Economic Viability of a Machine-Learning Expert Advisor for Small-Capital Retail Foreign Exchange Trading: Evidence from a Ten-Year EUR/USD Backtest

Zulkifli Bin Husin[1]*, Muhammad Amir Hakim Bin Ismail[2], Tan Wei Keong[2], Muhammad Luqman Bin Yasruddin[2]

[1]Faculty of Electronic Engineering & Technology, Universiti Malaysia Perlis, Arau, Perlis, Malaysia

[2]Faculty of Intelligent Computing, Universiti Malaysia Perlis, Arau, Perlis, Malaysia

***Corresponding Author**

DOI: https://doi.org/10.47772/IJRISS.2026.1015EC0111

Received: 23 October 2026; Accepted: 28 October 2026; Published: 03 November 2026

ABSTRACT

Retail participation in the foreign exchange (FX) market has grown with low-cost trading platforms, yet small-capital traders face high volatility, limited analytical capacity and emotional decision-making. Automated Expert Advisors (EAs) on MetaTrader promise disciplined execution, but evidence on their economic outcomes at very small account sizes is scarce. This study evaluates the economic performance and risk of an EA that converts four Technical Indicators, namely Candlestick Patterns (CP), Moving Average (MA), Relative Strength Index (RSI) and Average Directional Index (ADX), into trading decisions using a Support Vector Machine (SVM), a Probabilistic Neural Network (PNN) and Naive Bayes (NB) classifiers fused by a Fuzzy Inference System (FIS). EUR/USD hourly data from 2013 to 2023 were used, with a chronological 70/30 split and rolling-window cross-validation for model development and a held-out 2020–2022 test set. Economic outcomes were assessed with MetaTrader 4 (MT4) backtests in which each year from 2013 to 2022 started with a USD 30 balance. The integrated model reached 99.34% test accuracy in reproducing rule-defined trading signals. In the backtests, every year closed with positive net profit, from USD 35.10 to USD 69.90 (return on investment, ROI, of 117.0% to 233.0%), giving a cumulative USD 464.54 and an average annual ROI of 154.8%. However, profit factors were only 1.18 to 1.31, relative drawdown exceeded 30% in three years and reached 37.5% in 2019, and exit management relied on a trailing stop whose distance was fixed rather than optimised. The findings indicate that algorithmic decision support can generate positive simulated returns for small retail accounts, but that headline ROI overstates the risk-adjusted benefit. Retail traders, educators and regulators should weigh drawdown and position-scaling risk alongside return figures when assessing automated trading products.

Keywords: foreign exchange, retail trading, Expert Advisor, machine learning, risk management

INTRODUCTION

The foreign exchange (FX) market is the largest and most liquid financial market, and its turnover has increased further in periods of elevated volatility [10]. Exchange rates are, however, notoriously hard to predict [18], and the speed and volume of market data make manual analysis difficult for individual traders [12]. These conditions weigh most heavily on retail participants with small accounts, who typically lack institutional tools and are exposed to behavioural biases in discretionary trading [5, 14].

Automated trading systems, implemented on the MetaTrader 4 (MT4) platform as Expert Advisors (EAs), execute rule-based decisions without human intervention and are increasingly used by both institutions and individuals [16]. A growing literature combines Technical Indicators with machine learning to improve such decisions [4, 17, 21], but systematic reviews note that most studies report predictive accuracy without translating predictions into end-to-end, executable trading outcomes [7, 24]. From an economic standpoint, the relevant

question is not only whether a model classifies signals accurately, but what returns and risks a trader would actually face.

This study addresses that gap from the perspective of a small-capital retail trader. Building on our earlier EA prototypes [5, 14], we evaluate an EA that fuses three machine learning classifiers, namely a Support Vector Machine (SVM), a Probabilistic Neural Network (PNN) and Naive Bayes (NB), across four Technical Indicators, namely Candlestick Patterns (CP), Moving Average (MA), Relative Strength Index (RSI) and Average Directional Index (ADX), through a Fuzzy Inference System (FIS), and we examine its economic outcomes over ten years of EUR/USD data with a USD 30 starting balance. The contributions are: (i) a transparent, reproducible decision pipeline from price data to executable EA action; (ii) a ten-year, year-by-year assessment of returns, profit factor and drawdown at a realistic retail account size; and (iii) a critical discussion of how headline return on investment (ROI) figures can misrepresent risk for retail traders.

LITERATURE REVIEW

Classical evidence shows that exchange rates are difficult to forecast with structural models, and that apparent predictability is often unstable across periods [18]. Machine learning methods have therefore attracted attention in finance for their ability to capture non-linear patterns [12, 25]. Reviews of FX forecasting with machine learning report a wide spread of methods and data choices, with Technical Indicators remaining the most common inputs [4, 21].

