

Gold Price Forecasting Using Regression Techniques for Settling Economic and Stock Market Inconsistency

BY

Mosaddek Ali Mithu

ID: 171-15-8669

Kazi Motiour Rahman

ID: 171-15-8649

MD. Ruhul Amin Razu

ID: 171-15-8956

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

Supervised By

Abdus Sattar

Assistant Professor

Department of CSE

Daffodil International University



DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

MAY 2021

APPROVAL

This Project titled “**Gold Price Forecasting Using Regression Techniques For Settling Economic and Stock Market Inconsistency**”, submitted by **Mosaddek Ali Mithu, Kazi Motiour Rahman** and **Md Ruhul Amin Razu** 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 on May 2021.

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

Subhenur Latif
Assistant Professor

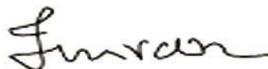
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



Internal Examiner

Md. Abbas Ali Khan
Senior Lecturer

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University



External Examiner

Shah Md. Imran
Industry Promotion Expert
LICT Project, ICT Division, Bangladesh

DECLARATION

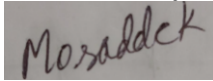
We hereby declare that, this thesis has been done by us under the supervision of **Abdus Sattar, Assistant Professor, Department of CSE** Daffodil International University. We also declare that neither this thesis nor any part of this thesis has been submitted elsewhere for award of any degree or diploma.

Supervised by:

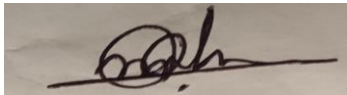


Abdus Sattar
Assistant Professor
Department of CSE
Daffodil International University

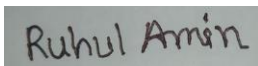
Submitted by:



Mosaddek Ali Mithu
ID: 171-15-8669
Department of CSE
Daffodil International University



Kazi Motiour Rahman
ID: 171-15-8649
Department of CSE
Daffodil International University



Md. Ruhul Amin Razu
ID: 171-15-8956
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final thesis successfully.

We really grateful and wish our profound our indebtedness to **Abdus Sattar**, Assistant Professor, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “Data mining” to carry out this thesis. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this thesis.

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 thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

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

ABSTRACT

Gold is one of the most important metals in a country's economy. For financial banks and stock exchanges, it is a regulatory factor. Gold has the potential to have a huge impact on the economy. Gold price volatility is a common occurrence in almost all countries. Our own country, Bangladesh, is no different. Because it is held as a reserve by the central bank, changes in its price can cause problems in the country's economy. In this paper, we present our models for predicting gold prices on a daily basis. We have used machine learning approach for achieving this goal. To forecast the daily gold price, we used Support Vector Regression (SVR), Random Forest Regressor (RFR), Decision Tree, Gradient Boosting, and XGBoost models. All the models that we have generated produce outcomes that are much acceptable. Amidst all the models we devised Random Forest Regressor (RFR) has generated the best outcome in all phases. The accuracy attained by the RFR algorithm is around 99% in all cases.

TABLE OF CONTENTS

CONTENTS	PAGE
Board of examiners	i
Declaration	ii
Acknowledgements	iii
Abstract	iv
CHAPTER	
CHAPTER 1: INTRODUCTION	1-4
1.1 Introduction	01
1.2 Motivation	02
1.3 Problem Definition	02
1.4 Research Question	02
1.5 Research Methodology	03
1.6 Research Objective	03
1.7 Report Layout	03
CHAPTER 2: BACKGROUND	5-7
2.1 Introduction	05
2.2 Related work	05
2.3 Bangladesh Perspective	07
CHAPTER 3: RESEARCH METHODOLOGY	8-14
3.1 Introduction	08
3.2 Experiment Data Set	08
3.3 Data Pre-Processing	10
3.4 Regression Techniques	11

CHAPTER 4: PERFORMANCE OF THE PROPOSED MODEL	15-25
4.1 Training, Testing and the Validation of the model	15
4.2 Model efficiency	15
CHAPTER 5: RESULT COMPARISON AND DISCUSSION	26-29
CHAPTER 6: CONCLUSION AND FUTURE WORK	30
REFERENCES	31

LIST OF FIGURES

FIGURES	PAGE NO.
Figure 3.1.1 Methodology flow chart	08
Figure 3.2.1 Price comparison of US dollar with BDT	08
Figure 3.2.2 Price comparison of Euro with BDT	09
Figure 3.2.3 Price comparison of GBP with BDT	10
Figure 3.4.1 Support Vector Regression	11
Figure 3.4.2 Random Forest Regression	12
Figure 3.4.3 Decision tree	13
Figure 3.4.4 Gradient boost	14
Figure 4.2.1 Accuracy plot (EUR)	15
Figure 4.2.2 Accuracy plot (GBP)	16
Figure 4.2.3 Accuracy plot (USD , EUR , GBP)	16
Figure 4.2.4 Accuracy plot (USD)	17
Figure 4.2.5 Accuracy plot (EUR)	18
Figure 4.2.6 Accuracy plot (GBP)	18
Figure 4.2.7 Accuracy plot (USD, EUR, GBP)	19
Figure 4.2.8 Accuracy plot (USD)	20
Figure 4.2.9 Accuracy plot (EUR)	20
Figure 4.2.10 Accuracy plot (GBP)	21
Figure 4.2.11 Accuracy plot (USD, EUR, GBP)	21
Figure 4.2.12 Accuracy plot (USD)	22
Figure 4.2.13 Accuracy plot (EUR)	23
Figure 4.2.14 Accuracy plot (GBP)	23
Figure 4.2.15 Accuracy plot (USD, EUR, GBP)	24
Figure 4.2.16 Accuracy plot (USD)	25
Figure 4.2.17 Accuracy plot (EUR)	25

Figure 4.2.18 Accuracy plot (GBP)	26
Figure 4.2.19 Accuracy plot (USD, EUR, GBP)	26
Figure 5.1 RFR prediction accuracy	28
Figure 5.2 Accuracy chart (USD)	29
Figure 5.3 Accuracy chart (EUR)	30
Figure 5.4 Accuracy chart (GBP)	30
Figure 5.5 Accuracy chart (USD, EUR, GBP)	31

LIST OF TABLES

TABLE NO.	PAGE NO.
Table 3.2.1 Daily price of gold	09
Table 3.3.1 Description of dataset	10
Table 5.1 Performance comparison of different models	27
Table 5.2 Performance comparison of different models with different currency	28

CHAPTER 1

INTRODUCTION

1.1 Introduction

In financial institutions and the stock exchange, gold prices have always been a significant factor. Gold serves as a reserve in every area, despite the fact that it generates valuable resources. Any country's central bank keeps gold as a reserve to ensure that it can be used to finance or trade on the global market, thus improving the country's economic situation. Gold has long been a popular investment option for people all over the world [1].

Since gold's price fluctuates so much, there is always a risk when investing in it. Throughout any financial crisis, the price of gold has remained high. A small change in the gold rate may result in a significant loss or benefit. In addition, the price of gold has a major impact on gold mining firms. Price fluctuations on a regular basis could spell disaster for these businesses. The central bank and investors can use gold production forecasting to make sound investments and mitigate potential risks.

Gold price fluctuations have a range of reasons, and the course of the gold price is dictated by a number of factors. The US dollar index (USDIX), the price of crude oil, the US Consumer Price Index (CPI), and the prices of US ten-year bond futures are all closely related to gold price adjustments. [2nd]. In terms of Bangladesh currency, we chose the US dollar, Euro, and pound during our study. If the value of the US dollar falls in comparison to other currencies, there is a chance that demand for imports will rise. In this context, as the value of the US dollar declines, gold can become more valuable. As a result, the price of gold tends to move in the opposite direction of the value of the US dollar. Nonetheless, fluctuations in the US dollar do not always translate into lower gold rates. Other currencies, such as the Euro, are also used to market the US dollar because it is consistently priced by its competitors in foreign exchange. As a result, the Euro falls when the US dollar increases. This suggests a similar relationship between the Euro and gold.

Support Vector Regression (SVR), Random Forest Regressor (RFR), Decision Tree, Gradient Boosting, and XGBoost models were used to forecast future gold prices in this study. To find the best model for accurately predicting the gold price, the accuracy of each model is compared. The thesis is divided into parts, the first of which includes a detailed description of the article's

introduction. The following is how the rest of the paper is organized: Section 2 discusses and explains previous related works, Section 3 discusses methodology and data, Section 4 discusses and explains the conclusions, and Section 5 concludes and makes recommendations

1.2 Motivation

We are currently facing dreadful circumstances. Due to the pandemic of COVID-19 economic imbalance has become a predicament for almost all countries. We must try our best to cope with this situation. We must make maximum effort to deal with this social and economical disaster. The price of gold can play an enormous role in the balancing of a countries economy. If the price gets out of hand in such a situation a country will have to face much misfortune. So if it is possible to learn of the price beforehand managing the economic balance will become more convenient. Doing such a thing manually is almost impossible. But that's where technology comes in. We have many potent Machine Learning algorithms now that can even do such a massive task easily. By making the best use of these algorithms we can smoothly detect the future price of gold. As a result, the proper authorities will be able to think of the steps to handle the situation if it were to get out of hand.

1.3 Problem Definition

Machine Learning has made our day to day challenges much more manageable. However, in order to use it properly, we must first understand where and how we can put it to use. If we are looking to attain a suitable solution, proper identification of the problem is a must. We need to understand everything we can about the problem if we want our approach to be as successful as possible. We must also properly point out the prerequisite of the analysis to bring out the most satisfactory conclusion.

1.4 Research Questions

We need to point out right question to find the right answer. The main questions that we have tried to find an answer to are given below:

1. Identifying current usage of Machine Learning in price prediction
2. The limitations that occur while attempting to predict price
3. How solution for such predicaments can be formed

1.5 Research Methodology

our workflow and data refinement is described in this segment. We have described data cleaning, attribute selection. Training models and employing the models are explained as well in this section. We have also talked about the outcome of the algorithms.

1.6 Research Objectives

We have much to gain from using machine learning in forecasting the price of gold. The objectives of our work are given below-

Economic objectives:

1. Predict the price of gold beforehand
2. Increase the usage of Machine Learning in stabilizing economic state
3. Reduce the risk of stock market instability

Technical objectives:

1. Find the best selection technique for further use in such cases
2. Identify the Best algorithm for price prediction

1.7 Research Layout

Chapter 1: will present the introduction, motivation, Problem Definition, Research Question, Research Methodology and the expected outcome of our project.

Chapter 2: will clarify the context of this research and related work, as well as the current status of the project from the perspective of Bangladesh and also about government priorities and regulations.

