INTERNATIONAL JOURNAL OF eBUSINESS AND eGOVERNMENT STUDIES
Vol: 17 **No:** 3 **Year:** 2025 **ISSN:** 2146-0744 (Online) (**pp.**123-148) **Doi:** 10.34109/ijebeg.2025170306
Received: 15.08.2025 | **Accepted:** 20.09.2025 | **Published Online:** 30.10.2025

-RESEARCH ARTICLE-

DEVELOPING A MACHINE LEARNING FORECASTING FRAMEWORK FOR EXCHANGE RATES IN THE CONTEXT OF CROSS-BORDER BUSINESS

Wei Chen

KMITL Business School, King Mongkut' s Institute of Technology
Ladkrabang, Thailand
ORCID: https://orcid.org/0009-0006-1305-2530
Email: 13188831790@163.com

Nuttawut Rojniruttikul*

KMITL Business School, King Mongkut' s Institute of Technology
Ladkrabang, Thailand
ORCID: https://orcid.org/0000-0003-3344-2406
Email: nuttawut.ro@kmitl.ac.th

—Abstract—

Traditionally, this matter was regarded as a largely technical undertaking; however, recent market disruptions have demonstrated that the issue extends far beyond purely technical considerations. This study addresses the subject from two interconnected perspectives. The first concern centers upon the capacity of different Machine Learning (ML) models to maintain performance during periods in which financial markets deviate from conventional behavioural patterns. The second concern focuses on the way professionals who depend upon such predictive outputs interpret, evaluate, and integrate these forecasts into routine operational decision-making. The empirical findings revealed that several ML approaches, particularly Long Short-Term Memory (LSTM) architectures and selected ensemble-based methods, adapted more consistently to abrupt market fluctuations than the econometric benchmark models employed within the study. Such resilience became especially apparent throughout the COVID-19 crisis, when exchange-rate dynamics departed substantially from the assumptions underpinning traditional forecasting frameworks. The interview findings produced a somewhat different perspective. Although most practitioners recognised the practical potential associated with ML-driven forecasting systems, their evaluations remained notably cautious. Collectively, these findings suggest that high

Citation (APA): Chen, W., Rojniruttikul, N. (2025). Developing a Machine Learning Forecasting Framework for Exchange Rates in the Context of Cross-Border Business. *International Journal of eBusiness and eGovernment Studies*, 17(3), 123-148. doi:10.34109/ijebeg.2025170306

predictive performance alone is insufficient to secure widespread organisational acceptance of ML applications within cross-border commercial activities. For these systems to become genuinely effective, they must integrate smoothly into existing organisational procedures and risk-management cultures while remaining comprehensible to end users. Consequently, effective forecasting frameworks should not only possess strong technical capability but must also remain interpretable, adaptable, and sufficiently resilient to withstand major structural transformations within financial markets.

Keywords: Machine Learning Forecasting for Exchange Rates, Cross-Border Business, COVID-19 Pandemic, Long Short-Term Memory Networks, Support Vector Regression, Random Forest Regression, and Gradient Boosting Regression.

INTRODUCTION

Background

International capital movements generate consequences that extend considerably beyond the boundaries of the foreign-exchange market itself. Persistent fluctuations in major currency pairs gradually shape international trade choices, influence investment allocation strategies, and affect the operational decision-making processes of multinational corporations. The present structure of global currency interdependence emerged largely after the collapse of the in 1971. Following the transition towards market-determined exchange-rate mechanisms, currency movements became increasingly connected to a broad combination of influences, including macroeconomic conditions, investor sentiment, geopolitical instability, and, more recently, algorithmic trading systems. In practice, these determinants rarely operate in isolation or remain aligned for extended periods. Instead, they interact in highly irregular ways, generating phenomena such as heavy-tailed return distributions, prolonged volatility persistence, and abrupt regime transitions. Such characteristics create substantial complications for conventional econometric frameworks, which typically depend upon restrictive assumptions concerning equilibrium stability and statistical distribution behaviour (Hamilton, 1989; Henrique et al., 2019; Sezer et al., 2020). Historically, forecasting approaches grounded primarily in economic theory have struggled to adapt effectively to real market behaviour. Financial economists have long debated the practical forecasting capacity of traditional theoretical models.

A highly influential contribution by Meese and Rogoff (1983) demonstrated that even theoretically sophisticated exchange-rate models frequently fail to outperform naïve forecasting approaches. More specifically, models derived from purchasing power parity and interest-rate differentials often perform little better than a random-walk specification across short- and medium-term forecasting horizons. Comparable conclusions have continued to emerge throughout subsequent research. Although

Vol: 17 **No:** 3 **Year:** 2025 **ISSN:** 2146-0744 (Online) (**pp.** 123-148) **Doi:** 10.34109/ijebeg.2025170306

time-series techniques such as Autoregressive Integrated Moving Average (ARIMA) models and their multivariate extensions occasionally generate modest forecasting improvements, their effectiveness often deteriorates during periods of policy intervention or financial instability, both of which have become increasingly common within modern currency markets (Henrique et al., 2019; Rossi, 2013).

The rise of ML approaches was driven largely by the shortcomings associated with conventional forecasting methodologies. Unlike traditional econometric systems that require relationships to be predetermined before estimation, ML techniques allow structural relationships to emerge directly from the underlying data. This capability enables the identification of nonlinear interactions, dynamic dependencies, and complex multivariable relationships that are often difficult to capture through conventional models. Deep-learning architectures, particularly LSTM networks, have received substantial academic and practical attention because they can model extended temporal dependencies without relying upon the rigid functional assumptions associated with classical econometrics. Their growing application reflects a broader attempt to construct forecasting systems capable of representing the unstable, nonlinear, and often erratic behaviour exhibited by real-world exchange rates (Fischer & Krauss, 2018; Sezer et al., 2020). The COVID-19 crisis imposed considerably greater forecasting challenges than many researchers had originally anticipated. During the early stages of 2020, exchange-rate movements shifted well beyond historically normal ranges, causing models calibrated upon pre-pandemic relationships to lose predictive reliability rapidly. Previously stable behavioural trends deteriorated abruptly, while established assumptions concerning market reactions to economic announcements no longer aligned with prevailing economic realities (Baker et al., 2020). The situation became even more difficult to interpret once central banks implemented extensive unconventional monetary measures, including negative interest-rate policies, large-scale asset-purchase programmes, and numerous emergency liquidity interventions. Consequently, the relationship between macroeconomic indicators and currency behaviour became substantially less stable and increasingly difficult to interpret. Under such conditions, forecasting robustness requires continuous recalibration because the structural relationships underpinning the models remain unstable for only limited periods (Aslam et al., 2020; Azzam et al., 2023).

At the same time, access to ML technologies expanded considerably. In principle, this development should have enabled smaller organisations to experiment independently with sophisticated forecasting infrastructures. Cloud-computing environments and open-source frameworks, including TensorFlow and PyTorch, increasingly provide smaller firms with tools previously available primarily to major financial institutions. The availability of pre-trained models has further lowered technical barriers to entry. Nevertheless, the practical environment remains unevenly distributed. Many organisations continue to face limited access to high-quality financial data, while smaller operational teams frequently lack sufficient computational resources to sustain complex modelling systems effectively. Perhaps the most persistent limitation concerns the shortage of specialists possessing expertise in both financial-market dynamics and modern ML methodologies. Consequently, technological accessibility alone has not guaranteed widespread adoption, as substantial skills gaps continue to obstruct implementation efforts despite increased technological availability (Gu et al., 2020; Henrique et al., 2019; Plakandaras et al., 2019). The present study therefore seeks to evaluate the performance of an ML-based exchange-rate forecasting framework during periods characterised by severe economic disruption, particularly conditions like those experienced throughout the pandemic era. The objective extends beyond the narrow pursuit of forecasting precision alone. Equally important is the examination of the practical characteristics required for a forecasting system to remain operationally useful under rapidly changing economic conditions. This becomes especially significant when the assumptions embedded within forecasting frameworks may require adjustment far sooner than anticipated because of continuous structural transformation within financial markets (Abedin et al., 2025; Tripathi et al., 2021; Yıldırım et al., 2021).

