IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Gold Price Prediction Using Machine Learning

Ranganath Maruti Bhajantri[1], Rohith F Puttappanavar[1], Shashank[1], Vishal[1], Dr. Shashikala[2]

1 Student, Department of CSE, Dayananda Sagar Academy of Technology and Management, Bangalore, India.

2 Professor,Department of CSE, Dayananda Sagar Academy of Technology and Management, Bangalore, India.

ABSTRACT

The yellow metal is all time favorite for everyone from the history to till date. Most of the exchange of commodities will happen through this yellow metal or gold. Any country economics is evaluated by the gold available in the reserve bank and even the common man keeps the gold as security purpose and uses as a ornament. The high demand of gold across the global made the gold rate changes very frequently. To invest in the gold the common man or any organizations need to check for its low value and then invest. The predication of gold rate depends on various attribute. This review paper gives a peek into the various predication methods its advantages and limitations which can be used by any one before investing.

The data science and machine learning approaches to anticipate the price of gold. We review historical gold price data and create and assess forecasting models using this data. We examine the information to find significant patterns and correlations that can be utilized to forecast gold prices in the future. We assess different machine learning models' reliability and accuracy in predicting gold prices. We outline the findings of our model evaluation and go over the implications for predicting gold prices. Finally, we evaluate our study's shortcomings and make suggestions for further research.

Keywords: Gold ETF, Linear Regression, Price Prediction, Machine Learning, Supervised Learning

INTRODUCTION

A metallic decoration, gold can also be traded, that is, it can be converted into money. It has served as a representation of worth, monarchy, and purity. Due to its favourable financial potential, this precious metal attracts a lot of interest from people at all phases of life. Over time, the gold's worth has fluctuated. Countless Kings and Emperors from all over India used a variety of monetary systems thousands of years ago, but only gold was regarded as a common exchange good. Gold's price is influenced by supply and demand, just like any other good. Gold is in high demand in India, particularly for jewellery.

Investors and traders alike have been interested in gold prices for a long time. Machine learning methods have gained popularity in recent years for forecasting gold prices. In this study, we investigate how machine learning can be used to forecast gold prices.

Previously, gold price predictions were made by hand. After the introduction of mutual funds, exchange traded funds (ETFs) emerged as the most innovative and widely used securities among Indian investors. Investors who find it challenging to master the trade trick of analyzing and choosing stocks from their portfolio now have access to ETF instruments, which fill a valuable niche.

More specifically, investors' attention has been greatly drawn to ETFs due to their low cost and proven track record of performance. Over time, investing in gold has taken on traditional forms, such as buying jewelry, as well as more

contemporary ones, such as buying gold coins and bars or Gold Exchange Traded Funds (Gold ETF).

The process of predicting the price of gold based on past data and current market trends is known as gold price prediction. A thorough understanding of the gold market and the related economic and political factors is necessary for this intricate process. A forecast for future gold prices is provided by gold price prediction models using a variety of techniques, such as technical analysis, fundamental analysis, and sentiment analysis. These models enable traders and investors to make well-informed decisions and more effectively control their risk exposure.

The Gold ETF is a form of Gold which can be used by anyone to keep the gold in the form of document. The Gold ETF represents the Gold Purity to 100%. The Gold ETF are issued by stock markets through Banks and maintained centrally. The Gold ETF can be continued in the form of document or it can be converted to physical form.

The goal of this research is to create a predictive model that can accurately predict the price of gold over an extended period. Soon, the model ought to be able to accurately forecast the direction of the gold price as well as the key variables that affect gold prices the most. To make more accurate predictions, the model should also be able to distinguish between short-term and long-term trends. As securities of every company, gold ETFs are listed and traded on the National Stock Exchange of India (NSE) and Bombay Stock Exchange Ltd. (BSE). Gold ETFs can be regularly bought and sold at market prices on the cash segment of the BSE and NSE, just like any other corporate stock. Purchasing gold through Gold ETFs entails doing so electronically. ETFs for gold can be bought and sold the same way stocks are. You receive the cash equivalent of the gold instead of the actual metal when you reclaim a Gold ETF. A dematerialized portfolio (Demat) and dealer manage gold ETF investments, making it a remarkably simple way to transact in physical gold.