On the trading side, multi-model and multi-agent systems have been proposed to support FX decisions [11, 19], and neural-network trading systems have been evaluated with economic metrics [15]. Yet reviews find that end-to-end automation, realistic cost modelling and risk reporting remain inconsistent [7, 24], and that evaluation is often vulnerable to look-ahead leakage [2, 26]. Our prior work showed that an EA combining three CP with SVM could classify EUR/USD buy and sell signals with 91.25% test accuracy [5], and that MA and RSI inputs with a feed-forward neural network could forecast GBP/USD H4 movements with a 79.70% total success rate [14]; the present study extends this to four indicator streams, three classifiers and a fuzzy-fused decision layer, and centres the analysis on economic outcomes and risk for small retail accounts.

Beyond finance, our group has used learned classifiers to replace manual, error-prone expert grading, including beef quality assessment from muscle colour [22], tomato maturity grading from skin colour [23] and fish disease detection [28] using computer vision with deep neural networks, and has developed an automated real-time water quality monitoring and control system in which sensor readings are processed by a microcontroller that actuates the controller automatically [27]. Together, these studies motivate the design used here: a classifier replaces subjective human judgement, and an automated controller acts on its output, which is the structure an EA requires when it turns indicator readings into trade orders.

METHODOLOGY

Research Design

The study follows an end-to-end design in which historical prices are converted into indicator features, classified by three learning models, fused into a single decision and executed by an MT4 EA (Figure 1). The study is restricted to the three classifiers evaluated here, namely SVM, PNN and NB; all results reported in this paper are produced by these three classifiers and by their FIS fusion.

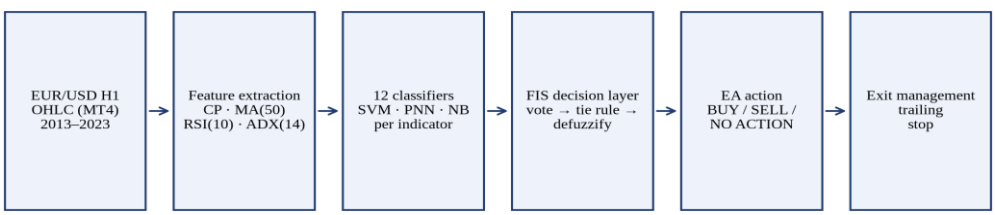


Figure 1: Proposed pipeline from MT4 price data to EA action.

Data and validation protocol

Historical EUR/USD one-hour (H1) OHLC data were exported from the MT4 History Center for 1 January 2013 to 1 January 2023, giving 87,600 hourly records. EUR/USD was selected because it is the most heavily traded currency pair [10]; the study is restricted to one pair and one timeframe to keep the experimental environment controlled. The series was split chronologically: approximately 70% (1 January 2013 to 31 December 2019) for model development and 30% (1 January 2020 to 1 January 2023) as a held-out test set used only for final reporting. Within the development segment, a rolling-window time-series cross-validation with ten folds was used (Figure 2): each fold trains on 11,600 hours and validates on the following 4,972 hours, and the origin then moves forward by 4,972 hours, so no later observation informs an earlier model [2, 26].

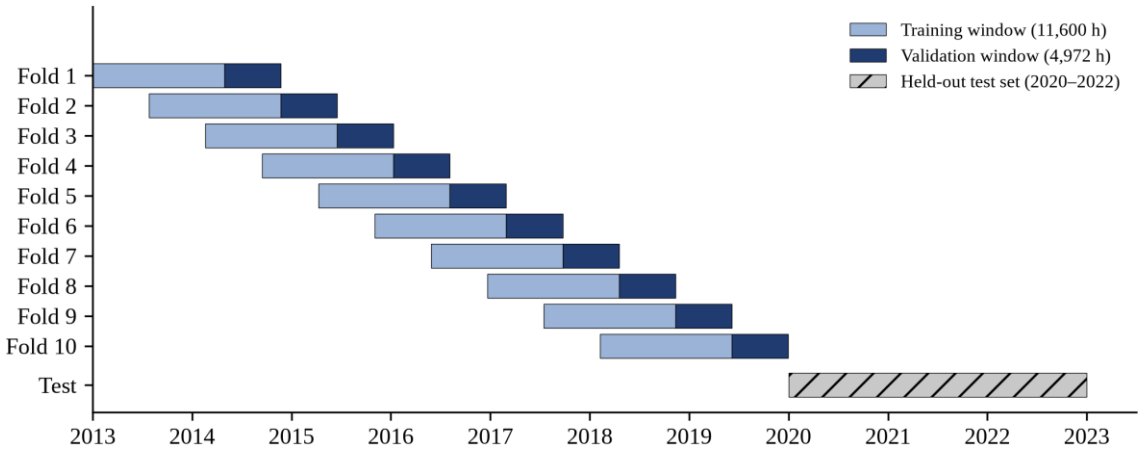


Figure 2: Rolling-window time-series cross-validation folds on the 2013–2019 development data and the held-out 2020–2022 test period.

