Tikrit Journal of Administrative and Economic Sciences, Vol. 19, No. 62, Part (2): 531-543
Doi: www.doi.org/10.25130/tjaes.19.62.2.29

Tikrit Journal of Administrative and Economics Sciences
مجلة تكريت للعلوم الإدارية والاقتصادية
ISSN: 1813-1719 (Print)

Employing Recurrent Neural Networks to Forecast the Dollar Exchange Rate in the Parallel Market of Iraq

Azhy Akram Aziz* [A], **Balsam Mustafa Shafeeq** [B], **Renas Abubaker Ahmed** [C], **Hindreen Abdullah Taher** [D]

[A] college of business and administration, Erbil Polytechnic University.
[B] Information Technology department, Technical College of Management, Middle Technical University
[C] Department of Statistics and Informatics, College of Administration & Economic, University of Sulaimani
[D] Department of Information Technology, college of Commerce, University of Sulaimani

Keywords:
Forecasting, Neural Networks, Activation Function, Exchange Rate.

A R T I C L E I N F O

Article history:

Received	19 Jun. 2023
Accepted	25 Jun. 2023
Available online	30 Jun. 2023

*Corresponding author:
Azhy Akram Aziz
college of business and administration, Erbil Polytechnic University.

Abstract: In this study, neural recurrent neural networks (RNN) have been used to forecast the price of dollars in Iraqi dinars, as it is clear that the government's efforts to control prices in the parallel markets, the commercial markets witnessed a decline in the exchange rate, but it rose again. Which indicates an economic problem that is still present in the country. Despite adjusting the exchange rate of the dinar, the dollar crisis in Iraq has not ended yet. Here we want to forecast the daily price of the dollar against Iraqi dinars for the common next 30 days. According to the results the RNN model have been performed for the data under consideration with different numbers of hidden layer and nodes. The best architecture for the RNN model was [1,10,1,1] using soft plus activation function, which gives the performance of 85% for the training dataset and (92% and 90%) for the testing and validation datasets respectively, with Mse (0.018, 0.000417, and 0.000477) for training, testing, and validation respectively at epoch 4. According to the results of the forecasted values which start from 15 May 2023 to 13 June 2023 the price of dollars will be between 1390 to 1435.

Tikrit Journal of Administrative and Economic Sciences, Vol. 19, No. 62, Part (2): 531-543
Doi: www.doi.org/10.25130/tjaes.19.62.2.29

توظيف الشبكات العصبية المتكررة للتنبؤ بسعر صرف الدولار بالسوق العراقي الموازي

ئەژی اكرم عزيز
قسم التقنية الاعلام، كلية التقنية الادارية أربيل
جامعة التقنية أربيل، اقليم كوردستان العراق

بلسم مصطفى شفيق
قسم تقنيات المعلوماتية
الكلية التقنية الادارية الجامعة التقنية الوسطى

ريناس ابوبكر احمد
قسم الإحصاء والمعلوماتية
كلية الإدارة الاقتصادية جامعة السليمانية

هندرين عبدالله طاهر
قسم تقنية المعلومات، كلية التجارة
جامعة السليمانية

المستخلص

في هذه الدراسة، تم استخدام الشبكات العصبية المتكررة للتنبؤ بسعر الدولار بالدينار العراقي، حيث يتضح أن جهود الحكومة للسيطرة على الأسعار في الأسواق الموازية، شهدت الأسواق التجارية تراجعا في سعر الصرف، لكنها ارتفع مرة أخرى، مما يشير إلى مشكلة اقتصادية لا تزال موجودة في البلاد. رغم تعديل سعر صرف الدينار، فإن أزمة الدولار في العراق لم تنته بعد. هنا نريد خلال الثلاثين يومًا القادمة أن ترتفع القيم المتوقعة للسعر اليومي للدولار مقابل الدينار العراقي أم لا، أفضل بنية لنموذج RNN كانت [1،10،1،1] باستخدام وظيفة التفعيل السوفت بلس، والتي تعطي أداء 85٪ لمجموعة بيانات التدريب و (92٪ و90٪) لمجموعات بيانات الاختبار والتحقق على التوالي، مع (Mse 0.018)، (0.000417، و0.000477) للتدريب والاختبار والتحقق على التوالي في الحقبة الرابعة. نتائج القيم المتوقعة والتي تبدأ من 15 مايو 2023 حتى 13 يونيو 2023 سيكون سعر الدولار بين 1390 و1435.

الكلمات المفتاحية: التنبؤ، الشبكات العصبية، وظيفة التنشيط، سعر الصرف.

