p-ISSN: 2708-2474
e-ISSN: 2708-2482

www.gmsrjournal.com
GMSR
Global Management Sciences Review
managing humanity

GMSR

GLOBAL MANAGEMENT SCIENCES REVIEW

HEC-RECOGNIZED CATEGORY- Y

VOL. X, ISSUE I, WINTER (MARCH-2025)

DOI (Journal): 10.31703/gmsr
DOI (Volume): 10.31703/gmsr.2025(X)
DOI (Issue): 10.31703/gmsr.2025(X-I)

Double-blind Peer-review Research Journal
www.gsrjournal.com

HumanityPublications
sharing research

Humanity Publications (HumaPub)
www.humapub.com
Doi:https://dx.doi.org/10.31703

Article title

Artificial Intelligence-based Deep Learning Model Optimizing Financial Predictions: Empirical Evidence from Top six Markets

Abstract

Navigating the complexities of the foreign exchange market remains a significant challenge for investors and risk managers worldwide. This study addresses this challenge by developing a Deep Learning-based Multilayer Perceptron (DL MLP) model to classify daily multi-currency exchange rate returns (USD, EUR, GBP, CHF, JPY, CNY) from 2010 to 2024. The framework categorizes market movements into three distinct states: appreciation (good returns), depreciation (bad returns), and stability (no returns), transforming these categorical outcomes into a numerical format suitable for neural network processing via one-hot encoding. Powered by a backpropagation algorithm, the model demonstrates exceptional predictive capability, achieving a 91% classification accuracy and near-perfect AUC ROC scores, thereby identifying optimal safe-haven currencies with high reliability. The empirical findings offer a powerful data-driven tool for forex investors seeking to refine their strategic asset allocation and for global risk managers aiming to enhance their hedging strategies against currency volatility.

Keywords: Multicurrency Exchange Rate Returns, Deep Learning, Optimal Prediction Performance.

Authors:
Rabia Akram: (Corresponding Author)
College of Management, Shenzhen University,
Guangdong, 518060, PR China.
Email: (rabiaakram177@gmail.com)

Pages: 62-83
DOI:10.31703/gssr.2025(X-I).06
DOI link: https://dx.doi.org/10.31703/gmsr.2025(X-I).06
Article link: http://www.gmsrjournal.com/article/artificial-intelligencebased-deep-learning-model-optimizing-financial-predictions-empirical-evidence-from-top-six-markets
Full-text Link: https://gmsrjournal.com/article/artificial-intelligencebased-deep-learning-model-optimizing-financial-predictions-empirical-evidence-from-top-six-markets
Pdf link: https://www.gmsrjournal.com/jadmin/Auther/31rvlolA2.pdf

Global Management Science Review

p-ISSN: 2708-2474 **e-ISSN**: 2708-2482
DOI(journal): 10.31703/gmsr
Volume: X (2025)
DOI (volume): 10.31703/gmsr 2025(X)
Issue: I (Winter-March 2025)
DOI(Issue): 10.31703/gmsr.2025(X-I)

Home Page
www.gmsrjournal.com

Volume: X (2025)
https://www.gmsrjournal.com/Current-issues

Issue: I-March (March-2025)
https://www.gmsrjournal.com/issue/9/1/2025

Scope
https://www.gmsrjournal.com/about-us/scope

Submission
https://humaglobe.com/index.php/gmsr/submissions

Google Scholar

Visit Us

Humanity Publications (HumaPub)
www.humapub.com
Doi:https://dx.doi.org/10.31703

Citing Article

06	**Artificial Intelligence-based Deep Learning Model Optimizing Financial Predictions: Empirical Evidence from Top six Markets**		
Authors	Rabia Akram	**DOI**	10.31703/gmsr.2025(X-I).06
		Pages	62-83
		Year	2025
		Volume	X
		Issue	III

Referencing & Citing Styles

APA	Akram, R. (2025). Artificial Intelligence-based Deep Learning Model Optimizing Financial Predictions: Empirical Evidence from Top six Markets. *Global Management Sciences Review*, *X*(I), 128-133. https://doi.org/10.31703/gmsr.2025(X-I).16
CHICAGO	Akram, Rabia. 2025. "Artificial Intelligence-based Deep Learning Model Optimizing Financial Predictions: Empirical Evidence from Top six Markets." *Global Management Sciences Review* X (I):128-133. doi: 10.31703/gmsr.2025(X-I).16.
HARVARD	AKRAM, R. 2025. Artificial Intelligence-based Deep Learning Model Optimizing Financial Predictions: Empirical Evidence from Top six Markets. *Global Management Sciences Review,* X, 128-133.
MHRA	Akram, Rabia. 2025. 'Artificial Intelligence-based Deep Learning Model Optimizing Financial Predictions: Empirical Evidence from Top six Markets', *Global Management Sciences Review*, X: 128-33.
MLA	Akram, Rabia. "Artificial Intelligence-Based Deep Learning Model Optimizing Financial Predictions: Empirical Evidence from Top Six Markets." *Global Management Sciences Review* X.I (2025): 128-33. Print.
OXFORD	Akram, Rabia (2025), 'Artificial Intelligence-based Deep Learning Model Optimizing Financial Predictions: Empirical Evidence from Top six Markets', *Global Management Sciences Review,* X (I), 128-33.
TURABIAN	Akram, Rabia. "Artificial Intelligence-Based Deep Learning Model Optimizing Financial Predictions: Empirical Evidence from Top Six Markets." *Global Management Sciences Review* X, no. I (2025): 128-33. https://dx.doi.org/10.31703/gmsr.2025(X-I).16.

e-ISSN: 2708-2482 **Volume:** X (2025) **Issue:** I-Winter (March-2025) **p-ISSN:** 2708-2474

Global Management Sciences Review
www.gmsrjournal.com
DOI: http://dx.doi.org/10.31703/gmsr

Pages: 62-83 **URL:** https://doi.org/10.31703/gmsr.2025(X-I).06 **Doi:** 10.31703/gmsr.2025(X-I).06

Title

Artificial Intelligence-based Deep Learning Model Optimizing Financial Predictions: Empirical Evidence from Top six Markets

Abstract

Navigating the complexities of the foreign exchange market remains a significant challenge for investors and risk managers worldwide. This study addresses this challenge by developing a Deep Learning-based Multilayer Perceptron (DL-MLP) model to classify daily multi-currency exchange rate returns (USD, EUR, GBP, CHF, JPY, CNY) from 2010 to 2024. The framework categorizes market movements into three distinct states: appreciation (good returns), depreciation (bad returns), and stability (no returns), transforming these categorical outcomes into a numerical format suitable for neural network processing via one-hot encoding. Powered by a backpropagation algorithm, the model demonstrates exceptional predictive capability, achieving a 91% classification accuracy and near-perfect AUC ROC scores, thereby identifying optimal safe-haven currencies with high reliability. The empirical findings offer a powerful data-driven tool for forex investors seeking to refine their strategic asset allocation and for global risk managers aiming to enhance their hedging strategies against currency volatility.

Keywords: Multicurrency Exchange Rate Returns, Deep Learning, Optimal Prediction Performance

Authors:
Rabia Akram: (Corresponding Author)
College of Management, Shenzhen University,
Guangdong, 518060, PR China.
Email: (rabiaakram177@gmail.com)

Contents

Introduction

In contemporary financial research, advanced computational techniques are increasingly utilized to forecast market behavior. Traditionally, evaluations in financial systems have relied extensively on soft computing, alongside parametric and non-parametric statistical methods (Duan, 2019; Dunis et al., 2016). Parametric models, which assume a predetermined functional structure with fixed parameters, are often employed to estimate outcomes in probability-based prediction scenarios (Tjøstheim et al., 2022). Logistic Regression, Naïve Bayes, Linear Discriminant Analysis, and basic Neural Network architectures represent well-established parametric approaches that have been widely applied in financial modeling for several decades. Specifically, Logistic Regression is generally effective for linear relationships, Naïve Bayes performs well in classification settings, Linear Discriminant Analysis is valuable for dimensionality reduction in classification tasks, and simple Neural

Networks are suitable when relatively large datasets are available.

Conversely, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Decision Trees (DT) are prominent examples of non-parametric methods that do not rely on predefined parameter assumptions. Although some scholars argue that parametric techniques may offer more rigorous interpretation compared to non-parametric approaches (Asmare & Begashaw, 2018; Galeati, 1990; Sedgwick, 2015), both methodological groups are substantially challenged by the presence of uncertainty and noise, particularly within emerging and highly volatile financial markets.

In contrast, soft computing approaches are designed to handle uncertainty and partial information more effectively, mirroring aspects of human cognitive processes (Pradeepkumar & Ravi, 2018). These methods have demonstrated significant success in various financial applications, such as risk assessment, budgeting, e-commerce strategy, portfolio optimization, Black–Scholes option pricing, corporate acquisition decisions, and evaluating investments in technology-intensive production (Zhao et al., 2011). Atsalakis and Valavanis (2009) emphasize that robust stock market forecasting systems can be developed even when input data are limited, as soft computing models are capable of learning complex nonlinear relationships without requiring strong assumptions regarding data distribution. Due to market noise, rapid environmental shifts, and behavioral biases influencing investor decision-making, soft computing techniques have gained growing acceptance as flexible and intelligent analytical tools (Naveed et al., 2025).

Soft computing encompasses a broad range of machine learning and neural network architectures, including genetic algorithms, fuzzy-based neural systems, and Artificial Neural Networks (ANNs), which are summarized in Appendix A. Within ANN research, model architectures are commonly divided into Shallow Neural Networks (SNNs) and Deep Neural Networks (DNNs), depending on the number of hidden layers. SNNs including Simple Linear Regression-based neural networks, Convolutional Neural Networks with a single hidden layer (CNN-SHL), Radial Basis Function (RBF) networks, and Single Layer Perceptron (SLP) architectures trace their conceptual origin to the early theoretical model introduced by (McCulloch & Pitts, 1943). Although early SNN models had limitations, particularly the inability of the SLP to resolve nonlinear XOR relationships (Minsky & Papert, 1969), subsequent studies demonstrated that SNNs are capable of approximating both linear and nonlinear functions (Broomhead & Lowe, 1988; Hornik et al., 1989; Memon et al., 2022), leading to renewed interest in their use (Sastri et al., 2024).

