Science World Journal Vol. 19(No 4) 2024 https://dx.doi.org/10.4314/swj.v19i4.9
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Full Length Research Article

MACHINE LEARNING PREDICTION MODELS OF BIRTH WEIGHT OF NEW BORN BABIES IN FCT ABUJA, NIGERIA

Alegbe A.A., *Adenomon M.O. & Maijamaa B.

Department of Statistics, Nasarawa State University, Keffi, Nigeria

*Corresponding Author Email Address: adenomonmo@nsuk.edu.ng

ABSTRACT

This research aimed at creating a machine learning model for predicting birth weight using the maternal risk factors that have been found to be associated with low birth weight. The data covered a period of ten years from 2010 to 2019 was utilized where the variables were extracted from the births recorded file. The study population included mothers between the age of 15 to 49 years. The machine learning algorithms employed were logistic regression, Decision trees, Random Forest, Support Vector Machines, Gradient Boosting and K -Nearest Neighbors, Neural Network, Gradient boosting and Linear Regression. The metrics used for classification method were Accuracy, Sensitivity, Specificity and Kappa. In terms of accuracy, the best machine learning model was the Decision tree with an accuracy of 0.9823. The other five models produced an accuracy that ranged between 0.9806 to 0.9822. Based on the kappa, decision tree again emerged to be the best with a value of 09631. The rest of the models had a kappa that ranged from 0.8859 to 0.9592. Sensitivity was also evaluated and Neural Network and support vector machine had the same sensitivity value of 0.9941 whereas the other models managed a recall score ranging from 0.9501 to 0.9853. Moreover, Specificity was also examined. Logistic Regression model had the best specificity value of 0.9908. The rest of the models ranged from 0.9378 to 0.9778. Furthermore, the ROC curves of all the tested models were plotted and the area under the curved evaluated. The decision tree had the highest area under the curve of 0.9896. The AUC of the other models ranged from 0.9440 to 0.9816. Therefore, from these results based on the performance metrics and ROC-AUC, decision tree emerged to be the most robust model for classification method. Furthermore, the ROC-AUC was used to test the classification ability of the models to differentiate between the low-birth-weight cases and the cases without low birth weight.

Keywords: Machine learning, birth weight, low birth weight, maternal risk factors, prediction, algorithm.

INTRODUCTION

Birth weight is one of the significant predictors of child mental development, future physical growth, and survival. It is one of the important risk factors for child morbidity and mortality, Daynia et al (2002). A new born's health is a primary factor that determines the overall health of a human being and its life expectancy. Therefore, its health should be monitored not only after birth but also when the baby is still growing in the womb. Birth weight is one of the crucial aspects to be observed. Low birth weight is among the main problems that new born's face. Low birth weight (LBW) is the weight at birth less than 2500g as defined by the World Health Organization. A global estimate of 15 to 20 percent of total live births are low birth weight representing over twenty million births every year. In Nigeria, LBW affects about 5–6 million children every year, Olu (2005). The incidence was 12.1% in Jos, Yilgwan et al (2009), 11.4% in Ogun, Olowonyo et al (2006), and 16.9% in Maiduguri, Takai et al (2014). Several methods have been used to measure and approximate birth weight in clinical practice including obstetric ultrasound, symphysio-fundal height measurements and abdominal palpation. However, these methods are associated with reliability and accuracy challenges therefore, calling for more robust methods hence these challenges called for the use of machine learning. Machine learning (ML) is a subject that focuses on the data analysis using various statistical tools and learning processes in order to gain more knowledge from the data by turning a huge amount of data into predictions. The objective of this study is to Evaluate the performance of the machine learning techniques on birth weight using classification models, to determine the most robust machine learning technique for predicting birth weight using classification models, to Identify the most important variables for predicting birth weight classification models. The binary logistic regression model was firstly employed on the train and the test data. Then, the random approach was also applied to the data set. The study explores the efficacy of models such as Random Forest, Linear Regression, Neural Network (ML Regressor), Support Vector Machine, Gradient Boosting, Decision Tree, and K-Nearest Neighbors.

CONCEPTUAL FRAMEWORK

Birth Weight

Birth weight is one of the significant predictors of child mental development, future physical growth, and survival. It is one of the important risk factors for child morbidity and mortality, Daynia et al (2002). According to the World Health Organization (WHO), low birth weight (LBW) is defined as an infant birth weight of less than 2500g. LBW can arise as a result of a baby being born too soon (at less than 37 weeks, also known as preterm birth) and/or being born too small for gestational age (small as a result of intrauterine growth restriction). This group of children is considered to have higher risk of neonatal, post-neonatal death, and morbidity, Gupta (2008).

