<u>Original Research Article</u>

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

Abstract

Aim: 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.

Data and Methods: 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 childhood mortality trends within the 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 most essential under-five mortality variables were first selected using a random forest classifier.

Results: 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. 1in every five children died before their 5th birthday in 2003, 1 in 6 in 2008, 1 in 8 in 2013, and 1 in 7 in 2018. The forecast result showed that the U5MR will 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. Logistic regression results showed that children who were put to breast 6-23 hours and 1-5 days after birth had a higher risk of dying before the age of five relative to those put to breast between 0-5 hours after birth. The accuracy of logistic regression on the test set was 60% and that of deep neural network was 74%, recall (sensitivity) for LR was 63% and DNN was 75%), Precision (LR=97%, DNN=95), F1 score (LR=76%, DNN=84%) and AUC (LR=79%, DNN=77%).

Conclusion: 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 logistic regression

Keywords: Under-five Mortality, Random Forest, Logistic Regression, Deep Neural Network, Trend, Forecast.

UNDER PEER REVIEW

1 Introduction

The under-five mortality rate 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 the developed nations. There has been a steady decline in under-five mortality (U5MR) in Nigeria from 135.2 in 2010 to 119.9 in 2018 (WHO, 2019). The U5MR in Nigeria 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 in 2018 had an under-five mortality rate of 33.8, Ghana recorded 47.9, while Ethiopia had 52.5 and Rwanda had 35.3. These figures show that Nigeria recorded under-five mortality higher than many African countries in 2018 (WHO, 2019). This high rate of under-five mortality in Nigeria is a clear indication that the medical system in the country is weak (Adegbosin, Stantic, and Sun, 2019).

According to WHO (2019), progress has been made globally in reducing child death since 1990. Globally, the total number of under-5 deaths has reduced from 12.6 million in 1990 to 5.3 million in 2018. About 15 000 children under-5 on average die every day compared with 34 000 in 1990. It is safe to report the child mortality rate from 1990 because the first Demographic and Health Survey in Nigeria was conducted in 1990. In the five years preceding the 1990 survey, nearly one in five children died before their fifth birthday (FOSNIRD, 1992). Over this period, under-five mortality declined slowly from 201 deaths to 192 deaths per thousand live births in Nigeria alone. The WHO (2019) stated that the whole world has been accelerating progress in reducing the under-5 mortality rate, but disparities still exist in under-5 mortality across regions and countries. Sub-Saharan Africa still records the highest under-5 mortality rate globally, with one child in 13 dying before their fifth birthday, 15 times higher than in high-income countries (UNICEF, 2019). WHO, (2019); Biradar, Patel & Prasad, (2019) 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-5 mortality rate of 122.1 per 1,000 live births in 2017, only less than Somalia and Chad in Africa, with a mortality rate of

UNDER PEER REVIEW

125.5 and 122.7, respectively. And in 2018, Nigeria was only second to Somalia with an under-5 mortality rate of 119.9 and 121.5 for Somalia (The World Bank, 2019). The average under-5 mortality rate in Africa between the period of 2003 and 2018 is presented in Figure 1 below:

