Bitlis Eren Üniversitesi Fen Bilimleri Dergisi
BİTLİS EREN UNIVERSITY JOURNAL OF SCIENCE
ISSN: 2147-3129/e-ISSN: 2147-3188
VOLUME: 13 NO: 4 PAGE: 1293-1303 YEAR: 2024
DOI: 10.17798/bitlisfen.1556171

Assessing the Effectiveness of Machine Learning Techniques for Silver Price Prediction: A Comparative Study*

Erhan ERGİN[1**], Binali Selman EREN[2]

[1]*Bitlis Eren University, Faculty of Economics and Administrative Sciences, Department of Business Administration, Bitlis, Türkiye.*
[2]*Bitlis Eren University, Vocational School of Tatvan, Department of Accounting and Tax, Bitlis, Türkiye.*
(ORCID: 0000-0001-6281-3654) (ORCID: 0000-0001-5136-6406)

Keywords: Silver Price, Financial Prediction, Machine Learning, Prediction Algorithms.

Abstract
Silver is considered an important asset in terms of economic indicators and a valuable investment asset in terms of the markets. Therefore, determining silver prices is critically important for both national economies and investors. However, the non-stationary and non-linear nature of silver prices makes predicting price movements challenging. The methods used for predicting silver prices must be suitable for capturing these volatile and complex behavioral characteristics. The silver market can be influenced by other commodities and investment assets. Factors affecting silver prices, such as gold prices, Brent crude oil prices, the US Dollar index, the VIX index, and the S&P 500 index, can play a significant role. In this context, these variables have been used as inputs for predicting silver prices in the study. Three different models have been developed to predict the prices one, two, and three days ahead. These models have been predicted using four different machine learning methods: linear regression, support vector regression (SMOReg), k-nearest neighbors (k-NN), and random forest (RF). The results show that the random forest and k-NN methods exhibit the highest performance. The random forest achieves the highest accuracy in the first two models, while k-NN excels in the third model. Linear regression and SMOReg methods are less successful compared to the others. Consequently, it can be concluded that random forest and k-NN methods can be preferred for long-term predictions, and that these results may provide valuable insights, especially for investors and decision-makers.

1. Introduction

For many years, individuals have utilized precious metals as a means of saving and investing, thereby making them a crucial component of the global financial system. The financialization of commodities markets permitted investors to diversify their portfolios and resulted in a considerable increase in investments associated with precious metals [1]. In both developed and developing economies, financial products employ precious metals as indispensable assets for industrial processes. Consequently, commodities play an important role in economic growth and development. Precious metals serve as a reliable source of stability during periods of political and economic uncertainty. Gold, silver, platinum, and palladium, in particular, are essential inputs in industrial production. This has led to a heightened interest among investors in these metals within the context of financial markets.

* This paper is an expanded and revised version of the abstract presented as an oral presentation at the 2nd International Rahva Conference on Technical and Social Research, held on December 3-4, 2022.

** Corresponding author: eergin@beu.edu.tr *Received: 25.09.2024, Accepted: 17.11.2024*

Silver is a precious metal of considerable importance to both the global financial system and the industrial production system. In addition to its status as an investment vehicle and a store of value, silver is employed as a precious metal in a number of other contexts, including industrial production, art, medicine, chemistry, and photography. In addition to its role as a store of value against economic uncertainty, inflation, and exchange rate fluctuations, silver provides portfolio diversification for investors and facilitates trading transactions due to its high liquidity. These features have made silver a valuable financial instrument for both individual and institutional investors [2].

Throughout history, the price of silver has been influenced by a variety of economic and social factors. Historically, medieval Europe relied heavily on silver, a precious metal, as an essential economic resource. The Industrial Revolution of the nineteenth century and the "Silver Break" of 1873 had a significant impact on the price of silver. In the 20th century, global political and economic developments rendered silver a valuable asset. During the "Silver Crisis" of 1980, extensive silver purchases led to regulations that artificially inflated prices, ultimately causing a rapid decline. This crisis revealed the effects of speculative bubbles and market manipulations, leading to tighter financial regulations. Subsequently, silver prices increased significantly due to factors such as economic uncertainties, inflation concerns, rising investment demand, increased industrial usage, and the impact of the 2008 global economic crisis. In recent years, global financial crises and the COVID-19 pandemic have impacted the value and price of silver. Figure 1 also shows the historical fluctuations in silver prices.

