FORECASTING GOLD PRICE DIRECTION: MACHINE LEARNING VERSUS TECHNICAL INDICATORS

Dr. Pooja PhD
Senior Lecturer in Computing Department
Westminster International University in Tashkent
Shakhzod Tulkinov MSc
Associate Lecturer in Finance and Accounting Department
Westminster International University in Tashkent

https://doi.org/10.5281/zenodo.21208574

Abstract. *Accurate forecasting of gold price movements remains a challenging task because of the complex and nonlinear nature of financial markets. This study evaluates the effectiveness of traditional technical indicators and ensemble machine learning algorithms for predicting the daily direction of XAU/USD price movements. Historical daily gold price data covering the period from July 2023 to March 2026 were collected and transformed into predictor variables, including the Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), MACD Signal, Bollinger Bands, and calendar-based features. Traditional technical indicator models were first established as baseline benchmarks and subsequently compared with three ensemble learning algorithms: Random Forest, XGBoost, and Gradient Boosting. Model performance was evaluated using a five-fold TimeSeriesSplit cross-validation procedure together with an independent test dataset. The empirical results indicate that ensemble machine learning algorithms generally outperform traditional technical indicator models. Among the baseline models, Bollinger Bands achieved the highest classification accuracy (57.53%), whereas Random Forest produced the strongest overall performance among the machine learning models with an accuracy of 59.47%, followed by Gradient Boosting (55.94%) and XGBoost (54.55%). The cross-validation results further demonstrated the stability of the ensemble learning algorithms across different chronological validation periods. Overall, the findings suggest that machine learning provides a more effective approach than standalone technical indicators for forecasting the direction of gold price movements. The proposed framework contributes to the growing application of artificial intelligence in financial forecasting and offers practical decision-support insights for investors, traders, and financial analysts.*

Keywords: *Gold price forecasting; XAU/USD; Machine learning; Random Forest; XGBoost; Gradient Boosting; Technical indicators; Financial forecasting.*

1. Introduction

Gold (XAU/USD) plays a significant role in the global financial system as both a commodity and a financial asset. Owing to its ability to preserve value during periods of economic uncertainty, inflation, and geopolitical instability, gold is widely regarded as a safe-haven asset and an effective hedge against financial risk (Baur and Lucey, 2010; Baur and McDermott, 2010). Unlike fiat currencies, which are influenced by inflationary pressures and monetary policy decisions, gold has historically maintained its purchasing power and continues to be an essential component of investment portfolios and central bank reserves (Ghosh et al., 2004; Beckmann, Berger and Czudaj, 2015). Consequently, accurate forecasting of gold price movements has become increasingly important for investors, portfolio managers, and financial institutions seeking to improve investment performance and risk management. Technical analysis remains one of the most widely adopted approaches for forecasting financial markets. Among the most commonly used technical indicators are the Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), and Bollinger Bands, which are designed to identify market momentum, trends, and potential price reversals. Despite their popularity among traders, these indicators often provide inconsistent forecasting performance when applied individually because they cannot fully capture the nonlinear, dynamic, and highly volatile behaviour of financial markets. Gold prices are influenced by numerous interacting factors, including market volatility, investor sentiment, macroeconomic conditions, and geopolitical developments, making directional forecasting particularly challenging. Recent advances in machine learning have provided new opportunities for improving financial forecasting by modelling complex nonlinear relationships that conventional statistical and technical analysis methods may fail to capture. Ensemble learning algorithms, such as Random Forest, XGBoost, and Gradient Boosting, have attracted considerable attention because of their ability to learn intricate patterns from historical market data while reducing prediction errors and improving generalisation. Previous studies have reported promising results for machine learning applications in financial forecasting; however, empirical evidence comparing ensemble machine learning algorithms with traditional technical indicator-based approaches for predicting gold price direction remains limited. Furthermore, relatively few studies have examined whether combining multiple ensemble learning techniques within a unified forecasting framework can further improve predictive performance. To address this gap, this study develops and evaluates a machine learning-based forecasting framework for predicting the daily direction of XAU/USD price movements using widely adopted technical indicators as predictor variables. Traditional technical indicator models are first established as baseline benchmarks and subsequently compared with Random Forest, XGBoost, Gradient Boosting, and a hybrid ensemble framework. Model performance is evaluated using five-fold time-series cross-validation and standard classification metrics, including accuracy, precision, recall, and F1-score. By providing a comprehensive comparison between conventional technical analysis and ensemble machine learning approaches, this study contributes to the growing literature on financial forecasting and offers practical insights for investors and financial analysts seeking more reliable data-driven forecasting methods.

