International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 • Website: www.ijfmr.com • Email: editor@ijfmr.com

Comparative Performance of Neural Network Architectures in Gold Price Prediction: A Study of GRU, N-BEATS, LSTM, and Hybrid Models

Mr. Gonese Tapiwa Anthony[1], Mr. Nana Adi Appiah[2], Dr. Dongxiao Ren[3]

[1,2,3]Zhejiang University of Science and Technology

ABSTRACT
This study assesses the efficacy of a variety of neural network models in predicting gold prices, with a particular emphasis on GRU, N-BEATS, N-BEATS-GRU, N-BEATS-LSTM, and LSTM. We evaluated the models using a comprehensive dataset and metrics such as MSE, MAE, MAPE, and RMSE. The GRU model demonstrated the highest MAPE of 3.10%, as indicated by the results. N-BEATS was the second-best-performing model. In stark contrast to expectations, hybrid models, notably N-BEATS-LSTM, underperformed. Compared to LSTM-based models, visual analysis demonstrated that GRU and N-BEATS more effectively captured short-term and long-term trends. The results indicate that simplified models, such as GRU, may be more effective than complex hybrids in predicting gold prices.

Keywords: Gold Price Prediction, Time Series Forecasting, Neural Networks, GRU, N-BEATS, LSTM, Hybrid Models

Introduction
Investors have historically considered gold to be a secure refuge, particularly during periods of economic instability, due to its inherent value. A diverse array of factors, including inflation rates, geopolitical events, global economic conditions, and currency fluctuations, influence its price fluctuations [1]. Investors, policymakers, and researchers are keenly interested in the precise prediction of gold prices, as it can offer valuable insights into market trends and facilitate strategic financial planning [2].
Fundamental analysis, which considers economic data and market sentiment, and technical analysis, which employs statistical tools and previous price patterns, have been the conventional methods of predicting gold prices [3].
Nevertheless, the application of sophisticated algorithms to enhance prediction accuracy has become increasingly prevalent as a result of the introduction of potent machine learning techniques [4]. There are a lot of promising results from recurrent neural networks (RNNs) like Gated Recurrent Units (GRU) and Long Short-Term Memory (LSTM), as well as new architectures like Neural Basis Expansion Analysis Time Series Forecasting (N-BEATS) [5].

LSTM and GRU models are particularly adept at identifying long-term dependencies and coping with sequential data, which makes them well-suited for time series forecasting tasks such as predict gold prices. N-BEATS, a recent addition to the time series forecasting framework, has outperformed other forecasting benchmarks by utilizing a deep learning technique that emphasizes interpretable basis expansion [8].

Current State of Research

Over the last few decades, generous scholars have worked tirelessly to develop better and more exact models for predicting the gold price [35]. Some researchers employ statistical methods such as the moving average (MA) [36], autoregressive conditional heteroskedasticity (ARCH) [37], generalized autoregressive conditional heteroskedasticity (GARCH) [38,39], threshold autoregressive conditional heteroskedasticity. The above models use precise formulas to explain a linear relationship between variables [40, 41]. The gold price series, on the other hand, exhibits feature such as fuzzy information, extended memory, nonlinearity, and complexity [42,43]. The standard statistical approach was incapable of producing a reasonably accurate prediction effect [44].

As a result, other researchers employ clever approaches to forecast the gold price. Intelligent approaches can overcome the limits of statistical methods and improve nonlinear system prediction. For example, researchers have discovered that machine learning methods like support vector regression (SVR) can outperform classical prediction models [45,46,47]. Meanwhile, deep learning models not only outperform machine learning models in terms of learning and adaptation, but they can also better forecast and evaluate time series [48].

Thus, deep learning models such as artificial neural networks (ANN) [49,50], back propagation neural networks (BPNN) [51], and other methods [52,53] have been widely used. According to Kristjanpoller and Minutolo [54], the error of solely utilizing the GARCH model to predict the gold price is usually rather significant. As a result, they employed the hybrid ANN-GARCH model to forecast gold spot and futures price fluctuations.

The results reveal that, when compared to the GARCH technique, the prediction of ANN-GARCH is generally enhanced, with a 25% reduction in average percentage error. This demonstrates the significant benefit of incorporating deep learning algorithms into standard models [55]. Furthermore, the deep learning method can improve the model's training speed and efficiency [56].

Long-term effect analysis, in particular, is critical for time series prediction and analysis [57]. For example, the recurrent neural network (RNN) can save previous period information to analyse the present period [58, 18]. Long short-term memory (LSTM) is a better version of RNN [60]. LSTM has both short-term and long-term memory [61]. As a result, the LSTM is ideal for forecasting time series. In contrast with traditional econometrics methods such as vector autoregressive (VAR), support vector machines (SVM) [62], random forest (RF) [63], multi-layer perceptron (MLP) and pseudo-random model, the prediction results of LSTM are more accurate [64,65,66]. A robust prediction model, on the other hand, must extract useful data features [67]. Convolutional neural networks (CNN) are capable of extracting feature information and anticipating prices [68].

In recent years, the field of time series forecasting has seen significant advancements, particularly with the integration of machine learning and deep learning techniques. Predicting gold prices, a critical task for investors and policymakers, has benefited from these advancements. Several models, such as N-BEATS, GRU, LSTM, and hybrid models, have been developed and tested to improve the accuracy and reliability of gold price predictions.

Applications in Financial Forecasting

The application of these advanced models to financial forecasting has been extensive. Studies have shown that models like LSTM and N-BEATS can effectively predict stock prices, exchange rates, and commodity prices, including gold [5][7][8]. The ability of these models to handle non-linearities and capture long-term dependencies makes them particularly suitable for financial time series forecasting.

Challenges and Future Directions

Despite the advancements, several challenges remain in the field of gold price prediction. These include the inherent volatility of financial markets, the need for large amounts of high-quality data, and the interpretability of complex models. Future research directions include improving model robustness, enhancing interpretability, and developing more sophisticated hybrid models that can better capture the intricate dynamics of financial time series.

In summary, the current state of research in gold price prediction using machine learning and deep learning models is robust and rapidly evolving. Models such as N-BEATS, LSTM, GRU, and their hybrids offer promising avenues for enhancing prediction accuracy and reliability, thereby aiding investors and policymakers in making informed decisions.

