International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 10 Issue 7 ǁ July 2022 ǁ PP. 361-370

Stock Price Prediction

Prof.Theja N,Abhirama K C,Charitha D S,Ganpath Singh B,Sanjana M S
AssistantProfessoratDepartmentofCSE,VidyaVikasInstituteofEngineeringandTechnology,Mysuru-570028,India
Final yearstudentsofDepartmentofCSE,VidyaVikasInstituteofEngineeringandTechnology,Mysuru-570028,India

ABSTRACT
A correct prediction of stocks can lead to huge profits for the seller and the broker. Frequently, it is brought out that prediction is chaotic rather than random, which means it can be predicted by carefully analyzing the history of respective stock market. Machine learning is an efficient way to represent such processes. It predicts a market value close to the tangible value, thereby increasing the accuracy. Introduction of machine learning to the area of stock prediction has appealed to much research because of its efficient and accurate measurements. The vital part of machine learning is the dataset used. The dataset should be as concrete as possible because a little change in the data can perpetuate massive changes in the outcome.
Index Terms—*Key Words: Long Short-Term Memory (LSTM), Convolution Neural Network (CNN), Machine Learning (ML), Trade Open, Trade Close, Trade Low, Trade High*

Date of Submission: 02-07-2022 Date of acceptance: 13-07-2022

I. INTRODUCTION

The price of the stocks is an important indicator for a company and many factors can affect their values. Different events may affect public sentiments and emotions differently, which may have an effect on the trend of stock market prices. Because of dependency on various factors, the stock prices are not static, but are instead dynamic, highly noisy and nonlinear time series data. Due to its great learning capability for solving the nonlinear time series prediction problems, machine learning has been applied to this research area. Learning-based methods for stock price prediction are very popular and a lot of enhanced strategies have been used to improve the performance of the learning-based predictors. However, performing successful stock market prediction is still a challenge.

Stock market analysis is a widely studied problem as it offers practical applications for signal processing and predictive methods and a tangible financial reward. Creating a system that yields consistent returns is extremely challenging and is currently an open problem as stock market prices are extremely volatile and vary widely both within a given stock and comparatively amongst many stocks. Further, stock market data is influenced by a large number of factors including foreign and domestic economies, trade agreements, wars, seasons, and even day of the week.

The financial market is a dynamic and composite system where people can buy and sell currencies, stocks, equities and derivatives over virtual platforms supported by brokers. The stock market allows investors to own shares of public companies through trading either by exchange or over the counter markets. This market has given investors the chance of gaining money and having a prosperous life through investing small initial amounts of money, low risk compared to the risk of opening new business or the need of high salary career.

Although humans can take orders and submit them to the market, automated trading systems (ATS) that are operated by the implementation of computer programs can perform better and with higher momentum in submitting orders than any human. However, to evaluate and control the performance of ATSs, the implementation of risk strategies and safety measures applied based on human judgements are required.

Many factors are incorporated and considered when developing an ATS, for instance, trading strategy to be adopted, complex mathematical functions that reflect the state of a specific stock, machine learning algorithms that enable the prediction of the future stock value, and specific news related to the stock being analyzed.

Time-series prediction is a common technique widely used in many real-world applications such as weather forecasting and financial market prediction. It uses the continuous data in a period of time to predict the result in the next time unit.

Many timeseries prediction algorithms have shown their effectiveness in practice. The most common algorithms now are based on Recurrent Neural Networks (RNN), as well as its special type - Long-short Term Memory (LSTM) and Gated Recurrent Unit (GRU). Stock market is a typical area that presents time-series data and many researchers study on it and proposed various models. In this project, LSTM model is used to predict the stock price.

II. LITERATURESURVEY

Every Software development requires the survey process. The Survey process is needed to get the requirement for the software. The Survey also consists of studying the present system and also studying about the tools needed for the development of the software. A proper understanding of the tools is very much essential. Following is an extract of the information of the material collected during literature survey. Literature survey is a methodology of identifying the problems in the existing system through research and proposing the development of the system to solve the problems of existing system.

Existing system:

In the existing system the user can't predict the price of the stocks manually. A stock market, equity market or share market is the aggregation of buyers and sellers of stocks (also called shares), which represent ownership claims on businesses; these may include securities listed on a public stock exchange, as well as stock that is only traded privately. It is not easy to predict the price of the stocks.

