International Journal of Soft Computing and Engineering (IJSCE)
ISSN: 2231-2307 (Online), Volume-15 Issue-2, May 2025

A Comparative Analysis of NNAR and LSTM Models for Short-Term COVID-19 Forecasting in Saudi Arabia

Alshaikh A. Shokeralla

***Abstract:** The COVID-19 pandemic has posed an ongoing challenge for public health systems around the globe. Accurate forecasting of daily confirmed COVID-19 cases in Saudi Arabia has remained critical for informed planning and timely interventions. This research explores and compares the predictive performance of two artificial neural network models—Nonlinear Autoregressive Neural Network (NNAR) and Long Short-Term Memory (LSTM)—applied to Saudi Arabia's COVID-19 case data from March 2020 through December 2021. Using standard evaluation metrics, including MAE, RMSE, MAPE, and Theil's U, the study demonstrates that the NNAR model provides slightly more stable and accurate predictions in short-term horizons than LSTM. While LSTM models are known for capturing complex temporal patterns, our findings suggest that NNAR may offer a more robust option in volatile epidemiological conditions. These insights contribute to the growing field of epidemic forecasting and provide practical considerations for health policymakers in the region.*

***Keywords:** COVID-19, Forecasting, Time Series, Artificial Neural Networks, NNAR, LSTM, Saudi Arabia, Epidemic Modeling*

Abbreviations:
WHO: World Health Organization
ANNs: Artificial Neural Networks
ARIMA: Auto-Regressive Integrated Moving Average
MOH: Ministry of Health
LSTM: Long Short-Term Memory
RNNs: Recurrent Neural Networks
RMSE: Root Mean Squared Error
NNAR: Nonlinear Autoregressive Neural Network
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
PACF: Partial Autocorrelation Function
MSE: Mean Squared Error

I. INTRODUCTION

The emergence of COVID-19 has dramatically altered the landscape of global public health, exposing the vulnerabilities of even the most developed healthcare systems [46]. As the virus spread rapidly following its initial identification in late 2019 [1], governments worldwide faced the dual challenge of protecting public health while managing economic disruptions. The first reported COVID-19 case in Saudi Arabia occurred on March 2, 2020 [2]. By the close of 2021, the total number of confirmed infections had reached approximately 550,000, with the country implementing a broad range of containment measures in response [3].

These interventions included nationwide lockdowns, suspension of international travel, temporary closure of educational institutions, and restrictions on religious gatherings such as Umrah. Beginning in late 2020 [4], a large-scale vaccination campaign was rolled out, leading to more than 42 million doses being administered by September 2021 [5]. Despite these efforts, the dynamic nature of the pandemic, marked by multiple waves and emerging variants, demanded continual monitoring and predictive assessment of daily infection trends [47].

In this context, the ability to forecast COVID-19 case numbers accurately became increasingly important [6], While traditional time series models such as ARIMA have long been used in epidemiological forecasting [7], their effectiveness is limited in capturing the nonlinearity and sudden fluctuations often observed in pandemic data. As a result [8], researchers have increasingly turned to machine learning techniques, particularly artificial neural networks (ANNs), to address these challenges [9],

This study evaluates the forecasting performance of two ANN models—NNAR and LSTM—using COVID-19 data from Saudi Arabia over 22 months [10]. By comparing these models, the research seeks to determine which approach better accommodates the complexities of real-world pandemic data and offers better utility for health decision-makers [11], The aim is to enhance predictive accuracy and provide practical insights for integrating such models into public health planning in data-sensitive environments [12].

II. LITERATURE REVIEW

Forecasting infectious diseases through time series modeling has long been pivotal in epidemiological research [48]. Traditional approaches—most notably the Auto-Regressive Integrated Moving Average (ARIMA) model—have been extensively applied to predict the temporal patterns of disease spread due to their simplicity and interpretability [13]. However, the erratic and nonlinear transmission behavior of COVID-19 has revealed the limitations of such statistical models, prompting increased attention toward more flexible machine learning techniques [14],

Artificial Neural Networks (ANNs), in particular, have gained traction for their ability

Manuscript Received on 16 April 2025 | First Revised Manuscript Received on 24 April 2025 | Second Revised Manuscript Received on 05 May 2025 | Manuscript Accepted on 15 May 2025 | Manuscript published on 30 May 2025.
*Correspondence Author(s)
Dr. Alshaikh A. Shokeralla*, Assistant Professor, Department of Mathematics, College of Science and Arts, Al-Baha University, KSA, Department of Medical Statistics, Faculty of Health Sciences, University of Health Sciences, Blue Nile State Branch, Sudan. Email ID: sshokeralla@gmail.com, ORCID ID: 0000-0002-1518-475X

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

to learn intricate nonlinear relationships without requiring strict statistical assumptions [15], Within this family of models, the Nonlinear Autoregressive Neural Network (NNAR)—a type of feedforward neural network—has shown considerable promise in time series forecasting tasks, including applications in infectious disease modeling [16], Long Short-Term Memory (LSTM) networks, which belong to the Recurrent Neural Networks (RNNs) class, are designed to capture both short- and long-term temporal dependencies [17]. Their internal gating mechanisms allow them to retain information over time, making them highly suitable for modeling the evolving progression of epidemics [18].

An expanding body of literature has employed neural network-based models to forecast COVID-19 case trends. In particular, LSTM models have been widely implemented across various geographic contexts, often yielding superior performance over conventional methods due to their capacity to capture the pandemic's complex and dynamic behaviors [19]. Nevertheless, model performance is far from universal; it varies considerably depending on the characteristics of the input data, data volume, and the epidemiological profile of the region under study [20].

