STOCK PRICE PREDICTION AND FORECASTING WITH MACHINE LEARNING

BY

Shaila Hossain Nodi
ID: 201-15-13715

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Ms. Nazmun Nessa Moon
Associate Professor
Department of CSE
Daffodil International University

Co-Supervised By

Ms. Sharmin Akter
Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2024

APPROVAL

This Project titled **"STOCK PRICE PREDICTION AND FORECASTING WITH MACHINE LEARNING"**, submitted by **Shaila Hossain Nodi**, ID No: 201-15-13715 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on *26/01/2024*.

BOARD OF EXAMINERS

Chairman

Dr. S.M Aminul Haque (SMAH)
Professor & Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

26.01.2024

Dr. Fizar Ahmed (FZA)
Associate Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Most. Hasna Hena (HH)
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

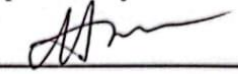
26/01/2024

Dr. Shamim H Ripon (DSR)
Professor
Department of Computer Science and Engineering
East West University

DECLARATION

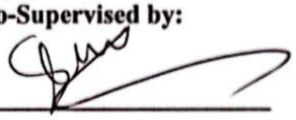
I hereby declare that; this project has been done by us under the supervision of **Ms. Nazmun Nessa Moon, Associate Professor, Department of CSE** Daffodil International University. I also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:



Ms. Nazmun Nessa Moon
Associate Professor
Department of CSE
Daffodil International University

Co-Supervised by:



Ms. Sharmin Akter
Senior Lecturer
Department of CSE
Daffodil International University

Submitted by:

Shaila Hossain Nodi

Shaila Hossain Nodi
ID: 201-15-13715
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

At first I would like to express my heartiest thanks and gratefulness to almighty God for His divine blessing makes me possible to complete the final year thesis successfully.

I am really grateful and wish my profound my indebtedness to **Ms. Nazmun Nessa Moon, Associate Professor**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of my supervisor in the field of "*Machine learning and Deep learning*" to carry out this project. Her endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts and correcting them at all stage have made it possible to complete this project.

I would like to express my heartiest gratitude to **Dr. Sheak Rashed Haider Noori**, **Professor and Head,** Department of CSE, for his kind help to finish my project and also to other faculty member and the staff of CSE department of Daffodil International University.

Finally, I must acknowledge with due respect the constant support and patients of my parents.

ABSTRACT

This project titled “STOCK PRICE PREDICTION AND FORECASTING WITH MACHINE LEARNING” which predict the price of stock for Bangladeshi data. Various methods such as analysis, time analysis and statistical analysis are used in estimating the market value. However, none of these methods have consistently proven to be reliable tools. Machine learning, a process that enables machines to learn from interactions in the world and expand from examples without explicit instruction, has the potential to be responsible for many important applications. Linear regression (LR) is an important technique in machine learning that can identify linear variables. Support Vector Machines (SVMs), on the other hand, have advanced features such as accuracy and prediction. LSTM has neurons that can precisely categorized features more accurately. I have implied three different model for each dataset. Each model shows satisfactory performance. Among them, linear regression performs the best for every dataset with accuracy 99.78%. SVM perform second best with accuracy 95.37%. LSTM model, performs 98.8% accuracy.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

TABLES **PAGE NO**

CHAPTER 1
INTRODUCTION

1.1 Introduction

The stock market is the worldwide financial market where companies trade stocks and securities. Supply and demand affect the stock price. Additionally, other features also affect the market. Understanding business growth is difficult and time consuming. Predicting the value of stocks covers many fields such as computer science, mathematics, economics, history and finance and is one of the most difficult and difficult subjects to study. Simple time series or regression techniques will be difficult due to the uncertainty of the stock market. Merchants and financial institutions have created various designs to market themselves or their customers; But only some of these models always produce above-average capital. In this study, I have analyzed historical data of companies' listed prices to predict the company's future market prices using machine learning algorithms. Start by using simple techniques like linear regression. Deep learning scientists attach importance to deep learning based on neural networks. Degraded performance can be resolved using neural network technology since they are discrete processes. Advantages of this technology; dynamic operation and self-learning organization, parallel operation process, good topology and very good ability.

1.2 Motivation

Investment decisions: Investors rely on market forecasts to help them decide whether to buy, sell or hold stocks. Investors can maximize their returns and manage their portfolios thanks to accurate forecasts.

Risk Management: The ability to predict business changes helps assess and manage risk. To reduce losses and preserve their cash, investors can adjust their investment strategies and portfolios according to the forecasts.

Trading Strategy: Traders develop short-term trading strategies based on market forecasts. Forecasting tools are useful for investors (such as swing traders and day traders) who want to profit from changes in prices and market patterns.

Corporations: Companies can use stock market forecasts to inform their options. Finance, strategy, mergers and acquisitions, and other business-related business decisions fall into this category.

Investor confidence: Market expectations will affect investor confidence. If the forecast is not good, investor sentiment will decrease, affecting market liquidity and overall market stability. Conversely, a good forecast will attract more investment.

Research and Development: Stock price predictions are often used by data scientists, financial institutions, and research professionals to develop financial models and machine learning. This is useful for general measurement and world finance.

1.3 Rationale of Study

Analysis of price forecasts leads to research in this field and contributes to the overall understanding of market conditions, volatility and market behavior. Considering that technological advances, especially in machine learning and artificial intelligence, are constantly changing financial analysis, it is important to understand the quality of cost estimates in order to keep up with new technologies and techniques. The principle of forecasting work stems from the impact of the financial market on various financial stakeholders. Investors and companies look for reliable sources and information to help them make investment decisions. Accurately predicting stock prices is crucial to managing

risk, optimizing investment strategies and meeting financial needs. Forecasting models play an important role in asset allocation as they allow business managers to adjust investment options based on the risks and prospects of the business.

1.4 Research Question

This project has been accomplished with hard work and a lot of dedication. This task was not easy to complete. During this project, many obstacles came to the path to make it fair, realistic, and accurate. Few questions are raised to understanding the concept and provide the answer:

1. Where did I find my data for this project?
2. How the data processing carried out?
3. Which ML model is used for this project?
4. Which algorithm is more effective?
5. What's the best recommended model accuracy?

1.5 Expected Outcomes

- This project outcome is to analyze and predict the stock price.
- To make an efficient system for making investment
- Express the level of uncertainty by presenting expected values or confidence intervals.
- To have better understanding of company's financial activities.
- Discuss the risks and uncertainties associated with the estimate.
- Enhance the contribution of academic knowledge.
- Provide findings from the model's back testing on historical data, if appropriate.
- Compare model forecasts to market or analyst requirements.

1.6 Report Layout

Chapter 1 contains the project introduction, motivation, research question, rationale of the study and expected outcome. This section covered the overall structure of the report. Chapter 2 covers the terminologies we need to follow for this work. It also contains the previous work that is been done on this topic. The challenges faced during developing the project and scope of the problem both explained in this section. Chapter 3 provides an overview of the theoretical analysis conducted in this study, explores statistical methods used to solve mathematical problems. Additionally, this chapter presents a programmatic approach to machine learning techniques. The next section describes the data collection and data preparation process. At the end of this section is the analysis of the confusion matrix, which evaluates the model and assigns labels to outliers. To ensure accuracy and precision in machine learning techniques, it is important that evaluation is carried out together with the application. This section includes the content and tools of the study, procedures, data collection procedures, data processing, recommendation models, training models and preliminary applications. Chapter 4 includes the presentation of experimental results, performance evaluation, and discussion of the results. This section contains some experimental images to facilitate the completion of the project. A comprehensive review of the results and machine learning techniques is also provided. Chapters 5 and 6 contain a summary of the research, information on future activities, and conclusions. This section provides a specific example to ensure that the project report meets the requirements. Additionally, the chapter concludes by explaining the limitations of the effort that may impact future professionals in this field.

CHAPTER 2
BACKGROUND

2.1 Introduction

Predicting future changes in stock prices in financial markets using various models, methods and techniques is called stock price forecasting. Predicting the future of stock prices involves analysis of historical data, current market conditions and other relevant factors. Traders and investors use price predictions to help them decide whether to buy, sell or hold stocks. Researchers and experts in the field of price forecasting work to predict future economic activity in the financial sector, which is a combination of finance, technology output, and measurement forecasting. The purpose of this forecast is to provide traders, investors and financial institutions with important information to maximize data, manage risk and make cuts. Estimating legal costs is still a complex and important area of research that relies on economic data, geography, and investor behavior, to name just a few of the many differences that continue to influence financial markets. Stock price prediction aims to help various stakeholders in the financial ecosystem. The main goal is to provide understandable information to investors so they can make informed decisions about buying, holding and selling stocks. Another important goal is optimization, which involves using cost estimates to influence the asset allocation strategy to create both mutually exclusive balance and reasonable risk data. Using these predictions, traders aim to create effective trading plans by discovering patterns, trends, and the best time to enter and exit the market. An important aspect of financial planning is stock price forecasting, which provides individuals and organizations with information for long-term financial management and retirement planning. From a research perspective, the aim is to improve financial literacy, improve our understanding of the market and improve our forecasting models. The continuous development of new technologies such as artificial intelligence and machine learning points to the goal of constantly improving the capabilities of cost estimation models. These objectives indicate the success of stock price prediction, but it is important to realize that there is always a lack of interest and difficulty in the financial market and 100% accuracy is very difficult to achieve. The consequences of trying to

predict stock prices are many and have a significant impact on many aspects of the financial environment. An equally important aspect of price forecasting is risk mitigation; This allows investors to stay informed about changes in the market and take preliminary measures to protect information from adverse events. Finally, the advancement and use of technologies such as machine learning and artificial intelligence have led to improved capability of cost estimation models. Although these results indicate good performance, it is important to understand that product price prediction is unreliable due to the complexity and power of market copy financing.

2.2 Related Works

This study presents a framework that combines PSO, LS-SVM, and other training methods to improve the accuracy of cost estimation. According to Smith et al. (2018) state that LS-SVM outperforms statistical models in stock price prediction. Their work demonstrates how well LS-SVM can handle noisy, high-dimensional data while preserving nonlinear relationships. Chen and Wang (2019) use PSO to improve the parameters of the LS-SVM model for stock price prediction. According to their study, combining PSO with LS-SVM improves prediction accuracy and overall ability in support. The best method is LS-SVM-PSO and the lowest error rate is an LS-SVM. The worst algorithm is ANN-BP [1]. According to this study, Li and Wang (2018) improved the robustness and accuracy of the regression line-based cost estimation model by using methods such as regularization and well-running. Support vector machines (SVMs) have become important in business forecasting because they can handle high data and nonlinear data. In previous studies, the use of vector systems for economic forecasting has yielded significant results. From here, when comparing LR and SVM, it is clear that SVM is better than LR [2].The goal of this paper is to use machine learning and support vector machines (SVM) to predict the direction of stock prices. It also tries to predict whether stock prices will rise or fall in the coming days. It discusses how support vector machines and machine learning can discover patterns and relationships in historical data to create predictions [3].