Research Scope

Defining the scope of a study or project is essential to maintaining its feasibility, as it establishes clear boundaries that prevent the investigation from extending into areas that cannot be addressed adequately within the available constraints. In this case, the structure of the work is organised into three interrelated sections, which assists in guiding the reader through the core concepts while simultaneously delimiting both the subject matter and the temporal coverage of the analysis. This structuring framework also serves a methodological function, as it explicitly articulates the research design choices and justifies why particular approaches are adopted over alternative options. When the boundaries of a study are clearly established at the outset, the research process becomes more coherent, and the resulting work is easier to navigate, evaluate, and defend academically. Likewise, when the scope and interpretation of findings are framed in broader, generalisable terms, their overall applicability and practical value are enhanced.

Conceptual Framework

The independent variables employed in this study can be broadly categorised into two main groups, although in practice the distinction between them is not always rigid. The first category comprises a range of ML forecasting techniques, including LSTM, SVR, Random Forest, Gradient Boosting, as well as additional methods that were initially examined during the preliminary phase but subsequently excluded from the final model set. The models illustrated in Figure 1 were not selected arbitrarily but

Vol: 17 **No:** 3 **Year:** 2025 **ISSN:** 2146-0744 (Online) (**pp.** 123-148) **Doi:** 10.34109/ijebeg.2025170306

based on performance merit, with each algorithm offering a distinct mechanism for handling the irregularities characteristic of financial time-series data.

LSTM can capture long-range temporal dependencies effectively, whereas SVR typically responds differently when dealing with noisy or sparse datasets. Tree-based approaches such as Random Forest and Gradient Boosting place greater emphasis on learning interactions among variables rather than imposing strict functional assumptions. Within this modelling framework, the selected techniques differ substantially in how they mitigate overfitting, the degree of interpretability of their internal structures, and their sensitivity to structural shifts in financial markets. Such classification is particularly valuable, as noted by Mussa, given the absence of any universally applicable model governing exchange-rate dynamics (Fischer & Krauss, 2018; Sezer et al., 2020).

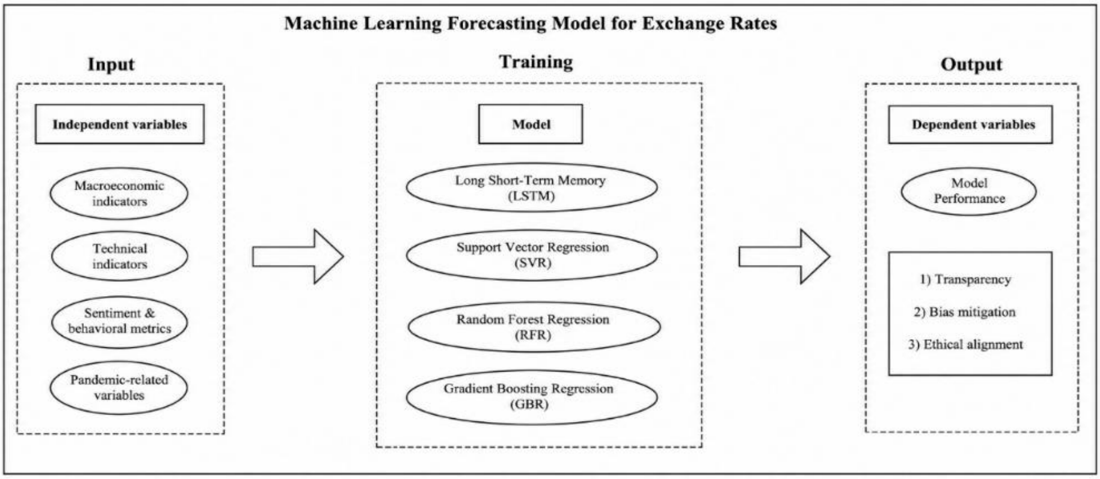

Figure 1: Research Conceptual Framework

The second group consists of factors that are external to the forecasting algorithms themselves yet remain essential for explaining why predictive performance becomes more difficult and unstable at certain points in time. These include conventional macroeconomic indicators such as interest rates, inflation rates, and current account balances, although their influence tends to vary depending on prevailing market conditions. In addition, technical indicators, including moving averages and the Relative Strength Index (RSI), introduce an additional analytical layer, while sentiment-based inputs such as news sentiment measures and geopolitical risk indicators further complicate the forecasting environment. Under such conditions, market responses to these inputs are often nonlinear and highly state dependent. During the COVID-19 period, for example, several macroeconomic indicators exhibited atypical behaviour; Gross Domestic Product (GDP) growth became significantly more volatile and, in many cases, lost much of its conventional explanatory power.

Vol: 17 **No:** 3 **Year:** 2025 **ISSN:** 2146-0744 (Online) (**pp.** 123-148) **Doi:** 10.34109/ijebeg.2025170306

In relatively stable economic environments, the reliance on alternative data sources and frequent re-weighting of model inputs would typically be less necessary (Aslam et al., 2020; Stillwagon & Sullivan, 2020). Beyond these variables, this study also considers the role of fairness considerations and interpretability requirements in shaping the acceptance of ML-based predictions. This is not merely a supplementary technical concern; rather, in cross-border and multi-regulatory contexts, the ability to explain why a model produces a given forecast is as important as the forecast itself. Accordingly, model evaluation in this analytical framework extends beyond conventional error-based metrics. It also involves assessing systematic bias within the models, examining the interpretability of outputs, and evaluating alignment with relevant institutional and normative decision-making expectations. Collectively, these dimensions influence both the perceived credibility and the eventual adoption of ML-based forecasting tools (Bussmann et al., 2021; Hadji Misheva et al., 2021).

LITERATURE REVIEW

Traditional Exchange Rate Forecasting Models

Linear Time Series Models

Exchange-rate forecasting was initially developed using relatively simple time-series approaches that sought to interpret price movements solely through historical patterns in the exchange rate. At this stage, broader macroeconomic structures were largely excluded from analytical consideration. The random walk model quickly emerged as a dominant benchmark, partly because it reflects the efficient-market intuition that tomorrow's exchange rate equals today's value plus an unpredictable shock. The robust empirical finding that structural economic models have generally failed to outperform this simple benchmark in out-of-sample forecasting further reinforced its credibility (Jung & Choi, 2021; Meese & Rogoff, 1983; Rossi, 2013). The introduction of ARIMA models added a greater degree of formal structure to the forecasting process. These models incorporate short-run dependencies and allow for gradual adjustments in the data-generating process, while the Box–Jenkins methodology provides systematic guidance for identifying appropriate specifications (Kyriazi & Thomakos, 2026). However, when applied to exchange-rate data, many ARIMA specifications tend to converge towards behaviour like a random walk, with only limited autoregressive and moving-average components included.

Subsequently, VAR models were introduced to better capture more complex interdependencies in which exchange rates co-move with other macroeconomic variables. The VAR framework is particularly advantageous because it does not require a priori assumptions regarding causal direction among variables. Instead, each variable is modelled as a function of its own lagged values as well as the lagged values of all other variables in the system, providing a representation that is closer to the observed dynamics of financial markets (Stock & Watson, 2002).

Nonlinear Time Series Models

Non-linear threshold autoregressive frameworks represent some of the earlier attempts to model exchange-rate dynamics within a regime-dependent structure. In essence, these models permit the behaviour of a time series to alter once a specified threshold is crossed (Uche et al., 2023). In the context of exchange rates, such thresholds are often interpreted as points at which deviations from purchasing power parity (PPP) become sufficiently pronounced for arbitrage incentives to emerge. Once deviations widen beyond a certain level, the speed of adjustment may increase markedly; however, when deviations remain small, transaction costs can dominate and slow down corrective forces.