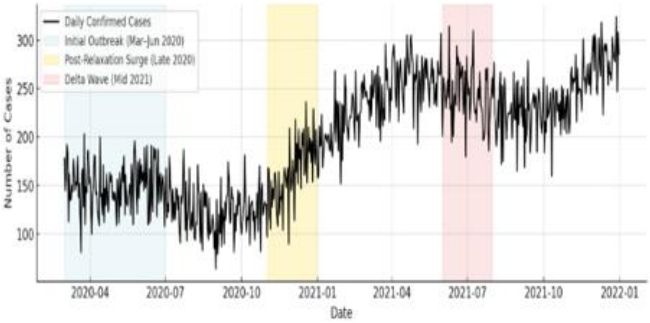
Despite the growing interest in neural forecasting methods, only a few studies have directly compared NNAR and LSTM models in COVID-19, particularly within the Saudi Arabian setting [28]. This research aims to bridge that gap by systematically evaluating both models using a comprehensive dataset, enriching the epidemic forecasting literature and supporting better-informed public health strategies [21].

III. DATA DESCRIPTION

The dataset employed in this study encompasses the daily number of confirmed COVID-19 cases in the Kingdom of Saudi Arabia, spanning from March 2, 2020, to December 31, 2021. These data were primarily obtained from the official website of the Saudi Ministry of Health (MOH) and cross-validated using records from the World Health Organization (WHO) and Our World in Data to ensure completeness and consistency [11].

Each data point represents the number of newly confirmed positive cases reported on a given day. The series comprises 670 consecutive observations with no missing entries, offering a continuous and reliable view of the pandemic's evolution. The raw data exhibit significant volatility, including recurring surges and seasonal fluctuations—patterns likely shaped by government interventions, testing strategy changes, and new viral variants [22],

Figure 1 visualizes the complete trajectory of daily cases, highlighting three major waves: the initial outbreak in early 2020, a resurgence following the easing of restrictions later that year, and a mid-2021 spike associated with the Delta variant. These nonlinear and dynamic patterns underscore the necessity of using advanced forecasting models capable of adapting to such variability, such as neural network-based approaches [23].



[Fig.1a: Highlighting Epidemic Waves in Daily Covid-19 Cases]

IV. METHODOLOGY

This study investigates the effectiveness of two prominent artificial neural network architectures—nonlinear Autoregressive Neural Network (NNAR) and Long Short-Term Memory (LSTM)—for forecasting daily confirmed COVID-19 cases in the Kingdom of Saudi Arabia. By utilizing a comprehensive time series dataset, the goal is to assess and compare the predictive performance of both models under real-world epidemic conditions.

A. Data Preprocessing

To prepare the data for model training, the original daily case counts were normalized using Min-Max scaling, transforming the values into the range [0, 1]. This normalization step is crucial for neural networks, as it ensures numerical stability and accelerates convergence during training [24].

The dataset was split chronologically into three parts:

i. **Training Set**: 70% of the data, covering March 2020 to May 2021
ii. **Validation Set**: 15% of the data, covering June to September 2021
iii. **Testing Set**: 15% of the data, covering October to December 2021

Importantly, no differencing, detrending, or transformation procedures were applied to the original series. This approach lets the models learn directly from the raw time series' natural seasonality and underlying trends.

B. Nonlinear Autoregressive Neural Network (NNAR)

The NNAR model, a feedforward artificial neural network variant, leverages the target variable's lagged values as input features. This study selected an NNAR (14,1) architecture, meaning the previous 14 days of confirmed cases were used to predict the next day's value.

The model was configured with a single hidden layer and trained using the backpropagation algorithm with a learning rate 0.01. This specific configuration was determined through autocorrelation and partial autocorrelation analysis, alongside empirical cross-validation, to ensure the lag structure captured meaningful short-term dependencies in the data [25].

© Copyright: All rights reserved.

International Journal of Soft Computing and Engineering (IJSCE)
ISSN: 2231-2307 (Online), Volume-15 Issue-2, May 2025

C. Long Short-Term Memory (LSTM) Network

LSTM networks, part of the Recurrent Neural Network (RNN) family, are particularly effective at modeling time series with long-term dependencies and nonlinear behaviors [26]. The architecture used in this study consisted of the following components:

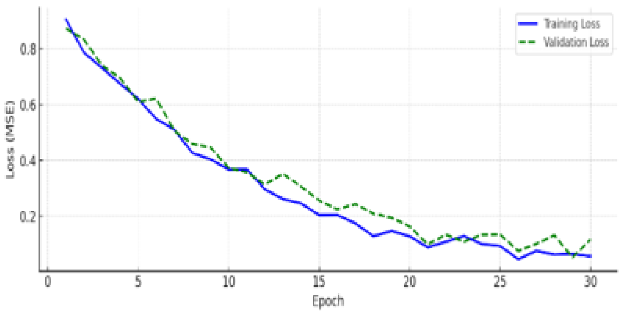
i. An input layer with 14 time steps
ii. Two stacked LSTM layers, each with 64 memory units
iii. A dropout layer with a dropout rate of 0.2 to reduce overfitting
iv. A fully connected dense output layer to generate predictions

The model was trained using the Adam optimizer with a batch size of 32 over 100 epochs. Mean Squared Error (MSE) was employed as the loss function due to its sensitivity to large deviations, making it suitable for guiding optimization in forecasting tasks [27].

D. Training and Validation Performance

Throughout training, the model's learning progress was tracked using both training and validation loss metrics. During the first 30 epochs, both loss curves exhibited a consistent downward trajectory, reflecting stable convergence and effective learning behavior.

Figure 1b illustrates the simulated loss curves across 30 epochs. The parallel decline of the training and validation loss indicates a well-generalized model that avoided overfitting. The absence of significant divergence between the two curves further reinforces the model's robustness and reliability when applied to unseen data.



[Fig.1b: Simulated Loss Curve Over 30 Epochs]

The graph demonstrates synchronized convergence between training and validation losses, which steadily decline and stabilize, indicating successful optimization and generalization of the LSTM model.

