IRAQI STATISTICIANS JOURNAL
https://isj.edu.iq/index.php/isj
ISSN: 3007-1658 (Online)

Forecasting Public Expenditures using ARDL–LSTM Hybrid Model

Sulaiman Hussien Jawad[1] Munaf Yousif. Hmood[2]

[1,2]Department of Statistics, College of Administration and Economics, University of Baghdad Iraq, Baghdad.

ARTICLE INFO

Article history:
Received 4 March 2026
Revised 4 March 2026
Accepted 31 March 2026
Available online 3 April 2026

Keywords:
ARDL,
LSTM,
ARDL-LSTM,
Public Revenues,
Public Expenditures.

ABSTRACT

Public expenditure is regarded as the main tool for regulating fiscal policy and, consequently, affects the overall economic position of the country. This highlights the importance of high-quality forecasting, particularly when the objective is fiscal sustainability and the efficient allocation of resources. As economic behavior becomes more intricate and the relationships among variables increasingly intertwined, linear models are no longer sufficient for accurate prediction

With the assumption that public expenditure is conditioned by public revenues, this work will endeavor to predict government expenditure through the application of a hybrid ARDL-LSTM model. The method combines Autoregressive Distributed Lag (ARDL) model to explain linear variations and fluctuations of the time series and a neural network model-Long Short-Term Memory (LSTM) to explain nonlinear trends. Based on the comparison criteria. The findings indicated that the hybrid ARDL-LSTM model is more effective in forecasting accuracy, based on the RMSE and MAPE criterion using both the estimation and testing samples, which confirms that the combination of linear and nonlinear model is more efficient in predicting economic time series.

1. Introduction

Government expenditure constitutes a fundamental pillar of fiscal policy and plays a crucial role in influencing economic activity within a country. Through public spending, governments finance development initiatives, deliver essential public services, and promote both economic and social stability. Consequently, obtaining reliable estimates and accurate forecasts of government expenditure is essential for designing an effective public budget that supports fiscal policy objectives and adequately responds to societal needs. Public expenditure is a dynamic financial time series since it directly depends on the amount of public revenues, and since oil prices have a severe impact on the level of the public revenue in Iraq. Oil prices are volatile, nonlinear and complicated by nature and driven by multiple factors including

Corresponding author E-mail address: suliaman.hussien83@gmail.com

https://doi.org/10.62933/xr5r0296

political events, economic conditions, financial market movements and investor expectations [1].

Traditional methodologies are used in forecasting time series, and these methodologies are represented by a variety of econometric (economic-statistical) models. These models have assumptions that must be verified, such as the normal distribution of residuals and the assumption of stationarity [2]. Nonetheless, these models have been used to predict different economic and financial series in the last few years, and the use of these models has been justified by the data available and its features, namely, its type, size, and short time period [3]. As technology advances and instruments to be used in data collection and processing are invented, economic data has grown in size, accuracy and details. Nevertheless, precise data and its magnitude presents a problem to conventional econometric models, including the Autoregressive Distributed Lag Model (ARDL).

ARDL methodology is regarded as one of the most outstanding econometric frameworks applied in co-integration testing, prediction and the dynamic relationship between variables. Pesaran et al. [4] developed it in 2001 but it was intended to overcome some of the weaknesses with traditional co-integration tests. These shortcomings are the need to have the same integration order of variables, the fact that such tests are not efficient at small samples, and failure to capture the short-run effect adequately.

The advantage of this methodology is that it is efficient with small sample sizes $(T < 50)$, it will be able to observe both the short-run and long-run relationships are available between the variables whether they are of the same order or mixed orders $I(0)$, $I(1)$, as long as none of the variables are of the second order $I(2)$ [5]. In addition, it provides flexibility in identifying the most effective lag structure of each variable in the model to make a more insightful understanding of the dynamic effect of independent variables on the dependent variable.

The ARDL approach however, presupposes linear relationship between the variables, which is not always the case particularly in the analysis of economic relationships. In that regard, ARDL may be regarded as a successful linear model of the description of a linear nature of the relationship between public revenues as an independent variable and public expenditure as a dependent variable. However, when the correlation is intricate, interactive, or nonlinear, it cannot offer highly precise forecasting or appropriate modeling, especially considering the complex nature of behavior of these variables in the economic environment.

