iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.62251

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Gold Price Prediction Using Machine Learning Techniques

Prabhat Kumar[1], Puran Singh[2], Dr. Sadhana Rana[3]
CSE Department, SRMCEM, Lucknow, India

Abstract: ***Gold holds critical esteem because it is commonly utilized to back exchange exchanges. In nations such as India, China, Indonesia, and others, gold isn't as it were cherished as gems but is additionally traded as blessings or utilized in marriage settlements. Furthermore, nations with significant gold saves are regularly seen as financially affluent. Within the show period, gold rises above its conventional part as unimportant decorations and is recognized as a important speculation. Central banks around the world hold gold as collateral for outside credits and as a implies to control swelling. The worldwide economy's status essentially impacts gold costs, given the rising request and supply of this valuable metal.***

Thus, numerous financial specialists are turning to gold ventures, propelled by the vacillations in its advertise esteem. Be that as it may, the instability in gold costs postures dangers for financial specialists. Subsequently, the objective of "Gold Cost Prediction" is to utilize different machine learning procedures to figure gold costs. This includes analyzing the complicated relationship between economic variables affecting gold rates. Utilized machine learning calculations incorporate Irregular Timberland Relapse, Choice Tree (DT), Back Vect or Relapse (SVR), Direct Relapse, and Fake Neural Arrange (ANN).

Keyword: Machine Learning, Stock Price, Prediction, Support vector machine (SVR), Artificial neural network (ANN), Random Forest (RF), Decision tree (DT)

I. INTRODUCTION

Gold, settled in as a valuable metal over various eras, expect a significant part as a pivotal money related resource for countries and a essential component in worldwide financial saves for exchange and money supporting [1]. The inherent characteristics of gold, characterized by its inactivity, constrained supply, and strenuous extraction prepare, have contributed essentially to its regarded esteem within the showcase. Be that as it may, past shortage and steadiness, the burgeoning request and supply of this precious metal are driven by another compelling figure – the burgeoning consideration it garners from investors. Gold has risen as a foundation resource within the speculation scene, gaining the moniker of "the extreme safe-haven asset" [2]. Speculators frequently turn to gold costs as a solid pointer of showcase changes. In spite of its seen soundness, the unusual undulations in gold rates present an component of chance, ingrains a sense of anxiety among speculators and underscoring the need for fastidious gold rate predictions. This inquire about endeavours to dig into the complex interconnects between particular financial advertise factors and the energetic domain of gold costs. The factors beneath investigation include the S&P500 file, silver cost, rough oil cost, and the US dollar trade value (USD). In a offered to improve prescient exactness, the think about utilizes five advanced machine learning techniques—Support Vector Relapse (SVR), Choice Tree (DT), Irregular Woodland, Straight Relapse, and Counterfeit Neural Arrange (ANN). The objective is to scrutinize and compare these strategies to discover the foremost capable show, hence contributing to a nuanced understanding of the complex elements administering gold cost developments.

II. PROBLEM STATEMENT

The expanding importance of gold as a valuable metal in worldwide money related scenes, coupled with its part as a urgent component in financial saves, prompts a require for comprehensive inquire about. Whereas gold's inactive nature, constrained supply, and complicated extraction forms contribute to its regarded advertise esteem, the multifaceted elements driving its request and supply warrant a nuanced examination. Past its conventional qualities, gold has advanced into a central venture resource, often hailed as "the extreme safe-haven asset." Be that as it may, the inborn chance related with the unstable rise and drop of gold rates raises concerns among speculators, requiring a cautious and prescient approach.

This survey paper points to address the basic issue of understanding the complicated interconnects between gold costs and particular financial showcase factors. The factors in center include the S&P500 record, silver cost, rough oil cost, and the US dollar trade esteem (USD). Given the advancing scene of gold as an venture, foreseeing its rates has gotten to be progressively complex, requiring progressed techniques.