2. Literature Review

2.1 Technical Indicators in Financial Forecasting

Technical analysis has long been employed by investors and traders to forecast future price movements based on historical market data. Unlike fundamental analysis, which evaluates intrinsic asset value using economic and financial information, technical analysis assumes that historical price movements and trading patterns contain valuable information about future market behaviour (Brock, Lakonishok and LeBaron, 1992). Among the most widely applied technical indicators are the Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), and Bollinger Bands, each of which captures different aspects of market dynamics, including momentum, trend direction, and volatility. Numerous empirical studies have investigated the forecasting ability of these indicators. Chong and Ng (2008) reported that combinations of RSI and MACD generated statistically significant trading returns in the London Stock Exchange. Similarly, Neely et al. (2014) demonstrated that technical indicators can improve directional forecasting when integrated with complementary analytical techniques. Bollinger Bands have also been recognised as effective tools for identifying periods of increased market volatility and potential price reversals (Sezer, Gudelek and Ozbayoglu, 2020). Collectively, these indicators continue to play an important role in financial forecasting because they transform historical price information into quantitative measures that can be incorporated into predictive models. Despite their widespread application, traditional technical indicators possess several limitations. Since they rely primarily on historical price information and predefined mathematical rules, their forecasting performance often deteriorates during periods of heightened market volatility or structural changes. Gupta (2025) argues that indicators such as RSI and MACD frequently generate false trading signals in rapidly changing market environments, while Lo, Mamaysky and Wang (2000) suggest that rule-based technical indicators are unable to fully capture the complex nonlinear relationships underlying financial markets. Consequently, researchers have increasingly explored machine learning techniques as more flexible alternatives capable of modelling complex financial data.

2.2 Machine Learning in Financial Forecasting

Recent developments in machine learning have transformed financial forecasting by enabling predictive models to identify nonlinear relationships and hidden patterns within large financial datasets. Unlike conventional statistical methods, machine learning algorithms require fewer assumptions regarding data distribution and can automatically learn complex interactions among explanatory variables. As a result, they have become increasingly popular in forecasting stock prices, exchange rates, commodity prices, and other financial assets. Among the most widely applied ensemble learning algorithms are Random Forest, Gradient Boosting, and XGBoost. Random Forest, introduced by Breiman (2001), improves predictive performance by combining multiple decision trees constructed using bootstrap sampling and random feature selection. Gradient Boosting, proposed by Friedman (2001), sequentially builds decision trees that minimise prediction errors from previous iterations, enabling the model to capture highly complex relationships. XGBoost further extends the Gradient Boosting framework through regularisation techniques and computational optimisations that improve both predictive accuracy and efficiency (Chen and Guestrin, 2016). Previous empirical studies consistently demonstrate the advantages of ensemble learning algorithms in financial forecasting. Patel et al. (2015) found that ensemble machine learning models outperformed several conventional classification methods in predicting stock market movements. Similarly, Fischer and Krauss (2018) reported that boosting-based algorithms achieved superior predictive performance for financial time-series forecasting. Gu, Kelly and Xiu (2020) further demonstrated that machine learning techniques effectively capture nonlinear relationships that traditional financial models often fail to identify. These findings suggest that ensemble learning provides a promising approach for improving forecasting accuracy in complex and volatile financial markets.

2.3 Research Gap

Although technical indicators and machine learning algorithms have been widely applied in financial forecasting, several important research gaps remain. First, much of the existing literature focuses on stock markets, cryptocurrencies, or foreign exchange markets, while comparatively fewer studies investigate the application of ensemble learning techniques to gold price direction forecasting. Second, previous research often examines either traditional technical indicators or machine learning models independently, making direct comparisons difficult because different datasets, feature sets, and validation procedures are employed. Consequently, the relative effectiveness of these approaches for forecasting daily XAU/USD price movements remains unclear. To address these gaps, this study evaluates and compares the forecasting performance of three widely used technical indicators, including RSI, MACD, and Bollinger Bands, with three ensemble machine learning algorithms: Random Forest, Gradient Boosting, and XGBoost. Using the same dataset, predictor variables, and five-fold time-series cross-validation procedure, this study provides a consistent empirical comparison of traditional technical analysis and modern machine learning techniques for gold price direction forecasting.

The findings contribute to the financial forecasting literature by identifying the most effective modelling approach for predicting daily gold price movements.

3. Methodology

3.1 Research Design

This study adopts a quantitative research design to evaluate the effectiveness of machine learning algorithms in forecasting the daily direction of gold (XAU/USD) price movements. A comparative experimental approach is employed to assess the predictive performance of traditional technical indicator-based models and ensemble machine learning models. Historical daily gold price data are transformed into technical indicators, which serve as predictor variables for model development. The analysis was conducted using Google Colab, a cloud-based Python development environment. Data preprocessing, feature engineering, model development, and performance evaluation were implemented using widely adopted Python libraries, including Pandas, NumPy, Scikit-learn, XGBoost, and Matplotlib. Model performance was evaluated using a five-fold time-series cross-validation procedure to preserve the chronological structure of the financial data and to ensure robust out-of-sample evaluation.

