استخدام الشبكات العصبية لتصنيف حالات وفيات كوفيد-19 في العراق

م. منى طاهر غافل

جامعة البصرة / كلية الإدارة والاقتصاد / قسم الإحصاء

muna.ghafil@uobasrah.edu.iq

الملخص:

استخدمت الشبكات العصبية في دراسة وتصنيف وفيات كوفيد-19 في العراق على أربع مراحل مختلفة، مما أدى إلى دقة تصنيف بلغت 96.7٪ للوفيات شديدة الخطورة و95.5 ٪ للوفيات منخفضة الخطورة. كانت النسبة المئوية للتصنيف الصحيح للوفيات شديدة الخطورة 86 ٪ عند 10٪ و91 ٪ للوفيات منخفضة الخطورة. وأخيرا، بلغ معدل دقة الشبكة 93.9٪. يوضح هذا البحث فعالية الشبكات العصبية في فهم وتصنيف حالات وفيات كوفيد-19 ودراسة تأثير الجائحة في العراق

الكلمات المفتاحية : طبقة الإدخال ، طبقة مخفية ، طبقة الإخراج ، الشبكات العصبية ، دالة سوفتماكس ، عبر دالة خطأ الانتروبيا.

Using Neural Networks to Classify COVID-19 Death Rates Cases in Iraq

Muna T. Ghafil

College of Administration and Economics, Statistics Department, Basrah University, Basrah, Iraq

Abstract:

Neural networks were used in the current study for the classification of Covid-19 death rates in Iraq at four different stages. This was led to a classification accuracy of 96.7% for high-severity deaths, and 95.5% for low-severity deaths. The percentage of correct classification of high-severity deaths, it was 86% at 10%, and 91% for low-severity deaths. Finally, the network accuracy rate reached 93.9%. This research paper shows the effectiveness of neural networks in understanding and classifying cases of Covid-19 deaths and studying the impact of the pandemic in Iraq.

Keywords: Input layer; Hidden layer; Output layer; Neural networks; Softmax function; Cross entropy error function**.**

1. INTRODUCTION

The Covid-19 virus had a significant impact on harming Iraqis in various fields of life, especially health, education, and economy, starting from its spread in February 2020, and on mitigating the effects of the epidemic, the Iraqi Ministry of Health directed efforts by developing specific mechanisms to confront the pandemic by the contexts defined by the World Health Organization, where artificial intelligence and neural networks, intense learning, played a significant role in confronting COVID-19. [Al-Zawahra, 2021:3-6]

The first infection with the Covid-19 virus in Iraq was on February 24 in Najaf Governorate, followed by the announcement of four infections in Kirkuk Governorate, and the next day an infection was announced in Baghdad Governorate. Thus, infections continued to appear and increase with the increase in deaths rates. With the tightening of precautionary measures that Iraq witnessed to control the spread of the virus, the number of infections decreased. Thus, the number of deaths rates decreased. [Obaid, 2020:69-74]

The research dealt with the use of neural networks in classifying COVID-19 death rates into four cases, high, continuous on high, low, and continuous on low, and when testing the neural network used in classifying COVID-19 deaths over 853 days. The results showed a validity rate of 93.9% in the network business, and the death rate was classified as high at 96.7%, while the deaths rates were low at 95.5%. We also noted that the proportion of correct classification for high COVID-19 deaths rates when the ratio is 10% is 86%, and the proportion of correct classification for low COVID-19 deaths rates when the ratio is 10% is 91%.

2. DESCRIPTION OF THE PROBLEM

The importance of the research centers on classifying Covid-19 death rates in Iraq into four cases: a low death rate and its continuity in declining, and a high death rate and its continuity in rising to estimate the severity of the epidemic. The research limits included the numbers of Covid-19 deaths rates within 853 days, which were documented in the Statistical Bulletin of the Iraqi Ministry of Health.

