

Forecasting the Consumer Price Index in Saudi Arabia using Time Series, Machine learning, and Hybrid Models: A Comparative Study

Alhanouf Almalki, Maryam Habadi, Dawlah Alsulami, Lamya Baharith*

Department of Statistics, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia

**Corresponding author: lbaharith@kau.edu.sa*

Abstract. The consumer price index (CPI) tracks the prices of a specified basket of consumer goods and services. It is a leading indicator of inflation and economic performance and plays a crucial role in economic and social planning, reinforcing the importance of accurate CPI forecasting. This article performs a comparative analysis to identify the most appropriate methodologies for forecasting the CPI in Saudi Arabia, using data from January 2011 to December 2024. The methods employed are autoregressive integrated moving average (ARIMA), machine learning models—artificial neural networks (ANN) and long short-term memory (LSTM)—and hybrid models ARIMA-LSTM and ARIMA-ANN. Three activation functions are used: Sigmoid, hyperbolic tangent (Tanh), and rectified linear unit (ReLU). Several accuracy measures evaluate predictive performance. The results demonstrate that the hybrid ARIMA-ANN model with ReLU activation function was the most effective model for forecasting the CPI in Saudi Arabia.

1. INTRODUCTION

The Kingdom of Saudi Arabia's Vision 2030 aims to create a vibrant economy by leveraging the kingdom's unique location and resources, expanding job opportunities for Saudis, improving the business environment to broaden economic activity and ensure stability, and supporting domestic firms in becoming competitive exporters to become a key center for global trade. The economy in Saudi Arabia is expected to grow by 3.6% in 2025 compared with 2024, due to the increase in the oil and non-oil related activities [1].

The Consumer Price Index (CPI) plays a central role in economic and social planning, as it evaluates monthly changes in the price of a specific collection of goods and services within a defined region. It serves as a crucial economic indicator for characterizing countries and is generally considered the official measure of inflation. When the CPI increases in a country, the inflation will rise and vice versa. Economic growth is indicated by a steady increase in the CPI, which

Received: Apr. 17, 2026.

2020 *Mathematics Subject Classification.* 62M10, 68T07, 91B84.

Key words and phrases. ARIMA; ANN; LSTM; time series; machine learning; hybrid models.

signals healthy economic demand. On the other hand, fluctuations in the CPI indicate economic instability; hence, people cannot afford basic goods and services [2].

Since 1999, the CPI in Saudi Arabia has fluctuated due to changes in both domestic economic and global factors, including oil prices, housing demand, global food shortages, and supply chain disruptions. Over the previous years, the Saudi CPI has varied, with the lowest being 67.00 points recorded in July 2001, followed by 97.39 in March 2019 [3,4]. Recently, the Saudi CPI has shown a steady upward trend, reaching 112.81 points in February 2025 and 113.10 points in March 2025. This growth is expected, as the essential categories (food, housing, water, electricity, gas, and other fuels) in the CPI, which account for 43%, rose by 8.0% in January of 2025 compared to the same period in 2024, as reported by the General Authority for Statistics. Consequently, the inflation rate increased from 2.0% in February to 2.3% in March, which is the highest since July 2023.

The CPI measures changes in the prices of goods and services consumed by individuals, indicating a society's purchasing power. Therefore, forecasting the CPI in Saudi Arabia is crucial, as it helps understand inflation trends and their impact on the economy. The CPI is also an important tool for policymakers, businesses, and investors to make informed financial and economic decisions based on well-researched forecasts, which helps reduce economic risks and support overall economic stability.

The performance of neural networks is influenced by several factors, including activation functions, which play a crucial role in training efficiency and prediction accuracy. These functions determine how neurons respond to input signals and enable networks to capture nonlinear relationships in data, which is essential for learning complex patterns. However, there is no universal rule for selecting the most suitable activation function, as its effectiveness depends on the specific application and the problem being addressed.

Motivated by the fact that forecasting the CPI provides clear insights into general economic trends and supports monetary policy decisions, including taxation, this article aims to:

- Model Saudi Arabia's CPI, as it is considered a benchmark for the stability of the Saudi economy.
- Compare the forecasting performance of statistical, machine learning, and hybrid models for the Saudi CPI.
- Evaluate the effectiveness of three activation functions in forecasting the CPI of Saudi Arabia to identify the most suitable approach.
- Provide insight into the expected CPI trend for Saudi Arabia in 2025.

This proactive approach helps decision makers mitigate inflation, contributing to economic stability and growth by ensuring sustainable development, resource management, and aligning with Vision 2030. The remainder of the article is organized as follows. Section 2 reviews previous studies. Section 3 describes the data and methods. Section 4 presents the results. Section 5 provides a discussion of the findings. Finally, conclusions are given in Section 6.

2. LITERATURE REVIEW

Recognizing the importance of the Consumer Price Index (CPI) in the modern economy, researchers have focused on modeling and forecasting it using a range of statistical and machine learning methods. [5] applied double exponential smoothing to monthly CPI data from March 2020 to August 2021 in Yogyakarta, Indonesia, for three months. The results showed that the double exponential smoothing performed well in predicting Yogyakarta's CPI. In the study of [6], a seasonal autoregressive integrated moving average (SARIMA) model was applied to data from January 2010 to December 2022 to forecast the monthly CPI of restaurants and hotels in South Korea. They used the mean absolute percentage error (MAPE) and the coefficient of determination (R^2) to assess the SARIMA model's forecasting efficiency. [7] aimed to forecast the CPI for food, beverages, and tobacco in Banten. They used data from January 2019 to April 2022 to determine the best ARIMA model for forecasting the CPI over the next eight months and reported that ARIMA (0,3,1) was the most suitable. [8] employed the ARIMA model to analyze and forecast the CPI in Somalia using data from November 2022 to November 2024, and selected the ARIMA (1,1,0) model as it yielded the lowest MAPE.

Due to the limitations of traditional statistical approaches in modeling complex nonlinear patterns in economic data, researchers have examined the benefits of machine learning (ML) over statistical models for forecasting CPI. [9] used Ecuadorian CPI data from January 2005 to June 2019 to predict the CPI through June 2020. They compared the efficiency of the statistical models (SARIMA and exponential smoothing) with ML models (ANN, LSTM, and support vector regression). This comparison showed that, based on MAPE values, support vector regression, followed by LSTM, achieved better prediction accuracy. [10] compared the ARIMA model and the ANN for predicting the Indonesian CPI using data collected between January 2009 and December 2017, with the aim of predicting the CPI for the next 12 months. They concluded that ANN produced better predictive accuracy than ARIMA. [11] compared SARIMA with neural network autoregression and multilayer perceptron (MLP). Using CPI data for Pakistan from 1960 to 2021, they forecasted the CPI from 2022 to 2031. The accuracy metrics RMSE, MSE, and MAPE showed that the MLP with 20 layers performed best. [12] compared the forecasting accuracy of ARIMA and four ML models: LSTM, Geometric Brownian Motion, gated recurrent unit (GRU), and adaptive neuro-fuzzy inference system using Malaysian CPI data from January 1968 to September 2022. The findings indicated that the GRU model has the lowest evaluation metrics values; therefore, it offered greater precision. [13] used CPI data in Portugal to compare the predictive accuracy of ARIMA and the random forest (RF) models. Data were collected from January 1948 to January 2023 for a one-year prediction (short term), from January 1948 to January 2019 for a five-year prediction (medium term), and from January 1948 to January 2014 for a ten-year prediction (long term). The results demonstrated the effectiveness of the RF in predicting CPI.

Researchers have also combined statistical time series models, which may have limitations in handling nonlinearity, with ML models that offer greater flexibility to address such complexities.

Both linear and nonlinear models may struggle to capture complex patterns accurately on their own; hybrid systems that integrate linear and nonlinear modeling techniques have been explored. [14] compared the accuracy of ARIMA, LSTM, PROPHET, and the hybrid ARIMA-LSTM for forecasting CPI inflation in twelve countries, including the United States, Canada, Australia, and India, using data from 1960 to 2021. The results demonstrated the superiority of the hybrid ARIMA-LSTM model. [15] adopted the ARIMA, backpropagation neural network (BPNN), ANN, hybrid ARIMA-BPNN, and hybrid ARIMA-ANN models to forecast the CPI in Yemen using data from January 2005 to December 2014. Based on multiple accuracy metrics, the ARIMA-ANN performed better than other models. [16] used CPI data from Vietnam (January 1995 to July 2022) to compare the predictive accuracy of the ARIMA, ANN, and hybrid ARIMA-ANN models. Based on the MSE values, the hybrid model outperformed individual models in multiple-step prediction. [17] compared ARIMA with the hybrid ARIMA-ANN model to forecast the CPI for the Philippines and its regions in 2022, using Akaike's Information Criterion (AIC), MAPE, and RMSE. The results indicated that the hybrid ARIMA-ANN model outperformed the traditional ARIMA model.

