

Article info

Received on: 22.05.2022

Accepted on: 24.06.2022

Published on: 31.07.2022

doi: <https://doi.org/10.52688/ASP21056>

Research Article

The use of Artificial Neural Networks in Predicting Patterns of Contemporary Trends Towards Childbearing - Baghdad Governorate as A model

Zainab Mohammed Ameen ^{1,*}, Ruqaya Ahmed M. Amin ², Lateef Hashim Kzar ³^{1,2} College of Arts, University of Al-Iraqia, Baghdad, Iraq³ College of Education for Humanities Sciences, University of Wasit, Kut, Iraq* zainab.m.ameen@aliraqia.edu.iq

ABSTRACT

The research came to reveal contemporary patterns of childbearing in light of the emerging changes that affected life in all its details, the research data was extracted from a questionnaire distributed electronically to cover the nine districts of Baghdad Governorate, and (546) electronic forms were collected during a period of three weeks during the period from (1-21) June 2022, during the curfew period that accompanied the outbreak of the Corona virus. After three weeks, the electronic questionnaire was closed and the results were collected and unpacked in Excel and transferred to a SPSS file to be analyzed later. The study came in two sections, the first: descriptive statistics and the characteristics of the research sample, which included a theoretical introduction and description of the research community with a review of descriptive statistics related to demographic, social and cultural variables. The second: It came under the title of Statistical Analysis, in which we reviewed in detail the analysis of multi-layer neural networks and their role in prediction by creating a scale consisting of 8 items (the answer in which was determined by five choices according to the Likert five-way scale) to demonstrate its ability to classify the dependent variable (contemporary patterns of reproduction) Where it was classified into two categories according to reproductive preferences, the first category (prefer to have 4 children or less / Modern style) and the second category (prefer to have 5 children or more / the traditional style), this classification was according to age, where the cases were classified into two categories, the first category (with Ages 40 years and less) and the second category (those aged 40 years and over). The predictive model gave accurate results that can be considered very good and reliable in classifying the phenomenon under study.

Keywords: Neural network MLP, interstitial links, hidden layer, future operating characteristics curve (area under the curve)

INTRODUCTION

Childbearing is one of the most dynamic human phenomena, as it changes according to place and time. Thus, we may find regionally different patterns between different countries and within one state, and within the same region we may find variations or differences in time, Thus, it is possible for us to show patterns that vary according to the generations being studied, procreation from phenomena that are affected by social, economic, cultural and civilizational changes imposed by the surrounding environment, these variables function differently according to the age variable. The study sheds light on these spatial and temporal changes and discrepancies, which are represented by the ages of the study population. The location of the study area in Baghdad governorate with its nine districts, which are located within two latitudes (33-34) north and longitudes (40-44) east, as shown in Figure (1).

Survey, Map Production Division, Baghdad Administrative Map at a scale of 1/500000. 2019.

Descriptive statistics and characteristics of the research sample:

The research data was taken from a questionnaire that was distributed electronically during the Corona pandemic period, and it was emptied to be the number of forms as shown in Table(1) .

*Corresponding author

Zainab Mohammed Ameen,
College of Arts, University of Al-Iraqia, Baghdad, Iraq
e-mail: zainab.m.ameen@aliraqia.edu.iq

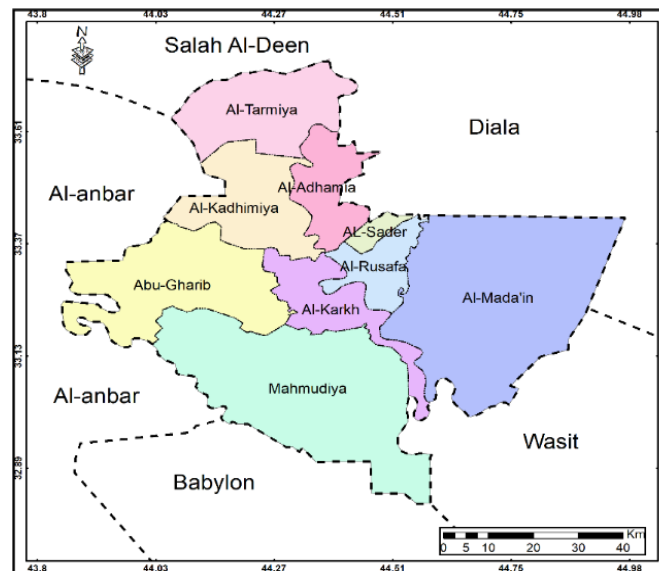


Figure 1: Area study, Source/ based on: Ministry of Water Resources, General Commission for

Table1: Distribution of the research sample to the districts of Baghdad governorate.

