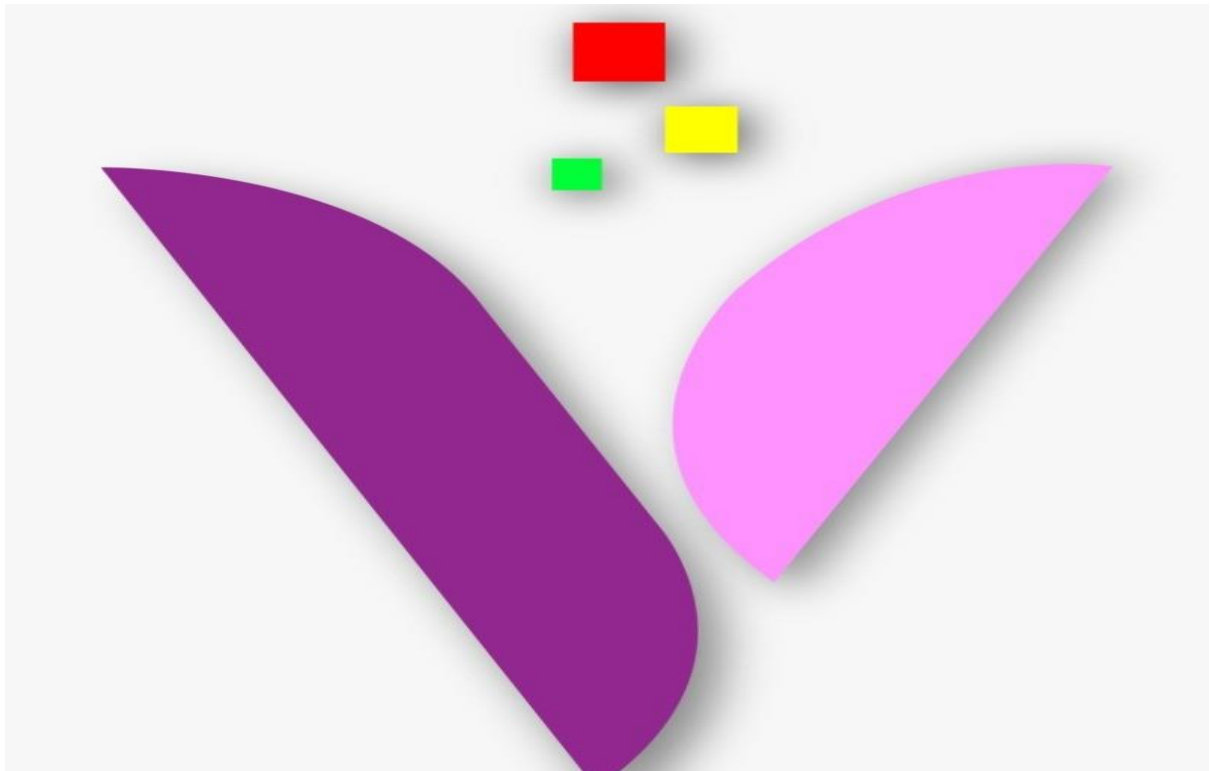# Evaluation and Prediction of Gold Prices

A project work made under the guidance of Vigor Council

Submitted by:

Saraansh Chikara

Acknowledgement

 I would like to express my deepest gratitude to several individuals who have contributed significantly to the completion of this project. First and foremost, I extend my sincere appreciation to my project supervisor, Dr. B.P Sharma. His guidance, unwavering support, and wealth of knowledge have been invaluable throughout the entire journey.

-Saraansh Chikara

Introduction

Most mined gold is stored as bullion. It is also, however, used extensively in jewellery, either in its pure form or as an alloy. The term 'carat' indicates the amount of gold present in an alloy. 24-carat is pure gold, but it is very soft. 18- and 9-carat gold alloys are commonly used because they are more durable.

The metal is also used for coinage, and has been used as standard for monetary systems in some countries.

Gold can be beaten into very thin sheets (gold leaf) to be used in art, for decoration and as architectural ornament. Electroplating can be used to cover another metal with a very thin layer of gold. This is used in gears for watches, artificial limb joints, cheap jewellery and electrical connectors. It is ideal for protecting electrical copper components because it conducts electricity well and does not corrode (which would break the contact). Thin gold wires are used inside computer chips to produce circuits.

Dentists sometimes use gold alloys in fillings, and a gold compound is used to treat some cases of arthritis.