Figure 1. Research framework of the proposed machine learning approach for gold price direction forecasting.

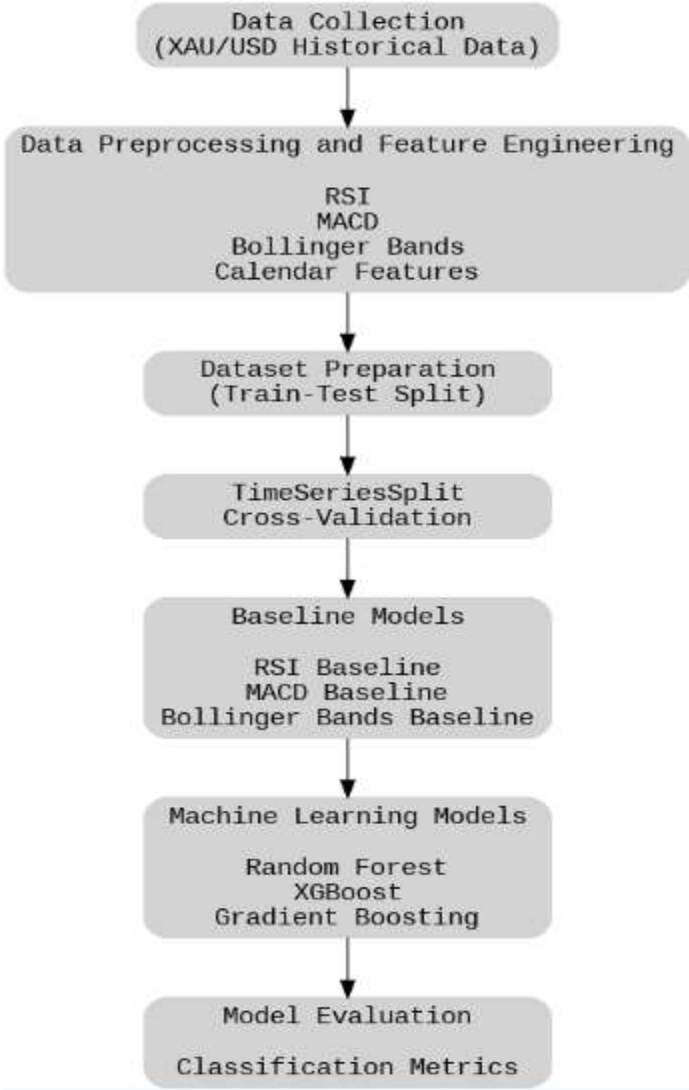

Figure 1 illustrates the overall research framework adopted in this study. The process begins with the collection of historical daily XAU/USD price data, followed by data preprocessing and feature engineering to prepare the dataset for model development. During this stage, technical indicators, including the Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), Bollinger Bands, and calendar-based features, were computed from historical price data and used as predictor variables. The processed dataset was subsequently divided into training and testing subsets using an 80:20 train-test split. To ensure robust model validation while preserving the chronological order of financial time-series data, a five-fold TimeSeriesSplit cross-validation procedure was applied during model development.

Traditional technical indicator models based on RSI, MACD, and Bollinger Bands were first implemented as baseline forecasting approaches. Their predictive performance was then compared with three ensemble machine learning algorithms: Random Forest, XGBoost, and Gradient Boosting. Model performance was evaluated using standard classification metrics, including accuracy, precision, recall, and F1-score. The proposed framework provides a systematic comparison between conventional technical analysis and ensemble machine learning techniques for forecasting the direction of daily gold price movements.

3.2 Data collection

This study utilised secondary data consisting of historical daily XAU/USD (Gold/US Dollar) price observations collected from the Myfxbook historical data platform. The dataset covers the period from 25 July 2023 to 1 March 2026 and includes daily open, high, low, close, and trading activity information. Myfxbook was selected because it provides reliable historical foreign exchange and commodity market data that are widely used by traders and financial analysts. The daily timeframe was selected as it provides an appropriate balance between capturing market dynamics and reducing the excessive noise commonly associated with high-frequency data. Furthermore, daily observations are widely adopted in technical analysis and machine learning studies of financial markets, making them suitable for evaluating the forecasting performance of the proposed models. Although the original dataset contained seven variables (Date, Open, High, Low, Close, Change (Pips), and Change (%)), only the Date and Close price variables were retained for subsequent feature engineering. The remaining explanatory variables were generated through technical indicator calculations based on historical closing prices.