III. METHODOLOGY

Proposed system:

This project is developed for stock market analysis and prediction of stock price. Stocks are categorized in various ways. One way is by the country where the company is domiciled. For example, Nestlé and Novartis are domiciled in Switzerland, so they may be considered as part of the Swiss stock market, although their stock may also be traded on exchanges in other countries, for example, as American depository receipts (ADRs) on U.S. stock markets. A stock exchange is an exchange (or bourse) where stock brokers and traders can buy and sell shares of stock, bonds, and other securities. Many large companies have their stocks listed on a stock exchange. This makes the stock more liquid and thus more attractive to many investors. Our framework will help the client in expectation of stock value as if the cost will go up or down and dependent on the forecast client can purchase or sell the stocks.

MACHINELEARNING

Machine learning (**ML**) is a field of inquiry devoted to understanding and building methods that 'learn', that is, methods that leverage data to improve performance on some set of tasks. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as training data, in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as in medicine, email filtering, speech recognition, and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks.

A subset of machine learning is closely related to computational statistics, which focuses on making predictions using computers, but not all machine learning is statistical learning. The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a related field of study, focusing on exploratory data analysis through unsupervised learning. Some implementations of machine learning use data and neural networks in a way that mimics the working of a biological brain. In its application across business problems, machine learning is also referred to as predictive analytics.

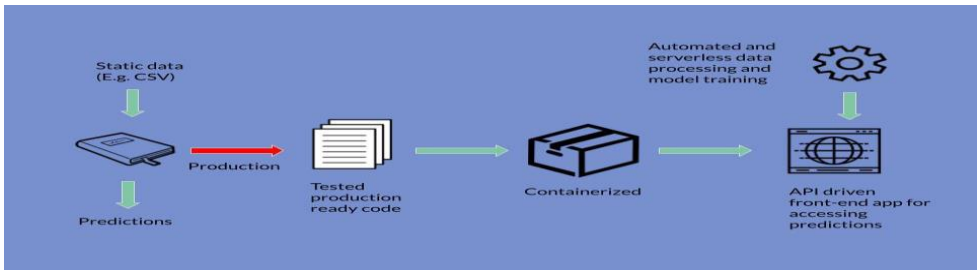

Fig1:-machinelearning model

Data flow diagram:

DFD graphically representing the functions, or processes, which capture, manipulate, store, and distribute data between a system and its environment and between components of a system. The visual representation makes it a good communication tool between User and System designer. Structure of DFD allows starting from a broad overview and expand it to a hierarchy of detailed diagrams. DFD has often been used due to the following reasons:

- Logical information flow of the system
- Determination of physical system construction requirements

- Simplicity of notation
- Establishment of manual and automated systems requirements

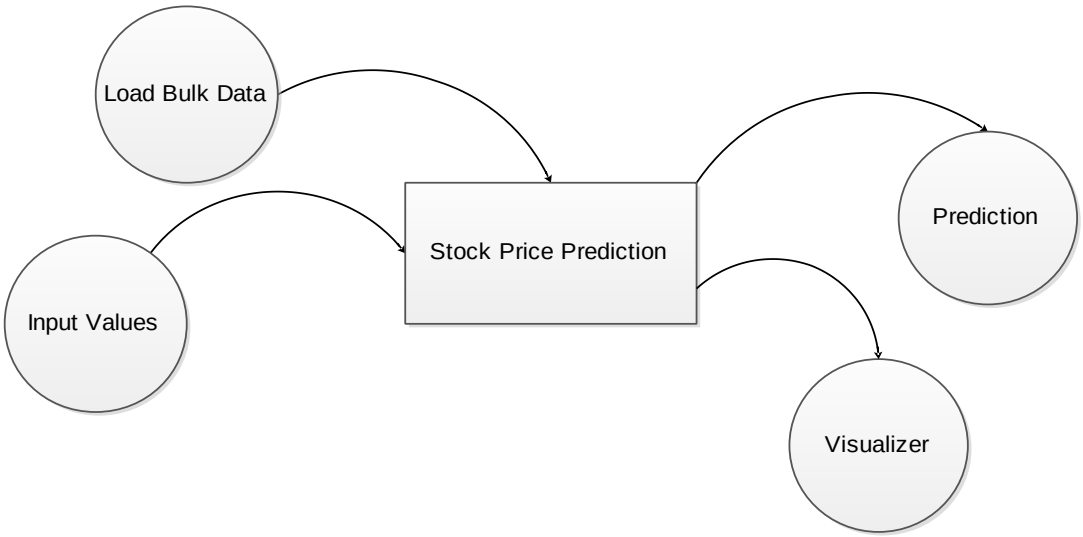

Fig2:-Level 0 Data Flow Diagram

Above diagram represents the Level 0 DFD diagram of the Stock Price Prediction. This diagram indicates the process of program. Processes are represented in the form circle that are performed in the system bulk data and input values will be given as input to the system and the output will be the prediction of stock price.