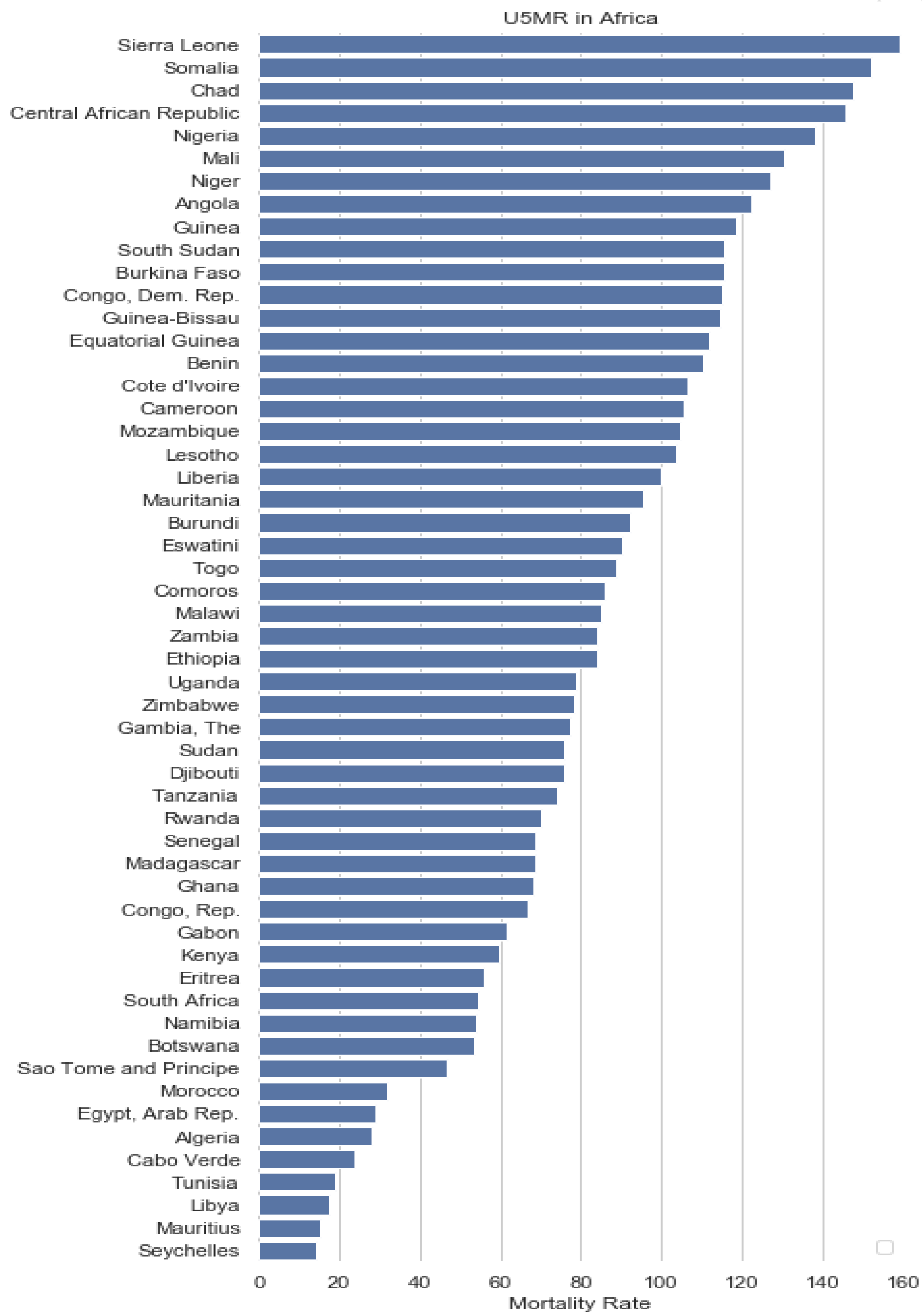


Figure 1: Mean Under-5 Mortality Rate in Africa between 2003 and 2018 (Author's computation, 2020)

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. And this has been attributed to a lot of factors; sociodemographic such as age, marital status, household wealth index, etc., the child factor such as birth order, sex, birth interval, duration of breastfeeding; as well as environmental/health-related factors such as the source of drinking water, toilet facilities, antenatal care, etc. (Jaiyeola *et al.,* 2016; Morakinyo & Fagbamigbe, 2017; Ezeh, Agho, & Dibley, 2015; Azuine, Murray, Alsafi & Singh, 2015).

Different researches have 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 (Mugwaneza *et al.,* 2011; Jaiyeola *et al.,* 2016; Adedini, Odimegwu, *et al.,* 2015). Sakr and Elshawi *et al*. (2017) stated 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*. (2017), machine learning algorithms automatically scan and analyze all predictor variables in a way that prevents overlooking important predictor variables even if it was unexpected. In this era of big data, machine learning can provide more accurate estimates of statistical analyses than the traditional methods (Wang, Li, and Reddy, 2017). This study, therefore, investigates the factors associated with under-five mortality in Nigeria using the 2018 Nigeria Demographic and Health Survey dataset.

2 Data and Methods

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 with an explicit goal of providing

UNDER PEER REVIEW

reliable information about maternal and child health and family planning services (The World Bank, 2013). 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 childhood mortality within the period, while machine learning was applied only to the 2018 dataset being 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 the children under five.

2.2 Variables Used in the Study

2.2.1 Predictor Variables

Demographic/Socioeconomic Factors

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

Environmental/Health Factors

Place of residence, source of drinking water, type of toilet facility, type of cooking fuel, age at 1st 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 (Breiman, 2001). Breiman (2001) stated that the random forest is one of the many machine-learning techniques useful for prediction and classification problems.