Under Vision 2030, Saudi Arabia diversified its economy and reduced its dependence on oil revenues, thereby significantly transforming the kingdom's economy, increasing the importance of modeling and forecasting the CPI. [18] applied several ARIMA models to forecast the annual CPI in Saudi Arabia using data from 1963 to 2017 and identified the ARIMA (0,2,1) model as the best performing. [2] predicted the monthly CPI in Saudi Arabia using data from May 2019 to July 2020 and concluded that the Holt linear model was suitable for both short- and long-term forecasting. [19] employed the SARMA (0,1,0)(12,0,12) model to forecast the monthly CPI data collected from January 2013 to May 2020 and showed its ability to capture seasonal patterns. [20] used data from January 2013 to November 2020 to forecast the CPI in Saudi Arabia through November 2024. They applied several ML models, including decision trees, random forests, neural networks, linear regression, k-nearest neighbor, and support vector machine. Based on MSE and the R^2 , they concluded that ANN provided the best prediction accuracy.

Although neural networks are widely used in time-series forecasting, few studies have explicitly examined the impact of activation function choice on forecasting performance. Existing research suggests that activation functions can significantly influence learning quality and prediction accuracy, highlighting their importance and their potential to improve forecasting results. [21] compared 12 activation functions for financial time series forecasting using MLP across 12 real financial time series. The study included standard sigmoid and tanh functions as well as the proposed complementary log-log, probit, and log-log functions. Neural networks were trained using conjugate gradient backpropagation (CGF) and Levenberg–Marquardt (LM) algorithms, with the AR model used only as a benchmark. The results showed that the choice of activation function affects forecasting accuracy, especially in small networks and for asymmetric time series.

In contrast, traditional functions generally require larger networks to achieve similar performance. [22] compared 10 activation functions, including conventional monotonic functions such

as tanh and linear activations, as well as non-monotonic, wavelet-based, and even functions, notably the sinc and Mexican hat functions. Experiments were conducted on synthetic time series, namely the Logistic Map, Mackey–Glass, NARMA10, and NARMA20, using Echo State Networks (ESNs). Prediction accuracy was evaluated using both normal and free-running prediction tasks. The results showed that activation functions with appropriate nonlinearity, particularly sinc and Mexican hat, outperformed others, highlighting the importance of activation function design for improving ESN predictability. [23] conducted an experimental study to evaluate the impact of activation function selection on the performance of MLP for electricity power production and consumption forecasting. They compared 11 activation functions, including Sigmoid, Tanh, ReLU, Leaky ReLU, ELU, SELU, Swish, and Gaussian, across different depths of hidden layers. Using four benchmark electricity-related datasets, three were multivariate, while the Individual Household Electric Power Consumption dataset was treated as a univariate time series. The results showed that the ReLU and Leaky ReLU outperformed other activation functions across all datasets. [24] applied an MLP to compare three activation functions—Sigmoid, Tanh, and ReLU—using three simulated nonlinear time series generated from a Threshold Autoregressive (TAR) model and two real-world datasets: the Canadian Lynx series and Wolf’s Sunspot data. They concluded that the Tanh activation function achieved the lowest error values across both simulated and real datasets, outperforming Sigmoid and ReLU. These findings indicate that selecting an activation function is critical for improving forecasting accuracy.

3. MATERIALS AND METHODS

In this study, open-source CPI data for Saudi Arabia are analyzed using various statistical, ML, and hybrid forecasting models.

3.1. Data. Monthly CPI data for Housing, Water, Electricity, Gas, and Other Fuels in Saudi Arabia from January 2011 to December 2024 are used. The data were obtained from the International Monetary Fund ¹. The data were divided so that the test data comprised the most recent 24 months, while the remaining data were used for training. Descriptive statistics of the CPI data are shown in Table 1, and the visual representation is in Figure 1.

TABLE 1. Descriptive statistics for Housing, Water, Electricity, Gas, and Other Fuels CPI in Saudi Arabia.

Mean	Median	Min.	Max.	SD	Skewness	Kurtosis
94.49	93.66	82.16	109.92	6.217	0.302	2.17

As shown in Table 1 and Figure 1, the CPI has fluctuated and reached the minimum in January 2011 and the maximum in December 2024.

¹<https://www.imf.org/en/Home>

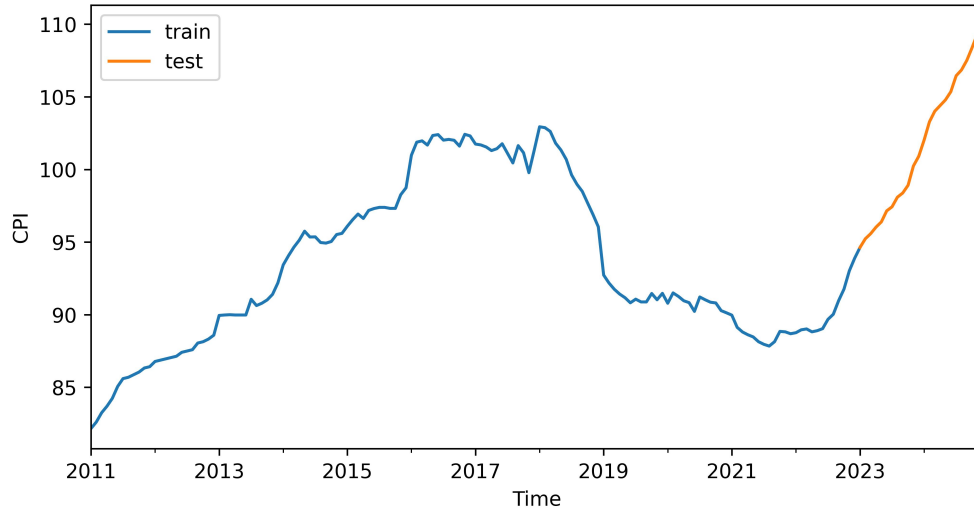


FIGURE 1. Time series plot (train and test data).

3.2. Autoregressive Integrated Moving Average (ARIMA) Model. [25] proposed a generalized model known as $ARIMA(p, d, q)$. This model integrates an autoregressive (AR) model of order p and a moving average (MA) model of order q , while incorporating an integrated component, represented by d to maintain the stationarity of the time series and eliminate any existing trend. The moving average model accounts for the possibility of a relationship between observations and residuals from previous periods, whereas the autoregressive model focuses on dependencies among current and past observations. The ARIMA model is represented as:

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j}, \quad (3.1)$$

where c is constant, ϕ_i and θ_j denote the coefficients of the autoregressive and moving average, respectively, and ϵ_t represents a series of Gaussian white noise, characterized by a mean of zero and a variance of σ^2 .

3.3. Machine Learning (ML) Approach. It is a branch of artificial intelligence that aims to create systems that learn from historical data, make predictions, identify patterns, and make logical decisions with minimal human oversight [26]. Deep learning is a more complex and specialized subset of ML, inspired by the human brain, that trains neural networks to classify sounds, text, or images using massive amounts of data [27]. ML is the field of study that enables computers to learn independently without explicit programming [28]. No superior type of ML models can solve all problems effectively, thus the choice of the most appropriate model for each problem depends on the characteristics and nature of the data [26, 29]. For neural network-based ML techniques, activation functions are an essential component of the system design. Mathematically, according to [30], activation functions compute the weighted sum of inputs and biases to activate neurons and prevent overfitting. They are essential for the management and decision-making processes

within neural networks. In addition, they regulate the transmission of data between cells, and are essential for enabling neurons to capture non-linearity and process complex patterns in data. The most frequently applied types of activation functions are as follows:

- **Sigmoid activation function:** It is a non-linear activation function designed to solve issues with the linear activation function. Its output ranges from 0 to 1, which makes it suitable for predicting probabilities. However, the Sigmoid function has several limitations, the most significant being the vanishing gradient problem and the nonzero centered [31]. The Sigmoid activation function can be represented mathematically in the following way:

$$f(y) = \frac{1}{1 + e^{-y}}, \quad (3.2)$$

where y denotes the input data.

- **Tanh activation function:** It is also a non-linear activation function with a range between 1 and -1, allowing it to solve the problem of non-centrality of the Sigmoid function. However, it also has been impacted by the vanishing gradient problem. The Tanh function is commonly used in recurrent neural networks for speech recognition tasks [31]. It can be represented as follows:

$$f(y) = \frac{\sinh(y)}{\cosh(y)} = \frac{e^y - e^{-y}}{e^y + e^{-y}}, \quad (3.3)$$

- **ReLU activation function:** It addresses the vanishing gradient problem found in earlier activation functions by returning the positive input value and outputting zero for negative values. It is a non-linear activation function widely utilized in deep neural networks [32]. It can be represented as follows:

$$f(y) = \begin{cases} 0, & y < 0. \\ y, & y \geq 0. \end{cases} \quad (3.4)$$

For more details, see [30].

