# A Predictive Analytics Approach to under-five Mortality measures for Improving Child Health Outcomes

## Oyamakin S. O., Onikede I. O. and  Baruwa M. D.
*Biostatistics Unit, Department of Statistics, University of Ibadan, Nigeria*
*fm_oyamakin@yahoo.com, olatunji5007@gmail.com, dbmoshood@gmail.com*

## ABSTRACT
*This study aims to develop a predictive model for under-5 mortality rates in Nigeria using a recently created database of mortality data from countries with reliable vital registration systems. The model will be based on a generalized linear model, random forest, k-means, and trend algorithms and will be customized to Nigeria's unique health and demographic characteristics. The goal is to create a model that can accurately estimate under-5 mortality rates for Nigeria, which can be used to inform evidence-based policy creation and decision-making. The model was able to accurately estimate under-5 mortality rates for Nigeria, with a root mean squared error (RMSE) of 0.00365558. The model was found to be highly accurate, while model's performance was evaluated using a hold-out test set, which showed a strong correlation between the predicted and actual values. The model was able to capture the underlying patterns and trends in the data, including the effects of seasonality and other factors on under-5 mortality rates. The study's findings have the potential to improve children's health outcomes and reduce child mortality rates in Nigeria. The predictive model developed in this study can be used to inform evidence-based policy creation and decision-making, and can help to identify areas where interventions may be needed to address high mortality rates.*
*Keywords: Child mortality, Algorithm, Classification Models, Predictive Models, Modeling*

---------------------------------------------------------------------------------------------------------------------------------------

---------------------------------------------------------------------------------------------------------------------------------------

## I.    INTRODUCTION

In 2020, an estimated 5 million children under the age of 5 died worldwide (WHO, 2016), mostly from preventable and treatable causes (Shegaw, 2002). About half of these deaths, 2.4 million, occurred in newborns within the first 28 days of life (WHO, 2014). Despite a global decrease in under-5 mortality rate to 37 deaths per 1000 live births, children in sub-Saharan Africa still have the highest mortality rates (WHO, 2015), with 74 deaths per 1000 live births - 14 times higher than the risk for children in Europe and North America. Preterm birth complications, birth asphyxia/trauma, pneumonia, diarrhoea, and malaria are the leading causes of death in children under 5 (WHO, 2015), all of which can be prevented or treated with affordable interventions in health and sanitation (UNICEF, 2016).
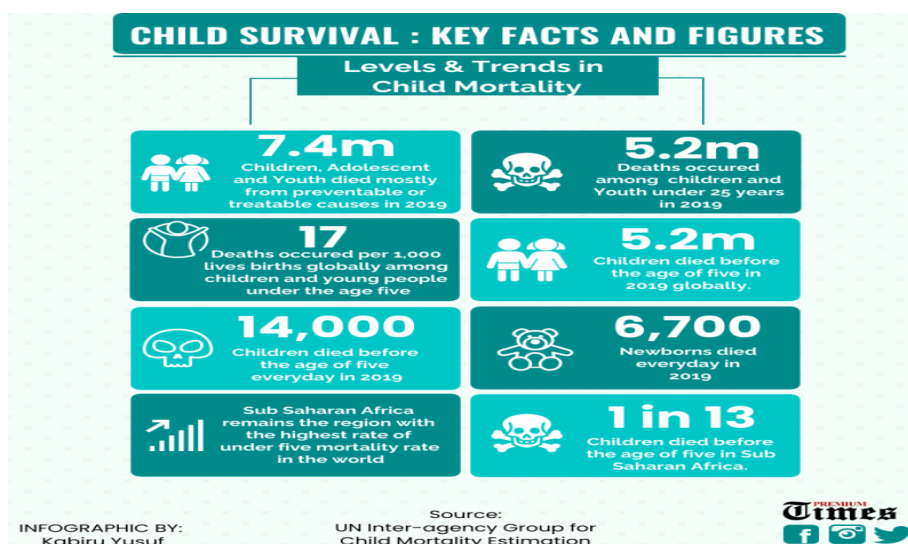


**Figure 1:  Key Facts and Figures on Child Survival in Sub-Sahara Africa**

---

There has been significant progress in reducing under-five mortality rates since 1990, with the total number of under-5 deaths worldwide declining from 12.6 million to 5 million in 2020, (UNICEF, 2016) as displayed in figure 1. The global under-5 mortality rate has decreased by 60%, from 93 deaths per 1000 live births in 1990 to 37 in 2020. The majority of under-five deaths occurred in just five countries: Nigeria, India, Pakistan, the Democratic Republic of the Congo, and Ethiopia, with Nigeria and India alone accounting for almost one-third of all deaths, (Adetoro, &Amoo, (2014).