RESEARCH DESIGN AND DATASET

The objective of this research article is to evaluate and contrast the efficacy of advanced machine learning models, including N-BEATS, GRU, and LSTM, as well as hybrid models that integrate N-BEATS with GRU and LSTM, in the prediction of gold prices. The project's objective is to combine the N-BEATS framework with the strengths of recurrent neural networks in order to identify the most accurate and reliable model for forecasting gold prices. Three distinct contributions are made by this work:

Model Implementation and Comparison: The performance of the N-BEATS, GRU, LSTM, N-BEATS-GRU hybrid, and N-BEATS-LSTM hybrid models for gold price prediction is implemented and rigorously tested [9][10].

Hybrid Model Development: Hybrid models that combine the interpretability and basis expansion capabilities of N-BEATS with the sequential learning strengths of GRU and LSTM are being developed and tested with the hypothesis that they could potentially provide improved predictive performance [11][12].

Empirical research: Conducting a comprehensive empirical analysis of historical gold price data to verify the efficacy of each model, thereby elucidating their practical utility and predictive power in financial forecasting [13][14].

Our objective is to enhance the comprehension of machine learning applications in financial markets and to contribute to the ongoing endeavors to enhance the accuracy and robustness of gold price predictions through this comprehensive study.

Comparative Studies

The efficacy of various machine learning models for time series forecasting has been the subject of numerous empirical studies. Ahmed et al. [12] conducted a thorough analysis of machine learning models, including neural networks, on a variety of time series datasets, emphasizing the advantages and disadvantages of each method. In the same vein, Siami-Namini et al. [10] conducted a comparison between ARIMA and LSTM models, illustrating the superiority of LSTM in capturing intricate patterns in time series data.

Dataset and Data Preprocessing

The dataset utilized in this investigation was acquired from the World Gold Council [19]. This dataset comprises three CSV files (daily, monthly, and yearly) that contain gold prices from January 1978 to May 2024. The United States dollar is the currency. The focus of this investigation was on monthly data that concluded in May 2024. The following considerations underpin my recommendation for utilizing the monthly dataset:

1. **Reduced Volatility and Noise**: Monthly data is less volatile and less chaotic than daily data. Adding disturbance to the dataset, intraday trading activities, news events, or short-term market movements can all affect daily gold prices. The predictive accuracy of the model is improved by aggregating data to monthly intervals, which allows it to concentrate on more significant trends and patterns [20].
2. **Smoothing Seasonal Fluctuations**: The model is able to more effectively capture the underlying trends by smoothing out short-term seasonal fluctuations using monthly data. This method is beneficial when the primary objective is to understand and anticipate the long-term fluctuations in gold prices. A more comprehensive understanding of the market dynamics is achieved by smoothing out these fluctuations [21].
3. **Computational Efficiency**: The computational cost of training models on daily data can be high, especially if the dataset spans multiple years. The quantity of data points is reduced by monthly data, which enables more computational efficiency and quicker model training and experimentation. This efficacy is essential for the development of iterative models and practical applications [22].
4. **Appropriateness for Medium to Long-Term Forecasting**: Monthly data is frequently suitable for predicting across medium to long time horizons. Despite the fact that daily data may be too detailed to predict monthly patterns, yearly data may overlook short-term fluctuations that are economically

significant. Consequently, monthly data achieves a balance by identifying substantial trends without being excessively detailed or overly broad [23].

5. **Data Availability and Consistency**: Daily gold price data may be unavailable for extended periods and may contain gaps or absent numbers. Monthly data is more likely to be accessible consistently over extended periods, which enables a more rigorous investigation. This consistency is essential for the creation of predictive models that are dependable and can be implemented over an extended period of time [24].
6. **Reduced Risk of Overfitting**: The model may learn noise in the data rather than genuine trends, which can result in overfitting when dealing with daily data.. By combining information, monthly data may reduce the risk of overfitting and enhance the model's generalization to previously unseen data. The model can enhance its performance on new, unseen data by emphasizing more stable and substantial trends [13].

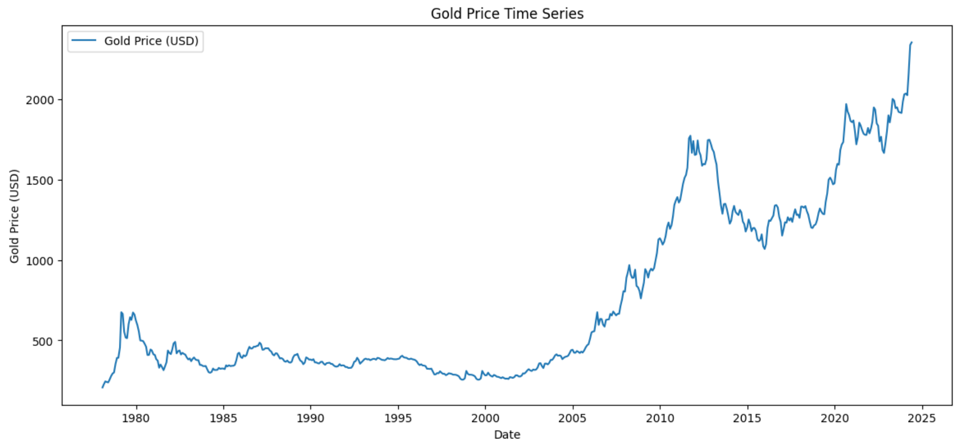

Figure 1. Graph of the Gold price over 1978 to 2024.

The gold price time series chart from 1978 to present shows an upward trend in gold prices, despite fluctuations. The chart shows significant volatility, indicating the dynamic nature of the gold market. The most pronounced peak occurred in 2011-2012, when gold prices surpassed $1,800 per ounce. The chart also shows price corrections or consolidations after the main peaks. Over 40 years, this historical perspective provides valuable insights into gold price fluctuations and future price behavior.

Models Involved in the Study and Hybrid Model Integration

This study encompasses a variety of sophisticated machine learning models that are designed to forecast gold prices, including N-BEATS, GRU, LSTM, and their hybrid versions, N-BEATS-GRU and N-BEATS-LSTM. The hybrid models are designed to capitalize on the synergistic benefits of each model's distinct strengths in time series forecasting.

