Solving Market Uncertainty By Predicting Potato Price In Bangladesh Using Regression Techniques

BY
Busra Islam
ID: 181-15-11092

Alima Islam Mim
ID: 181-15-11067

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

Supervised By
Ms. Rubaiya Hafiz

Senior Lecturer
Department of CSE
Daffodil International University

Co-Supervised By
Zerin NasrinTumpa

Lecturer
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

25 DECEMBER 2021

APPROVAL

This project titled "**Solving Market Uncertainty By Predicting Potato Price In Bangladesh Using Regression Techniques** ", submitted by **Busra Islam and Alima Islam Mim**, ID No: 181-15-11092 and 181-15-11067 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 (BSc) and approved as to its style and contents. The presentation has been held in 25 December 2021.

BOARD OF EXAMINERS

Dr. S.M Aminul Haque — **Chairman**
Associate Professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Naznin Sultana — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Raja Tariqul Hasan Tusher — **Internal Examiner**
Senior Lecturer
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Dewan Md. Farid — **External Examiner**
Professor
Department of Computer Science and Engineering
United International University

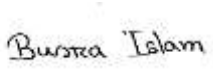
DECLARATION

We hereby declare that this thesis has been done by us under the supervision of **Ms. Rubaiya Hafiz** , **Senior Lecturer, Department of CSE**, and co-supervision of **Zerin Nasrin Tumpa, Lecturer**, and **Department of CSE** Daffodil International University. We also declare that neither this thesis nor any part of this thesis has been submitted elsewhere for the award of any degree or diploma.

Supervised by:

Ms. Rubaiya Hafiz
Senior Lecturer
Department of CSE
Daffodil International University

Submitted by:



Busra Islam
ID: 181-15-11092
Department of CSE
Daffodil International University

Alima Islam Mim
ID: 181-15-11067
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First of all, we want to render our gratitude to the Almighty Allah for the enormous blessing that makes us able to complete the final thesis successfully.

We are really grateful and express our earnest indebtedness to **Ms. Rubaiya Hafiz** Senior Lecturer, Department of CSE Daffodil International University, Dhaka, Bangladesh. Profound Knowledge & intense interest of our supervisor in the field of "Machine Learning & Deep Learning" make our way very smooth to carry out this thesis. Her remarkable patience and dedication, scholarly guidance, continual encouragement, vigorous motivation, direct and fair supervision, constructive criticism, valuable advice, great endurance during reading many inferior drafts and correcting the work to make it unique paves the way of work very smooth and ended with a great result.

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

We would like to express thankfulness to the fellow student of Daffodil International University, who took part in this discussion during the completion of this work.

We would like to express our immense thanks to the Ministry of Agriculture who provided us with the required raw data to make our work possible.

We would also like to thank the people who attend the survey done by us to collect the market real information.

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

ABSTRACT

Market uncertainty is a continuing problem in Bangladesh. As a result, the prices of our common materials fluctuate a lot. It has a significant impact on the components we use every day. In Bangladesh, potato is the third most commonly cultivated crop. In Bangladesh, it is served as the main meal. Bangladesh is a developing country. Potatoes are the third most popular vegetable in Bangladesh, after rice and wheat, with low-income individuals eating more potatoes than other vegetables. The price of potato affects whether people would eat or go hungry. In this era of artificial intelligence, we now have advance software that can extract information from data. Machine Learning is currently quite popular for predicting this sort of unpredictable fluctuation. We created our dataset using information obtained from Bangladesh's Ministry of Agriculture. We used six typical regression techniques to estimate the price of potato. We used Random Forest Regressor (RFR), Decision Tree, Gradient Boosting, Lasso regressor, Linear Regression and Neural Network Regressor models to predict the daily potato price. All of the models we've created yield results that are very satisfactory. Among all the models we created, the Random Forest Regressor (RFR) produced the best results in all stages.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLE

CHAPTER 1
INTRODUCTION

1.1 Introduction

The potato is the world's most significant nongrain food crop. Potatoes are a common item in our daily diet in Bangladesh. Bangladesh is the world's seventh-largest tuber producer. According to the Department of Agricultural Extension, it generated a record high of 1.09 crore tonnes last year. Bangladesh's yearly potato demand is estimated to be at 9 million tons. In Bangladesh, potato is a well-known and vital vegetable. It is used as the principal vegetable throughout the year. Due to a lack of preservation, most vegetables are only accessible seasonally. Low-income households eat potatoes more than other vegetables, and potato consumption is second only to rice and wheat in Bangladesh. In regular years, there is a significant price differential between harvest season and the other seasons.

Potato producers are suffering more and more every day, despite their huge yields, because they are trapped with surplus stocks and limited exports. With an annual average demand of roughly 70 lakh tonnes, the country saw a 40 lakh-tonne excess, the most of which was squandered. According to experts, not all of the surplus can be held in cold storages due to inadequate capacity. Farmers are left with little choice but to give these potatoes to cattle due to a lack of suitable uses for the surplus, according to producers.

Potato costs are low during the potato season, but they increase dramatically during other seasons. As we can see from the discussion above, the potato market is quite vulnerable. The goal of this article is to use machine learning to predict potato prices. To create an efficient dataset, a number of feature selection techniques are utilized. In this research, models such as Decision tree (DT), Random Forest Regressor (RFR), Lasso Regressor, Gradient Boosting, Neural Network Regressor and Linear Regression were used to predict future potato prices. The accuracy of each model is evaluated to discover the best model for properly forecasting the potato price.

Motivation

Potato is a common component in Bangladeshi food. In Bangladesh, we require nearly 9 million tons of potato each year. Over the potato season, prices are cheap, but they increase during the rest of the year. Potato prices will be 25 to 30 per kg at the start of the season in 2020. However, the price of potatoes towards the end of the season ranges from 40 to 50 per kg. The poor are particularly affected by rising potato prices. As a result, imports may help to lower prices, which will benefit consumers.

In this case, prediction is advantageous. Prediction may be used in the finance business to estimate future product pricing. We now have a wide range of effective Machine Learning algorithms that can manage such a large workload with easily. Using these algorithms to their full potential, we can predict future potato prices. As a consequence, the proper authorities will be able to implement a strategy for dealing with the situation if it increases.

1.2 Problem Definition

Machine learning will help in the creation of our agriculture industry. It is vital to identify the difficulties and related needs in this sector in order to give an effective solution. Understanding legislative action or regulations, as well as the software industry's needs and curriculum methodologies for using machine learning in the agriculture sector, is also required. We used data from Bangladesh's Ministry of Agriculture to discover concerns with agriculture and machine learning that are preventing market uncertainty from being reduced. If we forecast future prices, we can establish the market's demand-supply balance, and we may utilize this knowledge to maintain the market in balance.