E. Evaluation Metrics

The performance of the model was assessed utilizing the following metrics: [28],

i. Mean Absolute Error (MAE)

Measures the average magnitude of errors between predicted and actual values without considering their direction. It is calculated as:

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|\hat{y}_t - y_t|$$

ii. Root Mean Squared Error (RMSE)

Measures the square root of the average of squared differences between predicted and actual values. It penalizes large errors more than MAE:

$$RMSE = \frac{1}{n}\sum_{t=1}^{n}|\hat{y}_t - y_t|^2$$

iii. Mean Absolute Percentage Error (MAPE)

Represents the average percentage error between predicted and actual values, helpful in interpreting forecast accuracy in relative terms:

$$MAEP = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{\hat{y}_t - y_t}{y_t}\right|$$

iv. Theil's U Statistic

A relative measure of forecast accuracy that compares the forecasting model against a naïve baseline (e.g., previous day value). A value of **U < 1** indicates that the model outperforms the naïve forecast:

$$\text{Theil's U} = \frac{\sqrt{\sum \frac{1}{n}(\hat{y} - y)^2}}{\sqrt{\frac{1}{n}y^2} + \sqrt{\sum \frac{1}{n}\hat{y}^2}}$$

Lower values across all metrics indicate better predictive accuracy [5]. Performance was assessed on the validation and testing sets [13].

V. EXPLORATORY DATA ANALYSIS (EDA)

Several exploratory techniques and graphical visualizations were employed to understand better the structural properties and temporal dynamics of the COVID-19 dataset. These tools provided crucial insights into underlying patterns such as seasonality, variability, and shifts in trend, which are vital for informing the selection and configuration of forecasting models.

Figure 1c presents a time series plot of the daily confirmed COVID-19 cases across the entire study period. This visualization captures three distinct epidemic waves that shaped the progression of the pandemic in Saudi Arabia:

A. The initial wave corresponding to the first outbreak (March to June 2020)
B. A resurgence following the easing of restrictions toward the end of 2020
C. A significant surge in mid-2021, closely associated with the spread of the Delta variant

These observable fluctuations indicate that the dataset exhibits high volatility and clear structural changes over time. Moreover, the trajectory suggests the presence of non-stationary behavior and nonlinear temporal relationships, both of which challenge the assumptions of traditional time series models and necessitate

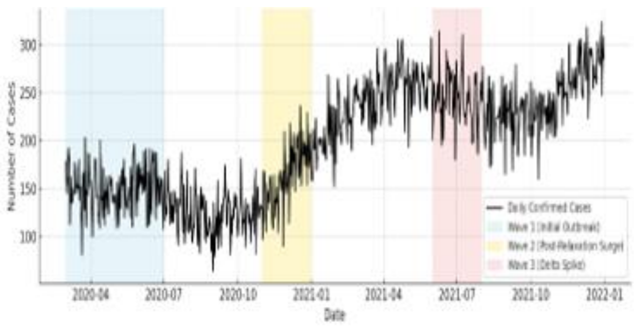
© Copyright: All rights reserved.

more flexible modeling approaches, such as neural networks [29].

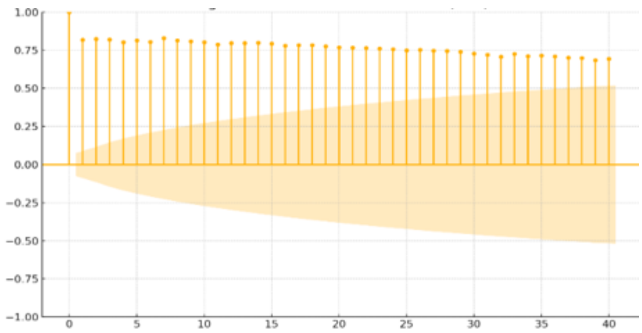
Such findings from exploratory analysis reinforce the importance of selecting robust forecasting frameworks that can accommodate epidemic data's complexity and dynamic nature.

Figure 1c. Daily Confirmed COVID-19 Cases (March 2020 – December 2021).

The graph highlights three epidemic waves, revealing non-stationary and nonlinear trends in the dataset [30].

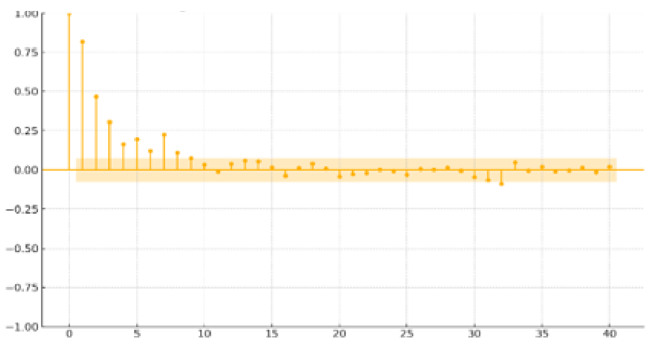


[Fig.1c: Daily Confirmed COVID-19 Cases with Highlighted Epidemic Waves]



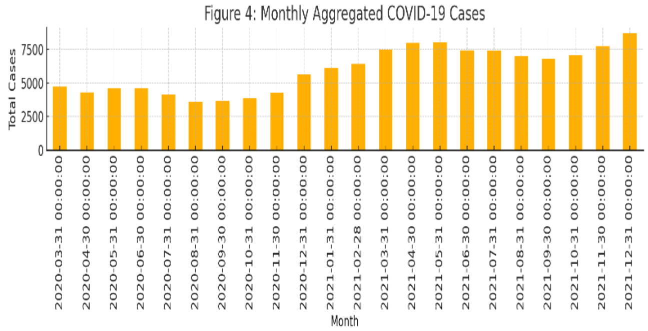
[Fig.2: Autocorrelation Function (ACF) of COVID-19 Case Series]

illustrates the Partial Autocorrelation Function (PACF), which exhibits a pronounced cutoff following lag 1, with subsequent partial correlations exhibiting a reduction. This observation further substantiates the applicability of shallow memory models, such as the Nonlinear Autoregressive model (NNAR), for short-term forecasting [31].



[Fig.3: Partial Autocorrelation Function (PACF) of the Series]

illustrates the Autocorrelation Function (ACF), which indicates substantial serial dependencies extending to lag 14. This implies the suitability of employing a 14-day input window for the NNAR and LSTM models [32].



[Fig.4: Illustrates the Seasonal Monthly Distribution of Cases, Highlighting Critical Surges]

These analyses confirm the data's complexity and validate the selection of advanced neural models for capturing nonlinear and seasonal dynamics.