Self-Exciting Threshold Autoregressive (SETAR) models extend this logic by using lagged values of the exchange rate itself as the switching mechanism between regimes. This feature allows SETAR specifications to capture momentum-like dynamics, depending on recent price movements. Empirical evidence suggests that such threshold effects are indeed present, particularly in real exchange-rate series. However, when parameter uncertainty is explicitly accounted for, the resulting improvements in forecasting performance tend to be limited, with only marginal gains observed (Rossi, 2013). Smooth Transition Autoregressive (STAR) models were introduced partly to address the restrictive nature of abrupt regime changes. In contrast to discrete switching structures, STAR-type models allow transitions between regimes to occur gradually rather than instantaneously (Maitra, 2025). The Exponential Smooth Transition Autoregressive (ESTAR) specification has been widely applied to real exchange-rate data, as it captures the idea that the corrective force towards equilibrium strengthens as deviations from equilibrium increase. Despite this theoretical advantage, greater model complexity does not consistently translate into superior forecasting accuracy, particularly at short horizons where linear approximations often perform comparably well (Stillwagon & Sullivan, 2020).

Markov-switching (MS) models provide an alternative mechanism for capturing regime changes by allowing the data to probabilistically determine latent states. These states are often interpreted as reflecting shifts in monetary policy regimes, variations in risk appetite, or episodic intervention in currency markets (Bollerslev, 1986). Empirical applications in exchange-rate modelling frequently reveal volatility clustering patterns, where periods of high volatility tend to be followed by further turbulent behaviour. Subsequent extensions of MS models permit transition probabilities to depend on observable economic conditions, thereby enabling time-varying probabilities of regime shifts and improving the interpretability of dynamic structural changes in exchange-rate behaviour.

Volatility Models

The literature broadly indicates that most research on exchange-rate volatility forecasting is grounded in the assumption that accurate volatility prediction is valuable, particularly for derivative pricing and risk management applications. Within this domain, the GARCH family of models is widely regarded as a standard analytical framework. As demonstrated by Nelson (1991), GARCH processes capture the empirical regularity that large absolute shocks in financial time series tend to be followed by further large absolute shocks, reflecting volatility clustering behaviour.

In its basic specification, the GARCH (1,1) model is often considered sufficient to represent a substantial proportion of volatility dynamics using a relatively parsimonious two-parameter structure. In practical applications, GARCH (1,1) has therefore been extensively employed in modelling exchange-rate volatility, largely because it effectively captures the persistent nature of volatility observed in empirical data. Subsequent methodological extensions have further refined this framework. Integrated GARCH (IGARCH) models allow volatility to exhibit near unit-root behaviour, while fractional variants such as Fractionally Integrated GARCH (FIGARCH) are designed to accommodate the long memory and very slow mean-reversion patterns observed in certain currency markets (Bollerslev, 1986). A further important development in this literature concerns asymmetric volatility effects, which refer to the phenomenon whereby positive and negative shocks exert different impacts on future volatility (Leippold et al., 2022). In equity markets, negative news—such as poor earnings announcements or adverse macroeconomic data—typically induces a stronger increase in volatility than equivalent positive news. Although exchange-rate markets are generally less directionally interpretable than equity markets, similar asymmetries are still observed. Extensions such as the Exponential GARCH (EGARCH) model introduced by Nelson (1991) incorporate such asymmetries in a relatively parsimonious form with minimal parameter constraints, while the Glosten–Jagannathan–Runkle GARCH (GJR-GARCH) specification introduces a more explicit threshold-based mechanism.

Empirical evidence suggests that in exchange-rate contexts, asymmetry is often driven more by the magnitude of shocks than by their direction. Unusually large movements—regardless of sign—tend to increase and prolong volatility. Although these enhanced specifications improve volatility forecasting and are particularly useful for option pricing and risk measurement, they are less frequently applied directly to forecasting exchange-rate levels themselves (Nelson, 1991).

Machine Learning Forecasting Model for Exchange Rates

Exchange-rate forecasting has experienced substantial improvements with the adoption of ML methodologies.

Table 1: Machine Learning Forecasting Model for Exchange Rates

Model	Contents	Mechanism	Importance
LSTM	A variant of the Recurrent Neural Network (RNN) particularly engineered to capture long-term dependencies in sequential input.	LSTM networks employ memory cells and gating techniques to identify patterns over long time periods, rendering them especially adept at capturing the dynamics of currency fluctuations. Their capacity to include both organized numerical data and unstructured information—such as news sentiment or geopolitical indicators—has rendered them a favoured option in contemporary FX forecasting literature.	Empirical evidence repeatedly demonstrates that LSTM models surpass traditional time series models, particularly in turbulent market conditions marked by sudden structural breaks or regime upheavals.
SVR	A supervised learning model employing kernel functions to address nonlinear interactions.	SVR maps input information into a high-dimensional space and constructs a regression function within a margin of tolerance, thus attaining significant resilience against outliers and noise.	SVR has demonstrated efficacy in short- and medium-term contexts, especially when the dataset is constrained in size or when macroeconomic factors interact in intricate, non-additive manners. Its efficacy and clarity render it a persuasive substitute for more computationally demanding deep learning models.
RFR	An ensemble machine learning approach predominantly employed for addressing regression issues, specifically forecasting continuous numerical variables such as real estate prices, temperature, or stock market values.	RFR functions by consolidating the outputs of several decision trees, each trained on random data subsets, thus mitigating overfitting and improving generalization.	This model is particularly adept at handling high-dimensional financial data that encompasses a combination of macroeconomic indicators, technical analysis measures, and other data sources. Furthermore, RFR's capacity to assess variable significance enhances interpretability, which is frequently absent in neural network models.
GBR	It encompasses widely utilized implementations such as XGBoost and LightGBM.	These models construct additive regression trees sequentially, with each iteration rectifying the residuals of its predecessor.	GBR excels in structured, tabular datasets and has shown strong performance in FX rate prediction when integrated with engineering characteristics, including lagged variables, volatility metrics, and sentiment indices. The adaptability of GBR in managing intricate feature interactions with minimal parameter adjustment.

These techniques enable the representation and approximation of complex, dynamic financial systems that are difficult to capture using conventional econometric approaches. Algorithms such as LSTM, SVR, RFR, and GBR each offer distinct modelling advantages, and the decision to apply them individually or within ensemble

Vol: 17 No: 3 Year: 2025 ISSN: 2146-0744 (Online) (pp. 123-148) Doi: 10.34109/ijebeg.2025170306

configurations can significantly enhance forecasting performance, as summarised in Table 1.

Ethics and Fairness in Machine Learning Applications

With ML increasingly integrated into finance, ethical considerations can no longer be treated as peripheral to technical model design. Beyond generating numerical forecasts, predictive systems now influence trading behaviour and shape how institutions interpret market dynamics. Consequently, issues of fairness and accountability arise not only after model deployment but throughout the entire modelling lifecycle, thereby influencing evaluation criteria from the outset.

Although bias is often the first concern raised, it conceals the multiple pathways through which it can emerge in ML systems. A significant portion originates from the underlying data, particularly when historical datasets contain periods of structural imbalance or extreme shocks that are no longer representative of current market conditions. Models trained on such data may implicitly encode these historical distortions, treating certain currencies as inherently more volatile simply because earlier crisis periods are overrepresented. This can produce a seemingly coherent structure that is misaligned with present-day dynamics. In such cases, the failure is not purely methodological but arises when corrective adjustments are not implemented (Goodell et al., 2021).

Fairness extends this discussion in a different direction. In ML contexts, fairness is often framed as the principle of generating comparable outputs for comparable inputs. In exchange-rate forecasting, this raises concerns about whether currencies with shorter or irregular historical datasets are structurally disadvantaged relative to those with long, stable time series. As highlighted in the literature, ensuring that predictive performance is both accurate and unbiased is essential, since otherwise decision-making processes may be distorted by systematically misleading outputs (Bussmann et al., 2021). Transparency presents an additional challenge. High-performing models, particularly deep learning architectures, often operate as opaque systems in which the internal reasoning behind predictions is not directly interpretable. Existing explainability techniques provide only partial insight into these processes. While analysts may tolerate a degree of opacity, regulators and risk managers generally require clearer justification of model outputs. A further complication is that such interpretability may be inherently limited, particularly when models are applied to market regimes that differ significantly from those on which they were trained (Hadji Misheva et al., 2021).

