



Gold Price Forecasting In India Using Arima Model On Cloud Platform

¹Sachin Kumar, ²Saurabh Pal, ³Sarika Srivastava, ⁴Vijendra Pratap Singh, ⁵Priya Jaiswal

Research Scholar, Professor, Professor, Assistant Professor, Research Scholar

Department of Computer Applications,

VBS Purvanchal University

Jaunpur, Uttar Pradesh, INDIA

Abstract: Gold is the most popular investment in the world because it has been shown to be the most effective safe haven in a lot of countries. It's delicate to use a style similar to specialized analysis to prognosticate the gold's value. Numerous vaticination problems that contain a time element bear time series soothsaying, which is an important content of machine literacy. This is a study of the gold rate that will prognosticate the gold price by using one of the time series styles, which is the Autoregressive Integrated Moving Average (ARIMA). The main point of the system is to prognosticate the gold price and fantasize about the prognosticated value in a line map. This will help druggies to know the unborn trend and can help them to make a decision on the right time to buy or vend the gold. The system will give other features to help druggies in gold investment, such as getting the latest gold news, a calculator of gold investment, and gold branch locales around India.

Index Terms: ARIMA, Gold, Vaticination

I. Introduction:

Gold is a kind of essence that can be veritably sensitive to the price change (1). When studying changes in world finance, changes in gold prices are a good form of expression. Other means like equities and currencies, in Corti and Holliday's view, are constantly identified with gold prices (2). Also, some indicators, similar to the bone indicator and the Shanghai Composite Index, also have the eventuality to impact gold prices (3). Investors should always grasp the changes in gold prices and make reasonable prognostications to reduce pitfalls (4). There have been several explorations on gold prices. Megan Potoski has tried to prognosticate how the current day's price fix affects the London PM price fix of gold the following day with machine literacy models(5).

This study is trying to figure out whether the ARIMA, Decision tree, and MLR model is suitable to prognosticate AU99.99's close price and which one has the most stylish performance. For ARIMA Model, Dr.M. Massarrat Ali Khan demonstrated how to prognosticate the price of gold with the traditional Box Jenkins approach (6). Meyler and his platoon developed a semi-automated algorithm to fit an ARMA model to stabilize time series data. By introducing the new algorithm, they prognosticated the Irish affectation rate (7). According to Tripathy, ARIMA () can be the most accurate model to read gold prices in India (8). Nyoni also used the ARIMA approach grounded on minimal AIC to study affectation in Kenya between 1960 and 2017(9). For the Decision tree Model, Rady and his platoon demonstrated how the tree-grounded model works in time series soothsaying (10). There have been numerous studies to prognosticate precious essence price movements in this way. Navin and Vadivu introduced the perpetration of using Decision tree Retrogression to prognosticate gold price (11). For Multi-Linear Regression Model, a multiple direct retrogression model was established for the stocks of China Citic Bank by Chen and the opening price of the

stock was prognosticated successfully(12). Shokri and his platoon estimated and read the global price of tableware using a combined multiple direct retrogression (MLR) which provides the alleviation to use MLR to prognosticate gold price (13).Section 2 introduces data sources and model principles. Section 3 is the analysis of the three models' performance grounded on R- sq, and RMSE. The section is the conclusion of the most suitable model.

II. Proposed Model:

To prognosticate time series data, the bus-accumulative integrated moving average model (ARIMA) is constantly employed. The advantage of ARIMA model vaticination is that it represents colorful types of time series, including autoregressive (AR), moving average (MA), and combinations of AR and MA (ARMA). The ARIMA model is represented by ARIMA (p, d, q), where " p " denotes the Autoregressive process, " d " denotes in which the order that the data are stationary, and " q " denotes the order in which the moving normal is applied. The ARIMA model can be epitomized in the following formula

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \dots - \theta_q e_t \quad (1)$$

Where "yt" stands for the factual value and "et" stands for the arbitrary error of time period "t"; the model parameters are explained by ϕ_i and θ_j . Integer's p and q indicate the orders of the model. In the event that either q or p is both 0 also the model changes to an AR or MA model, independently. Identification, parameter estimation, and individual testing are the three processes that are always involved in developing an ARIMA model.

III. Decision tree Model:

A Decision tree is considered a form of visualization containing two types of bumps root bumps and splint bumps. The splint bumps contain the results. For prognosticating gold prices, typically two types of Decision trees are used in studies the bracket tree and the retrogression tree. When the order to which the data belongs is the prognosticatedresult, also it'll be regarded as a bracket tree analysis. When the true value or number is prognosticated, also it'll be regarded as a retrogression tree analysis. The Decision Tree is always a way to dig out the anticipated data, and also can be applied to checking possible trends among colourful branches. While using the decision tree analysis, reducing the community, and opting for the features can be answered by presenting all the cases as trait values automatically (11).

