

## Multivariate Modeling and Analysis for Classification of Neonatal Congenital Malformations in Newborns

### نمذجة وتحليل متعدد المتغيرات لتصنيف اصابة الاطفال حديثي الولادة بالتشوهات الخلقية

Husam Abdulrazzak Rasheed<sup>1</sup>

حسام عبد الرزاق رشيد<sup>1</sup>

Statistics Department/collage  
of administration and  
Economics, Mustansiriyah  
University

قسم الاحصاء/ كلية الادارة والاقتصاد/  
جامعة المستنصرية/ العراق

[husamstat@uomustansiriyah.edu.iq](mailto:husamstat@uomustansiriyah.edu.iq)

Aseel Abdulrazzak Rasheed<sup>2</sup>

أسيل عبد الرزاق رشيد<sup>2</sup>

Statistics Department/collage  
of administration and  
Economics, Mustansiriyah  
University

قسم الاحصاء/ كلية الادارة والاقتصاد/  
جامعة المستنصرية/ العراق

[aseelstat@uomustansiriyah.edu.iq](mailto:aseelstat@uomustansiriyah.edu.iq)

Rawaa Salh AL-Saffar<sup>3</sup>

رواء صالح الصفار<sup>3</sup>

Statistics Department/collage  
of administration and  
Economics, Mustansiriyah  
University

قسم الاحصاء/ كلية الادارة والاقتصاد/  
جامعة المستنصرية/ العراق

[rawaaalsaffar@uomustansiriyah.edu.iq](mailto:rawaaalsaffar@uomustansiriyah.edu.iq)

#### Abstract

The health sectors are considered one of the most important sectors that provide health services to all members of society in the countries of the whole world through its health institutions in order to protect society and achieve welfare for all, we will look at one of the most important and dangerous diseases that affect newborns in all countries of the world, which is congenital malformations.

Congenital malformations are functional or structural abnormalities that develop during the life of the fetus inside the womb, that is, between pregnancy and birth, also known as a birth defect, congenital anomaly is identified before birth, at birth or later in childhood. Anomalies can be mild or severe. Some are life-threatening, and infected children live only a few months.

The causes of congenital malformations can be genetic, environmental or multifactorial, genetic causes of birth defects include chromosomal deviations such as those found in Down syndrome and defects of the single Mendelian gene. In these cases, defects are caused by too many or too many chromosomes, or are caused by structural problems with chromosomes.

Genetic causes also include dominant inheritance, where one parent passes through one defective gene, and recessive inheritance, where both parents, who do not have the disease, pass on the disease genes to the child, the risk of genetic birth defects is more prevalent in consanguineous marriages and women who give birth after the age of 35 and under 18.

A statistical analysis was conducted for newborns at Ibn Al-Balady Hospital in Baghdad, and a sample of 178 births was taken, divided into 99 births with congenital malformations, and 79 births without congenital malformations, using logistic regression and neural networks. When comparing the two methods, it was found that neural networks were superior in accurately classifying congenital malformations in newborns, and the most influential variable on the occurrence of congenital malformations was the mother's age.

**Keywords:** congenital malformations, logistic regression, neural networks, comparison criterion

#### المستخلص :

تعتبر القطاعات الصحية من أهم القطاعات التي تقدم الخدمات الصحية لكافة أفراد المجتمع في دول العالم أجمع من خلال مؤسساتها الصحية بهدف حماية المجتمع وتحقيق الرفاهية للجميع، سنتناول في هذا المقال أحد أهم وأخطر الأمراض التي تصيب الأطفال حديثي الولادة في كافة دول العالم، وهي التشوهات الخلقية.

التشوهات الخلقية هي تشوهات وظيفية أو بنوية تحدث أثناء حياة الجنين داخل الرحم أي بين الحمل والولادة، وتعرف أيضاً بالعيوب الخلقية، ويتم التعرف على التشوهات الخلقية قبل الولادة أو عند الولادة أو في وقت لاحق من الطفولة. يمكن أن تكون التشوهات خفيفة أو شديدة، وبعضها يهدد الحياة، ويعيش الأطفال المصابون بها بضعة أشهر فقط.