Those who survive are more likely to remain malnourished, have impaired immune function and face increased risk of health and developmental problems including lower intelligent quotients and cognitive disabilities leading to learning difficulties, school failure and poor job opportunities.

There is significant difference in the incidence of LBW between developed and developing countries and between various regions in a country. In developed countries, the occurrence is 7%, while in developing countries it is 15%, Ramakrishnan (2004). Globally, recent estimates suggest that there were 18 million of LBW babies born every year. Badshan et al (2008). In sub-Saharan Africa, the prevalence of LBW varies according to the regions. The prevalence of LBW in Ethiopia was 28.3% Assefa et al (2012) while there were

19% LBW infants per 1,000 live births in Zimbabwe, Fesura et al (2004).

Social – demographic factors of Nigerian mothers.

In the developing world, LBW stems primarily from poor maternal health and nutrition. Giving birth to a LBW baby is influenced by several determinants including maternal variables which are characterized as social-demographic factors like maternal age (<20 or >35 yrs), ethnicity, marital status, birth interval, educational level, place of residence and inadequate antenatal weight gain. A variety of maternal social-demographic factors are also known to increase the risk of low birth weight. Gupta et al (2007) Constitutional, gender and hereditary factors explain up to 40% of the variability of birth weight. Blanc et al (2005) Medical risk factors for LBW before pregnancy are chronic conditions like hypertension, renal insufficiency, cardio-respiratory, autoimmune, endocrine or infectious disorders. Takai (2014) conducted a prospective study of maternal risk factors for low-birth-weight babies in Maiduguri, Borno state, Nigeria. The study involved 854 pregnant women and their babies between February 2009 and July 2009. Social-demographic, obstetric, medical factors, obstetrics interventions and foetal birth weights were obtained and recorded. Association between variables were examined using student's *t*-test and Chi-squared test and multivariate logistic regression analysis a $P<0.05$ was considered significant. He concluded that the study showed high incidence of LBW of 16.9% which the study also indicated that poor maternal health associated with social demographic factors mentioned above could be a primary cause of Low birth weight in Nigeria.

Social – economic factors of Nigerian mothers.

More than 5-6 million (17%) babies are born with LBW in Nigeria, with nearly 100,000 resulting in fatality, Chidiebere et al., (2018). A study by Dahlui analyzed data from the 2013 Nigeria Demographic and Health Survey and reported the prevalence of LBW to be 7.3%. High incidences have been reported in states across the nation. For example, in Maiduguri, Takai, et al (2014) reported a high incidence (16.9%) of LBW; in Lagos, the incidence was 10.2%, Olusanya et al (2010), while an even higher incidence of LBW was reported in Kano at 32.1% (951 babies), which was associated with increased neonatal morbidity and mortality rates (Muktar-Yola et al 2007). In the aforementioned studies conducted in Nigeria, social – economic factors such as low income and standard of living, marital status, maternal factors, anaemia, poor nutritional status of mothers at conception, short maternal height and lack of antenatal care were highlighted as risk factors for LBW. Oladeinde, et al (2015) highlighted this as a major concern, especially since traditional birth attendants (TBAs) deliver a majority of pregnant women, and data on LBW of babies born in such centers is unaccounted for. In the North Central region where Abuja is located, studies have reported the prevalence of LBW to be 30.5% in Benue, Ochogah, et al (2018) and 16.9% in Jos Dahlui, et al (2016).

Obstetric factors of Nigerian mothers.

Obstetric factors category which describes aspects of pre-partum, pregnancy and childbirth. Four variables were included in this category. These are; the obstetric history which describes scenarios such as whether the respondent ever had a pregnancy that terminated in a miscarriage, abortion or still birth, the other variable in this category is number of antenatal visits, whether the mother took iron folic tablets while pregnant and the smoking status of the mother and the child's birth order. On the other hand, the dependent variable is low birth weight labeled as LBW which is binary in nature with two categories coded as 1 for the category with low birth weight and 0 for the category not having low birth weight.