Chapter 3: will describe the process of implementing the proposed models to predict the price of gold in Bangladesh

Chapter 4: will discuss the resolution of the models in price prediction.

Chapter 5: focuses on comparing the results and selection of the best performing algorithm.

Chapter 6: Contains conclusion for this research.

Chapter 7: exhibits all the references studied for this research.

CHAPTER 2

BACKGROUND

2.1 Introduction

Gold price is very impactful when it comes to economic state and stock exchange for a country. But no similar work has been conducted from Bangladesh perspective. So the background of the project is predicting the future price of gold.

2.2 Related Works

Machine learning is widely used for prediction and classification. There has also been some research into predicting the price of various objects. Below are several examples of jobs that are similar-

In this paper, predict the house price using Regression Techniques .This paper mainly focused on the comparison between different machines learning Regression based algorithm. They applied six different types regression algorithms are Multiple Linear Regression, Ridge Regression, LASSO Regression, Elastic Net Regression, Ada Boosting Regression, Gradient Boosting Regression and find which algorithm is best. MSE, RMSE and Root Mean Square Error used for calculated the accuracy of algorithms. Gradient Boosting Regression gives the best accuracy among these algorithms with Root Mean Square score 0.9177022, MSE 12037006 088.27804 and RMSE 10971390390[3].

In this research, they predict the price of used car using Regression models. They applied three difference regression algorithms Multiple Linear Regression, Random Forest Regression, Gradient Boosting Regression and compare their performance by the mean absolute error MSE value. Gradient Boosting Regression provided the best result with MSE value 0.28[4].

This research predict the Crypto-Currency price using Decision tree and Regression techniques. They applied Decision Tree and Linear Regression for bitcoin price prediction. Linear Regression gives the best result with the accuracy of 97.5% and Decision Tree 95.8% accuracy [5].

In this system they applied both Classification and Regression techniques for predict the stock price. They compared the performance of Classification and Regression algorithms with their accuracy. They applied different types of Classification and Regression algorithms and find out the accuracy. Random Forest Regression delivered the best result with the accuracy of 99.57% [6].

This paper they predict gold price using machine learning techniques. They used some indicator are stock market, crude oil price, rupee dollar exchange rate, inflation and interest rate for predict gold price. They applied three different regression based algorithms are Linear Regression ,Random Forest Regression and Gradient Boosting Regression. They collected monthly gold price data and time period is 2000-2018. Then divided into three period 2000-2018,2000-2011 and 2011-2018. Random forest regression gives better accuracy for 2000-2018 period and Gradient boosting regression provides for 2000-2011 and 2011-2018 periods. They measured algorithms performance by MAE, MSE, RMSE and Root Mean Square Error value [7].

In this paper, their aims at forecasting the future price of gold and diamonds using ensemble techniques and find out most accurate result. They applied regression models are linear , Random forest and later used three feature selection techniques are Chi-square Test ,PCA and RFE .Then compare the model with mean accuracy ,best accuracy and worst accuracy. And also compared with state of the art approach with gradient boosting, ada boosting and bagging [8].

In their work, they predict time-series gold price using Support Vector Regression (SVM) and Adaptive Neural Fuzzy Inference System (ANFIS) models. They compare their performance with their Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Nash–Sutcliffe model efficiency coefficient (E) and Mean absolute percentage error (MAPE) score . In ANFIS model ANFIS-GP provides best performance [9].

In their work, they proposed an Artificial Neural Network (ANN) model for forecasting gold price. Their aim to build a model which predict the gold price with maximum precision. They build a Long Term Short Memory(LSTM) model with the different layers , number of neurons, input shape and activation function. They used Root mean square error (RMSE) and Mean absolute error (MAE) for measuring the performance of their model [10].

M. M. Hasan et. al. [11] describe the fluctuation of rice price in Bangladesh. To minimize the fluctuation rate they tried to predict the future price of rice. For this purpose, they use traditional classification machine learning algorithms like KNN, Naïve Bayes, Decision Tree, SVM, and Random forest. Their highest accuracy rate was 98.17% that is achieved by Decision tree algorithm. In their work they classify rice price into three categories high mid and low range. And by Decision tree algorithm they tried to predict future price range of rice.

M. M. Hasan et. al. [12] showed that onion market instability solution by predicting onion price. For this purpose they used Machine learning algorithm to predict the future price. They used daily data of 2 years. Data collection Data Analysis Algorithm implementation and evaluation is their main 4 stapes to complete whole work. They used KNN, Naïve Bayes, Decision tree, SVM and Neural Network algorithm. The highest accuracy is gain by Neural network algorithm and accuracy rate is 98.17%.

Some of the works mentioned above are strikingly close to our own. These works have also forecasted the cost of various products. None of them, however, are focused on Bangladesh and gold price forecasting.

2.3 Bangladesh Perspective

The economic state of Bangladesh is not very stable at this time. Due to COVID-19, the economy has suffered much. if any other factors start acting to make it more unpredictable, it could lead to economic disaster. So we must try our best to keep other variables as stable as possible that affect the economy. Gold price is deeply associated with the economy. The government is taking necessary steps to keep all the other factors in control. If the price of gold can be predicted, then the government will be able to take the necessary steps required to keep it under control. It can lessen the economic struggle to some extent. Stock market equilibrium is greatly dependent on the gold price as well. So keeping the gold price stable will also have an immense effect on the stock market equilibrium.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

Using Convolutional Neural Network with 6 different classes of images this work achieved 93.46% accuracy. For our analysis, we followed the work flow that has been given below-

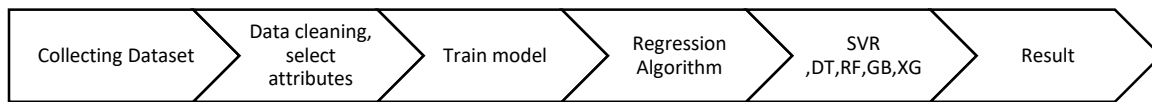


Figure 3.1.1: Methodology flow chart

3.2 Experiment Data Set

The aim of this research is to forecast the cost of gold using machine learning algorithms. The data used in this analysis are obtained from the website. The dataset contains data from January 2010 to October 2020 total of 3450 data.



Figure 3.2.1: Price comparison of US dollar with BDT

Figures 3.2.1, 3.2.2, 3.2.3 show that the price of gold in US dollars, euros, and pounds in comparison to the Bangladeshi taka is very high.

Day	Month	Year	USD	EUR	GBP	BDT_Price
4	1	2010	68.8750	153.385	110.859	31720
5	1	2010	68.8750	152.816	110.531	31720
6	1	2010	69.1050	152.299	109.924	31720
.....
30	10	2020	84.7836	160.500	115.000	65503
31	10	2020	84.8219	160.500	115.000	65503

Table 3.2.1: Daily price of gold

These data describe the gold price in various currencies on the same day. Our indicators for gold price are the US dollar, the Euro and the GBP with Bangladesh currency. Table 3.2.1 shows the dataset of daily price of gold in different currencies.

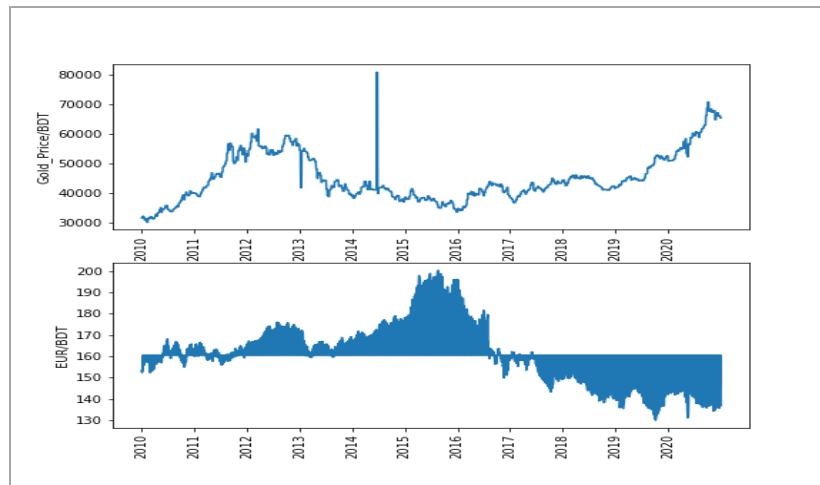


Figure 3.2.2: Price comparison of Euro with BDT

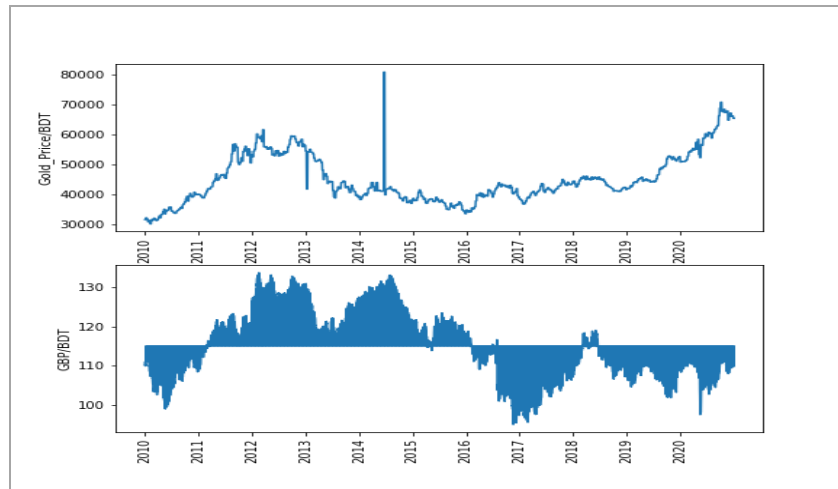


Figure 3.2.3: Price comparison of GBP with BDT

All the mentioned currencies have much deep relation to the price fluctuation of gold. They can be used to determine the future price of the gold.

3.3 Data Pre-Processing

Data processing has been carried out to ensure that data is well systematized. The issue of missing values was dealt with satisfactorily in order to complete the dataset. Table 3.3.1 shows the count, mean, std, minimum and maximum, used of data (25%,50%,75%). After the dataset is partitioned into training and test dataset where 30% data is utilized for test and 70% data is practiced for training the models

	Day	Month	Year	USD	EUR	GBP	BDT_Price
count	3952.0	3952.0	3952.0	3952.0	3952.0	3952.0	3952.0
mean	15.733806	6.452176	2014.926872	79.175841	160.833910	114.9	45121.7
std	8.793793	3.416093	3.122808	4.511768	13.202198	7.8	8157.7
min	1.000000	1.000000	2010.000000	68.825000	129.885000	94.9	30283.0
25%	8.000000	3.750000	2012.000000	77.600000	156.292000	109.9	39576.0
50%	16.000000	6.000000	2015.000000	78.619225	160.500000	115.000000	42792.0
75%	23.000000	9.000000	2018.000000	83.300000	165.914750	118.916000	51208.000000
max	31.000000	12.000000	2020.000000	85.360650	200.436000	133.773000	80985.000000

Table 3.3.1. Description of dataset

3.4 Regression Techniques

Support Vector Regression:

SVR allows one to restrict how much error is permissible in our model and will find a line that fits the results. The aim of SVR is to minimize the coefficients, specifically the l2-norm of the coefficient vector, rather than the squared error. Instead, we treat the error term in the constraints, where we set the absolute error to be less than or equal to a given margin, called the maximum error (epsilon). We may modify epsilon to achieve the model's desired efficiency.