1. Introduction

Research and development in the field of financial market analysis has been very popular, as it can represent exchange rate changes in a time series (D. Zhang, Q. Jiang and X. Li, 2005: 106-108). SM contributes significantly to the growth of economies worldwide. As a result of its importance in the industry, it influences the national economy. Stock market value forecasting has become challenging due to the nonlinear, uncertain nature of the relationship between inputs and outcomes. Selecting an appropriate training and prediction approach remains a significant issue (M. Adya and F. Collopy, 1998: 481-495). You can use numbers like daily highs and lows, stock volume, indicators of trends, the day's highest and lowest prices, lists, moving averages, etc (M. P. Wallace, 2008: 67-77). Attempts have been made in the past to find mathematical models that can properly distribute these inputs and accurately anticipate the desired outputs (R. Lawrence, 997). Numerous models serve as examples. However, the use of neural networks is becoming commonplace. Recurrent neural networks (RNNs) are utilized in deep learning and in the creation of models that mimic human brain neuron activity. They differ from other artificial neural networks in that they use feedback loops to digest a sequence of data that informs the final output, making them useful in situations where context is crucial to forecasting an outcome. This persistence of knowledge is made possible by feedback loops. Memory is a common term for this effect. Data processing components in artificial neural networks are integrated in a loosely human-like model (M. Thenmozhi, 2006: 59-69). They are built up from interconnected layers of artificial neurons (also known as network nodes) that can take in data and send it on to other nodes (Kh, A. M., & Kh, A. A., 2022,.95). The edges, or weights, between the nodes affect the strength of the signal and the network's final output (Askar, A. J., and et al., 2023). Some artificial neural networks only do one-way processing, from input to output. For example, image recognition systems rely on convolutional neural networks, which are a type of "feed-forward" neural network (Al-Mashhadani, B. N., 2022). However, RNNs have the capability of being layered, allowing for bidirectional processing of data (Abd-AlHassin, and et al. 2023).

2. Methodology:

2.1. Artificial Neural Networks (ANNs):

An Artificial Neural Network (ANN) is a type of information processing system inspired by neural network models. During training, ANNs can undergo modifications to their network architecture and data, making them a very adaptable system (Sutskever, Ilya, 2013). A computer model of the human brain's network of interconnected processing units, also known as an Artificial Neural Network (Taher, H. A., & Ahmed, N. M., 2023.2079).

2.2. Architecture of ANNs:

The ANN is a machine learning strategy that simulates the operation of a real brain through the use of a network of virtual neurons. Each neuron in an ANN will send in a number of factors (Ahmed, Nawzad, M, 2007). Applying an activation function to these inputs yields the neuron's output values. Layers of input, hidden processing, and output make up the ANN architecture (Ahmed, Farhad, 2014).

Doi: www.doi.org/10.25130/tjaes.19.62.2.29

2.3. Recurrent Neural Networks (RNNs):

In the fields of speech recognition and natural language processing, recurrent neural networks are a popular artificial neural network architecture. In order to predict what might happen next, recurrent neural networks are able to pick up on the sequential patterns in data (Sutskever, Ilya, 2013). The RNNs technically establish the standard that this work is based on, as stated by (Ahmed, Farhad, 2014). Assigned a persistent input the nets (X_1, X_2,,……..., X_T), the network calculates a continuance of hidden case (H_1,H_2,…..,H_T), and a continuance of estimation ($\hat{Y}_1$,$\hat{Y}_2$,……,$\hat{Y}_T$), by repetition the equations:

T_i=$W_{hx}X_i$+$W_{hh}H_{t-1}$ +B_h (1)

H_i= E (T_i) (2)

Z_i= $W_{yh}H_i$+ B_y (3)

$\hat{Y}_i$= $f(\hat{Y}_i)$ (4)

Where:

W_{hx}: is the weight matrices between factors layer and hidden layer.

W_{yh}: is the weight matrices between hidden layer and output layer.

W_{hh}: is the matrix of recurrent weights between the hidden layer and itself at adjacent time steps.

E,f: are the activation functions.

B_h: Bias of hidden layer.

B_y: Bias of output layer.