EMPIRICAL REVIEW/REVIEW OF PREVIOUS STUDIES

Amene et al (2023) published in the BMC pregnancy and Childbirth journal, employed a machine learning model to predict low birth weight (LBW). The researchers used eight statistical learning models, including logistics regression, decision tree classification, random forest classification, deep learning feed forward, extreme gradient boost model, light gradient boost model, support vector machine, and permutation feature classification with k-nearest neighbors. The study used a retrospective cohort design and included 8853 deliveries from Iranian Maternal and Neonatal Network (IMaN Net). The researchers used a predictive model built using eight statistical learning model and evaluated their performance using metrics such as AUROC, accuracy, precision, recall and F1 score. They found that extreme gradient boost model performed well in predicting LBW with an accuracy of 0.79,precision of 0.87, recall of 0.69 and F1 score of 0.77. Their key findings are; that gestation age and previous history of LBW were the top critical predictors, the extreme gradient boost model outperformed other models in predicting LBW and the study also highlighted the potential of machine learning approaches in predicting LBW and identifying high-risk pregnant patients early in their pregnancy. They concluded by highlights the importance of identifying high-risk pregnancy patients early in their pregnancy. However, the researchers noted that more research is needed to make a better conclusion on the performance of ML models in predicting birth weight which the study tend to fill the gap by identifying Random forest ac the best performing ML for Regression method and Decision tree on the hand for best performing ML for regression method.

Assefa et al (2012) conducted a research to predict birth weight using machine learning techniques specifically decision trees, neural networks, and support vector machines. The researchers used a data set of 1000 records from an Ethiopian hospitals, with features such as maternal age, gestational periods and medical history. They found decision trees performed best, with an accuracy 85.6% and mean absolute error (MAE) of 240.8 grams. Also, Neural networks achieved an accuracy of 82.1% and MAE of 261.4 grams and Support vector machine had an accuracy of 79.5% and MAE of 281.9 grams. The study concluded that machine learning techniques can accurately predict birth weight with decision trees showing the best performance. However, the study had a limited data set and did not consider feature engineering or hyper parameter turning. In addition, further research is needed to improve model performance and generalizability .This study could enhance Assefa et al research by employing a supervised machine learning technique models capable of capturing both classification and regression models that will be able to produce more accurate results. The metrics used for classification method were Accuracy, Sensitivity, Specificity and Kappa. In terms of accuracy, the best machine learning model was the Decision tree with an accuracy of 98% in line with the result of the previous researchers. The other five models produced an accuracy that ranged between 0.9806 to 0.9822. Based on the

 https://dx.doi.org/10.4314/swj.v19i4.9

kappa, decision tree again emerged to be the best with a value of 0.9631. The rest of the models had a kappa that ranged from 0.8859 to 0.9592. Sensitivity was also evaluated and Neural Network and support vector machine had the same sensitivity value of 0.9941 whereas the other models managed a recall score ranging from 0.9501 to 0.9853. This research concluded also that decision tree had the best performing metrics as having the highest level of accuracy and minimum degree of error using both classification and regression method.

Bekele (2022) investigated the prediction of low birth weight in Ethiopia with machine learning algorithms. The study implemented predictive LBW models based on the data obtained from the Ethiopia Demographic and Health Survey 2016. This study was employed to compare and identify the best-suited classifier for predictive classification among Logistic Regression, Decision Tree, Naive Bayes, K-Nearest Neighbor, Random Forest (RF), Support Vector Machine, Gradient Boosting, and Extreme Gradient Boosting. The study revealed that Random Forest was the best classifier and predicts LBW with 91.60 percent accuracy, 91.60 percent Recall, 96.80 percent ROC-AUC, 91.60 percent F1 Score, 1.05 percent Hamming loss, and 81.86 percent Jaccard score. The study therefore concluded that Random Forest and gradient boosting algorithms are effective in predicting birth weight, with potential application in prenatal care predicted the occurrence of LBW more accurately and effectively than other classifiers in Ethiopia Demographic Health Survey. However, the research has a limitation of generalizability to other populations and also did not evaluate model interpretability and explainability. In this study the results obtained from the birth weight prediction model demonstrate that the Random Forest algorithm outperforms other machine learning models, exhibiting the lowest Root Mean Squared Error (RMSE), highest R-squared, and lowest Mean Absolute Error (MAE). This suggests that Random Forest is a robust model for predicting birth weights. Additionally, the correlation matrix provides valuable insights into the relationships between different features, revealing, for instance, the impact of smoking status on birth status and the positive correlation between maternal age and gestation.
Sawe (2022) investigated machine learning prediction of low birth weight in Kenya using maternal risk factors, the study utilized secondary data from the 2014 Kenya Demographic Health Survey where the variables were extracted from the births recode file. The study population included mothers between the age of 15 to 49 years. The machine learning algorithms employed were logistic regression, decision trees, random forest, support vector machines, gradient boosting and extreme gradient boosting. Using performance evaluation metrics namely; accuracy, precision, recall, F1 score, and ROC-AUC, the random forest model was found out to be the most robust with 0.956679 accuracy, 0.956831 precision, 0.956679 recall an F1 score of 0.95666 and an AUC of 0.988. In addition, variable importance was performed using the random forest approach to ascertain the maternal risk factors that are the most important to predict low birth weight. It was found out that mother's weight was the most important variable for predicting low birth weight. The other important variables found were; mothers height, mother's age and the number of antenatal visits attended by the mother during pregnancy. Machine learning techniques are increasingly being used to provide information to guide health policy. The research was limited to a single data set and population and did not evaluate the impact of explanation on clinical decision-making. This study used a larger data set that was able to capture the full range of variability and avoided over fitting. This study emphases on the importance of early identification, target intervention and resource optimization of machine learning models in addressing the challenges of low birth weight in Nigeria. By implementing these recommendation healthcare system and policymakers can take proactive steps towards improving the country's maternal health outcomes which was not properly evaluated by Sharon.

