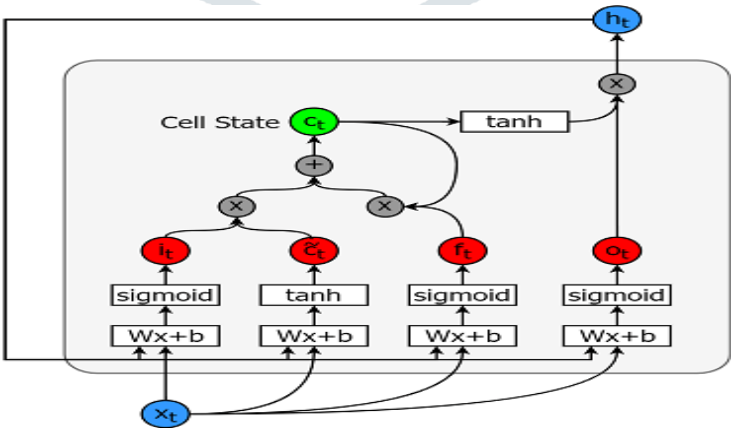

Figure 1: Essential Components

ARIMA (Autoregressive integrated moving average)

The ARIMA model was introduced by "Box and Jenkins" in 1970. It's conjointly noted as Box-Jenkins methodology that composes of set of activities for identifying and estimating ARIMA model with time series data. This model is wide utilized in monetary forecasting. ARIMA model has portrayed its economical ability to predict short term forecasts. An ARIMA display is a class of factual models for forecasting and determining time arrangement information. It expressly obliges a suite of standard structures in time arrangement information, and all things considered gives a straightforward yet amazing strategy for time series forecast. It is a speculation of the less difficult Autoregressive Moving Average and includes the idea of reconciliation. It has consistently outperformed varied difficult structural models in short-term forecasts. The AR a part of the ARIMA model shows that the perpetually evolving variable of interest is regressed on its own previous values. The MA a part of the ARIMA model shows that the regression error is really the linear combination of error terms whose values occurred throughout an equivalent amount of your time beside the assorted times within the past.

In ARIMA model, the longer term(future value) worth of a variable could also be a linear combination of past errors and past values expressed as follows;

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q}$$

Y is the actual value and ε_t is the random error at t, Φ_i and θ_j are the coefficients, p and q are integers that are often referred to as autoregressive and moving average, respectively.

PROPHET

A decomposable time series model has been depicted here, with three important primary models; 'trend', 'seasonality', and 'holidays'.

"$y(t)=g(t)+s(t)+h(t)+\varepsilon_t$"

g (t): piecewise linear or logistic growth curve for modelling non-periodic changes in data point

s (t): periodic changes (e.g. yearly/weekly seasonality)

h (t): effects of holidays with irregular schedules

εt: error term accounts for any uncommon changes not accommodated by the model

The Prophet provides intuitive parameters that square measure straightforward to tune. Indeed, even someone who desires ability in determining models can utilize this to form significant expectations for a variety of problems in a business scenario. Prophet is publish by the Facebook. This is not complicated compare to the ARIMA model. Trend is modelled by fitting a bit wise linear curve over the trend or the non-periodic a part of the time series analysis. In trend saturating growth and change points are affected in this model. Trend usually have law frequency of occurrence compare to seasonality. Seasonality to fit and forecast the effects of seasonality, in seasonality holidays and events are affect in this model. This depends on Fourier series to provide a versatile model. Seasonal effects s (t) square measure approximated by the subsequent function:

$$s(t) = \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi n t}{P}\right) + b_n \sin\left(\frac{2\pi n t}{P}\right) \right)$$

P is that the period amount (365.25 for yearly knowledge and seven (7) for weekly data)

Parameters [a1, b1… aN, bN] have to be compelled to be calculable for a given N to model seasonality.

Results:

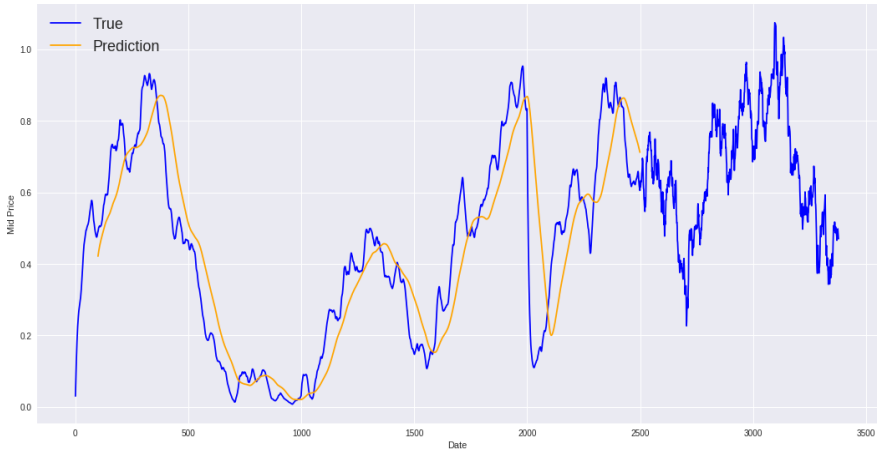

Figure 2: Actual Value and Predicted Value through standard average

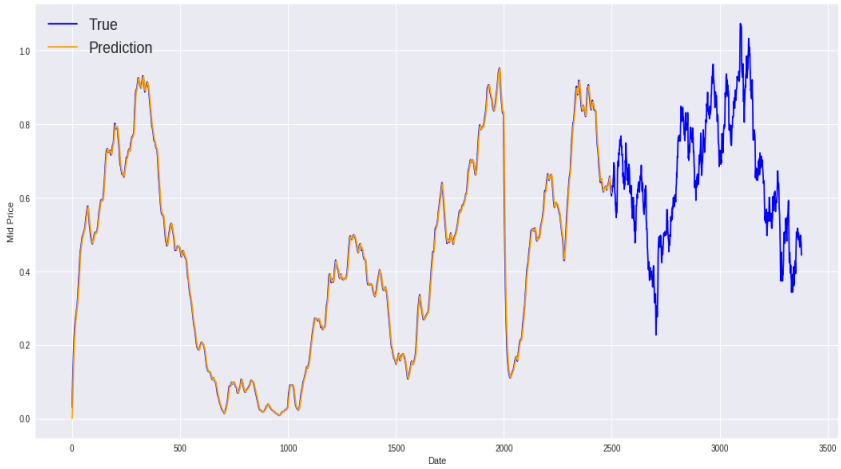

Figure 3: Actual Value and Predicted Value are similar

Table 3: Find MSE of LSTM

Step	Average Loss	Test MSE
1	5.806573	0.25331
2	3.263846	0.17737
3	1.300013	0.01815

Figure 4: Date vs Rate graph obtain prophet

Comparison of all Models

Table 4: Comparison of all Models

Models	Mean square error
LSTM	0.08508293939968078
ARIMA	0.8883784314259421
Prophet	4.158125578662469

Conclusion:

Stock market is a wide field for researchers. Lots of work has been done for predicting and forecasting price of the stock. Domain related knowledge is also required to build an effective model. From the literature survey we can analyse different machine algorithms like HMM, LSTM, ARIMA model, ANN and survey of their combination with various limitations. The work has been done on this three models like LSTM, ARIMA, and Prophet. Attributes such as Date, Time, Open, Close, High, Low, Eps and Volume work taken in consideration for analysing the data in different mention models to predict the stock price. It has been observed from the obtain results for each models accuracy that LSTM is better than other models.

References:

1. https://www.nasdaq.com/screening/company-list.aspx
2. Ayodele A. Adebiyi., Aderemi O. Adewum,Charles K. Ayo. (2014). “Stock Price Prediction Using the ARIMA Model”. *UKSim-AMSS 16th International Conference on Computer Modelling and Simulation*, (pp. 106-112).
3. San-hong Liu , Fang Zhou. (2010). “Stock Prediction Based on KNN-ANN Algorithm”. *IEEE*, 310-312.

4. Udit Agarwal, A Sai Sabitha. (2016). "Time Series Forecasting of Stock Market Index". *IEEE*, 1-6.
5. G.Mahalakshmi,Dr.S.Sridevi,Dr.S.Rajaram. (2016). "A Survey on Forecasting of Time Series Data". *IEEE*, 1-8.
6. Poonam Somani,Shreyas Talele,Suraj Sawant. (2014). "Stock Market Prediction Using Hidden Markov". *IEEE*, 89-92.
7. M.M. Mohamed, "Forecasting stock exchange movements using neural networks": empirical evidence from Kuwait, Expert Systems with Applications, vol. 27, no. 9, pp.6302–6309, 2010.
8. Mayankkumar B Patel, Sunil R Yalamalle. (2014). "Stock Price Prediction Using Artificial neural network". International Journal of Innovative Research in Science, 13755-13762.
9. ASHISH SHARMA,DINESH BHURIYA,UPENDRA SINGH. (2017). "Survey of Stock Market Prediction" Using. *International Conference on Electronics, Communication and Aerospace Technology*, 506-509.
10. Zhao, Lei ,Wang, Lin. (2015). "Price Trend Prediction of Stock Market" Using. *IEEE Fifth International Conference on Big Data and Cloud Computing*, 93-98.
11. https://www.nasdaq.com/screening/company-list.aspx
12. https://www.analyticsvidhya.com/blog/2018/05/generate-accurate-forecasts-facebook-prophet-python-r