3.3 Data Preprocessing and Feature Engineering

Prior to model development, the dataset underwent a series of preprocessing procedures to improve data quality and ensure compatibility with the machine learning algorithms. Data cleaning was performed using Microsoft Power Query, where missing values, duplicate observations, and inconsistencies were identified and removed where necessary. Non-data rows and irrelevant columns were excluded from the dataset to ensure a consistent tabular structure for subsequent analysis. In addition, the Date variable was standardised by removing the redundant time component and decomposed into three calendar-based variables: Year, Month, and Day.

These variables were included to capture potential seasonal and calendar-related patterns in gold price movements. Following data preprocessing, feature engineering was performed to generate explanatory variables from the historical XAU/USD price series. The feature set consisted of three widely used technical indicators: the Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD) together with its Signal Line, and Bollinger Bands, represented by the upper (BB_High) and lower (BB_Low) bands.

These indicators were selected because they capture complementary aspects of market behaviour, including momentum, trend direction, and price volatility. Together with the calendar-based variables, they formed the input features used for model development.

Figure 2. Data preprocessing and feature engineering workflow.

se	Change(Pips)	Change(%)	Year	Month	Day	RSI	MACD	MACD_Signal	BB_High	BB_Low
32	7064	1.46	2026	1	21	42.526874	-39.946594	-40.529254	5351.199364	4631.875636
55	8845	1.86	2026	1	20	40.555416	-55.554403	-43.534284	5355.820140	4611.429860
50	-703	-0.15	2026	1	19	37.823423	-74.971297	-49.821687	5356.357878	4569.000122
54	3861	0.83	2026	1	18	38.145612	-88.767995	-57.610948	5357.105448	4532.985552
65	-1452	-0.32	2026	1	16	35.954240	-104.699447	-67.028648	5357.244901	4486.592099

Figure 2 illustrates the preprocessing and feature engineering procedures applied to the original dataset. The resulting feature set was subsequently used to develop both the baseline technical indicator models and the ensemble machine learning models evaluated in this study.

3.4 Dataset Preparation

The forecasting task was formulated as a binary classification problem in which the objective was to predict the daily direction of XAU/USD price movements. The target variable was defined using the daily closing price, where a value of 1 represents an upward price movement and a value of 0 represents a downward price movement. This formulation reflects practical trading applications, where correctly identifying the direction of price movement is often more valuable than predicting the exact price level. The explanatory variables consisted of the engineered technical indicators (RSI, MACD, MACD Signal, BB_High, and BB_Low) together with the calendar-based variables (Year, Month, and Day). The dataset was divided into training and testing subsets using an 80:20 train-test split, resulting in 572 observations for model training and 143 observations for out-of-sample testing. This separation enabled an unbiased assessment of model generalisation on previously unseen data.

3.5 Model Development

To evaluate the effectiveness of different forecasting approaches, both traditional technical indicator models and ensemble machine learning algorithms were implemented. The baseline models were constructed using three widely adopted technical indicators: Relative Strength Index (RSI), Moving Average Convergence Divergence (MACD), and Bollinger Bands.

Each indicator generated trading signals according to its conventional decision rules and served as a benchmark for evaluating the performance of the machine learning models. Three ensemble learning algorithms, including Random Forest, XGBoost, and Gradient Boosting, were subsequently developed using the engineered technical indicators and calendar-based variables as predictor variables. These algorithms were selected because of their proven ability to capture complex nonlinear relationships in financial time-series data while providing robust predictive performance.

3.6 Model Validation and Performance Evaluation

To ensure reliable model evaluation and minimize the risk of overfitting, a five-fold TimeSeriesSplit cross-validation procedure was employed during model development.

Unlike conventional random cross-validation methods, TimeSeriesSplit preserves the chronological ordering of financial time-series data by training the models on historical observations and validating them on subsequent periods. This approach more closely reflects real-world forecasting conditions and prevents information leakage from future observations into the training process. Following cross-validation, the final models were evaluated using the independent testing dataset. Model performance was assessed using four standard classification metrics: accuracy, precision, recall, and F1-score. Accuracy measures the overall proportion of correctly classified observations, while precision evaluates the reliability of positive predictions.

Recall measures the ability of a model to correctly identify positive price movements, whereas the F1-score provides a balanced assessment by combining precision and recall into a single metric. The same validation procedure and evaluation metrics were applied to both the baseline technical indicator models and the ensemble machine learning algorithms, ensuring a consistent and objective comparison of their forecasting performance.

