



GOLD PRICE PREDICTION USING MACHINE LEARNING TECHNIQUE

¹Dr. K. Maheswari , ²Rajesh Kumar N

^{1,2}Assistant professor,

^{1,2} P.G. Student Department of Computer Science

^{1,2} K.R.College of arts and Science, Kovilpatti

Abstract

This paper predicts the gold EFT price based on the previous year's gold price data. Inventory forecasting plays a crucial role in the financial success of the business. Buying Gold ETFs implies buying gold in an electronic form. Buying and selling gold ETFs are just like to trading stocks. The claim of Gold ETF will not get actual gold but receive the cash equivalent. Machine learning technique is applied to predict financial variables and focused on predicting the gold price ETF using a linear regression algorithm. The dataset used is a numerical dataset. The main goal of this project is to forecast the rise and fall in the daily gold rates that can help investors to decide when to buy or sell the gold.

Keywords: Prediction, Machine learning , Regression and Gold ETF

1. Introduction

Machine learning involves programming computers to optimize performance criteria using sample data or past experience. We have a model defined except for some parameters and learning is the execution of a computer program to optimize the model's parameters using training data or past experience . Models can be predictive to make future predictions or descriptive to gain knowledge from data. But can a machine also learn from past experience or data like humans. The field of study known as machine learning deals with the question of how to build computer programs that automatically improve with experience.

Machine learning allows machines to automatically learn from data, improve performance from experiments, and predict things without explicit programming [8]. Machine learning gathers computations and statistics to create predictive models. Let's say we have a complex problem where we need to make a prediction, so instead of writing code for it we just need to feed data to general algorithms and with the help of algorithms [9]. In this calculation, the machine will build logic according to the data and predict the output. Machine learning has changed the way we think about problems. The block diagram below explains how the machine learning algorithm works. Gold is used to support commercial transactions around the world in addition to other payment methods. Many states have maintained and improved their gold reserves and are recognized as wealthy and progressive states. This project will benefit investors and will control the banks that decide when to invest in this commodity. Here, the commodity is called gold. Many multinational corporations and individuals have also invested in gold reserves.

2. Literature Review

Abhay Kumar Agarwal, [1] have proposed an Autoregressive Moving Average (ARMA) model, which is a statistical model that is frequently used for analyzing time series data. The model achieved an accuracy of 66.67%.

Manjula K. A, [2] have proposed an Artificial Neural Networks (ANN) for predicting the gold prices. The data for this study is collected from various sources over a period of eleven years. This data included variables such as oil price, S&P 500 index, USD exchange rates and other economic variables.

M.Sravani, [3] has done the work as the data collected from World Gold Council containing the daily prices of gold from January 1979 to December 2019. This study uses Autoregressive Integrated Moving Average (ARIMA) method and SVM to forecast the gold price. The accuracy achieved by ARIMA model is less than the accuracy achieved by SVM.

Navin, [4] has done the work as the decision tree and support vector regression algorithms for predicting and forecasting gold prices using the data containing five attributes namely open, close, low, high values and volume of gold. In this study, the decision tree suited best for feature selection whereas the SVR suited best for large training data.

P.Patil, [5] has done the work to used algorithms such as SVM, Logistic regression for prediction. The Logistic regression model achieved an accuracy of 61.92% outperforming the SVM model.

V.Preethi, [6] has done the work is intended to figure the gold value utilizing ARIMA model. For gauging it utilizes memorable information.

M. Krishna Reddy, [7] has done the work is depends on an investigation led to comprehend the connection between gold cost and chose

factors affecting it, to be specific financial exchange, unrefined petroleum value, rupee dollar conversion scale, swelling and loan cost.

3. METHODOLOGY

Data for this study are collected from source y-finance from January 2005 through current date. The sample dataset attributes are shown in table 1. GLD is the largest ETF with direct investments in actual gold. Linear regression in machine learning helps you find out patterns and relationships in data and make an educated decision or prediction. It is one of the most well-known and well understood algorithms in statistics and machine learning.

