**GOLD PRICE PREDICTION**

**PROJECT BASED LEARNING – VI (AIP106)**

**PROJECT REPORT**

Logo

Description automatically generated

**Submitted by**

**NAVYA (2110993819)**

**SHREYA (2110993836)**

**DEEPANSHI (2110993973)**

**BE-CSE (Artificial Intelligence)**

**Guided by**

**DR. SHRUTI ARORA**

**CHITKARA UNIVESITY OF ENGINEERING & TECHNOLOGY**

**CHITKARA UNIVERSITY, RAJPURA**

**APRIL, 2024**

# **ACKNOWLEDGEMENT**

With immense please we, Miss Navya, Miss Shreya and Miss Deepanshi presenting “Gold Price Prediction” project report as part of the curriculum of ‘BE - CSE (AI)’.

We would like to express our sincere thanks to Dr. Shruti Arora, for her valuable guidance and support in completing our project.

We would also like to express our gratitude towards our dean Dr. Sushil Kumar for giving us this great opportunity to do a project on “Gold Price Prediction”. Without their support and suggestions, this project would not have been complete.

Signature…..………….

Name: Navya

Roll No.: 2110993819

Signature…..………….

Name: Shreya

Roll No.: 2110993836

Signature…..………….

Name: Deepanshi

Roll No.: 2110993873

# 

**ABSTRACT**

Significantly, gold is among the precious metals that are utilized to finance trading purchases. In countries like India, China, Indonesia, and many more, gold is considered the ideal jewellery, in addition, gold is also served as a present/remembrance and even gold accessories are presented as marriage settlements. Moreover, the countries with large gold reserves are considered a booming nation. At present time, precious metal like gold, is not just considered ornaments or jewellery but are considered as an investment and are kept with all nations' central bank as an assurance for repayment of non-native loans, and also to manage inflation.

Due to the increasing demand and dispense of this asset in the market the state of the major economies throughout the globe has a considerable impact on gold prices. Due to the change in gold prices, more investors are now considering gold investments. But irregularity in the gold price in the market makes it riskier for the investor. Thus, the goal of “Gold price prediction” is to forecast gold’s price using a variety of Machine learning techniques, considering the relationship between several economic factors that influence gold rates. The machine learning algorithms that have been used are Random Forest regression.

# **TABLE OF CONTENTS**

ACKNOWLEDGEMENT

ABSTRACT

TABLE OF CONTENT

LIST OF FIGURES

**1.** INTRODUCTION

**2.** LITERATURE SURVEY

**3.** METHODOLOGY

**3.1** Research Methodology

**3.1.1** Assumptions for Random Forest

**3.1.2** Why Random Forest?

**3.1.3** Working process of Random Forest

**3.2** Description of algorithm

**4.** DATA SET

**5.** FLOW CHART

**6.** ACTUAL VALUE AND PREDECTED VALUE

**7.** EXPERIMENTAL INVESTIGATION

8. PROJECT HIGHLITES

9. CORRELATION ANALYSIS

10. CONCLUSION

11. REFERENCES

# **LIST OF FIGURES**

1. Dataset image
2. Test Data Prediction
3. Actual vs Predicted Graph
4. Read Data
5. Web Page Code
6. Model Training through Random Forest Regression
7. Distribution of the GLD Price Graph
8. Correlation Analysis of different attributes
9. Web page interface before prediction
10. Web page interface after prediction

# **1. INTRODUCTION**

Gold was the first well-known metal of our species. When we ponder the historical advancement of technology, we consider the development of iron and copper labor to be the greatest contributor to the economic and cultural advancement of our species, but gold came first. Gold has held its value and has been used as a means of assessing a country's financial strength. Big investors were attracted to this precious metal and invested large amounts in it.

In the early days, more money was invested in buying this basic product. Like most commodities, the price of gold is determined by supply and demand, including speculative demand. However, unlike most other raw materials, savings and disposal play a bigger role in influencing your price. Small investors also found this product to be a safe investment. Government investments in gold are largely determined by your financial conditions and interest rates as they are indicators of the strength of your economy. Activity is observed in the US, hence capital inflows into the gold market. Various phenomena are associated with gold rates and also affect the price. Gold spot prices are set twice a day based on the supply and demand of the gold market. A slight change in the price of gold can result in large gains or losses for both these investors and government banks.

The reason for the rise in the gold rate apparently lies in its incredible use and it is also a very rare metal to be found. There are several other reasons for the price intensification as well. Gold is used in various fields like finance, trade, mechanical. Industrial, dental and medical applications account for approximately 12% of the gold requirement. Gold has high thermal and electrical conductivity properties as well as high resistance to corrosion and bacterial colonization. It has fluctuated in recent years due to the constant expansion of the middle classes in emerging markets seeking western lifestyles. To extract a small amount of gold, a larger amount of gold ore could be used; along with a lot of staff associated with it. If the ore is of lesser quality, only 5 grams of gold can be extracted from a large ton of ore.

In fact, this metal can be ductile, fluffy, and malleable. easily in different ways. It becomes very flexible so that other metals can be added to make useful ornaments, and this valuable product is used for various purposes around the world. It has a decent cathode of power and heat. It also has the ability to keep through any atmosphere; not affected by moisture, air or the most dangerous or destructive elements. It is the best example of reliable materials. In addition, the recycling of used jewellery has grown into a multi-billion-dollar industry.

