An Artificial Intelligence Tool for Predicting the Number of Births According to the Characteristics of the Families

MostafaSayed Mostafa Abd – El Hameed El Misery*, Mohamed Solomon Jr**

Department of Statistics, Faculty of Economics and Political Science, Cairo University, Egypt

Department of Operations Research and Management, Faculty of Graduate Studies for Statistical Research, Cairo University, Egypt.

Abstract

In order to improve the economic and social qualities of Egyptian families and raise their standard of living, one of the most important goals of state policy is to identify the social, economic, and environmental characteristics as well as the factors that influence fertility and its decline trends at various levels. The current study aimed to identify the change in the number of children per woman in Egypt. An unconventional method that has not been used before in Egypt was applied to measure fertility levels in Egypt and predicts the factors that positively affect the reduction in the number of births per woman. There are many traditional methods that can measure fertility levels in Egypt and have been widely used in many literatures, but the current study uses the artificial intelligence method, specifically neural networks, to predict demographic and socio-economic variables that will have a significant impact on reducing the number of children per woman during her reproductive life. The artificial neural network methodology will analyze and predict the most important factors that can affect the number of children in Egypt. Data from - Health Survey for the Egyptian Households 2021 were used. The results of the study indicate that partner's occupation, wealth index quintile, ideal number of children, residency region, and partner's educational attainment are the most significant factors influencing the number of births.

KEYWORDS: Artificial Neural Networks, fertility level, Predicting, Multi-layer Perceptron algorithm, artificial intelligence.

1. Introduction

The number of births plays a crucial role in determining a country's population trends, affecting its total size, demographic structure, and makeup. It signifies the number of live births among women of reproductive age within marriages, as highlighted by Kiser and Hossain (2019). The increase in the total number of children born has been a major factor driving the swift growth of the global population (Gaigbe-Togbe et al., 2022). According to the 2015 revision of world population prospects, the global average fertility rate is currently 2.5 children per woman. Africa, which has a fertility rate of 4.7 children per woman, remains a significant contributor to the worldwide number of births (Lam and Leibbrandt, 2019). This trend is influenced by both immediate and significant factors, serving as a key determinant of a country's population dynamics (UN. Population, United Nations. Un, 2019). The total number of children ever born is a crucial aspect that influences the size, composition, and distribution of a nation's population. This measure is closely linked to numerous economic, social, and demographic elements, all of which significantly affect and predict this figure. Therefore, accurately forecasting the number of children ever born and understanding the factors that influence it is vital for the Egyptian government to develop effective policies and initiatives.

This study aims to predict the factors that influence the number of births in Egypt by employing a Neural Network as an artificial intelligence method. Over the years, the nation has experienced a decrease in fertility rates, declining from 6.8 births per woman in 1955 to 3.3 births per woman in 2020 (Worldometer, 2020). Zaki (2020) performed a descriptive analysis that outlined trends in the reliance on family planning services in the early 2000s. A significant finding from this study was the ongoing preference shown by women when choosing a service provider, even among those who indicated using services from multiple sectors.

Fertility rates have a significant impact on population growth. Fertility rates in Egypt increased beginning in 2006 and peaked in 2014 at 3.5 children per woman. After that, the number of children per woman decreased gradually to 3.4 in 2017 and then more quickly to 3.1 by the end of the reproductive years in 2018. From 3.5 in 2014 to 3.4 in 2017 and then to 3.1 in 2018, Egypt's overall fertility rate declined. The overall fertility rate decreased by a significant 11.1% in 2018 compared to a slight 3.4% decrease in 2017. At almost 8%, the decline from 2017 to 2018 was likewise noteworthy (Sayed, 2019). In order to predict the number of parameters that affect determining the number of births in Egypt using data from the 2021 Population Health Survey, this study will employ an unconventional approach, which is one of the most significant modern methods used in prediction today. It is also one of the most significant methods of artificial intelligence, specifically neural networks.

2. Related work

In order to complete a task, the scientific field of artificial intelligence (AI) imitates human conduct and intelligence (Ramakrishnan et al., 2021). Artificial intelligence can streamline decision-making processes and improve medical treatment. Within the healthcare sector, AI is utilized for diagnostics, early detection, monitoring, and predictive modeling. It enhances human intelligence by improving the accuracy of diagnostics, facilitating predictive analytics, assisting in the development of new medications, and increasing the efficiency of administrative tasks (Ramakrishnan et al., 2021). Artificial neural networks and other machine learning models are a subset of artificial intelligence. Machine learning is like teaching a computer to learn on its own. There are different types of machine learning, like supervised, semi-supervised, unsupervised, and reinforced. Basically, the computer uses special models and algorithms to figure things out. With supervised learning, the computer learns from labeled data, while unsupervised learning helps the computer find patterns in data that doesn't have labels. (Togunwa et al., 2023).