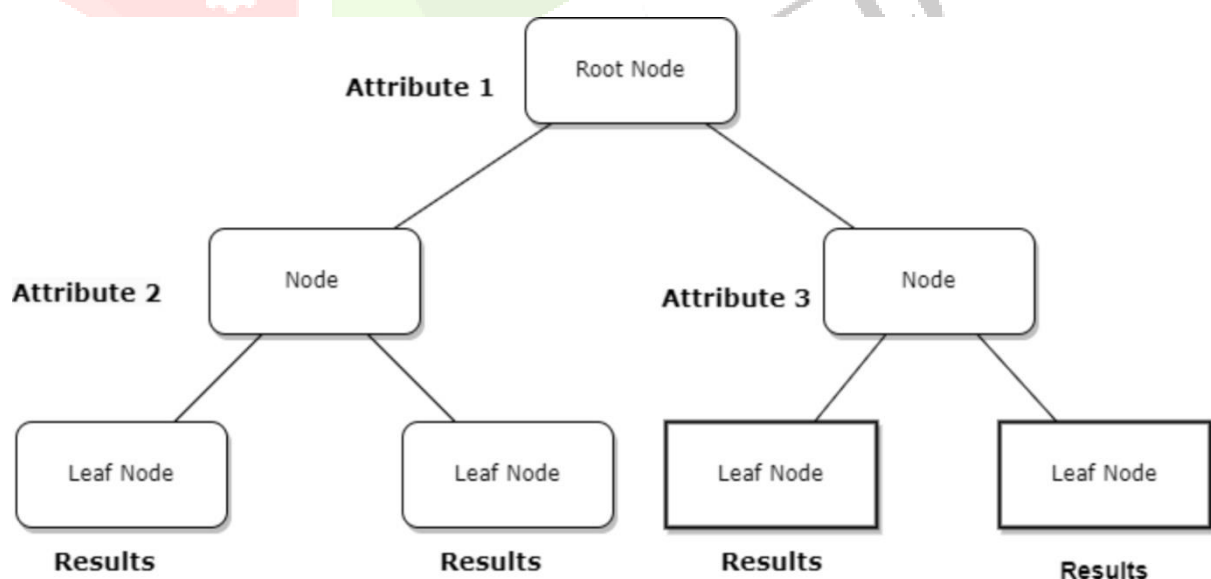


Fig 1: architecture of decision tree

In this study, the close price of AU99.99 for the following day is prognosticated using decision tree retrogression. The specific experimental design is as follows:

1. Use the former day's price to prognosticate the coming day's price
2. Conducting a5-fold time series cross-validation with the training set to get average RMSE and R-sq.
3. The model delicacy was estimated by RMSE and R- sq.

Table 1 lists all models of AIC and BIC and their values. The lowest AIC ARIMA model (2, 1, and 2) was selected.

Table 1. Different ARIMA models with their AIC and BIC value

MODEL	AIC	BIC	MODEL	AIC	BIC
ARIMA(0,1,0)	12560.02	12565.87	ARIMA(2,1,3)	12544.34	12579.42
ARIMA(0,1,1)	12559.47	12571.16	ARIMA(2,1,4)	12545.00	12585.93
ARIMA(0,1,2)	12558.50	12576.04	ARIMA(3,1,0)	12560.63	12584.01
ARIMA(0,1,3)	12560.50	12583.88	ARIMA(3,1,1)	12561.38	12590.61
ARIMA(0,1,4)	12560.48	12589.72	ARIMA(3,1,2)	12543.33	12578.41
ARIMA(1,1,0)	12559.30	12570.99	ARIMA(3,1,3)	12546.64	12587.56
ARIMA(1,1,1)	12560.08	12577.62	ARIMA(3,1,4)	12547.67	12594.45
ARIMA(1,1,2)	12560.50	12583.89	ARIMA(4,1,0)	12560.56	12589.79
ARIMA(1,1,3)	12561.33	12590.57	ARIMA(4,1,4)	12562.56	12597.64
ARIMA(1,1,4)	12562.42	12597.50	ARIMA(4,1,2)	12545.45	12586.38
ARIMA(2,1,0)	12558.71	12576.25	ARIMA(4,1,3)	12547.79	12594.57
ARIMA(2,1,1)	12560.68	12584.07	ARIMA(4,1,4)	12549.55	12602.17
ARIMA(2,1,2)	12542.49	12571.73			

IV. Data Collection:

For this design, it used secondary data that was published by the World Gold Council (WGC), which is the request development association for the gold 50 assiduity. The data is known as quantitative data because it's recorded as numerical. It handed numerous data from different ages of time but this design only focuses on the daily and yearly timeframe. The data handed since the time 1978 have different currencies. But not all data is used and it needs to go through the data-processing phases. This design doesn't use data since the time 1978 because the revolution of the internet still not be. For a daily timeframe, the gold price data collected from 1st January 2010 until 25th December 2020 contains about 575 rows of data. While yearly, the gold price data were collected from the time 1st of January 2000 until the 1st of December 2020 containing about 253 rows of data. It only has two columns which are the date and value of the gold. Indeed though the data only have two variables, it's enough to make a vaticination in time series. It's called univariate time series vaticinating because the model only used former values of the time series to prognosticate its future values.

