Journal of Economics and Administrative Sciences
2024; 30(144), pp. 501-516
P-ISSN 2518-5764
E-ISSN 2227-703X

Journal of Economics and Administrative Sciences (JEAS)

Available online at http://jeasiq.uobaghdad.edu.iq
DOI: https://doi.org/10.33095/wh488343

Building A hybrid Time Series Model Using ARDL With LSTM and GRU Models

Tabarak Yahya Abd*
Department of Statistics
College of Administration and Economics,
University of Baghdad, Iraq.
E-mail: tabarak.abd2101m@coadec.uobaghdad.edu.iq
Orcid: https://orcid.org/0000-0003-0615-0854
***Corresponding author**

Firas A. Mohammed Almohana
Department of Statistics
College of Administration and Economics,
University of Baghdad, Iraq.
E-mail: firasmohana@coadec.uobaghdad.edu.iq
Orcid: https://orcid.org/0000-0003-0615-0854

Received:7/2/2024 **Accepted:16/5/2024** **Published Online First:1/12/2024**

Abstract:
Purpose: The aim of the research is to utilize a hybrid model that combines the linear model represented by Autoregressive Distributed Lag (ARDL) and the nonlinear model represented by deep learning models, such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU).
Theoretical Framework: The theoretical framework integrates the linear component represented by the Autoregressive Distributed Lag (ARDL) model and the nonlinear component represented by deep learning models, namely Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) To create hybrid models.
Design/Methodology/Approach: The research methodology involves the use of EVIEWS 12 for analyzing standard data and integrating the ARDL model, as well as Python programming for building the proposed forecasting models. Weekly data from the Iraqi stock market, specifically from the banking and communications sectors spanning from 2017 to 2021, is utilized. The study compares the performance of the hybrid models ARDL_LSTM and ARDL_GRU with individual models using evaluation metrics such as root mean square error (RMSE) and mean absolute percentage error (MAPE).
Findings: The results indicate the superiority of the hybrid model ARDL_LSTM over other models due to its high accuracy and lower comparison measurement values.
Originality/Value: The originality of the research lies in its hybrid approach, combining ARDL with deep learning models like LSTM and GRU for time series forecasting. This approach adds value by addressing the limitations of individual models and improving forecasting accuracy in the context of the Iraqi stock market.
Keywords: Time Series Forecasting, Autoregressive Distributed Lag (ARDL), Deep Learning, Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Hybrid Models, Iraqi Stock Market, Banking Sector, Communications Sector.
JEL Classification: C22, C58, G1

Authors' individual contribution: Conceptualization—T.Y.A.; Methodology—T.Y.A.; Formal Analysis—T.Y.A.; Investigation—T.Y.A. and F.A.M.; Data Cu ration—T.Y.A.; Writing—Original Draft—T.Y.A.; Writing—T.Y.A.; Review &Edition—F.A.M.; Visualization—T.Y.A.; Supervision—F.A.M.; Project Administration—T.Y.A.

Declaration of conflicting interests: The Authors declare that there is no conflict of interest.

1.Introduction:

Time series is defined as a collection of data gathered sequentially over specific time intervals (Box et al., 2015). Mathematically, it is represented as a set of vectors ($x_t = x_1, x_2, x_3 \ldots\ldots, x_{t+1}$), where (t) represents the time variable and (x) is a variable generated randomly according to probability laws over time. Experimental studies indicate that non-linear models exhibit good performance in long-term forecasting, while linear models are suitable for short-term forecasting (Mousa & Mohammed, 2020). Rarely do real time series data consist solely of linear or non-linear components; rather, they often include both. Hence, the research problem lies in the fact that most time series consist of both linear and non-linear models, and therefore using a single model for prediction does not yield satisfactory results. Hence, a new technique is proposed, namely hybrid models, which are used in various fields to improve prediction performance or analysis. In this study, a hybrid model combining autoregressive distributed lag (ARDL) regression and deep learning models such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) was utilized to leverage the strengths of each model individually. This allows for dealing with both long-term and short-term time relationships, in addition to non-linear complexities.

2-Literature Review:

There are several studies that have addressed time series prediction, among them (Islam & Hossain, 2021), presented a study combining an LSTM model and a GRU model to form a hybrid LSTM_GRU network for predicting foreign exchange markets. The aim was to enhance the accuracy of currency price predictions. A comparison was made between the proposed hybrid model LSTM_GRU and the individual models LSTM and GRU using comparison metrics such as RMSE, MSE, and MAE. The results showed the superiority of the hybrid model over all individual models.(Mateus et al., 2021) . conducted a comparison between the Long Short-Term Memory (LSTM) model and the Gated Recurrent Unit (GRU) model for predicting energy time series. The results indicate that both LSTM and GRU models demonstrate effective predictive capabilities with slight differences in performance metrics. Overall, the GRU model shows better performance than the LSTM model.

(Omran et al., 2021) presented in it a comparison of deep learning models was presented, including Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), to predict the number of infections and deaths in Egypt, Kuwait, and Saudi Arabia from May 6, 2020, to December 6, 2020. They used evaluation metrics such as Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE) to assess the results. The study showed that the LSTM model performed better in predicting infections in the three countries, while the GRU model demonstrated superior performance in predicting deaths in Kuwait and Egypt.

(Nosier et al., 2022) conducted a study comparing prediction models using the Autoregressive Distributed Lag (ARDL) model and the Long Short-Term Memory (LSTM) deep learning model to forecast daily COVID-19 cases and deaths in Egypt from 2020 to 2021. The results showed that the LSTM model outperformed the ARDL model, exhibiting fewer prediction errors and thus better performance. (Zafar et al., 2022) focused on using a hybrid LSTM_GRU model to predict traffic speed in urban areas by integrating diverse data sources such as data from road sensors, geographic data, and online traffic information to improve the accuracy of traffic speed prediction. The results indicate that the proposed hybrid model yielded promising results, indicating its ability to predict speed with high accuracy.

P-ISSN 2518-5764
E-ISSN 2227-703X

(Ali et al., 2023) , Presented a study in which integrated time series models such as Auto Regressive integrated Moving Average (ARIMA), Long short-term memory (LSTM), and Gated Recurrent Unit (GRU).