Blanco et al. (2016) for the Data Structures, he focuses on forecasting student results. The model's architecture is composed of three layers, and each neurone has a unique activation function. The sigmoidal hyperbolic tangent function is employed by the input and hidden layers. To get the highest range amplitude in the output interval, the authors decided to use the linear function for the output layer. This also made it simpler to examine the data. The prediction effectiveness of the first subject in this investigation was above 78%, while that of the second subject was 75%Because both in-person and remote learners regularly need help to improve their performance, there was a need to improve AI-based educational systems (Sekeroglu et al., 2019). Therefore, models of long/short-term memory (LSTM) and backpropagation (BP) neural networks were used. The former utilizes a gradient descent strategy during the learning phase, propagating errors to adjust weights and minimize error values. In contrast, the latter retains past neural network inputs to achieve more accurate outcomes. Furthermore, According to Amazona and Hernandez (2019), it's best to use a deep learning neural network for making predictions. 98% recall, 97% F1-score and 98% precision were the deep learning results.

Daud et al. (2017) implemented a supervised learning approach to determine if students would follow through with their study plans or decide to drop out. They focused on the Support Vector Machine (SVM) model, which achieved an impressive 86% on the F1-score test, delivering the most favorable results.

Ma et al. (2018) also used the Support Vector Machine supervised learning model to forecast online student passing rates. This model's accuracy was 95% using the grid search technique and 50% pass/50% fail data.

3. Artificial Neutral Networks

An artificial neural network (ANN) functions similarly to a human brain, utilizing numerous components known as neurons to analyze and process data. Its remarkable capabilities have been demonstrated across various engineering fields (Moayedi et al., 2019; Koopialipoor et al., 2020; Gowid et al., 2019). One of the most impressive aspects of artificial neural networks is their ability to learn independently, which enables them to tackle complex challenges, even when faced with vast amounts of information (Feng L. and Lu J., 2019). An artificial neural network (ANN) consists of a vast number of interconnected artificial neurons, functioning as processing units. These units are categorized into input and output sections. The input units collect data utilizing an internal weighting system. Subsequently, the neural network within the hidden layer processes the acquired information, with the objective of generating an accurate output report.

To fine-tune these outputs, a technique known as backpropagation is employed, which helps minimize error values (Walczak S. and Cerpa N., 2003). This process results in the production of accurate outputs in the reports generated. Artificial neural networks (ANN) are generally categorized into two primary types: feed-forward networks and feedback (recurrent) networks (Chiang et al., 2004). In a feed-forward neural network, signals flow in a straight line from input to output, whereas feedback networks allow signals to move in both directions, creating potential loops in the connections, as illustrated in Figure 1 (Pekel E. and Kara S., 2017). The feed-forward neural network (FFNN) stands out as the most basic and widely utilized form of artificial neural network. It can be structured as a single layer, a multi-layer perceptron (MLP), or a radial basis function (RBF) (Pekel E. and Kara S., 2017).

3.1 Multilayer Perceptron (MLP)

The multilayer perceptron (MLP) is the most prevalent and widely utilized type of feed-forward neural network. (Al-Shamisi et al., 2013). Typically, a multilayer perceptron comprises three distinct layers: an input layer, a hidden layer, and an output layer (Pyo et al., 2017). Input signals x_i $(i = 1,2, \ldots, n)$ are transmitted to the neurons in the hidden layer by the neurons in the input layer, as illustrated in Figure 2. Each neuron in the hidden layer, denoted as (j) assigns weights to its incoming signals based on the respective connection strengths (w_{ji}) and subsequently sums the received signals (x_i). Finally, the outputs y_j of the hidden layer are then computed by the neurons using a summation function (f) (Al-Shamisi et al., 2013):

$$y_j = f\left(\sum_{i=1}^{n} w_{ji}\, x_i\right) \quad where\ f\ is\ threshold \qquad (1)$$

The way the output from neurons is calculated plays a crucial role in the training of the output layer in a Multi-Layer Perceptron. This procedure is dependent on the backpropagation algorithm, which functions as a method for supervised learning (Riedmiller, 1994). The training process is centered on adjusting the connection weights among the neurons.

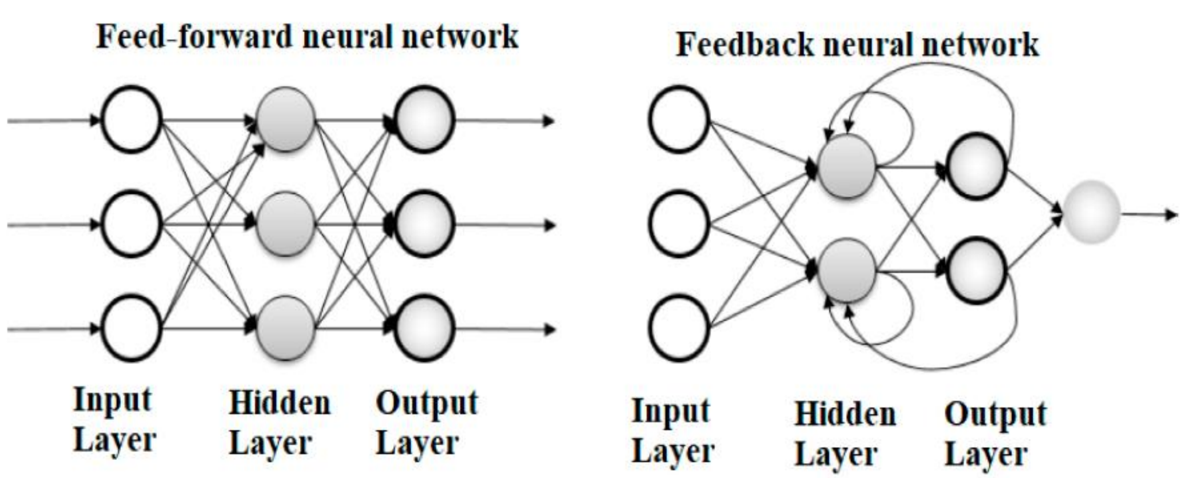

Figure 1. The architecture of ANN categorized into two main types: feed-forward neural networks and feedback (or recurrent) neural networks.

