

## Machine Learning-Based Exchange Rate Forecasting and Performance Evaluation: A Dual CNN-BiLSTM Framework with Explainable AI for the Pakistani Rupee

Walija Naseer<sup>1</sup> Muhammad Naveed Jan<sup>2</sup> Muhammad Shariq<sup>3</sup>

<sup>1</sup> Research Scholar, Department of Management Sciences, COMSATS University Islamabad, Abbottabad, Khyber Pakhtunkhwa, Pakistan. Email: [waliyanaseer3@gmail.com](mailto:waliyanaseer3@gmail.com)

<sup>2</sup> Assistant Professor, Department of Sciences, COMSATS University Islamabad, Abbottabad, Khyber Pakhtunkhwa, Pakistan. Email: [naveedjan@cuiatd.edu.pk](mailto:naveedjan@cuiatd.edu.pk)

<sup>3</sup> Assistant Professor, School of Management Sciences, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Topi, Khyber Pakhtunkhwa, Pakistan. Email: [muhammad.shariq@giki.edu.pk](mailto:muhammad.shariq@giki.edu.pk)

**ABSTRACT:** In this paper, we design and evaluate a novel dual hybrid architecture based on a combination of CNN and BiLSTM network (CNN-BiLSTM), for predicting exchange rates of the Pakistani Rupee (PKR) vis-à-vis the five key currencies: US Dollar (USD), Euro (EUR), British Pound (GBP), Chinese Yuan (CNY), and UAE Dirham (AED). Two different models with same architectural configuration but with independent training sets are designed in this research: one is a short-term model based on 22 technical indicators extracted from OHLC data of daily time horizon to predict tomorrow's exchange rate, and another is a long-term model based on 26 features derived from six monthly macroeconomic factors: foreign exchange reserves, KIBOR, net foreign assets, inflation, balance of trade, and external debt servicing to predict next month's exchange rate. With the 2010–2025-time span which includes 2018 devaluation, the COVID-19 pandemic, and the 2022-2023 currency crisis events, both models are validated using walk-forward cross-validation approach and explained by SHapley Additive exPlanations (SHAP). In terms of technical model, MAPE stands at 0.08%-0.36% while the directional accuracy reaches 67.5%-72.8%, while the MAPE of the macroeconomic model stands at 1.23%-2.64%, directional accuracy at 64%-72%. The most important short-term determinants recognized by SHAP are trend-strength (ADX, +DI), while for the long-term inflation rate, net foreign asset value, and KIBOR are identified as the determinants. The results in the simulated trading environment show that profits and positive alpha are generated over buy-and-hold benchmarks for both models and all five currency pairs, showing statistical and economic relevance and interpretability of the CNN-BiLSTM approach.

**KEYWORDS:** Exchange Rate Forecasting, Deep Learning, Convolutional Neural Network (CNN), Bidirectional Long Short-Term Memory (BiLSTM), Explainable Artificial Intelligence (XAI), Macroeconomic Factors, Technical Factors

**JEL Classification:** F31, F41, C45, C53, C52, G17.

**Pages:** 371 – 389

**Volume:** 5

**Issue:** 5 (May 2026)

### Corresponding Author

Muhammad Naveed Jan

✉ [naveedjan@cuiatd.edu.pk](mailto:naveedjan@cuiatd.edu.pk)

**Cite this Article:** Naseer, W., Jan, M. N., & Shariq, M. (2026). Machine Learning-Based Exchange Rate Forecasting and Performance Evaluation: A Dual CNN-BiLSTM Framework with Explainable AI for the Pakistani Rupee. *The Regional Tribune*, 5(5), 371-389.

<https://doi.org/10.55737/trt/v-v.420>

## Introduction

Forex forecasting holds an important position in finance because the changes in foreign currencies influence trade competitiveness, capital inflow, inflation rates, and the options for monetary authorities (Nsengiyumva et al., 2025; Akhtar et al., 2022). Foreign exchange market happens to be the largest and most liquid market in the whole world, and the difficulty in its forecasting is both empirical and theoretical (Yao & Tan, 2000). In the case of developing countries like Pakistan, exchange rate forecasting takes on an even more important meaning due to its interaction with debt service, import prices, and reserves (Akhtar et al., 2022).

There are two elements about the structure of exchange rates that make this difficult. First, the foreign exchange series is non-stationary with structural breaks, regime switching, and reversals making generalization of any model impossible across different time frames (Perla et al., 2023). Second, the behaviour of the exchange rate is determined by the joint effects of economics, politics, and psychology in a non-linear manner (Yao & Tan, 2000). The classical econometric model that includes autoregressive integrated moving average (ARIMA), vector autoregression (VAR), and generalized autoregressive conditional heteroskedasticity (GARCH) models is based on the assumption of linearity and stationarity, which is violated in the process of the exchange rate's behaviour (Nguyen et al., 2024). Although GARCH models are useful in modelling time-varying volatility, they do not have the capability of forecasting the directional movement of prices (Nguyen et al., 2024); however, recent research found that machine learning and deep learning models perform better than the classical models in this regard (Qureshi et al., 2023; Escudero et al., 2021).

This development has led to a considerable amount of work using Convolutional Neural Networks (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) networks for currency prediction purposes. CNN layers are responsible for capturing local and short-term patterns as well as de-noising financial time series data (Nguyen et al., 2024; Chandra et al., 2021), while BiLSTM layers process a time series from both ends, making it possible for the network to utilize both the past and future information in order to predict the non-stationary time series (Kim et al., 2025). Given that the two neural networks are able to compensate for each other's shortcomings, hybrid CNN-BiLSTM models have been found to perform better than individual components (Nsengiyumva et al., 2025; Ni et al., 2019).

In spite of this development, there are four interrelated weaknesses in the literature using AI techniques to forecast the exchange rates and in particular the exchange rate of PKR. The first one is the use of scattered inputs for such forecasts, i.e., either only the history of prices (Nguyen et al., 2024; Nsengiyumva et al., 2025), only selected macroeconomic indicators without any technical inputs (Qureshi, 2025), or only technical indicators without any structural indicators (Panopoulou & Souropanis, 2019). In case when both indicators have been used, their application was limited to directional forecasting of a certain currency pair (Yildirim et al., 2021) or did not yield a model separating short-term market dynamics from long-term structural dynamics (Abouzaid & Boussedra, 2025; Biswas et al., 2023). All of them do not comply with the Mundell-Fleming model, according to which the exchange rate is determined by both short-term market dynamics and long-term macroeconomic fundamentals.

Second, emerging market currencies, including the PKR, continue to be underserved: bibliometric analysis of exchange rate forecasting literature showed that there has been a bias towards developed market currency pair studies, with only minimal focus on South Asian and developing countries' currencies (Fang et al., 2023). In Pakistan's case, existing literature is mostly built on classical ARIMA/GARCH approaches (Rashid & Basit, 2021), despite the fact that volatility drivers vary between countries and require tailored models.

Third, deep learning forecasting models are often accused of being “black box” systems. Even though there have been some attempts to address the role of Explainable Artificial Intelligence (XAI) (Abouzaid & Boussedra, 2025; Nsengiyumva et al., 2025), tools for model interpretability, such as SHapley Additive exPlanations (SHAP), have not been widely used in the case of currency forecasting, and at least one comparative analysis has shown that simple and explainable models like logistic regression can produce more accurate results than complex neural networks, thereby questioning the use of black boxes in this sphere (Lopez-Herrera et al., 2025).

Fourthly, the empirical validation of forecasting models is lacking. Most of the studies provide statistical accuracy measures without assessing the profitability of the forecasts in terms of trading or policy relevance (Fang et al., 2023). However, some exceptions do apply (Chihab et al., 2019); however, these are limited to currency pairs of the developed countries.