LITERATURE SURVEY

The "GOLD PRICE PREDICTION USING MACHINE LEARNING" work by Bhavana Purra et al. [1] predicts the gold EFT price based on gold price data from the prior year. This project's goal is to forecast the price of gold so that investors can decide when the best time is to buy and sell the metal. Predicting the price of gold can be crucial to a company's success. The dataset containing the gold price from the prior year is examined to determine the gold price. The rise and fall of the gold price is influenced by a variety of factors, including the capital and real estate markets. Investors fear that the gold price could fall anytime so they think many times before investing in gold, it helps the investors by telling the suitable time for investing.

Gold price is predicted by using,
Analysis using linear regression: The model's R-square yields a result of 98.76%. R-squared typically ranges from 0% to 100%. A score of almost 100% means that the given model does a good job of explaining the prices of Gold ETFs. The analysis demonstrates that the proposed machine learning-based linear regression method is superior to conventional and existing predictive models.
This method's primary flaw is that linear regression is inappropriate for non-linear data. Non-linear data, such as categorical data, cannot be handled by it. It works best with evenly distributed data and assumes that the data is normally distributed.

Sandeep Patalay and et al, have proposed [2] “GOLD PRICE PREDICTION USING MACHINE LEARNING MODEL TREES”, It is thought that one of the commodities used to gauge global economic activity is gold. Numerous economic factors affect the price of gold, making it challenging for the average person to comprehend how difficult it is to predict gold prices. Predicting the price of gold is more crucial than anything else for businesses engaged in the gold trade. In this work, they create a system based on machine learning algorithms to forecast gold prices using historical data from commodities and stock market indicators that are related to gold. The system is based on the M5P model tree machine learning algorithm. This algorithm is used to train on historical commodity prices like crude oil and the S&P 500, which are important indicators of the global financial markets. With a forecasting accuracy of 85%, the analysis demonstrates that the M5P Machine Learning Algorithm outperforms the other approaches.
The S&P 500 index and crude oil prices are independent variables, and the gold price is the dependent variable. The gold price prediction is based on historical data.

The time series data under consideration is trained using the M5P algorithm.
The model created here is 85% accurate at predicting gold prices and will undoubtedly aid commodity investors in limiting their losses thanks to the premium prediction capabilities of the model.

These are the main drawbacks:
The ability of machine learning model trees to generalise from a small set of data is one of their limitations. A model may not be able to correctly forecast future gold prices that differ significantly from the past, for instance, if it is trained on a small sample of historical gold prices.The ability of machine learning model trees to accurately capture non-linear relationships between input variables and gold prices is a limitation. They might be unable to accurately depict the effects of monetary or political events, for instance.
The goal of this study's analysis was to improve the current knowledge-based and gold price prediction models so they could be applied to real-world situations. Traditional linear regression models and machine learning models like ANNs are both used in the price prediction models.

Mrs. B. Kishori and et al, have proposed [3] "GOLD PRICE FORECASTING USING ARIMA MODEL", A metallic ornament, gold can also be traded, that is, it can be converted into money. It has served as a representation of value, royalty, and purity. Due to its favourable investment potential, this precious metal attracts a lot of interest from people at all stages of life. Over time, the gold's worth has changed. Countless Kings and Emperors from all over India used a variety of monetary systems thousands of years ago, but only gold was regarded as a common exchange good. Gold's price is influenced by supply and demand, just like any other good. Gold is in high demand in India, particularly for jewellery. The ability to recognise shifts and variations in the price of gold is crucial for investors. so that investors can make wise decisions regarding their gold investment. This study aims to provide an understanding of timeseries ARIMA model-based gold price forecasting.
A model from Box-Jenkins modelling is called ARIMA. The four repeating steps of the Box-Jenkins methodology are model identification, parameter estimation, diagnosis and verification, and forecasting.

