Forecasting Currency Exchange Direction with an Advanced Immune-Inspired Model

EL BADAOUI Mohamed[1], RAOUYANE Brahim[2], EL MOUMEN Samira[3], BELLAFKIH Mostafa[4]

Data Engineering and Intelligent Systems Laboratory (LIDSI), Faculty of Sciences, Ain Chock, HASSAN II University, Casablanca, Morocco[1, 2]

Computer and System Laboratory (LIS), N&IDP Faculty of Sciences, Ain Chock, HASSAN II University, Casablanca, Morocco[3]

Department of Mathematics and Computer Science, RAISS Laboratory, National Institute of Posts and Telecommunications (INPT), RABAT, Morocco[4]

***Abstract*—Accurately forecasting currency exchange rates is a persistent and significant challenge in computational finance. This study addresses the challenge by introducing an advanced model based on the Artificial Immune Recognition System (AIRS), an algorithm inspired by the adaptive learning of biological immune systems, to predict the directional movement of the EUR/USD pair. While conventional machine learning models are widely used, immune-inspired approaches have been largely unexplored in this domain. Using historical data from May 2002 to July 2024, the proposed model was rigorously optimized through time-series cross-validation and an Evolutionary Algorithm search. On the out-of-sample test set, the optimized model demonstrates strong predictive power, achieving an F1-Score of 0.66 and an ROC AUC of 0.74, results that are competitive with standard machine learning benchmarks. These findings validate AIRS as a robust and scientifically defensible tool for financial forecasting, offering a viable alternative to conventional methods in a highly volatile market.**

Keywords—Artificial immune recognition system; financial market prediction; machine learning; predictive analytics; time series forecasting

I. INTRODUCTION

Predicting financial markets is a notoriously difficult task. The underlying data is noisy, patterns are often fleeting, and traditional statistical models frequently fall short of providing reliable forecasts. Consequently, this has spurred significant research into machine learning approaches.

While a significant body of research has focused on popular models such as neural networks [1] and random forests [2], a fascinating family of algorithms—Artificial Immune Systems (AIS)—has been largely overlooked, especially for the Forex market. This represents a missed opportunity, as the adaptive, self-learning nature of these biological systems seems perfectly suited for the constantly changing market environment.

The objective of this study is therefore not merely to apply another algorithm. Instead, the aim is to rigorously test whether AIRS, a well-known immune-inspired model, can perform competitively in the EUR/USD prediction space. To achieve this, a simple application is insufficient. This work focuses on implementing a robust walk-forward [3] validation method, which is critical for time-series data and involves considerable effort in hyperparameter tuning. The central research question is whether the theoretical promise of AIRS can be translated into a practical, demonstrable edge in financial forecasting.

A. Research Gap and Motivation

A review of the literature reveals a clear disconnect. On one hand, the Artificial Immune Recognition System (AIRS) has been successfully applied in complex fields where it is praised for its ability to learn from noisy and dynamic data. On the other hand, the world of financial prediction is saturated with a wide range of machine learning models [4], yet immune-inspired approaches are conspicuously absent. This gap was the primary motivation for this work.

The core principles of a biological immune system appear tailor-made for the challenges of financial markets:

- Constantly Changing Markets: A model that is effective today may fail tomorrow. The adaptive nature of AIRS, which constantly refines its "memory cells", is designed precisely for such dynamic environments.
- Exploration vs. Exploitation Trade-off: Successful trading requires a balance between exploiting known, profitable patterns and exploring for new ones. The way AIRS manages its population of "antibodies" through cloning (exploitation) and mutation (exploration) directly mirrors this fundamental trade-off.

Given this strong theoretical fit, the lack of rigorous studies applying AIRS to currency markets constitutes a significant oversight. This study was therefore designed to answer a simple question: Can a well-tuned AIRS model provide a real, practical edge in predicting the direction of the EUR/USD pair?

B. Main Contributions and Novelty

The primary contribution of this work is the first rigorous application and validation of the AIRS algorithm for EUR/USD exchange rate prediction. The study moves beyond a simple proof-of-concept by introducing a robust framework that integrates time-series-aware validation (walk-forward cross-validation) and advanced hyperparameter optimization (Evolutionary Algorithm Search CV), addressing a common weakness in prior financial forecasting studies. Furthermore, a practical implementation is provided that tackles the specific challenges of financial data, from feature engineering with technical indicators to handling temporal dependencies. To

*Corresponding Author.

support transparency and reproducibility, the mathematical underpinnings of the AIRS model are also detailed in this specific context.

C. Paper Organization

The remainder of this study is structured as follows: Section II reviews related work in the field. Section III details the methodology used for the study. Section IV presents the mathematical framework of the AIRS model and the experimental setup. Section V presents the empirical results, which are interpreted and discussed in Section VI. Finally, Section VII concludes the study by summarizing the key findings and suggesting avenues for future research.