Objective function and constraints are as follows:

Minimize:

$$\text{MIN } \frac{1}{2} \|\mathbf{w}\|^2$$

Constrain:

$$|y_i - w_i x_i| \leq \epsilon$$

Illustration:

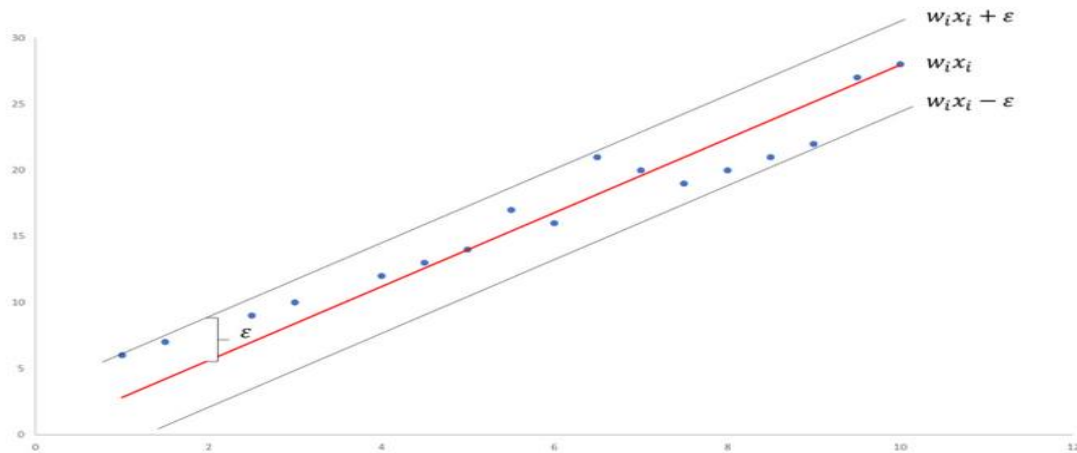


Figure 3.4.1: Support Vector Regression

Random Forest Regression:

Random Forest Regression is a supervised learning algorithm for regression that employs the ensemble learning process. The ensemble learning approach incorporates predictions from many machine learning models to provide a more reliable prediction than a single model.

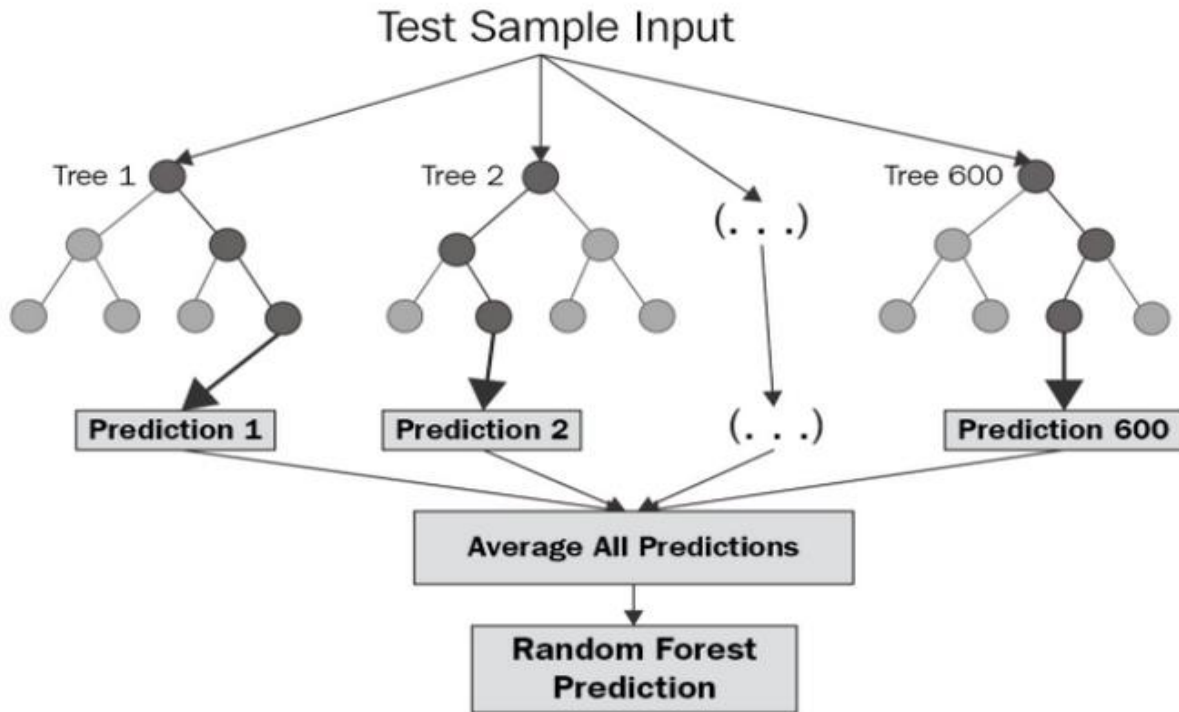


Figure 3.4.2: Random Forest Regression

The arrangement of a Random Forest is depicted in the diagram above. We can see that the trees are running in a straight line with no interaction. During preparation, a Random Forest constructs multiple decision trees and outputs the mean of the groups as the prediction of all the trees.

Decision tree:

A decision tree is a flowchart-like structure in which each internal node represents an attribute "test," each branch represents the test's result, and each leaf node represents the class name. The classification principles are described by the pathways from root to leaf. A decision tree and its closely related effect

diagram are used as a visual and empirical decision support method in decision analysis to quantify the predicted values of competing alternatives.

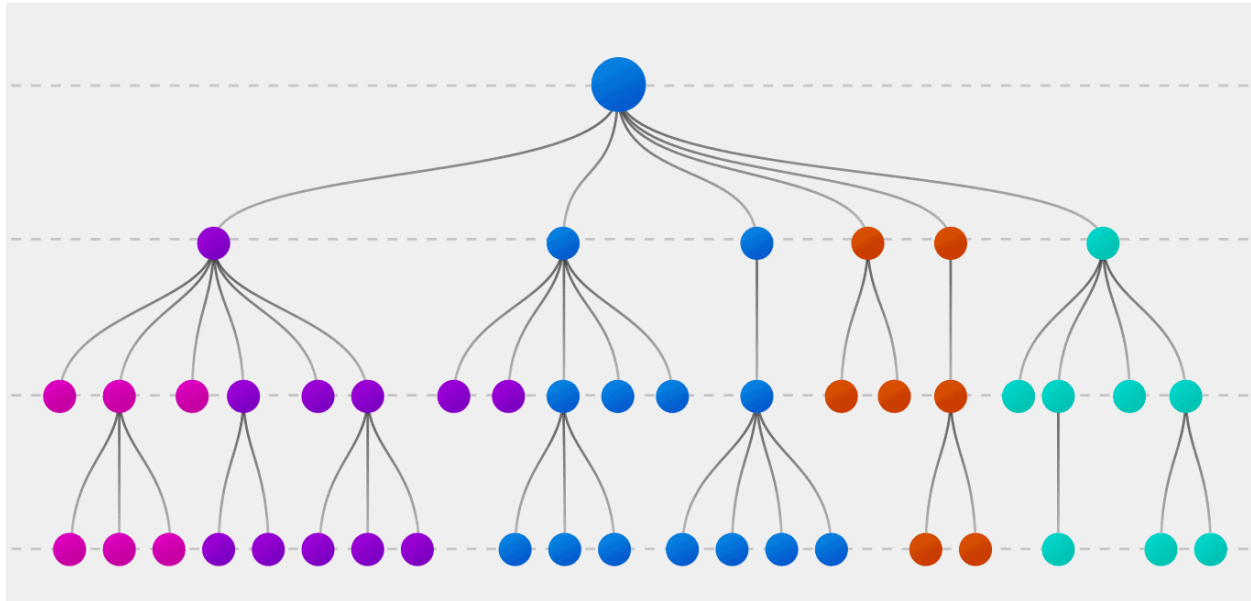


Figure 3.4.3: Decision tree

XGBoost:

XGBoost is a distributed gradient boosting library that has been developed for performance, flexibility, and portability. It uses the Gradient Boosting paradigm to apply machine learning algorithms. XGBoost is a parallel tree boosting algorithm that solves a wide range of data science problems quickly and accurately. It has the ability to solve challenges that go beyond billions of instances.

Gradient boost:

Boosting is a technique for turning vulnerable students into good ones. Each new tree in boosting is based on an updated version of the original data set. Gradient Boosting is a technique for gradually, additive, and sequentially training a large number of models.



Figure 3.4.4: Gradient boost

CHAPTER 4

PERFORMANCE OF THE PROPOSED MODEL

4.1 Training, Testing and the Validation of the model

SVR, RFR, Decision Tree, Gradienboosting, and Xgboosting are five models used in our study to examine the execution of forecasting gold price alteration. From our collection of 3450 data, we divided our data 70% for training and 30% for testing. So, 2415 number of data were used for training and 1035 were used for testing all the models.