1.4 Research Questions

The key questions that this study focuses on are given below:

- Where did we obtain the data?
- What is the most accurate approach to categorize potato prices?
- What is the people's profit from this effort?
- The difficulties that occur when attempting to predict pricing?
- Identifying recent machine learning technology for pricing prediction?

1.5 Research Methodology

In the methodology part of our study report, we collected data, examined it, categorized it, chose algorithms, implemented them, and finally assessed it. We've also talked about the algorithms' outcomes.

1.6 Research Objectives

Economic objectives:

1. Predict the price of potato in future.
2. Increase the use of machine learning to help the market.
3. Reduce the possibility of economic crisis.

Technical objectives:

1. Discover the optimal selection strategy to use in such cases.
2. Create an efficient model to remove market uncertainty.
2. Determine the most effective price prediction algorithm.
4. Encourage the government to react to changes in the market.

1.7 Research Layout

Our report's contents will have been distributed in the following categories:

Chapter 1: The introduction, motivation, problem definition, research question, research methodology, and projected outcomes will all be covered in Chapter 1.

Chapter 2: A background study and a brief review of important work in this area are included in Chapter 2. A list of noteworthy machine learning work, including prediction work, is provided below.

Chapter 3: The technique or workflow is described in detail in this chapter. In this part, you'll learn about how the research was carried out.

Chapter 4: This chapter is about analyzing the outcome. It includes the research findings as well as a graph.

Chapter 5: This chapter is part of the research's conclusion. The model's performance is presented in this section. This section also includes a comparison of accuracy. The model's web implementation and output are also included in this section. The chapter concludes with a discussion of the work's limitations. It was also encoded with the work that will be done in the future.

1.8 Expected Outcome

- We'll forecast the price of potato in the future.
- We'll make the market for potato less volatile.
- We will have presented a way to bring the market for potato a condition of balance.

CHAPTER 2

BACKGROUND STUDY

2.1 Introduction

Various machine learning approaches have been applied to make predictions. Prediction is one of the most often used applications of Machine Learning. These studies were focused on specific issues, and they utilized a range of machine learning approaches to resolve problems.

2.2 Related Works

Machine learning is widely applied for prediction and categorization in today's world. This section highlights the major exercises carried out by a few professionals in the recently referenced topic during the last many years.

The purpose of their research is to use a machine learning technique to predict rice prices. [2] The price was predicted using data from Bangladesh's Ministry of Agriculture website. Support Vector Machine (SVM), K-Nearest Neighbor (KNN), Nave Bayes, Decision Tree, and Random Forest were among the machine learning techniques used to produce this prediction. The random forest came out as the greatest performer. As a result, they applied the Random Forest to predict rice prices. They can predict the future potential price of rice using their prediction model.

Based on Virginia housing data, B. Park et al. [3] worked on a prediction of house prices. The researchers created a method for gathering datasets, cleaning datasets, and applying algorithms in their study. Data was gathered from three separate sources and carefully processed. Physical characteristics (which have sixteen attributes), public school ratting (which has three attributes), mortgage contract rate (which has eight attributes), and others (which have eight attributes) are the categories in which the data is organized. The dependent variable was Days on Market, which may be high or low. They employed the C4.5, Ripper, Nave Bayes, and AdaBoost algorithms to create models that would predict the outcome. Ripper had the best accuracy out of all of their algorithms.

They utilize regression techniques to predict the house price in this study. The main focus of this research was to compare different machine learning approaches using a regression-based technique. To determine which regression approach was the most effective, they used six different regression methods: Ridge Regression, LASSO Regression, Elastic Net Regression, Ada Boosting Regression, and Gradient Boosting Regression are all examples of multiple linear regression. MSE, RMSE, and Root Mean Square Error are used to determine the correctness of algorithms. MSE 12037006 088.27804 and RMSE 10971390390[4] have a Root Mean Square score of 0.9177022. Gradient Boosting Regression is the most accurate of these algorithms.

This study use Decision Tree and Regression approaches to forecast Crypto-Currency prices. This research first identifies the price trend on day-to-day variations in the Bitcoin price, as well as providing information on Bitcoin price patterns. The information includes open, high, low, and close price details for Bitcoin value up to the present day. For bitcoin price prediction, they used Decision Tree and Linear Regression. The greatest results come from Linear Regression, which has a 97.5 percent accuracy rate, and Decision Tree, which has a 95.8 percent accuracy rate [5].

M. M. Hasan et al. [6] shown how to manage onion market volatility by anticipating onion price. They did this by forecasting future prices using a machine learning technique. For two years, they used daily data. obtaining information The installation and assessment of Data Analysis Algorithms are the most critical four steps in completing the project. KNN, Nave Bayes, Decision Tree, SVM, and the Neural Network approach were all employed in this study. With a rate of 98.17 percent, the approach gives the highest level of accuracy.

For the prediction of daily gold prices, time-series gold price prediction models based on support vector regression and anfis models were created in this study. The gold values for training and testing were received from Australia's Perth Mint. Mean absolute error (MAE), Root Mean Square Error (RMSE), Nash–Sutcliffe model efficiency coefficient (E), and Mean absolute percentage error are the metrics used to compare the models (MAPE). The ANFIS-GP model performed marginally better than the ANFIS-SC model.[7]

They provide their methods for predicting gold prices on a daily basis in this study. They used a machine learning approach to do this. To forecast the daily gold price, they applied Support Vector Regression (SVR), Random Forest Regressor (RFR), Decision Tree, Gradient Boosting, and XGBoost models. Each model's accuracy is compared in order to find the best model for accurately predicting gold prices. All of the models they've designed have produced excellent outcomes. Out of all the algorithms, Random Forest Regression (RFR) fared the best. In any case, it has reached an accuracy of around 99 percent.[8]

In this study, they apply a machine learning technique to forecast gold prices. The study looked at monthly pricing data from January 2000 to December 2018. These data were analyzed using three machine learning algorithms: linear regression, random forest regression, and gradient boosting regression. Random forest regression has a higher prediction accuracy for the entire time, whereas gradient boosting regression has a higher prediction accuracy for the two periods individually.[9]

2.3 Bangladesh Perspective

For Bangladesh's lower-income people, the tremendous risk of decreasing market uncertainty has become a critical threat. Those who are financially disadvantaged cannot afford numerous things due to the increase in the food sector. Bangladesh's economy is now in a constant state of change. Because of the COVID-19 outbreak, the economy is slowing. Price fluctuation has made life more challenging for individuals on a tight budget. The government does not dedicate sufficient resources to identifying the reasons of market volatility. As a result, nobody knows about it. The government does not hold nearly enough seminars to educate the people about these concerns. The government is taking the required efforts to maintain control over all other issues. If the government can predict the price of potato, it will be able to take the necessary actions to keep it under control.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

The working method is divided into five stages: data collection, analysis, regression, Algorithm selection, and implementation. Figure 3.1.1 depicts the results of our research.

