Artificial Intelligence Applications And Innovations Vol. 2 No. 2 (2025)

Article history:
Received 5 July 2026
Revised 7 August 2026
Accepted 7 September 2026
Published online 1 April 2025

A Hybrid Pattern Extraction and Confidence-Weighted Volume Framework for Machine Learning-Based Forex Price Forecasting

Maysam Yazdanpanahi[1*], Morteza Zahedi[2] and Mohammad Mehdi Hosseini[3]

1. Department of Computer Engineering, Shahrood University of Technology, Shahrood, Iran.
2. Department of Computer Engineering, Shahrood University of Technology, Shahrood, Iran.
3. Department of Computer Engineering, Sha.C., Islamic Azad University, Shahrood, Iran.

[*]**Corresponding author's email address**: meysamyazdanpanahi@shahroodut.ac.ir

Article Info

Article type:
Original Research

How to cite this article:
Yazdanpanahi, M., Zahedi, M., & Hosseini, M.M. (2025). A Hybrid Pattern Extraction and Confidence-Weighted Volume Framework for Machine Learning-Based Forex Price Forecasting. *Artificial Intelligence Applications and Innovations, 2*(2), 48 - 73.
https://doi.org/10.61838/jaiai.2.2.4

A B S T R A C T

Since its inception, the Forex market has attracted various market players, including central banks, financial institutions, and even academic researchers, due to its high volatility and potential profitability following the technical analysis saying that "history tends to repeat itself", RWBCPE captures the historical patterns to anticipate future changes (up or down) within certain possible boundaries to avoid signals that lack confidence thresholds. RWBCPE underwent testing on the EUR/USD currency pair in the context of three different simple machine learning models: Random Forests, Support Vector Machines, and Long Short-Term Memory networks to demonstrate its strength and efficiency, even using simple, conventional machine learning and deep learning models. The framework achieved a higher accuracy than the traditional technical indicators, such as RSI and MACD. To evaluate the practical feasibility of the framework, we conducted a 25-year training set from 2000 to 2025 and a 1-year backtest from 2025 to 2026 with an initial balance of $10,000. The RWBCPE-driven strategy achieved 65.96% accuracy and 67.29% precision, growing the account balance to $34,844.75. The system also recorded a profit ratio of 2.17 and a maximum drawdown of 4.6%, demonstrating strong predictive performance and low-risk trading through confidence-weighted position size.

Keywords: *Algorithmic trading, Forex Price Forecasting, Risk Management, Feature Engineering, Binary-Coded Patterns.*

1. Introduction

1.1. Overview of the Forex Market

The combination of complexity and dynamism of the Forex market is precisely what drives academics to study financial trading [1] The largest, most liquid financial market is the Foreign exchange (Forex) market, with an average daily turnover of $7.5 trillion (as of 2022). Its high liquidity and accessibility attract a wide range of traders, but its non-stationarity, chaos, and opaque nature make price forecasting a substantial challenge [2, 3]. The Forex pricing is governed by a "black box" of a multitude of nonlinear and opaque system: economic, political and psychological [4]. The high volatility of such an ecosystem makes it essential to have deep expertise combined with cognitive dissonance and predictive value. A minor currency value shift can result in major ripples on the dependent instruments [5]. This makes the intelligent dissection of forecasting models a necessity. In this environment, human analysis is unable to produce the required optimal results, making automated decision support systems essential tools for lasting profitability.

1.2. Research Gap and Proposed Solution

Despite significant progress in Forex prediction through advanced feature engineering and machine learning, the existing literature suffers from two interrelated shortcomings:

Indicator Selection Ambiguity: Researchers routinely draw from a vast arsenal of technical indicators (e.g. moving averages, RSI, MACD, …). However, there is no consensus on which specific indicators or what minimal combination delivers robust out-of-sample performance. This inconsistency fuels overfitting and aggravates the curse of dimensionality when feature sets balloon without clear justification [2]. One such example is the work of [6], where the authors construct a framework that measures the usefulness of technical indicators such as the RSI, EMAs, and MACD for predicting movements in the S&P 500 and other indices. This underscores the substantial need for further optimization of technical prediction models to enhance their reliability, which is the central motivation of the present work.

Absence of Reliability Weighting: While many studies leverage pattern-based features to capture sequential price dynamics, very few integrate a dynamic measure of each pattern's historical effectiveness. Without a confidence or volume weighting mechanism, low-probability signals pollute the model input and erode trading profitability [7].

These shortcomings mean that many models are prone to overfitting on noisy data and fail to effectively capitalize on high-confidence trading opportunities. To bridge these gaps, this paper introduces the Return-Weighted Binary-Coded Pattern-Extractor (RWBCPE), a novel automated decision support system that standardizes pattern encoding and dynamically weighs each signal by its historical performance and acts as an input feature for selected predictive models.

RWBCPE outperforms conventional decision support tools through two key innovations: Firstly, it encodes sequences of past candle states into fixed-length binary patterns, standardizing feature representation and constraining dimensionality. Secondly it augments each binary pattern with a volume estimate derived from its historical return distribution, thereby embedding a reliability weight that filters out low-success signals.

By focusing on concise pattern codes instead of an open-ended list of indicators and by coupling them with return-based volume weights, RWBCPE simultaneously mitigates overfitting, tames the curse of dimensionality, and enhances both classification accuracy and financial returns, providing a robust automated decision support tool for Forex prediction tasks.

1.3. Main Contributions

This research advances Forex forecasting through four key innovations:

Return-Weighted Binary-Coded Pattern-Extractor (RWBCPE): a novel decision support system that encodes sequences of candle states, uptrend (1) and downtrend (0), over the previous L periods into an L-bit binary string. By formalizing core price-action principles and leveraging the hypothesis that "history tends to repeat itself", and weighting patterns by historical recurrence probability, RWBCPE enhances model predictive power.

Dynamic Volume Sizing: RWBCPE pairs each Buy/Sell signal with a "dynamic volume" estimation computed from the historical mean return of the corresponding binary patterns, so that position size reflects prediction confidence.

Confidence-Driven Signal Filtering: A dual-threshold "Hold" state suppressing trades when pattern reliability falls below a cost-adjusted threshold or when the sign of the pattern's historical mean return (μ_{R_K}) contradicts the sign of its historical trend bias (Δ_{Trend_K}). This dual filter reduces False Positives and False Negatives of the applied algorithm and respects transaction costs.

Multi-Objective Learning Framework: Simultaneous optimization of financial performance (net profit, account balance, drawdown, …) and classification metrics (accuracy, recall, F1-score, …), reconciling statistical accuracy with trading viability and framing Forex prediction as a true multi-objective problem.

1.4. Paper Organization

The rest of the paper is structured as follows: Section 2 reviews related literature on technical-analysis-driven models for Forex forecasting. Section 3 explains the details of experimental setup GUI, dataset acquisition, description and preprocessing summary and its partitioning and validation framework. It details the architectural framework of the Return-Weighted Binary-Coded Pattern-Extractor (RWBCPE) and trading workflow. Section 4 presents empirical results, including performance benchmarks against baseline methods and robustness analysis. Section 5 synthesizes these findings, discusses practical implications, examines limitations, and proposes future research directions. Section 6 concludes the paper with key takeaways and a summary of contributions.

2. Background and related works

Accurate financial market models are essential for investors, as they facilitate data-driven decision-making, reduce investment risks, and assist in identifying profitable stocks. The development of advanced models also allows for the incorporation of non-traditional data sources, such as historical stock prices and news, enhancing the predictive capabilities of these models [8, 9]. Research on financial market predictions has increasingly utilized ML and DL techniques, driven by advancements in big data and AI. Despite these efforts, accurately predicting price directions remains challenging due to the complexity, nonlinearity, and inherent irrationality of stock prices [9]. This section presents an extensive synthesis of prior literature on Forex market prediction using technical analysis-driven ML/DL algorithms. It outlines the strengths and limitations of existing methods while identifying the research gap this study aims to fill. Recent research emphasizes the importance of indicator selection in technical analysis for Forex forecasting.

The Forex market's profitability has drawn significant interest from institutions and researchers [10]. In financial prediction, data mining methods are often limited when handling high-dimensional datasets [7]. Techniques such as PCA, FRPCA, and KPCA are effective in restructuring such data. Reference [6] tackled the heterogeneity in stock distributions using multi-output transfer learning to forecast opening, closing, high, and low prices across multiple stocks. Reference [11] proposed a fundamental-based model integrating ML techniques to forecast the Canadian–U.S. exchange rate while ensuring model interpretability. Among the most accurate for one-day-ahead predictions on major pairs is the 2D-CNN-LSTM architecture put forth by [4] that employs lag features to encode deep learning-based trend information. Reference [4] also highlights the importance of data-preprocessing on their use of a ConvLSTM2D model that simultaneously forecasts multiple currency pairs to capture their spatiotemporal relationships. Reference [12] compared several ML algorithms (e.g. ANN, XGBoost, Decision Tree, SVR, RF, LSTM, and GRU) using multi-asset and ensemble training against the Black-Scholes model, finding ML approaches significantly more accurate. Reference [13] built a fusion model combining STOA, VMD, SVM, and BPNN for decomposing and predicting volatile financial data. Similarly, [14] tested various high-frequency trading models, concluding that FCNN and logistic regression enhanced trading reliability. Reference [15] introduced a hybrid gold-price forecasting model using CEEMDAN, VMD, WOA, and XGBoost, outperforming other methods in accuracy. Reference [16] presented the AHC algorithm, a bio-inspired ML technique based on AON. It surpasses the AHN and traditional ARIMA in market index prediction. Reference [17] focused on small-scale trader indicators and applied fuzzy rough sets to classify Forex trends beyond specific thresholds. Reference [18] modelled cryptocurrency markets via Deep Deterministic Policy Gradient, producing agents that better replicate market behaviour. Reference [19] proposed MMDPS, an ensemble system integrating fuzzy logic, chaos theory, DL, and wavelet denoising to improve exchange rate predictions across 15 datasets. Reference [20] used Q-learning to train trading agents that outperformed traditional strategies. [21] developed real-time decision-making agents that adapt strategies to market conditions for Forex trading. [22] combined ECNN, DAE, and LSTM in the SA-DLSTM model to incorporate sentiment analysis into stock prediction. [23] provides a rigorous agent-based version of Paul de Grauwe's chaotic exchange-rate model, showing how the interplay of technical and fundamental traders can replicate stylized facts such as bubbles, crashes, and volatility clustering. [24] created MAC, a classifier integrating numerical and news sentiment features, aligning media signals with stock movements. Reference [25] blended GNG and Triple Q-learning to develop dynamic agents for stock prediction using unsupervised learning. Reference [26] used SSA, PSO, and LSTM in a hybrid model, achieving strong accuracy during high volatility. [27] proposed MADDQN, a dual-agent reinforcement system optimizing risk-reward in trading. [10, 28, 29] emphasized DL's superiority over traditional models in stock forecasting.

[8, 9] highlighted the relevance of non-traditional data sources like news in enhancing market prediction models. [30] introduced the ESI model to simulate policy interventions in Forex and examine noise traders' effects on stability. [31] demonstrated how RL ensembles outperform Buy-and-Hold strategies in intraday trading. [32] completed a systematic multi-indicator analysis and noted that returns did not exceed market benchmarks; however, the overall results and forecasting precision were better when combining the RSI, EMAs, and MACD. Reference [9] optimized feature selection using GA-XGBoost, stressing the need for balanced feature engineering. Reference [28] developed NPMM labelling to manage small price change sensitivity and trained an XGBoost system for automated trading. Reference [33] simulated virtual markets with GAN-based agent models. Reference [34] integrated multimodal RL and sentiment analysis, showing how news dynamically affects financial outcomes. Reference [35] employed fuzzy candlestick pattern recognition for safer, more profitable trading. Reference [36] combined ensemble SVM and fuzzy NSGA-II to forecast EUR/USD trends using ML and DL. They optimized fuzzy trading rules using NSGA-II and demonstrated gains in trend accuracy. Reference [5] built a fuzzy multicriteria indicator fusion system for decision-making. Reference [1] recognized the complexity of market prediction and called for more effective ML models. For EUR/USD forecasting, Reference [37] proposed a hybrid model integrating Empirical Mode Decomposition, a stacked LSTM, and Particle Swarm Optimization. This framework outperformed benchmarks like XGBoost in profitability while minimizing trading risk. Reference [7] extracted new datasets using PCA-family methods for S&P 500 prediction. Reference [38] trained ANN models to recognize pattern curves from Forex time series. Reference [39] introduced RSLSTM-A, blending ResNet and LSTM for asset trading. Reference [40] applied the ARO algorithm to optimize hyper-parameters for deep LSTM models. [41] combined FRS, NSGA-II, and SW to predict crude oil futures. Reference [42] benchmarked ten ML methods, highlighting PKDE's effectiveness in predicting currency swings. Reference [43] demonstrated the effectiveness of combining traditional technical indicators with alternative data sources such as social sentiment and popularity metrics, using XGBoost and LightGBM to classify market trend direction across hundreds of securities. The study on high-resolution changes between the Korean won and the US dollar has been published by [44] and assesses the accuracy of Gumbel's nonparametric technique and periodicity filters in calculating alter jump probabilities based on the non-normative statistical methods and highlights the consequences of disregarding intraday periodicity in jump assessments [44]. Outside Forex, [45] introduce a neural network based Volume-Up (VU) strategy that forecasts abnormal events on the S&P 500 index by manipulating the volatility distribution to enhance prediction accuracy. Finally, a two-stage forecasting strategy integrating Bidirectional LSTM with a multi-head attention transformer network for price prediction and trend classification, respectively, and fine-tuned with the Flexible Fitness-Dependent Optimizer (FFDO) algorithm, has been suggested by [46].

3. Experimental Setup and Data Preprocessing

3.1. *Implementation Tools*

This study employed a multi-faceted approach for data extraction, processing, and modelling. A Python-based Audio-enhanced Graphical User Interface was developed to streamline experimentation and real-time evaluation. **Figure 1** shows the main interface of the Audio-enhanced GUI. The entire framework, including the GUI, was implemented using Python 3, leveraging a suite of open-source libraries for data handling, modelling, and visualization.

