Asian Journal of Probability and Statistics

Volume 24, Issue 2, Page 10-23, 2023; Article no.AJPAS.104745
ISSN: 2582-0230

Trend Analysis and Determinants of under-5 Mortality in Nigeria: A Machine Learning Approach

Solomon Ntukidem [a*], Angela Chukwu [a], Oluwafemi Oyamakin [a], Christabel James [b] and Ignace Habimana-Kabano [c]

[a] *Department of Statistics, University of Ibadan, Nigeria.*
[b] *Mathematics Unit, Nile University of Nigeria, Nigeria.*
[c] *The African Centre of Excellence in Data Science, University of Rwanda, Rwanda.*

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/AJPAS/2023/v24i2520

Open Peer Review History:
This journal follows the Advanced Open Peer Review policy. Identity of the Reviewers, Editor(s) and additional Reviewers, peer review comments, different versions of the manuscript, comments of the editors, etc are available here: https://www.sdiarticle5.com/review-history/104745

Original Research Article

Received: 12/06/2023
Accepted: 17/08/2023
Published: 29/08/2023

Abstract

The study aimed to examine the trend of the under-five mortality rate in Nigeria from 2003 to 2018 and the determinants of under-five mortality using the Nigeria Demographic and Health Survey (NDHS) data. The data for the study was the Nigeria Demographic and Health Survey data conducted in 2003, 2008, 2013, and 2018. These four surveys were used to study under-five mortality trends within the study period, while machine learning was applied only to the 2018 dataset being the latest in Nigeria. The data were partitioned into training and testing sets. 30% of the dataset was randomly selected for testing, while 70% was used in training the model. Before applying logistic regression and neural networks, the essential under-five mortality variables were first selected using a random forest classifier.
The trend showed that the mortality rates were 200.72, 156.86, 128.05, and 132.02 in 2003, 2008, 2013, and 2018 respectively, per 1,000 live births. This result means that one in every five children died before their

*Corresponding author: Email: ibomnteibom@gmail.com;

fifth birthday in 2003, one in six in 2008, one in eight in 2013, and one in seven in 2018. The forecast result indicated that the under-five mortality rate would likely be 102.17 in 2023. The variable importance result of the random forest showed that breastfeeding (when the child was put to the breast after birth) had the highest contribution to under-five mortality. The breakdown of breastfeeding from the logistic regression result showed that delaying the breastfeeding of a child to 6-23 hours in comparison with 0-5 hours after birth increases by 1.4 fold the likelihood of child death. The accuracy of logistic regression (LR) on the test set was 60%, and that of deep neural network (DNN) was 74%, recall (sensitivity) for LR was 63%, and DNN was 75%), Precision (LR=97%, DNN=95), F1 score (LR=76%, DNN=84%) and area under the curve (AUC) (LR=79%, DNN=77%).
Both logistic regression and deep neural network models performed very well in discriminative ability and accuracy. The deep neural network had a better performance than the logistic regression.

Keywords: Under-five mortality; random forest; logistic regression; deep neural network; trend, forecast.

1 Introduction

The under-five mortality rate (U5MR) is an indicator of the overall health condition of any nation. The under-five mortality rate in Sub-Saharan Africa has been very high compared to developed nations. Though there has been a steady decline in the under-five mortality rate in Nigeria from 135.2 in 2010 to 119.9 in 2018 [1], it is still much higher than what is obtained in developed countries like the United States of America, France, the United Kingdom, and Spain, with the U5MR less than seven in 2018. African countries like South Africa 2018 had an under-five mortality rate of 33.8; Ghana recorded 47.9; Ethiopia had 52.5; and Rwanda had 35.3. These figures show that Nigeria recorded an under-five mortality rate higher than many African countries in 2018 [1]. This high rate of under-five mortality in Nigeria is an indication of a lack of a strong medical system in the country [2].

According to WHO [1], progress has been made globally in reducing child death since 1990. Globally, the total number of under-five deaths has reduced from 12.6 million in 1990 to 5.3 million in 2018. On average, about 15,000 children under five die daily, compared with 34,000 in 1990. It is safe to report the child mortality rate from 1990 because Nigeria's first Demographic and Health Survey was conducted in 1990. In the years preceding the 1990 survey, nearly one in five children died before their fifth birthday [3]. Over this period, under-five mortality declined slowly from 201 to 192 per thousand live births in Nigeria. The WHO [1] states that the world has been accelerating progress in reducing the under-five mortality rate, but disparities still exist across regions and countries. Sub-Saharan Africa still records the highest under-five mortality rate globally, with one in 13 children dying before their fifth birthday, 15 times higher than in high-income countries [4]. WHO [1] and Biradar et al. [5] report that half of all under-five deaths in 2018 occurred in just five countries: India, Nigeria, Pakistan, Ethiopia, and the Democratic Republic of the Congo.

Empirical data shows that Nigeria recorded an under-five mortality rate of 122.1 per 1,000 live births in 2017, less than Somalia and Chad in Africa, with a mortality rate of 125.5 and 122.7, respectively. Furthermore, in 2018, Nigeria was only second to Somalia, with an under-five mortality rate of 119.9 and 121.5 for Somalia [6]. The average under-five mortality rate in Africa between the period of 2003 and 2018 is presented in Fig. 1.

From the above chart, Sierra Leone had the highest mean under-five mortality rate (159.2), followed by Somalia (152.1), and Nigeria is the fifth-highest (138).

The high under-five mortality rate has been problematic, especially in Africa and other developing countries. Furthermore, this has been attributed to many factors; sociodemographic such as age, marital status, household wealth index, etc., the child factor such as birth order, sex, birth interval, and duration of breastfeeding; as well as environmental/health-related factors, such as the source of drinking water, toilet facilities, antenatal care, etc. [7,8,9,10].