In 2020, under-five mortality rates ranged from 2 deaths per 1000 live births to 115 deaths per 1000 live births at the country level, with the highest mortality country having a risk of dying before the age of five about 65 times higher than the lowest mortality country, (Akinwande et. al., 2016).
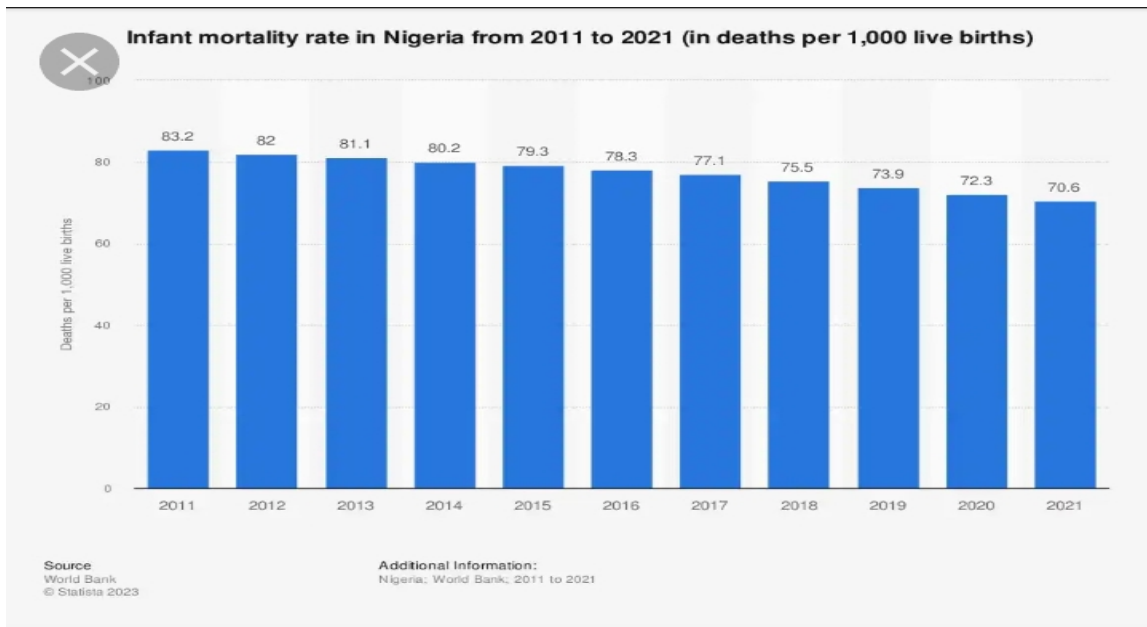


**Figure 2a: Infant mortality rate in Nigeria from 2011 to 2021 (in deaths per 1000 live births)**
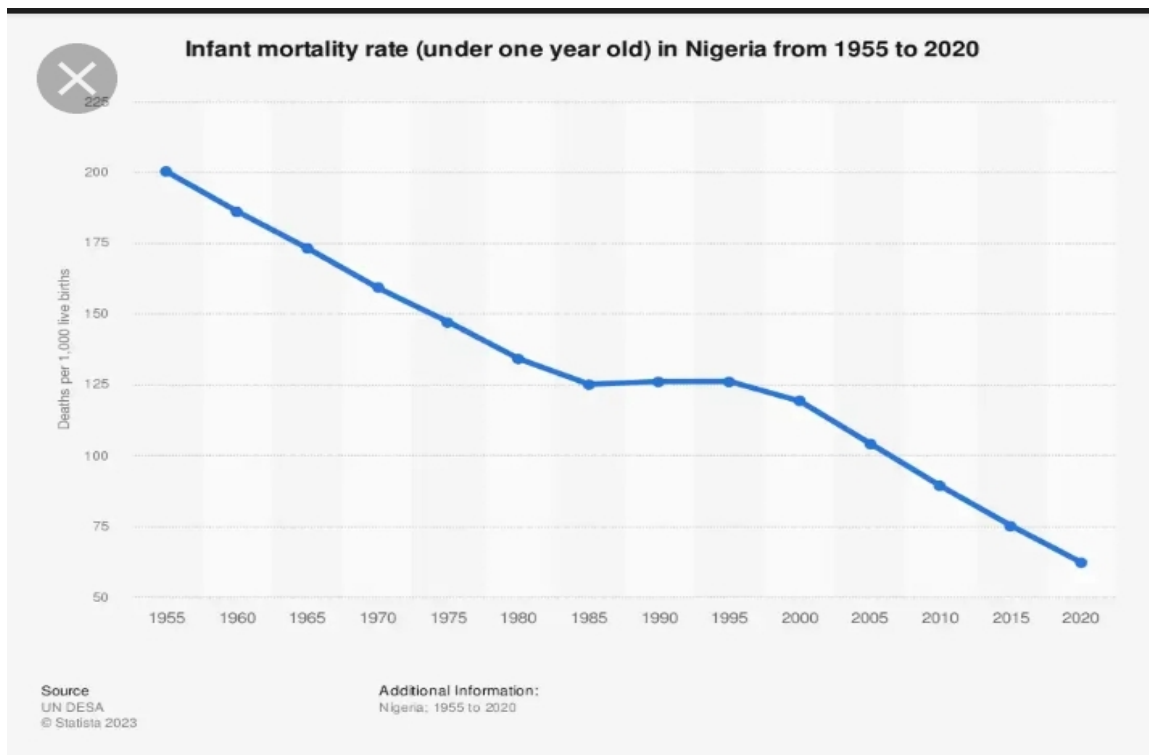


**Figure 2b: Infant mortality rate (under one year old) in Nigeria from 1955 to 2020**

According to the data presented in figure 2a, the Infant mortality rate in Nigeria for the years 2011 to 2021 showed the following values: 83.2 (2011), 82 (2012), 81.1 (2013), 80.2 (2014), 79.3 (2015), 78.3 (2016), 77.1 (2017), 75.5 (2018), 73.9 (2019), 72.3 (2020), and 70.6 (2021). These values represent the number of infant deaths per 1,000 live births. The data reveals a significant reduction of 15.1% in the infant mortality rate from 2011 to 2021 and as depicted in Figure 2b, which illustrates the infant mortality rate (under one year old) in Nigeria from 1955 to 2020, there is a clear and continuous downward trend in the infant mortality rate during this period. The data indicates a decline from approximately 200 deaths per 1,000 live births in 1955 to around 60 deaths per 1,000 live births in 2020. This trend highlights the significant improvement in infant survival and underscores the efforts and advancements made in child health and healthcare services in Nigeria over the years.