It combined the linear component of the ARIMA model with the nonlinear components of the LSTM and GRU models to enhance the prediction accuracy. A comparison was made between individual models and hybrid models such as ARIMA_LSTM and ARIMA_GRU, demonstrating the effectiveness and accuracy of the hybrid models used. (Ramzi & Warda 2023) presented a study on the impact of cash supply on the general stock price index in the Qatar Stock Exchange during the period 2000-2020 using the ARDL model. The results indicate a cointegration relationship between cash supply and the general stock price index. The study also found a significant negative relationship between cash supply and the general stock index in both the long and short terms. Additionally, there is a non-significant positive relationship between cash supply and the general stock index in the long term, with a significant positive relationship in the short term. (Rehman et al., 2023) . They presented a hybrid approach to deep neural network models, LSTM_GRU, for Parkinson's disease detection. The results indicate the effectiveness of the proposed hybrid model in identifying individuals with Parkinson's disease based on clinical and diagnostic data. The abstract emphasizes the importance of using deep learning techniques to enhance the diagnosis of complex diseases like Parkinson's.

3. Methodology:

3.1 Unit Root tests:

Unit Root tests are fundamental assessments for determining the stationary or not stationary of a time series, as the effectiveness of time series modeling requires stationary data (Bakshi et al., 2021) . Despite the existence of various tests, the Augmented Dickey-Fuller test will be employed in this study.

3.2 Augmented Dickey-Fuller (ADF) test:

The ADF test checks whether a unit root is present in a time series. If a unit root exists, it means the series is non-stationary, and includes three states shown in the following equations (Nkoro & Uko, 2016)

1_The model without intercept constant and time trend

$$\Delta Y_t = \lambda Y_{t-1} + \sum_{j=1}^{k} \rho_j \Delta Y_{t-j} + \varepsilon_t \qquad (1)$$

2_The model with intercept constant

$$\Delta Y_t = \alpha + \lambda Y_{t-1} + \sum_{j=1}^{k} \rho_j \Delta Y_{t-j} + \varepsilon_t \qquad (2)$$

3_The model with intercept constant and time trend

$$\Delta Y_t = \alpha + \beta T + \lambda Y_{t-1} + \sum_{j=1}^{k} \rho_j \Delta Y_{t-j} + \varepsilon_t \qquad (3)$$

Where:

λ: (ρ-1)

α: Constant

k: The lag of maximum deceleration

β: The time trend coefficient

T: The time trend, represented by the formula

$$\mathrm{T} = \left(\mathrm{t} - 1 - \frac{1}{2}\mathrm{n}\right) \quad , \mathrm{t} = 2,3 \ldots , \mathrm{n} \qquad (4)$$

After choosing one of the three models, the test is conducted under the hypothesis (Hamad et al., 2021.):

H_0: $\lambda = 0$ (There is a unit root, which means that the series is non-stationary).

H_1: $\lambda < 0$ (There is no unit root, which means that the series stationary).

After estimating any of the previously mentioned equations, it becomes essential to compute the Augmented Dickey-Fuller (ADF) statistic (Farhan, 2019)

$$\tau * \lambda = \frac{\hat{\lambda}}{S_{\hat{\lambda}}} \quad (5)$$

Where:

$S_{\hat{\lambda}}$: The standard deviation of the estimated parameter

If the absolute value of the statistic exceeds the critical value from the table, and if the p-value is less than 0.05, then the null hypothesis is rejected, while the alternative hypothesis is Accept. This implies the absence of a unit root, indicating that the time series is stationary.

3.3 Autoregressive Distributed Lag Model (ARDL):

This model was originally described by Pesaran & Shin (1998) and Pesaran et al. (2001) ،Among the advantages of this model (ARDL) (Pahlavani et al., 2005) It is not required that all study variables be at the same level of stationary. Some variables may remain stationary at their original level (I (0)), while others will stationary after undergoing the first difference (I (1)). It can be applied effectively even with small sample sizes (Jalaee et al., 2019).

At the same time, it estimates the long-term and short-term effects within one equation instead of requiring two separate equations. The ARDL (p,$q_1, q_2, \dots, q_k$) model is represented by the following equation (Javangwe et al., 2022):

$$\Delta Y_t = C + \alpha_1 Y_{t-1} + \alpha_2 X1_{t-1} + \alpha_3 X2_{t-1} + \cdots \dots . \alpha_{k+1} XK_{t-1} + \sum_{i=1}^{p-1} \emptyset_{1i} \Delta Y_{t-i} + \sum_{i=0}^{q1\_1} \emptyset_{2i} \Delta X1_{t-i} + \sum_{i=0}^{q2-1} \emptyset_{3i} \Delta X2_{t-i} + \cdots \dots + \sum_{i=0}^{qk-1} \emptyset_{k+1i} \Delta XK_{t-i} + e_t \quad (6)$$

Where:

Δ : The first difference.

α : Long-term relationship parameters.

$\emptyset$: Short-term relationship parameters.

e_t : Random error limit, which is white noise $e_t \sim IID(0, \sigma^2)$.

C: Fixed limit.

p: The lag period for the dependent variable.

$q_1, q_2, \dots \dots q_k$: Lag period for independent variable.

The optimum number of lags is chosen according to the Akaike information Criterion (ALC), the Schwarz Bayesian Criterion (SBC), and Hannan-Quinn (HQ) (Majed,2013).

