ISSN (1681- 6870)

A Comparison of the Logistic Regression Model and Neural Networks to Study the Determinants of Stomach Cancer

Dr. Haifa Taha Abed
haefaa_adm@uomustansiriyah.edu.iq
Al-Mustansiriyah University - Collage of Administration and Economics Department of Statistics - Baghdad - Iraq

Dr. Duraid Hussein Badr
duraidhussain@yahoo.com
Department of Statistics, Administration and Economics collage, University of Basrah - Iraq

Received 16/2/2020 **Accepted 28/6/2020**

Abstract: *The logistic regression model and the neural network model are among the best models in the bilateral response data and the classification of different medical conditions. Therefore this study addressed the comparison between the two models using statistical classification criteria (model accuracy, model sensitivity, model specificity, the model's false alarm rate, the area under the Receiver Operating Characteristic (Roc) curve, Wrong classification rate). After applying these criteria to the study data, we concluded that the neural network model is better than the logistic regression model, as it was reached through the final reconciliation of both models that the factor of the method of diagnosing stomach cancer has the obvious effect on the classification of the patient's condition, and this was confirmed by the relative importance of the factors studied using the neural network model, which showed that this factor reached its relative importance 100%, which is a very large percentage compared to other variables.*

Keywords: Artificial neural networks; Logistic regression; Classification; Model comparison; Model evaluation.

مقارنة بين نموذج الانحدار اللوجستي والشبكات العصبية لدراسة محددات سرطان المعدة

م.د. هيفاء طه عبد
haefaa_adm@uomustansiriyah.edu.iq
الجامعة المستنصرية - كلية الادارة والاقتصاد- قسم الاحصاء – العراق - بغداد

م.د. دريد حسين بدر
duraidhussain@yahoo.com
جامعة البصرة - كلية الادارة والاقتصاد - قسم الاحصاء العراق - البصرة

المستخلص

يعد انموذج الانحدار اللوجستي وانموذج الشبكات العصبية من افضل النماذج في بيانات الاستجابة الثنائية وفي تصنيف الحالات الطبية المختلفة. لذا تناولت هذه الدراسة المقارنة بين الانموذجين باستخدام معايير التصنيف الاحصائية (دقة النموذج، حساسية النموذج، خصوصية النموذج، معدل الانذارات الخاطئة للنموذج، المساحة تحت منحنى Roc، معدل خطأ التصنيف) وبعد تطبيق هذه المعايير على بيانات الدراسة توصلنا الى ان انموذج الشبكات العصبية افضل من انموذج الانحدار اللوجستي، كما تم التوصل من خلال التوفيق النهائي لكلا النموذجين الى ان عامل طريقة تشخيص مرض سرطان المعدة له الاثر الواضح في تصنيف حالة المريض، وهذا اوكد من خلال الاهمية النسبية للعوامل المدروسة باستخدام انموذج الشبكات العصبية التي بينت ان هذا العامل بلغت اهميته النسبية 100% وهي نسبة كبيرة جدا قياسا بالمتغيرات الاخرى .

الكلمات المفتاحية: الشبكات العصبية الاصطناعي، الانحدار اللوجستي، التصنيف، مقارنة النموذج ، تقييم النموذج.

1. Introduction

Logistic regression analysis examines the relationship between the dependent variable and a set of independent (illustrative) variables. Logistic regression is used when the dependent variable

ISSN (1681- 6870)

has only two values, such as 0 and 1 or Yes and No. Although the data type used for the dependent variable is different from multiple regression[5], the practical use of the procedure is similar. The binary logistic regression and the polynomial logistic regression are calculated for both digital and categorical independent variables. It reports on the regression equation as well as the quality of fit, probability ratios, confidence limits, probability, and deviation. It performs a comprehensive analysis of the remains including the remaining diagnostic reports. It can perform an independent variable subset search, find the best regression model with the fewest number of independent variables, and provide ROC curves to help determine the best classification breakpoint.

2. The concept of logistic regression

Logistic regression is one of the regression models in which the relationship between the dependent variable is (Dependent) Y and the independent (interpreted) variables X are nonlinear, often taking the model's response function The response function often takes the form S.[16]:

Let $X = x_1, x_2, \dots, x_q$

$$\theta_g = \begin{pmatrix} \theta_{g1} \\ \theta_{g2} \\ . \\ . \\ . \\ \theta_{gq} \end{pmatrix}$$

The logistic regression model is given by the equations:

$$ln\left(\frac{q_g}{q_1}\right) = \ln\left(\frac{q_g}{q_1}\right) + \theta_{g1}x_1 + \theta_{g2}x_2 + \cdots + \theta_{gq}x_q$$

$$= ln\left(\frac{q_g}{q_1}\right) + X\theta_g \tag{1}$$

Where $q_g = \Pr(Y = g/X)$

The term $ln\left(\frac{q_g}{q_1}\right)$ becomes zero and drops out ,when the prior properties are assume equal , the θ are population regression coefficients that are to be estimated from the data, In terms of probabilities, they are nonlinear. Then

The corresponding non-linear equations are[9]:

$$q_g = Prob(Y = g/X) = \frac{e^{x\theta_g}}{1+e^{x\theta_2}+e^{x\theta_3}+\cdots+e^{x\theta_G}} \tag{2}$$

Since $e^{x\theta_1} = 1$ because all regression coefficient is zero.