4. Results and Discussion

4.1 Five-Fold Time-Series Cross-Validation Results

```
=============5_FOLD TIME-SERIES CROSS-VALIDATION================
Random Forest Mean CV Accuracy:52.94%
XGBoost Mean CV Accuracy50.42%
Gradient Boosting Mean CV Accuracy53.78%
================================================================
=============5_FOLD TIME-SERIES CROSS-VALIDATION================
Random Forest Mean CV Accuracy:45.80%
XGBoost Mean CV Accuracy43.70%
Gradient Boosting Mean CV Accuracy48.32%
================================================================
=============5_FOLD TIME-SERIES CROSS-VALIDATION================
Random Forest Mean CV Accuracy:49.58%
XGBoost Mean CV Accuracy46.78%
Gradient Boosting Mean CV Accuracy49.30%
================================================================
=============5_FOLD TIME-SERIES CROSS-VALIDATION================
Random Forest Mean CV Accuracy:48.53%
XGBoost Mean CV Accuracy46.01%
Gradient Boosting Mean CV Accuracy50.84%
================================================================
=============5_FOLD TIME-SERIES CROSS-VALIDATION================
Random Forest Mean CV Accuracy:50.59%
XGBoost Mean CV Accuracy47.73%
Gradient Boosting Mean CV Accuracy51.43%
================================================================
```

Figure 3 presents the five-fold TimeSeriesSplit cross-validation results for the Random Forest, XGBoost, and Gradient Boosting models. Cross-validation was performed to evaluate model stability and generalisation under different chronological training and validation periods. Across the five validation folds, the ensemble learning algorithms achieved mean classification accuracies ranging from approximately 43% to 54%. Among the evaluated models, Gradient Boosting demonstrated the most consistent performance, achieving the highest accuracy in four of the five validation folds. Its classification accuracy ranged between 48.32% and 53.78%, indicating relatively stable predictive performance across different market periods. Random Forest produced comparable results, with accuracies ranging from 45.80% to 52.94%, whereas XGBoost recorded the lowest performance, achieving accuracies between 43.70% and 50.42%. The average cross-validation accuracy further confirms the superiority of Gradient Boosting over the other ensemble learning algorithms. Although the observed accuracy levels are moderate, they are consistent with the challenges associated with forecasting financial time-series, where price movements are influenced by dynamic market conditions, investor sentiment, and external economic factors. The variation in model performance across validation folds suggests that the predictive relationships within the data evolve over time, reinforcing the importance of time-series cross-validation for obtaining reliable estimates of model performance. Overall, the cross-validation results indicate that Gradient Boosting provides the most stable and reliable forecasting performance among the evaluated machine learning algorithms and was therefore expected to perform strongly during the final out-of-sample evaluation.

4.2 Classification Performance of Baseline Models

NEW RENAISSANCE international scientific journal
ResearchBib IF - 11.01, ISSN: 3030-3753, Volume 3 Issue 7

The classification performance of the three traditional technical indicator models is summarised in Figures 4–6. Overall, the forecasting results indicate that the predictive capability of individual technical indicators is limited when used as standalone forecasting models. Among the evaluated baseline models, Bollinger Bands achieved the highest overall classification accuracy (57.53%), outperforming both RSI (50.41%) and MACD (44.06%). The superior performance of Bollinger Bands suggests that volatility-related information provides greater predictive value for forecasting daily gold price movements than momentum- or trend-based indicators alone. The RSI model exhibited highly imbalanced classification behaviour. Although it achieved perfect precision for downward price movements, its extremely low recall indicates that the model rarely identified bearish market conditions. Instead, the model classified most observations as upward price movements, limiting its practical usefulness as a standalone forecasting tool. Similarly, the MACD model demonstrated the weakest overall performance among the baseline models. Both precision and recall remained relatively low across the two classes, indicating limited ability to distinguish between upward and downward market movements. Overall, these findings suggest that individual technical indicators provide only modest forecasting performance when applied independently. While Bollinger Bands produced comparatively stronger results, none of the traditional baseline models consistently captured the nonlinear dynamics of gold price movements, highlighting the need for more advanced machine learning approaches.

Figures 4 - 6. Classification Performance of RSI, MACD and Bollinger Bands Baseline Models

```
======================RSI BASELINE MODEL===================
Accuracy: 50.41%

Classification Report:
              precision    recall  f1-score   support

           0       1.00      0.03      0.06        63
           1       0.50      1.00      0.66        60

    accuracy                           0.50       123
   macro avg       0.75      0.52      0.36       123
weighted avg       0.75      0.50      0.35       123

============================================================
```

```
==================MACD BASELINE MODEL=====================
Accuracy: 44.06%

Classification Report:
              precision    recall  f1-score   support

           0       0.51      0.41      0.45       404
           1       0.39      0.48      0.43       311

    accuracy                           0.44       715
   macro avg       0.45      0.45      0.44       715
weighted avg       0.45      0.44      0.44       715

============================================================
```

```
====================BOLLINGER BANDS BASELINE====================
Accuracy:57.53%

Classification Report:
              precision    recall  f1-score   support

           0       0.78      0.34      0.47        41
           1       0.51      0.88      0.64        32

    accuracy                           0.58        73
   macro avg       0.64      0.61      0.56        73
weighted avg       0.66      0.58      0.55        73

================================================================
```

