

Real Time Stock Market Prediction Using LSTM, News Scrapper and Twitter Sentiment Analysis

Manpreet Kaur¹, Monika Kalra², Rakshya Jain³, Vaibhav Bahety⁴

^{1,2,3,4}Computer Science and Engineering Department, Chandigarh University, Mohali, India

Article Info

Volume 82

Page Number: 16750 - 16755

Publication Issue:

January - February 2020

Article History

Article Received: 18 May 2019

Revised: 14 July 2019

Accepted: 22 December 2019

Publication: 29 February 2020

Abstract

One of the hardest things to do is to guess how the stock market will do in future because prediction includes so many variables in feature matrix –physical factors ,physiological factors, gold price, crude oil price, sentiment of public ,company's performance in past etc. All these things together make share prices unpredictable and to get a high degree of precision very difficult.

I. INTRODUCTION

In the finance globe, one of the most significant operations is stock trading. We all know in our conservative society new ideas are not welcomed easily and same is case with the term which can hear everywhere that is Artificial Intelligence (AI). People say AI will take all our jobs and blame different other things but for us, we think totally opposite. It have many great application like one of them is our project which help you predict how the graph of a company in stock market will go. Whether towards increasing it stock price or decreasing it. And we don't say you can be god when you can predict future but you can definitely earn money. Github's repository is filled with many stock market predictors but none of them gives you the insights which we give that are a graph to show expected trend , a news analyser and displayer which helps you know the news about the company and last but not least Twitter's people reaction on the company. In this paper, we will concentrate on short-term stock price forecasting using stock price information from time series.

Machine Learning: Teaching a machine how to take decision which can be taken by a human mind or even those decisions which even a human mind can not take and doing it with a good accuracy and recall is known as Machine Learning. You can see it everywhere even at many toll booths Car's Number Plate is automatically read by camera or in meteorological department to predict weather of upcoming days.

Stock market: Every company wants to make profits and some companies to increase there profit and share it with common people launch there IPO i.e. Initial Public Offering and make them available to public and list themselves, so after that if company have profit its profit of its shareholder too and if have loss its loss of shareholders too so as a shareholder you must be very vigilant which companies shares to hold and which companies shares to sell. And the through which all these buying and selling happen is called stock market or also known as equity market or share market.

LSTM: Machine Learning solves a wide range of problems which we face on a regular basis or help us in easing our daily activities. But one of the problems which are considered to be most hard to

16750

solve with the help of machine learning is sequence forecast problems like what will be weather on the coming day after seeing previous 3 years weather or what will be the next word which you will type seeing on what you generally typed after that word last time or the one we are dealing with i.e. stock market prediction. In the present era seeing the current scenario it is empirically proven also that Long Short Term Memory are best in such type of forecasting problem discussed above and reason for that is recalling instances for long lengths of time and it makes LSTM brawnier than Recurrent Neural Networks and regular feed-forward neural systems . This is the reason mostly in sequence prediction LSTM is used and we are using it for the same reason. It has connection for feedback which helps in giving better prediction than others.

Regardless of its predominance, foreseeing Stock Market hasn't been possible to a great degree of prediction. If even anyone have found out the way to predict stock market he is not going to share this among all. A central objective of this paper is to understand on what factors can stock market works. The expectation is that with a more prominent comprehension of how the market moves, speculators will be better prepared to anticipate another monetary emergency. It is significant here to characterize the extent of the task. Albeit essential to any financial specialist working in reality, no endeavour is made in this project to build portfolio. But that can be the future scope of this project. This project will concentrate solely on anticipating the day by day pattern of individual stocks in terms of price. The other task will be to examine the correctness of the forecast.

II.RELATED WORKS

Ajit K R et al [1] has displayed estimating financial time series model utilizing a low multifaceted nature intermittent neural system and transformative learning approach. In the functional link artificial neural network, the concealed layers are wiped out by changing the input pattern to a high dimensional

space utilizing a lot of polynomial or trigonometric premise capacities offering to ascend to a less complex neural model. The point by point design and numerical displaying of various Fast Library for Approximate Nearest Neighbours (FLANNs) with trigonometric premise capacities have been depicted to anticipate the stock exchange return.