Table 1. Gold ETF Dataset Attributes

SI No	ATTRIBUTE NAME	DESCRIPTION	TYPE
1	Open	Opening price of the gold	Numerical
2	High	Higher range of the gold	Numerical
3	Low	Lower range of the gold	Numerical
4	Close	Closing price of the gold	Numerical
5	Volume	Total volume of the gold	Numerical

Regression is a method of modeling a target value based on independent predictions. This method is used for forecasting and finding out the cause and efficient relationship between variables. The flow of work is shown in figure1.

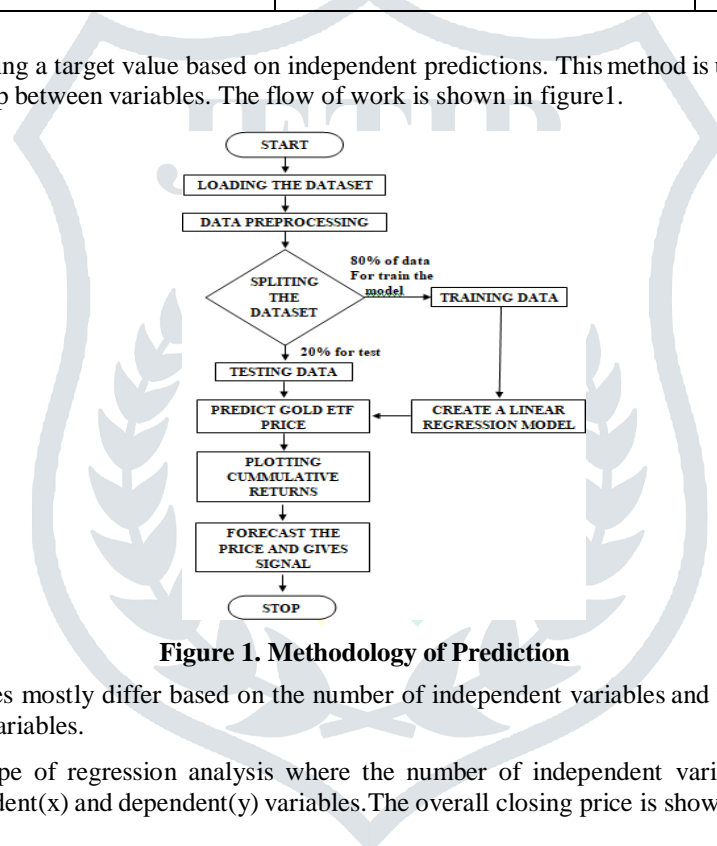


Figure 1. Methodology of Prediction

Usually, the regression techniques mostly differ based on the number of independent variables and the types of relationships between the independent and dependent variables.

Simple linear regression is a type of regression analysis where the number of independent variables is one and there is a linear relationship between the independent(x) and dependent(y) variables. The overall closing price is shown in figure 2.

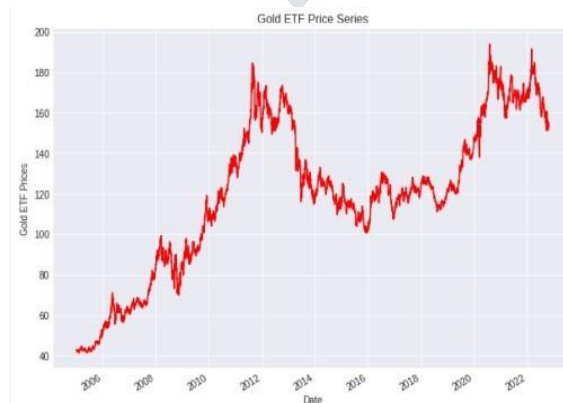


Figure 2. Gold ETF closing price

The red lines in the above graph are referred to as the best fit straight line. Based on the given data points, we try to plot a line that models the points the best. The line can be modeled based on the linear regression equation which is

$$y = a_0 + a_1 * x$$

1. Gathering the data from y-finances library and preparing the data by removing the missing values.

2. Now we split the gathered data into training and testing dataset.
3. Now using training data we create a linear regression model.
4. Using the testing data we test the created linear regression model.
5. Using the model now we predict daily gold ETF price.