Under-five mortality is often used as an indicator of child health care status (Adegbosin et. al., 2019) due to several reasons. Firstly, it provides an outcome-based measurement as opposed to input-based measurements such as the percentage of immunized children or the number of doctors in a country. Under-five mortality represents the final outcome of efforts to keep children healthy, (Adekola et. al., 2014) reflecting how many children under the age of five have survived. Secondly, high under-five mortality rates typically stem from multiple factors including inadequate nutrition, low immunization rates, poor maternal health and education, among others. As such, it serves as a valuable indicator of health inequity and systemic health challenges. Lastly, preventable deaths comprise the majority of under-five mortality rates. Therefore, under-five mortality serves as a reliable measure of access to quality health care, including critical services such as family planning, pre- and post-natal care, and disease prevention and management.

Making future predictions about unknown events with the help of techniques from data mining, statistics, machine learning, math modeling, and artificial intelligence is known as predictive analytics (Mahajan and Ganpati, 2014). A predictive algorithm is a computational model that utilizes statistical and mathematical techniques to identify patterns and make predictions about future outcomes. Predictive algorithms are commonly used in a wide range of industries, including finance, marketing, and healthcare, to name a few. With the help of past data, it makes predictions. We use predictive analytics in our day-to-day life without giving much thought. For example, predicting sales of an item (say chicken) in a market for a particular day. If it is Christmas day, the sales of chicken would be high! We can easily say that the sales of chicken would be higher on festive days than on regular days. Hence, the purpose of this article was to enhance health interventions and strengthen the health system using a Predictive Algorithm.

## II. METHODOLOGY

The use of a newly compiled database containing under-5 mortality information by detailed age in countries with high-quality vital registration systems adds to the significance and reliability of the study's findings. The extensive coverage of mortality levels and patterns in various populations enabled us to develop a robust and comprehensive model capable of accurately predicting under-5 mortality values for Nigeria. This predictive capability could have significant implications for understanding and addressing child mortality in Nigeria, a country with unique demographic and health challenges.

The availability of high-quality data and the development of a predictive model specific to Nigeria's context can support evidence-based decision-making and policy development, ultimately contributing to improved child health outcomes and mortality reduction efforts in the country.

Overall, this study's approach and the use of a comprehensive database represent a valuable contribution to the field of public health and demography, aiming to enhance our understanding of under-five mortality patterns and guide targeted interventions to save young lives. https://web.sas.upenn.edu/global-age-patterns-under-five-mortality/.

**Variable Used in the study -** Here's a summary of what each variable represents:
i. Country:25 countries
ii. Year interval:1921 – 2014
iii. sex: Male, Female both sexes combined
iv. age_d: Lower bound of the age interval in exact days. This variable denotes the starting age of a specific interval, with each interval typically covering one year (365.25 days).
v. n_d: Length of the age interval in exact days. This variable represents the duration of the age interval, typically covering one year or a fraction of a year.
vi. mx: Central death rate between age_d and age_d + n_d. The "mx" variable provides the central death rate, which indicates the mortality rate within the specific age interval.
vii. qx: Cumulative probability of dying from birth until age_d + n_d. This variable represents the cumulative probability of dying from birth to the end of the specified age interval.
viii. q5: Under-five mortality rate (q(5y)). The "q5" variable denotes the mortality rate for children under the age of five (under-five mortality rate).

    ix.    excl: Country-year excluded from the log-quad model due to data quality concerns. This variable takes a value of 1 to indicate that the specific country-year combination has been excluded from the log-quad model due to concerns related to data quality or reliability.

These variables help in assessing mortality trends, calculating probabilities, and identifying data quality issues for more accurate and reliable predictions and conclusions.

**Trend -** A trend in time series datasets refers to the general direction or movement observed over time. It is used to describe whether the data shows a consistent upward or downward pattern for some or all of the time series. In time series analysis, identifying and understanding trends are essential for making predictions, forecasting future values, and gaining insights into the underlying patterns and behaviors of the data. Trends can be either upward (indicating growth or increasing values) or downward (indicating decline or decreasing values) and can provide valuable information about the long-term behavior of the data. Analyzing trends is a fundamental aspect of time series analysis and is used in various fields, including economics, finance, environmental studies, and many others, to make informed decisions and projections based on historical data patterns.

**K – Means -** K-means is a widely used clustering algorithm that aims to partition a set of observations (data points) into k clusters, where each observation belongs to the cluster with the nearest mean (centroid). The algorithm iteratively updates the cluster centroids until convergence, creating Voronoi cells that divide the data space based on the distances to these centroids. The objective of k-means clustering is to minimize the within-cluster variances, which is equivalent to minimizing the squared Euclidean distances between data points and their respective cluster centroids. However, it's important to note that k-means does not minimize regular Euclidean distances, which would be a more complex optimization problem known as the Weber problem. In the Weber problem, the mean would optimize squared errors, while only the geometric median (e.g., k-medians) or geometric medoid (e.g., k-medoids) would minimize regular Euclidean distances.

Also, alternative clustering algorithms like k-medians and k-medoids exist to address cases where minimizing regular Euclidean distances is more appropriate or when the data contains outliers. Each of these algorithms has its strengths and weaknesses, and the choice of clustering method depends on the specific characteristics of the data and the objectives of the analysis. Given a set of observations ($\mathbf{x}_1$, $\mathbf{x}_2$, ..., $\mathbf{x}_n$), where each observation is a $d$-dimensional real vector, $k$-means clustering aims to partition the $n$ observations into $k$ ($\leq n$) sets for $\mathbf{S}$ = {$S_1$, $S_2$, ..., $S_k$} so as to minimize the within-cluster sum of squares (WCSS) (i.e. variance). This is explained analytically as follows;

where $\boldsymbol{\mu}_i$ is the mean (also called centroid) of points in $S_i$, i.e.

is the size of $S_i$, and is the usual $L^2$ norm. This is equivalent to minimizing the pairwise squared deviations of points in the same cluster:

The equivalence can be deduced from identity

Since the total variance is constant, this is equivalent to maximizing the sum of squared deviations between points in different clusters (between-cluster sum of squares, BCSS). This deterministic relationship is also related to the law of total variance in probability theory.(Kriegel, Hans-Peter; Schubert, Erich; Zimek, Arthur (2016) ).