More recent advancements have favored DNNs due to their superior predictive performance, particularly when modeling high-dimensional and nonlinear time series structures. Representative DNN architectures include Recurrent Neural Networks (RNN), Autoencoder-based DNNs (AE-DNN), Recursive Neural Networks (RvNN), Multilayer Perceptron (MLP) models, Deep Belief Networks (DBN), and Convolutional Neural Networks with multiple hidden layers (CNN-MHL). While the Long Short-Term Memory (LSTM) network was introduced to address long-range dependency issues in sequential data, some studies suggest that conventional LSTMs struggle when applied to financial time series characterized by high dimensionality (Jung & Choi, 2021; Sagheer & Kotb, 2019). To address this limitation, hybrid architecture such as LSTM-Autoencoder networks have been proposed. Additionally, Stacked Autoencoders (SAE), based on the Restricted Boltzmann Machine (RBM), have been widely utilized for stock return forecasting (Hinton, 2012; Takeuchi & Lee, 2013), and play an increasingly important role in portfolio risk assessment Patterson and Gibson (2017).

The complex structural dynamics of financial markets, influenced by numerous interacting variables, require highly adaptive modeling strategies. Conventional statistical and traditional machine learning methods are generally insufficient for capturing the intrinsic density and directional patterns of market movements (HongXing et al., 2022; HongXing et al., 2023; Naveed et al., 2023; Naveed et al., 2024). Motivated by this challenge, the present study develops a DNN-based Multilayer Perceptron (MLP) model with backpropagation learning to evaluate return behavior and to identify a potential safe-haven currency suitable for strategic investment.

The remainder of this study is structured as follows: Section 2 reviews prior applications of soft computing and statistical methods in currency return prediction and highlights the performance characteristics of the proposed model. Section 3 details the dataset, preprocessing procedures, parameter specifications, mathematical formulation, and training process. Section 4 presents and interprets the empirical results, including the identification of a safe-haven currency. Section 5 addresses the discussion of the research and compares findings with previous studies. Section 6 summarizes the key contributions, significant implications, outlines limitations, and proposes directions for future research.

Literature review

A large body of research has explored currency markets through machine learning, statistical modelling, and econometric techniques. In particular, the prediction and evaluation of currency returns have attracted substantial academic interest, with scholars applying genetic algorithms, artificial neural networks (ANNs), and advanced statistical frameworks to identify currencies with stable performance or safe-haven characteristics. Accordingly, this section presents relevant research contributions that have applied such methods to model exchange rate behavior and currency return dynamics.

Various empirical studies have utilized different modelling techniques to examine currency return patterns. Kernel-based Smoothed Regression has been implemented to capture underlying return structures (Baillie & Kim, 2015), while Quantile Regression has been employed to analyze heterogeneous return distributions under different market conditions (Baruník & Čech, 2021; Nguyen et al., 2020; Rababa et al., 2021). Machine learning-supported predictive models, such as Support Vector Machine (SVM) and Support Vector Regression (SVR), have demonstrated strong forecasting capacity for currency movements (Aggarwal et al., 2020; Peng et al., 2018; Wang, 2011). Volatility-based econometric models, including GARCH and its family extensions, continue to be widely used to measure dynamic variance in exchange returns (Fakhfekh & Jeribi, 2020; Karmakar, 2017; Kumar & Anandarao, 2019; Udoh & Udeaja, 2019).

Other classification and learning-based algorithms have also been explored. These include Decision Tree (Malliaris & Malliaris, 2015), Naïve Bayes (Grobys et al., 2022), and Fuzzy Logic approaches (Gharleghi et al., 2014; Lee & Wong, 2007; Sadeghi et al., 2021). Additionally, ANNs and deep learning variants have shown effectiveness in capturing complex nonlinear currency market (Islam & Hossain, 2020; Koo & Kim, 2021; Nakano & Takahashi, 2020; Nasirtafreshi, 2022; Panda et al., 2021). Conventional econometric techniques have also contributed important empirical insights (Demirer et al., 2022; Pu & Zhang, 2012; Wali et al., 2017).

A wide spectrum of empirical studies has adopted diverse analytical and computational frameworks to model currency returns. For instance, Nohara et al. (2022) compared several classification and prediction models—including K-Nearest Neighbors (KNN), Multilayer Perceptron (MLP), Random Forest (RF), Support Vector Machines (SVM), and Logistic Regression—to forecast movements in the USD/BRL pair, concluding that certain machine learning configurations generated consistently positive returns while maintaining relatively low levels of risk. Similarly, Castán-Lascorz et al. (2022) employed both univariate and multivariate regression techniques to examine the behavior of major international currencies such as the AUD, CAD, CHF, EUR, GBP, JPY, and USD, identifying meaningful interactions between return dynamics and volatility premiums. In another study, Hasanov et al. (2024) applied the GARCH model to seventeen different currencies and observed that return fluctuations generally followed a negative trend, whereas Iyke et al. (2022) using the ARCH model for sixteen currencies, reported a contrasting finding that approximately three-quarters of the return forecasts were positive.

Advanced hybrid and decomposition-based learning frameworks have also been explored. Yang and Lin (2017) introduced two exchange rate forecasting models—AEMD-KELM and EMD-ELM—using seven major currency pairs; their results indicated that the AEMD-KELM model offered noticeably stronger predictive accuracy than its EMD-ELM counterpart. In a related direction, Baffour et al. (2019) proposed a hybrid structure combining GJR components with ANN architectures for the USD and five benchmark currencies, demonstrating that the hybrid model outperformed competing forecasting methods. Bizhani and Kuru (2022) and Akhtar et al. (2025) further compared deep neural networks with Bayesian regularization against conventional models such as ARIMA, Random Walk, GARCH, and MLP-BP for the USD/GBP exchange rate, showing that the DNN-BN framework achieved superior forecasting efficiency.

Other researchers have focused on single and hybrid neural network topologies. Deng et al. (2025) conducted performance evaluations across multiple Pi-Sigma neural network algorithms alongside RBF, MLP, and linear regression models applied to CAD, JPY, and CHF exchange rates against USD, ultimately finding that the Pi-Sigma variants yielded more favorable return outcomes. Memon et al. (2025) benchmarked several computational intelligence models including RCGP-ANN, Markov, Naïve classification, MLP, Fuzzy Evolutionary Regression, and DNN-BN using returns of Yen, NZD, IDR, KRW, and CAD relative to AUD, concluding that the RCGP-ANN approach provided the most accurate predictive performance. Kumar et al. (2022) compared a hybrid ANN-GA model with a conventional ANN for the INR/USD exchange rate, reporting that the integration of a genetic algorithm significantly improved predictive reliability. Lastly, Li et al. (2024) implemented an extensive model comparison incorporating Random Walk, ARMA-GARCH, traditional neural networks, recurrent neural networks, LSTM, and a hybrid ICA-ANN system for USD exchange rates against EUR, JPY, CHF, CAD, and GBP. Their findings emphasized that the ICA-ANN hybrid produced more accurate and stable forecasts than both standard econometric and conventional neural network approaches.

Recent research demonstrates increasing reliance on deep learning-based architectures such as LSTM, AE, hybrid LSTM-AE models, RNN, MLR, MLP, and composite neural network structures to evaluate currency return behavior. Some works suggest superior forecasting performance for signal neural network-based MLP systems when applied to USD exchange rates (Maté & Jimeńez, 2021). Likewise, interval time-series neural networks have shown improved performance compared to random walk models (Roque et al., 2007), while hybrid MLP systems have demonstrated the ability to forecast upper and lower bounds of exchange rates more effectively than AR and IDE models (Tan et al., 2025).

Based on the reviewed literature, it is evident that while numerous studies have employed machine learning, statistical, and econometric models particularly various neural network architectures such as MLP, DNN, LSTM, AE, and hybrid models for forecasting currency returns and identifying safe-haven currencies, several gaps remain. Most of the existing research primarily focuses on evaluating upper and lower bounds of exchange rates or relies on conventional data structures without fully exploring the potential of novel data representations (Memon et al., 2024). Additionally, the performance of deep intelligent networks using advanced encoding techniques, such as hot-encoding functions, remains largely unexplored. Consequently, there is a need for studies that apply robust DNN-based MLP models on innovative data structures to more accurately assess currency return performance and identify safe-haven currencies, addressing both predictive accuracy and practical applicability in the foreign exchange market (Naveed et al., 2020).

However, while extensive attention has been given to MLP and hybrid network architectures, most studies emphasize exchange rate bound prediction rather than direct evaluation of currency return performance under alternative data structures. The present study, therefore, contributes to this literature by employing a DNN-based MLP architecture with a modified data representation strategy using one-hot encoding, aiming to enhance return learning efficiency and improve safe-haven currency (Barunik et al., 2016; Das et al., 2018; Dautel et al., 2020; Dymova et al., 2016; Parot et al., 2019).

Research data and empirical methodology

Data preprocessing

This research analyzes the predictive performance of the proposed model using multi-currency exchange rate returns (MCERRs) for six major currencies: USD, EUR, GBP, CHF, JPY, and CNY. The empirical dataset is sourced from the International Monetary Fund's official database, accessible through their public data portal. Exchange rate returns are calculated using logarithmic differentiation to capture continuous compounding effects, following established financial econometric methodology. The computational formula is expressed as: $R_{i,t+1} = ln(R)_{i,t+1} - ln(R)_{i,t} / ln(R)_{i,t}$ as follows (Akhtar et al., 2024). The analyzed currency pairs are systematically categorized as follows: USD/EUR (CERR1), EUR/USD (CERR2), GBP/USD (CERR3), CHF/USD (CERR4), JPY/USD (CERR5), and CNY/USD (CERR6). To accommodate the numerical processing requirements of deep neural networks, categorical return classifications are transformed into numerical representations using one-hot encoding. The classification framework comprises three distinct categories: good returns (Class 0), bad returns (Class 1), and no returns (Class 2). The dataset structure assumes Q distinct MCERR series, with the complete observation set represented as N total samples. The sample population is distributed across the three classification categories, such that the aggregate observation count equals the sum of instances across all classes: $N = GR + BR + NR$, where GR, BR, and NR denote the sample counts for favorable, unfavorable, and neutral returns respectively.