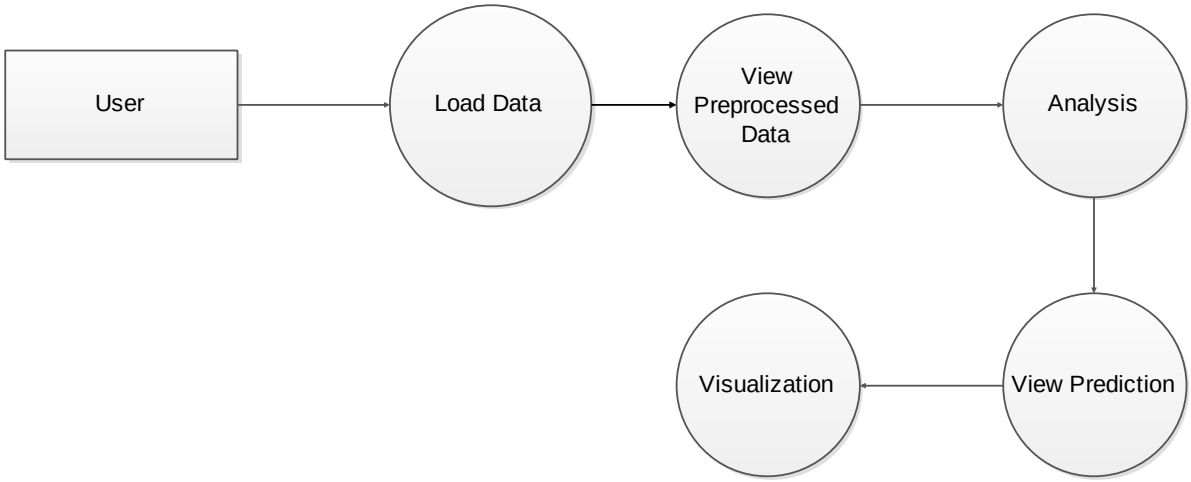

Fig 3:- Level 1 Data Flow Diagram – User

Above diagram represents the Level 1 DFD diagram of the Stock Price Prediction. This diagram indicates the process of program. Processes are represented in the form circle that are performed by the user, like loading the data required for the prediction after loading he can view the pre-processed data, data analysis, and get the prediction from the system he can view the output in the visualizer.

USECASE DIAGRAMS

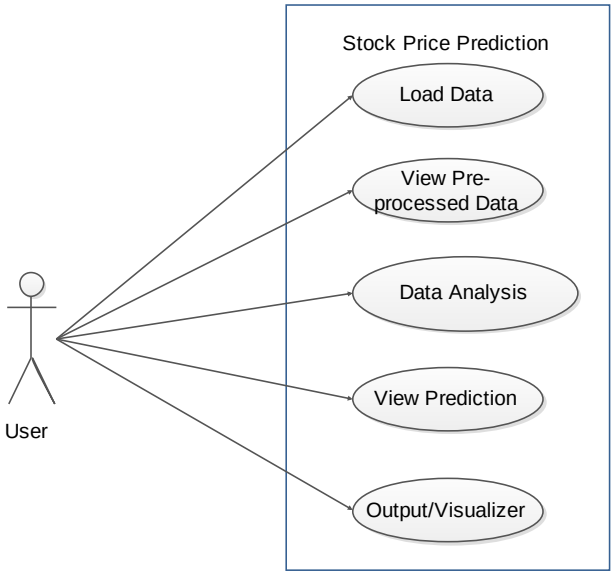

Fig4:- Use case diagram for user

Above diagram represents the user use case diagram of the Stock Price Prediction. This diagram indicates the use cases or functionalities of program. Those activities are load data, view preprocessed data and data analysis, view prediction in the visualization form.