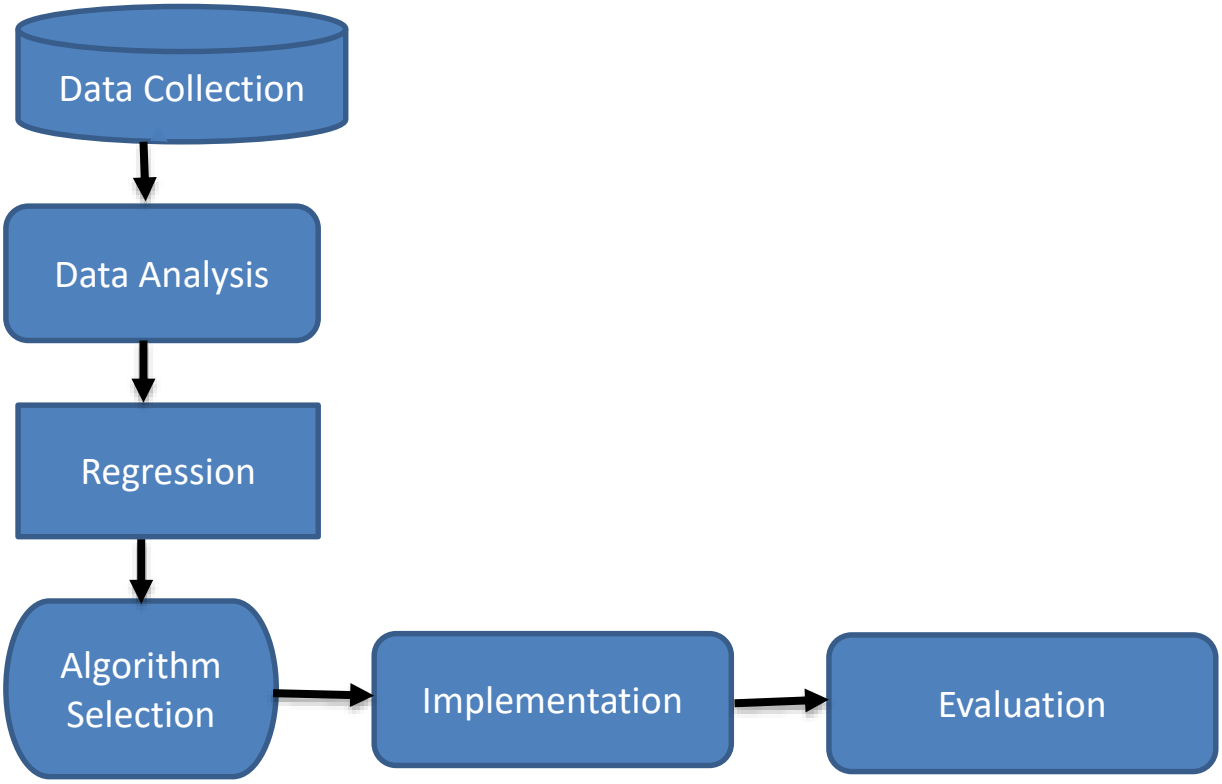


Figure 3.1: Methodology diagram

3.2 Data Collection

Data collecting is usually a challenging task in any research project. We got our potato pricing information from Bangladesh's Ministry of Agriculture's website. Our data is split into two sections. The first section was used for training and testing purposes. Another component was used to make predictions. We utilized 1064 daily prices from 2018, 2019 and 2020 for training and testing. For the projection, 30 daily prices from the month December in 2020 were used.

3.3 Data Analysis

Before we could apply our algorithms to the data we collected, we needed to examine it to understand what type of data we'd be dealing with. The price is always changing and fluctuates depending on the time of year or season. The data is disorganized, numerical, and time-series data, according to our research.

TABLE 3.1: DESCRIPTION OF ATTRIBURES

SL No.	Name of Attributes	Description
1	Month	The cost of potato varies depending on the season. As a result, one of our major qualities is months.
2	Year	We can see that the price of potato is increasing sometimes and sometimes decreasing from 2018 to 2020 in every month in this dataset. As a result, we chose year as our characteristic.
3	Day	We chose the day as our first characteristic because we were working with time-series data in our dataset.
4	Season	There is a strong link between the season and pricing. When the season is start, the price of potato is low, and when the season is off, the price of potato is high.
5	Location	Garlic prices vary depending on where you live. The cost of living in the city is constantly high. However, in the village, the price is cheap.
6	Price	The pricing is our most important feature. We wanted to be able to predict the price of potato.
7	Category	Category implies classification based on the price scale. We used the Category attribute to classify the dataset.

3.4 Regression

We can see Fig. 3.2.1 for a better understanding. The X-axis in this graph shows the month, while the Y-axis represents the price.

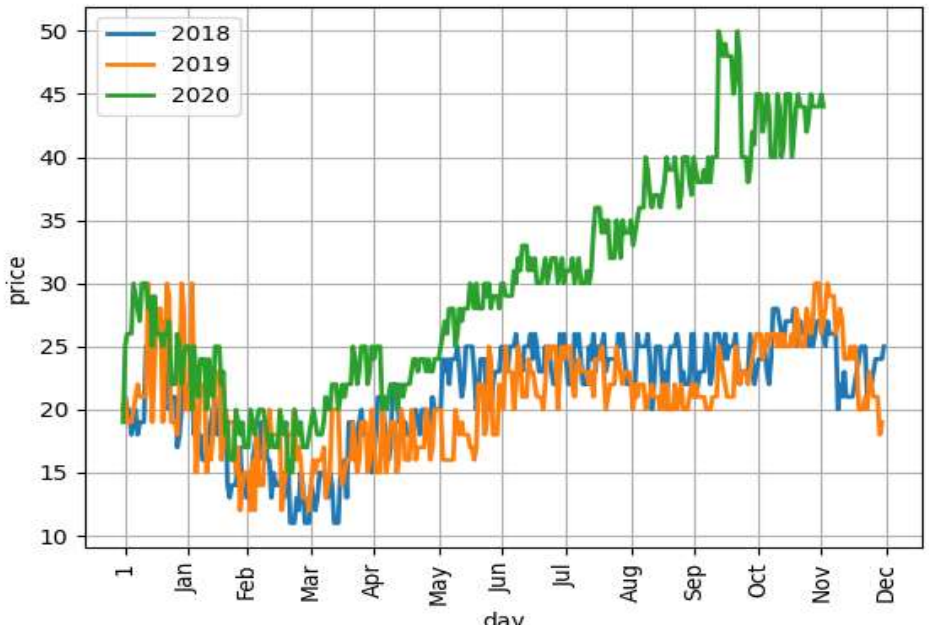


Figure 3.2: Monthly price of Potato

For the years 2018, 2019 and 2020 represents the correlation graph. Throughout the months of June to November in the year 2020, the variation rate was quite uncommon, as shown in Fig. 3.2.1. Potato price has risen suddenly since June 2020 and will continue to do so until November 2020.

