# Ascertaining and Resolving Market Instability by Foretelling Ginger Price in Bangladesh Using Machine Learning Approach

BY

Ahasan Habib Polton

152-15-6052

This Report Presented in Partial Fulfillment of the Requirements for

The Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

Ms. Tapasy Rabeya

Senior Lecture

Department of Computer Science and Engineering

Daffodil International University

Co Supervised By

Krishno Dey

Lecture

Department of Computer Science and Engineering

Daffodil International University


**DAFFODIL INTERNATIONAL UNIVERSITY**

**DHAKA, BANGLADESH**

**23 January 2023**

## APPROVAL

This Project/internship titled "Ascertaining and Resolving Market Instability by Foretelling Ginger Price in Bangladesh Using Machine Learning Approach", submitted by Ahasan Habib Polton, ID No: 152-15-6052 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 23 January 2023.

## BOARD OF EXAMINERS

**Chairman**

**Dr. Touhid Bhuiyan**
**Professor and Head**
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

**Internal Examiner**

**Dr. Md. Zahid Hasan**
**Associate Professor**
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

**Internal Examiner**

**Fahad Faisal**
**Assistant Professor**
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

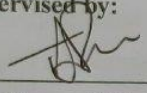
**External Examiner**

**Dr. Ahmed Wasif Reza**
**Associate Professor**
Department of Computer Science and Engineering
East West University

## DECLARATION

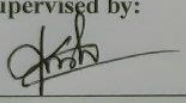
We hereby declare that, this project has been done by us under the supervision of Ms. Tapasy Rabeya, **Senior Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

**Supervised by:**



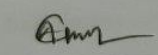
**Ms. Tapasy Rabeya**
**Senior Lecturer**
Department of CSE
Daffodil International University

**Co-Supervised by:**



**Krishno Dey**
**Lecturer**
Department of CSE
Daffodil International University

**Submitted by:**



Ahasan Habib Polton
ID: 152-15-6052
Department of CSE
Daffodil International University

# ACKNOWLEDGEMENT

First, we express our heartiest thanks and gratefulness to almighty God for His divine blessing making us possible to complete the final year project successfully.

We are really grateful and wish our profound our indebtedness to **Ms. Tapasy Rabeya**, Senior Lecture, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "Machine Learning and Artificial Intelligence" to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

We would like to express our heartiest gratitude to **Prof. Dr. Touhid Bhuiyan**, Head, of the Department of CSE, for his kind help to finish our project and also to other faculty members and the staff of CSE department of Daffodil International University.

We would like to thank our entire coursemates at Daffodil International University, who took part in this discussion while completing the coursework.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

# ABSTRACT

The unpredictability of markets and the fluctuation of prices is a feature that is commonplace throughout the entire world, but it is more widespread in Bangladesh. As a direct result of this, the costs of the things we use on a daily basis fluctuate in a predictable pattern. It suggests that there will be a significant effect on the component that we consume on a regular basis. In Bangladesh, ginger may be found in nearly all of the foods that we consume on a daily basis. For those who are living below the poverty line, maintaining a record of the price is of the utmost importance, but doing so manually is a laborious and time-consuming endeavor. Machine learning (ML) is the most effective method for performing this activity in a sufficient manner and for developing enormous techniques to forecast the price of commodities in order to deal with market volatility. In the course of our work, we aimed to make an estimate of the price of ginger by utilizing various ML applications. We constructed our dataset using the information that was available from the Ministry of Agriculture in Bangladesh. We used a variety of machine learning techniques, including K-Nearest Neighbor (KNN), Decision Tree, Neural Network (NN), Support Vector Machine (SVM), and Random Forest, to make a prediction about the price of ginger. To make an accurate forecast of ginger's expense, an accomplished algorithm of the highest caliber was used. I make an effort to determine whether the cost of ginger falls into the category of expensive (high), inexpensive (mid), or preferable by utilizing the methods that have been described previously (low).

# TABLE OF CONTENTS

## CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# CHAPTER 1

# INTRODUCTION

## 1.1 Introduction

Bangladesh is an agricultural country. Every year different types of crops and spices are produced in the country. It is known as the land of various kinds of spices. Among all the spices, ginger (an indispensable kitchen item) is one of the most quintessential and cultivated spices. It is used as a condiment more than spice. The country has a dictate for 0.65 million tons of ginger on an annual basis, of which it progenies 0.185 million tons regionally [1]. About 19,055 hectares of land of the country is employed to deliver ginger [2]. To consummation, the rest demand is conveyed mostly from India, Indonesia, China, and Myanmar. As the need is tremendous against the stock, the price is expanding posthaste. According to authoritative experts, trade for ginger is rising as a consequence, the cost of ginger diverges now and then. The price of ginger was 65 Taka on January 1st, 2020. The price fluctuates from 60 to 150 Taka in 2020s. By the end of 2021, the rate ended up to 200 Taka, which is moderately aerial from 2020. As the requirements of ginger risen day by day, the quantity persists the equivalent, proportionally, the price exaggerated. In 2022, it ascends to 200 to 220 Tk. The inflation and reduction of the costs of ginger are very incidental. The unexpected fluctuations in price placed destitute people of our nation in considerable contingency. The price movement affected regular survival inferior, particularly for the souls dwelling beneath the destitution boundary. Monetary forecasting is composite due to the disorderly nature of data ambiguity. A portion of other constituents like weather contingencies, storage constraints, potency, transportation, and stand apply-demand quotient address the prophecy even extra intriguing. At present, Artificial Intelligence (AI) rules over the world. AI and algorithms act like a human, and a portion of these algorithms concedes the scheme to learn automatically from the surroundings and apply the understanding to enhance enforcement externally performing unequivocally estimated. ML arranges the role of predicting the future of different sections like the weather forecast, share stocks, medical science, finance, transportation, consumer goods, and viand video games. Predicting future prices is one of the most prominent sectors for practice. M. Hasan et. al. [3] predicted future onion price using ML algorithms as it presents the safest process to divine the

future price. For monetary appraisal, ML implies a puissant influence on it. ML algorithms can ascertain the expected price of a commodity with even disorganized data. So, to foretell the price of ginger, we adopted machine learning methods on our ginger dataset to employ its adaptability. We used ML libraries like Scikit-Learn, Matplotlib, Numpy, Pandas, Tensorflow, Keras. These are convenient tools of ML algorithms. Various kinds of feature selection algorithms and extraction were applied to compose the dataset more expediently Besides, some algorithms provide optimal techniques to predict like reinforcement learning algorithms, Unsupervised learning algorithms, and supervised learning algorithms. As we are working on future price prediction, to perform it, we intend to use supervised machine learning because the diversity of time series data depends on period/seasons and regions as time-series forecasting can appear as supervised learning. The time-series method requires concocted data intents that happen over a particular time that is utilized to predict the future. Our main motive to cis categorize the predicted future price into three sections indicated as high, mid, low, and for a certain period to stable the market instability.

This exposition will be adapted as follows. In Section II review of related work is exhibited. In section III the proposed methodology stands illustrated. Section IV retains the expected results and correlates several outcomes that we acquired through various ML procedures to predict the ginger price. Finally, we terminated our paper with a conclusion bearing future works.

## 1.2 Motivation

Ginger is a widely used spice. There is a lot of demand for it in Bangladesh. There are very few recipes that do not use ginger in cooking. Morning tea with ginger has become a culture of Bangladeshis. So, there is a massive demand for it in the market. Hence price difference is observed. So, our model will be able to predict in advance what the price of ginger may change and accordingly increase the supply of ginger to keep the price within reach of the buyer. Ginger buyers will be able to buy ginger in this way. There will be no shortage of ginger in the market. Because the buyers will understand in advance what the price difference is and will buy ginger accordingly. Unscrupulous business people will not be able to profit much from this and destabilize the market by creating an artificial crisis of ginger by hoarding ginger. Because ordinary people, field-level ginger producers, and retail traders can understand the price difference of ginger in advance. We are able to make an accurate forecast of the future price of ginger by applying modern

technology. As a result, we contemplated applying a method that relies on machine learning in order to make adjustments.