QADA	sample size	QADA	sample size
Al-Rusafa	240	Mahmudiya	31
Al-Adhamia	29	Abu-Gharib	11
AL-Sader	63	Al-Tarmiya	18
Al-Karkh	115	Al-Mada'in	22
Al-Kadhimiya	17	Total	546

For many purposes, the age distributions of a population represents the most significant compositional variable. After population size, the age distribution is the most important factor in determining a society's character and for calculating many of the rates used by demographers. (Thomas,2018,p.78), Accordingly the research forms were classified according to the age groups shown in Table (2).

Table 2: Numerical and relative distribution of the research sample according to age groups

QADA	Under 22 years old	22- 29	30 - 35	36 - 40	More than 40 years	Total
Al-Rusafa	50	50	40	40	60	240
	20.8	20.8	16.7	16.7	25	100
Al-Adhamia	6	7	5	6	5	29
	21	24	17	21	17	100
AL-Sader	15	12	13	10	13	63
	24	19	21	15	21	100
Al-Karkh	64	12	13	19	7	115
	55.7	10.4	11.3	16.5	6.1	100
Al-Kadhimiya	3	5	4	2	3	17
	17.6	29.4	23.5	11.8	17.7	100
Mahmudiya	5	7	8	5	6	31
	16.1	22.6	25.8	16.1	19.4	100
Abu-Gharib	3	2	3	2	1	11

***Corresponding author**

Zainab Mohammed Ameen,
College of Arts, University of Al-Iraqia, Baghdad, Iraq
e-mail: zainab.m.ameen@aliraqia.edu.iq

	27.3	18.2	27.3	18.2	9	100
Al-Tarmiya	4	3	6	3	2	18
	22.2	16.7	33.3	16.7	11.1	100
Al-Mada'in	3	5	3	6	5	22
	13.6	22.7	13.6	27.4	22.7	100
Total	153	103	95	93	102	546
Ratio	28.0	18.9	17.4	17.0	18.7	100

The composition of the population by sex is no less important than the age structure since, Sex structure is without doubt one of the most important and relevant features of population composition. The interaction of sex structure with the three demographic processes of fertility, mortality and migration is of paramount importance for demographers, but the importance of sex structure extends far beyond demography. The division of labor in traditional societies is heavily differentiated on the basis of sex. For that matter, sex differentiation of one form or another is found in all known human societies (Poston, DeSalvo and Kincannon, Encyclopedia of Life Support Systems), The population in the research sample is distributed according to gender, as shown in Table 3.

Table 3: The numerical and relative distribution of the research sample according to gender (male-female)

QADA	Male	%	Female	%	Total
Al-Rusafa	120	50.0	120	50.0	240
Al-Adhamia	15	51.7	14	48.3	29
AL-Sader	34	54.0	29	46.0	63
Al-Karkh	66	57.4	49	42.6	115
Al-Kadhimiya	8	47.1	9	52.9	17
Mahmudiya	17	54.8	14	45.2	31
Abu-Gharib	6	54.5	5	45.5	11
Al-Tarmiya	8	44.4	10	55.6	18
Al-Mada'in	10	45.5	12	54.5	22
Total	284	52.0	262	48.0	546

Table 5: The numerical and relative distribution of the research sample according to the educational qualification