Since $e^{X\theta}$ may be represent as fallow:

$e^{X\theta} = e^{x_1\theta_1+x_2\theta_2+\cdots+x_q\theta_q}$

2.1.The Logit and Logistic Transformations

The logistic model is classified as nonlinear models that can be converted into linear models These models are called implicitly linear models. Statisticians usually tend to linearly convert these models to remove the curves of their parameters due to the negative effect of these curves in the case of their presence on the properties of the least-squares estimators of them and then the

ISSN (1681- 6870)

predicted response values, where these estimators are often biased and not distributed Naturally, its variations are not the smallest possible, which makes the results of the tests misleading, and there are several transformative procedures, including the logarithmic function that works to combine and compress high values in the data and expand and individualize very small values in them[13].

The Logit, the natural logarithm of Odds, is a linear structure of the explanatory variables, and the logistic regression model can be expressed in a linear relationship with the probability log, according to the following equation[6]:

$$L = Logit = Log\left(\frac{q}{1-q}\right) = \gamma + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_q x_q \quad (3)$$

$$-\infty < Logit < \infty$$

Log function has a major function, which allows the application of linear regression when analyzing relationships For data with two dependent variables, hence the logistic regression indicates to the regression models that include the log as a dependent variable in the equation, and calculates the amount of the change in the logarithm of the weighting coefficient of the dependent variable and not in the dependent variable itself as in an analysis Linear regression[5].

2.2.Estimation of logistic regression coefficients

To estimate the logistic regression coefficients, the Maximum Likelihood Method is used, and it is one of the most appropriate methods for all linear and non-linear models, and the Maximum probability method is an iterative method that depends on the frequency of mathematical operations several times until the best estimate of the coefficients is reached.

Let

$$\aleph_{gi} = prob(Y = g/X_i)$$

$$= \frac{e^{X_i\theta_g}}{e^{X_i\theta_1 + X_i\theta_2 + \cdots + X_i\theta_G}}$$

$$= \frac{e^{X_i\theta_g}}{\sum_{r=1}^{G} e^{X_I\theta_r}}$$

Where the $\aleph_{gi}$ is the probability of linear equations in the logit.

The Likelihood for a sample of N observation is:

$$L = \prod_{i=1}^{N}\prod_{g=1}^{G} \aleph_{gi}^{y_{gi}} \quad (4)$$

Where:

$$y_{gi} = \begin{cases} 1 & is\ one\ of\ the\ i^{th}\ observation\ is\ in\ outcome\ g \\ 0 & otherwise \end{cases}$$

Hence $\sum_{g=1}^{G} y_{gi} = 1$, The Logliklihood is:

ISSN (1681- 6870)

$$L = Ln(L) = \sum_{i=1}^{N}\sum_{g=1}^{G} y_{gi} Ln(\aleph_{gi})$$

$$= \sum_{i=1}^{N}\sum_{g=1}^{G} y_{gi} Ln\left[\frac{e^{X_i\theta_g}}{\sum_{r=1}^{G} e^{X_i\theta_r}}\right]$$

$$= \sum_{i=1}^{N}\left[\sum_{g=1}^{G} y_{gi}X_i\theta_g - Ln\left(\sum_{g=1}^{G} e^{X_i\theta_g}\right)\right] \tag{5}$$

To obtain estimates of the parameters that maximize the goal, and when the target function is derived relative to For the parameters to be estimated and made equal to zero . The resulting likelihood equations are[9]:

$$\frac{\partial L}{\partial \theta_{gm}} = \sum_{i=1}^{N} X_{mi}\left(y_{ig} - \aleph_{jg}\right)$$

Where g=1,2,…,G and m=1,2,…q. and When g=1, all coefficient are zero.

Since the nature of parameters is nonlinear, we use the Newton-Raphson method to solve these equations. the information matrix $I(\theta)$, which is formed from second partial derivatives. May be made form this method. The elements of the information matrix are given by[11]:

$$\frac{\partial^2 L}{\partial \theta_{jm}\partial \theta_{jm}} = -\sum_{i=1}^{N} X_{mi}X_{mi}\aleph_{jg}\left(1 - \aleph_{jq}\right)$$

$$\frac{\partial^2 L}{\partial \theta_{jm}\partial \theta_{jm}} = \sum_{i=1}^{N} X_{mi}X_{mi}\aleph_{jg}\aleph_{jg} \tag{6}$$

2.3. Check the suitability of the model as a whole

The first step in the process of appropriately assessing the model that has been reconciled is usually an indicative evaluation Variables as a whole in the model, determine whether the explanatory variables as a whole in the model are related Statistically significant with the dependent variable or not. There are several important measures that help in evaluating the final model[8] that has been reconciled to the data, which are $R2$ statistics, classification tables in addition to the ROC curve analysis[7]

2.4. statistics R^2

In the logistic regression model, substitute the coefficient of determination, which is used to determine the extent The suitability of the proposed regression models for the study data with Nagelkerke matchmaking statistics, Cox & snell R^2 which have the R^2 statistic aim is the same multiple linear regression[7].

ISSN (1681- 6870)

$$R^2 = 1 - \left(\frac{L_0}{L_1}\right)^{2/N} \tag{7}$$

$$R_t^2 = 1 - (L)^{\frac{N}{2}} \tag{8}$$

$$\breve{R}^2 = \frac{R^2}{R_t^2} \tag{9}$$

Where: L_0 is Maximum Likelihood Function Contains only the fixed term.
: L_1 IS Maximum Likelihood Function contains all explanatory variables.
: N , Sample size.