Figure 1. Main interface of the Audio-enhanced GUI of the experimental setup used for research navigation, evaluation, and result demonstration.

3.2. Data Acquisition and Description

This study employs historical OHLC (Open/High/Low/Close) data for some major currency pairs. Data were sourced from institutional APIs (MetaTrader 5) and public repositories (Yahoo Finance), ensuring both depth and liquidity coverage. The EUR/USD, GBP/USD, and USD/JPY currency pairs have been chosen for their significant importance in the global foreign exchange market, distinct volatility profiles, and abundant liquidity. Yahoo Finance offers extensive historical and current multi-asset data, whereas MetaTrader 5 serves as a sophisticated algorithmic trading platform with a rich array of financial data. A 25-year horizon (2000-2025) for the trainset captures diverse macroeconomic regimes and market conditions, providing a comprehensive basis for evaluating the proposed framework.

3.3. Dataset Preprocessing Summary

Preprocessing includes data cleaning, anomaly detection, imputation, denoising, scaling, and structuring time series for supervised learning suitable for ML/DL model training. To maintain the integrity of the data and its usability for the model, a multi-stage preprocessing pipeline was set.

In the provided dataset, inconsistencies like missing values, duplicates, and erroneous ticks were identified. Because of the liquidity and public domain nature of the forex data, the null values were negligible. The 25-year dataset was then subjected to statistical anomaly detection to identify incomplete/outlier records (e.g., holiday gaps, erroneous feeds), with contextually justified imputation or exclusion. To mitigate the impact of extreme short-term fluctuations, due to fundamental events, the procedure of truncating the scaled returns was applied to the price series (Eq. (3)). The returns of successive candles, obtained by differentiating consecutive close prices (Eq. (1)) are scaled in the trainset by dividing them by maximum value of the rolling window (Eq. (2)).

$$R_t = C_t - C_{t-1} \tag{1}$$

$$SR_t = \frac{R_t}{max(|R_{\acute{t}}|)} \quad \forall t \mid t \in \{1,2,\dots,N_{tr}\} \ \wedge \ \forall \acute{t} \mid \acute{t} \in [-len(Trainset), t-1] \tag{2}$$

These scaled returns were then bounded within ±5 times standard deviation (σ) of the series to stabilize the data distribution and to not overly affect the algorithm (Eq. (3)). This truncation is not a preprocessing step but is appropriate for evaluating the proposed trading system using the RWBCPE to temper the impact of exceptional conditions on other states.

$$SR_t^{trunc} = \begin{cases} SR_t & if \quad -5\sigma \le SR_t \le +5\sigma \\ -5\sigma & if \quad SR_t \le -5\sigma \\ +5\sigma & if \quad SR_t \ge +5\sigma \end{cases} \quad (3)$$

where σ is the standard deviation of the scaled returns. The scaling factor ($\max(|R_t|)$) is calculated separately for each trainset ($\mathcal{D}_{\text{train}}^{(k)}$) to avoid look-ahead bias.

3.4. *Dynamic Single Test Splitting (DSTS) protocol*

This research provides more methodologically sound and temporally unbiased validation than standard rolling-window techniques. The Dynamic Single Test Splitting (DSTS) protocol is in fact a rolling-origin evaluation with single-step forecasting as depicted in **Figure 2**.

It decomposes the test range of size N_{ts} into N_{ts} distinct partitions, each containing a test set consisting of precisely one subsequent candle. No candle appears in more than one test set so the protocol generates N_{ts} partitions, each with a mutually exclusive test candle as expressed in Eq. (4). Because a window of size 1 is used as the test set, N_{ts} iterations is required for proper validation across the entire test set.

$$\bigcup_{i=1}^{N_{ts}} \mathcal{D}_{\text{test}}^{(i)} = \mathcal{D}_{\text{test}}^{\text{full}}, \ \left|\mathcal{D}_{\text{test}}^{(i)}\right| = 1 \Longrightarrow \left|\mathcal{D}_{\text{test}}^{\text{full}}\right| = N_{ts} \quad (4)$$

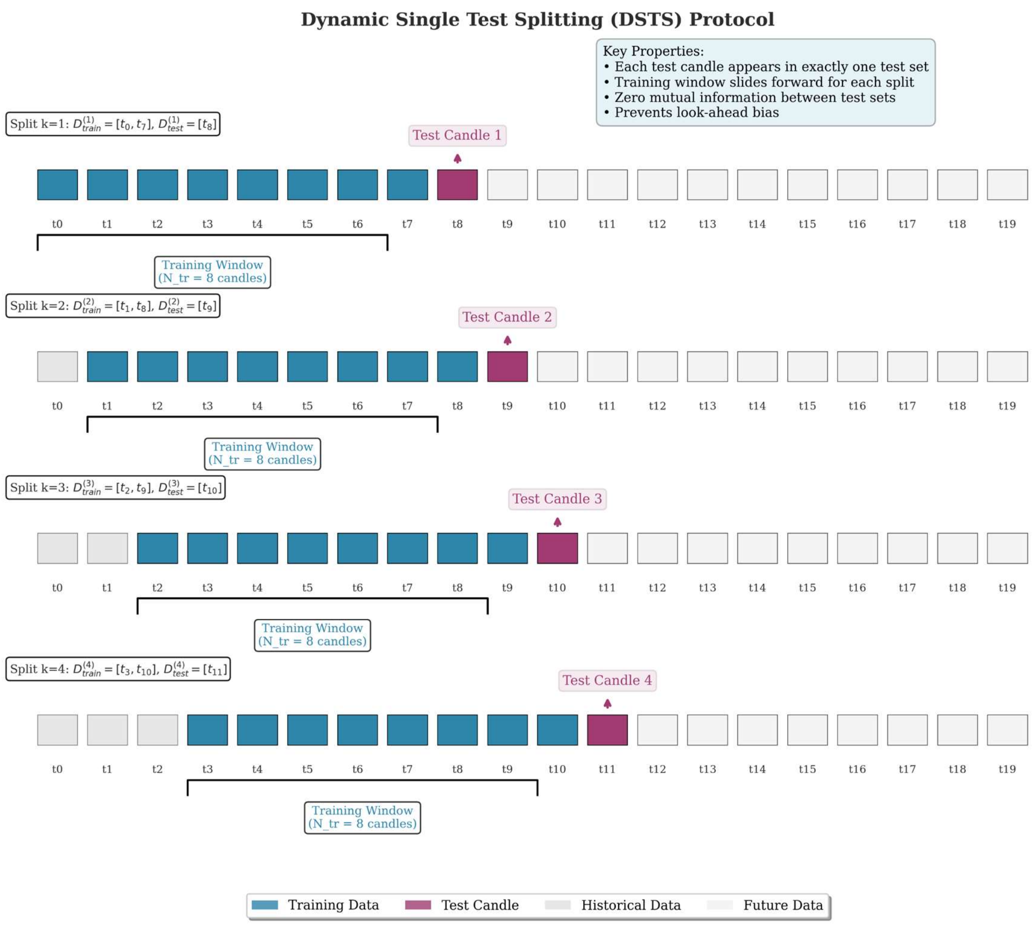

Figure 2. The Dynamic Single Test Splitting (DSTS) as a rolling train-test splitting.

A look-back window of consecutive time steps was allocated for training of size N_{tr}. This allocates the model's attention to the immediate test candles context, the precise candles just before the creation of the test candle, only when forecasting a single step ahead. For every test instance $k \in \{1,2,\dots,N_{ts}\}$, the DSTS protocol thoroughly defines clearly separated partitions (Eq. (5)).

$$\begin{aligned} \mathcal{D}_{\text{train}}^{(k)} &= \{(x_t, y_t) \ : \ t \in [t_k - N_{tr}, t_k - 1]\} \\ \mathcal{D}_{\text{test}}^{(k)} &= \{(x_{t_k}, y_{t_k})\} \qquad \forall k \mid k \in \{1,2,\dots,N_{ts}\} \end{aligned} \quad (5)$$

where t_k denotes the temporal index of the K-th test candle. The temporal isolation of each test instance allows DSTS to enhance model sensitivity to changing market conditions without look-ahead bias. This also controls overfitting due to effective sample isolation and ensures zero mutual information (Eq. (6)), zero test-set overlap (Eq. (7)), and antecedent-only training (Eq. (8)).

$$I\left(\mathcal{D}_{\text{test}}^{(i)}; \mathcal{D}_{\text{test}}^{(j)}\right) = 0 \qquad \forall i \neq j \tag{6}$$

$$\mathcal{D}_{\text{test}}^{(i)} \cap \mathcal{D}_{\text{test}}^{(j)} = \emptyset \qquad \forall i \neq j \tag{7}$$

$$\max\left(\mathcal{D}_{\text{train}}^{(k)}\right) < \min\left(\mathcal{D}_{\text{test}}^{(k)}\right) \tag{8}$$

In order to minimize data spillover, during the DSTS validation, test candles are appropriately removed from HPADs. In DSTS, for each test candle at time t, Eq. (8) states that HPADs would be built only using data till t-1, so the integration of DSTS with per-split HPAD reconstruction prevents feature leakage.

3.5. Synergistic Benefits of RWBCPE for ML-Based Trading

Algorithmic trading strategies typically fall into two categories: rule-based systems and machine learning (ML)-driven approaches. Rule-based trading relies on predefined heuristics (e.g. price action rules and candlestick formations) crafted by domain experts. In contrast, ML-based trading systems learn directly from historical data, enabling autonomous decision-making without explicit human-defined rules. Although this study does not directly implement traditional rule-based methods, the proposed Return-Weighted Binary-Coded Pattern-Extractor (RWBCPE) framework serves as a bridge between the two paradigms. By encoding historical candlestick patterns and weighting them based on their return profiles, RWBCPE effectively captures the essence of rule-based insights while embedding them into a data-driven ML framework. This integration allows ML models to benefit from the interpretability and domain relevance of rule-based features, while retaining the flexibility and adaptability of learning algorithms.

Leveraging RWBCPE enables machine/deep learning models to identify recurring patterns and their predictive significance, thereby enhancing both the accuracy and profitability of trading decisions. Moreover, this approach facilitates the discovery of latent relationships and profitable structures that may elude human analysts, offering a hybrid methodology that combines the strengths of expert-driven and data-driven systems in the context of Forex market prediction. Comparative analysis against conventional technical features demonstrates its superiority in enhancing predictive accuracy and financial profitability.

3.6. RWBCPE: Rationale and Architecture

The Return-Weighted Binary-Coded Pattern-Extractor (RWBCPE) is designed to algorithmically recognize and encode recurrent price action motifs in candle-based time series data, applying real-time weighting based on a dynamic confidence measure derived from its historical return profile. Its dual-module architecture balances the information-richness of pattern recognition with adaptation to market microstructure and shifting liquidity.

The RWBCPE employs the well-known axiom: "history tends to repeat itself." it captures recurrent candlestick configurations as binary-encoded sequences. By transforming successive candle states into patterns of length L, the model leverages past occurrences as a probability measure to forecast future trends. Each pattern's predictive potency is weighted by its historical bias and frequency, enabling the algorithm to distinguish robust signals from noise. This mechanism not only substitutes for handcrafted rule-based heuristics such as classic candlestick formations but also uncovers novel relationships hidden in high-dimensional historical data. Integrating this framework into ML/DL pipelines enhances both classification accuracy and trading profitability by dynamically adjusting reliance on patterns with proven performance.

3.7. Framework Overview

The experimental workflow encompasses dataset preparation and cleaning, novel indicator design, and details its rationale and structure, including its integral Pattern Extractor Module and Volume Estimator Module, and explains how they feed into signal filtering, money management, and machine-learning pipelines (**Figure 3**). It uses a Dynamic Single Test Splitting protocol and proposes a dynamic time-series validation, and adaptive trading components, designed for rigor and reproducibility in computational finance. It explains the signal filtering method and discusses the different states of issuing Hold signal and a dynamically adaptive money management system for position sizing based on drawdown exposure. The Return-Weighted Binary-Coded Pattern-Extractor (RWBCPE) encapsulates three core computational principles that underlie the framework's design:

Return-Weighted Component: This aspect dynamically modulates trading decisions by incorporating the historical performance of recurring candlestick patterns. Specifically, it utilizes the mean return (μ_{R_K}) of all prior instances matching the current pattern, where return is defined in Eq. (1). After formation of the recent pattern k, the indicator makes adjustments for the active trading position according to the historical occurrence mean return (μ_{R_K}). This weighting mechanism prioritizes patterns with demonstrated predictive power, formulated through the Confidence-Weighted Return ($CWR(t)$; Eq.(22)) and volume estimation ($\hat{V}_t$; Eq.(24)-(25)).

Binary-Coded Component: Price action is discretized into a machine-interpretable representation by encoding each candle's directional movement into binary states (Eq. (9), Eq. (10)). For a window of L consecutive candles, the sequence of states $[S_{t-L}S_{t-L+1}\cdots S_{t-2}S_{t-1}]$ is mapped to a unique integer value $L_{BCP_K}(t)$ (Eq. (12)). This binarization compresses continuous OHLC data into a maximum finite alphabet of 2^L pattern classes, resulting in efficient pattern matching while filtering market noise.

Pattern-Extractor Component: Historical Pattern Aggregation Databases (HPADs) systematically catalogue recurrence statistics for every possible L-bit patterns ($L \in [1, M]$). Each L_HPAD quantifies pattern frequency (f_{BCP_K};Eq. (15)), directional bias (Δ_{Trend_K}; Eq. (16)), and return characteristics (μ_{R_K};Eq. (19)) across decades of market data. This transforms raw price series into a searchable repository of technical patterns, replacing heuristic candlestick recognition with data-driven pattern retrieval and enabling conventional ML/DL models to access these newly defined databases as a source of their input features.