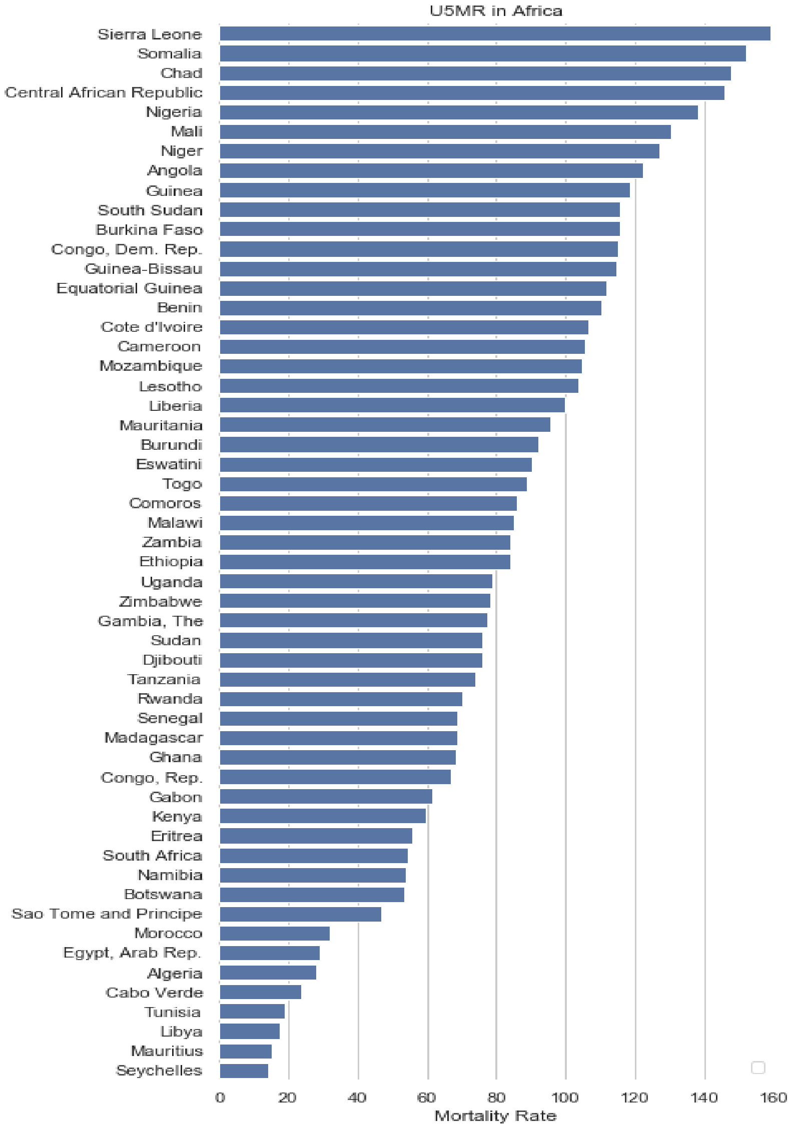


Fig. 1. Mean Under-five Mortality Rate in Africa between 2003 and 2018 (Author's computation, 2020)

Different research has been conducted to determine the factors associated with under-five mortality using different approaches, mainly traditional methods, and a few have used machine learning approaches to enhance the quality of prediction of the risk of mortality among children [11,7,12]. Sakr and Elshawi et al. [13] state that the traditional method of analyzing data investigates associations based on hypothesis testing, while machine learning clearly explains patterns in the set of input variables that identify the predicted variable. According to Sakr and Elshawi et al. [13], machine learning algorithms automatically scan and analyze all predictor variables to prevent overlooking essential predictor variables, even if it is unexpected. In this era of big data, machine learning can provide more accurate estimates of statistical analyses than traditional methods [14]. This study, therefore, examines the trend and the factors associated with under-five mortality in Nigeria using the Nigeria Demographic and Health Survey data.

2 Methodology

2.1 Data and source of data

The Nigeria Demographic and Health Survey (NDHS) was designed to provide data to monitor Nigeria's population and health situation to provide reliable information about maternal and child health and family planning services [15]. The data for the study is the Nigeria Demographic and Health Survey (NDHS) data conducted in 2003, 2008, 2013, and 2018. These four surveys were used to study the trend in under-five mortality within the period. At the same time, machine learning was applied only to the 2018 dataset being the most recent in Nigeria. The data was extracted from the child recode dataset. The extracted dataset was cleaned and recoded to suit the study. The study population comprised all children under the age of five.

2.2 Variables used in the study

2.2.1 Predictor variables

2.2.1.1 Demographic/socioeconomic factors

Mother's age, education, region, employment status, place of residence, religion, wealth index, marital status, and sex of a child.

2.2.1.2 Environmental/health factors

Place of residence, source of drinking water, type of toilet facility, type of cooking fuel, age at first birth, contraceptive use, when the child was put to the breast, birth order, preceding birth interval, months of breastfeeding, place of delivery, and delivery by cesarean section.

2.3 Method of analysis

2.3.1 Random forest

Random forest is a combination of tree predictors. Each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest [16]. Breiman [16] states that the random forest is one of the many machine-learning techniques useful for prediction and classification problems.

2.3.1.1 Random forest algorithm

The algorithm, as outlined by Xu [17], is as follows:

Suppose there exists a training set $\boldsymbol{C} = \{C_1, C_2, \dots, C_n\}$ with $C_i = (x_1, y_i)$ and an independent test case C_0 with predictor x_0.

1. The training set C is sampled with replacement to generate bootstrap resamples $\{B_1, B_2, \dots, B_m\}$.
2. For each resample $B_m, m = 1, 2, \dots, M$, a classification or regression tree T_m is grown
3. For predicting the test case C_0 with covariate x_0, the predicted value by the whole random forest is obtained by combining the results given by the individual trees. Let $\hat{f}_m^*(x_0)$ denote the prediction of C_0 by mth tress; the random forest is predicted as

$$RF = \begin{cases} \frac{1}{M}\sum_{m=1}^{M} \hat{f}_m^*(x_0) \ \ for\ regression\ problem \\ argmax_g\{\sum_{m=1}^{M} I[\hat{f}_m^*(x_0) = g]\} \ \ for\ classification\ problem \end{cases}$$

Where $I(.)$ denotes an indicator function

The random forest classification algorithm was used here since the outcome variable was categorical (either dead or alive).