3.3.1. Artificial Neural Networks (ANN). Artificial Neural Networks are widely used ML models inspired by the structure of the human brain. They have contributed significantly to time series forecasting in many fields, including medicine, computer science, statistics, and economics, due to their flexibility and ability to learn and model non-linear data [33,34]. The architecture of an ANN consists of three key layers: an input layer, where each input represents a specific time step; one or more hidden layers that process the input and learn underlying patterns; and an output layer that produces the final prediction. Each layer contains neurons connected by weights through which signals are transmitted. During training, the network adjusts the weights and biases using backpropagation based on the difference between the predicted and actual values. In forward propagation, each neuron takes the input from the previous layer, multiplies it by the corresponding weight, adds the bias, and then applies an activation function to the result. The

output of the activation function then becomes the input to the neuron in the next layer [35]. The output process of neuron j can be represented mathematically as follows:

$$y_j(n) = f\left(\sum_{i=1}^m w_{ji}(n) y_i(n) + b_j\right), \quad (3.5)$$

where $y_j(n)$ denotes the output of neuron j at iteration n , $f(\cdot)$ is the activation function, $w_{ji}(n)$ represents the weight from unit i to neuron j , and $y_i(n)$ is the input signal of unit i from the previous layer and b_j is the bias [36].

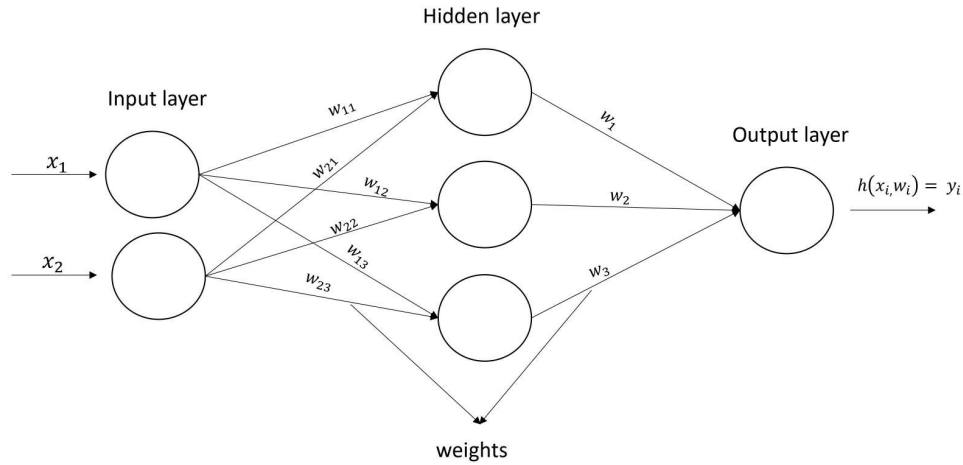


FIGURE 2. Architecture of the ANN.

3.3.2. Long Short-Term Memory Networks (LSTM). Long Short-Term Memory, first introduced by [37], is a specialized form of recurrent neural network (RNN) in which memory units replace summation units in hidden layers, enabling the network to learn long-term dependencies in information [38]. It is particularly effective in handling long sequences and can retain information for extended periods. Therefore, it can retain values from earlier stages for future reference and is designed to work with data sequences such as time series or texts [39]. The network internal architecture features three gates: input, forget, and output. The input gate controls which new information is stored in memory, whereas the forget gate decides which information to retain or discard. Lastly, the output gate controls the information passed to the next layer. These gates effectively address the problems of vanishing and exploding gradients that classic RNNs commonly encounter. Figure 3 displays the general architecture of the LSTM.

3.4. Hybrid Model. Hybrid approaches can significantly enhance forecasting accuracy by effectively capturing the linear and non-linear patterns in the data. Notably, [40] proposed hybrid models combining the traditional ARIMA model with an ANN. This method can be represented as follows:

$$y_t = L_t + NL_t, \quad (3.6)$$

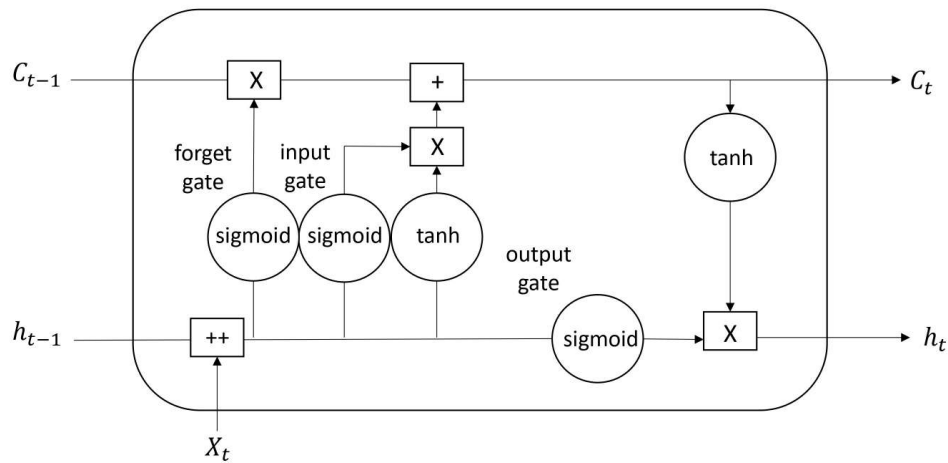


FIGURE 3. Architecture of the LSTM.

where L_t denotes the linear part and NL_t signifies the non-linear part. In the first step, the ARIMA model is applied, and the predicted values are then subtracted from the original series to obtain the residuals, which capture the nonlinear part that the ARIMA model failed to capture.

$$NL_t = y_t - \hat{L}_t, \quad (3.7)$$

In the next step, the ML models are applied the residuals N_t and the predicted values $\hat{N}_t$ are obtained. Finally, the predictions of the hybrid method can be obtained using an additive model [41] as follows:

$$\hat{y}_t = \hat{L}_t + \hat{N}_t, \quad (3.8)$$

The framework of the hybrid approach is presented in Figure 4.

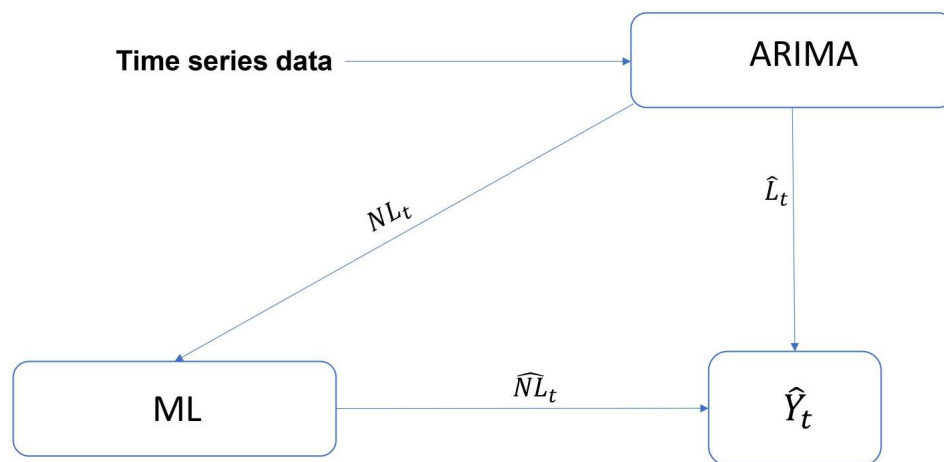


FIGURE 4. Framework of the hybrid models.

3.5. Accuracy Measures. Forecast accuracy measures are utilized to assess the effectiveness of various models in forecasting the CPI in Saudi Arabia. Each accuracy measure assesses the forecasting error from a different perspective. [42] recommend using the mean absolute error for comparing time-series predictions. The absolute error can be converted to absolute percentage error to eliminate scale dependency.

In this study, the prediction accuracy of the models was assessed using three measures: mean absolute error (MAE), mean squared error (MSE), and root mean square error (RMSE). The formulation of these accuracy measures is defined as follows:

$$MAE = \frac{1}{n} \sum_{t=1}^n |(Y_t - F_t)|, \quad (3.9)$$

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2, \quad (3.10)$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - F_t)^2}, \quad (3.11)$$

where Y_t is the actual value and F_t is the forecasting value for month, and n represents the test data size [43].