**2. LITERATURE SURVEY**

The journal “Gold price prediction” studied about machine learning system that can predict gold prices based on several other economic variables. By using this they compared the correlation between economic variables and gold price. To predict gold prices, they used several machine learning algorithms. They used linear regression, SVM, and Decision Tree to train the model for predicting gold rates. According to the second journal “Future gold prices can be predicted using machine learning techniques” one of the most significant metals in the world is gold.

Many countries maintain their gold reserves to be recognized as healthy and progressive countries, so, based on this they predicted the gold prices, so investors can invest in this commodity by analysing the proposed model and can get huge benefits. In her article "Modelling and Forecasting of Gold Prices on Financial Markets," V.K.F.B. Rebecca Davis makes use of the Autoregressive Moving Average (ARMA) model, a statistical tool that is often used to analyse time series data. The monthly prices of gold during a ten-year period are the data under consideration. The accuracy of the model was 66.67%. In their study titled "Predicting Future Gold Rates using Machine

Over the course of eleven years, the data for this study were gathered from diverse sources. This information includes factors like the price of crude oil, the S&P 500 index, USD exchange rates, and other economic factors. In their article "Prediction of the gold price with ARIMA and SVM," D Makala and Z Li use information gathered from the World Gold Council that includes daily gold prices from January 1979 to December 2019. To predict the price of gold, this study uses the Autoregressive Integrated Moving Average (ARIMA) approach with SVM. The ARIMA model's accuracy is lower than the SVM model's accuracy.

## 

**3. METHODOLOGY**

Gold Investments Since ancient times, the value of the precious metal has been high in various periods despite the economic and financial crisis. For the past few years, the value of the currency has fluctuated depending on the currency market, the price of crude oil and inflation. The gold rate is also unstable. Investors or customers have a higher risk of investing or buying them. Since gold is viewed as essential to the core and a liquid asset around the world, trading is very easy. Hence, it is widely used as a precious metal. The forecast for gold is of great help in planning and implementing future investment results. Predicting the gold rate is not only intended to give hope to people, but also to protect money in this scenario as its value fluctuates dramatically. Units are traded like stocks, and predicting stocks is always a challenging problem because their non-linearity makes them non-trivial.

As a result, shareholders will devote themselves to protecting themselves from political and monetary expansion and social fiasco. Gold is benevolent because there is no "crop rotation fluctuation in the market". Therefore, along with individuals in multinational corporations, they also invested in gold reserves. In the meantime, this precious metal has become more of a currency, so the government is likely to even increase the gold reserve. Usually, people assume that the benefit of gold is that it helps in tough times as there is a high level of liquidity control. According to government policy and gold regulators, the price of gold fluctuates daily. However, estimating the increases and decreases in monthly gold rates can help finance professionals decide when to buy or offer these products.

Another payment method is gold, which is used around the world for additional business transactions. In this modern scenario, the central banks of all countries keep expensive metals to ensure the payment of foreign debt, regulate inflation and also reflect financial strength. Alloy gold is required to have a more remarkable retention area in the placement of ultrasonic aluminium wire connections than ordinary alloy metallization. Understanding the factors affecting gold rates is imperative for any potential gold buyer so that they can more accurately predict interest rate patterns and, consequently, lead a company to a superior profit. Investors, researchers also have a great interest in understanding the gold business with the aim of a concrete study and an accurate forecast.

**3.1. RESEARCH METHODOLOGY**

Our project aims to improve the accuracy of gold price prediction by utilizing a dataset with various attributes in addition to the existing system. The Random Forest algorithm, which is a classifier that employs multiple decision trees on different subsets of the dataset, is utilized to enhance predictive accuracy. Instead of relying on a single decision tree, Random Forest calculates the average of predictions from each tree and makes the final output prediction based on the majority vote. A higher number of trees in the forest enhances accuracy and mitigates overfitting issues. This approach results in higher predictive accuracy.

**3.1.1 Assumptions for Random Forest**

The random forest algorithm utilizes multiple decision trees to predict the class of the dataset. Due to the use of multiple trees, it is possible that some trees may predict the correct output while others may not. However, by combining the predictions of all trees, the random forest algorithm produces the correct output. To improve the accuracy of the random forest classifier, there are two assumptions that need to be considered. Firstly, the feature variable of the dataset should have actual values to ensure the classifier predicts accurate results, rather than random or estimated results. Secondly, the predictions from each tree should have minimal correlations to avoid any biases and improve the overall accuracy of the classifier.

**3.1.2Why Random Forest?**

• The algorithm requires less time for training in comparison to other algorithms.

• It has a high accuracy rate in predicting output and performs efficiently even when dealing with large datasets.

• Additionally, it can maintain accuracy even when a considerable amount of data is missing.

**3.1.3 Working process of Random Forest**

Random Forest operates in two stages: the first stage involves creating the forest by combining multiple decision trees, and the second stage involves making predictions for each tree that was created in the first stage. The following steps and diagram illustrate the process:

Step 1: Choose K data points at random from the training set.

Step 2: Build decision trees using the selected data points (subsets).