## 1.3 Problem Definition

In today's world of cutting-edge information and communications technology, the concept of "machine learning" holds the ultimate importance (ICT). The development of our agriculture industry will be significantly facilitated by the application of machine learning, which will make its expansion a lot less difficult to achieve. In order to create a solution that can be put into action, it is vitally necessary first to describe the existing challenges and the requirements that are associated with them in this particular field. Only then can one begin to work on a solution. In order to be successful in the field of agriculture using machine learning, it is essential to have a solid understanding of not only the policies of the government but also the standards of the technology industry and the many different instructional options that are available.

## 1.4 Research Questions

The following are some of the most critical questions that this research will try to answer.

- It remains to be seen whether the reduction in ginger production due to climate has an impact on the price.
- With the increase in price, the demand for ginger in the market has to be determined.
- It should be shown whether the production of the country should be exported before meeting the market demand of the country.
- It has to be seen if there is any confusion at the field level.
- Is there any influence on dishonest traders?
- Is the correct price difference available?
- Which algorithm processes, if any, will be executed?

## 1.5 Research Methodology

The majority of the explanation for how we obtained data, preprocessed that data, categorized the data that was produced as a result of the data collection, selected algorithms, implemented those algorithms, and then evaluated those algorithms can be found in the section of our research article

titled "Methodology." You will find a definition of the output that the suggested model generates at the end of this section in this very document.

## 1.6 Research Objective

We plan to complete our goal of identifying, Ascertaining and Resolving Market Instability by Foretelling Ginger prices in Bangladesh by utilizing Data Mining and Machine Learning Techniques. This will be the method by which we will achieve our goal.

# CHAPTER 2

# BACKGROUND STUDY

## 2.1 Introduction

No amount of work or technology can determine the price of ginger or provide treatment in our location. As a consequence of this, the context in Bangladesh's agriculture industry and machine learning deteriorated. The subfield of artificial intelligence (AI) known as machine learning enables computers to teach themselves new skills and advance independently. Machine learning involves creating computer programs that can access data and self-educate. Although machine learning algorithms can learn, it's hard to define "learning" precisely. This is because there are multiple ways to get information from data, and they rely on how the machine learning algorithm is built. In general, learning requires a large amount of data that can predict responses to inputs. This is needed to learn. Each input-answer combination is an example of the algorithm, which learns faster with more. Because each input and response pair is within a problem domain represented by a line, cluster, or other statistical representation, this is true. This algorithm offers us the best results.

## 2.3 Background Study

To determine obstacles that remain obedient to forecasting, machine learning is universally accustomed. For exerting things approaching the price inconstancy of the commodities, a bunch of work has been performed by implementing ML. ML has executed this maneuver remarkably beneficial.

Z. H. Khan et. al. [4] outlined, numerous procedures have been performed to determine the rate in the share market, but none of these methods are established as a pleasantly satisfactory forecast tool. Artificial Neural Network (ANN) is a conventional technique to discover concealed and stored specimens in data, so they have adopted ANN to predict the rate. Here, Multilayer Feedforward networks as a network model and a Backpropagation algorithm for training sessions have been implemented. The prophesied rate and Error were 401.9 and 1.74% for 2 input datasets, 392.7 and 0.58% for 5 input datasets.

D. J. N. Srinivasan [5] stated a predicting system based on medium-term energy trade that will assist in prophesying energy demand and recognized services. The author has executed single double exponential smoothing, exponential smoothing, double moving average, Artificial Neural network, ARMA, and GMDH neural network. After executing all models, he affirmed that the Artificial Neural Network (ANN) archetypes gave more conventional results, especially GMDH illustrated adequately in constructing predictions that remained exceptionally further specific and more cramped lever concentrated than conventional time series and regression-based models. For the GMDH neural network, the prediction of the cumulative quantity series was 1.73.

R. Jammazi et. al. [6] operated on forming an oil price forecasting scheme. To get a protruding prediction of unrefined oil prices, a Hybrid model named HTW-BPNN and HTW-MBPNN has been staged. Three types of activation functions i.e. bipolar sigmoid, sigmoid, and hyperbolic tangent have been used to experiment with the model's compliance. Comparatively, the consequences of HTW-MBPNN accomplish improved than conventional BPNN.

D. Singhal et. al. [7] worked on the prognostication of market-clearing prices. They have suggested a neural network approach to foretell the market-clearing costs for day-ahead energy markets. There is a three-layer backpropagation network of the neural network model. The cost predicting rates applying the neural network pattern displays that the electricity price in the neutrality markets depends heavily on the trend in quantity charge and clearing price.

B. Park and associates [8] proposed using machine learning to forecast housing prices. There are 5359 townhomes analyzed in this study from Fairfax County, Virginia. These townhouses have been selected especially for you by the Multiple Listing Service (MLS) of the Metropolitan Regional Information Systems (MRIS). The study was carried out with the intention of making more accurate predictions of housing prices. The accuracy of home price predictions made using a variety of machine learning methods, such as C4.5, RIPPER, AdaBoost, and Naive Bayesian, has been analyzed. The ripper algorithm performed far better than the alternative methods.

F. A. de Oliveira et. al. [9] stated the forecast of the stock. In this article, the authors predict the current of the stock rather than continuously forecasting the hereafter price of the stock. The inclination can be interpreted as a model. They executed both short-term prophecies (day or week predictions) and also long-term prophecies. They perceived that the rearmost accomplished preferred outcomes with 79% accuracy. Based on the divined rendering the result assessment algorithm resolves to either buy, sell, or hold the stock.

S. Velankar et. al. [10] recommended a Machine Learning-based conformity to determine the indication of the regular price variation with the highest attainable precision. The authors have performed Bayesian Linear Regression and Generalized Linear Model (GLM) is also known as the Random Forest model. After the conclusive prophecy, the accuracy can be distinguished with various models.

S. J. I. J. I. C. T. Pudaruth [11] wanted to expand the system with the prediction of depreciated cars price utilizing ML algorithms in their paper. The author assembled data from various daily newspapers. To attain the forecast price of used cars, he implemented Linear regression, KNN, Naive Bayes, and a Decision Tree. The author applied the Confusion Matrix to analyze the accuracy of distinct algorithms.

N. Gandhi and his colleagues [12] made a prediction on the production of prices. Rice production in India is being pushed to its limits in order to keep up with the nation's burgeoning population, which is driving up demand for the staple food. In order for them to do their task in a timely manner, they relied heavily on the SVM technique, the SMO algorithm, and the confusion matrix. After first deleting any extraneous data that existed before to the implementation of the algorithms, they next protected the data and changed it into leveled data. Finally, they executed the algorithms. In addition to that, the SMO algorithm was employed in order to arrive at a conclusion regarding the matter at hand. After getting the findings of the suggested model, the situation was looked at using the confusion matrix..

The inconsistency of the market can be eliminated by making projections on the prices of goods. The significance of ML in the TAC market game was brought to light by M. P. Wellman and colleagues [13]. The implementation of price prediction relies heavily on historical averaging, competitive analysis, and machine learning (ML). The application of ML has been used in order to get the highest possible accuracy. In order to wrap up the survey, the author employed TAC02's concluding rounds. They ignored the introductory rates, thus they estimated that the EVPP would be 49%.

From the discussion of the preceding activities, we can discover a few correspondences and contrasts with our research. We have observed that there is no activity in foretelling ginger prices previously. So, we adjudicated to determine the market volatility of ginger prices by forecasting future prices.

## 2.3 Comparative Analysis and Summary

After reviewing a few research articles and projects, I chose machine learning methods because,

- They suit well with the natural environment, agriculture, and other classifications. It meets our goals best.
- ML is the most accurate image classification method, with 100% accuracy.
- It's easy to use and expandable.
- Using ML appropriately and with enough training can yield good results.