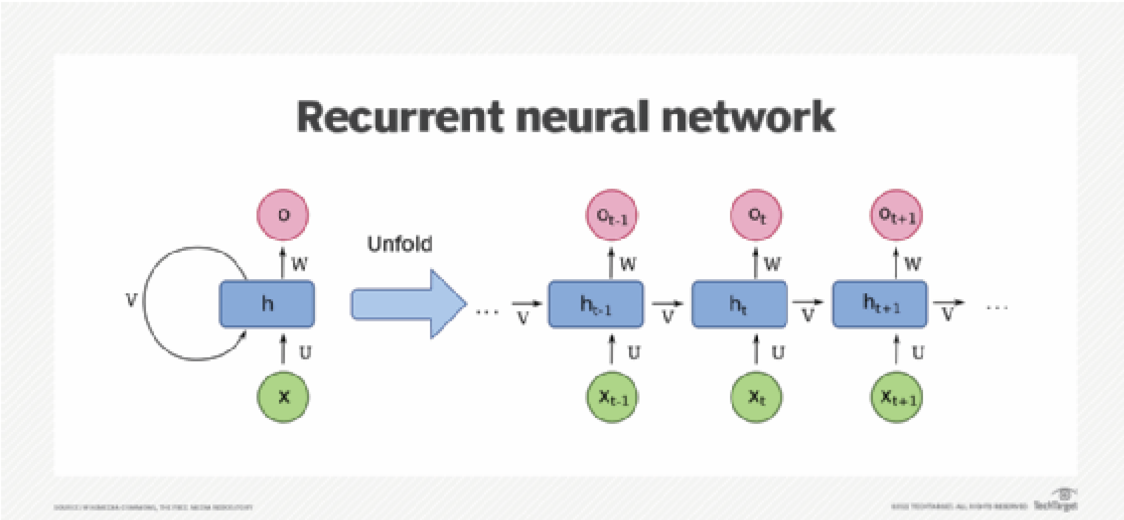

Figure (1): RNN where connections between units form a directed (Sutskever, Ilya, 2013)

Doi: www.doi.org/10.25130/tjaes.19.62.2.29

2.4. Activation Function:

The activation function is used in most of a neural network's output layers. The activation functions work to constrain the neural network's output to reasonable bounds. Synaptic connections carry the activation value to other units (Taher, H. A., & Ahmed, N. M., 2023 pp.2077-2087). Returning to squashing, which has a limited activation function range, is what gives neural networks the ability to deal with non-linear patterns; this process is also known as a (Transfer). The effectiveness of the neural network model can be improved by experimenting with different activation functions (Yu. Hao, 2010). The four activation functions tested are summarized in a table for ease of reference below.

Table (1): Represents the activation functions

N.	Name of Function	Function
1.	Sigmoid/Logistic	$f(x) = \frac{1}{1+e^{-x}}$
2.	Hyperbolic Tangent (Tanh)	$f(x) = \frac{(e^{x}-e^{-x})}{(e^{x}+e^{-x})}$
3.	Softplus	$f(x) = \ln(1+e^{x})$
4.	Growing Cosine Unit (Gcu)	$x\cos(x)$

2.5. Neural Network Training:

Training a neural network involves gauging how well it performs initially and exposing it to the pattern environment so that it becomes accustomed to it over time (Taher, H. A., & Ahmed, N. M., 2023 pp.2077-2087). Training can be divided into two categories: supervised and unsupervised (Sutskever, Ilya, 2013).

2.6. Weights:

When analyzing neural networks, each neuron is assigned a weight that determines how much of an impact it will have on the final result (Ahmed, Nawzad, M, 2007). Values of weights are linked in the same way that the weight of a real neuron corresponds to the strength of its synaptic connection, the weight values associated with each vector and node in a network reveal the relationship between their input and their output (Alex Graves, 2008). Biases are another name for the weight values of specific nodes. During training, the network learns to distinguish between different classes based on the usual features that they share in the input data, and these values are the result of this process.

$Net=\sum_{j=1}^{n} x_j w_{ij}$, $y_{j=} \sum_{j=1}^{n} x_j w_{ij}$ or $\hat{y}=xw$ (5)