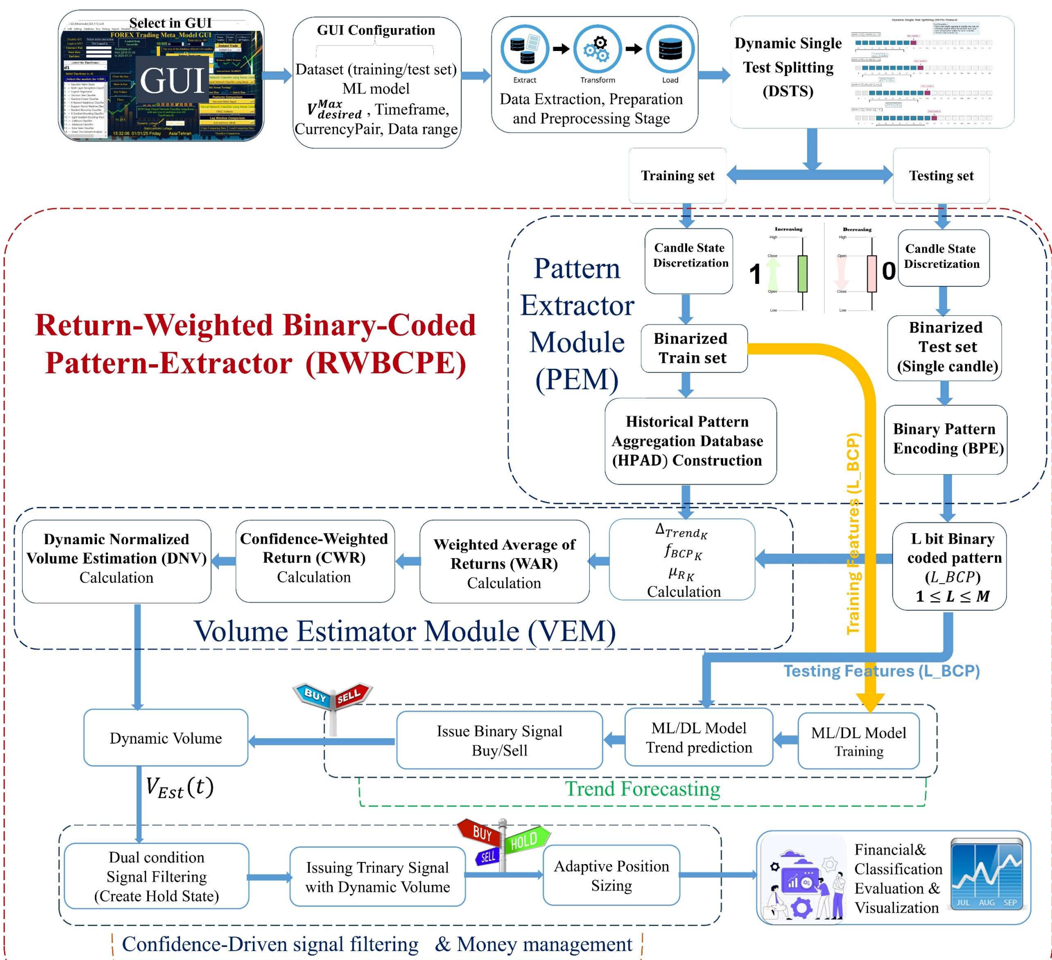

Figure 3. A comprehensive block diagram of the proposed RWBCPE framework, illustrating the data flow through the Pattern Extractor Module (PEM) and Volume Estimator Module (VEM).

This tripartite architecture weights decisions by historical returns, encodes market structure into binary sequences and extracts patterns from aggregated databases, accordingly establishes RWBCPE as a rigorous feature engineering framework for financial forecasting. Its integration into ML pipelines substitutes rule-based heuristics with statistically grounded signals while uncovering latent pattern-return relationships inaccessible to manual analysis. The RWBCPE architecture systematically transforms raw price movements into quantifiable trading signals through two integrated modules each of which comprises three main parts:

Module 1: PEM (Pattern Extractor Module) that converts sequential price actions into machine-interpretable features. This module comprises "*Candle State discretization* (binary classification)", "*Binary pattern encoding* (L-bit strings)" and "*Statistical aggregation* (HPAD databases)" parts.

Module 2: VEM (Volume Estimator Module) that dynamically estimates proper position sizes based on pattern confidence. It comprises "*WAR (weighted Average of Returns)*", "*CWR (Confidence-Weighted Return)*" and "*DNV (Dynamic Normalized Volume) Estimation*" parts.

This dual-module design addresses the challenge of extracting meaningful signals from high-noise Forex environments by compressing candle sequences into compact representations that preserve directional information while filtering non-predictive noise. The next subsections detail each step's implementation as well as the mathematical representations.

3.8. *The Pattern Extractor Module (PEM)*

The Pattern Extractor Module was created to automate pattern extraction. In the context of financially noisy time series, particularly FOREX, the objective is to transform local sequences of candles into compact representations that filter out non-predictive noise while retaining directional and structural details. PEM accomplishes this by converting sequences of candle stick price movements into a binary pattern, changing variable length OHLC data into a fixed length, history-rich feature vector that meets the requirements for ML input.

3.8.1. *Candle State Discretization (CSD)*

The directional movement of each candle is discretized into a binary state, denoted as S_t, using one of two methods: either successive closing price differentials (("Rise" ≡ 1 or "Fall" ≡ 0); Eq. (9)) or intrinsic open/close relationships (("Bullish" ≡ 1 or "Bearish" ≡ 0); Eq. (10)).

$$S_t = Rise_Fall = \begin{cases} 1 & if \quad C_t > C_{t-1}\,(Rise) \\ 0 & if \quad C_t \le C_{t-1}\,(Fall) \end{cases} \tag{9}$$

Where S_t denotes the binarized state at time t and C_t is the closing price at time t.

$$S_t = \text{Market} = \begin{cases} 1 & if \quad C_t > O_t\,(Bullish) \\ 0 & if \quad C_t \le O_t\,(Bearish) \end{cases} \tag{10}$$

Where O_t is the opening price at time t. In the "Rise_Fall" state (Eq. (9)), the focus is on analyzing closing prices in succession. In the "Market" state (Eq. (10)), one assesses momentum for the intra-candle period. This discretization converts continuous price series into binary sequences, forming the basis for pattern identification.

3.8.2. *Binary Pattern Encoding (BPE)*

Binary Pattern Encoding (BPE) is concerned with defining recurring patterns as sequences made up of previous candle states. A pattern of length L at time t consists of states in a sequence arranged in a chronological order as left-to-right (oldest on the left) and S_{t-1} signifies the candle right before the one in focus (Eq. (11)).

$$P_t = [S_{t-L} S_{t-L+1} \cdots S_{t-2} S_{t-1}] \qquad S_i \in \{0\,,1\} \qquad 1 \le L \le M \tag{11}$$

Where S_i represents the binary (0 or 1) state of the candle i. With each candle having a binary state, there are various sequences that can be generated. To turn these sequences into machine-learnable numerical characteristics, each sequence can be encoded into a unique integer named here the "L-bit Binary Coded Pattern", or L_{BCP_K} . It regards P_t as a binary number, where the most recent state (S_{t-1}) is the least significant bit (LSB), and the oldest state (S_{t-L}) is the most significant bit (MSB). The L_{BCP_K} value at time t ($L_{BCP_K}(t)$) is found in Eq. (12).

$$L_{BCP_K}(t) = \sum_{i=1}^{L} S_{(t-i)} \times 2^{i-1} \quad S_{t-i} \in \{0\,,1\} \forall L \in [1\,,M] \tag{12}$$

Where S_t is defined in Eq. (9) or Eq. (10) and M is the optimal maximum pattern length that can differ for different currency pairs and timeframes. The *L* in $L_{BCP_K}(t)$ varies from *L*=1 to *M* and generates a set of these features for various pattern lengths. The structure of L-bit strings (L_BCP) is explained in the following:

- L in L_BCP is the number of previous bits employed in the encoding procedure that can create up to 2^L states.
- LSB (the 0th order bit): Represents the most recent previous candle's Rise_Fall state.
- 1st order bit: Represents the penultimate candle's Rise/Fall state.
- This sequence continues to the MSB ((L-1)th order bit).

Each L_BCP can encode up to 2^L potential patterns comprising previous L states. It is evident that a 1-bit BCP (1_BCP) can hold just one bit of historical data and an L_BCP can hold L bits of it. So the number of bits in the string is crucial to retain as many effective candlestick patterns as possible in any market like Forex. Moreover, the number of all BCPs should be chosen with care to avoid the "curse of dimensionality" and overfitting and on the other hand to hold and include as many valuable and effective patterns as it can. These L_BCPs are part of the RWBCPE indicators that are to be fed to the selected model(s) as input features. To select the most suited numbers of indicators and also the maximum length (*M*) for selected currency-pairs and timeframes a validation range is used in which different sets of RWBCPE indicators are evaluated and the set with the best financial and classification results is selected.

3.8.3. *Historical Pattern Aggregation Database (HPAD)*

For each target currency-pair and timeframe, M databases are constructed using a trainset out of a large historical dataset available. These databases, termed "Historical Pattern Aggregation Databases" (HPAD), catalogue aggregated statistics for every possible L_BCP value for a given pattern with length *L*. These databases transform raw price series into quarriable knowledge bases, replacing heuristic pattern recognition with statistical aggregation. For example, for daily timeframe and considering the trainset in the 2000-01-01 to 2025-01-01 range (and the test-set in 2025-01-02 to 2026-01-01 range) there will be 6517 daily candles available for constructing the HPAD databases (active days of market for 25 years). Each *L*-HPAD has a tabular 2-D structure that is composed of up to 2^L rows, corresponding to each unique pattern (if available), and the following columns respectively:

L_BCP, that serves also as the index of each dataframe (Eq. (12)).

The total number of rising candles after a specific L_{BCP_K} pattern ($\mathrm{N}_{\mathrm{Rise}_K}$;Eq. (13)).

The total number of falling candles after a specific L_{BCP_K} pattern (N_{Fall_K} ;Eq. (14)).

Pattern frequency (f_{BCP_K}; Eq. (15))

Trend bias (Δ_{Trend_K}; Eq. (16))

Mean_Delta (μ_Δ; Eq. (17))

Sum_Return (σ_{R_K}; Eq. (18))

Mean_Return (μ_{R_K}; Eq. (19))

$\mathrm{N}_{\mathrm{Rise}_K}$: The total number of instances where the subsequent candle state is a "Rise" (S=1) after a specific L_{BCP_K} pattern occurred in all the trainset (Eq. (13)).

$$N_{Rise_K} = \sum_{i=1}^{len(DB)} (S_i) | L\_BCP_i = L_BCP_K \quad , \quad \forall\, L \in [1, M] \;\; \& \;\; \forall\, K \mid L\_BCP_K \in L\_HPAD \tag{13}$$

N_{Fall_K}: The total number of instances where the subsequent candle state is a "Fall" (S=0) after a specific L_{BCP_K} pattern occurred in all the trainset range (Eq.(14)):

$$N_{Fall_K} = \sum_{i=1}^{len(DB)} (\bar{S_i}) | L\_BCP_i = L_BCP_K \equiv \sum_{i=1}^{len(DB)} Not(S_i) | L\_BCP_i = L_BCP_K \tag{14}$$

Where:

- DB refers to the specific L_HPAD database for a given pattern length L, which is constructed exclusively from the training data within the current DSTS split to prevent look-ahead bias.

- i refers to the index of record in L_HPAD.

f_{BCP_K}: The total number of occurrences of a specific pattern L_BCP_K, that is the sum of N_{Rise_K} and N_{Fall_K} related to an L bit coded pattern as expressed in Eq. (15).

$$f_{BCP_K} = \mathrm{count}(BCP_K) = N_{Rise_K} + N_{Fall_K} = \sum_{i=1}^{len(Trainset)} (not(S_i) + S_i) | L_BCP_i = L_BCP_K) = \sum_{i=1}^{len(Trainset)} (\bar{S_i} + S_i) | L_BCP_i = L_BCP_K) \tag{15}$$

Δ_{Trend}: The net trend bias of the pattern BCP_K as expressed in Eq. (16).

$$\Delta_{Trend_K} = N_{Rise_K} - N_{Fall_K} \tag{16}$$

μ_Δ(Mean_Delta): The ratio of the trend bias of the pattern BCP_K, to the total number of its occurrences in the L_HPAD as expressed in Eq. (17).

$$\mu_\Delta = \frac{\Delta_{Trend}}{f_{BCP_K}} = \frac{N_{Rise_K} - N_{Fall_K}}{N_{Rise_K} + N_{Fall_K}} \tag{17}$$

R_t(Return): The difference of "Close" of two consecutive candles as expressed in Eq. (1).

σ_R(sum_Return_ K): Sum of Returns of all occurrences of a candle forming a specific pattern L_{BCP_K} in the whole trainset; Eq. (18).

$$\sigma_{R_K} = \sum_{t=1}^{len(Trainset)} (C_t - C_{t-1}))|L\_BCP_i = L_BCP_K \quad (18)$$

μ_{R_K} (Mean_Return_K): Average of "Returns" of all occurrences of a candle forming a specific pattern L_{BCP_K} in the whole trainset calculated by the sum of returns(σ_R) divided by the total number of L_BCP_K in the L_HPAD (f_{BCP_K}) as expressed in Eq. (19).

$$\mu_{R_K} = \frac{\sigma_R}{f_{BCP_K}} = \frac{\sum_{t=1}^{len(Trainset)}(C_t - C_{t-1})) \mid L\_BCP_i = BCP_K}{N_{Rise_K} + N_{Fall_K}} \quad (19)$$

While the offline construction of HPADs provides significant predictive power, it introduces a computational overhead that warrants consideration. For a training set of length (N_{tr}) and a maximum pattern length (M), the time complexity for database construction is ($\mathcal{O}(N_{tr} \times M)$), which for our primary configuration (($N_{tr} = 6517$), ($M = 11$)) requires approximately (7.17×10^4) operations per DSTS split that is in fact a negligible cost on modern hardware. However, the memory footprint scales with (2^L), making high-frequency data or large (M) values impractical. To mitigate this, we employ three implementation strategies: (1) Vectorized array operations for state transition counting, (2) Persistent caching of HPADs per currency-pair and timeframe to avoid redundant computation across DSTS splits, and (3) Pruning of pattern classes with zero occurrences. For real-time deployment, the online prediction phase is reduced to ($\mathcal{O}(1)$) lookups, ensuring that the framework remains viable for low-latency algorithmic trading.