This study uses Support Vector Machine (SVM) machine learning techniques to predict the prices of large- and small-cap stocks. RBF function (Radial Basis Function) is also used in this project. It is especially frequently used in support vector machine classification. The main purpose of this article is to investigate the possibility of using SVM, a popular learning algorithm, to predict product or model price. SVM works on many values from various financial markets around the world. This method does not have overfitting problems [4].The results show that the LSTM neural network model is able to predict stock prices when tracking techniques are included, despite limitations such as market forecasting. Its main purpose is to first analyze how time works by analyzing historical business data, and then deeply study the basic rules of the LSTM neural network model using the special knowledge of deep learning memory to predict market value [5].This article presents a new method for predicting stock prices using two well-known models: the Bidirectional Long-Term Memory (BI-LSTM) model and the Recurrent Neural Network (RNN) model, mainly called the Long Short-Term Memory (LWM) model. Since both RNN models (LSTM and BI-LSTM) are tuned according to hyperparameters, simulation results show that the proposed method can successfully predict future patterns. In order to develop more accurate models to predict future outcomes, the RMSE of LSTM and BI-LSTM models were measured with different time variables, hidden layers, layer thicknesses, and units used in the hidden process [6]. This work will introduce the concept of random forests, a popular machine learning method that excels at processing complex data and producing reliable predictions. Random Forest combines predictions from all decision trees in the set to create its prediction. Regression tasks use the average of the predictions, while classification tasks use the majority of votes to determine the final class. Random forest provides predictions from multiple decision trees to increase detail and overall performance. This article compares the performance of the algorithm with logistic regression. Finally, we can conclude that Random Forest performs better in terms of variance, mean score, mean square score, and log mean square error [7] . In the context of this study, most researchers investigate forecasting used in international business. It includes research using machine learning such as regression analysis, reinforcement learning, and support vector machine (SVM). It also discusses the use of benchmarking

techniques to evaluate the effectiveness and accuracy of algorithms, as well as the use of business models to evaluate their performance. This study uses individual forecasts (single forecast, long-term forecast, multiple forecast) to predict the daily trend of the Nasdaq index. The results include a high degree of accuracy and statistical data on the relationship between US stock indices and the global economy in predicting daily US stock market trends [8]. In summary, the literature review shows that machine learning techniques such as KNN and RF are promising in the field of business forecasting. These algorithms are used to improve forecast accuracy, identify important data, and detect nonlinear patterns. Their research shows that the choice of expertise helps determine the most important points for forecasting the market. Previous studies have demonstrated the effectiveness of RF in predicting stock prices. Their research shows that a combination of RF and electoral efficiency can improve representation [9]. This article aims to predict the future direction of stock prices. He acknowledges that sometimes the share price may be higher or lower than the current price. To achieve this prediction, this paper uses historical data to train an artificial neural network (ANN) model. ANN models have additional advantages such as improved accuracy for many variables due to the inability to follow the input and output graph. Inspired by the neuron model of the human brain, neural networks consist of interconnected cells, or "neurons." This article will provide an overview of ANN models and their capabilities [10]. The aim of this paper is to create a model which combines ANN and DT to enhance the rate of prediction accuracy in stock price. Here two combined models are compared, which are DT + ANN and DT + DT. DT+ANN model has 77% accuracy, which is higher than the single ANN and DT models [11]. In this work, Artificial Neural Network and Random Forest techniques have been utilized for predicting the next day closing price for five companies. ANN proves to be a better technique, giving better RMSE and MAPE values. ANN model gives RMSE (0.42), MAPE (0.77) and MBE (0.013) [12] . This study investigates the effectiveness and feasibility of using multiple regression methods in predicting stock prices. An attempt was made to predict the closing prices of shares of six different companies traded on the Dhaka Stock Exchange (DSE). Three years (2017-2019) historical business data of these organizations were used. The aim of this study is to create an additional learning-based model that can predict the closing

price of the product. Analyze and evaluate the performance of the forecast using standard deviation error (MAE). Experimental results demonstrate the ability to track rolling MAE scores [13]. In this study, an artificial neural network (ANN) model was proposed and implemented using multilayer feedforward backpropagation technology. The study analyzed data from 15 companies over a six-year period. To predict specific values, the proposed model was trained using four different methods: Bayesian regularization (BR), scaled conjugate gradient (SCG), quasi-Newton and Levenberg-Marquardt (LM). The parameters of these algorithms change during training. The accuracy of the LM algorithm reaches 95.64%, followed by BR (91.26%), SCG (88.91%) and Quasi Newton (84.20%). The results clearly show that the LM algorithm outperforms other methods in the following aspects: Accuracy. Additionally, the LM method offers a lower error rate, making it the best algorithm for prediction in our design [14]. This article presents a machine learning model that uses two different genes, linear regression and radial basis function (RBF), to predict stock prices through support vector regression (SVR). This study includes comparing the accuracy of various SVR kernels, as well as estimating and comparing the accuracy of SVR and LR. The accuracy of each algorithm was evaluated using the standard sum of squared error (SSE), which showed a significant improvement compared to previous studies. The linear regression model consistently achieved the highest accuracy with a prediction of 97.07%. Then, the accuracy of the SVR (linear) model and SVR (radial basis function) model reached about 97.06% and 96.82%, respectively [15]. With recent advances in deep learning and machine learning, various algorithms are now used to predict stock prices. This article will take an in-depth look at five different algorithms namely K-nearest neighbor algorithm, linear regression algorithm, support vector regression algorithm, decision tree regression algorithm and time short memory algorithm to predict the stock price of 12 famous Indian listed companies. In addition, this study also suggests other reliable and useful methods for forecasting the market economy. The main goal of this research is to develop machine learning (ML) models that can provide more cost-effective products. Using this technology, traders and investors can increase profitability by making timely and informed investment decisions [16]. In this study, Author evaluate the estimation of SVM, a type of fundamental research that estimates the value of products

by analyzing the value of transactions. SVM uses the financial data of the business as input. Based on our findings, they predict whether the stock price will rise or fall. As a result, they find that stock price prediction is poor over time and that using financial data as input for SVM leads to better predictions compared to professional forecasts of experts [17]. In this study, they present an analysis of the Hong Kong and Korean markets using machine learning combining analysis of support vector machine (SVM) and principal component analysis (PCA). Our goal is to predict whether stock markets and market prices will be bullish or bearish. The proposed method uses PCA, a special selection technique, to identify the main components of the business process. SVM is a classifier designed to predict future market trends by using these key points along with other financial metrics when training and forecasting [18]. The proposed algorithm uses market data and machine learning techniques, specifically the central link called short-term, to predict stock prices. This method uses stochastic gradient descent to adjust the weight of each data point. Our system will produce more accurate results than stock price prediction algorithms. This study investigates product forecasting and has the potential to cover a wide range of future products. Prediction accuracy can be further improved by training the model with additional data, improved power functions, additional layers, and more LSTM modules [19]. This study examines the comparison of estimates obtained from different tools. The accuracy of these comparisons forms the basis for evaluating the results. Classifiers are prepared by determining the accuracy: Random Forest (93%), Bag (93%), AdaBoost (82%), Decision Tree (79%), ANN (75%), K-NN (71%). SVM (61%), K-NN (56%), ANN (56%), Random Forest (53%), Bag (53%) and Decision Tree (49%). These findings clearly show that repeating output data can improve accuracy. Therefore, our research mainly aims to compare seven machine learning algorithms on four different types of products [20].

2.3 Research Summary

The opportunities in the stock market are the main topic of this study. Several Bangladeshi companies' stock prices are included in the dataset that I used. The dataset was gathered from a website in Bangladesh. I have employed a deep learning model (LSTM) and two

machine learning algorithms (LR & SVM). The SVM algorithm operates on a sizable dataset. Furthermore, overfitting is not a problem with SVM. Both models function incredibly well. Python is my chosen language to use for this project. I assess and forecast price using the entire model for my online application.

2.4 Scope of the Problem

The scope of the project's cost estimate includes a clear definition and objectives that explain the focus and limitations of the project. First, it is important to determine the time frame of your stock price forecast. This may take days, weeks, months or even years. It is important to determine the features or variables that will be used later in the prediction process. They may include historical stock prices, financial indicators, news sentiment and other trends. It should also describe the type of prediction. Estimates of actual stock prices, percentage changes, or direction may be included. It is also important to consider the complexity of the model to be used. This can range from simple linear regression models to more advanced machine learning or deep learning models. Finally, it is important to decide whether and how many external factors will be included in the model and how predictive they will be.

2.5 Challenges

Data collection is the most challenging part of this study. The online collection of Bangladeshi stock quotes is small. While this may seem interesting, there are many challenges to using machine learning and deep learning to successfully estimate the value of stock. Market volatility and confusing nonlinear data make forecasting difficult. The process continues to be complicated by many issues, such as optimization of the model, difficulty of feature selection, and risk of bias. It is also important to consider ethical issues such as fairness and transparency. Despite these challenges, machine learning provides valuable tools for better understanding business and investment planning. However, I must

not forget that finance requires human wisdom. For this reason, it is recommended that you consult an experienced advisor before making speculative investment decisions.

CHAPTER 3
RESEARCH METHODOLOGY

3.1 Introduction

In this section I will explain the process I used for this project. Some of the operations are performed such as downloading the data set, preparing the data, analyzing the data, selecting the model, training the model, verifying the results and making the prediction. Both machine learning models and deep learning models are used to predict the price of the product. Both models were trained carefully due to the target features in my dataset. Due to the time series data set, data needs to be prepared. Models are selected based on their performance. In figure 3.1 I demonstrate the full methodology at a glance.

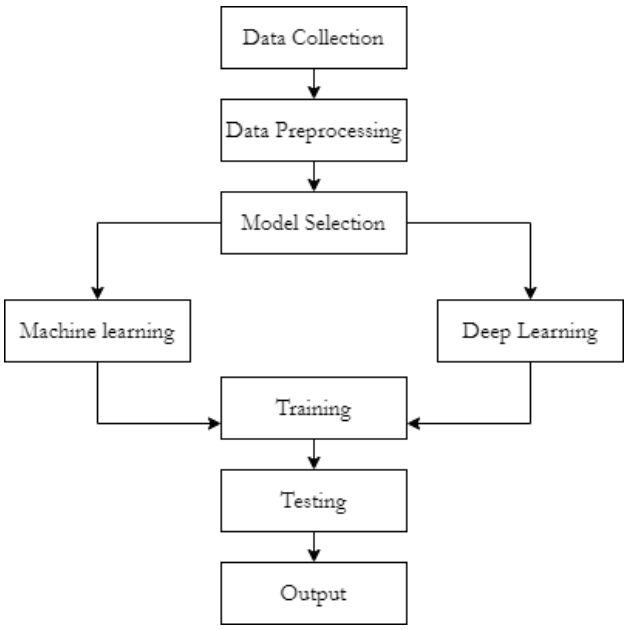


Figure 3.1: Methodology

3.2 Research Subject and Instrumentation

This project is a machine learning and deep learning model-based stock price prediction system. The primary goal is to create a system that will precisely evaluate how well these approaches forecast stock prices. Finding the information related to a specific business is the first step in continuing. Financial data from the past should be required for data. Examine the short- and long-term predictions based on the dataset. To predict stock price, I need to know market patterns and trends. Build hybrid models that combine deep learning and machine learning to increase accuracy. The models I use include Long Short-Term Memory (LSTM), Decision Tree (DT) and Support Vector Machine (SVM). This project was created using Python as the preferred language. The main platform for creating models and visualizations is Google Collaboratory. My evaluation of the model will be based on its accuracy**.**

3.3 Data Collection Procedure

I obtained the data for this study from the DSEstock website, which covers all companies from 2008 to 2022. Only three companies – Grameenphone, Beximco and BRAC – were selected as the preference study. Each company's dataset contains more than 3,000 records with appropriate dates and times.

3.3.1 Attributes

There are eleven characters in the file. Date and value (date, trading_code, last_traded_price, high, low, opening_price, closing_price, yesterdays_closing_price, trade, value_mn, volume) are attributes. Prices represent the highest and lowest prices for a specific period, while dates represent the data collection period. The first transaction price of the day is called the opening price, and the last transaction price is called the closing price. The company is identified by the trading code. The price at which the stock was last sold when the market closed for trading on a specific day is referred to as yesterday's traded price. During a regular trading session, the last traded price is the final price at which a stock is traded. Trade represents the total number of transactions that have taken place for

a particular security or asset within a designated timeframe. The equity value of a company's stock market, measured in millions, and volume indicates the number of shares available in the market. To build and evaluate the model, feature extraction method is used to extract feature.

3.3.2 Data Preprocessing

Data preprocessing is needed to go for further process. There are 11 features in the dataset, but we don't need all these features. I have used two different features selection technique, SelectKBest and Correlation based feature selection. By using SelectKBest features extraction technique, I got better score and I use it for data analysis and my web-based project. SelectKBest score and rank features based on the correlation of the features. In table 3.1, a small demonstration of the dataset of BRAC is given which shows all the attributes and value of the dataset. Here first 5 days stock value is taken from the dataset.

Table 3.1: Dataset of BRAC

Date	High	Low	Opening_price	Closing_ price	Yesterdays_closing _price	Last_traded _price
3/6/2008	1164	1142	1160	1150.5	1160.50	1142
3/9/2008	1198	1125.5	1150.25	1180.75	1150.50	1164
3/10/2008	1195	1141	1195	1170.75	1150.50	1164
3/11/2008	1170	1126.25	1152	1144	1180.75	1176.25
3/12/2008	1149	1110	1140	1130.5	1170.75	1143

In table 3.2, a small demonstration of the dataset of Beximco is given which shows all the attributes and value of the dataset. Here first 5 days stock value is taken from the dataset.

Table 3.2: Dataset of Beximco

Date	High	Low	Opening _price	Closing_price	Yesterdays_closing _price	Last_traded_price
3/6/2008	38.8	37.3	38	37.7	37.2	37.8
3/9/2008	38.8	37.5	38	37.5	37.7	37.6
3/10/2008	38.8	37.5	38	37.5	37.7	37.6
3/11/2008	38.8	37.5	38.3	38	37.5	38
3/12/2008	38	37.6	37.5	37.1	38	37.1

In table 3.3, a small demonstration of the dataset of Grameenphone is given which shows all the attributes and value of the dataset. Here first 5 days stock value is taken from the dataset.