أسباب التشوهات الخلقية قد تكون وراثية أو بيئية أو متعددة العوامل، الأسباب الوراثية للعيوب الخلقية تشمل الانحرافات الكروموسومية مثل تلك الموجودة في متلازمة داون و عيوب الجين المنجلي المفرد. في هذه الحالات، تحدث العيوب بسبب كثرة أو كثرة الكروموسومات، أو تحدث بسبب مشاكل بنوية في الكروموسومات.

الأسباب الوراثية تشمل أيضاً الوراثة السائدة، حيث يمرر أحد الوالدين جيناً معيباً واحداً، والوراثة المتنحية، حيث ينقل كلا الوالدين، اللذين لا يعانين من المرض، جينات المرض إلى الطفل، يكون خطر العيوب الخلقية الوراثية أكثر انتشاراً في حالات الزواج من الأقارب والنساء اللاتي يلدن بعد سن 35 عاماً وتحت 18 عاماً.

تم إجراء التحليل الإحصائي للأطفال حديثي الولادة في مستشفى (ابن البلدي) في بغداد، واخذ عينة من 178 حالة ولادة، مقسمة الى 99 حالة ولادة مصابة بالتشوهات الخلقية، و 79 حالة ولادة غير مصابة بالتشوهات الخلقية، باستخدام الانحدار اللوجستي والشبكات العصبية، وعند إجراء المقارنة بين الطريقتين، تم التوصل إلى أفضلية الشبكات العصبية في التصنيف الدقيق لحالات التشوهات الخلقية عند الأطفال حديثي الولادة، وكان المتغير الأكثر تأثيراً على حدوث التشوهات الخلقية هو متغير عمر الأم.

**الكلمات المفتاحية:** التشوهات الخلقية، الانحدار اللوجستي، الشبكات العصبية، معيار المقارنة

## 1. Introduction

Deformities are congenital defects that affect various organs in the human body, which the fetus is infected with during pregnancy, including those that affect the lungs due to incomplete growth, and those that affect the digestive system such as lip deformity or lack of formation of the anus and rectum, or may affect the stomach and duodenum, or may affect the gallbladder or lack of formation of the small or large intestine, and some of them affect the bones such as the upper and lower extremities and their poor growth, and some of them may affect the heart, brain, liver and other deformities in other organs of the body, and there are congenital deformities called dangerous, including hip dislocation or twisted foot or deficiency in the formation of some bones or their deformity ..... etc.

Some congenital deformities are visible to the eye, while others do not appear except through the necessary examinations using modern medical devices such as CT scans, ultrasound, etc., and the discovery of one congenital deformity can lead to the discovery of another congenital deformity.

Due to the importance and seriousness of this disease in many births and its widespread prevalence in many countries, its rates may increase if attention is not paid to reducing it. In most Arab countries, there are no sufficient statistics to determine the prevalence of these congenital malformations.

## 2. Research objective

To study congenital malformations by using two important statistical methods, neural networks and logistic analysis, on a sample of people with congenital malformations to classify them and then make a comparison between the two methods used in the research to obtain the best methods, while identifying the most important factors that affect the increase in the incidence of malformations

## 3. Theoretical Side

Multivariate statistical methods are considered important methods that many researchers have used to analyze some natural phenomena, in this research, we will discuss one of these methods, the neural network method and the logistic regression model.

### 3-1: Neural Networks

They are known as non-linear functions that are flexible, as they do not require many assumptions that govern the relationship between the dependent and independent variables, it also handles non-parametric and small-sized data and does not require normal distribution dependence, it has achieved great success in the fields of medicine because it provides a better level than traditional statistical models, Neural networks are divided into two types: (Abed, H., & Badr, D. 2021),[1]

Artificial Neural Networks (ANN), Neural Networks Algorithm (NNA)

### 3-1-1: Components of Artificial Neural Networks (Bradley. p. 1997), [3]

#### 1- Input Layer

The network is supplied and fed with data, the inputs may be in the form of raw data or other processing outputs, the processing units in the input layer do not perform any computational processing, but rather the data is transferred via connections to the hidden layer or to the output layer if there is no hidden layer, and any neural network contains only one layer of input units.

