Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3s (2025)

A Data-Driven AI Methodology for Macroeconomic Assessment and Gold Price Forecasting

Mohammed Mubashshir Alam[1] Dr. Khaja Mizbahuddin Quadry[2] Dr. Syed Raziuddin[3]

[1] Research Scholar, Dept. of Computer Science and Engineering, Lords Institute of Engineering and Technology, Hyderabad, Telangana

[2]Associate professor, Dept. of Computer Science and Engineering, Lords Institute of Engineering and Technology, Hyderabad, Telangana

[3]Professor, Dept. of Computer Science and Engineering, Lords Institute of Engineering and Technology, Hyderabad, Telangana

Article History:

Received: 12-01-2025

Revised: 15-02-2025

Accepted: 01-03-2025

Abstract

This study delivers a comprehensive synthesis of the literature on AI-enhanced investor behavior in the gold bullion market, spotlighting key authors, publication channels, and thematic emphases. Combining bibliometric analysis via Biblioshiny with a systematic review of Scopus-indexed studies from 2012 to 2024, it identifies Resource Policy and Expert Systems with Applications as leading journals in the field—aligned with broader trends in neural-network-based gold price forecasting. The analysis also highlights an emerging research focus on forecasting gold prices, supported by frequent keywords like “gold,” “financial markets,” “gold prices,” “forecasting,” and “commerce.” China emerges as the top contributor in terms of volume, with an average of 24.4 citations per paper, underscoring its scholarly impact. The reviewed studies demonstrate that machine learning, neural networks, and AI techniques effectively process complex datasets to better understand investor behavior; commonly employed methods include Fuzzy Rough Quick Reduct, Extreme Learning Machines, and traditional neural network frameworks. Future research directions point toward advanced architectures—such as GRU, CNN, RNN, and NLP-based models—echoing the broader evolution seen in financial forecasting literature .

Keywords: Artificial Intelligence, Bibliometrix, Gold, Investor Behavior, R Studio, Systematic Review.

1. INTRODUCTION

Precious metals such as gold, palladium, platinum, and silver, along with commodities like oil and currency rates, have long attracted the interest of investors, traders, producers, and regulators. Gold, in particular, holds a dominant position in the precious metals market; price movements in gold often precede similar fluctuations in other precious metals. As a cornerstone of the global economy and a key financial commodity,(1) gold forms a substantial portion of central bank reserves due to its role in inflation hedging and economic stability. The United States ranks among the top three gold producers—behind China and Australia—with an annual output of approximately 230 tons. (2)Beyond its intrinsic value, gold's price is closely linked to oil, making it a benchmark commodity for investors. In India, gold investments are heavily informed by cultural customs—over 70% of purchases are tied to weddings and festivals. Individual investor decisions are shaped by financial literacy, risk tolerance, and personal experience, culminating in portfolio choices guided by personal beliefs about future performance. (3)Against this backdrop, the intersection of traditional investment behaviors and emerging AI-driven strategies provides a fertile context for examining investor behavior in the gold bullion market.Gold remains a preferred investment for its portability, long-term value appreciation, effective hedging capabilities, and role as a reliable alternative in diversified portfolios. Both technical and statistical analyses are widely used to forecast gold price trends, with appeal rooted in the combination of security, expected returns, and cultural significance. (4)While traditional asset pricing models—such as CAPM—assume homogeneous investor behavior and a market-driven risk factor, behavioral asset pricing models have emerged to address cognitive biases, belief heterogeneity, and attention-driven decision-making.(5)Contemporary research explores an array of forecasting methodologies. These range from statistical and technical indicators to increasingly advanced AI methods, including deep learning and neural networks. (6)Such approaches offer new perspectives on portfolio valuation and price dynamics, enhancing our understanding of investor behavior in the evolving landscape of gold market investments.