3. Classification table

A table that shows the number of observation cases that have an attribute and the number of observation cases It does not possess that quality, as opposed to the number of expected cases that possess the quality and the number of expected cases That do not have that adjective, so the table shows the number of instances that were correctly classified And the number of cases that were classified incorrectly[7].

The basic idea of the analysis is to expect the cases to be classified correctly according to a criterion, because that gives evidence that the model matches the observation data, and table (1) shows the general form of the classification table.

Table (1): Classification table

Classification		Expected		Total
		Positive	Negative	
Observation	Positive	True positive TP	False negative FN	R
	Negative	False positive FP	True negative TN	-R
		S	-S	1

3.1.Receiver Operating Characteristic(ROC) Curve Analysis[1][3]

The area under the ROC curve, which ranges from zero to one correct, is given a measure of the model’s ability to distinguish between the states that possess the trait being examined and those that do not have the trait, and it is considered one of the best measures of classification accuracy. The area under the shell diameter is equal to 0.5. The higher the model's discriminatory power and the curve away from the shell diameter towards the upper left corner, the greater the area under the ROC curve until it reaches the value of one true. Which means complete discrimination of cases.

4. Artificial Neural Networks

Artificial neural networks are a structure with a parallel information structure, which consists This structure of processing units that process information and are called neurons or elements calculation, the signals between the neurons pass through the binding lines, and each neuron represents a local memory . Each connecting line is also attached to a weight (a specific numerical number that strikes with the signals entering the neuron)[2] and then applies to each neuron an activation function on the network input, which represents the sum of the weighted input signals to determine the resulting output signal.

Figure (1) shows the smart neuron k which has n of inputs $X_1, X_2, \ldots, X_n$ and each Entrance is weighted before it reaches the neuron k through the link weights $Z_1, Z_2, \ldots, Z_n$ also, the neuron has an activation or activation function that determines the output of the Y_K neuron and there are different

ISSN (1681- 6870)

types From the activation functions, but the exponential function sigmoid, which takes the values between o to 1, used a lot and is by the following formula[10]:

$$Y_k = \frac{1}{(1 + e^{net_k})} \tag{10}$$

Where: Y_k: It represents the neuron output k.

net_k : $\sum_l z_{lk} X_L$, Which represent Which represented the sum of the weights of the inputs to the neuron, and the aim of the activation function is to confirm the limited outputs of the neuron.

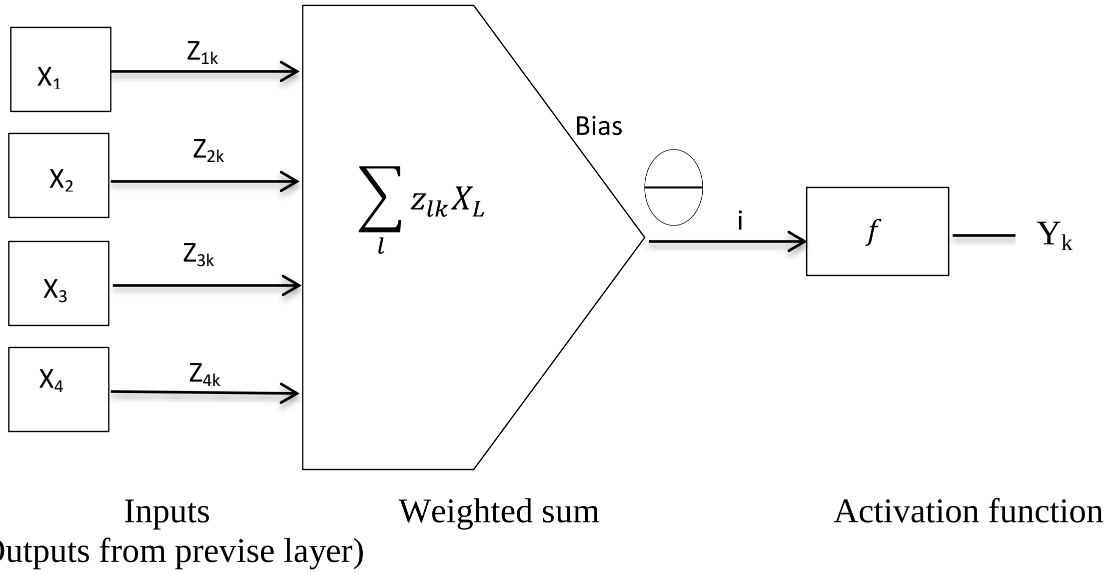

Figure (1): The neuron or treatment component

Each neuron network has its architecture, the number of neurons, and its method of interconnection. There are different types of network architectures and each type is used for a specific purpose, but the front-feeding algorithm of the neuronal network is famous and widely used.

4.1.Neural networks with forwarding feeding

The neural networks with the forward nutrition are divided between the input layer, the hidden layer, and the output layer. Each layer contains a different number of neurons, and each neuron in the input layer is bound to all the neurons in the hidden layer and each neuron in the hidden layer is associated with all the neurons present in the output layer as shown in figure (2).

This type of neural network has a forward path in feeding information and has no feedback track The number of neurons in the input layer and the output layer is determined by the number of inputs and outputs of the problem. Each neuron receives inputs from the rest of the neurons through the balanced links and after processing them, inputs to the following neurons and the activation level of the neuron in the input layer is determined according to the response to the inputs received and the activation level is a function of the weights accompanying these links.