©IJRASET: All Rights are Reserved

The inquire about utilizes five particular machine learning techniques—Support Vector Relapse (SVR), Choice Tree (DT), Irregular Woodland, Direct Relapse, and Manufactured Neural Arrange (ANN)—to analyze and compare their adequacy in predicting gold costs. By scrutinizing the exchange of these financial components, this audit points to contribute profitable experiences into the challenges and openings related with determining gold costs, supporting financial specialists and policymakers alike in making educated choices within the ever-evolving worldwide monetary scene.

III. LITERATURE REVIEW

The inquire about displayed within the diary [1], titled "Gold Cost Prediction," digs into the application of machine learning frameworks for determining gold costs based on different financial factors. The think about includes a comparative investigation of the relationship between these financial factors and gold costs. Utilizing straight relapse, Bolster Vector Machines (SVM), and Choice Tree calculations, the analysts prepared models to foresee gold rates.

In the moment diary [2], the center is on foreseeing future gold costs utilizing machine learning procedures. Recognizing gold's significance as one of the foremost profitable metals all inclusive, the think about highlights how nations keep up gold saves to extend a sound and dynamic picture. The proposed demonstrate points to help financial specialists in analyzing and possibly profiting from contributing in gold commodities.V.K.F.B.Rebecca Davis, in her article [3] titled "Modeling and Estimating of Gold Costs on Money related Markets," utilizes the Autoregressive Moving Normal (ARMA) show to analyze time arrangement information, particularly the month to month costs of gold over a ten-year period.

The model's precision is detailed at 66.67%.Iftikhar ul Sami and Khurum Nazir Junejo, in their think about [4] titled "Predicting Future Gold Rates utilizing Machine Learning Approach," utilize Manufactured Neural Systems (ANN) to forecast gold costs. Information crossing eleven a long time, counting variables such as unrefined oil costs, the S&P 500 record, USD trade rates, and other financial markers, contribute to their investigation.

The article by D Makala and Z Li [5], titled "Prediction of the Gold Cost with ARIMA and SVM," draws on data from the World Gold Chamber, covering day by day gold costs from January 1979 to December 2019. The consider utilizes the Autoregressive Coordinates Moving Normal (ARIMA) approach in conjunction with Bolster Vector Machines (SVM) for anticipating gold costs, noticing that the SVM demonstrate beats the ARIMA show in precision.

IV. METHODOLOGY

A. Random Forest Regressor

The Random Forest Regressor is grounded in the principle of ensemble learning, a technique that leverages the outputs of multiple machine learning algorithms or repetitions of similar algorithms. The rationale behind ensemble learning is that the combination of outputs from diverse models often yields more accurate results than relying on a single model alone [6]. In the context of Random Forest Regression, this technique employs multiple decision trees trained collectively to generate a consolidated prediction. Ensemble learning serves a crucial purpose in situations where individual models may fall short in delivering precise results. By aggregating the predictions from various models, Random Forest Regression enhances the overall accuracy and robustness of the outcome. Random Forest, a supervised learning technique, is proficient in handling both regression and classification tasks. Through ensemble learning, it excels in mitigating the limitations of individual models, ensuring a more reliable prediction by amalgamating the insights derived from multiple decision trees. This approach proves particularly valuable when faced with intricate datasets and real-world scenarios, where a singular model might not capture the full complexity of the underlying patterns.

B. Decision Tree

The decision tree organized a series of rules in the tree structure. A set of rules were arranged in a tree structure via a decision tree. It is one of the most useful techniques for supervised non-parametric learning. The decision tree has a few key terminologies. The Root node, which is at the top of the tree, depicts the overview of the entire community as it splits into sub-nodes. Decision nodes are referred to when sub-nodes can be split into further nodes. The node of a tree with no additional split nodes is called a leaf node

C. Support Vector Regressor

Support vector regression is a part of a support vector machine, and this concept is applied for both regression as well as classification of data. Support vector machine is one of the fast and non-linear supervised machine learning models, given a set of labeled training data. SVM will help to find the optimum hyperplane which categorized new examples in one-dimensional space. The hyperplane is a point in two-dimensional space.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

The hyperplane is a land in three-dimensional space, it is a surface that divides into two parts, where each class lies on either side. In two-dimensional linear separable cases, the data are separated by a line. In this case, the data is linearly separable [9].