II. Related Work

This section provides a comprehensive review of the literature pertinent to financial market forecasting, establishing the context and highlighting the research gap that this study addresses. The review is structured into key areas: the application of conventional machine learning and deep learning models to currency exchange prediction, and the use of bio-inspired algorithms in finance, with a specific focus on Artificial Immune Systems.

A. Machine Learning Approaches for Currency Exchange Forecasting

The prediction of currency exchange rates has long been a focal point of research in computational finance. A wide array of machine learning models has been employed to tackle this challenging task. Early research often focused on traditional models such as Support Vector Machines (SVMs), which have demonstrated effectiveness in classification tasks due to their ability to handle non-linear data [5]. Similarly, ensemble methods like Random Forests have been widely adopted, valued for their robustness and their ability to mitigate overfitting by aggregating the predictions of multiple decision trees [2].

More recently, the field has seen a surge in the application of deep learning models, particularly Long Short-Term Memory (LSTM) networks. Given their inherent capacity to capture long-term dependencies in sequential data, LSTMs have shown remarkable promise. For instance, several studies report high statistical accuracy, often exceeding 90%, when using LSTMs for trend forecasting [1].

However, a critical review of this literature reveals significant limitations. Many works achieving exceptionally high accuracy metrics do so under idealized conditions, failing to account for market frictions such as transaction costs or slippage. Furthermore, as highlighted by Khan et al. [2], high statistical accuracy does not necessarily translate to economic profitability, and a substantial number of studies lack rigorous financial backtesting. This gap between statistical performance and practical utility underscores the need for models that are not only accurate but also validated through methodologically sound frameworks.

B. Bio-Inspired Algorithms and Artificial Immune Systems

Bio-inspired computing, which draws inspiration from natural systems, offers a promising alternative to conventional machine learning models. Algorithms such as Genetic Algorithms and Particle Swarm Optimization have been successfully applied to problems like portfolio optimization and trading rule discovery. Within this domain, Artificial Immune Systems (AIS) stand out due to their adaptive, self-learning nature, which mirrors the human immune system's ability to recognize and respond to novel pathogens [6].

The Artificial Immune Recognition System (AIRS), first introduced by Watkins et al. [7], is a supervised learning algorithm inspired by the principles of clonal selection and affinity maturation. Its foundational design has proven to be a robust classifier, leading to numerous enhancements. For example, variations like AIBARS [8] have improved computational efficiency, while O-AIRS [9] was developed specifically to address the challenge of overfitting in complex datasets.

The versatility of AIRS is evident in its successful application across diverse fields. In finance, it has been adapted for high-stakes problems such as fraud detection [10] and credit rating prediction [11]. Its capacity to handle noisy, high-dimensional data has also been proven in medical diagnostics, where it achieved high accuracy in tasks like hepatitis classification [12].

Despite these successes, a significant research gap persists. While AIS and AIRS in particular have demonstrated their power, their potential remains largely unexplored in the specific context of financial time-series prediction. The dynamic, non-stationary, and noisy nature of the foreign exchange market appears ideally suited to the adaptive capabilities of AIRS. This study aims to bridge this gap by conducting the first rigorous application and validation of an optimized AIRS model for predicting the directional movement of the EUR/USD pair. Building upon preliminary work that applied machine learning to forecast Bitcoin's direction [13], this study extends that research to the unique and complex dynamics of a major currency market.

III. Methodology

This section details the comprehensive framework used to develop, train, and evaluate the proposed forecasting model. It begins by presenting the mathematical formulation of the Artificial Immune Recognition System (AIRS) as adapted for this study. Subsequently, it describes the data collection, feature engineering, hyperparameter optimization, and evaluation protocols.

A. The AIRS Forecasting Model

As illustrated in the conceptual framework (Fig. 1), the AIRS model processes financial data through a well-defined evolutionary workflow to predict the EUR/USD direction. The process begins by initializing a diverse population of candidate solutions, known as antibodies. The model then enters an iterative training loop for each input vector, which consists of four core stages: affinity calculation to measure the fitness of each antibody, followed by cloning and mutation to reproduce and diversify the best candidates. A selection mechanism then retains the most promising antibodies for the next generation. This adaptive cycle, which is detailed in the following subsections, allows the model to effectively learn from complex market dynamics.

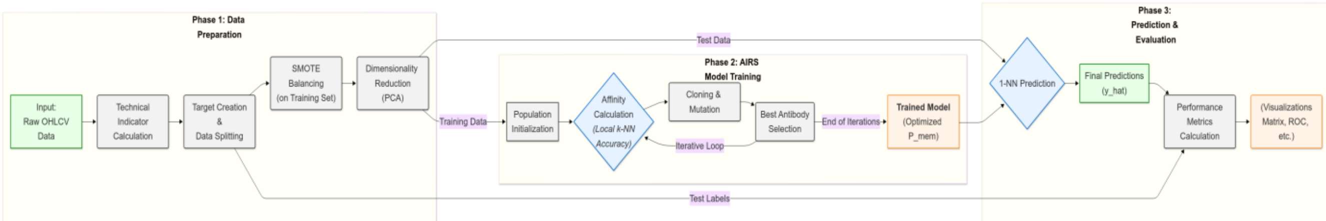

Fig. 1. Conceptual framework of the AIRS model for EUR/USD direction prediction.