#### 2- Output Layer

They are processing units through which the final output of the network is outputted. They may contain one or more processing units. The processing units receive signals coming to them from the hidden layer or from the input layer directly after performing the necessary processing. They may send the final output signal or return these outputs to the network again if the required processing is not completed. The network usually contains one output layer.

**3- Hidden Layer**

It is located between the input layer and the output layer. Some networks contain one or more hidden layers, and may not contain a hidden layer, as they receive signals from the input layer, process them, and send them to the output layer.

**4- Connections**

It is the one that connects the layers or the units within each layer and is responsible for transferring signals between the layers and processing units.

**5- processing Element**

It represents one of the basic components that make up all layers of the neural network, and each processing unit consists of the following:

**6- Weights Coefficients**

The neural network is based on the weighting of the element, which expresses the relative importance of the input element.

**7- Summation Function**

The first operation performed by the processing unit is to calculate the sum of the weighted inputs coming to the unit using the sum function. This function calculates the average of the weights for all the inputs coming to the processing unit by multiplying each input by its accompanying weight to calculate the sum as follows:

$$S_j = \sum_{i=1}^n x_i w_{ij} \quad (1)$$

Since:

$S_j$  : The result of the addition process for each processing unit, ( $j$ ).

$x_i$  : The value entered from unit ( $i$ ) to unit ( $j$ ).

$w_{ij}$  : The weight that connects unit ( $i$ ) to unit ( $j$ ) of the previous layer.

**8- Transfer Function**

It is known as the external stimulation function because the processing units receives inputs from other processing units or from other sources. It is the second process after addition and from the transformation functions: (Garson, G. D. (2014),[5]

**8-1 Sigmoid Function**

The tasks of this function are to convert the outputs to values between zero and one, which is called binary logistic activation, or to convert the outputs to values between (1+, 1-), which is called in this case the logistic regression function. It is one of the most used functions and is known as follows:

$$F(x) = \frac{1}{1 + e^{-s}} \quad (2)$$

Since:  $s$  is the weighted sum of the inputs plus the bias term, which is symbolized by the symbol  $\emptyset$ .

**8-2 Step Function**

The function converts the output function to values between (1,0) as in the binary system and takes the following form:

$$F(x) = \begin{cases} 1 & \text{if } s \geq 0 \\ 0 & \text{if } s < 0 \end{cases} \quad (3)$$

**8-3 linear function**

It works to provide outputs equal to the weighted inputs of the processing unit and usually uses linear units for linear approximation and takes the following form:

$$F(x) = s$$

where s is the output.

**8-4 Sign Function**

It is used in the classification and discrimination process and uses a specific value called  $\emptyset$  that takes the following form:

$$F(x) = \begin{cases} 1 & \text{if } s \geq \emptyset \\ -1 & \text{if } s < \emptyset \end{cases} \quad (4)$$

**8-5 Output Function**

After the addition process, the conversion process is carried out into values confined to a specific range, and an output function gives the final result of the problem.

**3-1-2: Neural Networks in Classification**

It has become possible to use neural networks (ANN) in the classification process by specifying an activation function specific to the classification. There are several activation functions that differ according to the output that is obtained in addition to the goals of the neural network that are to be achieved. The step function is used because it is suitable for the classification process and the outputs give two results, zero and one, as in equation (3).

The step function is used in the output layer units while the Sigmoid function is used in the hidden layer and takes as in equation (2).

The value of (s) can be extracted according to the following equation:

$$S = \sum_{i=1}^n x_i w_{ij} + \emptyset \quad (5)$$

Finally, the network outputs are either zero or one depending on the inputs. If the output is (0), the observation is classified in the first group, and if the output is (1), the observation is classified in the second group.