ISSN (1681- 6870)

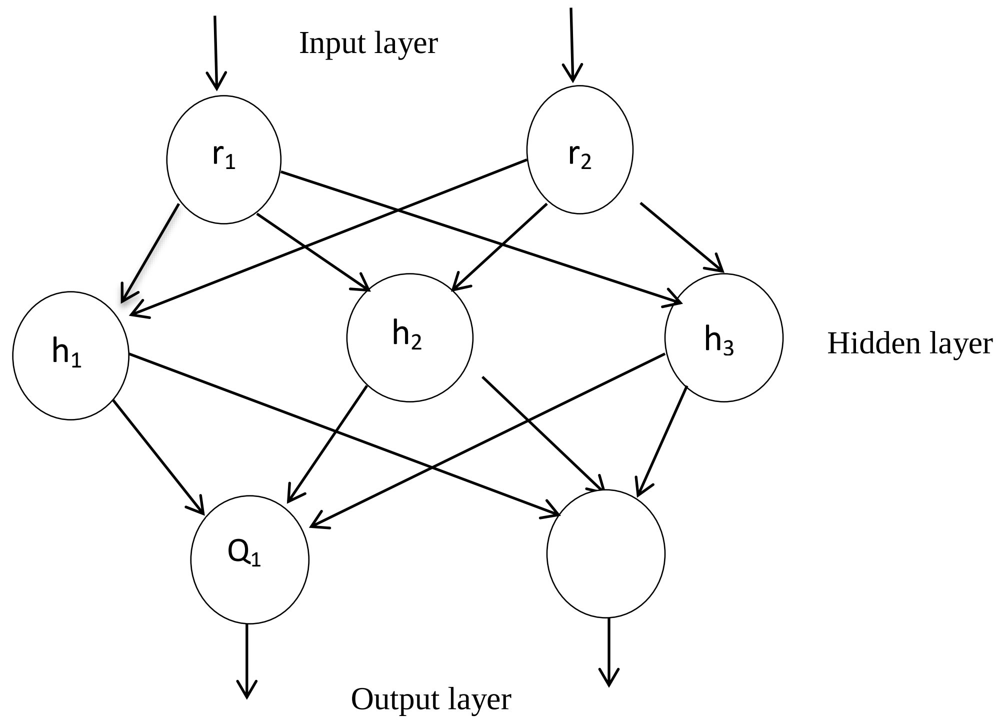

Figure (2): A typical structure of the frontal neural network

4.2. Retroactive learning

The reactionary learning algorithm is a supervised learning method that is used to train neural networks Front feeding In this way of learning, inputs are handled through the neural network, the outputs are compared with the real outputs, and errors, differences, are used to update the link weights to reach the lowest sum of squares[14].

Figure (2) is used to illustrate the way of retroactive learning, if they are values $X_l, l = 1,2, \dots, N$, The input vector and the output values required for these inputs are information and were ,N The size of the trained sample and the goal of the learning process Calculating the total weights Z_{lk}, Also, for neural network links that reduce the sum of squares of errors between the real and network outputs Q_J to the lowest possible level, the learning objective can be written as follows[15][17]:

$$\min \in = \sum_{J} \left(S_J - Q_J\right)^2 \tag{11}$$

The continuous change of weights helps to reach the smallest value of the error function. The reactionary learning process includes several steps, as follows:

1. Randomly change weights.
2. The value vector for the XL training is as an input vector to the input layer and by using an activation function, for example, the exponential function, the output of each neuron will be calculated as follows:

$R_k = f(net_k)$

ISSN (1681- 6870)

$$net_k = \sum_{l} Z_{lk} X_l \quad , l = 1,2, \dots , L, \quad k = 1,2, \dots , K$$

3. The neuron outputs in the hidden layer are used to calculate the outputs of the neurons in the output layer, which are the outputs of the neural network.

$$Q_j = f(net_j)$$

$$net_j = \sum_{k} Z_{kj} R_k \quad , \quad j = 1,2, \dots , J$$

4. The weights of the neural network are updated to reduce the error value calculated for each neuron of the output layer. First, the weights associated with the neurons in the hidden layer and the output layer must be updated using the following formula:

$$Z_{kj}(t+1) = Z_{kj}(t) + \Delta Z_{kj}(t) \tag{12}$$

$$\Delta Z_{kj}(t) = \delta \varphi_j R_k + \alpha Z_{kj}(t-1) \tag{13}$$

$$\varphi = (S_k - Q_k) \frac{\partial Q_k}{\partial net_k} \tag{14}$$

Where: $Z_{kj}(t+1)$, Total weights in the iteration (t+1).

$Z_{kj}(t)$, Total weights in the iteration(t).

$\frac{\partial Q_k}{\partial net_k}$, It is the derivative of the activation function used in the output layer.

δ *and* α, Constants are between 0,1, They are used to control and improve the efficiency of the training process δ Learning rate and α learning moment.

5. After updating the weights that link the neurons in the hidden layer to the output layer, the weights that link the inputs to the hidden layer are updated as follows:

$$Z_{lj}(t+1) = Z_{lj}(t) + \Delta Z_{lj}(t) \tag{15}$$

Where:

$$\Delta Z_{lj}(t) = \delta \sigma_j X_l + \alpha Z_{lj}(t-1) \tag{16}$$

$$\varphi_j = \frac{\partial R_j}{\partial net_j} \sum_{k} \varphi_k Z_{lk} \tag{17}$$

Where the term $\frac{\partial R_j}{\partial net_j}$, It is the derivative of the activation function of the neurons in the hidden layer.

