



A Novel Web Application for Stock Market Analysis and Prediction Using Machine Learning Methods.

¹Pradeep B M, ²Prasanna G

¹Assistant Professor, ²Assistant Professor

¹Department of Computer Science, ²Department of Computer Science

¹GMIT, Mandya, India, ²GMIT, Mandya, India.

Abstract: Stock Market Analysis and Prediction (SMAP) is a web-based application able to predict the stock prices of companies based on their market values and news sentiments surrounding the company. It is a portal where general stock market enthusiasts can keep track of their invested companies and are also able to instantly contact their brokers for purchases or sales of the stocks. The main application of this system however would be to predict the market values. Along with that, it has the features of a news portal and a general stock-related chatbot. Deep Neural Network (DNN), is used for stock market analysis and prediction. The algorithm's main goal is to learn market trends by training with past data and predicting future value. The calculated values of the computational analysis i.e., prediction is used to display nearly accurate result.

Index Terms – SMAP, News portal and general stock-related chatbot, Deep Neural Network (DNN).

I. INTRODUCTION

Stock analysis is the evaluation of a particular trading instrument, an investment sector, or the market as a whole. Stock analysts attempt to determine the future activity of an instrument, sector, or market. Stock market prediction is the act of trying to determine the future value of a company stock or other financial instrument traded on an exchange. The successful prediction of a stock's future price could yield significant profit. The paper entitled "Stock market Prediction and Analysis" is a web-based application. It predicts or forecasts the future of the stock market based on historical time series data. NEPSE historical time series data were scraped using scrapy and stored. Machine learning models for time series forecasting were used to train those historical data and the result is visualized on the web page for easy understanding and analysis of the stock market. In short, the system accepts the historical data set of the company which is processed on our local server and the result is displayed on a web browser. Since it is a web application, it can be accessible to everybody through the medium of the internet when it is hosted on a particular domain.

The advent of Machine Learning and its robust algorithms, the latest market analysis, and Stock Market Prediction developments have started incorporating such techniques in understanding the stock market data. In short, Machine Learning Algorithms are being used widely by many organizations in analyzing and predicting stock values. Machine Learning can be used to predict future market trends based on past historical data and can suggest to the users the general trend the market is about to follow. It can help traders keep track of their transactions in the market without any extra effort. It can also provide meaningful insights about those transactions and give a risk report of the portfolio. AI can also analyze stocks using technical and fundamental indicators and recommend favorable stocks to the user. Besides predicting stock market trends, AI can also predict general sectoral trends and can give a list of emerging sectors that could potentially become big in the future. AI can also be programmed to invest in new sectors where the investor does not have any knowledge or experience.

Some machine learning software's dedicated to speed trading i.e., to grab the opportunities at the earliest and to book a profit. AI and the trading community to increase earnings by scanning markets to locate optimal trading opportunities. Stock Price Prediction using machine learning helps you discover the future value of company stock and other financial assets traded on an exchange. There are other factors involved in the prediction, such as physical and psychological factors, rational and irrational behavior, and so on. All these factors combine to make share prices dynamic and volatile. This makes it very difficult to predict stock prices with high accuracy. The motivated idea is that, if we know all information about today's stock trading (of all specific traders), the price is predictable. Stock market prediction aims to determine the future movement of the stock value of a financial

exchange. The accurate prediction of share price movement will lead to more profit investors can make. The main objective is:

- To predict the future value of company stock and to analyze the current state of the market.
- To identify factors affecting the stock market.
- To visualize the share market with the help of interactive charts.

The scope of this project includes Stock Market Analysis and Prediction will be able to show live market status and classification of the polarity of financial news useful for new investors to invest in the stock market.

II. RELATED WORKS

1. Support Vector Machine

A Support Vector Machine (SVM) is a discriminative classifier that is formally defined by the separating hyperplane. In other words, in the given labeled training data (supervised learning), the algorithm outputs the optimal hyperplane which categorizes new examples. In the two-dimensional space, this hyperplane is a line dividing a plane into two parts where each class lay on either side.