Conversely, the application of Machine Learning (ML) has revolutionized macroeconomic forecasting to a large extent because the methods can address some of the flaws of conventional econometric models. ML techniques are able to discover nonlinear relationships and more complex patterns and relationships among variables, particularly in unstable economic environments, and do not necessitate the strong statistical conditions of the traditional econometric models [6] [7].

Moreover, theoretical and empirical evidence have shown that integrating the models of traditional economics with the models of machine learning an effective approach that can enhance the accuracy of predictions [8] [9].

To enhance the predictive accuracy of the public expenditure time series, this paper aims to combine the Autoregressive Distributed Lag (ARDL) model, applied to identify both linear and dynamic short-run and long-run effects, with one of the neural network models- Long Short-Term Memory (LSTM) to identify nonlinear features.

2. Long Short-Term Memory-LSTM:

Long short-term memory LSTM model is an extension of Recurrent Neural Networks (RNN), which was first proposed by Hochreiter and Schmidhuber in 1997 [10], and was later improved by Gers and Schraudolph in 2002 [11], and further refined by Greff et al. in 2016 [12].

The LSTM model was created to solve the vanishing gradient problem, which is caused by neural network weights variation as training progresses, and hence, makes traditional Recurrent Neural Networks (RNNs) unsuitable to learn long and sequential temporal dependencies in data. Contrarily, LSTM networks have the capacity to handle sequential data and they are able to remember information at a previous step along the sequence thus making them quite effective in predicting the next steps. This characteristic renders LSTM especially ideal in modeling and extracting long-term time information [13].

The (LSTM) model is characterized by having a more complex architectural structure compared with traditional neural networks. As shown in Figure (1), every LSTM unit in the model includes a Memory Cell, which is represented by the symbol (C_t) , as well as three types of gates, the forget gate(f_t), the input gate(i_t), and the output gate(o_t). All these gates work together to regulate the flow of information both in and out of the unit. This process allows to manage the long-term dependencies in the time series through selective information updating or information retention over time [14].

The mathematical formulas for the gates and the LSTM structure can be expressed as follows [15]:

$$f_t = \sigma(W_{fx} xt + W_{fh} h_{t-1} + b_f) \qquad (1)$$

$$i_t = \sigma(W_{ix} xt + W_{ih} h_{t-1} + b_i) \qquad (2)$$

$$o_t = \sigma(W_{ox} xt + W_{oh} h_{t-1} + b_o) \qquad (3)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \widetilde{C}_t \qquad (4)$$

$$\tilde{C}_t = \tanh(W_{cx} xt + W_{ch} h_{t-1} + b_c \qquad (5)$$

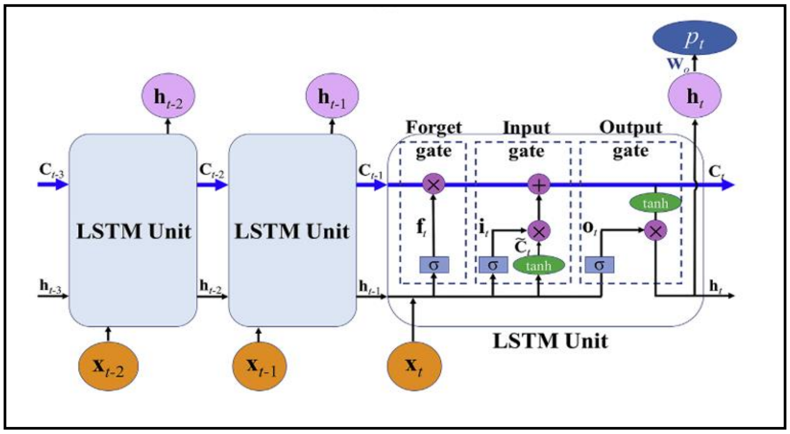
$$h_t = o_t \odot tanh\,(C_t) \qquad (6)$$

The (f_t, i_t, o_t) represents the forgetting gate, the input gate and the output gate respectively, $\odot$ refers Point-wise multiplication of two vectors, and the symbol (σ) represents the sigmoid activation function and is defined as follows:

$$\sigma(x)=1/1+e^{-x} \qquad (7)$$

$(W_{fh}, W_{ih}, W_{oh}, W_{ch})$ Represent the correlated weight matrices for the three gates and the cell state, respectively; $(W_{fx}, W_{ix}, W_{ox}, W_{cx})$ represent the input weight matrices for the three gates and the cell state, respectively; (b_f, b_i, b_o, b_c) are bias coefficients associated with each gate and the cell state and $(C_t, C_{t-1}, \widetilde{C}_t)$ represent the current cell state, the previous cell state, and the candidate cell state respectively.