N-BEATS Model

A deep learning model that is specifically designed for univariate time series forecasting is the Neural Basis Expansion Analysis Time Series Forecasting (N-BEATS) model. The time series is decomposed into interpretable components using a deep stack of fully connected layers with basis expansion in N-BEATS, which was developed by Oreshkin et al. [9]. Due to its capacity to extract intricate patterns and trends from the data, this model has demonstrated exceptional performance in a variety of forecasting competitions. The N-BEATS model has emerged as a highly effective tool for forecasting time series. To capture intricate patterns in time series data, N-BEATS employs a deep learning approach that emphasizes basis expansion and a backcast/forecast strategy. This model has exhibited exceptional performance in a variety of benchmarks, such as the M4 competition, a prominent forecasting competition in the field [14]. The interpretable nature of N-BEATS also enables a more comprehensive comprehension and insight into the forecasted time series.

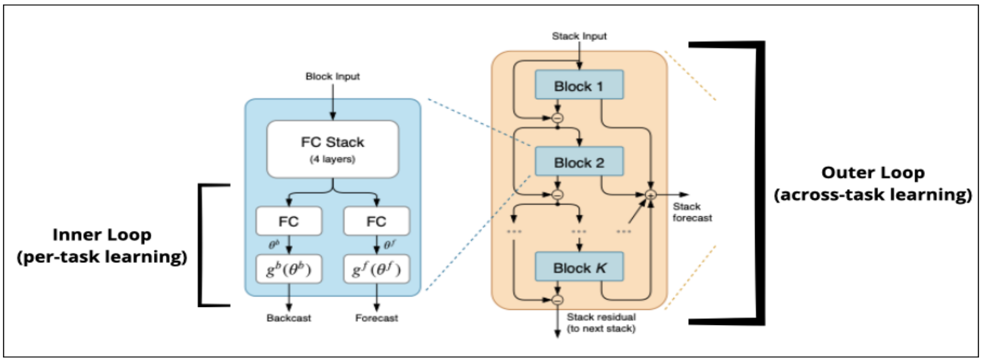

Figure 2. The NBEATS model architecture.

N-BEATS is composed of an array of fully connected (FC) blocks, each of which contains two sub-blocks: a forecast block and a backcast block. The backcast block endeavors to reconstruct the input series, capturing residuals that are transmitted to subsequent blocks, while the forecast block generates future predictions. This process of iterative refinement improves the model's capacity to grasp intricate patterns in the data.

GRU Model

Cho et al. [26] introduced the Gated Recurrent Unit (GRU) as a type of recurrent neural network (RNN). GRUs are engineered to reliably capture long-term dependencies and manage sequential data. They are computationally less expensive than LSTM networks and provide comparable performance, rendering them appropriate for time series forecasting tasks.

GRU Equations:

$Z_t = \sigma\left(W_z \cdot \left[h_{\{t-1\}}, x_t\right]\right)$

$r_t = \sigma\left(W_r \cdot \left[h_{\{t-1\}}, x_t\right]\right)$

$\widetilde{\{h\}}_t = \backslash\mathrm{tanh}\left(W \cdot \left[r_t * h_{\{t-1\}}, x_t\right]\right)$

$h_t = (1 - z_t) * h_{\{t-1\}} + z_t * \widetilde{\{h\}}_t$

LSTM Model

Another RNN type that was developed to resolve the vanishing gradient issue in conventional RNNs is the Long Short-Term Memory (LSTM) network, which was introduced by Hochreiter and Schmidhuber [18]. LSTMs are extremely effective for time series forecasting, as they have the ability to learn and remember long-term dependencies, which is essential for capturing temporal patterns over long periods.

LSTM Equation:

$f_t = \sigma(W_f \cdot [h_{\{t-1\}}, x_t])$

$i_t = \sigma(W_i \cdot [h_{\{t-1\}}, x_t])$

$\{\widetilde{C}\}_t = \backslash\text{tanh}(W_C \cdot [h_{\{t-1\}}, x_t])$

$C_t = f_t * C_{\{t-1\}} + i_t * \{\widetilde{C}\}_t$

$o_t = \sigma(W_o \cdot [h_{\{t-1\}}, x_t])$

$h_t = o_t * \tanh(C_t)$

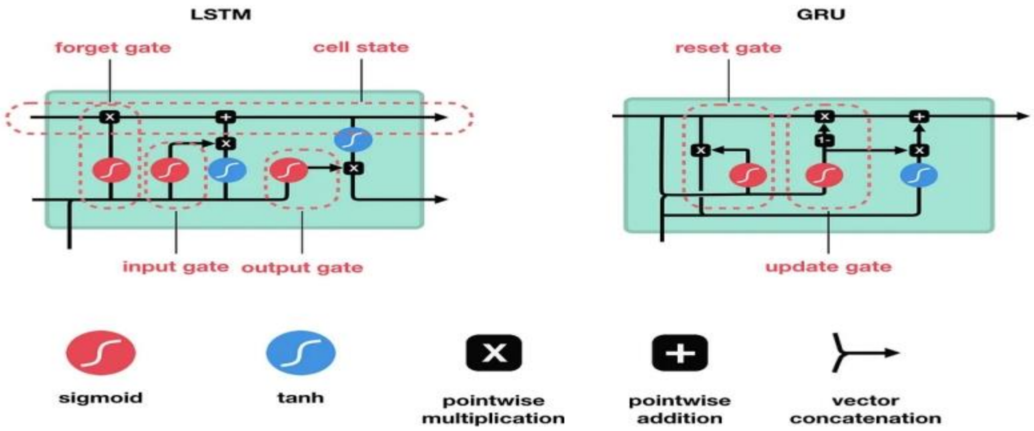

Figure 3. Architectural structures of LSTM and GRU

Recurrent Neural Networks (RNNs) have been extensively utilized for sequential data due to their ability to encapsulate temporal dependencies. Two popular variants of RNNs that are capable of learning long-term dependencies in data and resolve the vanishing gradient problem are Long Short-Term Memory (LSTM) networks, introduced by Hochreiter and Schmidhuber [18], and Gated Recurrent Units (GRU), proposed by Cho et al. These models have been extensively utilized in the prediction of financial time series, such as gold prices, as a result of their ability to manage sequential data [6][7][8].

Hybrid Models: N-BEATS-GRU and N-BEATS-LSTM

Hybrid models optimize forecasting capabilities by integrating the advantages of various architectures. The N-BEATS-GRU and N-BEATS-LSTM composites are suggested in this study to leverage the sequential learning strengths of GRU and LSTM networks with the interpretability and basis expansion capabilities of N-BEATS, respectively.