5. The practical frame

In this study, we will apply logistic regression and neural networks, and we will evaluate the classification accuracy based on some classification indicators in addition to the Roc curve, then we fit the final model for both applications to determine the most important factors affecting increase

ISSN (1681- 6870)

deaths stomach cancer, and we will use data on stomach cancer From the Iraqi Cancer Center at the Ministry of Health 2017.

5.1. Logistic regression model

We are interested in studying gastric cancer mortality, which is the dependent variable in the analysis that takes two values (1 if the person is deceased, 2 if the person is alive), and Table 1 shows the classification table for the logistic regression model.

Table (2) classification of logistic regression model

Is the person dead?		Prediction		Total
		No	Yes	
Observation	No	508	16	524
	Yes	39	151	190
	Total	547	167	714
a model Accuracy		92.29%		
model sensitivity		90.42%		
a model Privacy		92.87%		
The model's false alarm rate		7.13%		
The area under the Roc curve		91.11%		
Wrong classification rate		7.70%		

We note from Table (2) that the logistic regression model was able to accurately classify up to 92.29%, as it was able to classify 508 items out of 524 items, it was correctly classified by people still alive at a rate of 96.94%, while the model managed to classify 151 items Of the 190 people who died of stomach cancer, a correct classification with a rate of 79.47%, and based on the sensitivity of the model that reaches 90.42% indicates that the model can correctly predict depending on the explanatory variables into the model by 90.42% for people who died of stomach cancer, either Regarding the specificity of the model 92.87%, it indicates that the model You will prediction correctly based on the studied explanatory variables at 92.87% for people still alive. The model has also predicted people who considered them dead even though they are still alive at a rate of 7.13%, this is explained by the value of The model's false alarm rate for the model that reaches 7.13%, and the classification error rate the model 7.70%.

Through Figure (3), which shows the Roc curve in relation to the sensitivity of the model, the area under the curve is 91.11%, and it must be mentioned that the area under the curve is more than 50% the better the classification.

ISSN (1681- 6870)

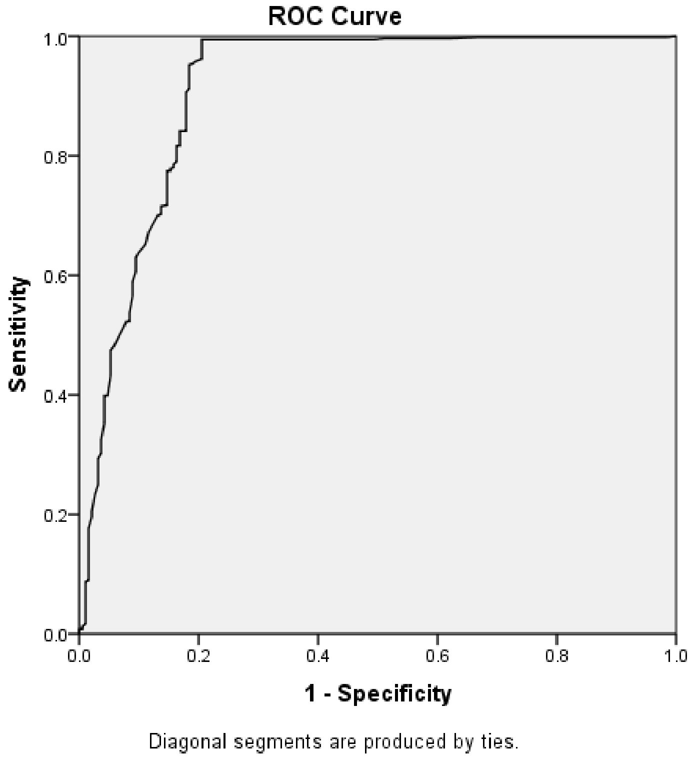

Figure (3): shows the Roc curve with respect to the sensitivity and specificity of the logistic regression

5.2. Neural network model

Table (3) shows the classification according to the neural network model

Table (3) classification of the neural networks model

Is the person dead?		**Prediction** No	**Prediction** Yes	**Total**
Observation	No	519	5	524
	Yes	37	153	190
	Total	556	158	714
a model Accuracy		94.12%		
model sensitivity		96.83%		
a model Privacy		93.34%		
The model's false alarm rate		6.65%		
The area under the Roc curve		93%		
Wrong classification rate		5.88%		

Through Table No. (3), we note that the neural network model has reached an accuracy of 94.12% and that the model was able to correctly classify 519 items out of 524 items from people who are still alive with a rate of 99.1%, and the group of people who died A correct classified, 153 items were classified out of 190 items, with a rate of 80.53%. As for the sensitivity of the model, which is 96.83%, it indicates that the model, and based on the studied explanatory variables, was able to correctly predict up to 96.83% for people who died of stomach cancer, while the specificity of the model was 93.34%, It indicates that the model was able to classify correctly depending on the studied explanatory variables is 93.34% for people who are still alive, and the rate of false alarms for the model was 6.65%, meaning that the model can predict persons dead but that they are still alive and that reaches 6.65%, while the wrong classification rate of the model has been It reached 5.88%.

As for the area under the Roc curve, it reached 93%, which is greater than 50%. The higher the area under the Roc curve, the better, as shown in Figure (4), which shows the Roc curve concerning the sensitivity of the model.