**3-2: Logistic Regression Model**

Logistic regression is a statistical method that aims to build a model that shows the relationship between a binary dependent variable and a set of independent variables, the dependent variable here may be nominal or categorized to take two or more values and logistic regression is characterized by the fact that this relationship between the variables is a non-linear function, logistic regression provides us with knowledge about the relationship between the binary dependent variable and the independent variables and the strength of that relationship. (Horton, N. & laird, N. (2001), [6]

One of the most important uses of logistic regression analysis is to predict whether an event will occur or not, as the dependent variable (Y) is binary, i.e. it takes only two values, each of which is coded (1) if the event occurs with a probability of (P) and (0) if the event does not occur with a probability of (1-P).

The logistic regression model (logit model) is expressed by the following formula:

$$\text{logit}(P_i) = \log_e \left( \frac{P_i}{1 - P_i} \right) = +B_1X_1 + B_2X_2 + \dots + B_kX_k \quad (6)$$

Express it as follows:

$$P(x, B) = \frac{e^{B_0+B_1X_1+B_2X_2+\dots+B_kX_k}}{1 + e^{B_0+B_1X_1+B_2X_2+\dots+B_kX_k}} = \frac{1}{1 + e^{B_0+B_1X_1+B_2X_2+\dots+B_kX_k}} \quad (7)$$

Whereas

$B_0$ : The cut off part.

$X$ : Independent variable

$B$ : The coefficient for each instance of the independent variable

$P$ : is the probability of the event of interest

$\frac{P}{1-P}$ : The ratio of the probability of an event occurring to not occurring is called (odds)

$\log\left(\frac{P}{1-P}\right)$ : Natural logarithm of the odds ratio

**3-2-1 Logistic Regression Model Hypotheses**

Logistic regression is a model with a basic hypothesis on the dependent variable, but the rest of the assumptions are a projection of assumptions that were inherent in ordinary regression models.

**(Ferrer, J. & wang. 1. (1999), [4]**

- 1- -The basic assumption on which logistic regression relies is about the dependent variable, i.e. the dependent variable  $Y$  is assumed to be a binary descriptive variable  $[1,0]$  and the probability of each value occurring is a fixed amount. While the independent variables have no assumptions about their nature, they may be continuous, discrete, binary descriptive, or multiple.
- 2- Logistic regression does not assume a linear relationship between the dependent variable and the explanatory variables:

$$P(Y) = \frac{\exp(B_0 + B_1X_1 + \dots + B_iX_i)}{1 + \exp(B_0 + B_1X_1 + \dots + B_iX_i)} \quad 0 < p \leq 1,$$

$$P(Y) = \frac{1}{1 + \exp(B_0 + B_1X_1 + \dots + B_iX_i)}, \quad i = 1, \dots, k \quad (8)$$

Since:

$B_i, B_0$ : Estimated data coefficients

$X$ : Independent variable

$P(Y)$ : Dependent variable

Exp: (e) is the base of the natural logarithm

- 3- The independence of observations and random errors and the independence of explanatory variables are assumed, i.e. there is no autocorrelation between independent variables, whereby variables that are perfectly correlated are omitted.
- 4- Independent variables are assumed to be measured without error, and non-significant variables are omitted. **(Hosmer, D. & lemeshow, S. 2000), [7]**

**3-2-2 Estimating the Logistic Regression Model**

Maximum Likelihood Estimation (MLE) is the appropriate method for estimating the parameters of a logistic regression model and is symbolized by the symbol  $B$ .

The maximum likelihood (MLE) method aims to maximize the logarithm of the likelihood so that it reflects the extent to which the observed values of the dependent variable can be predicted from the existing independent variables. The mathematical formula for the maximum likelihood function is written as follows: **(Jaccard, J. 2001), [9]**

$$L(B_i) = \prod_{i=1}^n p_i^{y_i} (1 - p_i^{y_i}) \quad (9)$$

By substituting a value from equation (11) and taking the natural logarithm of both sides, you get the following equation:

$$\ln(L(B_i)) = \sum_{i=1}^n y_i \{(B_0 + B_1X_1 + \dots + B_iX_i) - \sum_{i=1}^n \ln(1 + e^{(B_0+B_1X_1+\dots+B_iX_i)})\}$$

$i = 1, \dots, k$

To obtain the parameters that maximize the possibility function, we derive the logarithm of the possibility function with respect to the parameters to be estimated and then set it equal to zero, and solve the nonlinear equations that we obtained through an iterative algorithm.