Table 3.3: Dataset of Grameenphone

Date	High	Low	Opening _price	Closing_price	Yesterdays_closing _price	Last_traded_price
1/3/2010	192.3	187	187	189.3	187.5	189.8
1/4/2010	198.9	190	190	197.9	189.3	198.2
1/5/2010	202	195.2	199	196.5	197.9	195.7
1/6/2010	197.8	195	196	196.4	196.5	196
1/7/2010	197	193.5	197	193.8	196.4	193.6

3.3.3 Data Organizing

I have to split the data into two parts, testing dataset and validation dataset. These two parts is needed for training and evaluate my model. All three company will have separate dataset for this project. Data is well organized and accurate.

3.3.4 Data Storing

As this project based on Google Collaboratory, data I used for this project is better to store on Google drive. Because its faster and safer for storing the data. Also, data can be read direct in collaboratory from the drive. The data is stored in Excel format. The data can be used then for model training, analysis and prediction.

3.3.5 Machine Learning Algorithm

Based on the data characteristic Linear Regression, Support Vector Machine and LSTM model is applied. NumPy, Pandas, Matplotlib, scikit-learn and many more python library is used to visualize and manipulate the data.

3.4 Statistical Analysis

Initial market price refers to the initial price of the stock traded at the time the market opens for trading. The opening price should not be confused with the closing price of the previous day. Closing price is the last price of the security traded on the regular trading day.

In figure 3.2, the y axis defines the opening price in taka and x axis defines date. The graph shows the increase and decrease price of BRAC over time. The most recent opening price is around 500 taka.

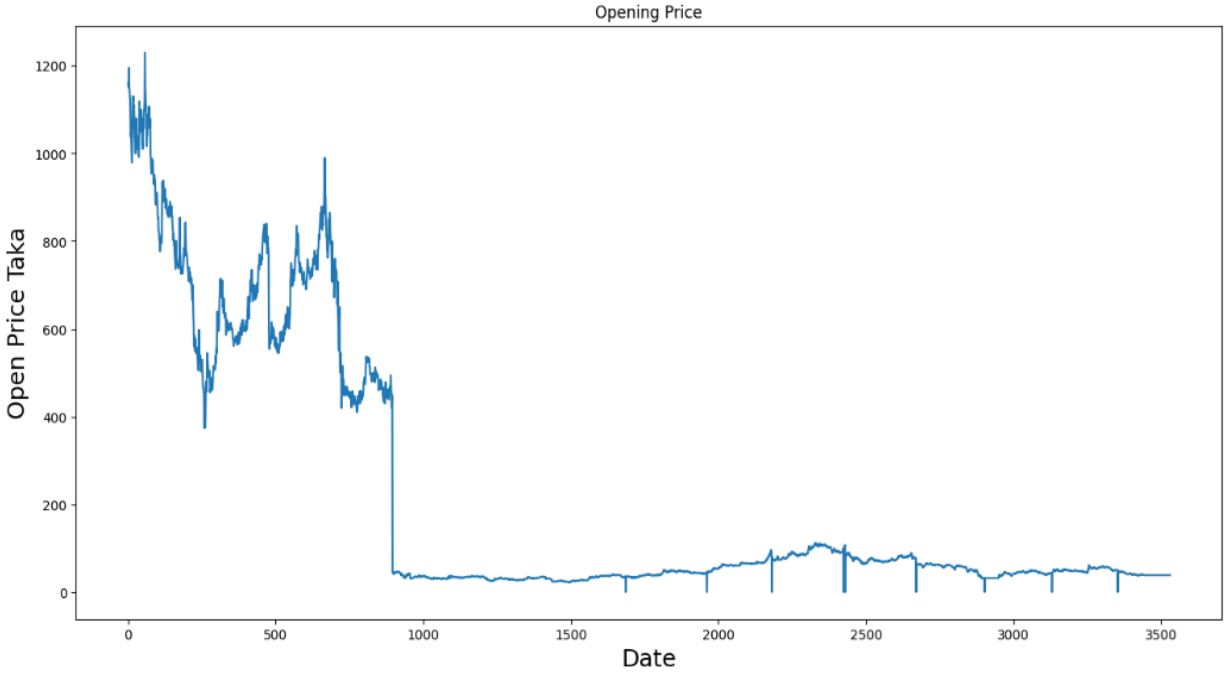


Figure 3.2: Opening price (BRAC)

In figure 3.3 the most recent closing price is 200 taka.

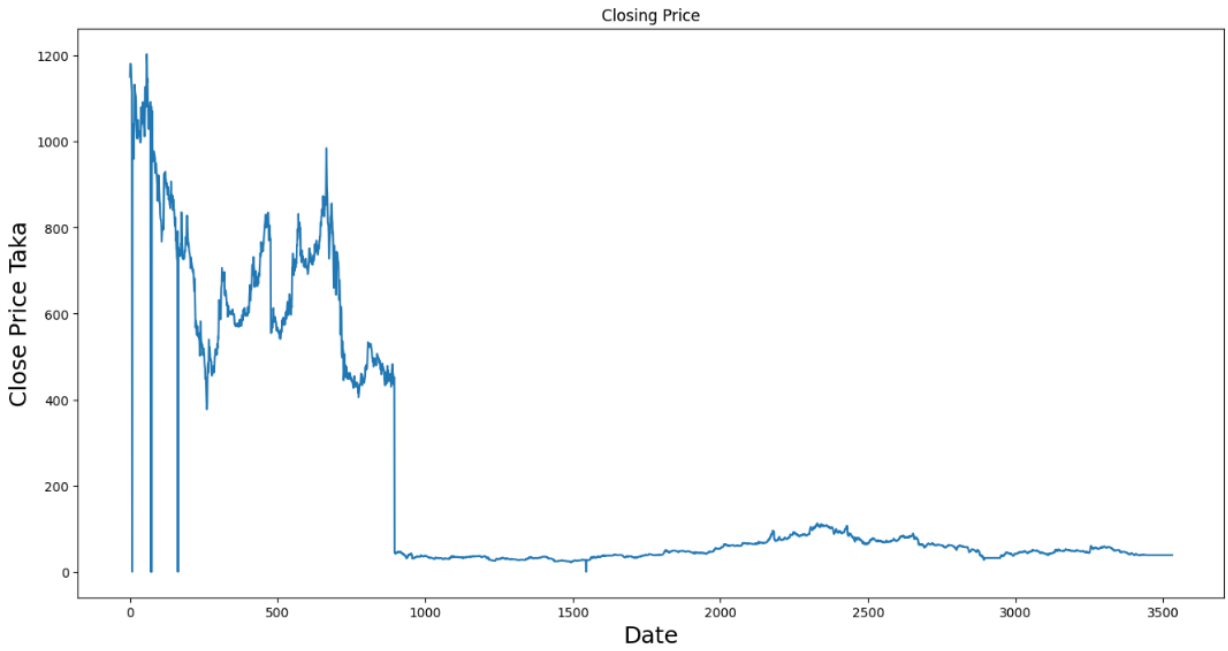


Figure 3.3: Closing Price (BRAC)

In figure 3.4, the blue line of the graph shows the opening price of Grameenphone. The graph starts at around 100 taka and raise up to 500 taka.

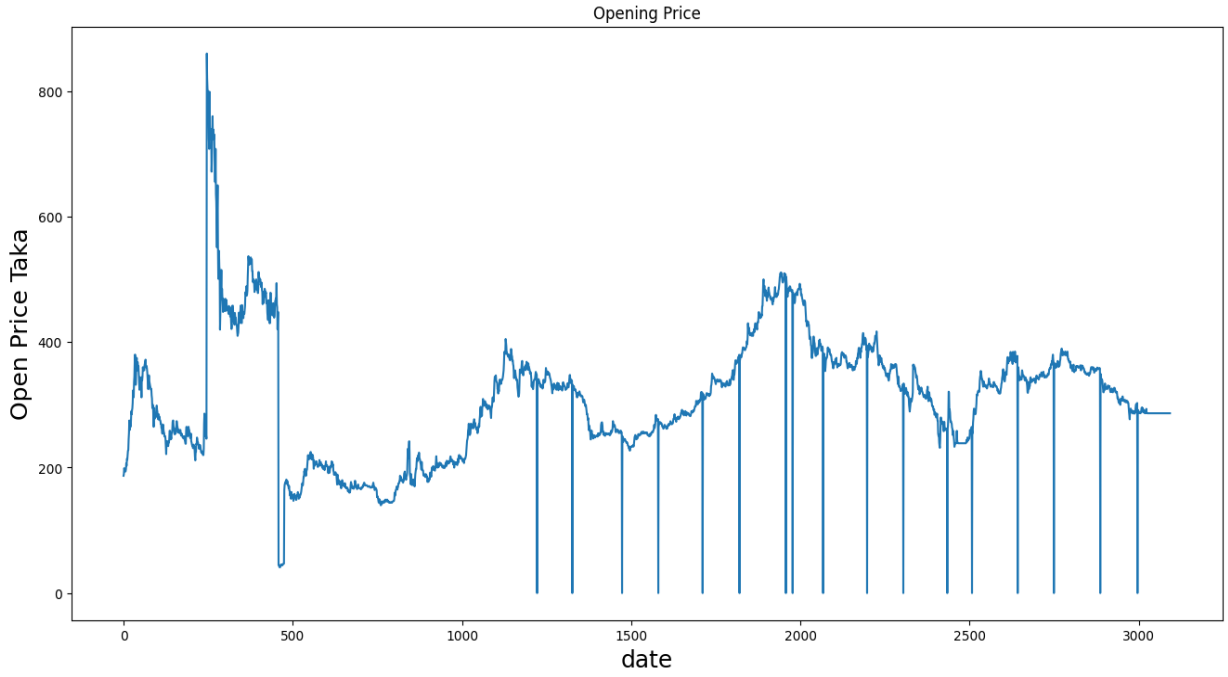


Figure 3.4: Opening Price (Grameenphone)

In figure 3.5, there have been period both increase and decreases in price. The most recent price is 300 taka.

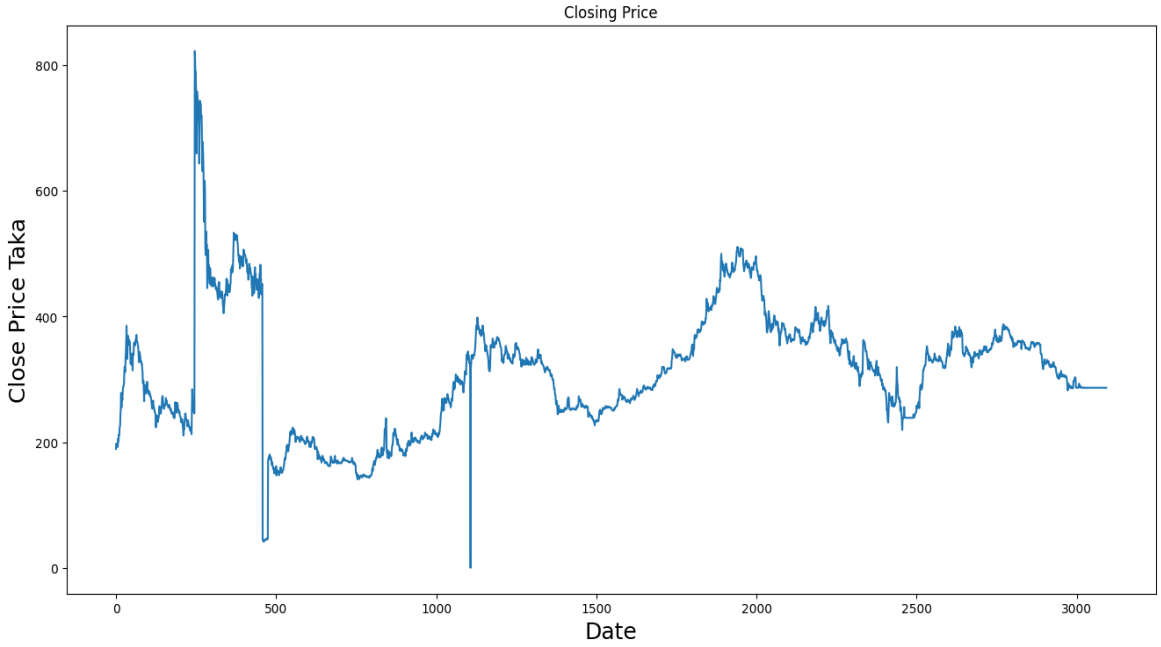


Figure 3.5: Closing Price (Grameenphone)

In figure 3.6, the graph demonstrates the opening price of Beximco. The blue line shows the increase and decrease of stock price. The most recent opening price is below 100.