This research paper bridges these four research gaps by designing and evaluating a dual hybrid CNN-BiLSTM approach for the PKR against the five most important currencies of the world, including the USD, EUR, GBP, CNY, and AED, covering 2010-2025 period. The proposed approach is based on two architecturally identical but independently trained models: a short-term model which relies on 22 daily technical indicators to forecast tomorrow’s exchange rate and a long-term model which utilizes 26 features based on 6 monthly macroeconomic indicators to forecast the exchange rate of the next month. SHAP is used in both models for identifying the most influential drivers of the forecasted results; furthermore, the forecasts generated by them will be backtested using trading simulations. Therefore, the research seeks to explore (i) the influence of technical and macroeconomic variables on the exchange rate dynamics of PKR at their individual horizons, (ii) the relative predictive ability of the dual CNN-BiLSTM model in terms of forecasting both types of horizons, and (iii) the applicability of the generated forecasts in the form of a simulated trading strategy.

The structure of this paper is organized as follows. The next section provides the literature review on exchange rate forecasting. Section 3 describes the theoretical background of the dual-model approach. Methodology is described in Section 4. Results are presented in Section 5. The next two sections cover the implications and conclusions respectively.

## Literature Review

### From Econometric Models to Artificial Intelligence

Economic forecasting using exchange rates has always been done using linear econometrics models. ARIMA forecasting model is used to make forecasts of changes in a time series through differencing of a non-stationary time series to make it stationary and has been extensively used for long-term forecasting (Akhtar et al., 2022). GARCH family of models are used in conjunction with ARIMA forecasting and are used to predict time-varying volatility (Nguyen et al., 2024). However, both of these models depend on the assumptions of linearity and stationarity, they are highly parametric and are prone to accumulate errors due to their long forecasting horizon (Nguyen et al., 2024). According to Sonmez & Birim (2024), such autoregressive models have limitations in predicting nonlinear dependencies.

This benefit was exhibited by early studies using ANNs. According to Panda and Narasimhan (2007), ANN models were more successful than econometric/time series models in exchange rate forecasting. Qureshi (2025) revealed that RF and SVM were superior to ARIMA and random walk in Pakistan, and RF proved to be the most accurate one. Regarding Pakistan, Qureshi et al. (2023) established that MLP and nonlinear ARIMA models were superior to linear econometrics models in the case of the real effective exchange rate, while Escudero et al. (2021)

discovered that LSTM was more accurate for EUR/USD forecasting in the short term while Elman RNN was better for longer horizons. According to a systematic review carried out by Sezer et al. (2020), advancements in computing power and the amount of data available have enabled deep learning to surpass classical machine learning in time series prediction in the financial sector, especially in terms of directional prediction. The ARIMA-ANN hybrid model of Perveen et al. (2024) and a bibliometric analysis by Fang et al. (2023) support the overall trend of a move towards hybrid and data-intensive prediction models in the field.

Evidence at an international level further confirms this trend. According to Dautel et al. (2020), who compared LSTM and GRU with feedforward networks and simple RNN for some major currency pairs, deep recurrent models yielded better performance in predicting prices, but simple models were capable of achieving the same performance levels as far as trading gains were concerned. Wang et al. (2021) have developed a CNN-TLSTM model for USD/CNY forecasting, which has performed better than MLP, CNN, RNN, LSTM, and CNN-LSTM models, and have attributed this improvement to the temporal gate regulation within the TLSTM structure and suggested inclusion of macroeconomic variables to further enhance the accuracy of their model.

### CNN, LSTM, and BiLSTM Architectures

The ability to detect localized features through convolution and pooling makes CNNs suitable for structured data analysis and has enabled them to adapt to financial time series in spite of having been developed to perform image recognition (Chandra et al., 2021). The local connectivity and weight sharing nature of CNNs reduces the number of parameters compared to fully connected networks without compromising on accuracy (Nsengiyumva et al., 2025), and they work best in detecting short-term dependencies associated with short-term market dynamics or events (Nsengiyumva et al., 2025). Being incapable of modeling long-range temporal dependencies by themselves, CNNs have been increasingly used together with recurrent neural networks. Ni et al. (2019) used CNN and RNN in the form of C-RNN model and obtained lower prediction error than using either of them alone for nine currencies. Nguyen et al. (2024) noted that CNN-LSTM model had 97% accuracy and 15% return in one experiment.

LSTMs have utilized gate-enabled memory cells such as input, output, and forget gates in order to counteract the vanishing/exploding gradient problems facing traditional recurrent networks to achieve the ability to remember over long sequences (Ayitey et al., 2022). This characteristic has allowed LSTM models to be more desirable for forecasting under volatile conditions, and stacked/deep LSTMs have been able to improve forecasting accuracy for highly volatile foreign exchange markets (Ayitey et al., 2022). Abouzaid & Boussedra (2025) revealed that LSTM was more efficient than MLP and RF in forecasting EUR/USD with an  $R^2$  value of 0.9234.

The BiLSTM network, first introduced by Schuster & Paliwal (1997; as cited in Sonmez & Birim, 2024), builds upon the basic LSTM network by processing a sequence of data points from both ends of the sequence, enabling the model to make use of forward and backward contexts in every instance (Nsengiyumva et al., 2025). Such a bidirectional mechanism is particularly suitable for exchange rates data with its heavy tails, volatility clustering and leverage effects that can only be accounted for on the basis of historical and contextual observations (Kim et al., 2025). García et al. (2024) have demonstrated that BiLSTM gives rise to better forecasts compared to unidirectional LSTM in 83% of currency pairs considered, with forecast errors reduced by 7% to 47%. Later studies extended the BiLSTM approach by combining it with feature selection techniques (Liu et al., 2024). In line with the current research problem, Nsengiyumva et al. (2025) have incorporated the capability of feature extraction provided by CNN and bidirectional temporal modeling in BiLSTM into a single model CNN-BiLSTM and showed that it performed better than individual CNN and BiLSTM models for multivariate exchange rate forecasts.

## The Differentiated Role of Technical and Macroeconomic Indicators

Another constant element in these papers is that there is a difference in the predictive capabilities of the technical and macroeconomic indicators due to their different horizons of predictions. Specifically, Qureshi (2025) discovered that foreign exchange reserves and interest rates have the highest effect on exchange rate among all fundamental indicators, which is explained by the role played by the reserves in terms of conveying the market's confidence and interest rates being responsible for regulating the capital inflow. Also, Korol (2014) revealed that a trade surplus is related to currency appreciation and a deficit is related to its depreciation; Nwanne and Onyekachi (2015) showed that the higher debt servicing abroad leads to higher demand of foreign currencies and depreciation of the local currency; Mancini Griffoli et al. (2015) identified net foreign assets as a stabilizing factor that helps in appreciating currency. Yildirim et al. (2021) considered this element directly by building LSTM models on both sets of indicators and integrating them using decision rules; the integrated model had 73.09% of directional accuracy for one day ahead.

From a technical standpoint, Panopoulou & Souropanis (2019) observed that metrics like Moving Averages, MACD, and Momentum are helpful in identifying short-term directional shifts, Wagdi et al. (2023) demonstrated that the impact of RSI, CCI, and other momentum metrics on neural network predictions is significant, and Yildirim et al. (2021) and Perla et al. (2023) found that metrics based on volatility and stochastic oscillator increase model flexibility during changes in the market environment. The two pieces of evidence – fundamentals affecting long-term structural movement and technical indicators reflecting short-term market sentiment – collectively explain why the design of the model should involve separate treatment of the two groups of indicators.

## Interpretability and the Practical Validation Gap

The second body of literature looks at the "black-box" argument against deep learning predictions. In their study comparing seven machine learning models on directional prediction, Lopez-Herrera et al. (2025) found that interpretable models, such as logistic regression, outperform neural networks, showing that increased complexity does not automatically mean better results. Using a cascading architecture designed to increase model explainability, Pornwattanavichai et al. (2022) created BERTFOREX; however, it does not reach the input level of interpretability. Using feature selection and visualization technique Grad-CAM on LSTM, CNN, and transformer models predicting RMB/USD, Meng et al. (2024) demonstrate how explainable deep learning can deliver accuracy and explainability simultaneously by discovering that trade volume and cross-currency rates are the most influential variables. However, SHAP-based explainability is rarely used in currency forecasting and never used for PKR forecasting.