**3-2-3 Wald test to detect the significance of logistic regression model parameters.**

It is a statistical test used in many statistical models, including the logistic regression model, where it tests the significance of the coefficients of the logistic regression model for each independent variable, by testing the null hypothesis, which states that the effect of the logistic regression coefficient is equal to zero, i.e.  $H_0: B = 0$ . (Mohammed, R. S., Rashid, A. A. R., & Abdul-Hassan, S. A. 2024), [10]

The Wald statistic follows a chi-square distribution  $\chi^2$  with one degree of freedom, and is written in the following mathematical form:

$$W = \frac{b}{S.E_b} \quad (10)$$

Since:

$b$ : It is the estimated value of the logistic regression coefficient for the independent variable.

$S.E_b$ : The standard error value of the logistic regression coefficient for the independent variable.

**3-2-4 Classification Table**

The classification table is a method for evaluating the accuracy of prediction in a logistic regression model, i.e. the efficiency of the model's classification, and it serves as a check on the quality of the model's fit to the data. (Al-Saffar, R. S., Rasheed, A. A. R., & Sadik, N. J. 2021),[2]

This analysis is based on the fact that if the model correctly predicts a classification of cases based on a specific criterion, this indicates that the model matches the observed data. Table (2-1) is the general form of classification tables. (Ibrahim, W. S. 2020), [8]

**Table (1-2) General form Classification Tables**

Classification		Expected		Total
		Positive	Negative	
Observation	Positive P	True Positive (Tp)	False Negative (Fn)	P
	Negative N	False Positive (Fp)	True Negative (Tn)	P/
Total		Q	Q/	1

Where P is called the model specificity and is the probability value that the predicted classification will be positive for cases that are actually positive, while P is called the model sensitivity and is the probability value that the predicted classification will be negative for cases that are actually negative.

**4. Applied side**

Data on congenital malformations of newborns were collected from Ibn Al-Balady Hospital in Baghdad. A sample of 178 births was taken, divided into 99 births with congenital malformations and 79 births without congenital malformations. Therefore, the variable Y took the following values: (0) no congenital malformations, (1) congenital malformations.

The research included a number of independent variables, which are considered among the most important factors that contribute to the infection of newborn children with congenital malformations, as follows:

X1	Mother's age	
X2	The mother has chronic diseases	(0) No chronic diseases (1) There are chronic diseases
X3	Mother's wight	
X4	The degree of kinship between the parents	(0) No kinship (1) There is kinship

X5	Baby weight	
----	-------------	--

**4-1 Data Analysis**

**4-1-1: Logistic Regression Model Analysis**

**Table (1) Estimated Logistic Regression Model Transactions**

Variables	B	S.E	Wald	df	Sig.	Exp(B)	95% C.I. for Exp	
							Lower	Upper
X1	.205	.051	16.188	1	0.000	1.228	1.111	1.357
X2(1)	1.370	.562	5.942	1	0.015	3.935	1.308	11.837
X3	.025	.062	.157	1	0.692	1.025	0.907	1.157
X4(1)	.577	.485	1.419	1	0.234	1.781	0.689	4.604
X5	.001	.000	8.522	1	0.004	1.001	1.000	1.002
Constant	-9.749	3.793	6.606	1	0.010	0.000		

We notice from Table (1), the variables (X1, X2, X5) that represent (wife's age, wife's weight and child's weight) have a significant effect in the model and this is clear from the Sig. value because it is less than 0.05, but this effect differs according to the risk ratio for each variable (and according to the Exp(B) value).

The variable X2 has the strongest effect on the incidence of birth defects in newborns, because the value of Exp(B) = 3.935, which is further than the other values from the correct (1), and this means that the diseases that the mother suffers from have the strongest effect on the incidence of birth defects in newborns. The variable of the mother's age (X1) came in second place in terms of effect, as the value of Exp(B) = 1.228, then the variable X5 came in third place, the weight of the child.