Sequence diagrams

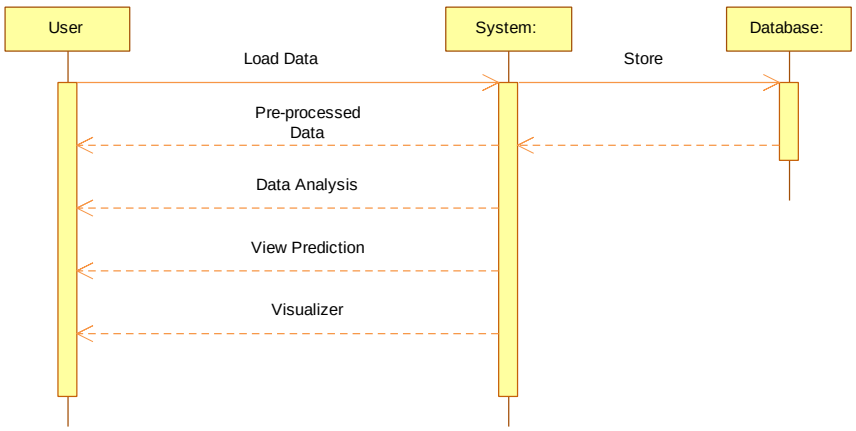

Fig 5:- Sequence diagram for User

Above diagram represents the user sequence diagram of the Stock Price Prediction. This diagram indicates the flow of program. Initially user will load the data into the system after loading the data to the system he can view the pre-processed data in the system and after that data analysis will be done based on this the prediction will be given to the user, user can view the output in the visualization.

Activity diagrams

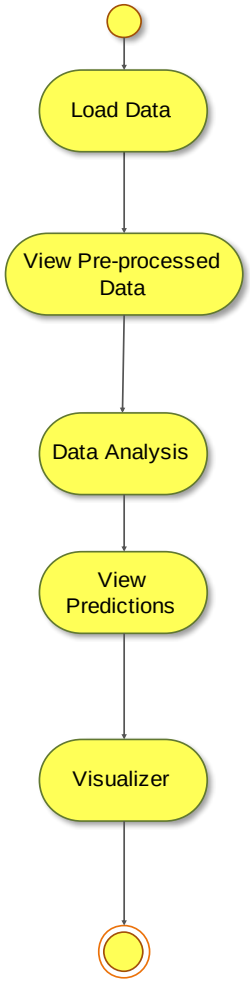

Fig 6:- Activity diagram for User

Above diagram represents the user activity diagram of the Stock Price Prediction. This diagram indicates the flow of program. Initially user will load the data which is used for the prediction and after loading the data he can view the pre-processed data in the system, then data analysis will happen and the user can view the prediction based on the given data he can view the output in the form visualization

Implementation:

The implementation stage involves careful planning, investigation of the existing system and it's constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods.

Implementation is the process of converting a new system design into operation. It is the phase that focuses on user training, site preparation and file conversion for installing a candidate system. The important factor that should be considered here is that the conversion should not disrupt the functioning of the organization.

Algorithms used:

We have used these algorithms – Linear regression, SVR, Random Forest, KNN, DT, Elastic Net, LSTM. Out of which KNN can be seen below.

KNN:

K-nearest neighbors (KNN) algorithm is a type of supervised ML algorithm which can be used for both classification as well as regression predictive problems. However, it is mainly used for classification predictive problems in industry.

K-nearest neighbours (KNN) algorithm uses 'feature similarity' to predict the values of new data points which

further means that the new data point will be assigned a value based on how closely it matches the points in the training set. We can understand its working with the help of following steps –

Step 1 − For implementing any algorithm, we need dataset. So during the first step of KNN, we must load the training data from the heart disease patients datalist and test data from the new patient entry info currently being taking tests.
Test data set – collected data samples.
Train data set – UCI & real time data from hospital.
Step 2 − Next, we need to choose the value of K i.e. the nearest data points. K can be any integer.
Here k is taken from test data set, which acts like a centroid point.
Step 3 − For each point in the test data do the following −

- 3.1 − Calculate the distance between test data and each row of training data with the help of any of the method namely: Euclidean, Manhattan or Hamming distance. The most commonly used method to calculate distance is Euclidean.
- 3.2 − Now, based on the distance value, sort the patients data based on parameters like age, gender, alcohol, bp, smoking parameters as preference of data in ascending order.
- 3.3 − Next, it will choose the top K rows from the sorted array.
- 3.4 − Now, it will assign a class to the test point based on most frequent segments from the patient data of these rows.