V. Discussion:

Through the analysis and comparison of the results of the three models, the most suitable model for prognosticating the close price of AU99.99 is the MLR model. Although ARIMA fits the literal data impeccably, it isn't suitable to read the accurate price. The experimental results show that ARIMA isn't ideal in prognosticating the ending price. The reasons are (1) the upward or downcast trend of the data isn't egregious. (2) ARIMA can only prognosticate the direct retrogression problem, so further variables need to be added. The Decision tree model shows the smallest RMSE in all the models, which means the value it could prognosticate, is accurate. Since there are several gaps between the vaticination line and the original data line, which implies that it's less sensitive to oscillations, it's still parlous for making an investment decision with this model. The MLR is from now on the stylish model and it can be used in practical situations. With an average R- sq of 0.97, it can prove that 97 of the test data can be explained by the model which is sufficient to show the vaticination delicacy and stability of the model.

VI. CONCLUSIONS:

Main Gold is a system that provides a vaticination model of time series which is an Autoregressive Integrated Moving Average (ARIMA). It predicts the value of gold for two timeframes which are daily timeframe and yearly timeframe. The prognosticated value is listed in a table and imaged in a line map to help investors to understand the unborn trend of gold. This process of vaticination was done by admin because it's nonstop data and needs to be brought if want to make a new vaticination. There are also other features handed to druggies similar as gold news. Admin will continue streamlining the rearmost news about gold to make sure druggies get the rearmost information about gold.

The system also handed a calculator to help investors to calculate their gold investment value. Druggies can choose their gold chastity and fit the quantum of gold in order to know the current investment. The computation is also shown to make druggies understand how the computation of the investment is done. Incipiently, it also handed a point like a list of gold branches around Malaysia. All countries are listed except Putrajaya. This list will help druggies to find the nearest position of gold branch 82 with help of a google chart position. The system also helps the admin by having a point for data visualization of each model with delicacy. Admin also can manage the gold news either to add new information or update or cancel the news.

References:

1. Ranson, D., Wainright, H.C. (2005), Why Gold, not Oil, is the Superior Predictor of Inflation. Gold Report, World Gold Council, November.
2. Corti, C.H., Holliday, R. (2010), Gold Science and Applications. USA: Taylor and Francis Group, LLC.
3. Chang Pengyuan, Yao Hongxin. Analysis of international gold price forecast [J]. Business Situation, 2019(7):64. (in Chinese) DOI:10.3969/j.issn.1673-4041.2019.07.055.
4. Shafiee, S., & Topal, E. (2010). An overview of global gold market and gold price forecasting. Resources policy, 35(3), 178-189.
5. Potoski, M. (2013). Predicting gold prices. CS229, Autumn.2
6. Khan, M. M. A. (2013). Forecasting of gold prices (Box Jenkins approach). International Journal of Emerging Technology and Advanced Engineering, 3(3), 662-670.
7. Meyler A, Kenny G, Quinn T. Forecasting Irish inflation using ARIMA models[J]. 1998.
8. Tripathy N. Forecasting gold price with auto regressive integrated moving average model[J]. International Journal of Economics and Financial Issues, 2017, 7(4): 324-329.
9. Nyoni T. Modeling and forecasting inflation in Kenya: Recent insights from ARIMA and GARCH analysis[J]. Dimorian Review, 2018, 5(6): 16-40.
10. Rady, E. H. A., Fawzy, H., & Fattah, A. M. A. (2021). Time series forecasting using tree based methods. J. Stat. Appl. Probab, 10, 229-244.
11. Navin G V. Big data analytics for gold price forecasting based on decision tree algorithm and support vector regression (SVR)[J]. International Journal of Science and Research (IJSR), 2015, 4(3): 2026-2030.
12. Lulu Chen. Stock price forecast based on multiple linear regression analysis -- A Case study of China Citic Bank [J]. Economic Research Guide, 2016(19):75-76. DOI: 10.3969/ j.issn.1673-291x.2016.19.032.
13. Shokri, B. J., Dehghani, H., & Shamsi, R. (2020). Predicting silver price by applying a coupled multiple linear regression (MLR) and imperialist competitive algorithm (ICA). Me-taheuristic Computing and Applications, 1(1), 1.