Step 3: Select the number of decision trees, N, that you want to build.

Step 4: Repeat Steps 1 and 2.

Step 5: When presented with new data points, find the predictions of each decision tree, and assign the new data points to the category with the highest number of votes.

**3.2. Description of algorithm**

Step 1: Start with a dataset containing various data points.

Step 2: Assign equal weights to each data point as input to the model.

Step 3: Identify the data points that are incorrectly classified and increase their weight.

Step 4: If the desired output is obtained, stop the process. Otherwise, go back to Step 2 and repeat the process.

**4. DATASET**

The provided dataset consists of daily financial market data spanning from January 2nd, 2008, to February 1st, 2008. It includes the following variables:

1. **Date**: The date of the recorded market data.
2. **SPX**: The closing price of the S&P 500 index, representing the performance of the US stock market.
3. **GLD**: The closing price of gold, a commonly traded precious metal often considered a safe-haven asset.
4. **USO**: The closing price of the United States Oil Fund, representing the performance of the oil market.
5. **SLV**: The closing price of silver, another precious metal often traded alongside gold.
6. **EUR/USD**: The exchange rate of the Euro against the US Dollar, indicating the strength of the Euro relative to the Dollar in international currency markets.

This dataset provides a snapshot of daily market movements and exchange rates, which can be used for various analyses such as trend identification, correlation studies between different assets, and potential insights into global economic dynamics during the specified time period.

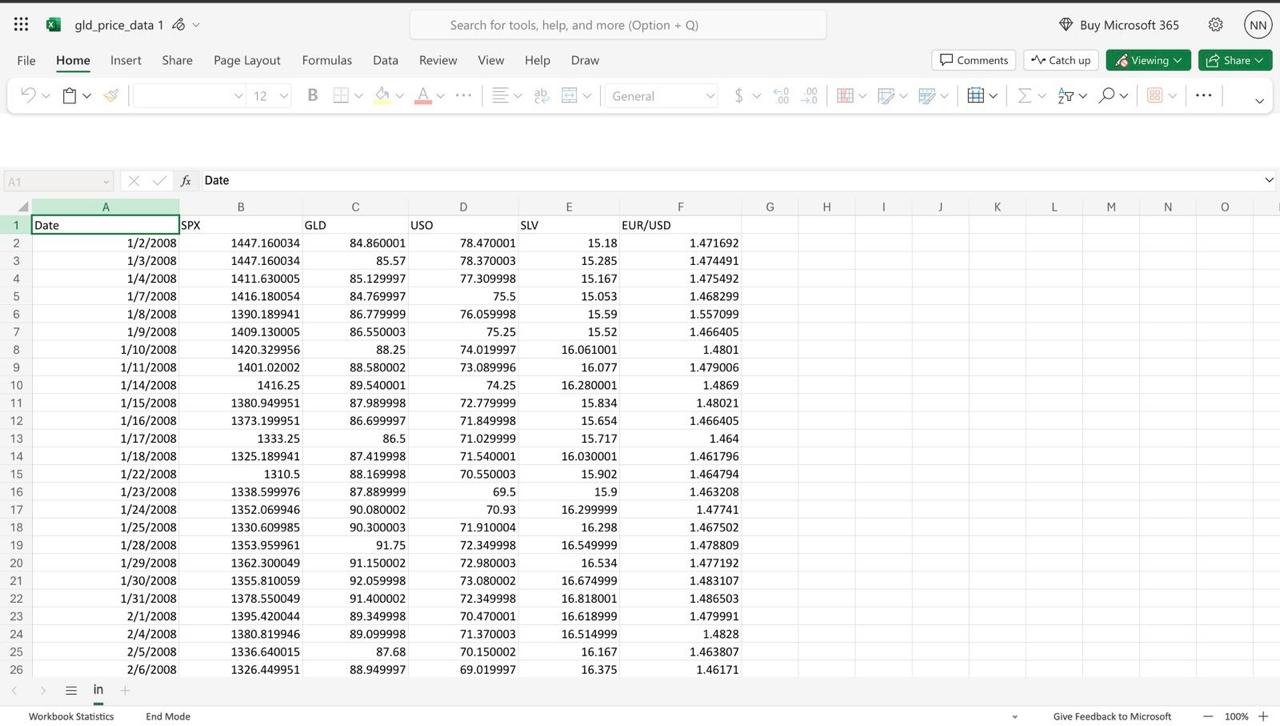


Fig1: Dataset Image

The provided values appear to be predictions generated by a model applied to test data. These predictions likely represent estimated values for a specific target variable, such as stock prices or currency exchange rates, based on input features.

It's essential to note that the accuracy and reliability of these predictions depend on various factors, including the quality of the training data, the chosen model architecture, and the features considered. Evaluating the model's performance metrics, such as mean squared error or mean absolute error, can provide insights into its predictive accuracy.

Additionally, further analysis could involve comparing these predicted values against the actual observed values to assess the model's effectiveness in capturing the underlying trends and patterns in the data. This comparison can help stakeholders understand the model's strengths and limitations and guide decision-making processes accordingly.