2.3.2 Logistic regression

The logistic regression model is given as $log\,\frac{\pi_i}{1-\pi_i} = \log\left(\frac{\frac{\exp(\beta_0+\beta_1X_1+\dots+\beta_iX_i)}{1+\exp(\beta_0+\beta_1X_1+\dots+\beta_iX_i)}}{1-\frac{\exp(\beta_0+\beta_1X_1+\dots+\beta_iX_i)}{1+\exp(\beta_0+\beta_1X_1+\dots+\beta_iX_i)}}\right)$

$$= \log\left[\exp(\beta_0 + \beta_1X_1 + \dots + \beta_iX_i)\right] = \beta_0 + \beta_1X_1 + \dots + \beta_iX_i$$

Where x_i are the predictor variables, β_i are the unknown regression coefficients to be estimated, π_i is the probability of dying before age five, and 1-π_i is the probability of not dying.

2.3.3 Neural network

The statistical neural network proposed by Anders (1996) is given as

$$y_i = aX + \sum_{h=1}^{H} \beta_h \, g\left(\sum_{i=0}^{I} \gamma_{hi} x_i\right) + e_t$$

Where y_i is the dependent variable, $X = (x_0, x_1, x_2, \dots, x_I)$ is a vector of independent variables, $(\alpha, \beta, \gamma) = w$ are network weights: 'α' is the weight of the input unit, 'β' is the weight of the hidden unit, 'γ' is the weight of the output unit. The stochastic term e_t is normally distributed (that is, $e_t \sim N(0, \sigma^2 I_n)$).

The network architecture used for this study comprised an input layer with 16 input units, two hidden layers with (13,10) neurons, and the output layer with one neuron.

2.4 Model performance

The measures employed in evaluating the performances of the models were accuracy (on the test set), the area under the curve (AUC) of the operating characteristic (ROC) curve, recall, precision, and f1 score.

Accuracy: This is the proportion of correctly classified instances.

Area Under the Receiver Operating Characteristic Curve (AUC-ROC): This is a commonly used metric for evaluating the performance of classification models. A higher AUC-ROC indicates that the model is good at distinguishing between positive and negative classes.

Recall: This is the proportion of true positives among all actual positive instances. A higher recall indicates that the model is good at identifying all positive instances.

Precision: This is the proportion of true positives (correctly classified instances) among all positive predictions. A higher precision indicates that the model is good at identifying true positives.

F1 Score: This is the harmonic mean of precision and recall. A higher F1 score indicates that the model has a good balance between precision and recall.

The measures were computed from a confusion matrix using the formula below;

$$Accuracy = \frac{True}{True + False} = \frac{TP + TN}{TP + TN + FP + FN}$$

Where TP is a true positive, TN is a true negative, FP is a false positive, and FN is a false negative.

$$Recall = \frac{TP}{TP + FN}$$

Recall (sensitivity) provides information about the performance of a classifier for false negatives.

$$Precision = \frac{TP}{TP + FP}$$

Precision provides information about the performance of a classifier for false positives.

$$F1\ score = 2 * \frac{recall * precision}{recall + precison}$$

3 Results

3.1 Descriptive analysis

Fig. 2 shows that 22.9% and 22.5% of mothers were between 25-29 and 30-34. Most of the women (65.5%) were in rural areas, and most of them (30.4%) were from the Northwest region of the country, followed by the Northeast (21.3%). Most of the mothers (45.4%) had no formal education, while only 7.8% of them had higher education, and 60.2% were Muslims. Most women (23.8%) fell in the wealth index classified as poorest, followed by those classified as poorer (22.85), while 4.1% of the women were classified as richest. Most women (66.9%) had their first child at 17-26 years, while 25.3% had their first child at less than 17 years of age.