Indicator features and labels

Four Technical Indicators, each describing a different aspect of price behaviour, were computed from the OHLC series (Table 1), following the indicator-based feature design of our earlier EA [14]. Eighteen two-candle CP capture reversal or continuation formations [5, 13]; a 50-period MA captures trend direction; a 10-period RSI captures momentum extremes; and a 14-period ADX captures trend strength. Each indicator state is encoded into a discrete class using the operational rules in Table 1, which also define the NO ACTION class needed when conditions are uncertain [9]. Labels are aligned with the decision time so that no future information is used.

Table 1: Indicator labelling rules

Indicator	**Condition → class**
CP	Bullish pattern → 0; Bearish pattern → 1; none → 2
MA	Close > MA → 0 (uptrend); Close < MA → 1 (downtrend)
RSI	RSI < 25 → 0 (oversold); RSI > 75 → 1 (overbought); 25–75 → 2 (no action)
ADX	ADX > 25 → 0 (trending); ADX ≤ 25 → 1 (sideways)

Classifiers and fuzzy decision layer

For each indicator stream, three supervised classifiers were trained: an SVM with a radial basis function kernel ($\sigma = 0.1$) [6], a PNN with a Gaussian smoothing parameter $\sigma = 0.5$ [20], and a Gaussian NB classifier without additional tuning. This yields twelve classifier outputs (3 classifiers × 4 indicators) per time step. Hyperparameters were fixed to allow a like-for-like comparison.

The twelve outputs are fused by a rule-based FIS, an approach that has been used to combine heterogeneous classifiers in other domains [1, 3, 8]. As shown in Figure 3, the FIS (1) groups outputs by indicator; (2) reduces each group to one linguistic decision by the mode of its three votes; (3) sets a group to NO ACTION when its three votes are all different; (4) aggregates the CP, MA and RSI group decisions by mode; and (5) applies a defuzzification rule that returns NO ACTION whenever BUY and SELL receive equal support. The ADX group decision (TRENDING or SIDEWAYS) is attached as market context and governs trade handling.

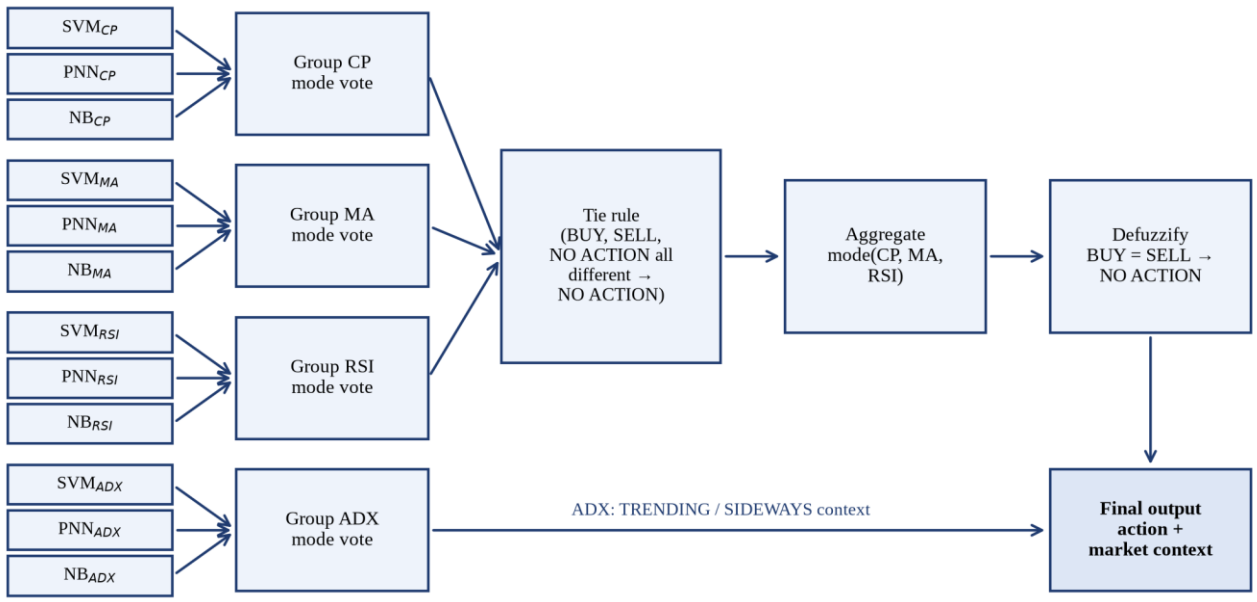


Figure 3: FIS fusing twelve classifier outputs into one trading decision.