2.1. Theoretical background

2.1.1. PREVIOUS STUDIES

1- Al-Rawi's study (2019), where he used neural networks by classifying groups of teachers in a set of variables to reach a model that helps in the future to place the academic staff in the correct class and compare it with evaluation. Obtained a set of criteria used in evaluating academic performance in Iraqi universities. [Al-Rawi, 2019:523-525]

2- Al-Nuaimi's study (2019) explained the importance of using neural networks in identifying audit risks that provide an objective scientific approach that enhances confidence in this profession, and he achieved this by following the descriptive and analytical approaches. [Al-Nuaimi,2019:21-23]

3- The study of Al-Ridha Anaz (2021) proposed an inverted neural network model with two hidden layers in which confirmed cases and deaths of COVID-19 in Iraq are expected. [Al-Ridha, 2021:2138-2141]

4- The study of Jaaf and Mohsen Al-Sunna (2021), where they adopted a set of feed-forward neural networks to predict the outbreak of COVID-19 in Iraq, and thus daily infections were expected by 87.6% and 82.3% for recovery, and 84.3% for deaths. [Jaff,2021:2-3]

5- The study of Sayed Ahmed and Al-Amin (2021) aimed to use artificial neural networks to study the factors affecting inflation. The study found that the exchange rate is one of the variables that most affect inflation, followed by the money supply, then the budget deficit, and finally, the domestic product. [Sayed Ahmed,2021:87-89]

This present research is unique from other papers used neural networks by transforming digital data into qualitative solutions and then studying them.

2.1.2. BUILDING A NEURAL NETWORK MODEL

To build the neural network model, the Multilayer Perception method was used to determine its weights through the SPSS statistical program in preparation for reducing the entropy error, where cross-entropy is used to determine the closeness between the actual results and the expected output, knowing that the researcher Sebastian Goldt from the University of Stuttgart confirmed that the entropy error is increasing continuously over time.

The model consists Y_i of the variables as well as the estimation model h_i in the neural network as in formula (2.1)

$$(2.1) \quad h_i = f(\sum_i w_{ji} y_i + b_i)$$

The output form equation follows formula (2.2)

$$(2.2) \quad X_k = g(\sum_j \tilde{w}_{kj} h_j + \tilde{b}_k)$$

Whereas,

g, f : Effective functions

$\tilde{w}_{kj}, w_{ji}$: Weight matrices

$\tilde{b}_k, b_y$: Sub parameters totals

FIGURE 1 represents the neural network model, located between the input and output models, and one hidden model. Through the study data, the neural network is trained to reduce errors by reducing the error squares of the function between the input and output models, where the data is divided into three groups 60% function training, 20% function testing, and 20% weight building.

During the data training process, function errors can be identified and controlled, which makes the classification highly efficient, but before the training process, the ordinary equation ((x-min) / (max-min)) is used to make the values range from zero to one, and the practical values to j of the output shown in formula (2.3)

$$(2.3) \quad O_j = \tanh(S_j) = \frac{e^{S_j} - e^{-S_j}}{e^{S_j} + e^{-S_j}}$$

Where the values of a function (2.3) range between -1 and 1, and the soft max function is used as a practical function, the output of the function with j values is as in formula (2.4)

$$(2.4) \quad O_j = \sigma(S_j) = \frac{e^{S_j}}{\sum_{k=1}^{m} e^{S_k}}$$

Where m is the number of outputs of the neural network, and through the SPSS program, the error function is calculated by cross entropy instead of the square error function, as in formula(2.5). Where m is the number of outputs of the neural network, and through the SPSS program, the error function is calculated by cross entropy instead of the square error function, as in formula (2.5)

$$(2.5) \quad E = \sum_{j=1}^{m} t_j \ln O_j$$

Whereas

m = the number of output groups.

t_j = the target values of the output parameters of the j sequence.