Figure 1. Silver prices (1970-2024)

Source: tradingeconomics.com

Because of these characteristics, market participants closely monitor silver prices. Predicting silver prices is crucial for trading decisions, portfolio management, industrial production planning, economic indicators, and the overall health of financial markets. The purpose of this study is to predict the price of silver using a variety of variables, including the US Dollar Index, VIX Index, S&P 500 Index, Brent Crude Oil, and gold prices. The central research question is to identify the most effective machine learning method for predicting silver prices with the minimum prediction error. Previous studies in this field have employed a range of methods, including linear regression, support vector regression (SMOReg), random forest, and K-Nearest Neighbors (k-NN), for financial prediction. A review of the literature reveals a limited number of studies on the use of machine learning methods for predicting silver prices. It is expected that the techniques and variables utilized in this study will enhance and broaden the existing body of research. The structure of the subsequent sections is as follows: The second section provides a summary of relevant literature; the third section details the dataset and method explanations; the fourth section presents the experimental results; and the final section discusses the conclusions and implications.

2. Literature Review

In recent years, machine learning techniques have been extensively utilized in the field of financial prediction. The literature in this field is growing rapidly, with new studies contributing to a continually expanding body of knowledge. A significant number of studies have employed machine learning techniques to predict the prices of precious metals. However, it has been observed that the number of studies that predict silver prices with machine learning methods is limited. This section presents an overview of studies that predict silver prices and prices of other precious metals using machine learning methods, with a particular focus on studies conducted in recent years.

Çelik and Başarır (2017) predict the prices of precious metals such as gold, silver, platinum, and palladium using artificial neural networks (ANN). They use five performance metrics to assess the accuracy of the predictions. The study analyzes five years of various financial data. The results show that the error rates in market predictions are at acceptable levels [3].

Goel et al. (2022) use machine learning methods to predict gold and silver prices in the Indian market. They conducted the analysis using CNN and

CNN-RNN hybrid models, utilizing data from January 2021 to August 2022. The results show that the RNN model is only successful in gold price prediction, while other models generally provide acceptable accuracy [4].

Üntez and İpek (2022) predict silver prices using LSTM and ARIMA deep learning methods. The results of the study reveal that the ARIMA method shows more realistic predicts [5].

Öndin and Küçükdeniz (2023) predict silver prices using Google Trends data and the Latent Dirichlet Allocation (LDA) method. They predict silver prices using random forests, gaussian processes, support vector machines, regression trees, and artificial neural networks. The regression tree method yields the most successful predict results [6].

Alparslan and Uçar (2023) use machine learning methods to examine gold and silver commodity price predictions for the COVID-19 pandemic period. The performance of the models is evaluated using MAE, MAPE, and RMSE metrics. The findings show that the Long Short-Term Memory (LSTM) model provides more accurate predictions in the pre-COVID-19 period, while Support Vector Regression (SVR) provides the best predict results for gold and LSTM for silver in the COVID-19 period [7].

Wang et al. (2023) develop an innovative model to predict silver prices. This model includes conventional neural networks (CNNs), the self-attention mechanism (SA), and the new gated unit (NGU). The results show that the CNN-SA-NGU model outperforms other models with a MAE of 87.90, the explained variance score (EVS) of 0.97, the r-squared (R^2) of 0.97, and a training time of 332.78 seconds [8].

Gono et al. (2023) use the XGBoost machine learning method to predict silver prices. They develop two main models to predict silver price fluctuations. In the first model, it predicts that prices will fall on the first two days, rise on the third day, fall again on the fourth day, and rise on the fifth and sixth days. The first model performs best with a MAPE value of 5.98% and a RMSE value of 1.6998. In the second model, it predicts that prices will fall on the first three days, then rise until the sixth day. The model's RMSE value is 1.6967, and its MAPE value is 6.06%. The proposed models perform best when compared to other ensemble models such as CatBoost and random forest [9].

Gür (2024) evaluates the performance of a hybrid model consisting of CNN, LSTM, GRU, and a combination of these models for silver price prediction. The models are trained with historical silver price data, and their prediction accuracies are compared. The findings show that the CNN-LSTM-GRU hybrid model produces more successful predictions [10].

Jin and Xu (2024) use 13 years of data to predict silver prices. In the study, models were developed using Gaussian process regression and Bayesian optimization techniques, and relatively high-level predictions were determined. For the silver price, RRSE was calculated as 0.2257%, RMSE as 0.0515, MAE as 0.0389, and correlation coefficient as 99.967% [11].

3. Material and Method

Machine learning is a subfield of artificial intelligence that enables computers to learn and act in a manner analogous to humans, with the assistance of algorithms and data. One of the fundamental tenets of machine learning is to achieve enhanced performance by minimizing human involvement [12]. In this context, machine learning empowers computers to improve their performance and make informed judgments through data acquisition. The main goal is to create models using data to generate forecasts and predictions. Machine learning algorithms construct a model by studying data and generating predictions based on it. The problem's specific attributes and data collection features inform algorithm selection and use. The field of machine learning has experienced significant advancements over time, particularly in the area of data interaction and modeling methods. From its earliest stages to its present state, machine learning has undergone a notable evolution, progressing from its basic forms to the more sophisticated deep learning techniques we observe today. These advancements have enabled the efficient application of machine learning across various fields. Machine learning provides an extensive array of applications in the financial industry, significantly contributing to the resolution of diverse financial issues. Machine learning methods are also extensively applied in the financial field. The analytical and predictive functionalities it offers in domains including credit scoring, algorithmic trading, portfolio management, and market predicting facilitate more effective and efficient financial decision-making [13].