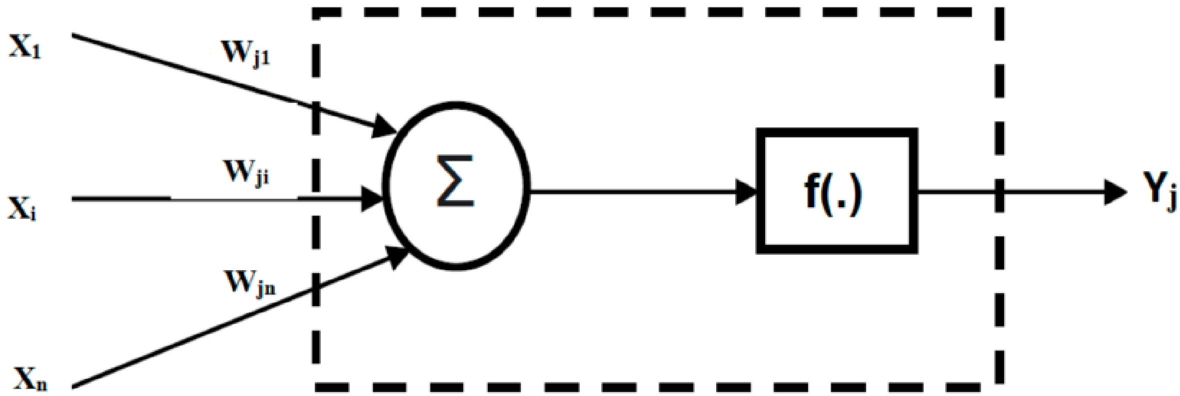

Figure 2. Process in multilayer perceptron.

3.2 Artificial Neural Network for Prediction

The artificial neural network technique was introduced in the 1980s as a distinct branch of artificial intelligence (AI) and has since been effectively utilized to address complex problems (Grossberg, 1988; Hertz et al., 1991).

As an advanced analytical tool, artificial neural networks (ANNs) are capable of accurately predicting output patterns by leveraging insights gained from prior learning experiences. Once properly trained, these neural networks identify similarities and generate predicted output patterns when presented with new data. The development of an ANN typically involves three steps: defining the network architecture, training, and testing. During the training phase, the neural network engages with a database to acquire knowledge, subsequently undergoing testing to yield more reliable results. The feed-forward back-propagation technique enables neural networks to effectively model the relationships between input and output patterns (Bahrami et al., 2011).

4. Methodology

4.1. Data

The study utilizes data from the Health Survey for the Egyptian Households 2021 [EGY-CAPMAS-EFHS2021], which provides comprehensive information on the characteristics of husbands and wives, as well as data on births. This survey features a self-weighted sample, enhancing its reliability. Additionally, it includes detailed information regarding birth history. The analysis specifically targets a sample of women aged 20-49 years who have been married at least once, have given birth to at least one child, and have only been married once. Based on these criteria, the study sample comprises 12004 women. The dataset contains 12004 rows and 33 columns, featuring 31 categorical variables and 2 numeric variables. The primary target variable is the total number of children ever born, categorized into three classes: 2, 3, or 4 children. Table 1 presents all features along with the target variable utilized in this analysis.

Table 1: all features of data with one target.

All the features with target (total children ever born)		
Region of residence	visited private doctor or clinic in the last 6 months	Who usually decides on spending husband's/partner's earnings
Type of place of residence	Ideal number of children	Final say on: Own health care
Wealth index quintile	Ever heard of premarital examination	Final say on: Making large household purchases
Current age of respondent	Ever breastfed	Final say on: Visits to family
Current marital status	Husband was relative through blood or marriage	Owns own house alone or jointly
Ever attended school	Partner ever attended school	Owns land alone or jointly
Total children ever born	Partner's level of education	Covered by any health insurance
Sex of child	Partner's occupation	Highest educational level
Single or Multiple Birth	Respondent's occupation	Currently using any method
Child is still alive	Paid in cash or kind	Ever used a contraceptive method
Visited health facility in last 12 months	Earns more than partner	Age of husband

4.2. Variables Utilized in the Construction of the ANN

In this section, we will discuss the various variables employed in the development of the artificial neural network. These variables play a crucial role in determining the network's architecture, performance, and overall effectiveness in learning and making predictions. The selection and definition of these variables are essential for the successful training and validation phases of the ANN.

Independent variables:

In table 1, all features in the table are independent except the total children ever born.

Dependent variable:

The total children ever born, which is dichotomous with three values: 2, 3 and 4. The values 2, 3 and 4 indicate the families of the study have 2 kids, 3 kids and 4 kids, respectively.