Since there isn't much evidence supporting linearity, the ARIMA model, which is based on a few variables, was created under the fundamental assumption that the data is organised in a linear way.The ability of ARIMA models to detect nonlinear relationships in the data is constrained. As a result, models that employ more advanced techniques, like neural networks, might be better suited for forecasting gold prices that are influenced by more complicated market conditions..
Time series data must be stationary for ARIMA models, which means that the mean and variance of the data must not change over time. The model might not be able to accurately depict the underlying patterns in the data if the data is not stationary.
One effective method for predicting gold prices is the ARIMA model. It is a statistical model that forecasts future values based on historical data. Time series data are modelled using the Autoregressive Integrated Moving Average (ARIMA) method to predict future values. ARIMA models use historical data to establish the model's parameters, which are then applied to forecast future values. Given how unpredictable and volatile gold prices are, the model is especially helpful in this market.

Iftikhar ul Sami and et al, have proposed [4] "PREDICTING FUTURE GOLD RATES USING MACHINE LEARNING APPROACH" The current market price at which a commodity is bought or sold for prompt payment and delivery is known as the spot price. It differs from the futures price, which is the cost at which the parties agree to conduct business later. Based on supply and demand in the gold market, gold spot rates are decided twice daily. Government banks and these investors could both make or lose a lot of money depending on how much the price of gold fluctuates. Investors can choose when to buy gold by predicting the rise and fall in the daily gold rates. Researchers have conducted several studies to predict gold rates.
While ANN has several tuning parameters, the two that have the biggest effects on performance are the number of layers and learning rate.
The outcome of running an ANN on the test set while changing these two parameters' values on the training set. Since it is the most thorough study to date, it considers a variety of economic indicators from different nations and businesses.
This work will not consider the:
Market sentiment: Machine learning algorithms are unable to consider market sentiment. When making investments, human investors take the

mood of the market into account, which is difficult for machines to replicate.
Randomness: A wide range of unpredictable factors can have an impact on gold prices. Machine learning algorithms struggle to predict future gold prices with accuracy because of this randomness.
According to the analysis of this work, future gold rates can be predicted using a machine learning method based on historical data. This strategy uses machine learning algorithms to find patterns in historical gold price data and then uses those patterns to forecast future gold prices. Utilizing the capabilities of artificial intelligence and combining data from various sources, this strategy has the potential to increase accuracy.

Manjula K A and et al, have proposed [5] "GOLD PRICE PREDICTION USING ENSEMBLE BASED MACHINE LEARNING TECHNIQUES", Saving and investing are regarded as essential components of human existence. Investment in finance refers to the acquisition of specific assets with the intention of profiting from them in the future. There are numerous investment opportunities, including those in stocks, cryptocurrencies, commodities, and real estate. Due to its rising price and widespread use in jewellery, gold is one of these and is also thought of as an investment. Based on changes in the stock market or capital market, the price of gold rises or declines. They considered 22 market factors that affect the price of gold when doing this research. The model in this work predicts the price of gold by analysing historical data sets using machine learning.
This study examines the relationship between the price of gold and the variables that affect that price. The three machine learning algorithms used in this study are gradient boosting regression, random forest regression, and linear regression. These algorithms support data analysis. We can determine the accuracy of these three methods under various circumstances by comparing them.
Five main factors that are thought to have an impact on the price of gold are outlined in this work. Stock market, crude oil price, rupee-to-dollar exchange rate, inflation, and interest rate are the variables considered for this study.
Mean Squared Error (MSE), Root Mean Square Error (RMSE), and Mean Absolute Error were used to assess the regression techniques' predictive power (MAE)
Since MAE is not differentiable, gradient descent optimization is challenging. MSE and RMSE square the errors before calculating the average, they are not resistant to outliers. Gradient boosting is a machine learning technique that generates a decision tree-based prediction model.