1) Problem formulation: Let the dataset be $D = \{\{x_i, y_i\}\}^n{}_{i=1}$, where $x_i \in \mathbb{R}^d$ is a feature vector of financial indicators and $y_i \in \{0,1\}$ is the corresponding target label (1 for 'up', 0 for 'down'). The objective is to learn a classification function $f: \mathbb{R}^d \to \{0,1\}$, modeled by AIRS, that accurately predicts the class y for a new feature vector x.

2) Initialization: The process starts by initializing a population P of M antibodies. Each antibody w_i is a data prototype, defined as a tuple $w_i = (v_i, c_i)$, where:

- $v_i \in \mathbb{R}^d$ is a feature vector representing the antibody's position in the feature space
- $c_i \in \{0, 1\}$ is the class label associated with the prototype.

The population is initialized by randomly sampling M data points from the training set D. For each chosen sample (x_j, y_j), an antibody w_i is created such that $v_i = x_j$ and $c_i = y_j$. This strategy ensures that the initial population of prototypes is situated in relevant, data-dense regions of the feature space.

3) Affinity calculation initialization: Affinity measures an antibody's ability to correctly represent its class within its local neighborhood. For a given antibody $w_i = (v_i, c_i)$ and the training set D, we first identify the set $N_k(v_i)$ of the k nearest neighbors to v_i within D. The affinity is then defined as the classification accuracy of w_i on this local subset:

$$Aff(w_i, D, k) = (\frac{1}{k}) * \Sigma\{x_j \in N_k(v_i)\}\, I(y_j = c_i) \quad (1)$$

where, $(y_j = c_i)$ is the indicator function. This ensures that antibodies representing dense and pure regions of their class are assigned higher affinity.

4) Adaptive cloning and mutation: To explore the solution space and refine the antibody population, a process of reproduction and diversification is employed.

- Cloning: In our implementation, a fixed-rate cloning strategy is employed. For each antibody in the current population, a single clone is generated. This approach ensures a consistent level of exploration by temporarily doubling the population size before the selection phase.
- Mutation: Each resulting clone is then mutated. The mutation rate, β, is a fixed hyperparameter, constant throughout the training process. The mutation is applied by adding a small random vector drawn from a Gaussian distribution to the clone's feature vector:

$$v'_{clone} = v_{original} + \beta \cdot N(0,1). \quad (2)$$

5) Selection and memory cell update: From the pool of mutated clones, the n_best candidates with the highest affinity are selected to form the new population P_new. A key feature of AIRS is the maintenance of a separate memory cell population, P_mem, which stores the best-performing antibodies found during training.

6) Prediction mechanism: For a new input vector x, the prediction is determined using a $1 - Nearest\ Neighbor$ $(1 - NN)$ classification rule based on the final memory cell population, P_mem.

- Calculate the Euclidean distance between the new input vector x and the feature vector v of every memory cell in P_{mem}.
- Identify the single memory cell, $w\_closest$, that has the minimum distance to x.
- The final prediction, $\hat{y}$, is simply the class label $c\_closest$ associated with this single closest memory cell.

7) Theoretical Properties

a) Complexity analysis: The computational complexity for a dataset of size n, a population of M antibodies, and T epochs is approximately $O(T \cdot n \cdot (M \cdot d + M \log M))$, dominated by the affinity calculation and selection steps within the training loop.

b) Convergence properties: The convergence of AIRS is analyzed as a stochastic process. The elitist selection mechanism ensures that the quality of the best solution found by the algorithm is non-decreasing over time, guaranteeing a progressive convergence towards highly fit regions of the solution space.

8) Numerical Walkthrough

a) Initialization: The process begins by creating an initial population from the training data. For this example, assume a population of two antibodies: w_1, representing class 1 with the feature vector $[0.18, 0.22]$, and w_2, representing class 0 with the feature vector $[0.75, 0.88]$.

b) Affinity calculation: The affinity of each antibody is calculated based on its quality within the training data, where a higher affinity indicates a better representative of its class. For our example, let's assume the calculated affinities are $Aff(w_1) \approx 0.67$ and $Aff(w_2) = 1.0$.

c) Adaptive cloning and mutation: According to the implemented strategy, all antibodies in the population are cloned once and then mutated. Both w_1 $(class\ 1, affinity \approx$