O_j = output values for j node values. [Zacharias,2016:18-20]

Moreover, that calculating the reverse generation in each repetition calculates the slope of the training errors as in formula (2.6)

$$(2.6) \quad \frac{\partial \mathrm{E}}{\partial \mathrm{W_{hj}}} = (\mathrm{O_j} - \mathrm{t_j})\mathrm{y_h}$$

To calculate the weights and nodes of the output model and the hidden model, we follow a formula (2.7)

$$(2.7) \quad \frac{\partial \mathrm{E}}{\partial \mathrm{W_{hj}}} = \sum_{\mathrm{j=1}}^{\mathrm{m}} (\mathrm{O_j} - \mathrm{t_j})\mathrm{y_h}\, \mathrm{W_{hi}}(1 - \mathrm{y_h})\mathrm{y_j}$$

Where each weight is updated in formula no. (2.8)

$$(2.8) \quad \Delta W_{ih} = -\gamma \frac{\partial E}{\partial w_{ih}}$$

Where w_{ih} takes on new values as shown in formula (2.9)

$$(2.9) \quad \Delta W_{ih} + \mathrm{W_{ih}} \rightarrow \Delta \mathrm{W_{ih}}$$

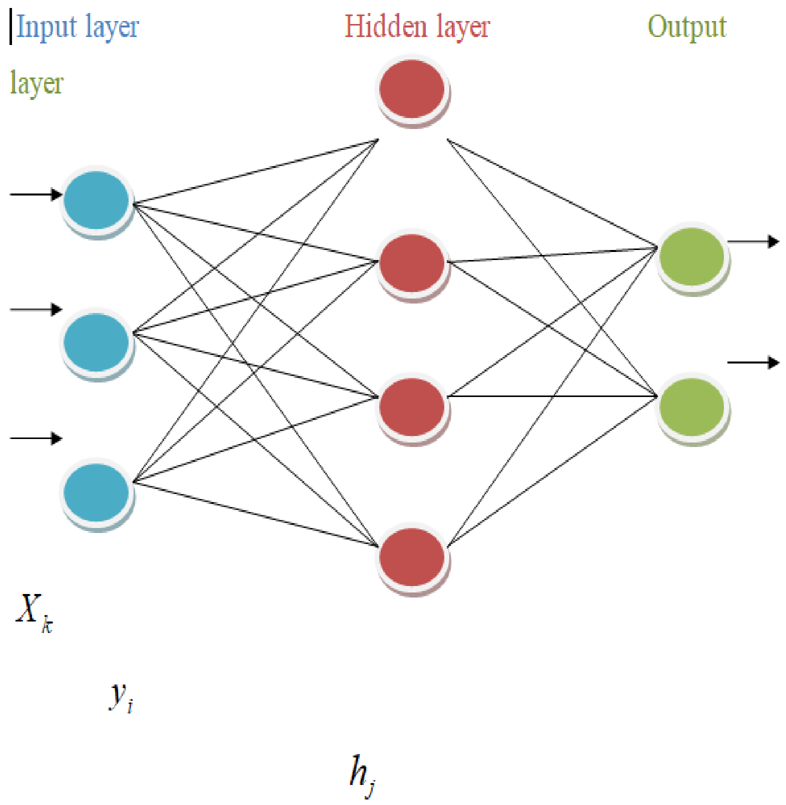

FIGURE 1. One hidden model of the neural network

3. PRACTICAL APPLICATION OF RESEARCH

Data on the deaths rates of COVID-19 in Iraq, classified by days, were collected from the official website of the Public

Health Departments of the Iraqi Ministry of Health, and after arranging and analyzing the data using Microsoft Office Excel 2007 and classifying it to low and high cases, and continue to decrease or increase during 853 days. **TABLE 1** shows the descriptive statistics of the data.

TABLE 1. Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Corona deaths in Iraq	853	.00	157.00	29.6495	27.68744
Valid N (list wise)	853				

The variables used to build the model are the independent variable representing the deaths rates for COVID-19 and the dependent variable representing four cases of deaths rates as follows:

0: Low death rates.

1: The death rate continues to decline.

2: High death rates.

3: The death rate continues to rise.

The network model was built, and its accuracy was tested using the statistical program SPSS version 20. It was based on applying the back propagation network and the method of scaling with a conjugate gradient to solve the simple linear equation by the iterative method, specifically on one layer.

