



Gold Price Prediction Using Machine Learning Algorithm

List of Authors

1. Shubham Kumar Singh

Department of Computer Science & Application, School of Engineering & Technology, Sharda University, Greater Noida, India, www.shubhamsingh6666@gmail.com

2. Nidhi Gupta

Department of Computer Science & Application, School of Engineering & Technology, Sharda University, Greater Noida, India, ng20102000@gmail.com

3. Smriti Baliyan

Department of Computer Science & Application, School of Engineering & Technology, Sharda University, Greater Noida, India, smritichaudhary2019@gmail.com

4. Pardeep Kumar Mishra

Department of Computer Science & Application, School of Engineering & Technology, Sharda University, Greater Noida, India, pradeepkumar.mishra@sharda.ac.in

Abstract-

Since ancient times, gold has been cherished for its value and worth. Back then, gold was primarily used for trading purposes and as a method of remuneration. But now, it is looked upon as an investment and is found to exhibit the wealth of a country. Expensive metals like gold, in critical times, are used to assure the reimbursement of money borrowed as well. Thereby, gold is not only found to behold the rich, but also the poor. During pandemic crises, like the Covid-19, investments on gold in early times, might have a beneficial impact. Hence, predicting gold rates with live data and investing in gold at the right time is quite useful. The gold rates have been predicted using a variety of machine learning methods like linear regression, decision trees, and random forests. We've deployed a model using a random forest regressor.

Keywords: Gold rates, prediction, random forest, predictive analytics.

DOI Number: 10.48047/NQ.2022.20.20.NQ109297

NeuroQuantology2022;20(20):2998-3003

1. Introduction

The first well-known metal used by humans was gold. When we contemplate the development of iron and copper labour, we believe it to be the biggest contributor to the economic and cultural growth of our species. However, gold arrived earlier. Gold's value has been steady, and its value has been used to assess a country's financial soundness. This precious metal caught the attention of big investors, who invested a lot of money in it. Early on, more money was used to purchase this essential good. Like the price of other commodities, the price of gold is affected by supply and demand, including speculative

demand. However, in contrast to most other raw materials, disposal and savings have a greater impact on your pricing. This commodity was regarded as a secure investment by smaller investors as well. Your economic health, as measured by your interest rates and financial situation, heavily influences government investments in gold. Capital is flowing into the gold market as a result of activity being seen in the US. Numerous factors are connected to gold rates and have an impact on the price. Gold current prices are set twice daily based on the market's supply and demand. For these investors and the government banks, even a



little fluctuation in the price of gold may have a significant positive or negative impact.

It appears that gold's extraordinary utility and rarity as a commodity are what are causing its price to climb. There are also a number of additional causes for the price increase. In a variety of industries, including banking, trade, and mechanics, gold is used. About 14% of the gold required is used for industrial, dental, and medicinal uses. A great resistance to corrosion and bacterial colonisation, as well as excellent thermal and electrical conductivity, are all characteristics of gold. Due to the ongoing growth of developing market middle classes desiring western lifestyles, it has varied in recent years. To extract a tiny amount of gold, a higher amount of gold might be used along with a lot of employees involved with it. Only 5 grams of gold can be retrieved from a huge tonne of ore if the ore is of poorer grade. In reality, this metal is capable of becoming malleable, fluffy, and ductile. readily in many ways. This priceless substance is used for many different things all throughout the world since it becomes extremely malleable, allowing other metals to be combined to create useful decorations. It has a respectable cathode of heat and power. Additionally, it can withstand any environment and is unaffected by air, moisture, or other potentially harmful or destructive substances. The finest illustration of trustworthy materials is this. Additionally, a multi-billion dollar business has developed around the recycling of worn jewellery.

2. Literature Review:

In study [1], Using machine learning techniques, we forecast gold prices based on 22 market indicators. Results indicate that we are quite accurate at predicting daily gold rates. For investors and central banks to choose when to invest in this commodity, our prediction models will be helpful. In another study [2], Although several techniques are used in the literature to anticipate the price of gold, to our knowledge, this is the first research to combine regression models for the prediction of the XAU/USD index. Five different contributions are made by this study. The findings of the experiment show