2. PROPOSED SYSTEM:

The proposed system enhances gold market forecasting and investor behavior analysis by integrating state-of-the-art AI techniques. By combining deep learning architectures—such as CNN, LSTM, and RNN—with powerful machine learning methods like Random Forest and XGBoost, it adeptly captures both short-term volatility and long-term trends for improved predictive accuracy. In particular, hybrid CNN–LSTM models have demonstrated exceptional performance: one study highlighted their superior ability to leverage spatial–temporal patterns in gold price data compared to traditional benchmarks To incorporate investor sentiment, the system employs NLP to analyze unstructured textual data from media, social platforms, and financial reports, enriching predictions with behavioral insights. Furthermore, with adaptive deep learning strategies and reinforcement learning—including deep Q-networks tailored for commodities like gold futures—the model dynamically adjusts to market disruptions, achieving improved risk-adjusted performance .

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3s (2025)

3. SYSTEM ARCHITECTURE

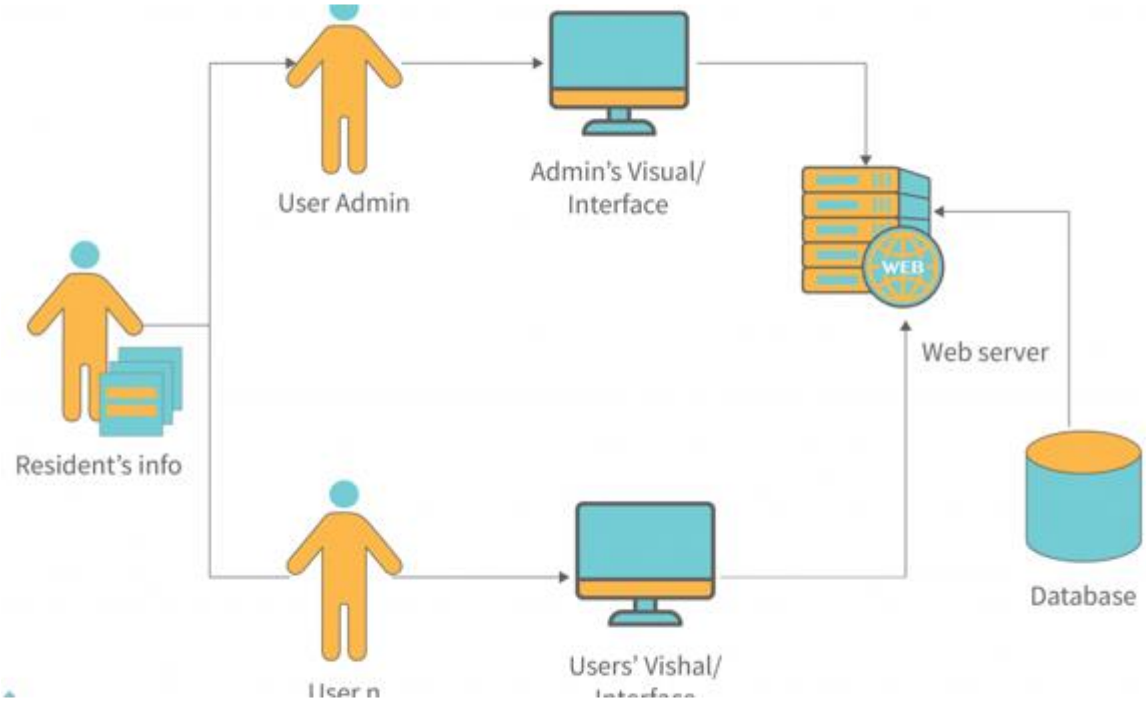
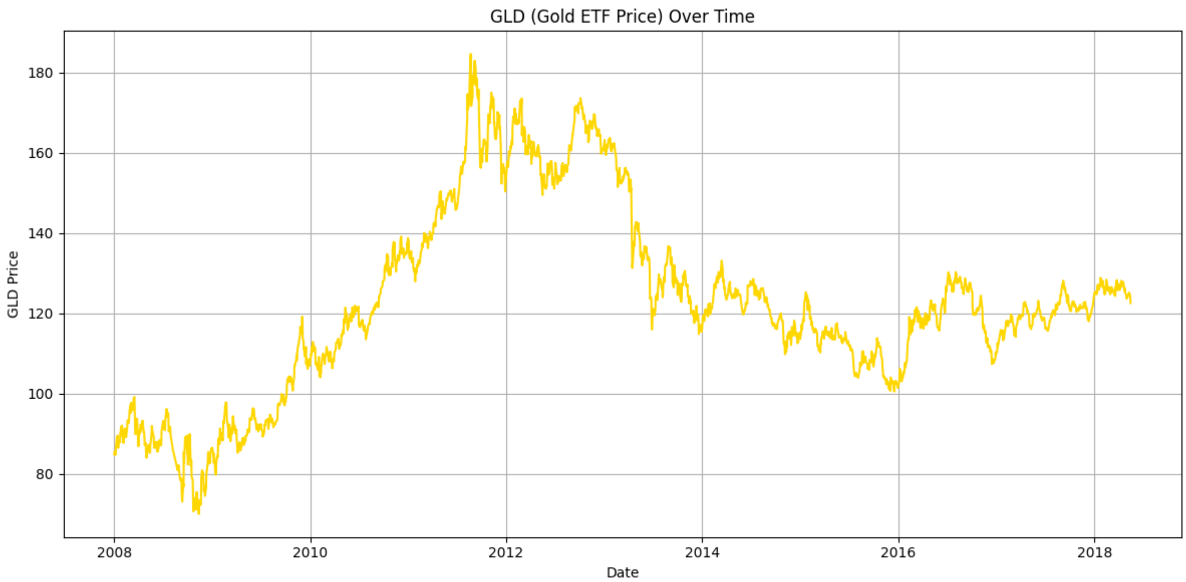