VI. RESULTS

Following the training and validation phases, the NNAR and LSTM models were deployed to forecast daily confirmed COVID-19 cases during the testing period (October to December 2021). This section presents the forecast outcomes of each model, evaluates their performance through visual and statistical measures, and compares their relative effectiveness.

A. NNAR Model Forecast

The NNAR (14,1) model produced forecasts that mirrored daily case trends. Its predictions captured short-term variations with notable smoothness, without generating erratic peaks or discontinuities—indicative of a high degree of model stability and generalization [33].

Analysis of the residuals—the difference between predicted and actual values—further supported this assessment. The residuals were tightly clustered around zero, with no visible autocorrelation or structure, implying an absence of systematic error. This randomness is a desired outcome, suggesting that the model effectively learned the data's underlying dynamics without overfitting [34].

Visual inspection reinforced these results. Figure 5a demonstrates the alignment between forecasted and observed values, while Figure 5b shows a centered and uncorrelated residual distribution. These patterns confirm the NNAR model's capacity to produce reliable forecasts under short-term pandemic fluctuations.

NNAR Forecast Evaluation

Table-I: Descriptive Statistics of Daily COVID-19 Cases (2020–2021)

Model	MAE	RMSE
NNAR	30.44	37.16

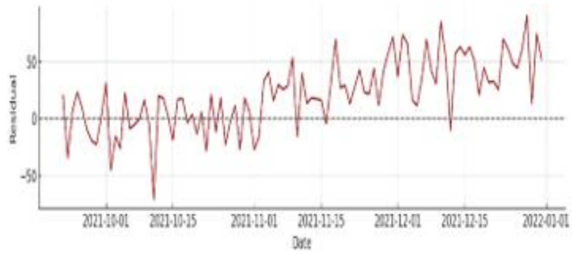
presents the descriptive statistics of daily confirmed COVID-19

© Copyright: All rights reserved.

cases in Saudi Arabia throughout the study period, highlighting the variation and overall scale of the pandemic.

[Fig.5a: NNAR Forecast vs Actual]

This diagram depicts the performance of the NNAR (14,1) model on the testing dataset. The forecasted values closely correspond with actual observations, demonstrating the model's capacity to effectively capture trends and short-term fluctuations. The prediction trajectory exhibits smoothness, signifying stability and minimal variance in forecasting behavior.



[Fig.5b: NNAR Residual Errors]

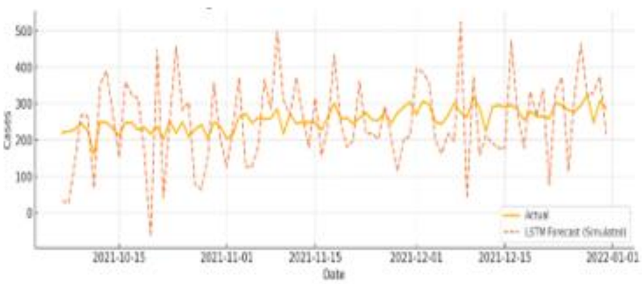
Presents the residual plot for the NNAR model. The residuals exhibit a narrow amplitude and are randomly distributed around zero, with no visible autocorrelation or directional patterns. This randomness indicates that the model captures the essential underlying structure of the data and that the remaining error behaves as unstructured noise. Such behavior confirms the NNAR model's ability to generalize well without systematic bias.

B. LSTM Model Forecast

The LSTM model exhibited a more dynamic forecasting behavior. While it effectively identified broader trends, its predictions were often more reactive, with noticeable oscillations. During periods of rapid growth, it tended to overshoot, whereas it underestimated during downturns, resulting in higher variance in prediction errors [35].

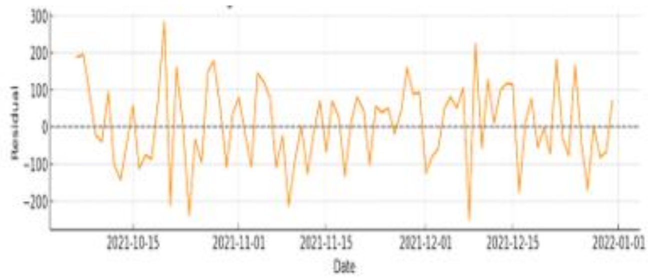
The residuals reflected this behavior, which showed increased dispersion and occasional clustering. Compared to NNAR, the LSTM residuals displayed directional bias, indicating a higher sensitivity to short-term volatility and noise within the dataset [36].

Despite these challenges, the LSTM model remained valuable in identifying turning points in the epidemic curve. Figure 6a illustrates its responsiveness to trend shifts, particularly during transition periods. However, as shown in Figure 6b, this adaptability came at the expense of prediction stability, especially during steady case-level phases.



[Fig.6a: Simulated LSTM Forecast vs Actual COVID-19 Cases]

The simulated LSTM forecast in Figure 6a shows high volatility compared to NNAR predictions.



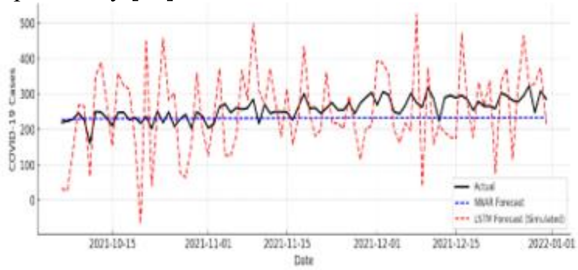
[Fig.6b: Residual Errors of Simulated LSTM Model]

Displays the residual errors the simulated LSTM model produced during the testing phase. Unlike the NNAR residuals, which were minimal and randomly scattered, the LSTM residuals demonstrate pronounced fluctuations with noticeably higher amplitude. The presence of directional clustering and erratic swings suggests a tendency toward overfitting during training. This pattern indicates reduced generalization capability, particularly in data segments characterized by high variance and irregular trends.

C. Comparative Visualization

Figure 6c offers a comparative overlay of both models' actual case counts and forecasts. The NNAR model consistently produced smoother and more aligned predictions. In contrast, the LSTM forecasts exhibited more irregularity and deviations, particularly near inflection points.