**Random Forests -** Random Forests are an ensemble learning technique that combines multiple decision trees to improve predictive accuracy and robustness. They are particularly well-suited for real-world datasets that may contain missing or error-prone data and may not have enough information for conventional methods. One of the strengths of Random Forests is their ability to handle large datasets efficiently. They can operate quickly over substantial amounts of data, making them suitable for big data applications. Random Forests have been widely used in various research projects and real-world applications across different domains due to their versatility and effectiveness.

The technique is based on Bagging (Bootstrap Aggregating), a method proposed by Leo Breiman in 1996. Random Forests use bootstrap sampling to generate multiple training sets, and each set is used to build a separate decision tree. The randomness is further enhanced by considering only a random subset of attributes at each node of the tree for splitting, reducing the risk of overfitting. The combination of multiple trees through voting or averaging helps to mitigate the risk of overfitting and provides more robust predictions. Overall, Random Forests have proven to be a valuable tool in the field of machine learning and data mining, offering improved accuracy and robustness for classification tasks across a wide range of real-world applications.
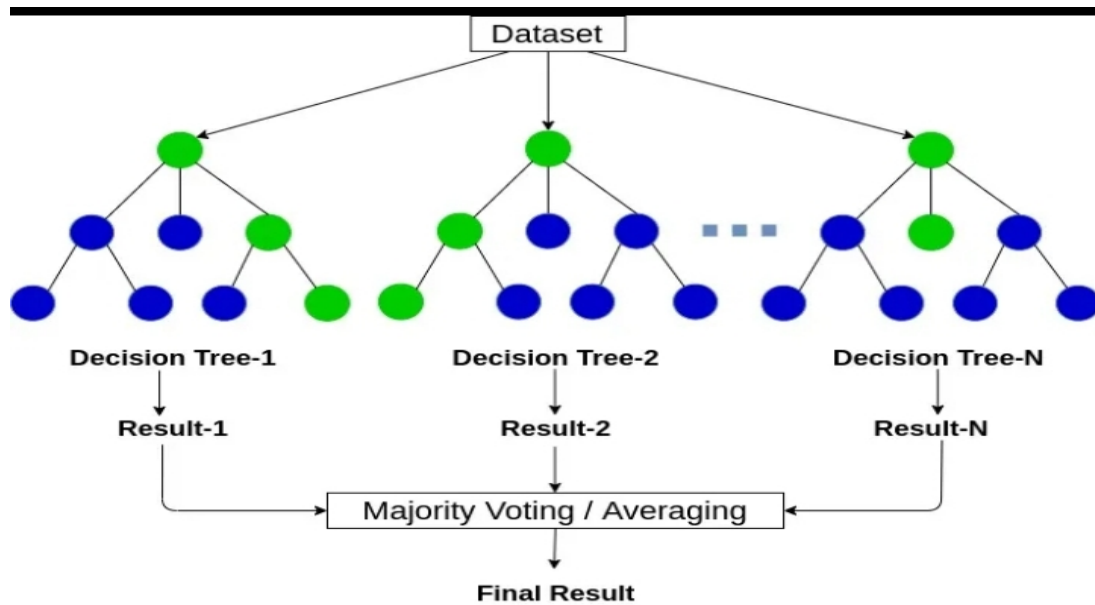
**Figure 3: Overview and structure of Random Forest Techniques for classification**

In Figure 3, the structure of a random forest is presented, illustrating the step-by-step process from the dataset to the final result through the decision trees. Dataset: The random forest process begins with a dataset that contains a collection of input features (independent variables) and corresponding target values (dependent variable). This dataset is used to train the random forest model. Random Subset and Bootstrap: For each decision tree within the random forest, a random subset of the dataset is selected through a process called bootstrapping. This involves sampling the dataset with replacement, creating different subsets for each tree. Decision Trees: The random forest consists of multiple decision trees. Each decision tree is constructed using the bootstrapped dataset, and at each node, the best split among the input features is determined based on specific criteria (e.g., Gini impurity or entropy for classification, mean squared error for regression). Ensemble Learning: The strength of a random forest lies in the ensemble learning technique. Instead of relying on the prediction of a single decision tree, the random forest combines predictions from all the decision trees. Voting (Classification) or Averaging (Regression): For classification tasks, the random forest employs majority voting, where each decision tree "votes" for the class of the input sample, and the final prediction is the class with the most votes. For regression tasks, the random forest takes the average of the individual decision tree predictions to produce the final result. Final Result: The combination of the decision trees' predictions through ensemble learning results in the final output of the random forest model. This final result is the model's prediction for a given input sample. The random forest's ability to aggregate predictions from multiple decision trees makes it a powerful and robust machine learning algorithm, suitable for various tasks like classification and regression.

**How Random Forests works -** Let us understand the working of Random Forest algorithm with the help of following steps −

**Step 1** − First, start with the selection of random samples from a given dataset.

**Step 2** − Next, this algorithm will construct a decision tree for every sample. Then it will get the prediction result from every decision tree.