4.2 Model efficiency

We have used SVR, RFR, Decision Tree, Gradienboosting, and Xgboosting. Their performance is given below-

Support Vector Regression:

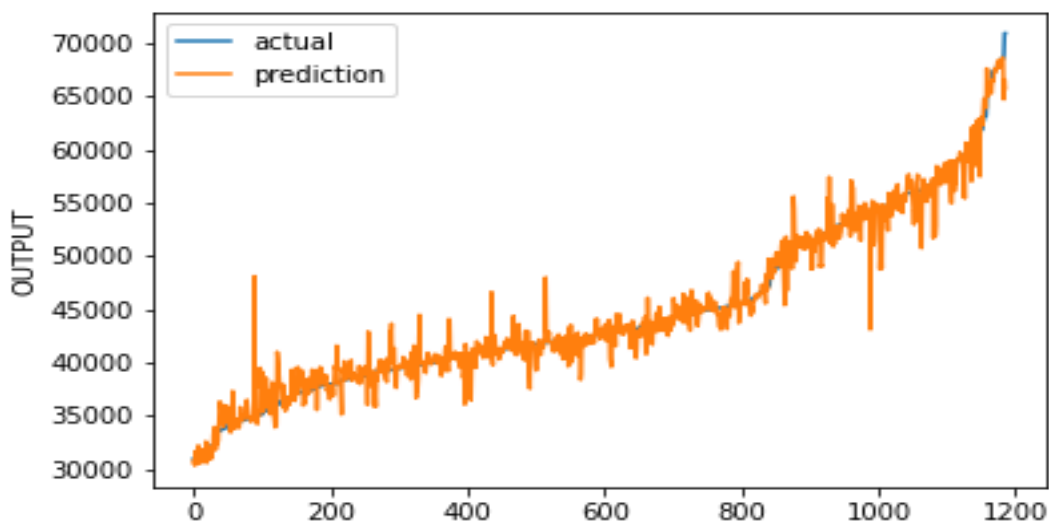


Figure 4.2.1: Accuracy plot (EUR)

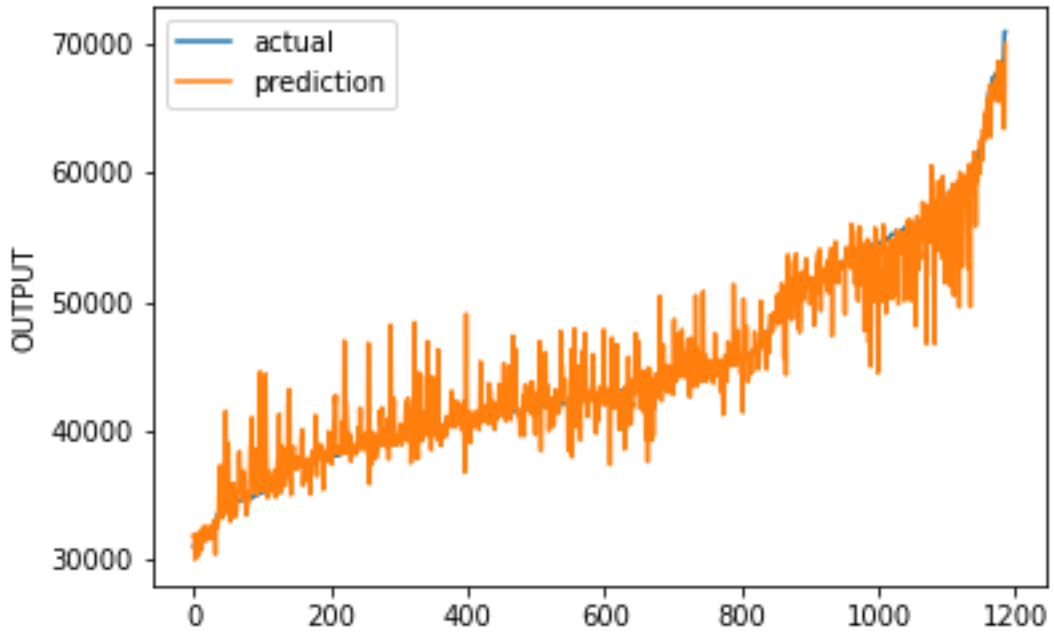


Figure 4.2.2: Accuracy plot (GBP)

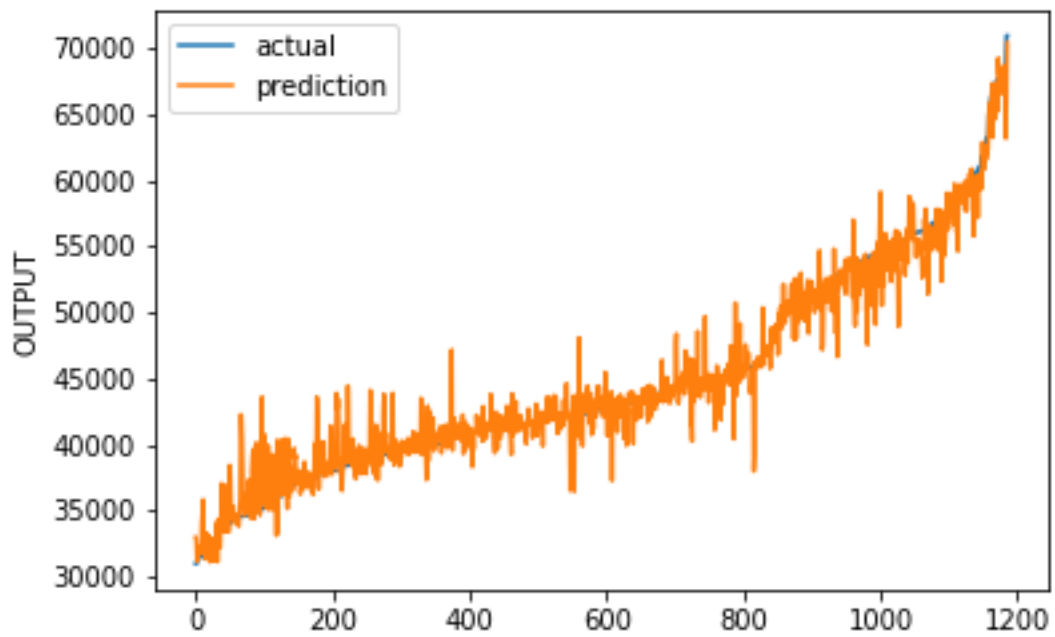


Figure 4.2.3: Accuracy plot (USD , EUR , GBP)

Random Forest Regression:

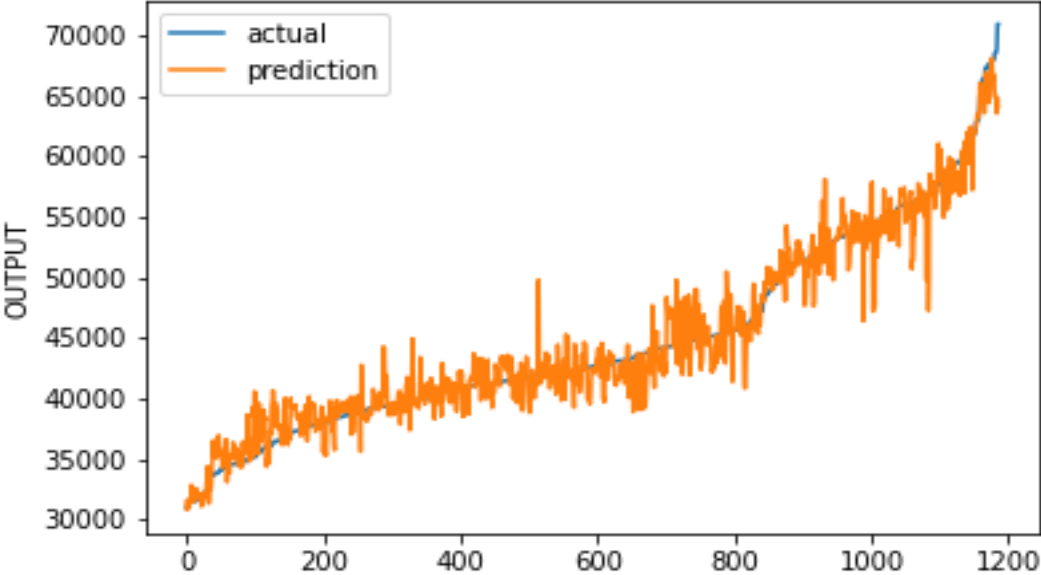


Figure 4.2.4: Accuracy plot (USD)

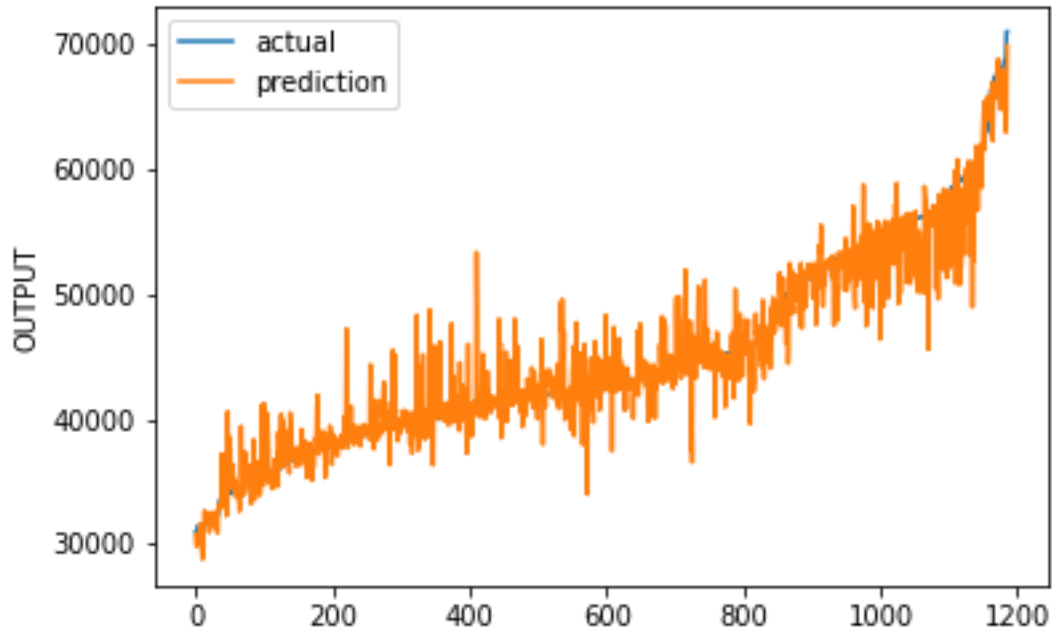


Figure 4.2.5: Accuracy plot (EUR)

