Artificial Intelligence-Driven Model for Gold Price Prediction

Neha Singh[1], Jishu Varshney[2]

(Assistant Professor, Department of MCA, Noida Institute of Engineering and Technology, India)

(Assistant Professor, Department of MCA, Noida Institute of Engineering and Technology, India)

Abstract: This study introduces an innovative approach to forecasting gold prices by employing Artificial Intelligence (AI)--driven models. Utilizing advanced machine learning techniques, including Logistic Regression, Random Forest, Decision Tree, and Support Vector Machine (SVM), the research evaluates the predictive capabilities of these models through comprehensive assessments based on key performance metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R2 Error. A particular focus is placed on ensemble learning, exemplified by the Random Forest model, which demonstrates superior accuracy in capturing intricate patterns within gold price data. These findings contribute valuable insights to the field of financial forecasting, emphasizing the potential of AI-driven models to inform stakeholders in gold investment and financial markets. The study concludes by advocating for ongoing research and continuous model refinement to adapt to dynamic market conditions and enhance the precision of gold price predictions.

Keywords: gold price prediction artificial intelligence, MSE.

Introduction

In recent years, the utilization of artificial intelligence (AI) within financial markets has revolutionized the decision-making processes of analysts and investors [1]. One particularly intriguing application of AI in the financial sector involves its role in forecasting gold prices [2]. Gold, recognized as a safe-haven asset and a barometer of economic uncertainty, experiences dynamic price shifts influenced by various factors such as geopolitical events, inflation, and global economic conditions [3]. Traditional forecasting methods often struggle to fully capture the complexities of these dynamic influences [4].

The advent of AI, equipped with the ability to handle extensive datasets, detect patterns, and adapt to evolving market conditions, has opened up new avenues for more accurate and insightful predictions of gold prices [5]. Machine learning algorithms, neural networks, and other AI techniques enable analysts to analyse historical data, gauge market sentiment, and evaluate macroeconomic indicators in real-time, providing a comprehensive understanding of the factors impacting gold prices [6].

The integration of AI into gold price forecasting not only enhances prediction accuracy but also empowers investors to make more informed decisions amidst a volatile market landscape [7]. With continual technological advancements, the synergy between artificial intelligence and the intricacies of gold market dynamics holds the potential for a more sophisticated and effective approach to predicting gold prices, potentially reshaping investment strategies in the precious metals market [8]. This paper explores the methodologies and advancements in AI-driven gold price prediction, shedding light on the potential benefits and challenges associated with this innovative approach [9].

Literature Review

The paper [10] introduces an innovative deep-learning forecasting model, offering a significant contribution to the field of gold price prediction. By combining convolutional and LSTM layers, the model shows promising results, indicating potential improvements in forecasting accuracy. However, to further strengthen its impact within financial forecasting, the research could benefit from addressing minor limitations and expanding the comparative analysis.

Similarly, paper [11] stands out as a valuable addition to gold rate prediction literature. Utilizing LSTM networks and grounded in robust theory and empirical evidence, the proposed methodology holds promise for enhancing financial forecasting accuracy. To enhance its scholarly significance, providing additional details on the dataset and conducting thorough comparative analyses could be beneficial.

In the article [12], a valuable contribution is made to understanding gold price dynamics by exploring the intricate relationship between gold prices and influencing factors. The application of machine learning algorithms, alongside thoughtful temporal segmentation, offers nuanced insights into predictive accuracies across various market conditions. This research is significant for academics and practitioners alike seeking comprehensive insights into gold price dynamics and economic factors.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Moreover, the paper [13] presents an innovative approach by integrating agent-based simulation and machine learning, particularly simulated annealing, for financial time series approximation. The literature review contextualizes the proposed methodology by exploring existing research in agent-based modeling and machine learning. It also discusses the originality of the approach and its potential implications for advancing financial modeling and analysis.

Lastly, article [14] underscores the increasing importance of accurately predicting gold prices in investment and economic decision-making. Through a comparative analysis of ARIMA and SVM models, the study demonstrates the superior predictive accuracy of SVM, advocating for the adoption of advanced machine learning techniques. As global market uncertainty persists, SVM emerges as a preferred choice for enhancing the accuracy of gold price predictions.