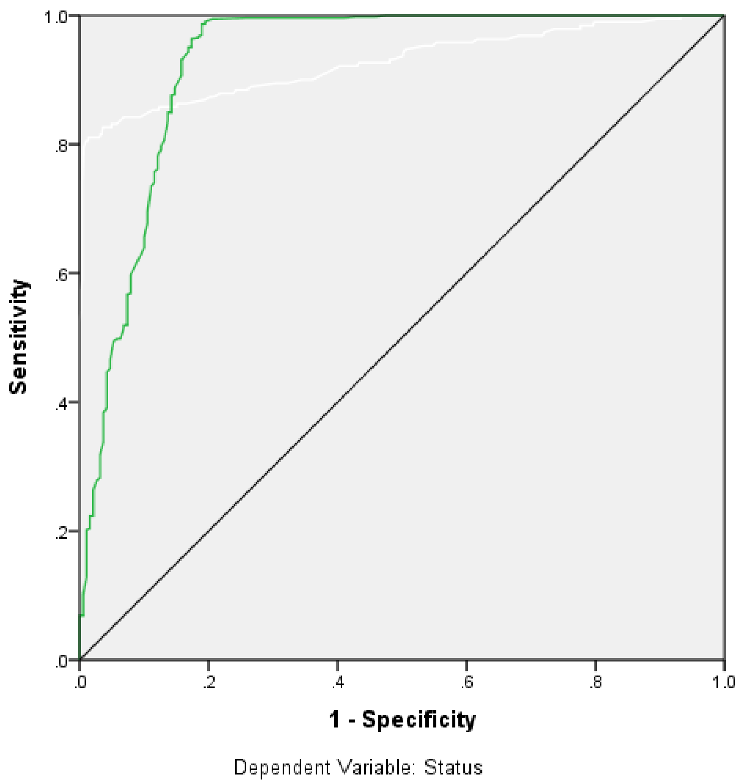

Figure (4) Roc curve, sensitivity and specificity of the neural network model

5.3. Comparison of models

After examining the classification according to the logistic regression model and the neural network model, it is necessary to compare the two models based on criteria (model accuracy, model sensitivity, model specificity, model false alarms rate, area under the Roc curve, wrong classification rate), and it turns out that the neural network model is more efficient than the logistic regression model, as the classification according to the neural network model was more accurate and more effective. Table (4) shows this according to the classification criteria.

Table (4): Classification criteria for logistic regression model and neural network model

Classification criteria	Logistic regression model	Neural network models
a model Accuracy	92.29%	94.12%
model sensitivity	90.42%	96.83%
a model Privacy	92.87%	93.34%
The model's false alarm rate	7.13%	6.65%
The area under the Roc curve	91.11%	93%
Wrong classification rate	7.70%	5.88%

ISSN (1681- 6870)

We note from Table (4) that the accuracy criterion for the model according to the model of neural networks is 94.12%, which is greater than the accuracy of the logistic regression model, which is 92.29%, and this gives preference to the model of neural networks, the more accurate the model is the best, the criterion for sensitivity of the model of neural networks was equal to 96.83% While the sensitivity of the logistic regression model is equal to 90.42%, this gives preference to the model of neural networks depending on the sensitivity of the model because if it is higher was better, the specificity of the neural network model is equal to 93.34%, which is greater than the specificity of the logistic regression model which is 92.87%, and this gives preference to the model For neural networks given that the specificity of the model the higher the specificity, the better, while the rate of false alarms for the model of neural networks is 6.65%, while the rate of false alarms for the logistic regression model 7.13%, this gives preference to the model of neural networks because the rate of false alarms the less the better. As for the area under the Roc curve, the neural network model was 93%, and the logistic regression model was 91.11%. This gives preference to the neural network model because the area under the Roc curve the higher the better, and finally the wrong classification rate, the neural network model was the 5.88% and the logistic regression model is 7.70%, and this gives preference to the neural network model at the expense of the logistic regression model. The lower the error rate of classification, the better.

5.4. Fit the final model

- **First: Fit the logistic regression model**

To verify the suitability of the model, χ^2 was used, which is statistically significant at the level of significance 0.05, which means that the statistical model was fitted, which includes five variables (class, age, profession, disease diagnosis method, degree disease), statistically significant in a decrease the value of the logarithm of the maximum likelihood of the model that includes a intercept term only from 707.880 to 247.742 in the case of the model that contains the explanatory variables studied in the model, and this also means that the model that includes the explanatory variables explains the classification of the condition to the deceased and is still alive, and he predicts that is better than the model Which does not contain explanatory variables, as shown in Table (5).

Table (5): Fit the logistic regression model

Model	-2 Log Likelihood	Likelihood Ratio Test		
		χ^2	df	Sig
Intercept only	707.880			
Final	247.742	460.139	5	0.000

It is also clear from Table No. (6) that the studied explanatory variables contributed to the interpretation of about 69% using the R2 Nagelkerke coefficient, and 47% using the R2 Cox & Snell coefficient, from the variables in the patient's case (dependent variable), and this indicates that there are a number Of the variables contributing to the interpretation of the dependent variable were not taken when studying the model.

Table (6): Measures the scientific significance of the studied model

-2 Log likelihood	R2 Nagelkerke	*R2* Cox & Snell
247.742	0.69	*0.47*

There is also another statistic for R2 denoting R_L^2 and calculated as follows:

$$R_L^2 = \frac{LL_G}{LL_0} = \frac{707.880 - 247.742}{707.880} = 0.65$$

ISSN (1681- 6870)

That is, the explanatory variables studied within the logistic regression model contribute 65% to the reduction of the logarithm of the weighting function of the model that includes the fixed term only, and this scale is good and important in explaining the scientific significance of the model because it relied on reducing the logarithm of the weighting function.