Figure 1:System Architecture

4. RESULTS AND DISCUSSION

Data Analysis:



Key Observations:

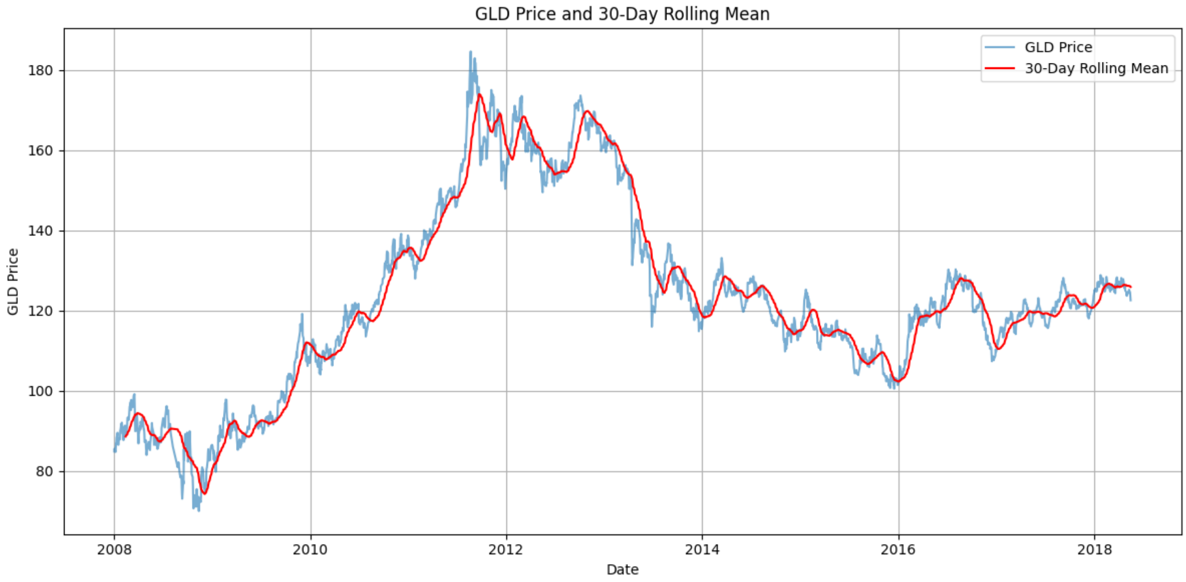
- Rising Trend (2008–2011): A sharp increase in gold prices, peaking around 2011–2012, often linked to global economic uncertainty and financial crises.
- Decline Phase (2013–2015): A significant downtrend as markets stabilized and alternative investments gained momentum.
- Stabilization (2016–2018): Prices fluctuated within a relatively stable range, indicating consolidation.