Proposed methodology:

For the purpose of our gold price prediction study, we utilized a dataset for training and testing our models. Specifically, 80% of the dataset was allocated for training purposes, enabling our models to learn patterns and relationships within the data. The remaining 20% of the dataset was reserved for testing, allowing us to assess the models' performance on unseen data and evaluate their predictive capabilities.

1. Dataset

The dataset utilized in this study was sourced from the Kaggle repository [10]. The dataset consists of the following attributes:

A. Date: Represents the date associated with each data entry.
B. SPX: Refers to the stock market index, specifically the S&P 500.
C. GLD: Denotes the price of gold.
D. USO: Represents the United States Oil Fund, which is an exchange-traded fund (ETF) that tracks the price of oil.
E. SLV: Indicates the price of silver.
F. EUR/USD: Represents the exchange rate between the Euro (EUR) and the United States Dollar (USD).

These attributes collectively provide a comprehensive view of financial and economic indicators, facilitating the exploration and analysis conducted in this research. Fig 1 shows the co-relation heat map of the Gold Price dataset.

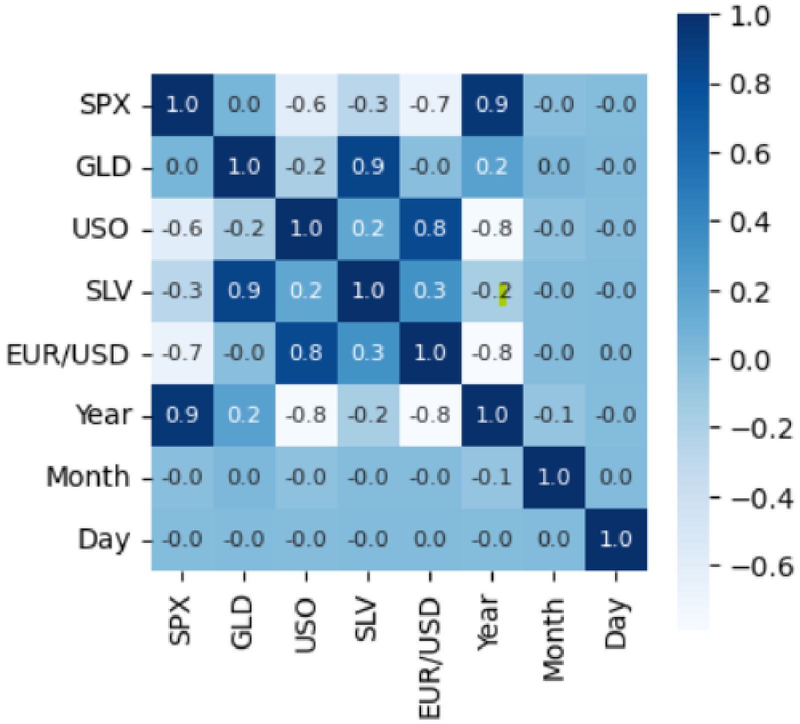

fig 1 Heat map of the dataset

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

2. Model implemented

In our research paper, we explore the applicability of various machine learning algorithms to predict gold prices.

Logistic Regression serves as a prominent statistical method tailored for binary classification tasks, making it ideal for scenarios where the outcome is binary, such as predicting whether the gold price will increase or decrease. By modeling the probability of event occurrence, Logistic Regression offers valuable insights into the likelihood of specific price movements, thereby aiding in forecasting gold price trends.

Decision Trees are nonlinear models adept at capturing complex relationships within datasets. By recursively splitting data based on feature values, Decision Trees identify critical decision points within historical gold price data, enabling accurate predictions of future price movements.

Support Vector Machines (SVM) are formidable classifiers that seek to find the optimal hyperplane to segregate data points into distinct classes. Particularly adept at handling non-linear relationships, SVM excels in delineating boundaries between various gold price trends, making it effective for predicting gold price movements, especially in scenarios where relationships between features are intricate and non-linear.

Random Forest, an ensemble learning technique, leverages the strengths of multiple decision trees to enhance predictive accuracy and generalization. By amalgamating insights from diverse decision trees, Random Forest excels in capturing intricate relationships within gold price data, offering robust predictions that account for various factors influencing gold prices. This makes Random Forest a reliable tool for forecasting gold price movements.