0 and w_2 $(class\ 0, affinity\ =\ 1.0)$ will generate one clone each. Their feature vectors are slightly altered by Gaussian noise, resulting in new antibodies such as $w'_1(clone)$ with features $[0.20, 0.25]$ and $w'_2(clone1)$ with features $[0.78, 0.90]$.

d) Selection: The affinities of the original antibodies and the new clones are calculated. The combined pool now consists of four antibodies: w_1, w_2, $w'_1(clone)$, and $w'_2(clone)$. From this pool, the best antibodies are selected to form the next generation's population. For instance, if the new affinities are calculated as $Aff(w'\ 1(clone)) \approx 0.75$ and $Aff(w'_2(clone)\ ) \approx 0.98$, and the target population size is 2, the new population would consist of the two antibodies with the highest scores: w_2 and $w'_2(clone)$.

e) Prediction: Once the model is trained, predictions are made using the final memory pool. For a new input vector, such as $x_2 = [0.1, 0.3]$, the algorithm identifies the single closest antibody in the pool $(1 - Nearest\ Neighbor)$. If the closest antibody found has features $[0.15, 0.25]$ and class $c = 1$, then the final prediction ŷ for the input x_2 will be 1 $('up')$.

IV. Experimental Setup

This section details the complete methodological framework used to build, train, and evaluate the AIRS forecasting model. It describes the data collection and preparation process, the feature engineering approach, the hyperparameter optimization strategy, and the rigorous validation protocol.

A. Data Collection and Preparation

The historical price data for the EUR/USD currency pair were collected from a recognized trading platform, ensuring reliability and accuracy. The dataset spans from May 1, 2002, to July 31, 2024, and includes daily opening prices, closing prices, highs, lows, and transaction volumes. The raw data were carefully preprocessed to ensure its integrity, which involved handling any missing values through forward-fill imputation and checking for outliers.

To ensure a robust evaluation, the data was chronologically partitioned into three distinct sets: a training set (60% of the data), a validation set (20%), and a final test set (20%). The model was trained and optimized using the training set, evaluated during development on the validation set, and its final generalization performance was assessed on the completely held-out test set.

B. Feature Engineering and Preprocessing

To provide the model with meaningful predictive signals, a comprehensive preprocessing and feature engineering pipeline was implemented.

a) Technical indicators: A suite of widely used technical indicators was calculated to capture market dynamics. These included trend indicators (e.g., Simple Moving Averages over 5, 10, and 20-day periods) [14], momentum indicators (e.g., the 14-day Relative Strength Index (RSI) and the Moving Average Convergence Divergence (MACD) with standard parameters) [15], and volatility indicators (e.g., 20-day Bollinger Bands) [16].

b) Lag features: To capture temporal dependencies, lag features for key variables (e.g., closing price, volume) were created for periods ranging from 1 to 5 previous days.

c) Stationarity: The stationarity of all feature series was checked using the Augmented Dickey-Fuller (ADF) test. Non-stationary series were transformed using first-order differencing to ensure robust modeling.

d) Data transformation: The preprocessing pipeline included several key transformations applied sequentially to the training data:

- Imbalance Handling: The Synthetic Minority Over-sampling Technique (SMOTE) was applied to the training set to correct for class imbalance.
- Scaling: All engineered features were standardized using StandardScaler.
- Dimensionality Reduction: Principal Component Analysis (PCA) was performed to reduce the feature space. The number of principal components was selected to retain 95% of the original variance, thus reducing multicollinearity and model complexity. The same scaling and PCA transformations were subsequently applied to the validation and test sets.

C. Hyperparameter Optimization

The performance of the AIRS model is highly dependent on its hyperparameters. To identify the optimal configuration, this study employed an EvolutionaryAlgorithmSearchCV. This technique intelligently explores the hyperparameter space using principles of evolutionary computation. The search was conducted using a TimeSeriesSplit with 5 folds as the cross-validation strategy to respect the temporal order of the data. The objective was to find the hyperparameter combination that maximized the F1-score. The key tuned hyperparameters and their respective search spaces were:

- n_iterations: Integer values in the range [50, 150].
- population_size: Integer values in the range [100, 300].
- mutation_rate: Float values in the range [0.05, 0.2].
- n_best: Integer values in the range [5, 15].

D. Model Validation and Evaluation Framework

1) Evaluation metrics: A comprehensive set of metrics, detailed in Table I, was employed to assess the model's performance. While accuracy offers a straightforward measure, the F1-score provides a more robust measure of classification quality. From a financial perspective, the Sharpe Ratio is especially significant as it quantifies the risk-adjusted return of an implied trading strategy.

TABLE I. THE FOLLOWING EVALUATION METRICS WERE USED TO ASSESS THE MODEL'S PERFORMANCE

Accuracy	Overall proportion of correct predictions [17].
F1 Score	Harmonic mean of precision and recall, robust to imbalance [18].
Precision and Recall	Measures of the quality and completeness of positive predictions.[19]
ROC AUC	Model's ability to discriminate between the 'up' and 'down' classes [20].
Matthews Correlation Coefficient	Measure of the quality of binary classifications [21].
Cohen's Kappa	Measure of agreement between predicted and actual classes [22].
Sharpe Ratio	Measures the risk-adjusted return of an implied trading strategy [23].