# CHAPTER 3

# RESEARCH METHODOLOGY

## 3.1 Introduction

The approach consists of a total of six steps, as shown in Figure 1, which are used to complete the analysis. The following are the steps that must be taken:

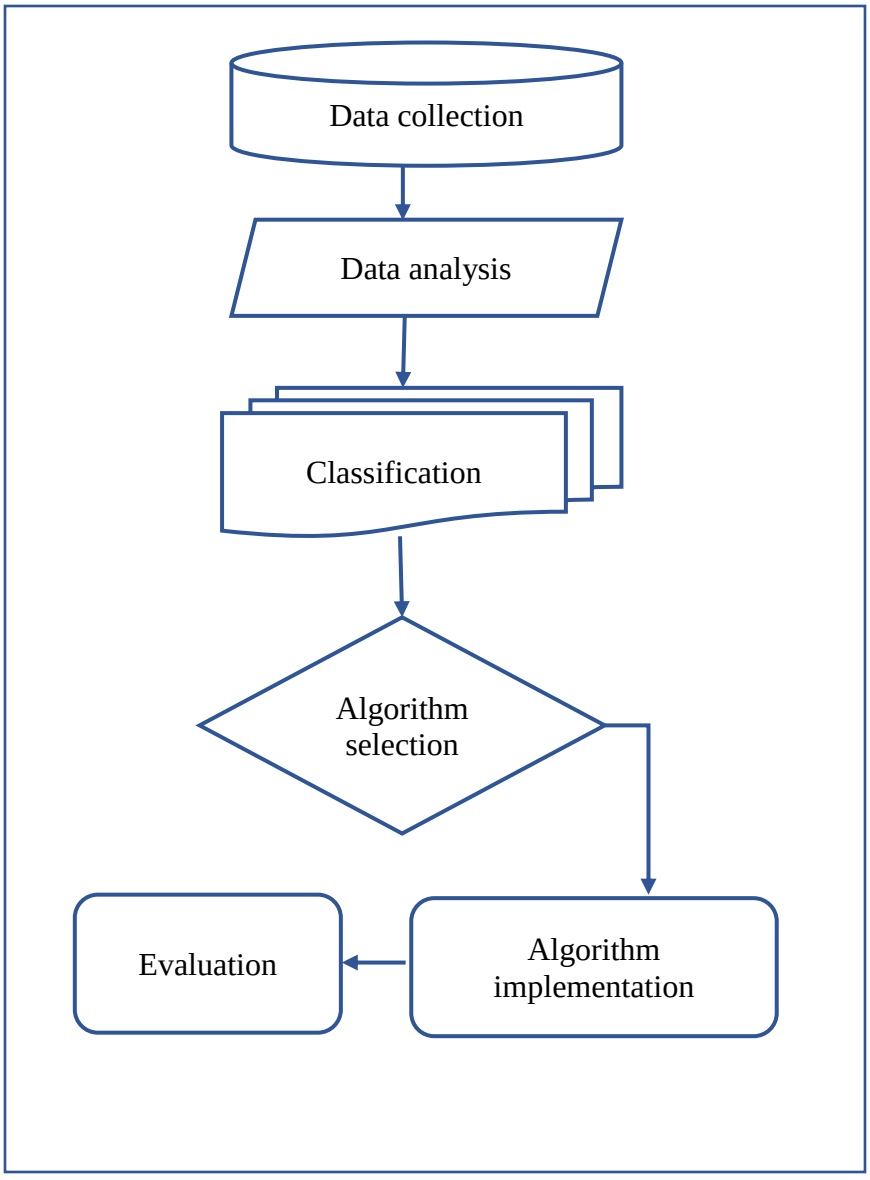



Figure 3.1: Methodology Diagram

## 3.2 Data Collection Procedure

The accumulation of data is a challenging task that each and every interpretation must take on. The website of Bangladesh's Ministry of Agriculture served as the source for our desired dataset, which we gathered. Following the completion of the data assembly, we then divided our dataset into two halves. The initial portion was used for the purposes of training and testing, and the remaining part was put to use in the forecasting process. In the first part of the analysis, we used 1460 daily prices from the years 2020 and 2021 in two separate cities: Dhaka and Chattogram. These were our bases for comparison. And for the part that deals with projections, 120 different daily prices for the year 2020 were employed from Dhaka city.

### Data Analysis

| Year | Date | Month | Season | Location | Price | Classes |
|---|---|---|---|---|---|---|
| 20 | 17 | 6 | 1 | 1 | 120 | 1 |
| 19 | 1 | 8 | 1 | 0 | 90 | 0 |
| 21 | 31 | 1 | 4 | 0 | 190 | 2 |
| 20 | 18 | 1 | 4 | 1 | 80 | 0 |
| 19 | 4 | 5 | 0 | 0 | 70 | 0 |

Table 3.1 Database Description

The sample of the set of data and the attributes we prefer for obtaining a suitable dataset is given in Table I. In our dataset, we perceived that the price of ginger is high and fluctuated more in the year 2021, so we picked the year as our attribute. As our database includes time-series data, we took the day as our principal attribute. The price relies on months. So we selected month as one of our key attributes. The different season holds different weather. And good or bad weather changes

the production of ginger and we collected the price of ginger from Dhaka and Chattogram. We observed the difference in prices, so we took season and location as our attributes. Our main attribute is the price. We desired to predict the future
price of ginger. The prices are categorized into three classes (high, low, and mid) regarding the price scale. So we decided to take the category attribute to classify our dataset. After a peculiar study, we apprehended that it is unstructured and time-series data. We remained concerned that time-series data is fluctuating over time.

## 3.3 Data Classification

Regarding the ginger price, I separated our dataset into the following three categories: According to our interpretation, these groupings were divided into three categories: desirable (low), economical (mid), and pricey (high).

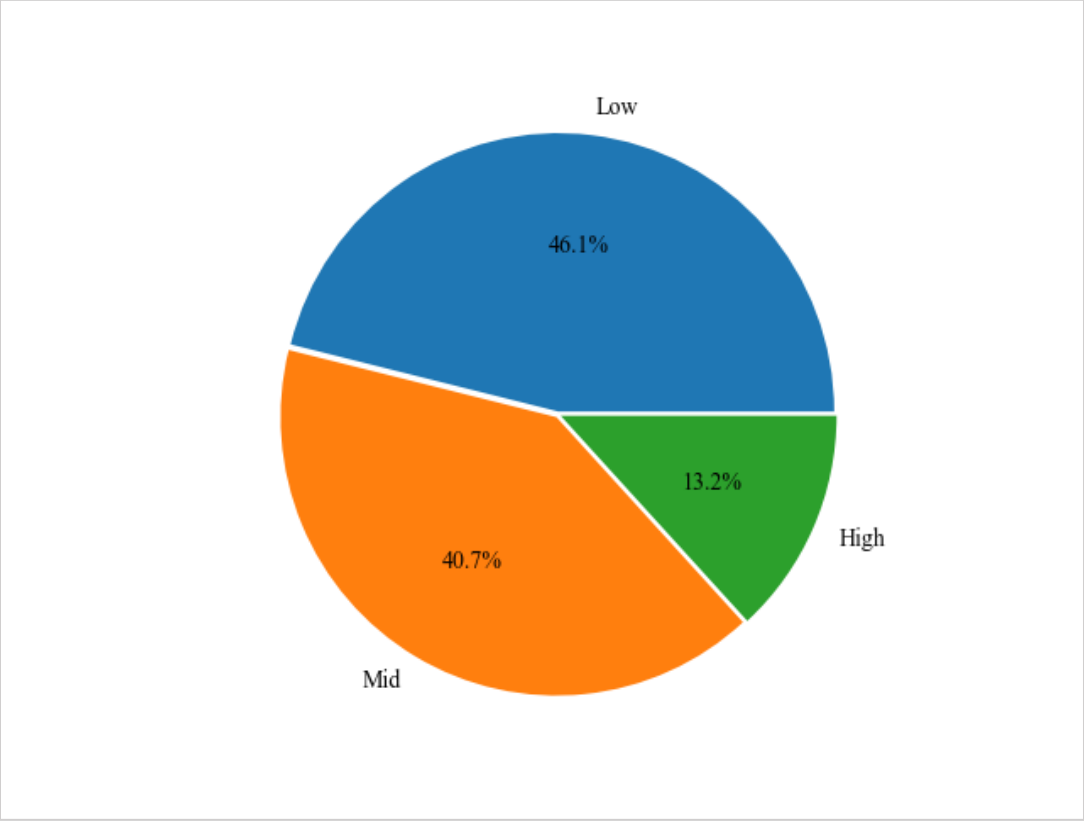



Figure 3.2: Classification of Database

Figure 3.2 includes a pricing scale that not only describes the categorization but also displays the percentage of each class that is present in the full dataset. The proper range for these three classes is 0-107, 108-163, and 164 and above respectively.