4.3. Design and Setup of the ANN

This section outlines the design and configuration of the artificial neural network. Key components of the setup include the architecture, which defines the number of layers and the number of neurons within each layer, as well as the choice of activation functions that influence the network's ability to learn complex patterns. Additionally, the training parameters, such as the learning rate, batch size, and number of epochs, are determined to optimize the learning process. The data preprocessing steps, including normalization and data splitting for training and testing, are also critical to ensure the effectiveness of the ANN. A systematic approach to the design and setup is vital for achieving reliable and accurate performance in the desired application.

The Multilayer Perceptron (MLP) Module of Orange 3 was employed to construct the neural network model and evaluate its accuracy. Orange 3 is a component-based data mining software that provides a wide array of data visualization, exploration, preprocessing, and modeling techniques. It features an intuitive user interface ideal for beginners, while also catering to advanced users with a module for Python programming. In the MLP neural networks, the back-propagation learning algorithm is utilized, leveraging sklearn's implementation of the Multi-layer Perceptron. This enables the model to effectively learn both non-linear and linear relationships. Additionally, the model is optimized to minimize the log-loss function using either the LBFGS optimization algorithm or stochastic gradient descent, thereby enhancing its predictive capabilities.

The data were randomly split into training (70%) and testing (30%) sets. The training dataset is utilized to establish the weights and construct the model, while the testing dataset is employed to evaluate errors and mitigate the risk of overfitting during the training phase. Before training commenced, all covariates were normalized by centering them to the mean and scaling them to a standard deviation of 1.

For the hidden layer, hyperbolic tangent (or tanh) was used as activation function. The activation of the *j*th output neuron is $O_j = \tanh(S_j) = \frac{e^{S_j} - e^{-S_j}}{e^{S_j} + e^{-S_j}}$ takes real numbers as arguments and returns real values between -1 and 1. The softmax function was used as activation function For the output layer. The activation of the *j*th output neuron is $O_j = \sigma(S_j) = \frac{e^{S_j}}{\sum_{k=1}^{m} e^{S_k}}$, where m is the number of output neurons. The softmax function accepts real numbers as inputs and transforms them into real values ranging between 0 and 1, with the sum of these values equating to 1. Given that the sum of the output activations is 1, the softmax layer can be conceptualized as a probability distribution. Consequently, the value O_j can be interpreted as the network's estimated probability (or pseudo-probability) regarding the classification of the input x .

Limited-memory BFGS (L-BFGS or LM-BFGS) is an optimization algorithm classified within the family of quasi-Newton methods. It serves as an approximation of the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm while requiring a restricted amount of computer memory. This feature makes it particularly suitable for large-scale optimization problems where memory efficiency is crucial (Liu and Nocedal, 1989). This algorithm is widely recognized for its effectiveness in parameter estimation within machine learning frameworks (Malouf, 2002; Andrew and Gao, 2007). The primary goal of the algorithm is to minimize a differentiable scalar function over unconstrained values of the real vector.

In a manner akin to the original BFGS algorithm, L-BFGS employs an estimate of the inverse Hessian matrix to navigate through the variable space. However, contrary to BFGS, which stores a dense approximation of the inverse Hessian proportional to the number of variables in the problem, L-BFGS retains only a limited set of vectors that implicitly represent this approximation. This characteristic results in a linear memory requirement, making the L-BFGS method especially beneficial for optimization problems involving a substantial number of variables. Instead of the inverse Hessian $\boldsymbol{H}_k$, L-BFGS maintains a history of the past *m* updates of the position $\boldsymbol{x}$ and gradient $\nabla f(\boldsymbol{x})$, where generally the history size *m* can be small (often). These updates are used to implicitly do operations requiring the $\boldsymbol{H}_k$ - vector product.

Stopping Rules:

200 iterations

When the softmax activation function is applied to the output layer, Orange 3 utilizes the cross-entropy error function rather than the squared error function, which is typically applied with other activation functions. The cross-entropy error function for an individual training example is mathematically expressed by the following formula:

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

where m is the number of output nodes/classes, t_j is the target value of output node j and O_j is the actual output value of output node j. The backpropagation algorithm, in each iteration (or epoch), calculates the gradient of the training error as $\frac{\partial E}{\partial w_{hj}} = (O_j - t_j)x_h$ for the weights linking nodes in the hidden and nodes in the output layer of the network, and $\frac{\partial E}{\partial w_{hj}} = \left(O_j - t_j\right)x_h w_{hj}(1 - x_h)x_i$ for the weights of links of the hidden and the input layer.

For each training example, every weight w_{ih} is updated by adding to it $\Delta w_{ih} = \gamma \frac{\partial E}{\partial w_{ih}}$, thus taking the new value: $\Delta w_{ih} \leftarrow w_{ih} + \Delta w_{ih}$.

5. Results of Artificial Neural Network

In this study, the objective was to examine whether a Multilayer Perceptron (MLP) neural network can assist instructors in accurately predicting the total number of children ever born, with possible outcomes of 2, 3, or 4. This investigation was conducted by analyzing data obtained from the Health Survey for Egyptian Households 2021 [EGY-CAPMAS-EFHS2021] utilizing the MLP Classifier.