UNDER PEER REVIEW

Random Forest Algorithm

The algorithm, as outlined by Xu (2013), 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 prediction of C_0 by mth tress, the random forest is predicted as

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

Where $I(.)$ denotes an indicator function

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

2.3.2 Logistic Regression

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

= log [exp $(\beta_0 + \beta_1X_1 + \cdots + \beta_iX_i)$] = $\beta_0 + \beta_1X_1 + \cdots + \beta_iX_i$

Where x_i are the predictor variable, 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

UNDER PEER PREVIEW

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

Where y_i is the dependent variable, $X = (x_0, x_1, x_2, \ldots, x_I)$ 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, and e_t is the stochastic term that 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 operating characteristic (ROC) curve, recall, precision, and f1 score

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 false positive, and FN is false negative

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

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

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

Precision provides information about the performance of a classifier for false positive

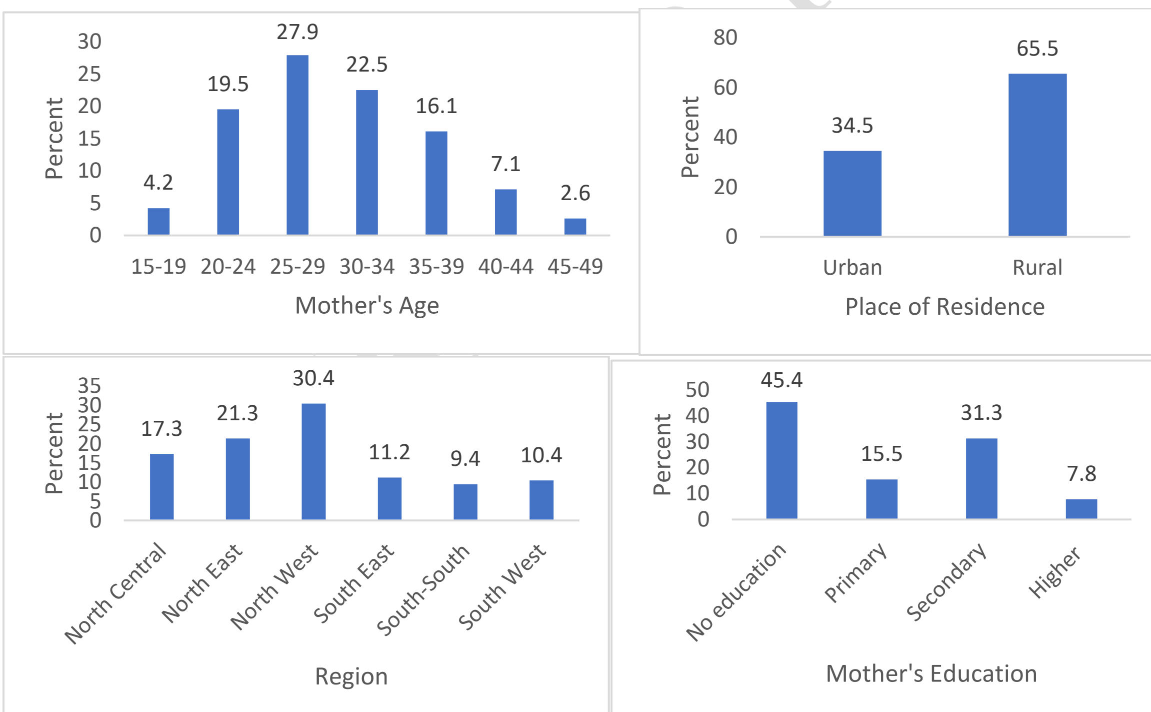
UNDER PEER REVIEW

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

3 Results

3.1 Descriptive Analysis

Figure 2 showed that 22.9% and 22.5% of mothers were between the ages of 25-29 and 30-34. Most of the women (65.5%) were in the rural area, and most women (30.4%) were from the Northwest region of the country, followed by the Northeast (21.3%). A greater percentage of the mothers (45.4%) had no education, while only 7.8% had higher education, and 60.2% were Muslims. Most women (23.8%) fell under the poorest wealth index, followed by poorer (22.85), while those in the richest wealth index were 14.1%. Most of the women (66.9%) had their first child at the age of 17-26 years, while 25.3% had their first child at less than 17 years of age.