Gold nanoparticles are increasingly being used as industrial catalysts. Vinyl acetate, which is used to make PVA (for glue, paint and resin), is made using a gold catalyst.

Gold is one of the few elements to occur in a natural state. It is found in veins and alluvial deposits. About 1500 tonnes of gold are mined each year. About two-thirds of this comes from South Africa and most of the rest from Russia.

Seawater contains about 4 grams of gold in 1,000,000 tonnes of water. Overall this is a huge amount of gold stored in the oceans but, because the concentration is so low, attempts to reclaim this gold have always failed.

➢ In this Project I have predicted the prices of gold using a model called Random Forest.

Random Forest
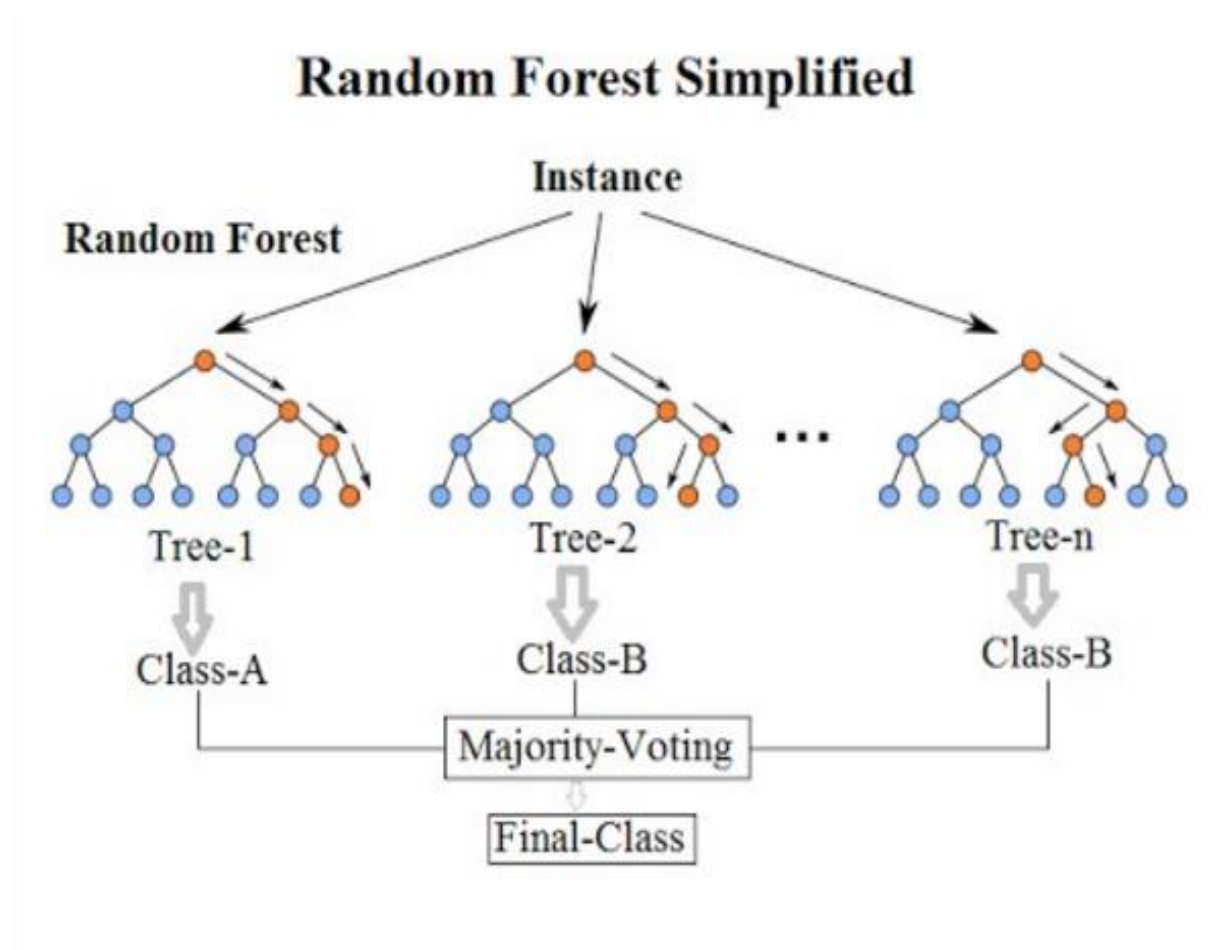
Random Forest is a popular ensemble learning method used in machine learning for both classification and regression tasks. It's based on the idea of building a multitude of decision trees during training and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual tree

Main Ideas:

1. Decision Trees: At the core of Random Forest are decision trees. Decision trees are hierarchical structures composed of nodes that represent decisions based on feature values, leading to outcomes. Each internal node represents a "test" on an attribute, each branch represents the outcome of the test, and each leaf node represents a class label or numerical value.