In addition, the price categories for the years 2020 and 2021 are depicted in Figure 3.3. Where the X-axis represents the years, and the Y-axis represents the days.

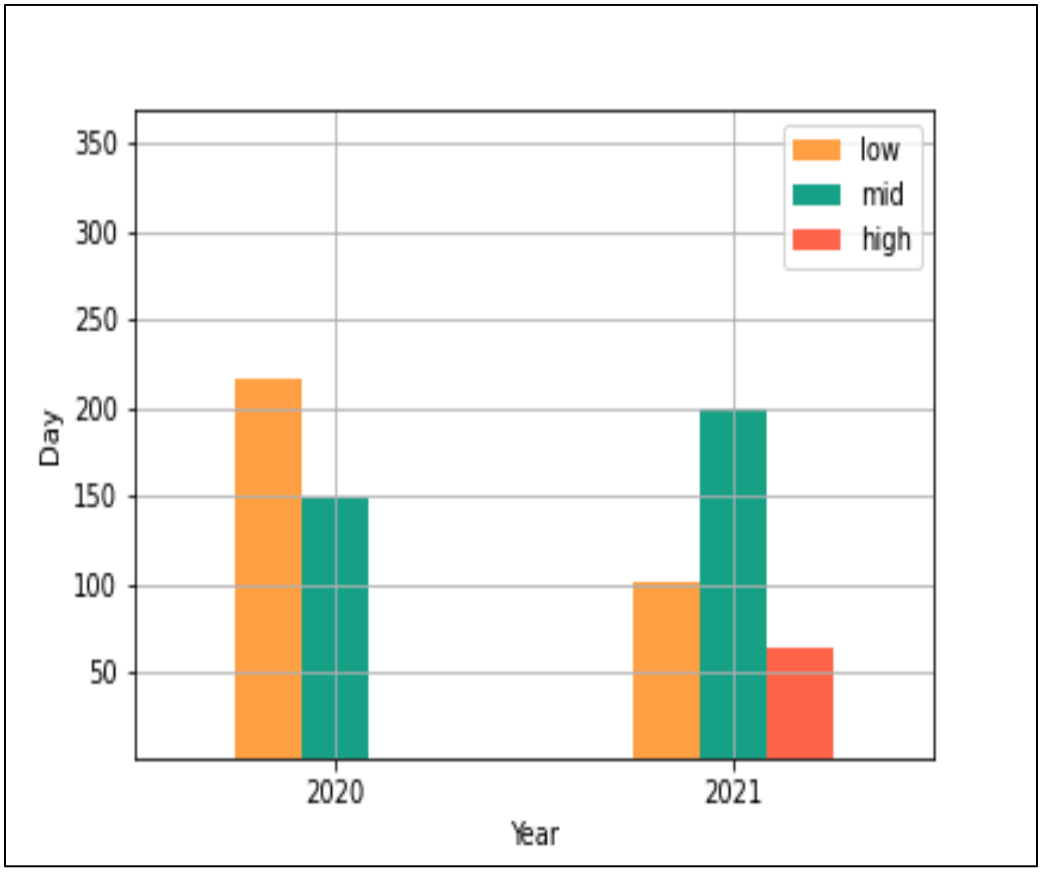



Figure 3.3: Yearly price range

Every single class, As shown in Figure 5, there is not high-class price for ginger in the year 2020, although there are numerous high-class prices in the year 2020. This indicated that the price of ginger is increasing day by day along with the increasing demand, as shown by the fact that it was demonstrated.

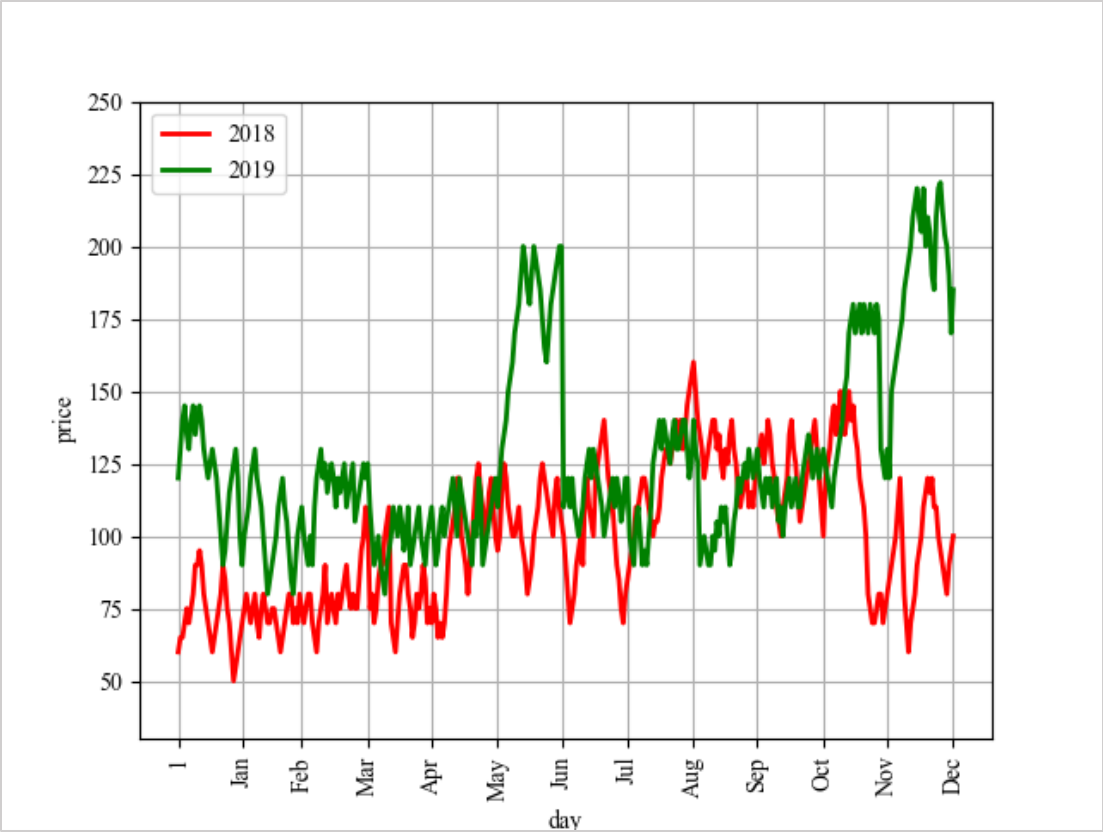



Figure 3.4: Monthly price of Ginger

The clear perspective of the variation rate over the years 2020 and 2021 is shown in Figure 3.3. The months are represented along the X-axis, while the cost of ginger is shown along the Y-axis. The year 2020 is represented by the red line, and the year 2021 is represented by the green line. We were aware that there was a significant amount of price volatility all during the year 2020, and that beginning in 2020, the cost of ginger was steadily climbing all the way up to the year 2022.

## Comparison price

The price of ginger is shown side-by-side in Figure 3.4 for the cities of Dhaka, the nation's capital, and Chattogram, its second-largest city. It displays the average price for the first six months of 2019 for each location. And we saw that the prices in the city of Dhaka were significantly more than those in Chattogram. The rate of fluctuation is extremely high in Dhaka, however, the pricing in Chattogram is relatively consistent.