3.5 Algorithm Implementation

We discussed the algorithm implementation procedure in this part. To finish this procedure, we must first complete the preceding one in order to create the requisite dataset. Because our job is in the regression form, we have five distinct regression methods. For regression,

we employ six algorithms: Decision tree, Gradient Boosting, Lasso regressor, Linear regressor, Neural Network regressor and Random Forest. The best suited parameter to achieve highest accuracy for various methods is shown in Table 3.2.

Table 3.2: Parameter usages

Algorithms	**Details**
Decision tree regressor	random_state=1
Gradient Boosting	n_estimators int, default=100
Lasso regressor	bool, default=True
Random Forest regressor	n_estimators=100
Linear regressor	fit_intercept=true
Neural Network regressor	hidden_layer_sizes=100

Table 3.2 displays the parameters and other items we used to implement the algorithms we picked.

3.6 Evaluation

Using real-time data estimates and an uncertainty matrix, we assessed our chosen RF approach. We started with 30 real data points, but our model was unable to learn from them.

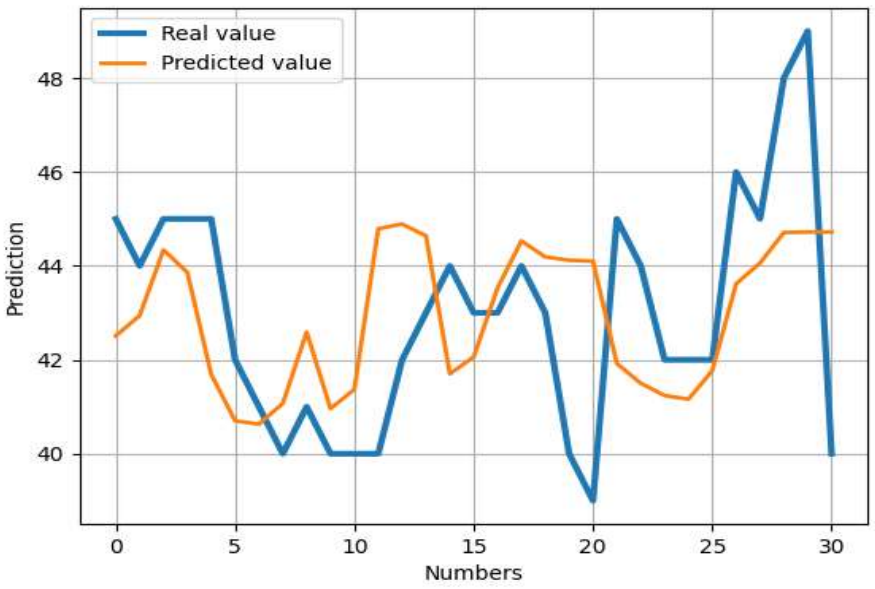


Figure 3.3: Comparison between Real and Predicted

The real and predicted prices for the months of December in the year 2020 were shown in Fig.3.3. The blue line shows the actual price, while the orange line shows the expected price, and we can see how near the orange line is to the blue line. It shows that our model is quite accurate at predicting prices. Only day 19 and 20 predict the opposite values. We may conclude from the preceding discussion that the algorithms utilized in this study worked pretty well.

CHAPTER 4

RESULT ANALYSIS

4.1 Introduction

This section of research analysis is highly dependent on empirical evidence and test results. What is the first result analysis when we investigate a subject? The outcomes section should be constructed in such a way that the outcomes are given without any interpretation or evaluation. The advice is also available in the section for academic papers. The test is shown and the results are announced. In a series of six algorithms, we looked at a variety of algorithms and will evaluate which ones are the best.

4.2 Experimental Result

Six models were utilized in our research to determine the execution of predicted potato actions: DT, Gradient boosting, Lasso, Linear, Neural Network and RF. The outcome is shown in Table 4.1.

Table 4.1: Performance comparison of different models

Models	MAE	MSE	RMSE
DT	1.6375	5.675	2.3822
Gradient boosting	1.5479	3.6896	1.9208
Lasso	4.0497	25.4007	5.0399
Linear	4.1329	24.9634	4.9963
Neural Network	3.71339	23.3604	4.8332
RF	**1.4207**	**3.5564**	**1.8858**

According to Table 4.1, the Random Forest Regressor appears to have the lowest mean square error of all the models on the cumulative dataset. In comparison to previous models, this model displays good approximation capabilities.

Table 4.2: Accuracy table

Test data usage rate		**30%**	**40%**	**50%**	**60%**	**70%**
Algorithms **Accuracy**	**DT**	**88.36**	86.18	84.14	84.59	85.35
	Gradient Boosting	**91.51**	91.01	90.55	89.55	89.32
	Lasso	45.01	46.31	46.85	**47.13**	46.39
	Linear	47.18	47.56	48.04	47.97	**48.46**
	Neural Network	51.91	**51.96**	50.30	41.92	41.26
	RF	**92.70**	92.50	92.21	90.95	90.92

The accuracy table is shown in Table 4.2. To assess which item performs best, we used 30 to 70% of the test data. red boxes reflect the test percentage for each algorithm that achieves the best accuracy. Except for the Lasso and Linear method, which performs best below 40% of the test results, as seen in this table. Linear offered 48.46 ppm precision at 70 ppm, whereas the RF delivered 92.70 ppm precision at 30 ppm. With yellow boxes on the table, RF still has the highest accuracy of all the algorithms.