(*hidden_layer_sizes*)*=(100,)*, ***activation****='tanh'*, ***solver****='L-BFGS-B'*, ***alpha****=0.0001*, ***batch_size****='auto'*, ***learning_rate****='constant'*, ***learning_rate_init****=0.001*, ***power_t****=0.5*, ***max_iter****=200*, ***shuffle****=True*, ***random_state****=None*, ***tol****=0.0001*, ***verbose****=False*, ***warm_start****=False*, ***momentum****=0.9*, ***nesterovs_momentum****=True*, ***early_stopping****=False*, ***validation_fraction****=0.1*, ***beta_1****=0.9*, ***beta_2****=0.999*, ***epsilon****=1e-08*, ***n_iter_no_change****=10*, ***max_fun****=15000***)**

In the Orange program, the MLP Classifier features a default configuration of 100 neurons in the hidden layer. The automatic architecture selection process determined the optimal number of nodes for this hidden layer, while the output layer was designed with three nodes to represent the dependent variable, total children ever born. The activation function utilized for the hidden layer was L-BFGS-B, whereas the softmax function was employed for the output layer. Due to the utilization of the softmax function, the cross-entropy loss function was implemented.

In Figure 1, the performance statistics are presented prior to any enhancements. The area under the receiver operating characteristic curve (AUC) quantifies the model's ability to differentiate between positive and negative classes. Classification accuracy (CA) reflects the proportion of instances that were correctly classified within the dataset. The F1 score is defined as the weighted harmonic mean of precision and recall. Precision (Prec) is the ratio of true positives to the total instances identified as positive; for instance, it indicates the proportion of instances with two children that were accurately classified as such. Recall measures the ratio of true positives to all actual positive instances in the dataset; for example, it accounts for the number of correctly identified sick individuals among all those diagnosed as sick. The Matthews correlation coefficient (MCC) evaluates both true and false positives and negatives, serving as a balanced metric applicable even when class sizes vary significantly. As shown in Figure 1, the performance statistics illustrate the model's effectiveness prior to improvements, which are subsequently detailed in Figure 2.

Scores

Model	AUC	CA	F1	Prec	Recall	MCC
Neural Network	0.998	0.976	0.976	0.976	0.976	0.963

Figure 1: the numbers of performance statistics before improvement

Scores

Model	AUC	CA	F1	Prec	Recall	MCC
Neural Network	1.000	1.000	1.000	1.000	1.000	1.000

Figure 2: the numbers of performance statistics after improvement

Tables 2 and 3 display a classification table, commonly referred to as a confusion matrix, for the categorical dependent variable indicating the total number of children ever born. The confusion matrix provides a detailed account of the number of instances corresponding to both the predicted and actual classes. The arrangement of elements within the matrix enables the mapping of corresponding instances to the output signal. This facilitates the identification of specific instances that were misclassified, as well as the nature of the misclassifications. For example, in table 2 in row 1, we can see the model predicted correctly 4658 in 2 children while predicted incorrectly 95 and 32 in 3 and 4 children, respectively. In row 2, the model predicted correctly 4440 in 3 children while predicted incorrectly 62 and 38 in 2 and 4 children, respectively. In row 3, the model predicted correctly 2616 in 4 children while predicted incorrectly 25 and 38 in 2 and 3 children, respectively.

In table 3, the confusion matrix after the improvement occurred on the neural network model is presented. In row 1, we can see the model predicted correctly 4785 in 2 children. In row 2, the model predicted correctly 4539 in 3 children while predicted incorrectly 1 in 2 children. In row 3, the model predicted correctly 2678 in 4 children while predicted incorrectly 1 in 3 children.

Table 2: confusion matrix (before improvement)

		Predicted			
Actual		**2**	**3**	**4**	
	2	4658	95	32	**4785**
	3	62	4440	38	**4540**
	4	25	38	2616	**2679**
		4745	**4573**	**2686**	

Table 3: confusion matrix (after improvement)

		Predicted			
Actual		**2**	**3**	**4**	
	2	4785	0	0	**4785**
	3	1	4539	0	**4540**

	4	0	1	2678	**2679**
		4786	**4540**	**2678**	

The ROC curve serves as a graphical representation of the relationship between sensitivity and specificity, illustrating classification performance across all possible cutoff thresholds. Figure 3 features the chart that depicts both sensitivity and specificity (defined as 1 minus the false positive rate) and the mean true positive rate, calculated from the combined training and testing samples. The 45-degree line that extends from the upper right corner to the lower left of the chart signifies the condition of random class selection. The farther the curve deviates from this 45-degree baseline, the higher the accuracy of the classification.We utilized the mean true positive rate to obtain vertical averages of the curves, thereby displaying the corresponding confidence intervals for all classes.