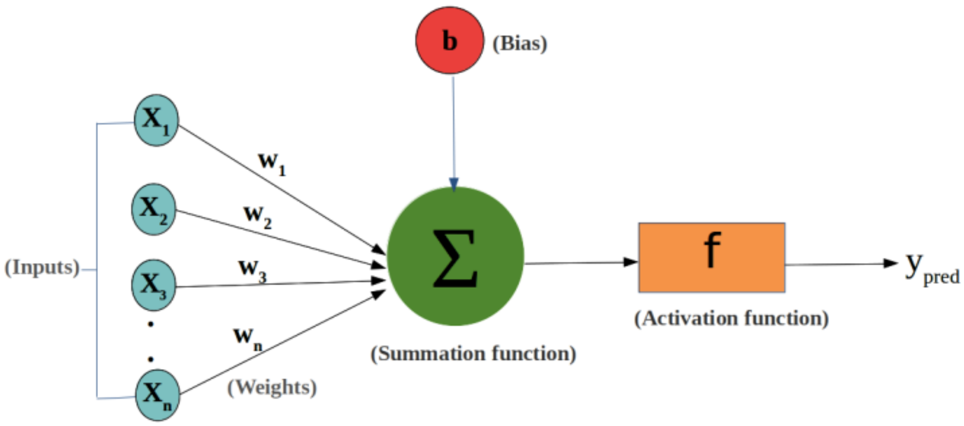

Figure (2): represents the weight values corresponding to the strength of synaptic connections (Taher, H. A., & Ahmed, N. M., 2023 pp.2077-2087).

2.7. The Gradient Descent Algorithm:

To locate a local minimum or maximum of a function, the first-order optimization process known as gradient descent (GD) can be applied. This strategy is frequently employed in machine learning (ML) and deep learning (DL) to minimize a cost/loss function (as in a linear regression, for instance) (Taher, H. A., & Ahmed, N. M., 2023 pp.2077-2087). This algorithm is typically taught first in almost all machine learning classes because of its significance and ease of implementation. Although this approach was proposed long before the advent of modern computers, it has undergone much improvement in the meantime, leading to various improved variations; still, for the sake of this article, a search for a stationary point may be broken down into two parts: the direction and the step size (Ahmed, Nawzad, M, 2007). The direction parameter specifies the next search direction, while the step size controls the depth of that search. Iteratively travelling in the direction x(k) by a step size tk from an initial arbitrary point x(0) to a final destination x(k+1) = x(k) + tk. x(k) is a convenient way to think about these kinds of procedures. When using gradient descent, the direction of travel is determined by the point's negative gradient, or x = - f (x). Therefore, the recursive rule below can be used to explain the gradient descent iterative search algorithm:

$X^{(k+1)} = x^{(k)} - t_k \vee f\,(x^{(k)})$ (6)

Doi: www.doi.org/10.25130/tjaes.19.62.2.29

Step size selection. How should we pick our step size tk if we know that we are already at position x(k) in our search for a stationary point? The value of the new point is minimized by selecting the step size so as to achieve our goal of minimizing the function. Find the step size that minimizes $f(x^{(k+1)})$. Since $x^{(k+1)} = x^{(k)} - t \vee f(x^{(k)})$ the step size t^*_k of this approach is:

$$t^*_k = \text{argmin}_{t\geq 0} f(x^{(k)} - t \vee f(x^{(k)})) \qquad (7)$$

For now we will assume that t^*_k can be computed analytically, and later revisit this assumption.

2.8. Performance measurements [1,5]

An estimator's mean squared error (MSE) or mean squared deviation (MSD) is the average square of the mistakes or deviations, or the gap between the estimate and the true value. An R2 of 1 shows aa perfect fit between the regression line and the data, while an R2 of 0 indicates no fit (Ahmed, B. K., 2020, pp.908-940), (Asraa, A., 2018)

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \qquad (8)$$

$$R^2 = 1 - \frac{SS_{residual}}{SS_{total}} \qquad (9)$$

3. Application:

3.1. Introduction:

In this Section the best architecture and best design for the recurrent neural network (RNN) is determined to view the contains of the suggested RNN about the number of layers of hidden layer and number of Nodes in each hidden layer. Also, to determine the best activation functions between layers that makes the Recurrent Neural Network to have the best performance for the data under consideration.

3.2. Data Description:

The data were taken from the dollars market in Sulaimani, which is about the average daily price of dollars in Iraqi dinar, form 16 November 2023 to 14 May 2023.

Doi: www.doi.org/10.25130/tjaes.19.62.2.29

3.3. Application of the RNN

The Matlab v.9.11 has been used to perform the model with initial weights ($-0.6 < w_i < 0.6$) the steps of fitting best RNN architecture for forecasting daily the price of dollars in Iraqi's dinars are shown below:

Table (2) Demonstrate the performance of the RNN model according to activations functions.

Activation function	Dataset	Mse	R2
Softplus	Training	0.0183	85.19%
Sigmoid	Training	0.0255	81.04%
Hyperbolic Tangent	Training	0.0281	77.41%
Growing Cosine Unit (Gcu)	Training	0.0322	72.07%

Source: Prepared by researchers based on the statistical program.

The above table shows the results of the training data set up to the activation functions, as it is clear that Softplus gives high performance among the others, so we will show up the details results depending on Softplus activation function.