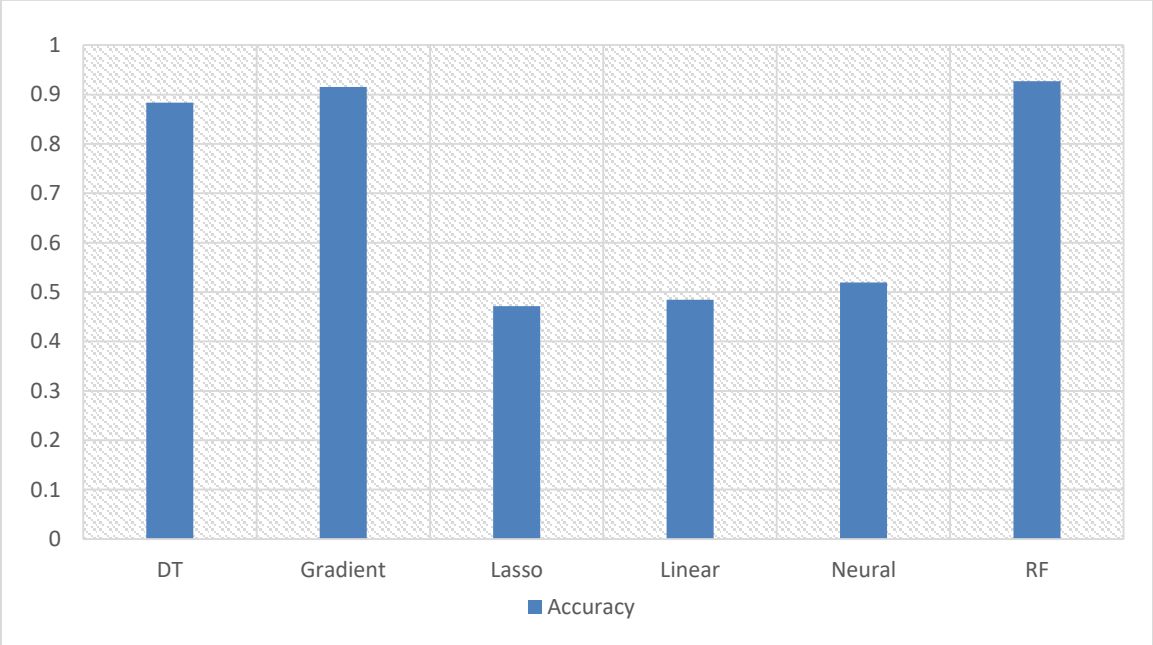


Figure 4.1: Comparison graph of Accuracy

4.2.1 Decision Tree

In the shape of a tree structure, a decision tree constructs regression or classification models. It gradually cuts down a dataset into smaller and smaller sections while also developing an associated decision tree. A tree containing decision nodes and leaf nodes is the end result. A decision node can have two or more branches each of which represents a value for the attribute being checked. A choice on the numerical aim is represented as a leaf node. The root node is the topmost decision node in a tree that corresponds to the best predictor. Both category and numerical data may be handled by decision trees. The best accuracy of the Decision tree method is 90.06 percent. Figure 4.2 shows that the Decision tree algorithm's best MAE, MSE, RMSE.

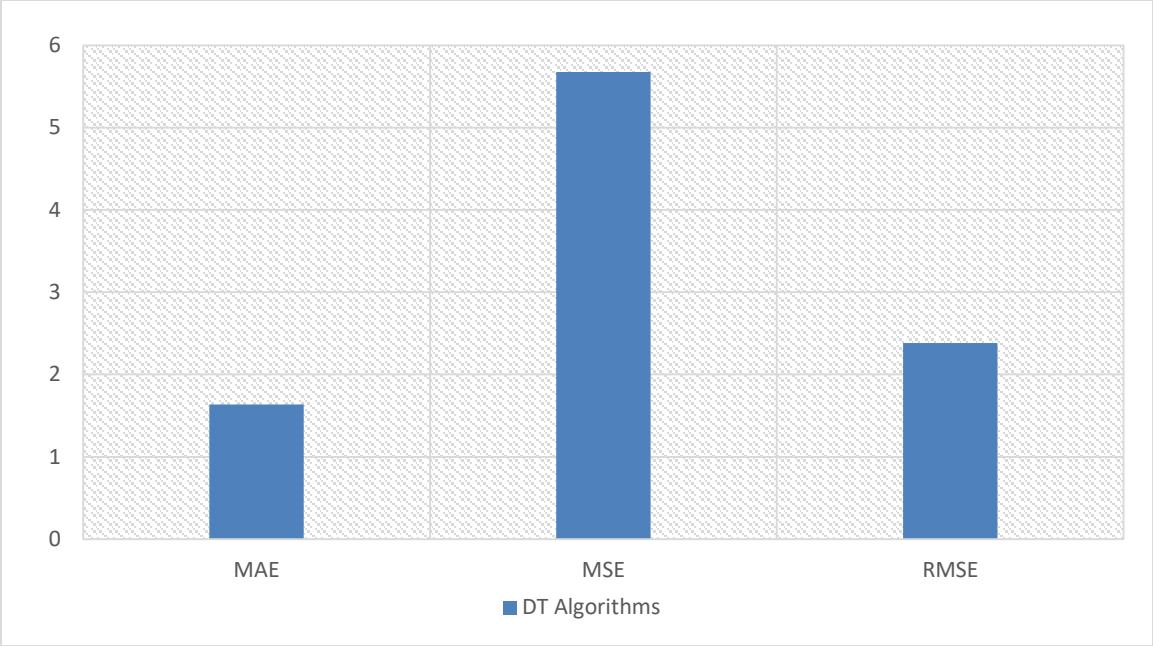


Figure 4.2: Different Score comparison graph of Decision Tree

4.2.2 Gradient Boosting

Gradient boosting is a machine learning approach that may be used for a variety of applications, including regression and classification. It returns a prediction model in the form of an ensemble of weak prediction models, most often decision trees. [1][2] The resultant approach is called gradient-boosted trees when a decision tree is the weak learner; it generally beats random forest. A gradient-boosted trees model is constructed in the same stage-wise manner as other boosting approaches, but it differs in that it allows optimization of any differentiable loss function. The best accuracy of the Gradient boosting method is 91.51 percent. Figure 4.3 shows that the Gradient boosting algorithm's best MAE, MSE, RMSE.

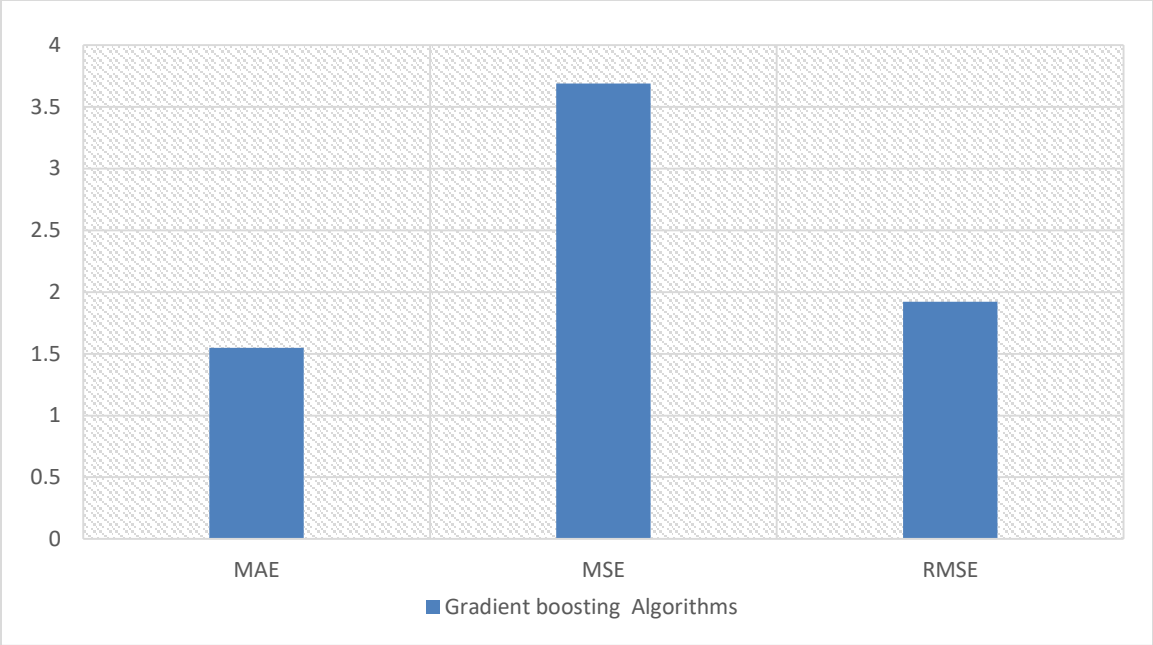


Figure 4.3: Different Score comparison graph of Gradient boosting.