that when used to estimate the price of the XAU/USD index, the stacking regression combination model produces significant results with a MAPE of 2.2036. In other study [3], This study compares Decision Tree (DT), Random Forest (RF), Gradient Boosted Trees (GBT), and ARIMA model to anticipate monthly gold prices using tree-based approaches for time series data. Based on RMSE=38.52, the findings showed that RF performed better in predicting future gold prices than DT, GBT, and ARIMA (0, 1, 1). In other study [4], To estimate highly skewed gold data in the vein-type at the Quartz Ridge region, this study employed five machine learning algorithms, including Gaussian Process Regression (GPR), Support Vector Regression (SVR), Decision Tree Ensemble (DTE), Fully Connected Neural Network (FCNN), and K-Nearest Neighbors (K-NN). In this study, more than only procedure is important for getting an accurate mineral estimation. It also pertains to how the data are separated and handled. In other study [5], This study forecasts the price direction of exchange-traded funds for gold and silver using a variety of machine learning tree-based classifiers, including bagging, stochastic gradient boosting, and random forests. The forecasts of the direction of the gold and silver prices made by decision tree bagging, stochastic gradient boosting, and random forests are substantially more precise than those made by logit models. In other study [6], The goal of the current study was to evaluate how well a hybrid model based on machine learning can forecast changes in the price of gold. A Particle Swarm Optimization (PSO)-based artificial neural network (ANN) model has been evaluated on the monthly gold price in India from January 2012 to June 2021. In other study [7], In this study, we attempt to forecast future gold prices using a well-known Deep Learning technique for time series analysis called Long Short-Term Memory (LSTM) networks. We discovered that the mean absolute percentage error (MAPE) value is 17.66144, while the root mean square error (RMSE) value for the prediction results is 39.94162. In other study [8], The methodology of the report makes use

of daily data from the World Gold Council between 1979 and 2019. The study's findings demonstrate that the SVM outperforms ARIMA when RMSE and MAPE are used as performance measuring methods, with an RMSE of 0.028 and MAPE of 2.5 for the SVM and 36.18 and 2897 for ARIMA, respectively.

3. Methodology

3.1. Gold Investment

Since the beginning of time, despite different economic and financial crises, the price of the precious metal has been high. The currency's value has changed over the past few years in response to the foreign exchange market, the cost of crude oil, and inflation. The price of gold fluctuates as well. The danger of investing in or purchasing them is higher for consumers or investors. Trading in gold is fairly simple since it is seen as being fundamental to society and a liquid asset everywhere. As a result, it is a valuable metal that is frequently utilised. The gold prediction is quite useful for making plans and carrying out future investment strategies. The goal of forecasting the gold price is to preserve financial assets in this situation, as gold's value is likely to see significant fluctuations. Units are exchanged similarly to stocks, and as stocks are non-linear, they are a difficult problem to anticipate. As a result, stockholders will focus their efforts on defending themselves against social catastrophe and monetary and political growth.

Because there is no "crop rotation volatility in the market," gold is advantageous. Consequently, they made investments in gold reserves in addition to those working for global organisations. The government may even expand the gold reserve in the meantime as this priceless metal has increasingly taken on the role of currency. Most people believe that one advantage of gold is that it provides support during difficult

In the present paper, The Consumer Price Index (CPI), the Federal Funds Rate (FFR), the Dollar Index U., and the U. were the six distinct variables used in the experiment. In this paper, Predict the gold price using random forest regression technique.

times due to its strict liquidity control. The price of gold varies daily, according to government policy and gold authorities. Finance experts may, however, plan when to acquire or offer these products by anticipating the monthly changes in gold prices. Gold is a different type of payment that is employed in more commercial transactions all over the planet. In the current context, all national central banks hold pricey metals to assure the repayment of foreign debt, control inflation, and serve as a barometer of financial health. The installation of ultrasonic aluminium wire connections requires alloy gold to have a more remarkable retention area than regular alloy metallization. Any prospective gold buyer must be aware of the elements influencing gold prices in order to better forecast interest rate trends and, as a result, steer their business toward greater profitability. With the purpose of conducting a thorough investigation and making an accurate forecast, investors and researchers are also very interested in knowing the gold business.

3.2. Dataset

Data for this study was gathered from Kagel, a live dataset that includes data for over 10 years prior to the analysis day in question. Data were obtained for features including date, SPX, GLD, USO, SLV and EUR/USD. The research also takes into account the price of every stock throughout this time. The below table shows the list of attributes of data which are extracted for the past 10 years in a manner of monthly.

