

Article

Neural-Networks-Based Gold Price Forecasting with Recursive Feature Elimination and Optuna Tuning

Alireza Panahi and Salim Lahmiri *

Department of Supply Chain and Business Technology Management, John Molson School of Business, Concordia University, Montreal, QC H3H 0A1, Canada; alireza.panahi@mail.concordia.ca

* Correspondence: salim.lahmiri@concordia.ca

Abstract

Background: Strategic resource planning is crucial for optimizing supply chain management and ensuring efficient operations. This study aims to enhance strategic planning in gold mines by leveraging advanced gold price forecasting models. By predicting future gold prices accurately, mining companies can better plan their extraction, processing, and distribution activities, thereby improving overall supply chain efficiency. **Methods:** Various advanced forecasting models are implemented, including backpropagation neural networks (BPNNs), convolutional neural network (CNN), long short-term memory (LSTM), bi-directional LSTM (Bi-LSTM), gated recurrent unit (GRU), and bi-directional GRU (Bi-GRU). The feature selection process is facilitated by recursive feature elimination (RFE), and Optuna is used to fine-tune neural network models. Evaluation is based on root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). **Results:** BPNN performed the best in terms of lowest RMSE (0.5928), MAE (0.4091), and MAPE (0.34%), whilst Bi-GRU was the poorest performer, as it achieved RMSE of 9.41, MAE of 8.1916, and MAPE of 6.94%. In addition, Optuna further improved each model's accuracy, except CNN, where the performance slightly decreased. **Conclusions:** Advanced forecasting neural systems underperformed the standard backpropagation neural networks. In this regard, BPNN proved to be highly effective in forecasting gold price, providing critical managerial implications for navigating the dynamic and volatile gold market for gold mining companies and investors.

Keywords: gold price; backpropagation neural networks; convolutional neural networks; long short-term memory neural networks; gated recurrent units; bi-directional GRU; bi-directional LSTM; recursive feature elimination; Optuna tuning; forecasting



Academic Editor: Abdulsalam Yassine

Received: 26 April 2026

Revised: 15 June 2026

Accepted: 3 July 2026

Published: 5 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Forecasting gold prices is a critical undertaking that holds substantial implications for various stakeholders in the financial landscape. The intrinsic value and historical significance of gold make it an essential natural resource commodity, impacting global economies, investment decisions, and risk management strategies. Precise gold price predictions empower investors, miners, and policymakers to make informed choices, seize opportunities, and mitigate potential risks.

The importance of forecasting gold prices is multifaceted. For investors, accurate predictions are crucial for optimizing portfolio diversification and safeguarding against inflation or economic uncertainties. Gold has long been considered a safe-haven asset, and during times of economic instability or market volatility, its value often increases as

investors seek stability. Therefore, being able to predict gold prices with precision allows investors to strategically adjust their portfolios, either by increasing their gold holdings to hedge against potential losses in other asset classes or by taking profits when gold prices are expected to peak.

Mining companies heavily rely on precise gold price forecasts to strategize long-term investment decisions and maximize returns from mining projects. The profitability of mining operations is closely tied to gold prices, as higher prices can justify the investment in new mining projects or the expansion of existing ones. Conversely, accurate forecasts can help mining companies avoid overinvestment during periods of expected price declines. This strategic planning is vital for maintaining financial stability and ensuring the efficient allocation of resources within the mining sector. Additionally, efficient resource planning for gold mines, including the assessment of production capacities and operational costs, is directly influenced by gold price forecasts. A well-planned production strategy can optimize output to match market demand, thus stabilizing prices and ensuring profitability. Fluctuations in gold prices can significantly impact the economic viability of mining projects, influencing decisions on when to ramp up or scale down production. In sum, gold price forecasting is intricately connected to key aspects of supply chain management, including demand forecasting, risk management, and inventory management. Precise predictions of gold prices enable businesses to optimize stock levels, ensuring that supply meets demand without excessive overstocking or increased holding costs.

In addition, financial institutions benefit from such predictions to develop effective risk management plans for commodities-related investments. Banks, hedge funds, and other financial entities engage in trading and investment activities involving gold and other commodities. By leveraging accurate gold price forecasts, these institutions can devise strategies to manage exposure to price fluctuations, hedge against potential losses, and enhance overall investment performance. This risk management capability is essential for maintaining the stability and profitability of financial institutions in a volatile market environment.

Recently, García-Gonzalo et al. [1] introduced three methods for constructing time series datasets—direct multi-step, recursive multi-step, and direct-recursive hybrid schemes—and applied two novel hybrid methods, NARX SVR and DE/GPR, for predicting spot gold prices. Differential evolution (DE) was used for hyperparameter tuning. Among the models tested, NARX DE/SVR achieved the best prediction accuracy, as measured by RMSE, using publicly available COMEX data. Li et al. [2] employed the wavelet method to analyze the time-dependent correlation between gold price (GP) and global supply chain pressure (GSCP). Findings revealed that GSCP positively influences GP across various time frames, indicating gold's ability to hedge against supply chain stresses. Additionally, GP leads GSCP in the medium and short terms, suggesting gold's role as a predictive indicator. The analysis showed that dollar- and renminbi-denominated gold are better hedges, while yen-denominated gold exhibited the weakest correlation with GSCP. Dynamic model averaging (DMA) [3] was another advanced technique explored for gold price prediction. This method demonstrated its superiority over linear models and Bayesian model averaging. The study identified the business cycle factor as a strong predictor of gold returns, showcasing the effectiveness of the DMA model in accounting for dynamic changes in the market environment. The ability of DMA to adaptively weight different models based on their predictive performance over time makes it a powerful tool for forecasting in volatile markets. In their research, Khani et al. [4] developed a model using long short-term memory networks (LSTMs) to predict stock market values considering the impact of COVID-19. The study employed multiple LSTM variations, including CNN-LSTM (convolutional neural networks) and bidirectional LSTM. The vector sequence output LSTM achieved superior

performance with MSE values of 0.0006, 0.0008, and 0.002 for 1-day, 2-day, and 30-day predictions, respectively, outperforming existing methods. The framework emphasized the importance of incorporating COVID-19 case data alongside traditional market indicators for accurate forecasting. Further, Roslan and Halim [5] utilized the geometric Brownian motion (GBM) model to forecast the world gold market price under the uncertain conditions of the COVID-19 pandemic in 2022. The model was validated using data from 2016 to 2021, achieving a mean absolute percentage error of 12.21%, indicating good accuracy. The results projected an increasing trend in gold prices for 2022, demonstrating the GBM model's robustness in unexpected situations. Very recently [6], an integrative system based on CNN, LSTM, and GRU to forecast gold was proposed where CNN and LSTM are employed in parallel to respectively learn local features and trends and generate the initial forecasts. Then, GRU learns and aggregates forecasts from CNN and LSTM to produce the final price prediction. In addition, the whale optimization algorithm was adopted to tune the parameters of CNN, LSTM, and GRU. It was found that the proposed integrative forecasting system outperforms baseline prediction systems.