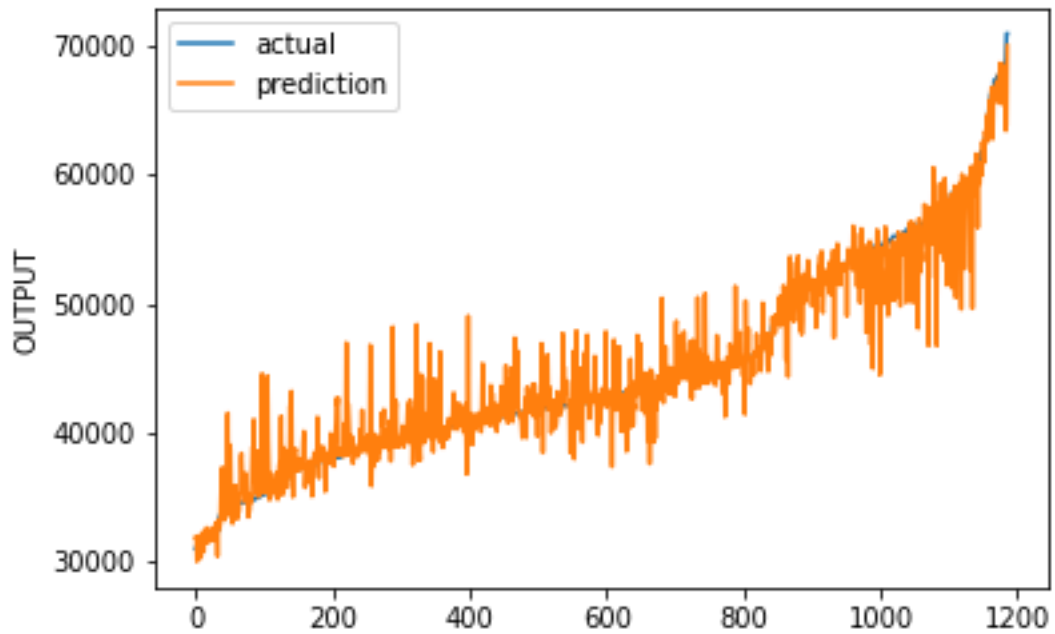


Figure 4.2.6: Accuracy plot (GBP)

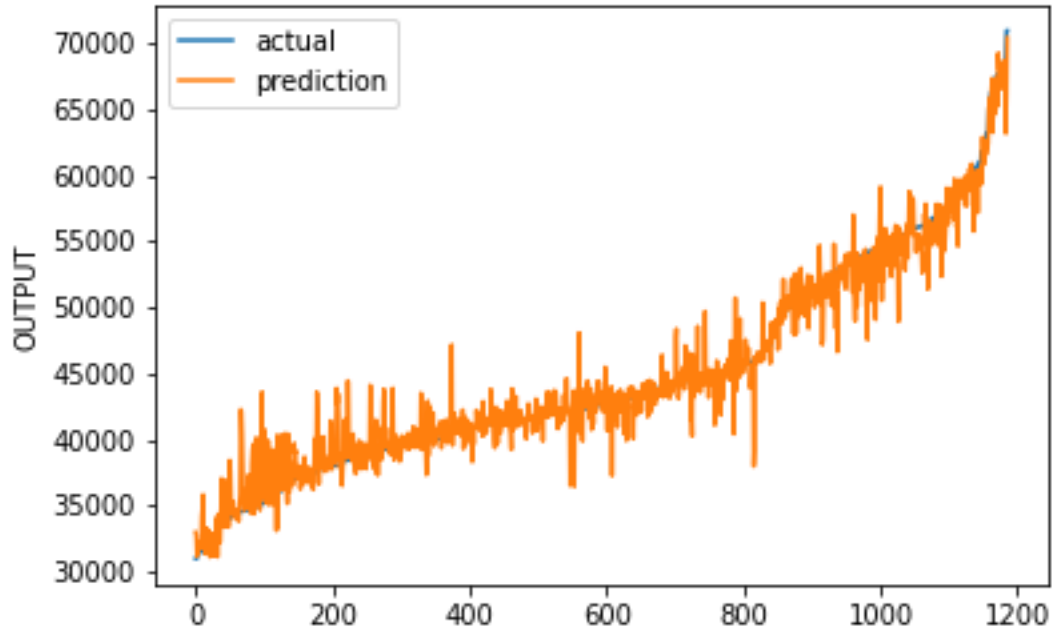


Figure 4.2.7: Accuracy plot (USD, EUR, GBP)

Decision tree:

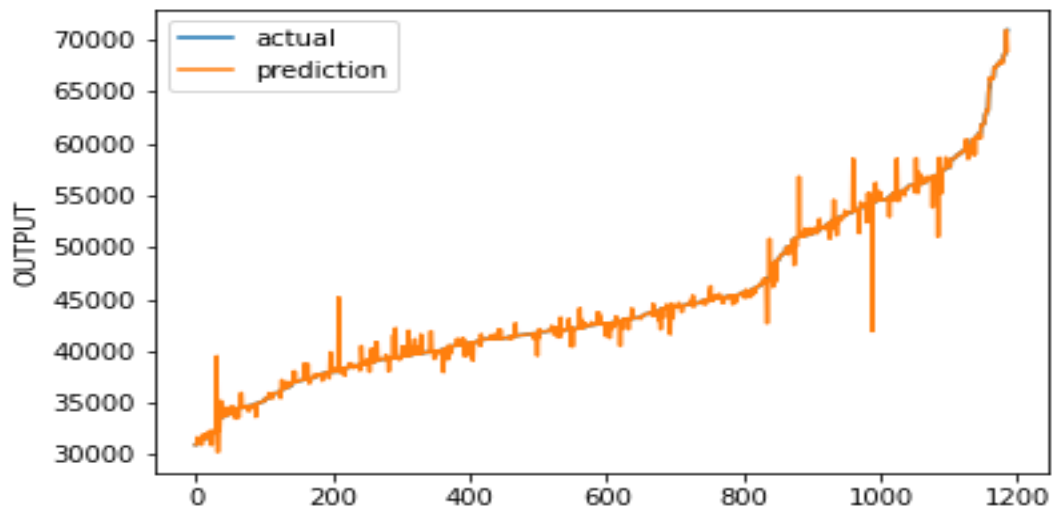


Figure 4.2.8: Accuracy plot (USD)

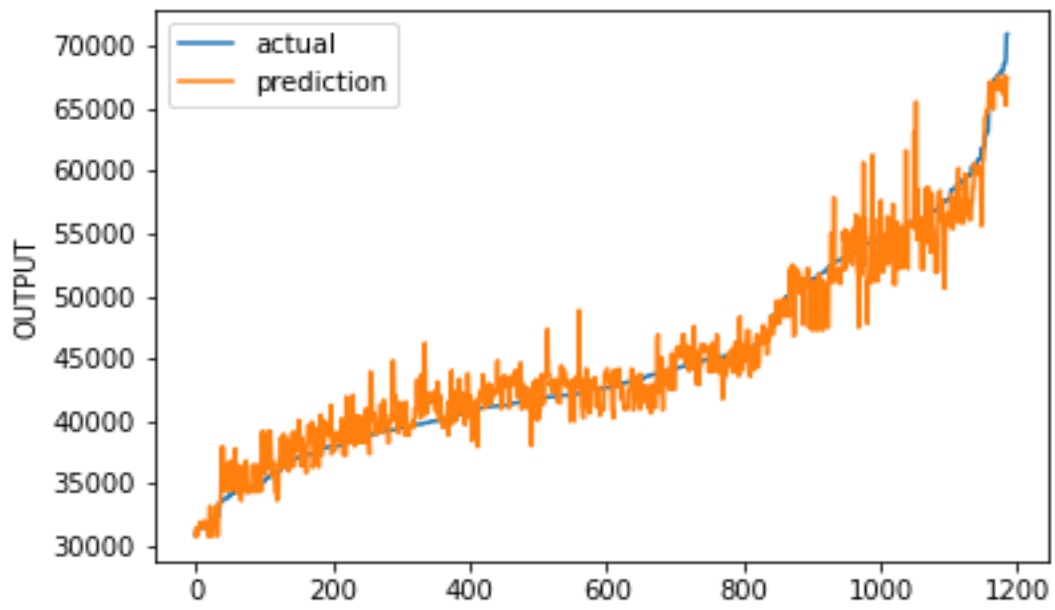


Figure 4.2.9: Accuracy plot (EUR)

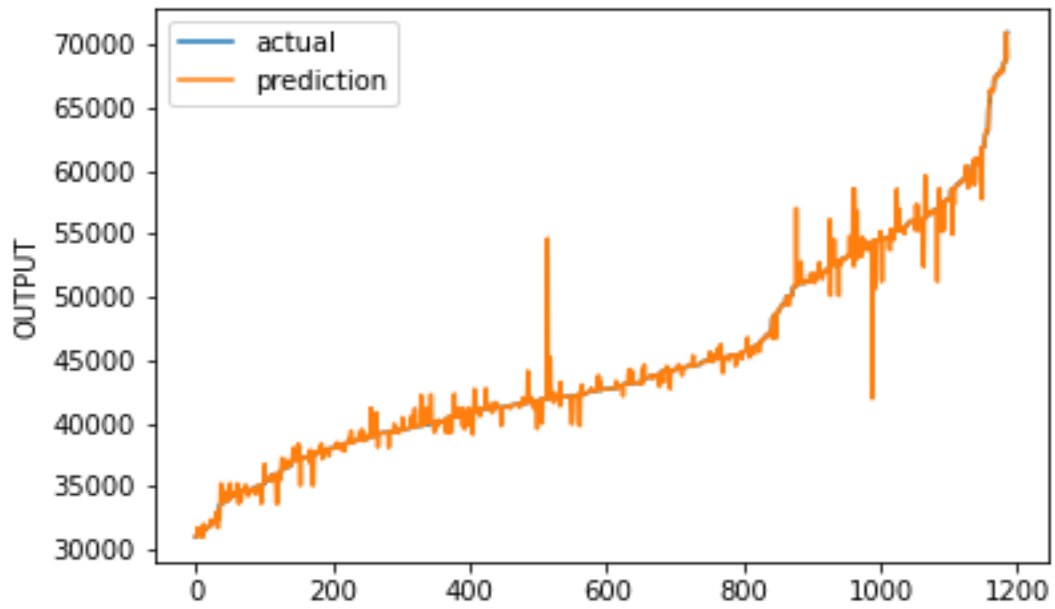


Figure 4.2.10: Accuracy plot (GBP)