Similar to the aforementioned issue of statistical accuracy of forecasters, the question of the practical usefulness of the economic signals generated by accurate forecasts also lacks development. Fang et al. (2023) pointed out that there is very limited research on the transformation of forecasting models into trading systems, with notable exceptions like the Random Forest Probit trading model of Chihab et al. (2019) only applying to developed markets without any relation to the PKR.

## Pakistan-Based Evidence

The studies conducted on Pakistan so far have been based primarily on the classical econometric models. In particular, Akhtar et al. (2022) examined the trend of depreciation of PKR against USD, ranging from 21.71 in 1990 to 176 by 2021, and found that ARIMA accounts for the trend, but the machine learning models are more accurate. Similarly, Qureshi et al. (2023) came to the same conclusion by comparing MLP and nonlinear ARIMA

models. Jabeen et al. (2022) utilized GARCH models to establish that announcements of news on macroeconomics have immediate impact on PKR volatility, whereas Rashid and Basit (2021) discovered that volatility of foreign reserves, government expenditure, and terms of trade affect PKR volatility in the selected Asian countries. Perveen et al. (2024) found that hybrid ARIMA-ANN models performed better than either individual model when it came to accounting for linear and non-linear aspects of PKR. Finally, Asadullah et al. (2021) combined univariate models with the multivariate NARDL model based on reserves, trade balance, inflation, and interest rate and found that a combination of NARDL and Naïve models provided the lowest MAPE.

The cumulative Pakistan-specific literature is evidence of a transition to AI techniques but not of the combination of hybrid deep learning frameworks with both technical and macroeconomic variables in one study nor of explainability techniques nor of forecast validation via simulation trading. It is here that the current study fills these gaps.

### Theoretical Framework

This research project uses the dual-horizon model which is based on the combination of two different theories – the theory of economics and fundamentals, the basis of which is the Mundell-Fleming Model, and the behavioural and price-pattern theory which can be explained with the help of Dow Theory and Elliott Wave Theory.

Mundell-Fleming model was designed by Robert Mundell and J. Marcus Fleming at the beginning of the 1960s and is the fundamental theory which establishes the link between the economic policy and the movements of exchange rates in an open economy. According to the theory, the exchange rates in the flexible exchange rates regime increase with the increase of the interest rates and decrease with the rise of the money supply. Such reasoning leads to the rationale behind the selection of the six macroeconomic variables included in the long-run model, which includes foreign exchange reserves and net foreign assets as proxies for external buffers and market confidence, KIBOR as a proxy for the interest-rate channel, inflation as a proxy for the purchasing-power channel, balance of trade as a proxy for the current-account channel, and external debt servicing as a proxy for the foreign currency demand channel.

Dow theory, created by Charles Dow and William Peter Hamilton and often viewed as a basis for technical analysis (Wagdi et al., 2023; Neely & Weller, 2011), provides the complementary short-horizon rationale. Though designed for stock market, its primary idea – that movements happen along definable trends and prices in themselves contain forecasting signals regardless of fundamentals – has been applied to other financial markets, such as foreign currency exchange (Neely & Weller, 2011). Survey data in this regard suggests that technical analysis surpasses fundamental one in terms of forecasting accuracy on short horizon (Neely & Weller, 2011), hence the choice of trend, momentum, and volatility indicators as inputs for short-term forecasting rather than any macroeconomic ones. Elliott wave theory proposed by Ralph Nelson Elliott in 1930s extends this idea further through the suggestion that price movements occur in five and three-wave cycles, hence containing forecasting signals on future movement or its reversal (D'Angelo & Grimaldi, 2017). In combination, the Dow Theory and Elliott Wave Theory support the employment of trend-following, momentum, and directional movement indicators in the technical-indicator framework for picking up signals of continuations and reversals of daily PKR price movements.

Dow Theory and Elliott Wave Theory are not alternative theories explaining the same phenomenon; rather, they describe behaviour at different levels, which represents the key theoretical premise underpinning the research design of this paper – combining daily technical indicators with monthly macroeconomic fundamentals in one and the same pool risks undermining the role of each, as they work at different frequencies and with different mechanisms. Building two CNN-BiLSTM models in parallel, one using only technical indicators at daily frequency,

and another using only macroeconomic indicators at monthly frequency, helps evaluate and interpret the role of each theoretical channel independently, whereas results from the two models help address the overarching theoretical question of whether the exchange rate dynamics of an emerging market like Pakistan can be jointly explained by the market microstructure dynamics and macroeconomic fundamentals.

Theoretical framework of a dual nature also has implications for the interpretation of cross-currency results presented in this paper. Since the PKR regime should be classified as a managed float regime and not a free float or hard peg (Corzo et al., 2025), short-term technical predictability is expected to be highest for those currencies whose regime characteristics support trendiness, such as the AED, which is pegged to the USD, and lower for those whose regime experiences wider swings as a result of policy divergence, such as the EUR and GBP. On the other hand, the macroeconomic channel would be informative against the USD as the main currency in which Pakistan’s foreign reserve and debt stocks and its trade transactions are denominated.

Methodology  
Research Design and Modelling Framework

This research applies a quantitative empirical approach using two parallel models with an identical architecture but trained individually. Model 1 – technical model, is based on 22 technical indicators generated by OHLC price and calculated daily and predicting the closing exchange rate for the following day. This model draws upon 4,171 days of trading (January 2010-December 2025), which are based on the Dow Theory and Elliott Wave Theory. Model 2 – macroeconomic model, is generated by 26 features extracted from 6 monthly macroeconomic variables predicting the closing exchange rate for the following month. The model draws upon 174 months of observations (July 2010-November 2025), based on the Mundell-Fleming model. The common architecture of models allows comparing two theoretical channels and applying them independently to five PKR currency pairs: USD, EUR, GBP, CNY, and AED.

Table 1  
Summary of the Dual CNN-BiLSTM Model Configuration

| Model                   | Input Variables                            | Frequency | Horizon    | Observations                   | Theoretical Basis               |
|-------------------------|--------------------------------------------|-----------|------------|--------------------------------|---------------------------------|
| Model 1 (Technical)     | 22 technical indicators from OHLC data     | Daily     | Next day   | 4,171 days (Jan 2010–Dec 2025) | Dow Theory; Elliott Wave Theory |
| Model 2 (Macroeconomic) | 26 features from 6 macroeconomic variables | Monthly   | Next month | 174 months (Jul 2010–Nov 2025) | Mundell–Fleming model           |

Data, Sampling, and Preprocessing

The data cover a period of 15 years (2010-2025), purposefully selected to subject the two models to several economic regimes, such as the post global financial crises recovery period, the 2018 PKR devaluation, the COVID-19 shock, and the 2022-2023 currency crises, instead of being confined to one homogeneous regime. The daily OHLC exchange rate data for all five currency pairs were obtained from Investing.com; after discarding non-trading days, the number of daily observations collected for each currency pair amounted to 4,171. The monthly data of the six macroeconomic variables FER, KIBOR, NFA, INF, BOT, and DS were procured from the State Bank of Pakistan (SBP) for July 2010-December 2025 (186 months); in this model, the monthly exchange rate is the closing OHLC

exchange rate observed on the last trading day of each month. The quarterly external debt data were converted into monthly using linear interpolation. A non-probability sampling method, purposive sampling, was employed where five currency pairs and six macroeconomic variables were selected based on theoretical importance to Pakistan, and observations in each series were arranged chronologically, in accordance with the sequence characteristic of time series data.