A systemic issue emerges when similar modelling frameworks are widely adopted across institutions. If multiple models rely on comparable datasets or shared assumptions, they may generate correlated errors, leading to synchronized misjudgements. Under such conditions, individual model weaknesses can be amplified at the system level, transforming institution-specific risks into broader market-wide vulnerabilities. These challenges point to the need for a more comprehensive governance framework that evolves alongside modelling practices. Mitigation strategies such as fairness evaluation, periodic model audits, and structured human oversight at critical decision points can help address the limitations of automated forecasting systems. Given that these concerns extend beyond purely technical dimensions, their resolution requires close integration with economic and institutional analysis rather than being treated as secondary considerations (Goodell et al., 2021).

There is little doubt that ML enhances exchange-rate analysis, particularly in capturing nonlinear relationships and processing large-scale datasets. However, these technical improvements do not diminish the importance of ethical and governance-related challenges. From the outset, models should be designed with fairness, transparency, and accountability embedded as core principles in order to support a stable and equitable global financial system (Bussmann et al., 2021; Hadji Misheva et al., 2021). From a practical standpoint, ML models may also provide faster adaptation to changing market conditions, thereby offering competitive advantages in cross-border decision-making. This perspective reflects less a purely theoretical argument and more an observation of how organisations manage uncertainty in practice, giving rise to hypotheses concerning strategic decision-making effectiveness (Abedin et al., 2025).

Finally, the performance of neural machine translation (NMT) models is influenced by multiple factors, including dataset characteristics, architectural design evolution, and evaluation methodologies. While empirical findings across studies are not fully consistent, there is broad agreement that ML-based predictive gains are not uniform. This variability supports additional hypotheses linking data quality and model design to forecasting outcomes (Dautel et al., 2020; Yıldırım et al., 2021).

Research Methodology

The study employs multiple supervised learning models, although not all are intended for identical analytical roles. One such model that was extensively evaluated is SVR, primarily due to its relatively stable behaviour in high-dimensional feature spaces following repeated testing. This outcome was not fully anticipated at the outset, as earlier expectations suggested that the introduction of nonlinear kernels might induce instability. However, this limitation proved less pronounced than initially assumed.

Tree-based ensemble methods operate according to a distinct modelling logic compared to other ensemble approaches. Among these, RF demonstrated strong robustness to noisy segments within the dataset, while its feature-importance outputs proved particularly valuable during the interpretative phase, even prior to formal

Vol: 17 **No:** 3 **Year:** 2025 **ISSN:** 2146-0744 (Online) (**pp.** 123-148) **Doi:** 10.34109/ijebeg.2025170306

comparative evaluation procedures (Fischer & Krauss, 2018; Henrique et al., 2019). GBR was incorporated at a later stage than the other modelling techniques. Its inclusion reflected observable improvements in handling unbalanced or weakly specified relationships within the data. However, these advantages only became apparent after several unsuccessful parameter configurations, which in turn prompted a reassessment of some initial modelling assumptions.

LSTM networks occupy a different position within the modelling framework. Due to their capacity to retain information over extended temporal horizons, they are particularly suited to periods in which policy announcements or market shocks are gradually incorporated into exchange-rate movements. In empirical testing, they produced more pronounced structural patterns in certain cases, while yielding less consistent outputs in others, suggesting that performance variability may be more closely linked to training window selection than to architectural design alone (Cao et al., 2020; Sun et al., 2020). Although the role of the gating mechanism in explaining these behavioural differences remains uncertain, the overall stability of the model outputs was sufficient to justify its inclusion in the primary analytical framework.

In performance evaluation, RMSE, MAE, MAPE, and Theil's U are commonly used as standard error metrics. However, these measures do not necessarily move in a consistent or aligned manner. At one stage, a model that initially appeared strong under RMSE performed poorly according to Theil's U, which prompted a re-examination of the rolling-window specification. This procedure introduces its own methodological challenges, as even minor adjustments to the cut-off date can materially alter the apparent forecasting performance. This sensitivity suggests that the metrics may, in part, capture not only predictive accuracy but also responsiveness to structural breaks in the underlying data-generating process. Given the instability observed across adjacent windows, several candidate parameter grids were abandoned mid-process, and time-series cross-validation was ultimately adopted for hyperparameter optimisation (Sezer et al., 2020; Yıldırım et al., 2021).

The overall research design adopts a mixed-methods approach. While the quantitative component provides statistical evidence through model estimation and evaluation, it does not fully capture institutional dimensions such as interpretability constraints, fairness considerations, and implementation costs, all of which emerged consistently in the interview data. The qualitative strand is therefore used to contextualise and interpret these limitations, particularly in relation to practitioners' reservations about models that perform well statistically but encounter resistance in practical application. Rather than fully integrating into a single unified framework, the two methodological strands remain analytically distinct, and this tension itself forms part of the study's contribution (Braun & Clarke, 2019; Hadji Misheva et al., 2021).

Vol: 17 **No:** 3 **Year:** 2025 **ISSN:** 2146-0744 (Online) (**pp.** 123-148) **Doi:** 10.34109/ijebeg.2025170306

The quantitative dataset spans the period from 2015 to 2023. This interval was selected deliberately to ensure coverage of both relatively stable pre-COVID-19 market conditions and the heightened volatility observed during and after the COVID-19 period. However, preparing the dataset required substantial pre-processing. Differences in frequency and irregular gaps across exchange-rate series, macroeconomic indicators, and pandemic-related variables created significant alignment challenges.

Table 2: Data Sources for Quantitative Research

Indicators	Data Sources and Collection
Exchange Rate Data	Major currencies (e.g., USD/EUR, USD/CNY, USD/JPY). These are collected on a daily frequency, which allows for high-resolution time series analysis and short-term volatility modelling. The exchange rate data are sourced from Bloomberg, Yahoo Finance, and the Federal Reserve Economic Data (FRED) platform, ensuring accuracy and consistency.
Macroeconomic Indicators, these include: 1) Interest Rates (short-term and long-term government yields) 2) Inflation Rates (Consumer Price Index) 3) Gross Domestic Product (GDP) Growth 4) Trade Balances and Current Account Data 5) Monetary Policy Announcements	We gather these variables from international sources including the International Monetary Fund (IMF), World Bank, OECD iLibrary and national central banks. To ensure comparability, the variable data is collected at the same intervals (weekly or monthly depending on release frequency) and transformed where necessary (e.g. seasonal adjusted or log-difference).
High-Frequency Market Sentiment Indicators, these include: 1) Google Trends Indexes on relevant keywords (e.g., “currency crash,” “FX risk”) 2) VIX (Volatility Index) as a proxy for global market uncertainty 3) News Sentiment Scores derived from financial headlines using natural language processing (NLP) techniques	These high-frequency indicators allow us to know investor psyche and media narratives, especially at the time of crisis such as COVID-19. The data was collected from Google Trends API, CBOE, and created databases like RavenPack or Thomson Reuters Refinitiv.
Pandemic-Specific Variables, these include: 1) Stringency Index from the Oxford COVID-19 Government Response Tracker 2) Daily case/death counts from the World Health Organization (WHO) 3) Economic Support Measures (e.g., stimulus packages, quantitative easing announcements)	These variables are intended to capture structural shocks and policy interventions that significantly influenced currency markets during the pandemic.

Missing observations and mismatched timestamps were addressed through interpolation techniques. In some cases, forward-filling was sufficient, whereas in others, spline-based smoothing was required to generate more consistent temporal structures. The resulting dataset was iteratively cleaned, validated, and refined before being transformed into feature matrices suitable for both econometric and ML modelling. A summary of data sources is presented in Table 2, although this summary does not fully reflect the extensive pre-processing required to reach an analytically usable dataset (Abedin et al., 2025; Aslam et al., 2020).