EA execution and evaluation

The fused decision is executed by the EA on every new H1 bar in a continuous monitor-and-act loop, comparable to the real-time monitoring and control systems developed in our earlier work [27]: BUY or SELL opens a position and NO ACTION holds. The EA holds at most one position at a time. Each position is opened with an initial stop loss of 30 pips, and a trailing stop of 20 pips is attached: once price has moved 20 pips in favour of the trade the stop is moved to the entry price, and thereafter it follows price at a fixed distance of 20 pips. The stop is only ever moved in the direction of the trade, so it can be tightened on favourable movement but never loosened on adverse movement; a trade is closed either at the initial stop or when price retraces 20 pips from its most favourable level. No position scaling, averaging or counter-positions are used.

Predictive performance is reported on the held-out test set using accuracy, false-positive rate, sensitivity, specificity and F1-score. Trading performance is reported from the MT4 Strategy Tester over 2013–2022, with each calendar year started independently from a USD 30 balance; we report net profit, ROI (= net profit / initial balance), profit factor (gross profit / |gross loss|), win ratio and maximum and relative drawdown.

RESULTS

Backtest profitability and risk

Table 2 and Figure 4 summarise the yearly MT4 backtests. Starting from USD 30 each year, the EA closed every year from 2013 to 2022 with a positive net profit, ranging from USD 35.10 (2021, ROI 117.0%) to USD 69.90 (2015, ROI 233.0%). The cumulative net profit across the ten independent yearly runs was USD 464.54, an average of USD 46.45 or an average annual ROI of 154.8% on the USD 30 base. Win ratios were between 78.8% and 83.5%.

Table 2: Yearly MT4 backtest results, EUR/USD H1, USD 30 starting balance each year.

Year	Net profit (USD)	ROI (%)	Profit factor	Win ratio (%)	Max DD (USD)	Rel. DD (%)	Max consec. losses
2013	55.31	184.4	1.27	82.1	12.68	21.1	3
2014	42.30	141.0	1.28	82.6	10.10	20.9	2
2015	69.90	233.0	1.30	82.3	9.16	18.3	3
2016	48.84	162.8	1.27	81.5	11.51	30.9	4
2017	43.00	143.3	1.31	83.5	9.98	32.3	2
2018	44.21	147.4	1.18	78.8	12.63	24.1	3
2019	39.58	131.9	1.18	80.6	15.74	37.5	3
2020	40.66	135.5	1.26	82.7	12.82	20.5	2
2021	35.10	117.0	1.26	81.3	6.44	17.0	3
2022	45.64	152.1	1.24	81.9	13.60	21.2	3
Total / mean	464.54	154.8 (mean)	1.25 (mean)	81.7 (mean)			

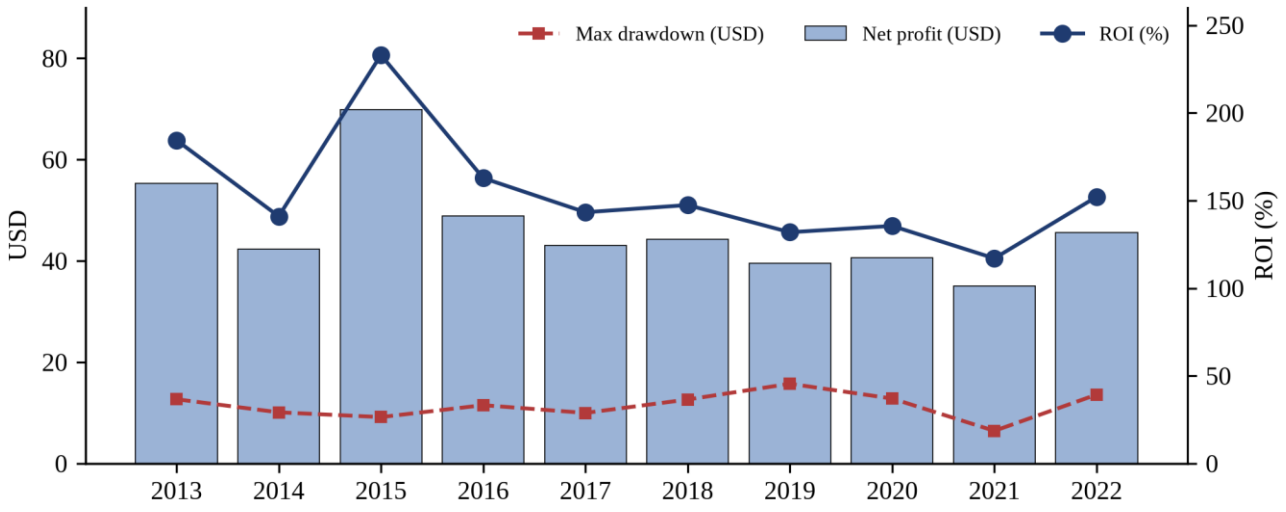


Figure 4: Yearly net profit, ROI and maximum drawdown from MT4 backtests (USD 30 start each year).