3.8.4. *Constructing L_HPADs & Interrelation with RWBCPE*

The starting point to create the RWBCPE indicator and to estimate the embedded volume is the construction of L_HPADs from historical data for a specific currency pair and timeframe. These databases are saved for subsequent use. The maximum value that is used to make databases is referred here and in equations as M that is a critical point. It is limited by the total number of records in the trainset. It's evident that selecting a large number M may decrease the number of its occurrences in the L_HPAD and increase the time complexity and the probability of overfitting. For example, for daily timeframe the trainset is selected from 2000-01-01 to 2025-01-01 that results in 6517 candles. So selecting an $M > 13$ is not rational because even for L=13 there are $2^{13} = 8192$ patterns, which is more than the number of unique patterns possible, and this will definitely lead to overfitting because of few number of samples for each pattern.

In each L_HPAD L stands for the number of previously encoded states (Rise_Fall) of historical data in that database (e.g. for M=12 there are 12 L_HPADs namely 1_HPAD, 2_HPAD,..., 12_HPAD). Utilizing L_BCPs as input Features: The L_BCP values, for $1 \leq L \leq M$, will be fed to the selected ML/DL models as input features to initiate the prediction process. The selected models can be trained on the entire range of these features or a selected subset.

By integrating Candle State Discretization (CSD), Binary Pattern Encoding (BPE), and HPADs construction processes, PEM processes raw candlestick sequences and creates binary encoded features with added statistical features. The outputs form the foundational elements of the RWBCPE indicator and act as inputs to the subsequent Volume Estimator Module (VEM) and the final feature vector construction stage for the downstream ML/DL models.

3.9. *The Volume Estimator Module (VEM)*

This framework stands out because it can adjust the trade volume depending on the historical confidence level of a detected pattern, allowing the elimination of low probability signals. In this research the estimated volume, $\hat{V}_t$, is calculated through a new proposed approach where a pattern's number of occurrences in database is combined with its average historical return.

3.9.1. *Weighted Average of Returns (WAR)*

First, the Weighted Average of Returns ($Weighted_{\mu_{R_K}}$) is calculated. The weight for each pattern is a function of its historical trend bias(Δ_{Trend_K}) and frequency (f_{BCP_K}), designed to evaluate consistency and significance. $Weighted_{\mu_{R_K}}$ is the Weighted Average of Returns of the candle forming the last bit of BCP_K pattern in the L_HPAD by scaling the μ_{R_K} with the harmonic averaging of Δ_{Trend_K} and f_{BCP_K} of a specific pattern; Eq. (20).

$$WAR = \mu_{R_K} \times \left(\frac{1}{\frac{1}{2} \times \left(\frac{1}{\Delta_{Trend_K}} + \frac{1}{f_{BCP_K}} \right)} \right) = \mu_{R_K} \times \frac{2 \times \Delta_{Trend_K} \times f_{BCP_K}}{\Delta_{Trend_K} + f_{BCP_K}} \quad (20)$$

This term gives higher weight to patterns that have both a strong historical directional bias and have occurred frequently. Merging Eq. (19) and Eq. (20) will result in Eq. (21).

$$WAR = \mu_{R_K} \times \frac{2 \times \Delta_{Trend_K} \times f_{BCP_K}}{\Delta_{Trend_K} + f_{BCP_K}} = \frac{2 \times \Delta_{Trend_K} \times \sigma_R}{\Delta_{Trend_K} + f_{BCP_K}} \quad (21)$$

3.10. Confidence-Weighted Return (CWR)

The Confidence-Weighted Return (CWR) is calculated for the set of patterns observed at time *t*. The CWR(t) is the sum of the rectified WAR (i.e. $Relu\left(Weighted_{\mu_{R_K}}\right)$) for each pattern length from L=1 to L=M that includes all L-bit patterns formed within the window of size L before the intended candle. The Confidence-Weighted Return at time t ($CWR(t)$) is calculated in Eq. (22):

$$CWR(t) = \sum_{K=K_1}^{K_L} ReLU(WAR) = \sum_{K=K_1}^{K_L} ReLU\left(\frac{2 \times \Delta_{Trend_K} \times \sigma_R}{\Delta_{Trend_K} + f_{BCP_K}}\right) \quad (22)$$

3.10.1. Dynamic Normalized Volume Estimation

The trade volume estimation is done by taking the desired lot size ($V_{desired}^{Max}$) from the GUI and scaling it by the ratio of the current pattern's CWR to the average CWR of the whole database (μ_{CWR}^{Total}). The equations for μ_{CWR}^{Total} and $\hat{V}_t$ are explained in Eq. (23) and Eq. (24) respectively.

$$\mu_{CWR}^{Total} = \frac{1}{len(DB)} \times \sum_{i=1}^{len(DB)} CWR(t) \quad (23)$$

$$\hat{V}_t = \max\left(0, min\left(V_{desired}^{Max}, \frac{V_{desired}^{Max} \times CWR(t)}{\mu_{CWR}^{Total}}\right)\right) = \mathrm{ReLU}(V_{desired}^{Max} \times min\left(1, \frac{CWR(t)}{\mu_{CWR}^{Total}}\right) \quad (24)$$

Where:

- $\hat{V}_t$ is the estimated volume for opening or modifying active positions.
- $V_{desired}^{Max}$ is the maximum desired volume and is the maximum value for trading volume in lots that will be set through the designed GUI.
- $len(DB)$ is the length or total records of the database.
- μ_{CWR}^{Total} is the average CWR observed across the entire historical database.

The expression $\frac{CWR(t)}{\mu_{CWR}^{Total}}$ in Eq. (24) serves as a normalization factor that highlights the historical pattern strength at baseline. This enables the model to flexibly enlarge the position size for signals that show strong confidence and diminish it for the ones that are feeble. Regarding the Volume Estimation process, the algorithm performs the following steps:

- It looks up all L_BCP records matching the current candlestick's pattern (L_BCP_K) within all the previously built L_HPADs (for $1 \leq L \leq M$).
- It takes the maximum desired Volume ($V_{desired}^{Max}$) value from the GUI and retrieves the corresponding f_{BCP_K}, Δ_{Trend_K} and μ_{R_K} values for every L_BCP_K by looking up in each L_HPAD.
- It calculates WAR, CWR for the intended test candle and estimates each trade volume $\hat{V}_t$ according to Eq. (25).

$$\hat{V}_t = ReLU\left(V_{desired}^{Max} \times \mathrm{min}\left(1, \frac{\sum_{K=K_1}^{K_L} ReLU\left(\frac{2 \times \Delta_{Trend_K} \times \sigma_R}{\Delta_{Trend_K} + f_{BCP_K}}\right)}{\frac{\sum_{i=1}^{len(DB)} \sum_{L=1}^{M} Relu\left(\frac{2 \times \Delta_{Trend_K} \times \sigma_R}{\Delta_{Trend_K} + f_{BCP_K}}\right)_i}{len(DB)}}\right)\right) \quad (25)$$

Eq. (25) constitutes the VEM module (the Volume Estimator Module) of RWBCPE indicator for opening new positions in the market or modifying the previously opened active positions. As can be seen in Eq. (25) the estimated volume can only be a positive value because of the first ReLU function in it. The internal ReLU functions are used to filter out patterns when the signs of Δ_{Trend_K} and μ_{R_K} contradict each other (i.e. when one is positive and the other is negative), as such patterns are deemed untrustworthy and the ReLU function filters them to avoid affecting the other terms of the equation.

The term $\frac{2 \times \Delta_{Trend_K} \times f_{BCP_K}}{\Delta_{Trend_K} + f_{BCP_K}}$ or $\frac{2 \times \Delta_{Trend_K} \times \sigma_R}{\Delta_{Trend_K} + f_{BCP_K}}$ in Eq. (21) that seems in a part of Eq. (25) is the Weighted Average of Returns that gives higher weight to patterns that have both a strong historical directional bias and have occurred frequently. It gives a weight to $\mathrm{Return_Mean}(\mu_{R_K})$ and determines how its value can be effective according to the repetition of that pattern (f_{BCP_K}) and how much similarity have been between previous occurrences of that pattern historically in the trainset. i.e., the higher the frequency of a pattern, the higher the probability of its accuracy and the higher Δ_{Trend_K} shows greater robustness of its output.

So, to sum things up, the denominator of the fraction in Eq. (25) is the μ_{CWR}^{Total} that shows a comprehensive mean of weighted returns of all occurrences of patterns in the database. It acts as a normalization factor. The nominator represents the baseline historical pattern strength. This allows the model to dynamically increase position size for high-confidence signals and reduce it for low-confidence ones. Altogether, Eq. (25) is a normalized equation because $0 \leq \frac{\hat{V}_t}{V_{desired}^{Max}} \leq 1$. So there is a maximum value limit for the estimated volume that can be defined by the trader according to the risk management and the amount of investment in the market. The final estimated trade volume($\hat{V}_t$) is calculated by scaling a predefined maximum desired lot size ($V_{desired}^{Max}$) against the ratio of the current pattern strength $CWR(t)$ to the average CWR observed across the entire historical database (μ_{CWR}^{Total}). The combination of estimated volume with the L_BCP_K string, incorporating both frequency and strength, gives the name Return-Weighted Binary-Coded Pattern-Extractor (RWBCPE) to the indicator. This structured approach ensures a comprehensive and adaptive framework for financial forecasting, leveraging the dynamic nature of RWBCPE indicators.

3.10.2. VEM interrelation with RWBCPE and Signal Generation

The VEM generates the estimated trading volume ($\hat{V}_t$) through WAR, CWR, and DNV. Along with pattern-based features from the PEM, this output finalizes the RWBCPE indicator construction. The estimated signal strength and volume is sent to the next stage, Signal Filtering, and Money Management, where they become executable trading actions.

3.11. Feature Vector Generation

At each forecasting timestamp t, the feature vector X_t synthesizes multi-scale pattern information through concatenation of L_BCP encodings across all window lengths L∈[1,M]. The final vector (X_t) comprises a selected subset of these L_BCPs features as expressed in Eq. (26).

$$X_t = \{n_BCP : n \in [1\ ,\ M]\} \tag{26}$$

Each n_BCP element represents a binary-coded integer (Eq. (5)) encapsulating the sequential states of n antecedent candles. Dimensionality optimization is achieved via correlation-based feature selection, retaining only L_BCP components satisfying Eq. (27)

$$|\rho(i_BCP, j_BCP)| \leq \alpha \qquad \forall\, i\ ,\ j \in [1\ ,\ M] \quad \&\quad i \neq j \tag{27}$$

where ρ denotes point-biserial correlation and α was tuned equal to 0.85 by a trial-and-error evaluation process. For the daily timeframe and the 25-year trainset (M=11) hierarchical dependencies ensure that cumulative state references total equals 55 (Eq. (28)). This multi-resolution representation captures hierarchical market structure:

- Short-term elements ($L \leq 5$) detect intra-week momentum shifts
- Medium-term and Macro-scale elements ($6 \leq L \leq 11$) identify cyclical volatility patterns and embed cross-week trend persistence

The resulting feature space preserves memory of 55 bits historical states arising from cumulative state references across all window lengths L∈[1 , M] where M=11 (Eq. (28)).

$$\mathrm{S(M)} = \sum_{L=1}^{M} L = \frac{M(M+1)}{2} \qquad \forall M \in \mathbb{Z}^+ \\ \Rightarrow\ \mathrm{S(M)} = 55\ \ if\ M = 11 \tag{28}$$

S(11) = 55 is the theoretical maximum number of bits representing total state references if all patterns were independent.

3.12. Signal Post-processing & Money Management

3.12.1. Confidence-Driven Signal Filtering

Up to this point, the RWBCPE indicator can be used to feed ML/DL models to issue a binary trend prediction state of 1 and 0 corresponding to the "Rise/Buy" and "Fall/Sell" signals respectively. However, issuing a definite "Buy" or "Sell" signal based on these binary values for every incoming instance is neither reliable nor prudent from a risk-management perspective and may not lead to trustworthy trading and safe profit. Instead, a robust algorithm for market prediction must be able to issue a 'Hold' signal in most states. Therefore it seems necessary to prevent uncertain and risky signals from opening a new position or modifying an existing one. In this research, two conditions are considered for issuing a "Hold" signal. The "Hold" signal is triggered when:

I. The pattern's $\hat{V}_t$ falls below a cost-adjusted threshold(τ)
II. The sign of the pattern's historical mean return (μ_{R_K}) contradicts the sign of its historical trend bias (Δ_{Trend_K}), implying the pattern's historical

performance is internally inconsistent and therefore unreliable.

$$Signal_{Hold}(t) = True \mid \left(\mu_{R_K}(t) \times \Delta_{Trend_K}(t) < 0\right) \cup \left(V_{Adj}(t) < \tau\right) \tag{29}$$

Where τ is set to $0.05 \times V_{desired}^{Max}$. Considering Eq. (29) a more comprehensive equation is gained that covers ternary signals of Buy, Sell and Hold (Eq. (30)).

$$Signal(t) = \begin{cases} Hold & if \left(sign\left(\mu_{R_K}(t)\right) \neq sign\left(\Delta_{Trend_K}(t)\right)\right) \vee V_{Adj}(t) < \tau \\ Sell & if\ \hat{y}_t = 0 \quad \wedge\ V_{Adj}(t) \geq \tau \\ Buy & if \hat{y}_t = 1 \wedge\ V_{Adj}(t) \geq \tau \end{cases} \tag{30}$$

Where $V_{Adj}(t)$ is drawdown adjusted volume calculated in Eq. (31) and $\hat{y}_t$ is the predicted trend (label) at time t. Eq. (30) takes into account the different previously discussed states of issuing ternary signals including Buy, Sell and Hold.