Recent works used to forecast gold price were based on statistical models [1], correlation analysis [2], dynamic model averaging [3], geometric Brownian motion [5], and deep learning models [4,6]. The latter studies [4,6] concluded that deep learning is effective and outperforms standard machine learning models. Indeed, deep learning can learn complex and deep patterns in the data to generate more accurate predictions. However, despite using deep learning like CNN, LSTM, and GRU, it is unclear which system is more appropriate compared to existing deep learning models. In addition, previous approaches still face critical challenges in capturing complex dynamics in gold price because of multiple variables including international economic conditions, political factors, and investor sentiment. Furthermore, identifying the most potential variables for gold price prediction by deep learning would improve accuracy while speeding up training and offering a way to explain the results.

To address these gaps, this study implements and compares various deep learning architectures trained with selected input variables selected from the universe of economic, political and market sentiment. As a result, the study aims to significantly improve both predictive accuracy and transparency, enabling more vigorous prediction and actionable insights for practitioners and policymakers.

The primary objective of this research is to implement and compare advanced forecasting models and integrate diverse data sources to enhance the accuracy of gold price predictions. This study aims to explore and implement machine learning techniques, such as deep learning models, including the standard backpropagation neural networks [7], convolutional neural networks (CNNs) [8], long short-term memory (LSTM) networks [9], and gated recurrent units (GRUs) [10]. These models are particularly well-suited for capturing complex, nonlinear relationships in time series data and have shown promise in various forecasting applications.

In this study, we rely on deep learning, as it has been found to be successful in various problems in supply chain and operations management. Deep learning has been effectively applied to a wide range of challenges within supply chain and operations contexts. For instance, it has been utilized for inventory optimization under demand uncertainty [11], KPI-based demand forecasting [12], time series forecasting in blood donation supply chains [13], and data-driven newsvendor problems [14]. It has also been adopted in multi-echelon network optimization under uncertainty [15], predicting supply disruptions [16], optimizing order picking in warehouses [17], and vehicle routing with time windows [18]. Further, it has supported tactical production planning decisions [19] and enhanced commodity hedging strategies through explainable artificial intelligence [20]. These diverse

applications confirm the effectiveness of deep learning in tackling complex, data-intensive, and dynamic problems in modern supply chain and operations management.

The key contributions of this research are as follows. First, by integrating deep learning models into predictive analytics, this research provides stakeholders such as manufacturers, jewellers, and investors with a data-driven approach to decision-making. The enhanced price predictions enable more efficient inventory management, mitigating risks associated with price volatility. Additionally, suppliers can optimize procurement strategies by aligning purchasing decisions with projected price trends, reducing costs and improving profitability. Ultimately, these contributions help streamline operations and reinforce resilience across the gold supply chain. Second, for the gold mining industry, reliable price forecasting is crucial for investment planning, resource allocation, and production scheduling. This study's deep learning models provide mining companies with more precise price predictions, facilitating strategic decisions on extraction rates, capital expenditures, and risk management. By anticipating market fluctuations, mining firms can adjust operations to optimize profitability and sustainability. Third, the study successfully developed and implemented cutting-edge deep learning architectures tailored to predict gold prices. The models include backpropagation neural networks (BPNNs) [7], convolutional neural networks (CNNs) [8], long short-term memory (LSTM) [9], and gated recurrent units (GRUs) [10], which were specifically designed to handle the nonlinear and dynamic complexities inherent in the gold market. In addition, advanced architectures of LSTM and GRU are considered in the study, including bidirectional LSTM (Bi-LSTM) [21], and bidirectional GRU (Bi-GRU) [22]. In this regard, the study conducts a comparative analysis of the BPNN, CNN, LSTM, Bi-LSTM, GRU, and Bi-GRU systems to highlight their unique advantages. BPNN is a universal function approximator, CNN, with its convolutional architecture, effectively identifies temporal patterns within data, LSTM exhibits superior capabilities in memorizing sequential data by capturing long-term dependencies, and GRU offers a computationally streamlined alternative to LSTM without sacrificing accuracy. This comparative evaluation provided critical insights into selecting the most appropriate model for specific forecasting scenarios. Fourth, this research lies in its utilization of an extensive dataset encompassing a diverse range of variables. These included traditional investment instruments, prices of precious metals, currency exchange rates, and company-specific metrics. By integrating this comprehensive set of inputs, the models were better equipped to capture the multifaceted dynamics of the gold market. This holistic approach ensured that the predictions reflected not only historical trends but also broader market influences, enhancing the relevance and robustness of the forecasting results. Fifth, recursive feature elimination (RFE) [23] for dimensionality reduction is adopted to systematically remove noisy and redundant variables so that the models focus on the most significant predictors, thereby improving accuracy. Moreover, the dimensionality reduction achieved through RFE would enhance computational efficiency, accelerate the training process and maintain the integrity of the predictions. Sixth, to further refine the models' performance, the study employed Optuna [24], an advanced hyperparameter optimization framework based on sampling and pruning algorithms, to find the best parameters. Through these contributions, the research sets a benchmark for utilizing advanced deep learning techniques for gold price forecasting, offering both methodological rigor and actionable insights for future studies and industry practices.

The rest of the paper is organized as follows. Section 2 presents the methods and Section 3 describes the data and provides the experimental results. Finally, a discussion and conclusion are provided in Section 4.

2. Materials and Methods

2.1. General Framework

The selected artificial neural networks models for this study include BPNN, CNN, LSTM, Bi-LSTM, GRU, and Bi-GRU. A recursive feature elimination algorithm [23] is adopted to select the most suitable inputs to train the neural networks. In addition, each artificial neural network uses hyperparameter tuning based on Optuna [24] to generate a set of hyperparameters that yield an optimal model to minimize the loss function. The hyperparameter optimization technique used is Bayesian optimization with the algorithm tree-structured Parzen estimator (TPE), which builds a probability model of the objective function using the best-suited hyperparameters that are chosen to input into the true objective function.

Overall, the experimental architecture is illustrated in Figure 1, outlining a comprehensive framework for our investigations. First, the data is normalized in [0,1], which is then split into 80% for training and 20% for testing, and then a 30-day rolling window is applied to the dataset. The data pre-processing stage, including a feature selection step and model tuning, was applied to each model during the training phase. Each model, trained independently, is assessed based on three standard evaluation metrics: root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). The artificial neural networks, the optimization technique, and performance measures are described next.

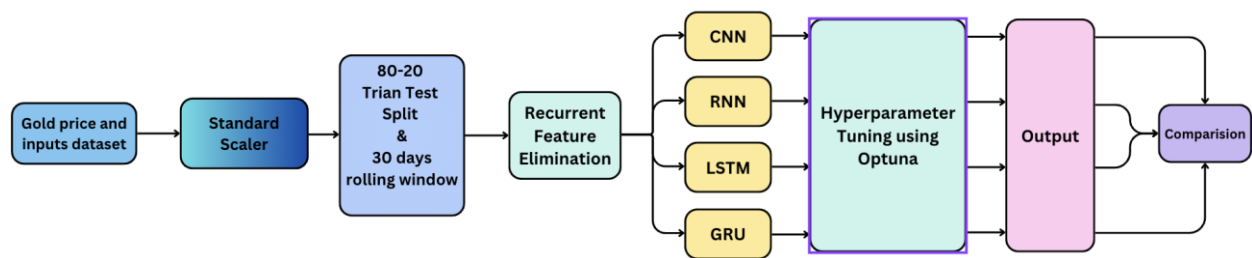


Figure 1. Flowchart of the forecasting systems and experiments.