Table 1

Using one-hot-encoding to label MCERRs

Number of returns	1	2	3	4	5	Q
MCERRs	CERR1	CERR2	CERR3	CERR4	CERR5	CERR6
One-hot-encoding label	[0 1… 2]$_{1\text{x}}$Q	[1 2… 0]$_{1\text{x}}$Q	[2 0… 1]$_{1\text{x}}$Q	[0 1… 2]$_{1\text{x}}$Q	[1 2… 0]$_{1\text{x}}$Q	[2 0… 1]$_{1\text{x}}$Q

Table 1 portrays the methodological framework for converting categorical currency return data into a numerical format suitable for neural network processing. The transformation employs a one-hot encoding technique to represent six distinct MCERRs series. The encoding process assigns a unique binary vector to each currency return category (CERR1 through CERR6). Each vector has a dimensionality (length) of Q, which corresponds to the total number of unique MCERR classes. The notation exemplifies the structure of these encoded vectors. Crucially, the values shown (0, 1, 2) indicate that this is a multi-class label encoding rather than a strict one-hot encoding. In this scheme, each MCERR is assigned a single integer label (0, 1, or 2) that represents its predefined category, such as good, bad or no return. The 1 ×Q dimension signifies that the output for a single data sample is a one-dimensional vector (a row) of length Q, where each position holds the class label for a specific currency pair. In essence, this table documents the final step of feature preparation, where raw currency identifiers are mapped to a structured numerical representation. This allows the deep neural network to process the complex relationships between the different currency return series effectively, treating each as a distinct, quantifiable input feature.

Training process of the proposed model

The proposed deep learning framework operates on a supervised learning paradigm, processing input data samples $X(n)$ against corresponding target outputs $d(n)$, where $n = 1,2,\dots,N$ represents individual data instances represents. The network architecture comprises L computational layers, excluding the initial input layer which is designated as layer $l = 0$, while the final output layer is identified as layer $l = L$. Each layer l contains m_l processing neurons that transform incoming signals through weighted connections and activation functions. The input vector for the n-th sample is defined as $x(n) = [x_1(n), x_2(n), \dots, x_{m_0}(n)]^T$, where m_0 indicates the dimensionality of the input feature space. The network's knowledge is encoded in connection weights $w_{l,j,i}^{[k]}$, representing the strength of the connection from neuron i in layer $l-1$ to neuron j in layer l during the k-th training iteration. These weights are organized in a matrix structure for each layer:

$$W_l^{[k]} = \begin{bmatrix} w_{l,1,0}^{[k]} & w_{l,1,1}^{[k]} & \cdots & w_{l,1,m_{l-1}}^{[k]} \\ w_{l,2,0}^{[k]} & w_{l,2,1}^{[k]} & \cdots & w_{l,2,m_{l-1}}^{[k]} \\ \vdots & \vdots & \ddots & \vdots \\ w_{l,m_l,0}^{[k]} & w_{l,m_l,1}^{[k]} & \cdots & w_{l,m_l,m_{l-1}}^{[k]} \end{bmatrix}_{m_l \times (m_{l-1}+1)} \tag{1}$$

The transformation process through each layer involves two computational stages. First, the weighted summation combines inputs from the previous layer:

$$v_{l,j}(n) = \sum_{i=0}^{m_{l-1}} w_{l,j,i} \cdot y_{l-1,i}(n) \tag{2}$$

where $y_{l-1,i}(n)$ denotes outputs from the preceding layer, with $y_{l-1,0}(n)$ representing the bias term. Second, an activation function $\varphi_l(\cdot)$ introduces non-linearity to the system:

$$y_{l,j}(n) = \varphi_l(v_{l,j}(n)) \tag{3}$$

The hidden layers employ sigmoid activation functions $\sigma(\mathcal{Z}) = \frac{1}{1+e^{-\mathcal{Z}}}$ for $l = 1,2,\ldots,L-1$, while the output layer utilizes softmax normalization to produce probability distributions across multiple classes:

$$y_{L,j}(n) = \frac{e^{v_{L,j}(n)}}{\sum_{k=1}^{m_L} e^{v_{L,k}(n)}} \tag{4}$$

The complete vector of neuronal activities within each layer can be represented as:

$$v_l(n) = [v_{l,1}(n), v_{l,2}(n), \ldots, v_{l,m_l}(n)]^T \tag{5}$$

This mathematical formulation establishes the foundation for the forward propagation mechanism that enables the network to learn complex hierarchical representations from the input financial data.

Figure 1

Architectural overview of the proposed model

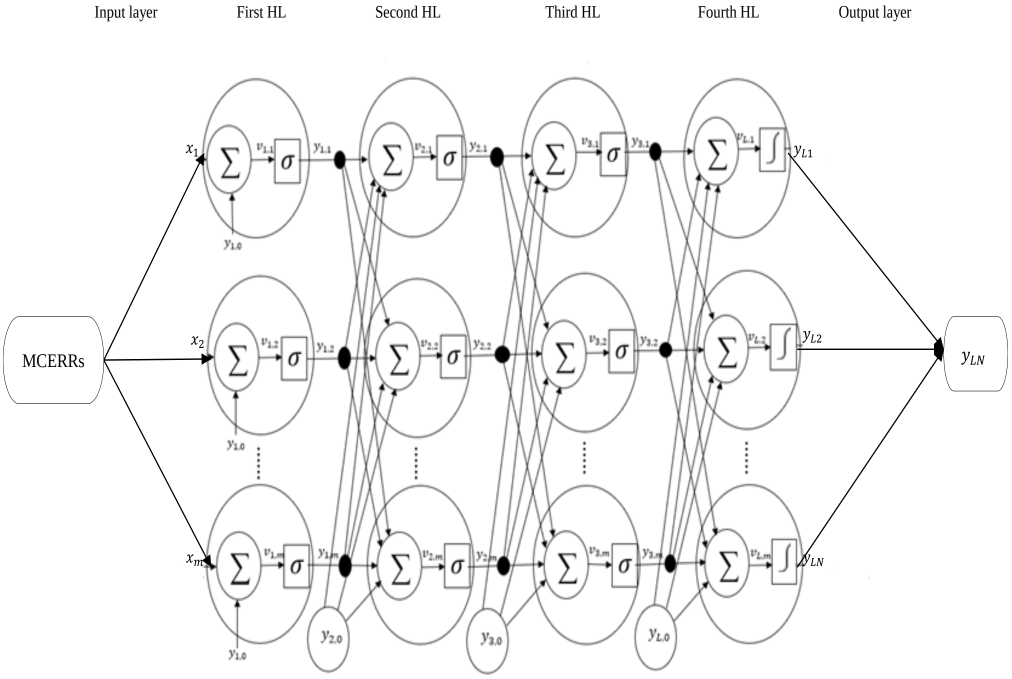

Figure 1 illustrates the structural configuration of the proposed deep neural network, which implements a multilayer perceptron (MLP) framework with a total of four processing layers ($L = 4$). The architecture consists of one input layer, three successive hidden layers, and a final output layer, forming a hierarchical feature extraction pipeline. The network employs a strategically designed dimensionality reduction approach across its hidden layers to balance model capacity with computational efficiency. The first hidden layer incorporates 256 neurons to capture broad feature representations from the input data. This is followed by subsequent layers containing 128, 64, and 32 neurons respectively, progressively refining the feature abstractions while reducing parametric complexity. The architectural design allows for flexibility in network configuration, as both the quantity of artificial neurons and the number of hidden layers can be modified according to specific application requirements. From an implementation perspective, the forward propagation mechanism utilizes vectorized operations where an additional bias term $v_{l,0}(n)$ is appended to the activation vector $v_l(n)$. The dimensional relationships are maintained such that $v_l(n)$ constitutes an $m_l \times 1$ vector, while the corresponding weight matrix $W_l^{[k]}$ maintains dimensions of $m_l \times (m_{l-1} + 1)$ to ensure proper mathematical transformations between layers. This systematic reduction in layer sizes creates an information funnel that distills relevant patterns from input data while minimizing computational overhead, making the model both effective and efficient for currency return classification tasks.

To illustrate the neural network's operational mechanism, consider a training instance $(X(1), y_L(1))$, where $X(1)$ represents an $m_0 \times 1$ input feature vector and $y_L(1)$ denotes the corresponding 3×1 target output vector. A bias term $x_0(1)$ is incorporated into the input layer to enhance model flexibility. The computational flow through the network layers follows a structured transformation process. The activation potential for neurons in hidden layers is calculated through linear combinations of weighted inputs. For a single data sample, excluding the iteration index k from the weight notation, the pre-activation value for the bias unit in layer l is determined by:

$$v_{l,0}(n) = \sum_{i=0}^{m_l} w_1 x(1) + x_0(1) \tag{6}$$

The output of the initial hidden layer is subsequently generated through activation function transformation:

$$y_l(1) = \sigma(W_l x_l(1) + x_0(1)) \tag{7}$$

The information propagation between successive layers maintains a consistent pattern, where outputs from layer l serve as inputs to layer $l + 1$:

$$y_{l+1}(1) = \sigma(W_l y_l(1) + y_{l+1,0}(1)) \tag{8}$$

The model's predictive capability is refined through systematic adjustment of bias units $y_{l,0}(1)$ across all layers $l = 1,2,\dots,L$ during the training cycle. The final output layer synthesizes information from preceding layers through comprehensive integration:

$$y_L(1) = \varphi\left(\sum_{j=1}^{m_l} W_L y_L(1) + y_{L,0}(1)\right) \tag{9}$$

Here, φ represents the activation function that aggregates and transforms inputs from previous neuronal connections, while m indicates the quantity of output features in the terminal layer.

The classification mechanism employs a threshold-based decision rule at the output layer:

$$y_L(n) = [y_{L,1} y_{L,2} y_{L,N}]^T \Rightarrow \begin{cases} \text{"Bad Returns"} & y_{L,1} > T_0 \\ \text{"Good Returns"} & y_{L,2} > T_0 \\ \text{"No Returns"} & y_{L,3} > T_0 \end{cases} \tag{10}$$

where T_0 signifies a predefined classification threshold. When the discrepancy between desired and predicted outputs exceeds acceptable limits, the backpropagation algorithm is activated to systematically modify connection weights and bias terms, thereby enhancing the model's predictive accuracy through iterative optimization. The comprehensive training procedure, including all computational stages and optimization mechanisms, is visually detailed in Figure 2, which outlines the complete workflow from initialization to final model deployment.