TABLE 2 includes the data used to build the neural network model, as it shows the deaths rates classified according to 853 days, where the classification was done using 586 days, or 68.7% of the data, to train the neural network from the total data, in order to reach the correct classification of deaths rates. As for the second part of the data, 267 days, or 31.3% of the total data, were used to perform the necessary tests to recognize the significance of the neural network, and the last part represents the weights. In this part, the neural network weights used to classify death rates by days were calculated, noting that no day was excluded, as all death rates classified by days were included in this study. **TABLE 3** shows one hidden model of the neural network, which contains only five units to access the output with a hyperbolic tangent function, and we also note that there is one dependent variable, which contains 4 cases of classification with a Soft max function and a Cross entropy error function. **FIGURE 2** shows the diagram of the input node and six hidden nodes in addition to the output nodes, as these nodes represent the rise and fall of deaths, and each of them continues to do so.

TABLE 2. Cases Processing Summary

		N	Percent
Sample	Training	586	68.7%
	Testing	267	31.3%
Valid		853	100.0%
Excluded		0	
Total		853	

TABLE 4 shows the results of applying the neural network, where the expected entropy error appears in training the network by 82.003, and this is due to the large number of daily observations within 853 days. The entropy error decreased to 35.562 when testing the network, and the percentage of false prediction of deaths was based on the training and testing sample at 4.9% and 6.1%, respectively.

TABLE 3. Network Information

Input Layer	Covariates	1	COVID-19 deaths in Iraq
	Number of Units[a]		1
	Rescaling Method for Covariates		Standardized
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1[a]		5
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	COVID-19 deaths rates
	Number of Units		4
	Activation Function		Softmax
	Error Function		Cross-entropy