2.2. Backpropagation Neural Network

Backpropagation neural network (BPNN) [7] consists of layers of interconnected neurons, where each connection has an associated weight. These networks are designed to recognize patterns and make predictions by learning from data. The information flows in one direction from the input to the output with cycles. BPNN learns to map inputs to outputs through a process known as backpropagation, which involves adjusting the weights of connections based on the error between the predicted and actual outputs. The activation function σ applied to each neuron's output introduces nonlinearity into the model, allowing ANNs to capture complex relationships within the data. The basic mathematical formulation for the output y of a neuron in a BPNN layer can be expressed as:

$$y = \sigma(\sum_i w_i x_i + b) \quad (1)$$

where w_i are the weights, x_i are the input features, and b is the bias term. The activation function σ (for instance, the ReLU) determines the output of the neuron based on the weighted sum of inputs.

The fully connected BPNN processes the input as a static feature set. It starts with a dense layer with 256 units followed by another dense layer with 128 units. To prevent overfitting, a dropout layer is introduced. A dense layer with 64 units further refines the output, and the model finishes with a single-unit output layer.

2.3. Convolutional Neural Networks

Convolutional neural network (CNNs) [8] were designed for image analysis and classification. Renowned for their hierarchical feature extraction capabilities, CNNs employ convolutional layers that apply learnable filters to detect local patterns within the input data. These filters progressively capture features, starting from basic edges and textures and evolving to more complex representations in deeper layers. In addition, one-dimensional (1D) CNN is a specialized deep learning architecture adapted for processing sequential data, particularly time series, signals, and textual data. Unlike traditional CNNs primarily designed for spatial data like images, 1D CNN is specifically structured to capture dependencies and patterns along a single axis, such as a temporal sequence.

The fundamental concept of 1D CNN involves employing one-dimensional convolutional filters that move across the input sequence, extracting local patterns and capturing relationships between adjacent data points. These filters slide through the sequential input, performing convolutions to detect temporal patterns, transitions, and dependencies.

In each CNN layer, the forward propagation is expressed as follows:

$$x_k^l = b_k^l + \sum_{i=1}^{N_{l-1}} \text{conv1d}(w_{ik}^{l-1}, s_i^{l-1}) \quad (2)$$

where x_k^l is defined as the input, b_k^l is defined as the bias of the k^{th} neuron at layer l , s_i^{l-1} is the output of the i^{th} neuron at layer $l - 1$, and w_{ik}^{l-1} is the kernel of the i^{th} neuron at layer $l - 1$ to the k^{th} neuron at layer l . $\text{conv1d}(\dots)$ is used to perform ‘in-valid’ 1D convolution without zero-padding.

Like the standard CNN, the 1D CNN architecture consists of convolutional layers followed by pooling layers that reduce dimensionality, thereby extracting and emphasizing crucial features from the input sequence. These layers enable hierarchical feature extraction, progressively learning higher-level representations as information flows through the network. The advantages of 1D CNN lie in their ability to automatically learn relevant features from raw sequential data without relying on manual feature engineering. They excel in capturing local patterns and dependencies, making them valuable for various tasks such as time series forecasting, speech recognition, sentiment analysis in text, and bioinformatics.

CNN captures local patterns in sequential data. The model begins with a 1D convolutional layer (128 filters) that scans the input sequence, followed by a max pooling layer that down-samples the input. Another convolutional layer (64 filters) followed by max pooling extracts hierarchical features. After flattening the data, a dense layer (64 units) refines the extracted features, leading to the final output layer.

2.4. LSTM and Bi-LSTM

Long short-term memory (LSTM) [9] is a variant of the standard recurrent neural network (RNN) crafted to overcome the challenge of capturing long-range dependencies in sequential data. LSTMs have gained widespread recognition for their efficacy in handling sequential information. LSTMs feature a unique memory cell structure with three gates—input, forget, and output—that regulate the flow of information. These gates enable LSTMs to selectively retain or discard information over time steps, allowing them to effectively capture and utilize long-term dependencies in sequential data without facing the vanishing gradient problem encountered by traditional RNNs. Each memory cell is built around a central linear unit with a fixed self-connection (the CEC). In addition to net_{cj} , c_j gets input from a multiplicative unit out_j and from another multiplicative unit in_j . in_j 's activation at time t is denoted by $y^{in_j}(t)$, and out_j 's by $y^{out_j}(t)$. We have

$$Y^{out_j}(t) = f_{out_j}(net_{out_j}(t)) \quad (3)$$

$$y^{in_j}(t) = f_{in_j}(net_{in_j}(t)) \quad (4)$$

The input gate decides the relevance of new information, the forget gate manages the retention or deletion of outdated data, and the output gate controls the output based on the current cell state and input. These mechanisms, facilitated by an activation function (for instance, a sigmoid), empower LSTM to retain essential information, enabling it to capture intricate patterns and dependencies across sequences. The ability to comprehend contextual information and long-range dependencies makes it invaluable in time series forecasting and sequential data classification. However, LSTMs might demand substantial computational resources, particularly in training larger networks or with extensive datasets.

In addition, the bi-directional LSTM (Bi-LSTM) [21] architecture encompasses both forward and backward LSTMs, functioning in parallel to process input feature vector sequences in opposing directions. This bidirectional approach optimally leverages contextual insights, encompassing not just the attributes of the present time step but also those of preceding and subsequent time steps.

In this study, the first LSTM layer with 128 units returns the entire sequence, followed by another LSTM layer with 64 units that condenses the sequence into a final hidden state. A dense layer with 128 units and ReLU activation, followed by a dropout layer (to prevent overfitting), refines the features before outputting a final prediction. Also, the Bi-LSTM captures information from both directions within a sequence. The first Bi-LSTM layer (128 units) processes the input in both forward and reverse directions, followed by a dropout layer to add regularization. A second Bi-LSTM layer (64 units) further refines the sequence representation. This architecture is finished with a dense output layer.

2.5. GRU and Bi-GRU

GRU [10] consists of a simplified architecture with two gates—an update gate and a reset gate—designed to manage the flow of information throughout the network. Unlike LSTMs, GRUs integrate the reset and update gates, allowing for a more straightforward and computationally efficient structure. The update gate in GRUs controls the blending of new input information with the previous memory state, determining the extent to which the new information is incorporated. Meanwhile, the reset gate regulates the amount of past information to forget or retain based on the current input.

The simplicity of the GRU architecture, with its reduced number of gates, results in a more straightforward training process and computational efficiency compared to LSTMs. Despite their simpler design, GRUs have demonstrated competitive performance in various sequential modelling tasks. Their capability to capture dependencies and context across sequences makes them valuable in scenarios where computational efficiency and streamlined architecture are essential.

The gated recursive convolutional neural network (grConv) consists of four weight matrices W^l , W^r , G^l and G^r . At each recursion level $t \in [1, T - 1]$, the activation of the j^{th} hidden unit $h_j^{(t)}$ is computed by:

$$h_j^{(t)} = \omega_c h_j^{(t)} + \omega_l h_{j-1}^{(t-1)} + \omega_r h_j^{(t-1)} \quad (5)$$

where ω_c , ω_l and ω_r are the values of a gate that sum to 1, ensuring a weighted combination of the current and previous hidden states.

GRU is a simple alternative to LSTM. It starts with a GRU layer containing 128 units that processes input sequences and returns all the hidden states. This is followed by another GRU layer with 64 units, but this time only the last hidden state is returned, condensing the sequence into a single vector. After the GRU layers, a fully connected (dense) layer

with 64 units and ReLU activation refines the output before the final layer produces a single prediction.