4.2.3 Lasso Regressor

Lasso is a regression analysis approach used in statistics and machine learning that does both variable selection and regularization to improve the predictability and interpretability of the produced statistical model. Lasso was created with linear regression models in mind. This straightforward example exposes a great deal about the estimator. These include the linkages between lasso coefficient estimates and so-called soft thresholding, as well as its relevance to ridge regression and optimal subset selection. It also indicates that if variables are collinear, the coefficient estimates do not need to be unique (as in traditional linear regression). The best accuracy of the Lasso regression method is 47.13 percent. Figure 4.4 shows that the Lasso regressor algorithm's best MAE, MSE, RMSE.

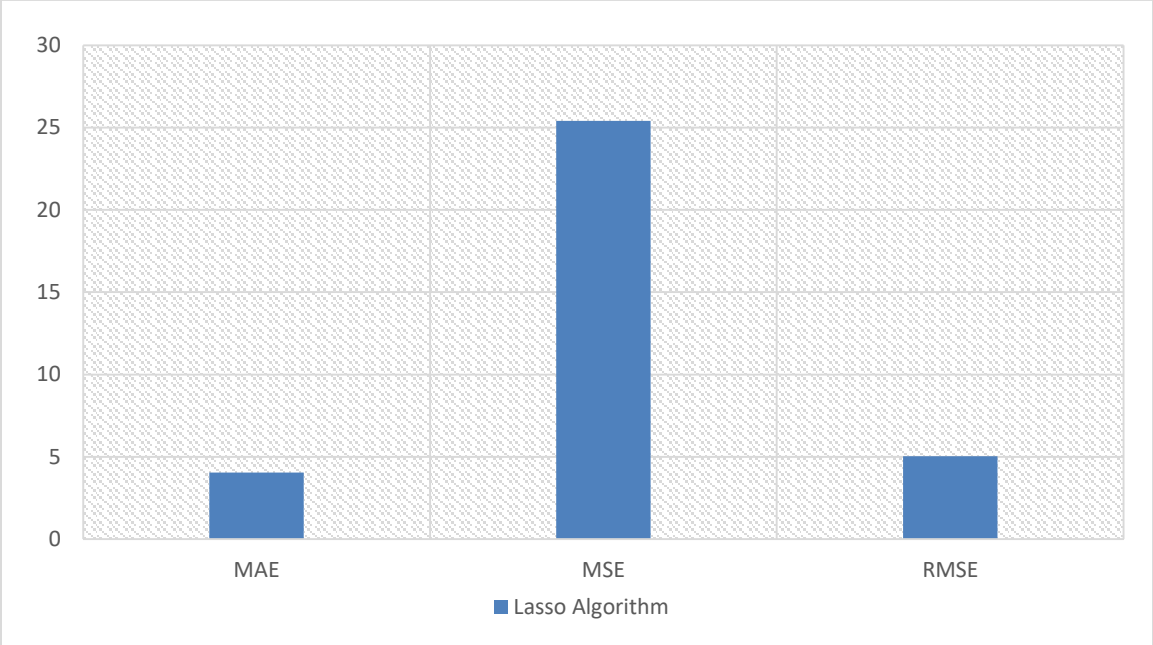


Figure 4.4: Different Score comparison graph of Lasso Regressor

4.2.4 Linear Regression

Linear Regression is a supervised Machine Learning model that determines the best fit linear line between the independent and dependent variables, or the linear connection between the two variables. The basic goal of a Linear Regression model is to determine the best-fit linear line and the appropriate intercept and coefficient values such that the error is minimized. The best accuracy of the Linear Regression method is 48.46 percent. Figure 4.5 shows that the Linear Regression algorithm's best MAE, MSE, RMSE.

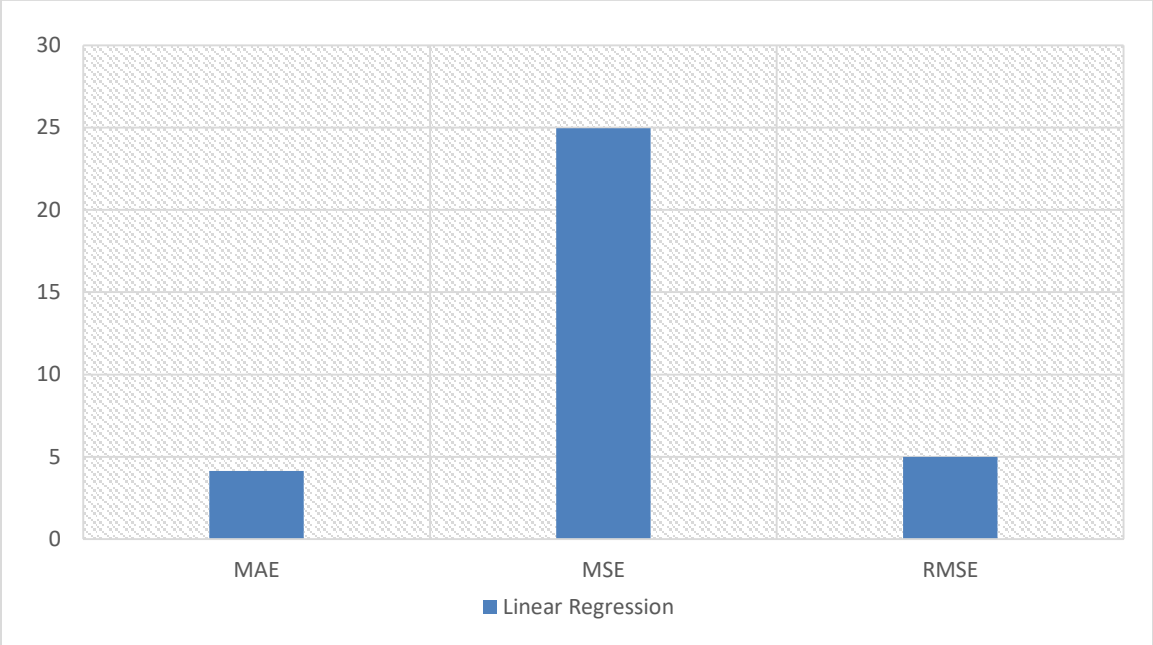


Figure 4.5: Linear Regression Score Comparison

4.2.5 Neural Network

In the financial sector, neural networks help with time-series forecasting, algorithmic trading, securities categorization, credit risk modeling, and the creation of proprietary indicators and price derivatives. Layers of linked nodes make up a neural network. Each node is a perceptron, which works in the same way as a multiple linear regression. The perceptron converts the signal from a multiple linear regression into a nonlinear activation function. The best accuracy of neural network method is 51.96 percent. Figure 4.6 shows that the neural networks algorithm's best MAE, MSE, RMSE.