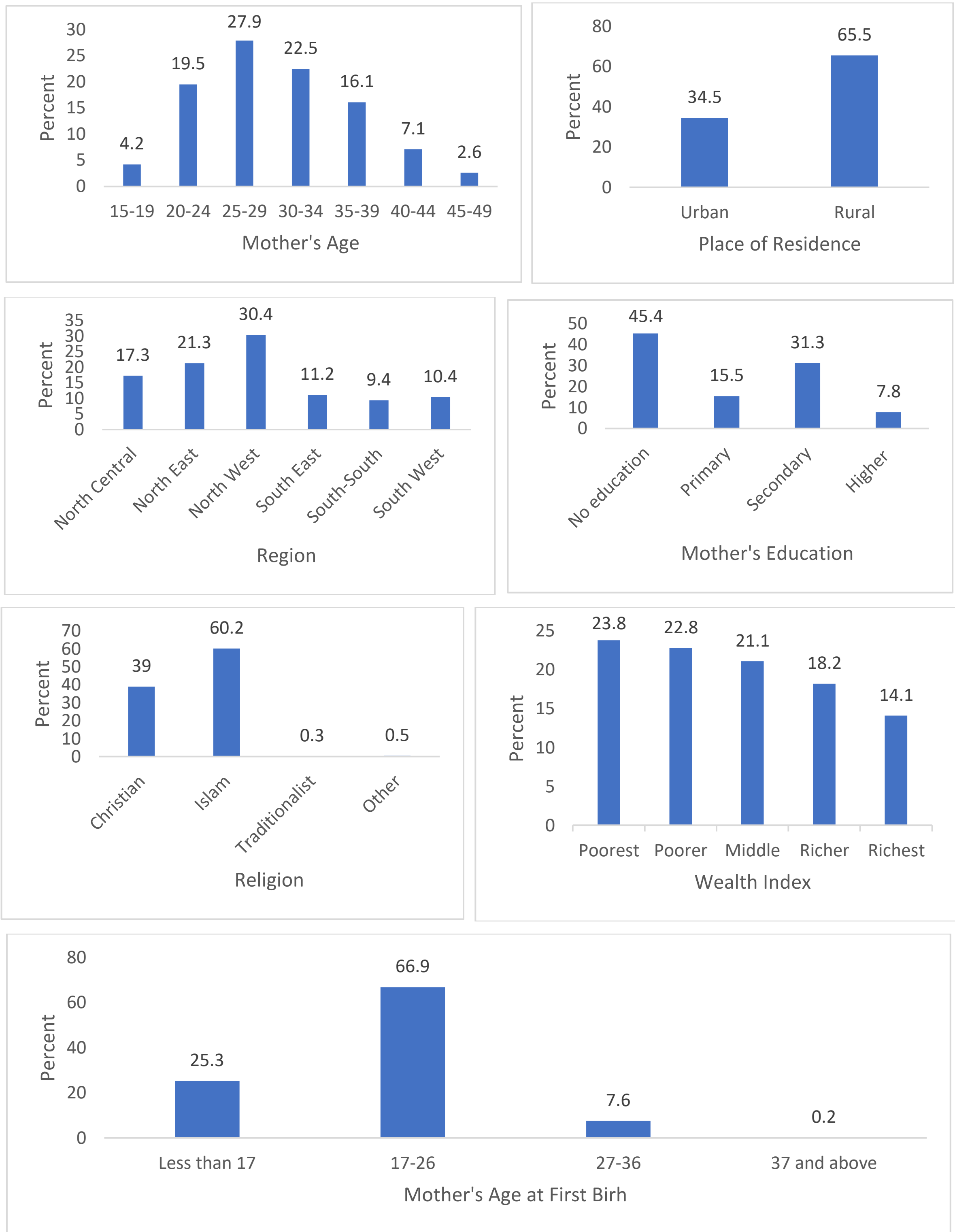


Fig. 2. The descriptive analysis of some selected variables of mother

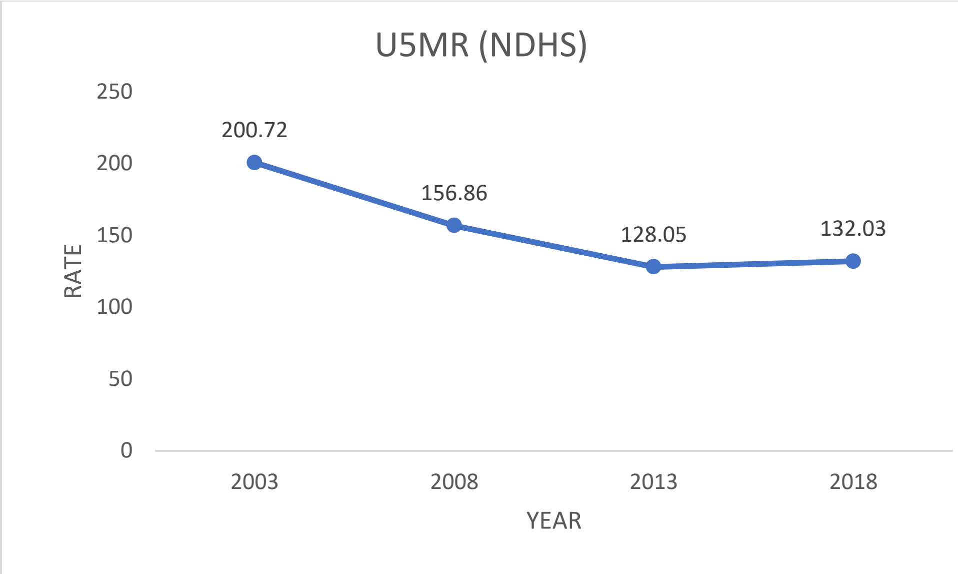
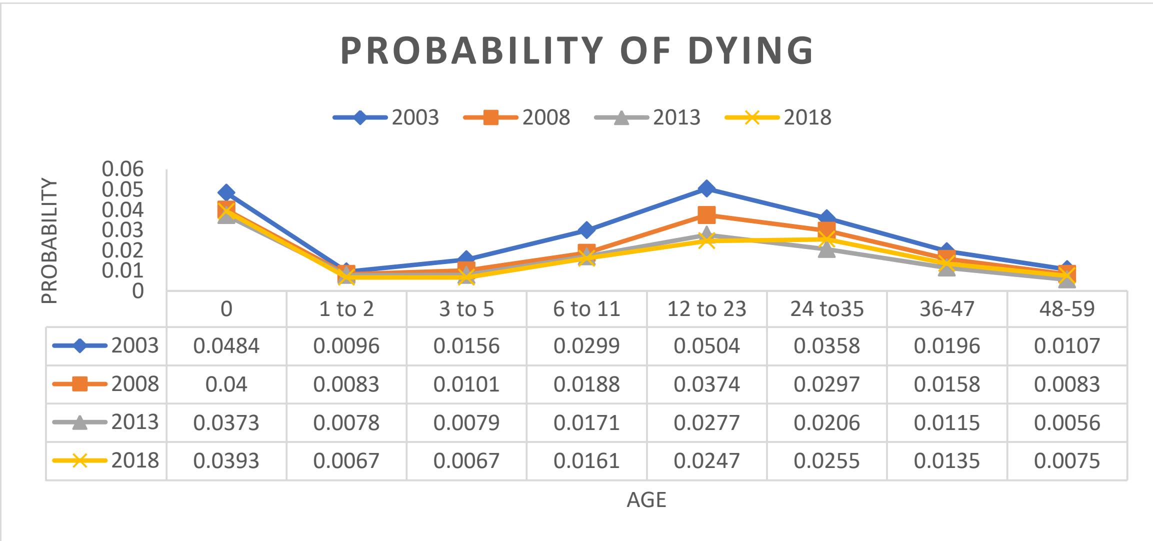


Fig. 3. Trend in under-five mortality rate from 2003 to 2018 in Nigeria

The result shows that the under-five mortality rates were 200.72, 156.86, 128.05, and 132.02 in 2003, 2008, 2013, and 2018 respectively, per 1,000 live births. This result shows that one in every five children died before their fifth birthday in 2003, one in six in 2008, one in eight in 2013, and one in seven in 2018. This gives the rate of change in under-five mortality of 21.9% reduction in 2008, 18.4% reduction in 2013, and 3.1% increment in 2018. This result shows that the under-five mortality rate declined from 2003 to 2013 but increased in 2018. The mortality rate for 2018 was higher than in 2013 but less than in 2008 and 2003.