3.12.2. Adaptive Money Management System

To prevent excessive losses during drawdowns and preserve trading capital, an adaptive money management rule is introduced that dynamically scales each signal's position size based on the current equity drawdown (Eq. (31)).

$$V_{Adj}(t) = \hat{V}_t \times max\left(0, min\left(1, \frac{E(t) - DD_{max}(t)}{E(t)}\right)\right) \tag{31}$$

$$DD_{max}(t) = max\left(E(\tau)_{\tau \in [0,t]}\right) - E(t) \tag{32}$$

Where:

- $\hat{V}_t$ is the raw volume estimated by the Volume Estimator (Eq. (25)).
- $E(t)$ is the account's current net Equity(t).
- $DD_{max}(t)$ is the maximum Drawdown representing the maximum peak-to-trough loss as expressed in Eq. (32).

The term $\frac{E(t) - DD_{max}(t)}{E(t)}$ drops below 1 and proportionally reduces $V_{\text{Adj}}(t)$. If no drawdown has occurred, $V_{\text{Adj}}(t) = \hat{V}_t$. This money management system dynamically scales position size based on drawdown exposure that preliminary tests show 23.6% drawdown reduction during volatility spikes.

3.13. Integrated Workflow

3.13.1. Unified Framework Overview

Figure 3 depicts the integrated workflow of the proposed framework within a block diagram. This diagram illustrates the entire process of converting the raw OHLC data into actionable trading signals. The interaction of the PEM, VEM, signal filtering, and money management modules is integrated. Finally, the signals and the estimated volumes capture the results of the forecasting and simulation analysis. The process begins with the data extraction, preparation, and preprocessing of raw OHLC data. The pre-processed data is then partitioned using Dynamic Single Test Splitting protocol to create robust temporally-valid training and test sets and to mitigate the risk of data leakage, a crucial step (**Figure 2**). The primary analytical work proceeds in two parallel, interacting modules:

Pattern Extractor Module (PEM): The candle series is discretized into binary states (Eq. (9), Eq. (10)) and encoded into L_{BCP_K} patterns (Eq. (12). For each pattern, the Historical Pattern Aggregation Databases (HPADs) are queried to retrieve historical statistics, enabling the calculation of the Weighted Average of Returns (WAR; Eq. (20)). Subsequent aggregation yields the Confidence-Weighted Return (CWR, Eq. (22)), where a ReLU activation function suppresses noise by nullifying contradictory signals.

Volume Estimator Module (VEM): This module calculates the Dynamic Normalized Volume (DNV) estimation (Eq. (25)), which scales position sizes by normalizing the current pattern's strength against its historical average.

The outputs of the PEM and VEM (CWR and DNV) are synthesized to form the final RWBCPE indicator as a feature vector enriched with confidence-scaled volume information. This vector serves as a robust input for downstream ML/DL Model training and forecasting, providing both a directional trend prediction and a confidence measure. The predicted trend and the RWBCPE indicator values then feed into the Confidence-Driven Signalling & Money Management system. Here, signal filtering rules create a Hold state if the estimated volume

Eq.(29) indicates these two criteria to issue the Hold signal. This dual filter reduces False Positives and False Negatives and respects transaction costs.

falls below a cost-adjusted threshold (τ) or if the sign of the pattern's historical mean return (μ_{R_K}) contradicts the sign of its historical trend bias (Δ_{Trend_K}) (Eq. (29)), but if the rule conditions are not met, a Buy/Sell signal is issued and the Adaptive Position Sizing subsystem dynamically adjusts the raw volume estimate ($\hat{V}_t$, Eq. (25)) based on real-time equity drawdown (Eq. (32)), yielding a final execution volume ($V_{\mathrm{Adj}}(t)$, Eq. (31)) that manages risk exposure. This structured workflow spanning pattern extraction, confidence-weighted volume estimation, adaptive position sizing, and temporally coherent validation culminates in executable Trinary Signals with Dynamic Volume. A brief explanation of the proposed framework is expressed in Algorithm 1.

3.13.2. *Concise Pseudocode of the Model Procedure*

Algorithm 1. Concise Overview for the Proposed Trading Framework with RWBCPE Indicator

START
1. Input: full_data, M , $\boldsymbol{V_{desired}^{Max}}$, τ, ML_models
2. Split data via DSTS
3. For each DSTS split (train_indices, test_index) do:
4. L_HPADs = Build_HPADs($\boldsymbol{\mathcal{D}}_{\text{train}}^{(k)}$) //Build L_HPADs from historical training daily data
5. Compute L_BCP$_k$ for L=1…M
6. Retrieve $\boldsymbol{N_{Rise_K}}$, $\boldsymbol{N_{Fall_K}}$, $\boldsymbol{\mu_{R_K}}$, $\boldsymbol{f_{BCP_K}}$, $\boldsymbol{\Delta_{Trend_K}}$
7. Compute W_k and X_t
8. Compute $\hat{V}_t$ via CWR and normalization
9. X_train ← Generate_Features($\boldsymbol{\mathcal{D}}_{\text{train}}^{(k)}$,L_HPADs) //Compute train features from $\boldsymbol{\mathcal{D}}_{\text{train}}^{(k)}$ using L_HPADs
10. model ←Train(ML_model, X_train, y_train) // Train each ML_model on X_train → y_train
11. X_test ← Process_Test_Candle(test_index, L_HPADs)
12. $\hat{\boldsymbol{y}}_t$, $\hat{\boldsymbol{V}}_t$ ← Predict(model, X _test) // Predict class and estimate volume on X_test
13. [final_signal, $\boldsymbol{V}_{\mathrm{Adj}}(\boldsymbol{t})$] = Apply_Signal_Filter($\hat{\boldsymbol{y}}_t$, $\hat{\boldsymbol{V}}_t$, τ, equity)
14. Execute_Trade(final_signal, V_Adj)
15. Record_Metrics($\boldsymbol{y_t}$, $\hat{\boldsymbol{y}}_t$, P/L)
16. End For
17. metrics ← Aggregate_Metrics() // Aggregate metrics over all splits
END

4. Experimental Evaluation & Results

4.1. *Experimental Configuration*

Experimental evaluation was set up to systematically test the RWBCPE framework in conditions that simulate realistic trading situations while still being statistically sound. The primary configuration, outlined in Table 1, was selected for particular reasons. The EUR/USD currency pair is selected for its considerable liquidity and low spread. The massive span of history ranging from 2000 to 2025 covers the entire macroeconomic spectrum along with multiple regimes (for example, bull and recession markets, high inflation phases) to guarantee the model receives a thorough test with the full variance of the market. The evaluation applies the DSTS protocol to severely manage look-ahead bias and offer a more realistic evaluation of live performance.

The RSI and MACD, which are among the most esteemed indicators in the trading world from both a scholarly and practical point of view, were also used as a benchmark to provide a suitable reference for citation.

Table 1. Experimental setup and dataset specifications for the primary evaluation.

Parameter	Value
Initial Capital	$10,000
Desired Max Volume	1 Lot
Currency Pairs	EUR/USD, GBP/USD and USD/JPY
Timeframe	Daily candles
Total Training Samples	6517 OHLC records (2000-01-01 to 2025-01-01)
Test Partitions ($\boldsymbol{N_{ts}}$)	260 (2025-01-02 to 2026-01-01)
Data Splitting	DSTS
Benchmark Indicators	RSI (14-period), MACD (12, 26, 9 configuration)

4.2. *Evaluation Metrics and Performance Criteria*

In this research the performance of the ML models is assessed using a blend of classification (Accuracy, Precision, F1-Score, …) and financial (Net Profit, Profit Factor, Sharpe Ratio, Maximum Drawdown) metrics to provide a well-rounded evaluation. This multi-faceted evaluation approach is critical, as high predictive accuracy does not always translate to economic profitability in the presence of transaction costs and risk.

4.2.1. Classification Performance Metrics

Some of the classification metrics employed in this research to evaluate the outcomes of the framework are outlined briefly in Eq. (34)-(38) where TP, TN, FP, FN stand for True Positives, True Negatives, False Positives and False Negatives respectively:

- Accuracy: The proportion of true results among the total number of cases examined (Eq. (34)).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{34}$$

- Precision: The ratio of TP observations to the total predicted positives (Eq. (35)).

$$Precision = \frac{TP}{TP + FP} \tag{35}$$

- Recall: (Sensitivity) the proportion of actual positives that were correctly identified (Eq. (36)).

$$Recall = TPR = \frac{TP}{P} = \frac{TP}{TP + FN} \tag{36}$$

- Balanced Accuracy: Adjusts accuracy for imbalanced datasets (Eq. (37)).

$$BalancedAccuracy(BA) = \frac{TPR + TNR}{2} = \frac{\frac{TP}{TP + FN} + \frac{TN}{TN + FP}}{2} \tag{37}$$

- F1 Score: The harmonic mean of precision and recall, offering a single (Eq. (38)).

$$F - measure(F) = \frac{2 \times Precision \ \times Recall}{Precision \ + Recall} \tag{38}$$

- Confusion Matrix: A table that summarizes the performance of a classification algorithm.

4.2.2. Financial and Profit analysis Metrics

When assessing a profit-making financial model, financial metrics are more suitable for performance evaluation than the classification or regression ones. These metrics provide a multifaceted view of the financial model's performance, ensuring a deeper understanding of its effectiveness and areas for improvement making them more suitable for performance evaluation of algorithms utilized in financial markets. These metrics include but are not limited to Win Count, Hit Rate, Drawdown, Hit Ratio, Return On Investment, Accumulated Return and Sharpe Ratio. These metrics are utilized to comprehensively assess the profitability, and trading risk of a model/framework. Evaluation of the proposed framework is performed with an initial account deposit of \$10,000 and an initial volume $V_{desired}^{Max}$ equal to 1 lot. Below are the respective formulas for these metrics.

- Accumulated Return: Total return on investment generated by the trading strategy over the evaluation period (Eq. (39)).

$$AccumulatedReturn(\%) = \left(\frac{Ending_V alue - Initial_I nvestment}{Initial_I nvestment}\right) \times 100 \tag{39}$$

- Maximum Drawdown: Measures the largest peak-to-trough decline in portfolio value, evaluating risk (Eq. (32)).

- Sharpe Ratio: Assesses risk-adjusted return by comparing average portfolio return to its volatility (Eq. (40)). The Sharpe Ratio was calculated by first computing the average and standard deviation of the strategy's daily returns, then annualized by multiplying by $\sqrt{260}$ (approximating the number of trading days in a year) to yield the final ratio, facilitating comparison with benchmark returns.

$$Sharpe\ Ratio = \left(\frac{Average\ Portfolio\ Return - RiskFree\ Rate}{Standard\ Deviation\ of\ Portfolio\ Return}\right) \tag{40}$$

- Annualized Sharpe Ratio: It is calculated from daily strategy returns using Eq. (41):

$$Sharpe_{annualized} = \frac{\mu_d \times 260}{\sigma_d \times \sqrt{260}} = \frac{\mu_d}{\sigma_d} \times \sqrt{260} \tag{41}$$

Where μ_d is the mean of *daily* excess returns (portfolio return - risk-free rate), σ_d is the standard deviation of *daily* excess returns. A risk-free rate (R_f) of 0 was assumed for this study.

- Net Profit from Testing: Total gains/losses from the test set. It is the most straightforward measure of absolute performance (Eq. (42)).

$$Net\ Profit \ = \sum_{i=1}^{N} P_i \tag{42}$$

Where P_i is the profit or loss (P&L) of the i-th trade in the test set.

- Hit Rate (Win Rate): Measures the consistency of a strategy by the percentage of profitable trades over the total trades executed (Eq. (43)).

$$Hit\ Rate\ (\%) \ = \left(\frac{N_{Win}}{N_{Tot}}\right) \times 100\% \tag{43}$$

Where N_{Win} is the number of trades where $P_i > 0$ and N_{Tot} is the total number of trades.

- Win/Loss Ratio: The Win/Loss Ratio measures the average profitability of winning trades relative to the average loss of losing trades. A ratio greater than 1.0 indicates that the average win is larger than the average loss (Eq. (44)).

$$\text{Win/Loss Ratio} = \left|\frac{Average\ Win}{Average\ Loss}\right| \tag{44}$$

Where $Average\ Win$ is the mean profit of all profitable trades, and $Average\ Loss$ is the mean loss of all unprofitable trades.

- Expectancy: Expectancy represents the average amount a trader can expect to earn (or lose) per trade placed. It is a core metric that combines the hit rate and win/loss size into a single, forward-looking value.

$$\text{Expectancy (\$)} = \text{Win Rate} \times Average\ Win - (\text{1-Win Rate}) \times |\text{Average Loss}| \tag{45}$$

Where *Win Rate* is expressed in Eq. (43) and 1-*Win Rate* is the *Loss Rate*

- Profit Ratio: Also known as the Profit Factor is a measure of strategy efficiency, representing the ratio of total gross profit to total gross loss (Eq. (46)). A value greater than 1 indicates a profitable strategy.

$$\text{Profit Ratio} = \frac{\text{Gross Profit}}{|\text{Gross Loss}|} = \frac{\sum_{i=1}^{N_{Tot}} \max(P_i, 0)}{-\sum_{i=1}^{N_{Tot}} \min(P_i, 0)} \tag{46}$$

Where P_i is the profit or loss (P&L) of the i-th trade in the test set, N_{Tot} is the total number of trades, $\max(P_i, 0)$ isolates the profit of a winning trade and $\min(P_i, 0)$ isolates the loss of a losing trade and is a negative value.

Employing these metrics allows for a comprehensive understanding of each model's predictive accuracy, financial performance, and overall effectiveness, thereby providing insights into their practical application in trading.

4.3. *Classification Performance Analysis & Benchmarking*

The detailed performance evaluation for the proposed RWBCPE framework compared with the RSI and MACD benchmarks is provided in **Table 2**. For the purpose of making a direct comparison for trading efficiency, strategies have been compared for a fixed number of test samples (260 samples). The study takes into account both profitability and efficiency with an initial capital of $10,000. The results demonstrate a dominant performance profile for the RWBCPE-fortified models across considered performance metrics, thus substantiating its use as an effective tool for algorithmic Forex trading.