QADA	Primary	%	high school	%	diploma	%	Bachelor	%	Higher than Bachelor's	%	Total
Al-Rusafa	137	57.1	68	28.3	12	%	21	8.8	2	0.8	240
Al-Adhamia	16	55.2	6	20.7	2	5.0	3	10.3	2	6.9	29
AL-Sader	30	47.6	19	30.2	7	6.9	6	9.5	1	1.6	63
Al-Karkh	40	34.8	38	33.0	20	11.1	15	13.0	2	1.7	115
Al-Kadhimiya	10	58.8	3	17.6	1	17.4	2	11.8	1	5.9	17
Mahmudiya	20	64.5	5	16.1	3	5.9	2	6.5	1	3.2	31
Abu-Gharib	7	63.6	2	18.2	1	9.7	1	9.1	0	0.0	11
Al-Tarmiya	10	55.6	4	22.2	1	9.1	2	11.1	1	5.6	18
Al-Mada'in	10	45.5	2	9.1	5	5.6	4	18.2	1	4.5	22
Total	280	51.3	147	26.9	52	22.7	56	10.3	11	2.0	546

The population of the research sample is distributed according to marital status into two categories (married and single) as shown in Table (4), from which it is noted that the highest percentage in the research sample is for married people in all districts of Baghdad Governorate.

***Corresponding author**

Zainab Mohammed Ameen,
College of Arts, University of Al-Iraqia, Baghdad, Iraq
e-mail: zainab.m.ameen@aliraqia.edu.iq

Table 4: Numerical and relative distribution of the research sample according to marital status

QADA	married	%	single	%	Total
Al-Rusafa	66	82.8	174	17.2	240
Al-Adhamia	24	81.0	5	19.0	29
AL-Sader	51	72.2	12	27.8	63
Al-Karkh	83	64.7	32	35.3	115
Al-Kadhimiya	11	67.7	6	32.3	17
Mahmudiya	21	72.7	10	27.3	31
Abu-Gharib	8	72.2	3	27.8	11
Al-Tarmiya	13	59.1	5	40.9	18
Al-Mada'in	13	53.1	9	46.9	22
Total	290	82.8	256	17.2	546

The modern pattern of childbearing is affected by the scientific background (the educational qualification) of the respondents. Most demographic studies have found a direct relationship between education and childbearing. While developing countries go through the demographic transition from high birth rates and high death rates to low birth rates and low death rates, female education becomes an important factor in determining population growth. It has been observed around the world that women who are given primary education and especially women who are given secondary education have a lower total fertility rate than women who have not received any formal schooling (Jain 1981. 577–595). Distribution of the research sample individuals are distributed according to the academic qualification into five categories, as shown in Table (5), where the highest percentage in the research sample was those who completed the primary school only. The study developed a scale consisting of 8 items to measure contemporary patterns of childbearing, and the answer levels were determined by (5) according to the five-point Likert scale. The results of the scale are summarized in Table 6.

STATISTICAL ANALYSIS

Artificial neural networks are called by this name because they are a network of internal communication units, as these units are inspired by the study of biological nervous systems and are symbolized by an acronym ANN, which is an acronym for the beginning of its words “Artificial Neural Networks” and it is one of the applications of artificial intelligence “AI”

The idea of artificial neural networks is the process of simulating data to reach a model for this data for the purpose of analysis, classification, prediction or any other processing without the need for a proposed model for this data.

Neural networks are defined as an interconnected group of neurons, as the outputs of a particular neuron are linked through weights to the inputs of another neuron to be the results of inputs to a third neuron, and then artificial neural networks can be defined as a structure of information that is processed in parallel, and it consists of a group of processing elements that are linked together by a group of connectors and the processing elements apply an activation function to their inputs to get the outputs (Ashour, 2018. P. 22)

Multi-layered Neural Network (MLP usins) analysis for the purpose of predicting contemporary trends in childbearing according to the dependent variable Class Child, i.e. the type of childbearing according to the number of children desired to be born, and it was of two types, the first: - the type that prefers to have 4 children or less, the second: - who prefers to have children 5 children or more, this classification of the scale is according to the age factor, where the two samples were classified into two categories: the first: - 40 years and less, the second: more than 40 years.