The risk profile is less comfortable than the returns suggest. Figure 5 shows that gross losses were large relative to gross profits: profit factors computed from gross values were only 1.18 to 1.31, so gross losses were 76–85% of gross profits. Maximum drawdown reached USD 15.74 (37.5% relative drawdown) in 2019, and relative drawdown exceeded 30% in three of ten years. The strategy therefore wins often but by small amounts, while individual losing trades reached USD 20.96 in 2015.

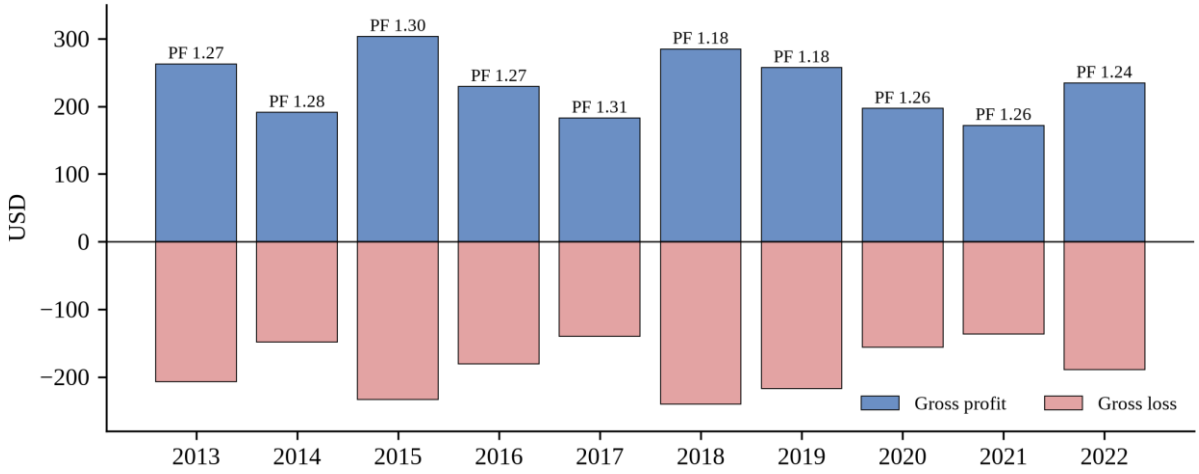


Figure 5: Yearly gross profit, gross loss and profit factor (PF) from MT4 backtests.

Predictive performance

Table 3 and Figure 6 report test-set accuracy. Among single classifiers, SVM was the strongest for every indicator input (96.06–97.58%), followed by PNN (95.17–97.40%) and NB (94.87–95.56%). The FIS-integrated configuration recorded the highest accuracy for every input, from 97.72% for ADX to 99.34% for the integrated indicator set. Figure 7 reports the gain over the strongest single classifier on the same input: 1.14 to 1.76 percentage points, with the largest gain on the integrated indicator set and the smallest on RSI, which suggests that fusion contributes most when the individual streams disagree.

Table 3: Test-set (2020–2022) classification accuracy by classifier and indicator input

Classifier	CP (%)	MA (%)	RSI (%)	ADX (%)	Integrated (%)
SVM	97.46	96.85	96.91	96.06	97.58
PNN	96.84	95.76	96.72	95.17	97.40
NB	95.47	94.99	94.92	94.87	95.56
SVM+PNN+NB+FIS	98.81	98.26	98.05	97.72	99.34