Step 4 – End
In our project we compare k value with train data set using KNN methodology likes Euclidean, Manhattan, where we cluster the values based on the nearest distance. Here we get nearest matches of heart disease preference to test data from train data set in turn we predict the content values from it

PSEUDOCODE FOR KNN:

Nearest-neighbor algorithm

a) A pseudo code for the nearest neighbor algorithm is

ALGORITHM *Nearest-neighbor*$(D[1...n,1...n],s)$
//Input: A $n\times n$ distance matrix $D[1...n,1...n]$ and an index s of the starting city.
//Output: A list Path of the vertices containing the tour is obtained.
for $i \leftarrow 1$ to n do Visited $[i] \leftarrow$ false
Initialize the list Path with s
Visited $[s] \leftarrow$ true
Current $\leftarrow s$
for $i \leftarrow 2$ to n do
    Find the lowest element in row current and unmarked column j containing the element.
    Current $\leftarrow j$
    Visited$[j] \leftarrow$ true
    Add j to the end of list Path
Add s to the end of list Path
return Path

Implementation methodology:
The project is implemented in modular approach. Each module is coded as per the requirements and tested and this process is iterated till the all the modules have been thoroughly implemented.

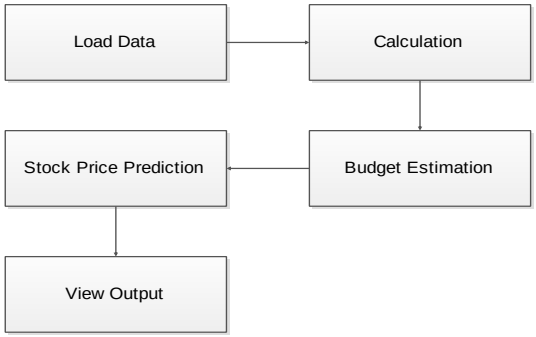

Fig 7:- Activity diagram for User

Bulk data will be given as input to the system using the given data the system will calculate and budget the estimation and based on that estimation the system will predict if the stock price will go in profit or in loss and the user can view this as output, the system will greatly help the user while buying or selling stocks.