In the Bi-GRU, the layers are bidirectional, meaning they process the input sequence in both forward and backward directions, capturing information from both the past and the future within a sequence. The first Bi-GRU layer has 128 units and returns sequences, which allows the second Bi-GRU layer (64 units) to refine the entire sequence further. The model concludes with a dense layer that produces the final prediction.

2.6. Recursive Feature Elimination

We employ recursive feature elimination (RFE) [23] to enhance model performance by systematically identifying and retaining the most significant features while eliminating those contributing least to the predictive power. The process is particularly advantageous in high-dimensional datasets, mitigating overfitting, reducing computational costs, and improving interpretability. The generic algorithm is as follows.

- i. The RFE process begins with the complete set of features from the dataset, constructing an initial model to evaluate and rank feature importance. Feature importance is assessed based on coefficients in linear models or importance scores in tree-based models.
- ii. Once all features are ranked, the feature with the lowest importance is removed from the dataset. The model is then re-trained using the reduced set of features, allowing it to recalibrate its predictions with a more focused input set.
- iii. This process of ranking, elimination, and re-training is repeated iteratively, with the weakest feature removed in each iteration.
- iv. The iterations continue until either the desired number of features is achieved or further eliminations cease to improve model performance.

The advantages of RFE are manifold. By reducing the number of features, it decreases the risk of overfitting, where the model might otherwise learn noise and irrelevant patterns from the data. This simplification also improves model accuracy by ensuring that only the most pertinent features are used in predictions. Moreover, by reducing the dimensionality of the dataset, RFE lowers computational costs, making the model more efficient and faster to train. These benefits make RFE a particularly valuable technique for our study, where we are dealing with high-dimensional data and seek to optimize model performance and interpretability.

2.7. Hyperparameter Tuning Using Optuna

Optuna [24] is hyperparameter optimization protocol used to automate the search for optimal hyperparameters. It is designed to be both flexible and efficient, enabling the exploration of complex search spaces while offering a user-friendly interface for defining and executing optimization tasks. The process begins with the definition of an objective function, which is the function to be minimized or maximized. In machine learning, this typically corresponds to the model's loss function.

The next step involves defining the search space, which includes the hyperparameters to be tuned and their respective ranges or distributions. For neural networks, this might involve parameters such as the kernel size, which determines the size of the convolutional filters in CNNs; the number of filters, which specifies how many convolutional filters are present in each layer of a CNN; and the number of units, which refers to the number of neurons in each layer of a BPNN or LSTM.

Optuna employs a tree-structured Parzen estimator (TPE) for Bayesian optimization, which models the distribution of the objective function and selects hyperparameters that are likely to yield better results. During trial execution, Optuna runs multiple trials, each

with a different set of hyperparameters, and evaluates their performance using the objective function. A notable feature of Optuna is its ability to prune unpromising trials early, thereby conserving computational resources and accelerating the optimization process. After a predefined number of trials or upon achieving a satisfactory result, Optuna identifies the best set of hyperparameters that optimize the objective function. This flexibility and efficiency make Optuna a powerful tool for hyperparameter tuning, particularly in complex models where manual tuning would be impractical.

2.8. Performance Measures and Parameter Setting

We employ a rolling window approach to form the set of inputs to predict gold price, where a window of 30 days moves iteratively through the time series data. The rolling window methodology involves using a fixed window size of 30 days to train the model on historical data and then assess its performance on the subsequent period. The window moves forward one step at a time, incorporating new observations while dropping the oldest, enabling the model to iteratively learn from past 30-day intervals and forecast the next step. Within this context, metrics such as root mean squared error (RMSE), mean absolute deviation (MAE), and mean absolute percentage error (MAPE) are used to evaluate the accuracy of each artificial neural network.

RMSE calculates the square root of the average of squared differences between predicted $\hat{y}$ and actual y over n observations. It is given by:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (6)$$

In addition, MAE computes the average of absolute differences between predicted and actual values, focusing on the average magnitude of errors, regardless of their direction. It is expressed as follows:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (7)$$

Finally, MAPE measures the average percentage difference between predicted and actual values over n observations. It is given by:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (8)$$

The lower the performance metric, better is the performance of the predictive neural networks. Finally, Table 1 provides the search space for parameter setting.

Table 1. Search space for hyperparameter setting.

Model	Parameters	Value and Range
BPNN	Hidden layers	(1, 5)
	Network neurons	(10, 100)
	Activation	ReLU, sigmoid, tanh
CNN	Filters	(16, 256, 16)
	Kernel size	[3, 5]
	Dense units	(8, 128, 8)
	Learning rate	(0.0001, 0.01, Log-uniform)
Uni/BiLSTM	Units	(16, 256, 16)
	Learning rate	(0.0001, 0.01, Log-uniform)
Uni/BiGRU	Units	(16, 256, 16)
	Learning rate	(0.0001, 0.01, Log-uniform)

3. Results

The dataset employed in this study encompasses a rich collection of financial and economic time series data obtained from Yahoo Finance [25] spanning from 18 November 2011 to 1 January 2019. The data encompasses a total of 1718 data points, each containing information across 80 distinct attributes. This comprehensive dataset offers valuable insights into the relationships between various financial instruments and economic indicators over a substantial time period.

A core component of data collection includes metrics related to traditional investment instruments. The inclusion of the Standard and Poor's (S&P500) index and the Dow Jones Industrial Average provides a benchmark for overall stock market performance. Additionally, the dataset incorporates US Bond rates (specifically, 10-year rates) to understand the prevailing interest rate environment, which can significantly impact investment decisions. Furthermore, the presence of the Euro USD exchange rate data allows for the examination of currency fluctuations and their potential influence on international financial markets.

The dataset delves deeper into precious metals by including the prices of silver, platinum, palladium, and rhodium. These commodities are often viewed as hedges against inflation and economic uncertainty, and their price movements can offer valuable insights into investor sentiment and risk perception. The inclusion of the US Dollar Index, a measure of the US dollar's relative strength against a basket of major currencies, further enriches the dataset by enabling analysis of how the value of the US dollar interacts with other financial instruments.

Beyond traditional financial instruments, the dataset incorporates data from the Eldorado Gold Corporation, a gold mining company. This inclusion allows for the exploration of the relationship between the performance of a specific gold mining company and the broader gold market. Additionally, the presence of data for Gold Miners ETF (Exchange Traded Fund) facilitates the examination of how a basket of gold mining companies performs in comparison to individual companies within the sector.

A particularly noteworthy aspect of the dataset pertains to the inclusion of historical data for a Gold Exchange Traded Fund (ETF) obtained from Yahoo Finance. Indeed, Gold ETF freely available on Yahoo Finance is chosen as it tracks the spot price of gold. Gold ETF is a new investment product used to track the performance characteristics of physical gold after providing for required expenses. It is created and redeemed by authorized participants and traded like stocks on the exchange market. Because of the arbitrage activity assumed by the authorized participants, the value of gold ETF is aligned with that of gold. This data encompasses seven distinct columns: date, open, high, low, close, adjusted close, and volume. It is crucial to distinguish between the "close" and "adjusted close" price points. The closing price represents the price of the ETF at the close of the trading day. In contrast, the adjusted closing price incorporates factors such as dividends, stock splits, and new stock offerings to determine a more comprehensive value. As such, the "adjusted close" price serves as the target variable within the dataset, representing the value that the model aims to predict.