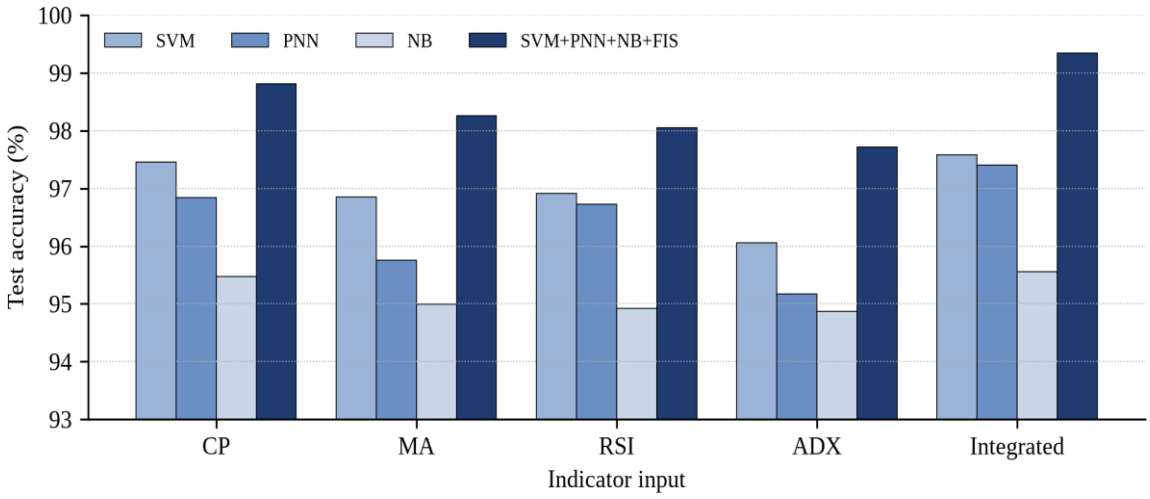


Figure 6: Test-set accuracy (2020–2022) by classifier and indicator input.

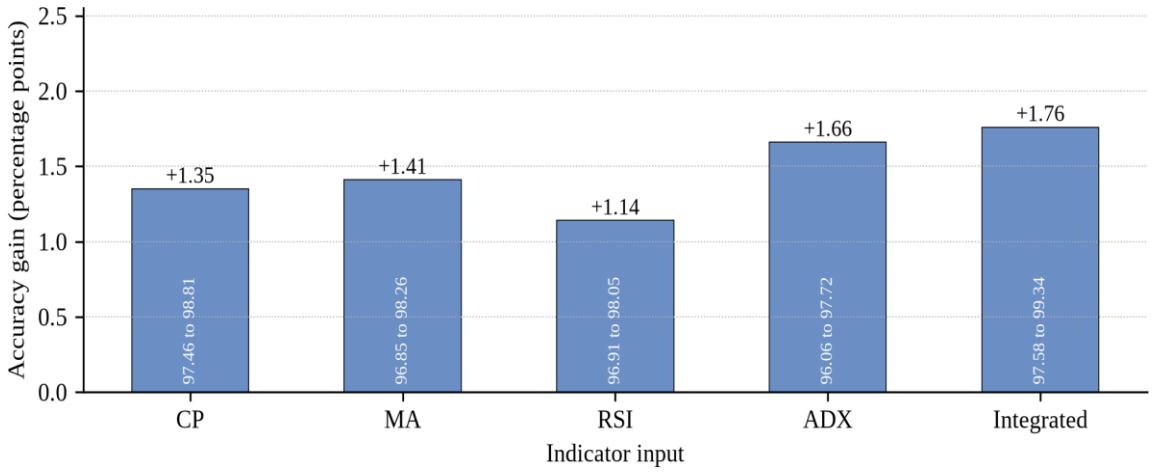


Figure 7: Accuracy gain of the FIS-integrated configuration over the strongest single classifier for each indicator input.

DISCUSSION

From a retail investor's perspective, the results are mixed. On the positive side, the EA produced positive simulated returns in every year, across contrasting macroeconomic regimes: the 2015 ECB quantitative easing, the 2016–2020 Brexit period, the 2020 COVID-19 shock and the 2022 tightening cycle. This consistency suggests that disciplined, rule-based execution can protect small traders from some discretionary errors, in line with arguments for algorithmic execution [16] and with the emotion-free trading motivation of our earlier EAs [5, 14].

On the other hand, the ROI figures must be interpreted carefully. Because the base balance is only USD 30, absolute profits are small (USD 35–70 per year) and percentage returns appear very large. The profit factor of about 1.2–1.3 shows that the edge per unit of risk is thin, and relative drawdowns above 30% in several years imply that a trader would regularly see a third of the account at risk. Because the Trailing Stop is only ratcheted and never widened, part of any open profit is given back on each retracement, and the fixed trailing distance means the exit rule was not tuned to the volatility regime. For retail traders, educators and regulators, the key implication is that marketing of automated trading products should report drawdown, profit factor and position-scaling rules alongside ROI.