According to the SPSS Statistical Program Manual, "neural" is an adjective for the term nervous system. A neural network, also referred to as an artificial neural network, behaves similarly to neurons in the human brain. He acquires knowledge through the process of learning, just as the human mind learns through past events. It is a function of prediction variables (independent input vectors) that reduces the prediction error of the target variable (dependent output variable). A neural network is a highly parallel distributed processor that has a natural tendency to store empirical knowledge and make it available for use. Similar to the human brain in two respects (Guide SPSS, 2011)

Knowledge is gained through the network by the learning process.

Communication forces between neurons known as network weights are used to store knowledge.

Discuss the results of the analysis of multi-layer artificial neural networks

In neural networks, there is an input layer and a hidden layer, and the processing takes place to get the output layer out with the emergence of interfaces (interconnections) of the processors between the input and the hidden layers on the one hand, and between the hidden and output layers on the other hand. These links are loaded with weights and these weights are modified.

*Corresponding author

Zainab Mohammed Ameen,
College of Arts, University of Al-Iraqia, Baghdad, Iraq
e-mail: zainab.m.ameen@aliraqia.edu.iq

Iteratively through training algorithms, which then work to reduce the error and make the model predictions more accurate. The classification of the sample according to the desired number of children (the dependent variable) was descriptive statistics as in Table (7).

Table 7: Descriptive statistics for the dependent variable

Class _child					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	4 children or less	470	86.1	86.2	86.2
	5 or more children	75	13.7	13.8	100.0
	Total	545	99.8	100.0	
Missing	System	1	.2		
Total		546	100.0		

The cases are divided into the SPSS program according to the treatment models and as shown in Table 8.

Table 8: Summary of sample division according to processing models

Case Processing Summary			
		N	Percent
Sample	Training	360	66.8%
	Testing	179	33.2%
Valid		539	100.0%
Excluded		7	
Total		546	

The first section has 360 cases, with a percentage of 66.8%, and the second section is a test, and 179 cases have been placed, with a rate of 33.2%, excluding 7 cases. The network information is clear from Table 9.

Table 9: Network Information

Input Layer	Factors	1	Age class
	Covariates	1	Did you have the opportunity to work that you prefer and love? (V1)
		2	The security situation and its stability in my country affect the number of children desired to have (V2)
		3	In light of the current situation, I expect a good future for my children (V3)
		4	The financial resources that I receive each month are sufficient for me and meet my needs and the needs of my family (V4)
		5	I have an urgent desire to immigrate to search for a better future for me and my current or future children (V5)
		6	It is necessary for parents to have good knowledge and a real desire to family planning (V6)
		7	Do you think that the number of children should take into account the level of income and financial resources of the family? (V7)
		8	There is no age limit for parents to have children (V8)
	Number of Units ^a		10
Rescaling Method for Covariates		Standardized	
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1 ^a		6
Activation Function		Hyperbolic tangent	
Output Layer	Dependent Variables	1	Class _child
	Number of Units		2
	Activation Function		SoftMax
	Error Function		Cross-entropy

a. Excluding the bias unit

***Corresponding author**

Zainab Mohammed Ameen,
College of Arts, University of Al-Iraqia, Baghdad, Iraq
e-mail: zainab.m.ameen@aliraqia.edu.iq

The factor (two categories of cases according to age) and the constant (Bias) above. The layers in the middle are the hidden number 5 with the constant (Bias) on top, and after all the explicit manipulations of the interconnections we have two output layers that are either Class1 or Class2. The interfaces show us in the figure that those in blue have interlocking weights less than 0 and those in gray have interlocking weights greater than. The error percentage in the two models, Training and Testing, appears from the Model Summary in Table 10.

Table 10: Model Summary

Training	Cross Entropy Error	129.375
	Percent Incorrect Predictions	13.6%
	Stopping Rule Used	1 consecutive step(s) with no decrease in error ^a
	Training Time	0:00:00.11
Testing	Cross Entropy Error	63.195
	Percent Incorrect Predictions	15.1%
Dependent Variable: Class_child		
a. Error computations are based on the testing sample.		

The error rate in the Training model is 13.6% and in the Testing model it is 15.1%, and the thickness of the lines expresses the strength of the links, the thicker it is, the greater the binding strength and vice versa, and this is better shown in the parameter estimation Table (Estimation of interconnection coefficients in a neural network) e.g. Table 11.