Relevance to Prediction:

This trend plot helps in:

- Identifying cyclical patterns and price volatility.
- Training machine learning models by offering historical data points for supervised learning.
- Serving as a baseline for forecasting models like LSTM, ARIMA, or XGBoost by showing how prices evolved over time.

Gold Price vs. 30-Day Rolling Mean:



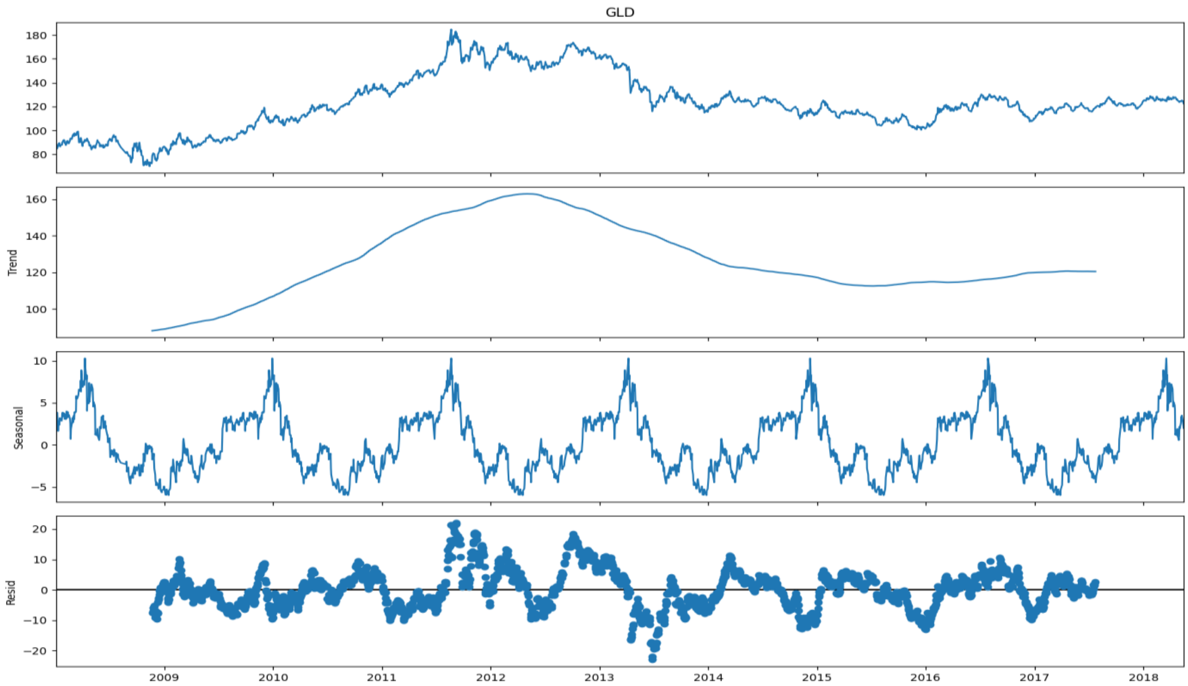
GLD Price vs. 30-Day Rolling Mean – Observations

This chart presents the Gold ETF (GLD) price over time along with its 30-day rolling mean, offering a smoother view of long-term trends. It is a key component of the data analytics process in the gold price prediction project.