2. Ensemble Learning: Random Forest employs an ensemble of decision trees. Instead of relying on a single decision tree, it builds multiple decision trees during training. This ensemble approach helps to improve the predictive performance and robustness of the model.

3. Randomization: Randomness is introduced in two key ways:

   - Bootstrap Aggregating (Bagging): Random Forest builds each tree using a random subset of the training data with replacement. This process is known as bootstrapping. By using random subsets of data, it introduces diversity among the trees.
   - Feature Randomness: At each node of the decision tree, a random subset of features is considered for splitting. This helps to decorrelate the trees and make them less sensitive to noise in the data.

4. Voting or averaging: For classification tasks, Random Forest combines the predictions of individual trees by taking a majority vote (mode) of the class labels. For regression tasks, it averages the predictions made by individual trees.

Main Work/Contribution

The main task with this dataset was to use these features to build a predictive model that can accurately determine what was the predicted price for gold.We can solve this problem by using Random Forest algorithm

1)Dataset
Contents of the Dataset include:

- Date - date (mm/dd/yyyy format)
- SPX - stands for The Standard and Poor's 500 index, or simply the S&P 500. It is a stock market index used for tracking the stock performance of 500 of the largest companies listed on stock exchanges in USA
- GLD - gold price
- USO - stands for "The United States Oil Fund ® LP (USO)". It is an exchange-traded security whose shares may be purchased and sold on the NYSE Arca
- SLV - silver price
- EUR/USD - Euro to US dollar exchange ratio

2)Data Cleaning

Data cleaning is an important step in any data analysis process. It involves preparing the data for analysis by removing irrelevant data or modifying data that is incorrect, incomplete,

duplicated. In our dataset, the number of null values was 0 so there was no need for data cleaning.

3)Data Visualisation

Data visualization is the graphical representation of information and data. It uses visual elements like charts, graphs, and maps to provide an accessible way to see and understand trends, outliers, and patterns in data.

➢ Corelation Matrix
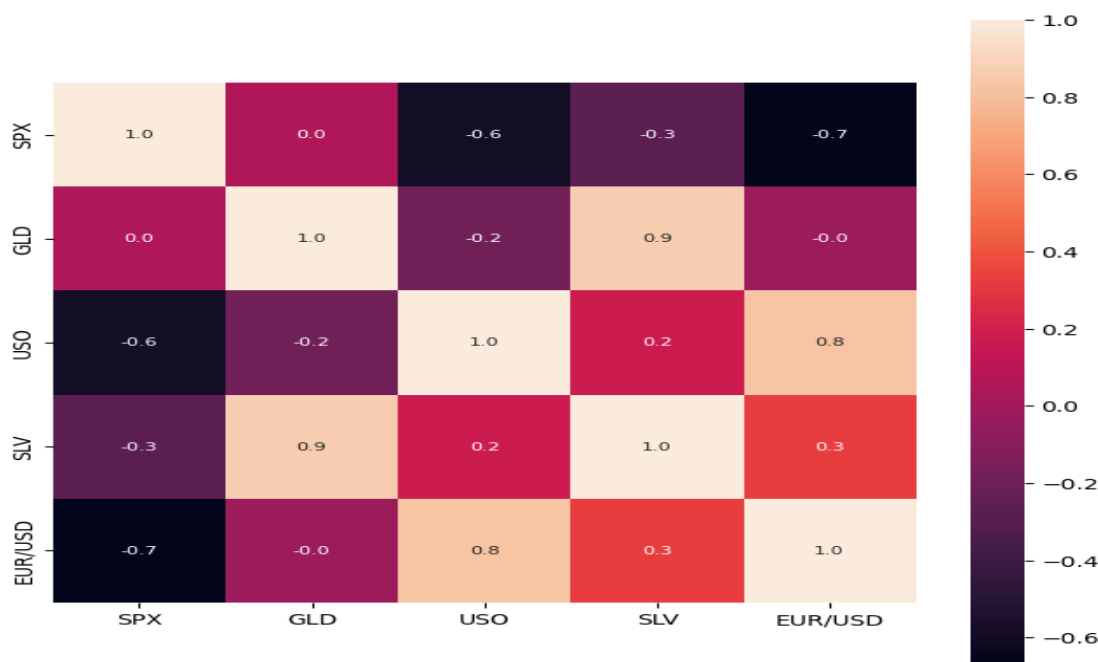
We observe that Gold Price and SLV have a high corelation