As shown in Table No. (7), which shows estimates of the model parameters, standard error, degree of significance of factors, Wald test and value ((Odds Ratio Exp (B)), that the moral variable method of diagnosing the disease at the level of significance is 0.05, and the patient's age variable is significant at the level of Significance 0.1, and therefore it can be concluded that these variables have a significant effect on the dependent variable.

Also, the Wald statistics for the variable of the diagnostic method reached 219.48, which means that the logistic regression model for the diagnostic method differs from zero. It also means that this variable has a high ability to classify predictive power, and it is also considered the best moral variable to distinguish between the patient's condition. Also, the patient's age variable comes in the second stage, because the value of the Wald statistic was 2.709 at the level of significance 0.1, and this means that adding the patient's age variable to the variable of the diagnostic method contributes to increasing the ability of the model to predict the classification of the patient's condition.

Table (7): Estimating the parameters of the logistic regression model

Variable	B	S.E	Wald	df	Sig.	EXP(B)	95% CI for EXP(B)	
							Lower	Upper
Gender	-0.245735	.391	.396	1	.529	.782	.364	1.681
Age	0.296001	.180	2.709	1	.100	1.344	.945	1.913
Occup	-0.144475	.138	1.099	1	.294	.865	.661	1.134
Basis	-1.865843	.126	219.481	1	.000	.155	.121	.198
Grad	0.060236	.135	.199	1	.655	1.062	.815	1.384

- **Second :Fit the final model using the neural network model**

It must be mentioned first that there are names of some terms that differ in neural networks from those in the commonly used statistical models. Explanatory variables are called inputs and dependent variables are called training values or target values. As for the predicted values, they are called outputs. Finally, weights, whose values are positive, and the sum of them is equal to one called the parameters, and secondly, the use of neural networks for data is 100% available for Training and 0% for the Testing.

The inputs of the neural network were represented in the five studied explanatory variables, and the outputs are the dependent variable (patient's condition). After application it was found that the network training outputs according to Figure (5) that the network had been trained through stimulus functions (hidden layers) and 4 hidden neurons of the model were estimated, and from During which the model was able to explain the studied phenomenon, as some of these weights were less than zero as shown in Table (8).

ISSN (1681- 6870)

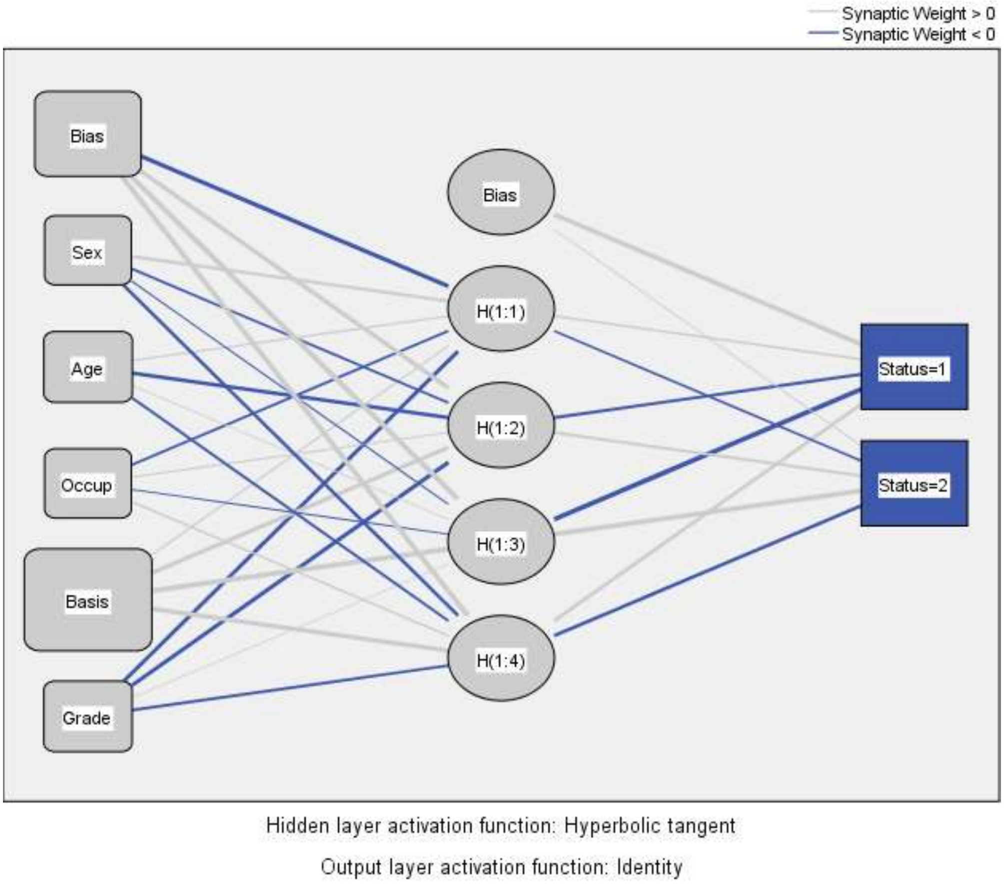

Figure (5): Network training outputs

Table (8): Estimation of model parameters resulting from network training

Predictor		Predicted					
		Hidden Layer 1				**Output Layer**	
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	[Status=1]	[Status=2]
Input Layer	(Bias)	-.672-	.755	2.884	1.079		
	Sex	.272	-.151-	-.033-	-.319-		
	Age	.055	-.411-	.011	-.247-		
	Occup	-.166-	.043	-.018-	.072		
	Basis	.048	.645	1.693	.855		
	Grade	-.442-	-.577-	.001	-.259-		
Hidden Layer 1	(Bias)					.978	.017
	H(1:1)					.105	-.111-
	H(1:2)					-.268-	.265
	H(1:3)					-1.017-	1.020
	H(1:4)					.395	-.392-

Table (9) shows the degree of importance for each factor affecting the patient’s condition using the neural network model. It turned out that the method of diagnosing the disease is the most influential factor of 100%. As for the following factors, they are of little effect compared to the method of diagnosing the disease, the patient’s age by 8.1%, then the Degree of illness by 7.6%, class, and profession by 3.8%.