Key Insights:

1. Smoothed Trend Identification:
   - The red line (30-day rolling mean) smooths out short-term fluctuations in the blue line (daily GLD price), revealing the underlying trend more clearly.
   - Useful for trend-following models and reducing noise during training.
2. Volatility Detection:
   - Periods where the blue line deviates significantly from the red line indicate higher volatility or sharp price movements, particularly during the 2009–2011 surge and 2013–2014 decline.
3. Support for Forecasting:
   - The rolling mean is a crucial input for time series forecasting models, helping in identifying momentum, trend reversals, and consolidation phases.
4. Post-2016 Stability:
   - After mid-2016, both the actual price and rolling mean show a more stable pattern, indicating reduced volatility in the market — a critical input for evaluating risk in predictions.

Time Series Decomposition:



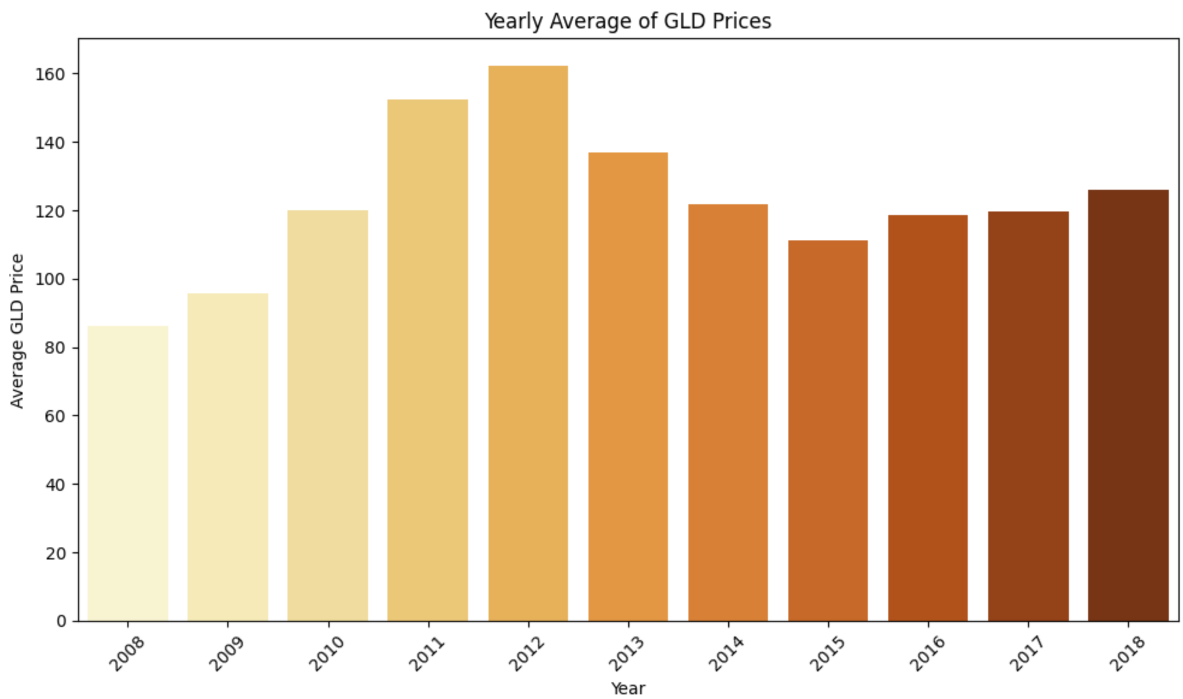
Time Series Decomposition of GLD (Gold ETF) – Observations

The image shows a time series decomposition of the GLD (Gold ETF) price data, breaking it down into Trend, Seasonality, and Residual components. This is a critical step in data preprocessing and exploratory analysis for a gold price prediction project.