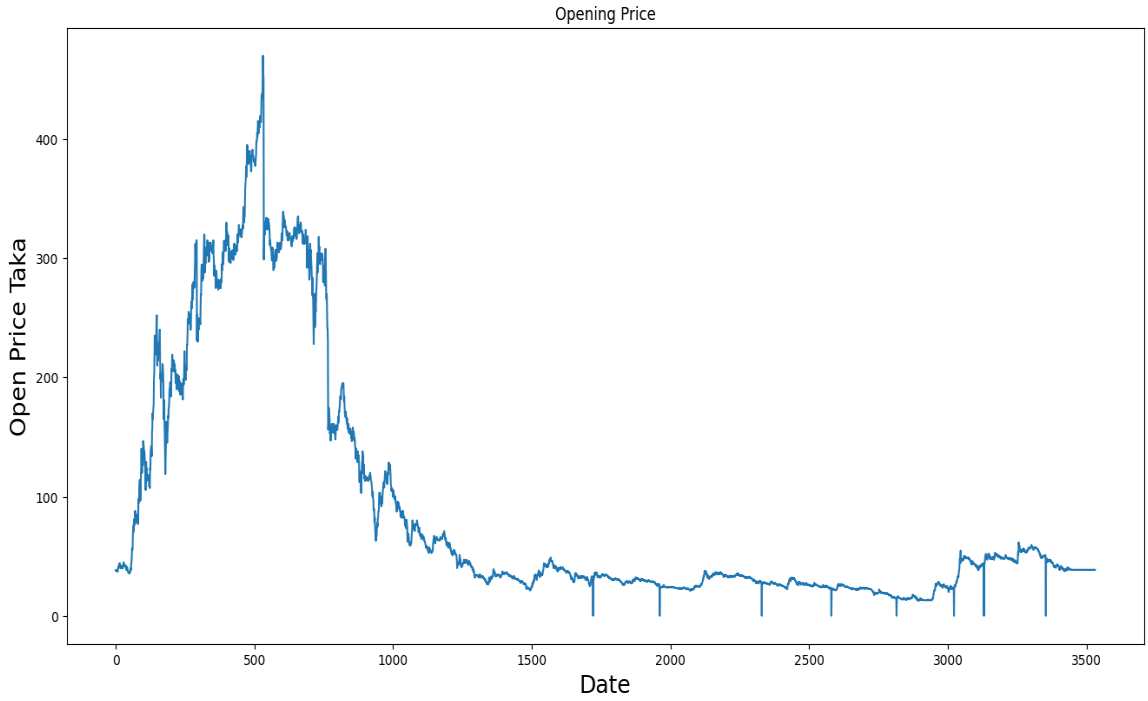


Figure 3.6: Opening Price (Beximco)

In figure 3.7, the graph shows that the highest closing price touched around 500 taka. The most recent closing price is around 100 taka.

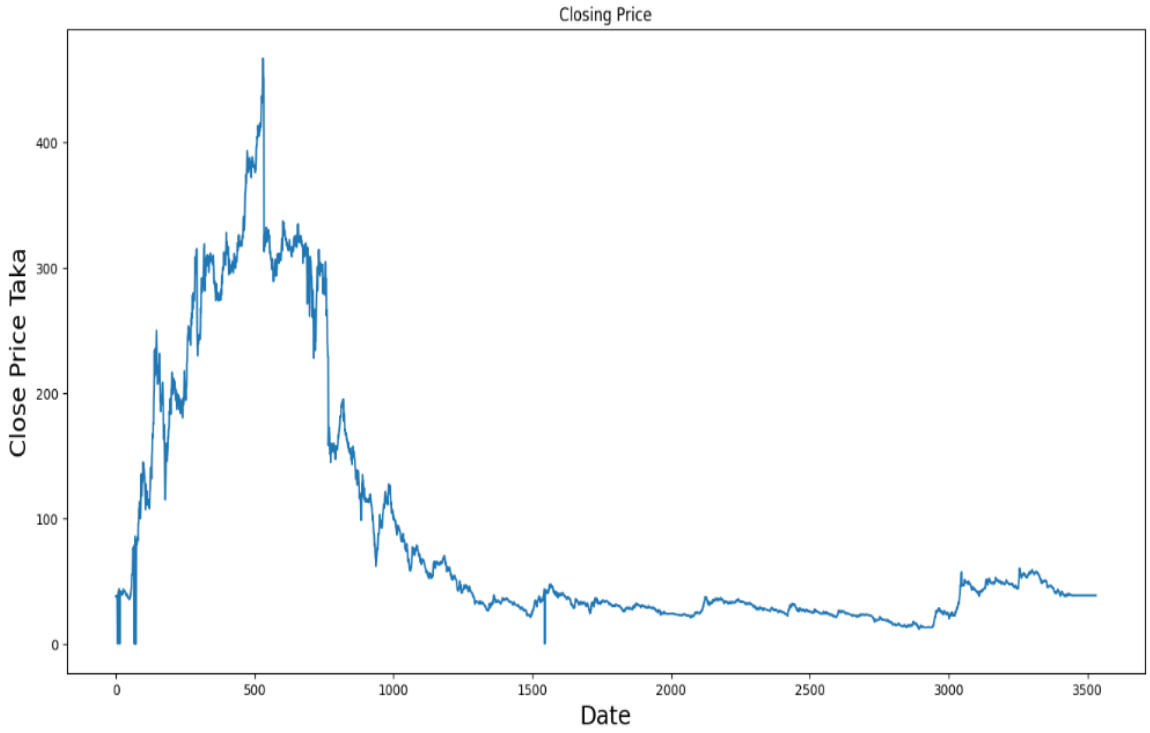


Figure 3.7: Closing price (Beximco)

In figure 3.8, the graph shows the moving average of BRAC. In this graph 5, 20 and 50 days of moving average is shown in three different color line.

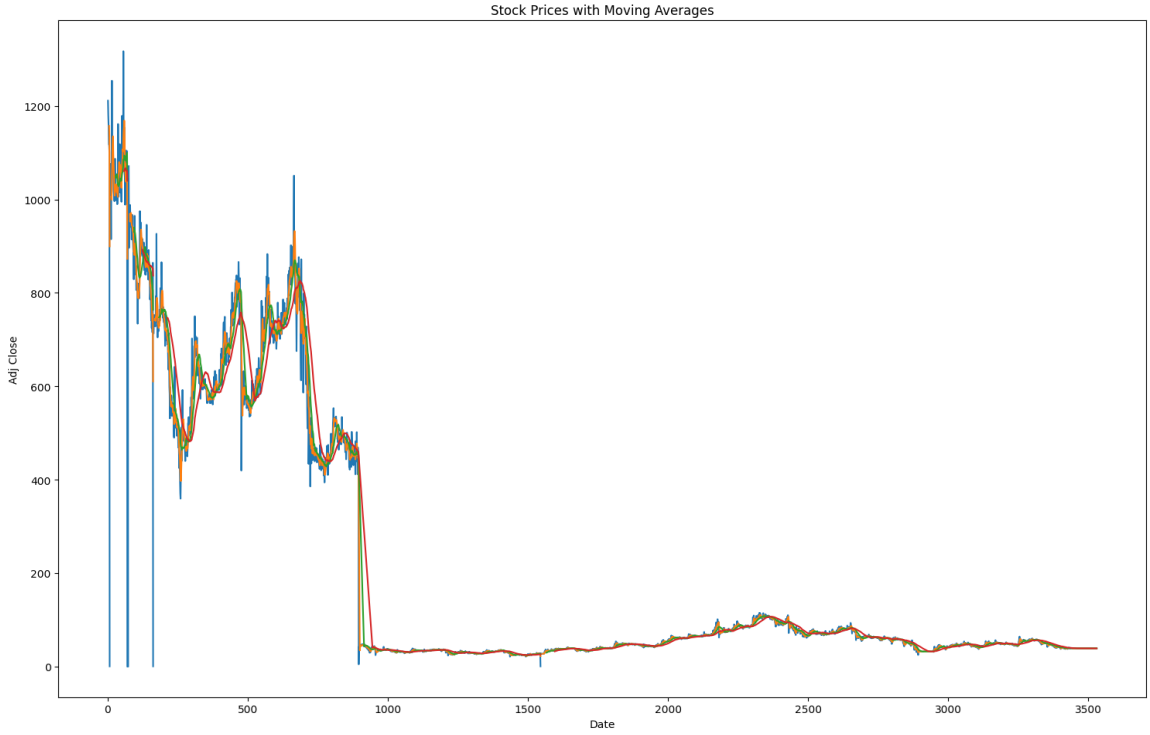


Figure 3.8: Moving Average (BRAC)

In figure 3.9, the graph demonstrates the moving average of Grameenphone for 5, 20 and 50 days. The curve shows the downfall and the uprising of prices.

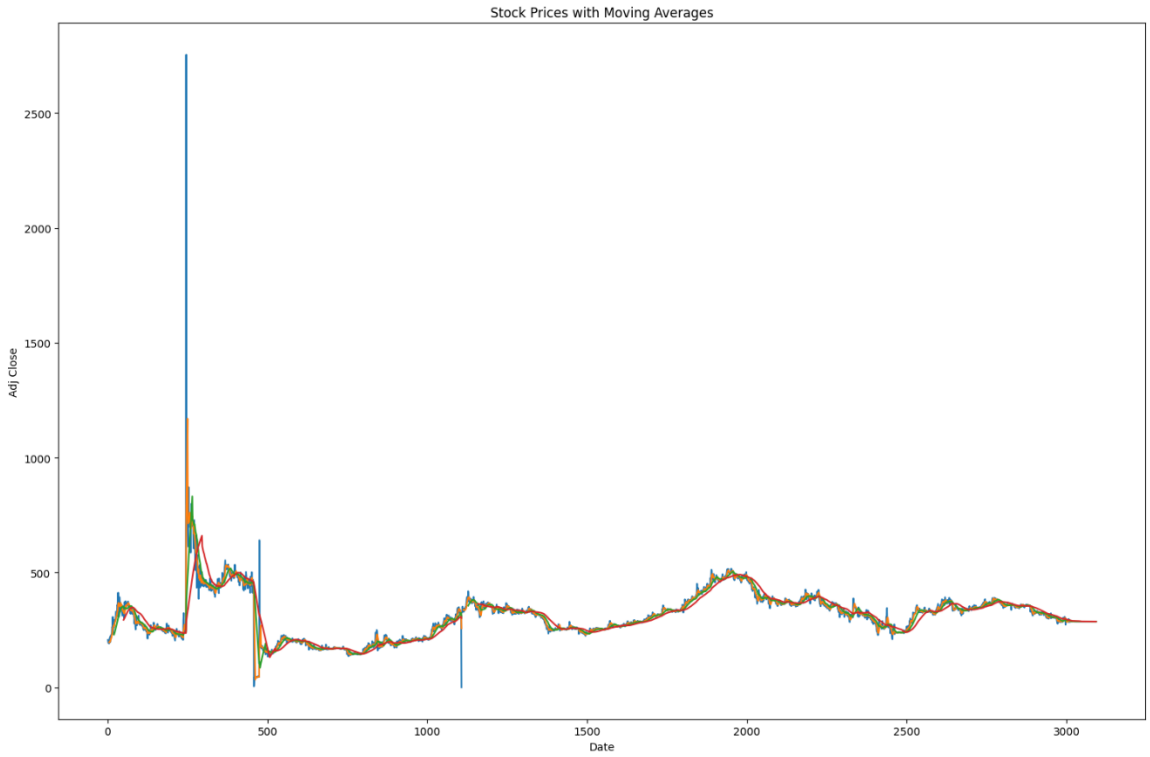


Figure 3.9: Moving Average (Grameenphone)

In figure 3.10, the graph demonstrates 5, 20 and 50 days of moving average of Beximco. Moving average is used to determine data point.