The list of predictors includes opening value of the S&P500, highest value of the S&P500, lowest value of the S&P500, closing value of the S&P500, adjusted closing value of the S&P500, trading volume of the S&P500, opening value of the Dow Jones index, highest value of the Dow Jones, lowest value of the Dow Jones, closing value of the Dow Jones, adjusted closing value of the Dow Jones, trading volume of the Dow Jones, opening stock price of Eldorado Gold, highest stock price of Eldorado Gold, lowest stock price of Eldorado Gold, closing stock price of Eldorado Gold, adjusted closing stock price of Eldorado Gold, trading volume of Eldorado Gold stock, EUR/USD exchange rate, opening value of the EUR/USD, highest value of the EUR/USD, lowest value of the EUR/USD,

Directional trend of the EUR/USD, current price of Brent crude oil futures, opening price of Brent crude oil futures, highest price of Brent crude oil futures, lowest price of Brent crude oil futures, trading volume of Brent crude oil futures, directional trend of Brent crude oil futures, current price of WTI crude oil, opening price of WTI crude oil, highest price of WTI crude oil, lowest price of WTI crude oil, directional trend of WTI crude oil prices, current market price of silver futures, opening price of silver futures, highest price of silver futures, lowest price of silver futures, trading volume of silver futures, directional trend of silver futures, market price of US bond instruments, opening price of US bonds, highest price of US bonds, lowest price of US bonds, directional trend of US bond yields or prices, market price of platinum, opening price of platinum, highest price of platinum, lowest price of platinum, directional trend of platinum prices, market price of palladium, opening price of palladium, highest price of palladium, lowest price of palladium, directional trend of palladium prices, current price of rhodium, current price of the US Dollar Index, opening price of the US Dollar index, highest price of the US Dollar, lowest price of the US Dollar, trading volume of the US Dollar, directional trend of the US Dollar, opening price of the GDX (The VanEck Gold Miners ETF, a major exchange-traded fund designed to track the performance of companies involved in the global gold mining industry), highest price of the GDX, lowest price of the GDX, closing price of the GDX, Adjusted closing price of the GDX, trading volume of the GDX, opening price of the USO oil (The United States Oil Fund is an exchange-traded security that is designed to track the daily price movements of light, sweet crude oil), highest price of the USO oil, lowest price of the USO oil, closing price of the USO oil, adjusted closing price of the USO oil, and trading volume of the USO oil.

In conclusion, the dataset employed in this study offers a rich tapestry of financial and economic time series data. The comprehensive nature of the data, encompassing traditional investment instruments, precious metals, currency exchange rates, and company-specific information, provides a fertile ground for exploring various research questions related to financial modelling and market dynamics. The inclusion of both closing and adjusted closing prices for the Gold ETF further enhances the dataset's value, allowing for the development of models that predict the true economic value of the ETF. This comprehensive dataset paves the way for in-depth analysis of financial markets and the complex relationships between various economic factors.

In this study, the RFE algorithm selected the following features: Gold ETF open, Gold ETF low, Gold ETF volume, trading volume of the S&P 500, trading volume of the Dow Jones, opening stock price of Eldorado Gold, closing stock price of Eldorado Gold, trading volume of Eldorado Gold stock, EUR/USD exchange rate, trading volume of Brent crude oil futures, trading volume of the US Dollar, opening price of the GDX, closing price of the GDX, closing price of the GDX, trading volume of the GDX, and trading volume of the USO oil. The performance results of each neural system tuned by Optuna is provided in Table 2.

Table 2. Performance of artificial neural networks when tuned by Optuna.

Model	RMSE	MAE	MAPE
BPNN	0.5928	0.4091	0.34%
CNN	2.1241	1.7431	1.42%
LSTM	8.7046	7.3769	6.27%
Bi-LSTM	7.314	5.8623	5.00%
GRU	7.8233	6.4011	5.45%
Bi-GRU	9.41	8.1916	6.94%

Accordingly, the BPNN model surpassed all others, achieving an RMSE of 0.5928, MAE of 0.4091, and MAPE of 0.34%. BPNN's simplicity and robustness in learning complex, nonlinear relationships within the data facilitated this outstanding performance. Its ability to generalize well from the training data to unseen data is reflected in the minimal error metrics. BPNN's architecture, while straightforward compared to more complex models like LSTM and GRU, proved that it is highly effective in this context. The CNN model delivered impressive results with an RMSE of 2.1241, MAE of 1.7431, and MAPE of 1.42%. CNN's convolutional layers are adept at identifying local patterns and hierarchies in the data, which can be crucial for detecting trends and short variations. The significant reduction in error metrics highlights CNN's ability to process the multi-dimensional feature space effectively, making it a highly capable model for predicting gold prices.

In addition, LSTM reported an RMSE of 8.7046, MAE of 7.3769, and MAPE of 6.27%. LSTM is designed to capture long-term dependencies in sequential data, addressing some limitations of the standard recurrent neural network. The Bi-LSTM model outperformed LSTM as it achieved an RMSE of 7.314, MAE of 5.8623, and MAPE of 5.00%. By processing the input data in both forward and backward directions, the bidirectional architecture provided a more nuanced understanding of temporal dependencies. This enhanced capability allowed the model to capture more intricate patterns in the data, leading to better predictive performance compared to the unidirectional LSTM and GRU models.

Furthermore, GRU achieved an RMSE of 7.8233, MAE of 6.4011, and MAPE of 5.45%. The moderate performance of the GRU model indicates that it can capture some temporal patterns in the data, but its ability to predict gold prices accurately is limited compared to BPNN and CNN. The Bi-GRU model achieved an RMSE of 9.41, MAE of 8.1916, and MAPE of 6.94%, performing worse than its unidirectional counterpart, GRU. Despite its bidirectional architecture, which processes information in both forward and backward directions to capture more comprehensive temporal relationships, the model exhibited higher error rates. This counterintuitive result suggests that the additional complexity of the bidirectional architecture did not translate to improved performance in this case, possibly due to overfitting or challenges in optimizing the model's parameters effectively.

To show the effect of tuning the neural systems by Optuna on their respective performance, Figure 2 exhibits the rate of change in each value of RMSE, MAE, and MAPE obtained by each system before and after tuning. After hyperparameter tuning using Optuna, a noticeable performance improvement was observed across models. For instance, RMSE, MAE, and MAPE dropped by 33%, 56%, and 51% under BPNN. For Bi-LSTM, RMSE decreased by 15%, and both MAE and MAPE dropped by 19%. However, the performance of CNN did not improve with Optuna. For example, RMSE, MAE, and MAPE increased by 5%, 6%, and 1% respectively.