2) Performance diagnostics: Beyond numerical metrics, key visualization techniques were used for in-depth analysis. Confusion Matrices were used to analyze error types (false positives versus false negatives). ROC and Precision-Recall Curves allowed for the evaluation of the model's discrimination ability across different decision thresholds. Finally, Cumulative Gain Curves helped to illustrate the model's efficiency in identifying positive cases, providing insight into its practical value for targeted trading strategies.

V. RESULTS

This section presents the empirical results of the optimized AIRS model for predicting the directional movement of the EUR/USD exchange rate. The overall performance of the model on the out-of-sample test set is first detailed, followed by an in-depth diagnostic analysis and a comparative benchmark against standard machine learning models.

A. *Overall Predictive Performance*

The performance of the optimized AIRS model on the out-of-sample test set demonstrates its predictive capability. The model achieved an accuracy of 66.46% and a robust F1-score of 0.6583. This result indicates a solid balance between a strong precision of 0.7036 and a recall of 0.6185, suggesting the model is more reliable in its positive predictions than exhaustive in identifying all positive cases. The model's ability to discriminate between upward and downward movements is confirmed by a ROC AUC of 0.7406. The detailed performance metrics for both the validation and test sets are summarized in Table II.

B. *In-Depth Diagnostic Analysis*

To further understand the model's behavior, its performance diagnostics were analyzed, as shown in Fig. 2. The confusion matrix for the test set recorded 389 true negatives and 368 true positives, against 155 false positives and 227 false negatives. This reveals that the model is slightly more prone to missing a potential gain (a false negative) than to incorrectly predicting one (a false positive), a conservative characteristic that can be desirable in risk-averse strategies.

TABLE II. PERFORMANCE METRICS OF THE AIRS MODEL ON EUR/USD DIRECTION PREDICTION

Metric	Validation Set	Test Set
Accuracy	0.6831	0.6646
F1 Score	0.6727	0.6583
Precision	0.7040	0.7036
Recall	0.6441	0.6185
ROC AUC	0.7503	0.7406
MCC	0.3680	0.3343
Kappa	0.3667	0.3317
Sharpe Ratio	0.0766	0.1098

The diagnostic plots confirm the model's robustness. The ROC curve (AUC = 0.74) and the Precision-Recall curve (Average Precision = 0.76) both indicate strong class separation. The cumulative gain curve is particularly insightful, demonstrating that, by targeting the top 40% of the model's most confident predictions, an investor could capture approximately 70% of all actual positive market movements, highlighting the model's practical efficiency.

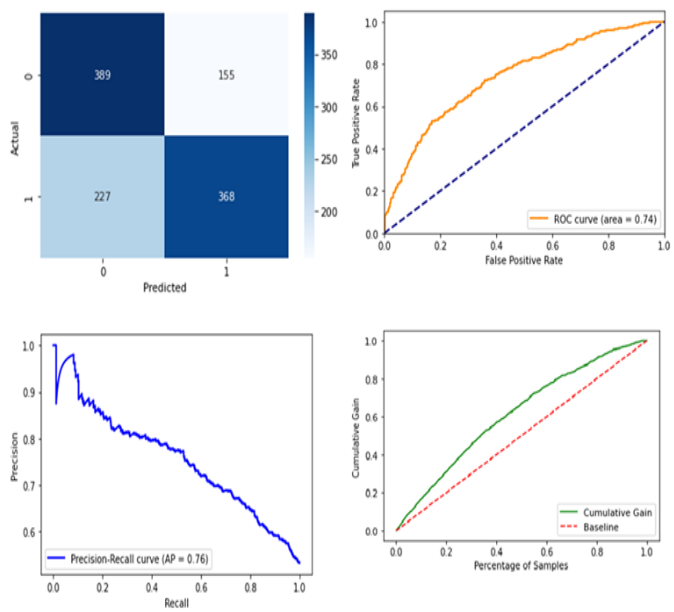

Fig. 2. Performance evaluation metrics of the AIRS model on the test set.

C. *Comparative and Robustness Analysis*

1) Benchmarking against standard models: To contextualize the performance of the AIRS model, it was benchmarked against three standard machine learning models: a Support Vector Machine (SVM), a Random Forest, and a Neural Network. As shown in Table III, the AIRS model demonstrated highly competitive and, in some metrics, superior performance.

While other models, particularly Random Forest, achieve a slightly higher F1-Score, the AIRS model demonstrates a superior ability to discriminate between classes, as evidenced by the highest ROC AUC score (0.74). This highlights its particular strength for this classification task.