Many algorithms have been developed for machine learning. Some of these algorithms are particularly suitable for financial predicting and have high performance, as reported in the literature. This study employed linear regression, support vector regression (SMOReg), random forest, and K-Nearest Neighbor (k-NN) methods commonly utilized in the literature to predict silver prices [14].

Linear regression is a statistical method used to establish a relationship between a dependent variable and one or more independent variables. The purpose of linear regression is to find the best-fit line representing the linear relationship between the dependent variable and the independent variables [15].

In the simple linear regression algorithm, a single attribute is used to predict the response that emerges from this attribute. It is assumed that these two variables are linearly related. Therefore, a function is tried to be found that predicts the response value (y) as accurately as possible as a function of the attribute or independent variable (x) [16].

$$h(x_i) = \beta_0 + \beta_1 x_i + \varepsilon_i \tag{1}$$

where,
$h(x_i)$ represents the predicted response value for the i th observation.
β_0 is regression coefficient (intercept/constant)
β_1 is regression coefficient (slope)
ε_i, represents the error term.

Support Vector Machines (SVM) are algorithms that can be used to solve both classification and regression problems [17]. SVM was proposed for solving classification problems in the study conducted by Cortes and Vapnik [18]. Smola and Scholkopf [19] and Shevade et al. [20] developed the Support Vector Regression method to address regression problems involving SMOs [21]. Regression analysis is a model that deals with the relationship between variables in the prediction problem. SMOreg addresses the chaotic and inefficient issues found in standard SMO and utilizes optimization guidelines for binary problems. By employing two threshold parameters, SMOreg operates more efficiently than the original SMO. The SMOreg algorithm further enhances the convergence value observed in SMO, thereby demonstrating improved efficiency and performance [22]. SMOreg uses structural risk minimization constraints as a model, which makes it a reliable way to predict regression and deal with small-sample data that is not linear. Consequently, SMOreg is capable of accurately predicting time series data [23].

The k-Nearest Neighbors (k-NN) algorithm is a method used for classification problems. When k-NN classifies a new data point, it does so based on the labels of the k nearest neighbors in the existing dataset [24]. Additionally, regression problems commonly use it due to its low computational cost and ease of interpretation. In regression tasks, k-NN determines the predicted value of a new data point by calculating the average of the target values of the k nearest neighbors. This process involves calculating the distances between data points, selecting the k nearest neighbors, and averaging the target values of these neighbors.

Random Forest (RF) is a machine learning technique that utilizes multiple decision trees and is employed in both regression and classification problems [25]. The Random Forest algorithm is a bagging-based machine learning method [26]. This algorithm constructs multiple decision trees to improve accuracy and stability by combining them. In Random Forest, each tree's growth increases the degree of randomness in the model. Specifically, rather than searching for the best feature during node splits, the best feature is selected from a randomly chosen subset of features. This approach enhances model diversity, resulting in improved overall performance [27].

3.1. Evaluation Metrics

Correlation Coefficient: The correlation coefficient is a statistical measure that quantifies the direction and strength of the relationship between a dependent variable and one or more independent variables. The correlation coefficient can assume a value between -1 and 1. A correlation coefficient of -1 indicates an inverse relationship between the two variables; a correlation coefficient of 0 indicates the absence of a relationship between the two variables; and a correlation coefficient of 1 indicates a perfect relationship between the two variables [28].

Mean Absolute Error (MAE): Mean absolute error attempts to find out how far the predicted values are from the actual values. The formula for Mean Absolute Error (MAE) is:

$$MAE = \frac{1}{n}\sum_i^n |P_i - A_i| \tag{2}$$

where:
n is the number of observations.
P_i represents the actual values.
A_i represents the predicted values.
$|P_i - A_i|$ is the absolute error for each observation.
In essence, MAE represents the average of the absolute differences between the actual and predicted values, thereby providing a measure of the degree to which the model's predictions align with the actual data.

Root Mean Square Error (RMSE): The root mean square error (RMSE) is a statistical measure that quantifies the discrepancy between the predicted and actual values. It provides a means of assessing the magnitude of the error in estimating the distance

between the two sets of values. The RMSE is calculated by taking the standard deviation of the prediction errors, thereby providing a measure of the spread of the prediction errors.
The formula for Root Mean Square Error (RMSE) is:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i}^{n}(P_i - A_i)^2} \tag{3}$$

where:
n is the number of observations.
P_i represents the actual values.
A_i represents the predicted values.
$(P_i - A_i)^2$ is the squared error for each observation.
The RMSE is useful for assessing the accuracy of a model; it provides a measure of how much error is present in the predictions and tends to give more weight to larger errors due to the squaring step.