4.3 Classification Performance of Machine Learning Models

The classification results of the three ensemble learning algorithms are presented in Figures 7–9. Overall, the machine learning models demonstrated stronger forecasting performance than most of the traditional technical indicator models, confirming their ability to capture complex nonlinear relationships within financial time-series data.

Among the evaluated algorithms, Random Forest achieved the highest overall classification accuracy (59.47%), followed by Gradient Boosting (55.94%) and XGBoost (54.55%).

Random Forest exhibited particularly strong performance in identifying downward price movements, achieving a recall of 0.92 for the negative class. However, its ability to detect upward price movements remained relatively weak, indicating a tendency to favour bearish market predictions. Gradient Boosting and XGBoost produced more balanced classification results across the two classes. Although their overall accuracies were slightly lower than Random Forest, both models demonstrated improved consistency between precision and recall, suggesting a more balanced forecasting capability. The results indicate that ensemble learning algorithms generally outperform traditional technical indicator models because they learn nonlinear interactions among multiple predictor variables rather than relying on predefined trading rules. Nevertheless, forecasting daily gold price direction remains challenging, as reflected by the moderate accuracy levels achieved by all evaluated models.

Figures 7-9 Classification Performance of Machine Learning Models

```
=============== RANDOM FOREST RESULTS ====================
Overall Test Accuracy:  59.44%

Detailed Classification Report:
              precision    recall  f1-score   support

           0       0.61      0.92      0.73        86
           1       0.46      0.11      0.17        57

    accuracy                           0.59       143
   macro avg       0.53      0.51      0.45       143
weighted avg       0.55      0.59      0.51       143

============================================================
```

```
====================XGBOOST RESULTS ======================
Overall Test Accuracy:  54.55%

Deatiled Classification Report:
              precision    recall  f1-score   support

           0       0.62      0.65      0.63        86
           1       0.42      0.39      0.40        57

    accuracy                           0.55       143
   macro avg       0.52      0.52      0.52       143
weighted avg       0.54      0.55      0.54       143

============================================================
```

```
Gradient Boosting Results
Accuracy: 55.94%
Classification Report:
              precision    recall  f1-score   support

           0       0.62      0.69      0.65        86
           1       0.44      0.37      0.40        57

    accuracy                           0.56       143
   macro avg       0.53      0.53      0.53       143
weighted avg       0.55      0.56      0.55       143
```