**Step 3** − In this step, voting will be performed for every predicted result.

**Step 4** − At last, select the most voted prediction result as the final prediction result.

---
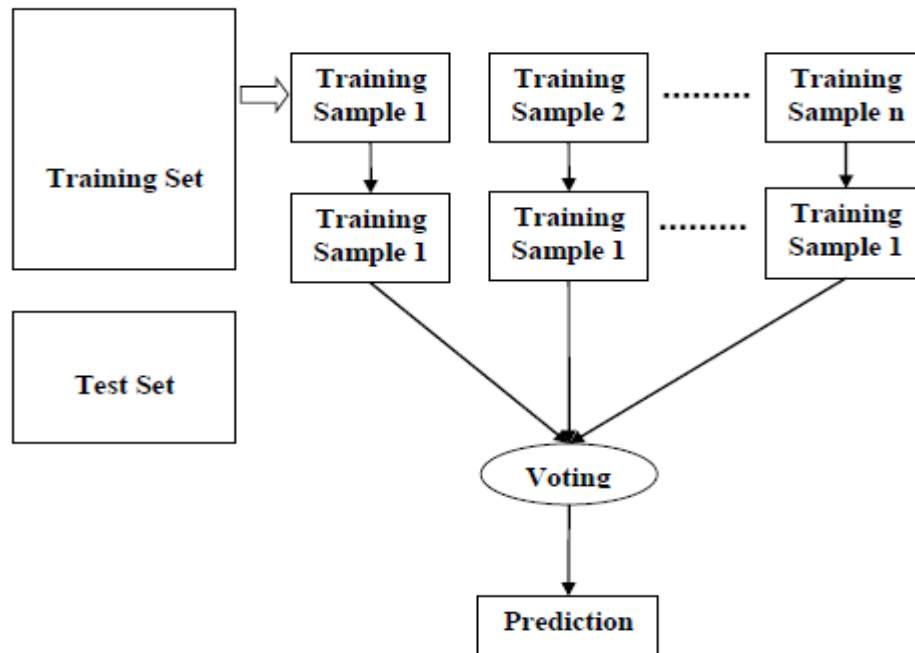
**Figure 4: Operational workings of Random Forest Classification Algorithms**

In Figure 4, the working of the random forest classification algorithm is depicted, outlining the process from the training set and test set to the final prediction.

Training Set: The random forest classification algorithm begins with a training set, which is a subset of the dataset containing labeled examples. Each example in the training set consists of input features (independent variables) and their corresponding class labels (dependent variable).

Random Subset and Bootstrap: For each decision tree within the random forest, a random subset of the training set is chosen using bootstrapping. This involves sampling the training set with replacement to create different subsets for each tree.

Decision Trees: The random forest comprises multiple decision trees, where each tree is constructed using one of the bootstrapped subsets of the training set. At each node of a decision tree, the best split among the input features is determined based on criteria such as Gini impurity or entropy, aiming to improve class separation.

Ensemble Learning: The power of the random forest lies in ensemble learning. Instead of relying on a single decision tree's prediction, the random forest combines the predictions from all the decision trees in the forest.

Voting: For classification tasks, the random forest employs majority voting. Each decision tree "votes" for the class of a given input sample in the test set. The class with the most votes across all decision trees becomes the final prediction.

Test Set: The test set is a separate subset of the dataset that is not used during training. It serves as an independent evaluation set to assess the performance of the random forest model.

Prediction: Using the trained random forest model, the algorithm makes predictions on the test set. For each example in the test set, the random forest classifies the input sample into one of the possible classes based on the majority voting outcome of the decision trees.

Final Result: The combination of predictions from all decision trees through majority voting results in the final output of the random forest classification algorithm. This final result is the model's classification prediction for each example in the test set.

The random forest classification algorithm's ability to aggregate predictions from multiple decision trees helps improve accuracy and generalization, making it a popular choice for various classification tasks.

**Overview of Generalized Linear Model -** The mean, $\mu$, of the distribution depends on the independent variables, X, through:

where $E(Y|X)$ is the expected value of Y conditional on X; $X\beta$ is the linear predictor, a linear combination of unknown parameters $\beta$; g is the link function.

In this framework, the variance is typically a function, V, of the mean:

It is convenient if V follows from an exponential family of distributions, but it may simply be that the variance is a function of the predicted value. The unknown parameters, β, are typically estimated with <u>maximum likelihood</u>, maximum <u>quasi-likelihood</u>, or <u>Bayesian</u> techniques.

## III. DISCUSSION OF RESULTS

In table 1 below, it is indicated that the mean of the under-five mortality rate is 0.4263429, and the standard deviation is 0.57176496. The mean, also known as the average, represents the central tendency of the under-five mortality data. It is calculated by summing up all the individual under-five mortality values and dividing by the total number of observations. In this case, the mean is approximately 0.4263429. The standard deviation measures the spread or dispersion of the under-five mortality data points around the mean. A higher standard deviation indicates that the data points are more spread out, while a lower standard deviation suggests that the data points are closer to the mean. In this case, the standard deviation is approximately 0.57176496, which implies that the under-five mortality rates have a moderate level of variability around the mean value of 0.4263429.

**Table 1:    Continuous Variable Information of under-five mortality**
**Continuous Variable Information**

| | | N | Minimum | Maximum | Mean | Std. Deviation |
|---|---|---|---|---|---|---|
| Dependent Variable | under five mortality | 96691 | .000000 | .336540 | .04263429 | .057176496 |