Relative Absolute Error (RAE): The relative absolute error is calculated by taking the sum of the difference between the predicted and actual values and dividing it by the sum of the difference between the actual value and the mean of the actual values.
The formula for Relative Absolute Error is:

$$RAE = \frac{1}{n}\sum_{i}^{n}\frac{\sum_{i}^{n}(P_i - A_i)}{\sum_{i}^{n}(P_i - A_m)} \tag{4}$$

where:
n is the number of observations.
P_i is the actual value for observation i.
A_i is the predicted value for observation i.
A_m is total of actual values.
RAE is a valuable tool for assessing the efficacy of a model by providing a benchmark against which its performance can be evaluated. It is particularly useful for comparing the added value of more complex models relative to a straightforward reference model.

Root Relative Squared Error (RRSE): The Root Relative Squared Error (RRSE) formula is utilized to assess the efficacy of a predictive model, particularly in regression tasks. It entails a comparison between the error associated with a model's predictions and the error associated with a baseline model that predicts the mean of the actual values.
The formula for RRSE is:

$$\text{RRSE} = \sqrt{\frac{1}{n}\sum_{i}^{n}\frac{\sum_{i}^{n}(P_{ij} - A_i)^2}{\sum_{i}^{n}(P_i - A_m)^2}} \tag{5}$$

where:
n is the number of observations.
P_{ij} is the predicted value of dataset j for observation i.
A_i is the predicted value for observation i.
A_m is total of actual values.

The RRSE is a useful tool for evaluating the relative performance of a regression model in comparison to a basic model that solely predicts the mean of the actual values. By normalizing the squared error with respect to the error of the mean prediction, the RRSE provides a measure of improvement or deterioration in predictive performance.

3.2. Data Set

The data set of the study consists of daily price data for silver, US dollar index, VIX index, S&P 500 index, Brent oil and gold between 02/01/2008-28/06/2024. The variables included in the study's daily closing prices include 4090 days of data. The dataset of the study is obtained from Yahoo Finance [29]. The open-source software WEKA, utilized in the field of data mining, has been selected for the analysis of the working data. The variables that are thought to affect silver prices and the studies in the literature on these variables are shown in Table-1.

Table 1. Variables

Variables	Abbreviation	Unit	Variable Type	Frequency	References
Silver	XAG	Ons/USD	Output	Daily	
USD Dollar Index	DXY	Index	Input	Daily	[30]
Volatility Index	VIX	Index	Input	Daily	[31];[32]; [33];[34];
S&P 500 Index	SPX	Index	Input	Daily	[30]
Brent Oil	XBR	Barrel/USD	Input	Daily	[35];[36]; [37];[38]
Gold	XAU	Ons/USD	Input	Daily	[39];[40];[41]; [42];43

4. Results and Discussion

In this section of the study, the historical prices of silver during the research period are presented. This section also includes descriptive statistics for the variables, a correlation matrix, and the study's analysis results.

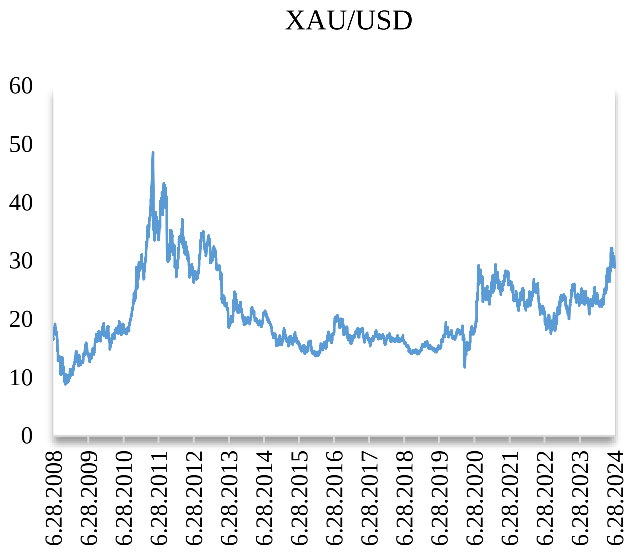

Figure 2. Silver prices in the research period
Source: created by the authors

Figure 2 illustrates that after the 2008 global economic crisis, silver prices entered a distinct upward trend. During this period, the price of silver reached approximately 50 USD per ounce, marking a historical peak. In the following years, particularly until 2020, silver prices exhibited relatively low volatility. However, in 2020, with the increased perception of global risk caused by the COVID-19 pandemic, there was a significant surge in silver demand, leading to a rise in silver prices. In 2021, prices remained elevated but showed some fluctuations. In 2022 and 2023, price variability continued due to economic recovery, monetary policies implemented by central banks, and other economic factors. In the first six months of 2024, silver prices fluctuated due to global economic developments, ranging from between 22 and 32 USD.