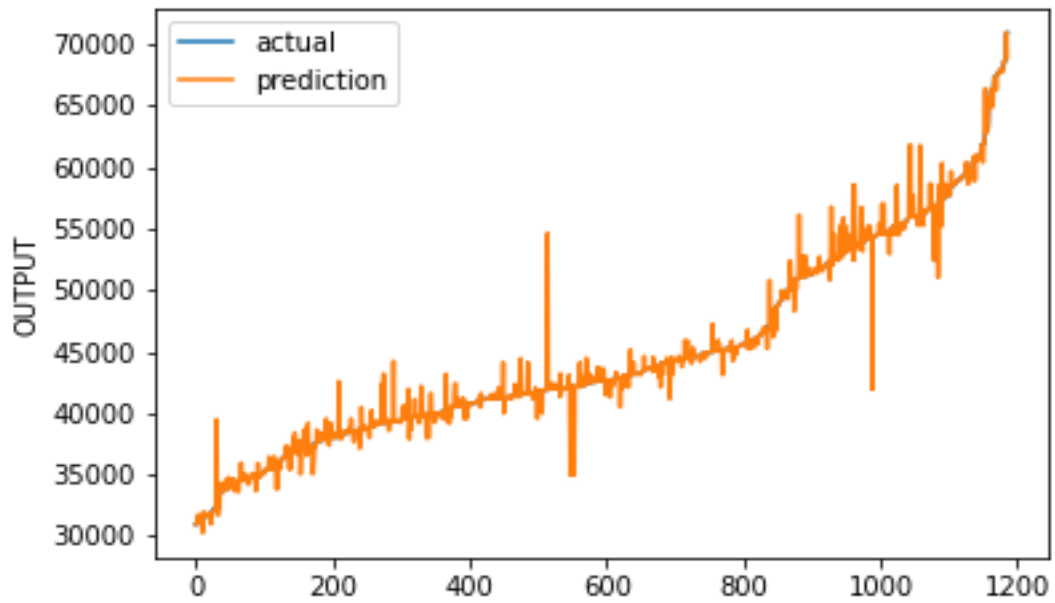


Figure 4.2.11: Accuracy plot (USD, EUR, GBP)

XGBoost:

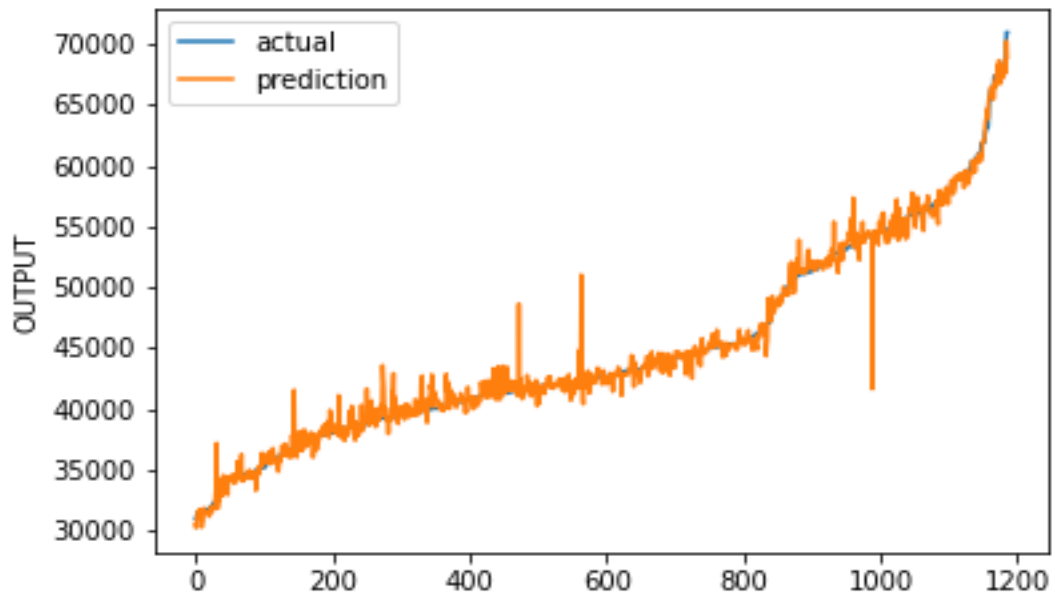


Figure 4.2.12: Accuracy plot (USD)

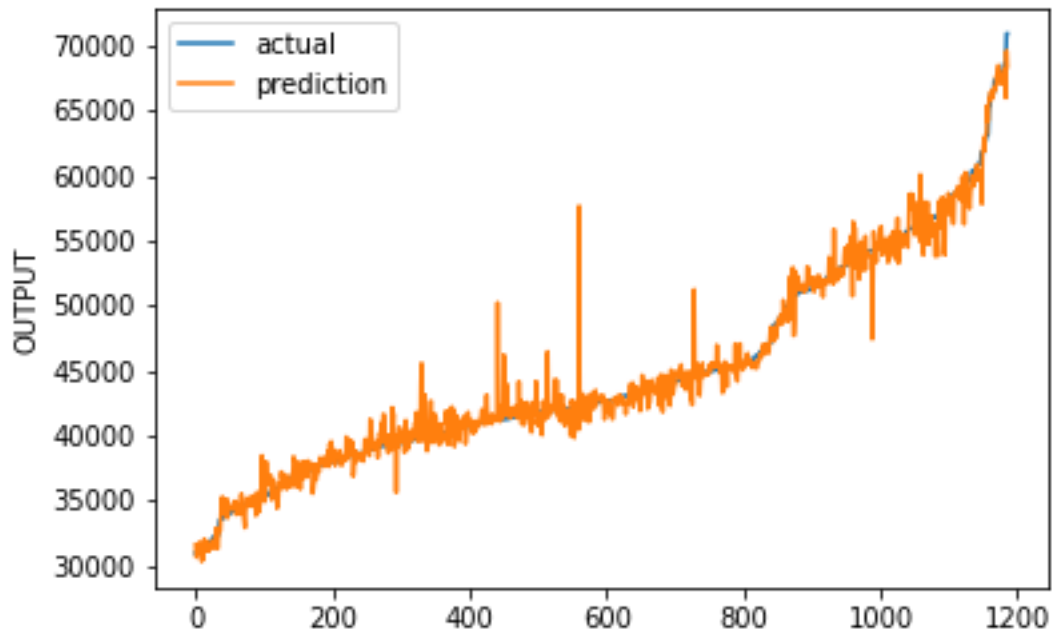


Figure 4.2.13: Accuracy plot (EUR)

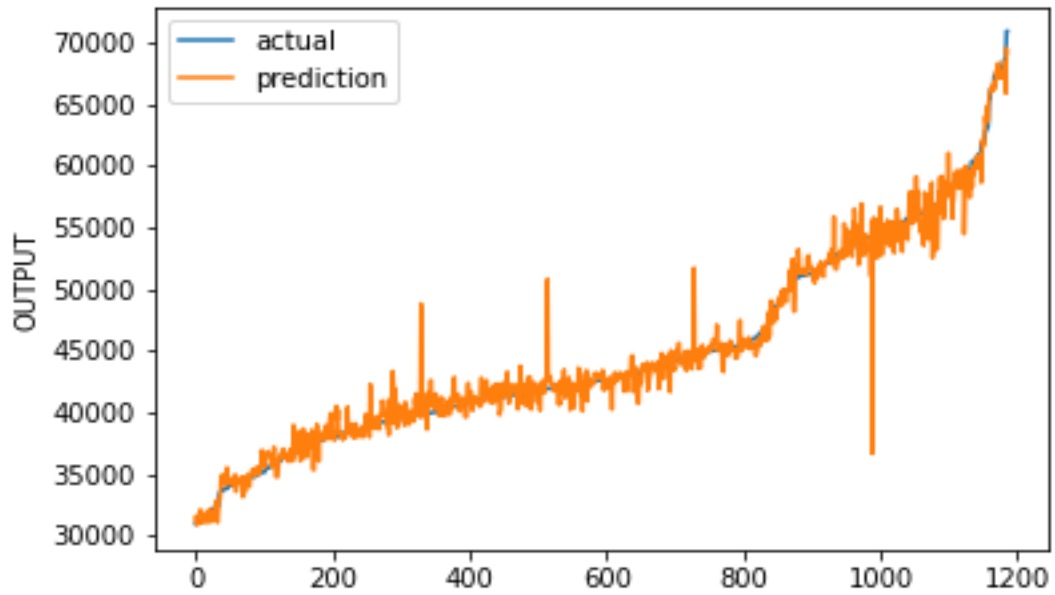


Figure 4.2.14: Accuracy plot (GBP)

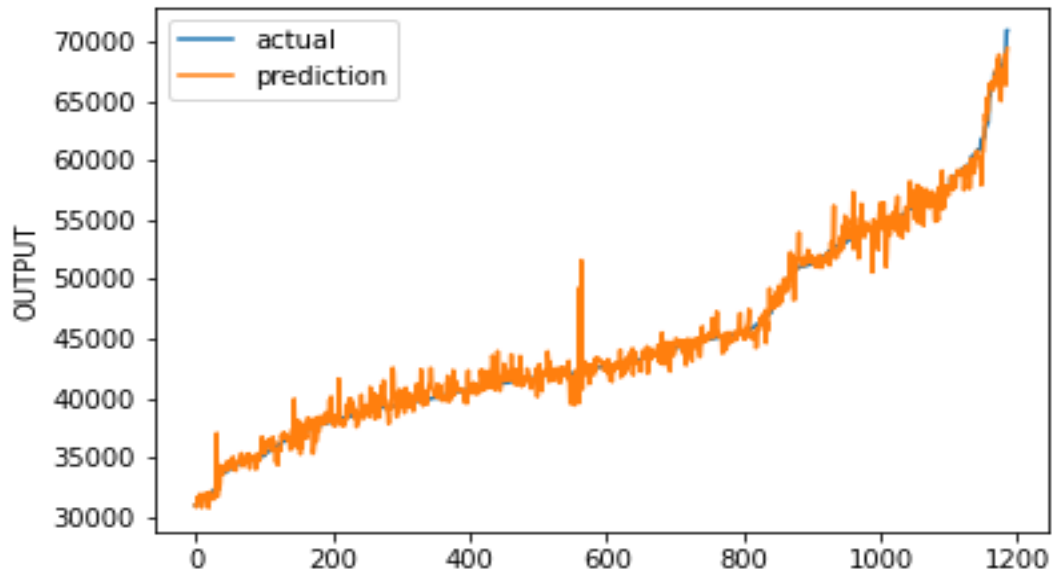


Figure 4.2.15: Accuracy plot (USD, EUR, GBP)

Gradient boost:

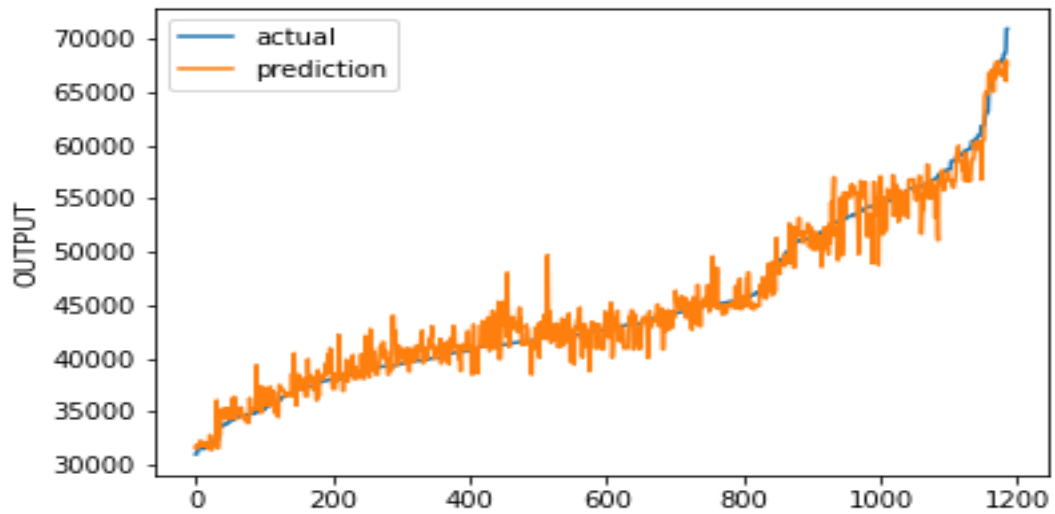


Figure 4.2.16: Accuracy plot (USD)