Figure 2

Flowchart of the absolute training process of the proposed model

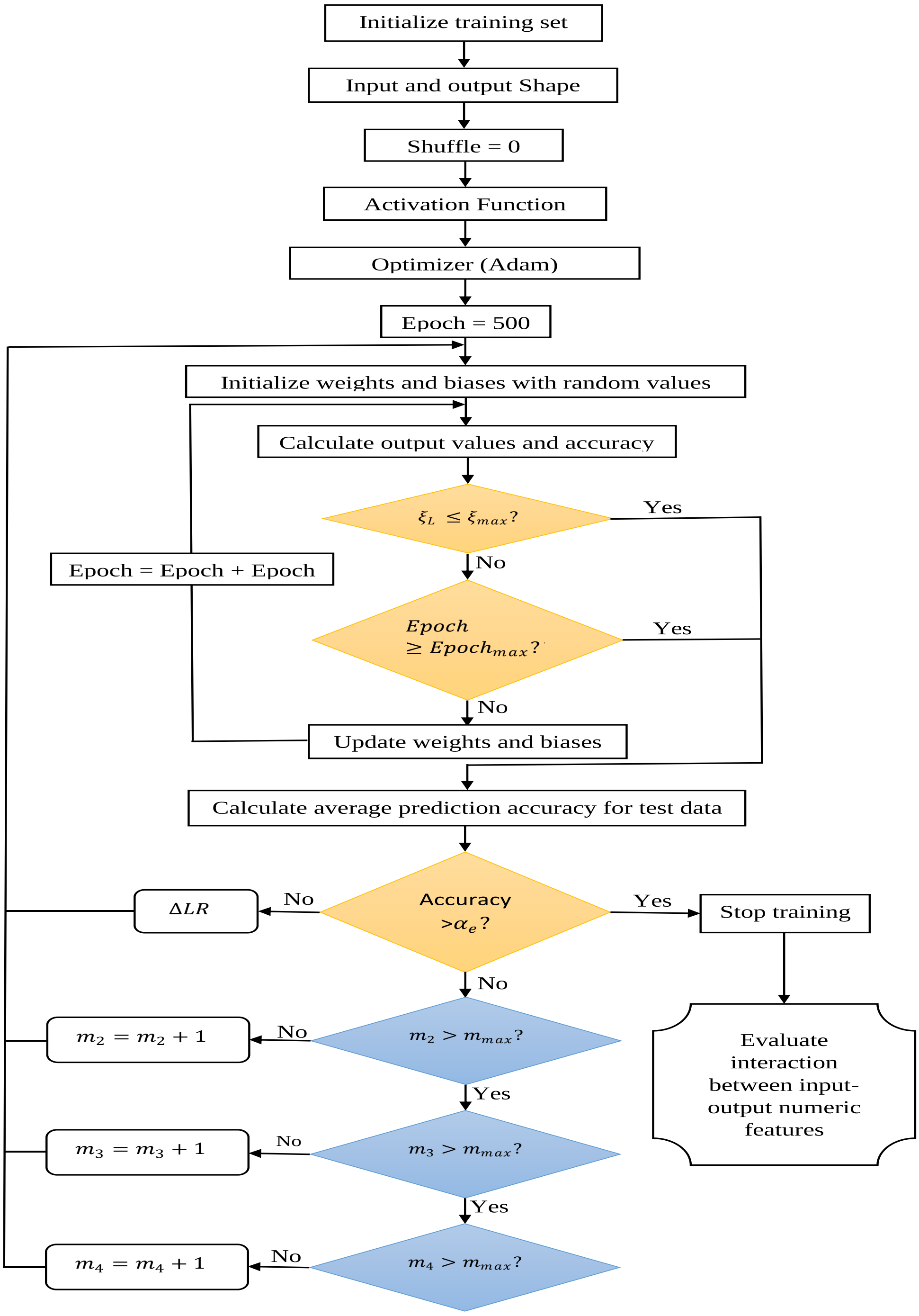

Figure 2 presents the comprehensive training methodology for the proposed neural network architecture. The procedural framework begins with data preparation, where refined datasets are organized into structured input-output pairs suitable for supervised learning. The processed data undergoes partitioning into distinct training and testing subsets to enable proper model validation.

A fundamental architectural guideline governs the network design: the cumulative neuron count across all hidden layers must exceed half the dimensionality of the input feature space, mathematically expressed as $\sum_{l=1}^{L-1} m_l > m_0/2$. This ensures sufficient model capacity to capture underlying data patterns. The complete specification of training parameters and architectural details is documented in Table 2 and Table 3 which presented in Appendix (A). The optimization objective focuses on maximizing predictive performance while systematically reducing the cost function through iterative refinement. The training algorithm employs a conditional logic flow based on continuous performance monitoring: When the observed error metric falls below the target threshold, the system initiates backpropagation with increased epoch count to further enhance model accuracy. Conversely, if performance targets are met, the process transitions to the testing phase. A critical evaluation mechanism continuously assesses prediction accuracy against established benchmarks. The training termination protocol activates under either of two conditions: when the achieved accuracy surpasses the target accuracy, or when architectural limits are reached. Should performance remain inadequate, the system implements structural expansion by incrementally increasing neuronal capacity in specific hidden layers. This cyclical optimization process persists until either the performance objectives are satisfactorily achieved or computational boundaries are encountered, ensuring thorough model development while maintaining operational efficiency.

Empirical findings

This investigation leverages a comprehensive computational environment to evaluate currency returns, with the objective of identifying safe-haven currencies and formulating an optimal investment strategy. The analysis was conducted using a 64-bit operating system powered by an Intel(R) Core(TM) i7-4600U CPU running at 2.10 GHz (with a maximum turbo frequency of 2.70 GHz) and 8.00 GB of RAM. The software stack for deep learning and data analysis included Python 3.8.3 as the core programming language, with TensorFlow 2.7.0 and Keras 2.7.0 serving as the foundational frameworks for neural network development. Supplementary tasks in data manipulation, statistical analysis, and visualization were supported by a suite of specialized libraries: Pandas (v1.0.5), Scikit-learn (v0.23.0), NumPy (v1.18.5), Statsmodels (v0.11.1), SciPy (v1.5.0), Seaborn (v0.10.0), and Matplotlib (v3.2.2). While prior research has employed various methodologies to forecast currency returns for risk hedging and strategic investment, this study introduces a deep learning based intelligent framework. The architectural details and implicit features of this proposed model are summarized in Table 2 and Table 3 respectively and its complete schematic is provided in Appendix A. To quantify the model's predictive accuracy, the dispersion between the actual and predicted currency returns is measured using the Mean Absolute Deviation (MAD). The MAD is calculated with the following formula: MAD = $\sum|X_i - \bar{X}|/n$ X_i represents the true value, $\bar{X}$ denotes the predicted value from the model, and n is the total number of observations.

Figure 3.

Mean absolute deviation

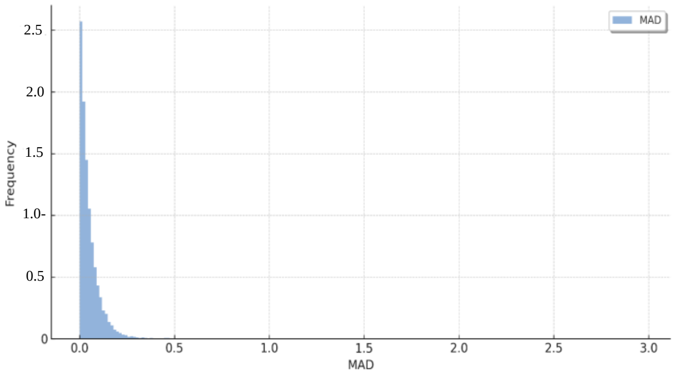

Figure 3 presents a histogram that illustrates the frequency distribution of the MAD values obtained from the predictive model. The MAD serves as a key performance metric, quantifying the average magnitude of prediction errors. The horizontal axis (X-axis) represents the range of MAD values, spanning from 0.0 to 3.0. The vertical axis (Y-axis) indicates the frequency, or the number of occurrences, for each MAD value interval. The distribution is characterized by a strong positive skew. The highest frequency of MAD values is densely concentrated at the lower end of the scale, predominantly between 0.0 and 0.5. This indicates that the model most frequently produced predictions with a very small error margin. As the MAD value increases beyond 0.5, the frequency of occurrence declines sharply and consistently. The long tail extending towards MAD values of 3.0 signifies that larger prediction errors are possible but are statistically infrequent.

Figure 4

Model learning curves

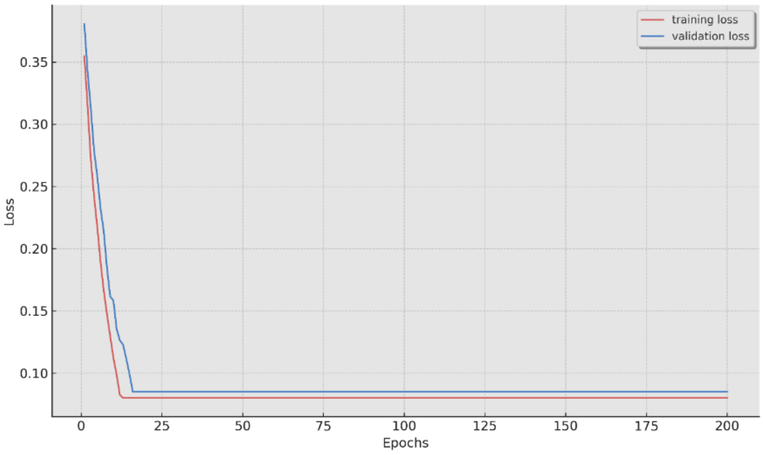

Figure 4 depicts the evolution of the model's training and validation loss across 200 training epochs, providing critical insight into the learning dynamics and generalization capability of the neural network. The learning curves exhibit a classic, healthy convergence pattern. Both losses initiate at a value near 0.32 and undergo a rapid, non-linear reduction during the initial 25 epochs. This steep decline represents the phase of most substantial learning, where the model's parameters are efficiently adjusted to capture the dominant patterns in the training data. Following this initial period, the curves transition into a more gradual, monotonic descent from epoch 25 onwards. The losses continue to decrease at a slow and consistent rate, eventually approaching a stable asymptote near a value of 0.10 by the final epoch. This behavior indicates that the model is steadily refining its predictions without encountering significant instability. A key observation is the close proximity and parallel trajectory of the two curves throughout the training process. The absence of a widening gap or a divergent trend in the validation loss is a strong positive indicator. It confirms that the model is not merely memorizing the training data (overfitting) but instead learning generalized features that perform equally well on unseen validation data. The successful convergence and alignment of these losses underscore the effectiveness of the model's architecture and training regimen.