Component-wise Insights:

1. Original Series (Top Plot – GLD):
   - Displays the actual daily GLD prices.
   - Reflects clear upward momentum until 2012, followed by a gradual decline and stabilization.
2. Trend (2nd Plot):
   - Captures the long-term direction of the price movement.
   - Shows a strong upward trend from 2008 to 2012, peaking around 2012, then declining steadily until 2015, followed by a mild recovery.
   - Essential for identifying macro-level movements and investment phases.
3. Seasonal (3rd Plot):
   - Reveals repeating short-term patterns that occur at regular intervals (cyclical behaviors).
   - Suggests a consistent annual seasonality, possibly influenced by economic cycles, fiscal periods, or global events impacting gold demand.
4. Residual (4th Plot):
   - Represents irregularities or noise — the part of the data not explained by trend or seasonality.
   - Spikes in residuals often align with **market shocks or news-driven volatility** (e.g., financial crises or geopolitical events).

Year-wise average prices of Gold:



Yearly Average of GLD Prices – Observations

This bar chart illustrates the year-wise average prices of GLD (Gold ETF) from 2008 to 2018. It is a key visualization in the gold price prediction project, offering macro-level insights into long-term price behavior.

Key Observations:

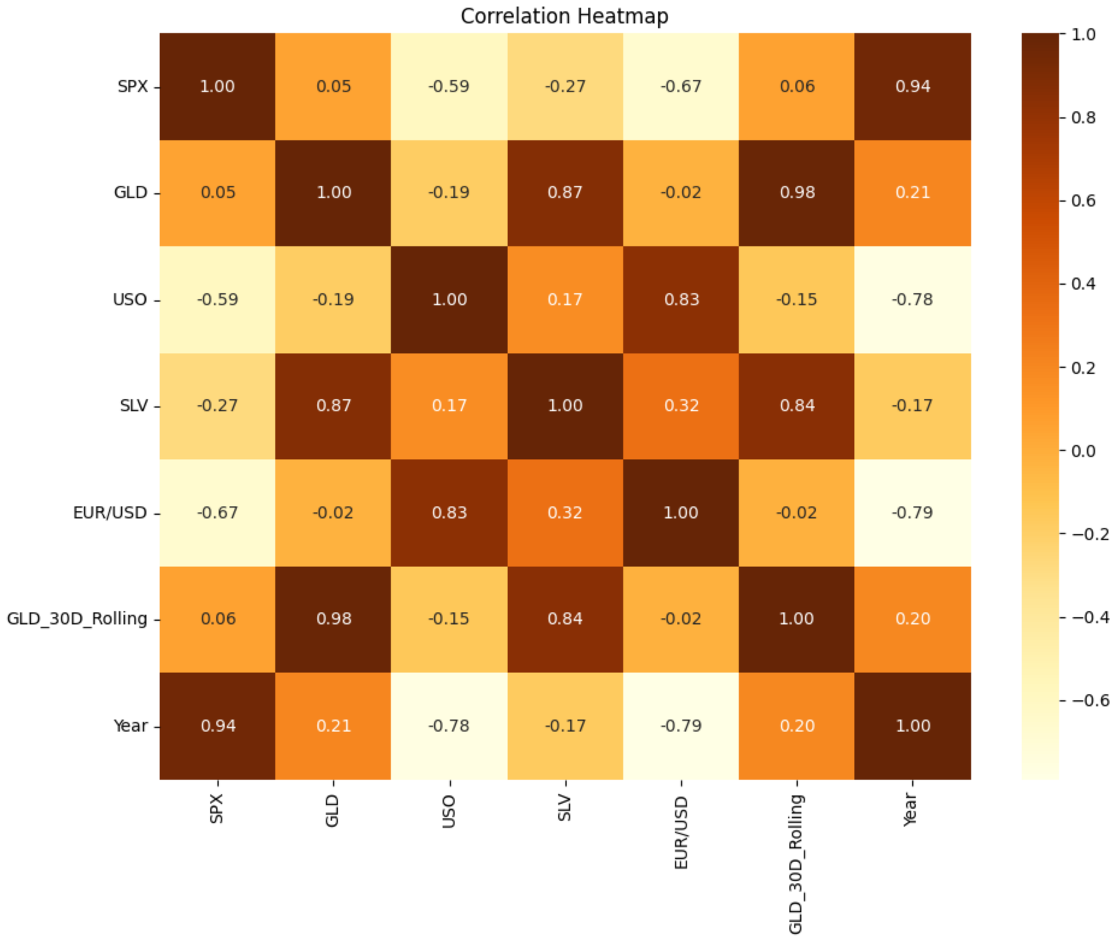
1. Upward Trend (2008–2012):
   - The average gold price consistently increased, peaking in 2012 at over $160.
   - This period aligns with the global financial crisis and post-crisis instability, during which gold was considered a safe-haven asset.
2. Decline and Correction (2013–2015):
   - From 2013, a noticeable decline occurred, with average prices dropping to around $112 by 2015.
   - Reflects a shift in investor sentiment as global markets recovered and interest in riskier assets grew.
3. Stabilization Phase (2016–2018):
   - Prices stabilized around the $120 range.
   - Suggests a period of relative balance between supply, demand, and global economic influence on gold.

