

Analyzation and Visualization of Stock

**Anurag Kumar¹, Shivam Kumar², Shubham Kumar Yadav³, Vishal Kumar⁴,
Uzma Sulthana⁵**

Student, Department of Information Science and Engineering Atria Institute of Technology India^{1,2,3,4}

Assisant Professor, Department of Information Science and Engineering,

Atria Institute of Technology Bengaluru, Karnataka, India⁵

Abstract -Model used is ARIMALSTM abbreviated as Auto Regressive Moving Averages and Long-short-term Memory. This model has been collaborated with Deep learning Reinforcement to make use of its library FinRL. There are several proposed models in recent time like RNN, KNN etc. yet need to be improved and modified. Our proposed model uses mathematical approach and calculation to know the market behavior -this is the analyzation step, neural network having time-series visualization capability to present the analyzed data as form of graphs plotted by using python framework and its libraries.

Algorithms implemented such as ARIMA, LSTM for the current trends to be integrated with the past observed data. Flask, a python framework, has been included to make a reactive web page by assembling all required tools like wordpress for the database purpose and tensorflow, etc. Purpose is to analyze and visualize the stock prices graphically of Indian and International Firms over several timestamps and we have also provided additional features to know about each and every stock using Technical Indicators.

Keywords- Stock, analyzation, visualization, forecasting, prediction, Models.

I. INTRODUCTION

Machine Learning is a domain of the Artificial Intelligence which builds the concepts of predicting the new datasets which are results uses mathematical approach and calculation to know the market actions.. The contribution of the Machine Learning in the field of the Stock Visualization is incredibly significant. In today's world, people want second income from market by putting believe in the reputed firms. Investing Money in portions is a best alternative for people to make profits. Increase in the market tends to attract more investors, so it requires a platform where exchanges should be done. Exchanges in the terms of people buying company equity as company stock and selling them to take out profits. Important factor for entire stock market is the graph. We have implemented several tools (wordpress, tensorflow, etc), methodologies (Regression, Clustering, etc), algorithms (ARIMA, LSTM), python framework(flask).

II. PURPOSE

With the help of the machine learning, we have built a platform where stock prices are plotted graphically to know market movements over time frames such as 1day,1month,1year and5year.Making easier for the people to know about their stock on daily basis to prevent their losses. This visualization is effective as it has social and financial indicators. These indicators have been brought up by tweepy- "a python library" through which twitter API can be accessed. we have also implemented Prediction based algorithms to predict stock prices for upcoming timeframe as. Opening and closing of a future work. Opening and closing of a particular stock, its volume, how much fluctuation has been occurred all these parameters have been presented as a numerical chart which is dynamic in nature. A large collection of companies database has been obtained from yahoo finance.

Function:

1. Manage uninterrupted connection
2. Manage user Information
3. Manage Information
4. Manage opening and closing List
5. Manage standard deviation Results
6. Manage Financial trends
7. Manage volume change percentage
8. Technical factor detail
9. Dynamic stock price

Features:

1. Providing the smooth interface.
2. Have interactive data visualization of analyzed stocks.
3. Have collections of dynamic pages.
4. It provides the dynamic value of a list of stocks having technical values like low, high, change, change %, open, previous, etc.
5. yahoo finance a list of companies database having parameters such as NYQ (New York Stock Exchange), NMS (national Market System), etc.
6. The database comes with the csv format which can be easily modified if needed.
7. User can select a list of their interested stock and can also waitlist them.
8. Search bar is provided to choose an individual stock since list is quite bulky.
9. A single integrated page having analyzed value and its associated graph of different time frames. So that User won't find major difficulties
10. in understanding the basics.
11. Convolutional Neural network has been utilized for the accurate time series forecasting.

III. LITERATURE REVIEW

To reach final goal, various researches in the field of stock analyzation and visualization have been done implementing various algorithms, technologies, methodologies. we are going to make a glance to those works through this paper.