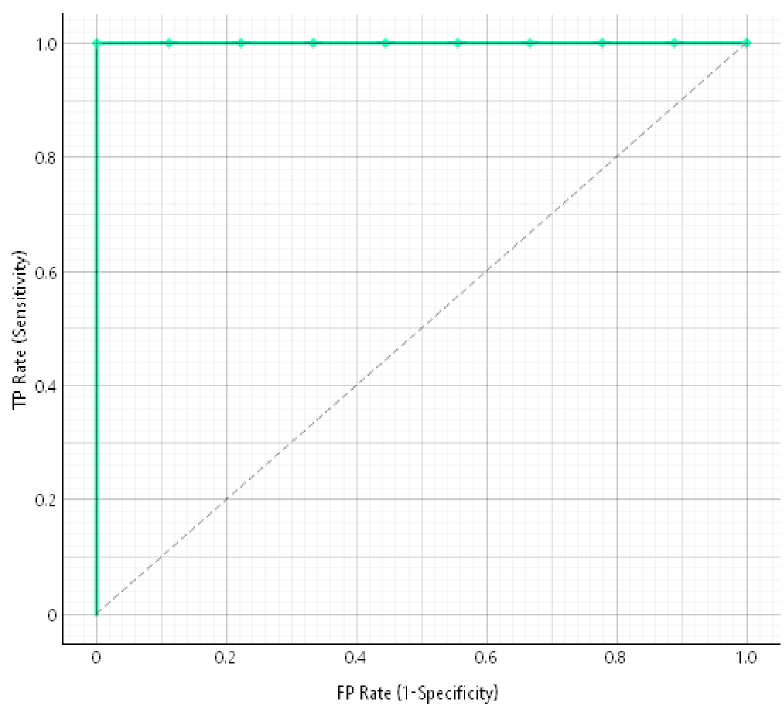

Figure 3: ROC curve

The area value indicates that when selecting a family from the two-children category, a family from the three-children category, and another family from the three-children category at random, there exists a probability of 1.00 that the model-predicted pseudo-probability for the first family being classified in the two-children category will be greater than the model-predicted pseudo-probability for the second family classified in the three-children category, and so on.

In figure 4, the Feature Importance widget elucidates the workings of the classification and neural network model. It accepts a trained model along with reference data as input. Using this information, it computes the contribution of each feature to the model's predictions by measuring the increase in prediction error that occurs after permuting the values of a feature. This process disrupts the relationship between the feature and the target variable, allowing for a clearer understanding of each feature's impact. There are 32 features in our study; we displayed top 5 features only to demonstrate the idea.

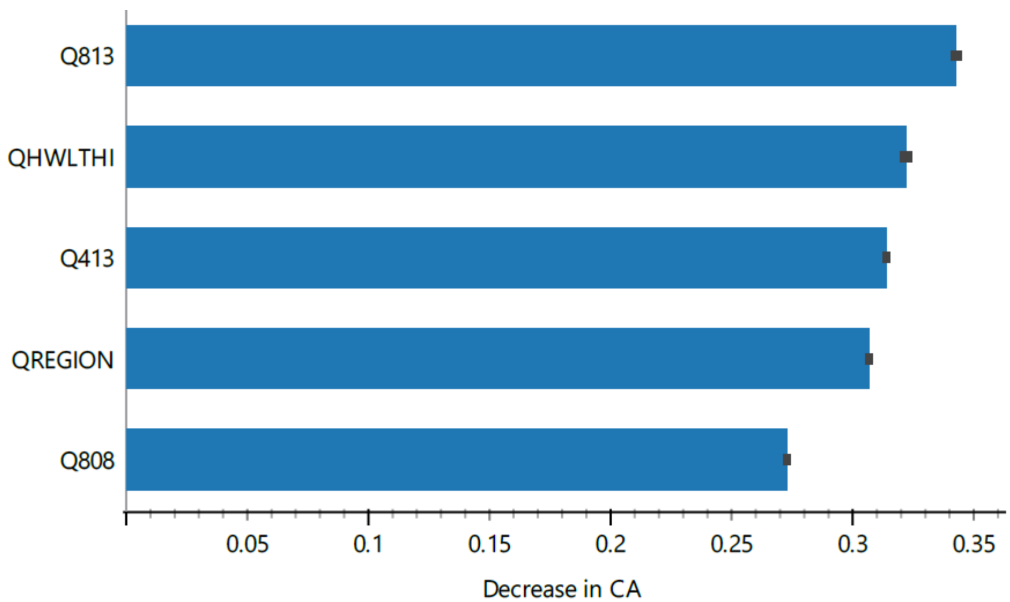

Figure 4: independent variable importance chart for top 5 features only

The current study is concerned with using neural networks to predict the variables that have a significant impact on reducing the number of births per woman. Through the analysis, it was found that from the previous figure, the most important variables that affects the number of births is (Q813: Partner's occupation, QHWLTHI: Wealth index quintile, Q413: Ideal number of children, QREGION: Region of residence, and Q808: Partner's level of education. This is consistent with the results of many previous studies, such as: Mason and Taj, 1987; Krafft, 2016; Jejeebhoy, 1995;

Dyson and Moore, 1983; McDonald, 2000a) indicate that a significant factor influencing fertility is the socio-economic status of women, which is predominantly assessed in the literature through education and/or participation in the workforce. Also, the wealth index was significantly associated with the number of children born. (Ahmed, & Ali, 2015).