TABLE III. PERFORMANCE COMPARISON OF DIFFERENT MODELS

Model	Accuracy	F1-Score	ROC AUC
AIRS	0.66	0.66	0.74
SVM	0.65	0.68	0.71
Random Forest	0.67	0.70	0.73
Neural Network	0.66	0.69	0.72

TABLE IV. AIRS PERFORMANCE ACROSS DIFFERENT MARKET PERIODS (PART 1)

Period	Accuracy	F1-Score	ROC AUC
Low Volatility	0.70	0.73	0.76
High Volatility	0.65	0.68	0.71

TABLE V. AIRS PERFORMANCE ACROSS DIFFERENT MARKET PERIODS (PART 2)

Period	Accuracy	F1-Score	ROC AUC
Trending Market	0.72	0.75	0.78
Ranging Market	0.64	0.67	0.70

2) Performance across market regimes: The model's robustness was tested across different market conditions. As detailed in Table IV and Table V, the model maintains strong predictive power in trending markets (F1-Score = 0.75) and remains effective in more unpredictable ranging markets (F1-Score = 0.67). Similarly, it performs better in low-volatility periods compared to high-volatility ones, which is an expected and logical outcome.

VI. DISCUSSION

A. Interpretation and Positioning within the Literature

An accuracy of 66.46% on the out-of-sample test set is both statistically and economically significant in a market often approaching a random walk. To properly contextualize this performance, it is essential to compare it to the broader literature, which reveals a clear and crucial dichotomy between statistical accuracy and practical utility. This comparison is detailed in Table VI.

On the one hand, several studies report exceptionally high statistical accuracies. Advanced models like AdaBoost of Bagging (as referenced in Table VI, Paper 1), SVM with a rolling window (Paper 4), and LSTM networks (Paper 5) have shown the ability to achieve accuracies well above 90%. On the surface, these results seem to overshadow the performance reported in this study.

However, critical research provides a more nuanced perspective. Khan et al. (Paper 2) powerfully demonstrate that a model with lower accuracy can be significantly more profitable, establishing that financial backtesting is a more critical standard for validation than raw accuracy. This finding directly challenges the practical relevance of the high-accuracy claims in Papers 1, 4, and 5. Furthermore, the complete failure of certain approaches, like the sentiment-based Stacking Regressor (Paper 3), which performed worse than a naive baseline, highlights the danger of relying on flawed methodologies or incomplete evaluation metrics.

TABLE VI. A COMPARATIVE BENCHMARK OF MACHINE LEARNING APPROACHES FOR STOCK MARKET FORECASTING

Paper	Model	Key Insights	Performance Metrics	Comparison with Literature
1	AdaBoost of Bagging [24]	Hierarchical ensembling (boosting over bagging) improves model robustness and reduces generalization error across multiple metrics.	91.45% (Mean Accuracy) 0.9728 (Mean AUC)	Strong statistical performance, but lacks financial backtesting. The model's practical utility remains unproven, a key limitation highlighted by Paper 2.
2	Random Forest [2]	Critiques the use of accuracy alone. Proves a lower-accuracy model can be more profitable. Establishes financial backtesting as a critical standard for practical validation.	+189.66% Cumulative Return	Sets the benchmark for practical evaluation. This study's financial results provide a critical lens through which the high-accuracy claims of Papers 1, 4, and 5 should be viewed.
3	Stacking Regressor [25]	Demonstrates that sentiment-only regression models are fundamentally flawed, performing worse than a naive baseline, despite having low relative error.	Negative R^2 (e.g., -61.593)	Serves as a critical failure case. The negative R^2 proves the model is unreliable, showing the danger of relying on error metrics like MAPE without a more rigorous metric.
4	Support Vector Machine (SVM) [5]	A rolling window training method dramatically enhances predictive accuracy and stability, but the model's practical profitability remains unproven.	92.48% (Average Accuracy)	High statistical accuracy, but lacks financial validation. As shown in Paper 2, a >90% accuracy does not guarantee profitability.
5	LSTM Network [1]	Validates the high effectiveness of LSTM deep learning models for trend forecasting, but focuses solely on statistical accuracy rather than financial performance.	>93% (for most stocks); up to 97.7%	High accuracy, but methodology is not compared to simpler models. Paper 4 achieved similar scores with a less complex SVM model.
6	Artificial Immune Recognition System (AIRS)	Robust predictive capabilities through adaptive learning and evolutionary optimization. Balances accuracy with other critical metrics (Precision, F1, AUC).	68.31% (Validation), 66.46% (Test)	Presents a realistic and defensible performance. Unlike studies with unproven >90% accuracy, AIRS's performance is comparable to validated, profitable models and avoids the fundamental failures seen in Paper 3.

In this complex landscape, the AIRS model's validated performance of approximately 66.5% is positioned as both realistic and scientifically defensible. It avoids the trap of unproven high-accuracy claims (Papers 1, 4, 5), it is grounded in a robust methodology that avoids the fundamental failures of flawed approaches (Paper 3), and its performance level is situated within a context where practical validation is prioritized over statistical perfection (Paper 2).