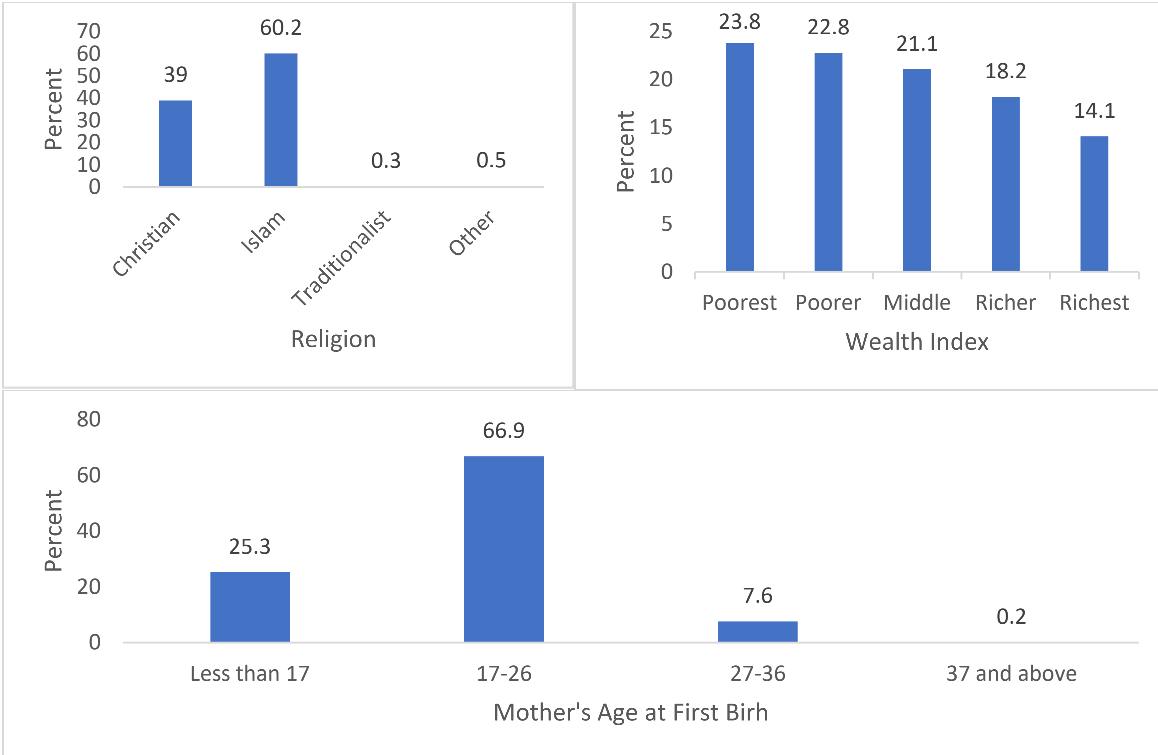


Figure 2: The Descriptive Analysis Some Selected Variables of Mother

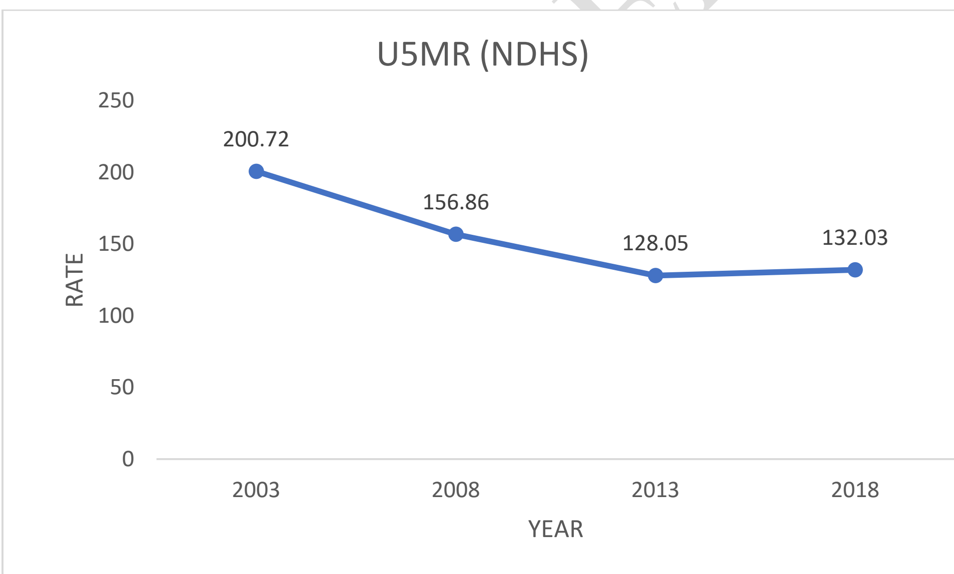
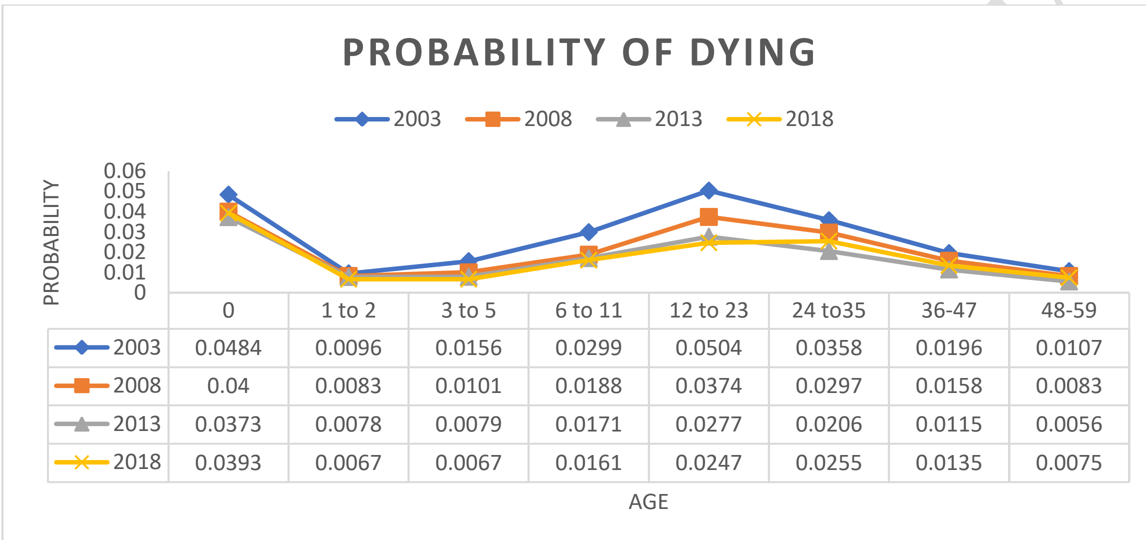