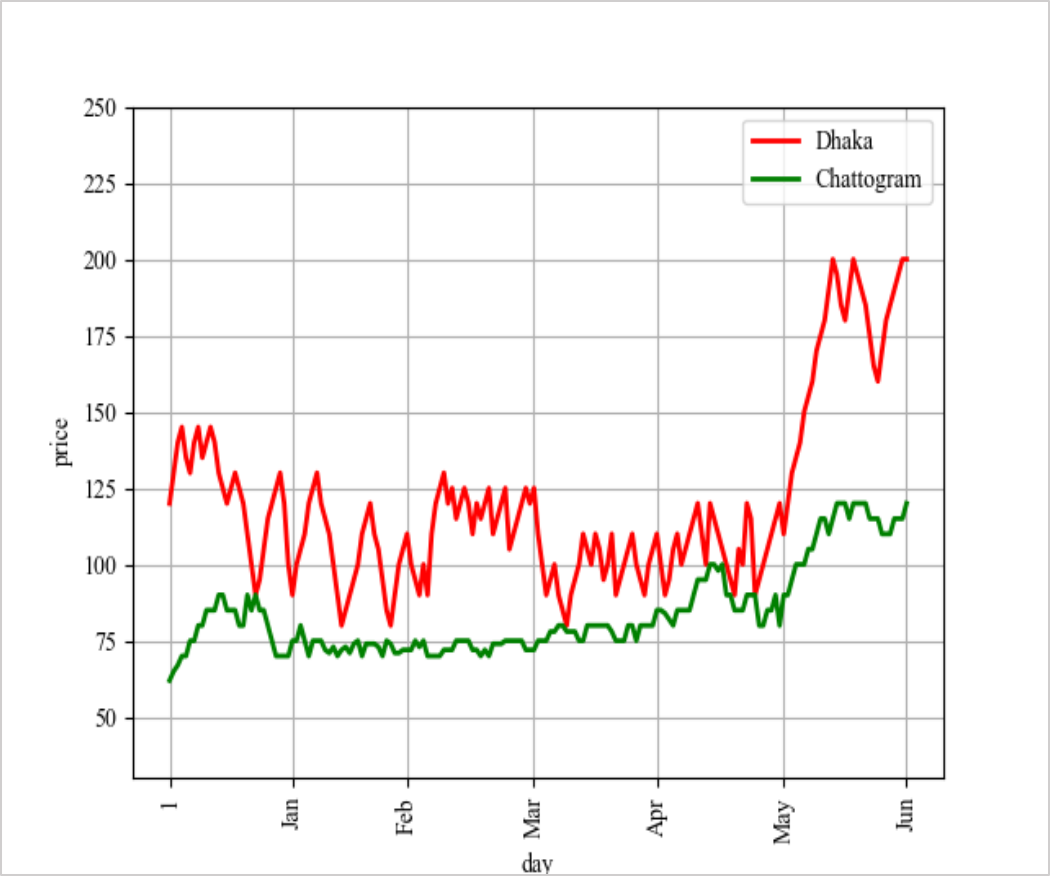



Figure 3.5: Comparison between Dhaka and Chattogram price

## 3.4 Algorithm selection

During the course of our investigation, we came to the conclusion that the classification method would best suit our needs. This was made possible by the partitioning of our complete dataset into three distinct categories: high, medium, and low. In order to obtain initial assurance, we made use of a total of five diverse and fantastic ML methods. These algorithms included KNN, Neural Network, SVM, Decision Tree, and Random Forest. We were able to determine the algorithm that works best with our model and has the maximum level of efficiency by following the aforementioned approach.

## 3.5 Algorithm Implementation

Following the application of those algorithms, we found that Random forest produced the highest level of precision, including an accuracy rate of 92.58% by utilizing 30% of the available testing data. The remaining four algorithms, on the other hand, were carried out fairly adequately. Due to the fact that random forests produced the most consistent results.

Because of the unpredictability of the market, we came to the conclusion that the best way to forecast the price of ginger was to use the algorithm that was just described.

| Algorithms | Details |
|---|---|
| KNN | K=3,p=2,random_state=0 |
| Decision Tree | random_state=0 |
| SVM | Kernel=linear |
| NN | max_iter=1000, alpha=1 |
| Random Forest | n_estimators=100 |

Table 3.2: Parameter of algorithm

Table 3.2 determines the parameters and the several things that we employed for implementing the selected algorithms.

# CHAPTER 4

# EXPERIMENTAL RESULTS AND DISCUSSION

## 4.1 Introduction

Repair orders, delivery dates, new users to add, and uninstalled apps are all examples of tasks that are now being monitored by the RA system. One type of analysis is the one that looks at the outcome after resources have been allocated. Because of the formatting requirements of the impacts section, the results are presented without evaluation or interpretation. As with the rest of the term paper, assistance is available. The findings are shared, and the analysis is demonstrated. We will discuss our findings and recommendations after doing an analysis of multiple separate algorithms. Resolution, consistency, recall, and f1 were prioritized when establishing the criteria for this data collection.

## 4.2 Experiment Result

TABLE 4.1 ACCURACY TABLE

| Data usage rate | Algorithms | | | | |
|---|---|---|---|---|---|
| | KNN | Decision Tree | SVM | Neural Network | Random Forest |
| 30% | 81.16% | 88.84% | 78.14% | 75.29% | 93.58% |
| 40% | 78.25% | 87.59% | 76.47% | 73.35% | 89.13% |
| 50% | 79.36% | 84.90% | 79.85% | 70.14% | 87.61% |
| 60% | 76.18% | 81.11% | 72.22% | 68.11% | 85.79% |
| 70% | 73.42% | 79.16% | 71.32% | 66.39% | 84.21% |

those algorithms contained inside Table III relating to the association for purposes of subsequent research. In the study that we did, we used five different kinds of training sets in order to enumerate the accuracy with a data usage rate ranging from 30–70%. In Table III, each column is denoted by a yellow mark that corresponds to the level of effectiveness at which those particular algorithms operate at their best. Following the interpretation of these algorithms, a fascinating result was found to have occurred. Random forest and Decision tree both achieved immediate performance at a high level of efficiency, but KNN, SVM, and NN only reached an average level of accuracy. According to a proclamation made by L. J. M. l. Breiman [14], a random forest is a collection of Tre forecasters that are organized in such a way that the values of a random vector are inspected independently and with the same apportionment concerning each tree in the forest. This allows each individual tree to depend on the values of the random vector. By utilizing the Random forest algorithm, our model was able to achieve the highest level of forecasting performance. It achieved an efficiency of 93.58% despite having a data utilization rate of just 30% and a training data percentage of 70%, as shown by the red rectangle box in Table I. S. B. J. A. According to I. R. Kotsiantis [15], the decision tree is a sequential model, which ineluctably combines a series of straightforward investigations. An individual test assigns a numerical characteristic that is inconsistent with a threshold value or a nominal characteristic to a group of potential outcomes. The decision tree produced the second-highest accuracy and had an efficiency of 88.84% while only using 30% of the available data. The researchers G. Guo et al. [16] made the assumption that KNN is an instance-based learning approach that stores all of the training data for later distribution. The k-Nearest-Neighbours (KNN) approach is a non-parametric classification strategy that is common but useful in a variety of contexts. As a result, KNN was able to achieve an accuracy of 81.16 percent by utilizing a data utilization rate of 30 percent. According to Y. Zhang [17], a Support vector machine (SVM) is a two-dimensional representation of the ideal surface that arose from the linearly detachable instance. This representation includes the fact that it does prefer existing approaches on a number of accomplishments. After then, the SVM achieved a precision of 79.85% while utilizing only 50% of the available data. According to M. KangaraniFarahani and colleagues [18] research, NN is a preferable method for analyzing distribution difficulties. The primary application for NN is in the dynamic implementation of complex prophesy functions, and NN also works in harmony with the process of denoising data. As a consequence of this, the Neural

Network achieved a prediction rate of 75.29% while only recognizing 30% of the test data patterns. KNN achieved an efficiency rating of 74.17% while only utilizing 30% of available data.