Amin HM [2] has proposed predicting the stock market using an artificial neural network. In this paper, numerous networks for National Association of Securities Dealers Automated Quotations (NASDAQ) list forecasting were created and approved for two information dataset(for example four earlier days and nine earlier days). Later, the upgraded organized structure for both sorts of dataset was chosen by their capacities in prediction. The model uses the values of the NASDAQ exchange price of last four and nine working days along with all days of the week as input parameters.

G. Peter Zhang [3] has proposed time series forecasting using the hybrid Auto-Regressive Integrated Moving Average (ARIMA) and neural network model. The most important and challenging task is to accomplish sensibly exact forecasts of time series. The two effective forecasting models are ARIMA and Artificial Neural Network (ANN). ARIMA is suitable for linear data generations function, while ANN is more suitable for non-linearly generated time series. The proposed word describes the hybrid forecasting method which applies ARIMA and ANN to linear and non-linear components respectively.

R. Roman J [4] informs us regarding Back propagation and Recurrent Neural Networks in financial analysis of different stock exchange returns. He proposed another technique to select between different stock exchanges for investing when the networks being regularly utilized are not extremely precise. The new proposed strategy for structuring portfolios crosswise over international stock exchange prediction has demonstrated to be successful on genuine financial exchange

information for five nations more than four years. The quality of this methodology is that it doesn't just rely on the exactness of forecasting of networks.

H EngjianJia [5] in her paper has explored the effectiveness of long short term memory networks which were trained by back propagation through time for stock price prediction. A wide scope of various architecture LSTM systems is developed, trained and tested in the paper.

Erkam G et al [6] proposed a model that evaluates the effectiveness of various neural networks. These are very effective as well as dynamic in the prediction of stock-market trends. Multi-Layer Perceptron (MLP), Dynamic Artificial Neural Network (DAN2) and the hybrid neural networks are the models that are analysed. These models use GARCH to obtain new input variables.

Hirota M et al [7] have presented a neural network model for technical analysis of the stock market. This model helps in buying and selling timing prediction for stock index.

Andrej K et al [8] has presented an analysis of representations, predictions and error types of LSTM, in order to put some light on performance and limitations of LSTM which are unknown at present.

Jian-Z W et al [9] have presented a new approach based on namely Wavelet De-noising based Back Propagation (WDBP) neural network model to predict the stock prices.

III. PROPOSED WORK

Here below it is given prediction of two companies just by using LSTM.

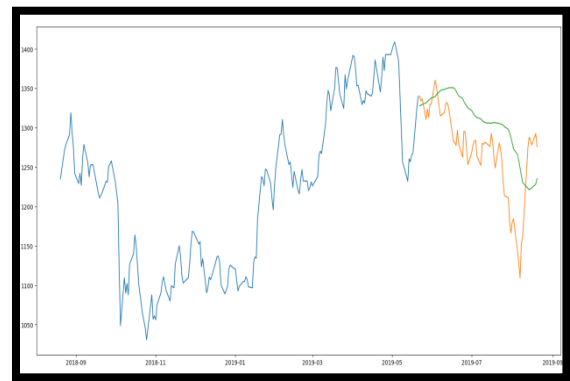


Fig1. Reliance Industry Predictions

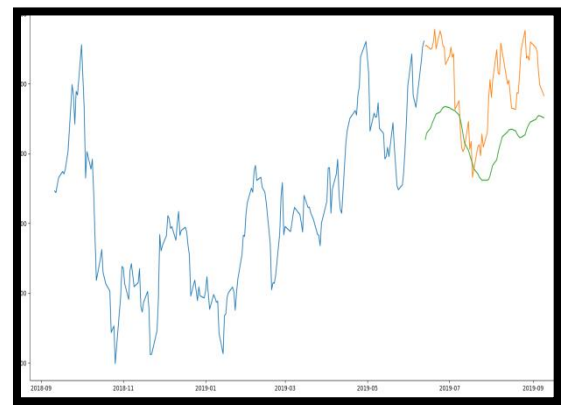
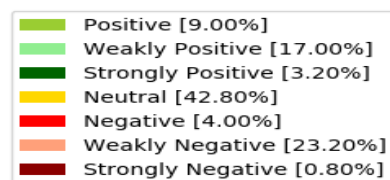