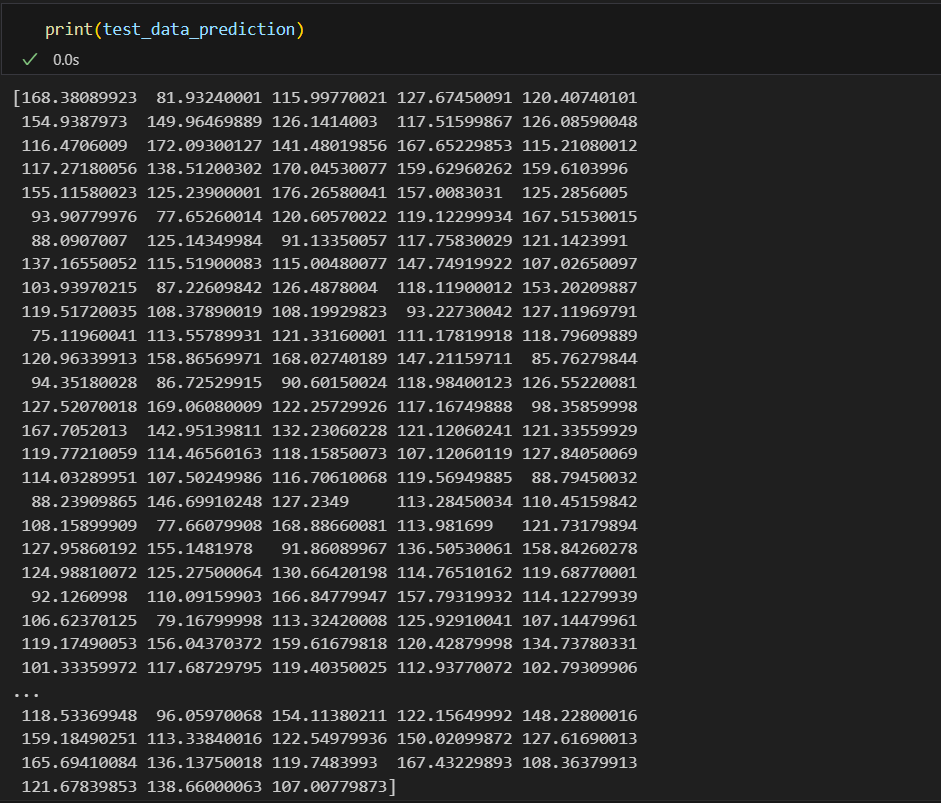
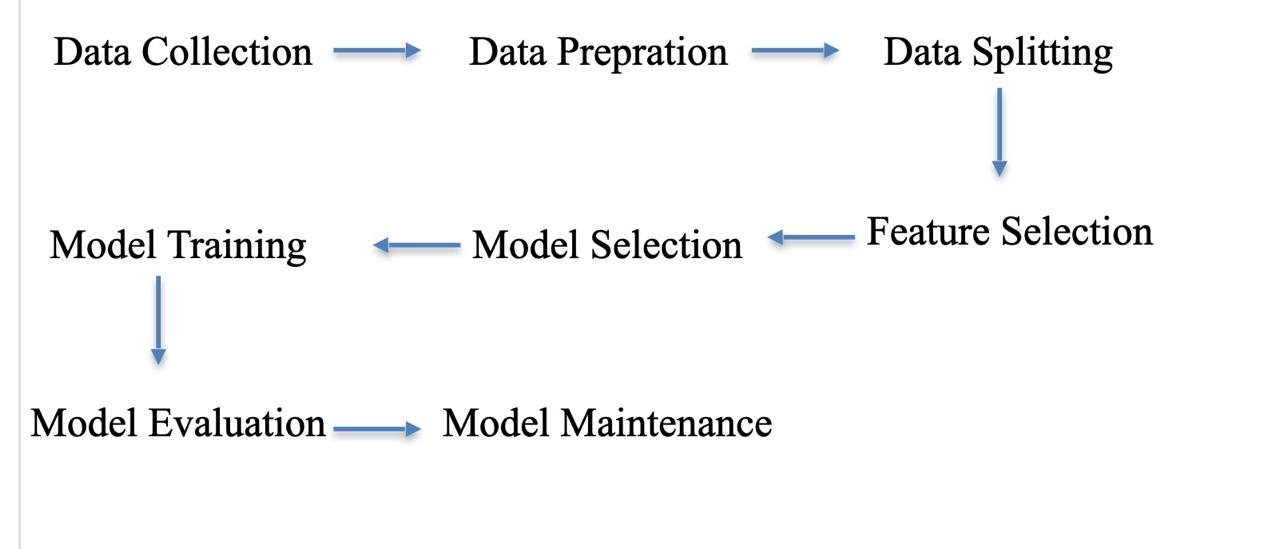


Fig2: Test Data Prediction

**5. Top of Form**

**FLOWCHART**



**6. ACTUAL VALUE AND PREDECTED VALUE**

The link between real and anticipated values obtained from our machine learning model is depicted in the graph. The target variable's related values are indicated by the y-axis, while the x-axis shows the instances or time periods. The observed ground truth data points gathered over the designated time period are represented by actual values, which are represented by a solid line. The model's estimates for the identical cases are displayed by anticipated values, which are indicated by a dashed line or different markers.