4. RESULTS

In this section, the results of employing various models on Saudi Arabia's CPI data will be outlined and discussed.

4.1. Time series models.

4.1.1. Data preprocessing. The initial phase involves preprocessing time series data to ensure accurate and consistent results. The CPI data were checked for outliers and missing values. Furthermore, stationarity, a fundamental requirement for time series models, has been assessed using the Phillips-Perron (PP) test proposed by [44]. This test suggests applying the first difference to ensure stationarity. Identifying seasonality is also critical in time series models to ensure accurate forecasts; thus, the Kruskal-Wallis was used to examine the seasonal effect. The results indicate that there is no evidence of a seasonal effect. Based on the Box-Jenkins procedure, the optimal ARIMA model for the CPI is ARIMA (1,1,1). The model is expressed as follows:

$$\nabla y_t = 0.8726 \nabla y_{t-1} - 0.6316 \varepsilon_{t-1} + \varepsilon_t, \quad (4.1)$$

Figures 5 and 6 show the ACF and PACF plots of the Saudi Arabian CPI training data.

Figure 7 shows the ARIMA(1, 1, 1) predicted vs. actual values for the test set, and Figure 8 shows the forecasted values along with the historical data.

4.2. Machine Learning (ML) Models.

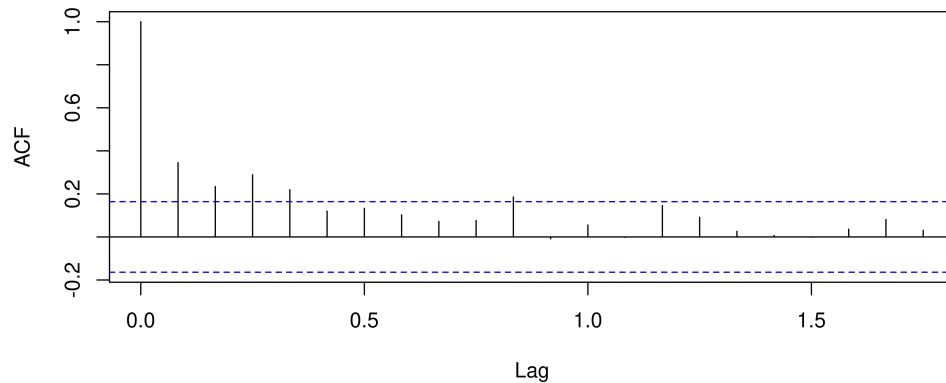


FIGURE 5. Autocorrelation Function (ACF) plot.

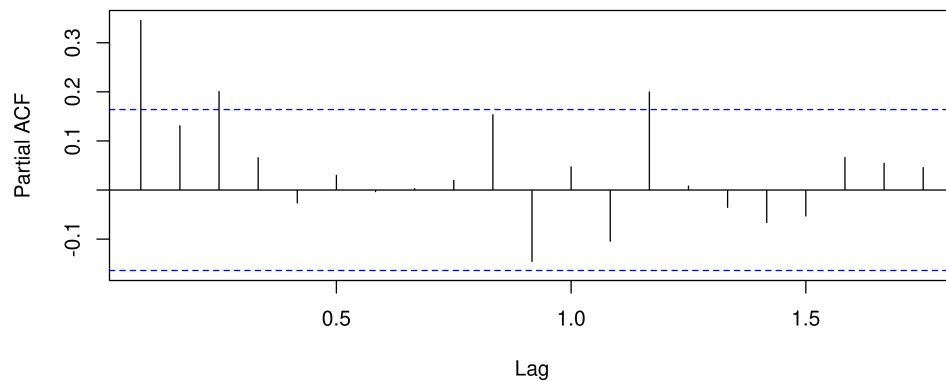


FIGURE 6. Partial Autocorrelation Function (PACF) plot.

4.2.1. Data preprocessing. In most ML models, normalization is crucial for accurate, efficient learning during training, resulting in more reliable predictions. In this study, the min-max normalization method was applied to scale the CPI data to a specified range, typically 0 and 1, preserving the relationship among data points while ensuring that every feature contributes equally to the analysis. Accordingly, the CPI data were normalized and divided into features and lagged features to capture temporal dependencies. Targets were generated for both the training and test datasets using a sliding window approach. The model employs a look-back value to predict the following observation from the previous four. For all ML models, the layers and nodes were iteratively added and removed to achieve the lowest values for the accuracy measures. The Sigmoid, Tanh, and ReLU activation functions were employed and compared to identify the most suitable option. The models were trained on the training set for up to 500 epochs using different batch sizes. Data

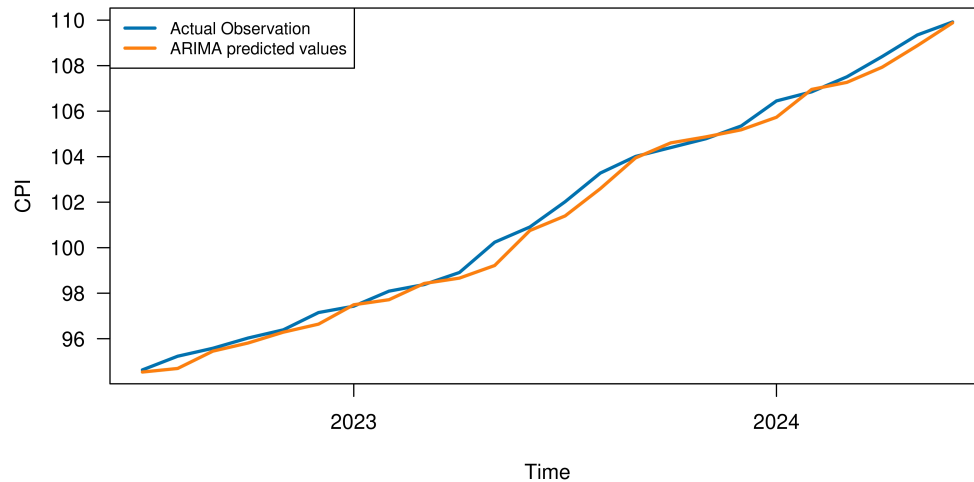


FIGURE 7. Actual (Test Data) and Predicted CPI Values Using the ARIMA(1,1,1) Model.

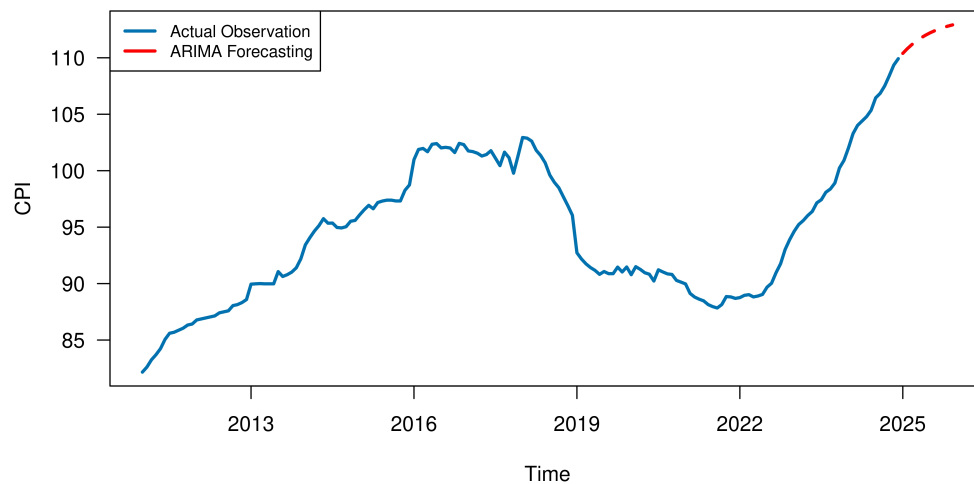


FIGURE 8. Actual Data and Forecasted CPI Values Using the ARIMA(1,1,1) Model.

shuffling was not applied because the dataset is a time series; shuffling would disrupt the temporal order, which is essential for capturing time dependencies in forecasting models. The following subsection provides a detailed overview of the construction of each model.

4.2.2. Artificial Neural Network. In ANN models, the layer and neuron configurations were determined empirically to optimize the accuracy metrics. The ANN model with the ReLU activation function consists of two hidden layers with 21 and 8 neurons, respectively, and an output layer with one neuron, using a batch size of 14 during training. In contrast, the ANN models employing Tanh and Sigmoid activation functions adopt a simpler structure: each model consists of a single hidden layer with 19 and 25 neurons, respectively, and uses batch sizes of 16 and 10, respectively.

This structure enables the models to capture complex patterns in the data while producing continuous output. Table 2 summarizes the structures of all ANN models with different activation functions.