N-BEATS-GRU Hybrid Model

The N-BEATS-GRU hybrid model combines the N-BEATS architecture with a GRU layer. The N-

BEATS component is responsible for the decomposition of the time series into interpretable components, which captures overall trends and seasonality [9]. The sequential dependencies in the residuals or errors left by the N-BEATS component are subsequently modeled using the GRU layer. This hybrid approach enables the model to learn both global patterns and local temporal dependencies effectively, potentially enhancing forecast accuracy.

N-BEATS-LSTM Hybrid Model

In the same vein, the N-BEATS-LSTM hybrid model integrates an LSTM layer with the N-BEATS architecture. The N-BEATS model initially captures the primary trends and seasonal components of the time series [9]. The LSTM layer is then supplied the residuals from the N-BEATS component, which is particularly adept at capturing long-term dependencies and sequential patterns [18]. The objective of this combination is to improve the model's capacity to learn intricate temporal dynamics, thereby delivering more precise and resilient forecasts.

The trend in time series forecasting is to combine various models in order to capitalize on their distinct capabilities. Hybrid models that incorporate N-BEATS with RNNs, including GRU and LSTM, have demonstrated optimistic outcomes. These hybrid models are designed to combine the sequential learning strengths of GRU and LSTM with the interpretability and basis expansion capabilities of N-BEATS, with the potential to provide improved predictive performance. For example, hybrid methods that integrate recurrent neural networks and exponential smoothing have demonstrated enhanced accuracy in a variety of forecasting tasks [13].

Integration of Hybrid Models

The hybrid models' integration procedure is comprised of two primary steps:

Decomposition with N-BEATS: The N-BEATS model is initially applied to the time series in order to extract the primary components, including seasonality and trend. By isolating the residuals that represent the series' more difficult-to-predict components, this phase simplifies the time series.

Sequential Modeling with GRU/LSTM: The residuals from the N-BEATS model are subsequently transferred to the GRU or LSTM layer. This layer concentrates on the acquisition of sequential patterns within these residuals, which are frequently the irregular components and short-term fluctuations that the N-BEATS model failed to fully capture.

Through the integration of these methodologies, the hybrid models can capitalize on the interpretability and trend-capturing capabilities of N-BEATS, as well as the robust sequence learning capabilities of GRU and LSTM networks. It is anticipated that this hybrid strategy will improve the overall predictive performance, particularly in the area of capturing both long-term trends and short-term fluctuations in gold prices.

Data Preprocessing and Model Training

The process of data preprocessing is essential for the development of any machine learning model, as it guarantees that the data is in a format that is suitable for training and enhances the model's performance. A series of preprocessing procedures were implemented in this investigation, which was succeeded by the models' training.

Data Loading and Formatting

The dataset utilized in this investigation was downloaded from the website of the World Gold Council and imported into a Pandas Data Frame. In order to facilitate time series analysis, the date column was converted to datetime format and the data was configured to utilize the date as the index [19].

Data Normalization

The MinMaxScaler from scikit-learn was employed to normalize the data in preparation for training. Normalization is indispensable for neural networks, as it reduces the data to a scale that the model can more efficiently manage, typically between 0 and 1. This procedure expedites convergence during training and guarantees that the model weights are updated adequately [6].

Train-Test Split

The dataset was subsequently partitioned into training and testing sets, with 80% of the data being utilized for training and the remaining 20% for testing. This division is a standard procedure for assessing the model's performance on unseen data, thereby ensuring that the model generalizes effectively and is not overfitted to the training data [10].

Creating Datasets for LSTM and GRU Models

It is crucial to generate datasets that contain a sequence of previous observations in order to forecast the subsequent value for sequential models such as LSTM and GRU. A function was developed to generate these datasets, which consist of sequences of n_steps observations. The models are able to learn from the temporal dependencies in the data as a result of this approach [18].

Environment

The time series forecasting models were trained in this study using a univariate dataset that contained the variables 'Date' and 'USD'. The dataset was loaded, the data was normalized, and sequences were generated for model input during the data preprocessing procedure. The models were able to learn effectively from the temporal patterns in the gold price data as a result of these steps.

Hardware and Software Specifications

The experiments were conducted on Google Colab, a cloud-based environment that offers access to high-performance GPUs. The computational power available in the cloud was leveraged to considerably reduce the training time for the neural network models in this setup.

Python Libraries and Frameworks

Python 3.8.7 was employed to implement the algorithms and train the model. The subsequent libraries were implemented:

PyTorch 1.8.1: PyTorch 1.8.1 was employed to construct and train neural network models, including N-BEATS, GRU, LSTM, and hybrid models. PyTorch's extensive support for neural network operations and dynamic computational graph rendered it appropriate for this endeavor [27].

scikit-learn 0.24.1: The scikit-learn library was employed to perform data preprocessing tasks, including normalization and train-test division. The gold price data was scaled to a range of 0 to 1 using the MinMaxScaler from scikit-learn [28].

matplotlib 3.3.4: This library was employed to visualize the data and the model's performance. The model predictions were evaluated and the trends and patterns in the gold price data were better understood through the use of graphs and plots generated with matplotlib [29].

RESULTS

The gold price dataset used in this study consists of 557 data points representing monthly gold prices in USD from 1950 to 2023. Below is a detailed statistical summary of the dataset:

	USD
count	557.000000
mean	757.954093
std	544.707055
min	207.800000
25%	352.700000
50%	423.400000
75%	1236.000000
max	2352.140000

The average gold price over this period is $757.95. This provides a central value around which the gold prices have fluctuated. The standard deviation is $544.71, indicating high variability in the monthly gold prices. This suggests significant fluctuations in the price of gold over the years. At the 25th percentile, the gold price is $352.70. This means that 25% of the observed gold prices were below $352.70. The median gold price is $423.40, indicating that half of the observed prices are below this value and half are above. The median provides a robust central tendency measure, especially in the presence of outliers. The price at the 75th percentile is $1236.00, showing that 75% of the prices were below this value, while the top 25% were higher. The descriptive statistics reveal a right-skewed distribution of gold prices, indicated by the fact that the mean ($757.95) is greater than the median ($423.40). This skewness suggests that there are some significantly high values (outliers) that raise the average gold price. The large standard deviation ($544.71) and the wide range between the minimum ($207.80) and maximum ($2352.14) prices highlight the extensive spread and variability in the data. The IQR, which is the range between the 25th percentile ($352.70) and the 75th percentile ($1236.00), is $883.30. This further indicates the wide dispersion and suggests that the middle 50% of the gold prices are spread over a large range.