A clear evaluation of the model's effectiveness in identifying the underlying patterns and trends in the data is made possible by this visual display. A high degree of precision and dependability in the model's predictions is indicated by a close match between the actual and projected values. On the other hand, considerable differences or deviations between the two lines might point to places where the model has trouble correctly predicting the target variable.

By examining this graph, one can gain important knowledge about the machine learning model's efficacy and make recommendations for future improvements or modifications that will improve its predictive power.

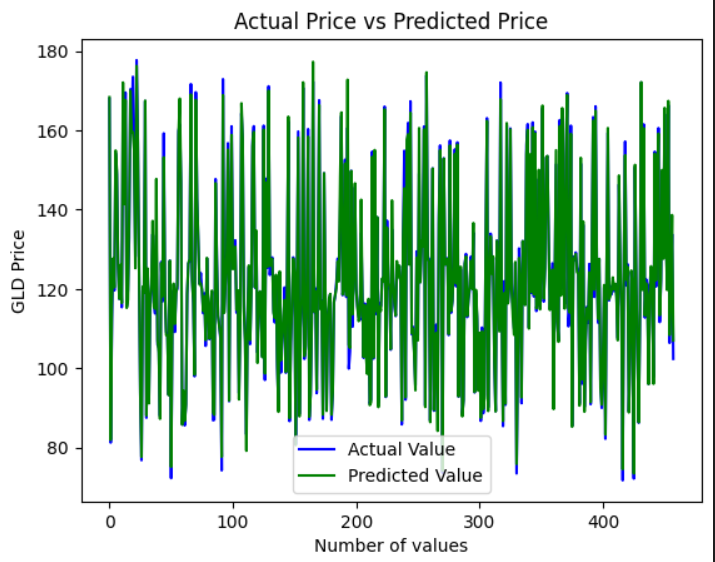


Fig3: Actual vs Predicted Graph

**7. EXPERIMENTAL INVESTIGATION**

The model is trained on processor Intel, core i5 and RAM 8GB. The system type is windows 10 ultimate of 64 bit operating.

The result has a good amount of accuracy. This model is tested with various No. of data and the results are good.