A. "Stock Market Prediction Using Machine Learning"

(Author-V Kranti Sai Reddy, Srinidhi Institute of Science and Technology, Hyderabad, India)

This paper deals with the scope of Machine Learning. complexity of the data requirements increasing day by day. So, it's quite not possible for every model to have pre-defined datasets. Machine Learning tries to reduce the data volume by suggesting several performing algorithms such as Support Vector Machine. Stock market Prediction is an estimate of the stock prices over different time intervals. Stockbrokers use Fundamental and Technical fundamentals are to achieve several prediction techniques. Python has been used as programming language to make such prediction of small and large capitals. Since Implementation of python libraries, framework and tools have been carried, so python is better than other languages for this job. ML based approach is designed to get the intelligence from the provided stock data and make estimations for the data visualization.

B. "The Correlation Coefficient of Stock Prices Prediction with ARIMALSTM HYBRID MODEL"

(Author- Hyeong Kyu Choi, Business Administration Korea University Seoul, Korea)

This paper illustrates that portfolio management is an essential task. For that Price Coefficient Prediction is important to correlate the two individual stock for future. LSTM (long-short term Memory) and RNN (recurrent Neural network) are applied for very same purpose. LSTM holds the Correlation Coefficient between the two individual stock and RNN capable of providing temporal dependencies. To include linear and non-linear properties in System, ARIMA model is chosen. ARIMA takes input of mixed values let's say linear and non -linear and passes the filtered output as input to LSTM. Their combined functionalities come out to be more efficient and more superior than several other integrated financial models such as Constant correlation, single indexed and multi indexed, etc. So, considering to implement ARIMALSTM model approach worth quite a lot.

C. "Deep Learning for stock exchange Prediction Using Technical Indicators and Financial News Article"

(Author-Manuel R. Vargas, Carlos E. M. dos Anjos, Gustavo L G. Bichara, Alexandre G. Evsukoff)

Deep learning has use case for analyzing market movements on daily basis. Such deep learning models takes financial news title and technical indicator as inputs. Input comprises of different sets of technical indicators. comparison is made between such sets. Two algorithms namely CNN and RNN are priority based. Former is used for catching the semantics whereas the latter is for catching context information. CNN particularly very helpful in the technical indicator's comparisons but financial news articles serve as input for RNN. There are varieties of technical indicators forming a sets namely stochastic(%K), stochastic(%D), Rate of change, William(%D) and on balance volume, Bollinger bands, relative strength index, exponential moving averages, etc.

D. "It explained about stock prices prediction from financial news articles using some analysis"

(Authors-Shilpa Gite, Hrituja Khatavkar, Ketan Kotecha, Shilpi Srivastava, Priyam Maheshwari and Neerav Pandey)

Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune. The overall description of the paper is that Stock Market is volatile as it changes with sentiments following articles, financial news and media releases. Some classification is required to prevent these improper and illegal market influences. Models and techniques are presented to first investigate such market polarizing factors. Situational procedure has been done to test these models preventing them from being manipulated from external sources. SVM, Naive Bayes, nearest Neighbors are such type of algorithms. SVM comes out to be more effective as it filters out the unofficial and unregular sentimental trends while classifying these polarized input as it yields more accurate, more precise outcomes.

E. "Hybrid ARIMABPNN Model for Time Series Prediction of the Chinese Stock Market"

(Authors-Li Xiong, Yue Lu)

Time-series patterns are complex behind stock analysis and visualization. Linear time-series are easier as they suggest moving averages over equal time intervals. For non-linear tendencies of time frames, there are algorithms which are quite useful like ARIMA, BPNN, etc. These algorithms are integrated with the Neural Network to classify the nature of dynamic graphs. A particular time-frame data is sliced over significant time to train these models. Nature of the model is hybrid because one algorithm or model will be insufficient to layout different time-series. ARIMA is compatible for Linear and BPNN for non-linear time-series forecasting.