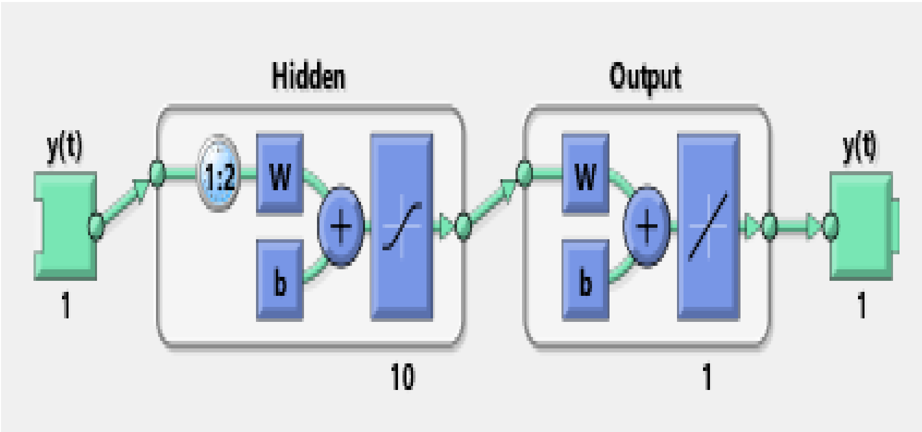

Figure (3): Displays the architecture of the neural networks model (Matlab v.9.11).

Figure (3) represent the best architecture RNN model (1-5-10-1). The best Network is [1-10-1-1], depend on Maximum fitness, Minimum AIC and mean square error for (training, testing, validation). By using applying the Softplus activation function between layers and change the number of nodes between layers.

Doi: www.doi.org/10.25130/tjaes.19.62.2.29

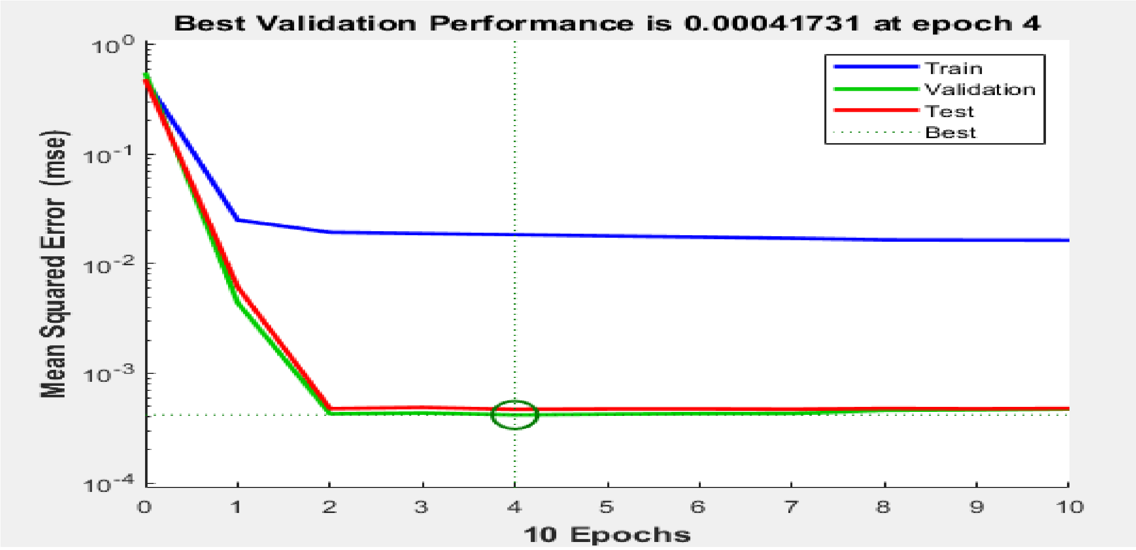

Figure (4): represent the training performance (Matlab v.9.11).

It is clear that the global solution is achieved at epoch 4with minimum Mse and higher R^2.

Source: Prepared by researchers based on the statistical program

Table (3): Represents the performance of the model.

Datasets	Observations	Mse	R^2
Training	143	0.0183	85.19%
Testing	18	0.000417	92.77%
Validation	18	0.000477	90.72%

Summing to up table, which is represents the Mse and R^2 for training, testing and validation data sets.