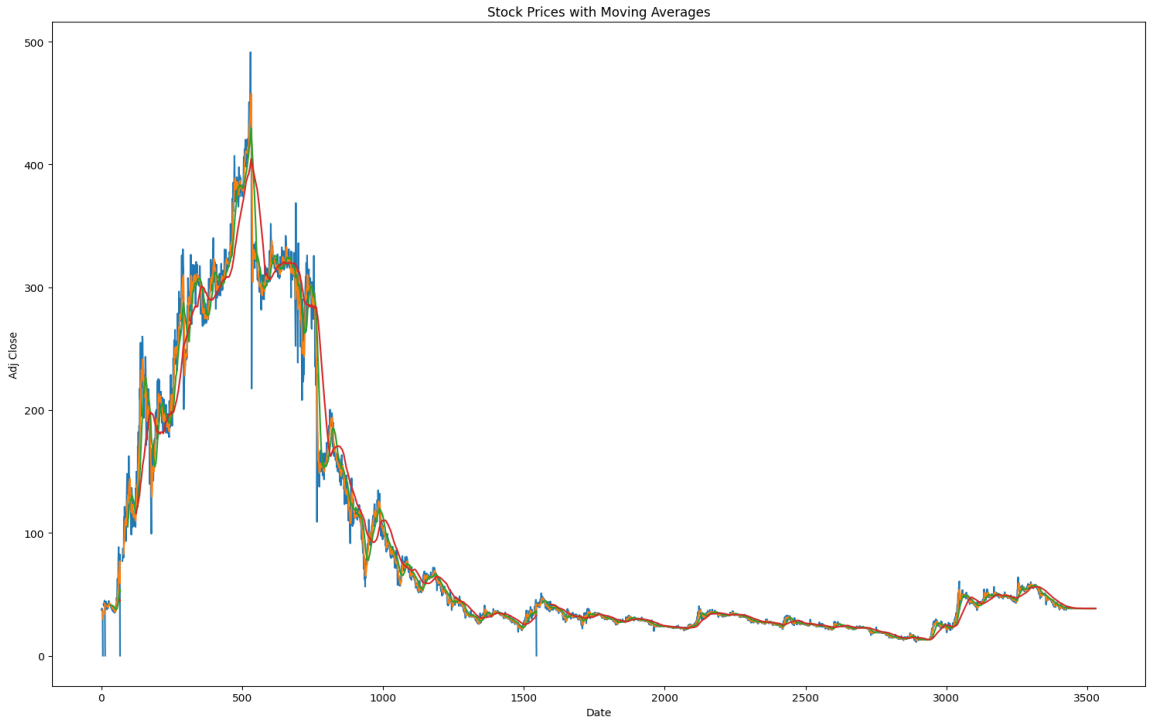


Figure 3.10: Moving Average (Beximco)

By calculating a series of averages from various subsets of the entire data set, a moving average is a technical analysis technique that is frequently used to examine data points. In essence, it is a method for emphasizing the underlying trend or cycle in a time series and mitigating short-term variations. It averages the price after a certain time period rather than daily average of price.

.

In figure 3.11 the scatter plot represents the correlation between two stocks closing price. Each point of the plot defines a single day. The closer the points are to the diagonal line; they are considered as correlated.

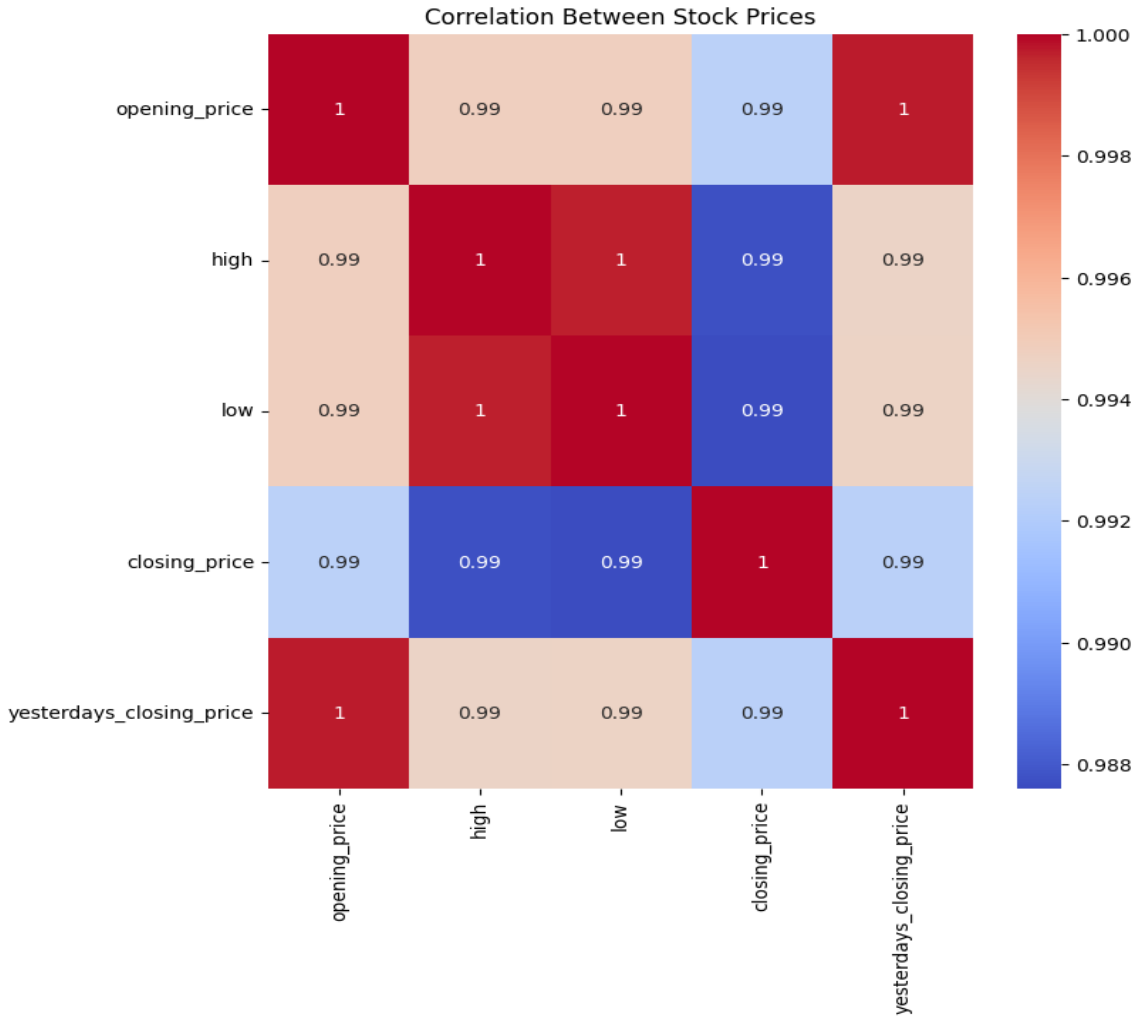


Figure 3.11: Heatmap (BRAC)

In figure 3.12, the correlation given below is the two-stock price. Coefficient of 1 means two stocks are perfectly correlated and they move the same direction. The two stocks are not correlated when coefficient is 0.

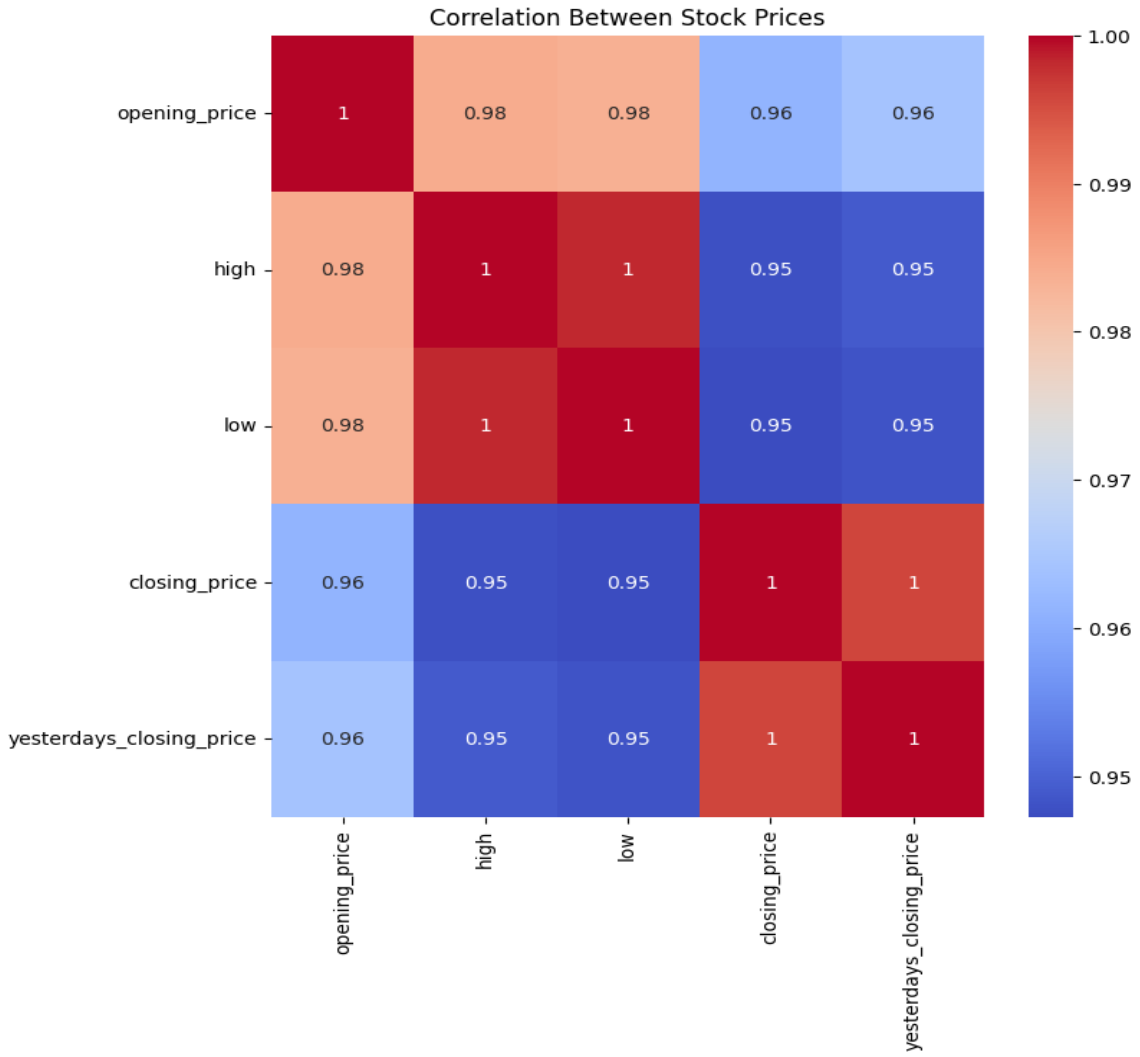


Figure 3.12: Heatmap (Grameenphone)

The scatterplot represents with 2 color, red and blue which is at upper and lower left and right corner of the heatmap. This means there are more stock that are correlated than the stock that are less correlated.

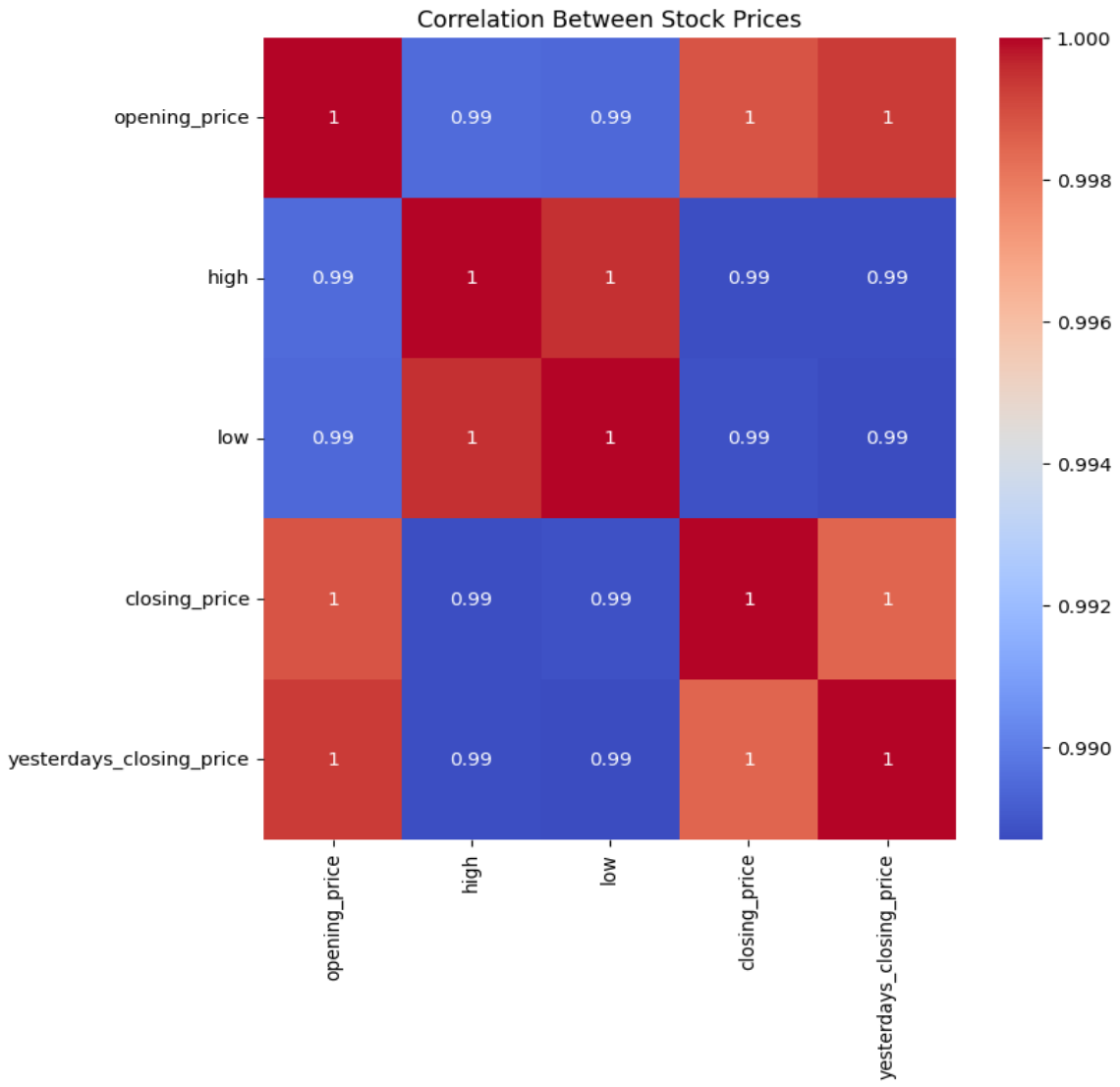


Figure 3.13: Heatmap (Beximco)

In figure 3.14, the histogram shows the stock price of BRAC for a certain period of time. the x axis represented the closing price and the y axis represent the days the stock close at that price.