F. "Indian stock market prediction using artificial neural networks on tick data"

(Author-Dharmaraja Selvamuthu, Vineet Kumar and Abhishek Mishra Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India)

The work of the paper is briefing the importance of the stock. In recent times, the stock sector has directly or indirectly impacted people. Stock of companies can be traded at a reasonable price agreed by that firm, so it is important to predict the stock prices. Price prediction is not an easy task but with the improvement in applied approach, approximate estimation can be made. Using these predictions, people are making a fair amount of money. Event representation in terms of syntactic and semantic should be learnt. These methods represent the Script Event Prediction. Such prediction shows the reality of the event participants, so this way of making predictions is more accurate than sentiment extraction from raw data. Raw data-based prediction lacks common knowledge, so the aim of this paper is to leverage such uncommon external factors.

G. "Forecasting the Stock Market Index Using Artificial Intelligence Techniques"

(Author- Lufuno Ronald Marwala A, Johannesburg)

The summary of the paper is that Efficient Market Hypothesis is the innovative step in the proximity of forecasting and prediction of the stock. Using previous or historical data as a basis for a satisfactory output is not enough. Some proposed techniques for the applied approach are neural network, support vector machine and neuro-fuzzy system. Random walk is the key feature of the volatile market. Artificial intelligence has a set of tools which are major contributors in the time-series forecasting. Machine Learning along with deep learning and Artificial Intelligence, are the more advanced approaches for any type of hypothesis, forecasting and prediction. Two disciplines of AI served as quality-based are Moving averages and random walk. Random walk outnumbered the other models as it is good enough for both linear and non-linear results. Closing prices of the Johannesburg stock index are fed as input to the experiments performed. Frequency and cost of continuous trading help in analyzing the market movement.

H. "A Deep Reinforcement Learning Library for Automated Stock Trading in Quantitative Finance":

(Author-Xiao-Yang Liu¹ Hongyang Yang, Qian Chen⁴, Runjia Zhang, Liuqing Yang, Bowen Xiao, Christina Dan)

Focus of this survey is on Quantitative Finance. Proposed model is DRL (Deep Reinforcement Learning). It is quite basic and easy to learn for beginners. The important factor or aspect for this approach is training. Training of DRL increases accuracy rate at which a user can set trading cost, trading time. This initializes a process which allows the people to set their own strategy. Market and trading constraints such as liquidity, cost, investor degree of investment provides enhancement in the DRL library FinRL. Key features of the DRL are that it delivers time level granularity, etc. an environment where various stocks types (NASDAQ-100, DJIA, S&P 500, HSI, SSE 50, and CSI 300) can be bought or

sold. Layered architecture-based algorithms help in debugging workloads and provides reproducibility. single-trading, multiple-trading, portfolio management are customer key feature interface.

IV. DESIGN OVERVIEW

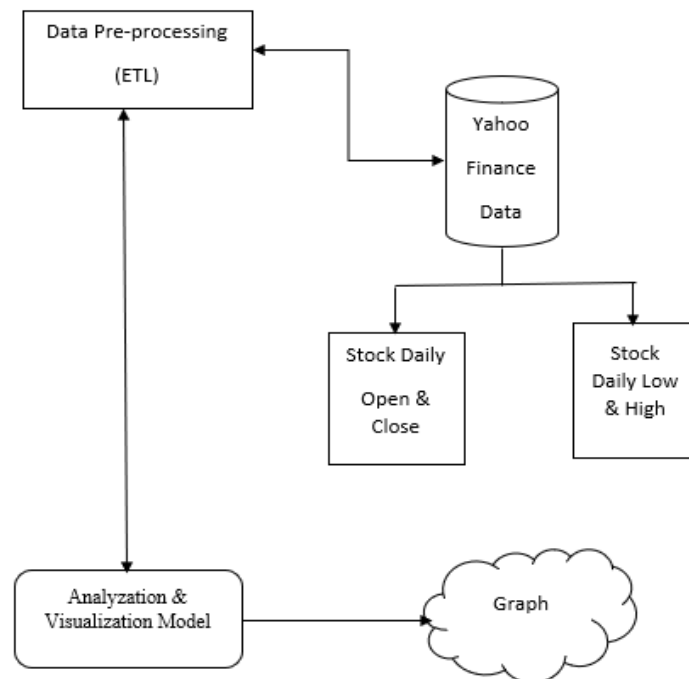


Fig 4.1

Here, we are taking inputs from yahoo finance dynamically. Several preprocessing steps involving filtration of Datasets obtained. ARIMA model organize the raw datasets and reduces the volume of datasets by considering Only the recent datasets required enough to train the model. This residual output are then passes to LSTM Whose efficiency is good then it is integrated with sentimental analysis. Which are important external factors Because it considers the market constraints and event representation. Several times of cyclic operation is formed as required to customize for starting the analyzation phase.