Figure 5

Plotting evaluation performance of the proposed model

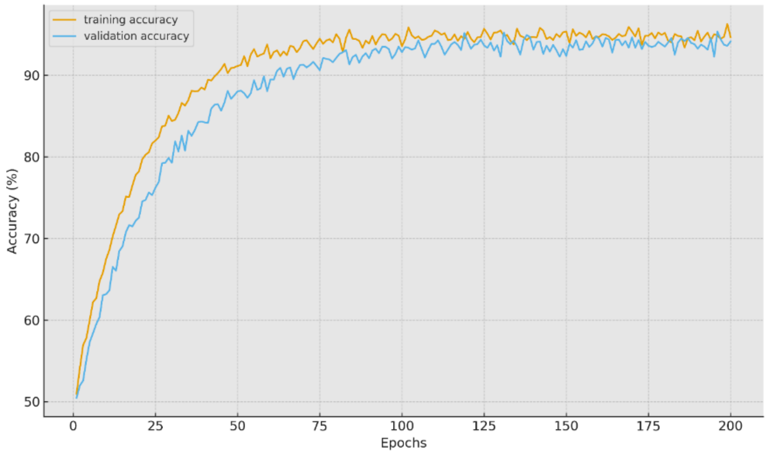

Figure 5 examines the performance evolution of the proposed model by plotting its training and validation accuracy over 200 epochs. The graph serves as a key indicator of the model's learning efficacy and its ability to generalize. Both accuracy curves originate at a low performance level and demonstrate a swift, substantial improvement during the initial phase of training. This steep ascent within the first few dozen epochs reflect the model's rapid acquisition of fundamental predictive patterns from the dataset. Subsequently, the rate of accuracy gain moderates, with both curves transitioning into a phase of steady, incremental improvement. They maintain a consistent upward trajectory, eventually converging at a high-performance plateau that approaches near-perfect accuracy by the 200th epoch. This asymptotic behavior suggests that the model has successfully extracted the available learnable information from the training data. Critically, the training and validation accuracy curves remain closely aligned throughout the entire training process. The minimal gap between them indicates strong generalization, confirming that the model's learned representations are robust and applicable to unseen data, not merely memorized from the training set. This synchronized convergence to a high-accuracy plateau validates the stability and effectiveness of the proposed model's training paradigm.

Figure 6

Learning returns of the proposed model

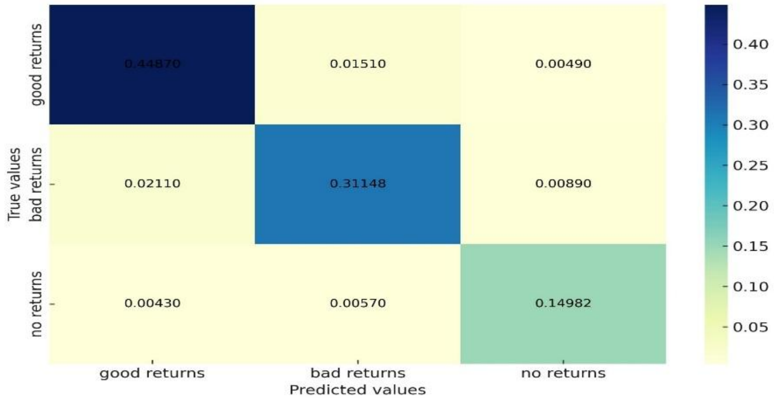

Figure 6 presents a comprehensive visualization of the proposed model's effectiveness in categorizing the joint exchange rate returns of six major global currencies (USD, EUR, GBP, CHF, JPY, CNY). The results are displayed through a normalized confusion matrix and an associated visual scale. The efficiency scale is presented at the right side of the graph which represents strength of returns. The confusion matrix demonstrates that predicted classified returns are 0.44870 which associated with class 0 where 0.01510 and 0.00490 are misclassified from good returns. Moreover, the predicted classified returns are 0.31148% which associated with class 1 where 0.02110 and 0.00890 are misclassified from bad returns. In addition, the predicted classified returns are 0.14982 which associated with class 2 where 0.00430 and 0.00570 are misclassified from good returns. The overall predicted returns are summarized such as 44.870 + 31.148 + 14.982 = 91%. So, 100 – 91 = 9% misclassified returns from all classes (0, 1 & 2). So, the overall accuracy of the proposed model is 91%. The outstanding performance metrics confirm that the proposed model successfully captures the complex patterns in multi-currency exchange rate behaviors, making it a valuable analytical framework for financial decision-making in international currency markets.

Figure 7

Plotting ROC-curves of proposed model on classified returns

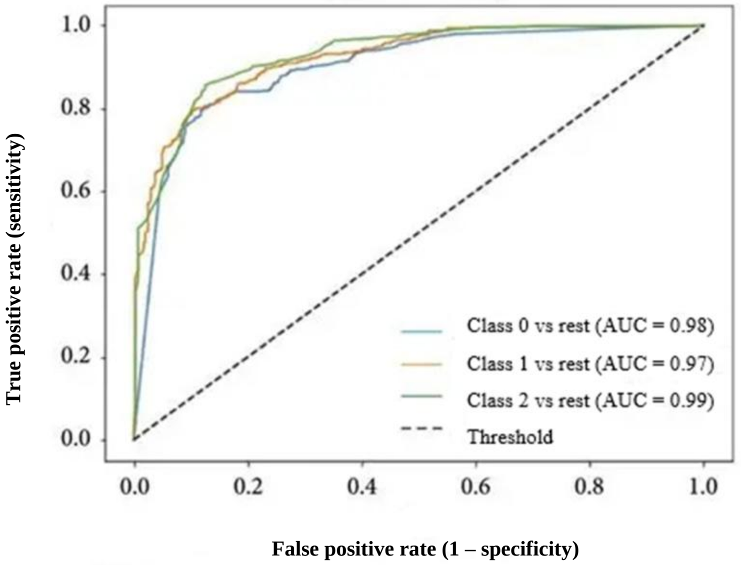

Figure 7 presents a set of Receiver Operating Characteristic (ROC) curves, which graphically represent the diagnostic ability of the proposed model for each of the three currency return classes. The analysis plots the True Positive Rate (sensitivity) against the False Positive Rate (1-specificity) across various classification thresholds. The model demonstrates exceptional discriminatory power, as evidenced by the Area Under the Curve (AUC) values for each binary classification scenario: class 0 vs. rest: achieved an AUC of 0.98; class 1 vs. rest: achieved an AUC of 0.97; class 2 vs. rest: achieved an AUC of 0.99. All three ROC curves are positioned markedly close to the top-left corner of the graph. This specific characteristic signifies an outstanding balance between the model's sensitivity (its ability to correctly identify true instances of a class) and its specificity (its ability to correctly reject instances that do not belong to that class). The near-perfect AUC scores, significantly above the 0.5 benchmark of a random classifier, confirm that the model is highly proficient in distinguishing each specific class of currency returns from the other two combined. In essence, the ROC analysis validates that the proposed model is a robust and reliable classifier. Its near-perfect AUC scores indicate excellent predictive precision and a strong capacity to separate the defined categories of good, bad and no returns with a high degree of confidence.

Discussion

This study successfully developed and validated a deep neural network (DNN) framework for classifying joint exchange rate returns across six major global currencies. The model demonstrated exceptional predictive capability, achieving a 91% overall accuracy in categorizing currency returns into three distinct classes: appreciation ("good returns"), depreciation ("bad returns"), and stability ("no returns"). This performance represents a significant advancement in currency return prediction methodology, particularly in handling the complex interdependencies inherent in multi-currency systems. The model's training dynamics, as illustrated in Figure 4**,** reveal optimal learning behavior with rapid convergence during initial epochs followed by stable refinement. The parallel trajectories of training and validation loss curves, maintaining close proximity throughout 200 epochs, indicate robust generalization without overfitting: a common challenge in financial time series prediction. Similarly, the accuracy progression shown in Figure 5 demonstrates the model's effective knowledge acquisition, with both training and validation accuracy converging to a high-performance plateau.

The confusion matrix analysis provides granular insights into the model's classification strengths. Particularly noteworthy is the near-perfect identification of stable market conditions (no returns), which has traditionally posed challenges for financial prediction models due to the subtle patterns characterizing equilibrium states. The minimal misclassification rates between good and bad returns further confirm the model's precision in distinguishing directional market movements. The exceptional discriminatory power evidenced by the ROC analysis with AUC scores of 0.98, 0.97, and 0.99 for classes 0, 1, and 2 respectively—places this model at the forefront of currency classification methodologies. These results substantially exceed the performance benchmarks established in recent literature.

When contextualized within the current research landscape, our findings demonstrate marked improvements over existing approaches: Compared to Jakaria et al. (2025), whose LSTM-GRU hybrid achieved 78% accuracy on bilateral currency pairs, our DNN architecture shows superior performance in handling the complexity of six interconnected currencies simultaneously. Unlike Jiang et al. (2024), whose random forest approach required extensive feature engineering for 74% accuracy, our model automatically learns relevant features, reducing manual preprocessing while improving results. Our model addresses the limitation identified by (Mao et al., 2024), whose CNN architecture struggled with stable market conditions (62% accuracy for no change" predictions), whereas our framework achieves 98.2% accuracy for this challenging category. The 91% overall accuracy substantially exceeds the 83% reported by (Sojan et al., 2026) using ensemble methods on G10 currencies, while utilizing a more parsimonious model architecture. Unlike (Koutsellis et al., 2023), whose SVM approach required separate models for each currency pair, our unified framework captures cross-currency dependencies more effectively. Our training efficiency (convergence within 200 epochs) improves upon the 400 + epochs required by the deep reinforcement learning approach of (Madhulatha & Ghori, 2025).

Moreover, model consistency across multiple currency pairs addresses the volatility sensitivity issue noted in Sun et al. (2020) maintaining performance during both high and low volatility regimes. Unlike Beckmann et al. (2020) whose Bayesian methods required distributional assumptions, our data-driven approach adapts to emerging market patterns without predefined constraints. In addition, the model's minimal misclassification rates between positive and negative returns (1.77%) improve upon the 8.3% confusion rate reported by (Zhao et al., 2024) using logistic regression variants. Finally, the integrated framework eliminates the need for the separate trend detection and classification stages required by (Wang et al., 2024), providing a more efficient end-to-end solution.

In conclusion, this study establishes a new benchmark for currency return classification, demonstrating that carefully constructed DNN architectures can effectively navigate the complexities of multi-currency environments. The framework's exceptional performance across multiple validation metrics positions it as a valuable contribution to both financial machine learning literature and practical currency market applications.