Support Vector Machine (SVM) is considered to be one of the most suitable algorithms available for time series prediction. The supervised algorithm can be used in both, regression and classification. The SVM involves plotting data as a point in the space of n dimensions. These dimensions are the attributes that are plotted on particular coordinates. The SVM algorithm draws a boundary over the data set called as the hyper-plane, which separates the data into two classes as shown in Fig 1.

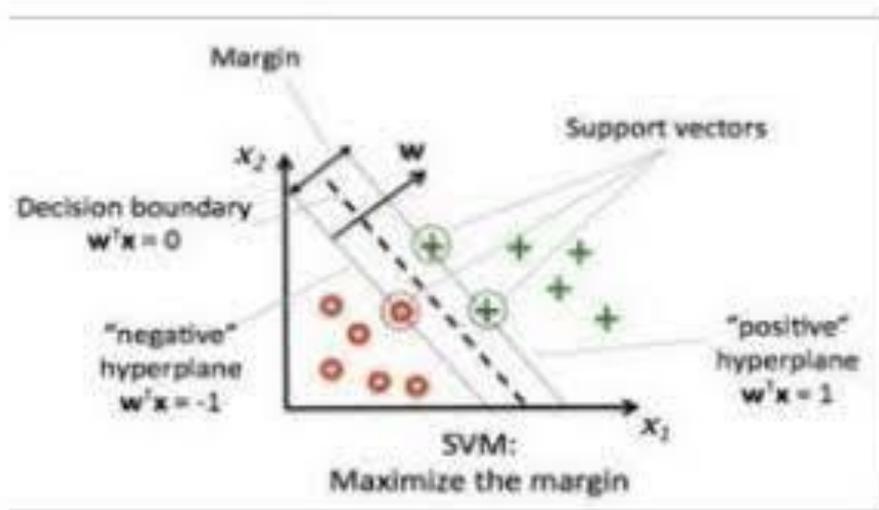


Fig -1: The Support Vector Machine Decision-Making Boundary.

2. Radial Basis Function (RBF)

In machine learning, the radial basis function kernel, or RBF kernel, is a popular kernel function used in the various kernelized learning algorithms. In particular, it is most commonly used in support vector machine classification. A radial basis function is a real-valued function whose value depends only on the distance from the origin, so that; alternatively on the distance from some other point, called a center, so that. Any function which satisfies the property is a radial function.

RBF = Local Response Function

The RBF Kernel is nothing more than a low-band pass filter, which is well known in Signal Processing as a tool to smooth images. RBF Kernel acts as the prior that selects out smooth solutions. The Radial basis function kernel also called the RBF kernel, or Gaussian kernel is a kernel that is in the form of a radial basis function (more specifically, a Gaussian function). The RBF kernel is defined as

$$KRBF(x, x') = \exp[-y \|x - x'\|^2]$$

Where y is the parameter that sets the “spread” of the kernel.

III. DESIGN AND IMPLEMENTATION

3.1 Algorithm Design

Algorithms are the operational infrastructure of every project; the algorithms determine how and how the program operated and generated results based on the calculations. An effective algorithm must encompass all the data variables available for computation and in return generate an efficient flow as well as true results of the processing afterward. When it comes to predictive analysis there is a myriad of choices over the internet that operates in statistical data to generate associative output. Choosing between these numerous algorithms itself needs a good amount of study on the topics and also a deep analysis of the predictions being made from the system. Since, in this case, multiple dependent variables are key points in prediction, we have adopted the algorithm of DNN.

1	Date	Open	High	Low	Close	Adj Close	Volume
2	#####	1269.89	1276.82	1256.44	1261.15	1261.15	1764600
3	#####	1242.71	1250	1205	1212.16	1212.16	2482400
4	#####	1241.11	1279.88	1237.2	1258.41	1258.41	2314800
5	#####	1265.74	1288.15	1260.53	1271.17	1271.17	1710100
6	#####	1255	1277.71	1244	1276.6	1276.6	1870100
7	#####	1292	1294.1	1265.06	1270.86	1270.86	2209300
8	#####	1283.2	1284.76	1230.38	1232.59	1232.59	4035000
9	#####	1345	1360.15	1326.73	1342.18	1342.18	5417900
10	#####	1331.36	1350	1321.5	1346.7	1346.7	2792100
11	#####	1324.09	1351.43	1309.66	1317.32	1317.32	2443600
12	#####	1308.13	1324.47	1296.01	1322.9	1322.9	1540300
13	#####	1337.5	1371.01	1335.03	1349.02	1349.02	1983300
14	#####	1358	1368.69	1345.13	1345.43	1345.43	1586600
15	#####	1361.31	1376	1352.54	1369.28	1369.28	1532600
16	#####	1381.82	1396.64	1372.01	1384.34	1384.34	1666300
17	#####	1376.79	1415.53	1375.5	1403.59	1403.59	1635900
18	#####	1408.22	1414.33	1374.51	1375.18	1375.18	1695900
19	#####	1376.16	1385.33	1328.01	1348.33	1348.33	2018400
20	#####	1333.52	1357.37	1323.3	1356.86	1356.86	1833000