Senthilkumar et al (2015) used data mining techniques in predicting infants at risk of low birth weight and its factors. Compared to other classification methods, the classification tree produced the best results; AUC of 93.80 percent, the prediction accuracy of 89.95 percent, specificity of 72.88 percent, an F-value of 93.04 percent and precision of 88.81 percent.
Indonesia has not been left behind either. Two studies in Indonesia conducted by Eliyati et al (2019) utilized the Indonesia Demographic and Health Survey data to predict and classify low birth weight using machine learning techniques. The first study compared binary logistic regression and random forest in prediction and classification. Random forest proved to be the best model in both tasks. Using the same data with the same variables, another study used Support Vector Machines (SVM) for the classification of LBW. The results revealed that SVM with four kernel functions (hyperbolic tangent, polynomial, linear and radial) were fit for binary classification of LBW. Moreover, their average predictive performance was satisfactory since the predictive error was below 10 percent. This research concluded that the Support Vector Machines based on linear kernel competed well with the binary logistic regression for classification Low Birth Weight data in Indonesia.

MATERIALS AND METHODS
This section describes the fundamental procedures and techniques that were applied to meet the objectives of the research. This consists of the data that was used, pre-processing techniques, handling data imbalance, machine learning models training, hyperparameter tuning and evaluation of performance metrics, variable importance and the software tools for the whole process. The first step began with the dataset which contains the variables that were used in the research for prediction. The data was then subjected to a pre-processing process where exploratory analysis and cleaning of the data to prepare it for analysis is performed. After preprocessing, the data was balanced using the Synthetic Minority Oversampling Technique (SMOTE) in order to avoid the models from skewing results towards the majority class. Before the actual modelling, the data was subjected to cross validation in order to estimate how the models would perform before applying hyper-parameter tuning. Model building was then performed using the classification and regression machine learning algorithms because the problem under study entails a target variable which is categorical and also continuous in nature. After building the models, they were subjected to performance evaluation using evaluation metrics including accuracy ,Kappa, Sensitivity, Specificity and ROC-AUC. A comparative analysis of the models was then done based on the evaluation metrics to get the most robust model. Variable importance was then performed to identify the independent variables that contributed most to the performance of the most robust model.
The study analyzed the secondary data collected from the

Science World Journal Vol. 19(No 4) 2024 https://dx.doi.org/10.4314/swj.v19i4.9
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

University of Abuja Teaching Hospital as a hospital-based cross-sectional study undertaken in the labour room and postnatal ward of Obstetrics and Gynecology. The data covered the period of ten years from 2010 to 2019 to predict the birth weight of new born babies in FCT Abuja. The study variables which were extracted from report data include both the dependent and independent variables. The dependent variable is birth weight which is a binary variable consisting of two categories i.e, low birth weight and those without low birth weight. The independent variables, are the maternal risk factors for low birth weight, are place of residence, mother's education level, age of the mother, smoking status, mother's obstetric history, gestation period, Birth Status, Birth order, wealth index, number of antenatal visits during pregnancy. The dynamics among the variables were analyzed using seven distinct machine learning algorithms for both regression and classification methods. The analysis encompasses a detailed examination of each model's performance metrics, revealing insights into their predictive capabilities. Furthermore, the study explores the correlations between key features and birth weight, shedding light on the factors influencing the outcomes. The accuracy measures of different classification algorithms with their performance metrics are depicted in the Table 4.2.The accuracy of a model on a given test set is the percentage of test set that are correctly classified by the classifier Measures are defined Jiawei et al (2010), Hain et al (2013) and Senthikumar et al (2013) .