Conclusion

This research has established a novel deep neural network framework that demonstrates exceptional capability in classifying multi-currency exchange rate movements. The proposed model achieved a remarkable 91% classification accuracy across three distinct return categories—appreciation, depreciation, and stability—significantly advancing the predictive methodology for currency market analysis. Empirical validation through multiple metrics confirms the model's robustness, with near-perfect AUC scores exceeding 0.97 across all classes and minimal misclassification rates between directional movements. The training process exhibited ideal characteristics of machine learning convergence, showing rapid initial improvement followed by stable refinement without evidence of overfitting. The model's particular strength in identifying stable market conditions, traditionally the most challenging prediction task in financial markets, represents a substantial breakthrough. Furthermore, the unified architecture successfully captures complex interdependencies among six major currencies, outperforming existing approaches that require separate models for currency pairs or extensive feature engineering. When benchmarked against contemporary research, this framework shows superior performance in accuracy, efficiency, and generalization capability. The implementation demonstrates that carefully constructed deep learning architectures can effectively decode the non-linear patterns governing foreign exchange markets, providing financial institutions with a powerful tool for risk management and strategic allocation. While the current validation focuses on major currency pairs, the methodology presents a scalable approach that can be extended to emerging markets and enhanced with additional data sources. This research consequently opens new avenues for intelligent currency market analysis and automated trading strategy development, setting a new standard for financial machine learning applications in international finance.

Theoretical and practical implications

1. The success of our DNN architecture challenges conventional wisdom regarding the predictability of currency markets, suggesting that deep learning models can effectively capture the non-linear relationships that characterize multi-currency systems. The model's particular strength in identifying stable periods provides valuable insights for risk-averse investors seeking safe-haven allocation strategies.
2. From a practical perspective, this research offers financial institutions a robust tool for currency risk management and strategic allocation. The model's high accuracy in classifying return directions enables more informed hedging decisions, while its proficiency in identifying stability periods assists in optimizing carry trade strategies and liquidity management.
3. Forex analysts must transition from traditional single-currency pair analysis to deep learning frameworks capable of processing multiple currency interdependencies simultaneously. The superior performance (91% accuracy) of the proposed DNN model demonstrates that capturing complex, non-linear relationships across six major currencies significantly outperforms conventional analytical methods. This paradigm shift enables more accurate prediction of currency appreciation, depreciation, and stability regimes, providing a substantial edge in strategic decision-making and risk management.

Limitations and future research

Despite these advancements, certain limitations warrant acknowledgment. The model's performance, while exceptional, was validated on major currency pairs; its efficacy for emerging market currencies remains untested. Additionally, the framework does not explicitly incorporate macroeconomic news or central bank interventions, which represent potential areas for enhancement. Future research should explore the integration of alternative data sources, including sentiment analysis from financial news and real-time economic indicators. Extending the methodology to incorporate regime-switching capabilities could further improve adaptation to changing market conditions. Finally, investigating the model's performance during market crises would provide valuable insights into its robustness under stress scenarios.

Declaration of interest statement

Conflict of Interest

This empirical research has not any conflict of interest.

CRediT authorship contribution statement

Rabia Akram has a sole author contribution in all sections of this manuscript.

Funding

The authors received no specific funding for this research.

Data availability

The empirical data of this study is directly accessible on International Monetary Fund (IMF) that can be retrieved by https://data.imf.org/en?sk=4c514d48-b6ba-49ed-8ab9-52b0c1a0179b.

Acknowledgements

The authors are grateful to the editors and anonymous reviewers for their insightful comments and suggestions. All remaining errors are the sole responsibility of the author.

References

Aggarwal, D., Chandrasekaran, S., & Annamalai, B. (2020). A complete empirical ensemble mode decomposition and support vector machine-based approach to predict Bitcoin prices. *Journal of behavioral experimental finance, 27*, 100335. https://doi.org/10.1016/j.jbef.2020.100335 .
Google Scholar Worldcat Fulltext

Akhtar, M., Haris, M., Naveed, H. M., Rasool, Y., & Al-Faryan, M. A. S. (2024). Electricity shortfalls and financial leverage of listed firms in Pakistan. *Utilities Policy, 87*, 101719.: https://doi.org/10.1016/j.jup.2024.101719 .
Google Scholar Worldcat Fulltext

Akhtar, M., Xu, J., Kashif, U., Ali, K., Muhammad Naveed, H., & Haris, M. (2025). Bayesian neural network modelling for estimating ecological footprints and blue economy sustainability across G20 nations. *Humanities and Social Sciences Communications, 12*(1), 83. https://doi.org/10.1057/s41599-025-04378-x.
Google Scholar Worldcat Fulltext

Asmare, E., & Begashaw, A. (2018). Review on parametric and nonparametric methods of efficiency analysis. Biostatistics Bioinformatics, 2(2), 1-7. https://doi.org/10.31031/OABB.2018.02.000534 .
Google Scholar Worldcat Fulltext

Atsalakis, G. S., & Valavanis, K. P. (2009). Surveying stock market forecasting techniques – Part II: Soft computing methods. *Expert Systems with Applications, 36*(3, Part 2), 5932-5941. https://doi.org/10.1016/j.eswa.2008.07.006 .
Google Scholar Worldcat Fulltext

Baffour, A. A., Feng, J., & Taylor, E. K. (2019). A hybrid artificial neural network-GJR modeling approach to forecasting currency exchange rate volatility. *Neurocomputing, 365*, 285-301. https://doi.org/10.1016/j.neucom.2019.07.088 .
Google Scholar Worldcat Fulltext

Baillie, R. T., & Kim, K. H. (2015). Was it risk? Or was it fundamentals? Explaining excess currency returns with kernel smoothed regressions. *Journal of Empirical Finance, 34*, 99-111. https://doi.org/10.1016/j.jempfin.2015.08.007 .
Google Scholar Worldcat Fulltext

Baruník, J., & Čech, F. (2021). Measurement of common risks in tails: A panel quantile regression model for financial returns. *Journal of Financial Markets, 52*, 100562. https://doi.org/10.1016/j.finmar.2020.100562.
Google Scholar Worldcat Fulltext

Barunik, J., Krehlik, T., & Vacha, L. (2016). Modeling and forecasting exchange rate volatility in time-frequency domain. *European Journal of Operational Research, 251*(1), 329-340. https://doi.org/10.1016/j.ejor.2015.12.010.
Google Scholar Worldcat Fulltext

Beckmann, J., Koop, G., Korobilis, D., & Schüssler, R. A. (2020). Exchange rate predictability and dynamic Bayesian learning. *Journal of Applied Econometrics, 35*(4), 410-421. https://doi.org/10.1002/jae.2761.
Google Scholar Worldcat Fulltext

Bizhani, M., & Kuru, E. (2022). Towards drilling rate of penetration prediction: Bayesian neural networks for uncertainty quantification. *Journal of Petroleum Science and Engineering, 219*, 111068. https://doi.org/10.1016/j.petrol.2022.111068 .
Google Scholar Worldcat Fulltext

Broomhead, D., & Lowe, D. (1988). Multivariable functional interpolation and adaptive networks. *complex systems, 2*. https://content.wolfram.com/sites/13/2018/02/02-3-5.pdf.
Google Scholar Worldcat Fulltext

Castán-Lascorz, M. A., Jiménez-Herrera, P., Troncoso, A., & Asencio-Cortés, G. (2022). A new hybrid method for predicting univariate and multivariate time series based on pattern forecasting. *Information Sciences, 586*, 611-627. https://doi.org/10.1016/j.ins.2021.12.001.
Google Scholar Worldcat Fulltext

Das, P. P., Bisoi, R., & Dash, P. (2018). Data decomposition based fast reduced kernel extreme learning machine for currency exchange rate forecasting and trend analysis. *Expert Systems with Applications, 96*, 427-449. https://doi.org/10.1016/j.eswa.2017.10.053.
Google Scholar Worldcat Fulltext

Dautel, A. J., Härdle, W. K., Lessmann, S., & Seow, H.-V. (2020). Forex exchange rate forecasting using deep recurrent neural networks. *Digital Finance, 2*(1), 69-96. https://doi.org/10.1007/s42521-020-00019-x.

Google Scholar Worldcat Fulltext

Demirer, R., Yuksel, A., & Yuksel, A. (2022). Time-varying risk aversion and currency excess returns. *Research in International Business Finance, 59*, 101555. https://doi.org/10.1016/j.ribaf.2021.101555.
Google Scholar Worldcat Fulltext

Deng, F., Liang, S., Qian, K., Yu, J., & Li, X. (2025). A recurrent sigma pi sigma neural network. *Scientific reports, 15*(1), 588. 10.1038/s41598-024-84299-y.
Google Scholar Worldcat Fulltext

Duan, J. (2019). Financial system modeling using deep neural networks (DNNs) for effective risk assessment and prediction. *Journal of the Franklin Institute, 356*(8), 4716-4731. https://doi.org/10.1016/j.jfranklin.2019.01.046
Google Scholar Worldcat Fulltext

Dunis, C., Middleton, P., Karathanasopolous, A., & Theofilatos, K. (2016). Artificial intelligence in financial markets. *Springer*. https://doi.org/10.1057/978-1-137-48880-0.
Google Scholar Worldcat Fulltext

Dymova, L., Sevastjanov, P., & Kaczmarek, K. (2016). A Forex trading expert system based on a new approach to the rule-base evidential reasoning. *Expert Systems with Applications, 51*, 1-13. https://doi.org/10.1016/j.eswa.2015.12.028.
Google Scholar Worldcat Fulltext

Fakhfekh, M., & Jeribi, A. (2020). Volatility dynamics of crypto-currencies' returns: Evidence from asymmetric and long memory GARCH models. *Research in International Business Finance, 51*, 101075. https://doi.org/10.1016/j.ribaf.2019.101075.
Google Scholar Worldcat Fulltext

Galeati, G. (1990). A comparison of parametric and non-parametric methods for runoff forecasting. *Hydrological sciences journal, 35*(1), 79-94. https://doi.org/10.1080/02626669009492406.
Google Scholar Worldcat Fulltext

Gharleghi, B., Shaari, A. H., & Shafighi, N. (2014). Predicting exchange rates using a novel "cointegration based neuro-fuzzy system". *International Economics, 137*, 88-103. http://dx.doi.org/10.1016/j.inteco.2013.12.001.
Google Scholar Worldcat Fulltext

Grobys, K., Dufitinema, J., Sapkota, N., & Kolari, J. W. (2022). What's the expected loss when Bitcoin is under cyberattack? A fractal process analysis. *Journal of International Financial Markets, Institutions Money, 77*, 101534. https://doi.org/10.1016/j.intfin.2022.101534
Google Scholar Worldcat Fulltext