Source: Prepared by researchers based on the statistical program

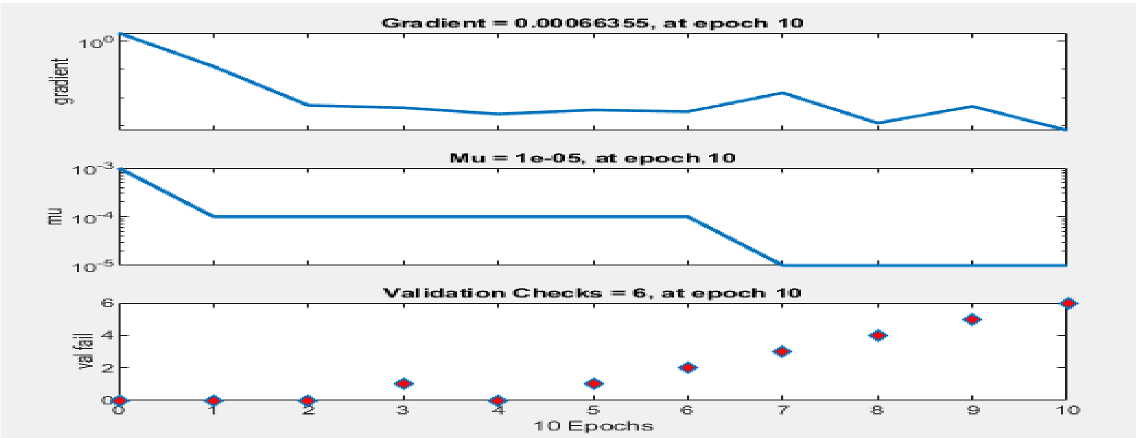

Figure (5): show the training state (Matlab v.9.11).

The variation of the gradient error (0.00066355) and validation checks at epoch (10) equal to (6).

Source: Prepared by researchers based on the statistical program

Tikrit Journal of Administrative and Economic Sciences, Vol. 19, No. 62, Part (2): 531-543
Doi: www.doi.org/10.25130/tjaes.19.62.2.29

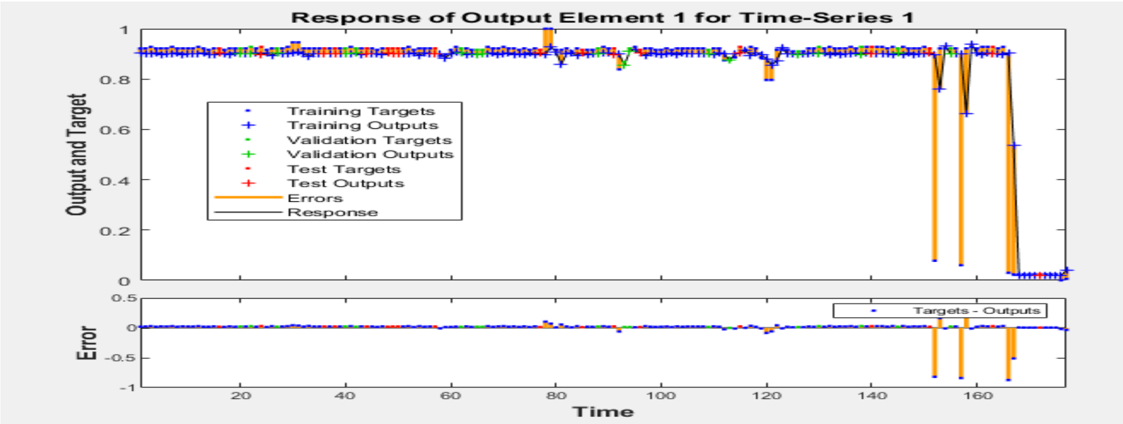

Figure (6): represents the Actual data, prediction and the residuals (Matlab v.9.11).

Source: Prepared by researchers based on the statistical program

$$b_1^T = \begin{pmatrix} 2.95164 \\ -1.53044 \\ 1.61271 \\ 1.12750 \\ 0.33014 \\ -1.70073 \\ -2.07889 \\ -0.61847 \end{pmatrix} \quad W_1^T = \begin{pmatrix} -2.25597 \\ 0.84346 \\ 1.13152 \\ 5.43744 \\ -1.54857 \\ -3.74371 \\ -1.81927 \\ -1.92238 \\ 0.27728 \\ -2.63861 \\ -0.22434 \\ -2.59909 \\ 1.52976 \\ 0.25321 \\ 2.48402 \\ 0.91161 \\ -0.67220 \\ -1.37233 \\ -1.60841 \\ -1.26843 \\ -3.00771 \\ -5.64612 \\ -2.63983 \\ 1.32044 \end{pmatrix} \quad b_2 = \begin{pmatrix} 0.56074 \end{pmatrix} \quad W_2^T = \begin{pmatrix} -0.02369 \\ 2.02628 \\ -1.07723 \\ -0.47024 \\ 0.72676 \\ -1.17070 \\ 0.89083 \\ 1.13417 \end{pmatrix}$$

The above matrixes are the values of the biases and the weight of the RNN best model, the b_1 and W_1 are for the first layer and the b_2 and W_2 are for the second layer.