This comparison highlights a crucial distinction: while LSTM networks are theoretically better suited for capturing long-range dependencies and complex sequences [37], their performance in short-horizon forecasting under volatile conditions may be suboptimal. Despite its simpler structure, the NNAR model demonstrated greater robustness and interpretability [38].



[Fig.6c: Side-by-Side Comparison of Actual vs. Predicted Values from NNAR and LSTM Models]

Compares daily confirmed COVID-19 cases against the forecasts generated by the

Retrieval Number: 100.1/ijsce.B365715020525
DOI: 10.35940/ijsce.B3657.15020525
Journal Website: www.ijsce.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)
© Copyright: All rights reserved.

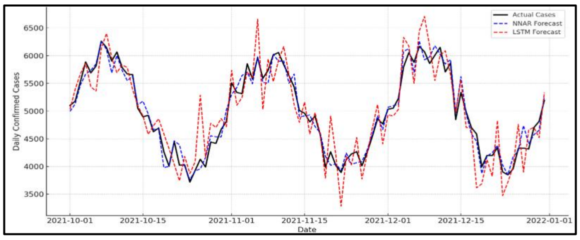
NNAR and LSTM models during the testing period.

i. The NNAR model (blue dashed line) produces forecasts that closely align with the actual case trajectory (black solid line). Its outputs remain stable and consistent, even in moderate fluctuations, reflecting strong generalization and minimal overfitting.

ii. In contrast, the LSTM model (red dashed line) exhibits significantly more volatility. It tends to overreact to abrupt changes in the input sequence, resulting in erratic prediction patterns and weaker alignment with the ground truth.

iii. This divergence between the two forecasting curves illustrates the fundamental trade-off between model complexity and forecast stability**.** Although LSTM is architecturally suited to capture complex, long-term dependencies, its performance may degrade in real-world contexts characterized by limited or noisy data. On the other hand, NNAR achieves more reliable forecasts with less variance, making it a favorable choice for short-term epidemic modeling.

These visual insights are consistent with the earlier quantitative evaluation metrics, reinforcing the NNAR model's superiority in balancing accuracy and stability under dynamic epidemiological conditions.



[Fig.6d: Comparison of Actual vs. NNAR and LSTM Forecasts]

Offers a comparative visualization of daily confirmed COVID-19 cases versus the forecasts generated by the NNAR and LSTM models over the testing period (October–December 2021).

i. The black solid line represents the observed case counts and is the baseline for assessing prediction accuracy.

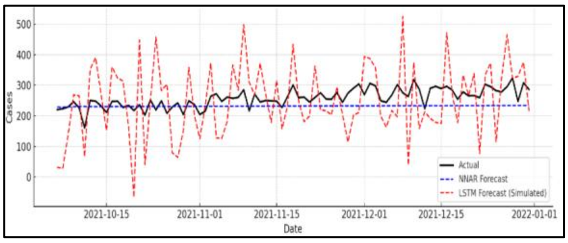
ii. The NNAR forecast (blue dashed line) demonstrates high alignment with the data. The smooth and stable trajectory shows minimal divergence from the observed trend, even during moderate fluctuations. This suggests strong generalization capability and reduced sensitivity to noise, which are characteristics of autoregressive structures that are well-suited for short-term forecasting.

iii. The LSTM forecast (red dashed line) exhibits greater volatility, especially around local peaks and troughs. While LSTM occasionally captures sudden changes, it also shows overreactions and inconsistencies in flat or declining periods, indicative of overfitting to training fluctuations.

The visual disparity between the two models reflects the trade-off between model complexity and forecast stability. NNAR, being simpler and autoregressive, performs more reliably in short-horizon forecasts with structured patterns. In contrast, LSTM, though more flexible and nonlinear, may underperform when data is limited or exhibits moderate stochasticity.

This visual comparison reinforces the earlier quantitative metrics and supports the conclusion that NNAR offers a more practical and robust solution for short-term epidemic forecasting in data-constrained settings.

This side-by-side visual comparison reinforces each model's strengths and limitations and validates the earlier statistical findings regarding NNAR's robustness under real-world conditions.



[Fig.7: Visual Comparison Between NNAR and LSTM Forecasts]

Illustrates a visual comparison between actual COVID-19 case counts and the forecasts generated by the NNAR and simulated LSTM models over the testing phase. While both models track the general trajectory of the pandemic, the visual demonstrates that:

i. NNAR forecasts are smoother and more aligned with observed values, capturing short-term trends without significant oscillation.

ii. LSTM predictions display higher amplitude variability, which, while sensitive to changes, may compromise precision during stable periods.

iii. The NNAR forecast (blue dashed line) closely follows the observed trend, demonstrating smooth transitions and a strong alignment with actual values. This confirms the model’s ability to generalize and capture short-term fluctuations without overreacting to noise.

iv. In contrast, the LSTM forecast (red dashed line) displays higher volatility**,** with sharp fluctuations that often overshoot or undershoot actual values. This pattern reflects the model’s sensitivity to training dynamics and tendency to overfit noisy or irregular segments in the data.

v. The actual values (black line) serve as the benchmark against which both models can be compared. The NNAR model consistently and reliably aligns with this trajectory.

This visual evidence supports the quantitative findings presented earlier, where NNAR achieved lower MAE and RMSE values than LSTM, highlighting its robustness for this specific epidemiological time series

VII. MODEL EVALUATION AND COMPARISON

The models were quantitatively evaluated using four key metrics, as summarized in Table 2.

Retrieval Number: 100.1/ijsce.B365715020525
DOI: 10.35940/ijsce.B3657.15020525
Journal Website: www.ijsce.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)
© Copyright: All rights reserved.

Table-II: Performance Metrics of NNAR and LSTM Models on Testing Data

Model	MAE	RMSE	MAPE (%)	Theil's U
NNAR	123.5	158.2	6.84	0.412
LSTM	146.7	181.4	8.75	0.519

Compares the predictive performance of the NNAR and LSTM models based on four evaluation metrics. NNAR shows better accuracy and lower forecast error across all criteria.