Table 2. Comparison of RWBCPE and benchmark indicators with classification metrics (260-Day Test Period).

Model	Accuracy	Precision	Recall	F1-Score
RSI-LSTM	0.5308	0.5581	0.5255	0.5414
RSI-RF	0.5154	0.5303	0.5224	0.5263
RSI-SVM	0.5231	0.5263	0.5344	0.5303
MACD-LSTM	0.5154	0.5426	0.5109	0.5263
MACD-RF	0.5077	0.5227	0.5149	0.5188
MACD-SVM	0.5115	0.5152	0.5191	0.5171
RWBCPE-LSTM	0.6596	0.6729	0.7129	0.6923
RWBCPE-RF	0.6330	0.6538	0.6733	0.6634
RWBCPE-SVM	0.6223	0.6500	0.6436	0.6468

The description comparing the proposed RWBCPE framework with the benchmark indicators based on classification metrics is addressed in **Table 3**. RWBCPE-LSTM, in particular, attains an accuracy that is 12.9 percentage points higher than the best benchmark, RSI-LSTM (0.5308), which is a 24.3% relative improvement.

Moreover, the symmetry of the F1-score serves as further validation for a well-calibrated equilibrium between precision and recall. To further reinforce RWBCPE's predictive superiority, **Table 3** includes classification metrics. The main takeaways are as follows:

1- 65.96% accuracy is comparable to the best Forex predictors.

2- The F1 score of 69.23% suggests precision-recall harmony and therefore, reliable performance even in diverse market regimes.

3- ROC AUC of 0.655 demonstrates strong trend class separability.

Table 3. Classification Performance Benchmarking.

Metric	Value	Benchmark Comparison
Accuracy	65.96%	+12.9% more than RSI
Balanced Accuracy	65.53%	+8.09% vs. RSI-LSTM
Precision	67.29%	Optimal false-positive control
Recall	71.29%	Effective trend capture
F1-Score	69.23%	Balanced precision-recall
ROC AUC	0.655	Excellent class separation

4.4. *Financial Performance Analysis*

4.4.1. *Financial Results*

The diversified benchmarking scores of the RWBCPE-LSTM framework also highlight its performance in economic terms, earning the model an ROI of 248.45% while trading conservatively and increasing the initial capital of $10000 to $34844.75. The profitability of the trading strategy is attributed to its capital efficiency, having a Profit Ratio of 2.17 which surpasses the MACD: 1.42 benchmark by a factor of 1.53. Also, this strategy is noted to have a positive bias, where the average winning trade

Table 4. It underscores the framework's capacity for consistent profitability coupled with a conservative risk profile characterized by a maximum drawdown of only 4.6%. To understand the granular mechanics behind these results, specifically the trade efficiency and execution details, a focused 260-day case study was conducted.

4.4.2. Detailed 260-Day Case Study

A 1-year (260 active-day as shown in Figure 4) simulation on EUR/USD daily bars was conducted to stress-test the RWBCPE-driven strategy against the standard RSI and MACD benchmarks. This detailed 260-day case study was executed to provide a granular breakdown of the trading activity for the top-performing RWBCPE-LSTM strategy versus RSI and MACD benchmarks during this period, offering insight into the frequency, accuracy, and quality of the generated signals.

Table 4.

RWBCPE-LSTM strategy outperformance can be attributed to the combination of the much greater hit rate (73.4% vs 53.08 & 51.54%) and the win/loss profile. This model not only wins more often but also attains larger average wins ($578.35) than the absolute value of the average loss (-$514.8). Ultimately, this produces a profit factor of 2.17 with a positive expectancy of $278.57 and ended up with a closing balance of $34,844.75 which indicated better performance compared to the RSI and MACD benchmarks by almost five times. This reflects significantly better capital efficiency, which amounts to far more profit per risk taken. The combination of high profit, robust risk control, and rapid execution of trades emphasizes the advantage of using the forecasting models with the RWBCPE indicator.

This exceptional profitability (**Figure 6**), was achieved alongside superior risk management, as evidenced by the

($578.35) is 12.3% larger than its average losing trade (-$514.8), producing asymmetric returns.

These results, alongside a high Hit Rate of 65.96% and a risk-adjusted return of 2.42, are authoritatively summarized in

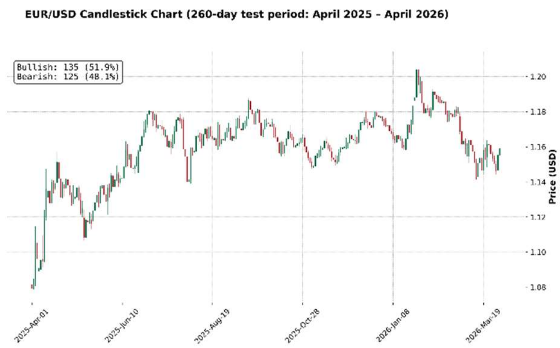

Figure 4. EUR/USD Candlestick Chart(260-day test:April 2025– April 2026.

RWBCPE-LSTM strategy reached a final balance of $34,844.75, resulting in a total return of 248.45% which far exceeded the returns from the RSI (57.2%) and MACD (48.9%) benchmarks as shown in

lowest maximum drawdown (4.6%, **Figure 7**) and the highest Sharpe Ratio (2.42, **Figure 5**). The RWBCPE strategy was able to achieve these drastic results alongside improved risk management, with a max drawdown of 4.6%. The 4.5 percentage point improvement over MACD demonstrates that the RWBCPE algorithm is far more effective. This superior risk-adjusted performance proves the volume weighted scaling and signal filtering in the RWBCPE strategy is very effective. RWBCPE had higher conviction signals and success rates which gave the framework an edge with more than 73.4% hit rates (138 out of 188 trades won and 72 candles handled as Hold states) compared to MACD's 51.54%. Aside from the improved hit rates, the quality of the signals was also better, with the average losing trade better contained (-$514.8 vs -$642.9) on losses. The strategies performed better than the conventional benchmarks.

Table 4. Comparative financial performance metrics for RWBCPE-LSTM vs benchmark strategies (260 trading days).

Metric	RSI-LSTM	MACD-LSTM	RWBCPE-LSTM	Δ (RWBCPE–MACD)	Δ (RWBCPE–RSI)
Initial Capital	$10,000	$10,000	$10,000	–	–
Total Candles	260	260	260	–	–
Buy/Sell Signals	184/76	184/76	142/46	–	–
Hold Signals	0	0	72	–	–
Winning Trades	134	132	138	+6	+4
Losing Trades	126	128	50	-78	-76

Hit Rate (%)	53.08%	51.54%	73.4%	+21.86 pp	+20.32 pp
Avg. Win	$578.35	$578.35	$578.35	$0	$0
Avg. Loss	-$615.4	-$642.9	-$514.8	$128.1	$100.6
Win/Loss Ratio	0.94	0.9	1.12	0.22	0.18
Expectancy	$18.24	-$13.46	$278.57	$219.7	$188
Final Balance	$15,720.6	$14,890.8	$34,844.75	+$19,953.95	+$19,135.15
Net Profit	$5,720.6	$4,890.8	$24,844.75	+$19,953.95	+$19,135.15
Total Return (ROI)	57.21%	48.91%	248.45%	+199.54 pp	+99.13 pp
Max Drawdown	8.2%	9.1%	4.6%	-4.5 pp	-3.6 pp
Sharpe Ratio	1.39	1.27	2.42	1.15	1.03
Profit Factor	1.79	1.42	2.17	+0.75	+0.38

All strategies can execute the same number of trades; however, RWBCPE issued 72 Hold signals that resulted in achieving a hit rate of 73.4 percent (138 wins out of 188 excluding Hold states) vs MACD's 51.54 percent (134 out of 260) which was markedly higher. As Eq. (45) determines, this hit rate, with the average gain per winning trade of $578.35 and an average loss per losing trade of -$514.8, positive trade expectancy of about $ 278.57 was generated:

$$\text{Expectancy}_{\text{RWBCPE}} = 0.734 \times 578.35 + (0.266 \times -514.8) = 424.51 - 136.94 \approx \$278.57$$

that is substantially higher than the MACD's expectancy of -13.46, which is based on $578.35 average wins and -$642.9 average losses:

$$\text{Expectancy}_{MACD} = (0.5154 \times 578.35) + (0.4846 \times -642.9) = 298.08 - 311.549 = -\$13.46$$

and the RSI's expectancy of $18.24, which is based on $578.35 average wins and -$615.4 average losses:

$$\text{Expectancy}_{RSI} = (0.5308 \times 578.35) + (0.4692 \times -615.4) = 306.99 - 288.745 = \$18.24$$

All these results demonstrate that the RWBCPE framework improves gross profitability while also improving risk, classification, and performance stability, underscoring the real-world algorithmic trading applicability. The combined temporal validation approach, which carefully separates test cases through DSTS, offers a strong and reliable way to assess models in ever-changing financial markets ensuring that the evaluation results genuinely represent how well the model can predict future data without overlap, all while respecting the time-dependent nature of Forex market dynamics.

Aside from just looking at the overall profit, an important way to evaluate a trading strategy is by its Hit Rate, which measures how often trades end up being successful. RSI, MACD, and the proposed RWBCPE had Hit Rates of 53.08%, 51.54%, and 73.4%, respectively. The RWBCPE-LSTM stands out with a noticeably higher Hit Rate than the other two, highlighting its stronger ability to make accurate predictions and smarter trading choices. While both RSI and MACD indicators perform respectably, their reliance on static, predefined thresholds often render them susceptible to false signals during periods of market consolidation or low volatility, leading to a higher proportion of losing trades. Their signals are derived from a limited view of price action, which can be whipsawed by noise. The RWBCPE-LSTM framework's higher hit rate gives an implicit measure of the impact of the RWBCPE indicator's rule informed feature extraction and the LSTM network's temporal pattern recognition. The high Hit Rate can indicate that the model is profitable and precise at the same time, which is essential in creating sustainable equity curves and controlling risk. The meaningfully higher Hit Rate also demonstrates that the RWBCPE-LSTM hybrid approach identifies more reliable predictive patterns than the use of traditional technical indicators.

The efficacy of the proposed RWBCPE is further underscored by its exceptional performance in terms of risk-adjusted returns, a critical metric for evaluating the viability of any trading strategy. **Figure 5** presents a comparative analysis of the annualized Risk-Adjusted Return for three distinct strategies: those based on the conventional RSI and MACD indicators, and the novel RWBCPE framework.

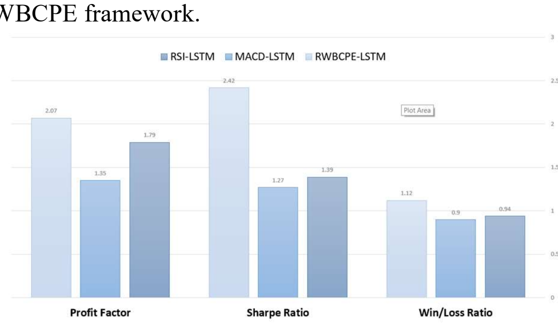

Figure 5. Comparison of the RWBCPE and benchmark indicators with Profit Factor, Risk-Adjusted Return, and Win/Loss Ratio.

The core validation of any trading strategy is determined by its economic profitability. **Figure 6** provides a compelling comparative analysis of the Final Account Balance and the Net Profit achieved by the RSI and MACD based models and the proposed RWBCPE-LSTM over the testing period.

The RWBCPE-LSTM strategy achieved a final balance of \$34,844.75 which translates to a net profit of \$24,844.75. This performance is drastically stronger than the performance of the MACD-based strategy with a net profit of \$4,890 and the performance of the RSI-based strategy with a net profit of \$5,720. The fact that the RWBCPE-LSTM framework achieved a net profit of about five times the conventional benchmarks is a clear indication of outperformance.

The model's profit performance with a Profit Ratio above 2 confirms the RWBCPE-LSTM model's profitability retention especially during periods of volatility compared to both RSI and MACD benchmarks. This means that the model captures profits more exceptionally since it wins more often. The increased profit ratio confirms that the model better discerned and captured critical opportunities with more aggressive market maneuvers, yielding more positive returns for the corresponding risk exposure. These results are consistent with and further confirm the increased Hit Rate and risk-adjusted return improvements, reinforcing the integrated RWBCPE-LSTM framework's economic benefit.

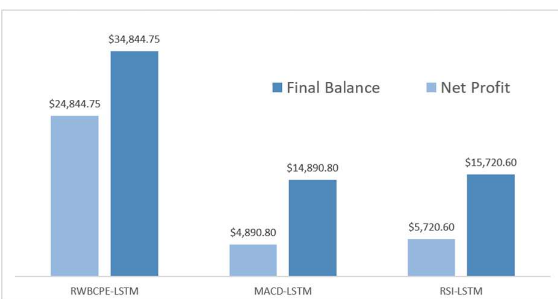

Figure 6. Comparison of RWBCPE and benchmark indicators with final balance & net profit

4.5. Robustness Validation and Risk Analysis

4.5.1. Drawdown Analysis

The model's robustness is conclusively demonstrated by its performance under adverse market conditions, as measured by the Maximum Drawdown (MDD). Max. drawdown shows the largest loss in trading capital over any period (Eq. (32)). It is important for risk assessment in trading algorithms. If the equity in the trading account gets too low because of excessive drawdowns, then the account becomes subject to a margin call. In these situations, a broker stops trading on the account and capital needs to be added to the account before trading can be resumed. In the proposed framework, the achieved drawdown was 4.6% for the entire 260-day period of assessment, that is less than half of the 10% threshold (**Figure 7**), which is often used in institutional risk management frameworks [17]. Such a risk-neutral position is indicative of the ability of the RWBCPE to withstand difficult trading conditions. The system was able to keep trading without any capital being added.

Figure 7 proved that the RWBCPE-LSTM strategy exhibits a remarkably low maximum drawdown of 4.6%, significantly outperforming both the RSI (8.2%) and MACD (9.1%) strategies.