Table 3.1.1 Company data

3.2 System Architecture

The system architecture is simply the overall design of the system. The readily set system design parameters are especially useful for the micro process of system development, converting the product from a blueprint to an actual application. This document contains the overall design of the system. The system will be constructed in 3-Tier Architecture:

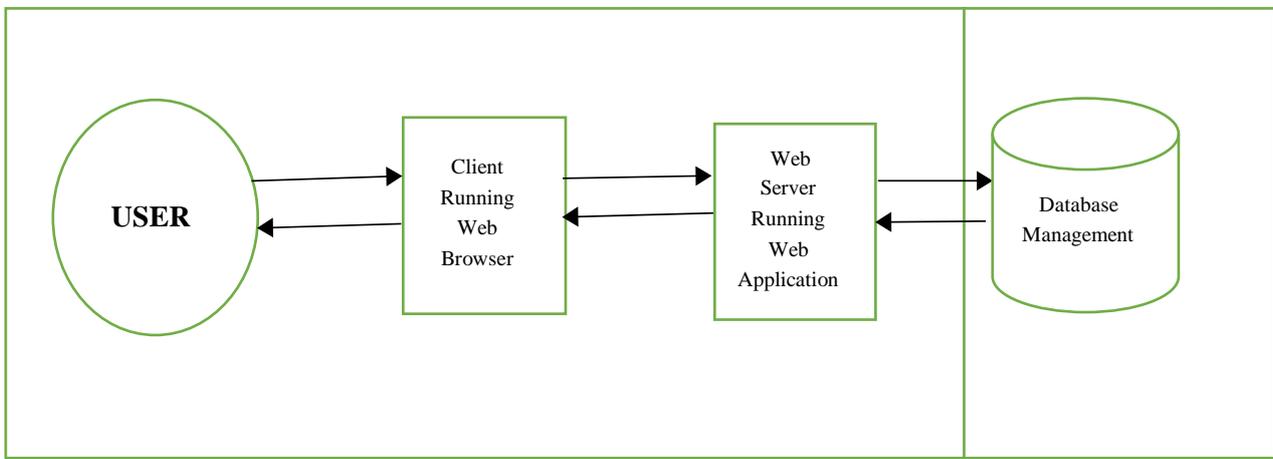


Fig 3.2.1 System Architecture

3.3 System Flow Diagram

A system flow chart simply describes a working method of a system in which users choose a company whose value is to be predicted. Then the DNN algorithm runs which simply generates a result that is shown properly in the charts.