The R2 rating is used as assessment metrics. R-squared is a statistical degree of the way near the information are to the geared up regression line. It is likewise called the coefficient of dedication, or the coefficient of more than one dedication for more than one regression. R-squared is usually among zero and 100%. zero% suggests that the version explains none of the variety of the reaction information round its mean. 100% suggests that the version explains all of the variability of the reaction information round its mean.

Plotting overall performance of Gold ETF closing from 2005-01-03 to upto date

In this work, closing price is set to be a target variable. To predict the future closing price by using the previous day price to know whether buy/sell for the next day. So that the overall closing price of Gold ETF is plotted from 2005-01-03 to upto date.

The linear regression model can predict the closing price of the testing data. The below graph show the comparison of the actual price and predicted price using linear regression model is shown in figure 3.

Gold ETF Price (y) = 1.33 * 5 Days Exponential Moving Average (x1) + -0.33 * 12 Days Exponential Moving Average (x2) + 0.32(constant).



Figure 3. Expected and Actual Gold ETF

To evaluate its results, measure the cumulative profit of this technique which is shown in figure 4. The following steps are taken to calculate the total profit.

- Generate regular percentage change of gold price.
- Construct trade symbol "1", if the expected price of the next day is higher than the expected price of the current day.
- Calculate the profit of the strategy using trading signals to multiply the daily percentage change.

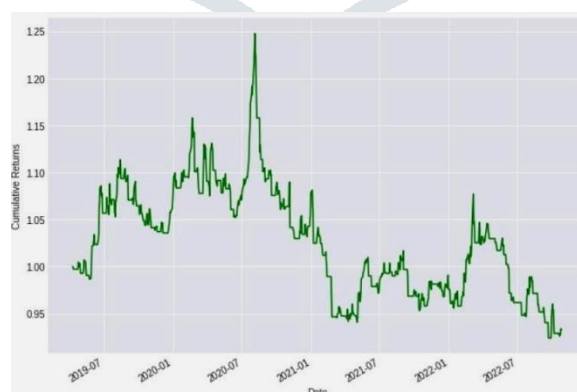


Figure 4 . Cumulative returns of Gold ETF

4. RESULT AND ANALYSIS

A Gold ETF is an exchange-traded fund (ETF) to reveal the residence rate of the real gold. These are defensive investing vehicles focused on gold, so they invest in gold bullion.

In general, Gold ETFs are units of actual gold and can be in paper or dematerialized shape. Each unit of Gold ETF is equivalent to 1 gram of gold, which is supported by actual gold of extremely high purity. Gold ETFs integrate simplicity with equity ownership with the facility to trade in cash.

Let find the market movement by using the two regression line. The slow regression line is the period of 12 days and the fast regression line is the period of 5 days. These two regression lines are only used to identify the short term trend changes.

If the actual price is breaks upwards then the trend is indicated to uptrend. If the actual price is breaks downwards then the trend is indicated to downtrend.

After these crossover the price movement is predicted to be on the same trend while the next crossovers are made. The fast regression line is more reactive than the slow regression for market movements.

After finding the trend, the next step is to predict the daily closing price. So that we additionally using the two regression lines in which the period of 8 & 9. Then generate the signals as buy/sell for future price movement and return the status of the signal is profit/loss. Here we show GOLD ETF price for future which are predicted using the linear regression model is shown in figure 5.



Figure 5 . Gold ETF regression line Crossover

The future prediction of gold ETF per day is shown in figure 6.