D. Linear Regression

Linear regression is the linear estimation of a normal connection between two or more variables.
Regression models are extremely important since they are among the most popular methods for reasoning and prediction and the activity goes like this [10].

$\boldsymbol{Y} = \boldsymbol{F}(\boldsymbol{X1}, \boldsymbol{X2}\cdots\boldsymbol{Xk})$ …… (1)

Independent variables labeled as X1, and X2, are the predictors. Y is the function of the X variable. The regression model is a linear estimation of this function. T.

E. Artificial Neural Network

A neural network means, the connection of neurons. It constructs the base of deep learning. The working of Human neural networks and Artificial neural networks are almost the same, although there is a difference in their mechanism.
An artificial neural network works in three layers, the Input layer, and the output layer, and in between there exists the Hidden layer. The input layer receives the data, the output layer provides the final result, and the hidden layer is in charge of handling all of the weight distribution, calculating bias, and calculations
[11][12]. The formula is mentioned below:

$\boldsymbol{\Sigma i}=\boldsymbol{1}$

$\boldsymbol{n}\ \boldsymbol{W} * \boldsymbol{X}i + \boldsymbol{B}$ …… (2)

$\boldsymbol{y1}=\boldsymbol{Activation}(\boldsymbol{W1}*\boldsymbol{X1}+\boldsymbol{W2}*\boldsymbol{x2}+\cdots\boldsymbol{Wn}*\boldsymbol{Xn})$ …… (3)

V. CHALLENGES AND LIMITATIONS

A. Data Quality and Availability

1) *Challenge:* The quality and availability of historical data on gold prices and related economic variables can vary, impacting the robustness and generalizability of the findings.
2) *Limitation:* Limited access to comprehensive, high-quality datasets may restrict the depth of analysis and hinder the establishment of universally applicable trends.

B. Model Sensitivity

1) *Challenge:* Machine learning models, including those reviewed here, often exhibit sensitivity to changes in input parameters, potentially resulting in variations in predictive accuracy.
2) *Limitation:* The sensitivity of models may lead to challenges in model stability and hinder their performance in dynamic market conditions.

C. Overfitting and Generalization

1) *Challenge:* Overfitting, where a model performs well on training data but poorly on unseen data, remains a persistent challenge in machine learning applications.
2) *Limitation:* The risk of overfitting could impact the practical utility of the models discussed, emphasizing the need for robust validation strategies.

D. Interpretability

1) *Challenge:* Some machine learning algorithms, such as Artificial Neural Networks (ANN), lack interpretability, making it challenging to decipher the rationale behind specific predictions.
2) *Limitation:* Limited interpretability may pose challenges in understanding the factors influencing gold price predictions, limiting the practical insights for stakeholders.

E. Model Complexity

1) *Challenge:* Certain machine learning algorithms, particularly ensemble methods like Random Forest, can be complex, making it challenging to communicate their functionality and outcomes effectively.

2) *Limitation:* The complexity of models may hinder their adoption and comprehension, especially for non-expert users and decision-makers.

F. External Factors

1) *Challenge:* External factors, such as geopolitical events or economic policy changes, are often beyond the scope of machine learning models but can significantly impact gold prices.
2) *Limitation:* The inability of models to account for unforeseen external factors limits their predictive accuracy in dynamic and evolving economic landscapes.

G. Model Evaluation Metrics

1) *Challenge:* Variability in the selection of evaluation metrics across different studies may hinder direct comparisons and standardization.
2) *Limitation:* Lack of standardized metrics may complicate the identification of the most effective machine learning models for predicting gold prices.

Addressing these challenges and limitations is crucial for advancing the field and enhancing the practical applicability of machine learning models in predicting gold prices. Future research should aim to overcome these obstacles, fostering the development of more robust and reliable predictive models.