Supervised machine learning is a type of machine learning that learns the relationship between input and output. The inputs are known as features or 'X variables'(The independent variables, are the maternal risk factors for low birth weight, are place of residence, mother's education level, age of the mother, smoking status, mother's obstetric history, gestation period, Birth Status, Birth order, wealth index, number of antenatal visits) and output is generally referred to as the target or 'y variable'(Birth weight). The type of data which contains both the features and the target is known as labeled data. Supervised machine learning learns patterns and relationships between input and output data. It is defined by its use of labeled data. A labeled data is a dataset that contains a lot of examples of Features and Target. Supervised learning uses algorithms that learn the relationship of Features and Target from the dataset. This process is referred to as Training or Fitting. There are two types of supervised learning algorithms the Classification and the Regression Algorithms

Classification is a type of supervised machine learning where algorithms learn from the data to predict an outcome or event in the future. Classification algorithms are used for predicting discrete outcomes, if the outcome can take two possible values such as True or False, Default or No Default, Yes or No, it is known as Binary Classification. When the outcome contains more than two possible values, it is known as Multiclass Classification. There are many machine learning algorithms that can be used for classification tasks. Some of them are: Logistic Regression,DecisionTree Classifier,K Nearest Neighbor Classifier,Random Forest Classifier,Neural Networks.

Regression is a type of supervised machine learning where algorithms learn from the data to predict continuous values such as sales, salary, weight, or temperature.There are many machine learning algorithms that can be used for regression tasks. Some of them are:Linear Regression,Decision Tree Regressor,K Nearest Neighbor Regressor,Random Forest Regressor, andNeural Networks.The main difference between supervised and unsupervised machine learning and the justification for choosing supervised learning is that supervised learning uses labeled data. Labeled Data is a data that contains both the Features (X variables) and the Target (y variable). When using supervised learning, the algorithm iteratively learns to predict the target variable given the features and modifies for the proper response in order to "learn" from the training dataset. This process is referred to as Training or Fitting. Supervised learning models typically produce more accurate results than unsupervised learning but they do require human interaction at the outset in order to correctly identify the data. If the labels in the dataset are not correctly identified, supervised algorithms will learn the wrong details.

There are two prominent use-cases for supervised learning i.e. Classification and Regression. In both the tasks a supervised algorithm learns from the training data to predict something. If the predicted variable is discrete such as "Yes" or "No", 1 or 0,, or "Low birth" or "Normal birth" in this study", then a classification algorithm is required. If the predicted variable is continuous like gestation period,Age of mothers, Education level, birth status etc., then the Regression algorithm is required.

Support Vector Machines

Support vector machines (SVM) are introduced by Cortes and effective method for binary classification, regression or ranking function and it is based on statistical earning. It is very popular used by the researcher in health care for classification due to many attractive features, handling complex non-linear data points. Its accuracy is high and less prone to over more fitting than other well-known classifier. Neelamegam et al (2013), Hetal et al (2012), Santhosh (2010) Hain et al (2013). It is a good classifier, does not require a priori knowledge, even the input space is very high. Neelamegam et al (2013).

Model specification for support vector machine using binary classification

For a binary classification problem, we are given a training dataset

$$\{(x_i, y_i)\}_{i=1}^{N}$$

where $x_i \epsilon \mathbb{R}^d$ represents the feature vector (socioeconomic factors) and $y_i \ \epsilon \{-1, +1\}$ represents the class label (birth weight category: low or normal). The goal of SVM is to find a hyperplane that best separates the two classes.

The decision function for SVM is defined as: $f(x) = W^T X + b$

Here, W is the weight vector, and b is the bias term.

Soft Margin (Handling Misclassifications)

In practice, data may not be perfectly separable. To handle this, we introduce slack variables $\xi_i \geq 0$ for each data point and solve the following optimization problem:

$$min_{w,b,\xi} \frac{1}{2} ||W||^2 + C \sum_{i=1}^{N} \xi_i$$

subject to the constraints: $y_i(W^T X_i + b) \geq 1 - \xi_i \quad \forall_i$

$$\xi_i \geq 0 \quad \forall_i$$

Here, $C > 0$ is a regularization parameter that controls the trade-off between maximizing the margin and minimizing the classification error.