The core dependent variable is defined as either the daily exchange-rate level or its logarithmic return for selected currency pairs, including USD/EUR, USD/CNY, and USD/JPY. The log-return transformation is widely adopted in financial modelling due to its ability to stabilise variance and improve stationarity properties of the time series. It is computed as follows:

$$R_t = \text{In}\left(\frac{p_t}{p_{t-1}}\right)$$

Where p_t denotes the exchange rate at time t. In some specifications, binary direction indicators (upward/downward movement) are also constructed to enable classification-based modelling.

To enhance model robustness and predictive capacity, the following techniques are applied:

1. Lagged Variables: All independent variables are lagged across multiple time steps (e.g., t−1, t−5, t−10) to capture temporal dependencies.
2. Differencing: First-order differencing is applied to non-stationary series.
3. Normalization/Scaling: All features are standardised using Min-Max scaling or Z-score normalisation to ensure comparability across models.
4. Interaction Terms: Key interaction features (e.g., interest rate × inflation) are constructed to capture nonlinear interdependencies.
5. Principal Component Analysis (PCA): Applied as an optional dimensionality reduction technique to mitigate multicollinearity and reduce noise (Henrique et al., 2019).

Phase 2: Qualitative Research

In addition to the quantitative modelling work, a qualitative strand of research is incorporated based on semi-structured interviews conducted with practitioners from finance, international trade, and, in a limited number of cases, policy advisory roles (Braun & Clarke, 2019). The inclusion of this component was partly necessitated by analytical limitations observed in the preceding modelling phase, as several findings could not be adequately explained through numerical indicators alone and instead required an understanding of how forecasts are interpreted and operationalised within institutional settings. The interview protocol was initially aligned with the overarching research objectives; however, after the first set of interviews, several refinements were

Vol: 17 No: 3 Year: 2025 ISSN: 2146-0744 (Online) (pp. 123-148) Doi: 10.34109/ijebeg.2025170306

introduced. This adjustment arose because participants frequently raised concerns—particularly around interpretability, internal validation procedures, and organisational constraints—that were not originally foregrounded in the design.

Although the interviews were structured around key thematic areas, their progression was often non-linear in practice. As discussions developed, participants initially engaged with questions regarding the reliability of ML-based models but frequently shifted towards broader operational issues, such as the difficulty of communicating model outputs to non-technical stakeholders or the extent to which a single ambiguous result can delay or disrupt decision-making processes. Responses also revealed heterogeneous forecasting practices: while some practitioners continued to rely heavily on traditional econometric approaches, others alternated between conventional and algorithmic methods depending on prevailing market conditions. Notably, these qualitative reflections occasionally diverged from the performance hierarchies indicated by the quantitative results, reinforcing the importance of this analytical strand (Braun & Clarke, 2019).

A recurring discussion focused on trade-offs between accuracy, transparency, and ethical accountability. Although not initially central, it became a key theme during interviews. Participants noted that the statistically best-performing model is not always the most usable in practice, while fairness concerns—especially for emerging-market currencies—are increasingly treated as institutional requirements rather than abstract issues. The COVID-19 period also reshaped expectations: some demanded greater model flexibility, while others became less tolerant of black-box systems after the breakdown of economic relationships.

Table 3: Data Sources for Quantitative Research

Interview Format	Semi-Structured Interview	Approximately 10 to 15 interviews are targeted to achieve theoretical saturation, where additional interviews yield diminishing new insights.
Participants	Corporate finance executives managing foreign exchange exposure, institutional traders operating in currency markets, and economic policy advisors engaged in macroeconomic planning.	
Analytical Strategy	Thematic analysis is a prevalent qualitative research method that entails detecting, classifying, and analysing recurring patterns within interview transcripts. Initial codes are generated both inductively from the data and deductively from existing theoretical frameworks. The codes are subsequently categorized into higher-order themes that encapsulate the predominant storylines and conflicts found in the responses.	

Table 3 summarises the data sources and interview metadata, though it simplifies a more iterative process of continuous refinement and expert dialogue.

The purpose of the qualitative narratives in this study is not primarily to validate the quantitative models, but rather to develop a deeper understanding of the practical

Vol: 17 **No:** 3 **Year:** 2025 **ISSN:** 2146-0744 (Online) (**pp.** 123-148) **Doi:** 10.34109/ijebeg.2025170306

implications underlying why different models exhibit distinct behaviours under conditions of market stress. Interview evidence indicates that stakeholders rarely rely on single performance metrics such as RMSE or MAE in isolation; instead, they tend to consider multiple indicators simultaneously, or in some cases rely on alternative evaluative heuristics altogether. Some participants even acknowledged selecting models that align more effectively with internal reporting structures, despite being aware that alternative models may offer superior technical performance.

As a result, the qualitative component assumed a somewhat different function from what was initially envisaged. Rather than serving as a supplementary validation tool, it evolved into an analytical lens through which issues of interpretability, institutional routines, and organisational risk culture are seen to outweigh purely statistical considerations. The integration of the quantitative and qualitative strands is best understood as a form of thematic triangulation rather than a strict methodological synthesis. Certain themes demonstrated strong alignment with the quantitative findings, others revealed clear inconsistencies, while some reconfigured the interpretation of the numerical results altogether. Convergence between the two strands did not emerge through imposed alignment but developed organically through the analytical process (Braun & Clarke, 2019).

RESULTS

Quantitative Results

The study compares three classical econometric models (ARIMA, GARCH, VAR) with four ML approaches (SVR, RFR, LSTM, GBR). Forecast accuracy is evaluated using RMSE; however, as becomes evident in later analysis, RMSE alone does not fully capture the underlying behaviour and performance dynamics of the models.

Model Performance Trends (For Hypothesis 1, 2, 5)

As shown in Table 4, classical econometric models perform relatively well during stable or low-volatility periods but exhibit a marked deterioration under turbulent market conditions. In the pre-COVID period, RMSE values for ARIMA and GARCH were 0.84 and 0.79 respectively. However, model accuracy declined significantly following the onset of the pandemic shock, with RMSE rising to 1.32 and the corresponding index reaching 1.41. Among the classical models, VAR experienced the most pronounced deterioration during periods of stress. This outcome is consistent with the structural limitation that VAR models rely on fixed lag relationships that do not adapt efficiently when underlying economic linkages break down due to structural shocks.

In contrast, ML models demonstrate a comparatively smaller decline in performance. Owing to their capacity to learn sequential dependencies rather than relying solely on

Vol: 17 **No:** 3 **Year:** 2025 **ISSN:** 2146-0744 (Online) (**pp.** 123-148) **Doi:** 10.34109/ijebeg.2025170306

fixed lag structures, LSTM achieves pre-COVID RMSE values of 0.65, which increase moderately to 0.74 during the COVID-19 period. RFR and GBR also maintain relatively stable error levels across both regimes. Overall, these results indicate that ensemble-based learners tend to exhibit greater robustness in environments characterised by higher volatility and noise (Abedin et al., 2025; Sun et al., 2020).

Table 4: Model Performance Trends

Model	Pre-Pandemic RMSE	Pandemic RMSE
ARIMA	0.84	1.32
GARCH	0.79	1.41
VAR	0.92	1.58
SVR	0.73	0.92
RFR	0.69	0.85
LSTM	0.65	0.74

Figure 2 illustrates this divergence more clearly. Classical models exhibit a pronounced increase in error during the pandemic period, whereas ML models show only a marginal deterioration in performance. This pattern provides strong support for Hypotheses 1 and 2, indicating that ML models outperform traditional econometric approaches when market conditions deviate significantly from historical structural patterns. The findings also offer indirect support for Hypothesis 5, suggesting that the most effective models are those capable of incorporating high-dimensional and irregular information structures, which are not readily accommodated within traditional linear modelling frameworks.