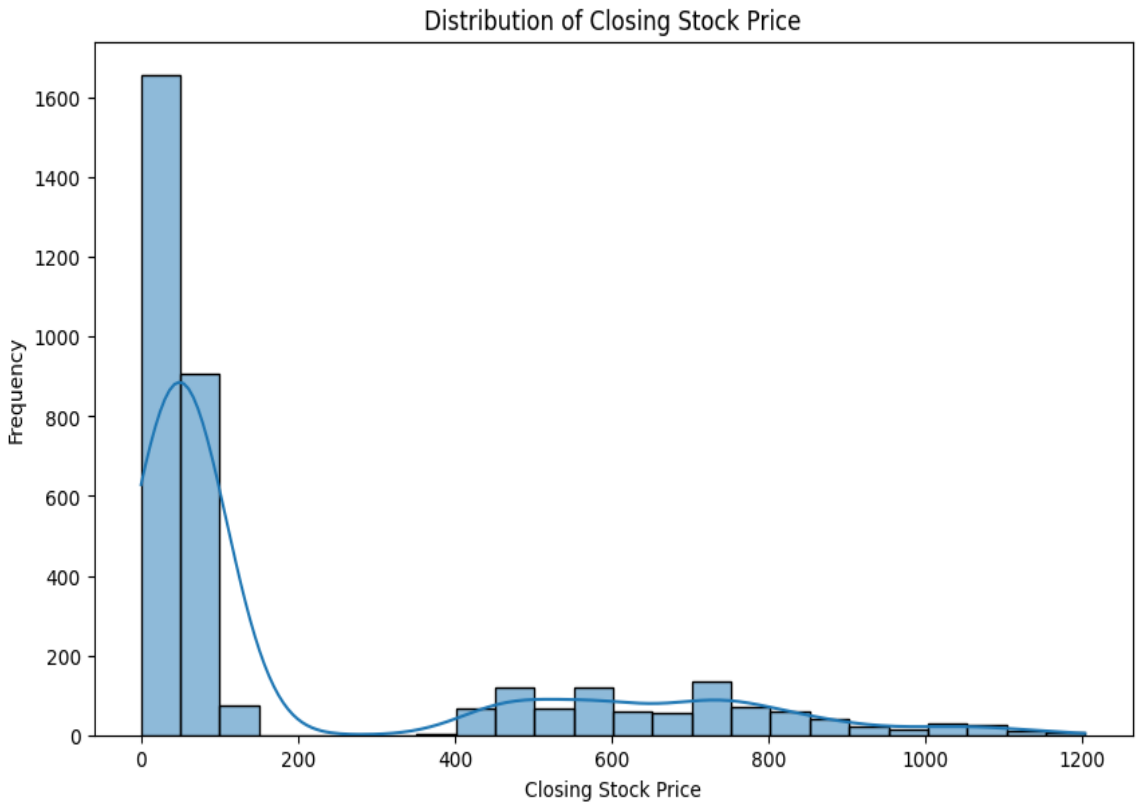


Figure 3.14: Distribution of closing stock price (BRAC)

In figure 3.15, the histogram is roughly symmetric means the stock price is near the average closing price. The histograms bumps and peaks represent the few days when the stock closed at those prices more often.

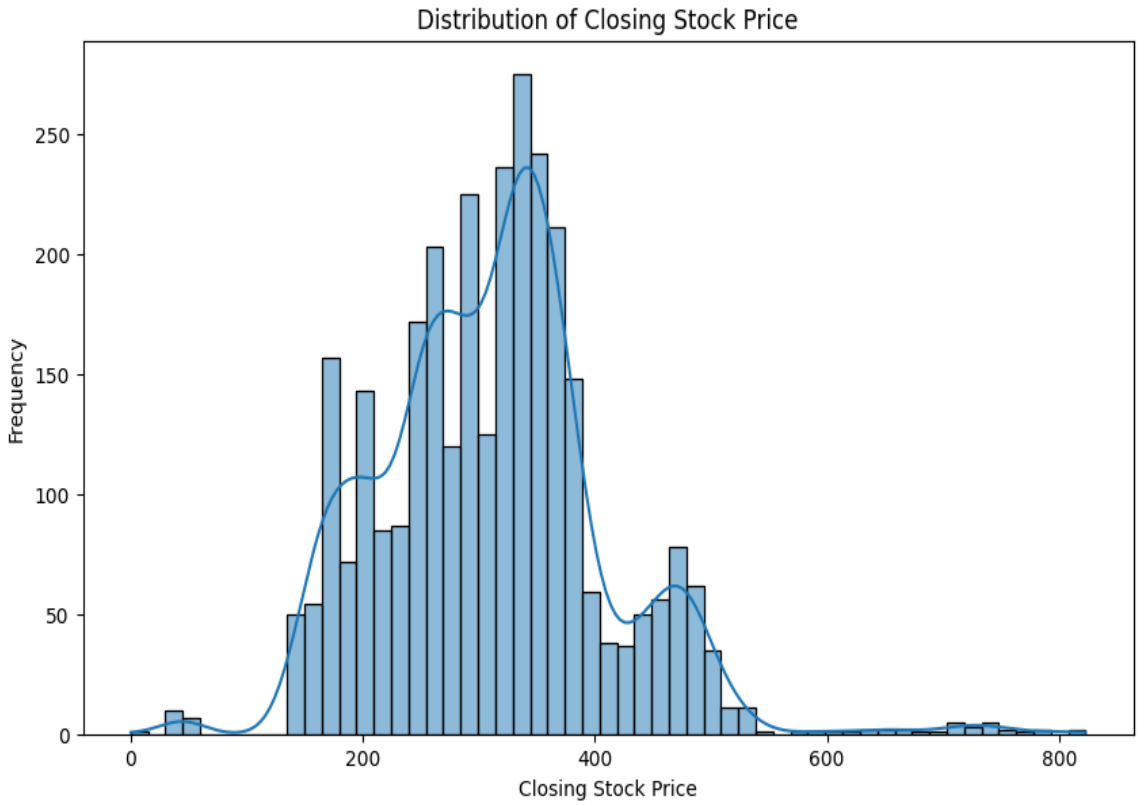


Figure 3.15: Distribution of closing stock price (Grameenphone)

In figure 3.16, the specific number on axis will depend on the amount of stock and the time period. The histogram will tell you whether the stock price is near average or higher or lower the average.

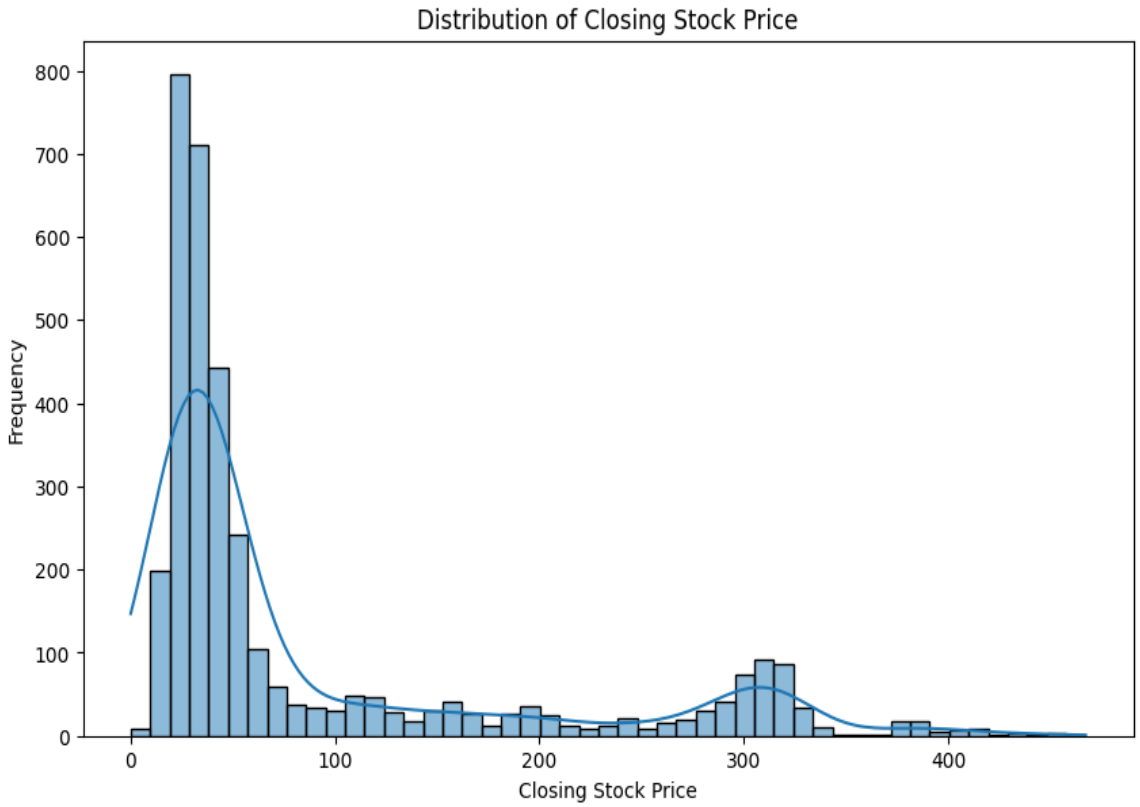


Figure 3.16: Distribution of closing stock price (Beximco)

Allocation or arrangement of closing costs relates to the amount or arrangement of an item's closing costs over a specific period of time. This classification provides great insights into stock price dynamics, volatility and patterns. Analyzing the distribution of closing prices is important for investors, traders, and analysts because it helps understand the risks and benefits associated with specific items.

3.5 Implementation Requirements

This information should be used for verification after completing the next task. I have to finish main divisions to reach the goal. These steps must be followed to meet the objective.

- Data collection
- Data pre-processing
- Data analysis
- Visualization of data
- Data Preparation
- Machine learning model creation
- Deep learning model creation
- Discuss about accuracy and achievement
- Creating a web prototype

In figure 3.17, the implementation requirements are shown in a diagram. The requirements will be followed for the further procedure.

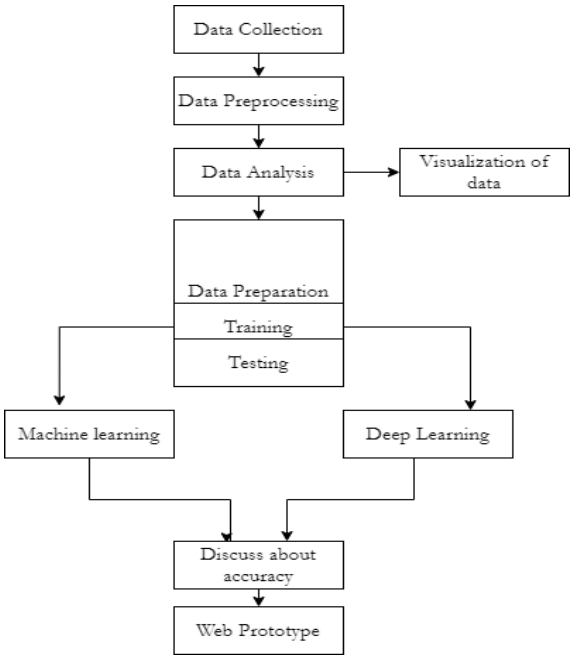


Figure 3.17: Implementation Requirement

I started developing the code after extracted the dataset. I evaluate the accuracy of all three model and came to a conclusion. But I used all the 3 model for prediction for my project.