As part of a data quality test, it has been discovered that 10-13% of OHLC rows contained a structural inconsistency where the close price was fractionally out of the high-low range, which is an artefact caused by the differences between intra-day interbank quotes and end-of-day fixing prices, and would have distorted range indicators such as ATR, Bollinger Bands, and the Stochastic Oscillator. It was fixed by setting  $Low = \min(Low, Close)$   $High = \max(High, Close)$

to make sure that  $High \geq Close \geq Low$  for all 4,171 observations. There were no missing values in the daily series after excluding non-trading days, and the only missing data adjustment was the conversion from quarterly to monthly data for external debt. Observations associated with the 2022-2023 PKR crisis have not been excluded as outliers as they are indeed structural market events; however, these data points are weighted down in gradient descent using the Huber loss function in lieu of MSE, acting as MSE for small residuals and as MAE for large residuals.

All computations were performed using Google Colab in Python, with Pandas and NumPy for data management, ta for construction of technical indicators, TensorFlow and Keras (with Adam optimization and early stopping), scikit-learn's StandardScaler for normalization, scikit-learn metrics and Matplotlib for evaluation and visualization, and SHAP for explainability.

## Variables and Feature Engineering

There are 22 independent variables in the technical model that fall into four broad classes: trend (SMA, EMA, MACD Line/Signal/Histogram, ADX, +DI, -DI), momentum (Momentum, Stochastic %K/%D, Rate of Change, Commodity Channel Index, Williams %R, RSI, TRIX), volatility (ATR, Bollinger %B, Bollinger Band Width), and return-based own-momentum features (1-, 5-, and 10-day logarithmic returns). Instead of absolute price levels, all features were converted to stationary and bounded versions. For instance, SMA and EMA were used in percentage terms relative to the closing price, ATR was scaled to the price level, Bollinger Bands were in bounded form, and MACD indicators were scaled by the closing price in order for the model to learn the relative price changes rather than the absolute price levels.

The 26 attributes used in the macroeconomic model are based on six attributes with their present levels, one- and three-month lags (to capture the one-to-three-month transmission lag of changes in SBP policy and balance-of-trade effects), and log-change attributes, together with one- and three-month momentum attributes of the exchange rate. Foreign exchange reserves are the total foreign currency reserves of the central bank; KIBOR is the reference interbank rate that represents liquidity conditions; net foreign assets consist of foreign assets minus foreign liabilities; inflation is the percent change in the consumer price index; balance of trade is the difference between exports and imports; and external debt service consists of principal and interest payment to foreign creditors (Mancini Griffoli et al., 2015; Korol, 2014; Nwanne & Onyekachi, 2015).

In both models, the dependent variable is the log return, not the price level,

$$y_t = \ln\left(\frac{close_{t+1}}{close_t}\right)$$

Log returns are close to stationarity, symmetry, and additivity, which makes optimization easier and prevents non-stationarity that would occur due to predicting PKR levels in absolute value throughout the sample period where there was significant depreciation of the currency. The predicted log returns are then transformed to PKR levels as follows, meaning that all presented error measures are in PKR values, or percentages in the case of MAPE and SMAPE.

$$\widehat{Close}_{t+1} = Close_t \times e^{\hat{y}_t}$$

### Data Splitting, Normalization, and Sequence Construction

Temporal splits were performed on both the datasets. In the case of technical modeling, 70% (Jan-2010-Dec-2022) of the data set was used for training, 6% (Jan-December 2023) for validation, while the rest 15% (Jan-2024-Dec-2025) was kept aside for testing. In case of macroeconomic modeling, 65% (Oct-2010-Dec-2020) of the data set was kept for training, 13% (January 2021-Dec 2022) for validation, and the remaining part (July 2023-Nov 2025) was kept for testing. Both the input and output variables were z-score standardized (StandardScaler) in each walk-forward validation/test process, where the mean and standard deviation were calculated using the train fold data alone. The reason for not opting for the Min-Max scaler is that the post training data could lie out of range in Min-Max Scaling which we observed during USD/PKR appreciation between 2022 and 2023.

Due to the requirement that the CNN-BiLSTM model needs three-dimensional tensor (samples, timesteps, features), a sliding window approach is employed to transform the standardized features into overlapping sequences. In terms of the technical model, the lookback window used is 10 days due to the fact that most of the momentum and oscillators indicators used (RSI, Stochastic Oscillator, Williams %R, ATR) themselves have lookback windows of 10 to 14 days, giving a total of 3,370 training, 240 validation, and 502 testing sequences for each currency pair.

### Model Architecture and Training

The models can be presented in a linear equation format just for the sake of clarifying the set of predictors, although non-parametric in nature:

Macroeconomic model:

$$ER_{t+1} = \beta_0 + \beta_1 \times FER_t + \beta_2 \times KIBOR_t + \beta_3 \times NFA_t + \beta_4 \times INF_t + \beta_5 \times BOT_t + \beta_6 \times DS_t + \epsilon_t$$

Technical model:

$$\begin{aligned} ER_{t+1} = & \beta_0 + \beta_1 \times SMA\ Dev_t + \beta_2 \times EMA\ Dev_t + \beta_3 \times MACD_t \\ & + \beta_4 \times MACD\ Signal_t + \beta_5 \times MACD\ Hist_t \\ & + \beta_6 \times ADX_t + \beta_7 \times (+DI)_t + \beta_8 \times (-DI)_t + \beta_9 \times TRIX_t \\ & + \beta_{10} \times CCI_t + \beta_{11} \times ATR\ \%_t + \beta_{12} \times BB\ width_t + \beta_{13} \times BB\ \%_t \\ & + \beta_{14} \times MOM_t + \beta_{15} \times STOCH\ \%K_t + \beta_{16} \times STOCH\ \%D_t \\ & + \beta_{17} \times ROC_t + \beta_{18} \times \%R_t + \beta_{19} \times RSI_t + \beta_{20} \times Log\ Ret1_t \\ & + \beta_{21} \times Log\ Ret5_t + \beta_{22} \times Log\ Ret10_t + \epsilon_t \end{aligned}$$

The real forecasting models are non-parametric hybrid networks which have the following structure:

$$ER_{t+1} = f_{\theta} \left( BiLSTM \left( CNN(\hat{Z}_t) \right) \right)$$

in which  $\hat{Z}_t$  is a windowed input tensor of either the set of macroeconomic or technical features, the CNN layer finds local temporal features, the BiLSTM layer learns forward and backward sequence dependencies and the

$f_{\theta}$  is a dense regression layer providing final output. The regularization techniques used for both models include: dropout layers (ratio=0.4) after each of the main blocks, L2 weight regularization with coefficient 0.0001 for all convolutional and recurrent layers, Layer Normalization was chosen over Batch Normalization due to the stability of it on financial time series, after each Conv1D layer. Only the number of dimensions of the input and the number of units in BiLSTM differ in these two models: the technical model has around 38,000 parameters, whereas the macroeconomic model, owing to its limited training set of 113 sequences, has approximately 22,000 parameters. Both models employ Adam optimization method with early stopping and ReduceLROnPlateau, as well as Huber loss instead of MSE.

### Validation, Explainability, and Evaluation Metrics

These models undergo validation using walk-forward cross-validation with six folds where the training window progressively increases over time and the model gets tested on unseen future data. This method of validating the model is more appropriate for non-stationary financial time-series as compared to doing a simple train/test split. For the macroeconomic model, the test window used during each fold of walk-forward validation is 12 months, due to its low observation frequency. The walk-forward validation windows contain the devaluation phase of PKR from 2021 to 2022.