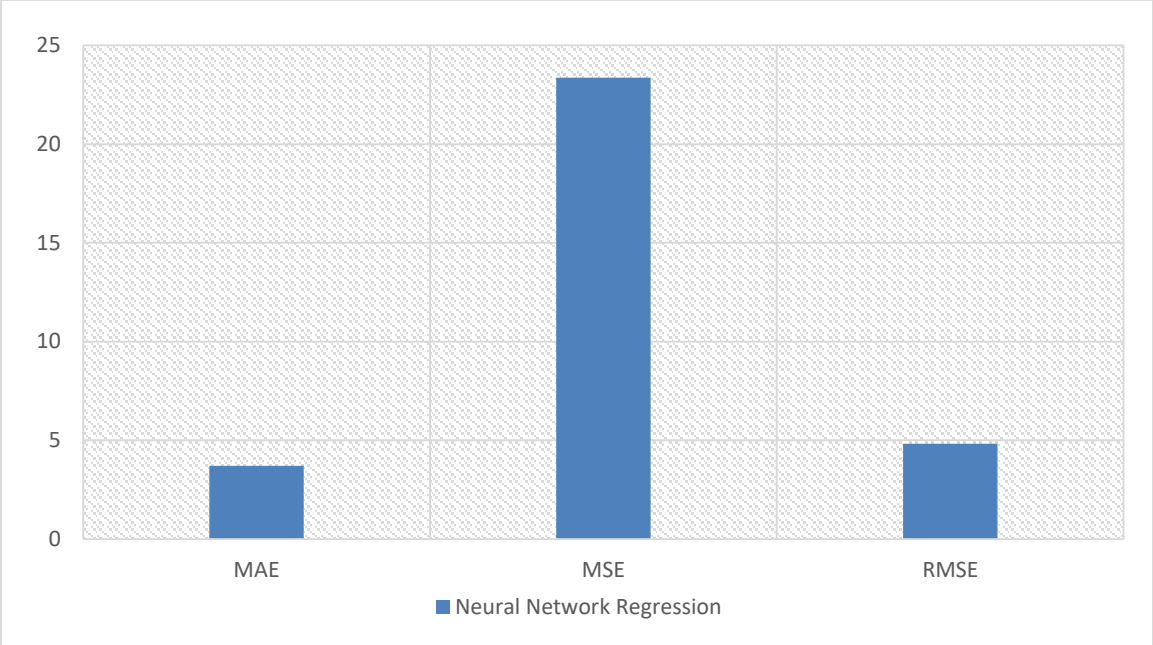


Figure 4.6: Different Score comparison graph of Neural Network Algorithm

4.2.6 Random Forest

Random forests are an ensemble learning approach for classification, regression, and other problems that works by training a large number of decision trees. For classification tasks, the random forest's output is the class chosen by the majority of trees. The mean or average forecast of the individual trees is returned for regression tasks. Random decision forests adjust for decision trees' characteristic of over fitting to their training set. Random decision forests outperform decision trees in general. We'll explore at how the RF Algorithm works, how it varies from other algorithms, and how it's used in this paper. The best accuracy of Random forest method is 92.70 percent. Figure 4.7 shows that the Random forests algorithm's best MAE, MSE, and RMSE.

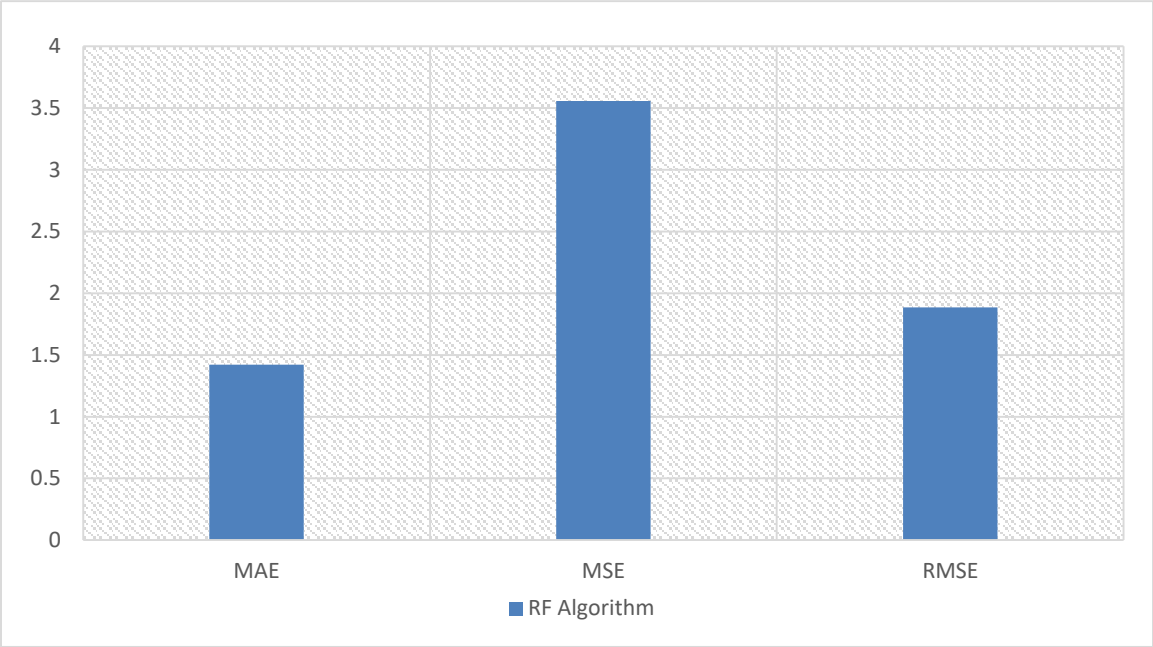


Figure 4.7: Different Score comparison graph of Random forest Algorithm

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Conclusion

To get the best result, we used six regression Machine Learning algorithms in our suggested model. With a little trial and error, we looked at six of the most popular machine learning methods. Despite the fact that DT, RF and GB algorithms performed similarly. The random forest came out as the greatest performer. We used the Random Forest to predict future price of potato because it had the greatest accuracy. It is more practical to decrease market imbalance by predicting future prices. This effort will enable the government to take actions to control market stability.