	0	1 to 2	3 to 5	6 to 11	12 to 23	24 to35	36-47	48-59
2003	0.0484	0.0096	0.0156	0.0299	0.0504	0.0358	0.0196	0.0107
2008	0.04	0.0083	0.0101	0.0188	0.0374	0.0297	0.0158	0.0083
2013	0.0373	0.0078	0.0079	0.0171	0.0277	0.0206	0.0115	0.0056
2018	0.0393	0.0067	0.0067	0.0161	0.0247	0.0255	0.0135	0.0075

Fig. 4. The probability of dying before the age five

Fig. 3 shows that the probability of dying before one month is higher than other age (in months) groups, followed by 12 to 23 months. Except for 2008, the probability of dying between 24 to 35 months is higher than 12 to 23.

Fig. 5 shows the forecast rate of 102.17 Under-five death per 1,000 live births for 2023. This indicates a possible reduction in under-five mortality in Nigeria, holding other factors constant.

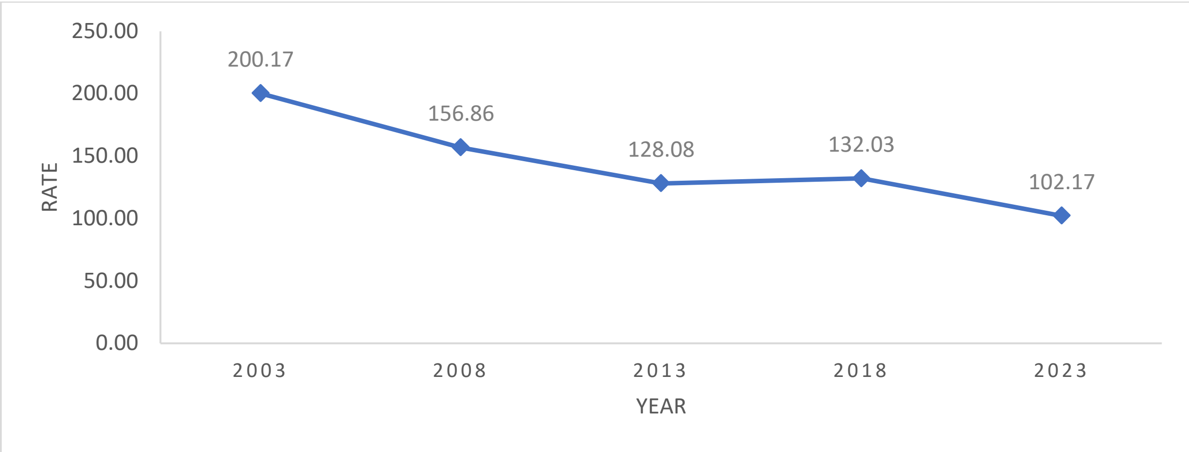


Fig. 5. Under-five mortality forecast for 2023

3.2 Random forest

Random forest was used to select variables based on their contributions to a child being alive or dead, as presented in Fig. 5. The result shows the variables contributing to under mortality in their order of importance. When a child was put to breast had the highest contribution to under-five mortality, followed by the mother's age, toilet facility, source of drinking water, and preceding birth interval. Predictors with the lowest contribution were mothers' marital status, frequency of reading newspapers (media exposure), delivery by caesarian section, and smoking of cigarettes.

3.2.1 Performance of random forest

Performance of random forest shown in Fig. 7.

3.3 Logistic regression

The result in Table 1 shows the effect of the explanatory variables on the outcome variable (under-five mortality). The outcome variable was coded 0 and 1 (0 for the alive, 1 for the dead). The level with the highest frequency was used as the reference category (RC) for the explanatory variables. The mothers' age bracket of 25-29 was used as the reference category. Children born to women who were 15-19 years were 1.23 times more likely to die before the age of 5 than those in the reference category (*OR=1.23, P=0.0278*). Children born to women between 45 and 49 were 1.9 times more likely to die before age five compared to mothers between 25-29 (*OR=1.90, P<0.001*). For the region (Zone), children born in the North East were about 1.4 times more likely to die before reaching five years compared to those in the North West (*OR=1.36, P<0.001*). Children in the rest of the regions had less risk of dying than in the North West. The risk of children born in Urban areas dying before age five was significantly less than those born in rural areas (*OR=0.66, P<0.001*). The risk of children born to mothers who had primary (*OR=0.76*), secondary (*OR=0.52*), or higher (*OR=0.40*) education dying before the age of five was significantly less than those born to mothers who had no formal education *(P<0.001).* Children born to mothers in the wealth index classified as middle class, richer, and richest had a significantly lower risk of dying before the age of five than those born to the women classified as poorest (*OR = 0.73, 0.52, 0.38, P<0.001*). The risk of female children dying before five years was significantly less than that of male children (*OR=0.89, P=0.0013*). Children born to mothers who were Catholics (*OR=0.49*) and other Christians (*OR = 0.60*) had a significantly lower risk of dying before the age of five compared to children born to Muslim mothers (RC).