**8. Project Highlights**

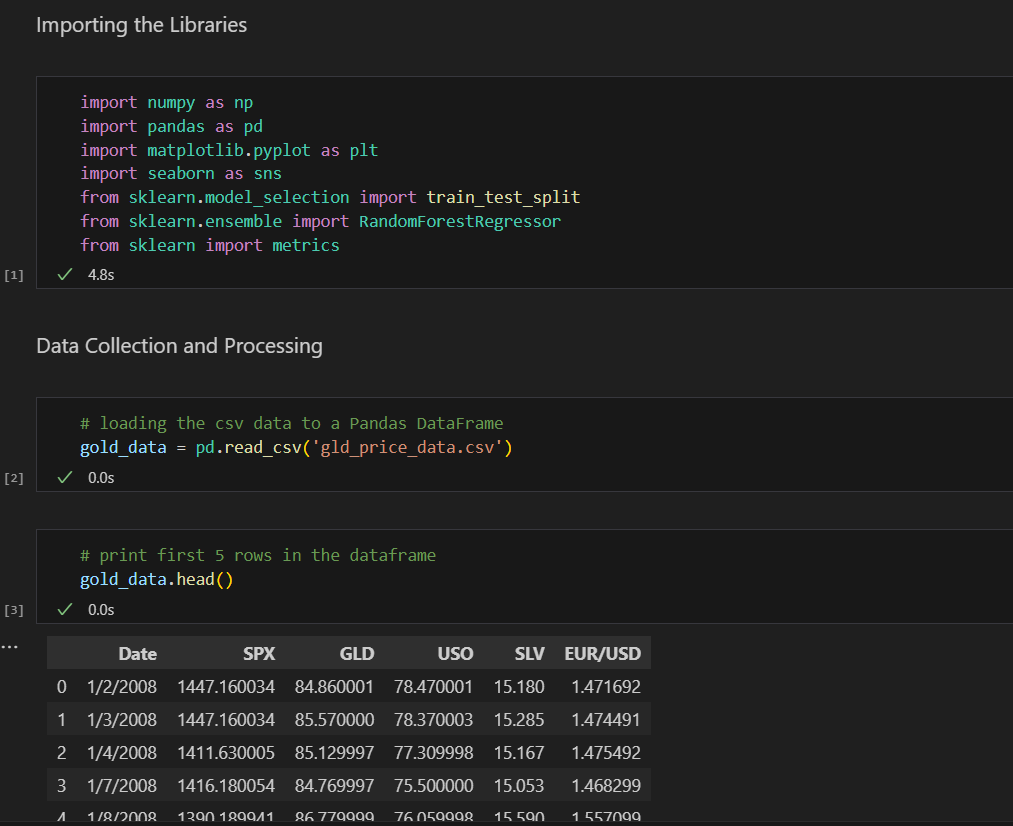
****

Fig4: Read Data

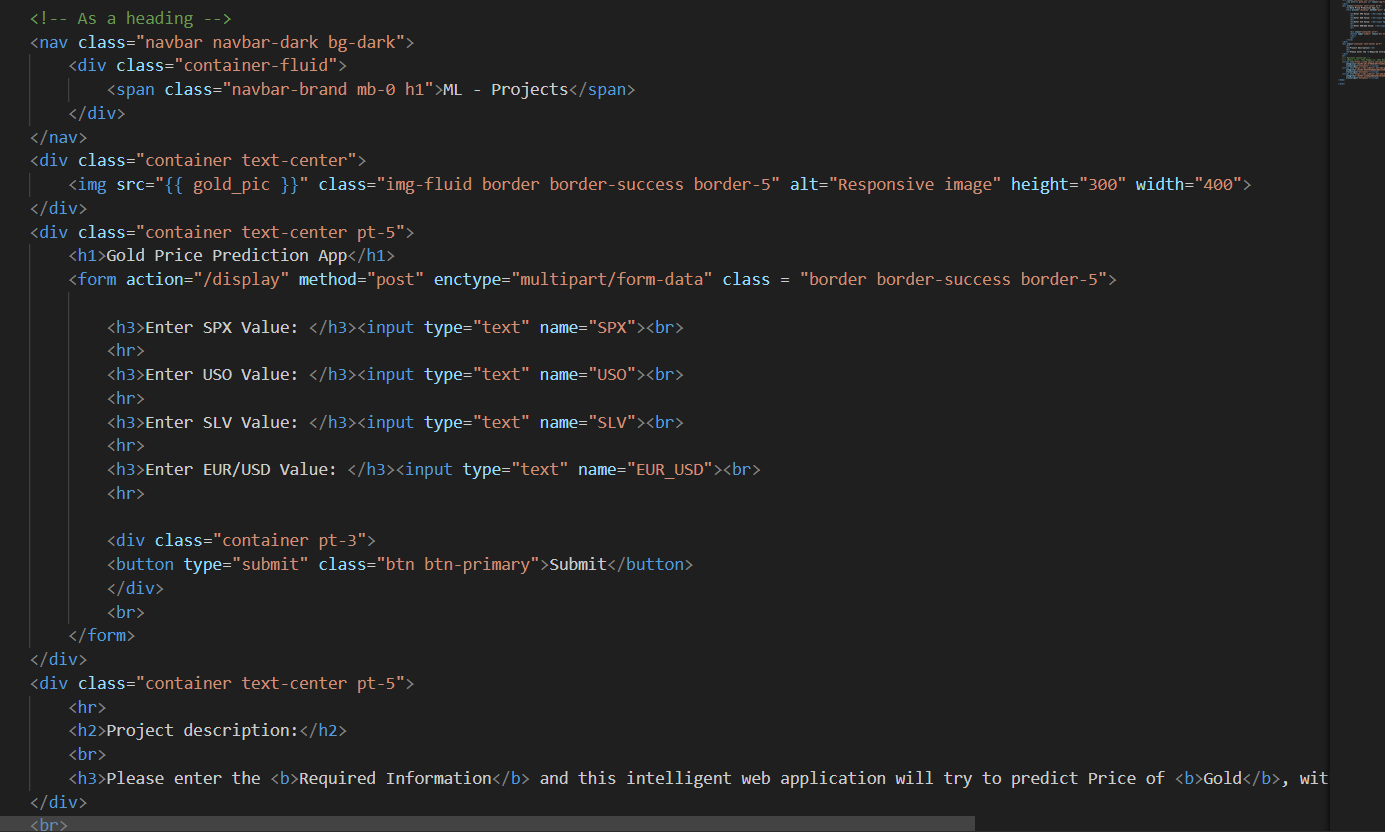


Fig5: Web page code

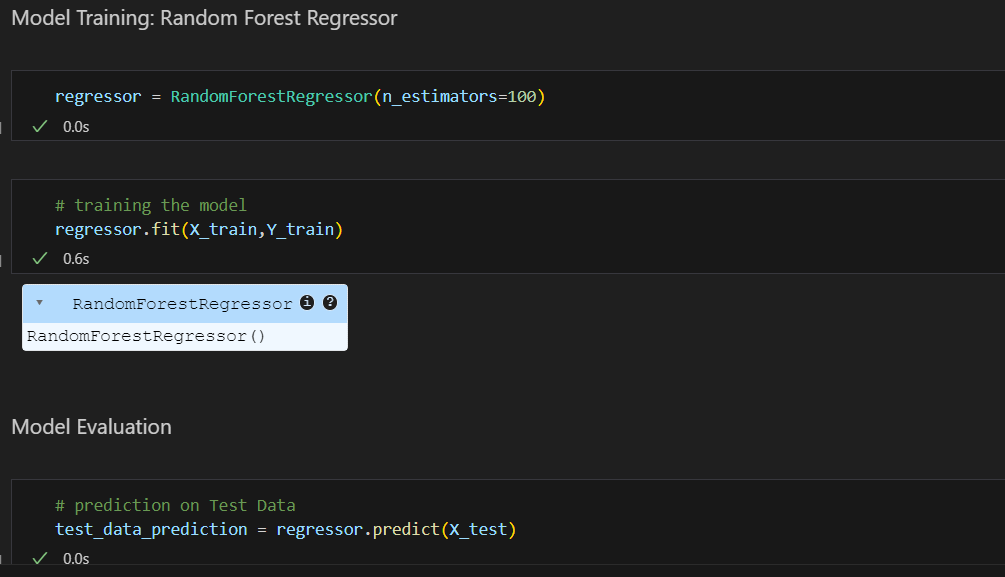


Fig6: Model Training through Random Forest Regression