The long-term relationship will be estimated using the ARDL(p,$q_1, q_2, \dots, q_k$) model as follows (Abdelkader & Hamza, 2021.):

$$Y_t = C + \sum_{i=1}^{p} \emptyset_{1i} Y_{t-i} + \sum_{i=0}^{q1} \emptyset_{2i} X1_{t-i} + \sum_{i=0}^{q2} \emptyset_{3i} X2_{t-1} + \cdots \dots + \sum_{i=0}^{qk} \emptyset_{k+1,i} XK_{t-i} + \varepsilon_t + \quad (7)$$

After concluding the existence of a common integration relationship, it is possible to calculate the Error correction formula for the previously estimated equation (7), where the residuals are calculated from the selected appropriate model in the estimation process

$$EC_t = Y_t - C - \sum_{i=1}^{p} \emptyset_{1i} Y_{t-i} + \sum_{i=0}^{q1} \emptyset_{2i} X1_{t-i} + \sum_{i=0}^{q2} \emptyset_{3i} X2_{t-i} + \cdots + \sum_{i=0}^{qk} \emptyset_{k+1,i} XK_{t-i} + \varepsilon_t \quad (8)$$

After adding the calculated residuals (EC_{t-1}) to the previous equation, we obtain:

$$\Delta Y_t = C + \sum_{i=1}^{P} \emptyset_{1i} Y_{t-i} + \sum_{i=0}^{q1} \emptyset_{2i} X1_{t-i} + \sum_{i=0}^{q2} \emptyset_{3i} X2_{t-i} + \cdots + \sum_{i=0}^{qk} \emptyset_{k+1,i} XK_{t-i} + \lambda EC_{t-1} + U_t \quad (9)$$

P-ISSN 2518-5764
E-ISSN 2227-703X

3.4 Long short-term memory (LSTM):

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) used in deep learning. It was First introduced by the scientists Hochreiter & Schmidhuber in 1997 and is designed to address the vanishing gradient problem often faced by traditional RNN During training (Ahmadzadeh et al., 2022) . It has achieved significant success across various fields, especially in the prediction of time series (Elsworth & Uttel, 2020) . An LSTM cell consists of several components, including (Choi, 2018):

• Input Gate: The input gate manages the input information.
• Cell State: The cell state traverses the entire network and has the capability to incorporate or remove information with the assistance of gates.
• Forget Gate: This gate determines the fraction of information to retain or discard.
• Output Gate: The output gate produces the output generated by the LSTM.
• Sigmoid Layer: The sigmoid layer generates values between zero and one, indicating the extent to which each component should be allowed to pass through.
• Tanh Layer: The tanh layer generates a new vector that will be added to the current state.

The cell state is updated based on the outputs from these gates. Mathematically, this can be represented using the following equations (Selvin et al., 2017).

$$f_t = \sigma\left(w_f . [h_{t-1}, x_t] + b_f\right)$$
$$\mathrm{j}_t = \sigma(w_i . [h_{t-1}, x_t] + b_i)$$
$$\tilde{C}_t = \tanh(w_c . [h_{t-1}, x_t] + b_c)$$
$$c_t = \mathrm{f}_t * \left(c_{t-1} + i_c * \tilde{C}_t\right)$$
$$o_t = \sigma(w_0 . [h_{t-1}, x_t] + b_0)$$
$$h_t = o_t * \tanh(\mathrm{c}_t)$$

Where:

x_t : input vector
h_t: output vector
c_t: cell state vector
f_t: forget gate vector
j_t: input gate vector
o_t: output gate vector
W and b: are the parameter matrix and vector.

3.5 Gate Recurrent Unit (GRU):

Cho et al.(2014) (Chung et al., 2014) proposed a simplified version of the LSTM cell called Gated Recurrent Units (GRU), which requires less training time while improving network performance Functionally , GRU and LSTM operate similarly (Fu et al., 2016) , but the GRU cell utilizes a single hidden state that integrates the forget gate and input gate into a unified update gate , Additionally , The GRU unit consists of gating units that regulate the flow of information within the unit, without separate memory cells unlike LSTM. In contrast to LSTM, the GRU exposes the entire state at each time step and computes a linear sum between the current state and the newly computed state (Salehinejad et al., 2017). The Two gates of GRU include the update gate and the rest gate (Arunkumar et al., 2022). GRU formulation can be given by following equations(Shewalkar et al., 2019):

$$z_t = \sigma(W_{xz} X_t + W_{hz} h_{t-1} + b_z)$$
$$r_t = \sigma(W_{xr} x_t + W_{hr} h_{t-1} + b_r)$$
$$\grave{h}_t = \tanh(\mathrm{W}_{xh} x_t + W_{hh}(r_t \odot h_{t-1}) + b_h)$$
$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \grave{h}_t$$

r_t, z_t, x_t, h_t are the reset gate, update gate, input vector and output vector, respectively , W denotes the weight matrices, b the biases, sigm(σ) is the sigmoid activation, and tanh is the hyperbolic tangent activation function.

3.6 Hybrid Model:

The construction of hybrid models is based on the premise that the time series consists of two components, one being linear (L_t)and the other non-linear (N_t) over time (t), and is represented as follows:

$$y_t = L_t + N_t \quad (10)$$

L: For the linear component (ARDL model) in the time series,
N_t: For the nonlinear component (LSTM, GRU) in the time series.
To build the hybrid model, the residuals from the ARDL model will be used to construct LSTM and GRU models According to the following steps:
1. Building an ARDL model for the time series phenomenon observed at the study location, denoted as y.
2.Acquiring forecasted values $\hat{L_t}$ from the ARDL model, which represents the linear component of the time series.
3. Creating the non-linear component $\hat{N_t}$ of the Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks using the residuals obtained from the ARDL model and generating forecasts for the future.

$$e_t = f(e_{t-1}, e_{t-2}, \ldots\ldots, e_n) + \varepsilon_t$$

The estimated values are obtained by disregarding random errors ε_t :

$$\hat{e_t} = f(e_{t-1}, e_{t-2}, \ldots\ldots, e_n)$$

4. The forecasted values for the hybrid model are derived using the following equation.

$$\hat{y} = \hat{L_t} + \hat{N_t}$$

3.7 Comparison Criteria:

1. Root Mean Square Error (RMSE):

A standard method used to measure the standard deviation of a dataset, representing the square root of the squared differences between actual values and predicted values (Chicco et al., 2021) .

$$\text{RMSE}=\sqrt{\frac{\sum_{t=1}^{N}(y_t-\hat{y}_t)^2}{N}} \quad \ldots (11)$$