Model interpretability is achieved via SHapley Additive exPlanations (SHAP), which is a game theory approach that associates each prediction to the effect of each variable used as an input, providing a means for either global or local interpretation. Model forecast accuracy is measured by Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), Symmetric MAPE (SMAPE), Directional Accuracy (DA), and Pearson correlation coefficient ( $r$ ) after back-transforming log returns predictions to PKR. Lastly, in both models' forecasts, a simulation-based trading strategy is employed to create a buying signal in case the forecast exchange rate exceeds the prevailing exchange rate, and a selling/holding signal otherwise, from an initial capital of PKR 1,000,000 with a transaction cost of 0.1%, and measured through Total Return, Annualized Return, Alpha, Sharpe Ratio, Maximum Drawdown, number of trades, and win rate.

## Results And Discussion

### Descriptive Characteristics of the Test Samples

Within the test period for the technical model (January 2024 to December 2025, 512 daily observations per currency pair), exchange rate volatility varies tremendously among the various currencies. The EUR/PKR and GBP/PKR have the highest standard deviation (14.71 and 13.12 PKR) and daily return standard deviation (0.898% and 0.903%), respectively, given that both euro and sterling follow a free-floating system due to the monetary policy differences between SBP, the European Central Bank, and the Bank of England. AED/PKR and USD/PKR have much lower standard deviations (0.66 and 2.31 PKR, respectively), owing to the pegging of the former to the latter. For the entire 174-month period covered by the macroeconomic model, the inflation rate is the most positively skewed measure (skewness 1.70, range 1.32%-38.00%) due to the inflationary experience of 2022-2023; KIBOR attains a maximum value of 23.22% in comparison to an average of 11.12%, which is a consequence of SBP's response to tighten the monetary stance; the standard deviation of NFA as compared to its mean is the highest for the data set; foreign exchange reserves vary between \$7.99 billion and \$27.07 billion; and the balance of trade is consistently negative.

### Short-Term Technical Model: Forecasting Accuracy

Table 2 provides the results for the last test-set performance for the technical model. The minimum value of MAPE is 0.0849% for USD/PKR while the maximum value is 0.3574% for EUR/PKR. The directional accuracy ranges

from 67.5% (EUR/PKR) to 72.8% (AED/PKR), both of which are greater than the 50% benchmark set by pure guessing.

Table 2

Final Test-Set Performance — CNN-BiLSTM Technical Indicators Model (Jan 2024–Dec 2025, N = 512)

| Currency | MAE (PKR) | RMSE (PKR) | MAPE (%) | DA (%) | r      |
|----------|-----------|------------|----------|--------|--------|
| AED/PKR  | 0.0750    | 0.1731     | 0.0984   | 72.8   | 0.9646 |
| CNY/PKR  | 0.0643    | 0.1109     | 0.1652   | 69.1   | 0.9826 |
| EUR/PKR  | 1.1064    | 1.5363     | 0.3574   | 67.5   | 0.9945 |
| GBP/PKR  | 1.2952    | 1.7674     | 0.3562   | 69.9   | 0.9909 |
| USD/PKR  | 0.2377    | 0.3999     | 0.0849   | 71.2   | 0.9851 |

USD/PKR and AED/PKR register the lowest forecast error and maximum directional accuracy, a logically consistent result due to the USD peg of the AED, which makes the two pairs react in the same way to technical indications and also makes them trendier than freely floating European currencies. EUR/PKR and GBP/PKR display higher MAEs (1.11 and 1.30 PKR respectively) but lower percentage errors (0.36% and 0.36% respectively) along with the maximum Pearson correlations (0.9945 and 0.9909 respectively), which implies that the model closely follows the depreciating trend of the PKR with respect to these currencies despite their higher price levels.

The ordering holds true according to walk-forward cross-validation through six folds, where AED/PKR and USD/PKR again have the lowest average MAPE at 0.3060% and 0.3116%, respectively, compared to 0.4807% and 0.5628% for EUR/PKR and GBP/PKR, respectively; yet, average errors are consistently higher than the test set error because the walk-forward intervals include the 2022-2023 PKR crisis, in which case USD/PKR appreciated from around PKR 180 to PKR 285 in the same period. Instead of being evidence of poor performance by the model, decreasing and then increasing accuracy in each fold proves that the model is able to adapt to the crisis data, which is the very purpose of walk-forward validation. Training and validation loss lines converge without a consistent gap between them, and early stopping occurs at around 40-60 epochs, while residuals are symmetric around 0, which suggests that the regularization strategy comprising of dropout, L2, and Layer Normalization effectively prevents overfitting.

### Long-Term Macroeconomic Model: Forecasting Accuracy

The results of the final test-set performance for the macroeconomic model are displayed in Table 3. The values for MAPE are 1.2268% (AED/PKR) and 2.6355% (CNY/PKR); those for the measure of directional accuracy are 64.0% (AED/PKR, USD/PKR) and 72.0% (EUR/PKR, GBP/PKR); and those for Pearson correlation are 0.5664 (CNY/PKR) and 0.8359 (EUR/PKR).

Table 3

Final Test-Set Performance — CNN-BiLSTM Macroeconomic Variables Model (Jul 2023–Nov 2025, N = 26)

| Currency | MAE (PKR) | RMSE (PKR) | MAPE (%) | DA (%) | r      |
|----------|-----------|------------|----------|--------|--------|
| AED/PKR  | 0.9484    | 1.6290     | 1.2268   | 64.0   | 0.6402 |
| CNY/PKR  | 1.0266    | 1.2619     | 2.6355   | 68.0   | 0.5664 |
| EUR/PKR  | 6.2946    | 8.6717     | 2.0638   | 72.0   | 0.8359 |
| GBP/PKR  | 8.1635    | 10.9436    | 2.2875   | 72.0   | 0.7210 |
| USD/PKR  | 3.6349    | 5.8053     | 1.2786   | 64.0   | 0.6644 |

True to expectation because of the nature of the macroeconomic model being a monthly, structural model, both error size and correlation coefficient are lower compared to the technical model because the macroeconomic model is meant to account for movements at the regime level as opposed to monthly accuracy. Nevertheless, the directional accuracy for both models is about the same (64-72% versus 67.5-72.8%). The latter is the most important one from an economic perspective when making decisions concerning positioning. CNY/PKR shows the poorest fit, just like expected due to the Yuan's managed float regime and other external factors not captured by Pakistan's macroeconomic regime. EUR/PKR and GBP/PKR show the best fit.

The results from walk-forward testing show that there is consistency between the average MAPE value from six folds (1.90%-2.67%) and the MAPE value in the final test set. However, when analyzing the MAPE value in the 2022 PKR crisis-fold separately (see Table 4), it is evident that the MAPE value in this one fold is higher by 2.9% to 4.2%.

**Table 4**

*Fold 6 (2022 PKR Crisis) vs. Folds 1–5 Average MAPE — Macroeconomic Model*

| Currency | Fold 6 MAPE (%) | Folds 1–5 Avg. (%) | Difference (pp) |
|----------|-----------------|--------------------|-----------------|
| AED/PKR  | 4.77            | 1.52               | +3.25           |
| CNY/PKR  | 6.11            | 1.89               | +4.22           |
| EUR/PKR  | 5.94            | 1.88               | +4.06           |
| GBP/PKR  | 5.05            | 2.19               | +2.86           |
| USD/PKR  | 4.84            | 1.32               | +3.52           |

This decay can be precisely linked to the structural break caused by the structural break resulting from the movement in USD/PKR exchange rate from about PKR 176 to PKR 224 during the course of a year on account of the uncertainty surrounding the IMF program, the global commodity price shock following the Russia-Ukraine war, low reserves, and political instability. The forecast accuracy improves as the post-crisis data becomes available in the training window, again, conforming to the principles of walk forward framework rather than showing any structural weaknesses in the framework itself. Similar to the technical model, the losses during training and validation for the macroeconomic model become close, and the early stopping happens in around 40-130 iterations.