The predicted values for all activation functions are presented in Figure 9, while the forecasted values for the best-performing model ANN with ReLU are shown in Figure 10. As shown in these figures, the models effectively capture the underlying trends and patterns in the data. Moreover, the ANN model with the Relu activation function produces more stable predicted values, suggesting that the choice of activation function has a notable impact on forecasting performance.

TABLE 2. Architecture of ANN models.

Activation Function	Input Layer (Features)	Hidden Layers	Output Layer
ReLU	4	21–8	1
Tanh	4	19	1
Sigmoid	4	25	1

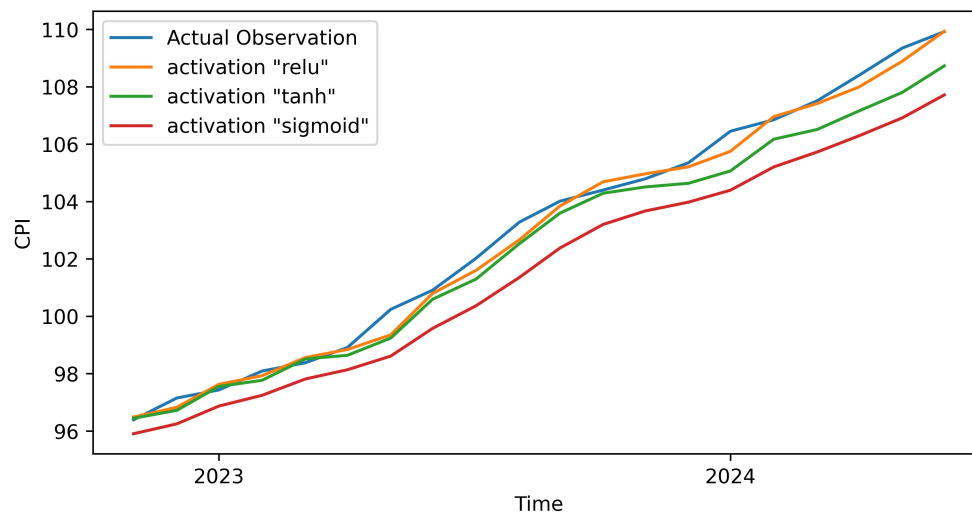


FIGURE 9. Actual (Test Data) and Predicted CPI Values Using the ANN Models.

4.2.3. Long Short-Term Memory. For the LSTM models, the following configurations are employed: for the ReLU and Tanh activation functions, a single hidden layer with 17 nodes and an output layer was used, with batch sizes of 16 and 8, respectively. For the sigmoid activation function, a single hidden layer with 18 nodes and an output layer was used, using a batch size of 18. The ReLU-based LSTM model used the *LeCun_uniform* weight initialization, whereas the models with Tanh and Sigmoid activation functions used the *glorot_uniform* initializer. Table 3 summarizes the architectures of different LSTM models with different activation functions. The

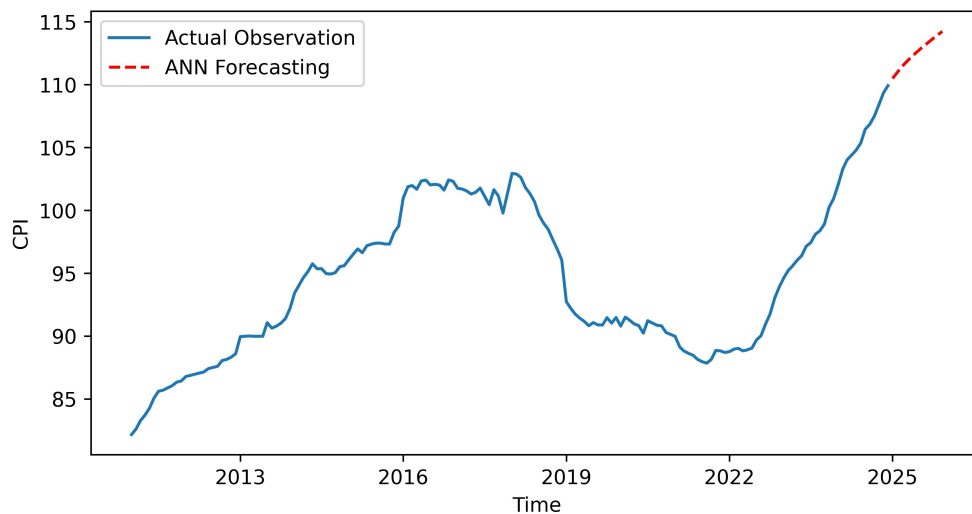


FIGURE 10. Actual Data and Forecasted CPI Values Using the ANN (ReLU) Model.

predicted values of the LSTM models are presented in Figures 11 , while the forecasted values of the best-performing model with the ReLU activation function are shown in Figure 12.

TABLE 3. Architecture of LSTM models.

Activation Function	nput Layer (Features)	Hidden Layers	Output Layer
ReLU	4	17	1
Tanh	4	17	1
Sigmoid	4	18	1

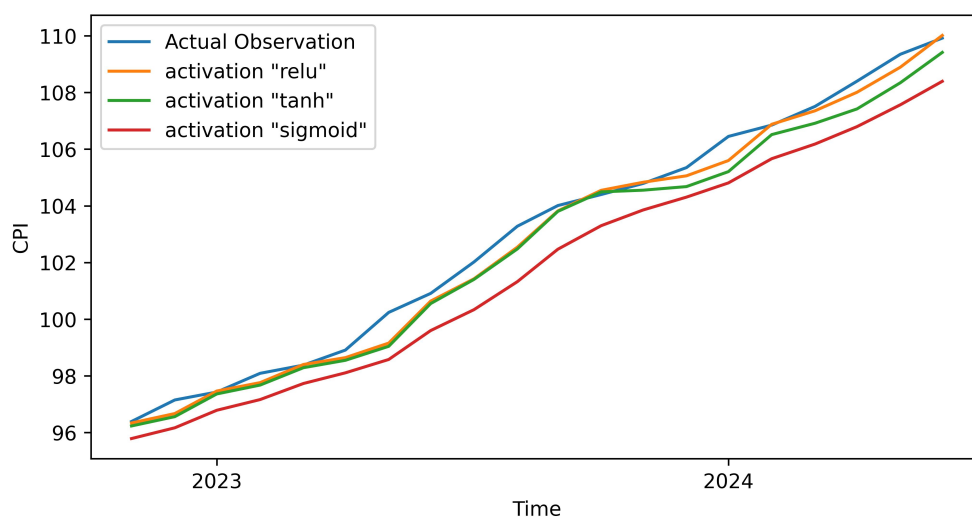


FIGURE 11. Actual (Test Data) and Predicted CPI Values Using the LSTM Models.

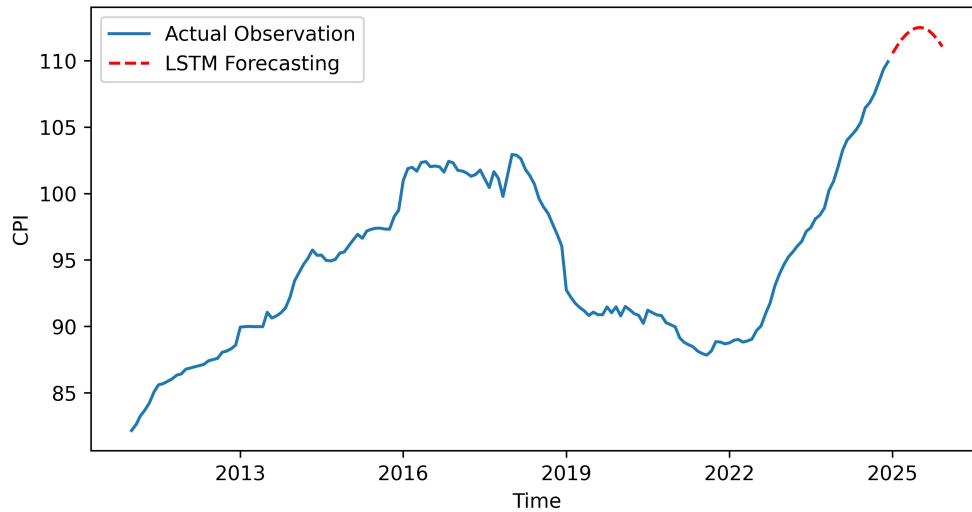


FIGURE 12. Actual Data and Forecasted CPI Values Using the LSTM (ReLU) Model.