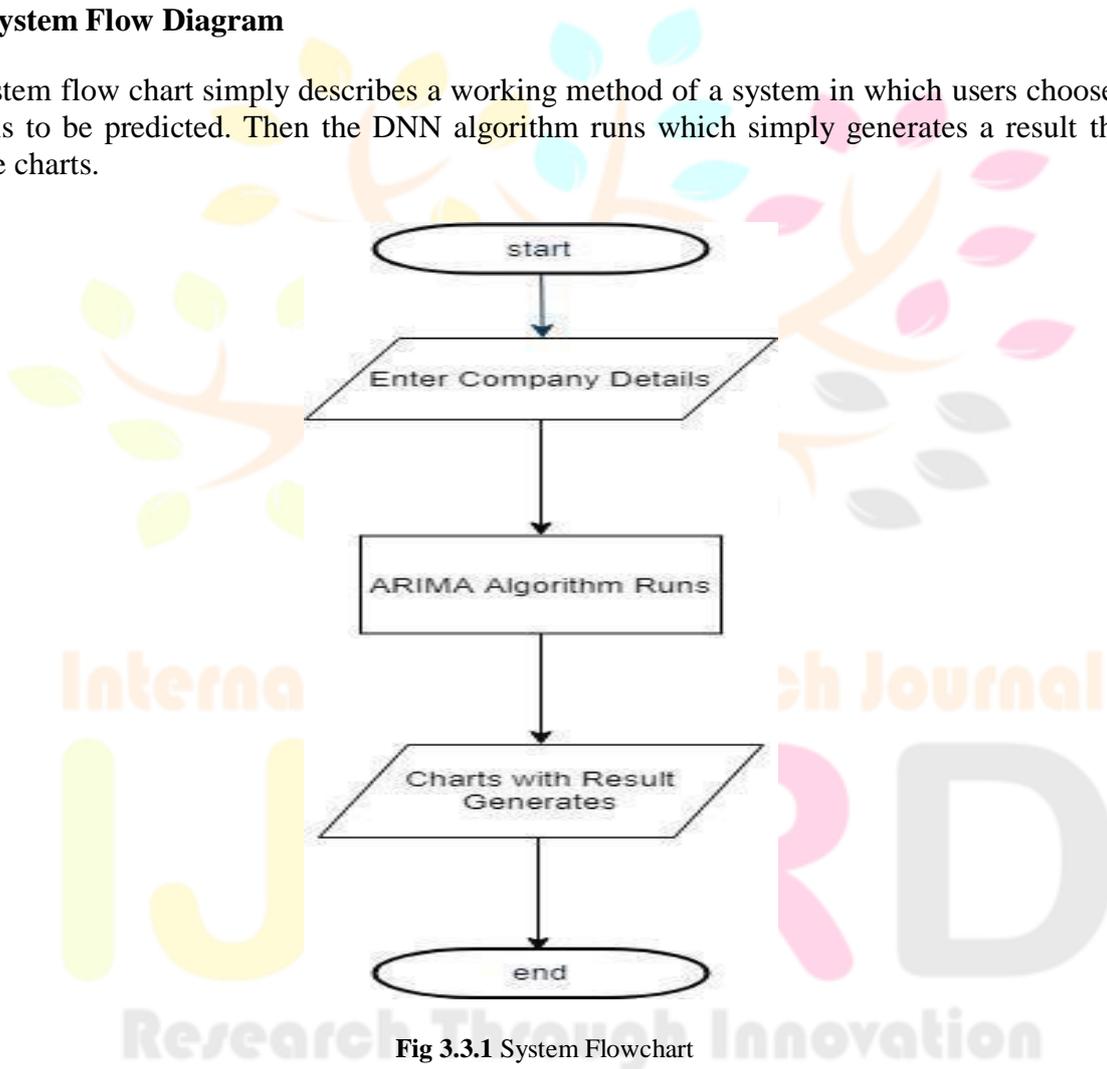


Fig 3.3.1 System Flowchart

3.4 Data Flow Diagram

A data flow diagram maps out the flow of information for any process or system. Figure 4.4.1 shows the level 0 DFD which simply shows how users interact with the system to get the desired result.

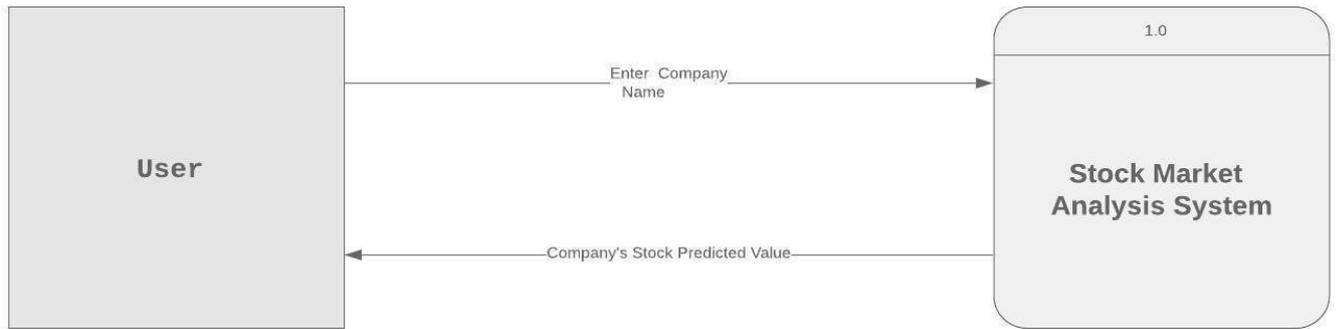


Fig 3.4.1 DFD Level-0

3.5 Use Case Diagram

Actor 1(User Description):