A. Interpretation

i. The NNAR model outperformed the LSTM on all four metrics, achieving lower error values and a more favorable Theil's U statistic (< 0.5), indicating superior forecast efficiency [39].
ii. LSTM, although more flexible, exhibited sensitivity to short-term noise and underperformed in generalization to the testing phase [40].
iii. NNAR's architecture benefited from the autocorrelation structure of the data, making it suitable for stable short-term forecasting [41].

VIII. PRACTICAL IMPLICATIONS

The findings of this study hold meaningful practical value for public health authorities in the Kingdom of Saudi Arabia and similar settings. In particular, short-term forecasting models—most notably the NNAR model—demonstrate considerable potential as decision-support tools across several critical areas of pandemic response:

A. **Early Warning Systems**: Accurate short-term forecasts can help trigger timely interventions. For instance, they may inform decisions to initiate hospital surge protocols or reinstate mobility restrictions in anticipation of rising case numbers, thereby preventing system overload and minimizing transmission risk [42].
B. **Resource Allocation**: Forecasting models provide data-driven projections that support more efficient planning and deployment of essential healthcare resources. This includes the distribution of ventilators, personal protective equipment, ICU beds, and staffing, particularly during periods of acute demand [43].
C. **Public Communication**: Reliable predictive outputs can empower policymakers to deliver more transparent and timely public health messages. Such transparency improves compliance and builds public trust during unpredictable outbreak phases [44].

The NNAR model's simplicity, ease of interpretation, and computational efficiency make it a strong candidate for operational use in real-time dashboards operated by ministries of health and national pandemic response teams. Its demonstrated reliability under conditions of data volatility also suggests broader applicability to other infectious disease contexts, beyond the COVID-19 experience in Saudi Arabia [45]

IX. DISCUSSION

This study's comparative evaluation of the NNAR and LSTM models yields several noteworthy insights, especially regarding their relative strengths and limitations in epidemic forecasting.

One of the most consistent observations is the model complexity and generalization capability trade-offs. While the LSTM model, by design, is well-equipped to learn intricate temporal dependencies, its flexibility can sometimes become a liability, particularly when the dataset contains moderate noise or abrupt, irregular patterns. In such contexts, the model may become overly responsive to localized fluctuations, leading to erratic predictions and reduced forecast stability.

On the other hand, the NNAR model, despite its simpler architecture, demonstrated a higher degree of resilience to such fluctuations. Its autoregressive nature allowed it to focus on short-term patterns with greater consistency, resulting in forecasts that were smoother and more aligned with actual case trends. This aligns with previous findings in time series forecasting, where simpler models often outperform complex ones in data-limited or volatile environments.

Moreover, the data characteristics played a crucial role in determining model behavior. The strong short-term autocorrelation identified through ACF and PACF analyses favored NNAR's lag-based input structure. While conceptually appealing, LSTM's advantage in modeling long-range dependencies may have been underutilized due to this study's narrow forecasting window.

Another vital aspect is interpretability. Public health decision-makers often require accurate predictions and transparent and understandable models. NNAR's relative simplicity makes it more suitable for real-world applications where clarity and computational efficiency are as important as performance metrics.

Nevertheless, the findings should be contextualized within the study's limitations. For example, the models did not incorporate external variables such as mobility data, vaccination rates, or government intervention indices, which could potentially enhance predictive accuracy. Additionally, the performance of both models might vary in the presence of new viral variants or shifts in testing policies, factors not explicitly modeled in this analysis.

Future research should therefore explore the integration of exogenous variables and hybrid modeling strategies. Such efforts may offer a more holistic and adaptable forecasting framework capable of accommodating the dynamic nature of epidemic progression across different regions and stages.

X. CONCLUSION

This study sought to evaluate and compare the predictive performance of two artificial neural network models—Nonlinear Autoregressive Neural Network (NNAR) and Long Short-Term Memory (LSTM)—in forecasting daily confirmed COVID-19 cases in Saudi Arabia. Drawing on an extended time series dataset spanning from March 2020 to December 2021, the analysis revealed distinct differences in how each model responded to the dynamic and nonlinear nature of the pandemic.

The NNAR model demonstrated superior performance across all evaluation metrics, including MAE, RMSE, MAPE, and

© Copyright: All rights reserved.

Theil's U statistic. Its predictions were more stable and closely aligned with case trajectories, particularly in short-term forecasting scenarios. The LSTM model, despite its architectural complexity and theoretical strength in capturing long-term dependencies, tended to volatility and overfitting, especially when dealing with data segments characterized by noise or sudden variation.

These findings underscore the importance of selecting forecasting models that are appropriately matched to the data characteristics and practical application contexts. In settings where data volatility is moderate and the forecasting horizon is relatively short, simpler autoregressive models like NNAR may offer more reliable and interpretable outputs than their more complex deep learning counterparts.

Beyond model performance, this study also contributes to the ongoing conversation about the role of artificial intelligence in epidemic modeling. As health authorities continue to rely on data-driven tools for planning and intervention, it becomes increasingly important to ensure that models are accurate, transparent, adaptable, and easy to implement within existing public health infrastructure.

XI. RECOMMENDATIONS

Based on the findings of this research, several recommendations are proposed for both public health practitioners and data scientists:

A. For Public Health Authorities

i. Adopt NNAR-based forecasting models in real-time surveillance systems to support early intervention strategies and hospital resource allocation.
ii. Integrate model outputs into public health dashboards to improve situational awareness and facilitate data-driven decision-making.
iii. Continuously update and retrain models to reflect new data trends, emerging variants, and changes in public behavior or policy interventions.

B. For Data Scientists and Model Developers

i. Consider model simplicity and interpretability when deploying forecasting tools in high-stakes environments, particularly where explainability is valued.
ii. Experiment with hybrid architectures, combining the strengths of NNAR, LSTM, and other machine learning approaches, to achieve balanced performance under varying data conditions.
iii. Include exogenous variables (e.g., vaccination rates, mobility patterns, public holidays) to enhance model accuracy and relevance.

C. For Future Research

i. Conduct comparative studies across different countries or regions to assess the generalizability of model behavior in diverse epidemiological contexts.
ii. Explore fairness and bias considerations in AI-driven forecasting, especially in datasets with uneven coverage or varying quality across geographic or demographic segments.
iii. Encourage open-source sharing of models and datasets**,** promoting transparency, reproducibility, and collaborative improvement in epidemic forecasting.