Vidya G S and et al, have proposed [6] "GOLD PRICE PREDICTION AND MODELLING USING DEEP LEARNING TECHNIQUES", the demand for gold is increasing daily. Gold is one of the best assets to invest in, according to gold trend. Data modelling and prediction use a variety of models. The price of gold is not linear. For better investments, gold price prediction is necessary. RNNs are the most effective among Convolutional Neural Networks (CNNs) for time series estimation and prediction. CNNs are one of the best methods for resolving nonlinearities in data.
Various commodities are priced differently because of numerous factors. Market swings, monetary policy, epidemics, climatic conditions, etc. are a few of the things that have a big impact on it. The ability to extract nonlinearity from data and analyse behaviour is exquisite today, allowing for the execution of sound financial plans.
Artificial neural networks (ANNs) have been used in this study. To implement computing behaviour, ANN imitates human brain mechanisms. Convolutional neural networks (CNNs) can recognise nonlinear patterns in data characteristics.
The choice of the best analysis strategy and the application of asset pricing theory in nonlinear circumstances is the main issue in this work.
Studying the behaviour of financial data is challenging and computationally complex.
To predict the nature of the gold rate, an LSTM network is created and designed based on the analysis of this work. This study shows that the proposed model performs better than established techniques like ARIMA, covariance matrix estimation, deep regression, SVR, and CNN.
Chintya Christina and et al, have proposed [7] "GOLD PRICE PREDICTION USING TYPE-2 NEURO-FUZZY MODELING AND ARIMA", Due to its rising value, gold is regarded as an asset for investing. It's crucial to predict and forecast to get the most profit and benefits from the investment. Based on historical data that was analysed with specific algorithms, this work makes predictions about the price of gold. These data were examined, analysed, and linked to the changes in time. The time factor is taken into consideration, so the analysis's findings will indicate a result that is uncertain and could change in the future. It shows that the obtained

forecasting result is not always 100% accurate. However, this does not mean that forecasting is useless; on the contrary, forecasting has been used as the cornerstone for planning and making decision in relation to gold investment. In this study, the effectiveness of type-2 neuro-fuzzy modelling, the most popular method for price prediction, will be compared to ARIMA. This study aims to forecast the price of gold using type-2 neuro-fuzzy modelling and compare the precision of the prediction using ARIMA and type-2 neuro-fuzzy modelling.

In this study, type-2 neuro-fuzzy modelling is used to predict the price of gold. Self-Constructing Clustering is used to divide the historical gold price data into several clusters and generate some type-2 fuzzy rules.

Using a hybrid learning algorithm, the parameters are optimised, Least-Square Estimation (LSE) is used to optimise the consequent parameters in type-2 Neuro-Fuzzy modelling, while Particle Swarm Optimization (PSO) is used to optimise the antecedent parameters.

Error prediction using ARIMA and type-2 Neuro Fuzzy differs little in the first scenario but significantly in the second.

For type-2 neuro-fuzzy systems to accurately model the system under study, several fuzzy rules and membership functions must be created.

Chunlin Xin and et al, have proposed [8], "COMPETITIVE ANALYSIS OF GOLD TRADING AND ITS RISK REWARD MODEL", Trading in gold has become a significant method of investing. Investors try to buy gold at a low price and sell it at a high price to maximise their profit. In real life, buying and selling gold is a common online decision-making problem. Only by mastering the purchase price and selling price can investors achieve higher returns when dealing with the erratic fluctuations of gold. Gold's long-term maintained value and added value are determined by its natural and social characteristics; as a result, the price of gold usually rises again after falling for a while. The value of the US dollar, economic conditions, stock market fluctuations, inflation, global political unrest, gold reserves, and many other variables can have a significant impact on gold's price. Investors use gold price forecasting to increase profits as they attempt to reduce risks that could result in losses. When gold prices rise in the future, we research the online strategy and present the risk-reward model.