The user must sign up to have full access to the system. Users log in through their username and password. Users are prohibited from some features if they are not logged in to the system. But the user which doesn't have the account also has the access to view market information. An authorized user can calculate predictions of different companies, use feedback features, and be updated on different stock news.

Actor 2(Admin Description):

Admin is responsible for verifying user registration and is capable of user management in the system. Market information is updated in the system by the user. All the information about the stock is handled by the admin.

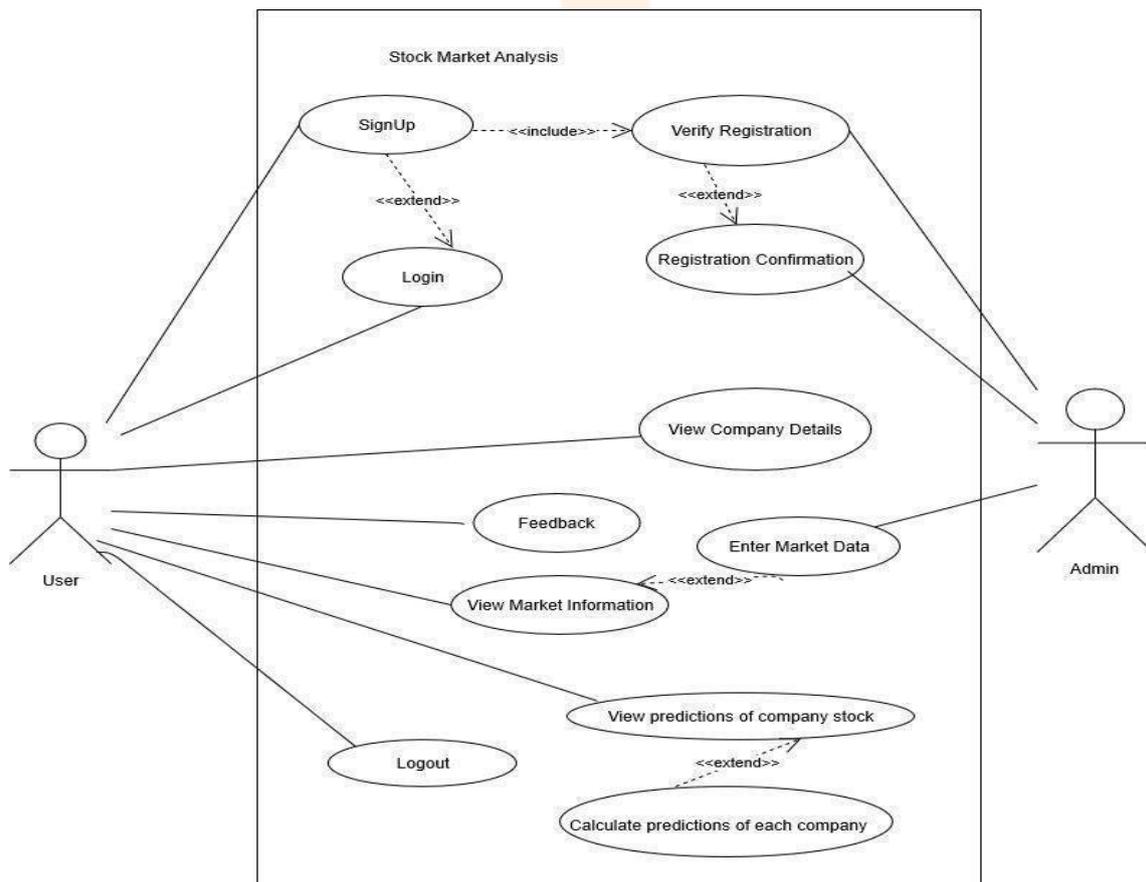


Fig 3.5.1: Use Case Diagram

V. RESULTS

Fig 5.1: Enter the name of the company stock symbol

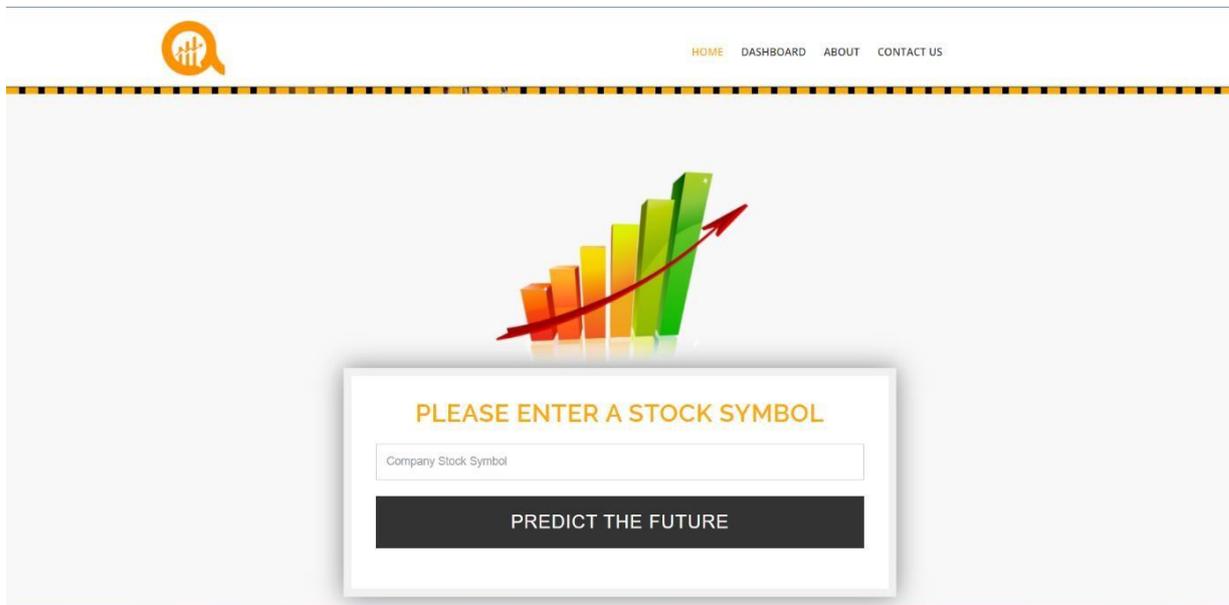


Fig 5.2: Today's prediction of Google stock.