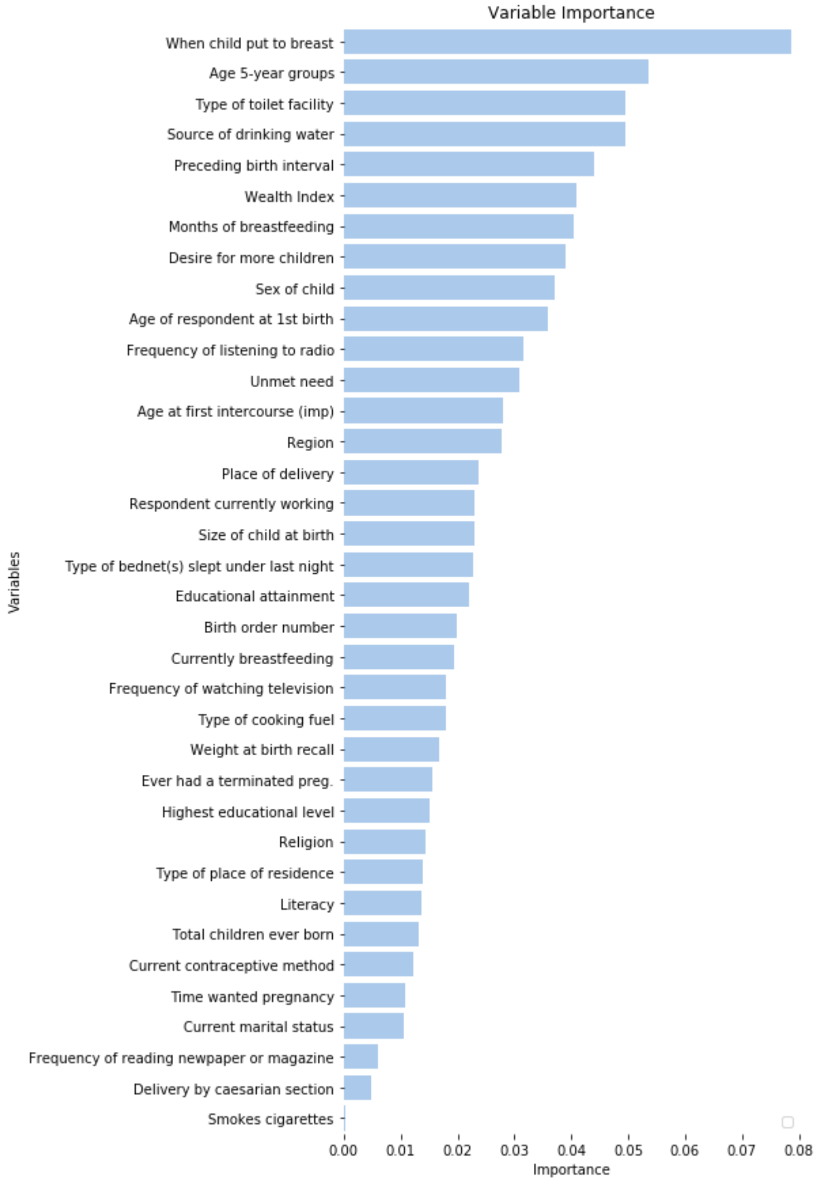


Fig. 6. Feature selection for under-five mortality

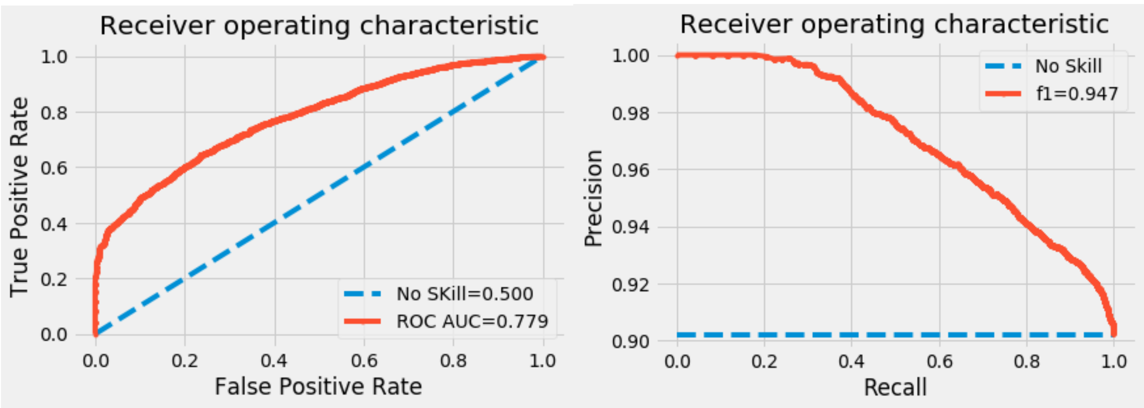


Fig. 7. Graphical representation of true positive rate vs false positive rate and precision vs recall of Random Forest performance

Children who drank well water were 1.29 times more likely to die before the age of five compared to those who drank borehole water (RC) (*OR=1.29, P<0.001*), and those who drank stream/river water were about 1.15 times at risk of dying before the age of five compared to the children who drank borehole water (*OR=1.15, P=0.019*). Children who were put to the breasts 6-23 hours after birth were 1.38 times more likely to die before the age of five compared to those put to breast between 0-5 hours (*OR=1.38, P<0.001*), and children who were put to breast between 1-5 days had a higher risk of dying before the age of five years than those put to breast between 0-5 hours after birth (*OR=1.29, P<0.001*).

The children who were between positions 6-10 in the family were about 1.5 times more likely to die before the age of five relative to those whose birth order number was between 1 and 5 (*OR=1.50, P<0.001*) and those whose birth order number was 11 and above were 2.22 times more likely to die before the age of five compared to the birth order numbers 1-5 (*OR=2.22, P<0.001*). The risk of children whose birth interval was less than 24 months dying before age five was 1.8 times more than those whose birth interval was between 24-59 months (*OR=1.82, P<0.001*). Mothers who were less than 17 years at first birth were 1.4 times more likely to have children dying before the age of five compared to those whose age at first birth was 17-26 years (*OR=1.42, P<0.001*). Children born in a government hospital (*OR=0.68, P<0.001*) or private hospital (*OR=0.59, P<0.001*) were significantly less likely to die before the age of five compared to those born at home.

Table 1. Logistic regression analysis of factors affecting under-five mortality

Explanatory variables	**OR**	**P-value**	**95% Confidence Interval**	
			Lower	**Upper**
Age (age 25-29 is the RC)				
15-19	1.230	0.0278	1.0229	1.4801
20-24	1.195	0.0012	1.0727	1.3306
30-34	1.050	0.3672	0.9441	1.1683
35-39	1.1572	0.0269	1.0317	1.2980
40-44	1.1622	0.0532	0.9979	1.3536
45-49	1.9041	<0.001	1.5641	2.3182
Region (NW is the RC)				
North Central	0.7756	<0.001	0.6870	0.8756
North East	1.3624	<0.001	1.2385	1.4987
South East	0.6336	<0.001	0.5460	0.7352
South-South	0.5319	<0.001	0.4495	0.6295
South West	0.5441	<0.001	0.4633	0.6389
Place of Residence				
Urban	0.6569	<0.001	0.6052	0.7129
Rural (RC)				
Education				
No education (RC)				
Primary	0.7567	<0.001	0.6818	0.8398
Secondary	0.5198	<0.001	0.4749	0.5689
Higher	0.4040	<0.001	0.3383	0.4835
Wealth Index				
Poorest (RC)				
Poorer	0.9502	0.2984	0.8631	1.0462
Middle	0.7311	<0.001	0.6589	0.8113
Richer	0.5241	<0.001	0.4651	0.5906
Richest	0.3819	<0.001	0.3301	0.4418
Sex of Child				
Male (RC)				
Female	0.8875	0.0013	0.8251	0.9546
Source of drinking water				
Borehole (RC)				
Piped water	0.9706	0.6602	0.8497	1.1087
Well	1.2857	<0.001	1.1760	1.4056