	Date	SPX	GLD	USO	SLV	EUR/USD
0	1/2/2008	1447.160034	84.860001	78.470001	15.180	1.471692
1	1/3/2008	1447.160034	85.570000	78.370003	15.285	1.474491
2	1/4/2008	1411.630005	85.129997	77.309998	15.167	1.475492
3	1/7/2008	1416.180054	84.769997	75.500000	15.053	1.468299
4	1/8/2008	1390.189941	86.779999	76.059998	15.590	1.557099

Figure 1. Dataset of last five row

	Date	SPX	GLD	USO	SLV	EUR/USD
2285	5/8/2018	2671.919922	124.589996	14.0600	15.5100	1.186789
2286	5/9/2018	2697.790039	124.330002	14.3700	15.5300	1.184722
2287	5/10/2018	2723.070068	125.180000	14.4100	15.7400	1.191753
2288	5/14/2018	2730.129883	124.489998	14.3800	15.5600	1.193118
2289	5/16/2018	2725.780029	122.543800	14.4058	15.4542	1.182033

Figure 2. Dataset of first five row

The above listed dataset consists of 2290 rows and 6 columns. Where GLD columns are used to predict the price of gold. And the 4 column is used to show that also the attribute can affect the price of gold. Here, we do preprocessing and cleaning techniques to remove the unwanted attributes. Gold prices change day- to-day and the daily changes have been recorded in the analysis. The current price of gold is higher in comparison with the past year. As we see huge variations in price, we decided to split the dataset in a sequential split instead of random sampling. Therefore, the recent 30% data has been used as the test set, and the remaining 70% of data has been used for training. There is a major fluctuation in gold price over the past years, so recent historical data would be more indicative of the future trends.

3.3. Machine Learning Implementation

The prediction and accuracy will not change because we receive the input as tickle data, fit the models, and then deliver the predicted result as tickle.

3.3.1 Splitting of training and testing data.

To increase accuracy, the training set has been segmented into 70% and the testing set into 30% using the history of the dataset. The machine learning algorithms listed below are employed to forecast future values that are virtually precise.

3.3.2 Decision Tree.

Finding the values in the features that split the data into classes allows decision tree classifiers to create nodes that include a high percentage of data points from a specific or single class. It might be seen as a nonlinear model made up of several linear bounds. This model classifies the points based on features after receiving both labels and features. The overfitting has reduced precision.

3.3.3 Random forest model.

A bagging methodology called random forest constructs many decision trees and combines them to increase prediction accuracy. Although merely a bagging classifier, random forest has virtually identical hyperparameters to a decision tree classifier. In a random forest, as the trees grow, extra unpredictability is added to the model that we fit to the training set of data. In Accordance with our dataset, the random forest model gives the best accuracy. The Random Forest method may be used by the stock market to determine a stock's behaviour, and the prediction model can be used to estimate the predicted loss or profit.

4. Model

With the help of the random forest regressor, we could get almost the exact predicted value. In the below graph we can see, the actual value is in green colour and predicted value is in yellow colour, and the predicted value is very close to the actual value or in other words it's very close to each other.



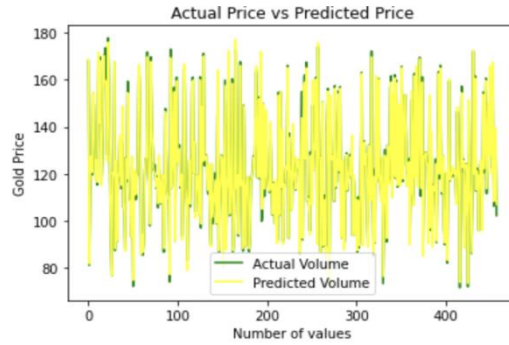


Figure 3. Actual Price vs Predicted Price

5. User Interface Creation

In order to make the analysis and prediction more interactive to the people, a user interface has been built. This user interface has been created with the help of html, css and javascript. We have used html for creating the web app. index.html have been used to show the index page and display.html have

been used to show the page along with the predicted value. And we have python to build the app to predict the price of the gold and we have also imported different python libraries to build the application. And we have also used flask to deploy our model and create a user interface also. And we have been deployed using the flask.

3002

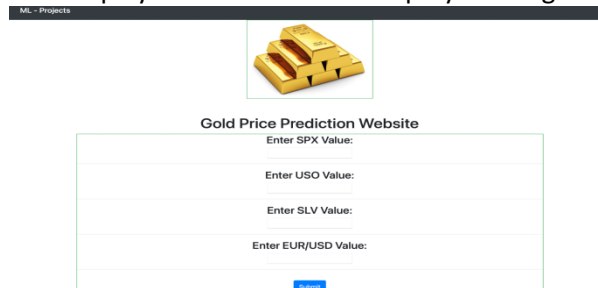


Fig 4. Index Page

6. Discussion

In the GUI below, We have entered some value from the dataset as a reference to predict the price of the gold. Where the values have been following, SPX value is

1447.160034, USO value is 78.470001, SLV value is 15.18 and EUR/USD value is 1.471692. We have taken this value as an input. And after clicking the submit button, we will get the result as an output.