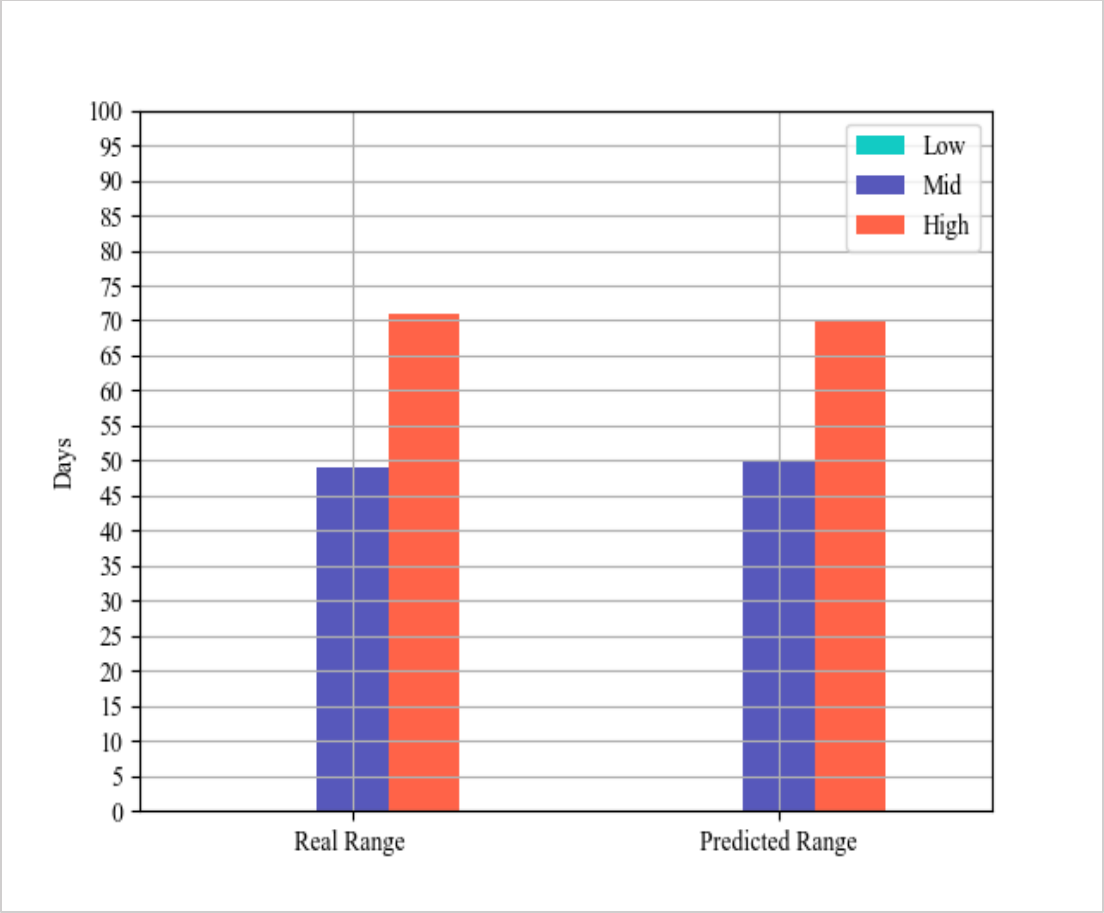



Figure 4.1: Comparison between real and predicted class

We obtained the first four months' worth of actual price data for the year 2020 in order to evaluate how well our model forecasts future prices. This analysis will take place during the year 2020. We may observe a comparison of the actual class with the projected class for the first four months of the year 2020 in the figure titled "Figure 4.1." Surprisingly, our model was only off by one prediction for the entirety of the set of projections, which indicates how exact its predictions may be. When we found out that our model had properly predicted 119 out of 120 classes, it felt like we had made a magical discovery. Our model accurately predicted that there would be no low class in the actual data, which brings us to the most notable finding, which is that there is no such thing as a low class.

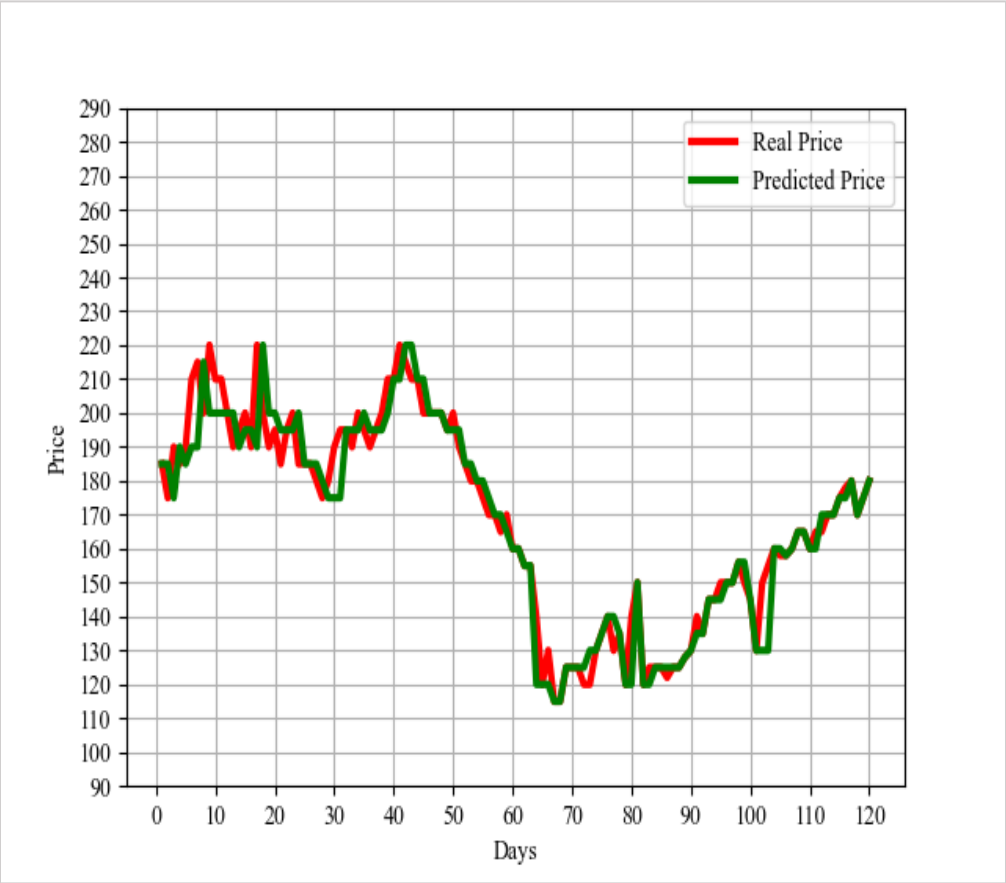



Figure 4.2: Comparison between real and forecasting prices.

Figure 4.2 illustrates both the initial price as well as the price that is forecasted to be in effect for the first four months of the year 2020. The price that was really paid is shown by the red line, while the price that our model projects will be reached is depicted by the green line. Because the green line greatly covered the red line in Figure 4.2, we can see that our model was able to produce an accurate estimate of the price. This is proven by the fact that the green line was superimposed on top of the red line. The ability of our model to reliably anticipate future prices is significantly impacted, in a way that is both positive and bad, as a result of this.

We are confident in asserting, on the basis of the research that came before this one, that the specified algorithms that were used in this investigation were executed in an equitable and agreeable manner. Because of this, the significance of the work that we do is brought into sharper focus.

## 4.3 Evaluation

After obtaining the data from a convenient source, we performed an analysis on them, and then we categorized them into a variety of categories. After that, we tested five different machine learning algorithms on our dataset to determine which algorithm is the most effective.