Several limitations bound these findings. First, class labels are derived from deterministic indicator rules (Table 1), so the reported accuracy measures how faithfully the classifiers and FIS reproduce rule-defined trading signals; it is not a direct measure of forecasting the direction of future returns, and the high values should be read in that light. Second, profitability comes from historical simulation under one Strategy Tester configuration; spreads, slippage, execution latency and broker differences beyond that configuration were not modelled, and no live or forward test was performed. Third, the initial stop and trailing distance were fixed at 30 and 20 pips and were not optimised or sensitivity-tested, so the reported outcomes depend on these two exit parameters and on the order of intrabar price movement. Fourth, the results cover one currency pair and one timeframe, and no statistical comparison against a buy-and-hold or random-entry benchmark was performed.

CONCLUSION

This study evaluated a machine-learning EA for small-capital retail FX trading over ten years of EUR/USD data. The fuzzy-fused classifier pipeline reproduced rule-defined trading signals with up to 99.34% test accuracy, and annual backtests from a USD 30 balance were profitable in every year, with an average annual ROI of 154.8%. These returns, however, came with thin profit factors and relative drawdowns of up to 37.5%, indicating that headline ROI overstates the risk-adjusted benefit for retail traders. Future work should include live or forward testing with realistic spreads and slippage, risk-adjusted metrics such as the Sharpe and Calmar ratios, benchmark

comparisons, optimisation and sensitivity analysis of the trailing-stop distance, and extension to other currency pairs and timeframes.

Ethical Considerations

This study uses publicly available historical market price data only and involves no human or animal subjects; ethical approval was therefore not required.

Conflict Of Interest

The authors declare no conflict of interest.

Data Availability

The EUR/USD H1 price data were obtained from the MT4 History Center. Processed datasets and backtest reports are available from the corresponding author on reasonable request.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to all individuals and parties who contributed, directly or indirectly, to the completion of this study.

REFERENCES

1. Abbas, B., Alyas, T., Abbas, Q., Alqahtany, S., Alghamdi, T., et al. (2025). A hybrid support vector machine and neural network model with fuzzy logic fusion for smart city traffic prediction. Scientific Reports, 15(1), 34758. https://doi.org/10.1038/s41598-025-18472-2
2. Arratia, A., El Daou, M., Kagerhuber, J., & Smolyarova, Y. (2025). Examining challenges in implied volatility forecasting: A critical review of data leakage and feature engineering combined with high-complexity models. Computational Economics, 68(4), 3323–3345. https://doi.org/10.1007/s10614-025-11172-z
3. Atacak, İ. (2023). An ensemble approach based on fuzzy logic using machine learning classifiers for Android malware detection. Applied Sciences, 13(3), 1484. https://doi.org/10.3390/app13031484
4. Ayitey Junior, M., Appiahene, P., Appiah, O., & Bombie, C. N. (2023). Forex market forecasting using machine learning: Systematic literature review and meta-analysis. Journal of Big Data, 10(1), 9. https://doi.org/10.1186/s40537-022-00676-2
5. Bin Ismail, M. A. H., Bin Husin, Z., Keong, T. W., & Bin Yasruddin, M. L. (2022). Prediction of expert advisor trading system using an artificial intelligence system. In 2022 IEEE IAS Global Conference on Emerging Technologies (GlobConET) (pp. 233–237). IEEE. https://doi.org/10.1109/GlobConET53749.2022.9872367
6. Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine Learning, 20(3), 273–297. https://doi.org/10.1007/BF00994018
7. Dakalbab, F., Talib, M. A., Nasir, Q., & Saroufil, T. (2024). Artificial intelligence techniques in financial trading: A systematic literature review. Journal of King Saud University – Computer and Information Sciences, 36(3), 102015. https://doi.org/10.1016/j.jksuci.2024.102015
8. Dey, R., Piri, J., Acharya, B., Das, P. K., Gerogiannis, V. C., & Kanavos, A. (2025). A hybrid evolutionary fuzzy ensemble approach for accurate software defect prediction. Mathematics, 13(7), 1140. https://doi.org/10.3390/math13071140
9. Dezhkam, A., Manzuri, M. T., Aghapour, A., Karimi, A., Rabiee, A., & Shalmani, S. M. (2023). A Bayesian-based classification framework for financial time series trend prediction. The Journal of Supercomputing, 79(4), 4622–4659. https://doi.org/10.1007/s11227-022-04834-4
10. Drehmann, M., & Sushko, V. (2022). The global foreign exchange market in a higher-volatility environment. BIS Quarterly Review, December 2022, 33–48. https://www.bis.org/publ/qtrpdf/r_qt2212f.htm