6. Conclusion

Egypt's total fertility rate has declined significantly recently. From 5.3 children in 1980 to 3.5 children in 2000, the rate began to decline. It fell by 41% to around 3.1 children in 2005, then increased to 3 children in 2008 before rising again to 3.5 in 2014. The Central Agency for Public Mobilization and Statistics reported that Egypt's total fertility rate rose again to 3.24 in 2020. The "2 is Enough" initiative, launched in 2018, caused the birth rate to drop to 2.9 in 2022. The country has made efforts to reduce fertility rates, but the decline has varied by region. The demographic and socioeconomic characteristics of women in each region affect fertility patterns and levels. The new study attempts to evaluate programs such as "Decent Life," which targets high-needs communities in informal settlements and rural areas. The aim is to develop appropriate population plans, propose solutions that help achieve the goal, and direct programs to reach the target fertility rate of 2.1 children per woman by 2030. This study aims to clarify the main elements that can contribute to or benefit from reducing the number of children per woman. The network analysis method, which is an unusual approach that has not been used before in analysis, was used in order to benefit from the characteristics of artificial intelligence in the study, which is used in prediction. The network analysis method was used to predict the main factors that can contribute to a family having two children.

The "Two is Enough" program aims to decrease the fertility rate from 3.5 children per woman to 2.4 by the year 2030. This initiative is specifically directed at economically disadvantaged families who participate in the Takaful program, an income support initiative. The reproductive choices of both men and women are largely determined by the social, educational, cultural, and economic circumstances they encounter. Furthermore, policies that focus on improving women's conditions—by enhancing their income, advancing their education, and promoting their empowerment—are critical for achieving the necessary fertility reductions to effectively manage population growth in a voluntary and sustainable manner. In this study, the aim was to examine whether a multilayered neural network (MLP) could help teachers correctly predict the outcome of the total number of children born (2, 3, or 4), by analyzing data from the 2021 Egyptian Household Health Survey [EGY-CAPMAS-EFHS2021]. The Orange 3 MLP module was used to build the neural network model and test its accuracy. This model optimizes the logarithmic loss function using LBFGS or stochastic gradient descent.

References

1- Ahmed, H. M. M., & ALI, H. M. H. (2015). Using count Regression models to determinate factors influencing fertility of Sudanese women. Glob. Econ. Manag. Bus. Res, 6(4), 277–285.

2- Amazona, M. V., and Hernandez, A. A. (2019). "Modelling student performance using data mining techniques." in *Proceedings of the 2019 5th International Conference on Computing and Data Engineering—ICCDE' 19*. New York, NY, USA: Association for Computing Machinery. 36–40. doi: 10.1145/3330530.3330544

3- Al-Shamisi, M.H.; Assi, A.H.; Hejase, H.A. (2013). Artificial Neural Networks For Predicting Global Solar Radiation In Al Ain City – UAE, Int. J. Green Energy, 10, 443–456.

4- Andrew, G; Gao, J (2007). "Scalable training of L_1-regularized log-linear models". Proceedings of the 24th International Conference on Machine Learning. doi:10.1145/1273496.1273501. ISBN 9781595937933. S2CID 5853259.

5- Blanco, J., Lovelle, S., Fernandez, R., and Perez, E. (2016). Predicción de resultados académicos de estudiantes de informática mediante el uso de redes neuronales. *Ingeniare. Revista chilena de ingeniería* 24, 715–727. doi: 10.4067/S0718-33052016000400015.

6- Bahrami, A.; Monjezi, M.; Goshtasbi, K.; Ghazvinian, A. (2011). Prediction of rock fragmentation due to blasting using artificial neural network, Eng. Comput., 27, 177–181.

7- Central Agency for Public Mobilization and Statistics (2021). Arab Republic of Egypt - Health Survey for the Egyptian Households 2021.

8- Chiang, Y.-M.; Chang, L.-C.; Chang, F.-J. (2004). Comparison of static-feedforward and dynamic-feedback neural networks for rainfall-runoff modeling, J. Hydrol, 290, 297–311.

9- Daud, A., Aljohani, N. R., Abbasi, R. A., Lytras, M. D., Abbas, F., and Alowibdi, J. S. (2017). "Predicting student performance using advanced learning analytics." in *Proceedings of the 26th International Conference on World Wide Web Companion—WWW '17 Companion*.

Republic and Canton of Geneva, CHE: International World Wide Web Conferences Steering Committee. 415–421. doi: 10.1145/3041021.3054164.

10- Dyson, T., Moore, M. (1983). On kinship structure, female autonomy, and demographic behavior in India. Population and Development Review, 9, pp. 35–60.

11- Gaigbe-Togbe V., L. Bassarsky, D. Gu, T. Spoorenberg, L. Zeifman, (2022) World Population Prospects, Available at: [Internet]. 1–54 pp. Available from: https://population.un.org/wpp/publications/files/wpp2019_highlights.pdf.

12- Feng, L.-H.; Lu, J. (2010). The practical research on flood forecasting based on artificial neural networks, Expert Syst. Appl., 37, 2974–2977.

13- Gowid, S.; Mahdi, E.; Alabtah, F. (2019). Modeling and optimization of the crushing behavior and energy absorption of plain weave composite hexagonal quadruple ring systems using artificial neural network, Compos. Struct. 229, 111473.

14- Grossberg, S. (1988). Nonlinear neural networks: Principles, mechanisms, and architectures, Neural Networks, 1, 17–61.