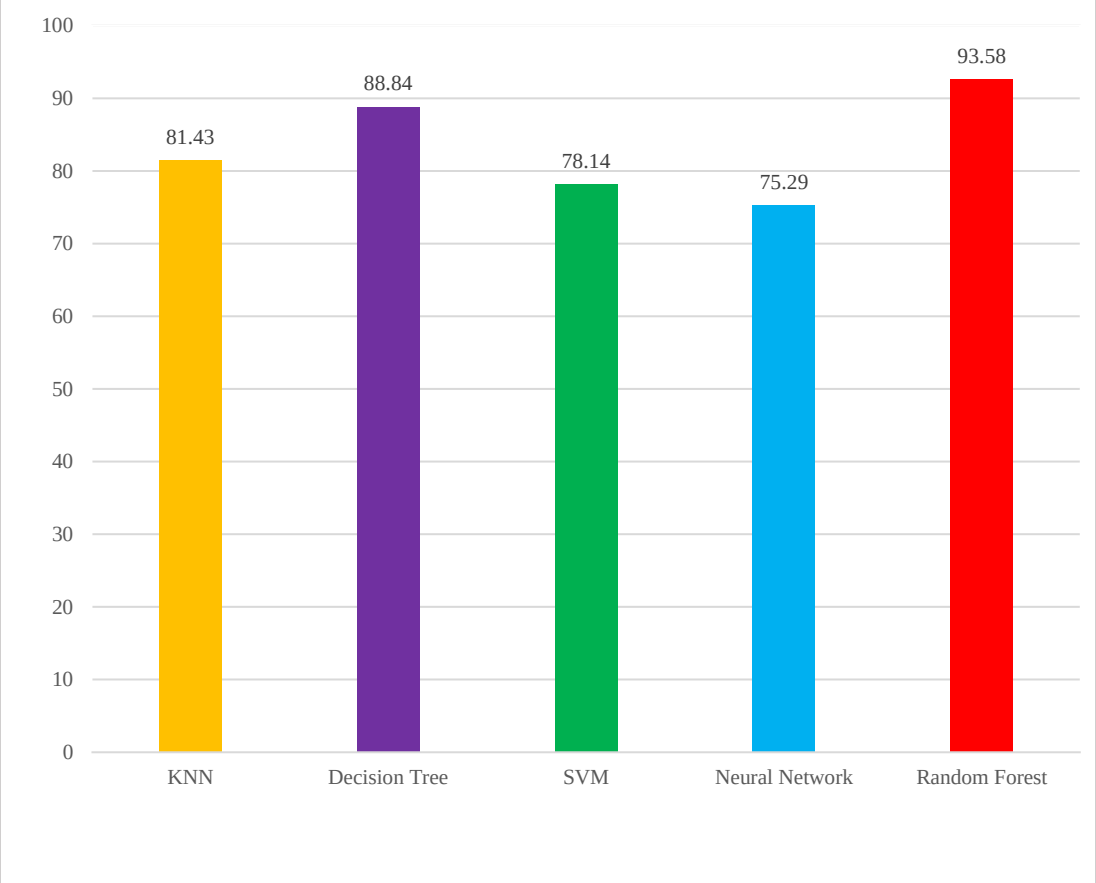



Figure 4.3: Highest accuracy of applied algorithms

Figure 4.3 illustrates the best level of success that the applied algorithm was able to achieve with a variety of data consumption rates. Since random forest performed the best, we chose to use this algorithm to predict what the price of ginger will be in order to alleviate the instability of the market. We experimented with Random forest using a variety of data utilization rates, and we found that using 30% of the available data produced a maximum efficiency of 93.58%.

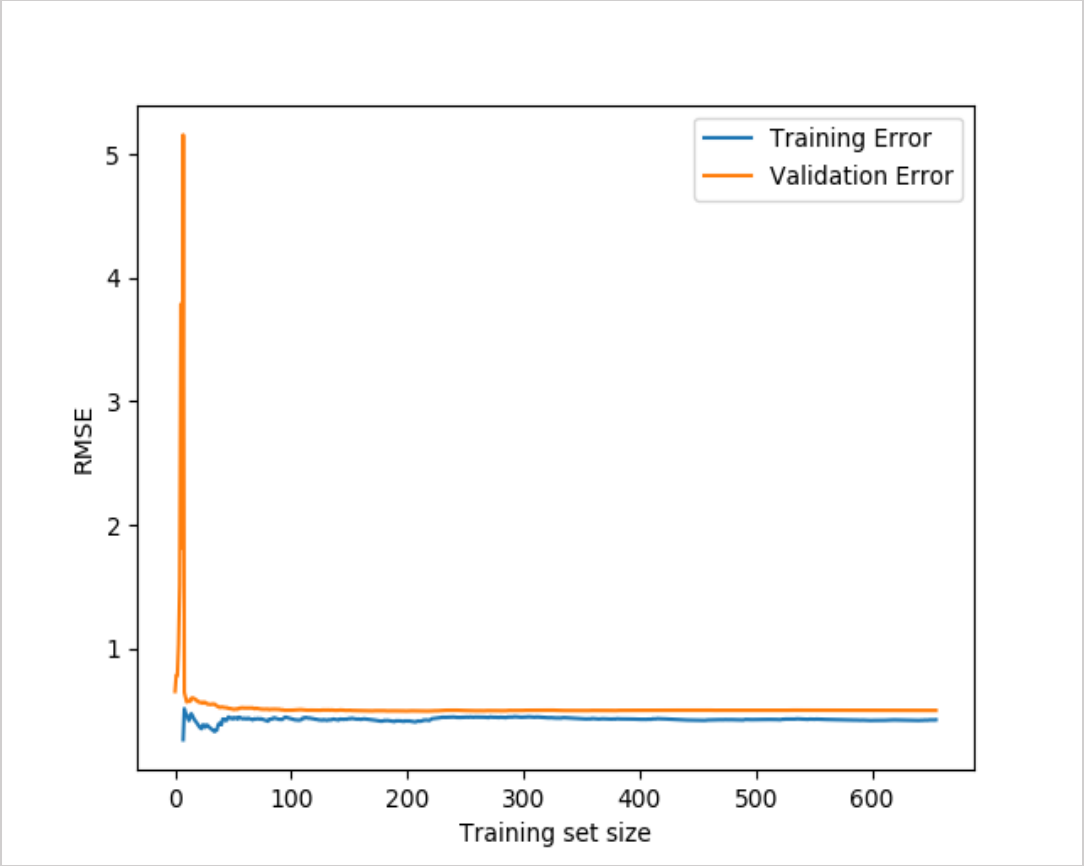



Figure 4.4: Training vs validation graph

Figure 4.4 provides a breakdown of the learning rate that the Random Forest algorithm achieves. We had the impression that the difference between the training error and the validation error was not that great. While increasing the quantity of the training data, it took into account that both mistakes were consistent. So, we can safely assume that our model has a sufficient amount of information from our dataset

# CHAPTER 5
# CONCLUSION AND FUTURE WORK

## 5.1 CONCLUSION

Forecasting the price and expected demand for a particular product will be adequate to reduce market volatility. And it additionally encourages management to convey significant determinations to stabilize the market. To accomplish the stated model, we operated five distinct established ML algorithms to acquire the best potential result. The highest level of efficiency, 93.58%, was achieved using the RF procedure. The final result is that our government and its people will understand how much ginger to produce and take the necessary steps to meet the demand and price of ginger on the market. This outcome guides such a circumstance as the following. We have strived to obtain the most desirable result, yet there have a few constraints. The main impediment to our work is the insufficient number of data. Instead of assembling entire country data, we only obtained two years of data from Dhaka and Chattogram.

## 5.2 FUTURE WORK

Moving forward to this work,

- We expect to expand our dataset by accumulating data for the whole country.
- We will further try to assemble data covering the wider time range.
- Another precedence will be to build an intelligence system to support the government and other people comprehend the future price and demand of all the groceries in the market.
- This project only can work with two of the biggest cities in Bangladesh next my task will work with other divisions of Bangladesh.
- I also proprovideweb service with importance notifications.

## 5.3 Limitation

We exerted a lot of effort in an attempt to find out how to arrive at the most positive conclusion possible, but there were still a few roadblocks on our path. One of the limits was that the data was

not easily accessible (i.e., soil data). We might have been able to get further data, but it would have been difficult to do so because of restrictions that were placed on the majority of online government sites. We might have been able to obtain additional data.