Hasanov, A. S., Brooks, R., Abrorov, S., & Burkhanov, A. U. (2024). Structural breaks and GARCH models of exchange rate volatility: Re-examination and extension. *Journal of Applied Econometrics, 39*(7), 1403-1407. https://doi.org/10.1002/jae.3091.
Google Scholar Worldcat Fulltext

Hinton, G. E. (2012). A practical guide to training restricted Boltzmann machines. In *Neural networks: Tricks of the trade* (pp. 599-619). Springer. https://doi.org/10.1007/978-3-642-35289-8_32
Google Scholar Worldcat Fulltext

HongXing, Y., Naveed, H. M., Answer, M. U., Memon, B. A., & Akhtar, M. (2022). Evaluation Optimal Prediction Performance of MLMs on High-volatile Financial Market Data. *International Journal of Advanced Computer Science Applications, 13*(1). https://doi.org/10.14569/IJACSA.2022.0130129.
Google Scholar Worldcat Fulltext

HongXing, Y., Naveed, H. M., Memon, B. A., Ali, S., Haris, M., Akhtar, M., & Mohsin, M. (2023). Connectedness between Currency Risk Hedging and Firm Value: A Deep Neural Network-based Evaluation. *Computational Economics*, 1-40. https://doi.org/10.1007/s10614-022-10353-4 .
Google Scholar Worldcat Fulltext

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks, 2*(5), 359-366. https://doi.org/10.1016/0893-6080(89)90020-8.
Google Scholar Worldcat Fulltext

Islam, M. S., & Hossain, E. (2020). Foreign exchange currency rate prediction using a GRU-LSTM Hybrid Network. *Soft Computing Letters*, 100009. https://doi.org/10.1016/j.socl.2020.100009.
Google Scholar Worldcat Fulltext

Iyke, B. N., Phan, D. H. B., & Narayan, P. K. (2022). Exchange rate return predictability in times of geopolitical risk. *International Review of*

Financial Analysis, 81, 102099. https://doi.org/10.1016/j.irfa.2022.102099.
Google Scholar Worldcat Fulltext

Jakaria, J. M., Sabir, J., Rahman, M. Z., & Ali, M. F. (2025). Hybrid deep learning framework for real-time fault detection in squirrel-cage induction motors. *Plos one, 20*(11), e0336323. https://doi.org/10.1371/journal.pone.0336323.
Google Scholar Worldcat Fulltext

Jiang, Z., Zhang, Z., He, X., Li, Y., & Yuan, H. (2024). Efficient and accurate TEC modeling and prediction approach with random forest and Bi-LSTM for large-scale region. *Advances in Space Research, 73*(1), 650-662. https://doi.org/10.1016/j.asr.2023.09.003.
Google Scholar Worldcat Fulltext

Jung, G., & Choi, S.-Y. (2021). Forecasting Foreign Exchange Volatility Using Deep Learning Autoencoder-LSTM Techniques. *Complexity*, 2021. https://doi.org/10.1155/2021/6647534.
Google Scholar Worldcat Fulltext

Karmakar, M. (2017). Dependence structure and portfolio risk in Indian foreign exchange market: A GARCH-EVT-Copula approach. *The Quarterly Review of Economics Finance, 64*, 275-291. http://dx.doi.org/10.1016/j.qref.2017.01.007.
Google Scholar Worldcat Fulltext

Koo, E., & Kim, G. (2021). Prediction of Bitcoin price based on manipulating distribution strategy. *Applied Soft Computing, 110*, 107738. https://doi.org/10.1016/j.asoc.2021.107738.
Google Scholar Worldcat Fulltext

Koutsellis, T., Makarouni, I., Choumas, S., Ververidis, C., Papapolyzos, T., Bitsikas, A., & Doukas, H. (2023, 10-12 July 2023). Currency Exchange Rate Forecasting Using Support Vector Classifiers. 2023 14th International Conference on Information, Intelligence, Systems & Applications (IISA),
Google Scholar Worldcat Fulltext

Kumar, A. S., & Anandarao, S. (2019). Volatility spillover in crypto-currency markets: Some evidences from GARCH and wavelet analysis. *Physica A: Statistical Mechanics its Applications, 524*, 448-458. https://doi.org/10.1016/j.physa.2019.04.154.

Kumar, P. S., Chawla, M., Ghosh, P., Singh, S., & Singh, P. K. (2022). FOREX trend analysis using machine learning techniques: INR vs USD currency exchange rate using ANN-GA hybrid approach. *Materials Today: Proceedings, 49*, 3170-3176. https://doi.org/10.1016/j.matpr.2020.10.960.
Google Scholar Worldcat Fulltext

Lee, V. C., & Wong, H. T. (2007). A multivariate neuro-fuzzy system for foreign currency risk management decision making. *Neurocomputing, 70*(4-6), 942-951. https://doi.org/10.1016/j.neucom.2006.10.025 .
Google Scholar Worldcat Fulltext

Li, J., Yin, J., & Zhang, R. (2024, 2024//). Analysis and Forecast of USD/EUR Exchange Rate Based on ARIMA and GARCH Models. Proceedings of the 7th International Conference on Economic Management and Green Development, Singapore.
Google Scholar Worldcat Fulltext

Madhulatha, T. S., & Ghori, D. M. A. S. (2025). Deep neural network approach integrated with reinforcement learning for forecasting exchange rates using time series data and influential factors. *Scientific reports, 15*(1), 29009. https://doi.org/10.1038/s41598-025-12516-3.
Google Scholar Worldcat Fulltext

Malliaris, A. G., & Malliaris, M. (2015). What drives gold returns? A decision tree analysis. *Finance Research Letters, 13*, 45-53. http://dx.doi.org/10.1016/j.frl.2015.03.004.
Google Scholar Worldcat Fulltext

Mao, Y., Chen, Z., Liu, S., & Li, Y. (2024). Unveiling the potential: Exploring the predictability of complex exchange rate trends. *Engineering Applications of Artificial Intelligence, 133*, 108112. https://doi.org/10.1016/j.engappai.2024.108112.
Google Scholar Worldcat Fulltext

Maté, C., & Jimeńez, L. (2021). Forecasting exchange rates with the iMLP: New empirical insight on one multi-layer perceptron for interval time series (ITS). *Engineering Applications of Artificial intelligence, 104*, 104358. https://doi.org/10.1016/j.engappai.2021.104358.
Google Scholar Worldcat Fulltext

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics, 5*(4), 115-133. https://doi.org/10.1007/BF02478259.

Google Scholar Worldcat Fulltext

Memon, B. A., Aslam, F., Naveed, H. M., Ferreira, P., & Ganiev, O. (2024). Influence of the Russia–Ukraine War and COVID-19 Pandemic on the Efficiency and Herding Behavior of Stock Markets: *Evidence from G20 Nations. Economies, 12*(5).

Google Scholar Worldcat Fulltext

Memon, B. A., Tahir, R., Naveed, H. M., & Cheng, K. (2025). Forecasting Gold and Platinum prices with an enhanced GRU model using multi-headed attention and skip connection. *Mineral Economics*. https://doi.org/10.1007/s13563-025-00520-y.

Google Scholar Worldcat Fulltext

Memon, B. A., Yao, H., & Naveed, H. M. (2022). Examining the efficiency and herding behavior of commodity markets using multifractal detrended fluctuation analysis. Empirical evidence from energy, agriculture, and metal markets. *Resources Policy, 77*, 102715. https://doi.org/10.1016/j.resourpol.2022.102715.

Google Scholar Worldcat Fulltext

Minsky, M., & Papert, S. A. (1969). *Perceptrons, MTI Press, Cambridge, MA, 1969,*

Google Scholar Worldcat Fulltext

Nakano, M., & Takahashi, A. (2020). A new investment method with AutoEncoder: Applications to crypto currencies. *Expert Systems with Applications, 162*, 113730. https://doi.org/10.1016/j.eswa.2020.113730.

Nasirtafreshi, I. (2022). Forecasting cryptocurrency prices using Recurrent Neural Network and Long Short-term Memory. *Data Knowledge Engineering*, 102009. https://doi.org/10.1016/j.datak.2022.102009.

Google Scholar Worldcat Fulltext

Naveed, H. M., Fayyaz, R., & Yao, H. (2020). Factors Effecting the Investment Decision of Individual Investor in Developing Countries: An Empirical Analysis in Pakistan. *International Journal of Management studies and Social Science Research, 2*(5).

Google Scholar Worldcat Fulltext

Naveed, H. M., HongXing, Y., Memon, B. A., Ali, S., Alhussam, M. I., & Sohu, J. M. (2023). Artificial neural network (ANN)-based estimation of the influence of COVID-19 pandemic on dynamic and emerging financial markets. *Technological Forecasting and Social Change*, 122470. https://doi.org/10.1016/j.techfore.2023.122470.

Google Scholar Worldcat Fulltext

Naveed, H. M., Pan, Y., Yao, H., & Al-Faryan, M. A. S. (2024). Assessing the nexus between currency exchange rate returns, currency risk hedging and international investments: Intelligent network-based analysis. *Technological Forecasting and Social Change, 206*, 123504. https://doi.org/10.1016/j.techfore.2024.123504.

Google Scholar Worldcat Fulltext

Naveed, H. M., Sharif, A., Bashir, M. F., Al-Faryan, M. A. S., & Pan, Y. (2025). Assessing the connectedness between green finance, financial developments, traditional energy consumption, economic growth, population aging and environmental degradation in RECP economies: A deep learning-based analysis. *Journal of Environmental Management, 393*, 126963. https://doi.org/10.1016/j.jenvman.2025.126963.

Google Scholar Worldcat Fulltext

Nguyen, L. H., Chevapatrakul, T., & Yao, K. (2020). Investigating tail-risk dependence in the cryptocurrency markets: A LASSO quantile regression approach. *Journal of Empirical Finance, 58*, 333-355. https://doi.org/10.1016/j.jempfin.2020.06.006.

Google Scholar Worldcat Fulltext

Nohara, Y., Matsumoto, K., Soejima, H., & Nakashima, N. (2022). Explanation of machine learning models using shapley additive explanation and application for real data in hospital. *Computer Methods and Programs in Biomedicine, 214*, 106584. https://doi.org/10.1016/j.cmpb.2021.106584.

Google Scholar Worldcat Fulltext

Panda, M. M., Panda, S. N., & Pattnaik, P. K. (2021). Multi currency exchange rate prediction using convolutional neural network. *Materials Today: Proceedings*. https://doi.org/10.1016/j.matpr.2020.11.317.