In figure 5 below, the trend pattern of under-five mortality is presented based on the data, which reveals a consistent downward movement over the specified period. The graph included data points representing different years or time intervals, and the corresponding under-five mortality rates for each period. These rates show a declining trend over time, indicating a reduction in the number of deaths among children under five years old. The data demonstrates that efforts and interventions aimed at improving child health and reducing child mortality have been effective, resulting in a positive trend towards better survival rates for young children. The downward movement in under-five mortality is a promising sign of progress and signifies the potential impact of various healthcare initiatives and policies implemented during this time period.
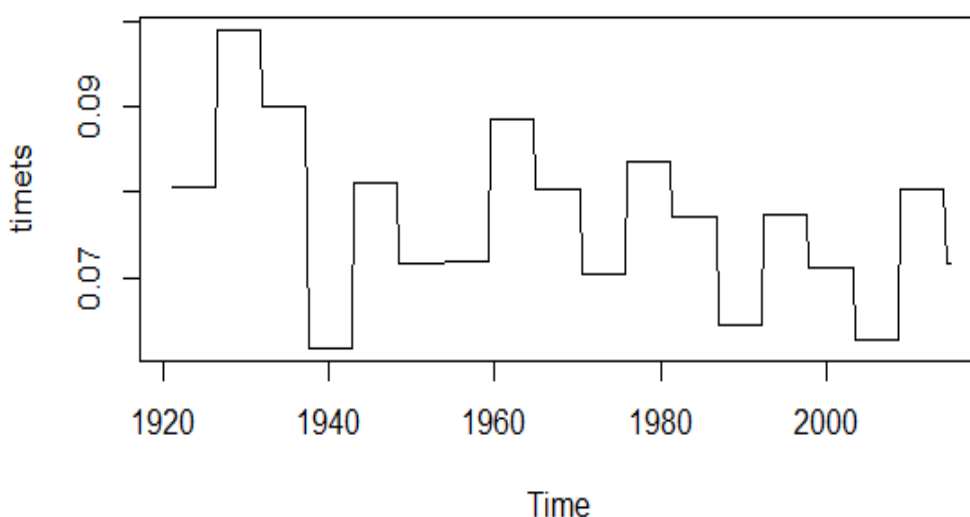


**Figure 5: Trend patterns based on under five mortality data**

The Augmented Dickey-Fuller (ADF) Test is a statistical test commonly used to determine whether a time series is stationary or not. Stationarity means that the statistical properties of the time series, such as the

mean and variance, do not change over time. A stationary time series is easier to model and analyze, making it a desirable characteristic in many cases. If the ADF test indicates that the time series is stationary, it suggests that there is no significant trend or systematic patterns in the data over time. This is often interpreted as a stable and consistent behavior of the variable being measured.

In the context of Figure 5, the time series representing under-five mortality demonstrates a stationary trend according to the ADF test results. This stationary behavior could imply that the under-five mortality rate remains relatively consistent and does not exhibit a significant increasing or decreasing trend over the analyzed period. i.e. Augmented Dickey-Fuller Test data: timets Dickey-Fuller = -3.5371, Lag order = 7, p-value = 0.03922 alternative hypothesis: stationary.



**Figure 6: The Cluster Plot of under-five mortality data**

In Figure 6, it is illustrated that the observations were assigned to different clusters as follows: The first cluster contains 16 observations, the second cluster contains 13 observations, the third cluster contains 13 observations while the fourth cluster contains 8 observations. Clustering is a technique used in unsupervised learning to group similar data points together based on their features or characteristics. Each cluster represents a distinct group with similar attributes, and the number of observations assigned to each cluster is indicated in the figure.

By clustering the data in this way, it becomes easier to identify patterns, similarities, or differences among the observations, which can be useful for various data analysis and pattern recognition tasks.

**Table 2: The mean number of variables in each cluster using under five mortality data**

| Variables | Cluster 1 | Cluster 2 | Cluster 3 | Cluster 4 |
|---|---|---|---|---|
| Under-five mortality rate | 0.15 | 0.09 | 0.02 | 0.024 |
| Lower bound of age interval in exact days | 319.3 | 0.38 | 216.86 | 1103.2 |
| Length of age interval in exact days | 76.4 | 7.0 | 38.9 | 365.3 |
| Central death rate between age_d and age_d + n_d | 0.08 | 0.95 | 0.02 | 0.00094 |
| Cumulative probability of dying from birth until age_d + n_d | 0.08 | 0.02 | 0.015 | 0.024 |
| Country-year excluded from log-quad model due to data | 0.98 | 0.40 | 0.000015 | 0.049 |

| quality concerns (value = 1) | | | | | |
|---|---|---|---|---|---|

Table 2 above shows that the mean number of Under-five mortality rate in cluster 1 is 0.15, cluster 2 is 0.09, cluster 3 is 0.02 and cluster 4 is 0.024.Also, the mean number of Lower bound of age interval in exact days in cluster 1 is 319.3, cluster 2 is 0.38, cluster 3 is 216.86 and cluster 4 is 1103.2.Also,he mean Length of age interval in exact days in cluster 1 is 76.4, cluster 2 is 7.0, cluster 3 is 38.9 and cluster 4 is 365.3.Also, the mean Central death rate between age_d and age_d + n_d in cluster 1 is 0.08, cluster 2 is 0.95, cluster 3 is 0.02 and cluster 4 is 0.00094.Also, the mean Cumulative probability of dying from birth until age_d + n_d in cluster 1 is 0.08, cluster 2 is 0.02, cluster 3 is 0.015 and cluster 4 is 0.024. Lastly the mean Country-year excluded from log-quad model due to data quality concerns (value = 1) in cluster 1 is 0.98, cluster 2 is 0.4, cluster 3 is 0.000015 and cluster 4 is 0.049.