4.4 Comparative Analysis

Figure 10. Comparison of Baseline and Machine Learning Models

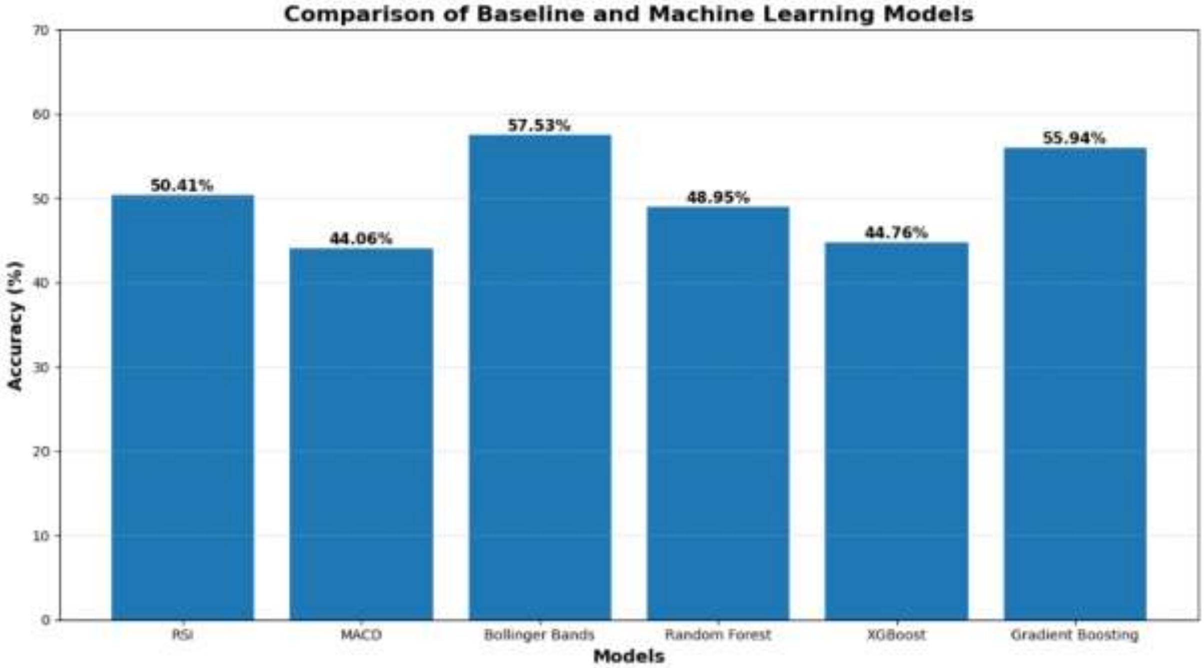

Figure 10 summarizes the classification performance of all forecasting models evaluated in this study. Among the traditional technical indicator models, Bollinger Bands achieved the strongest performance with an accuracy of 57.53%, whereas RSI and MACD achieved accuracies of 50.41% and 44.06%, respectively. Among the machine learning algorithms, Random Forest produced the highest overall classification accuracy (59.47%), marginally outperforming Bollinger Bands by approximately 2 percentage points. Gradient Boosting and XGBoost also achieved competitive performance, although both recorded slightly lower accuracies than Random Forest.

These findings suggest that machine learning algorithms provide measurable improvements over conventional technical analysis for forecasting gold price direction. Unlike traditional technical indicators, which rely on fixed trading rules, ensemble learning algorithms can exploit complex nonlinear relationships among multiple market variables, enabling more accurate classification of future price movements. Although the improvement in predictive accuracy is relatively modest, the consistent superiority of ensemble learning models demonstrates their potential for supporting data-driven investment decision-making in the gold market.

5. Legal and Ethical Considerations

This study used publicly available historical XAU/USD market data obtained from recognised financial data platforms. No proprietary, confidential, or personal data were collected or processed; therefore, ethical approval was not required. All data sources, preprocessing procedures, modelling techniques, and evaluation methods are transparently reported to ensure reproducibility and compliance with accepted standards of academic integrity. The study was conducted independently without financial sponsorship or conflicts of interest.

6. Limitations

Several limitations should be acknowledged. First, the proposed forecasting framework relies exclusively on historical price data and technical indicators and therefore does not incorporate macroeconomic variables, such as inflation, interest rates, exchange rates, or geopolitical events, which may substantially influence gold price movements.

Second, the analysis is based on a specific historical period (July 2023 to March 2026), and the findings may not fully generalise to different market conditions or economic cycles.

Third, although a five-fold TimeSeriesSplit cross-validation procedure was employed to reduce overfitting, the relatively limited sample size may constrain the robustness of the machine learning models. Finally, the proposed framework is static and does not adapt to real-time changes in market conditions. Future research should incorporate macroeconomic and sentiment-based variables, utilise longer historical datasets, and investigate adaptive forecasting models capable of continuously updating their predictions.

7. Conclusion

This study evaluated the effectiveness of traditional technical indicators and ensemble machine learning algorithms for forecasting the daily direction of gold (XAU/USD) price movements. Using historical market data, technical indicators including RSI, MACD, MACD Signal, and Bollinger Bands were engineered as predictor variables and applied to both traditional rule-based forecasting models and three ensemble learning algorithms: Random Forest, XGBoost, and Gradient Boosting. The empirical results demonstrate that ensemble machine learning algorithms generally outperform traditional technical indicator models.

Among the baseline approaches, Bollinger Bands achieved the highest classification accuracy (57.53%), whereas Random Forest produced the strongest overall performance among the machine learning models with an accuracy of 59.47%. The five-fold TimeSeriesSplit cross-validation results further demonstrated that the ensemble learning models provided relatively stable forecasting performance across different chronological validation periods. Overall, the findings indicate that machine learning algorithms are more effective than standalone technical indicators for predicting the direction of daily gold price movements.

Although the improvement in predictive accuracy is moderate, the results highlight the potential of ensemble learning techniques as practical decision-support tools for investors and financial analysts. Future research should investigate the integration of macroeconomic variables, market sentiment indicators, and adaptive learning techniques to further improve forecasting performance.

8. Conflict of Interest

The author declares that there are no financial, commercial, or personal relationships that could have influenced the research reported in this study. No conflicts of interest exist regarding the publication of this manuscript.

9. Data Availability Statement

The historical XAU/USD price data used in this study were obtained from publicly available financial data sources. The processed dataset and Python code used for data preprocessing, model development, and performance evaluation are available from the author upon reasonable request.