Table (4) Shows the forecasted values.

Periods	**Forecsted**	**Confidence Intervls**	
		95% lower	**95% Uper**
15 May. 2023	1421.79	1408.24	1435.34
16 May. 2023	1433.1	1419.55	1446.65
17 May. 2023	1421.7	1408.15	1435.25
18 May. 2023	1421.7	1408.15	1435.25
19 May. 2023	1421.7	1408.15	1435.25
20 May. 2023	1420.7	1407.15	1434.25
21 May. 2023	1421.7	1408.15	1435.25
22 May. 2023	1420.7	1407.15	1434.25
23 May. 2023	1390.3	1376.75	1403.85
24 May. 2023	1420.7	1407.15	1434.25
25 May. 2023	1420.7	1407.15	1434.25
26 May. 2023	1420.67	1407.12	1434.22
27 May. 2023	1416.17	1402.62	1429.72
28 May. 2023	1431.1	1417.55	1444.65
29 May. 2023	1416.17	1402.62	1429.72
30 May. 2023	1435.1	1421.55	1448.65
31 May. 2023	1416.17	1402.62	1429.72
1 Jun. 2023	1416.2	1402.65	1429.75
2 Jun. 2023	1416.17	1402.62	1429.72
3 Jun. 2023	1417.2	1403.65	1430.75
4 Jun. 2023	1420.52	1406.97	1434.07
5 Jun. 2023	1417.2	1403.65	1430.75
6 Jun. 2023	1435.1	1421.55	1448.65
7 Jun. 2023	1417.2	1403.65	1430.75
8 Jun. 2023	1417.2	1403.65	1430.75
9 Jun. 2023	1393	1379.45	1406.55
10 Jun. 2023	1391	1377.45	1404.55
11 Jun. 2023	1391.1	1377.55	1404.65
12 Jun. 2023	1390	1376.45	1403.55
13 Jun. 2023	1390	1376.45	1403.55

Source: Prepared by researchers based on the statistical program.

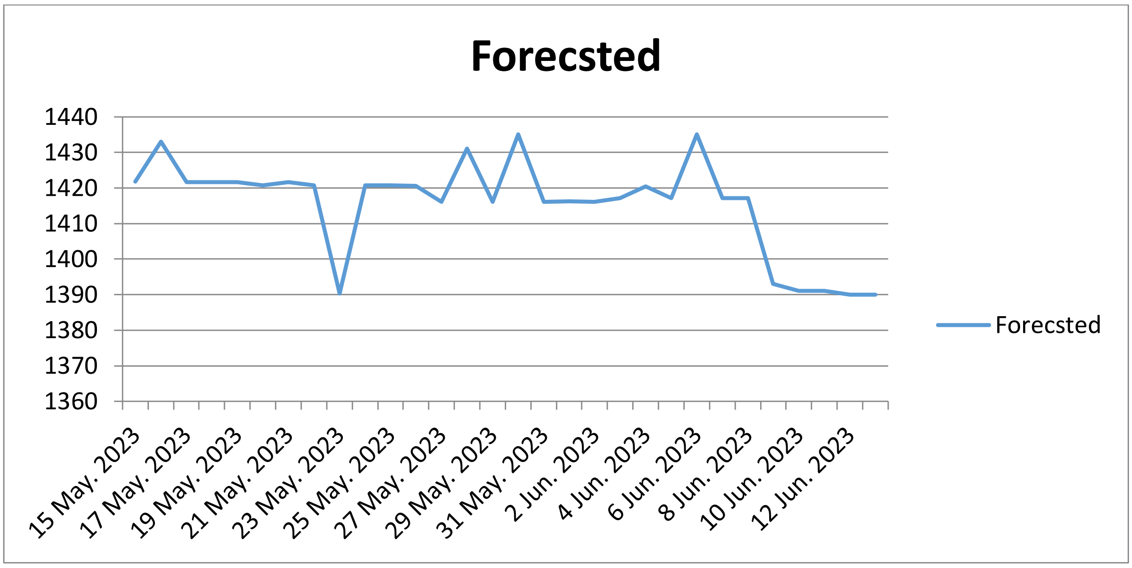

Figure (7): Demonstrate the forecasted values for thirty periods of time (Matlab v.9.11).