### Explainability: SHAP Results

Based on the SHAP analysis of the USD/PKR technical model, the Average Directional Index (ADX), Stochastic %D, Positive Directional Indicator (+DI), Commodity Channel Index (CCI), and Bollinger Band Width are the top five most important features, respectively. As seen from the beeswarm distribution, high ADX and directional-movement values mean high forecast reliability, and the three return-based features – 1-day, 5-day, and 10-day log returns – have less influence on the forecasts of the model compared to other features. Thus, the model's forecasts seem to be based mainly on the strength and directional movement, rather than the continuation of returns, providing an evidence for the basis of the technical model according to the Dow Theory, and supporting the results obtained by Panopoulou and Souropanis (2019).

The SHAP analysis for the USD/PKR macroeconomic model indicates that current inflation is the major factor, followed by one-month lagged Net Foreign Assets, three-month lagged inflation, three-month lagged KIBOR, and current KIBOR; rate-of-change factors in the balance of trade, inflation, KIBOR, and external debt add additional explanatory power. These findings are in line with the Mundell-Fleming transmission mechanism described in Section 3: Inflation depreciates the currency through the purchasing-power channel; KIBOR's counterbalancing

factor works through the capital-flow channel, while the importance of lagged variables specifically reflects the delayed impact of monetary policy and reserve position adjustments on the exchange rate. The importance of lagged Net Foreign Assets further underlines the importance of Pakistan's external position, which depends on the inflows from the IMF, foreign financing, and debt servicing, as the first order determinant of PKR stability.

### Simulated Trading Performance

However, the total return of the technical strategy simulated on a daily basis (refer to Table 5) shows positive results for all five currency pairs, starting from +11.60% for USD/PKR and up to +130.94% for GBP/PKR, positive alpha for all cases, starting from +11.76% for USD/PKR and ending with +124.36% for GBP/PKR, as well as Sharpe ratio from 1.789 to 4.5

**Table 5**

*Simulated Daily Trading Performance — CNN-BiLSTM Technical Indicators Model (512 Trading Days)*

| Currency | Total Return (%) | Ann. Return (%) | Alpha (%) | Sharpe | Max DD (%) | Win Rate (%) |
|----------|------------------|-----------------|-----------|--------|------------|--------------|
| AED/PKR  | +22.68           | +10.59          | +22.83    | 1.869  | -2.46      | 78.91        |
| CNY/PKR  | +43.35           | +19.39          | +40.84    | 2.846  | -1.72      | 85.48        |
| EUR/PKR  | +117.64          | +46.63          | +109.58   | 4.106  | -3.05      | 85.04        |
| GBP/PKR  | +130.94          | +50.98          | +124.36   | 4.534  | -2.32      | 90.51        |
| USD/PKR  | +11.60           | +5.55           | +11.76    | 1.789  | -1.63      | 82.73        |

The biggest profits are obtained in the case of EUR/PKR and GBP/PKR, which were the two currency pairs showing the strongest depreciation of PKR over the analysis period, and hence presenting the greatest possibilities for trend-following; however, the system is also highly profitable on the relatively stable USD/PKR and AED/PKR currency pairs.

The trading approach under the macroeconomic model simulation on a monthly basis (Table 6) is benchmarked against a negative buy-and-hold approach across all the five currencies tested, indicating the PKR gained value or remained constant vis-à-vis each of the currencies for the most part but still managed to deliver positive returns on a cumulative basis (+7.73% to +46.45%), with an alpha varying between +16.09% and +48.63%, positive Sharpe ratio (0.491-1.251), maximum drawdown (-4.63%), and win rates ranging between 75.00% and 100.00% with only 4-12 transactions during the 26-month period.

**Table 6**

*Simulated Monthly Trading Performance — CNN-BiLSTM Macroeconomic Variables Model (26 Months)*

| Currency | Total Return (%) | Ann. Return (%) | Alpha (%) | Sharpe | Max DD (%) | Win Rate (%) |
|----------|------------------|-----------------|-----------|--------|------------|--------------|
| AED/PKR  | +7.73            | +3.50           | +16.09    | 0.491  | -2.71      | 75.00        |
| CNY/PKR  | +14.01           | +6.24           | +18.01    | 0.833  | -3.39      | 100.00       |
| EUR/PKR  | +37.94           | +16.00          | +38.41    | 0.979  | -4.63      | 83.33        |
| GBP/PKR  | +46.45           | +19.25          | +48.63    | 1.251  | -4.26      | 100.00       |
| USD/PKR  | +9.15            | +4.12           | +17.52    | 0.569  | -2.56      | 100.00       |

This selectiveness in the model is due to the model's purpose of identifying regimes in the economy rather than sending frequent trading signals, which is why the limited number of monthly trades makes the percentage wins

numbers more of an indication than a definite number, but the fact remains that despite the low number of trades the positive alpha values on all five currencies show that the signals from the macroeconomic model do have some economic significance.

### Cross-Model Comparison

Table 7 shows the findings. On the issue of forecasting accuracy (MAPE 0.08%-0.36% compared to 1.23%-2.64%; Pearson r 0.96-0.99 compared to 0.57-0.84), the technical model is undoubtedly superior. In addition, the technical model is superior on trading performance at its own frequency level

**Table 7**

*Cross-Model Comparison — Technical vs. Macroeconomic CNN-BiLSTM Models*

| Metric                   | Technical Model                    | Macroeconomic Model                 |
|--------------------------|------------------------------------|-------------------------------------|
| MAPE range (%)           | 0.08–0.36                          | 1.23–2.64                           |
| Directional Accuracy (%) | 67.5–72.8                          | 64.0–72.0                           |
| Pearson r                | 0.96–0.99                          | 0.57–0.84                           |
| Best Alpha (%)           | +124.36 (GBP/PKR)                  | +48.63 (GBP/PKR)                    |
| Best Sharpe Ratio        | 4.534 (GBP/PKR)                    | 1.251 (GBP/PKR)                     |
| Max Drawdown range (%)   | –1.63 to –3.05                     | –2.56 to –4.63                      |
| Win Rate range (%)       | 78.91–90.51                        | 75.00–100.00                        |
| Test period              | Jan 2024–Dec 2025 (daily, N = 512) | Jul 2023–Nov 2025 (monthly, N = 26) |

But such a comparison should be noted to come with a caveat regarding how the two models differ in terms of frequency, sampling size (512 daily observations compared to 26 monthly), and purpose, meaning a numerical comparison of the total returns achieved is not as relevant as the purpose for which each of the two models has been built. What is more theoretically important in comparison is directional accuracy, which proves to be quite similar for the two models (67.5-72.8% for technical/Dow vs. 64.0-72.0% for macroeconomic/Mundell-Fleming). Despite the vast difference in sampling volumes and indicators, both channels, technical/Dow and macroeconomic/Mundell-Fleming, provide comparable directional information on the PKR. Such results are consistent with those found by Yildirim et al. (2021), who showed that macroeconomic and technical LSTM forecast combination outperformed each of the forecasts taken alone, but extend the results by proving that the two channels have separate informational value, even when modeled independently through SHAP analysis.

### Synthesis

The findings herein show better performance compared to the best benchmark results available in the literature. For instance, Abouzaid and Boussedra (2025) found  $R^2 = 0.9234$  for an LSTM model in forecasting the EUR/USD exchange rate, which is quite similar to the Pearson correlations of 0.96-0.99 established for the technical model in the current study even though it is a less explored emerging market currency compared to the EUR/USD. Likewise, Nsengiyumva et al. (2025) found that a CNN-BiLSTM model outperforms CNN and BiLSTM models for the Rwandan Franc, while in the current paper, this has been shown to be applicable even in forecasting the PKR using a different set of inputs, either technical or macroeconomic, on both daily and monthly intervals.