Explanatory variables	**OR**	**P-value**	**95% Confidence Interval**	
			Lower	**Upper**
Stream/river	1.1450	0.0190	1.0225	1.2821
Other	0.7167	<0.001	0.6120	0.8259
Religion				
Islam (RC)				
Catholic	0.4852	<0.001	0.4129	0.5702
Other Christians	0.6011	<0.001	0.5512	0.6556
Traditionalist	0.5836	0.1424	0.2842	1.1983
Other	0.1535	0.0013	0.0489	0.4815
When a child put to the breast				
0-5hrs (RC)				
6-23hrs	1.3771	<0.001	1.1916	1.5916
1-5days	1.2899	<0.001	1.1594	1.435
6-23days	0.8022	0.5736	0.3724	1.7281
I cannot tell exactly	9.3533	<0.001	8.1875	10.6851
Birth Order				
1-5(RC)				
6-10	1.4965	<0.001	1.3793	1.6237
11 and above	2.2185	<0.001	1.7752	2.7724
Age at first birth				
17-26 (RC)				
Less than 17	1.4198	<0.001	1.311	1.5375
27-36	0.7892	0.0033	0.6739	0.9242
37 and above	0.5691	0.2728	0.2078	1.5586
Birth Interval				
24-59months (RC)				
Less than 24months	1.8230	<0.001	1.6715	1.9882
60months and above	0.7634	0.0018	0.6443	0.9045
Don’t know	1.1380	0.0010	1.0310	1.2562
Place of delivery				
Homes (RC)				
Govt hospital/clinic	0.6789	<0.001	0.6217	0.7414
Private hospital/clinic	0.5879	<0.001	0.5164	0.6695
other	0.6909	0.0301	0.4947	0.9650

3.4 Performance of Logistic Regression and Deep Neural Network

Logistic Regression and Deep Neural Network shown in Fig. 8.

Table 2. The performance indices of logistic regression and deep neural network

Performance indices	Logistic Regression	Deep Neural Network
Accuracy	60%	74%
AUC	79%	77%
F1 score	76%	84%
Recall	63%	75%
Precision	97%	95%

4 Discussion

A progressive nation and society invest heavily in children's health since the future depends so much on them. Past and present study findings have revealed the state of children’s health and the mortality rate. The trend in under-five mortality is clearly shown in Fig. 3. While there was a steady decline in the mortality rate, measured per 1,000 live birth [1], from 2003 to 2013 (200.17 to 128.08), the mortality rate recorded in the year 2018, which is the most recent NDHS data in Nigeria (132.03) was higher than that of the previous survey (2013, 128.08). The result indicates that proper attention has not been given to children's health, and the factors responsible for under-five mortality pointed out in previous studies [12,7]) were not taken cognizance of by parents and authorities. These factors, also evident in the present study, are breastfeeding, birth interval, mother's age, etc. To improve the quality of life of children and secure a better future for them and society, health professionals need to carry out awareness campaigns to educate mothers on how to safeguard their children's lives.

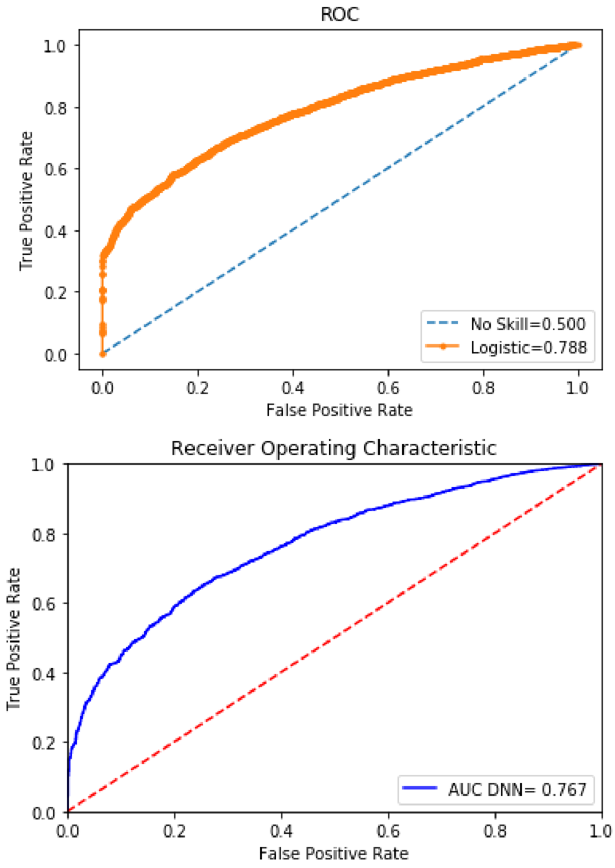


Fig. 8. Graphical representation of true positive rate vs false positive rate of logistic regression and deep neural network performance

The present study goes beyond examining past and present events surrounding under-five mortality in Nigeria to forecast for 2023. The finding revealed a possible reduction in the under-five mortality rate by 22.6%, holding factors like war, outbreak of diseases, and others constant. The result means the mortality rate would reduce from 132.03 in 2018 to 102.17 in 2023.

The variable importance analysis conducted using random forest shows that the time the child was first breastfed after birth was the most crucial contributor to under-five mortality [7,2]. This result shows that breastfeeding is crucial to the child's health, as also pointed out by Azuine et al. [10]. The second most important variable of under-five mortality was the mother's age, followed by toilet facility, source of drinking water, preceding birth interval, and wealth index. Parents' socioeconomic status, demographics, and care for the children play a predominant role in determining the health status of the children.