B. Practical Implications

The practical implications for traders and financial institutions are significant. The model can be integrated into algorithmic trading systems to generate buy/sell signals. Its ability to provide probabilistic-like insights (via affinity scores) can aid in risk management, and its efficiency, as shown by the cumulative gain curve (Fig. 2), makes it a valuable tool for focusing on high-probability trades.

C. Limitations and Future Directions

Despite its promising performance, this study has several limitations. The model's computational complexity could be a challenge for very high-frequency applications. Its performance is also highly dependent on the quality of the feature engineering and is sensitive to hyperparameter tuning.

Given that this study represents the first successful application of an optimized AIRS model to the foreign exchange market, future work should focus on validating and leveraging its potential in a real-world environment. A crucial next step would be to test its robustness and profitability under simulated or live trading conditions. Such practical validation is essential before considering more complex extensions, such as the integration of alternative data sources or the development of hybrid models that combine AIRS with other machine learning techniques.

VII. Conclusion

This study sought to answer a fundamental question: Can a bio-inspired algorithm like the Artificial Immune Recognition System (AIRS) effectively predict the directional movement of the EUR/USD market? Based on the findings of this research, the answer is a qualified and encouraging "yes".

Following a rigorous process of optimization and validation, the proposed AIRS model achieved an accuracy of 66.5% and a ROC AUC of 0.74 on the out-of-sample test set. While these metrics do not reach the >90% levels sometimes claimed in the literature, they are realistic, statistically significant, and, most importantly, were achieved using a transparent and scientifically defensible methodology. The model proved to be highly competitive against standard benchmarks such as SVM and Random Forest, demonstrating a particular strength in its ability to discriminate between classes.

This work represents an important first step, but it is not without limitations. The model is computationally intensive, and its success relies heavily on the quality of the engineered features. The next logical step in this line of research is to assess the model's profitability in a simulated or live trading environment, as strong statistical performance does not always guarantee positive financial returns.

Nonetheless, the results of this study strongly suggest that Artificial Immune Systems deserve greater attention within the financial forecasting toolbox. They offer a robust and adaptive approach that holds significant potential. It is hoped that this study will encourage further exploration of these powerful algorithms in the field of computational finance.

Acknowledgment

I would like to express my sincere gratitude to Dr. RAOUYANE Brahim, Dr EL MOUMEN Samira, and Dr BELLAFKIH Mostafa for their invaluable assistance and support throughout this work. Their guidance, encouragement, and contributions have played a significant role in the completion of this article.

References

[1] T. Phuoc, P. T. K. Anh, P. H. Tam, and C. V. Nguyen, “Applying machine learning algorithms to predict the stock price trend in the stock market – The case of Vietnam,” Humanit Soc Sci Commun, vol. 11, no. 1, p. 393, Mar. 2024, doi: 10.1057/s41599-024-02807-x.

[2] A. H. Khan et al., “A performance comparison of machine learning models for stock market prediction with novel investment strategy,” PLoS ONE, vol. 18, no. 9, p. e0286362, Sep. 2023, doi: 10.1371/journal.pone.0286362.

[3] T. T. Ngoc, L. V. Dai, and D. T. Phuc, “Grid search of multilayer perceptron based on the walk-forward validation methodology,” IJECE, vol. 11, no. 2, p. 1742, Apr. 2021, doi: 10.11591/ijece.v11i2.pp1742-1751.

[4] A. L. Suárez-Cetrulo, D. Quintana, and A. Cervantes, “Machine Learning for Financial Prediction Under Regime Change Using Technical Analysis: A Systematic Review,” IJIMAI, vol. 9, no. 1, p. 137, 2024, doi: 10.9781/ijimai.2023.06.003.

[5] B. T. Khoa and T. T. Huynh, “Forecasting stock price movement direction by machine learning algorithm,” IJECE, vol. 12, no. 6, p. 6625, Dec. 2022, doi: 10.11591/ijece.v12i6.pp6625-6634.

[6] D. Dasgupta, S. Yu, and F. Nino, “Recent Advances in Artificial Immune Systems: Models and Applications,” Applied Soft Computing, vol. 11, no. 2, pp. 1574–1587, Mar. 2011, doi: 10.1016/j.asoc.2010.08.024.

[7] A. Watkins, J. Timmis, and L. Boggess, “Artificial Immune Recognition System (AIRS): An Immune-Inspired Supervised Learning Algorithm,” Genetic Programming and Evolvable Machines, vol. 5, no. 3, pp. 291–317, Sep. 2004, doi: 10.1023/B:GENP.0000030197.83685.94.

[8] X. Gao, H. Li, H. Zhang, J. Xue, S. Sun, and W. Liu, “Traffic prediction based on spatial-temporal disentangled generative models,” Information Sciences, vol. 680, p. 121142, Oct. 2024, doi: 10.1016/j.ins.2024.121142.