Model train data sent to LSTM model and test data for prediction. combination of these outputs is again sent back to LSTM and thento ARIMA model which has the has libraries especially used for plotting purpose. Finally, graph is generated. DSL enriched library namely FinRL has tools which helps in the multiprocessing. Sentimental analysis required as it affects the overall stock Market because it has sorts of technical factors and companies' meetings, new roadmap news which attract customers People invest increases and decreases which is sole reason for market fluctuations.

V. SYSTEM ARCHITECHTURE

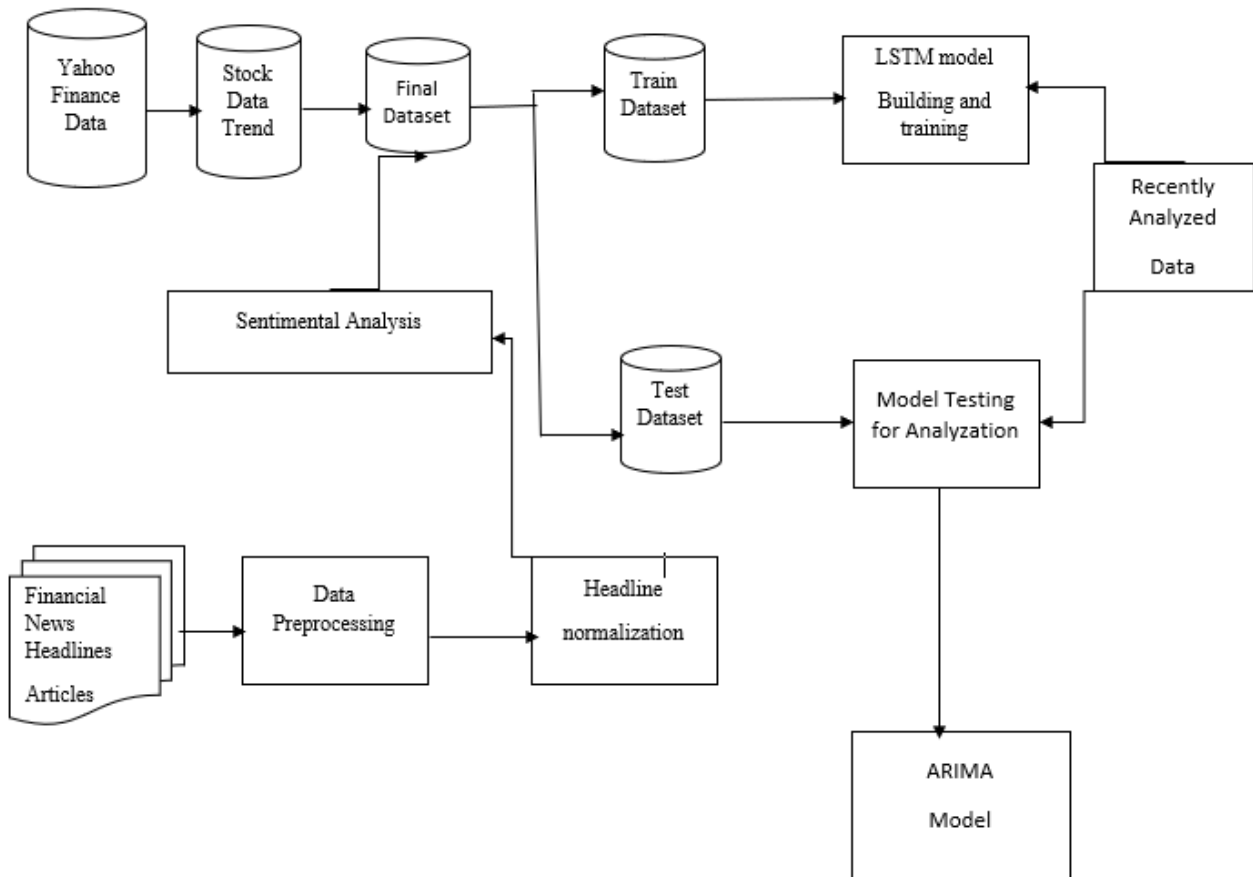


Fig 4.2

VII. IMPLEMENTATION

Here are the steps included which has increases the accuracy to produce a satisfactorily result.

1. Datasets has been taken from yahoo finance.
2. These datasets have been preprocessed and integrated with financial new articles and technical articles to follow the market random walk.
3. Test and Train model is then implemented to know data behavior to reduce datasets and keep only required sets.
4. ARIMA model has produced minimum squared results
1. Which is residual for the LSTM.
5. LSTM keeps on analyzing the datasets to get the nearest possible values.
6. Again, ARIMA uses this result to plot the time-series graph.

VIII. REQUIREMENTS

1. Wordpress
2. Flask
3. Tenserflow
4. Keras
5. Yahoo Finance
6. Alphavantage
7. Scikit- Learn
8. Tweepy
9. Python
10. Php
11. HTML

- 12. CSS