5.2 Recommendations

There are a few excellent ideas for this:

- This paper contains a very little quantity of data.
- To improve the study's outcomes by increasing the accuracy of data collecting.
- The use of Deep Learning would be preferable.

5.3 Future Work

The future guidance on the development of this work is given bellow:

- In the future, we will work with a huge dataset.
- We want to create an intelligence system that can predict potato prices in the future.
- In the future, we will create an intelligence system based on deep learning techniques.

REFERENCE

[1] https://www.thedailystar.net/backpage/potato-production-in-bangladesh-glut-badly-hurts-growers-1748572?fbclid=IwAR0RYoD_ch4GL2nrGfOfySTUfBObQnhI_wQ0CsFY6fi006TnFKno0n9Ahxs.

[2] M. M. Hasan, M. T. Zahara, M. M. Sykot, A. U. Nur, M. Saifuzzaman and R. Hafiz, "Ascertaining the Fluctuation of Rice Price in Bangladesh Using Machine Learning Approach," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020, pp. 1-5, doi: 10.1109/ICCCNT49239.2020.9225468.

[3] 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.

[4] Madhuri, C. R., Anuradha, G., & Pujitha, M. V. (2019, March). House price prediction using regression techniques: A comparative study. In 2019 International Conference on Smart Structures and Systems (ICSSS) (pp. 1-5). IEEE..

[5] K. Rathan, S. V. Sai and T. S. Manikanta, "Crypto-Currency price prediction using Decision Tree and Regression techniques," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019, pp. 190-194, doi: 10.1109/ICOEI.2019.8862585.

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

[7] Dubey, A. D. (2016, January). Gold price prediction using support vector regression and ANFIS models. In 2016 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-6). IEEE.

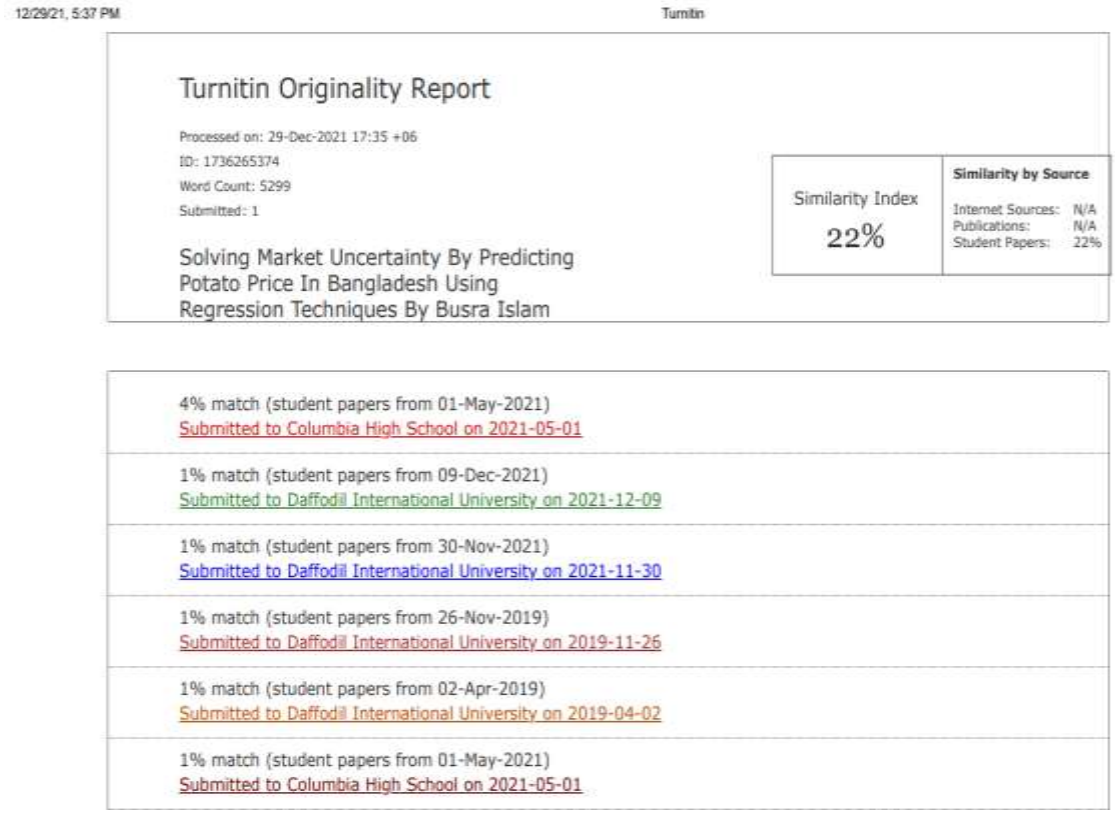
[8] M. A. Mithu, K. M. Rahman, R. A. Razu, M. Riajuliislam, S. I. Momo and A. Sattar, "Gold Price Forecasting using Regression Techniques for Settling Economic and Stock Market Inconsistency," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021, pp. 1-4, doi: 10.1109/ICCCNT51525.2021.9579755.

[9] K. A. Manjula and P. Karthikeyan, "Gold Price Prediction using Ensemble based Machine Learning Techniques," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), 2019, pp. 1360-1364, doi: 10.1109/ICOEI.2019.8862557.

APPENDIX

We had several challenges while doing the research, the first of which was determining the methodological strategy for our study. It wasn't typical work, and there hadn't been much done in this area before. As a result, we were unable to obtain much assistance from any source. We also began manually collecting data. We were able to achieve it after a long period of hard work.

PLAGIARISM REPORT



12/29/21, 5:37 PM Turnitin

Turnitin Originality Report

Processed on: 29-Dec-2021 17:35 +06
ID: 1736265374
Word Count: 5299
Submitted: 1

Solving Market Uncertainty By Predicting Potato Price In Bangladesh Using Regression Techniques By Busra Islam

Similarity Index	Similarity by Source
22%	Internet Sources: N/A Publications: N/A Student Papers: 22%

4% match (student papers from 01-May-2021)
Submitted to Columbia High School on 2021-05-01

1% match (student papers from 09-Dec-2021)
Submitted to Daffodil International University on 2021-12-09

1% match (student papers from 30-Nov-2021)
Submitted to Daffodil International University on 2021-11-30

1% match (student papers from 26-Nov-2019)
Submitted to Daffodil International University on 2019-11-26

1% match (student papers from 02-Apr-2019)
Submitted to Daffodil International University on 2019-04-02

1% match (student papers from 01-May-2021)
Submitted to Columbia High School on 2021-05-01