ISSN (1681- 6870)

Table (9): The relative importance of each factor affecting the patient's condition

Variable	Importance	Normalized Importance
Gender	.031	3.8%
Age	.066	8.1%
Occup	.031	3.8%
Basis	.812	100.0%
Grade	.061	7.6%

6. Conclusions

After applying the logistic regression model and the neural network model on the sample taken from the Iraqi Cancer Center at the Ministry of Health for the year 2017, the following conclusions were reached:

1. Using the evaluation criteria (model accuracy, model sensitivity, model specificity, The model's false alarm rate, The area below the Roc curve, Wrong classification rate), a comparison is reached between the logistic regression model and neural network model, preference of the neural network model for application to study data.
2. The results also showed the efficiency of the inferred model in predicting the classification of the patient's condition.
3. The results indicated that the factor of the disease diagnosis method is a variable with a significant effect, with a level of significance less than 5%, using the logistic regression model.
4. The explanatory variables studied in the inferred model succeeded. An explanation of 0.47% to 0.69% of the changes in the dependent variable, according to the value of the extracted R2, but according to the results there are other variables not included in the model that may affect, for example, the economic condition of the patient, is there A member of his family has stomach cancer ... and others.
5. The efficiency of the correct classification was 100%, which is a high percentage indicating the relative importance of the explanatory variables, as it was found that the neural network has appropriately learned and trained.
6. The results indicated that the most influencing factors on the patient's condition according to the logistic regression model (method of diagnosing the disease, patient's age, grade, gender, and profession).
7. The factor of the method of diagnosing the disease was the factor 100% the most influencing the dependent variable, and the gender and profession workers were the least influencing the dependent variable by 3.8%.

Reference

[1] Andrew. E , 1997, "The use of the Area Under the ROC Curve in the Evaluation of Machine Learning Algorithms", Pattern Recognition, Vol. 30, No. 7, pp. 1145-1159.

[2] Andy .Th , 2017,"An introduction to neural networks for beginners", Adventures in Machine Learning, (online), available at: https://adventuresinmachinelearning.com/wp-content/uploads/2017/07/An-introduction-to-neural-networks-for-beginners.pdf

[3] Camilla .C · Longobardi1. M ,2015, "Some mathematical properties of the ROC curve and their applications", Ricerche di Matematica, Vol. 64, No. 2.

[4] Chao-Ying. J, Kuk. L, Gary. M. , 2002, "An Introduction to Logistic Regression Analysis and Reporting", The Journal of Educational Research, Vol. 96, No. 1.

[5] Draper, R. and Smith, H., 1981, "Applied Regression Analysis", 2nd Edition, John Wiley, New York .

ISSN (1681- 6870)

[6] Hajmeer. M, Basheer, I, 2003, “Comparison of logistic regression and neural network-based classifiers for bacterial growth”, Food Microbiology, Volume 20, Issue 1, February, Pages 43-55

[7] Hosmer, D. and Lemeshow, S. (2000), Applied Logistic Regression, 2nd edition. New York: Johnson Wiley & Sons, Inc .

[8] Joseph. M. H ,2015, "Practical Guide to Logistic Regression", CRC Press, Arizona State University, USA.

[9] King, G. and Zeng, L. 2001, "Logistic Regression in Rare Events Data", Society for Political Methodology. February , (online), available at: https://gking.harvard.edu/files/0s.pdf

[10] Kukrej.H, Bharath. N, Siddesh. C, Kuldeep. S, 2016, "An Introduction To Artificial Neural Network", International Journal Of Advance Research And Innovative Ideas In Education, Vol.1 Issue 5.

[11] Logistic regression, NCSS Statistical Software NCSS.com, (online), available at: https://ncss-wpengine.netdna-ssl.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Logistic_Regression.pdf

[12] Michael .G, Jonas B, Jakob .F, Ulf . E, Lars. E, Mattias. O, 2006, Comparison between neural networks and multiple logistic regression to predict acute coronary syndrome in the emergency room, Artificial Intelligence in Medicine , Vol. 38, No. 3, 305—318.

[13] Michael. P. L, 2008, "Logistic Regression" , Circulation, Vol. 117, No. 18.

[14] Raghavendra B.K., Srivatsa.S.K, 2011, "Evaluation of Logistic Regression and Neural Network Model With Sensitivity Analysis on Medical Datasets", International Journal of Computer Science and Security (IJCSS), Vol. 5, Issue 5.

[15] Saeed. M, and Norman .C, 2000, Neural Network Versus Econometric Models in Forecasting Inflation, Journal of Forecasting J. Forecast. 19, 201-217..

[16] Logistic Regression, Chapter 321, NCSS Statistical Software NCSS.com, (online), Available at: pdfslide.net/document/chapter-321/Logistic-Regression-ncss-321-logistic-regression,html.

[17] Yan Lai Mei and Choon Ong Hong, 2000, "Neural Network Forecasting on electricity consumption in Malaysia", USM, Malaysia.