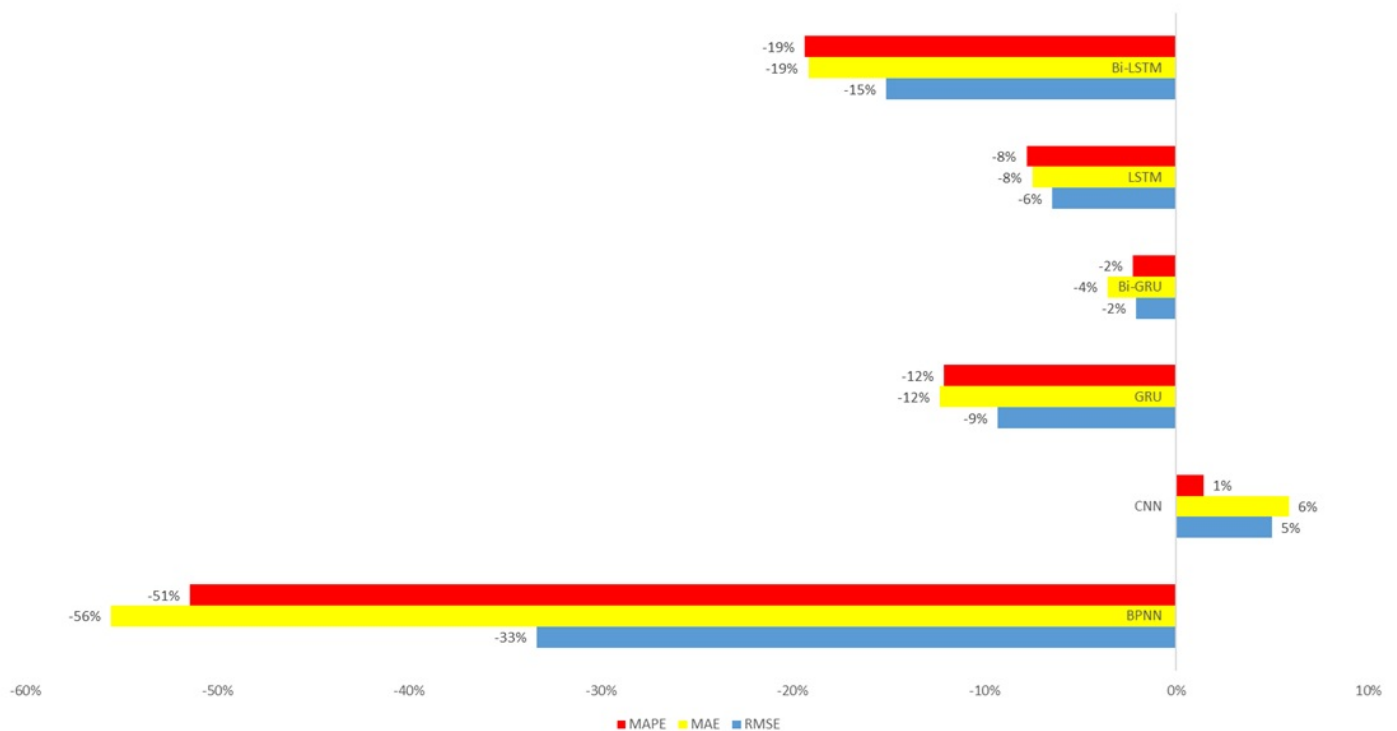


Figure 2. Change in performance measures after applying Optuna. A negative (positive) value indicates a decrease (increase) in the error after tuning the model by using Optuna.

4. Discussion and Conclusions

Accurate gold price predictions serve as a crucial tool for investors to optimize portfolio diversification and safeguarding against inflation or economic uncertainties. Also, mining companies heavily rely on precise gold price forecasts to strategize long-term investment decisions and maximize returns from mining projects. In addition, financial institutions, including banks, hedge funds, and other entities, benefit from accurate gold price predictions to develop effective risk management plans for commodities-related investments. In this regard, the main purpose of the current study is to implement and compare various gold price forecasting systems used to train different artificial neural networks with features selected by the RFE algorithm and using Optuna for optimization. To this end, BPNN, which is a standard artificial neural network, and five deep learning neural networks (CNN, LSTM, Bi-LSTM, GRU, Bi-GRU) are implemented, tested, and compared. It is worth mentioning that in this study, we rely on artificial neural networks, specifically deep learning systems, as they were found to be effective in various forecasting tasks [26–38].

In addition, the recursive feature elimination process effectively identified key features that significantly influence gold prices, such as trading volumes, market indices, and prices of related commodities. This selection ensured that the models were trained with the most relevant data to enhance their predictive accuracy.

The comparative analysis of these models based on RMSE, MAE, and MAPE, reveals several key insights. Firstly, BPNN's superior performance suggests that a simpler architecture can outperform more complex ones, especially when the dataset is well-structured and the feature selection is effective. BPNN's low error rates across all metrics indicate its strong ability to capture the underlying patterns in the data without overfitting. The impressive results of CNN demonstrate the power of convolutional layers in identifying local patterns and trends. Despite being originally designed for image data, CNN's adaptability to time series forecasting underscores its versatility and effectiveness. The model's ability

to process the multi-dimensional feature space efficiently contributed to its high predictive accuracy compared to LSTM, Bi-LSTM, GRU, and Bi-GRU. The performance of LSTM highlights the importance of bidirectional processing in capturing temporal dependencies. Indeed, the Bi-LSTM model outperformed the unidirectional LSTM, indicating that processing information in both directions enhances the model's understanding of temporal dynamics. However, the overall performance of LSTM and Bi-LSTM was still inferior to that of CNN and BPNN, suggesting that while LSTM-based neural networks are powerful for sequential data, they may not always be the best choice for every series forecasting task. GRU-based systems, both unidirectional and bidirectional, exhibited moderate performance, with the bidirectional version surprisingly underperforming compared to the unidirectional version. This result suggests that the additional complexity introduced by the bidirectional architecture may not always lead to improved performance, possibly due to overfitting or difficulties in optimization. In addition, Optuna improved significantly the accuracy of all neural networks considered in this study, except that of CNN, for which MAPE deteriorated by 1% after tuning. The results from the optimization suggest that fine tuning of the neural networks plays an important role in enhancing their performance in forecasting gold price. This finding is in accordance with recent works [26–47] used to show that fine tuning enhances performances of neural networks in forecasting tasks.

It is worth mentioning that like BPNN, all deep learning models implemented in our study are also trained by using the backpropagation algorithm. The latter effectively calculates gradients of the loss function with respect to all model parameters based on the chain rule to propagate the error from the output layer back to the hidden layers. Both standard BPNN and deep learning architectures can learn nonlinearity in data, including trends and short and sudden variations. In addition, deep learning systems are more able to focus on learning deep features such as granular variations and characteristics. The fact that BPNN outperforms deep learning systems can be explained by trend dominance in the data that can be easily learnt by BPNN, while deep learning architectures focus more on deep variations. In addition, compared to BPNN, deep learning architectures are less efficient in terms of computation as they perform multiple computations through layers; furthermore, they require larger datasets to extract deep features and improve learning and accuracy.

In sum, BPNN outperformed advanced deep learning models in this study due to several key advantages. First, BPNN has a simpler architecture, reducing the risk of overfitting when working with smaller datasets. In addition, its relatively low computational complexity compared to deep models such as CNN, LSTM, and GRU allows for faster training and easier implementation, making BPNN more practical for real-time applications in financial forecasting. In this study, BPNN's ability to efficiently capture historical patterns without requiring extensive data pre-processing or advanced feature extraction techniques contributed to its superior performance. This simplicity, combined with its accuracy, makes BPNN a reliable tool for gold price prediction in environments with constrained computational resources or limited technical expertise. However, it should be noted that the outperformance of BPNN compared to deep learning systems is limited to the gold data, current sample, and experimental setup adopted in our work. Furthermore, the additional complexity of the bidirectional architecture in GRU did not translate to improved performance in this case, possibly due to overfitting or challenges in optimizing the model's parameters effectively. Moreover, Optuna is effective in tuning artificial neural networks as accuracy notably increases.