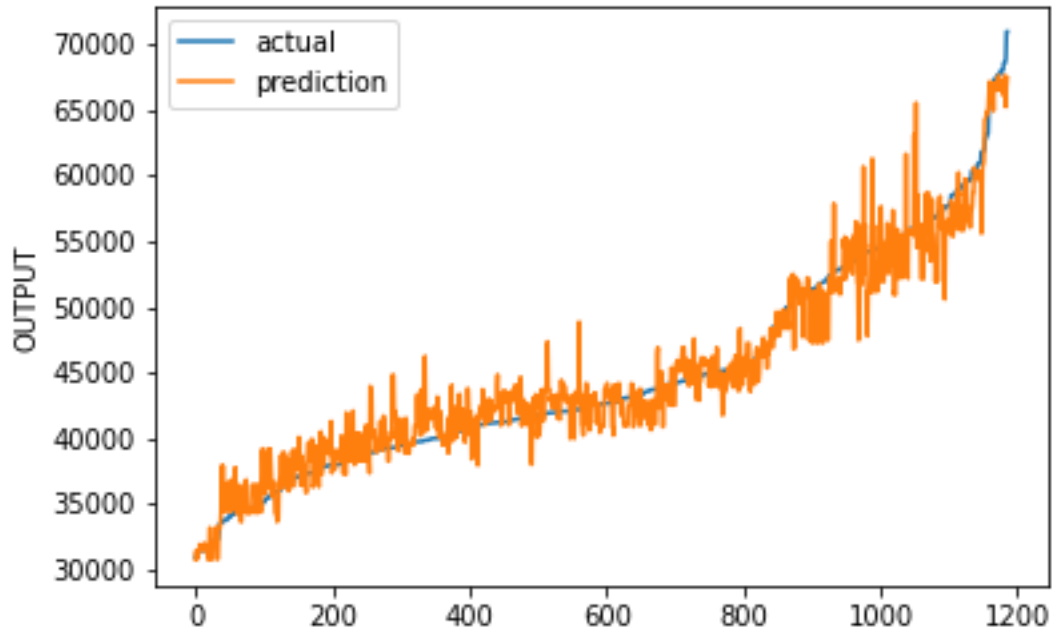


Figure 4.2.17: Accuracy plot (EUR)

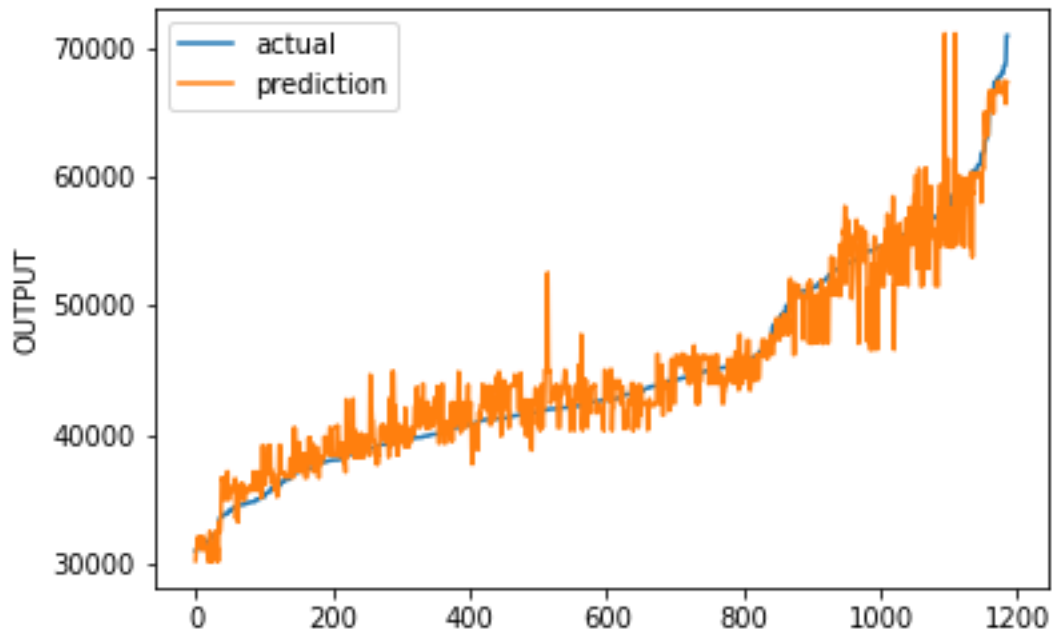


Figure 4.2.18: Accuracy plot (GBP)

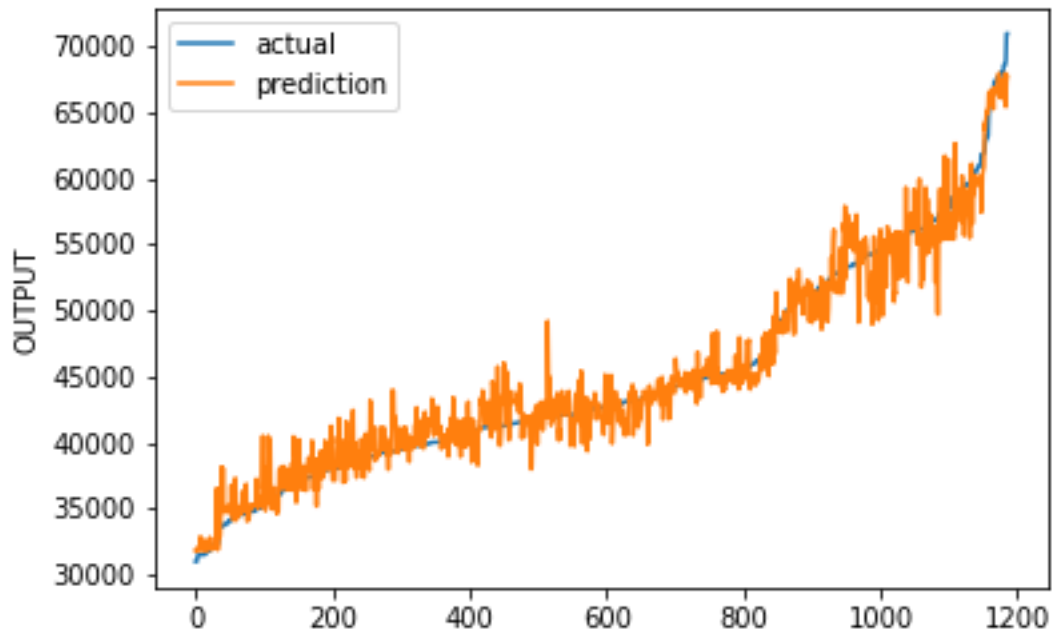


Figure 4.2.19: Accuracy plot (USD, EUR, GBP)

CHAPTER 5

RESULT COMPARISON AND ANALYSIS

SVR, RFR, Decision Tree, Gradienboosting, and Xgboosting are five models used in our study to examine the execution of forecasting gold actions. Table 5.1 presents the result.

Models	MAE	MSE	RMSE
SVR	1043.18	2665137.89	32.29
RFR	437.28	777916.65	20.91
Decision Tree	386.07	1118381.99	19.64
Gradienboosting	1160.54	2492841.86	34.06
XGBoost	628.76	867895.26	25.07

Table 5.1: Performance comparison of different models

In terms of the cumulative dataset, the Random Forest Regressor appears to have the lowest mean square error of all the models, according to Table 5.1. This model has promising approximating capabilities as compared to other models.

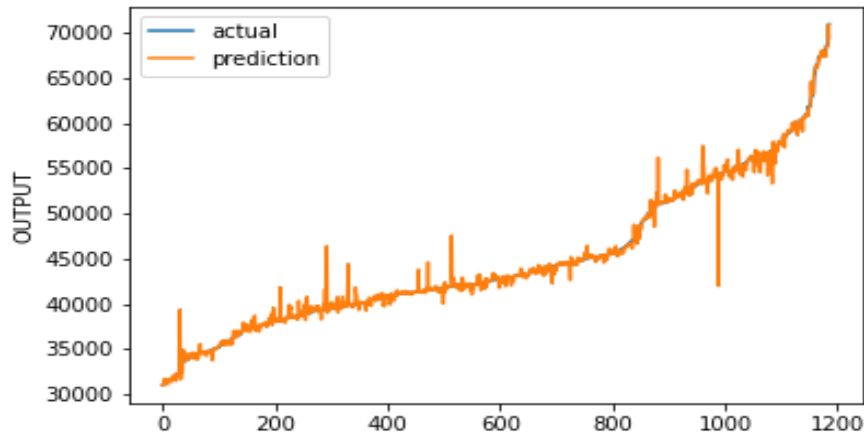


Figure 5.1: RFR prediction accuracy

XG boosting, on the other hand, has the lowest RMSE value. Despite the fact that the RFR algorithm has the lowest MAE, it outperformed XG boosting in terms of precision.

Models	US dollar			Euro			Gbp		
	MAE	MSE	RMS E	MAE	MSE	RMS E	MAE	MSE	RMS E
SVR	883.7	1970842.4	29.7	1487.8	5264016.6	38.5	1297.6	4473327.9	36.0
RFR	322.2	547929.9	17.9	373.9	655800.3	19.3	370.7	643645.3	19.2
Decision Tree	231.2	571984.3	15.2	284.7	755805.4	16.8	262.8	716881.5	16.2
Gradient boosting	1067.5	2058822.6	32.6	1338.4	3190612.7	36.5	1433.7	3839329.5	37.8
XGBoost	549.6	810541.3	23.4	673.4	1210380.0	25.9	714.5	1346281.6	26.7

Table 5.2: Performance comparison of different models with different currency

Table 5.2 shows that the RFR algorithm provides promising accuracy for each currency individually. The RFR algorithm's prediction accuracy is relatively high, as seen in the graph.