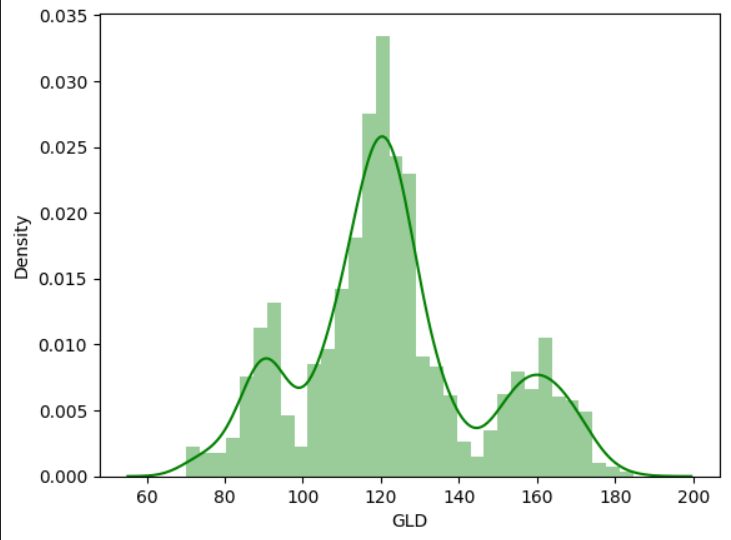


Fig7: Distribution of the GLD Price

**9. CORRELATION ANALYSIS**

275 Correlation is used to check the relationship among all the different variables that we have. For this paper, the correlation analysis has been shown.