The comparative analysis of five distinct models for gold price prediction provided substantial insights into their relative performance. The models that were assessed were GRU, N-BEATS, N-BEATS-GRU Hybrid, LSTM, and N-BEATS-LSTM Hybrid. Conventional metrics, such as Mean Squared Error (MSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE), were employed to assess their efficacy. Below is a thorough analysis of their performance.

GRU Model

The GRU model had the lowest MSE, indicating the smallest average squared errors between the predicted and actual gold prices. The GRU's MAE was also the lowest, showing that, on average, its predictions deviated by $50.38 from the actual prices. The GRU model achieved the best MAPE, with an average error of 3.10% relative to the actual values. The lowest RMSE among the models, indicating typical prediction errors of around $67.71.

The GRU model exhibited superior performance in all metrics (MSE: 4584.6374, MAE: 50.3836, MAPE: 3.10%, RMSE: 67.7100). Below is the graph of the actual gold prices and the predicted GRU prices.

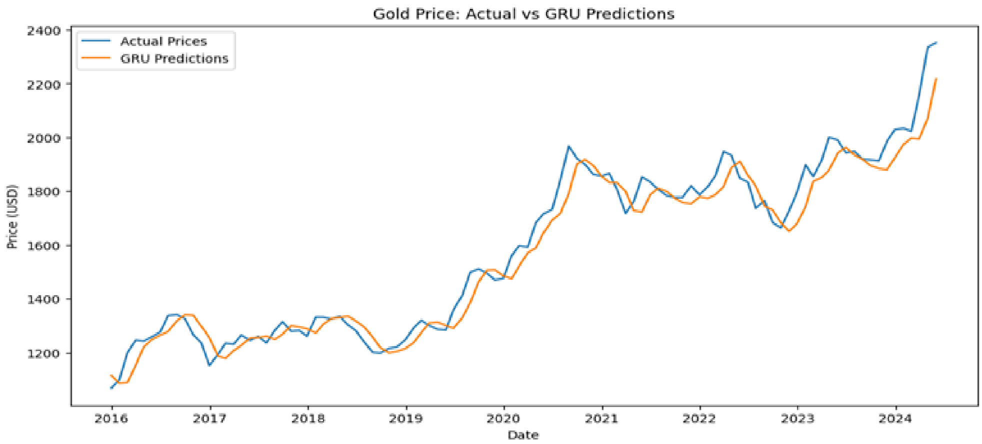

Figure 4. Graph of Actual vs GRU predictions.

The N-BEATS model followed closely behind the GRU, with a higher but still relatively low MSE. The average error was $60.29, showing greater deviation from actual prices compared to GRU. A MAPE of 3.73%, indicating a slightly higher relative error than the GRU model. An RMSE of 76.6820, reflecting typical prediction errors of about $76.68.

It was followed closely by the N-BEATS model (MSE: 5880.2638, MAE: 60.2945, MAPE: 3.73%, RMSE: 76.6820). Below is the graph of the actual gold prices and the predicted NBEATS prices.

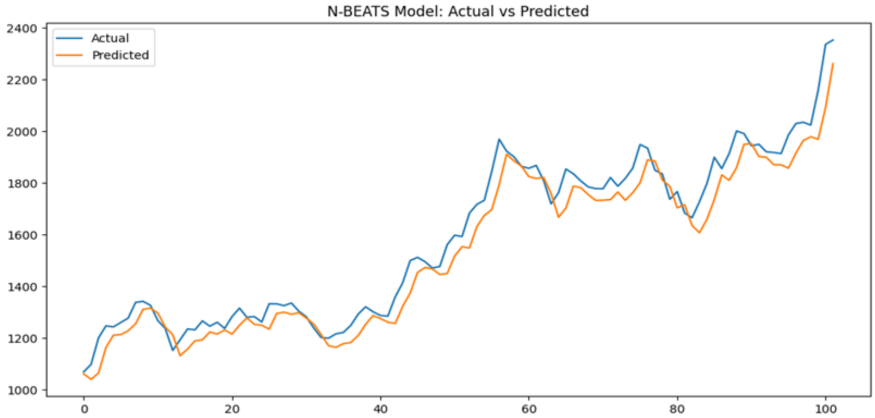

Figure 5. Graph of Actual vs NBEATS predictions.

The hybrid model (NBEATS-GRU model) exhibited higher MSE, indicating larger average squared errors. The MAE was $69.78, reflecting significant deviations from actual prices. A MAPE of 4.02%, indicating higher relative errors. The RMSE of 96.4374 signifies larger typical prediction errors. Below is the graph of the actual gold prices and the predicted NBEATS-GRU hybrid model prices.

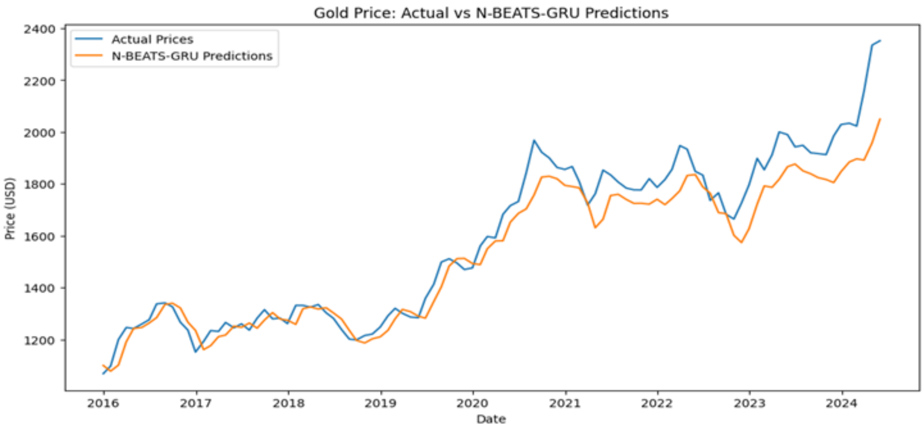

Figure 6. Graph of Actual vs NBEATS-GRU predictions.

The LSTM model had an even higher MSE, showing substantial prediction discrepancies. An MAE of $78.37, indicating greater average errors. The MAPE of 4.79% shows higher relative errors compared to other models. An RMSE of 105.1172, indicating typical prediction errors of around $105.12.