IV. Experiential Results

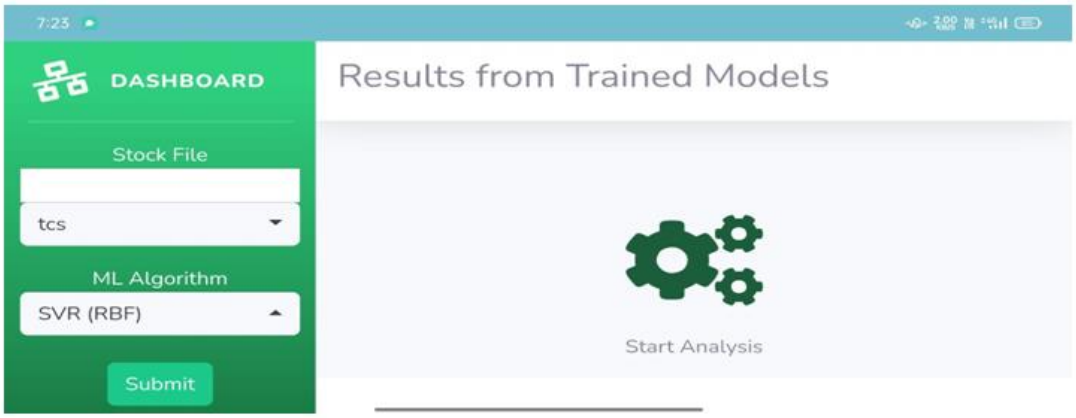

Fig.8 Set of options to choose stock and algorithms

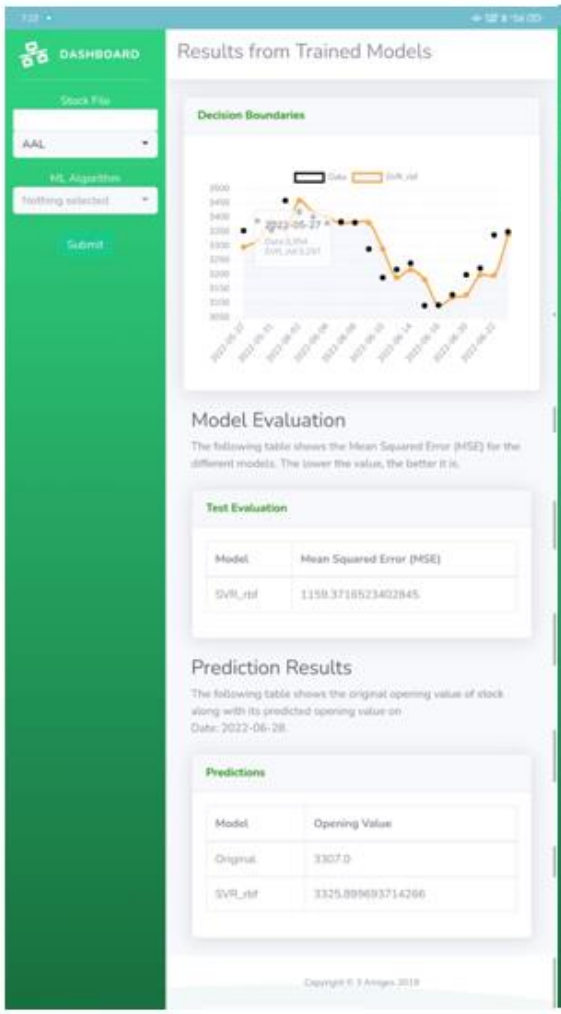

Fig.9 TCS stock prediction using SVR algorithm

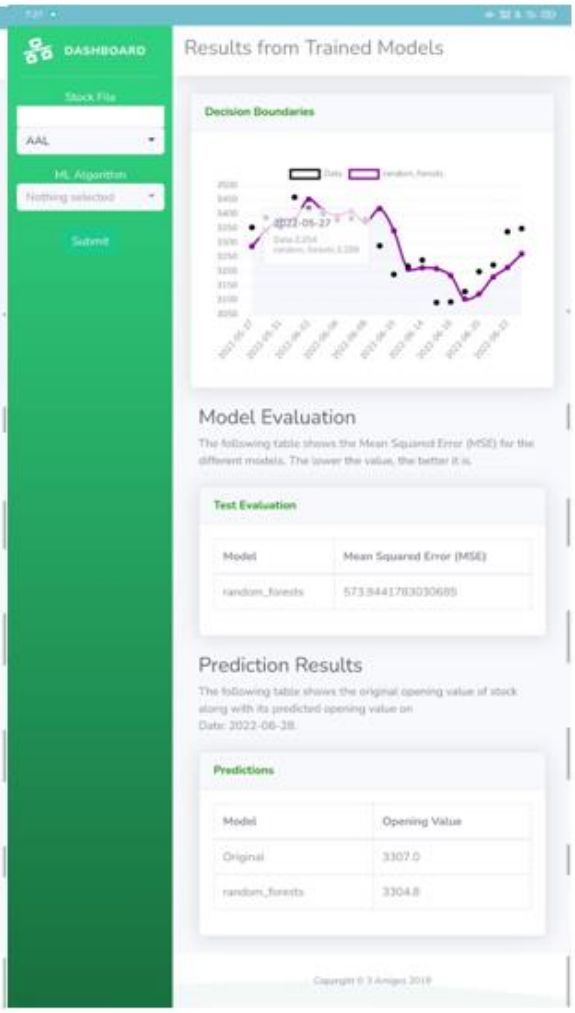

Fig.10 TCS stock prediction using Random Forest algorithm

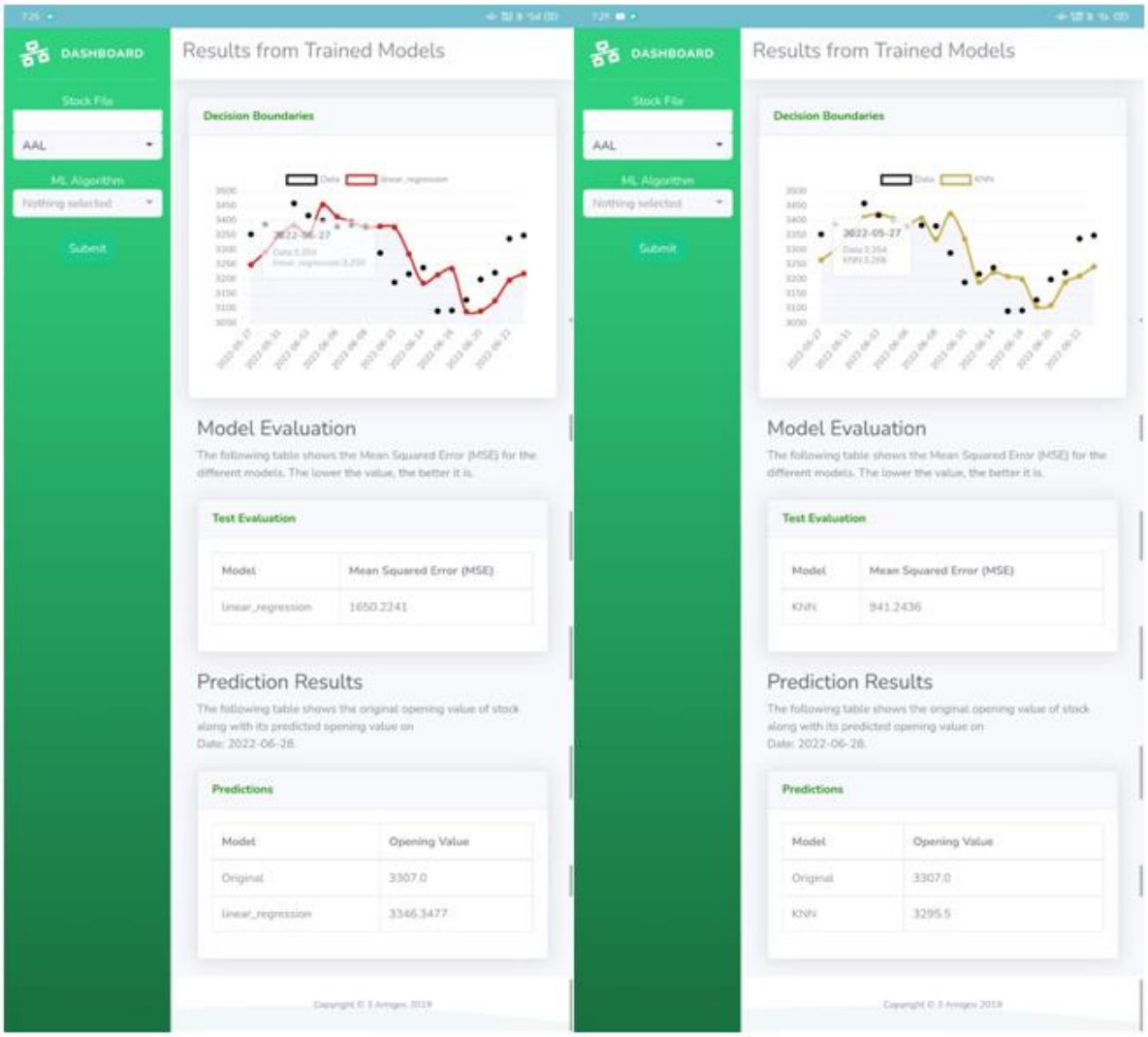

Fig.11 TCS stock prediction using linear regression algorithm

Fig.12 TCS stock prediction using KNN algorithm

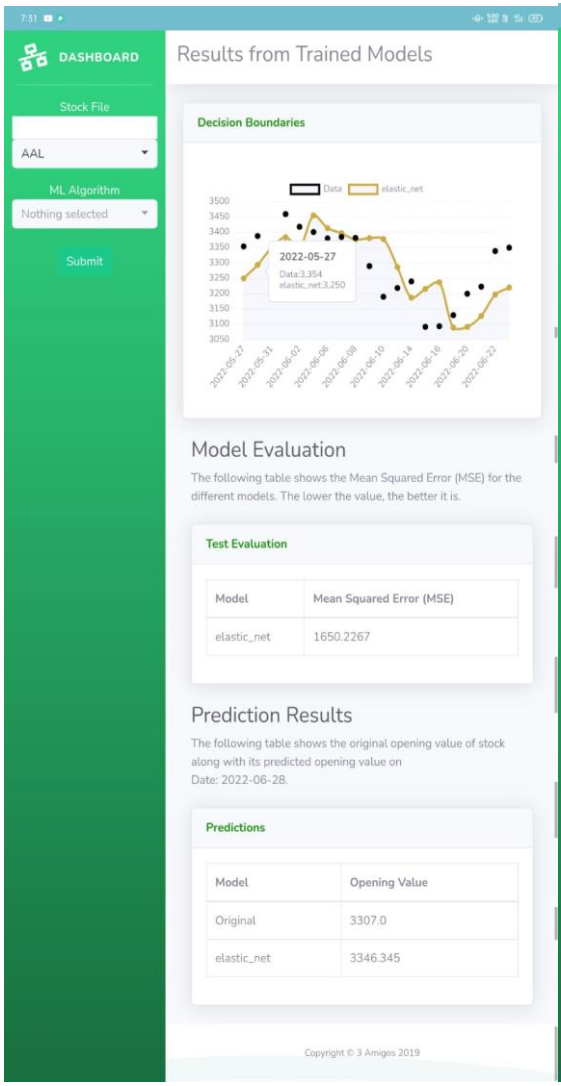

Fig.13 TCS stock prediction using elastic net algorithm

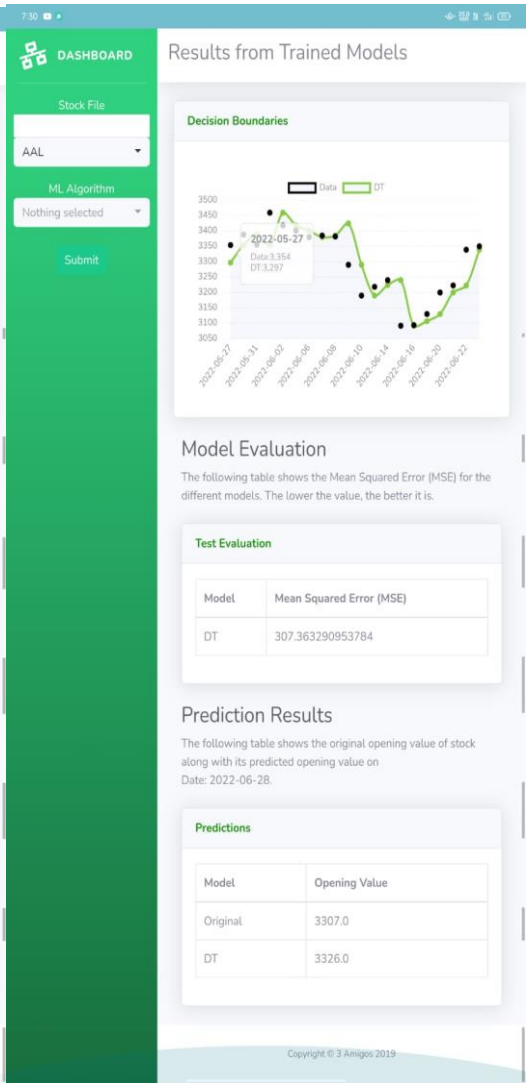

Fig.14 TCS stock prediction using DT algorithm

V. CONCLUSIONANDFUTUREENHANCEMENT

In the future, more studies on the use of different algorithms in RSs can be done to observe the implications of their use, performance, and utility. Moreover, RS development lacks studies analyzing early stages, such as requirements and design, and late stages, such as maintenance. Open questions in these stages must be investigated to improve the knowledge about the field.