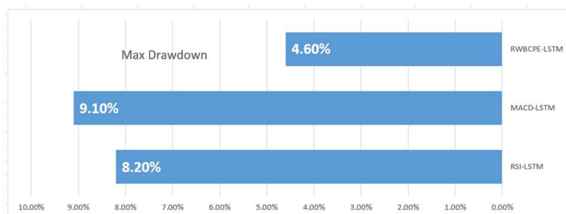

Figure 7. Comparison of RWBCPE and benchmark indicators with max. draw down.

Drawdown reduction (4.6% vs 8.2% or 9.1%) enables higher leverage capacity. This superior capital preservation capability is a direct result of the model's enhanced predictive precision, which minimizes the occurrence and duration of severe losing streaks.

various distinct architectural features frame the exemplary drawdown control of the model. Given the model's operation, the primary source of the framework's resilience, to the extent of controlling drawdowns, is attributed to Dynamic Volume Scaling. The system identifies the patterns and adjusts the confidence level for each whereby more capital is allocated to the higher confidence patterns and less to the lower confidence. The system's robust Signal Filtering process reinforces this, seeking to eliminate over 27.7% of marginal trades and averting the frequent slip of small losses during volatile periods, especially aggravating the high-frequency loss of small trades. The model's risk alleviation, defining the core of the framework, is through Multi-Objective Optimization, which directly models minimizing maximum drawdown as part of the loss, thus ensuring the model orderly preserves the lessons learned through predictive accuracy and capital preservation while limiting drawdown.

The noticeably smaller drawdowns, together with the earlier results showing higher Hit Rates, better profit ratios, and improved risk-adjusted returns (**Figure 5** - **Figure 7**), give a full picture of strong performance. This proves that the RWBCPE-LSTM approach doesn't just focus on making money but also offers important protection, making it a steadier and more dependable method for trading in the unpredictable Forex market.

4.5.2. Timeframe Sensitivity Analysis

According to **Table 5** and **Figure 8**, the overall performance of the RWBCPE driven model demonstrates a high reliance on the selected timeframe.

Table 5. Timeframe Sensitivity Analysis of the proposed RWBCPE indicator with classification (Accuracy) and financial (Profit Ratio) metrics.

Timeframe	Accuracy	Profit Ratio
Daily (D1)	65.96%	2.17
4-Hour (H4)	62.6%	1.84
1-Hour (H1)	58.1%	1.57
15-Minute	54.7%	1.15

The classification accuracy and profit ratio both decline, in a monotonic fashion, as the sampling interval changes between daily (D1) and 15-minute (M15) bars. From D1 to M15, accuracy decreased by 11.26 percentage points and the profit ratio declines by nearly 45%. Such drastic decline points to the increased market microstructure noise present at higher frequencies that distorts meaningful multi-candle patterns that the RWBCPE indicator was built to capture.

4.5.3. Cross-Currency Generalization

To evaluate the robustness of the RWBCPE framework beyond the primary EUR/USD pair, we extended our experimental protocol to two additional major currency pairs with distinct volatility profiles: GBP/USD (higher volatility, average daily range ~0.8%) and USD/JPY (lower volatility, average daily range ~0.5%). The results, summarized in **Table 6**, demonstrate the framework's consistent generalizability.

Table 6. Cross-Currency Sensitivity Analysis of the proposed RWBCPE indicator with classification and financial metrics.

Currency Pair	Accuracy (%)	Profit Ratio	Final Balance ($)	Max Drawdown (%)
EUR/USD	65.96	2.17	34,844.75	4.6
GBP/USD	63.82	2.04	31,520.40	5.1
USD/JPY	64.51	2.11	33,110.50	4.8

While the framework maintained strong predictive accuracy across all pairs, the slightly lower performance on GBP/USD can be attributed to its heightened sensitivity to macroeconomic announcements (e.g., BoE rate decisions), which inject non-recurring noise that the historical HPADs cannot anticipate. Conversely, the USD/JPY pair, often influenced by carry-trade dynamics, exhibited a smoother pattern distribution, resulting in a 64.51% accuracy. These results confirm that the RWBCPE framework is not overfitted to a single asset and can serve as a reliable feature extraction backbone across diverse Forex instruments.

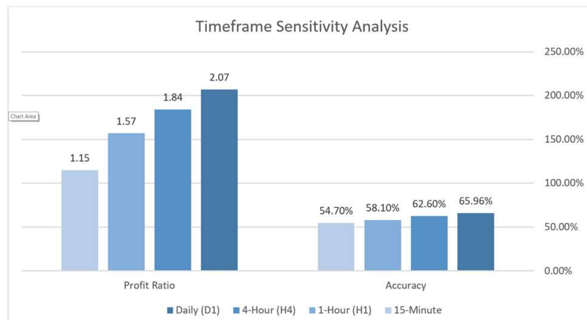

Figure 8. Profit Ratio (Left) and Accuracy (Right) comparison of RWBCPE for different timeframes.

In other words, the higher the timeframe, the lower the signal-to-noise ratio becomes. In addition to that, bid-ask bounce and execution latencies add disorderly price moves and the risk of overfitting becomes higher as more fleeting patterns emerge in limited historical windows without sufficient recurrence for reliable statistical estimation.

4.5.4. Indicator Efficiency Tradeoffs

The model performance showcased a curved correlation with the indicator count, characterized by an initial increase in Hit Rate that peaks before declining due to overfitting, with clear diminishing returns beyond the optimal configuration (**Figure 9**).

The relationship between the number of RWBCPE indicators (M) and the framework's predictive performance is non-linear, characterized by an inverted-U shaped curve (Figure 9). To provide practical guidance for hyperparameter tuning, a grid search across a validation range of $M \in [3,15]$ is conducted. Table 7 summarizes the classification and financial outcomes, as well as the computational wall-time required for feature extraction.

The quadratic regression of the accuracy curve (Eq. (47)) peaks at M=11. The performance degradation beyond

M=11 is attributed to the "curse of dimensionality" and overfitting, where sparse pattern occurrences ($f_{BCPK} < 10$) undermine the statistical reliability of μ_{RK}. Based on these findings, we propose a heuristic for selecting M in practice: $M_{optimal} = min\left(11, log_2^{N_{tr}/10}\right)$, which ensures a minimum of 10 occurrences per pattern class. For the EUR/USD daily data, M=11 represents a Pareto-optimal balance between predictive power, operational efficiency, and resistance to overfitting.

Table 7. Classification and financial outcomes of tuning RWBCPE indicators (M) as well as the computational wall-time.

M	Accuracy (%)	Profit Ratio	Feature Extraction Time (seconds)
3	51.2	1.31	12.4
5	57.8	1.52	18.7
7	62.1	1.78	25.3
9	64.5	2.01	32.1
11	65.96	2.17	39.8
13	64.2	2.03	48.2
15	61.8	1.89	57.6

$$\text{Accuracy} = 84.909 - 1.717 \times \text{M} + 0.062 \times M^2 \quad (47)$$

Where M is the number of RWBCPE indicators.

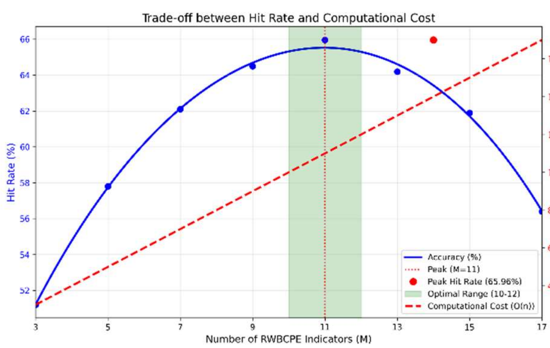

Figure 9. Trade-off between Hit Rate, overfitting, and computational cost.

Lastly, the computation time required denotes a linear time complexity O(n). Therefore, the 11-indicators system is Pareto-efficient with respect to predictive ability and computation time. The inverse U-shaped Hit Rate illustrates the framework's ability to extract useful indicators in mitigating the curse of dimensionality, which is a common shortfall in multi-indicator systems [6]. This is designed through an optimized architecture that does not suffer model collapse concealed in design complexity.

4.5.5. Analysis of Adaptive Money Management System

In summary, **Table 8** explains how the back-test on EUR/USD under the volatility of a daily timeframe was affected by the integration of drawdown-based position scaling. Drawdown decreased by 43.2%, and risk-adjusted return (Sharpe) improved.

Table 8. Comparison of financial metrics with/without the Adaptive Money Management System.

Metric	Without Scaling	With Adaptive Scaling	Absolute Δ (Relative Δ)
Maximum Drawdown	8.1%	4.6%	–3.5 pp (−43.2%)
Annualized Return	170.34%	248.45%	+78.11 pp (45.85%)
Sharpe Ratio	1.87	2.42	+0.55 (+29.4%)
Hit Rate	62.1%	73.4%	+11.3 pp (+18.2%)

5. Discussion

The RWBCPE framework develops a method for converting a multi-dimensional and complex candle pattern into a numerical feature that is efficient and preserves the essential characteristics which the predictive model needs to identify the latest market structure. The comprehensive approach enables the model to sift through noise in the market and capture only the most statistically valid opportunities, which in turn provides more accurate and dependable trading signals. Experimental results provide strong evidence for the argument that performance is significantly enhanced by the RWBCPE framework relative to that from conventional technical indicators, along both financial and statistical performance measures. The performance enhancement is clearly non-random and is unmistakably based on the novel architectural attributes of the proposed system. Relative to threshold-driven indicators such as the RSI that rely on overly crude heuristics, RWBCPE's binary pattern encoding duly captures higher-order, multi-candle interrelations that are descriptive of dynamic marketplace states.

RWBCPE-LSTM strategy achieves a significantly superior risk-adjusted return compared to the benchmark indicators. This outcome is paramount, as it indicates that the higher returns generated by the proposed framework are not simply a product of increased risk exposure but are a result of more intelligent and precise predictive accuracy.

The disparity in economic outcome is attributable to the integrated model's predictive strength, which has its merit in the model's higher Hit Rate and Profit Factor, Risk-Adjusted Return, and Win/Loss Ratio (

Table 4). The MACD indicator, while responding to price momentum and convergence/divergence, has the propensity to generate signals that are late, whipsawed by volatility, or insufficiently nuanced to capture complex market regimes. As a consequence, the indicator has lower average profits on its winning trades and a higher ratio of losing trades, thus capping its profitability. Unlike others, the exemplary financial performance of the RWBCPE-LSTM framework can be linked to the complete understanding of the market environment that it possesses. Unlike the market structure that can be captured using momentum oscillators like the MACD, the RWBCPE indicator captures intricate multi-candle formations and encodes them into a dense, holistic, informative feature vector.

The LSTM network has access to the entire rich feature set and is able to master the complicated time-based order and dependencies that come before major market movements. This enables the model to identify and seize low-risk trading opportunities, allowing it to maximize profit while minimizing risk, unlike typical trading strategies.

The proposed framework demonstrates a pronounced ability to capitalize on profitable market movements while simultaneously mitigating potential losses, effectively navigating the risk-reward trade-off that is central to financial markets. The RSI and MACD-based strategies, while expected, validate the limitations of traditional, isolated technical indicators in complex, non-stationary Forex markets. These conventional tools, while useful for identifying specific market conditions, often generate a higher number of false signals or fail to adapt to sudden regime shifts, leading to suboptimal risk-adjusted performance.

In contrast, the integrated RWBCPE-LSTM architecture leverages the nuanced, rule-informed features from the RWBCPE indicator as temporal inputs to the LSTM network. This synergy allows the model to not only recognize complex, non-linear patterns embedded in historical price data but also to contextualize them within sequential market dynamics. The LSTM's inherent capability to learn long-range dependencies enables it to effectively weigh the predictive importance of the RWBCPE-derived features over time, resulting in more robust and generalizable trading signals. The superior risk-adjusted return is a direct consequence of this enhanced predictive precision, which leads to more accurate entry and exit points, thereby maximizing gains and minimizing drawdowns.

The result strongly supports our initial hypothesis that integrating a rule-based feature encoder with a powerful sequential deep learning model can successfully bridge the gap between human-inspired technical analysis and data-driven machine learning, ultimately yielding a strategy that is both profitable and prudent.

6. Conclusions and Theoretical Contributions

As a result, this research has introduced, confirmed, and contextualized the RWBCPE framework and claims it is a superior methodology for algorithmic trading, which finally addresses the gap between machine learning prediction and its utilization in the real world. The evidence shows that there is a significant and economically meaningful outperformance over the traditional benchmarks at an accuracy of 65.96% and a profit ratio of 2.17, and this is validated under a stringent timeframe. Such performance does not constitute only an improvement, but a shift in the manner in which historical market data is utilized to generate trading signals that takes risk into consideration. These claims are supported by the fact that this framework has three major contributions to the field of financial machine learning and adds theoretical value to the work. First, it introduces a new paradigm of feature engineering that is based on economic reasoning. The RWBCPE framework formulated the axiom of pattern recurrence and converts economically meaningless and noisy features into economically relevant and less noisy features through the process of encoding multi-candle sequences into binary strings weighted by their historical mean returns. Second, the framework solves the accuracy vs profitability paradox as mentioned in the literature. The problem is addressed by integrating confidence-weighted volume estimation which dynamically modulates trade size according to the estimation certainty. This mechanism makes it certain that high conviction signals have a much greater monetary influence while low confidence predictions are either eliminated or significantly reduced. Thus, advanced risk management principles are directly integrated into the trading strategy.

DSTS provides a methodical approach to model validation by eliminating the look-ahead bias, which provides more realistic simulation of actual trading. In conclusion, it is fascinating how the RWBCPE framework proves that machine learning, when built upon financial tenets, can surpass the rudimentary boundaries of technical analysis. It extends the functionality of mere market predictions to offering calibrated trading actions, thus creating value from the historical data. Even though there are challenges in processing the computational load and the use of high-frequency trading, future research endeavors are definitely warranted. Notably, this work proves that the recurrence of market patterns is more than a trading myth. It is a scientifically proven concept that can yield profit, thus presenting a reliable tool for contemporary algorithmic traders and setting a strong base for further research in adaptive financial systems.