A risk-reward model aids investors in objectively evaluating the potential risk and reward of various investments, enabling them to allocate their capital more wisely and intelligently.

A risk-reward model can assist investors in identifying and mitigating potential risks as well as in making more informed decisions about how much risk to take on by comparing the risk and reward of various investments.

It restricts itself to this parameter:

Information that is lacking: The accuracy of a risk-reward model depends on the data used to create it. The model might not accurately depict the potential risk and reward of an investment if the data is lacking or unreliable.

Focus on the short term: A risk-reward model is frequently concerned with the performance of an investment over the short term and may overlook the investment's potential over the long term.

Zhanhong He and et al, have proposed [9], "GOLD PRICE FORECAST BASED ON LSTM-CNN MODEL", they initially suggest a novel method for predicting the price of gold that combines attention mechanisms, long short-term memory neural networks (LSTM), and convolutional neural networks (CNN) (denoted to LSTM-Attention- CNN model). Three parts make up the LSTM-Attention-CNN model: the LSTM component, the Attention Mechanism component, and the CNN component. The estimation of the relationship between potential predictors and expected returns is greatly influenced by a variety of factors, making the financial domain extremely complex and nonlinear. They initially suggest an LSTM-CNN model with an Attention Mechanism for daily gold price forecasting in this work. They carry out extensive experiments, and the outcomes demonstrates that the given model performs better than other traditional models. However, computing a set of input data, such as economic data, accounting data, daily transaction records, etc., mapping to a particular return is at the core of deep learning. Therefore, theoretically, we can still find the relationships for the corresponding returns regardless of how irregular or non-linear the data we input.

Processing of sequential data: A recurrent neural network (RNN) type called LSTM is especially effective at handling sequential data, such as text or time series data. It can remember pertinent

information for a longer period and capture long-term dependencies in the data.

High performance: LSTM-CNN models can perform well on a variety of tasks involving natural language processing. They have proven to be successful at tasks like sentiment analysis, text classification, and language translation.

Complexity: For those with little machine learning experience, designing, and implementing LSTM-CNN models can be challenging. To achieve good performance, they require careful tuning of the various hyperparameters.

Data requirements: To train, LSTM-CNN models need a lot of labelled data, which can be time-consuming and expensive to acquire. The data must also be accurate and pertinent to the task being modelled.

REFERENCES

1. Gold Price Prediction using Machine Learning, Chodavrapu pragna, Bhavana Purra, Adari Viharika, Garabhapu Roshni, Prof. B. Prajna: 2022
2. Gold Price Prediction using Machine Learning Model Trees, Sandeep Patalay, Dr. Madhusudhan Rao Bandlamudi : 2021
3. Gold Price Forecasting using Arima Model, Mrs. B. Kishori, V. Preethi: 2018
4. Predicting Future Gold Rates using Machine Learning Approach, Iftikhar ul Sami, Khurum Nazir Junejo: 2017
5. Gold Price Prediction using Ensemble based Machine Learning Techniques, Karthikeyan. P, Manjula Ka: 2019
6. Gold Price Prediction and Modelling using Deep Learning Techniques, Hari V S, Vidya G S: 2020
7. Gold Price Prediction Using Type-2 Neuro-Fuzzy Modeling and ARIMA, Chintya Christina, Rian Febrian Umbara :2015
8. Competitive Analysis of Gold Trading and its Risk reward model, Chunlin Xin, Miaomiao Zhang: 2011
9. Gold Price Forecast based on LSTM-CNN Model, Zhanhong He, Junhao Zhou, Hong-Ning Dai, Hao Wang: 2019