Table 2. Descriptive statistics

	DXY	VIX	SPX	XBR	XAU	XAG
Min.	71.33	9.14	676.53	19.33	713.30	8.79
Max.	114.11	82.69	5487.03	146.08	2433.90	48.58
Mean	90.30	19.99	2450.50	78.63	1439.60	20.86
Std.	9.84	9.12	1193.38	24.97	340.24	6.27
Obs.	4090	4090	4090	4090	4090	4090

Table 2 presents the descriptive statistics of the variables. According to the descriptive statistics table, the highest volatility is observed in the VIX, with a standard deviation (9.12) significantly higher than its mean (19.99). The SPX shows significant fluctuation, with a wide range and a high standard deviation (1193.38).

Correlation analysis is effectively used in machine learning-based modeling studies to identify input variables with high correlation in order to eliminate the problem of multicollinearity. A correlation coefficient greater than 0.8 indicates a significant relationship between the variables [44]. In this study, correlation analysis was conducted to examine whether there is a multicollinearity relationship among the variables. The analysis concluded that there are no barriers to including all selected input variables in the model. The results of the correlation analysis among the input variables are presented in Table 3. According to the findings, there is a strong positive relationship (0.7794) between DXY and SPX, indicating that the U.S. Dollar Index is in alignment with the S&P 500 Index. A moderately positive relationship (0.4443) is observed between DXY and XAU.

Table 3. Correlation table for the variables

	DXY	VIX	SPX	XBR	XAU	XAG
DXY	1					1
VIX	-0.1352	1				-0.1352
SPX	0.7794	-0.2140	1			0.7794
XBR	-0.5106	-0.1390	-0.1687	1		-0.5106
XAU	0.4443	-0.1383	0.7477	0.1537	1	0.4443

The strong positive correlation (0.7477) between SPX and XAU suggests that these two variables generally move together. On the other hand, there is a negative relationship (-0.5106) between DXY and XBR, indicating that the U.S. Dollar Index moves in the opposite direction of Brent crude oil prices.

Table 4. Analysis results

Linear Regression			
Error Terms	t+1	t+2	t+3
Correlation Coefficient	0.9210	0.9183	0.9167
MAE	1.7980	1.8168	1.8312
RMSE	2.5324	2.5734	2.5975
RAE%	34.6658	35.0281	35.3059
RRSE%	38.9844	39.6149	39.9858
SMOReg			
Error Terms	t+1	t+2	t+3
Correlation Coefficient	0.9181	0.9151	0.9136
MAE	1.7497	1.7671	1.7793
RMSE	2.6594	2.702	2.7307
RAE%	33.7345	34.0712	34.3061
RRSE%	40.9398	41.5948	42.0372
k-NN			
Error Terms	t+1	t+2	t+3
Correlation Coefficient	0.9954	0.9937	**0.9926**
MAE	0.4192	0.4696	**0.4713**
RMSE	0.6320	0.7316	**0.7878**
RAE%	8.0816	9.0535	**9.0873**
RRSE%	9.7298	11.2619	**12.1281**
Random Forest			
Error Terms	t+1	t+2	t+3
Correlation Coefficient	**0.9960**	**0.9941**	0.9937
MAE	**0.3969**	**0.4458**	0.4745
RMSE	**0.5877**	**0.7144**	0.7342
RAE%	**7.6518**	**8.5953**	9.1479
RRSE%	**9.0468**	**10.9971**	11.3019

In the current study, three distinct models have been developed to predict silver prices. As presented in Table 4, the first model is denoted as t+1, the second as t+2, and the third as t+3. The primary objective in constructing these models is to predict silver prices for one, two, and three days into the future. Each model uses the input variable values from day t, while the output variable-silver price-is represented by its value on days t+1, t+2, and t+3 for the first, second, and third models, respectively. The price predictions from these models are evaluated using four different machine learning algorithms. The use of multiple machine learning methods is intended to ensure the comparability and reliability of the results, facilitating the identification of the model that incorporates the most accurate method. The dataset is split into 80% training data and 20% test data. Predictions are made exclusively with the test data, and the models' performance is assessed based on these results. Model comparison is conducted through the error terms derived from the test data, method producing the lowest error being deemed the most successful.

Table 4 presents the prediction performance of four different machine learning methods individually:

Linear Regression

Correlation Coefficient: The correlation coefficient decreases slightly over time (0.9210→0.9167), indicating a gradual decline in the model's predictive accuracy as time progresses.
MAE and RMSE: Both MAE (1.7980→1.8312) and RMSE (2.5324→2.5975) show a slight increase over time. As errors increase, the model's predictions become less precise.
RAE and RRSE: Percentage errors also increase over time (RAE%34.66→%35.30, RRSE%38.98→%39.98), suggesting a decline in model performance over time.