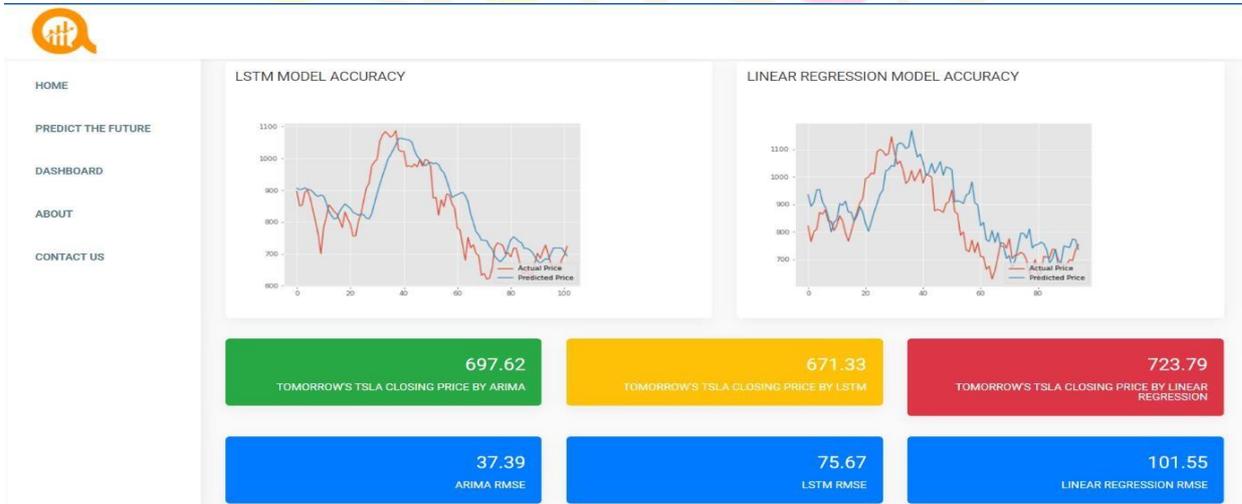


Fig 5.3: Accuracy of Actual price and predicted price.

PREDICTED GOOGL PRICE FOR THE NEXT 7 DAYS

#	Close
	2344.699116769323
	2289.1642531919624
	2284.5912627272864
	2376.365212923736
	2402.910738095535
	2488.3068110477952
	2498.446457528234

Fig 5.4: Predicted google price for the next 7 days.

VI. FUTURE ENHANCEMENT

In the Future more companies along with multiple news sources can be included. The current system is built using the Auto regressive integrated model to increase accuracy, and different combinations of DNN order were generated. By selecting the best DNN order we can obtain accuracy up to 90% or higher. A system is never fully completed as we can enhance the system in the future using different methods.

Some of the future enhancements that can be done to the system are:

- We can predict the stock value based on additional parameters such as opening values, turn-over, etc.
- We can add different additional features like alerting the user about the price rise/fall of the different company's stock.
- We can further integrate different algorithms to enhance the accuracy of the system.

VI. CONCLUSION

The final system is a web-based application, which can visualize historic time series data and future predictions. The web-based application in FLASK, with the implementation of database and visualization tools, can show the interactive plots of the scores. Finally, we were able to achieve our objectives through the build system. The system can predict the value of company stock according to the data provided to the system to train it. We can analyze the current state of the current market. The time series stat model DNN has been implemented & achieved a high accuracy rate. Our system can predict all the company stock values taking the closing value only. Besides reaching our main objective to predict the value we can add different features to our system. Although we have reached our objectives we are not fully able to get the accuracy completely. We can achieve accuracy up to 95% maximum and 90% minimum. We will be adding other features in the future to increase accuracy.

REFERENCES

1. James, Chain. (August 5, 2018). Stock Analysis <https://www.investopedia.com/terms/s/stock-analysis.asp>
2. Nicola W Burton (January 2018). A random walk down wall street. Retrieved from https://www.researchgate.net/publication/325247657_Burton_G_Malkiel's_A_random_walk_down_wall_street
3. Singh, Aishwarya. (October 25, 2018). Stock Prices Prediction Using Machine Learning and Deep Learning Techniques. Retrieved from <https://www.analyticsvidhya.com/blog/2018/10/predicting-stock-price-machine-learning-deep-learning-techniques-python/>
4. Marijana Zekic. (Unknown). Neural Network application in stock market predictions-A methodology analysis. <https://www.semanticscholar.org/paper/Neural-Network-Applications-in-Stock-Market-Zekic/9da27fa48f23766009a35dc83b6bea5901562c2f?navId=citing-papers>
5. KangZhang (6 February 2019). Stock Market Prediction Based on Generative Adversarial Network. <https://www.sciencedirect.com/science/article/pii/S1877050919302789>
6. Wei Huang, Yoshiteru Nakamori, Shou-Yang Wang, "Forecasting stock market movement direction with support vector machine", Computers & Operations Research, Volume 32, Issue 10, October 2005, Pages 2513–2522.
7. Zhen Hu, Jibe Zhu, and Ken Tse "Stocks Market Prediction Using Support Vector Machine", 6th International Conference on Information Management, Innovation Management and Industrial Engineering, 2013.M
8. N. Ancona, Classification Properties of Support Vector Machines for Regression, Technical Report, RIIESI/CNRNr. 02/99.
9. K. Jae Kim, "Financial time series forecasting using support vector machines," Neurocomputing, vol. 55, 2003.
10. Debashish Das and Mohammad shorif Uddin data mining and neural network techniques in stock market prediction: a methodological review, international journal of artificial intelligence & applications, vol.4, no.1, January 2013