Source: Prepared by researchers based on the statistical program

4. Conclusions:

As conclusion we get the best architecture was [1,10,1,1], and Softplus gives best result compared to the others aforementioned activation functions, this study used RNN to forecast the price of dollars in Iraqi dinars. The Mse were (0.018, 0.000417, and 0.000477) for training, testing, and validation respectively at epoch 4. From 15 May 2023 through 13 June 2023, the predicted values place the value of the US dollar between 1390 and 1435.

References

1. Ahmed, B. K., Rahim, S. A., Maaroof, B. B., & Taher, H. A. (2020). Comparison Between ARIMA And Fourier ARIMA Model to Forecast the Demand of Electricity in Sulaimani Governorate. Qalaai Zanist Journal, 5(3), 908-940.
2. Ahmed, Nawzad, M, (2007), “Forecasting water Inflow and Electric Power Generation in Darbandi-khan Station by Using FFNN, and Adaptive Method in Time series”, Doctorate Thesis, Statistics Dept., Administration and economics college, Sulaimani, Iraq.
3. Ahmed, Farhad, A, (2014), “Using FFNN and NARX models to Forecast daily temperature in Sulaimani”, MSc Thesis, Statistics Dept., Administration and economics college, Sulaimani, Iraq.
4. Alex Graves, (2008), “Supervised Sequence Labelling with Recurrent Neural Networks”, Technische Universität München Fakultät für Informatik, PhD. Thesis.
5. Asraa, A., Rodeen, W., & Tahir, H., (2018). Forecasting the Impact of Waste on Environmental Pollution. International Journal of Sustainable Development and Science, 1(1), 1-12.

6. D. Zhang, Q. Jiang and X. Li, "Application of Neural Networks inFinancial Data Mining," International Journal of Information and Mathematical Sciences, vol. 1, pp. 106–109, Spring 2005.
7. M. Adya and F. Collopy, "How Effective are Neural Networks at Forecasting and Prediction? A Review and Evaluation," Journal of Forecasting, vol. 17, pp. 481–495, September – November 1998.
8. M. P. Wallace, "Neural Networks and Their Application to Finance," Business Intelligence Journal, vol. 1, pp. 67–77, July 2008.
9. M. Thenmozhi, "Forecasting Stock Index Returns Using Neural Networks," Delhi Business Review, vol. 7, pp. 59–69, July December 2006.
10. R. Lawrence, "Using Neural Networks to Forecast Stock Market Prices," Department of Computer Science, University of Manitoba, Winnipeg, Canada, Tech. Rep., Dec. 1997, Available: https: //people.ok.ubc.ca/rlawrenc/research/Papers/nn.pdf.
11. Sutskever, Ilya, (2013), “Training Recurrent Neural Networks”, a thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy Graduate Department of Computer Science University of Toronto.
12. Taher, H. A., & Ahmed, N. M., (2023). Using Bayesian Regression Neural Networks Model to Predict Thrombosis for Covid-19 Patients. resmilitaris, 13(1), 2077-2087.
13. Yu. Hao, (2010), “Levenberg–Marquardt Training”, Bogdan M. Wilamowski, Auburn University.
14. Kh, A. M., & Kh, A. A., (2022). application Deep learning to Predict crypto currency Prices and their relationship to Market adequacy (applied research bitcoin as an example). Финансы: теория и практика, 26(4), 95-108.
15. Abdalhammed M.K., Ghazal A.M., Ibrahim H.M., Ahmed A.K. Application Deep Learning to Predict Crypto Currency Prices and their Relationship to Market Adequacy (Applied Research Bitcoin as an Example). Finance: Theory and Practice. 2022;26(4):95-108. https://doi.org/10.26794/2587-5671-2022-26-4-95-108
16. Askar, A. J., Battal, A. H., & Hamad, A. A., (2023). Forecasting foreign reserves in the Central Bank of Iraq until 2025. Tikrit Journal of Administration and Economics Sciences, 19(61 part 1).
17. Al-Mashhadani, B. N., (2022). The Role of the Internal Auditor in Predicting Expected Credit Losses Suggested Procedures for Implementation in Banks Listed in the Iraq Stock Exchange. Tikrit Journal of Administration and Economics Sciences, 18(57 part 1).
18. Abd-AlHassin, S. K., & Abd-Alreda, N. Y., (2023). Comparison between prediction of demand for electric power using artificial intelligence techniques and traditional methods (ARIMA) A case study in the diesel station north of Al-Amarah to generate electric power. Tikrit Journal of Administration and Economics Sciences, 19(61 part 2).