SMOReg

Correlation Coefficient: Similar to linear regression, the correlation coefficient decreases over time (0.9181→0.9136).
MAE and RMSE: Over time, there have been slight increases in MAE (1.7497 → 1.7793) and RMSE (2.6594→2.7307). Although errors remain relatively low, they increase over time.
RAE and RRSE: Percentage errors increase over time (RAE %33.73→%34.31, RRSE %40.94→%42.04), indicating a decline in performance.

k-NN (k-Nearest Neighbors)

Correlation Coefficient: The correlation coefficient remains very high compared to other models (0.9954→0.9926). Although there is a very slight decrease over time, the model's accuracy remains quite stable.
MAE and RMSE: MAE (0.4192→0.4713) and RMSE (0.6320→0.7878) show a slight increase over time, but the errors are very low compared to other models.
RAE and RRSE: Percentage errors (RAE %8.08→%9.09, RRSE %9.73→%12.13) also increase over time but remain much lower than those of other models.

Random Forest

Correlation Coefficient: is very high (0.9960→0.9937), with a slight decrease over time, but the accuracy remains strong.
MAE and RMSE: MAE (0.3969→0.4745) and RMSE (0.5877→0.7342) increase over time, however these values remain relatively low compared to other models.
RAE and RRSE: Percentage errors (RAE % 7.65→% 9.15, RRSE % 9.05→% 11.30) are low, with a moderate increase.

k-NN and Random Forest exhibit the lowest error rates and highest correlation coefficients for t+1, t+2, and t+3 predictions. Although both models experience a minimal decrease in performance over time, they remain superior. In contrast, linear regression and SMOReg models show deteriorating performance and increasing errors over time, particularly in terms of MAE and RMSE. As time progresses (towards t+3), k-NN and Random Forest models provide better predictions, whereas linear regression and SMOReg models perform weaker.

5. Conclusion and Suggestions

In this study, silver prices are predicted using machine learning methods. In this context, four different machine learning algorithms commonly used in the literature are utilized. These algorithms are linear regression, support vector regression (SMOReg), k-nearest neighbors (k-NN), and random forest. Through these four algorithms, the predictive capabilities and performance of the models over time are demonstrated. This allows for a comparative assessment of the methods' performance in predicting silver prices. In general, Random Forest and k-NN (k-Nearest Neighbors) algorithms, respectively, perform better compared to other methods. These models exhibit higher correlation coefficients and significantly lower error rates than the other models. Random Forest produces the most accurate predictions in the first two models, while k-NN shows the highest prediction accuracy in the third model. Over time, the performance decline in these models remains minimal, and their predictive accuracy remains generally high. The high correlation coefficients and low error rates indicate that these two algorithms provide long-term reliability and accuracy in predicting silver prices. Based on these findings, it is recommended that k-NN and Random Forest algorithms be preferred for long-term predictions, especially in financial analysis, risk management, and investment strategies.

However, the linear regression and SMOReg models have shown a decline in performance over time. There has been a noticeable increase in correlation coefficients, MAE (mean absolute error), and RMSE (root mean square error) values. These increases indicate that these models' prediction accuracy decreases over time, resulting in a reduction in their overall performance. This suggests that the results produced by the linear regression and SMOReg algorithms may be less effective for long-term predicts. Therefore, it may be more appropriate to use these methods for shorter-term predictions or in periods of low volatility. Nevertheless, considering that SMOReg delivers acceptable results in certain cases in the short term, its performance could be improved through further model development and optimization.

In future studies, incorporating additional variables that influence silver prices into the model has the potential to enhance prediction accuracy. Specifically, integrating external factors such as geopolitical events, central banks' monetary policies, and macroeconomic indicators into the model could enrich the algorithms with more comprehensive datasets, leading to more consistent results. In addition, analyzing the interplay between these external factors and silver prices can provide deeper insights into market behavior and trends. Furthermore, exploring more advanced modeling techniques, such as deep learning, may yield superior outcomes compared to existing algorithms, thereby contributing to a more accurate predicting of financial markets.

Finally, we recommend updating the modeled data more frequently and periodically testing the accuracy of the results to improve the algorithms' applicability in real market conditions. This approach would enable the continuous optimization of the models, allowing them to better adapt to changing market dynamics. We anticipate that with more advanced parameter optimization techniques and larger datasets, successful algorithms like k-NN and Random Forest could achieve even higher performance. Moreover, the creation of models with the capacity to adapt to real-time market fluctuations may serve to enhance their validity and effectiveness across a range of economic conditions. This would enable investors and analysts to make more informed decisions.

Contributions of the authors

Erhan Ergin: Data Collection, Literature Review, Data Analysis, Model Development, Editing.

Binali Selman Eren: Conceptualization, Model Development, Manuscript Writing, Interpretation of Results.

Conflict of Interest Statement

There is no conflict of interest between the authors.