a. Excluding the bias unit

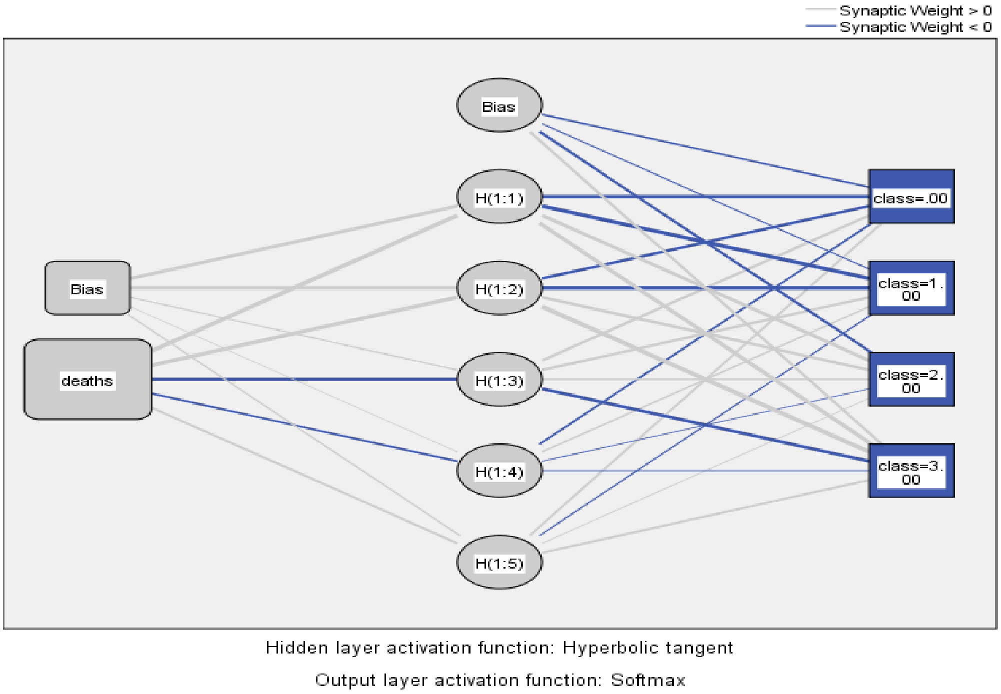

FIGURE 2. Hidden layer activation function

TABLE 5 shows the interlocking weights that were calculated using the data under study, and **TABLE 6** shows the percentage of correct classification and classification errors, where the training side shows the highest percentage of correct classification in the case of the persistence of the high deaths rates, which is 100%, followed by the percentage of stability in Low deaths rates which is 96.3%, As for the experimental aspect, it is shown that the highest percentage of correct classification in the case of high deaths rates stability is also 100%, followed by the percentage of stability in low deaths rates which is 98.6% so that the percentage of correct classification at each start for high or low deaths rates is 18.8%, 0%, respectively from the training side and, 14.3%, 0% respectively from the testing side. **FIGURE 3** shows a Faulty neural network prediction, where values greater than 0.5 show the correct prediction, while the error is less than 0.5. We note that a small percentage of low death rates are classified as high, and the same is true for persistently low death rates. **FIGURE 4** shows us the ROC curve, which shows the sensitivity measures for determining the best classification. **TABLE 7** shows the area under this curve, where the death rates were classified as low to 95.5% and continue to decrease at 99.5%. As for the death rate is classified as high is 96.7%, and the death rate is classified as continuing to rise at 99.7%.

TABLE 4. Model Summary

Training	Cross Entropy Error	82.003
	Percent Incorrect Predictions	6.1%
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:01.14
Testing	Cross Entropy Error	35.562
	Percent Incorrect Predictions	4.9%

Dependent Variable: COVID-19 deaths rates

a. Error computations are based on the testing sample.

TABLE 5. Parameter Estimates

Predictor		Predicted								
		Hidden Layer 1					Output Layer			
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	[class=.00]	[class=1.00]	[class=2.00]	[class=3.00]
Input Layer	(Bias)	3.889	4.160	.241	.041	.265				
	deaths	5.578	5.149	-2.966-	-.447-	.618				
Hidden Layer 1	(Bias)						-.360-	-.200-	-1.247-	1.341
	H(1:1)						-3.648-	-4.828-	3.528	4.698
	H(1:2)						-2.752-	-5.399-	2.806	5.632
	H(1:3)						1.540	1.581	.354	-3.448-
	H(1:4)						-.399-	.354	-.045-	-.187-
	H(1:5)						.589	-.240-	.089	.684

TABLE 6. Classification

Sample	Observed	Predicted				
		Low death rate	The death rate continues to decline	high death rate	The death rate continues to rise	Percent Correct
Training	Low death rate	3	13	0	0	18.8%
	The death rate continues to decline	7	180	0	0	96.3%
	high death rate	0	0	0	16	0.0%
	The death rate continues to rise	0	0	0	367	100.0%
	Overall Percent	1.7%	32.9%	0.0%	65.4%	93.9%
Testing	Low death rate	1	6	0	0	14.3%
	The death rate continues to decline	1	72	0	0	98.6%
	high death rate	0	0	0	6	0.0%
	The death rate continues to rise	0	0	0	181	100.0%
	Overall Percent	0.7%	29.2%	0.0%	70.0%	95.1%