1. Hardware and software
    - Operating system: Windows
    - HDD (1 TB or More)
    - RAM (Minimum 4 GB)
2. Tool
    - PyCharm Environment
    - Pycharm
    - Google Colab

CHAPTER 4
EXPERIMENTAL RESULT AND DISCUSSION

4.1 Introduction

The development process of stock price prediction system described here. Data collection, data preprocessing, data analysis, model selection, model implementation, evaluation and deployment of the model is discussed in this section.

4.2 Experimental Result

The development process of stock price prediction system described here. Data collection, data preprocessing, data analysis, model selection, model implementation, evaluation and deployment of the model is discussed in this section. For my project, I have used two machine learning model and deep learning model. I will shortly describe the outcome of my model. I have used 3 datasets of different company.In table 4.1, it represents the time period of each dataset. BRAC dataset contains data from 2008 to 2022, Grameenphone from 2010 to 2022 and Beximco is from 2008 to 2022.

Table 4.1: Dataset and year

Company	Year
BRAC	2008-2022
Grameenphone	2010-2022
Beximco	2008-2022

In table 4.2, The accuracy for BRAC dataset of LR is 98.36% which has the better accuracy, SVM 89.99% and LSTM 95.99%.

Table 4.2: Accuracy for BRAC

Dataset	Algorithm	Accuracy
BRAC	Linear Regression	98.36 %
BRAC	SVM	89.99 %
BRAC	LSTM	95.99 %

In table 4.3, the maximum accuracy for Grameenphone is from Linear regression which is 99.39%, SVM 81.5% and LSTM 98.8 %.

Table 4.3: Accuracy for Grameenphone

Dataset	Algorithm	Accuracy
Grameenphone	Linear Regression	99.39 %
Grameenphone	SVM	81.5 %
Grameenphone	LSTM	98.8 %

In table 4.4, Linear Regression perform the best for Beximco with the accuracy of 99.78%. SVM and LSTM perform 95% and 92.18% respectively.

Table 4.4: Accuracy for Beximco

Dataset	Algorithm	Accuracy
Beximco	Linear Regression	99.78 %
Beximco	SVM	95.37 %
Beximco	LSTM	88.62 %

4.3 Descriptive Analysis

I developed a web-based application for stock price prediction. The application was build using python. The web-based application has a homepage and a project page in figure 4.1. I can select dataset and algorithm from the interface. The result and analysis are visible on the screen.

Figure 4.1: Homepage

In figure 4.2, the project page is shown where we can select the dataset and algorithms. It also shows the detailed information of the analysis and prediction. The visual representation and prediction part will change per dataset.

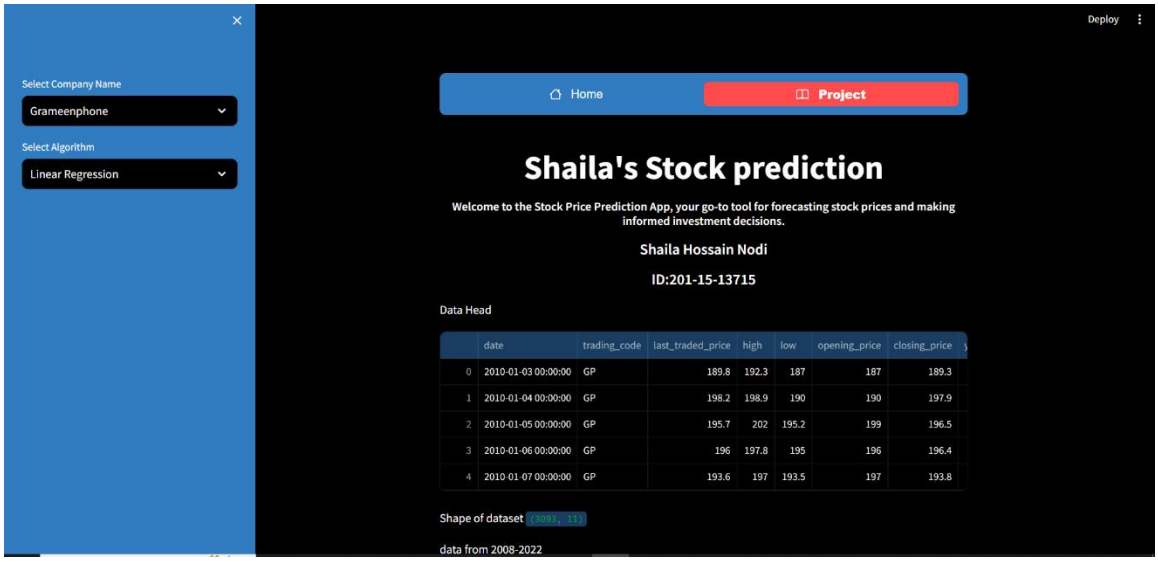


Figure 4.2: Project page

In figure 4.3 I can select company name and algorithm. The both can be selected according to the need of the user. The company name contains the dataset from 2008 to 2022. More data can be added in the company name section.

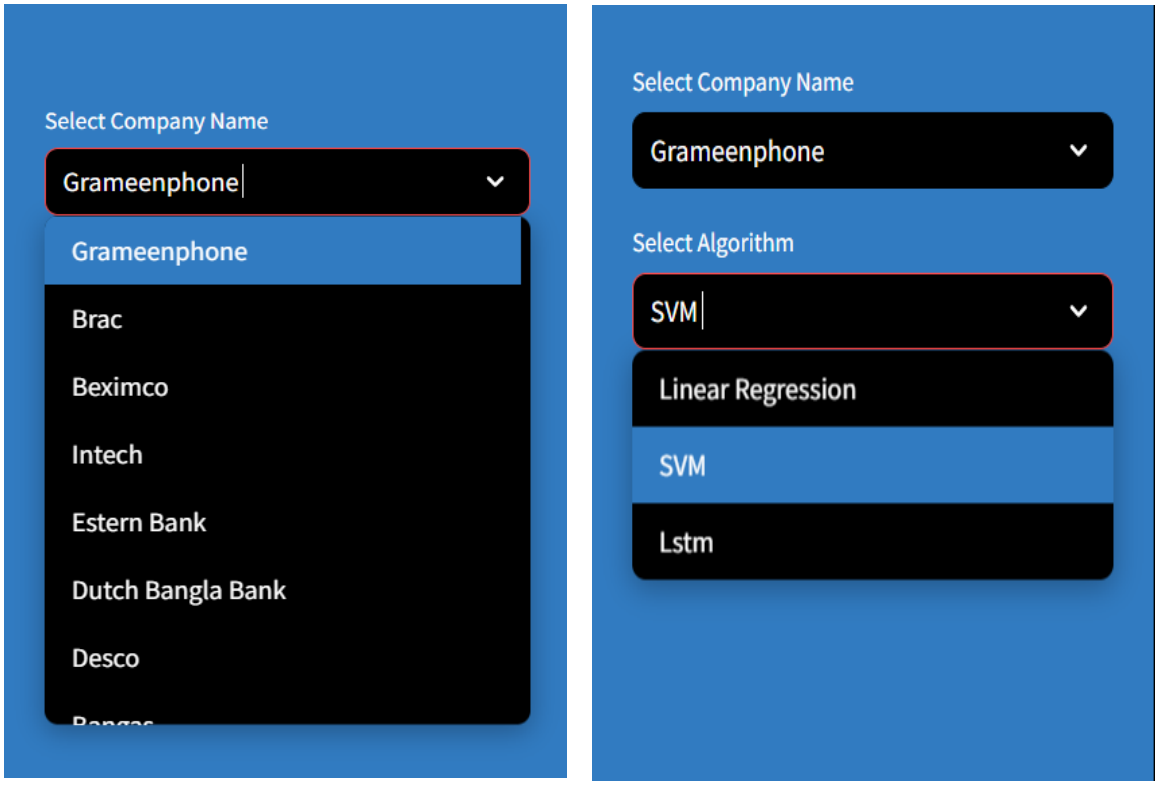


Figure 4.3: Data and model selection

.

In figure 4.4, The first 5 data and the shape of the data is printed on the screen. The interface will change the information according to the company name.

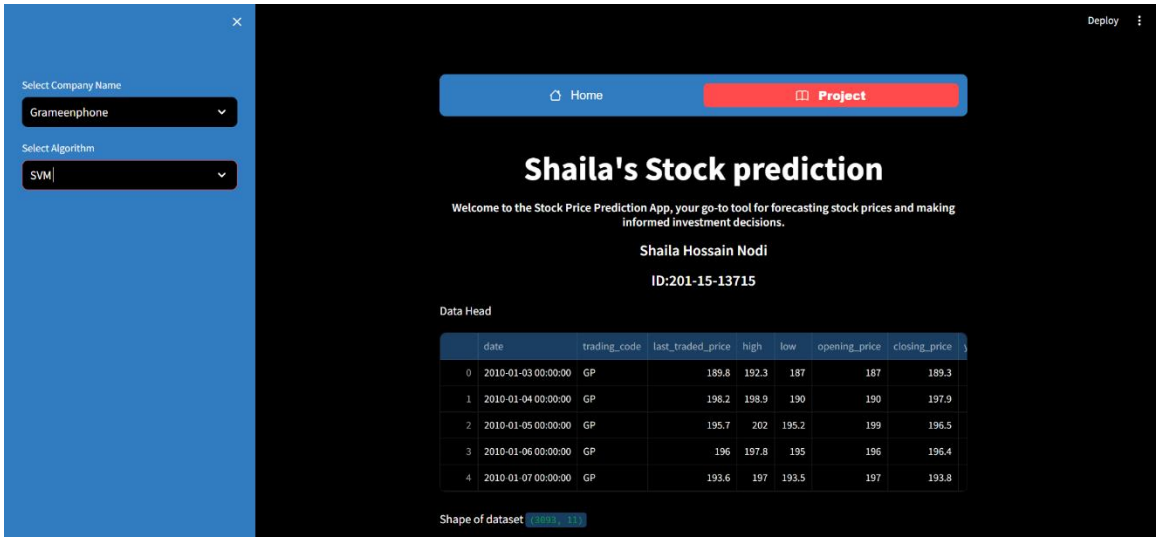


Figure 4.4: Data sample

In figure 4.5, the interface shows the description of the dataset where I can see the total count, mean, std, min, max, 25%, 50%, and 75% respectively.

	last_traded_price	high	low	opening_price	closing_price	yesterdays_closing_price	trade
count	3,093	3,093	3,093	3,093	3,093	3,093	
mean	309.5167	313.2473	306.2546	310.9762	312.1002	312.2628	1,143
std	100.8932	102.7812	99.8165	100.3493	97.0478	97.3492	1,179
min	0	0	0	0	0	41.6	
25%	250	252.5	247	250	251.3	251.3	
50%	316.4	318.7	313.1	316.8	317.3	317.3	
75%	360.9	364.3	359	361.9	361.3	361.3	
max	820.25	860	820	860	822.75	856.25	

Figure 4.5: Data Description

In figure 4.6, the interface shows the graph of opening price for the selected dataset. The x-axis represents the date and y-axis represent the price of stock.

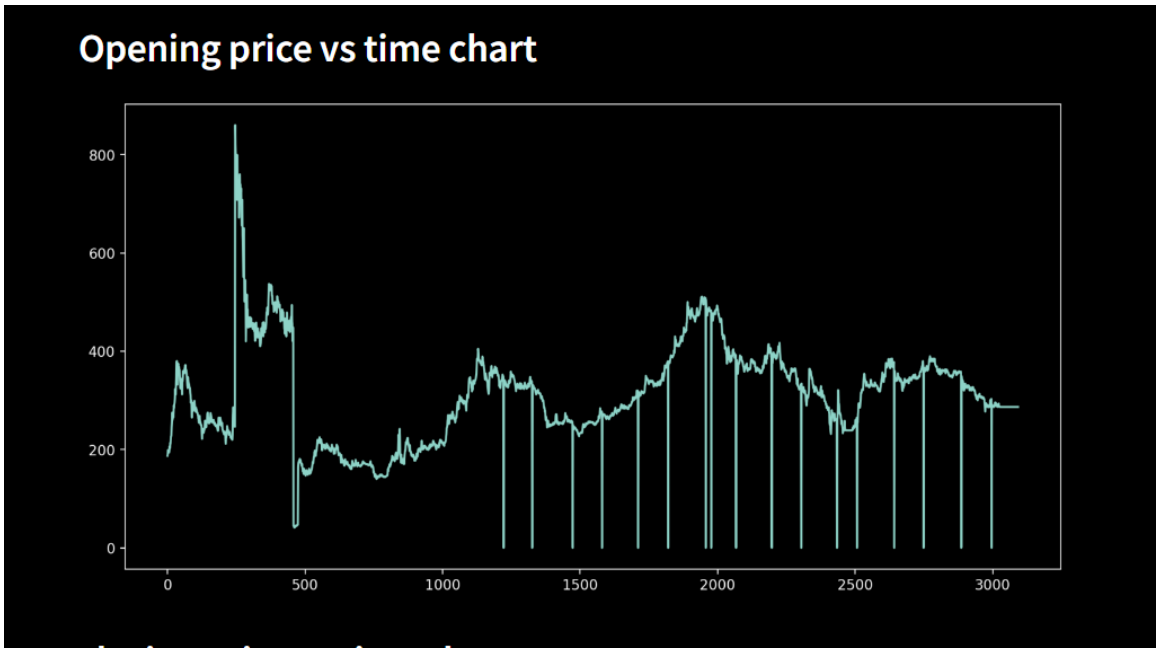


Figure 4.6: Opening price vs Time chart

In figure 4.7, the interface shows the graph of closing price for the selected dataset. The x-axis represents the date and y-axis represent the price of stock.