Statement of Research and Publication Ethics

The study is complied with research and publication ethics.

References

[1] L. A. Smales and B. M. Lucey, "The influence of investor sentiment on the monetary policy announcement liquidity response in precious metal markets," *Journal of International Financial Markets, Institutions and Money*, vol. 60, pp. 19-38, 2019.

[2] A. Charles, O. Darné, and J. H. Kim, "Will precious metals shine? A market efficiency perspective," *International Review of Financial Analysis*, vol. 41, pp. 284-291, 2015.

[3] U. Çelik and C. Başarır, "The prediction of precious metal prices via artificial neural network by using RapidMiner," *Alphanumeric Journal*, vol. 5, no. 1, pp. 45-54, 2017.

[4] S. Goel, M. Saxena, P. K. Sarangi, and L. Rani, "Gold and silver price prediction using hybrid machine learning models," in *2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC)*, 2022, pp. 390-395.

[5] A. Üntez and M. İpek, "Developing financial forecast modeling with deep learning on silver/ons parity," Journal *of Advanced Research in Natural and Applied Sciences*, vol. 8, no. 1, pp. 35-44, 2022.

[6] S. Öndin and T. Küçükdeniz, "Latent Dirichlet allocation method-based nowcasting approach for prediction of silver price," Accounting, vol. 9, no. 3, pp. 131-1, 2023.

[7] S. Alparslan and T. Uçar, "Comparison of commodity prices by using machine learning models in the COVID-19 era, " *Turkish Journal of Engineering*, vol. 7, no. 4, pp. 358-368, 2023.

[8] H. Wang, B. Dai, X. Li, N. Yu, and J. Wang, "A novel hybrid model of CNN-SA-NGU for silver closing price prediction," *Processes*, vol. 11, no. 3, p. 862, 2023.

[9] D. N. Gono, H. Napitupulu, and Firdaniza, "Silver price forecasting using extreme gradient boosting (XGBoost) method," *Mathematics*, vol. 11, no. 18, p. 3813, 2023.

[10] Y. E. Gür, "Comparative analysis of deep learning models for silver price prediction: CNN, LSTM, GRU, and hybrid approach," *Akdeniz İİBF Dergisi*, vol. 24, no. 1, pp. 1-13, 2024.

[11] B. Jin and X. Xu, "Gaussian process regression-based silver price forecasts," *Journal of Uncertain Systems*, vol. 18, no. 1, 2024. [Online]. Available: https://doi.org/10.1142/S1752890924500132.

[12] M. M. Taye, "Understanding of machine learning with deep learning: architectures, workflow, applications and future directions," *Computers*, vol. 12, no. 5, p. 91, 2023.

[13] R. Gupta, A. Sharma, and T. Alam, “Building Predictive Models with Machine Learning,” in *Data Analytics and Machine Learning: Navigating the Big Data Landscape*, Singapore: Springer Nature Singapore, 2024, pp. 39-59.

[14] E. Koçoğlu and F. Ersöz, “Data mining application with machine learning algorithms to manage interest rate risk,” *Business & Management Studies: An International Journal*, vol. 10, no. 4, pp. 1545-1564, 2022.

[15] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis*. Hoboken, NJ, USA: John Wiley & Sons, 2021.

[16] S. Savaş and S. Buyrukoğlu, Eds., *Teori ve Uygulamada Makine Öğrenmesi*, 1st ed. İstanbul, Türkiye: Beta Basım Yayım Dağıtım A.Ş., 2022.

[17] J. Cervantesa, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, “A comprehensive survey on support vector machine classification: Applications, challenges, and trends,” *Neurocomputing*, vol. 189, pp. 189-215, 2020.

[18] C. Cortes and V. Vapnik, “Support vector networks,” *Machine Learning*, vol. 20, no. 3, pp. 273-297, 1995.

[19] A. J. Smola and B. Scholkopf, “A tutorial on support vector regression,” *Statistics and Computing*, vol. 14, pp. 199–222, 2004.

[20] S. K. Shevade, S. S. Keerthi, C. Bhattacharyya, and K. R. K. Murthy, “Improvements to the SMO algorithm for SVM regression,” *IEEE Transactions on Neural Networks*, vol. 11, no. 5, pp. 1188-1193, 2000.

[21] I. H. Witten, E. Frank, M. A. Hall, C. J. Pal, and M. Data, “Practical machine learning tools and techniques,” in *Data Mining*, vol. 2, no. 4, pp. 403-413, 2005. Amsterdam: Elsevier.

[22] C. Li and L. Jiang, “Using locally weighted learning to improve SMOreg for regression,” in *Pacific Rim International Conference on Artificial Intelligence*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 375-384.

[23] L. Wang, L. Tan, C. Yu, and Z. Wu, “Study and application of non-linear time series prediction in ground source heat pump system,” in *2012 2nd International Conference on Consumer Electronics, Communications and Networks (CECNet)*, 2012, pp. 3522-3525.