These recommendations aim to bridge the gap between technical model development and real-world application, ensuring that forecasting tools effectively support public health preparedness and response efforts.

DECLARATION STATEMENT

After aggregating input from all authors, I must verify the accuracy of the following information as the article's author.

- **Conflicts of Interest/ Competing Interests:** Based on my understanding, this article has no conflicts of interest.
- **Funding Support:** This article has not been funded by any organizations or agencies. This independence ensures that the research is conducted with objectivity and without any external influence.
- **Ethical Approval and Consent to Participate:** The content of this article does not necessitate ethical approval or consent to participate with supporting documentation.
- **Data Access Statement and Material Availability:** The adequate resources of this article are publicly accessible.
- **Authors Contributions:** Alshaikh A. Shokeralla conceptualized the research idea, collected and prepared the data, and implemented the modeling framework. The author also contributed to statistical analysis, data visualization, and the drafting and revision of the manuscript. Finally, the author reviewed and approved the final version of the paper.

REFERENCES

1. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. DOI: https://doi.org/10.1162/neco.1997.9.8.1735
2. Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175. DOI: https://doi.org/10.1016/S0925-2312(01)00702-0
3. Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). OTexts. https://otexts.com/fpp2/
4. Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time Series Analysis: Forecasting and Control* (4th ed.). Wiley. DOI: https://doi.org/10.1002/9781118619193
5. Benvenuto, D. et al. (2020). Application of the ARIMA model on the COVID-19 epidemic dataset. *Data in Brief*, 29, 105340. DOI: https://doi.org/10.1016/j.dib.2020.105340
6. Liu, D. et al. (2021). Forecasting COVID-19 dynamics and interventions in the US using deep learning. *Nature Communications*, 12, 1–10. DOI: https://doi.org/10.1038/s41467-021-22457-w
7. Shastri, S. et al. (2020). Time series forecasting of COVID-19 using deep learning models. *Chaos, Solitons & Fractals*, 140, 110227. DOI: https://doi.org/10.1016/j.chaos.2020.110227
8. Ribeiro, M. H. D. M. et al. (2020). Short-term forecasting COVID-19 cumulative confirmed cases. *Chaos, Solitons & Fractals*, 135, 109853. DOI: https://doi.org/10.1016/j.chaos.2020.109853
9. Chimmula, V. K. R., & Zhang, L. (2020). Time series forecasting of COVID-19 in Canada using LSTM. *Chaos, Solitons & Fractals*, 135, 109864. DOI: https://doi.org/10.1016/j.chaos.2020.109864
10. WHO Coronavirus Dashboard. https://covid19.who.int
11. Saudi Ministry of Health COVID-19 Dashboard. https://www.emro.who.int/right-teasers/saudi-arabia/sau-statistics.html
12. Iqbal, M. W. et al. (2020). Prediction of COVID-19 using ARIMA and LSTM models. *SN Computer Science*, 2(5), 1–7. DOI: https://doi.org/10.3390/math10214001
13. Rustam, F. et al. (2020). COVID-19 future forecasting using machine learning models. *IEEE Access*, 8, 101489–101499. DOI: https://doi.org/10.1186/s12874-022-01755-x
14. Bandara, K. et al. (2020). Forecasting COVID-19 cases using time series models and ML. *BMC Medical Informatics and Decision Making*, 20, 341.

© Copyright: All rights reserved.