It is evident from Figure 11 that the models are able to capture the trends and behaviors in the CPI data. The predicted values from the ReLU and Tanh activation functions exhibit consistent, coherent patterns, suggesting that the choice of activation function influences the LSTM model's forecasting performance. However, the forecasted results shown in Figure 12, obtained using the best LSTM model with ReLU, indicating that the LSTM model is less effective than the ANN model for long-term CPI forecasting.

4.3. Hybrid ARIMA-ANN Model. In developing the hybrid ARIMA-ANN model, we utilized the ANN with a ReLU activation function model. The network consists of two hidden layers with 21 and 8 neurons, and an output layer with a single neuron. The model was trained for 300 epochs with a batch size of 32. The predicted and forecasted values of the ARIMA-ANN model are presented in Figures 13 and 14, respectively.

Figures 13 and 14 illustrate the prediction and forecasting results of the ARIMA-ANN model, respectively. As shown in Figure 13, the predicted values closely follow the observed CPI data with minimal deviations, indicating that the model effectively captures the underlying patterns. Figure 14 presents the future CPI forecasts generated by the model, reflecting a consistent continuation of the learned trend.

4.4. Hybrid ARIMA-LSTM Model. In constructing the ARIMA-LSTM models, a LSTM with a ReLU activation function was selected. The model consists of a single hidden layer with 17 nodes and an output layer with a single node. Training was carried out over 300 epochs with a batch size of 32. Figures 15 and 16 show the predicted and forecasted values for ARIMA-LSTM model, respectively.

Figures 15 and 16 illustrate the prediction and forecasting results of the ARIMA-LSTM model, respectively. As shown in Figure 15, the predicted values align closely with the actual observations,

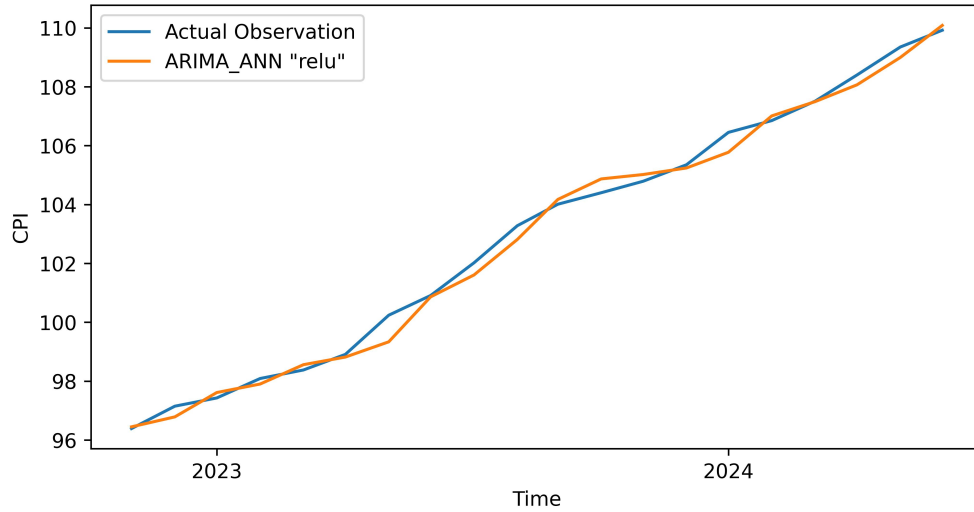


FIGURE 13. Actual (Test Data) and Predicted CPI Values Using the ARIMA-ANN Models.

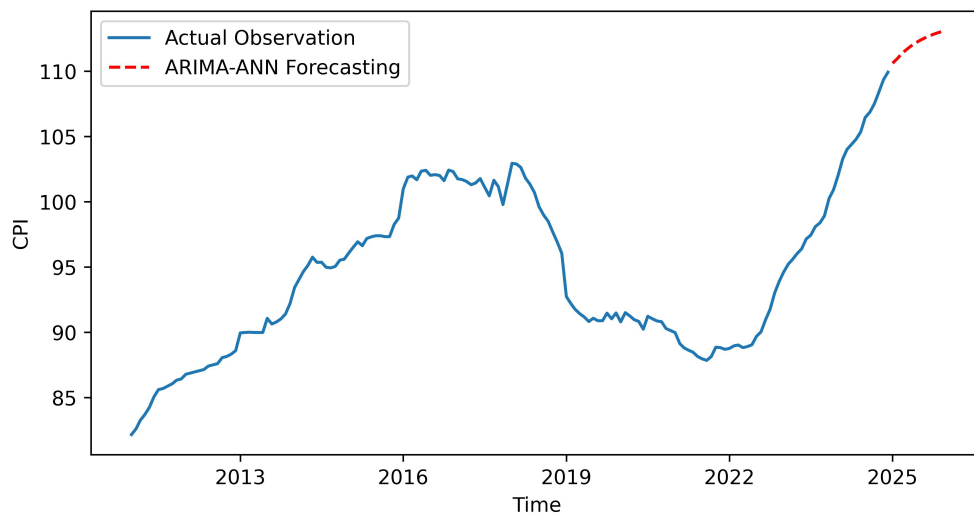


FIGURE 14. Actual Data and Forecasted CPI Values Using the ARIMA-ANN Model.

indicating satisfactory performance during the prediction phase. Figure 16 presents the model's future forecasts based on the learned temporal patterns.

5. DISCUSSION

Table 4 presents the forecast accuracy measures for all models. According to Table 4, the hybrid ARIMA-ANN model with a ReLU activation function outperformed all other models across the three prediction error metrics: RMSE, MSE, and MAE. The ARIMA-LSTM model with the ReLU activation function follows closely behind in performance.

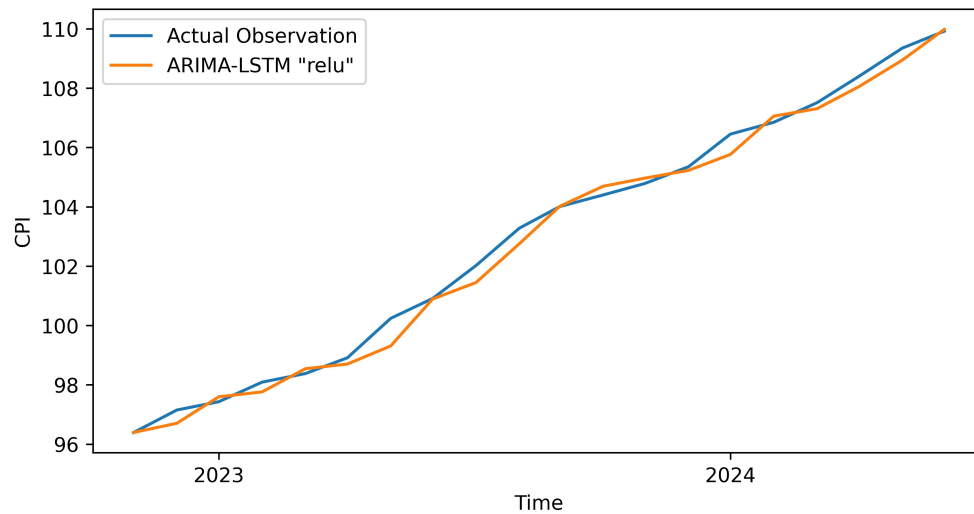


FIGURE 15. Actual (Test Data) and Predicted CPI Values Using the ARIMA-LSTM Models.

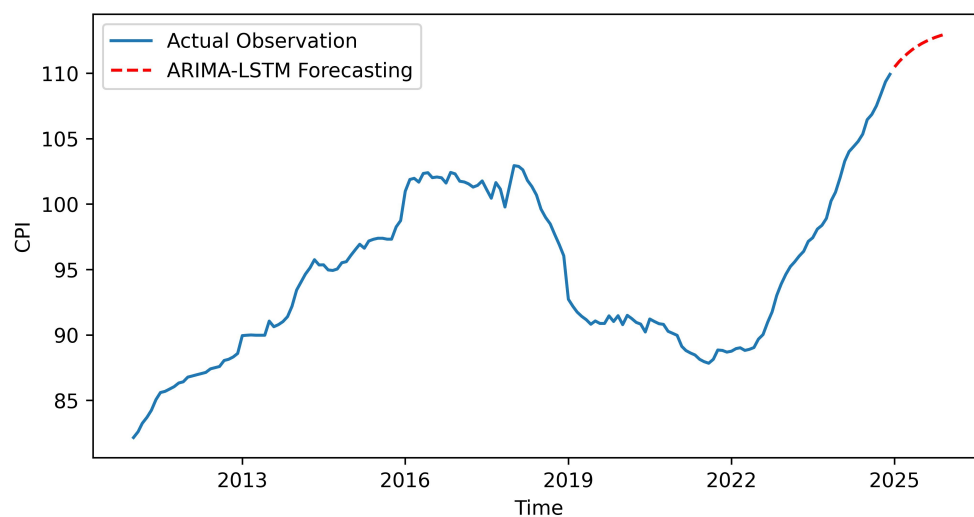


FIGURE 16. Actual Data and Forecasted CPI Values Using the ARIMA-LSTM Model.