[24] P. Cunningham and S. Delany, “k-Nearest neighbour classifiers - A tutorial,” *ACM Computing Surveys*, vol. 54, no. 6, pp. 1-25, 2021.

[25] Y. C. Liang, Y. Maimury, A. H. L. Chen, and J. R. C. Juarez, “Machine learning-based prediction of air quality,” *Applied Sciences*, vol. 10, no. 24, p. 9151, 2020.

[26] L. Breiman, “Random forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[27] A. Ibrahim, R. Kashef, and L. Corrigan, “Predicting market movement direction for Bitcoin: A comparison of time series modeling methods,” *Computers & Electrical Engineering*. 2021. [Online]. Available: https://doi.org/10.1016/j.compeleceng.2020.106905.

[28] B. Marapelli, “Software development effort duration and cost estimation using linear regression and k-nearest neighbors machine learning algorithms,” *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 9, no. 2, pp. 2278-3075, 2019.

[29] finance.yahoo.com (Accessed Date: 1 July 2024)

[30] D. G. Baur and L. A. Smales, "Hedging geopolitical risk with precious metals," *Journal of Banking & Finance*, vol. 117, p. 105823, 2020.

[31] A. A. Salisu, A. E. Ogbonna, and A. Adewuyi, "Google trends and the predictability of precious metals," *Resources Policy*, vol. 65, p. 101542, 2020.

[32] T. L. D. Huynh, "The effect of uncertainty on the precious metals market: New insights from transfer entropy and neural network VAR," *Resources Policy*, vol. 66, p. 101623, 2020.

[33] M. B. Tuncel, Y. Alptürk, T. Yılmaz, and İ. Bekci, "Korku endeksi (VIX) ile kıymetli madenler arasındaki ilişki üzerine ekonometrik bir çalışma," *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, vol. 35, no. 3, pp. 1069-1083, 2021.

[34] M. Dumlupınar and T. Kocabıyık, "Yatırım aracı olarak kıymetli metaller: Kıymetli metallerin fiyatını etkileyen unsurlar ve kıymetli metallerde nedensellik ilişkisi," *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, vol. 28, no. 2, pp. 241-266, 2023.

[35] M. Bildirici and C. Türkmen, "The chaotic relationship between oil return, gold, silver, and copper returns in Turkey: Non-linear ARDL and augmented non-linear Granger causality," *Procedia-Social and Behavioral Sciences*, vol. 210, pp. 397-407, 2015.

[36] I. Arif, L. Khan, and K. M. Iraqi, "Relationship between oil price and white precious metals return: A new evidence from quantile-on-quantile regression," *Pakistan Journal of Commerce and Social Sciences*, vol. 13, no. 2, pp. 515-528, 2019.

[37] M. Kamışlı, S. Kamışlı, and F. Temizel, "Emtia fiyatları birbirlerini etkiler mi? Asimetrik frekans nedensellik analizi," *Uluslararası Yönetim İktisat ve İşletme Dergisi*, vol. 13, no. 13, pp. 1079-1093, 2017.

[38] A. A. Salisu, A. E. Ogbonna, and A. Adewuyi, "Google trends and the predictability of precious metals," *Resources Policy*, vol. 65, p. 101542, 2020.

[39] B. Elmas and M. Polat, "Gümüş fiyatları ve Dow Jones endeksi'nin altın fiyatlarına etkisi üzerine eşbütünleşme ve nedensellik analizi," *Bingöl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, vol. 3, no. 6, pp. 33-48, 2013.

[40] A. F. Erol and S. Aytekin, "Altın fiyatlarını etkileyen faktörlerin nedensellik analizi ile incelenmesi," in *International Congress of Management, Economy and Policy (ICOMEP'20)*, 2020.

[41] B. R. Mishra, A. K. Pradhan, A. K. Tiwari, and M. Shahbaz, "The dynamic causality between gold and silver prices in India: Evidence using time-varying and non-linear approaches," *Resources Policy*, vol. 62, pp. 66-76, 2019.

[42] A. Açacak, E. Gülsar, and E. Meriç, "Kıymetli madenlerin birbirleriyle ilişkisi: Asimetrik nedensellik," *Anadolu Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, vol. 21, no. 1, pp. 28-37, 2020.

[43] M. Aslan and L. Sizer, "Dow Jones endeksi, gümüş fiyatları ve altın fiyatları arasındaki ilişkinin ekonometrik açıdan incelenmesi," *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, vol. 23, no. 4, pp. 1035-1048, 2023.

[44] M. Uluskan and H. D. Şenli, "YSA sınıflandırma modellerinde korelasyon-hipotez testi tabanlı filtreleme yoluyla girdi seçimi," *Nicel Bilimler Dergisi*, vol. 6, no. 1, pp, 68-102, 2024.