Table 11: Estimation of parameters of interconnections in a neural network

Predictor		Predicted						Output Layer	
		Hidden Layer 1							
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	H(1:6)		
Input Layer	(Bias)	.205	-.188-	.410	.011	-.087-	.075		
	[Age_clacc=1]	.427	-.158-	-.431-	.616	-.058-	.461		
	[Age_clacc=2]	-.021-	-.053-	.326	.100	.479	-.106-		
	Did you have the opportunity to work that you prefer and love?	-.029-	.141	-.454-	-.452-	-.380-	.420		
	The security situation and its stability in my country affect the number of children desired to have	.022	.281	-.256-	.277	-.429-	.273		
	In light of the current situation, I expect a good future for my children	.156	-.418-	.123	-.211-	.310	.480		
	The financial resources that I receive each month are sufficient for me and meet my needs and the needs of my family	.464	-.054-	.481	.264	.054	-.245-		
	I have an urgent desire to immigrate to search for a better future for me and my current or future children	.134	-.122-	.263	.283	.372	-.192-		
	It is necessary for parents to have good knowledge and a real desire to family planning	-.092-	-.305-	.487	.053	.179	-.063-		
	Do you think that the number of children should take into account the level of income and financial resources of the family	.055	-.183-	-.449-	.081	.621	.312		
There is no suite age for parents to have children	-.053-	.004	-.160-	-.355-	-.170-	-.281-			
Hidden Layer 1	(Bias)						.817	-.834-	
	H (1:1)						.199	-.050-	
	H (1:2)						-.218-	.012	
	H (1:3)						-.380-	-.290-	
	H (1:4)						.604	-.391-	
	H (1:5)						.285	-.595-	
	H (1:6)						-.385-	-.249-	

The upper half of the table shows the connections between the In Put Layer and the Hidden Layer from 1-6, the lower half of the table shows the connections between the Hidden Layer and the Output Layer, and between Class child1 and Class child2, these connections are positive and some are negative. Processing function algorithms are responsible for extracting these links. One of the important outputs in the results of the analysis is the Classification table 12, which shows the validity of the predicted values. For example, in Observed, the predicted values in Training for the first category (4 children and less) are 310 correct predictions and only 2 false predictions. In the same way, we interpret the classification table data to show us the final percentage of correct predictions. It is 86.4% for the Training model and 84.9% for the Testing model, which are very well accepted rates.

***Corresponding author**

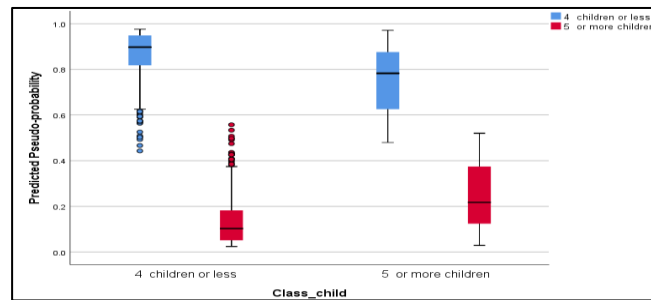
Zainab Mohammed Ameen,
College of Arts, University of Al-Iraqia, Baghdad, Iraq
e-mail: zainab.m.ameen@aliraqia.edu.iq

Table 12: Predicting of childbearing patterns

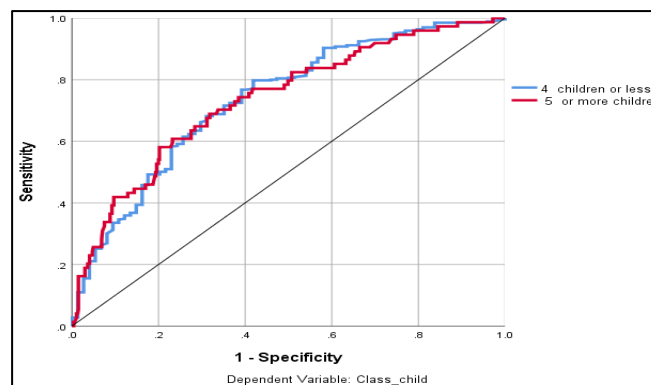
Classification				
Sample	Observed	Predicted		
		4 children or less	5 or more children	Percent Correct
Training	4 children or less	310	2	99.4%
	5 or more children	47	1	2.1%
	Overall Percent	99.2%	0.8%	86.4%
Testing	4 children or less	152	1	99.3%
	5 or more children	26	0	0.0%
	Overall Percent	99.4%	0.6%	84.9%