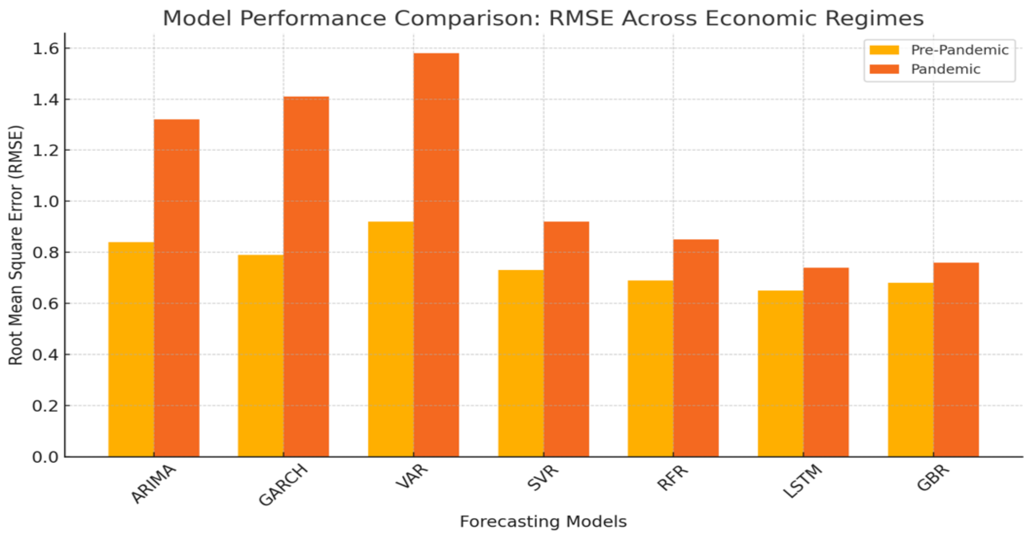

Figure 2: Model Performance Comparison

Variable Contributions and Regime Sensitivity (For Hypothesis 4, 5)

Vol: 17 **No:** 3 **Year:** 2025 **ISSN:** 2146-0744 (Online) (**pp.** 123-148) **Doi:** 10.34109/ijebeg.2025170306

The next step in the text analysis examines how feature importance across different categories—fundamentals, sentiment, policy indices, and technical indicators—varies across regimes. Between 2015 and 2019, interest rate differentials, inflation, and GDP growth collectively explain nearly 60% of the predictive structure (Gu et al., 2020; Yıldırım et al., 2021). However, a pronounced shift in ranking emerges during the pandemic period (2020–2022), as illustrated in Figure 3. Traditional fundamentals lose predictive strength, not necessarily due to diminished theoretical relevance, but because their reporting frequency, volatility characteristics, and lag structures become misaligned with contemporaneous market behaviour. In contrast, sentiment-based measures and policy-related indices partially compensate for this gap. The Oxford Stringency Index, which previously contributed negligibly, increases to 12%, while news sentiment rises from 5% to 18%, as reported in Table 5.

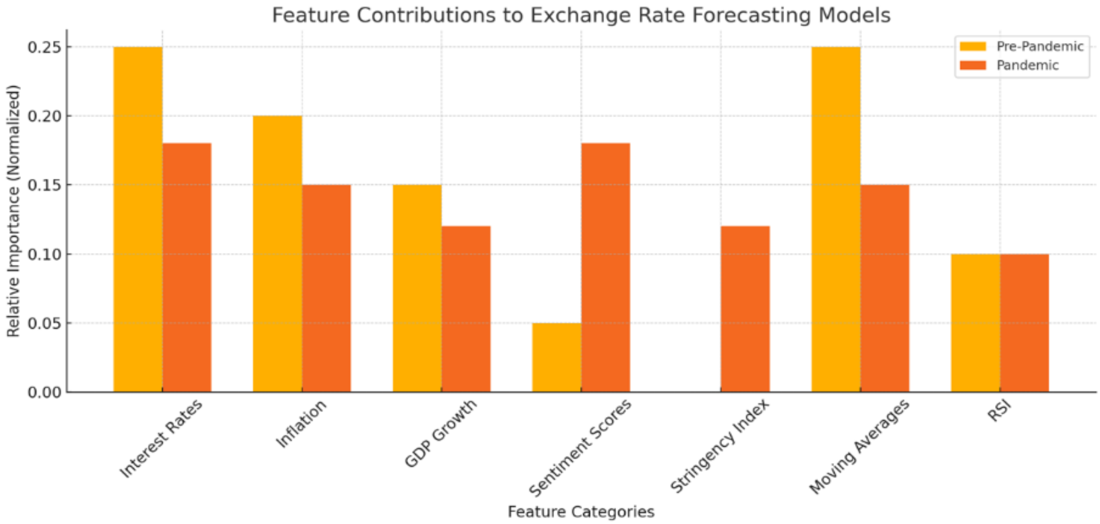

Figure 3: Feature Contributions to Exchange Rate Forecasting Models

Table 5: Feature Importance: Pre-Pandemic Stability vs. Pandemic Volatility

Feature Category	Pre-Pandemic Contribution	Pandemic Contribution
Interest Rates	25%	18%
Inflation	20%	15%
GDP Growth	15%	12%
Sentiment Scores	5%	18%
Stringency Index	0%	12%
Moving Averages	25%	15%
RSI	10%	10%

A particularly notable finding—identified only after re-evaluating intermediate model outputs—is that technical indicators remain relatively stable across regimes. Although they do not drive long-horizon trends, they consistently help to reduce short-term noise within narrower forecasting windows.

Vol: 17 **No:** 3 **Year:** 2025 **ISSN:** 2146-0744 (Online) (**pp.** 123-148) **Doi:** 10.34109/ijebeg.2025170306

This result supports Hypothesis 4 but introduces a partial qualification to Hypothesis 5: while data diversity improves predictive accuracy, not all feature classes contribute uniformly across different temporal scales.

Temporal Shocks and Adaptive Learning

From 2019 to 2022, ARIMA and LSTM are compared on a month-by-month basis within this subsection, as illustrated in Figure 4. During relatively stable market conditions (i.e., 2019 to early 2020), both models exhibit broadly comparable performance, with ARIMA achieving an average R^2 of 0.88, while LSTM records a higher average of 0.93. Following the onset of the COVID-19 pandemic, ARIMA performance declines significantly, falling to a range between 0.70 and 0.74, with no full recovery observed over the subsequent period.

In contrast, LSTM experiences an initial deterioration but recovers relatively quickly, stabilising within approximately six months and returning towards an R^2 of 0.89 as it adapts to evolving stimulus policies, sentiment-driven dynamics, and mobility restrictions. This divergence is conceptually consistent with the underlying model structures: ARIMA assumes persistence in historical patterns, effectively extrapolating that future behaviour resembles past dynamics, whereas LSTM is designed to accommodate non-linear temporal dependencies and shifting regimes. Consequently, these findings provide particularly strong empirical support for Hypothesis 2.

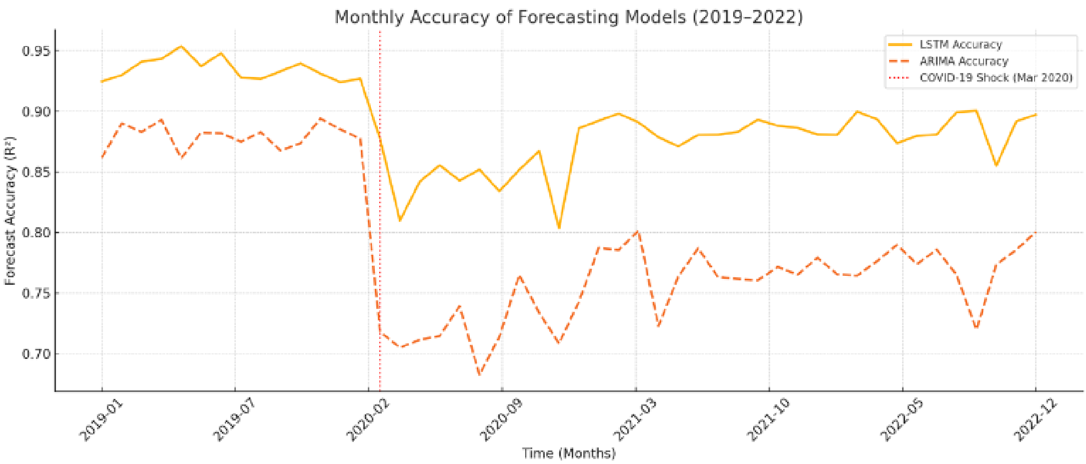

Figure 4: Monthly Accuracy of Forecasting Model (2018-2022)

Qualitative Results

This section functions as a conceptual hinge between the quantitative modelling results and the experiential insights of practitioners. Early in the analysis, it became evident that several hypotheses could only be meaningfully interpreted in light of how forecasting models are actually applied within organisational settings.