Dependent Variable: COVID-19 deaths rates

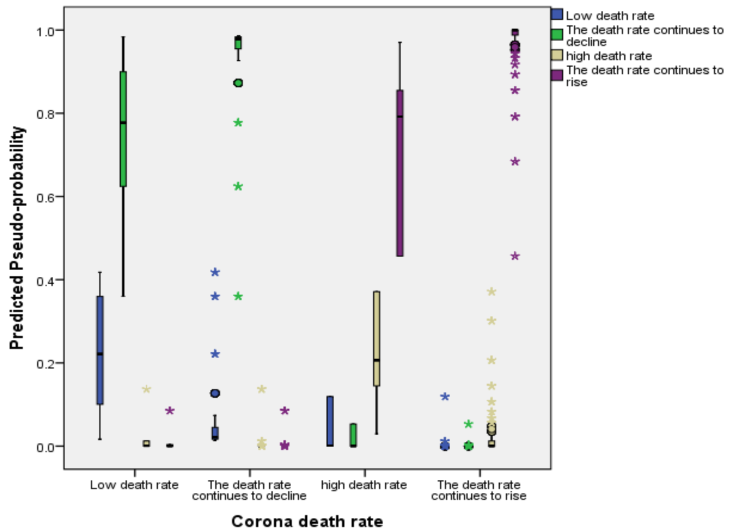

FIGURE 3: Faulty prediction of a neural network

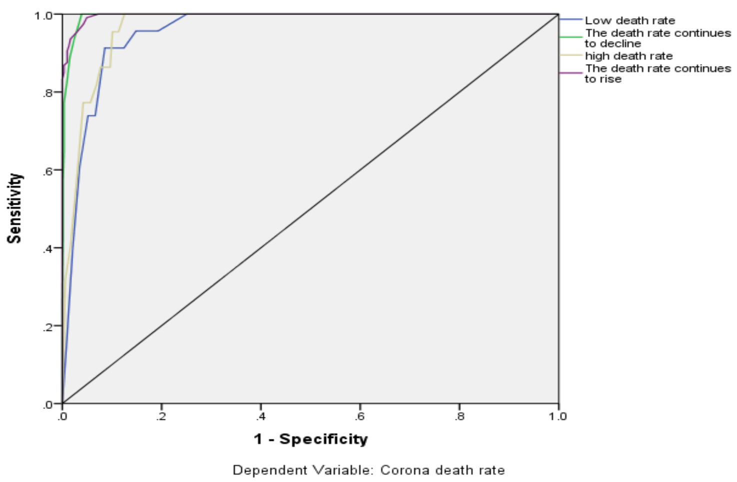

FIGURE 4: The ROC curve (sensitivity measures for determining the best classification)

TABLE 7. Area Under the Curve

		Area
COVID-19 deaths rates	Low deaths rates	.955
	The death rates continue to decline.	.995
	high deaths rates	.967
	The death rates continue to rise.	.997

FIGURE 5 represents the correct classifications of COVID-19 deaths rates obtained from the neural network application against the correct classifications of COVID-19 deaths rates that appear by chance, the correct classification percentage of low COVID-19 deaths rates when the percentage is 10% is 91%, While the percentage of correct classification when COVID-19 deaths remain low is 65%, and that is when the percentage is 20%. Also, the correct classification rate for high COVID-19 deaths is 86% when the ratio is 10%, and the correct classification rate when COVID-19 deaths remain high is 78% when the ratio is 50%.

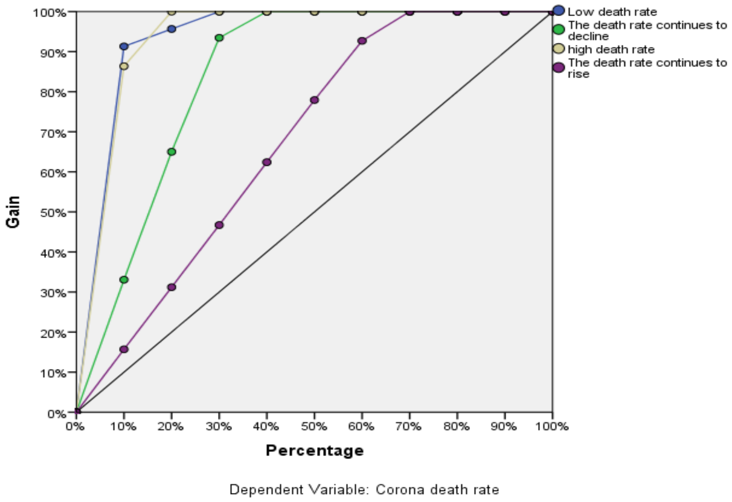

FIGURE 5: Represents the correct classifications of COVID-19 death rates.

FIGURE 6 shows the use of part of the data to benefit from using the model compared to not using it. At 10%, the value of the function is 9.1 for low death rates, 3.3 for continuously decreasing death rates, 8.6 for high death rates, and 1.5 for continuously increasing death rates.