References

1. Amini, M. and Kalantari, M. (2024) ‘Gold price prediction using deep learning models: A hybrid CNN–BiLSTM approach’, *Expert Systems with Applications*, 237, 121615.
2. Baur, D.G. and Lucey, B.M. (2010) ‘Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold’, *Financial Review*, 45(2), pp. 217–229.
3. Baur, D.G. and McDermott, T.K. (2010) ‘Is gold a safe haven? International evidence’, *Journal of Banking & Finance*, 34(8), pp. 1886–1898.
4. Beckmann, J., Berger, T. and Czudaj, R. (2015) ‘Does gold act as a hedge or a safe haven? New evidence from time-varying coefficient models’, *Economic Modelling*, 48, pp. 16–24.
5. Bollinger, J. (2001) *Bollinger on Bollinger Bands*. New York: McGraw-Hill.

6. Breiman, L. (2001) ‘Random forests’, *Machine Learning*, 45(1), pp. 5–32.
7. Brock, W., Lakonishok, J. and LeBaron, B. (1992) ‘Simple technical trading rules and the stochastic properties of stock returns’, *The Journal of Finance*, 47(5), pp. 1731–1764.
8. Chen, T. and Guestrin, C. (2016) ‘XGBoost: A scalable tree boosting system’, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794.
9. Chong, T.T.L. and Ng, W.K. (2008) ‘Technical analysis and the London Stock Exchange: Testing the MACD and RSI rules using the FT30’, *Applied Economics Letters*, 15(14), pp. 1111–1114.
10. Fischer, T. and Krauss, C. (2018) ‘Deep learning with long short-term memory networks for financial market predictions’, *European Journal of Operational Research*, 270(2), pp. 654–669.
11. Friedman, J.H. (2001) ‘Greedy function approximation: A gradient boosting machine’, *The Annals of Statistics*, 29(5), pp. 1189–1232.
12. Ghosh, D., Levin, E.J., Macmillan, P. and Wright, R.E. (2004) ‘Gold as an inflation hedge?’, *Studies in Economics and Finance*, 22(1), pp. 1–25.
13. Gu, S., Kelly, B. and Xiu, D. (2020) ‘Empirical asset pricing via machine learning’, *The Review of Financial Studies*, 33(5), pp. 2223–2273.
14. Hoseinzade, E. and Haratizadeh, S. (2019) ‘CNN-based stock market prediction using technical indicators’, *Expert Systems with Applications*, 136, pp. 46–57.
15. Jiang, W. (2021) ‘Applications of machine learning in stock market prediction: A survey’, *Journal of Physics: Conference Series*, 1861(1), 012021.
16. Kim, K.J. (2003) ‘Financial time series forecasting using support vector machines’, *Neurocomputing*, 55(1–2), pp. 307–319.
17. Konur, O., Göçken, M. and Dosdoğru, A.T. (2024) ‘The role of technical indicators in machine learning-based financial forecasting’, *Journal of Open Innovation: Technology, Market, and Complexity*, 10(2), 100215.
18. Krauss, C., Do, X.A. and Huck, N. (2017) ‘Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500’, *European Journal of Operational Research*, 259(2), pp. 689–702.
19. Krollner, B., Vanstone, B. and Finnie, G. (2010) ‘Financial time series forecasting with machine learning techniques: A survey’, *European Symposium on Artificial Neural Networks*, pp. 389–394.
20. Lo, A.W., Mamaysky, H. and Wang, J. (2000) ‘Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation’, *The Journal of Finance*, 55(4), pp. 1705–1765.
21. Natekin, A. and Knoll, A. (2013) ‘Gradient boosting machines, a tutorial’, *Frontiers in Neurorobotics*, 7(21), pp. 1–21.
22. Neely, C.J., Rapach, D.E., Tu, J. and Zhou, G. (2014) ‘Forecasting the equity risk premium: The role of technical indicators’, *Management Science*, 60(7), pp. 1772–1791.
23. Nguyen, T.H. (2024) ‘Forecasting gold price using technical analysis indicators’, *International Journal of Financial Studies*, 12(1), 15.
24. Patel, J., Shah, S., Thakkar, P. and Kotecha, K. (2015) ‘Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques’, *Expert Systems with Applications*, 42(1), pp. 259–268.
25. Sezer, O.B., Gudelek, M.U. and Ozbayoglu, A.M. (2020) ‘Financial time series forecasting with deep learning: A systematic literature review’, *Applied Soft Computing*, 90, 106181.
26. Wilder, J.W. (1978) *New Concepts in Technical Trading Systems*. Greensboro, NC: Trend Research.
27. Zhang, X., Aggarwal, V. and Qi, G. (2017) ‘Stock price prediction via discovering multi-frequency trading patterns’, *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 2141–2149.