The Logistic regression result clearly shows that children born to women in the age bracket of 45 to 49 had a higher tendency to die before the age of five compared to those born to women between 25-29 years [18]. Children born to women classified as the poorest had a higher risk of dying before the age of five than those in the other categories [19]. Children who were breastfed between 6-23 hours and 1-5 days after birth had a higher risk of dying before the age of five compared to those put to breast between 0 to 5 hours after birth [20-22].

The model's accuracy on the test set for logistic regression was 60%, while that of the deep neural network was 74%. This finding shows that the accuracy of logistic regression was less than that of the neural network model. The area under the curve (AUC) in the Receiver Operating Characteristic curve for logistic regression was 79%, while that of the neural network was 77%. For logistic regression and the neural network models, the F1 scores were 64% and 84%, respectively. Recall for logistic regression was 63%, while that of the neural network was 75%. Logistic regression had a precision of 97%, while the neural network had a precision of 95%. The result shows that the neural network model slightly outperforms the logistic regression model.

5 Conclusion

The first five most important predictors of the under-five mortality rate in Nigeria are breastfeeding (when a child is put to the breast), mother's age, toilet facility, source of drinking water, and preceding birth interval. Both logistic regression and deep neural network models performed well in discriminative ability and accuracy, but deep neural networks performed better than logistic regression. Despite the increase in the under-five mortality rate recorded in the 2018 survey, the forecast made, however, suggests a decline in the mortality rate in 2023, where the next survey is expected to be conducted.

Competing Interests

Authors have declared that no competing interests exist.

References

[1] Adedini SA, Odimegwu C, Imasiku EN, Ononokpono DN. Ethnic differentials in under-five mortality in Nigeria, Ethnicity & Health. 2015;20(2):145-162.

[2] Adegbosin AE, Stantic B, Sun J. Predicting Under-five mortality across 21 Low and Middle-Income Countries using Deep Learning Methods; 2019.
Available:http://dx.doi.org/10.1101/19007583

[3] Adetoro GW, Amoo EO. A statistical analysis of child mortality: Evidence from Nigeria. Journal of Demography and Social Statistics. 2014;(March 2014):110-120.

[4] Azuine RE, Murray J, Alsafi N, Singh GK. Exclusive breastfeeding and under-five mortality, 2006-2014: A cross-national analysis of 57 low- and-middle income countries. International Journal of MCH and AIDS. 2015;4(1):13-21.

[5] Biradar R, Patel KK, Prasad JB. Effect of birth interval and wealth on under-5 child mortality in Nigeria. Clinical Epidemiology and Global Health. 201;97(2019):223-238.

[6] Breiman L. Random Forest, Machine Learning. 2001;45:5–32.

[7] Ezeh OK, Agho KE, Dibley MJ, Hall JJ, Page AN. Risk factors for postneonatal, infant, child and under-5 mortality in Nigeria: A pooled cross-sectional analysis. BMJ Open. 2015;2015(5):1-9.

[8] Federal Office of Statistics/Nigeria and Institute for Resource Development - IRD/Macro International; 1992.

[9] Jaiyeola MO, Oyamakin SO, Akinyemi JO. Adebowale SA, Chukwu AU, Yusuf OB. Assessing infant mortality in Nigeria using artificial neural network and logistic regression models. British Journal of Mathematics & Computer Science. 2016;19(5):1-14

[10] Morakinyo OM, Fagbamigbe AF. Neonatal, infant and under-five mortalities in Nigeria: An examination of trends and drivers (2003-2013). PLoS ONE. 2017;12(8):1-21.

[11] Mugwaneza P, Shema NU, Ruton H, et al. Under-two child mortality according to maternal HIV status in Rwanda: assessing outcomes within the National PMTCT Program Pan African Medical Journal. 2011;9(37):1-13.

[12] Sakr S, Elshawi R, Ahmed A, Qureshi W, Brawner C, Keteyian S, Blaha M, Al-Mallah M. Comparison of machine learning techniques to predict all-cause mortality using fitness data: the Henry Ford exercise testing (FIT) project. BMC Medical Informatics and Decision Making. 2017;2017(17):174.

[13] The World Bank. Nigeria - Demographic and Health Survey 2013; 2013.
Available: https://datacatalog.worldbank.org/dataset/nigeria-demographic-and-health-survey-2013

[14] UNICEF. Under-five Mortality; 2019.
Available:https://data.unicef.org/topic/child-survival/under-five-mortality/

[15] Wang P, Li Y, Reddy CK. Machine learning for survival analysis: A survey. ACM Computing Surveys. 2017;1(1):1–39.

[16] WHO. Children: reducing mortality; 2019.
Available:https://www.who.int/news-room/fact-sheets/detail/children-reducing-mortality.

[17] World Bank. The mortality rate is under-5 (per 1,000 live births); 2019.
Available:https://data.worldbank.org/indicator/SH.DYN.MORT?end=2018&start=1960&view=chart

[18] Xu R. Improvements to random forest methodology, Graduate Theses and Dissertations. 13052; 2013.
Available:https://lib.dr.iastate.edu/etd/13052

[19] Yaya S, Ekholuenetale M, Tudeme G, Vaibhav S, Bishwajit G, Kadio B. Prevalence and determinants of childhood mortality in Nigeria, BMC Public Health. 2017;17:485

[20] Nigeria Demographic and Health Survey 1990. Columbia, Maryland, USA: Federal Office of Statistics/Nigeria and IRD/Macro International.
Available:Retrieved from https://dhsprogram.com/pubs/pdf/FR27/FR27.pdf

[21] Nanga S, Lotsi A. A comparison between logistic regression and neural networks in modeling mortality amongst children under five years in Ghana. IOSR Journal of Mathematics (IOSR-JM). 2016;12(5):26-31.

[22] Yaya S, Ekholuenetale M, Tudeme G, Vaibhav S, Bishwajit G, Kadio B. Prevalence and determinants of childhood mortality in Nigeria, BMC Public Health. 2017;17:485.

Peer-review history:
The peer review history for this paper can be accessed here (Please copy paste the total link in your browser address bar)
https://www.sdiarticle5.com/review-history/104745