Fig2, Tata Consultancy Services Predictions

But now with help of a sentiment analysis program we will show you the relations of sentiments on twitter and share price.



How people are reacting on RIL by analyzing 500 Tweets.

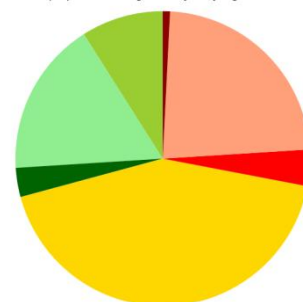


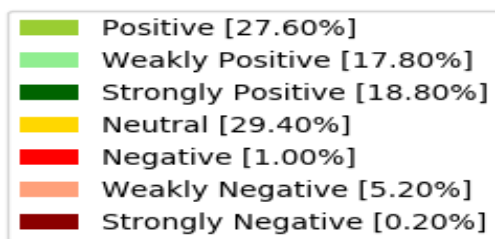
Fig3. Twitter Sentiment Analysis Of Reliance Industry

Reliance Industries Limited (RELIANCE.NS)
NSE - NSE Real Time Price. Currency in INR

1,222.20 -0.30 (-0.02%)
At close: 9 September 3:30PM IST

Fig4. Reliance Industry Share Price

What insights we can gain here from this Pie Chart of reviews for RIL and the Graph of RIL is showing downwards and the Strongly Negative views and Negative Views compared to TCS are more in per cent. So most probably on 10th of September RIL share are about to go down .



How people are reacting on tcs by analyzing 500 Tweets.

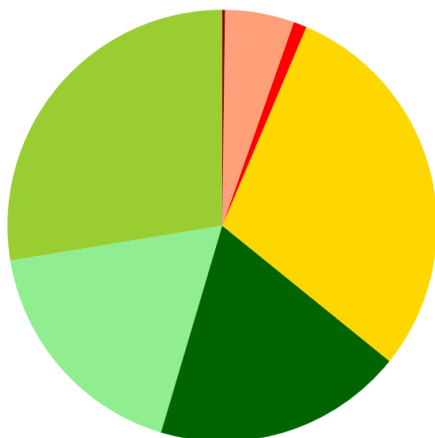


Fig5. Twitter Sentiment Analysis Of Tata Consultancy Services

Tata Consultancy Services Limited (TCS.NS)
NSE - NSE Real Time Price. Currency in INR

2,182.85 -14.65 (-0.67%)
As of 9 September 3:29PM IST. Market open.

Fig6. Tata Consultancy Services Share Price

What the insights we can gain here from this Pie Chart of reviews for TCS and the Graph of RIL is showing upwards and the Strongly Negative views and Negative Views compared to RIL are less in per cent. So most probably on 10th of September TCS share are about to go up.

Reliance Industries Limited (RELIANCE.NS)
NSE - NSE Real Time Price. Currency in INR

1,222.20 -0.30 (-0.02%)
At close: 9 September 3:30PM IST

Fig7. New Reliance Industry Share Price

Tata Consultancy Services Limited (TCS.NS)
NSE - NSE Real Time Price. Currency in INR

2,182.85 -14.65 (-0.67%)
As of 9 September 3:29PM IST. Market open.

Fig 8. New Tata Consultancy Services Share Price

One more thing which we can consider as a factor which will affect the stock market price will be news. As news generally have a huge impact on company's stock rates. For that apart from twitter sentiment analysis and LSTM we will have a third attribute in our feature matrix which we will pass to our LSTM an that we will get from our news scrapping in range between -1 to 1 like the figure given below.