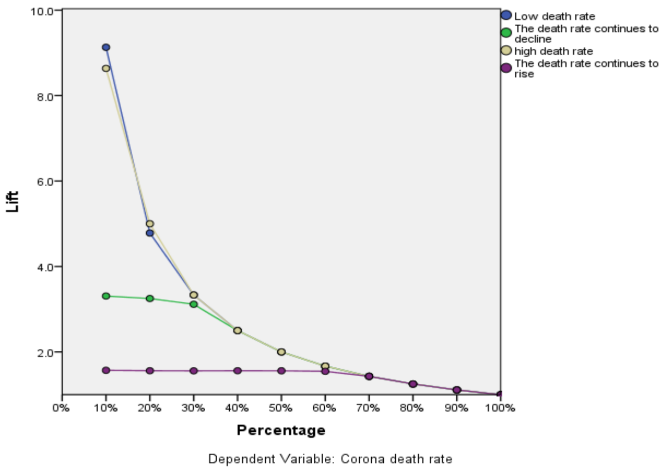

Figure 6: Shows the use of part of the data to benefit from using the model compared to not using it.

4. CONCLUSIONS

The current study concluded group of conclusions can be summarized as in below:

1 - Since neural network models train data, they are considered robust models.

2- The idea of using a neural network is to train and simulate it with accurate data, thus becoming more assertive in obtaining the estimated values with the lowest possible average error.

3 – It has been noted that fluctuation in the death rates of Covid-19 in Iraq, between an increase and a decrease during the 853 days since the beginning of the epidemic.

4 - When testing the neural network used to classify COVID-19 death rates within 853 days, as it was displayed on the network, which in turn was able to classify deaths rates into cases on which the network was trained in the training phase, where the accuracy rate in the network's work was 93.9%.

5- The percentage of death rates classified as high was 96.7%, while the percentage of those classified as low 95.5%.

6- The correct classification of high COVID-19 death rates when the ratio is 10% is 86%, and the correct classification of low COVID-19 death rates when the ratio is 10% is 91%.

7- It is possible to use the neural network model under study to predict future COVID-19 death rates and prepare for COVID-19 virus mutations that society may be exposed to.

ACKNOWLEDGMENTS

Finally, I thank the Iraqi Ministry of Health for providing the data in its daily statistical bulletin on COVID 19.

REFERENCES

1. Al-Nuaimi, Anmar Hani, (2019), "Using Neural Networks to Identify Audit Risks / A Case Study in the Federal Office of Supreme Audit," Higher Diploma Research, Department of Accounting, College of Administration and Economics, University of Mosul.

2. Al-Rawi, Omar Fawzi (2019), "Building a model to identify teachers' evaluations through a set of variables using neural networks," Tikrit Journal of Administrative and Economic Sciences / Volume 15 / Issue 48 / Part 2/College of Administration and Economics/ Tikrit University/ Iraq.

3. Al-Ridha, Moatasem Yaseen, Anaz, Ammar Sameer, Al-Nima, Raid Rafi Omar,(2021), "Expecting confirmed and death cases of covid-19 in Iraq by utilizing back propagation neural network", Bulletin of Electrical Engineering and Informatics Vol. 10, No. 4, August 2021, pp. 2137~2143 ISSN: 2302-9285, DOI: 10.11591/.v10i4.2876.

4. Al-Zawahra Saado (2021), "Artificial Intelligence in Combating Corona: Research: Current and Future," Atom and Development Bulletin, Volume 33, Issue 2.

Jaff, Ahmed, Mohsen, Thakir Muhammad, Al-Jumaili, Diaa, 5.Al-Loughani, Muhammad (2021), "Integration of data science and neural network-based modeling to predict the outbreak of Covid 19 in Iraq," Journal of Biomedical Informatics, 103766, 118.

6. Obaid, Duha Saad, Hussein, Heba Ali (2020), “The Crisis of Dealing with the Corona Pandemic...Iraq as a Model,” Hammurabi Magazine, Issue 34-33, Year 8.

7. Sayed Ahmed, Ibrahim Mohamed Ibrahim, Al-Amin, Rajaa Mustafa Salih, Abbas, Magdy Abd El-Ilah Mohamed (2021), “The use of artificial neural network models in regression analysis to study the factors affecting inflation in Sudan in the period from 1991-2018”, Journal of Human and Natural Sciences, HNSJ Volume 2, Issue 6.

8. Zacharias, Nick Z.,(2016) "Predicting Student Academic Performance in blended learning using artificial neural networks", International Journal of Artificial Intelligence and Applications (IJAIA), Vol. 7, No. 5, September.