The forget gate defines how much historical information to forget of the previous cell stateC_{t-1}, the input gate defines how much it should extract out of the current candidate cell state $\widetilde{C}_t$ and pass it to the current cell state C_t , and the output gate defines how much the cell outputs should be passed to the rest of the network.



3. Proposed model ARDL-LSTM

The ARDL technique can be regarded as an appropriate econometric model according to which the relationship between the independent variable (public expenditure, EXP) and the dependent variable (public revenues, REV) can be estimated, and, can also be employed to predict the dependent variable. It is a relatively flexible methodology in the sense that it has the capacity to incorporate lagged effects that affect the dependent variable, thereby making possible the capturing of both the past expenditure effects and the contribution of revenues in the current period (direct effect) and the immediate past period (lagged effect) to expenditure in the long term. ARDL model may be expressed as follows [17]:

$$Y_t = a_o + a_1 t + \sum_{i=1}^{p} \theta_i Y_{t-i} + \sum_{j=1}^{k} \sum_{I_j=0}^{q_j} B_{j,I_j} Z_{j,t-I_j} + e_t \qquad (8)$$

Where:
(Y_t): refers to the dependent variable that is denoted as public expenditures, (Y_{t-i}) is the lags of the dependent variable.
(Z_t): refers to the independent variable (public revenues), (Z_{t-j}) is the lags of the independent variable.
(β_j): refers to the coefficient of the independent variable.
(a_0): refers to the constant term of the model.
(a_1): refers to the coefficient of the trend.
(e_t): refers to the random error

(p, q) : are the lags of the dependent and independent variables, respectively.

In formula (8), the linear behavior of public expenditures is modeled in relation to revenues. This model is insufficient to describe and model the non-linear behavior of the expenditure series. The part not explained linearly is expressed as(e_t). Consequently, the estimated time series of the state's public expenditures $(\mathrm{Y_t})$ can be described as follows [8] :

$$\hat{Y}_t = \hat{L}_t + \hat{N}_t \qquad (9)$$

Where $(\hat{\mathrm{L}}_\mathrm{t})$ represents the linear part and is obtained from the traditional (ARDL) model, while $(\widehat{N}_t)$ represents the nonlinear part of the expenditure series and is obtained from the analysis of the residuals resulting from formula (8) using the (LSTM) model, such that:

$$e_t = f(e_{t-1}, e_{t-2}, \ldots, e_{t-T}) + \varepsilon_t \qquad (10)$$

Where (f) represents a non-linear function and (ε_t) represents the random error.

From estimating the residuals of formula (10), the nonlinear part $(\hat{N}_t)$ of the public expenditure series can be predicted. Thus, the predicted series will be the sum of the linear predictions using (ARDL) and the nonlinear predictions using (LSTM) according to the following formula:

$$EXP_{t+1} = EXP_{t+1}^{ARDL} + e_{t+1}^{LSTM} \tag{11}$$

The procedural sequence of the proposed model can be illustrated in the following Flowchart:

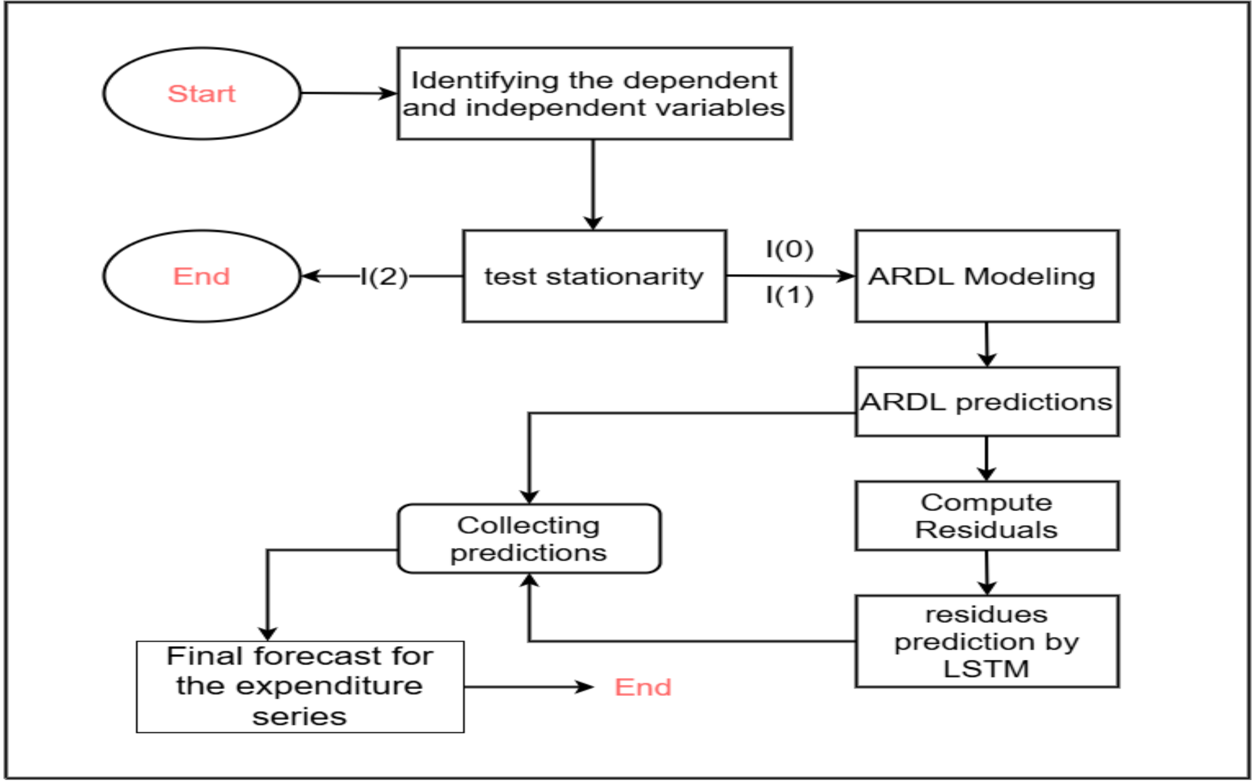


Fig. 2. Flowchart for the ARDL-LSTM hybrid model

4. Application of the proposed models:

I. Data:

The sample size of the study was 144 observations of monthly data between the years 2013–2024, treated as dynamic monthly balances rather than cumulative values. The dependent variable is the public expenditure of the state, and the independent variable is the public revenues of the state. All values are expressed in trillion Iraqi dinars.

II. Result and discussion:

As stationarity is a prerequisite for constructing time series models or dynamic regression models such as the ARDL model, the KPSS test was employed to examine the public revenue and public expenditure series. The results indicate that the computed LM statistics for both variables at level exceed the critical values; therefore, the null hypothesis of stationarity is rejected. Conversely, after taking the first difference, the test statistics fall below the critical values, implying that the null hypothesis cannot be rejected at the first difference. Accordingly, as reported in Table (1), both series are integrated of order one $I(1)$.

Table (1) presents the results of the KPSS stationarity test for the revenue and public expenditure variables.

Variables	Level		First Diff.		Integration Degree
	LM-Stat.	Critical Values	LM-Stat.	Critical Values	
Revenues	0.261130	0.146	0.272740	0.463	I(1)
Expenditures	0.271878	0.146	0.267802	0.463	I(1)

After verifying the order of integration of the variables, the Autoregressive Distributed Lag (ARDL) model was estimated to examine the nature of the relationship between public revenues and public expenditures.