2. Mean Absolute Percentage Error (MAPE):

MAPE quantifies the percentage of absolute error concerning forecasted values and serves as an indicator of prediction precision. Lower MAPE values signify higher prediction accuracy (Wang et al., 2021).

$$\text{MAPE}=\frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i-\hat{y_i}}{y_i}\right| \ldots(12)$$

Where:
y_i:The actual value
$\hat{y}_i$:The predicted value
N: The number of the observation

4. The Application:

A dataset of stocks in the Iraqi Stock Market was applied over a five-year period, weekly, spanning from 2017 to 2021, comprising 261 observations for the banking sectors as dependent variable and the telecommunications sectors as an independent variable. Graphs and unit root tests were utilized to ensure the stationary of the time series. Additionally, a common integration test using the Autoregressive Distributed Lag (ARDL) model was conducted. LSTM and GRU models were employed for predictive purposes, relying on E-views 12 for standard data analysis and Python programming language for building the prediction models, including LSTM and GRU.

4.1 stationary Test:

To analysis the data, the time series stationary of the banking and telecommunications sectors is tested. This aims to clarify whether the time series requires appropriate transformations to ensure its stationary or the stationary of its variance before conducting any analysis. Figures 1 and 2 illustrate the time series of closing prices for the communication and banking sectors at the level.

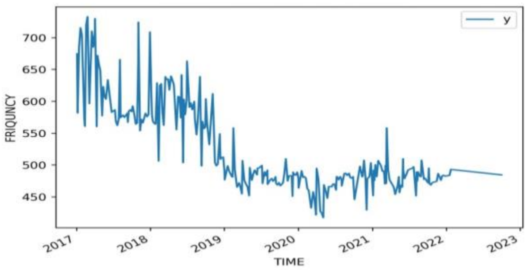

Figure 1: Displays the time series of the closing index for the telecommunications sector at the level.

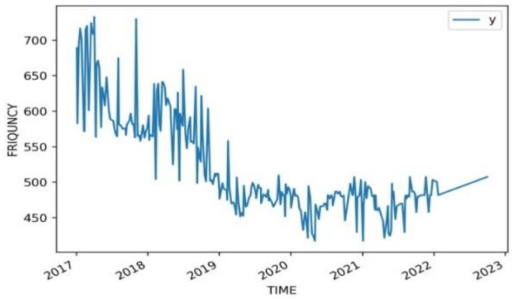

Figure 2: Displays the time series of the closing index for banking sector at the level.

Figures 1 and 2 demonstrate the non-stationary in time series at the level within the banking and telecommunications sectors. As a result, an expanded Dickey-Fuller test will be carried out to confirm the stationary of the time series.

Table 1: Augmented Dickey-fuller (ADF) test for unit Root for both sectors at the level

sectors	p-valu	The decision
Banking sectors	0.05 < 0.0952	Non-stationary
telecommunications sectors	0.05 < 0.7054	Non-stationary

The results of Table 1 indicate that the time series of the closing index in the banking and telecommunications sectors is unstable at the level where the probability value is greater than 0.05. This indicates acceptance of the null hypothesis of the existence of a unit root, which indicates the non-stationary of the time series. To achieve consistency in the mean, the first difference of the time series will be taken as show.

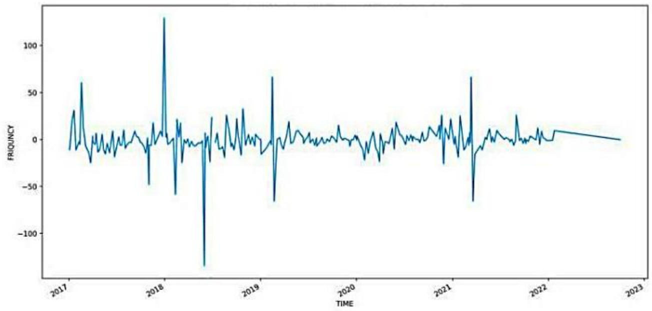

Figure 3: Time series of the closing index for the telecommunications sector at the first difference.

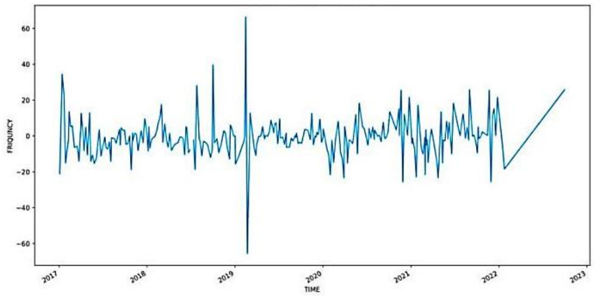

Figure 4: Time series of the closing index for the Banking sector at the first difference.

Figures 3 and 4 indicate that the time series for the banking and telecommunications sectors are stationary around their arithmetic mean after taking the first difference. An extended Dickey-Fuller test will be conducted to ensure the stationary of the series, as illustrated.

Table 2: Augmented Dickey-fuller (ADF) test for unit Root for both sectors at the first difference

sectors	p-valu	The decision
Banking sectors	0.05 >0.0000	stationary
telecommunications sectors	0.05 >0.0000	stationary

The results of Table 2 indicate the stationary of the time series for both sectors after taking the first difference, as the p-value is less than 0.05. This suggests accepting the alternative hypothesis of the absence of a unit root and the stationary of the time series.

4.2 Co-integration test using the Autoregressive Distributed Lag (ARDL) model:

After confirming the stationary of the time series at the first difference, cointegration will be tested using E-views 12. Prior to model estimation, it is necessary to specify the lag periods for the model variables, as illustrated in Table 3.