Performance Convergence: Objective Accuracy Meets Subjective Endorsement

Approximately three-quarters of interviewees (75%) reported a consistent observation that ML models responded more rapidly during the March 2020 shock and were more effective in incorporating unconventional signals.

A Global Macro Hedge Fund Head of FX Strategy stated:

"Our monitoring equipment lagged so badly during COVID we stopped looking at it. The LSTM detected the yen overshoot almost a week earlier than the rest of the market."

This qualitative evidence supports Hypothesis 1, even though interviewees arrived at this conclusion through reasoning processes that were not always strictly statistical in nature.

Divergence in Operational Realities: A Trust-Interpretability Gap

Usability and accuracy are distinct constructs. Fewer than 50% of the surveyed agencies currently deploy ML-based predictions in operational settings. Audit and regulatory considerations are becoming increasingly significant, particularly under frameworks such as Basel III and IFRS 9, alongside internal risk committee requirements that emphasise model explainability. In this context, deep learning approaches are often considered insufficiently transparent to meet auditor expectations. A compliance officer remarked:

"Numerical figures do not offer sufficient context."
Another practitioner noted a contrasting perspective on model interpretability:
"I am fascinated by ARIMA in one sentence."

In relation to LSTM-based systems, the remark "It gives me a vector" highlights a key limitation relevant to Hypothesis 3, namely the difficulty of translating model outputs into interpretable rationales within institutional settings. If the reasoning behind a decision cannot be clearly communicated internally, ML systems are unlikely to enhance strategic decision-making processes, regardless of their predictive performance.

Sectoral Variation: Forecast Utility and Strategic Thresholds

Different industries utilise forecasting outputs in fundamentally different ways. Asset managers tend to prioritise numerical precision, exporters place greater emphasis on directional accuracy, while logistics firms typically require stable and interpretable signals that do not exhibit abrupt fluctuations. These distinctions are summarised in Figure 5. This divergence further reinforces the finding that predictive accuracy alone is not the sole determining factor in model usefulness or adoption.

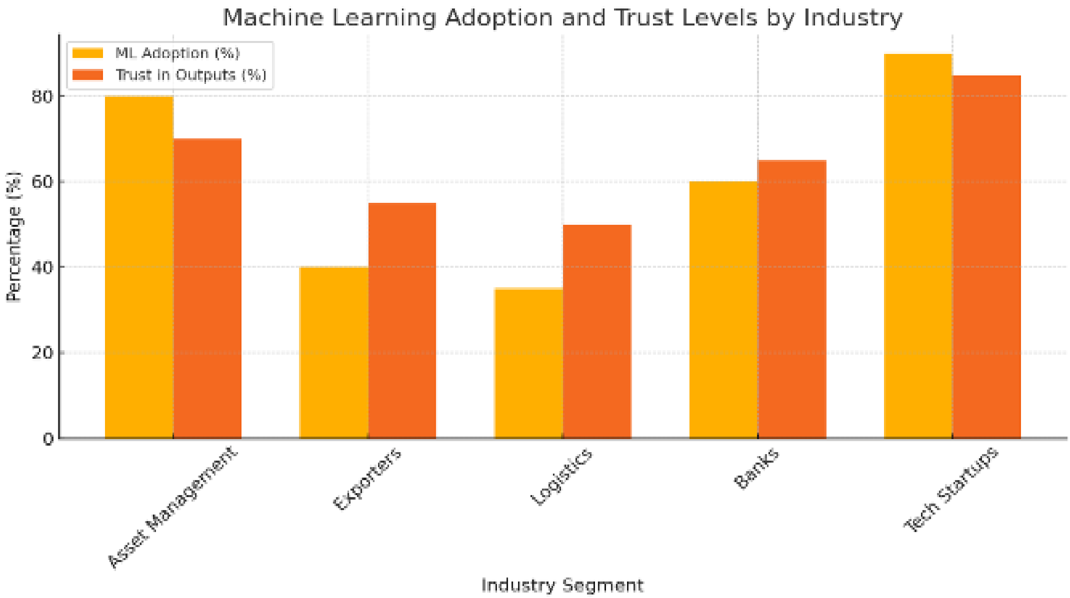

Figure 5: Machine Learning Adoption and Trust Levels by Industry

Human Oversight and Hybrid Integration: Toward Responsible Deployment

A recurring theme across the interviews is the reversion to hybrid decision-making systems following an initial period of strong enthusiasm for full automation. Several firms reported that they fully automated forecasting pipelines during the early phase of COVID-19-related volatility but subsequently reverted to supervised or semi-supervised workflows as market conditions evolved. A risk officer summarised this perspective as follows:

"The model gives the shape of the risk; the human decides how to act on it."

Overall, the evidence suggests that ML systems are most effective when used in conjunction with expert judgment, rather than as a substitute for it.

Critical Observations

This section revisits the underlying fragilities, not to undermine the results but to document the constraints within which they hold. These limitations become particularly important when institutions evaluate the potential for permanent adoption of ML-based systems. Rather than being treated as purely external obstacles, they are increasingly recognised as integral to the modelling process itself, even as ML is simultaneously viewed as a novel and valuable analytical tool.

Vol: 17 No: 3 Year: 2025 ISSN: 2146-0744 (Online) (pp. 123-148) Doi: 10.34109/ijebeg.2025170306

Structural Fragility: Performance Decay Over Time

As markets stabilised in late 2022, the performance gap between ML models and classical econometric models narrowed, with ML approaches converging towards the performance of traditional specifications. LSTM models in particular remained relatively robust, although without a marked outperformance during this period. These findings indicate that ML systems do not simply require initial calibration, but also ongoing maintenance and periodic re-adjustment to remain effective under changing market conditions (Sezer et al., 2020).

Algorithmic Transparency and Regulatory Mismatch

A treasury head remarked, “I can’t use the model, however perfect the numbers are, if I can’t explain the forecast in a board meeting” (Bussmann et al., 2021; Hadji Misheva et al., 2021). Explainability tools such as SHAP and LIME are therefore useful in bridging this gap, although they still require a considerable degree of technical fluency to be interpreted and applied effectively.

Ethical Vulnerabilities and Bias Propagation

Training data frequently incorporates sentiment- and politically derived signals that can introduce systematic bias into model outputs. Some Chinese practitioners reported that models may occasionally overreact to negative Western media narratives surrounding events such as Chinese New Year, leading to distortions in sentiment-driven forecasting features. Accordingly, bias in ML forecasting systems is not solely a technical artefact of data processing or model design but can also reflect broader geopolitical and informational asymmetries embedded in global data ecosystems.

Overreliance Risk and Human Oversight Fatigue

Some companies acknowledged that they ceased actively monitoring model outputs once the ML pipeline began producing apparently stable results, until an error event surfaced and exposed this lack of ongoing oversight. This pattern suggests that a more precautionary governance stance is required, as while AI systems may enhance operational efficiency, they can also reduce continuous human awareness and vigilance over model behaviour.