11. Hernes, M., Korczak, J., Krol, D., Pondel, M., & Becker, J. (2024). Multi-agent platform to support trading decisions in the FOREX market. Applied Intelligence, 54(22), 11690–11708. https://doi.org/10.1007/s10489-024-05770-x
12. Hoang, D., & Wiegratz, K. (2023). Machine learning methods in finance: Recent applications and prospects. European Financial Management, 29(5), 1657–1701. https://doi.org/10.1111/eufm.12408
13. Hu, W., Si, Y.-W., Fong, S., & Lau, R. Y. K. (2019). A formal approach to candlestick pattern classification in financial time series. Applied Soft Computing, 84, 105700. https://doi.org/10.1016/j.asoc.2019.105700
14. Ismail, M. A. H., Yasruddin, M. L., Husin, Z., & Tan, W. K. (2022). Automated trading system for forecasting the foreign exchange market using technical analysis indicators and artificial neural network. In 2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA) (pp. 63–68). IEEE. https://doi.org/10.1109/CSPA55076.2022.9781856
15. Lim, H., Jeong, S., Oh, K., & Lee, H. (2022). Neural network foreign exchange trading system using CCS-IRS basis: Empirical evidence from Korea. Expert Systems with Applications, 205, 117718. https://doi.org/10.1016/j.eswa.2022.117718
16. Markets Committee. (2020). FX execution algorithms and market functioning (Markets Committee Papers No. 13). Bank for International Settlements. https://www.bis.org/publ/mktc13.htm
17. Mostafavi, S. M., & Hooman, A. R. (2025). Key technical indicators for stock market prediction. Machine Learning with Applications, 20, 100631. https://doi.org/10.1016/j.mlwa.2025.100631
18. Rossi, B. (2013). Exchange rate predictability. Journal of Economic Literature, 51(4), 1063–1119. https://doi.org/10.1257/jel.51.4.1063
19. Sevastjanov, P., Kaczmarek, K., & Rutkowski, L. (2024). A multi-model approach to the development of algorithmic trading systems for the Forex market. Expert Systems with Applications, 236, 121310. https://doi.org/10.1016/j.eswa.2023.121310
20. Specht, D. F. (1990). Probabilistic neural networks. Neural Networks, 3(1), 109–118. https://doi.org/10.1016/0893-6080(90)90049-Q
21. Suárez Cetrulo, A. L., Quintana, D., & Cervantes, A. (2024). Machine learning for financial prediction under regime change using technical analysis: A systematic review. International Journal of Interactive Multimedia and Artificial Intelligence, 9(1), 137–148. https://doi.org/10.9781/ijimai.2023.06.003
22. Tan, W. K., Husin, Z., & Ismail, M. A. H. (2020). Feasibility study of beef quality assessment using computer vision and deep neural network (DNN) algorithm. In 2020 8th International Conference on Information Technology and Multimedia (ICIMU) (pp. 243–246). IEEE. https://doi.org/10.1109/ICIMU49871.2020.9243353
23. Tan, W. K., Ismail, M. A. H., Husin, Z., & Yasruddin, M. L. (2022). Automated tomato grading system using computer vision (CV) and deep neural network (DNN) algorithm. In 2022 IEEE 12th Symposium on Computer Applications & Industrial Electronics (ISCAIE) (pp. 22–27). IEEE. https://doi.org/10.1109/ISCAIE54458.2022.9794557
24. Vancsura, L., Tatay, T., & Bareith, T. (2025). Navigating AI-driven financial forecasting: A systematic review of current status and critical research gaps. Forecasting, 7(3), 36. https://doi.org/10.3390/forecast7030036
25. Wang, G., Ma, J., Wang, Y., Tao, T., Ren, G., & Zhu, H. (2023). SUDF-RS: A new foreign exchange rate prediction method considering the complementarity of supervised and unsupervised deep representation features. Expert Systems with Applications, 214, 119152. https://doi.org/10.1016/j.eswa.2022.119152
26. Yang, X., Li, J., & Jiang, X. (2024). Research on information leakage in time series prediction based on empirical mode decomposition. Scientific Reports, 14(1), 28362. https://doi.org/10.1038/s41598-024-80018-9
27. Yasruddin, M. L., Ismail, M. A. H., Husin, Z., & Tan, W. K. (2022). Development of automated real-time water quality monitoring and controlling system in aquarium. In 2022 IEEE 12th Symposium on Computer Applications & Industrial Electronics (ISCAIE) (pp. 241–245). IEEE. https://doi.org/10.1109/ISCAIE54458.2022.9794472
28. Yasruddin, M. L., Ismail, M. A. H., Husin, Z., & Tan, W. K. (2022). Feasibility study of fish disease detection using computer vision and deep convolutional neural network (DCNN) algorithm. In 2022

IEEE 18th International Colloquium on Signal Processing & Applications (CSPA) (pp. 272–276). IEEE. https://doi.org/10.1109/CSPA55076.2022.9782020