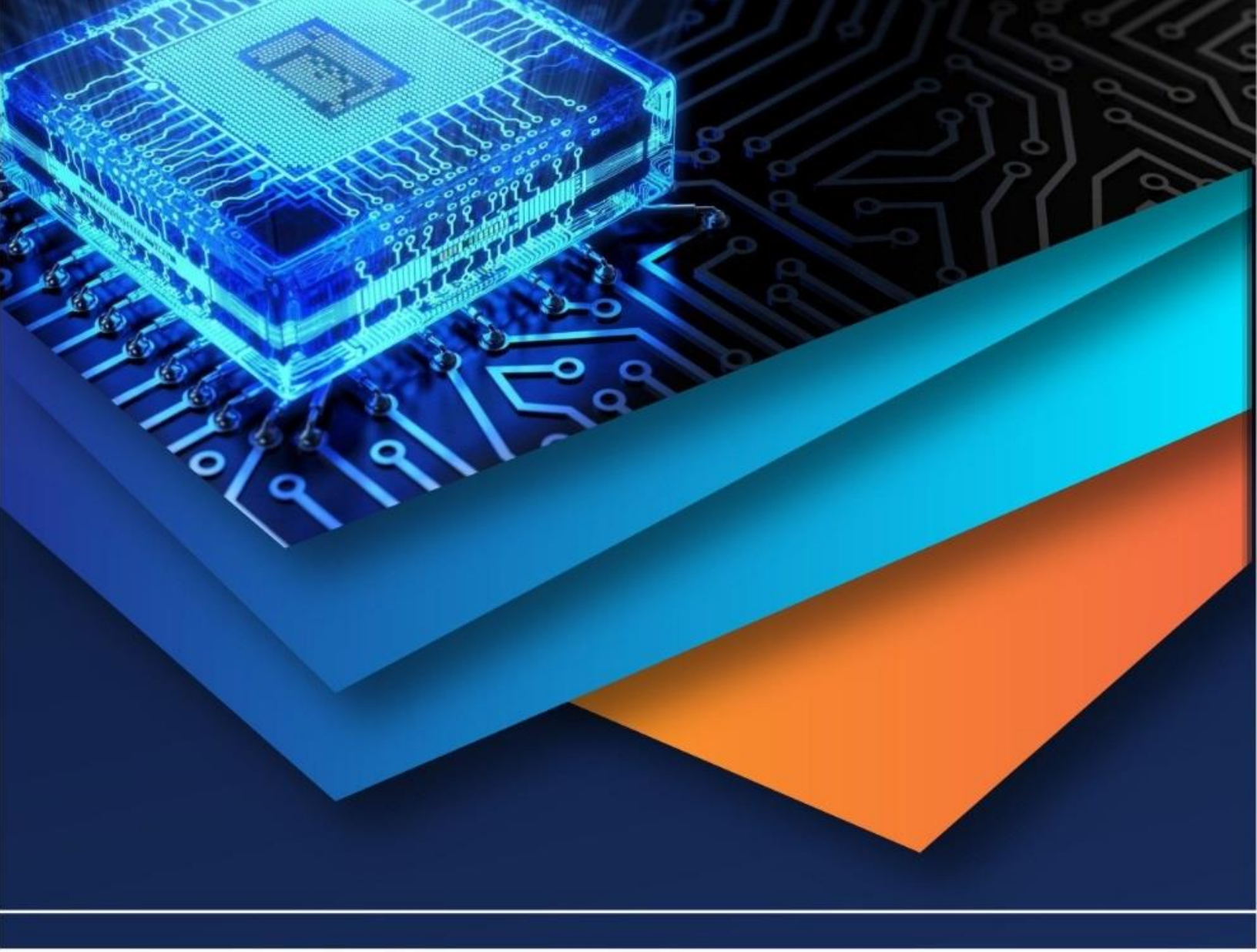
VI. CONCLUSION

In synthesizing the different considers inspected in this audit paper on the expectation of gold costs utilizing machine learning procedures, a comprehensive understanding develops with respect to the complicated relationship between financial factors and gold cost elements. The amalgamation of experiences from different machine learning calculations, counting Arbitrary Woodland Relapse, Bolster Vector Machines (SVM), and Fake Neural Systems (ANN), sheds light on the advancing scene of gold price forecasting. The investigation of outfit learning, eminently through Irregular Woodland Relapse, underscores its essential part in improving prescient exactness. By saddling the collective control of numerous choice trees, this method addresses the confinements of person models, giving a vigorous system for anticipating gold costs in the midst of the complexities of budgetary markets. Yet, the audit has lit up a few challenges and restrictions inalienable within the current body of investigate. From information quality concerns to issues of show affectability and interpretability, these challenges require a cautious approach in applying machine learning procedures to gold cost expectation. The heterogeneity in assessment measurements over thinks about assist emphasizes the require for standardized strategies to encourage significant comparisons. As this audit paper concludes, it gets to be clear that whereas machine learning offers a effective toolset for foreseeing gold costs, future research must explore these challenges. Securing high-quality datasets, creating interpretable models, and building up standardized assessment measurements are fundamental for progressing the unwavering quality and pertinence of machine learning within the domain of gold cost prediction. This survey serves as a establishment for future examinations, encouraging analysts to dive more profound into refining strategies and overcoming the recognized challenges. As the intrigue field of machine learning and back proceeds to advance, the experiences gathered from this survey contribute to the progressing exchange on leveraging progressed computational strategies for more exact and noteworthy expectations within the unstable world of gold markets.