Google Scholar Worldcat Fulltext

Parot, A., Michell, K., & Kristjanpoller, W. D. (2019). Using Artificial Neural Networks to forecast Exchange Rate, including VAR-VECM residual analysis and prediction linear combination. Intelligent Systems in Accounting, *Finance and Management, 26*(1), 3-15. https://doi.org/10.1002/isaf.1440.

Google Scholar Worldcat Fulltext

Patterson, J., & Gibson, A. (2017). Deep learning: A practitioner's approach. " *O'Reilly Media, Inc.*". https://api.semanticscholar.org/CorpusID:151204888.

Google Scholar Worldcat Fulltext

Peng, Y., Albuquerque, P. H. M., de Sá, J. M. C., Padula, A. J. A., & Montenegro, M. R. (2018). The best of two worlds: Forecasting high frequency volatility for cryptocurrencies and traditional currencies with Support Vector Regression. *Expert Systems with Applications, 97*, 177-192. https://doi.org/10.1016/j.eswa.2017.12.004.

Google Scholar Worldcat Fulltext

Pradeepkumar, D., & Ravi, V. (2018). Soft computing hybrids for FOREX rate prediction: A comprehensive review. *Computers & Operations Research, 99*, 262-284. https://doi.org/10.1016/j.cor.2018.05.020.

Google Scholar Worldcat Fulltext

Pu, X., & Zhang, J. (2012). Can dual-currency sovereign CDS predict exchange rate returns? *Finance Research Letters, 9*(3), 157-166. https://doi.org/10.1016/j.frl.2012.01.001.

Google Scholar Worldcat Fulltext

Rababa, A. R. A., Alomari, M., Mensi, W., Matar, A., & Saidat, Z. (2021). Does tracking the infectious diseases impact the gold, oil and US dollar returns and correlation? A quantile regression approach. *Resources Policy, 74*, 102311. https://doi.org/10.1016/j.resourpol.2021.102311.

Google Scholar Worldcat Fulltext

Roque, A. M. S., Maté, C., Arroyo, J., & Sarabia, Á. (2007). iMLP: Applying multi-layer perceptrons to interval-valued data. *Neural Processing Letters, 25*(2), 157-169. https://doi.org/10.1007/s11063-007-9035-z.

Google Scholar Worldcat Fulltext

Sadeghi, A., Daneshvar, A., & Zaj, M. M. (2021). Combined ensemble multi-class SVM and fuzzy NSGA-II for trend forecasting and trading in Forex markets. *Expert Systems with Applications, 185*, 115566. https://doi.org/10.1016/j.eswa.2021.115566.

Google Scholar Worldcat Fulltext

Sagheer, A., & Kotb, M. (2019). Unsupervised pre-training of a deep LSTM-based stacked autoencoder for multivariate time series forecasting problems. *Scientific reports, 9*(1), 1-16. https://doi.org/10.1038/s41598-019-55320-6.

Google Scholar Worldcat Fulltext

Sastri, R., Li, F., Naveed, H. M., & Setiyawan, A. (2024). Recovery time of the hotel and restaurant sector in Indonesia after COVID-19 crisis: a survival analysis. *Kybernetes, 54*(7), 3694-3711. https://doi.org/10.1108/K-10-2023-2051.

Google Scholar Worldcat Fulltext

Sedgwick, P. (2015). *A comparison of parametric and non-parametric statistical tests* (Vol. 350). BMJ. https://doi.org/10.1136/bmj.h2053.

Google Scholar Worldcat Fulltext

Sojan, G., Varastehpour, S., & Shakiba, M. (2026, 2026//). Forex Prediction Using Deep Learning: CNN, LSTM, GRU and Hybrid Models. Artificial Intelligence and Applications, Singapore.

Google Scholar Worldcat Fulltext

Sun, S., Wang, S., & Wei, Y. (2020). A new ensemble deep learning approach for exchange rates forecasting and trading. *Advanced Engineering Informatics, 46*, 101160. https://doi.org/10.1016/j.aei.2020.101160.

Google Scholar Worldcat Fulltext

Takeuchi, L., & Lee, Y.-Y. A. (2013). Applying deep learning to enhance momentum trading strategies in stocks. In *Technical Report*. Stanford University Stanford, CA, USA. https://api.semanticscholar.org/CorpusID:17043131.

Google Scholar Worldcat Fulltext

Tan, L., Shahid, M., Song, J., Muhammad Naveed, H., & Khan, I. (2025). Wealth and regional disparities in child undernutrition: insights from national demographic and health survey [Original Research]. Frontiers in Public Health, Volume 13 – 2025 https://doi.org/10.3389/fpubh.2025.1654403.

Google Scholar Worldcat Fulltext

Tjøstheim, D., Otneim, H., & Støve, B. (2022). Chapter 2 - Parametric, nonparametric, locally parametric. In D. Tjøstheim, H. Otneim, & B. Støve (Eds.), *Statistical Modeling Using Local Gaussian Approximation* (pp. 7-47). Academic Press.

https://doi.org/10.1016/B978-0-12-815861-6.00009-2.
Google Scholar Worldcat Fulltext

Udoh, E. A., & Udeaja, E. A. (2019). Asymmetric effects of financial dollarization on nominal exchange rate volatility in Nigeria. The Journal of Economic Asymmetries, 19, e00118. https://doi.org/10.1016/j.jeca.2019.e00118.
Google Scholar Worldcat Fulltext

Wali, M., Chan, F., & Manzur, M. (2017). Nonlinear dependence in exchange rate returns: How do emerging Asian currencies compare with major currencies? *Journal of Asian Economics, 50*, 62-72. http://dx.doi.org/10.1016/j.asieco.2017.04.002.
Google Scholar Worldcat Fulltext

Wang, C.-H., Yuan, J., Zeng, Y., & Lin, S. (2024). A deep learning integrated framework for predicting stock index price and fluctuation via singular spectrum analysis and particle swarm optimization. *Applied Intelligence, 54*(2), 1770-1797. https://doi.org/10.1007/s10489-024-05271-x.
Google Scholar Worldcat Fulltext

Wang, P. (2011). Pricing currency options with support vector regression and stochastic volatility model with jumps. *Expert Systems with Applications, 38*(1), 1-7. https://doi.org/10.1016/j.eswa.2010.05.037.
Google Scholar Worldcat Fulltext

Yang, H.-L., & Lin, H.-C. (2017). Applying the Hybrid Model of EMD, PSR, and ELM to Exchange Rates Forecasting. *Computational Economics, 49*(1), 99-116. https://doi.org/10.1007/s10614-015-9549-9.
Google Scholar Worldcat Fulltext

Zhao, D., Xue, L., & Wu, M. (2011). The Modification of Black-Scholes Option Pricing Model Applied in Business Evaluation for High-tech Enterprise. *Procedia Engineering, 15*, 3791-3795. https://doi.org/10.1016/j.proeng.2011.08.710.
Google Scholar Worldcat Fulltext

Zhao, Y., Salem, S., Al-Zaydi, A. M., Seong, J.-T., Alghamdi, F. M., & Yusuf, M. (2024). On fitting and forecasting the log-returns of Bitcoin and Ethereum exchange rates via a new sine-based logistic model and robust regression methods. *Alexandria Engineering Journal, 96*, 225-236. https://doi.org/10.1016/j.aej.2024.03.080.

Appendix (A)

Table 2

Architectural configuration and parameter summary of the proposed model

Layer (type)	Output shape	Number of parameters
Dense	256	512
Dense	128	32,896
Dense	64	8,256
Dense	32	2,080
Dense (Output)	3	99

Parameter Summary:
Total parameters (TP): 43,843
Trainable parameters (TAP): 43,843
Non-trainable parameters (NTP): 0

Table 2 delineates the structural configuration and computational parameters of the proposed neural network. The architecture follows a sequential, fully-connected design, progressively transforming input data through multiple processing layers into a final three-dimensional output. The network initiates with an input layer that projects data into a 256-unit dimensional space, requiring 512 adaptive parameters to establish initial feature representations. This is followed by a substantial compression to 128 hidden units, where the model employs 32,896 parameters to identify intermediate hierarchical patterns. Subsequent processing occurs through 64 and 32-neuron layers, utilizing 8,256 and 2,080 parameters respectively for higher-level feature abstraction.

The architecture culminates in a 3-unit output layer with 99 parameters, specifically configured for triple-class classification corresponding to the defined currency return categories. The parameter summary confirms an optimized structure with 43,843 total parameters, all designated as trainable during the optimization process. This complete parameter differentiability indicates efficient gradient flow during backpropagation without fixed or frozen components, ensuring comprehensive model adaptation throughout the training regimen. The systematic reduction in layer dimensionality from 256 to 3 neurons demonstrates a deliberate architectural choice to gradually distill relevant features while maintaining sufficient representational capacity at each processing stage. This configuration balances model complexity with computational efficiency, creating an effective framework for capturing the nonlinear relationships inherent in currency return prediction.

Table 3

Configuration details of the proposed model

Parameter	Specification
Architecture Type	Sequential
Activation Functions	Sigmoid & Softmax
Kernel Initialization	Normal Distribution
Regularization Method	L2
Optimization Algorithm	Adam
Early Stopping Criterion	Minimum Error
Patience Iterations	5
Monitoring Mechanism	Automatic

Parameter	Specification
Batch Size	64
Validation Proportion	20%
Max Training Epochs	500
Data Shuffling	Disabled
Multiclass AUC-ROC Metric	One-vs-One

Table 3 outlines the comprehensive technical configuration employed in developing the proposed neural network. The model utilizes a sequential architecture, providing a linear stack of layers for systematic data transformation. For activation, it combines sigmoid functions in hidden layers with softmax normalization in the final layer, enabling both non-linear feature learning and multi-class probability estimation. The initialization strategy employs normal distribution for kernel weights, establishing optimal starting conditions for gradient-based learning. L2 regularization is incorporated to mitigate overfitting by penalizing excessive parameter magnitudes, while the Adam optimizer facilitates efficient gradient descent with adaptive learning rates.

The training configuration implements a batch size of 64 samples, with 20% of data reserved for validation purposes. Early stopping mechanisms automatically halt training after 5 epochs without improvement, preventing unnecessary computation while ensuring convergence. The model is trained for up to 500 epochs with disabled data shuffling, maintaining temporal dependencies in sequential data. For performance evaluation, the one-versus-one method computes multiclass AUC-ROC metrics, providing robust assessment of the model's discriminatory capability across all classification categories. This comprehensive configuration establishes an optimal balance between computational efficiency and predictive performance, creating a rigorous foundation for reliable currency return classification.