Figure 3: Trend in Under-five Mortality Rate from 2003 to 2018 in Nigeria

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 5th birthday in 2003, 1 in 6 in

2008, 1 in 8 in 2013, and 1 in 7 in 2018. This gives the rate of change in under-five mortality (21.9% reduction in 2008, 18.4% reduction in 2013, and 3.1% increment in 2018). This result showed that the under-five mortality rate declined from 2003 to 2013 and 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

Figure 4: The Probability of Dying Before the Age Five

Figure 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 the ages of 24 to 35 months is higher than 12 to 23.

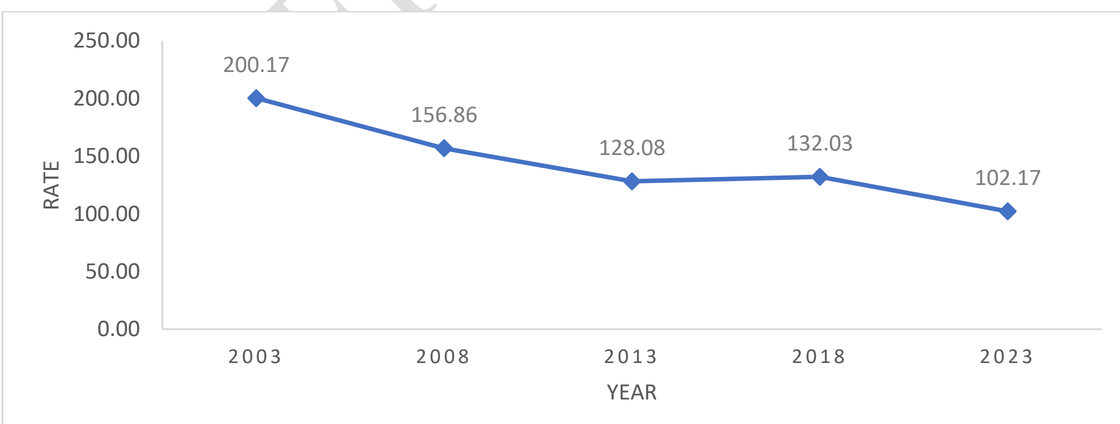


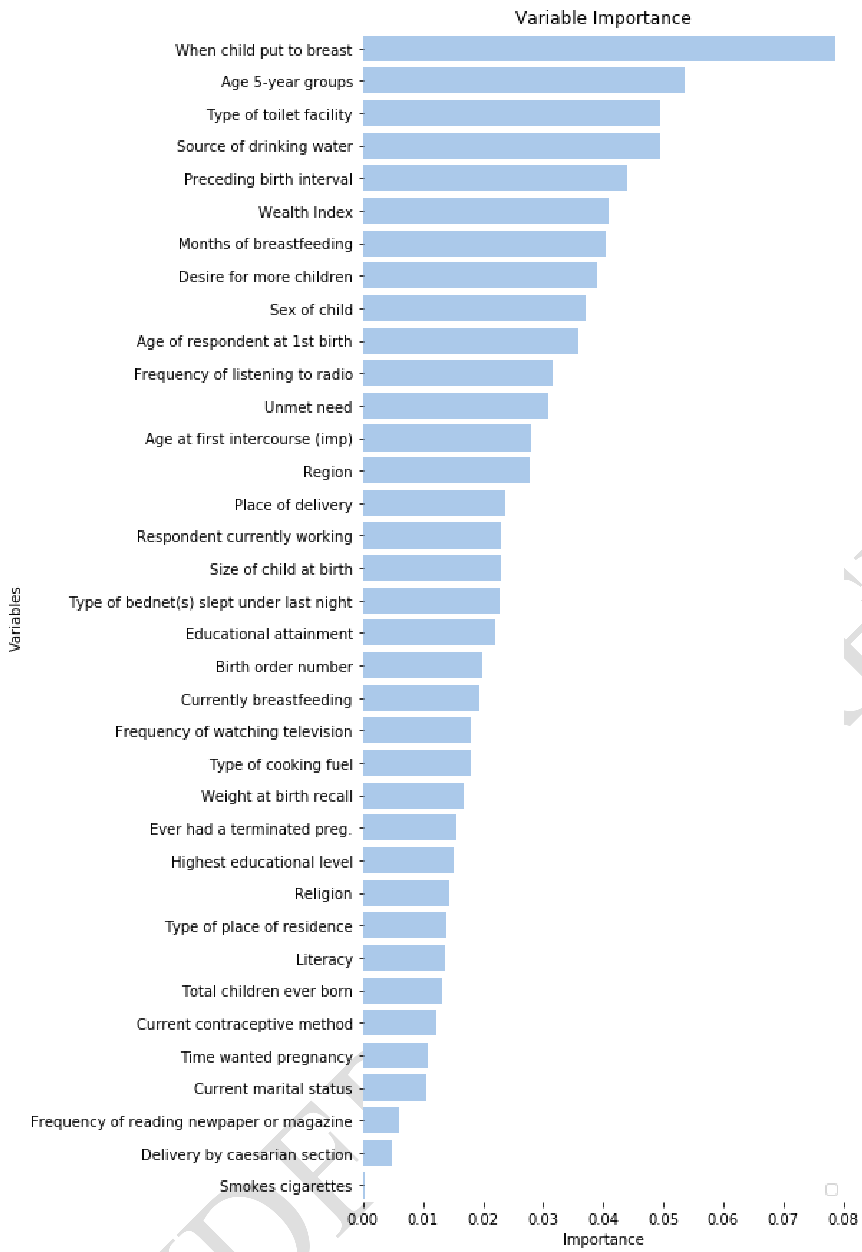
Figure 5: Under-five Mortality Forecast for 2023