Overall, the results indicate that hybrid models consistently outperform both standalone statistical and single machine learning models, whose performance varies across accuracy measures. Moreover, the ReLU activation function enhances forecasting accuracy, yielding superior results compared with alternative activation functions, and therefore is the most suitable choice for modeling the CPI data.

Based on these results, the hybrid ARIMA-ANN model with a ReLU activation function is the most effective approach for forecasting the CPI of housing, water, electricity, gas, and other fuels in Saudi Arabia. These results highlight the strong potential of hybrid modeling techniques for improving accuracy in economic forecasting applications.

TABLE 4. Accuracy measures for housing, water, electricity, gas, and other fuels CPI.

Methodology	Activation Function	RMSE	MSE	MAE
ARIMA	–	0.4019	0.1615	0.307
ANN	ReLU	0.362	0.131	0.281
ANN	Tanh	0.778	0.605	0.635
ANN	Sigmoid	1.524	2.324	1.412
LSTM	ReLU	0.437	0.191	0.326
LSTM	Tanh	0.635	0.403	0.526
LSTM	Sigmoid	1.309	1.715	1.245
ARIMA-ANN	ReLU	0.354	0.126	0.278
ARIMA-LSTM	ReLU	0.377	0.142	0.293

Based on these results, the strong performance of hybrid models, specifically the ARIMA-ANN model with the ReLU activation function, has significant implications for economic policymaking and financial planning in Saudi Arabia. Accurate CPI forecasting helps improve inflation monitoring and supports policymakers in making informed and forward-looking economic decisions. This finding aligns with the goals of Saudi Vision 2030, which focuses on economic diversification, financial stability, and the use of data-driven governance.

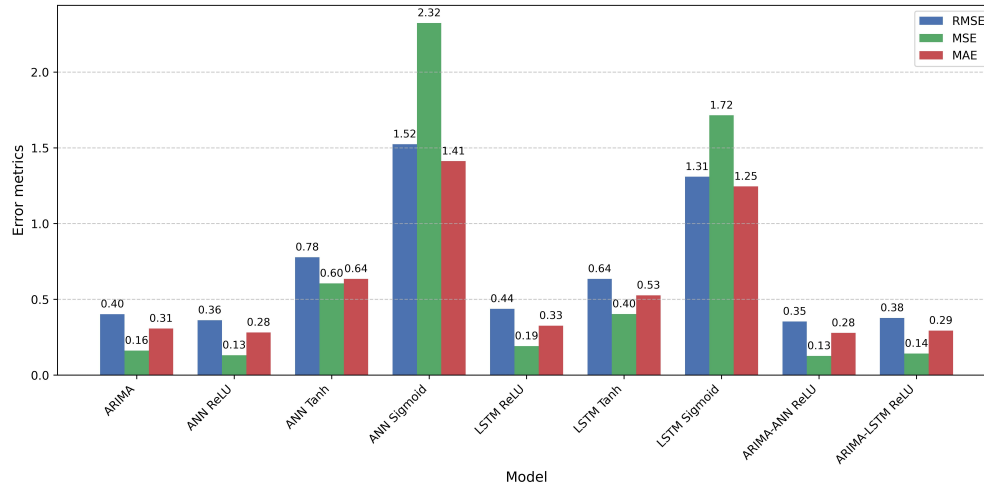


FIGURE 17. Comparison of Forecasting Models Using Different Accuracy Metrics.

Figure 17 compares the forecasting performance of different statistical, machine learning, and hybrid models using three accuracy measures: RMSE, MSE, and MAE. The results highlight differences in model performance across the selected error metrics.

Table 5 presents the monthly forecasted CPI for housing, water, electricity, gas, and other fuels in Saudi Arabia for 2025 generated by the best-performing forecasting model, ARIMA-ANN. The results provide insights into the expected trend of CPI over the forecast horizon.

TABLE 5. Forecasted CPI values for 2025 using the best-performing model ARIMA-ANN.

Date	Forecasted Value
Jan 2025	110.605
Feb 2025	110.949
Mar 2025	111.307
Apr 2025	111.620
May 2025	111.890
Jun 2025	112.129
Jul 2025	112.337
Aug 2025	112.519
Sep 2025	112.678
Oct 2025	112.817
Nov 2025	112.938
Dec 2025	113.044

6. CONCLUSIONS

This article compared statistical time series model (ARIMA), machine learning (ML) models (ANN and LSTM), and hybrid models (ARIMA-ANN and ARIMA-LSTM), using three activation functions (ReLU, Tanh, and Sigmoid) to forecast the CPI for housing, water, electricity, gas, and other fuels in Saudi Arabia over 12 months. The analysis was conducted using monthly data collected over 14 years. Forecasting performance was evaluated using three accuracy metrics: RMSE, MSE, and MAE.

The results indicated that the hybrid model effectively combines the strengths of ARIMA in capturing linear patterns and ML models in modeling nonlinear patterns. The hybridization approach improved forecasting accuracy compared to single statistical or ML models. Among all models, the hybrid ARIMA-ANN with ReLU activation function achieved the best performance, followed by the ANN with ReLU activation function. As the complexity in the time series increases, the efficiency of using statistical or ML models alone decreases relative to hybrid models.

Accurate CPI forecasting is essential for effective economic decision-making and long-term planning. Reliable forecasts help track inflation trends and design proactive monetary and fiscal policies that maintain economic stability. Forecasting the CPI's components also enables policymakers to enhance economic resilience and support sustainable development in key sectors of the economy, in line with the objectives of Saudi Vision 2030.

Conflicts of Interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

REFERENCES

- [1] R.A. Khan, I. Ghulam, Saudi Arabia's Economic Outlook Report 2025, SNB Capital, 2025.
- [2] A. Ali, A. Mahgoub, Prediction of CPI in Saudi Arabia: Holt's Linear Trend Approach, *Res. World Econ.* 11 (2020), 302–309. <https://doi.org/10.5430/rwe.v11n6p302>.
- [3] CEIC Data, Saudi Arabia Consumer Price Index (CPI), 2023.
- [4] General Authority for Statistics, General Authority for Statistics, 2023. <https://www.stats.gov.sa>.
- [5] S. Febriyanti, W.A. Pradana, J.S. Muhammad, E. Widodo, Forecasting the Consumer Price Index in Yogyakarta by Using the Double Exponential Smoothing Method, *Param. J. Stat.* 2 (2021), 1–7. <https://doi.org/10.22487/27765660.2021.v2.i1.15641>.
- [6] S. Zhang, Z. Lin, W. Yhang, Forecasting CPI of Restaurants and Hotels in Korea Using the Seasonal Autoregressive Integrated Moving Average (SARIMA) Model, *Int. J. Tour. Hosp. Res.* 37 (2023), 85–94. <https://doi.org/10.21298/ijthr.2023.4.37.4.85>.
- [7] I. Mahuda, S.D. Rahmawati, S. Sukmawati, S. Abdullah, Implementation of ARIMA Method to Forecast CPI for COICOP of Food, Beverage and Tobacco in New Normal Period, *J. Bayesian: J. Ilm. Stat. Ekon.* 4 (2024), 165–176.
- [8] A. Hussein, A. Abdillahi, Modelling and Forecasting Somalia's Consumer Price Index Using the ARIMA Model, *Am. J. Theor. Appl. Stat.* 14 (2025), 89–98. <https://doi.org/10.11648/j.ajtas.20251402.14>.
- [9] J. Riofrío, O. Chang, E.J. Revelo-Fuelagán, D.H. Peluffo-Ordóñez, Forecasting the Consumer Price Index (CPI) of Ecuador: a Comparative Study of Predictive Models, *Int. J. Adv. Sci. Eng. Inf. Technol.* 10 (2020), 1078–1084. <https://doi.org/10.18517/ijaseit.10.3.10813>.
- [10] I.Y. Purbasari, F.T. Anggraeny, N.A. Ardiningrum, Time-Series Modeling for Consumer Price Index Forecasting Using Comparison Analysis of AutoRegressive Integrated Moving Average and Artificial Neural Network, in: *Proceedings of the International Conference on Culture Heritage, Education, Sustainable Tourism, and Innovation Technologies*, SCITEPRESS - Science and Technology Publications, pp. 599–604, 2020. <https://doi.org/10.5220/0010369205990604>.
- [11] M. Qureshi, A. Khan, M. Daniyal, K. Tawiah, Z. Mehmood, A Comparative Analysis of Traditional SARIMA and Machine Learning Models for CPI Data Modelling in Pakistan, *Appl. Comput. Intell. Soft Comput.* 2023 (2023), 3236617. <https://doi.org/10.1155/2023/3236617>.
- [12] Y.C. Cham, M.H. Muhammed Nor, B.K.B. Lee, Evaluation of Machine Learning Techniques for Forecasting Malaysia's Consumer Price Index: a Comparative Study, *J. Qual. Meas. Anal.* 20 (2024), 199–214. <https://doi.org/10.17576/jqma.2003.2024.14>.
- [13] M.C.F.A. Ferreira, Forecasting Inflation: Can Machine Learning Outperform Econometric Models?, Master's Thesis, Universidade NOVA de Lisboa, 2024.
- [14] H. Jamil, Inflation Forecasting Using a Hybrid ARIMA–LSTM Model, PhD Thesis, Laurentian University, 2022.
- [15] M. Hadwan, B.M. Al-Maqaleh, F.N. Al-Badani, R.U. Khan, M.A. Al-Hagery, A Hybrid Neural Network and Box-Jenkins Models for Time Series Forecasting, *Comput. Mater. Contin.* 70 (2022), 4829–4845. <https://doi.org/10.32604/cmc.2022.017824>.
- [16] T.T.H. Le, T.N. Nguyen, A Hybrid Model Based on ARIMA and Artificial Neural Network to Forecast Consumer Price Index: The Case of Vietnam, in: N. Ngoc Thach, V. Kreinovich, D.T. Ha, N.D. Trung, (eds) *Optimal Transport Statistics for Economics and Related Topics*, Studies in Systems, Decision and Control, Vol 483, Springer, Cham, (2024). https://doi.org/10.1007/978-3-031-35763-3_32.
- [17] J.P.O. Echevarria, P.J.B. Aranas, Forecasting the Consumer Price Index in the Regions of the Philippines using Machine Learning for Time Series Models, *J. Artif. Intell. Mach. Learn. Neural Netw.* 3 (2023), 11–22. <https://doi.org/10.55529/jaimlnn.36.11.22>.
- [18] T. Nyoni, Predicting Consumer Price Index in Saudi Arabia, MPRA Paper, (2019). <https://ideas.repec.org/p/pramprapa/92422.html>.