The following conclusions emerge from the interpretation of the findings in light of the theoretical framework and the research gaps identified in the literature. The first is that the SHAP findings empirically back the hypothesis that technical and macroeconomic measures need to be treated as two distinct theoretical mechanisms working on different time frames, contrary to the current practice followed by many scholars who use these measures as part of

one general model (Nguyen et al., 2024; Nsengiyumva et al., 2025): it is the trend strength and movement measures that dominate the short-term model just as predicted by Dow Theory. Second, the two-model approach solves the problem of the lack of explainability that was pointed out by Lopez-Herrera et al. (2025) and Pornwattanavichai et al. (2022): the model is not a black box, but rather one that provides forecasts based on statistically derived factors that are also economically relevant and meaningful. Third, the backtesting results solve the problem of practical validation raised by Fang et al. (2023): both models are able to convert the statistically accurate and directionally relevant forecast into profitable trades, even though the currency involved in the analysis has been underexplored in the existing literature.

### Practical and Policy Implications

The outcomes have clear implications for various groups of market players based on the two-horizon nature of the developed framework. For traders, short-term portfolio managers, and organizations facing daily FX risks, the developed technical CNN-BiLSTM model can be used as a signal generation mechanism that provides trading rules that lead to win rates exceeding 78% and Sharpe ratios up to 4.534 while maintaining maximum drawdowns not worse than -3.05% on all the five currency pairs. It is important to note that despite the model being efficient when applied to USD/PKR and AED/PKR currency pairs with limited volatility, which demonstrates the usefulness of the forecasting model outside obvious large currency movements.

Whereas the model represents the State Bank of Pakistan as well as macroeconomic policymakers with regards to EUR/PKR and GBP/PKR, their use of the model in such a case would be more like a structural monitoring tool as opposed to a trading tool. This follows from the SHAP rankings which show that current inflation, lagged NFA and lagged KIBOR are the key determinants of PKR exchange rate movements each month. This means that monetary policy, which entails the regulation of the monetary policy, inflation and reserves of the bank, is the most likely way through which the authorities could manipulate PKR in a manner that ensures stability in its exchange rates, and the one and three month lags identified in the SHAP analysis are useful in indicating the period through which the effects of monetary policies on the PKR are felt.

The trading simulation of the macroeconomic model is very revealing for treasuries and institutional investors who have medium-term currency exposure because it was back-tested during a time period when being long or holding those five currencies passively would result in losses; the positive alpha generated by the model even during this time period, reaching 48.63% in case of GBP/PKR, clearly indicates that structurally and macroeconomically driven exposure can indeed add value, especially when a naive currency exposure does not.

In general, there is an implication on the methodological front with the application of the dual model approach: by splitting short-horizon technical inputs and long-horizon macroeconomic inputs into two distinct models as opposed to building a combined model with the two sets of indicators, the contributions of each input channel can be individually analyzed and dealt with, while at the same time SHAP explainability provides the added benefit of not having to treat the output of either model as a black-box input. It is through such a design and explainability aspect of the framework that makes it applicable for both trading purposes and monitoring purposes.

### Conclusion

In this research, a new forecasting framework based on a combination of a short-term and a long-term forecasting model, where the former model is based on a hybrid CNN-BiLSTM approach and uses 22 daily technical indicators, and the latter model is based on a similar approach and uses 26 features of six macroeconomic variables, was designed

and evaluated. The testing period was 15 years (from 2010 to 2025), including the periods of 2018 devaluation, COVID-19 shock, and 2022-2023 currency crisis, during which the technical model exhibited MAPE of 0.08-0.36% and directional accuracy of 67.5-72.8%, and the macroeconomic model exhibited MAPE of 1.23-2.64% and directional accuracy of 64-72%, both models showing stability through walk-forward cross-validation despite the most drastic structural break.

Results from SHAP analysis proved that the two models were based on theoretically different variables, the trend strength and directional movement indicator variables for the short-term model, and inflation, net foreign assets in the past, and KIBOR in the past for the long-term model, directly proving the validity of Dow theory and Mundell-Fleming model, respectively, and proved that the complexity of deep learning predictions can be achieved without losing economic intuition. The simulation also proved that both models were profitable and generated positive alpha against the passive benchmark in all five currency pairs, including in the window of macroeconomic-model validation where passively holding currencies would result in loss.

Collectively, these results close the four gaps mentioned in the literature review: the proposed framework does not split but rather combines technical and macroeconomic information by incorporating each one of these aspects into a model appropriate for a corresponding time horizon; hybrid deep learning is used on an emerging-market currency which has been under-researched in the AI-forecasting literature; the use of SHAP for interpretability allows the model to be transparent rather than a black box; finally, forecast validation is done using simulation of trading and thus done economically rather than statistically. Thus, the main contribution of this research cannot be reduced only to outperforming classical benchmarks in forecasting ability using hybrid deep learning – a fact already well-established by now – but to the fact that splitting the forecasting of technical and macroeconomic aspects into two models provides a more comprehensive and transparent way of analyzing PKR exchange rate dynamics.

Some possible directions for future research would be to build on the existing approach by adding more exogenous factors, such as oil prices, political risk indexes, and measures of policy uncertainty; by experimenting with more sophisticated network architectures, including attention layers, Temporal Fusion Transformers, and hybrid models of the technical and macroeconomic models; using more explainability techniques, such as LIME or Grad-CAM; by considering more realistic trading costs and bid-ask spreads; and by checking how well the CNN-BiLSTM architecture works in other emerging market currencies, besides Pakistan.

## References

- Abouzaid, O., & Boussedra, F. (2025). Artificial intelligence and exchange rate forecasting: Assessing predictive accuracy and macroeconomic sensitivity. *Frontiers in Applied Mathematics and Statistics*, 11. <https://doi.org/10.3389/fams.2025.1654093>
- Akhtar, S., Ramzan, M., Shah, S., Ahmad, I., Khan, M. I., Ahmad, S., El-Affendi, M. A., & Qureshi, H. (2022). Forecasting exchange rate of Pakistan using time series analysis. *Mathematical Problems in Engineering*, 2022, 1–11. <https://doi.org/10.1155/2022/9108580>
- Asadullah, M., Bashir, A., & Aleemi, A. R. (2021). Forecasting exchange rates: An empirical application to Pakistani rupee. *Journal of Asian Finance, Economics and Business*, 8(4), 339–347. <https://doi.org/10.13106/jafeb.2021.vol8.no4.0339>
- Ayitey, M., Appiahene, P., & Appiah, O. (2022). Forex market forecasting with two-layer stacked long short-term memory neural network (LSTM) and correlation analysis. *Journal of Electrical Systems and Information Technology*, 9(1). <https://doi.org/10.1186/s43067-022-00054-1>
- Biswas, A., Uday, I., Rahat, K., Akter, M., & Mahdy, M. (2023). Forecasting the USD/BDT exchange rate with deep learning models: Inclusion of macroeconomic factors influencing the currency exchange rates. *PLOS ONE*, 18(2), e0279602. <https://doi.org/10.1371/journal.pone.0279602>
- Chandra, R., Goyal, S., & Gupta, R. (2021). Evaluation of deep learning models for multi-step ahead time series prediction. *IEEE Access*, 9, 83105–83123. <https://doi.org/10.1109/ACCESS.2021.3085085>
- Chihab, Y., Bousbaa, Z., Chihab, M., Bencharef, O., & Ziti, S. (2019). Algo-trading strategy for intraweek foreign exchange speculation based on random forest and probit regression. *Applied Computational Intelligence and Soft Computing*, 2019, 1–13. <https://doi.org/10.1155/2019/8342461>
- Corzo, T., Martin-Bujack, K., Portela, J., & Rodríguez-Gallego, A. (2025). Exchange rate regime changes and market efficiency: An event study. *Journal of International Financial Markets, Institutions and Money*, 100, 102132. <https://doi.org/10.1016/j.intfin.2025.102132>
- D'Angelo, E., & Grimaldi, G. (2017). The effectiveness of the Elliott Waves Theory to forecast financial markets: Evidence from the currency market. *International Business Research*, 10(6), 1–18. <https://doi.org/10.5539/ibr.v10n6p1>
- Dautel, A., Härdle, W., Lessmann, S., & Seow, H. (2020). Forex exchange rate forecasting using deep recurrent neural networks. *Digital Finance*, 2(1–2), 69–96. <https://doi.org/10.1007/s42521-020-00019-x>
- Escudero, P., Alcocer, W., & Paredes, J. (2021). Recurrent neural networks and ARIMA models for Euro/Dollar exchange rate forecasting. *Applied Sciences*, 11(12), 5658. <https://doi.org/10.3390/app11125658>
- Fang, S., Wei, Y., & Wang, S. (2023). 30 years of exchange rate analysis and forecasting: A bibliometric review. *Journal of Economic Surveys*, 37(4), 1123–1150. <https://doi.org/10.1111/joes.12572>
- Frenkel, J. A., & Razin, A. (1987). The Mundell–Fleming model: A quarter century later (NBER Working Paper No. 2321). National Bureau of Economic Research. <https://doi.org/10.3386/w2321>
- García, F., Guijarro, F., Oliver, J., & Tamošiūnienė, R. (2024). Foreign exchange forecasting models: LSTM and BiLSTM comparison. *Engineering Proceedings*, 68(1), 19. <https://doi.org/10.3390/engproc2024068019>
- Jabeen, M., Rashid, A., & Ihsan, H. (2022). The news effects on exchange rate returns and volatility: Evidence from Pakistan. *International Journal of Finance & Economics*, 27(1), 745–769. <https://doi.org/10.1002/ijfe.2176>
- Kim, D., Kim, D., & Choi, S. (2025). Modeling stylized facts in FX markets with FINGAN-BiLSTM: A deep learning approach to financial time series. *Entropy*, 27(6), 635. <https://doi.org/10.3390/e27060635>