Table 3: Lags for the dependent and independent variables

Model Selection Criteria Table
Dependent Variable :BANK
Date:04/16/24 Time:17:23
Sample:1 260
Included Observations: 252

Model	Log L	AIC	BIC	HQ	Specification
3	-894.539911	7.202698	7.384771	7.275960	ARDL(8,2)
2	-893.809350	7.204836	7.400916	7.283734	ARDL(8,3)
1	-893.164402	7.207654	7.417739	7.292188	ARDL(8,4)
4	-896.236217	7.208224	7.376292	7.275851	ARDL(8,1)
33	-902.028430	7.214511	7.312551	7.253960	ARDL(2,2)
32	-901.104996	7.215119	7.327164	7.260204	ARDL(2,3)
31	-900.121537	7.215250	7.341301	7.265971	ARDL(2,4)
39	-904.455951	7.217904	7.287933	7.246082	ARDL(1,1)
26	-899.686476	7.219734	7.359791	7.276090	ARDL(3,4)
38	-903.837096	7.220929	7.304963	7.254743	ARDL(1,2)
27	-900.911696	7.221521	7.347572	7.272242	ARDL(3,3)
28	-901.971286	7.221994	7.334040	7.267079	ARDL(3,2)
34	-904.095719	7.222982	7.307016	7.256795	ARDL(2,1)
37	-903.380553	7.225242	7.323282	7.264692	ARDL(1,3)
21	-899.588740	7.226895	7.380957	7.288886	ARDL(4,4)
23	-901.711039	7.227865	7.353916	7.278586	ARDL(4,2)
22	-900.727078	7.227993	7.368049	7.284349	ARDL(4,3)
36	-902.732323	7.228034	7.340080	7.273119	ARDL(1,4)
29	-903.957192	7.229819	7.327859	7.269268	ARDL(3,1)
6	-897.218050	7.231889	7.427969	7.310788	ARDL(7,4)
7	-898.253406	7.232170	7.414244	7.305433	ARDL(7,3)
8	-899.408397	7.233400	7.401468	7.301027	ARDL(7,2)
16	-899.487907	7.234031	7.402099	7.301658	ARDL(5,4)
18	-901.506052	7.234175	7.374232	7.290531	ARDL(5,2)
17	-900.574466	7.234718	7.388780	7.296709	ARDL(5,3)
24	-903.622088	7.235096	7.347141	7.280181	ARDL(4,1)
13	-900.693517	7.235663	7.389725	7.297654	ARDL(6,2)
12	-899.799518	7.236504	7.404572	7.304131	ARDL(6,3)
11	-898.821038	7.236675	7.418749	7.309938	ARDL(6,4)
9	-901.287853	7.240380	7.394442	7.302371	ARDL(7,1)
14	-902.305976	7.240524	7.380580	7.296880	ARDL(6,1)
19	-903.365282	7.240994	7.367045	7.291715	ARDL(5,1)
5	-908.687471	7.299107	7.453169	7.361098	ARDL(8,0)
40	-916.779470	7.307774	7.363796	7.330316	ARDL(1,0)
35	-916.665793	7.314808	7.384836	7.342986	ARDL(2,0)
15	-913.198894	7.319039	7.445090	7.369759	ARDL(6,0)
30	-916.237867	7.319348	7.404482	7.353162	ARDL(3,0)
10	-912.389734	7.320553	7.460610	7.376909	ARDL(7,0)
25	-915.804631	7.323846	7.421886	7.363295	ARDL(4,0)
20	-915.034227	7.325668	7.437714	7.370753	ARDL(5,0)

The results in Table 3 indicate the automated determination of lags by E-views 12, Based on the minimum value of the criteria (AIC,BIC,H-Q). The optimal ARDL model was identified as ARDL (8,2), with 12 lags defined as the maximum of the variables.

Table 4: Statistic indicators

R-Squared	0.980274
F-statistic	989.7427
Prob(F-statistic)	0.000000
Akaike info criterion (AIC)	7.202698
Bayesion information criterion (BIC)	7.384771
Hannan -Quinn criter (H.Q)	7.279560

The results in Table 4 reveal that the R-Squared value reached 98%, indicating a high explanatory power for the model. Additionally, the model is statistically significant, as evidenced by the Prob (F statistic)=0.000000 , which is less than the significance level of 5% .

4.3 Boundary test:

The bounds integration test will be performed at 1%, 5% and 10% significance level to evaluate the suitability of the bounds for integration, as shown in Tables 5 and 6.

Table 5: F_BOUND TEST

Test Statistic	Value	Sig	I(0)
F-statistic	12.73795	10%	5.59
		5%	6.56
		1%	8.74

Table 6: T_BOUND TEST

Test Statistic	Value	Sig	I (0)
T-statistic	-4.915195	10%	-3.13
		5%	-3.41
		1%	-3.96

Table 5 indicates the significance of the F test with a value of 12.73795, which is greater than the maximum limit of integration (8.74). This means rejecting the null hypothesis and accepting the alternative hypothesis of the existence of joint integration between the banking and telecommunications sectors. Likewise, in Table 6, the value of t-statistic is (-4.915195), which is less than the critical value of integration at (-3.96), which indicates acceptance of the alternative hypothesis of cointegration between variables

4.4 Long-run coefficients estimation:

After confirming the existence of a long-term relationship between the variables, long-term futures coefficients will be estimated. However, before initiating this estimation process, some diagnostic tests must be conducted to ensure the model's quality, as outlined in the table 7

Table 7: Model quality test

Test	F-statistic	Obs*R-squared	Prob.F	Prob.Chi-Squer
Heteroskedasticity Test: Breusch-Pagan-Godfrey	1.121595	4.826951	0.2556	0.2498
Breusch-Godfrey serial Correlation LM Test	1.960254	4.826951	0.1729	0.0985
Heteroskedasticity Test: ARCH	2.073067	8.1827586	0.0850	0.0851

The results of the Breusch-Pagan-Godfrey test and the ARCH test, as shown in Table 7, indicate no heteroscedasticity issue. This is because the Chi-squared and F-statistic values for each test are greater than the critical values at a 5% significance level, suggesting the acceptance of the null hypothesis of no heteroscedasticity. As for the Breusch-Godfrey serial correlation LM test, it suggests the absence of serial correlation, as the Chi-squared and F-statistic values are greater than the critical values at a 5% significance level, indicating the acceptance of the null hypothesis that residuals do not suffer from serial correlation. Also, the CUMSUM and CUMSUM SQ tests to ensure the structural stationary of the model, as shown in the figure 5.