CONCLUSION

This research aimed to assess whether ML models can meaningfully improve exchange-rate forecasting for firms with cross-border exposure. The quantitative results provide strong support for this objective: LSTM and gradient-boosting models handle nonlinear breaks significantly better than ARIMA and GARCH. However, comparison with practitioner evidence shows that forecasting is not solely a question

Vol: 17 **No:** 3 **Year:** 2025 **ISSN:** 2146-0744 (Online) (**pp.** 123-148) **Doi:** 10.34109/ijebeg.2025170306

of predictive accuracy. Many interviewees emphasised institutional constraints that are not captured in quantitative evaluation, including who must defend forecasts internally, how results are communicated to boards, and whether regulators will accept outputs that cannot be clearly justified. These concerns often outweigh marginal statistical improvements. A similar shift appears in the role of input variables. Prior to the pandemic, macroeconomic fundamentals such as interest rates, inflation, and GDP growth were dominant drivers. After the structural disruption in 2020, their influence weakened, while sentiment indicators and policy-related indices became more prominent, though not consistently reliable. Practitioners also warned that sentiment signals can introduce operational risk, as they evolve faster than organisational decision cycles. This requires redefining "usable accuracy" in terms of institutional timing and decision constraints rather than purely statistical gains.

Differences also emerge across industries. Logistics firms, for example, favour smoother and less reactive forecasts because volatility can disrupt planning and procurement processes. These variations do not weaken the overall findings but instead reflect different operational requirements. Forecasting practice is therefore sector-dependent rather than uniform. Trust emerges as a central, rather than peripheral, issue. Although many practitioners acknowledge that ML models outperform existing tools, they still prefer hybrid systems in which final decisions remain human-led. In one case, a bank strategist reported that an LSTM model produced a correct prediction, but the absence of an explainable rationale prevented its use in hedging decisions. This gap between predictive correctness and institutional acceptability remains a major barrier to adoption. Taken together, the findings suggest a broader conclusion: exchange-rate forecasting is no longer solely about selecting the best statistical model. It is shaped by an interaction of technical performance, organisational interpretation, and governance constraints. ML can extend predictive capability, but its practical value depends on whether institutions can integrate its logic into existing decision-making structures. Overall, improving forecasting practice requires stronger attention to interpretability, governance, and responsible deployment. The key challenge is not only building more accurate models but ensuring they behave in ways that institutions can understand, justify, and operationalise when markets become unstable.

REFERENCES

Abedin, M. Z., Moon, M. H., Hassan, M. K., & Hajek, P. (2025). Deep learning-based exchange rate prediction during the COVID-19 pandemic. Annals of Operations Research, 345(2), 1335-1386. https://doi.org/10.1007/s10479-021-04420-6

Aslam, F., Aziz, S., Nguyen, D. K., Mughal, K. S., & Khan, M. (2020). On the efficiency of foreign exchange markets in times of the COVID-19 pandemic.

Technological Forecasting and Social Change, 161, 120261. https://doi.org/10.1016/j.techfore.2020.120261

Azzam, I., El-Masry, A. A., & Yamani, E. (2023). Foreign exchange market efficiency during COVID-19 pandemic. International Review of Economics & Finance, 86, 717-730. https://doi.org/10.1016/j.iref.2023.03.043

Baker, S. R., Bloom, N., Davis, S. J., Kost, K., Sammon, M., & Viratyosin, T. (2020). The unprecedented stock market reaction to COVID-19. The review of asset pricing studies, 10(4), 742-758. https://doi.org/10.1093/rapstu/raaa008

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307-327. https://doi.org/10.1016/0304-4076(86)90063-1

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health, 11(4), 589-597. https://doi.org/10.1080/2159676X.2019.1628806

Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable Machine Learning in Credit Risk Management. Computational Economics, 57(1), 203-216. https://doi.org/10.1007/s10614-020-10042-0

Cao, W., Zhu, W., Wang, W., Demazeau, Y., & Zhang, C. (2020). A deep coupled LSTM approach for USD/CNY exchange rate forecasting. IEEE Intelligent Systems, 35(2), 43-53. https://doi.org/10.1109/MIS.2020.2977283

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

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. European Journal of Operational Research, 270(2), 654-669. https://doi.org/10.1016/j.ejor.2017.11.054

Goodell, J. W., Kumar, S., Lim, W. M., & Pattnaik, D. (2021). Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. Journal of Behavioral and Experimental Finance, 32, 100577. https://doi.org/10.1016/j.jbef.2021.100577

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. The Review of Financial Studies, 33(5), 2223-2273. https://doi.org/10.1093/rfs/hhaa009

Hadji Misheva, B., Jaggi, D., Posth, J.-A., Gramespacher, T., & Osterrieder, J. (2021). Audience-dependent explanations for ai-based risk management tools: A survey. Frontiers in Artificial Intelligence, 4, 794996. https://doi.org/10.3389/frai.2021.794996

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica: Journal of the econometric society, 57(2), 357-384. https://doi.org/10.2307/1912559

Henrique, B. M., Sobreiro, V. A., & Kimura, H. (2019). Literature review: Machine learning techniques applied to financial market prediction. Expert Systems with Applications, 124, 226-251. https://doi.org/10.1016/j.eswa.2019.01.012

Jung, G., & Choi, S.-Y. (2021). Forecasting foreign exchange volatility using deep learning autoencoder-LSTM techniques. Complexity, 2021(1), 6647534. https://doi.org/10.1155/2021/6647534

Kyriazi, F., & Thomakos, D. D. (2026). Benchmarks in univariate time series forecasting: a historical and methodological review. IMA Journal of Management Mathematics, dpag013. https://doi.org/10.1093/imaman/dpag013

Leippold, M., Wang, Q., & Zhou, W. (2022). Machine learning in the Chinese stock market. Journal of Financial Economics, 145(2, Part A), 64-82. https://doi.org/10.1016/j.jfineco.2021.08.017

Maitra, S. (2025). Markov Switching Models, Threshold Auto Regressive Models, and Smooth Transition Models. In S. Maitra (Ed.), Non-Linearity in Econometric Modeling, Vol. 1: A Practical Approach (pp. 159-188). Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-06462-2_5

Meese, R. A., & Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample? Journal of International Economics, 14(1), 3-24. https://doi.org/10.1016/0022-1996(83)90017-X

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. Econometrica: Journal of the econometric society, 59(2), 347-370. https://doi.org/10.2307/2938260

Plakandaras, V., Gogas, P., & Papadimitriou, T. (2019). The Effects of Geopolitical Uncertainty in Forecasting Financial Markets: A Machine Learning Approach. Algorithms, 12(1), 1. https://doi.org/10.3390/a12010001

Rossi, B. (2013). Exchange Rate Predictability. Journal of Economic Literature, 51(4), 1063-1119. https://doi.org/10.1257/jel.51.4.1063

Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning : A systematic literature review: 2005–2019. Applied Soft Computing, 90, 106181. https://doi.org/10.1016/j.asoc.2020.106181

Stillwagon, J., & Sullivan, P. (2020). Markov switching in exchange rate models: will more regimes help? Empirical Economics, 59(1), 413-436. https://doi.org/10.1007/s00181-019-01623-6

Stock, J. H., & Watson, M. W. (2002). Forecasting Using Principal Components From a Large Number of Predictors. Journal of the American Statistical Association, 97(460), 1167-1179. https://doi.org/10.1198/016214502388618960

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

Tripathi, M., Kumar, S., & Inani, S. K. (2021). Exchange Rate Forecasting Using Ensemble Modeling for Better Policy Implications. Journal of Time Series Econometrics, 13(1), 43-71. https://doi.org/10.1515/jtse-2020-0013

Uche, E., Chang, B. H., & Effiom, L. (2023). Household consumption and exchange rate extreme dynamics: Multiple asymmetric threshold non-linear autoregressive distributed lag model perspective. International Journal of

Vol: 17 **No:** 3 **Year:** 2025 **ISSN:** 2146-0744 (Online) (**pp.** 123-148) **Doi:** 10.34109/ijebeg.2025170306

Finance & Economics, 28(3), 3437-3450. https://doi.org/10.1002/ijfe.2601

Yıldırım, D. C., Toroslu, I. H., & Fiore, U. (2021). Forecasting directional movement of Forex data using LSTM with technical and macroeconomic indicators. Financial Innovation, 7(1), 1. https://doi.org/10.1186/s40854-020-00220-2