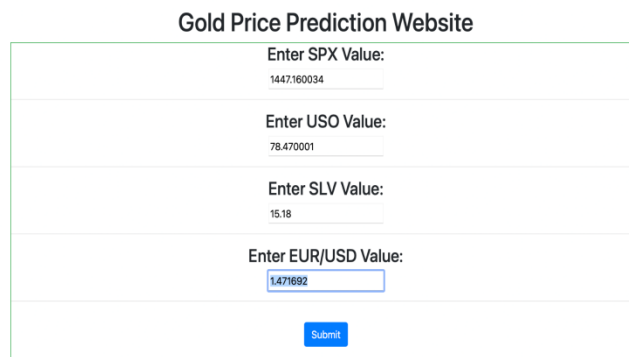


Fig 5. Input Value



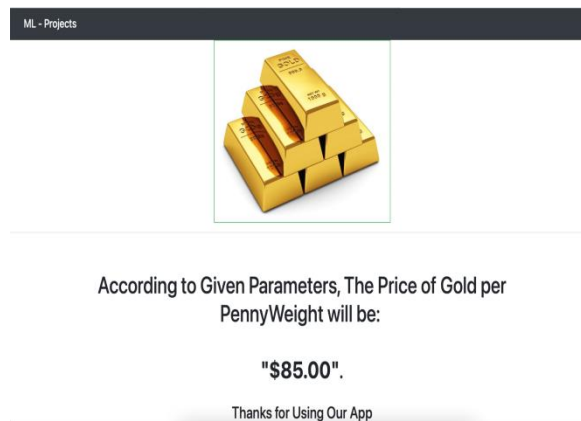


Fig 6. Predicted Output.

6. Conclusion

As gold proves to be a viable source of investment, the investment opportunities have been expanded to huge numbers, and there arose a need for predicting the future highs and lows that the commodity(gold) might hit. After learning different machine learning models, we have predicted that random forest regressor has a higher accuracy in forecasting future gold rates than linear regression, decision tree regression, and other machine learning prediction models. Deploying a model that might help those who are considering investing in gold could help more people globally

7. Future Enhancement

Our initial step will be to update our model, adding decorative share pricing, and build a User interface application (GUI). Where users will enter the month in respective date format to predict the price of the gold and finally the user will get the value of gold respectively. And our second step would be to build an application to make it more user pleasant. We also intend to release it as a play store app in order to reach a wider audience.

References

- [1] ul Sami, I. and Junejo, K.N., 2017. Predicting future gold rates using a machine learning approach. *International Journal of Advanced Computer Science and Applications*, 8(12).
- [2] Kilimci, Z.H., Ensemble Regression-Based Gold Price (XAU/USD) Prediction. *Journal of Emerging Computer Technologies*, 2(1), pp.7-12.
- [3] Rady, E.H.A., Fawzy, H. and Fattah, A.M.A., 2021. Time series forecasting using tree based

methods. *J. Stat. Appl. Probab*, 10, pp.229-244.

[4] Zaki, M.M., Chen, S., Zhang, J., Feng, F., Khoreshok, A.A., Mahdy, M.A. and Salim, K.M., 2022. A Novel Approach for Resource Estimation of Highly Skewed Gold Using Machine Learning Algorithms. *Minerals*, 12(7), p.900.

[5] Sadorsky, P., 2021. Predicting gold and silver price direction using tree-based classifiers. *Journal of Risk and Financial Management*, 14(5), p.198.

[6] Sadorsky, P., 2021. Predicting gold and silver price direction using tree-based classifiers. *Journal of Risk and Financial Management*, 14(5), p.198.

[7] P. K. Sarangi, R. Verma, S. Inder and N. Mittal, "Machine Learning Based Hybrid Model for Gold Price Prediction in India," *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, 2021, pp. 1-5, doi: 10.1109/ICRITO51393.2021.9596391

[8] Hansun, S. and Suryadibrata, A., 2021. Gold price prediction in COVID-19 era. *Int J Comput Intell Control*, 13(2), p.1.

[9] Makala, D. and 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.

[10] Manoj, J. and Suresh, K.K., 2019. Forecast model for price of gold: Multiple linear regression with principal component analysis. *Thailand Statistician*, 17(1), pp.125-131.