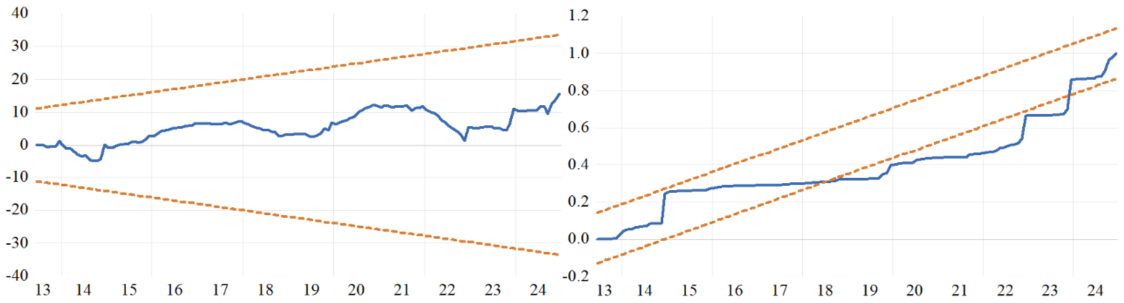
To determine the optimal lag order, several candidate models were compared using the Akaike Information Criterion (AIC). The results indicated that the most appropriate specification is ARDL(1,3) under the case without a constant and without a trend, with an AIC value of (5.68).

Following model estimation, the Bounds Test was conducted. The results revealed the existence of co-integration between the two variables, as the calculated F-statistic exceeded the upper bound critical values at all conventional significance levels, confirming the presence of a long-run equilibrium relationship between them.

Diagnostic tests were also conducted to verify the validity of the classical regression model assumptions. As shown in Table (2) and Figure (3), the results indicate that some of these assumptions are not fully satisfied.

The Breusch-Godfrey test revealed the presence of significant serial correlation in the model residuals. The Breusch-Pagan-Godfrey and White tests indicated heteroskedasticity, while the Jarque-Bera test rejected the null hypothesis of normality of the residuals. Furthermore, the Ramsey RESET test provided evidence of potential nonlinear components not explicitly captured in the linear specification.

Overall, these findings suggest that although the linear model captures the fundamental relationship between the variables, the residuals still contain underlying dynamic and nonlinear patterns.



b) CUSUM Square

a) CUSUM

Figure (3) illustrates the reliability test of the estimated model.

After estimating the ARDL model and extracting the linear component of the series, the residuals are obtained, representing the linearly unexplained part. Since the ARDL model captures the dynamic linear behavior, any remaining pattern in the residuals may reflect nonlinear behavior that can be modeled using LSTM. Accordingly, the forecasted series is formulated as the sum of two components: a linear component derived from the ARDL model and a nonlinear component generated by the LSTM model. The residuals are modeled as a nonlinear function of their lagged values, and the final forecast is reconstructed by combining the two forecasts. This approach is consistent with the theoretical formulation that considers the nonlinear component to be extracted from the ARDL residuals via LSTM and then added to the linear component to obtain the overall prediction.

The LSTM model was used to correct the residuals of the ARDL model, with the model trained on the residual series after the removal of the linear component. When applying this in the empirical analysis, a group of hyperparameters intended for use by all models in time series modeling were taken up to provide both learning stability and performance efficiency. The data were split in training sample (70 percent) and test sample (30 percent) based on their chronological order without mixing. The dynamic dependence of the residuals was captured by a time window of four periods (lag = 4), whereas the number of units in the

LSTM layer was 64 units, which was not excessive in terms of complexity.

The Adam optimization algorithm was used to train the model with the learning rate of 0.001. A loss function was a mean squared error (MSE). The batch size was used as 64 and the upper limit of the training epochs was 200. Besides this, an early stopping mechanism was applied to avoid Overfitting and improve the predictive power of the model.

To ensure training stability, the residuals were standardized before being fed into the LSTM model. The predicted values were then transformed back to their original scale prior to being combined with the ARDL model outputs to obtain the final forecast of the hybrid model. Out-of-sample forecasting was conducted for seven future periods using the recursive multi-step forecasting approach. The forecasting results are presented as follows:

Table (2) shows the predicted values

2025	Y_{Actual}	Y_{ARDL}	$Y_{ARDL\text{-}LSTM}$
Jan	9.07557	6.03361	5.56653
Feb	8.89180	8.57501	8.81947
Mar	10.17206	8.96849	9.75915
Apr	9.49242	8.84337	9.47173
May	9.34740	8.73856	9.15390
Jun	9.73296	12.27824	12.76965
July	20.62115	12.09205	12.59544