DOI: https://doi.org/10.1186/s12911-020-01395-3
15. Petropoulos, F., & Makridakis, S. (2020). Forecasting COVID-19. *PLOS ONE*, 15(3), e0231236. DOI: https://doi.org/10.1371/journal.pone.0231236
16. Chakraborty, T., & Ghosh, I. (2020). Real-time forecasts and risk assessment of COVID-19. *Chaos, Solitons & Fractals*, 135, 109850. DOI: https://doi.org/10.1016/j.chaos.2020.109850
17. Sagheer, A., & Kotb, M. (2019). Forecasting petroleum production with LSTM. *Neurocomputing*, 323, 203–213. DOI: https://doi.org/10.1016/j.neucom.2018.09.082
18. Salgotra, R. et al. (2020). Forecast of COVID-19 in India using genetic programming. *Chaos, Solitons & Fractals*, 138, 109945. DOI: https://doi.org/10.1016/j.chaos.2020.109945
19. Kim, S. et al. (2021). Deep learning for real-time COVID-19 forecasting. *Journal of Biomedical Informatics*, 117, 103773. DOI: https://doi.org/10.1016/j.neucom.2022.09.005
20. Khan, M. A. et al. (2022). Ensemble learning for COVID-19 prediction. *Computer Methods and Programs in Biomedicine*, 214, 106584. DOI: https://doi.org/10.3390/s22062224
21. Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*, 72(1), 37–45. DOI: https://doi.org/10.1080/00031305.2017.1380080
22. Al Zahrani, S., Al Sameeh, F. A. R., Musa, A. C. M., & Shokeralla, A. A. (2020). Forecasting diabetes patients attendance at Al-Baha hospitals using autoregressive fractional integrated moving average (ARFIMA) models. *Journal of Data Analysis and Information Processing*, 8, 183–194. DOI: https://doi.org/10.4236/jdaip.2020.83011
23. Wu, J. T. et al. (2020). Forecasting COVID-19 spread. *The Lancet*, 395(10225), 689–697. DOI: https://doi.org/10.1016/S0140-6736(20)30260-9
24. Wang, P. et al. (2021). Forecasting COVID-19 with Prophet model. *Scientific Reports*, 11, 20513. DOI: https://doi.org/10.1155/2022/1556025
25. Tuli, S. et al. (2020). Predicting COVID-19 growth using ML and cloud. *Scientific Reports*, 10, 21061. DOI: https://doi.org/10.1016/j.iot.2020.100222
26. Bezerra, A. K. L., & Santos, É. M. C. (2020). Prediction of COVID-19 in Sudan with ARIMA & Holt-Winters. *International Journal of Development Research*, 10(8), 39408–39413. https://www.journalijdr.com/sites/default/files/issue-pdf/19811.pdf
27. Hazarika, B. B. et al. (2022). Deep learning for COVID-19 forecasting: Review. *Computer Methods and Programs in Biomedicine*, 214, 106585. DOI: https://doi.org/10.1016/j.cmpb.2021.106585
28. Raza, S. E. A. et al. (2021). Comparative analysis of ARIMA and LSTM in Gulf. *J. King Saud Univ. Comput. Inf. Sci.* DOI: https://doi.org/10.1016/j.jksuci.2021.03.005
29. Shokeralla, A. A. et al. (2023). The effect of symptoms on survival time of COVID-19 patients in Sudan. *International Journal of Statistics in Medical Research*, 12, 249–256. DOI: https://doi.org/10.6000/1929-6029.2023.12.29
30. Daqqa, I. et al. (2024). Predictive modeling of breast cancer incidence. *Journal of Statistics Applications & Probability*, 14(2), 183–189. DOI: https://doi.org/10.18576/jsap/140204
31. Fath-Elrhman, E. I. et al. (2020). Modeling Sudan's inflation rate using neural networks. *International Journal of Engineering, Science and Mathematics*, 9(10), 1–11. https://www.indianjournals.com/ijor.aspx?target=ijor:ijesm&volume=9&issue=10&article=001&type=pdf
32. Ghosh, I., & Chakraborty, T. (2020). A deep learning-based framework to predict COVID-19 outbreak. *Neural Networks*, 132, 294–302. DOI: https://doi.org/10.1016/j.neunet.2020.09.004
33. Yang, Z., Zeng, Z., Wang, K., Wong, S.-S., Liang, W., Zanin, M., ... & He, J. (2020). Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *Journal of Thoracic Disease*, 12(3), 165–174. DOI: https://doi.org/10.21037/jtd.2020.02.64
34. Khosravi, A., Nahavandi, S., Creighton, D., & Atiya, A. F. (2011). Comprehensive review of neural network-based prediction intervals and new advances. *IEEE Transactions on Neural Networks*, 22(9), 1341–1356. DOI: https://doi.org/10.1109/TNN.2011.2162110
35. Gu, Q., Ding, Y., & Han, J. (2008). Learning representative graphical structures for time series modeling. *IEEE Transactions on Knowledge and Data Engineering*, 20(3), 376–392. DOI: https://doi.org/10.1109/TKDE.2007.190712
36. Ceylan, R. (2020). Estimation of COVID-19 prevalence in Italy, Spain, and France. *Chaos, Solitons & Fractals*, 138, 109924. DOI: https://doi.org/10.1016/j.chaos.2020.109924
37. Deb, S., & Majumdar, M. (2021). A time series method for modeling and forecasting the number of COVID-19 cases. *Heliyon*, 7(3), e06504. DOI: https://doi.org/10.1016/j.heliyon.2021.e06504
38. Liu, Q., Liu, Z., Li, D., & Gao, Q. (2020). Modeling and forecasting confirmed cases of COVID-19 in China using ARIMA model. *Journal of Infection and Public Health*, 13(6), 817–822. DOI: https://doi.org/10.1016/j.jiph.2020.04.003
39. Islam, A. R. M. T., Hasan, M. M., Rahman, M. M., & Rahman, M. A. (2021). Modeling COVID-19 outbreaks in the Middle East using hybrid neural networks. *Environmental Science and Pollution Research*, 28(25), 31869–31885. DOI: https://doi.org/10.1007/s11356-020-10827-8
40. Anaraki, A. K., & Shahraki, A. (2021). Forecasting COVID-19 in Iran using hybrid artificial intelligence models. *Soft Computing*, 25(20), 13071–13087.
41. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. https://www.deeplearningbook.org/
42. Li, R. et al. (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science*, 368(6490), 489–493. DOI: https://doi.org/10.1126/science.abb3221
43. Bandyopadhyay, S. et al. (2021). Forecasting COVID-19 using hybrid models. *Pattern Recognition Letters*, 145, 1–8. DOI: https://doi.org/10.1016/j.patrec.2020.11.008
44. Islam, A. R. M. T. et al. (2020). Time series analysis and forecasting of COVID-19 outbreak in Gulf countries. *International Journal of Environmental Health Research*, 31(6), 682–696. DOI: http://dx.doi.org/10.4018/IJEHMC.20220701.oa3
45. Adhikari, R., & Agrawal, R. K. (2013). A combination of artificial neural network and ARIMA models for time series forecasting. *Neural Computing and Applications*, 23(1), 1–11. DOI: http://dx.doi.org/10.1007/s00521-013-1386-y
46. R Rajkumar, M V Sudhamani, Content based Image Retrieval System using Combination of Color and Shape Features, and Siamese Neural Network. (2019). International Journal of Innovative Technology and Exploring Engineering, 9(2S), 71–77. DOI: https://doi.org/10.35940/ijitee.b1053.1292s19
47. Magapu, H., Krishna Sai, M. R., & Goteti, B. (2024). Human Deep Neural Networks with Artificial Intelligence and Mathematical Formulas. International Journal of Emerging Science and Engineering, 12(4), 1–2. DOI: https://doi.org/10.35940/ijese.c9803.12040324
48. Niharika, & Malhotra, S. (2020). Sentiment Analysis using Artificial Neural Network. International Journal of Recent Technology and Engineering (IJRTE), 8(5), 3267–3273. DOI: https://doi.org/10.35940/ijrte.e6450.018520

AUTHOR'S PROFILE

Dr. Alshaikh A. Shokeralla, Ph.D., Assistant Professor of Mathematics at Al-Baha University, Kingdom of Saudi Arabia, is a dedicated academic passionate about applied mathematical sciences. With a strong background in mathematical statistics, stochastic modeling, and simulation, he has contributed extensively to research in epidemiological forecasting, time-series analysis, and machine learning applications in health data. Dr. Shokeralla has published several peer-reviewed articles in reputable international journals and collaborates on interdisciplinary projects. He is committed to academic excellence, innovation in teaching, and mentoring future researchers.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of the Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)/ journal and/or the editor(s). The Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP) and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.