The two variables (mother's weight and kinship between spouses) did not have a significant effect on the incidence of congenital malformations in newborns, and this is clear from the Sig. value because it is greater than 0.05.

**Table (2) Data Classification Using Logistic Regression Model**

Observed	Y			Percentage Correct
	There are no birth defects	There are birth defects		
Y	There are no birth defects	20	15	57.1
	There are birth defects	11	90	89.1
Overall Percentage				80.9

We note from Table (2) that the percentage of correct classification for the absence of congenital malformations for newborn children was 57.1, the percentage of correct classification for children with congenital malformations was 89.1, and the classification percentage for the overall model was 80.9, and this percentage is considered good in such studies.

**4-1-2: Neural Network Model Analysis**

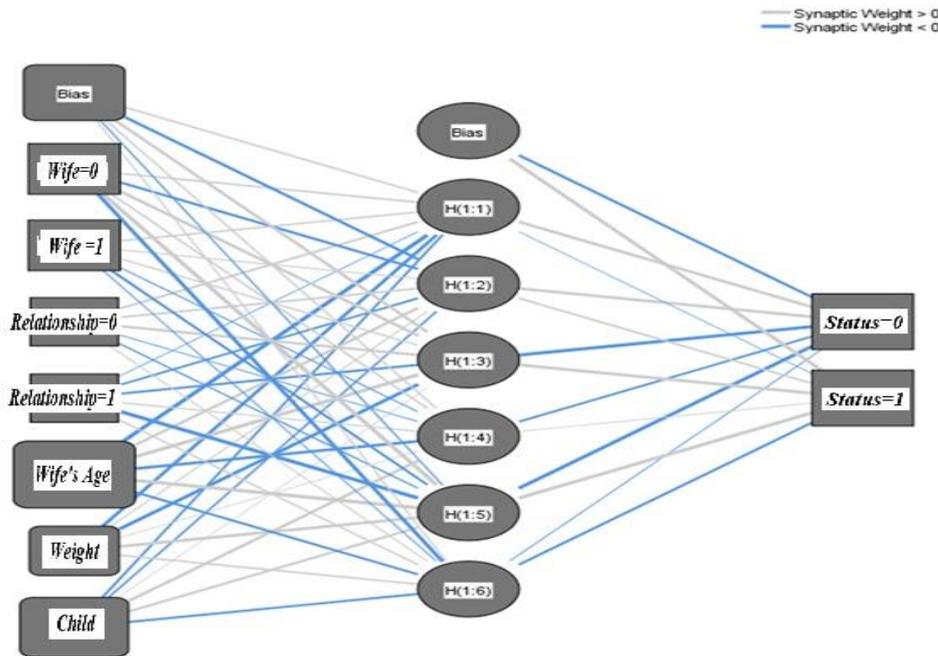


Figure (1) Neural Network Drawing

It is clear from Figure (1) which shows the architecture of the neural network which consists of 3 layers, where the first layer (inputs) consists of 7 units in addition to the bias unit Bias, while the second layer in the middle (hidden) consists of 6 units, while the third layer (outputs) consists of two results of the network which are ((0) there are no congenital malformations, (1) there are congenital malformations).

Table (3) Data Classification Using Neural Networks

Sample	Observed	Predicted		
		There are no birth defects	There are birth defects	Percent Correct
Training	There are no birth defects	9	17	34.6%
	There are birth defects	3	61	95.3%
	Overall Percent	13.3%	86.7%	77.8%
Testing	There are no birth defects	5	4	55.6%
	There are birth defects	4	33	89.2%
	Overall Percent	19.6%	80.4%	82.6%

Dependent Variable: Y

We note from the table above that the correct classification of the absence of congenital malformations is (34.6%) in the training sample, while it was (55.6%) in the testing sample, and the percentage of correct classification of the presence of congenital malformations was 95.3% in the training sample, while in the testing sample it was 89.2%. We can conclude from the testing sample that the percentage of correct classification of data using neural networks was 82.6%, and this percentage is considered good for classifying new vocabulary.