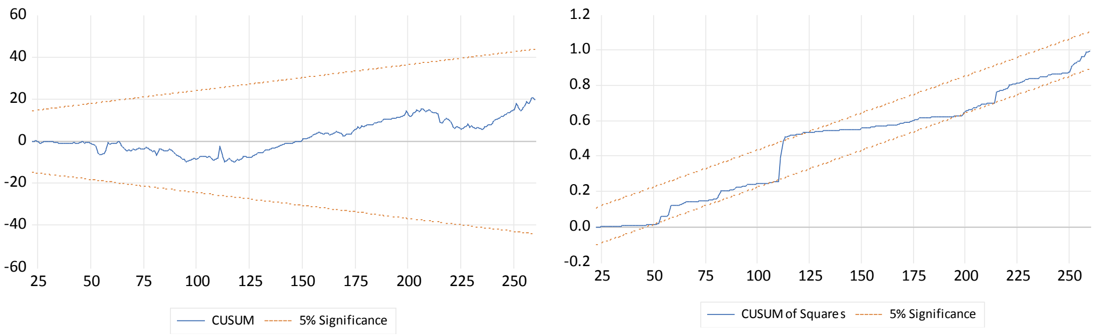

Figure 5: CUSUM and CUSUM Square

From Figure (5), it is evident that the blue line falls between the boundaries at a 5% significance level, indicating the stationary of the model. Therefore, it allows for estimating both long-term and short-term coefficients, As a shown in table (8)

Table 8: Result of estimating model parameters in the long run

Variable	Coefficient	Std .Error	t-statistic	Prob.
TEL	0.856763	0.081477	10.51537	0.0000
EC=BANK-(0.8568*TEL)				

Form Table 8 is it evident that the communications sector has a positive and statistically significant impact (prob=0.0000) on the banking sector. A 1% change in the communications sector leads to on 85% change in the banking sector.

4.5 Error correction model estimation:

After confirming the presence of long-term cointegration, we can now deduce an error correction model, as illustrated in the following table:

Table 9: Error correction

Variable	Coefficient	Std .Error	t-statistic	Prob
C	18.52242	4.254883	4.353214	0.0000
TREND @	-0.020998	0.010413	-2.016553	0.0449
D(BANK(-1))	-0.094632	0.063558	-1.488905	0.1378
D(BANK(-2))	0.029057	0.054963	0.528663	0.5975
D(BANK(-3))	0.0342562	0.054904	0.624026	0.5332
D(BANK(-4))	0.037470	0.0555387	0.676508	0.4994
D(BANK(-5))	0.100113	0.055478	1.804549	0.0724
D(BANK(-6))	-0.058962	0.054778	-1.076391	0.2828
D(BANK(-7))	0.145526	0.053637	2.713171	0.0071
D(TEL)	0.344856	0.037189	9.272962	0.0000
D(TEL(-1))	0.068281	0.043833	1.557777	0.1206
Coint Eq(-1)*	-0.228456	0.045168	-5.057915	0.0000

The error correction term coefficient was statistically significant (prob = 0.0000) and negative, indicating the presence of a short-term to long-term equilibrium relationship.

4.5 Long Short-Term Memory (LSTM):

Python was used to build the LSTM model, Using the library (Tensor flow) and (Keras). The model consisted of four hidden layers, one input layer, and one output layer. Using the activation function (Relu) in the hidden layers and using the linear activation function (Linear Activation function) in the output layer. Through several attempts, the model parameters were obtained as (Optimizer (Adam), epoch=500, Learning Rate=0.01). figure 6 illustrates the extent of convergence between the actual data and the predicted data.

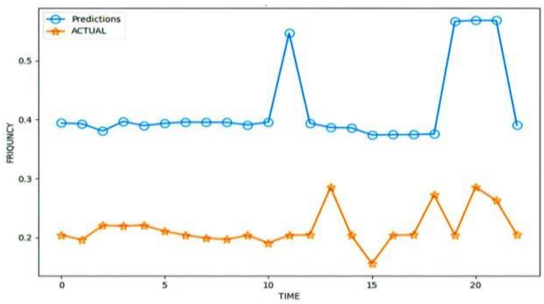

Figure 6: The actual values and predicted values for the LSTM model

4.6 Gated Recurrent Unit (GRU):

Using python and the library (Tensor flow) and (Keras), a GRU model was built with 6 hidden layers, an input layer, and an output layer, using the activation function (Sigmoid Activation Function) . After several attempts, the parameters (Optimizer (Adam), epoch=500, Learning Rate=0.01) were obtained, and figure 7 illustrates the extent of convergence between the actual and predicted value.

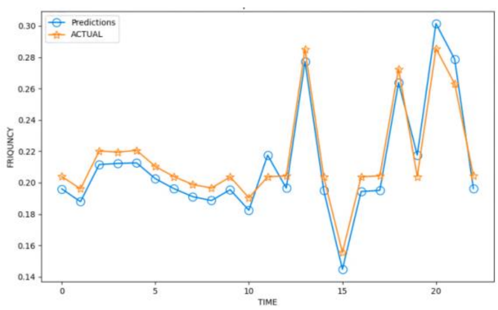

Figure 7: The actual values and predicted values for the GRU model

4.7 Prediction using the hybrid model (ARDL_LSTM):

The LSTM model is build based on the remainders of the ARDL model, and by combining the prediction of the two models, the hybrid model ARDL-LSTM is built, Figure 8 illustrates actual values alongside predicted values.

 P-ISSN 2518-5764
E-ISSN 2227-703X

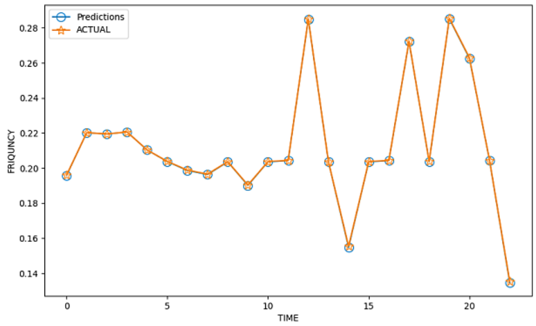

Figure 8: The actual values and predicted values for the ARDL-LSTM model

4.8 Prediction using the hybrid model (ARDL_GRU):

The closing index for the banking sector will be predicted using the hybrid model ARDL_GRU, which was constructed by relying on the residuals of the ARDL model to build the GRU model. The predictions from both models were collected, and Figure 9 illustrates the actual and predicted values.