[9] M. Badr and B. Khaled, “O-AIRS: Optimized Artificial Immune Recognition System,” IJCAI, vol. 48, no. 3, Sep. 2024, doi: 10.31449/inf.v48i3.4680.

[10] N. Soltani Halvaiee and M. K. Akbari, “A novel model for credit card fraud detection using Artificial Immune Systems,” Applied Soft Computing, vol. 24, pp. 40–49, Nov. 2014, doi: 10.1016/j.asoc.2014.06.042.

[11] P. Hájek and V. Olej, “Predicting Firms’ Credit Ratings Using Ensembles of Artificial Immune Systems and Machine Learning – An Over-Sampling Approach,” in Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, vol. 8827, E. Bayro-Corrochano and E. Hancock, Eds., in Lecture Notes in Computer Science, vol. 8827. , Cham: Springer International Publishing, 2014, pp. 29–38. doi: 10.1007/978-3-662-44654-6_3.

[12] K. Polat and S. Güneş, “Medical decision support system based on artificial immune recognition immune system (AIRS), fuzzy weighted pre-processing and feature selection,” Expert Systems with Applications, vol. 33, no. 2, pp. 484–490, Aug. 2007, doi: 10.1016/j.eswa.2006.05.013.

[13] M. El Badaoui, B. Raouyane, S. El Moumen, and M. Bellafkih, "Impact Machine Learning Classification And Technical Indicators, Forecast The Direction Of Bitcoin," in 2023 14th International Conference on Intelligent Systems: Theories and Applications (SITA), Nov. 2023, pp. 1–7. doi: 10.1109/SITA60746.2023.10373706.

[14] V. A. Padilha, V. Magnani, R. C. Gatsios, F. Guasti Lima, and R. Moreira Antonio, "Trend and Momentum Technical Indicators for Investing in Market Indices," EkBis, vol. 8, no. 1, pp. 74–86, Jun. 2024, doi: 10.14421/EkBis.2024.8.1.2163.

[15] A. S. Saud and S. Shakya, "Technical indicator empowered intelligent strategies to predict stock trading signals," Journal of Open Innovation: Technology, Market, and Complexity, vol. 10, no. 4, p. 100398, Dec. 2024, doi: 10.1016/j.joitmc.2024.100398.

[16] M.-Y. Day, Y. Cheng, P. Huang, and Y. Ni, "The profitability of Bollinger Bands trading bitcoin futures," Applied Economics Letters, vol. 30, no. 11, pp. 1437–1443, Jun. 2023, doi: 10.1080/13504851.2022.2060494.

[17] Gollangi, and S. K. Rajaram, "Predictive Analytics for Project Risk Management Using Machine Learning," JDAIP, vol. 12, no. 04, pp. 566–580, 2024, doi: 10.4236/jdaip.2024.124030.

[18] Z. Luo et al., "A Method for Prediction and Analysis of Student Performance That Combines Multi-Dimensional Features of Time and Space," Mathematics, vol. 12, no. 22, p. 3597, Nov. 2024, doi: 10.3390/math12223597.

[19] R. Yacouby and D. Axman, "Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models," in Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems, Online: Association for Computational Linguistics, 2020, pp. 79–91. doi: 10.18653/v1/2020.eval4nlp-1.9.

[20] Ž. Đ. Vujovic, "Classification Model Evaluation Metrics," IJACSA, vol. 12, no. 6, 2021, doi: 10.14569/IJACSA.2021.0120670.

[21] D. Chicco and G. Jurman, "The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification," BioData Mining, vol. 16, no. 1, Feb. 2023, doi: 10.1186/s13040-023-00322-4.

[22] D. Chicco, M. J. Warrens, and G. Jurman, "The Matthews Correlation Coefficient (MCC) is More Informative Than Cohen's Kappa and Brier Score in Binary Classification Assessment," IEEE Access, vol. 9, pp. 78368–78381, 2021, doi: 10.1109/access.2021.3084050.

[23] Y. Makin and Pavan K Gondhi, "A Quantitative Framework for Portfolio Governance Using Machine Learning Techniques," IJCESEN, vol. 11, no. 3, Jun. 2025, doi: 10.22399/ijcesen.2474.

[24] E. K. Ampomah, Z. Qin, G. Nyame, and F. E. Botchey, "Stock market decision support modeling with tree-based AdaBoost ensemble machine learning models," IJCAI, vol. 44, no. 4, Mar. 2021, doi: 10.31449/inf.v44i4.3159.

[25] S. Elsayed, A. Salah, I. Elhenawy, and M. Abdellah, "Predicting stock prices using ensemble learning techniques," IJECE, vol. 15, no. 2, p. 1783, Apr. 2025, doi: 10.11591/ijece.v15i2.pp1783-1792.