In figure 5, the result showed the forecast rate of 102.17 Under-five death per 1,000 live birth for 2023. This indicates a possible reduction of under-five mortality in Nigeria, holding other factors constant.

3.3 Random Forest

Random forest was used to select variables based on their contributions to a child being alive or dead, as presented in Figure 5. The result showed the variables contributing to under mortality in their order of importance. When child put to breast had the highest contribution to under-five mortality followed by 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.

UNDER PEER REVIEW



UNDER PEER REVIEW

Figure 6: Feature Selection for Under-five

3.3.1 Performance of Random Forest

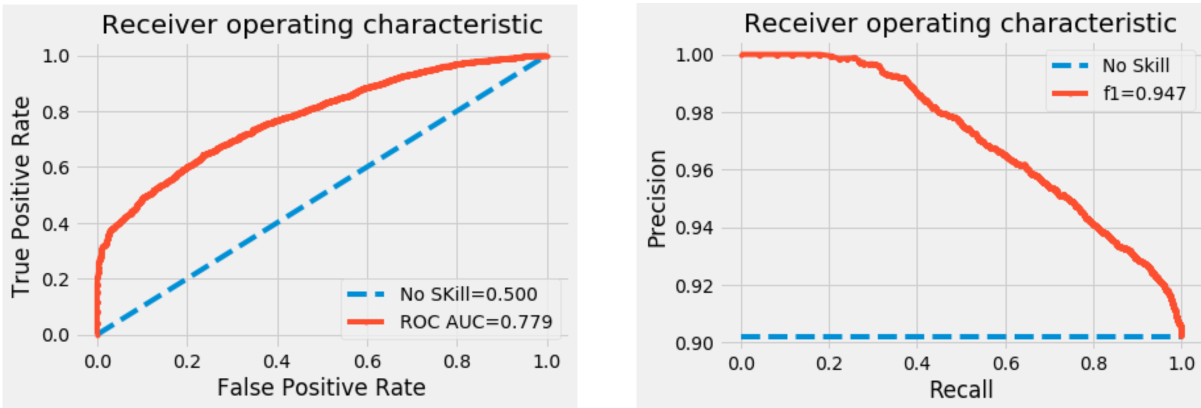


Figure 7 : Graphical representation of true positive rate vs false positive rate and precision vs recall of Random Forest performance.