Date	2022-10-17	2022-10-18	2022-10-19	2022-10-20	2022-10-21	2022-10-24	2022-10-25	2022-10-26	2022-10-27	2022-10-28
Pair	Gold ETF	Gold ETF	Gold ETF	Gold ETF	Gold ETF	Gold ETF	Gold ETF	Gold ETF	Gold ETF	Gold ETF
Heiken_Close	154.247501	153.027503	151.952499	151.967499	153.067497	153.554996	154.145	155.067497	154.690002	153.040001
Buy	False	False	False	False	True	True	False	False	False	False
Sell	False	False	False	False	False	False	False	False	False	False
Future_Close_Price	153.027503	151.952499	151.967499	153.067497	153.554996	154.145	155.067497	154.690002	153.040001	NaN
B_Status (PL)	False	False	False	False	True	True	False	False	False	False
S_Status (PL)	False	False	False	False	False	False	False	False	False	False

Figure 6. Predicting future Gold ETF closing price

Finally the signals are calculated and returns the accuracy of the linear regression model. This model is improved to attain the accuracy percentage of 83 for predicting the future closing price based on previous day closing price of Gold ETF. The trading results are shown in figure 7.

Gold ETF Trading Results
 Statistical Report
 Total Signals : 18
 Profit Signals : 15
 Loss Signals : 3
 Accuracy : 83 %

Figure 7. Accuracy of future Gold ETF

5. CONCLUSION

Gold has been one of history's most significant commodities. Maintaining central banks' gold reserves is essential to maintaining the world's existing economic system. Some big firms and investors are now spending large amounts of money in gold. While forecasting the rate of gold is not very easy, it will allow investors and central banks to determine better when to sell and buy them and thus maximize their income.

Furthermore, an attempt has been made in this study by using machine learning algorithms to accurately predict the gold prices and when to sell them and purchase them. This research was done in order to clarify the gold ETF price predictions using machine learning using Python. The research was carried out for data between January 2005 and upto date. At first I will create a machine learning linear regression model. Then train this machine learning algorithm by giving information from past gold ETF prices. After this trained model is used for future prediction. Finally concluded that machine learning algorithm with linear regression analysis are very useful in gold price prediction and the model's accuracy is 97.77 percent.

There are still many technical indicators and feature variables which we have not included in our project, may be there are some other indicators which we haven't explored would perform better. There are lots of Machine Learning algorithms which we haven't tried and may be neural network or LSTM would perform better than our solution. For this research work, 17 years of Gold ETF data is used.

If we would increase the data, the performance of our solution models may be improved. In future this work can be continue with neural network to find the better predictive rate and automation of trade taken with better entry and exit point in a real platform.

REFERENCES

1. Dr. Abhay Kumar Agarwal, Swati Kumari, "Gold Price Prediction using Machine Learning", International Journal of Trend in Scientific Research and Development (ijtsrd), 2020.
2. Manjula K. A., Karthikeyan P, "Gold Price Prediction using Ensemble based Machine Learning Techniques", Third International Conference on Trends in Electronics and Informatics, 2019.
3. Yakubu S. Baguda† and Hani Meateg Al- Jahdali, "An Intelligent Gold Price Prediction Based on Automated Machine and k-fold Cross Validation Learning", IJCSNS International Journal of Computer Science and Network Security, VOL.21 No.4, April 2021
4. B. Prajna*5, "GOLD PRICE PREDICTION USING MACHINE LEARNING", International Research Journal of Modernization in Engineering Technology and Science, Volume:04/Issue:05/May-2022
5. M.Sravani, Ch.Abhilash, T.Divya, Ch.Vasthav, D.Priyanka, "GOLD PRICE PREDICTION", INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS, Volume 9/issue 6 June 2021
6. Musli Yanto , Sigit Sanjaya , Yulasmai , Dodi Guswandi , Syafri Arlis, "Implementation multiple linear regresion in neural network predict gold price", Indonesian Journal of Electrical Engineering and Computer Science, Vol. 22, No. 3, June 2021
7. Perry Sadorsky, "Predicting Gold and Silver Price Direction Using Tree-Based Classifiers", journal of risk and financial management, 29 April 2021
8. Hansun, S. and Suryadibrata, A. (2021). Gold Price Prediction in COVID-19 Era. International Journal of Computational Intelligence in Control12(2), pp.1-4
9. P. Sai Shankar, M. Krishna Reddy, "Forecasting Gold Prices in India using Time series and Deep Learning Algorithms", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249- 8958 (Online), Volume-10 Issue-5, June 2021.