Accuracy comparison:

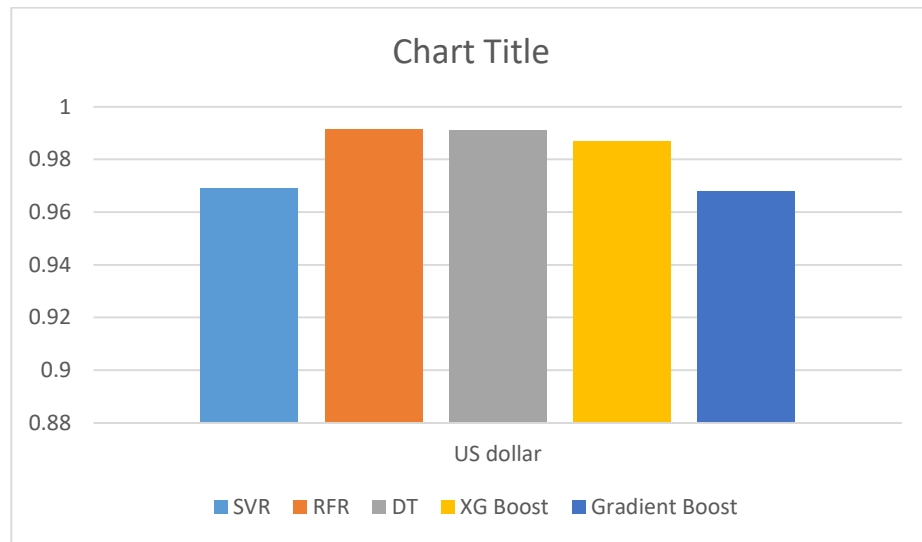


Figure 5.2: Accuracy chart (USD)

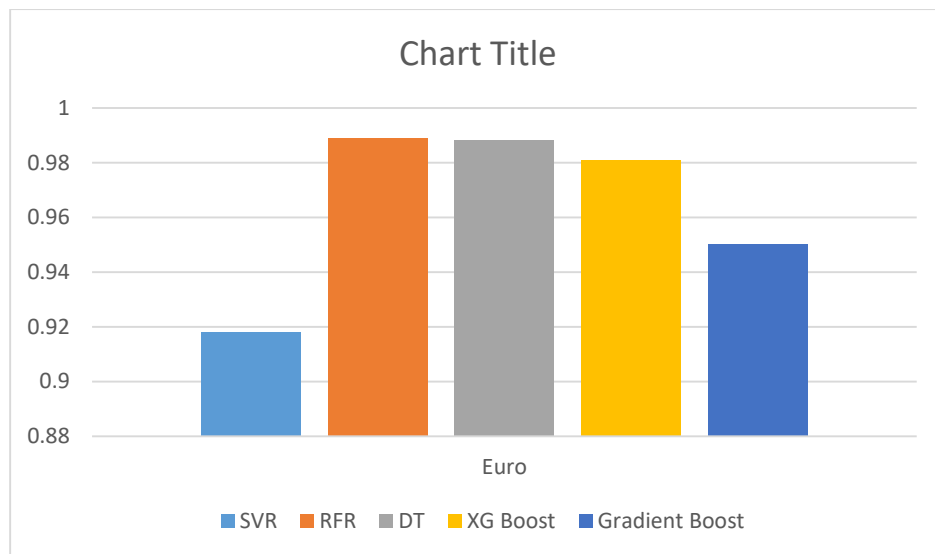


Figure 5.3: Accuracy chart (EUR)

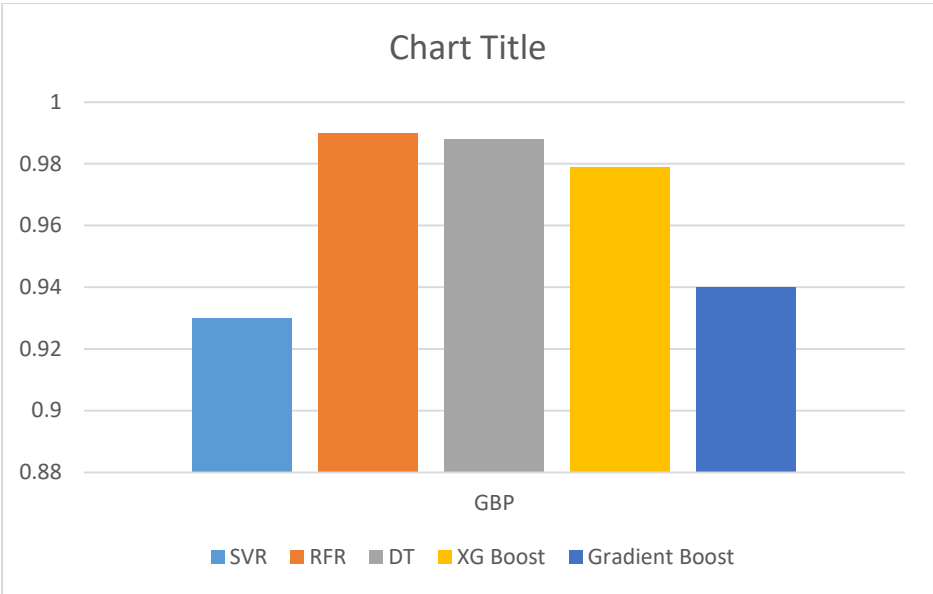


Figure 5.4: Accuracy chart (GBP)

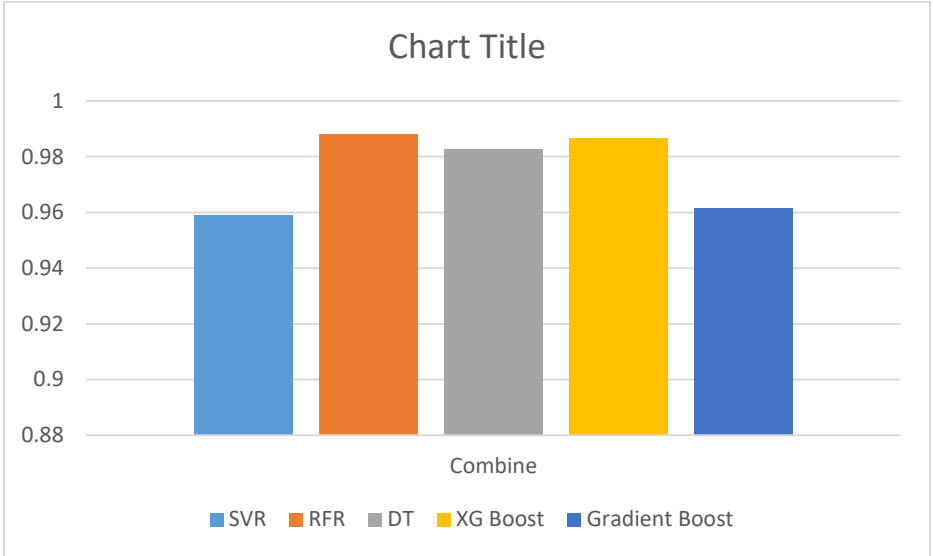


Figure 5.5: accuracy chart (USD, EUR, GBP)

CHAPTER 6

CONCLUSION AND FUTURE WORK

Accurate gold price forecasting can help to reduce stock market and economic instability to some extent. The ability to reliably predict gold prices will have a significant effect on a country's economic stability. About any algorithm has produced satisfactory results. Random Forest Regression (RFR) performed the best out of all the algorithms. In any event, it has achieved a level of precision of about 99 percent. A bigger dataset in the future could improve the accuracy of the algorithms even further. The use of Neural Network Algorithms could improve the accuracy of gold price forecasting. We want to develop an intelligence framework that can forecast potential gold prices and thereby support our country's economy.

REFERENCE

- [1]. Makala, D., & Li, Z. (2021, February). Prediction of gold price with ARIMA and SVM. In *Journal of Physics: Conference Series* (Vol. 1767, No. 1, p. 012022). IOP Publishing.
- [2]. Aruna, S., Umamaheswari, P., & Sujipriya, J. (2021). Prediction of Potential Gold Prices using Machine Learning Approach. *Annals of the Romanian Society for Cell Biology*, 1385-1396.
- [3]. 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.
- [4]. Monburinon, N., Chertchom, P., Kaewkiriyaa, T., Rungpheung, S., Buya, S., & Boonpou, P. (2018, May). Prediction of prices for used car by using regression models. In *2018 5th International Conference on Business and Industrial Research (ICBIR)* (pp. 115-119). IEEE.
- [5]. Rathan, K., Sai, S. V., & Manikanta, T. S. (2019, April). Crypto-currency price prediction using decision tree and regression techniques. In *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 190-194). IEEE.
- [6]. Ravikumar, S., & Saraf, P. (2020, June). Prediction of Stock Prices using Machine Learning (Regression, Classification) Algorithms. In *2020 International Conference for Emerging Technology (INCET)* (pp. 1-5). IEEE.
- [7]. Manjula, K. A., & Karthikeyan, P. (2019, April). Gold Price Prediction using Ensemble based Machine Learning Techniques. In *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 1360-1364). IEEE.
- [8]. Pandey, A. C., Misra, S., & Saxena, M. (2019, August). Gold and diamond price prediction using enhanced ensemble learning. In *2019 Twelfth International Conference on Contemporary Computing (IC3)* (pp. 1-4). IEEE.
- [9]. 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.
- [10]. Salis, V. E., Kumari, A., & Singh, A. (2019). Prediction of gold stock market using hybrid approach. In *Emerging Research in Electronics, Computer Science and Technology* (pp. 803-812). Springer, Singapore.
- [11]. Hasan, M. M., Zahara, M. T., Sykot, M. M., Nur, A. U., Saifuzzaman, M., & Hafiz, R. (2020, July). Ascertaining the Fluctuation of Rice Price in Bangladesh Using Machine Learning Approach. In *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-5). IEEE.
- [12]. 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.

Gold Price Forecasting

ORIGINALITY REPORT

16%

SIMILARITY INDEX

11%

INTERNET SOURCES

5%

PUBLICATIONS

10%

STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Daffodil International University Student Paper	3%
2	towardsdatascience.com Internet Source	1%
3	theses.gla.ac.uk Internet Source	1%
4	Submitted to University of Reading Student Paper	1%
5	Submitted to CSU Northridge Student Paper	1%
6	Submitted to University of Hertfordshire Student Paper	1%
7	Avinash Chandra Pandey, Shubhangi Misra, Mridul Saxena. "Gold and Diamond Price Prediction Using Enhanced Ensemble Learning", 2019 Twelfth International Conference on Contemporary Computing (IC3), 2019 Publication	1%