- Korol, T. (2014). A fuzzy logic model for forecasting exchange rates. *Knowledge-Based Systems*, 67, 49–60. <https://doi.org/10.1016/j.knosys.2014.06.009>
- Liu, S., Huang, Q., Li, M., & Wei, Y. (2024). A new LASSO-BiLSTM-based ensemble learning approach for exchange rate forecasting. *Engineering Applications of Artificial Intelligence*, 127, 107305. <https://doi.org/10.1016/j.engappai.2023.107305>
- López-Herrera, F., Jiménez, J. G. M., & Santiago, A. R. (2025). Directional forecasting for eight forex pairs against the US dollar using machine learning techniques. *Discover Artificial Intelligence*, 5, Article 224. <https://doi.org/10.1007/s44163-025-00424-4>
- Mancini Griffoli, T., Meyer, C., Natal, J.-M., & Zanetti, A. (2015). Determinants of the Swiss franc real exchange rate. *Swiss Journal of Economics and Statistics*, 151(4), 303–328. <https://doi.org/10.1007/BF03399419>
- Meng, S., Chen, A., Wang, C., Zheng, M., Wu, F., Chen, X., Ni, H., & Li, P. (2024). Enhancing exchange rate forecasting with explainable deep learning models. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.2410.19241>
- Neely, C. J., & Weller, P. A. (2011). Technical analysis in the foreign exchange market (Working Paper No. 2011-001B). Federal Reserve Bank of St. Louis. <http://research.stlouisfed.org/wp/2011/2011-001.pdf>
- Nguyen, P. D., Thao, N. N., Kim, T., Nguyen, H.-C., Mach, B.-N., & Nguyen, T. Q. (2024). Deep learning-based predictive models for forex market trends: Practical implementation and performance evaluation. *Science Progress*, 107(3). <https://doi.org/10.1177/00368504241275370>
- Ni, L., Li, Y., Wang, X., Zhang, J., Yu, J., & Qi, C. (2019). Forecasting of Forex time series data based on deep learning. *Procedia Computer Science*, 147, 647–652. <https://doi.org/10.1016/j.procs.2019.01.189>
- Nsengiyumva, E., Mung’atu, J., & Ruranga, C. (2025). A comparative study of multivariate CNN, BiLSTM and hybrid CNN–BiLSTM models for forecasting foreign exchange rate using deep learning. *Cogent Economics & Finance*, 13(1), 2526148. <https://doi.org/10.1080/23322039.2025.2526148>
- Nwanne, T. F. I., & Onyekachi, E. R. (2015). Assessing the effect of external debt servicing and receipt on exchange rate in Nigeria. *International Journal of Economics and Finance*, 7(9), 278–285. <https://doi.org/10.5539/ijef.v7n9p278>
- Panda, C., & Narasimhan, V. (2007). Forecasting exchange rate better with artificial neural network. *Journal of Policy Modeling*, 29(2), 227–236. <https://doi.org/10.1016/j.jpolmod.2006.01.005>
- Panopoulou, E., & Souropanis, I. (2019). The role of technical indicators in exchange rate forecasting. *Journal of Empirical Finance*, 53, 197–221. <https://doi.org/10.1016/j.jempfin.2019.07.004>
- Perla, S., Bisoi, R., & Dash, P. K. (2023). A hybrid neural network and optimization algorithm for forecasting and trend detection of Forex market indices. *Decision Analytics Journal*, 6, 100193. <https://doi.org/10.1016/j.dajour.2023.100193>
- Perveen, N., Tariq, K., & Ahmad, H. S. (2024). Comparative studies of hybrid ARIMA and artificial neural network (ANN) techniques for predicting exchange rate in Pakistan. *Bulletin of Business and Economics*, 13(2), 630–636. <https://doi.org/10.61506/01.00372>
- Pornwattanavichai, A., Maneeroj, S., & Boonsiri, S. (2022). BERTFOREX: Cascading model for forex market forecasting using fundamental and technical indicator data based on BERT. *IEEE Access*, 10, 23425–23437. <https://doi.org/10.1109/access.2022.3152152>
- Qureshi, A. M. (2025). Machine learning prediction of exchange rates using macroeconomic indicators. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.5287273>

- Qureshi, M., Ahmad, N., Ullah, S., & Mustafa, A. R. U. (2023). Forecasting real exchange rate (REER) using artificial intelligence and time series models. *Heliyon*, 9(5), e16335. <https://doi.org/10.1016/j.heliyon.2023.e16335>
- Rashid, A., & Basit, M. (2021). Empirical determinants of exchange-rate volatility: Evidence from selected Asian economies. *Journal of Chinese Economic and Foreign Trade Studies*, 15(1), 63–86. <https://doi.org/10.1108/JCEFTS-04-2021-0017>
- 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>
- Sönmez, F. E., & Birim, Ş. Ö. (2024). *Forecasting exchange rate depending on the data volatility: A comparison of deep learning techniques*. <https://doi.org/10.21203/rs.3.rs-4218174/v1>
- Wagdi, O., Salman, E., & AlBanna, H. (2023). Integration between technical indicators and artificial neural networks for the prediction of the exchange rate: Evidence from emerging economies. *Cogent Economics & Finance*, 11(1), 2255049. <https://doi.org/10.1080/23322039.2023.2255049>
- Wang, J., Wang, X., Li, J., & Wang, H. (2021). A prediction model of CNN-TLSTM for USD/CNY exchange rate prediction. *IEEE Access*, 9, 73346–73354. <https://doi.org/10.1109/ACCESS.2021.3080459>
- Yao, J., & Tan, C. L. (2000). A case study on using neural networks to perform technical forecasting of forex. *Neurocomputing*, 34(1–4), 79–98. [https://doi.org/10.1016/S0925-2312\(00\)00300-3](https://doi.org/10.1016/S0925-2312(00)00300-3)
- Yıldırım, D. C., Toroslu, İ. H., & Fiore, U. (2021). Forecasting directional movement of Forex data using LSTM with technical and macroeconomic indicators. *Financial Innovation*, 7(1), 1–22. <https://doi.org/10.1186/s40854-020-00220-2>