**Table 3: Random Forest Results outputs using under five mortality data set**

| Random Forest Algorithm | Estimates |
|---|---|
| randomForest(formula = q5 ~ ., data = u5m)<br>Type of random forest: regression<br>Number of trees: 500<br>No. of variables tried at each split: 2 | Mean of squared residuals: 1.351154e-05<br>% Var explained: 99.59<br>RMSE = 0.00365558 |

From the output in table 3, it is evident that the model utilizing 500 trees achieved the lowest test mean squared error (MSE). The test MSE is a measure of how well the model performs on unseen data, and a lower test MSE indicates better generalization and predictive accuracy. Additionally, the root mean squared error (RMSE) of the model was recorded as 0.00365558. The RMSE is a commonly used metric to assess the accuracy of regression models. It represents the average difference between the predicted values and the actual observed values for the target variable, in this case, the under-five mortality rate (q5). A lower RMSE indicates that the model's predictions are, on average, closer to the actual observed values, suggesting a higher level of accuracy and reliability in predicting the under-five mortality rate. Consequently, the model utilizing 500 trees is considered to be the most effective in terms of minimizing prediction errors and providing a better fit to the data.
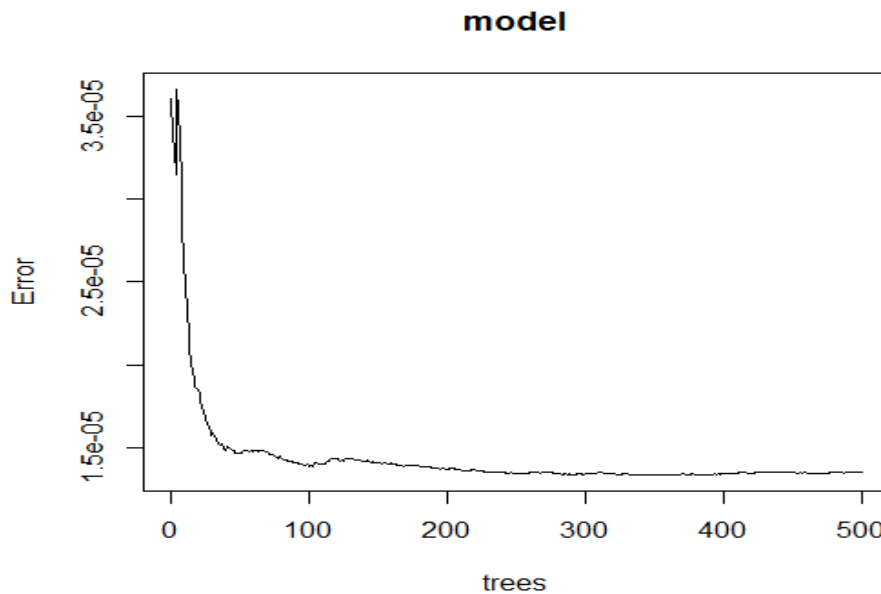


**Figure 7: Plot of the test MSE based on the number of trees used**

In Figure 7, it is depicted that the model with the lowest least mean square error (MSE) was achieved using 500 trees. The figure represents the performance of a random forest model with different numbers of decision trees. The mean square error is a metric used to measure the accuracy of regression models, and a lower MSE indicates better model performance and better fit to the data. By varying the number of trees in the random forest model and measuring the MSE for each configuration, it is observed that the model with 500 trees yielded the lowest MSE among all the tested options. This suggests that using 500 trees in the random forest Algorithm

led to the most accurate predictions and the best overall model performance for the under five mortality dataset used in this study.

## model
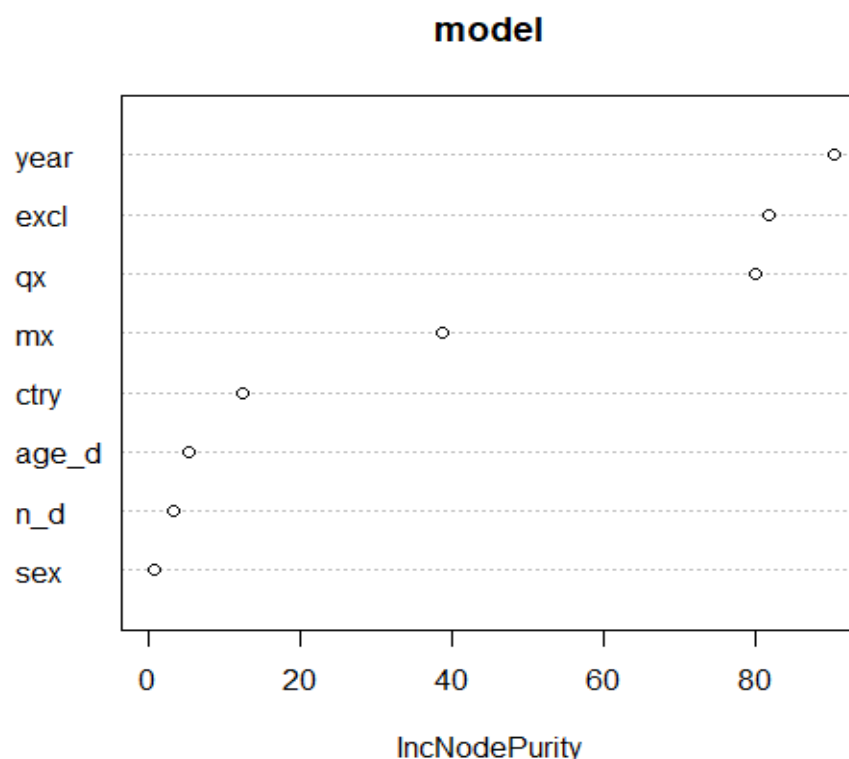


**Figure 8: Variable important Plot of under-five mortality**

Figure 8 above shows that the x axis displays the average increase in node purity of the trees based on splitting on the various predictors displayed on the y axis. From the plot we can see that year is the most important predictor variable, followed closely by excl country year excluded from log-quadratic model due to data quality concern while sex is the least predictor variable.