The main objective of our dissertation is to develop a more adaptive and effective stock prediction system by applying machine learning techniques. The survey papers prove the successfully of our proposed approach. A systematic prediction tool is developed could be used to assist invested make more accurate decision in their stock market investment. Our prediction system integrates the stock movement forecasting and stock price forecasting.

In addition, numbers of visualization are provided to enhance our system. Various external online data sources. The models using the features from these external sources along with the traditional stock market data improve the performance for the stock market prediction.

REFERENCES

[1]. M. Usmani, S. H. Adil, K. Raza and S. S. A. Ali, "Stock market prediction using machine learning techniques," 2016 3rd International Conference on Computer and Information Sciences (ICCOINS), Kuala Lumpur, 2016, pp. 322-327.

[2]. K. Raza, "Prediction of Stock Market performance by using machine learning techniques," 2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT), Karachi, 2017, pp. 1-1.

[3]. H. Gunduz, Z. Cataltepe and Y. Yaslan, "Stock market direction prediction using deep neural networks," 2017 25th Signal Processing and Communications Applications Conference (SIU), Antalya, 2017, pp. 1-4.

[4]. M. Billah, S. Waheed and A. Hanifa, "Stock market prediction using an improved training algorithm of neural network," 2016 2nd

International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), Rajshahi, 2016, pp. 1-4.
[5]. H. L. Siew and M. J. Nordin, "Regression techniques for the prediction of stock price trend," 2012 International Conference on Statistics in Science, Business and Engineering (ICSSBE), Langkawi, 2012, pp. 1- 5.
[6]. K. V. Sujatha and S. M. Sundaram, "Stock index prediction using regression and neural network models under non normal conditions," INTERACT-2010, Chennai, 2010, pp. 59-63.
[7]. S. Liu, G. Liao and Y. Ding, "Stock transaction prediction modelling and analysis based on LSTM," 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, 2018, pp. 2787-2790.
[8]. T. Gao, Y. Chai and Y. Liu, "Applying long short term memory neural networks for predicting stock closing price," 2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, 2017, pp. 575-578.
[9]. K. A. Althelaya, E. M. El-Alfy and S. Mohammed, "Evaluation of bidirectional LSTM for short-and long-term stock market prediction," 2018 9th International Conference on Information and Communication Systems (ICICS), Irbid, 2018, pp. 151-156