The primary goal of our work was to implement and compare various deep learning systems in forecasting gold price, and the standard BPNN was adopted as the main baseline model. The latter also outperformed standard machine learning models in gold

price prediction including support vector regression (SVR), random forest, and extreme gradient boosting (XGBoost). For instance, while BPNN achieved 0.5928, 0.4091, and 0.34% in terms of RMSE, MAE, and MAPE respectively, SVR achieved 0.7746, 0.6273, and 0.56%, random forest achieved 0.7659, 0.6185, and 0.54%, and XGBoost achieved 0.7626, 0.6163, and 0.50%. These findings confirm the superiority of the standard BPNN in forecasting gold price based on the data sample considered in the study.

The managerial implications of our study are as follows. The accurate forecasts generated by BPNN can guide when to lock in supply agreements or explore alternative sourcing options, especially during market uncertainty. In addition, BPNN-based gold price predictions can aid financial planning and budgeting for procurement teams by forecasting capital needs for production-linked inputs, reducing inventory overstock and minimizing working capital lockup. Furthermore, BPNN's accurate predictions enable better resource planning by aligning production schedules with market demand. This alignment reduces the risks of overproduction during price drops or underproduction during price surges, optimizing inventory levels and financial performance. Moreover, reliable price predictions assist in strategic decision-making, such as timing the sale of gold reserves to maximize revenue or securing favorable financing terms, inform long-term investment decisions in mining infrastructure, technology, and exploration activities, and help companies allocate resources more efficiently. Finally, BPNN-based accurate forecasts can further enhance risk management by supporting hedging strategies, protecting companies from the financial impacts of price volatility.

The study has limitations that warrant discussion. One critical limitation is the reliance on historical data to predict future gold prices. While historical trends offer valuable insights, they fail to account for unpredictable market disruptions, geopolitical events, or macroeconomic shifts. Consequently, the models may produce inaccurate predictions during periods of significant market volatility. Another limitation is the dependence on the quality and relevance of input features. As market conditions evolve, these features may lose their predictive power, which requires constant monitoring and retraining of the models. Additionally, the complexity of financial markets introduces numerous variables that may not be fully captured by the models, even with advanced algorithms. Finally, implementing these models in real-world gold mining operations poses challenges such as data quality variability, computational resource constraints, and the need for skilled personnel to manage and update the models and data effectively.

While the results of this research are promising, there are several avenues for further exploration. First, future research could integrate external factors such as geopolitical events, currency exchange rates, and market sentiment analysis (for example, social media trends) to further enhance forecasting accuracy. These variables often have a profound impact on commodity prices and would provide a more comprehensive view of price fluctuations. Second, implementing real-time forecasting systems using online learning techniques could help gold mining companies and investors adjust more dynamically to market changes. This would be particularly beneficial in environments where price shifts occur rapidly and unpredictably. Third, expanding the scope of forecasting beyond gold to include other commodities such as silver or platinum could provide a holistic view of how the broader precious metals market behaves, offering companies more strategic insights. Fourth, developing scenario-based forecasting systems that simulate different market conditions, such as a sudden gold price surge or drop, could prepare companies to respond to various contingencies and improve their resilience.

Author Contributions: A.P.: Conceptualization, methodology, data curation, writing—review and editing. S.L.: Conceptualization, review. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data source is provided in reference [25].

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BPNN	Backpropagation neural networks
CNN	Convolutional neural networks
GRU	Gate recurrent units
Bi-GRU	Bidirectional GRU
LSTM	Long short-term memory neural networks
Bi-LSTM	Bidirectional LSTM
RMSE	Root mean of squared errors
MAE	Mean absolute error
MAPE	Mean absolute percentage error
RFE	Recursive feature elimination

References

1. García-Gonzalo, E.; García-Nieto, P.J.; Valverde, G.F.; Fernández, P.R.; Lasheras, F.S.S.; Gómez, S.L.S. Hybrid DE-Optimized GPR and NARX/SVR Models for Forecasting gold spot Prices: A Case Study of the Global Commodities market. *Mathematics* **2024**, *12*, 1039. [CrossRef]
2. Li, J.; Wang, Y.; Song, Y.; Su, C.W. How resistant is gold to stress? New evidence from the global supply chain. *Resour. Policy* **2023**, *85*, 103960. [CrossRef]
3. Aye, G.; Gupta, R.; Hammoudeh, S.; Kim, W.J. Forecasting the price of gold using dynamic model averaging. *Int. Rev. Financ. Anal.* **2015**, *41*, 257–266. [CrossRef]
4. Khani, M.M.; Vahidnia, S.; Abbasi, A. A Deep Learning-Based Method for Forecasting Gold Price with Respect to Pandemics. *SN Comput. Sci.* **2021**, *2*, 335. [CrossRef]
5. Roslan, N.H.A.; Halim, N.A. Forecasting world gold price in year 2022 using Geometric Brownian Motion model. *AIP Conf. Proc.* **2024**, *2905*, 050003. [CrossRef]
6. Lahmirmi, S. A hybrid CNN-LSTM-GRU system tuned by whale optimization algorithm for enhanced gold and Brent price prediction. *Financ. Res. Open* **2026**, *2*, 100100. [CrossRef]
7. Rumelhart, D.E.; McClelland, J.L. *Parallel Distributed Processing, Explorations in the Microstructure of Cognition*; Foundations, MIT Press: Cambridge, MA, USA, 1986; Volume 1.
8. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [CrossRef] [PubMed]
9. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
10. Cho, K.; van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 25–29 October 2014; pp. 1724–1734.
11. Seyedan, M.; Mafakheri, F.; Wang, C. Order-up-to-level inventory optimization model using time-series demand forecasting with ensemble deep learning. *Supply Chain Anal.* **2023**, *3*, 100024. [CrossRef]
12. Tadayonrad, Y.; Ndiaye, A.B. A new key performance indicator model for demand forecasting in inventory management considering supply chain reliability and seasonality. *Supply Chain Anal.* **2023**, *3*, 100026. [CrossRef]
13. Abbasi, M.; Vahidnia, M.H.; Khani, A. Multivariate time-series blood donation/demand forecasting for supply chain management. *J. Adv. Intell. Manuf.* **2022**, *3*, 45–56.
14. Zhang, C.; Tian, Y.-X. Solving data-driven newsvendor problem with textual reviews through deep learning. *Soft Comput.* **2023**, *27*, 4967–4986. [CrossRef]
15. Moslehi, S.; Ghasemi, P.; Aghaei, M. Deep reinforcement learning and optimization approach for multi-echelon supply chain with uncertain demands. *Comput. Ind. Eng.* **2024**, *183*, 108122.
16. Guner, M.; Sahin, G. A mixed methods approach to analyze and predict supply disruptions. *Eur. J. Oper. Res.* **2022**, *300*, 121–135.
17. Wang, Y.; Liu, Z.; Ma, X. Deep reinforcement learning for dynamic order picking in warehouses. *Comput. Oper. Res.* **2025**, *155*, 105140.