The hybrid models, contrary to expectations, did not outperform their simpler counterparts. The N-BEATS-GRU hybrid ranked third (MSE: 9300.1735, MAE: 69.7807, MAPE: 4.02%, RMSE: 96.4374), followed by LSTM (MSE: 11049.6155, MAE: 78.3674, MAPE: 4.79%, RMSE: 105.1172). Below is the graph of the actual gold prices and the predicted LSTM model prices.

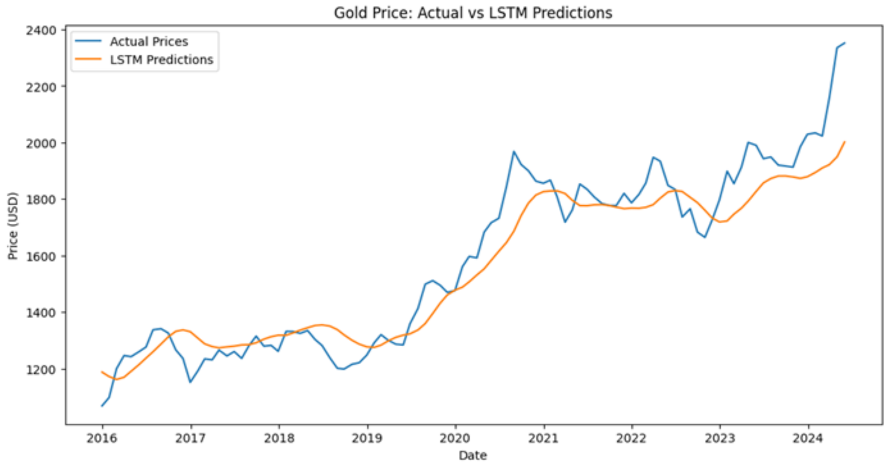

Figure 7. Graph of Actual vs LSTM predictions.

The N-BEATS-LSTM hybrid had the highest MSE, indicating the poorest performance with the largest average squared errors. The highest MAE at $138.70, reflecting significant average deviations. The highest MAPE of 8.40%, showing large relative prediction errors. The highest RMSE of 158.7741, indicating the largest typical prediction errors.

The N-BEATS-LSTM hybrid showed the poorest performance (MSE: 25209.2187, MAE: 138.6964, MAPE: 8.40%, RMSE: 158.7741). Below is the graph of the actual gold prices and the predicted NBEATS-LSTM hybrid model prices.

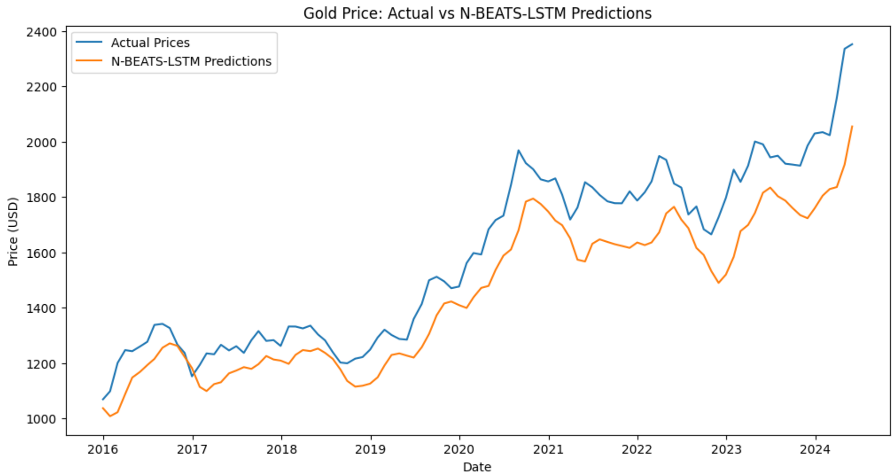

Figure 6. Graph of Actual vs NBEATS-LSTM predictions.

The high values across these metrics indicate that the model struggled to accurately predict gold prices. High MSE and RMSE values suggest frequent and large-in-magnitude errors, indicating substantial mispredictions. High MAE values support the finding that the model's predictions are significantly off from actual prices, showing a consistent pattern of error rather than occasional large outliers. High MAPE values indicate a considerable relative error, which can be problematic in financial forecasting.

Potential causes for poor performance include model complexity, overfitting, incompatibility of components, and insufficient or suboptimal hyperparameter tuning. The high error metrics suggest that simpler models like GRU or standalone N-BEATS might be more effective, challenging the assumption that combining models always leads to better performance.

Visually, the GRU model most accurately captured both short-term fluctuations and long-term trends in gold prices. The N-BEATS model also showed strong performance in this regard. The LSTM-based models tended to smooth out short-term variations more than their GRU counterparts.

Conclusion and Discussion:

This study offers valuable insights into the effectiveness of a variety of neural network architectures in the prediction of gold prices. The GRU model's exceptional performance is consistent with the results of other time series forecasting studies, including that conducted by Chung et al. [30], which illustrated the GRU's capacity to capture long-term dependencies. The N-BEATS model's exceptional performance supports the conclusions of Oreshkin et al. [9], who introduced this architecture for interpretable time series forecasting. It is intriguing that the hybrid models did not outperform their simpler counterparts, which contradicts the widely held belief that an increase in model complexity results in improved performance. In machine learning, the principle of Occam's razor, as discussed by Raschka [32], posits that simplified models frequently generalize more effectively to unseen data. This is consistent with this.

The findings of Hua et al. [31] in their study on stock market prediction are consistent with the inferior performance of LSTM in comparison to GRU. They discovered that GRU frequently outperforms LSTM as a result of its simplified structure and fewer parameters, which can result in improved generalization, particularly when training data is scarce.

The significance of model selection in financial forecasting is underscored by the substantial disparity in MAPE between the best-performing (GRU, 3.10%) and worst-performing (N-BEATS-LSTM, 8.40%) models. Sezer et al. [33] have emphasized that this distinction could have substantial implications in real-world trading scenarios in their review of deep learning models in financial market predictions.

Nevertheless, it is crucial to emphasize that these findings are unique to the gold price dataset and prediction assignment in question. The performance of various models can fluctuate in accordance with the specific characteristics of the time series under analysis, as underscored by Zhang et al. [34].