The following values represent the criteria for comparing the two methods, both within and outside the sample:

Table (3) shows the values of the comparison criteria

Methods	In-Sample		Out-Sample	
	RMSE	MAPE	RMSE	MAPE
ARDL	4.05879	0.34758	3.60195	0.18539
ARDL-LSTM	3.99094	0.30524	3.50837	0.16564

Table (3) indicates that the hybrid ARDL-LSTM model performs better than the traditional ARDL model during both the in-sample and out-of-sample periods, based on the RMSE and MAPE measures. Although the numerical improvement is relatively modest, it is consistent across all evaluation criteria, indicating the capability of the hybrid model to capture the nonlinear patterns in the residuals that the linear model could not fully represent. For visual comparison between the two approaches, the following figure presents the actual expenditure series along with the fitted values of the two models, ARDL and ARDL-LSTM

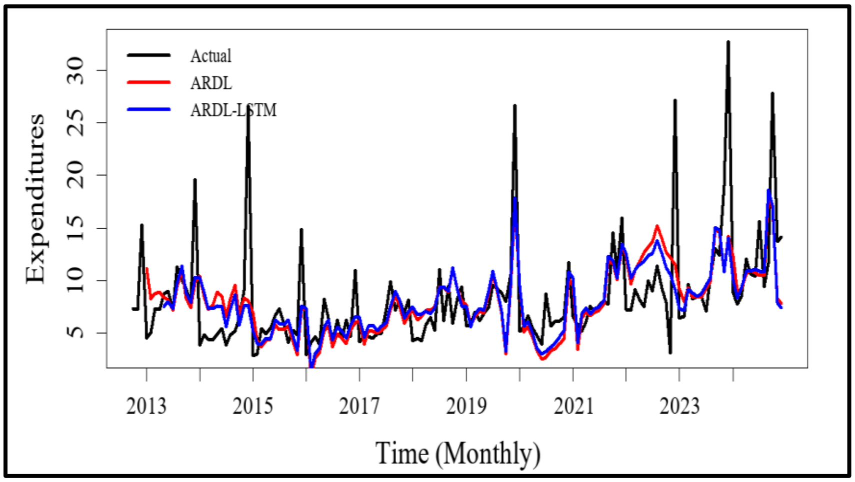


Figure (4) shows a diagram of the expenditure series with estimated values.

5. Conclusion:

The empirical analysis and diagnostic tests indicate that the traditional ARDL model, despite its ability to represent the relationship between the variables, did not fully satisfy all the assumptions of the classical regression model. The tests revealed the presence of certain issues in the model residuals, which may affect estimation efficiency and predictive accuracy.

In contrast, the hybrid ARDL–LSTM model demonstrated a consistent improvement in the forecasting accuracy measures (RMSE and MAPE), both in-sample and out-of-sample. Although the magnitude of improvement was numerically modest, it was consistent across different evaluation metrics. This reflects the hybrid model's superior ability to capture complex patterns in the data compared to the traditional linear model.

Accordingly, it can be concluded that employing hybrid approaches that combine econometric techniques with deep learning methods contributes to enhancing predictive performance and improving the overall model fit to the data.

References

[1] M. Thenmozhi, "Forecasting stock index returns using ARIMA-SVM , ARIMA-ANN , and ARIMA-random forest hybrid models Manish Kumar *," *Int. J. Banking, Account. Financ.*, vol. 5, no. 3, pp. 284–308, 2014, doi: 10.1504/IJBAAF.2014.064307.

[2] F.-M. Tseng, G.-H. Tzeng, H.-C. Yu, and B. J. C. Yuan, "Fuzzy ARIMA model for forecasting the foreign exchange market," *Fuzzy Sets Syst.*, vol. 118, pp. 9–19, 2001, doi: https://doi.org/10.1016/S0165-0114(98)00286-3.

[3] M. L. De Prado, "Beyond Econometrics: A Roadmap Towards Financial Machine Learning," *SSRN Electron. J.*, pp. 1–32,

2020, doi: 10.2139/ssrn.3365282.

[4] M. H. Pesaran, Y. Shin, and R. J. Smith, “Bounds testing approaches to the analysis of level relationships,” *J. Appl. Econom.*, vol. 16, pp. 289–326, 2001, doi: 10.1002/jae.616.