3.4 Logistic Regression

The result in Table 1 showed the effect of the explanatory variables on the outcome variable (under-five mortality rate). The outcome variable was coded 0 and 1 (0 for alive, 1 for the dead). For the levels in the explanatory variables, the level with the highest frequency was used as the reference category (RC). For the age of mothers, the age bracket of 25-29 was used as the reference category. Children born to women in the age bracket of 15-19 were 1.23 times more likely to die before the age of 5 years compared to 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 reaching five than those born 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 the children born in the Urban areas dying before the age of 5 was significantly less than those born in the rural area (*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 age five was significantly less than those born to mothers who had no education *(P<0.001).* Children born to mothers who were in the wealth index of middle class, richer and richest had a significantly lower risk of dying before the age of five compared 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

UNDER PEER REVIEW

(*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 of 5 compared to children born to Muslim mothers (RC).

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

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

UNDER PEER REVIEW

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
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 child put to breast				

UNDER PEER REVIEW

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
Can't 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

3.5 Performance of Logistic Regression and Deep Neural Network

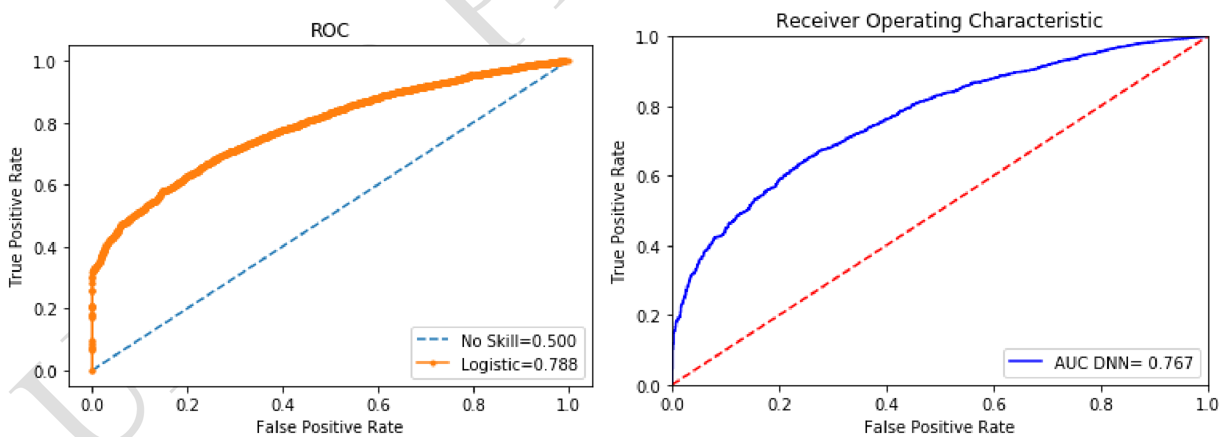


Figure 8 : Graphical representation of true positive rate vs false positive rate of Logistic Regression and Deep Neural Network Performance.

Table 2 : Logistic Regression and Deep Neural Network Analysis of various Performance indices

Performance indices	Logistic Regression	Deep Neural Network

UNDER PEER REVIEW

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 was clearly shown in figure 3. While there was a steady decrease in mortality rate, measured per 1,000 live birth (WHO, 2019) from 2003 to 2013 (200.17 to 128.08), the mortality rate recorded by the most recent survey (2018, 132.03) was higher than that of the previous surveys (2013, 128.08). This indicates that proper attention is not given to children’s health, and the factors responsible for under-five mortality pointed out in previous studies (Adedini et al., 2015; Jaiyeola et al., 2016) were not taken cognisance of by parents and authorities. These factors also evident in the present study are breastfeeding, birth interval, mother’s ages, etc. To improve the quality of life of the children and secure a better future for them and society, health professionals need to carry out awareness campaigns to educate the mothers on what to do to safeguard their children's lives.

The variable importance analysis conducted using random forest showed that the time the child was put to the breast after birth was the most important contributor to under-five mortality (Jaiyeola *et al.,* 2016, Adegbosin *et al.,* 2019). This shows that breastfeeding is very important to child’s health, as also pointed out by Azuine et al. (2015). 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 showed that children born to women aged 45 to 49 had a higher tendency to die before the age of five compared to those born to women

UNDER PEER REVIEW

aged 25-29 (Adetoro, and Amoo, 2014). Children born to women classified as poorest had a higher risk of dying before the age of five than those in the other categories (Yaya *et al.,* 2017). Children who were put to breast 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.