# REFERENCE


[1] M. H. Opu. (June 19th, 2019). Ginger price doubles in a month. Available:https://www.dhakatribune.com/business/commerce/2019/06/19/ginger-price-doubles-in-a-month

[2] M. K. Pasha. ( 15 June 2014). Ginger. Available: http://en.banglapedia.org/index.php?title=Ginger

[3] M. M. Hasan, M. T. Zahara, M. M. Sykot, R. Hafiz, and M. Saifuzzaman, "Solving Onion Market Instability by Forecasting Onion Price Using Machine Learning Approach," in 2020 International Conference on Computational Performance Evaluation (ComPE), 2020, pp. 777-780: IEEE.

[4] Z. H. Khan, T. S. Alin, and M. A. J. I. J. o. C. A. Hussain, "Price prediction of share market using artificial neural network (ANN)," vol. 22, no. 2, pp. 42-47, 2011

[5] D. J. N. Srinivasan, "Energy demand prediction using GMDH networks," vol. 72, no. 1-3, pp. 625-629, 2008.

[6] R. Jammazi and C. J. E. E. Aloui, "Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modeling," vol. 34, no. 3, pp. 828-841, 2012.

[7] D. Singhal, K. J. I. J. o. E. P. Swarup, and E. Systems, "Electricity price forecasting using artificial neural networks," vol. 33, no. 3, pp. 550-555, 2011.

[8] B. Park and J. K. J. E. S. w. A. Bae, "Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data," vol. 42, no. 6, pp. 2928-2934, 2015.

[9] F. A. de Oliveira, L. E. Zárate, M. de Azevedo Reis, and C. N. Nobre, "The use of artificial neural networks in the analysis and prediction of stock prices," in 2011 IEEE International Conference on Systems, Man, and Cybernetics, 2011, pp. 2151- 2155: IEEE.

[10] S. Velankar, S. Valecha, and S. Maji, "Bitcoin price prediction using machine learning," in 2018 20th International Conference on Advanced Communication Technology (ICACT), 2018, pp. 144-147: IEEE.

[11] S. J. I. J. I. C. T. Pudaruth, "Predicting the price of used cars using machine learning techniques," vol. 4, no. 7, pp. 753-764, 2014.

[12] N. Gandhi, L. J. Armstrong, O. Petkar, and A. K. Tripathy, “Rice crop yield prediction in india using support vector machines,” in 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE). IEEE, 2016, pp. 1–5.

[13] M. P. Wellman, D. M. Reeves, K. M. Lochner, and Y. J. J. o. A. I. R. Vorobeychik, "Price prediction in a trading agent competition," vol. 21, pp. 19-36, 2004.

[14] L. J. M. l. Breiman, "Random forests," vol. 45, no. 1, pp. 5-32, 2001.

[15] S. B. J. A. I. R. Kotsiantis, "Decision trees: a recent overview," vol. 39, no. 4, pp. 261-283, 2013.

[16] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN model-based approach in classification," in OTM Confederated International Conferences" On the Move to Meaningful Internet Systems", 2003, pp. 986-996: Springer.

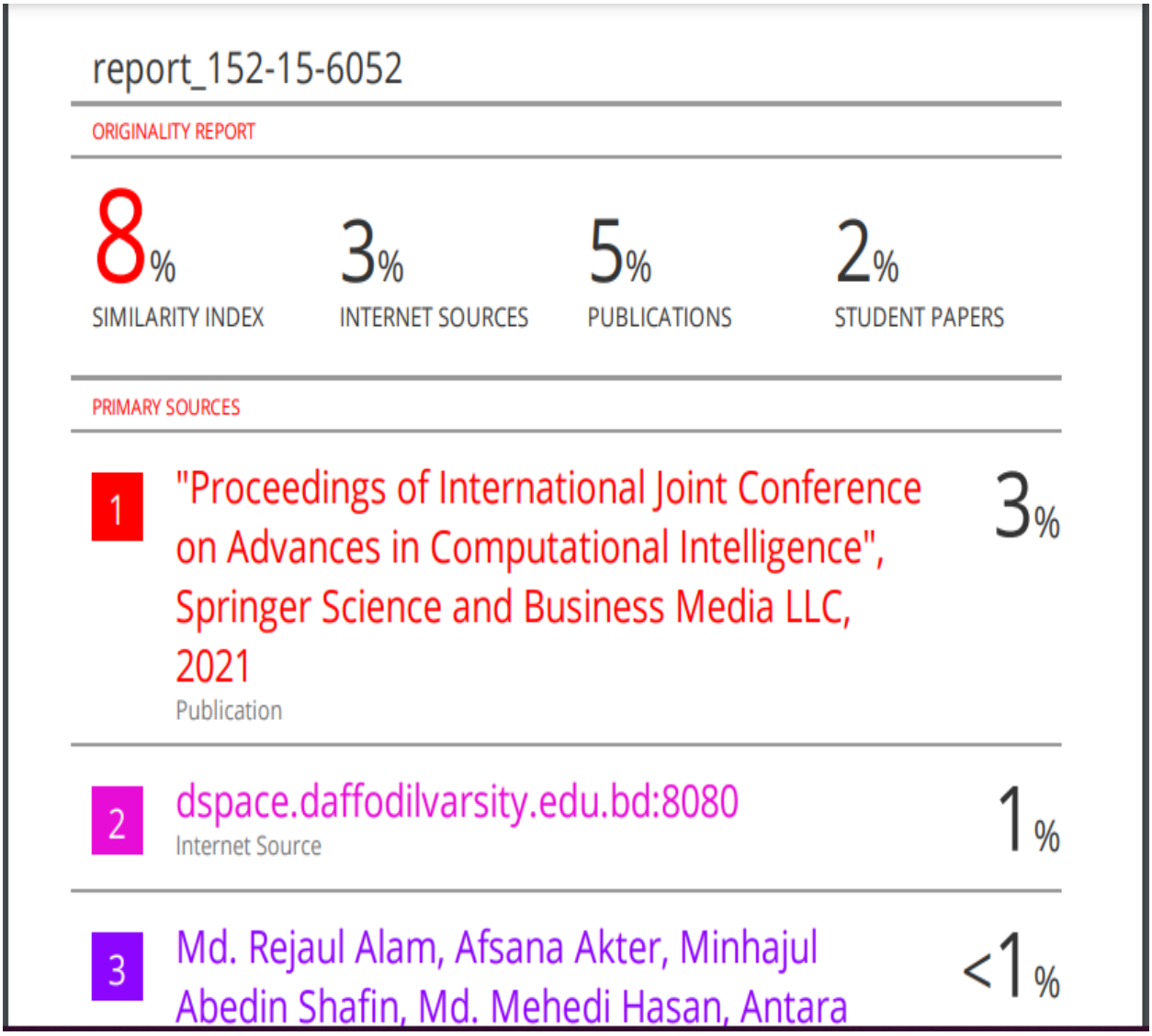
[17] Y. Zhang, "Support vector machine classification algorithm and its application," in International Conference on Information Computing and Applications, 2012, pp. 179-186: Springer.

[18] M. KangaraniFarahani and S. Mehralian, "Comparison between artificial neural network and neuro-fuzzy for gold price prediction," in 2013 13th Iranian Conference on Fuzzy Systems (IFSC), 2013, pp. 1-5: IEEE.

# APPENDIX

The first problem we had when doing the analysis was establishing the analytical technique for our investigation. It wasn't standard job, and little had been done in this subject previously. As a result, we weren't able to get much help from any source. We also started gathering data by hand. After a lengthy time of hard labor, we might be able to achieve it.

## PLAGIARISM REPORT



report_152-15-6052

ORIGINALITY REPORT

| 8% | 3% | 5% | 2% |
|---|---|---|---|
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|---|---|---|
| 1 | "Proceedings of International Joint Conference on Advances in Computational Intelligence", Springer Science and Business Media LLC, 2021<br>Publication | 3% |
| 2 | dspace.daffodilvarsity.edu.bd:8080<br>Internet Source | 1% |
| 3 | Md. Rejaul Alam, Afsana Akter, Minhajul Abedin Shafin, Md. Mehedi Hasan, Antara | <1% |