Key	Type	Size	Value
01 Mar 2019	list	4	[-0.9191, -0.9191, -0.9191, -0.9191]
07 Aug 2019	list	4	[-0.4824, -0.4824, -0.4824, -0.4824]
08 Feb 2019	list	4	[-0.9364, -0.9364, -0.9364, -0.9364]
08 Jul 2019	list	4	[-0.186, -0.186, -0.186, -0.186]
09 Feb 2019	list	4	[-0.9251, -0.9251, -0.9251, -0.9251]
09 Jul 2019	list	4	[-0.186, -0.186, -0.186, -0.186]
14 Aug 2019	list	4	[0.7886, 0.7886, 0.7886, 0.7886]
25 Jul 2019	list	4	[-0.9859, -0.9859, -0.9859, -0.9859]
27 Mar 2019	list	4	[0.7825, 0.7825, 0.7825, 0.7825]
28 Mar 2019	list	4	[0.7825, 0.7825, 0.7825, 0.7825]

Fig 9. Raw News Probability

And it will convert it into meaningful input by taking the average of values and putting date in a more easily readable format by computer.

Key	Type	Size	Value
2019-02-09	float	1	-0.936
2019-02-10	float	1	-0.925
2019-03-02	float	1	-0.919
2019-03-28	float	1	0.782
2019-03-29	float	1	0.782
2019-07-09	float	1	-0.186
2019-07-10	float	1	-0.186
2019-07-26	float	1	-0.986
2019-08-08	float	1	-0.482
2019-08-15	float	1	0.789

Fig 10. Edited News Probability

NOVELTY: Our novelty in this project is our innovation that we should use the LSTM which is being used in stock market prediction a lot as it gives best result among all other algorithms or models and plus that if we add an attribute of sentiments in the feature matrix it will surely increase the performance and help us choose in a better way and with one more attribute in the feature matrix of new scrapper it will definitely boost our result.

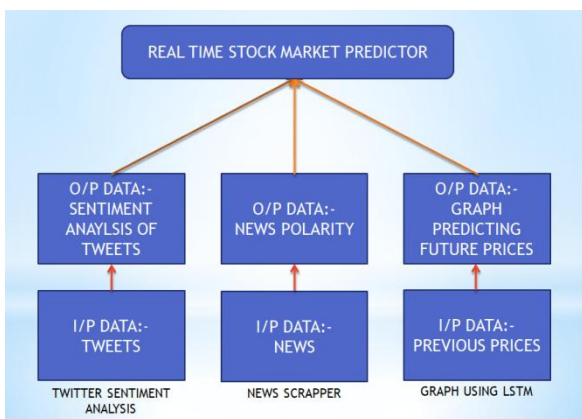


Fig 11. Basic Representation of our Innovation

Another best thing about our project is if you were offline for several days you can be updated with the news about the company as we display the news and you even don't have to apply your brain in thinking whether it is good or bad news as we give you a -1 to 1 rating where -1 is worst and 1 is best.

```

*****
-----
Date      || News Sentiment Value
-----
| 2019-10-30 || 0.82 |
-----
| 2019-08-15 || 0.925 |
-----
| 2019-08-08 || 0.633 |
-----
| 2019-07-26 || -0.976 |
-----
| 2019-07-10 || 0.763 |
-----
| 2019-07-09 || 0.697 |
-----
| 2019-03-29 || 0.851 |
-----
| 2019-03-28 || 0.821 |
-----
| 2019-03-02 || -0.76 |
-----
| 2019-02-10 || -0.899 |
-----
*****
    
```