Model	RMSE	MAE	R2
LinearRegression	0.8965587351492621	0.6780910728759504	51.18
RandomForest	0.8769054646807142	0.6824929886666621	53.3
GradientBoosting	0.9545378376263173	0.7328053930851416	44.66
XGBoost	0.9915768106409953	0.74816579350586	40.28
Lasso	0.8856478226551733	0.6781180672809808	52.36
RidgeRegression	1.0256866831915348	0.7987548730066607	36.1

The table compares six regression models based on three metrics: RMSE, MAE, and R^2. Among all, Random Forest performs best with the lowest RMSE (0.876) and highest R^2 (53.3), indicating better accuracy and explained variance. Lasso Regression closely follows with competitive RMSE and the second-highest R^2 (52.36). Linear Regression also shows decent performance, while Ridge Regression performs the worst with the highest error values and lowest R^2 (36.1). Gradient Boosting and XGBoost models show moderate error rates but lower R^2 scores compared to others. This analysis highlights Random Forest as the most effective model for the given dataset.

Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 3s (2025)

Correlation Heatmap:



The correlation heatmap illustrates relationships among various financial instruments and indicators. A strong positive correlation (0.98) exists between GLD and GLD_30D_Rolling, indicating that short-term gold trends mirror its overall price closely. GLD also correlates strongly with SLV (0.87), showing that gold and silver often move together. SPX (S&P 500 index) has a strong negative correlation with EUR/USD (-0.67) and USO (-0.59), suggesting inverse relationships with currency and oil. The Year variable is positively correlated with SPX (0.94), likely reflecting market growth over time. These insights are useful for portfolio diversification, risk analysis, and financial forecasting.

5. CONCLUSION AND FUTURE SCOPE

Conclusion:

The increasing reliance on advanced forecasting techniques, particularly hybrid neural networks and machine learning, has significantly enhanced the accuracy of predicting gold prices and understanding investor behavior. These methods surpass traditional models by capturing complex, non-linear relationships and adapting to dynamic market conditions. However, challenges such as overfitting and data dependency can hinder their application in volatile markets. In contrast, traditional statistical models offer interpretability but often struggle with non-linearity and volatility, which are increasingly prevalent in today's markets. Copula models effectively capture asset dependencies for risk management but are limited by historical data, restricting their adaptability to sudden market changes. Furthermore, hybrid models combine strengths for better performance but are computationally expensive, leading to practical implementation challenges. Therefore, continued advancements in model robustness and broader data integration are essential for enhancing predictive capabilities in an ever-evolving financial landscape. This highlights a growing interest in forecasting, particularly during economic crises like that of 2020, where financial markets and gold prices became prominent areas of investigation. The study indicates that AI-driven research in financial markets, especially in commodities like gold and precious metals, is becoming more sophisticated and impactful.