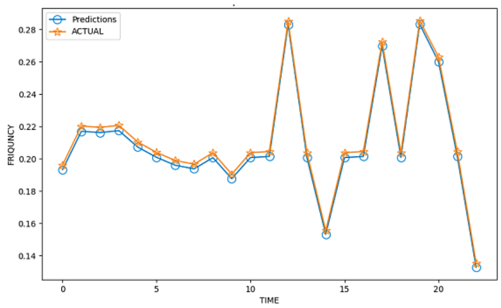

Figure 9: The actual values and predicted values for the ARDL-GRU model

5. Result Evaluation:

At this stage, the models will be evaluated, and the best model will be selected based on the evaluation metrics (RMSE, MAPE), where the best model is considered to have the lowest values for both of these metrics. Table 8 illustrates the comparison metric results for the predictive models.

Table 10: Comparison metrics for predictive models

Model	MAPE	RMSE
ARDL	1.3622	10.72307
LSTM	0.965760	0.044344
GRU	0.044130	0.002039
ARDL-LSTM	0.001257	0.000065
ARDL-GRU	0.014067	0.000609

The table above (10) indicates that the best model is the ARDL_LSTM, as it possesses the lowest values for the comparison metrics compared to the other used models, according to the time series data used.

6. Discussion:

1-The research has found that the time series under study has become stationary at first differences, meaning that we can apply the Auto Regressive Distributed Lag (ARDL) model for cointegration.
2-The cointegration test was applied using the Auto Regressive Distributed Lag (ARDL) approach on the financial securities dataset of the banking and communication sectors. It was found that the ARDL (8,2) model is the best choice due to its lower evaluation criteria values.
3-Python programming language was utilized to build proposed prediction models using LSTM and GRU architectures. After several attempts and experiments, the models were constructed with three hidden layers in addition to input and output layers. The Sigmoid activation function was employed for all layers except for the LSTM model, where the ReLU activation function was applied in the hidden layer, and the Linear activation function in the output layer.
4-The hybrid models (ARDL-LSTM and ARDL-GRU) were built by combining the estimated values from the ARDL (8,2) model with its residuals estimated from the LSTM and GRU models
5-The hybrid model ARDL-LSTM outperformed others due to having lower values of comparison metrics such as MAPE and RMSE, followed by the other hybrid model ARDL-GRU.

7. Conclusion:

The most important conclusions reached through the results in the application are:
1-The research results indicate that the studied phenomenon is a nonstationary time series based on the available original data, hence the application of first differencing to stationary the series.
2-According to the cointegration analysis using the bounds approach, the study confirmed the existence of a long-term equilibrium relationship between the dependent variable (banking sector) and the independent variable (telecommunications sector).
3-The model ARDL (8,2) is considered the best, as it possesses the lowest for evaluation criteria
4-The research indicates the superiority of the ARDL-LSTM model, as it possesses lower values for comparison criteria such as RMSE and MAE.

Authors Declaration:

Conflicts of Interest: None
-We Hereby Confirm That All The Figures and Tables In The Manuscript Are Mine and Ours. Besides, The Figures and Images, which are Not Mine, Have Been Permitted Republication and Attached to The Manuscript.
- Ethical Clearance: The Research Was Approved by The Local Ethical Committee in The University.

Reference:

Abdelkader, S., & Hamza, T. (2021). Comparison of ARDL And Artificial Neural Networks Models for Foreign Direct Investment Prediction in Algeria. Journal of Finance, Investment and Sustainable Development 6(2),388-400.

Ahmadzadeh, E., Kim, H., Jeong, O., Kim, N., & Moon, I. (2022). A deep bidirectional LSTM-GRU network model for automated ciphertext classification. IEEE Access,10,3228-3237 https://ieeexplore.ieee.org/abstract/document/9668927

Ali, N., S., M., & Mohammed, F., A. (2023). The use of ARIMA, LSTM and GRU models in time series hybridization with practical application .International Journal Of Nonlinear Analysis and Applications, 14(1), 2008–6822. https://doi.org/10.22075/ijnaa.2022.7110

Abdelkader, S., & Hamza, T. (2021). Comparison of ARDL And Artificial Neural Networks Models for Foreign Direct Investment Prediction in Algeria. Journal of Finance, Investment and Sustainable Development 6(2),388-400.

Ahmadzadeh, E., Kim, H., Jeong, O., Kim, N., & Moon, I. (2022). A deep bidirectional LSTM-GRU network model for automated ciphertext classification. IEEE Access,10,3228-3237 https://ieeexplore.ieee.org/abstract/document/9668927

Ali, N., S., M., & Mohammed, F., A. (2023). The use of ARIMA, LSTM and GRU models in time series hybridization with practical application .International Journal Of Nonlinear Analysis and Applications, 14(1), 2008–6822. https://doi.org/10.22075/ijnaa.2022.7110

Arunkumar, K. E., Kalaga, D. V, Mohan, C., Kumar, S., Kawaji, M., & Brenza, T. M. (2022). Comparative analysis of Gated Recurrent Units (GRU), long Short-Term memory (LSTM) cells, autoregressive Integrated moving average (ARIMA), seasonal autoregressive Integrated moving average (SARIMA) for forecasting COVID-19 trends. Alexandria Engineering Journal, 61, 7585–7603. https://doi.org/10.1016/j.aej.2022.01.011

Bakshi, S. S., Jaiswal, R. K., & Jaiswal, R. (2021). Efficiency Check Using Cointegration and Machine Learning Approach: Crude Oil Futures Markets. Procedia Computer Science, 191, 304–311. https://doi.org/10.1016/j.procs.2021.07.038

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time series analysis: forecasting and control. John Wiley & Sons.