- [19] A. Mahgoub, T. Alam, Modelling and Predicting the Consumer Price Index in Saudi Arabia, *J. Infrastruct. Policy Dev.* 8 (2024), 5270. <https://doi.org/10.24294/jipd.v8i9.5270>.
- [20] F. Sibai, A. El-Moursy, A. Sibai, Forecasting the Consumer Price Index: a Comparative Study of Machine Learning Methods, *Int. J. Comput. Digit. Syst.* 15 (2024), 487–497. <https://doi.org/10.12785/ijcds/150137>.
- [21] G.S. da S. Gomes, T.B. Ludermit, L.M.M.R. Lima, Comparison of New Activation Functions in Neural Network for Forecasting Financial Time Series, *Neural Comput. Appl.* 20 (2010), 417–439. <https://doi.org/10.1007/s00521-010-0407-3>.
- [22] H. Chang, S. Nakaoka, H. Ando, Effect of Shapes of Activation Functions on Predictability in the Echo State Network, *arXiv:1905.09419*, 2019. <https://doi.org/10.48550/arXiv.1905.09419>.
- [23] A. Salam, A.E. Hibaoui, A. Saif, A Comparison of Activation Functions in Multilayer Neural Network for Predicting the Production and Consumption of Electricity Power, *Int. J. Electr. Comput. Eng.* 11 (2021), 163–172. <https://doi.org/10.11591/ijece.v11i1.pp163-170>.
- [24] N.H.A. Rahman, C.H. Yin, H.S. Zulkafli, Activation Functions Performance in Multilayer Perceptron for Time Series Forecasting, *AIP Conf. Proc.* 3158 (2024), 070001. <https://doi.org/10.1063/5.0223864>.
- [25] G.E.P. Box, G.M. Jenkins, G.C. Reinsel, *Time Series Analysis: Forecasting and Control*, Wiley, 1970.
- [26] M. Wazid, A.K. Das, V. Chamola, Y. Park, Uniting Cyber Security and Machine Learning: Advantages, Challenges and Future Research, *ICT Express* 8 (2022), 313–321. <https://doi.org/10.1016/j.ict.2022.04.007>.
- [27] S.M. Mian, M.S. Khan, M. Shawez, A. Kaur, Artificial Intelligence (AI), Machine Learning (ML) & Deep Learning (DL): A Comprehensive Overview on Techniques, Applications and Research Directions, in: *2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, IEEE, pp. 1404–1409, 2024. <https://doi.org/10.1109/icscss60660.2024.10625198>.
- [28] A.L. Samuel, Some Studies in Machine Learning Using the Game of Checkers, *IBM J. Res. Dev.* 3 (1959), 210–229. <https://doi.org/10.1147/rd.33.0210>.
- [29] L.C.M. Dafico, E. Barreira, R.M.S.F. Almeida, R. Vicente, Machine Learning Models Applied to Moisture Assessment in Building Materials, *Constr. Build. Mater.* 405 (2023), 133330. <https://doi.org/10.1016/j.conbuildmat.2023.133330>.
- [30] C. Nwankpa, W. Ijomah, A. Gachagan, S. Marshall, Activation Functions: Comparison of Trends in Practice and Research for Deep Learning, *arXiv:1811.03378*, 2018. <https://doi.org/10.48550/arXiv.1811.03378>.
- [31] Y.A. LeCun, L. Bottou, G.B. Orr, K.R. Müller, Efficient BackProp, in: G. Montavon, G.B. Orr, K.R. Müller, (eds) *Neural Networks: Tricks of the Trade*, Lecture Notes in Computer Science, Vol 7700, Springer, Berlin, Heidelberg, (2012). https://doi.org/10.1007/978-3-642-35289-8_3.
- [32] V. Nair, G.E. Hinton, Rectified Linear Units Improve Restricted Boltzmann Machines, in: *Proceedings of the 27th International Conference on Machine Learning*, pp. 807–814, 2010.
- [33] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, MIT Press, 2016.
- [34] S. Haykin, *Neural Networks: a Comprehensive Foundation*, Prentice Hall, 1999.
- [35] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning Representations by Back-Propagating Errors, *Nature* 323 (1986), 533–536. <https://doi.org/10.1038/323533a0>.
- [36] M.H. Sazli, A Brief Review of Feed-Forward Neural Networks, *Commun. Fac. Sci. Univ. Ank. Ser. A2 A3 Math. Stat.* 50 (2006), 11–17. <https://doi.org/10.1501/0003168>.
- [37] S. Hochreiter, J. Schmidhuber, Long Short-Term Memory, *Neural Comput.* 9 (1997), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>.
- [38] A. Graves, Long Short-Term Memory, in: *Supervised Sequence Labelling with Recurrent Neural Networks*, Studies in Computational Intelligence, Vol 385, Springer, Berlin, (2012). https://doi.org/10.1007/978-3-642-24797-2_4.
- [39] R. Zhang, H. Song, Q. Chen, Y. Wang, S. Wang, et al., Comparison of ARIMA and LSTM for Prediction of Hemorrhagic Fever at Different Time Scales in China, *PLoS One* 17 (2022), e0262009. <https://doi.org/10.1371/journal.pone.0262009>.

- [40] G.P. Zhang, Time Series Forecasting Using a Hybrid ARIMA and Neural Network Model, *Neurocomputing* 50 (2003), 159–175. [https://doi.org/10.1016/s0925-2312\(01\)00702-0](https://doi.org/10.1016/s0925-2312(01)00702-0).
- [41] L. Wang, H. Zou, J. Su, L. Li, S. Chaudhry, An ARIMA-ANN Hybrid Model for Time Series Forecasting, *Syst. Res. Behav. Sci.* 30 (2013), 244–259. <https://doi.org/10.1002/sres.2179>.
- [42] M.V. Shcherbakov, A. Brebels, N.L. Shcherbakova, A.P. Tyukov, T.A. Janovsky, et al., A Survey of Forecast Error Measures, *World Appl. Sci. J.* 24 (2013), 171–176.
- [43] C. Chen, J. Twycross, J.M. Garibaldi, A New Accuracy Measure Based on Bounded Relative Error for Time Series Forecasting, *PLoS One* 12 (2017), e0174202. <https://doi.org/10.1371/journal.pone.0174202>.
- [44] P.C.B. Phillips, P. Perron, Testing for a Unit Root in Time Series Regression, *Biometrika* 75 (1988), 335–346. <https://doi.org/10.1093/biomet/75.2.335>.