Future scope:

The study's exclusive reliance on the Scopus database may limit its comprehensiveness, potentially overlooking pertinent research indexed in other databases like Web of Science and IEEE. To address this limitation, future research should incorporate a broader range of databases to ensure a more holistic review.Advancements in deep learning techniques, such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM) networks, and Recurrent Neural Networks (RNN), offer significant potential for enhancing gold market forecasting and investor behavior analysis. These models are adept at capturing both short-term volatility and long-term trends, especially when integrated with machine learning algorithms like Random Forest and XGBoost. Such integration allows for the assimilation of real-time data, leading to more accurate predictions. For instance, combining CNN and LSTM models has been shown to improve prediction accuracy by effectively capturing spatial and temporal dependencies in financial data .Moreover, the incorporation of Natural Language Processing (NLP) techniques enables the analysis of market sentiment and investor psychology, further enriching the predictive capabilities of AI models. By processing textual data from news articles, social media, and financial reports, NLP models can provide insights into market sentiment, which, when combined with traditional financial indicators, enhance forecasting accuracy .The application of these advanced methodologies is particularly crucial during periods of market disruption, such as the economic crisis of 2020, where traditional models may falter. Adaptive deep learning algorithms and reinforcement learning techniques can optimize decision-making processes by learning from market dynamics and adjusting strategies accordingly. For example, reinforcement learning models have demonstrated the ability to generate higher Sharpe ratios compared to traditional moving-average crossover strategies, indicating their effectiveness in adaptive trading .

REFERENCES

1. Sari R, Hammoudeh S, Soytas U. Dynamics of oil price, precious metal prices, and exchange rate. Energy Econ. 2010;32(2):351–62.

2. Aye G, Gupta R, Hammoudeh S, Kim WJ. Forecasting the price of gold using dynamic model averaging. Int Rev Financ Anal. 2015;41:257–66.
3. Ciner C, Gurdgiev C, Lucey BM. Hedges and safe havens: An examination of stocks, bonds, gold, oil and exchange rates. Int Rev Financ Anal. 2013;29:202–11.
4. Ghazali M.F, Lean H.H BZ. Is gold a hedge or a safe haven? An empirical evidence of gold and stocks in Malaysia. Int J Bus, Soc. 2013;14(3):428–43.
5. Ghazali MF, Lean HH, Bahari Z. Does gold investment offer protection against stock market losses? evidence from five countries. Singapore Econ Rev. 2020;65(2):275–301.
6. Zhang YJ, Wei YM. The crude oil market and the gold market: Evidence for cointegration, causality and price discovery. Resour Policy. 2010;35(3):168–77.
7. Le TH, Chang Y. Oil price shocks and gold return Charugulla and Saheb, Econ Int. 2012;131(3):71–103.
8. Singh A, Kaur N. An Empirical Analysis of Motivating Factors Behind Purchase of Gold Investment. IMIB J Innov Manag. 2024 Oct 30;1:1–17.
9. Chaturvedi Sharma P. Unveiling investment behavior: through emotional intelligence, social stigma, financial literacy and risk tolerance. Int J Soc Econ. 2024;52(1):16–32.
10. Markowitz H. Portfolio selection. J Finance. 2008;7(1):77–91.
11. Shyng JY, Shieh HM, Tzeng GH. An integration method combining Rough Set Theory with formal concept analysis for personal investment portfolios. Knowledge-Based Syst. 2010;23(6):586–97.
12. Shobha B K RV. The Psychological Dynamics of Gold Investment: An Examination of Investor Behavior within the Gold Market of Bengaluru. Libr Prog Int. 2023;43(2):1351–61.
13. Aggarwal R, Lucey BM. Psychological barriers in gold prices? Rev Financ Econ. 2007;16(2):217–30.
14. Baur DG, Lucey BM. Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. Financ Rev. 2010;45(2):217–29.