[5] N. Ozturk, A. Prof, K. Karagoz, and A. Prof, “Relationship Between Inflation and Financial Development: Evidence from Turkey Etkileri: Türkiye İçin Bulgular,” *Int. J. Alanya Fac. Bus.*, vol. 4, no. 2, pp. 81–87, 2012, [Online]. Available: https://www.ajindex.com/dosyalar/makale/acarindex-1423869086.pdf

[6] M. E. P´erez-Pons, J. Parra-Dominguez, S. Omatu, E. Herrera-Viedma, and J. M. Corchado, “MACHINE LEARNING AND TRADITIONAL ECONOMETRIC MODELS : A SYSTEMATIC MAPPING STUDY,” *J. Artif. Intell. Soft Comput. Res.*, vol. 12, no. 2, pp. 79–100, 2022, [Online]. Available: https://www.researchgate.net/publication/358879543_Machine_Learning_and_Traditional_Econometric_Models_A_Systematic_Mapping_Study

[7] D. Stempie and R. Slepaczuk, “Hybrid Models for Financial Forecasting : Combining Econometric , Machine Learning , and Deep Learning Models,” *arXiv.org*, pp. 1–30, 2022, [Online]. Available: https://arxiv.org/abs/2505.19617

[8] C. H. Aladag, E. Egrioglu, and C. Kadilar, “Improvement in Forecasting Accuracy Using the Hybrid Model of ARFIMA and Feed Forward Neural Network,” *Am. J. Intell. Syst.*, vol. 2, no. 2, pp. 12–17, 2012, doi: 10.5923/j.ajis.20120202.02.

[9] M. A. Khan, H. Ali, H. Shabbir, F. Noor, and M. D. Majid, “Impact of Macroeconomic Indicators on Stock Market Predictions : A Cross Country Analysis,” *J. Comput. Biomed. Informatics*, vol. 08, no. 01, 2024, doi: doi.org/10.56979/801/2024.

[10] S. Hochreiter and J. Schmidhuber, “LONG SHORT-TERM MEMORY,” *Neural Comput.*, vol. 9, no. 8, pp. 1–32, 1997, [Online]. Available: https://www.bioinf.jku.at/publications/older/2604.pdf

[11] F. A. Gers and N. N. Schraudolph, “Learning Precise Timing with LSTM Recurrent Networks,” *J. Mach. Learn. Res.*, vol. 3, pp. 115–143, 2002, [Online]. Available: https://www.jmlr.org/papers/volume3/gers02a/gers02a.pdf

[12] K. Greff, R. K. Srivastava, J. Koutn, and B. R. Steunebrink, “LSTM : A Search Space Odyssey,” *Trans. neural networks Learn. Syst.*, vol. 28, no. 10, pp. 2222–2232, 2016, doi: 10.1109/TNNLS.2016.2582924.

[13] Y. Yu, X. Si, C. Hu, and J. Zhang, “A Review of Recurrent Neural Networks : LSTM Cells and Network Architectures,” *Neural Comput.*, vol. 31, pp. 1235–1270, 2019, doi: https://doi.org/10.1162/neco_a_01199.

[14] H. Abbasimehr and R. Paki, “Improving time series forecasting using LSTM and attention models,” *J. Ambient Intell. Humaniz. Comput.*, no. 0123456789, 2021, doi: 10.1007/s12652-020-02761-x.

[15] S. M. Al-selwi, M. Fadzil, and S. Jadid, “Journal of King Saud University - Computer and Information Sciences RNN-LSTM : From applications to modeling techniques and beyond — Systematic review ☆,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 36, no. 5, p. 102068, 2024, doi: 10.1016/j.jksuci.2024.102068.

[16] X. Wei, L. Zhang, H. Yang, L. Zhang, and Y. Yao, “Geoscience Frontiers Machine learning for pore-water pressure time-series prediction : Application of recurrent neural networks,” *Geosci. Front.*, no. January, 2020, doi: 10.1016/j.gsf.2020.04.011.

[17] M. Eren, “Fuzzy Autoregressive Distributed Lag model-based forecasting,” *Fuzzy Sets Syst.*, vol. 459, pp. 82–94, 2023, doi: 10.1016/j.fss.2022.06.003.