REFERENCES

[1] M. Sravani, Ch. Abhilash, T. Divya, Ch. Vasthav, D. Priyanka, “Gold price prediction”, International journal of creative research thoughts (IJCRT), Issue- 6 (June 2021).
[2] Iftikhar ul Sami, Khurum Nazir Junejo, “Predicting Future Gold Rates using Machine Learning Approach”, International Journal of Advanced Computer Science and Applications, Issue-12(2017).
[3] V. K. F. B. Rebecca Davis, "Modeling and Forecasting of Gold Prices on Financia Markets," American International Journal of Contemporary Research, 2014.
[4] Iftikharul Sami and Khurum Nazir Junejo, "Predicting Future Gold Rates using Machine Learning Approach", International Journal of Advanced Computer Science and Applications, 2017.
[5] D Makala and Z Li, “Prediction of gold price with ARIMA and SVM”, Journal of Physics: Conference Series, 2021.
[6] S. Chakravarty, B. K. Paikaray, R. Mishra and S. Dash, "Hyperspectral Image Classification using Spectral Angle Mapper," 2021 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE), 2021, pp. 87-90, doi: 10.1109/WIECONECE54711.2021.9829585.
[7] Manjula K A, Karthikeyan P,” Gold Price Prediction using Ensemble based Machine Learning Techniques”, IEEE 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Issue-23-25 (April 2014).
[8] Z. Ismail, A. Yahya, A. Shabri,” Forecasting Gold Prices Using Multiple Linear regression Method”, American Journal of Applied Sciences, Issue-2009.

[9] L.A Sjaastad, F. Scacciavillani," The price of gold and exchange rate", Journal of International Money and Finance, Issue-December, 1996
[10] loannis E. Livieris, EmmanuelPintelas, Panagiotis Pintelas," A CNN-LSTM model for gold price time-series forecasting", Neural computing and Applications, Issue- 21(November 2019).
[11] Dr. Abhay Kumar Agarwal, Swati Kumari, "Gold Price Prediction using Machine Learning", International Journal of Trend in Scientific Research and Development (ijtsrd), 2020.
[12] Sami Ben Jabeur, Salma Mefteh-Wali, Jean-Laurent Viviani," Forecasting gold price with the XGBoost algorithm and SHAP interaction values, Annals of Operations Research, Issue-24(June2021).



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)