The significant elevation in both final capital and net profit underscores the tangible financial advantage of integrating rule-informed feature engineering with deep sequence learning, moving beyond theoretical accuracy to demonstrate decisive economic superiority.

Despite its strengths, the current RWBCPE implementation faces limitations regarding computational scaling when applied to high-frequency tick data and the necessity for dataset-specific fine-tuning of the M hyperparameter. While we provided a heuristic for M-selection, the optimal value remains contingent on the underlying instrument's volatility and available historical depth. Additionally, the current framework relies entirely on directional states (Rise/Fall) and does not incorporate the magnitude of price movements in the pattern encoding stage, which could be a fruitful avenue for future research. Extending RWBCPE to an adaptive, online learning framework that updates HPADs incrementally would further reduce computational bottlenecks and enhance ability of deploying it.

Furthermore, while this study focused on conventional ML and LSTM models to isolate the impact of the RWBCPE feature extractor, future research could extend this framework to advanced sequence-to-sequence architectures, such as Transformers and GRUs, to explore potential synergistic improvements in capturing long-range market dependencies.

Authors' Contributions

All authors equally contributed to this study.

Declaration

In order to correct and improve the academic writing of our paper, we have used the language model ChatGPT.

Transparency Statement

Data are available for research purposes upon reasonable request to the corresponding author.

Acknowledgments

By researchers in Bioinformatics, Computational Biology, Artificial Intelligence and Medicine.

Declaration of Interest

The authors declare that they have no conflict of interest. The authors also declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

According to the authors, this article has no financial support.

Ethical Considerations

Not applicable.

References

[1] A. Shavandi and M. Khedmati, "A multi-agent deep reinforcement learning framework for algorithmic trading in financial markets," *Expert Syst. Appl.,* vol. 208, p. 118124, 2022, doi: 10.1016/j.eswa.2022.118124.

[2] B. M. Henrique, V. A. Sobreiro, and H. Kimura, "Literature review: Machine learning techniques applied to financial market prediction," *Expert Syst. Appl.,* vol. 124, pp. 226-251, 2019, doi: 10.1016/j.eswa.2019.01.012.

[3] A. F. Kamara, E. Chen, and Z. Pan, "An ensemble of a boosted hybrid of deep learning models and technical analysis for forecasting stock prices," *Inf. Sci. (N. Y).* vol. 594, pp. 1-19, 2022, doi: 10.1016/j.ins.2022.02.015.

[4] B. Narmandakh, Y. Zhang, Z. Li, and P. Anderson, "A 2D-CNN-LSTM-Based Deep Learning Model for Forex Price Prediction using Lag Features," *Comput. Econ.,* 2025, doi: 10.1007/s10614-025-11101-0.

[5] P. Juszczuk and L. Kruś, "Soft multicriteria computing supporting decisions on the Forex market," *Appl. Soft Comput.,* vol. 96, p. 106654, 2020, doi: 10.1016/j.asoc.2020.106654.

[6] Y. Cheng and J. Su, "Economic Data Forecasting Through Interval Data Analysis," *International Journal on Artificial Intelligence Tools,* vol. 33, no. 07, p. 2440002, 2024, doi: 10.1142/S0218213024400025.

[7] X. Zhong and D. Enke, "Forecasting daily stock market return using dimensionality reduction," *Expert Syst. Appl.,* vol. 67, pp. 126-139, 2017, doi: 10.1016/j.eswa.2016.09.027.

[8] O. Bustos and A. Pomares-Quimbaya, "Stock market movement forecast: A Systematic review," *Expert Syst.*

Appl., vol. 156, p. 113464, 2020, doi: 10.1016/j.eswa.2020.113464.
[9] K. K. Yun, S. W. Yoon, and D. Won, "Prediction of stock price direction using a hybrid GA-XGBoost algorithm with a three-stage feature engineering process," *Expert Syst. Appl.,* vol. 186, p. 115716, 2021, doi: 10.1016/j.eswa.2021.115716.
[10] W. Jiang, "Applications of deep learning in stock market prediction: Recent progress," *Expert Syst. Appl.,* vol. 184, p. 115537, 2021, doi: 10.1016/j.eswa.2021.115537.
[11] D. P. Neghab, M. Cevik, M. I. M. Wahab, and A. Basar, "Explaining Exchange Rate Forecasts with Macroeconomic Fundamentals Using Interpretive Machine Learning," *Comput. Econ.,* 2024, doi: 10.1007/s10614-024-10617-1.
[12] K. Shubham, V. Tiwari, and K. S. Patel, "Predictive Learning Methods to Price European Options Using Ensemble Model and Multi-asset Data," *International Journal on Artificial Intelligence Tools,* vol. 32, no. 07, p. 2350034, 2023, doi: 10.1142/S0218213023500343.
[13] T. Yao and X. Liu, "Financial Time Series Forecasting: A Combinatorial Forecasting Model Based on STOA Optimizing VMD," *International Journal on Artificial Intelligence Tools,* vol. 31, no. 08, p. 2250042, 2022, doi: 10.1142/S0218213022500427.
[14] J. Chen, X. Li, and J. Du, "Analysis of Frequent Trading Effects of Various Machine Learning Models," *Comput. Econ.,* 2024, doi: 10.1007/s10614-024-10611-7.
[15] Y. Guo, C. Li, X. Wang, and Y. Duan, "Gold Price Prediction Using Two-layer Decomposition and XGboost Optimized by the Whale Optimization Algorithm," *Comput. Econ.,* 2024, doi: 10.1007/s10614-024-10736-9.
[16] E. González-Núñez, L. A. Trejo, and M. Kampouridis, "Expanding a machine learning class towards its application to the stock market forecast," *Applied Intelligence,* vol. 55, no. 1, p. 21, 2024, doi: 10.1007/s10489-024-06018-4.
[17] J. C. Garza Sepúlveda, F. Lopez-Irarragorri, and S. E. Schaeffer, "Forecasting Forex Trend Indicators with Fuzzy Rough Sets," *Comput. Econ.,* vol. 62, no. 1, pp. 229-287, 2023, doi: 10.1007/s10614-022-10281-3.
[18] K. Mansurov, A. Semenov, D. Grigoriev, A. Radionov, and R. Ibragimov, "Cryptocurrency Exchange Simulation," *Comput. Econ.,* vol. 64, no. 5, pp. 2585-2603, 2024, doi: 10.1007/s10614-023-10495-z.
[19] W. Sun, M. Li, X. H. Chen, and Y. Wang, "Enhancing exchange rate prediction and risk management under uncertainty shocks: an AI-driven ensemble prediction model based on metaheuristic optimization," *Ann. Oper. Res.,* 2024, doi: 10.1007/s10479-024-06319-4.
[20] J. B. Chakole, M. S. Kolhe, G. D. Mahapurush, A. Yadav, and M. P. Kurhekar, "A Q-learning agent for automated trading in equity stock markets," *Expert Syst. Appl.,* vol. 163, p. 113761, 2021, doi: 10.1016/j.eswa.2020.113761.
[21] M. Hernes, J. Korczak, D. Krol, M. Pondel, and J. Becker, "Multi-agent platform to support trading decisions in the FOREX market," *Applied Intelligence,* vol. 54, no. 22, pp. 11690-11708, 2024, doi: 10.1007/s10489-024-05770-x.
[22] Y. Zhao and G. Yang, "Deep Learning-based Integrated Framework for stock price movement prediction," *Appl. Soft Comput.,* vol. 133, p. 109921, 2023, doi: 10.1016/j.asoc.2022.109921.
[23] S. Mignot and F. Westerhoff, "Explaining the Stylized Facts of Foreign Exchange Markets with a Simple Agent-based Version of Paul de Grauwe's Chaotic Exchange Rate Model," *Comput. Econ.,* 2024, doi: 10.1007/s10614-024-10546-z.
[24] Y. Ma, R. Mao, Q. Lin, P. Wu, and E. Cambria, "Multi-source aggregated classification for stock price movement prediction," *Information Fusion,* vol. 91, pp. 515-528, 2023, doi: 10.1016/j.inffus.2022.10.025.
[25] Y. Wu, Z. Fu, X. Liu, and Y. Bing, "A hybrid stock market prediction model based on GNG and reinforcement learning," *Expert Syst. Appl.,* vol. 228, p. 120474, 2023, doi: 10.1016/j.eswa.2023.120474.
[26] C. H. Wang, J. Yuan, Y. Zeng, and S. Lin, "A deep learning integrated framework for predicting stock index price and fluctuation via singular spectrum analysis and particle swarm optimization," *Applied Intelligence,* vol. 54, no. 2, pp. 1770-1797, 2024, doi: 10.1007/s10489-024-05271-x.
[27] Y. Huang, C. Zhou, K. Cui, and X. Lu, "A multi-agent reinforcement learning framework for optimizing financial trading strategies based on TimesNet," *Expert Syst. Appl.,* vol. 237, p. 121502, 2024, doi: 10.1016/j.eswa.2023.121502.
[28] Y. Han, J. Kim, and D. Enke, "A machine learning trading system for the stock market based on N-period Min-Max labeling using XGBoost," *Expert Syst. Appl.,* vol. 211, p. 118581, 2023, doi: 10.1016/j.eswa.2022.118581.
[29] M. M. Kumbure, C. Lohrmann, P. Luukka, and J. Porras, "Machine learning techniques and data for stock market forecasting: A literature review," *Expert Syst. Appl.,* vol. 197, p. 116659, 2022, doi: 10.1016/j.eswa.2022.116659.
[30] E. Drakonakis and S. Kotsios, "Stochastic Exchange Rate Dynamics, Intervention Dynamics and the Market Efficiency Hypothesis," *Comput. Econ.,* 2024, doi: 10.1007/s10614-024-10581-w.
[31] S. Carta, A. Ferreira, A. S. Podda, D. Reforgiato Recupero, and A. Sanna, "Multi-DQN: An ensemble of Deep Q-learning agents for stock market forecasting," *Expert Syst. Appl.,* vol. 164, p. 113820, 2021, doi: 10.1016/j.eswa.2020.113820.
[32] N. Sukma and C. S. Namahoot, "Enhancing Trading Strategies: A Multi-indicator Analysis for Profitable Algorithmic Trading," *Comput. Econ.,* vol. 65, no. 6, pp. 3807-3840, 2025, doi: 10.1007/s10614-024-10669-3.
[33] M. Ayitey Junior, P. Appiahene, O. Appiah, and C. N. Bombie, "Forex market forecasting using machine learning: systematic literature review and meta-analysis," *J. Big Data,* vol. 10, no. 1, p. 9, 2023, doi: 10.1186/s40537-022-00676-2.
[34] Y. F. Chen and S. H. Huang, "Sentiment-influenced trading system based on multimodal deep reinforcement learning," *Appl. Soft Comput.,* vol. 112, p. 107788, 2021, doi: 10.1016/j.asoc.2021.107788.
[35] R. Naranjo, J. Arroyo, and M. Santos, "Fuzzy modeling of stock trading with fuzzy candlesticks," *Expert Syst. Appl.,* vol. 93, pp. 15-27, 2018, doi: 10.1016/j.eswa.2017.10.002.
[36] A. Sadeghi, A. Daneshvar, and M. M. Zaj, "Combined ensemble multi-class SVM and fuzzy NSGA-II for trend forecasting and trading in forex markets," *Expert Syst. Appl.,* vol. 185, p. 115566, 2021, doi: 10.1016/j.eswa.2021.115566.
[37] M. G. Papatsimpas and K. E. Parsopoulos, "Hybrid decomposition and deep learning approach for data-driven FOREX forecasting optimization," *Journal of Global Optimization,* 2025, doi: 10.1007/s10898-025-01505-5.
[38] L. C. O. Tiong, D. C. L. Ngo, and Y. Lee, "Forex prediction engine: framework, modelling techniques and implementations," *International Journal of Computational Science and Engineering,* vol. 13, no. 4, pp. 364-377, 2016, doi: 10.1504/IJCSE.2016.080213.
[39] L. K. Felizardo and et al., "Outperforming algorithmic trading reinforcement learning systems: A supervised approach to the cryptocurrency market," *Expert Syst. Appl.,* vol. 202, p. 117259, 2022, doi: 10.1016/j.eswa.2022.117259.

[40] B. Gülmez, "Stock price prediction with optimized deep LSTM network with artificial rabbits optimization algorithm," *Expert Syst. Appl.,* vol. 227, p. 120346, 2023, doi: 10.1016/j.eswa.2023.120346.

[41] S. Deng, C. Xiao, Y. Zhu, J. Peng, J. Li, and Z. Liu, "High-frequency direction forecasting and simulation trading of the crude oil futures using Ichimoku KinkoHyo and Fuzzy Rough Set," *Expert Syst. Appl.,* vol. 215, p. 119326, 2023, doi: 10.1016/j.eswa.2022.119326.

[42] F. Kamalov and I. Gurrib, "Machine learning-based forecasting of significant daily returns in foreign exchange markets," *International Journal of Business Intelligence and Data Mining,* vol. 21, p. 465, 2022, doi: 10.1504/IJBIDM.2022.126505.

[43] A. Frattini, I. Bianchini, A. Garzonio, and L. Mercuri, "Financial technical indicator and algorithmic trading strategy based on machine learning and alternative data," *Risks,* vol. 10, no. 12, p. 225, 2022, doi: 10.3390/risks10120225.

[44] C. D. Yi, "Nonparametric Realized Volatility and Jumps in High Frequency Foreign Exchange Rates with a Gumbel Distribution," *Comput. Econ.,* 2025, doi: 10.1007/s10614-025-11153-2.

[45] E. Koo and G. Kim, "A New Neural Network Approach for Predicting the Volatility of Stock Market," *Comput. Econ.,* vol. 61, 2022, doi: 10.1007/s10614-022-10261-7.

[46] S. Dash, P. K. Sahu, and D. Mishra, "Enhancing Forex market trend analysis with bidirectional long short-term memory and transformer attention network," *J. Korean Stat. Soc.,* vol. 54, no. 3, pp. 825-857, 2025, doi: 10.1007/s42952-025-00322-6.