Fig 12. Representation of dates and sentiment value

```

29 Oct 2019 https://www.businesstimes.com.sg/transport/tata-motors-shares-surge-as-china-revival-narrows-loss
14 Aug 2019 https://www.businesstimes.com.sg/technology/ibm-indias-tata-join-us-tech-platforms-governing-council
07 Aug 2019 https://www.businesstimes.com.sg/companies-markets/singapore-airlines-tata-sons-jv-vistara-to-boost-fleet-to-41-by-year-end
25 Jul 2019 https://www.businesstimes.com.sg/transport/jaguar-owner-tata-motors-has-fourth-loss-in-five-quarters-as-hurdles-mount
09 Jul 2019 https://www.businesstimes.com.sg/technology/tata-pays-500b-rupees-to-complete-sale-of-mobile-business
08 Jul 2019 https://www.businesstimes.com.sg/technology/tata-settles-us73b-of-dues-to-complete-sale-of-mobile-business
28 Mar 2019 https://www.businesstimes.com.sg/transport/tata-gic-to-buy-80b-rupee-stake-in-delhi-airport-operator
27 Mar 2019 https://www.businesstimes.com.sg/transport/tata-gic-group-buy-stake-in-delhi-airport-operator-for-80b-rupees
01 Mar 2019 https://www.businesstimes.com.sg/transport/indias-tata-motors-denies-reports-of-jlr-unit-sake-sale
    
```

Fig 13. Representation of dates with news

IV.RESULTS

Final result is that we are giving user best material so that he can choose wisely whether he want to buy the stock , sell the stock or have the stock with him.

MODULE	OUTPUT
News-scrapper	Tells him about big news related to company till the last year.
Sentiment Analyser	Tells him about mood of public for that company by analysing tweets.
Trend Predictor	Tell him expected trend according to Machine Learning(LSTM)

*NOTE- In order to check the accuracy of our LSTM made graph we compared our predicted data of 5 days with original data of 5 days and you can the trend matches remarkably.

V.CONCLUSIONS

In spite of its commonness, Stock Market expectation stays a cryptic and experimental craftsmanship. Hardly any individuals, assuming any, are happy to share what effective systems they have. A main objective of this undertaking is to add to the scholarly comprehension of securities exchange forecast. People are saying that Artificial intelligence is going to destroy job opportunities and all but think for a while why we will need jobs if everyone can be an entrepreneur. As the type of program which we have made can be corrected and as new machine learning models are coming in market every year who know we have a special model for Stock and commodities market prediction and people can freely use that and can earn money. And if they have jobs they can also take it as a side business. You will be shocked to know that in a country with such a huge population we only mere 1% of people invest in stock market which is a really small percentage as it calculates up to only 1.3

Creore and by the boom in digitalization as Paytm and many other brands have already given a platform to directly invest in stocks I think this number will increase and with increase in this number need of projects like this will also increase as there is great novelty in our project as we are considering main three factors through which you can decide whether to buy or sell this stock by taking a decision using News Scrapper, our graph and our twitter sentiment analyzer as twitter is a place of reaction of our public and we can know peoples reaction about the stock.

VI.REFRENCES

- [1] Rout A. K., Dash P. K., Dash R., and BisoiR.. "Forecasting financial time series using a low complexity recurrent neural network and evolutionary learning approach."(2015)
- [2] Moghaddam A. H., Moghaddam M. H., and Esfandyari M."Stock market index prediction using an artificial neural network." (2016).
- [3] [9] Zhang G. P. (2003). "Time series forecasting using a hybrid ARIMA and neural network model."(2003)
- [4] Roman J., and JameelA.. "Backpropagation and recurrent neural networks in financial analysis of multiple stock market returns." (1996)
- [5] JiaH.. "Investigation into the effectiveness of long short term memory networks for stock price prediction."(2016)
- [6] Guresen E., Kayakutlu G., and Daim T. U.. "Using artificial neural network models in stock market index prediction." (2011)
- [7] Mizuno H., Kosaka M., Yajima H. and Komoda N. "Application of neural network to technical analysis of stock market predic- tion." (1998)
- [8] Karpathy A., Johnson J. and Fei-FeiL.. "Visualizing and understanding recurrent networks." (2015)
- [9] Wang J. Z., Wang J. J., Zhang Z. G. and Guo S. P. "Forecasting stock indices with back propagation neural network." (2011)