Future research could explore several avenues:

- Investigating the impact of different input features and time scales on model performance.
- Exploring ensemble methods combining the strengths of GRU and N-BEATS models.
- Applying attention mechanisms to hybrid models to potentially improve their performance.
- Conducting a more in-depth analysis of model interpretability, particularly for the GRU and N-BEATS models.
- Extending the study to other financial time series to assess the generalizability of these findings.

In summary, the GRU model exhibited the most effective performance in this gold price prediction task. However, the selection of a model for similar financial forecasting problems should be based on the specific characteristics of the data and the forecasting requirements. The unexpected underperformance of hybrid models also underscores the necessity of conducting a thorough evaluation and comparison of various architectures, rather than presuming that more complex models will necessarily produce superior results.

REFERENCE

1. Baur, D. G., & Lucey, B. M. (2010). Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. Financial Review, 45(2), 217-229.
2. Wang, Y., & Chueh, Y. L. (2013). Dynamic transmission effects between the interest rate, the US dollar, and gold and crude oil prices. Economic Modelling, 30, 792-798.
3. Shahbazi, M., & Attari, M. (2014). The effect of macroeconomic variables on the gold price in Iran. Journal of Money and Economy, 9(2), 149-176.
4. Ghosh, D., Levin, E. J., Macmillan, P., & Wright, R. E. (2004). Gold as an inflation hedge? Studies in Economics and Finance, 22(1), 1-25.
5. Kaastra, I., & Boyd, M. (1996). Designing a neural network for forecasting financial and economic time series. Neurocomputing, 10(3), 215-236.
6. Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2015). Time series analysis: forecasting and control. John Wiley & Sons.
7. Wang, J., & Wang, J. (2015). Forecasting stock market indexes using principle component analysis and stochastic time effective neural networks. Neurocomputing, 156, 68-78.
8. Nelson, D. M., Pereira, A. C. M., & de Oliveira, R. A. (2017). Stock market's price movement prediction with LSTM neural networks. International Joint Conference on Neural Networks (IJCNN), 1419-1426.
9. Oreshkin, B. N., Carpov, D., Chapados, N., & Bengio, Y. (2019). N-BEATS: Neural basis expansion analysis for interpretable time series forecasting. arXiv preprint arXiv:1905.10437.
10. Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2018). A comparison of ARIMA and LSTM in forecasting time series. 17th IEEE International Conference on Machine Learning and Applications (ICMLA), 1394-1401.
11. Lim, B., Zohren, S., & Roberts, S. (2020). Temporal fusion transformers for interpretable multi-horizon time series forecasting. arXiv preprint arXiv:1912.09363.
12. Ahmed, N. K., Atiya, A. F., Gayar, N. E., & El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. Econometric Reviews, 29(5-6), 594-621.
13. Smyl, S. (2020). A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. International Journal of Forecasting, 36(1), 75-85.
14. Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2020). The M4 competition: 100,000 time series and 61 forecasting methods. International Journal of Forecasting, 36(1), 54-74.
15. Bandara, K., Shi, P., & Bergmeir, C. (2020). Forecasting across time series databases using recurrent neural networks on groups of similar series: A clustering approach. Expert Systems with Applications,

140, 112896.
16. Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting, 14(1), 35-62.
17. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
18. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural computation, 9(8), 1735-1780.
19. World Gold Council Dataset. Retrieved on 15th June, 2024 from https://www.gold.org/goldhub/data/gold-prices#registration-type=google&just-verified=1.
20. Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2015). Time series analysis: forecasting and control. John Wiley & Sons.
21. Wang, J., & Wang, J. (2015). Forecasting stock market indexes using principle component analysis and stochastic time effective neural networks. Neurocomputing, 156, 68-78.
22. Nelson, D. M., Pereira, A. C. M., & de Oliveira, R. A. (2017). Stock market's price movement prediction with LSTM neural networks. International Joint Conference on Neural Networks (IJCNN), 1419-1426.
23. Siami-Namini, S., Tavakoli, N., & Siami Namin, A. (2018). A comparison of ARIMA and LSTM in forecasting time series. 17th IEEE International Conference on Machine Learning and Applications (ICMLA), 1394-1401.
24. Ahmed, N. K., Atiya, A. F., Gayar, N. E., & El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. Econometric Reviews, 29(5-6), 594-621.
25. the quick brown fox jumps rights over the lazy dog. the quick brown fox jumps rights over the lazy dog. Cho, K., Merrienboer, B. V., Bahdanau, D., & Bengio, Y. (2014). On the properties of neural machine translation: Encoder-decoder approaches. arXiv preprint arXiv:1409.1259.
26. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., ... & Chintala, S. (2019). PyTorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems, 32.
27. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12, 2825-2830.
28. Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. Computing in Science & Engineering, 9(3), 90-95.
29. Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555.
30. Hua, Y., Zhao, Z., Li, R., Chen, X., Liu, Z., & Zhang, H. (2019). Deep learning with long short-term memory for time series prediction. IEEE Communications Magazine, 57(6), 114-119.
31. Raschka, S. (2018). Model evaluation, model selection, and algorithm selection in machine learning. arXiv preprint arXiv:1811.12808.
32. Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep

learning: A systematic literature review: 2005–2019. Applied Soft Computing, 90, 106181.