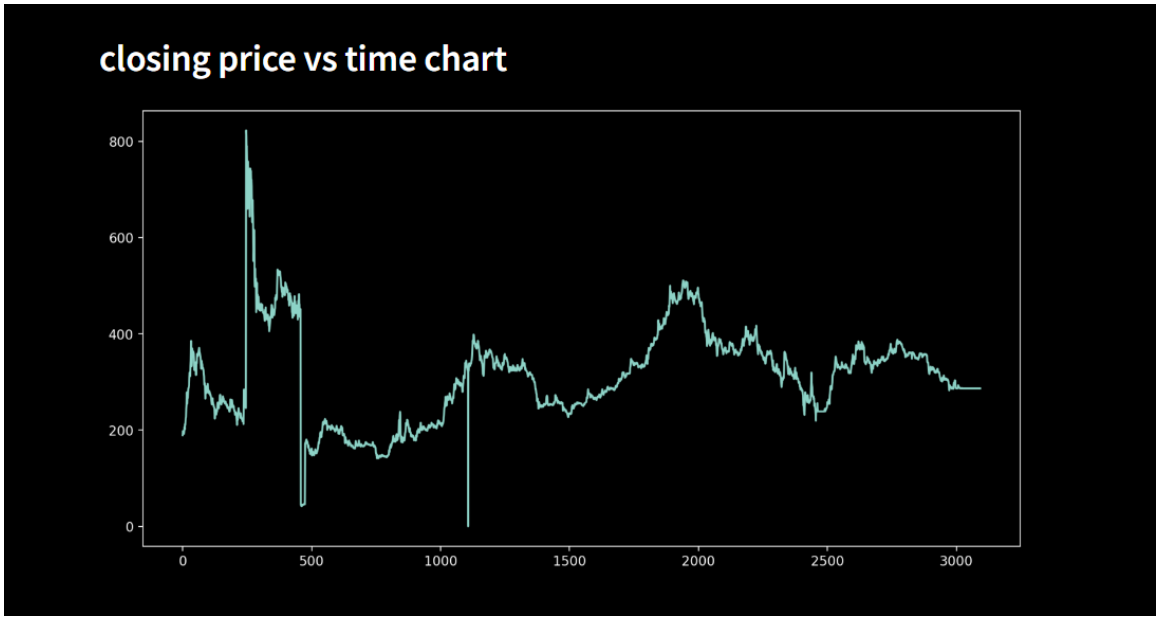


Figure 4.7: Closing price vs Time chart

In figure 4.8, the interface shows the graph of actual price and predicted price for the selected dataset. The selected machine learning model predict the price and make a comparison graph. The x-axis represents the date and y-axis represent the price of stock. It also shows the model accuracy and predicted price for next days.

Figure 4.8: Actual vs Predicted price

4.4 Discussion

- By using the interface, I can select the dataset and algorithm. The model is successfully showing the analysis and prediction for selected dataset. It also can print the model accuracy and price prediction for next day.

- In figure 4.8, actual price vs predicted price is shown for Grameenphone using support vector machine. The algorithm can be changed and the interface will print the output.

CHAPTER 5
IMPACT ON SOCIETY, ENVIRONMENT AND SUSTAINABILITY

5.1 Impact on Society

Using machine learning to predict stock prices can have a huge impact on people in a variety of ways. On the positive side, it provides investors with information to make better decisions, thus improving business management and reducing risk. Accurate forecasts can make financial markets more efficient by quickly incorporating new information into stock prices. These technological developments may lead to the emergence of financial instruments and services. However, it is important to consider the possible downsides. The problem is that powerful institutions with very good educational standards can control the economy. It is also possible to rely too much on these criteria and ignore other important factors when making an investment decision. Factors such as potential unemployment in financial markets, ethical dilemmas, and biases in information and decision-making processes must be taken into account. To address these issues, it is important to prioritize disclosure and transparency of machine learning models, establish guidelines and quality control, measure and evaluate performance standards, and promote education about the risks and limitations of these technologies.

5.2 Impact on Environment

Using machine learning in stock price prediction will have a direct or indirect impact on the environment. On the other hand, the needs to run and train machine learning models can lead to more effort and the potential to produce more carbon. Training complex models often requires large calculations and can have environmental impacts if not optimized for energy efficiency. Hardware that supports machine learning, including servers and data centers, can require a lot of power, especially when using non-renewable power sources. The increasing demand for processing power through machine learning will worsen environmental problems. On the plus side, advances in machine learning algorithms have the potential to improve many industries, including business. These technologies can contribute to ecologically sustainable businesses and activities by improving capital

allocation through the development of accurate forecasts and informed decision-making. Many measures have been taken to reduce the environmental impact of machine learning. These include developing energy-saving algorithms, using renewable energy in data centers, and focusing on best practices in the technology industry. Despite concerns about the environmental impacts of machine learning, continued efforts to promote sustainable practices and improve energy efficiency can help reduce these impacts. Prioritizing knowledge of ecological methods when creating and using machine learning, such as stock price prediction algorithms, is crucial for business technology.

5.3 Ethical Aspects

Using machine learning to predict stock prices has the potential to have a huge impact on people, both positive and negative. On the positive side, it improves business management and reduces risk by providing investors with the tools to make informed decisions. This approach can improve the overall results of the financial market by combining important information in stock prices. It can also spur innovation in financial services. But there are concerns that organizations with strong educational standards may dominate the market. Over-reliance on these models will lead to over-reliance on investment decisions. Ethical issues such as unemployment in the financial sector need to be taken into account, as well as impartial information and decision-making processes. To solve these problems, it is important to clarify the machine learning model, use appropriate rules, maintain it regularly, and make the public aware of the risks involved.

CHAPTER 6

SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH

6.1 Summary of the Study

I have used different dataset with different model to determine the best outcome. Among all the model linear regression outperform with all the dataset. I have successfully predicted the stock price and the figure is given here. Overall, the model is helpful to discover the movement of stock market.

6.2 Conclusion

The goal of my project was building a system that can predict stock price based on different dataset. I successfully developed my model that can analyze time series data and can predict future stock price. The system can precisely predict the price and it can visualize the accuracy level. I also develop a web application where I can choose dataset and preference model to predict the price and visualize the data. The project is successfully running and meet the requirement goal.

6.3 Recommendations

This system can predict the stock price with high accuracy and has a useful web application. The interface is built in a way that a user can easily use the system for prediction. Economist and analyzer can also modify and use this model based on their study. I recommend to use multiple machine and deep learning model. Deep learning model can handle large number of features and continuous data more precisely. More dataset and algorithm can be added in the web application.

6.4 Implication of further study

- Working with more data and features to predict more accurately
- More model and feature selection techniques could be implemented.
- Different programming languages and platforms could be more useful.
- The web interface could fetch data automatically every day from the internet.
- The system could predict the price for more days or month.

References

[1] O. S. S. A. S. Osman Hegazy, "A Machine Learning Model for Stock Market," *IJCST,* vol. 4, no. 12, pp. 17-23, 2013.

[2] S. V. R. A. K. Vaishnavi Gururaj, "Stock Market Prediction using Linear Regression and Support Vector," *International Journal of Applied Engineering Research ISSN,* vol. 14, pp. 1931-1934, 2019.

[3] B. S. Madge S, "Predicting stock price direction using support vector machines," *Independent work report spring,* p. 45, 2015.

[4] R. VK, "Stock market prediction using machine learning," *International Research Journal of Engineering and Technology (IRJET),* vol. 5, no. 10, pp. 1033-5, 2018 Oct.

[5] W. D., "Prediction of stock price based on LSTM neural network," in *In2019 international conference on artificial intelligence and advanced manufacturing (AIAM)*, 2019 Oct 16.

[6] M. M. A. A. Sunny MA, "Deep learning-based stock price prediction using LSTM and bi-directional LSTM model," in *In2020 2nd novel intelligent and leading emerging sciences conference (NILES)*, 2020 Oct 24 .

[7] S. a. C. C. K. Kompella, "Stock market prediction using machine learning methods," *International Journal of Computer Engineering and Technology,* vol. 3, no. 2019, p. 10, 2020.

[8] J. H. Z. T. Shen S, "Stock market forecasting using machine learning algorithms," in *Department of Electrical Engineering, Stanford University*, Stanford, 2012.

[9] O. D. a. H. S. A. Madeeh, "An efficient prediction model based on machine learning techniques for prediction of the stock market," *In Journal of Physics: Conference Series,* vol. 1804, p. 012008, 2021.

[10] M. P. N. G. M. a. T. V. Gurjar, "Stock market prediction using ANN," *International Research Journal of Engineering and Technology ,* vol. 3, pp. 2758-61, 2018.

[11] W. S. Tsai CF, "Stock price forecasting by hybrid machine learning techniques," *InProceedings of the international multiconference of engineers and computer scientists ,* vol. 1, p. 60, 2009.

[12] C. D. T. V. K. A. Vijh M, "Stock closing price prediction using machine learning techniques," *Procedia computer science,* vol. 167, pp. 599-606, 2020.

[13] M. T. A. G. M. Z. M. R. M. M. N. M. a. M. T. I. Mahtab, "Stock Price Prediction: An Incremental Learning Approach Model of Multiple Linear Regression," *AIUB Journal of Science and Engineering (AJSE) ,* vol. 3, no. 2021, pp. 159-166, 2022.

[14] T. Z. K. M. S. E. a. M. U. Us-Salehin, "Stock market prediction in bangladesh perspective using artificial neural network," 2022.

[15] A. M. H. S. I. M. H. F. S. A. Das S, "Bangladeshi Stock Price Prediction and Analysis with Potent Machine Learning Approaches," in *InCyber Security and Computer Science: Second EAI International Conference, ICONCS 2020*, Dhaka, Bangladesh, 2020.

[16] M. A. G. a. A. C. Bansal, "Stock market prediction with high accuracy using machine learning techniques," *Procedia Computer Science* , vol. 215, pp. 247-265, 2022.

[17] Y. J. Heo J, "Stock price prediction based on financial statements using SVM," *International Journal of Hybrid Information Technology,* vol. 9(2), pp. 57-66, 2016.

[18] Y. a. I.-C. C. Wang, "Market index and stock price direction prediction using machine learning techniques: an empirical study on the KOSPI and HSI," *arXiv preprint arXiv:1309.7119* , pp. 1-13, 2013.

[19] P. M. Pramod BS, "Stock price prediction using LSTM," *Test Engineering and Management,* vol. 83, pp. 5246-51, 2020.

[20] A. e. a. Subasi, "Stock market prediction using machine learning," *Procedia Computer Science* , vol. 194, pp. 173-179, 2021.

STOCK PRICE PREDICTION AND FORECASTING WITH MACHINE LEARNING

ORIGINALITY REPORT

7%	6%	1%	4%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	3%
2	dspace.daffodilvarsity.edu.bd:8080 Internet Source	2%
3	Submitted to Stadio Holdings Student Paper	<1%
4	www.jku.at Internet Source	<1%
5	espace.curtin.edu.au Internet Source	<1%
6	eudl.eu Internet Source	<1%
7	www.investopedia.com Internet Source	<1%
8	Pilar Gómez, Angela Nebot, Sabrine Ribeiro, René Alquézar, Francisco Mugica, Franz Wotawa. "Local Maximum Ozone Concentration Prediction Using Soft	<1%

Computing Methodologies", Systems Analysis Modelling Simulation, 2003
Publication

9	Hershey, Marc Alan. "Applications of Photoplethysmogram Signal Analysis for Hemodynamic Modeling", Villanova University, 2023 Publication	<1%
10	Ubaida Fatima, Sadia Kiran, Muhammad Fouzan Akhter, Muhammad Kumail, Jaweria Sohail. "Unveiling the Optimal Approach for Credit Card Fraud Detection: A Thorough Analysis of Deep Learning and Machine Learning Methods", Research Square Platform LLC, 2024 Publication	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		