Dependent Variable: Class_child

The classification table (Table 12) can be translated into the following graph:-

**Figure 2: Predicting of childbearing patterns.**

It appears from the figure and from the left box plot that the cases that were predicted in the first reproductive category remained in the same observed category, i.e. a correct expectation, and the same is true in the interpretation of the rest of the figures. Another important output is to draw the future operating characteristics curve, Area Under the Curve (AUC) Figure (3) shows the plot of the receiver operating characteristics curve. The greater the area under the curve, the greater the probability that the positive test results will be correct. If the area under the curve reaches the correct one, this means that all positive test results are correct. This is what we find in the figure. Whereas, the figure confirms the model's ability to classify.

**Figure 3: Receiver Operating Characteristics Curve (Area Under the Curve).**

*Corresponding author

Zainab Mohammed Ameen,
College of Arts, University of Al-Iraqia, Baghdad, Iraq
e-mail: zainab.m.ameen@aliraqia.edu.iq

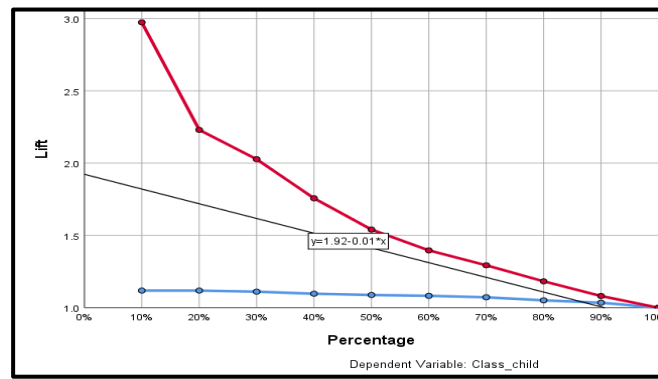


Figure 4: Expectations Percentages.

The median diagonal line in black is a line without prediction, but the blue line is for the expectations of the first category (4 children or less), the distance between the blue and black line is the percentage of gain (gain chart) in the case of using the predictive model, and the same is said on the distance between the black line, the second category is represented by the red line. The other output, which is no less important than the above, is the relative importance of the variables used in building the model, which explain the dependent variable, and it is clear from Table 13. When summing the Importance column, the result should be 1.

Table 13: The relative importance of the variables.

Independent Variable Importance		
	Importance	Normalized Importance
Age Class	.027	12.8%
Did you have the opportunity to work that you prefer and love?	.198	94.0%
The security situation and its stability in my country affect the number of children desired to have	.051	24.2%
In light of the current situation, I expect a good future for my children	.066	31.2%
The financial resources that I receive each month are sufficient for me and meet my needs and the needs of my family	.115	54.5%
I have an urgent desire to immigrate to search for a better future for me and my current or future children	.144	68.2%
It is necessary for parents to have good knowledge and a real desire to family planning	.082	38.9%
Do you think that the number of children should take into account the level of income and financial resources of the family	.211	100.0%
There is no suite age for parents to have children	.107	50.9%

A visual comparison of the importance of the variables is shown in Figure 5.