The model's accuracy on the test set for logistic regression was 60%, while that of the deep neural network was 74%. This 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 the neural network was 77%. F1 scores for the two models were 64% and 84%, respectively, for logistic regression and neural network. Recall for logistic regression was 63%, while that of the neural network was 75%. Logistic regression had a precision of 97%, while the neural networks had a precision of 95%. The result showed that the neural network model slightly outperformed 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.

UNDER PEER REVIEW

REFERENCES

Adedini, S. A., Odimegwu, C., Imasiku, E. N. and Ononokpono, D. N. (2015). Ethnic differentials in under-five mortality in Nigeria, *Ethnicity & Health,* 20(2):145-162.

Adegbosin, A. E., Stantic, B. and Sun, J. (2019). Predicting Under-five mortality across 21 Low and Middle Income Countries using Deep Learning Methods, retrieved from http://dx.doi.org/10.1101/19007583

Adetoro, G. W. & Amoo, E. O. (2014). A Statistical Analysis of Child Mortality: Evidence from Nigeria, *Journal of Demography and Social Statistics*, (March, 2014):110-120

Anders, U. (1996). Statistical Model Building for Neural Networks. AFIR Colloqium.Nunberg, Germany. http://www.actuaries.org/AFIR/Colloquia/Nuernberg/Anders.pdf

Azuine, R. E., Murray, J., Alsafi, N. and Singh, G. K. (2015). 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* 4(1):13-21

Biradar, R., Patel, K. K. and Prasad, J. B. (2019). Effect of birth interval and wealth on under-5 child mortality in Nigeria, *Clinical Epidemiology and Global Health*, 7(2019):223-238

Breiman, L. (2001). Random Forest, *Machine Learning,* 45, 5–32

Ezeh, O. K., Agho, K. E., Dibley, M. J., Hall, J. J. and Page, A. N. (2015). Risk factors for postneonatal, infant, child and under-5 mortality in Nigeria: a pooled cross-sectional analysis, *BMJ Open*, 2015(5):1-9

Federal Office of Statistics/Nigeria and Institute for Resource Development - IRD/Macro International. 1992. *Nigeria Demographic and Health Survey 1990*. Columbia, Maryland, USA: Federal Office of Statistics/Nigeria and IRD/Macro International. Retrieved from https://dhsprogram.com/pubs/pdf/FR27/FR27.pdf

Jaiyeola, M. O., Oyamakin, S. O., Akinyemi, J. O. Adebowale S. A., Chukwu, A. U. and Yusuf, O. B. (2016). Assessing Infant Mortality in Nigeria Using Artificial Neural Network and Logistic Regression Models, *British Journal of Mathematics & Computer Science*, 19(5):1-14

Morakinyo, O. M. and Fagbamigbe, A. F. (2017). Neonatal, infant and under-five mortalities in Nigeria: An examination of trends and drivers (2003-2013), *PLoS ONE* 12(8):1-21

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

Nanga, S. and Lotsi, A. (2016). A Comparison Between Logistic Regression and Neural Networks in Modeling Mortality Amongst Children Under Five Years in Ghana, *IOSR Journal of Mathematics (IOSR-JM)*, 12(5):26-31

Sakr, S., Elshawi, R., Ahmed, A., Qureshi, W., Brawner, C., Keteyian, S., Blaha, M. and Al-Mallah, M. (2017). 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) 17:174.

The World Bank, (2013). Nigeria - Demographic and Health Survey 2013. Retrieved from. https://datacatalog.worldbank.org/dataset/nigeria-demographic-and-health-survey-2013

UNICEF (2019). Under-Five Mortality. Retrieved from https://data.unicef.org/topic/child-survival/under-five-mortality/

UNDER PEER REVIEW

Wang, P. Li, Y. and Reddy, C. K. (2017). Machine Learning for Survival Analysis: A Survey. *ACM Computing Surveys*, 1(1):1–39.

WHO (2019). Children: reducing mortality. Retrieved from https://www.who.int/news-room/fact-sheets/detail/children-reducing-mortality.

World Bank (2019). Mortality rate, under-5 (per 1,000 live births). Retrieved from https://data.worldbank.org/indicator/SH.DYN.MORT?end=2018&start=1960&view=chart

Xu, R. (2013). Improvements to random forest methodology, *Graduate Theses and Dissertations.* 13052. https://lib.dr.iastate.edu/etd/13052

Yaya, S., Ekholuenetale, M., Tudeme, G., Vaibhav, S., Bishwajit, G. and Kadio, B. (2017). Prevalence and determinants of childhood mortality in Nigeria, *BMC Public Health* 17:485

UNDER PEER REVIEW