**Table (4) Independent Variable Importance**

Variables	Importance	Normalized Importance
X2(1)	0.114	26.6%
X4(1)	0.071	16.5%
X1	0.430	100.0%
X3	0.092	21.4%
X5	0.292	67.9%

We note from Table (4) that the most important variable affecting the classification using neural networks is the mother’s age, as its percentage reached 100%, followed by the child’s weight, as its percentage reached 67.9%, then the diseases that the mother suffers from, as its percentage reached 26.6%, while the two variables (kinship between the spouses and the mother’s weight) had percentages of (16.5%, 21.4%) respectively, and this percentage is considered to have a small effect on the children’s infection with congenital malformations.

**4-1-3: Comparison Between Logistic Regression Models and Neural Networks**

**Table (5) Comparison of models by classification accuracy**

Models	Sensitivity	Specificity	Accuracy
Logistic regression	57.1%	89.1%	80.9%
Neural networks	55.6%	89.2%	82.6%

We note from Table (5) that the accuracy of the correct classification of the neural network model analysis is the best, as the model’s accuracy in classification reached 82.9%.

**5. Conclusions and Recommendations**

1. We note that the results of the logistic regression and neural network models are identical in terms of the (risk factors) that have a significant impact on the model, as we note that the variables that affect the incidence of congenital malformations in children are (diseases that the mother suffers from during pregnancy, the mother’s age, and the child’s weight).
2. In the logistic regression, the order of risk factors in terms of importance was (diseases suffered by the mother in first place, followed by the mother's age in second place, and then the child's weight in third place).
3. In the neural networks, the sequence of risk factors in terms of importance was (the mother's age in first place, followed by the child's weight in second place, then the diseases the mother suffers from in third place, then the mother's weight in fourth place, and finally the kinship between the spouses in fifth place).
4. Neural networks give an order of importance to independent variables.
5. We notice through the logistic regression model that there is a direct relationship between all independent variables and the dependent variable.
6. We recommend conducting similar studies by introducing other variables that affect the incidence of congenital malformations in newborns.
7. We recommend using neural network models in classifying data in different fields, such as economic and social, as they have high accuracy in data analysis.

**Reference**

[1] Abed, H., & Badr, D. (2021). A Comparison of the Logistic Regression Model and Neural Networks to Study the Determinants of Stomach Cancer. Journal of Al-Rafidain University College For Sciences (Print ISSN: 1681-6870, Online ISSN: 2790-2293), (2), 251-266

- [2] Al-Saffar, R. S., Rasheed, A. A. R., & Sadik, N. J. (2021). Nonlinear modeling of the factors affecting the disease Covid-19 by adopting the method of sampling by reference. *Int. J. Agricult. Stat. Sci.* Vol, 17(1), 2301-2307.
- [3] Bradley. p. (1997). "The use the area under the roc curve in the evaluation of machine learning algorithms", pattern recognition.
- [4] Ferrer, J. & wang. 1. (1999). "Comparing the classification accuracy among nonparametric, Paramtric Discriminant analysis and Logistic Regression Method.
- [5] Garson, G. D. (2014). *Logistic regression: Binary and multinomial*. Asheboro, NC
- [6] Horton, N. & laird, N. (2001). "Maximumlielihood analysis of logistic regression models with in complete covariate date and auxiliary information *Biometrics*", Vol.57, pp. 34-42, march 2001
- [7] Hosmer, D. & lemeshow, S. (2000). "Applied logistic regression". 2nd edition. Now yorh: Johnson Wiley Sons, Ine. Vol.30, no.7, pp. 1145-1159
- [8] Ibrahim, W. S. (2020). To determine the most important factors affecting pancreatic cancer in Iraq using the logistic regression model. *Periodicals of Engineering and Natural Sciences (PEN)*, 8(1), 277-283
- [9] Jaccard, J. (2001). "Interaction effects in logistic Regresstion. Sage university paper series on quantitative applications in the social Sciencies", No.7 -135.
- [10] Mohammed, R. S., Rashid, A. A. R., & Abdul-Hassan, S. A. (2024). Classification of Children's Participation in the Processions of Imam Hussein (Field Study). *AL ARBA'IN*, (28)