15- Hertz, J.; Krogh, A.; Palmer, R.G.; Horner, H. (1991). Introduction to the Theory of Neural Computation, Phys. Today, 44, 70.

16- Jeejeebhoy, S. (1995). Women's Education, Autonomy, and Reproductive Behaviour: Experience from Developing Countries. Oxford University Press, Oxford.

17- Koopialipoor, M.; Fahimifar, A.; Ghaleini, E.N.; Momenzadeh, M.; Armaghani, D.J (2020). Development of a new hybrid ANN for solving a geotechnical problem related to tunnel boring machine performance, Eng. Comput., 36, 345–357.

18- Kiser H., Hossain M.A., (2019) Estimation of number of ever born children using zero truncated count model: evidence from Bangladesh Demographic and Health Survey, Health Inf. Sci. Syst. 7 (1) 1–10.

19- Krafft, C. (2016). Why is fertility on the rise in Egypt? the role of women's employment opportunities. Economic Research Forum Working Paper Series 1050, Economic Research Forum, Giza, Egypt.

20- Lam D., Leibbrandt, J, (2019) The Demography of the Labor Force in Sub-saharan Africa: Challenges and Opportunities, GLM|LIC, IZA Institute of Labor Economics, UKaid,.

21- Liu, D. C.; Nocedal, J. (1989). "On the Limited Memory Method for Large Scale Optimization". Mathematical Programming B. 45 (3): 503–528. CiteSeerX 10.1.1.110.6443. doi:10.1007/BF01589116. S2CID 5681609.

22- Moayedi, H.; Mosallanezhad, M.; Rashid, A.S.A.; Jusoh, W.A.W.; Muazu, M.A (**2019**). A systematic review and meta-analysis of artificial neural network application in geotechnical engineering: Theory and applications, Neural Comput. Appl. 32, 495–518.

23- Moayedi, H.; Mosallanezhad, M.; Rashid, A.S.A.; Jusoh, W.A.W.; Muazu, M.A. (2019). A systematic review and meta-analysis of artificial neural network application in geotechnical engineering: Theory and applications. Neural Comput. Appl., 32, 495–518.

24- Ma, X., Yang, Y., and Zhou, Z. (2018). "Using machine learning algorithm to predict student pass rates in online education." in Proceedings of the 3rd International Conference on Multimedia Systems and Signal Processing - ICMSSP '18. New York, NY, USA: Association for Computing Machinery. 203–207. doi: 10.1145/3220162.3220188.

25- Macdonald, P. (2000a). Gender equity in theories of fertility transition. Population and Development Review, 26, no. 3, pp. 427–439

26- Malouf, R. (2002). "A comparison of algorithms for maximum entropy parameter estimation". Proceedings of the Sixth Conference on Natural Language Learning (CoNLL-2002). pp. 49–55. doi:10.3115/1118853.1118871.

27- Mason, K., TAJ, A. (1987). Differences between women's and men's reproductive goals in developing countries. Population and Development Review, 13, pp. 611–638.

28- Pekel, E.; Kara, S.S. (2017). A Comprehensive Review for Artifical Neural Network Application to Public Transportation, Sigma J. Eng. Nat. Sci., 35, 157–179.

29- Pyo, S.; Lee, J.; Cha, M.; Jang, H. (2017). Predictability of machine learning techniques to forecast the trends of market index prices: Hypothesis testing for the Korean stock markets. PLoS ONE **2017**, 12, e0188107.

30- Ramakrishnan, R., Rao, S., and He, J. (2021), Perinatal health predictors using artificial intelligence: A review, *Women's Health*, 17, 1–7, DOI: 10.1177/17455065211046132.

31- Riedmiller, M. (1994). Advanced supervised learning in multi-layer perceptrons—From backpropagation to adaptive learning algorithms, Comput. Stand. Interfaces, 16, 265–278.

32- Sayed, Hussein Abdul Aziz (2019). Trends of Fertility Levels in Egypt in Recent Years. Preliminary assessment based on population census data for 2017, United Nations Population Fund, Cairo

33- Sekeroglu, B., Dimililer, K., and Tuncal, K. (2019). "Student performance prediction and classification using machine learning algorithms." in *Proceedings of the 2019 8th International Conference on Educational and Information Technology—ICEIT 2019.* New York, NY, USA: Association for Computing Machinery. 7–11. doi: 10.1145/3318396.3318419

34- Togunwa, T.O, Babatunde, A.O and Abdullah, K. (2023). Deep hybrid model for maternal health risk classification in pregnancy: synergy of ANN and random forest. *Frontier Artificial Intellelligence*. 6(1213436), 1-11. doi: 10.3389/frai.2023.1213436.

35- UN. Population | United Nations. Un, 2019.

36- Worldometer (2020). Egypt demographics. https://www.worldometers.info.

37- [5] Walczak, S.; Cerpa, N. (2003). Artificial Neural Networks. In Encyclopedia of Physical Science and Technology, 3rd ed.; Academic Press: Cambridge, MA, USA, pp. 631–645.

38- Zaki,H.(2020) The Pattern of Women's Reliance on Family Planning Providers in Egypt Advances in Reproductive Sciences 08(01):82-95