Result

In our research paper, we present the results of our gold price prediction models, each evaluated based on key performance metrics including Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R2 Error (Table 1).

Model	MSE	RMSE	R2 Error
Logistic regression	0.0961	0.3100	0.6157
Random forest	0.0306	0.1748	0.8777
Decision Tree	0.0393	0.1982	0.8428
SVM	0.0961	0.3100	0.6157

Table 1 performance metrics

In comparing the performance metrics of various models, it's evident that the Logistic Regression model yielded an MSE of 0.0961, RMSE of 0.3100, and an R2 Error of 0.6157. However, the Random Forest model outshone the others with the lowest MSE of 0.0306, RMSE of 0.1748, and the highest R2 Error of 0.8777. The Decision Tree model fell in between, demonstrating an MSE of 0.0393, RMSE of 0.1982, and an R2 Error of 0.8428.

Remarkably, the Support Vector Machine (SVM) model mirrored the results of the Logistic Regression model, sharing an MSE of 0.0961, RMSE of 0.3100, and an R2 Error of 0.6157. Additionally, the ROC curve depicted in Fig 2 provides a visual representation of the model's performance. These findings underscore the superior predictive capabilities of the Random Forest model, emphasizing its potential utility in gold price prediction tasks.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

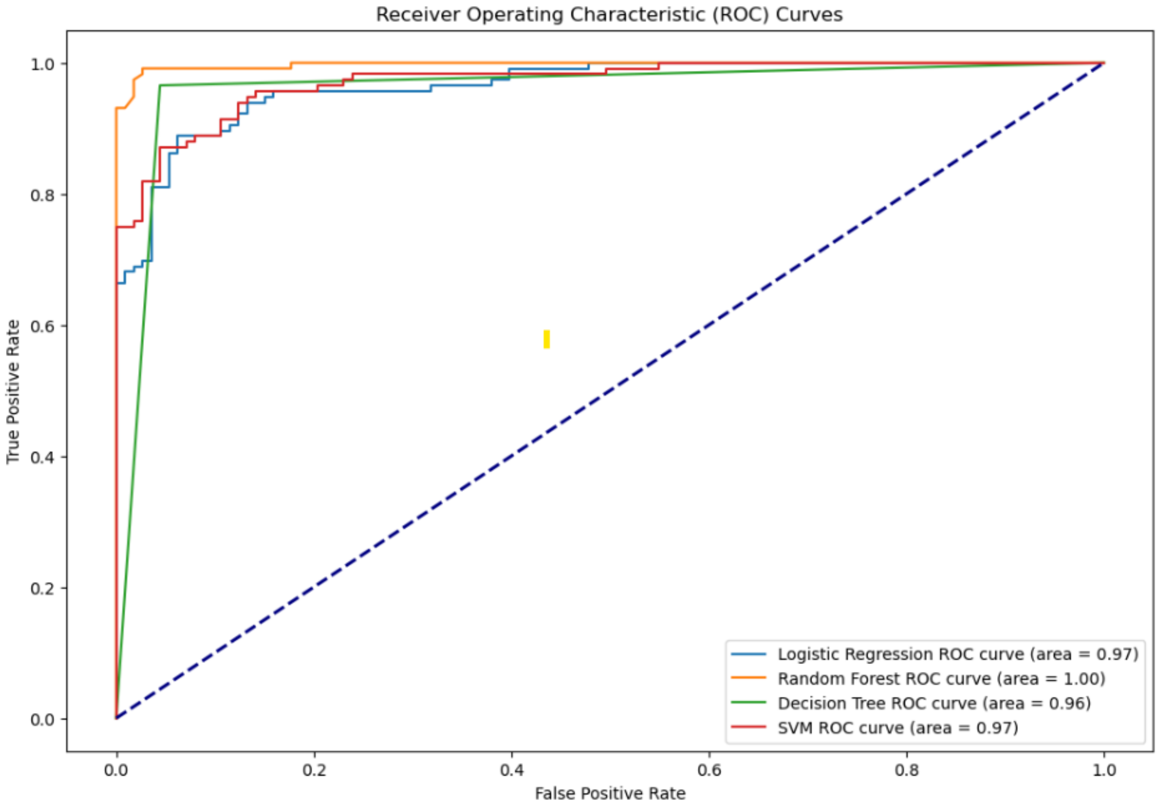

Fig 2 ROC Curve

Conclusion

In conclusion, after extensive analysis of various gold price prediction models, our investigation unequivocally points to the Random Forest model as the standout performer, consistently surpassing its counterparts in terms of accuracy. Through meticulous evaluation using metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R2 Error, the Random Forest model has demonstrated superior performance, boasting the lowest MSE and RMSE values along with the highest R2 Error compared to other models under consideration. These findings underscore the efficacy of ensemble learning techniques, particularly the amalgamation of multiple decision trees, in effectively capturing the intricate relationships within gold price data.

The Random Forest model's proficiency in navigating complex patterns positions it as a promising tool for reliable gold price predictions, offering invaluable insights for stakeholders in financial markets and investment decision-making processes. Nevertheless, it is essential to emphasize the importance of ongoing research and refinement of predictive models to adapt to evolving market dynamics and further enhance the precision of gold price forecasts. Continued efforts in this direction will be instrumental in ensuring the continued relevance and effectiveness of predictive modeling in the realm of gold price forecasting.