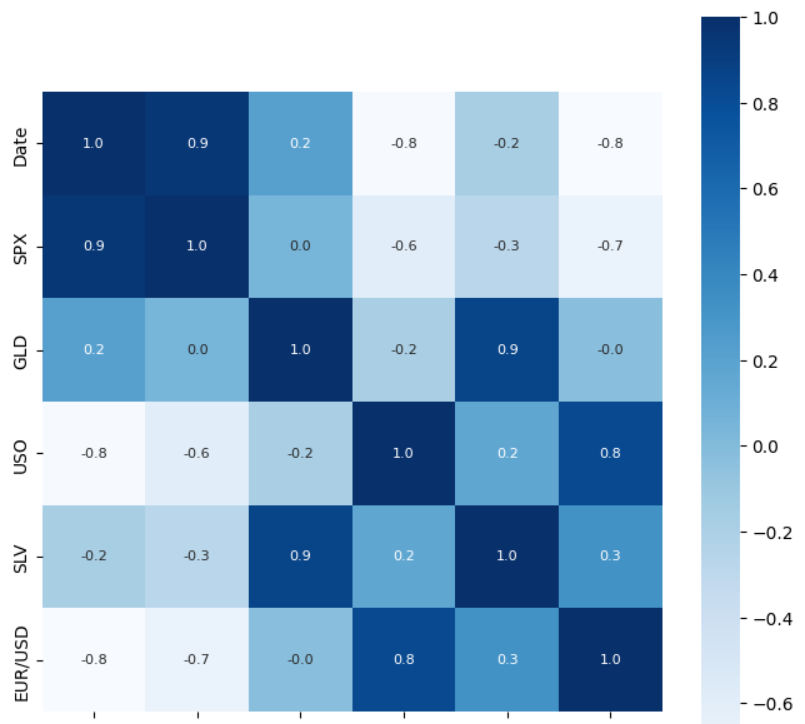


Fig8: Correlation Analysis of different attributes

**10. Web Page**

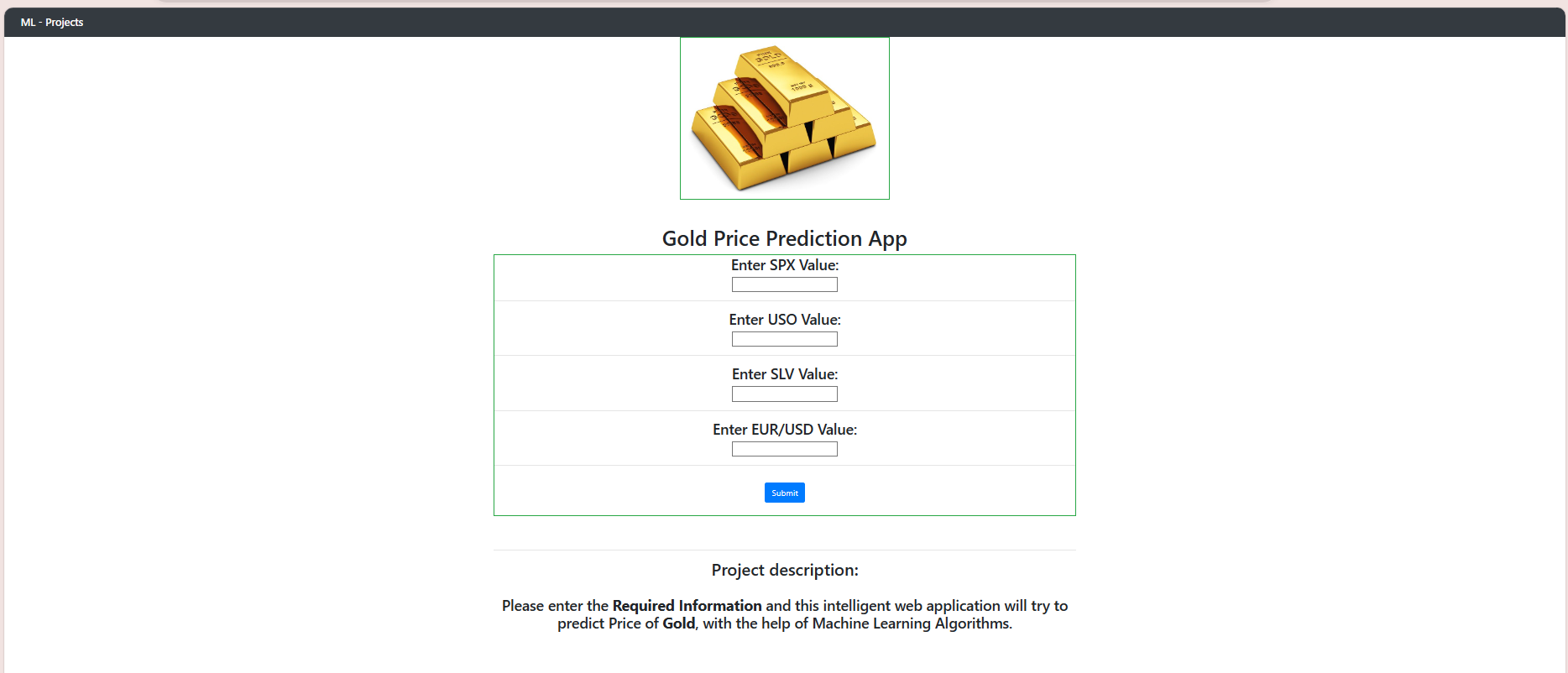


Fig9: Web page interface before prediction

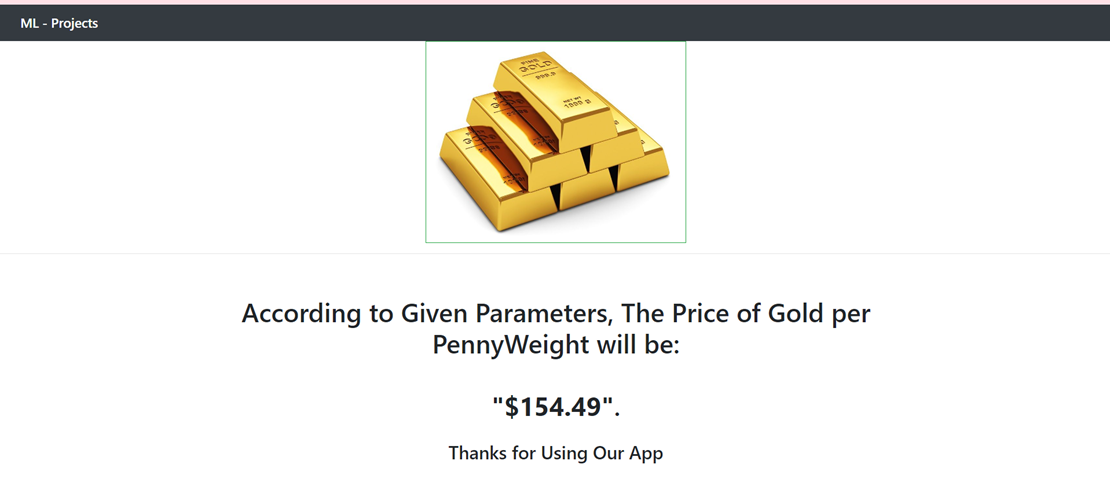


Fig10. Web page interface after prediction

**11. CONCLUSION**

Random Forest is a powerful algorithm that can handle a large number of features and nonlinear relationships between them. The model was able to capture the patterns in the historical data and make accurate predictions on the test set. The feature importance analysis indicated that the most important factors affecting the price of gold were the USD/INR exchange rate, the S&P 500 index, and the crude oil price.

The model's accuracy could be improved further by incorporating additional relevant features and data sources. Overall, the Random Forest algorithm showed promising results in predicting the price of gold and could be a useful tool for investors and analysts looking to make informed decisions in the gold market. However, as with any predictive model, it is important to note that the predictions are not guaranteed to be accurate and should be used in conjunction with other forms of analysis and market research.

**12.** **REFERENCES**

* [**https://www.geeksforgeeks.org/gold-price-prediction-using-machine-learning**](https://www.geeksforgeeks.org/gold-price-prediction-using-machine-learning)
* [**https://www.kaggle.com/code/sid321axn/gold-price-prediction-using-machine-learning**](https://www.kaggle.com/code/sid321axn/gold-price-prediction-using-machine-learning)
* [**https://www.analyticsvidhya.com/blog/2021/07/building-a-gold-price-prediction-model-using-machine-learning**](https://www.analyticsvidhya.com/blog/2021/07/building-a-gold-price-prediction-model-using-machine-learning)