**Table 4: Goodness of Fit of under-five mortality    model used**

|  | Value | df | Value/df |
|---|---|---|---|
| Deviance | 316.094 | 96690 | .003 |
| Scaled Deviance | 96691.000 | 96690 |  |
| Pearson Chi-Square | 316.094 | 96690 | .003 |
| Scaled Pearson Chi-Square | 96691.000 | 96690 |  |
| Log Likelihood[b] | 139494.076 |  |  |
| Akaike's Information Criterion (AIC) | -278984.151 |  |  |
| Finite Sample Corrected AIC (AICC) | -278984.151 |  |  |
| Bayesian Information Criterion (BIC) | -278965.193 |  |  |
| Consistent AIC (CAIC) | -278963.193 |  |  |

Dependent Variable: under five mortality
Model: (Intercept)
a. Information criteria are in small-is-better form.
b. The full log likelihood function is displayed and used in computing information criteria.

In Table 4.3, the goodness of fit measures for the under-five mortality model are presented as follows: Deviance: The deviance is a measure of the discrepancy between the observed data and the model's predictions. In this case, the deviance is calculated as 3160.94. Scaled Deviance: The scaled deviance is a standardized version of the deviance and is often used in comparison with other models. In this context, the scaled deviance is reported as 96691. Pearson Chi-Square: The Pearson chi-square is another measure of how well the model fits the data. It quantifies the difference between the observed and expected values based on the model. In this case, the Pearson chi-square is reported as 316.09. AIC (Akaike Information Criterion): The AIC is a metric used for model selection, taking into account both the goodness of fit and the model's complexity. A lower AIC value

indicates a better trade-off between model fit and complexity. In this case, the AIC is reported as -278984.151. Overall, these goodness-of-fit measures provide important insights into how well the under-five mortality model performs in representing the observed data. A good model should have low deviance, scaled deviance, and Pearson chi-square values, while also having a lower AIC value, indicating a better fit to the data with a balanced level of model complexity.

**Table 5: Tests of Model Effects based on under-five mortality data**

| Source | Type III | | |
|---|---|---|---|
| | Wald Chi-Square | df | Sig. |
| (Intercept) | 53761.756 | 1 | .000 |

Dependent Variable: under five mortality
Model: (Intercept)

In the context of Table 5, the Wald chi-square value of 53761.756 represents the result of testing the significance of the coefficients in the logistic regression model using under-five mortality data. A higher Wald chi-square value indicates that the coefficients are more significant and have a stronger impact on the outcome variable (under-five mortality rate) in the logistic regression model. Conversely, a lower value would suggest that the coefficients are less significant and may not have a substantial effect on the outcome variable.

**Table 6: Model Parameter Estimates based on under-five mortality data**

| Parameter | B | Std. Error | 95% Wald Confidence Interval | | Hypothesis Test | | |
|---|---|---|---|---|---|---|---|
| | | | Lower | Upper | Wald Chi-Square | df | Sig. |
| (Intercept) | .043 | .0002 | .042 | .043 | 53761.756 | 1 | .000 |
| (Scale) | .003[a] | 1.4868E-005 | .003 | .003 | | | |

Dependent Variable: under five mortality
Model: (Intercept)
a. Maximum likelihood estimate.

In Table 6, it is stated that the under-five mortality model parameter estimate for the intercept is 0.043. The under-five mortality model parameter estimates of 0.043 represents the coefficient associated with the intercept term in the logistic regression model. The intercept term is the constant term in the model that captures the baseline or average log-odds of the event of interest (in this case, under-five mortality) when all the independent variables (input features) are equal to zero. The coefficient of 0.043 indicates the estimated increase in the log-odds of under-five mortality when all the independent variables are set to zero.

## IV. CONCLUSIONS

This study provided reliable and genuine information that could be useful for the relevant stakeholders and decision – makers to project the under-five mortality rate globally. It is also meant to alert the Nigeria government and stakeholders in Child health to prioritize investments in evidence-based interventions so as to further improve child's physical health and overall well-being thereby ensuring a brighter and healthier future for the next generation. A strong recommendation for the government is to prioritize the well-being of children under five years old, aligning with other nations and the international community that have consistently demonstrated their dedication to improving child health. However, this commitment must not remain merely symbolic; instead, it should be translated into concrete and impactful measures aimed at preventing avoidable deaths, illnesses, disabilities, and impaired psychosocial development among children. By doing so, we can work towards the ultimate objective of improving the overall quality of life for children on a global scale.

## REFEREENCES

[1]. Mahajan and A. Ganpati, (2014). "Performance evaluation of rule-based classification algorithms," International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Vol, vol. 3, pp. 3546-3550.
[2]. Shegaw, A. (2002)."Application of data mining technology to predict child mortality patterns: the case of Butajira Rural Health Project (BRHP)," aau,
[3]. 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 https://dx.doi.org/10.1101/19007583
[4]. Adekola, P. O., Allen, A. A. and Akintunde, E. A. (2014). Environmental factors affecting infant mortality in Ibadan North Local Government Area of Nigeria, African Journal of Social Sciences, 4(1):53-67
[5]. Adetoro, G. W. &Amoo, E. O. (2014). A Statistical Analysis of Child Mortality: Evidence from Nigeria, Journal of Demography and Social Statistics, 10-120
[6]. Akinwande, M. O., Ibrahim, A. and Ibrahim A. A. (2016). Analyzing Infant and Child (Under-five) Mortality in Zaria: A Regression Analysis Approach, African Research Review, 10(3):72-95
[7]. U. N. C. Fund, The State of the World's Children (2016): Executive Summary. A Fair Chance for Every Child. UNICEF
[8]. W. H. Organization, (2014) "WHO methods and data sources for country-level causes of death 2000– 2012," Geneva, Switzerland: WHO

[9]. W. H. Organization, Health in (2015). From MDGs, millennium development goals to SDGs, sustainable development goals. World Health Organization

[10]. W. H. Organization, (2016). World Health Statistics 2016: Monitoring Health for the Sustainable Development Goals (SDGs). World Health Organization