4)Data Modelling

The classification algorithm used is Random Forest. When we train the model, the dataset was split into a 80:20 ratio, allocating 80 percent of the data for training the model and reserving the remaining 20 percent for evaluating model accuracy

Random Forest

R squared error : 0.9900176651446791

Actual Price vs Predicted Price
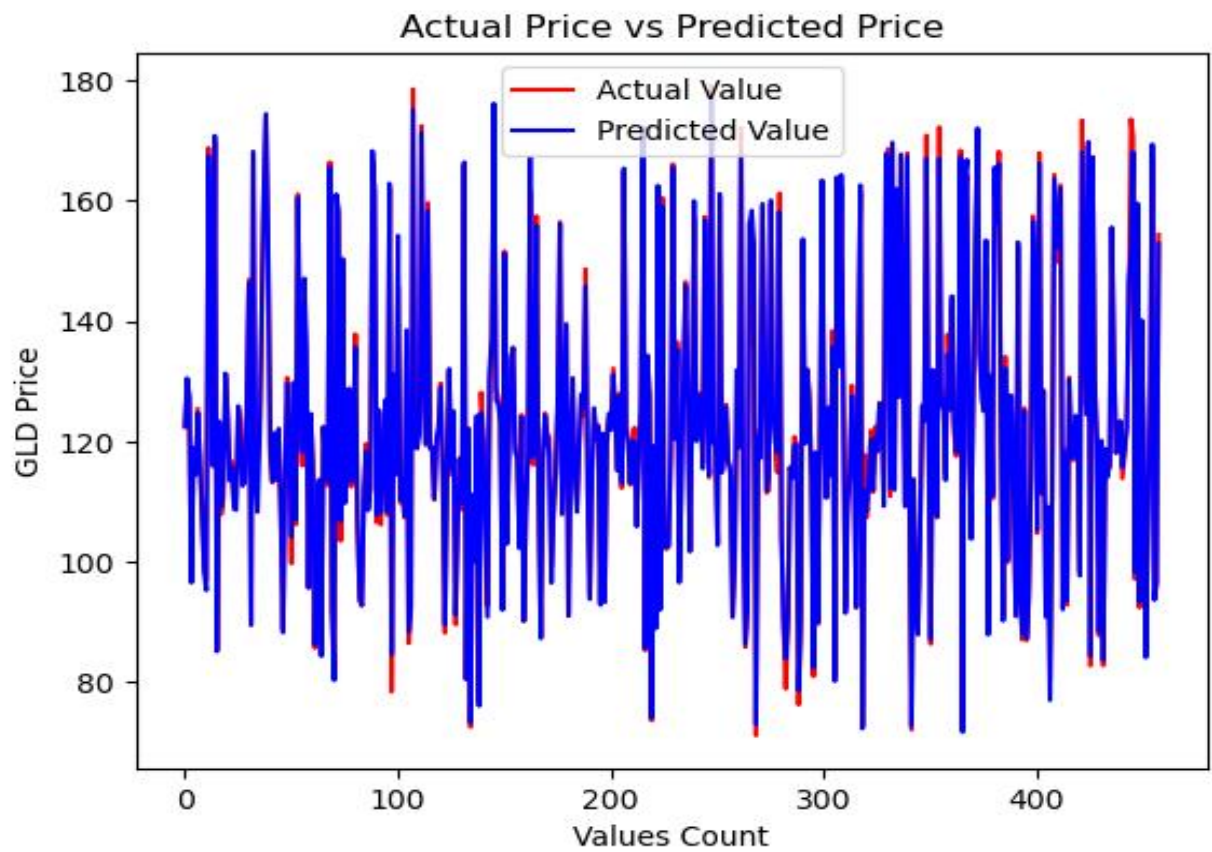
Conclusion

Based on the this dataset which includes the SPX index, gold (GLD), oil (USO), silver (SLV) prices, and the EUR/USD exchange rate from 01-02-2008 to 01-11-2008, a Random Forest model could be used to predict gold prices.

According to our models, Random Forest had R squared error as : 0.990017665144679 which is a very high accuracy for predicting the gold prices.

References

1) https://www.kaggle.com/

2) https://chatgpt.com/

3) https://www.bankbazaar.com/