Reference

1. Dunis, C., Middleton, P. W., Karathanasopolous, A., & Theofilatos, K. (2016). *Artificial intelligence in financial markets*. London: Palgrave Macmillan.
2. Agrawal, A., Gans, J., & Goldfarb, A. (2022). *Power and prediction: The disruptive economics of artificial intelligence*. Harvard Business Press.
3. Baugi, A., & Zhang, E. (2024). Beyond the Crisis: A Safe Haven Analysis: Empirical Insights into the Divergence of Gold and Bonds for Portfolio Hedging.
4. Huang, S. C., & Wu, C. F. (2018). Energy commodity price forecasting with deep multiple kernel learning. *Energies*, *11*(11), 3029.
5. Cao, Longbing. "Ai in finance: challenges, techniques, and opportunities." *ACM Computing Surveys (CSUR)* 55, no. 3 (2022): 1-38.
6. Khattak, B. H. A., Shafi, I., Khan, A. S., Flores, E. S., Lara, R. G., Samad, M. A., & Ashraf, I. (2023). A Systematic Survey of AI Models in Financial Market Forecasting for Profitability Analysis. *IEEE Access*.

7. Chopra, R., & Sharma, G. D. (2021). Application of artificial intelligence in stock market forecasting: a critique, review, and research agenda. *Journal of risk and financial management*, *14*(11), 526.
8. Rezaei, Rasa, and Zhao Fijin. "Navigating the Golden Seas: A Comprehensive Review of Business Management in the Gold Market." (2024).
9. Kishor, R., Patle, B., & Bhojane, P. (2023). The AI vantage: an exposition of global dynamics. *Journal of Mechatronics and Artificial Intelligence in Engineering*, *4*(2), 58-82.
10. *Livieris, I.E., Pintelas, E. & Pintelas, P. A CNN–LSTM model for gold price time-series forecasting. Neural Compute & Applic* ***32****, 17351–17360 (2020).* *https://doi.org/10.1007/s00521-020-04867-x*
11. *Vidya G.S, Hari V.S,2020 IEEE Recent Advances in Intelligent Computational Systems (RAICS), Doi:* 10.1109/RAICS51191.2020.9332471
12. *K.**L** Manjula, Karthikeyan,* 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Doi:10.1109/ICOEI.2019.8862557
13. Neri, F. (2019). Combining Machine Learning and Agent-Based Modeling for Gold Price Prediction. In: Cagnino, S., Mordonini, M., Pecori, R., Roli, A., Villani, M. (eds) Artificial Life and Evolutionary Computation. WIVACE 2018. Communications in Computer and Information Science, vol 900. Springer, Cham. https://doi.org/10.1007/978-3-030-21733-
14. D Makala and Z Li 2021, International E- Conference on Data Analytics, Intelligent Systems and Information Security & ICDIIS, DOI 10.1088/1742-6596/1767/1/012022
15. https://www.kaggle.com/datasets/aakash013/gold-price-prediction-dataset/code.