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Computer Science, 7, 1–24. https://doi.org/10.7717/PEERJ-CS.623

Choi, C. (2018). Time Series forecasting with Recurrent Neural Network in presence of missing Data. Munin uit No https://munin.uit.no/handle/10037/14887

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. ArXiv Preprint ArXiv:1412.3555.

Elsworth, S., & Uttel, S. G. (2020). Time Series Forecasting Using LSTM Networks: A Symbolic Approach. https://github.com/nla-group/ABBA-LSTM.

Farhan, A. H. (2019). Using dickey _ fuller expanded test for testing variables of investment function in Iraq. Journal of Economics and Administrative Sciences, 25(114), 1–19.

Fu, R., Zhang, Z., & Li, L. (2016). Using LSTM and GRU neural network methods for traffic flow prediction. In 2016 31st Youth academic annual conference of Chinese association of automation (YAC) (pp. 324-328). IEEE.

Hamad,M.J , Mhmood, M.M. and Hassan ,F.F.(2021)" Measuring and analyzing the impact of external debt on the Gross Domestic product in Morocco for the period 1990-2017 using the ARDL model" Journal of Economics and Administrative sciences, 27(125) , pp.462-476, https://www.iasj.net/iasj/download/e7ac60c7d22e9607

Islam, M. S., & Hossain, E. (2021). Foreign exchange currency rate prediction using a GRU-LSTM hybrid network. Soft Computing Letters, 3, 100009. https://doi.org/10.1016/j.socl.2020.100009

Jalaee, S. A., Lashkary, M., & GhasemiNejad, A. (2019). The Phillips curve in Iran: econometric versus artificial neural networks. In Heliyon (Vol. 5, Issue 8). Elsevier Ltd. https://doi.org/10.1016/j.heliyon.2019.e02344

Javangwe, K. Z., & Takawira, O. (2022). Exchange rate movement and stock market performance: An application of the ARDL model. Cogent Economics and Finance, 10(1). https://doi.org/10.1080/23322039.2022.2075520

Majed,H.H.(2013)"Using time series methods to address seasonal variations in the consumer price" Journal of Economics and Administrative sciences Vol.19 , No.74,pp.360-380, https://www.jeasiq.uobaghdad.edu.iq/index.php/JEASIQ/article/view/1439

Mateus, B. C., Mendes, M., Farinha, J. T., Assis, R., & Cardoso, A. M. (2021). Comparing LSTM and GRU models to predict the condition of a pulp paper press. Energies, 14(21). https://doi.org/10.3390/en14216958

Mousa, M. A., & Mohammed, F. A. (2020). Applying some hybrid models for modeling bivariate time series assuming different distributions for random error with a practical application. Journal of Economics and Administrative Sciences, 26(117). https://www.iasj.net/iasj/article/177097

Nkoro, E., & Uko, A. K. (2016). Autoregressive Distributed Lag (ARDL) cointegration technique: application and interpretation. Journal of Statistical and Econometric Methods, 5(4), 63–91.

Nosier, S., El-Shobaky, S., & Salah, R. (2022). Comparing Econometrics Approach Vs. Deep Learning Approach in Forecasting Covid-19 Infections and Deaths Horizon in Egypt. Scientific journal of the Faculty of Economic Studies & Political Science, https://doi.org/10.21608/ESALEXU.2022.247220

Omran, N. F., Abd-El Ghany, S. F., Saleh, H., Ali, A. A., Gumaei, A., & Al-Rakhami, M. (2021). Applying Deep Learning Methods on Time-Series Data for Forecasting COVID-19 in Egypt, Kuwait, and Saudi Arabia. Complexity, 2021. https://doi.org/10.1155/2021/6686745

Pahlavani, M., Wilson, E., & Worthington, A. C. (2005). Trade-GDP Nexus in Iran: An Application of the Autoregression Distributed Lag (ARDL) Model. American journal of Applied Sciences 2(7), 1158–1165. https://ro.uow,edu.au/commpapers/144

Ramzi ,B.&Warda,A.(2023)"The Impact Of Money Supply on the general Stock Price Index on the Qatar Stock Exchange – on econometric study using the ARDL model for the period (2000-2020)".Journal of Finance , Investment and sustainable Development , 8(1),150-170 , https://www.asjp.cerist.dz/index.php/en/article/224252

Rehman, A., Saba, T., Mujahid, M., Alamri, F. S., & El Hakim, N. (2023). Parkinson's Disease Detection Using Hybrid LSTM-GRU Deep Learning Model. Electronics (Switzerland), 12(13). https://doi.org/10.3390/electronics12132856

Salehinejad, H., Sankar, S., Barfett, J., Colak, E., & Valaee, S. (2017). Recent Advances in Recurrent Neural Networks. http://arxiv.org/abs/1801.01078

Selvin, S., Vinayakumar, R., Gopalakrishnan, E. A., Menon, V. K., & Soman, K. P. (2017). Stock price prediction using LSTM, RNN and CNN-sliding window model. 2017 International Conference on Advances in Computing, Communications and Informatics, ICACCI 2017, 2017-January, 1643–1647. https://doi.org/10.1109/ICACCI.2017.8126078

Shewalkar, A., nyavanandi, D., & Ludwig, S. A. (2019). Performance Evaluation of Deep neural networks Applied to Speech Recognition: RNN, LSTM and GRU. Journal of Artificial Intelligence and Soft Computing Research, 9(4), 235–245. https://doi.org/10.2478/jaiscr-2019-0006

Wang, Y., Xu, C., Ren, J., Li, Y., Wu, W., & Yao, S. (2021). Use of meteorological parameters for forecasting scarlet fever morbidity in Tianjin, Northern China. Environmental Science and Pollution Research, 28(6), 7281–7294. https://doi.org/10.1007/s11356-020-11072-9

Zafar, N., Haq, I. U., Chughtai, J. U. R., & Shafiq, O. (2022). Applying Hybrid LSTM-GRU Model Based on Heterogeneous Data Sources for Traffic Speed Prediction in Urban Areas. Sensors, 22(9). https://doi.org/10.3390/s22093348