18. Ghezavati, V.R.; Shafiee, M.; Aghajani, A. A deep learning approach for vehicle routing problem with time windows. *Comput. Oper. Res.* **2025**, *155*, 105115.
19. Dehnavi, E.; Niaki, S.T.A. A deep learning approach for tactical production planning. *Comput. Oper. Res.* **2025**, *155*, 105116.
20. Rettinger, M.; Minner, S.; Birzl, J. Understand your decision rather than your model prescription: Towards explainable deep learning approaches for commodity procurement. *Comput. Oper. Res.* **2024**, *175*, 106905. [\[CrossRef\]](#)
21. Graves, A.; Schmidhuber, J. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Netw.* **2005**, *18*, 602–610. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Vukotić, V.; Raymond, C.; Gravier, G. A Step Beyond Local Observations with a Dialog Aware Bidirectional GRU Network for Spoken Language Understanding. *Proc. Interspeech* **2016**, 3241–3244. [\[CrossRef\]](#)
23. Guyon, I.; Weston, J.; Barnhill, S.; Vapnik, V. Gene selection for cancer classification using support vector machines. *Mach. Learn.* **2002**, *46*, 389–422.
24. Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*; ACM: New York, NY, USA, 2019; pp. 2623–2631. [\[CrossRef\]](#)
25. Yahoo Finance. Available online: <https://ca.finance.yahoo.com/> (accessed on 7 October 2023).
26. Benghiat, S.; Lahmiri, S. Tuning for Precision Forecasting of Green Market Volatility Time Series. *Stats* **2026**, *9*, 12. [\[CrossRef\]](#)
27. Kachwaha, S.; Lahmiri, S. Improving Deep Learning Models by Bayesian Optimization to Predict Crude Oil Prices. *Algorithms* **2025**, *18*, 762. [\[CrossRef\]](#)
28. Taganova, G.; Tussupov, J.; Abdildayeva, A.; Kaldarova, M.; Kazi, A.; Simpson, R.C.; Zakirova, A.; Nurbekov, B. HST-MB-CREH: A Hybrid Spatio-Temporal Transformer with Multi-Branch CNN/RNN for Rare-Event-Aware PV Power Forecasting. *Algorithms* **2026**, *19*, 94. [\[CrossRef\]](#)
29. Taganova, G.; Zakirova, A.; Abdildayeva, A.; Nurbekov, B.; Akhayeva, Z.; Azykanov, T. Sub-Hourly Multi-Horizon Quantile Forecasting of Photovoltaic Power Using Meteorological Data and a HybridCNN-STTransformer. *Algorithms* **2026**, *19*, 123. [\[CrossRef\]](#)
30. Luo, S.; Meng, X.; Pang, X.; Li, H.; Zheng, Z. Optimized Hybrid Deep Learning Framework for Short-Term Power Load Interval Forecasting via Improved Crowned Crested Porcupine Optimization and Feature Mode Decomposition. *Algorithms* **2025**, *18*, 659. [\[CrossRef\]](#)
31. Wang, Z.; Lahmiri, S.; Bekiros, S. Machine Learning Systems Tuned by Bayesian Optimization to Forecast Electricity Demand and Production. *Algorithms* **2025**, *18*, 695. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Aikodon, N.; Ortega-Martorell, S.; Olier, I. Predicting Decomensation Risk in Intensive Care Unit Patients Using Machine Learning. *Algorithms* **2024**, *17*, 6.
33. Zhang, S.; Li, S.; Li, X.; Yao, Y. Representation of Traffic Congestion Data for Urban Road Traffic Networks Based on Pooling Operations. *Algorithms* **2020**, *13*, 84. [\[CrossRef\]](#)
34. Yaghi, M.A.; Al-Omari, H. Physics-Aware Deep Learning Framework for Solar Irradiance Forecasting Using Fourier-Based Signal Decomposition. *Algorithms* **2026**, *19*, 81. [\[CrossRef\]](#)
35. Davila Hernandez, C.; Ho, J.; Kim, D.; Oubeidillah, A. Machine-Learning-Based Model for Hurricane Storm Surge Forecasting in the Lower Laguna Madre. *Algorithms* **2023**, *16*, 232. [\[CrossRef\]](#)
36. Antonio Figueroa, A.E.; Lahmiri, S. Tuning Deep Learning for Predicting Aluminum Prices Under Different Sampling: Bayesian Optimization Versus Random Search. *Entropy* **2026**, *28*, 145. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Foroutan, P.; Lahmiri, S. Deep learning systems for forecasting the prices of crude oil and precious metals. *Financ. Innov.* **2024**, *10*, 111. [\[CrossRef\]](#)
38. Bahareh, A.; Lahmiri, S. Hybrid deep learning and GARCH-family models for forecasting volatility of cryptocurrencies. *Mach. Learn. Appl.* **2023**, *12*, 100465. [\[CrossRef\]](#)
39. Zheng, C.; Eskandari, M.; Li, M.; Sun, Z. GA-Reinforced Deep Neural Network for Net Electric Load Forecasting in Microgrids with Renewable Energy Resources for Scheduling Battery Energy Storage Systems. *Algorithms* **2022**, *15*, 338. [\[CrossRef\]](#)
40. Zhang, Y.; Xu, H.; Li, Y.; Lin, G.; Zhang, L.; Tao, C.; Wu, Y. An Integer-Fractional Gradient Algorithm for Back Propagation Neural Networks. *Algorithms* **2024**, *17*, 220. [\[CrossRef\]](#)
41. Purnamasari, P.D.; Ratna, A.A.P.; Kusumoputro, B. Development of Filtered Bispectrum for EEG Signal Feature Extraction in Automatic Emotion Recognition Using Artificial Neural Networks. *Algorithms* **2017**, *10*, 63. [\[CrossRef\]](#)
42. Zhao, S.; Ma, Y.; Leng, D. An Intelligent Artificial Neural Network Modeling of a Magnetorheological Elastomer Isolator. *Algorithms* **2019**, *12*, 195. [\[CrossRef\]](#)
43. Wei, C.; Liu, T.; Zhang, B.; Wang, X. A Feature Selection Method for Yarn Quality Prediction Based on SHAP Interpretation. *Algorithms* **2026**, *19*, 266. [\[CrossRef\]](#)
44. Chen, J.; Zhu, H.; Shu, T.; Cao, C.; Deng, Y.; Cheng, Q. Balanced Grey Wolf Optimizer Algorithm for Backpropagation Neural Networks. *Mathematics* **2026**, *14*, 554. [\[CrossRef\]](#)

45. Qin, S.; Cao, Z.; Wang, F.; Ngu, S.S.; Kho, L.C.; Cai, H. Design of Optimal Pitch Controller for Wind Turbines Based on Back-Propagation Neural Network. *Energies* **2024**, *17*, 4076. [[CrossRef](#)]
46. Pan, Y.; Yu, Y.; Zhou, J.; Qin, W.; Wang, Q.; Wang, Y. Research on Performance Predictive Model and Parameter Optimization of Pneumatic Drum Seed Metering Device Based on Backpropagation Neural Network. *Appl. Sci.* **2025**, *15*, 3682. [[CrossRef](#)]
47. Shao, J.-J.; Li, L.-B.; Yin, G.-J.; Wen, X.-D.; Zou, Y.-X.; Zuo, X.-B.; Gao, X.-J.; Cheng, S.-S. Prediction of Compressive Strength of Fly Ash-Recycled Mortar Based on Grey Wolf Optimizer–Backpropagation Neural Network. *Materials* **2025**, *18*, 139. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.