- 13. JavaScript

IX. ALGORITHM

Algorithm 1 Finding optimal ARIMA model

```

1: procedure FINDOPTIMALARIMA
2:    $aic \leftarrow inf$ 
3:   for  $p \leftarrow 0$  to 3 do
4:     for  $d \leftarrow 0$  to 2 do
5:       for  $q \leftarrow 0$  to 3 do
6:          $model \leftarrow fit(arima(p, d, q, allow\_drift \leftarrow True,$ 
7:            $allow\_mean \leftarrow True), x)$ 
8:          $aic\_curr \leftarrow compute\_AIC(model)$ 
9:         if  $aic\_curr < aic$  then
10:           $model\_opt \leftarrow model$ 
11:           $aic \leftarrow aic\_curr$ 
12:   return  $model\_opt$ 
    
```

X. RESULTS

Snapshot1:

Name	Value	Change	Chg%	Open	High	Low	Prev
AAPL	131.56 ^U	1.50	1.15%	130.07	133.08	129.81	130.06
GOOGL	2142.87 ^U	22.20	1.05%	2120.67	2173.99	2100.92	2120.67
AMZN	106.22 ^U	2.56	2.47%	102.80	106.98	102.51	103.66
HDFCBANK	1290.10 ^U	8.95	0.70%	1275.00	1297.45	1271.75	1281.15
TATAMOTORS	389.15 ^D	-3.85	-0.98%	388.50	394.35	382.55	393.00
ICICIBANK	688.10 ^U	9.70	1.43%	675.00	692.45	670.35	678.40
MSFT	247.65 ^U	2.68	1.09%	244.70	250.50	244.03	244.97
JPM	113.03 ^D	-0.40	-0.35%	113.54	114.63	112.16	113.43
META	163.74 ^U	2.87	1.78%	161.68	165.90	159.90	160.87
TWTR	37.78 ^U	0.42	1.12%	37.32	38.05	37.02	37.36
ZEEL	215.25 ^U	1.55	0.73%	213.70	217.80	209.70	213.70
NTPC	140.45 ^D	-2.60	-1.82%	140.20	144.40	139.55	143.05
COALINDIA	182.70 ^U	2.65	1.47%	180.00	184.75	179.40	180.05
YESBANK	12.45 ^D	-0.06	-0.48%	12.57	12.67	12.35	12.50

Fig. 1

Snapshot2:

TICKER	LAST	CHG %	CHG	TECHNICAL RATING	VOL	VOLUME/PRICE	MKT CAP	P/E	EPS (TTM)	EMPLOYEES	SECTOR
NATIONSD	5918.05	-4.87%	-302.95	Sell	289	1.71M	122.505B	481.17	12.30	-	Finance
AETHER	763.15	-2.17%	-16.95	Strong Sell	9.488K	7.241M	95.721B	-	-	-	Process Industries
KAMAHOLD	10400.00	6.39%	40.50	Sell	171	1.778M	66.867B	7.93	1310.96	3	Process Industries
LLOYDSME	132.25	-0.86%	-1.15	Sell	43.058K	5.694M	58.633B	46.98	2.92	-	Non-Energy Minerals
EURKAFORBE	293.00	-1.33%	-3.95	Strong Sell	17.709K	5.189M	57.076B	-	-	-	Consumer Durables
SINDHUTRAD	34.30	4.10%	1.35	Sell	98.966K	3.395M	52.888B	-	-0.50	-	Transportation
EKI	7041.95	0.60%	42.00	Sell	3.725K	26.231M	48.531B	12.63	557.43	-	Utilities
APOLLOTRI	772.90	-4.72%	-38.25	Strong Sell	43.94K	33.961M	46.997B	33.64	22.97	-	Non-Energy Minerals
IRILON	372.40	-2.49%	-9.50	Sell	4.908K	1.828M	33.533B	30.29	12.30	-	Process Industries
PARADEEP	38.90	-3.23%	-1.30	Strong Sell	344.358K	13.396M	31.684B	-	-	-	Process Industries

Fig2

Snapshot3:



Fig. 3

Snapshot4:

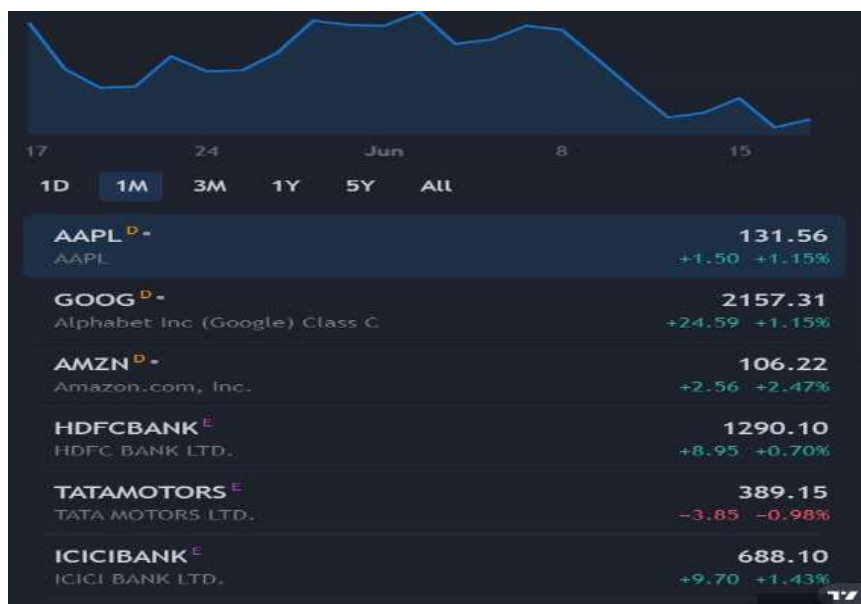


Fig. 4

Snapshot5:

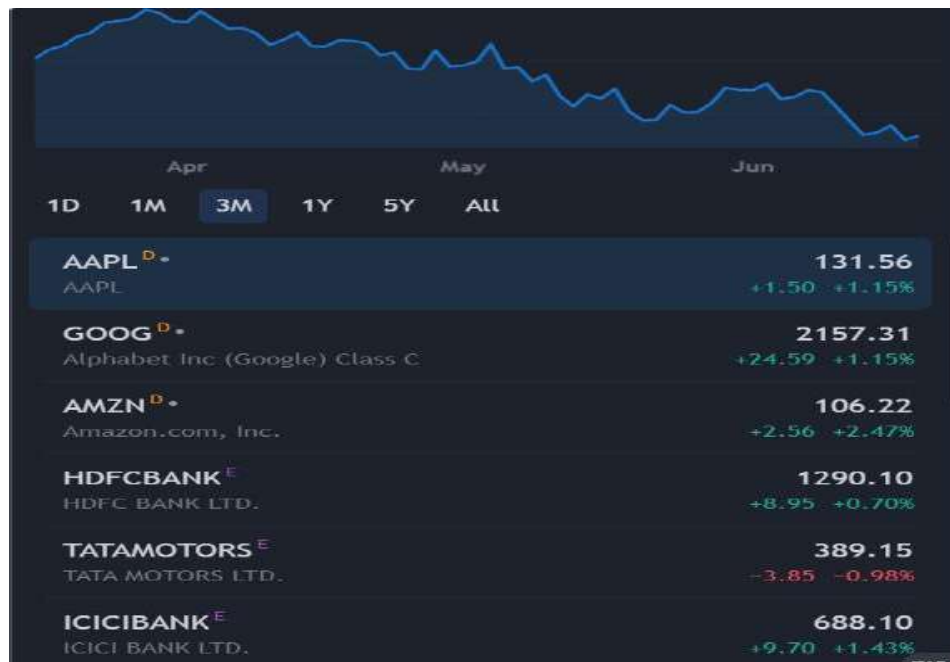


Fig 5

Snapshot6:



Fig 6

XI. CONCLUSION & FUTURE WORK

1. We have proposed a model to make some improvement in the existing models.
2. Stock Market Analyzation and visualization Web App is designed on the basis of the Sentiment Analysis and Tweet API.
3. Flask and Wordpress serves the basis of Web page Front End.

4. Algorithms used for the prediction are LSTM, ARIMA, Linear Regression.
5. Stock is predicted as requested from user which falls under NASDAQ or NSE.
6. Previous datasets are used to plot the time-series graph.
7. Our future work is to include minute frequencies for better prediction results.

XII. REFERENCES

- [1] Ajith A., Baikunth N. & Mahanti P. K. (2008). Hybrid intelligent systems for stock market analysis. In proceedings Hybrid intelligent systems for stock market analysis in proceedings of international conference on computer science.
- [2] Ajith A., Sajith N., & Sarathchandran P. (2011). Modelling chaotic behaviour of stock indices using Intelligent Paradigms. Neural, Parallel & Scientific Computations Archive.
- [3] Andreou A. S., Neocleous C., Schizas C. N., & Toumpouris C. (2015). Testing the predictability of the Cyprus Stock Exchange: The case of an emerging market. Proceedings of the International Joint Conference on Neural Networks.
- [4] Atiya A., Noha Talaat & Samir Shaheen, (2020). An efficient stock market forecasting model using neural networks neural. In proceedings of the IEEE international conference on neural networks.
- [5] Stock Price Prediction Using LSTM on Indian Share Market by Achyut Ghosh, Soumik Bose¹, Giridhar Maji, Narayan C. Debnath, Soumya Sen.
- [6] Murtaza Roondiwala, Harshal Patel, Shraddha Varma, "Predicting Stock Prices Using LSTM" in Undergraduate Engineering Students, Department of Information Technology, Mumbai University, 2015.
- [7] Ishita Parmar, Navanshu Agarwal, Sheirsh Saxena, Ridam Arora, Shikhin Gupta, Himanshu Dhiman, Lokesh Chouhan Department of Computer Science and Engineering National Institute of Technology, Hamirpur – 177005, INDIA - Stock Market Prediction Using Machine Learning.
- [8] Dharmaraja Selvamuthu , Vineet Kumar and Abhishek Mishra Department of Mathematics, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India - Indian stock market prediction using artificial neural networks on tick data.
- [9] Huicheng Liu Department of Electrical and Computer Engineering Queen's University, Canada - Leveraging Financial News for Stock Trend Prediction with Attention-Based Recurrent Neural Network.
- [10] Hyeong Kyu Choi, B.A Student Dept. of Business Administration Korea University Seoul, Korea = Stock Price Correlation Coefficient Prediction with ARIMA-LSTM Hybrid Model.