33. Zhang, L., Aggarwal, C., & Qi, G. J. (2020). Stock price prediction via discovering multi-frequency trading patterns. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2220-2230).
34. Kroner, K. F., Kneafsey, K. P., & Claessens, S. (1995). Forecasting volatility in commodity markets. J. Forecast., 14(2), 77–95.
35. Aye, G., Gupta, R., Hammoudeh, S., & Kim, W. J. (2015). Forecasting the price of gold using dynamic model averaging. International Review of Financial Analysis, 41, 257–266. https://doi.org/10.1016/j.irfa.2015.03.010
36. Rahimi, Z. H., & Khashei, M. (2018). A least squares-based parallel hybridization of statistical and intelligent models for time series forecasting. Computers & Industrial Engineering, 118, 44–53. https://doi.org/10.1016/j.cie.2018.02.023
37. Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of econometrics, 31, 307–327. https://doi.org/10.1016/0304-4076(86)90063-1
38. Tully, E., & Lucey, B. M. (2007). A power GARCH examination of the gold market. Research in International Business and Finance, 21(2), 316–325. https://doi.org/ 10.1016/j.ribaf.2006.07.001
39. Chen, Y. H., & Chang, F. J. (2009). Evolutionary artificial neural networks for hydrological systems forecasting. Journal of Hydrology, 367(12), 125–137. https:// doi.org/10.1016/j.jhydrol.2009.01.009
40. Clements, M. P., Franses, P. H., & Swanson, N. R. (2004). Forecasting economic and financial time-series with non-linear models. International Journal of Forecasting, 20 (2), 169–183. https://doi.org/10.1016/j.ijforecast.2003.10.004
41. Alameer, Z., Elaziz, M. A., Ewees, A. A., Ye, H., & Jianhua, Z. (2019). Forecasting gold price fluctuations using improved multilayer perceptron neural network and whale optimization algorithm. Resources Policy, 61, 250–260. https://doi.org/10.1016/j.resourpol.2019.02.014
42. Anastasakis, L., & Mort, N. (2009). Exchange rate forecasting using a combined parametric and nonparametric self-organising modelling approach. Expert Systems with Applications, 36(10), 12001–12011. https://doi.org/10.1016/j.eswa.2009.03.057
43. Kristjanpoller, W., & Minutolo, M. C. (2015). Gold price volatility: A forecasting approach using the Artificial Neural Network-GARCH model. Expert Systems with Applications, 42(20), 7245–7251. https://doi.org/10.1016/j.eswa.2015.04.058
44. Chen, Y., & Tan, H. (2017). Short-term prediction of electric demand in building sector via hybrid support vector regression. Applied Energy, 204, 1363–1374. https://doi. org/10.1016/j.apenergy.2017.03.070
45. Hao, W., & Yu, S. (2006). Support Vector Regression for Financial Time Series Forecasting. IP International Federation for Information Processing. https://doi.org/10.1007/0-387-34403-9_115
46. Levis, A., & Papageorgiou, L. (2005). Customer demand forecasting via support vector regression analysis. Chemical Engineering Research and Design, 83(8), 1009–1018. https://doi.org/10.1205/cherd.04246

47. Chen, S., & He, H. (2018). Stock Prediction Using Convolutional Neural Network. IOP Conference Series Materials Science and Engineering. https://doi.org/10.1088/1757-899X/435/1/012026
48. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85–117. https://doi.org/10.1016/j.neunet.2014.09.003
49. Tealab, A. (2018). Time series forecasting using artificial neural networks methodologies: A systematic review. Future Computing and Informatics Journal, 3(2), 334–340. https://doi.org/10.1016/j.fcij.2018.10.003
50. Zhou, S., Lai, K. K., & Yen, J. (2012). A dynamic meta-learning rate-based model for gold market forecasting. Expert Systems with Applications, 39(6), 6168–6173. https://doi.org/10.1016/j.eswa.2011.11.115
51. Di Persio, L., & Honchar, O. (2018). Multitask machine learning for financial forecasting. International Journal of Circuits, Systems and Signal Processing, 12, 444–451.
52. Pradeepkumar, D., & Ravi, V. (2017). Forecasting financial time series volatility using Particle Swarm Optimization trained Quantile Regression Neural Network. Applied Soft Computing, 58, 35–52. https://doi.org/10.1016/j.asoc.2017.04.014
53. Kristjanpoller, W., & Minutolo, M. C. (2015). Gold price volatility: A forecasting approach using the Artificial Neural Network-GARCH model. Expert Systems with Applications, 42(20), 7245–7251. https://doi.org/10.1016/j.eswa.2015.04.058
54. Kristjanpoller, R. W., & Hernandez, P. E. (2017). Volatility of main metals forecasted by a hybrid ANN-GARCH model with regressors. Expert Systems with Applications, 84, 290–300. https://doi.org/10.1016/j.eswa.2017.05.024
55. Sezer, O. B., & Ozbayoglu, A. (2018). Algorithmic financial trading with deep convolutional neural networks: Time series to image conversion approach. Applied Soft Computing, 70, 525–538. https://doi.org/10.1016/j.asoc.2018.04.024 Berat Sezer
56. Cao, J., & Wang, J. (2019). Stock price forecasting model based on modified convolution neural network and financial time series analysis. International Journal of Communication Systems, 32(12). https://doi.org/10.1002/dac.3987
57. Di Persio, L., & Honchar, O. (2017). Analysis of recurrent neural networks for short-term energy load forecasting. AIP Conference Proceedings, 1906, Article 190006.
58. Kamal, I., Bae, H., Sunghyun, S., & Yun, D. H. (2020). DERN: Deep Ensemble Learning Model for Short- and Long-Term Prediction of Baltic Dry Index. Appl. Sci, 1504. https://doi.org/10.3390/app10041504
59. Zhang, Y., Xiong, R., He, H., & Pecht, M. (2018). ong short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries. IEEE Transactions on Vehicular Technology, 5695–5705. https://doi.org/10.1109/TVT.2018.2805189
60. Choudhury, S., Ghosh, S., Bhattacharya, A., Fernandes, K. J., & Tiwari, M. K. (2014). A real time clustering and svm based price-volatility prediction for optimal trading strategy. Neurocomputing, 131, 419–426. https://doi.org/10.1016/j.neucom.2013.10.002

61. Pierdzioch, C., & Risse, M. (2020). Forecasting precious metal returns with multivariate random forests. Empir. Econ., 58, 1167–1184. https://doi.org/10.1007/s00181-018-1558-9
62. Kaushik, M., & Giri, A. (2020). Forecasting foreign exchange rate: A multivariate comparative analysis between traditional econometric, contemporary machine learning & deep learning techniques. Quantitative finance, statistical finance. Cornell University.
63. Nelson, D., Pereira, A., & de Oliveira, R. (2017). Stock market's price movement prediction with LSTM neural networks. 2017 International joint conference on neural networks (IJCNN), IEEE, 1419-1426. https://doi.org/10.1109/IJCNN.2017.7966019.
64. Nikou, M., Mansourfar, G., & Bagherzadeh, J. (2019). Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms. Intelligent Systems in Accounting, Finance and Management, 164–174. https://doi.org/10.1002/isaf.1459
65. Pang, X., Zhou, Y., & Wang, P. (2018). An innovative neural network approach for stock market prediction. https://doi.org/10.1007/s11227-017-2228-y
66. Hoseinzade, E., & Haratizadeh, S. (2019). CNNpred: CNN-based stock market prediction using a diverse set of variables. Expert Systems with Applications, 129, 273–285. https://doi.org/10.1016/j.eswa.2019.03.029