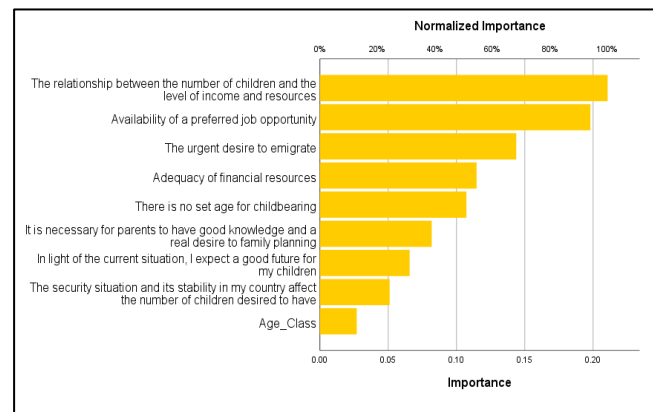


Figure 5: The relative importance of the study variables.

Figure 6 shows the two childbearing patterns that were created in the previous table, which is illustrated by the dominant pattern in the previous pattern (4 children and less).

***Corresponding author**

Zainab Mohammed Ameen,
College of Arts, University of Al-Iraqia, Baghdad, Iraq
e-mail: zainab.m.ameen@aliraqia.edu.iq

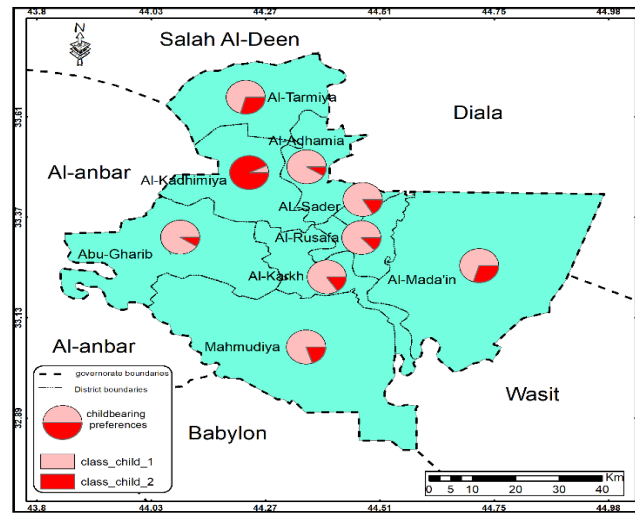


Figure 7: Preferred reproductive patterns (number of children) Among the study sample.

CONCLUSIONS

The study gave a very good model that shows the validity of the expected values, as the observations that the expected values in the (training model) for the first category (4 children or less) are 310 correct predictions and only two false predictions, and therefore the correct prediction rate is 99.4% which is a very high percentage. The correct predictions for the second category (5 children and over) were 86.4% for the training model and 84.9% for the test model.

1. The study gave very good predictive values for the classification of the dependent variable (number of children favorite).
2. The analysis of artificial neural networks can be used more broadly in geographical studies in general and demographic studies in particular in analyzing behaviors related to the main components of population growth (births - deaths - migration) these components that are the soul of population and demographic studies
3. The main purpose of forecasting or prediction is the possibility of controlling the phenomena under study, depending on the accuracy of prediction models. Therefore, the attempt to control the population increase in line with the capabilities of the country cannot be feasible unless the applications of artificial intelligence are adopted, especially artificial neural networks.

REFERENCES

- [1] Ashour, Marwan Abdel Hamid, Artificial Neural Networks and Time Series Prediction Methods, 1st Edition, Al Dhakera for Publishing and Distribution, Baghdad, 2018.
- [2] IBM SPSS Neural Networks 20 (2011) : Guide from IBM Software Group, chapter 5 ,Application Of Neural Network.
- [3] Ministry of Water Resources, General Authority for Survey, Map Production Division, Map of Baghdad Administrative Province at a scale of 1/500000, 2019.
- [4] Richard K.Thomas, Concepts, Methods and Practical Applications in Applied Demography (An Introductory Textbook), Springer,2018.
- [5] Dudley L. Poston, Jr., Bethany S. DeSalvo, and Heather Terrell Kincannon
- [6] Jain, A. (1981). The effect of female education on fertility: A simple explanation.
- [7] Demography, 18(4), 577-595. doi:10.2307/2060948.

*Corresponding author

Zainab Mohammed Ameen,
College of Arts, University of Al-Iraqia, Baghdad, Iraq
e-mail: zainab.m.ameen@aliraqia.edu.iq