Non-Linear SVM (Kernel Trick)

For non-linear decision boundaries, SVM can be extended using kernel functions $K(x_i, x_j)$ that implicitly map the input features to a higher-dimensional space. The optimization problem becomes:

Science World Journal Vol. 19(No 4) 2024
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

https://dx.doi.org/10.4314/swj.v19i4.9

$$min_\alpha \frac{1}{2}\sum_{i,j=1}^{N} \alpha_i\alpha_j\, y_i y_j K(x_i, x_j) - \sum_{i=1}^{N} \alpha_i$$

Subject to:

$$\sum_{i=1}^{N} \alpha_i y_i = 0$$

$$0 \leq \alpha_i \leq C \quad \forall_i$$

Here, $\alpha i \alpha i$ are the Lagrange multipliers.

By applying SVMs to socioeconomic data, we can effectively classify and predict Birth Weight categories in babies, helping to identify key factors that influence infant health and guide interventions to improve outcomes.

Logistic Regression

Logistic regression is a special case of generalized linear modelling, also called a logistic model or logit model and is extensively used for binary classification method in the medical, social sciences, marketing applications. It is used based on the assumption when the outcome of a situation is not linearly associated to the explanatory variables. It allows probabilistic interpretation; easily we can update the model for the new data, unlike decision trees or SVM and ease of interpretation. It has some drawbacks, not suitable for high-dimensional problems, it is slower than SVM, non-linearities and identifying interaction is difficult. The dependent variable is restricted to discrete number. It accepts large number of explanatory variables; in many situations it is not. The researcher should decide whether to use logistic regression for the classification if the data set is a large size. It is applied to the studies that using between subject design. It may be suitable in the fields of medicine and psychology, in fact; it is not a choice always. Raghavendra et al (2011)

Logistic Regression Equation

The odd is the ratio of something occurring to something not occurring. it is different from probability as the probability is the ratio of something occurring to everything that could possibly occur. so odd will be:

$$\frac{p(X)}{1-P(x)} = e^{z}$$

Applying natural log on odd. then log odd will be;

$$\text{Log}\left[\frac{p(X)}{1-P(x)}\right] = z$$

$$\text{Log}\left[\frac{p(X)}{1-P(x)}\right] = w.X + b$$

$$\frac{p(X)}{1-P(x)} = e^{w.X+b}$$

..... Exponentiate both sides

$$P(x) = e^{w.X+b} \,.\, (1 - P(x))$$

$$P(x) = e^{w.X+b} - e^{w.X+b}\, P(x)$$

$$P(x) + e^{w.X+b} \,.\, P(x) = e^{w.X+b}$$

$$P(x)\,(1 + e^{w.X+b}) = e^{w.X+b}$$

$$P(x) = \frac{e^{w.X+b}}{1+e^{w.X+b}}$$

then the final logistic regression equation will be:

$$P\ (X;b,w) = \frac{e^{w.X+b}}{1+e^{w.X+b}} = \frac{1}{1+e^{w.X+b}}$$

Neural Network

Neural Networks are a complex non-linear modelling method and able to learn complex relationship between dependent and independent variables without any external assistance based on a model of the neural architecture of the human brain. Neural Network is a successful technique applied for the real-world application like accounting and auditing, finance, management, decision making, marketing, production, biology, psychology, handwritten character recognition, pathology, statistics, mathematics, computer science, medical research and many more. Neural Network is a popular data mining tool because of its predictive power even in the complicated domain compared with statistical techniques. It will handle both categorical and continuous data types and tolerate noisy data. It supports parallelization techniques, which will speed up the computation process. Identifying patterns is very difficult and requires long learning time if the input features are large. Neural etworks are also called as "black boxes" due to its poor transparency to explain the process of neural networks built. Identifying the required number of parameters for modelling neural network is trial and error design like, network topology or structure, number of hidden layers, number of units in each hidden layer an in output layer Hetal et al (2013), Rabindra et al (2014),Moawia et al (2010),Gireesh et al (2012),Mozziyar et al (2013) . Neural network performance can be highly automated, minimizing human involvement. It is very flexible with incomplete, noisy and missing data. It does not make any prior assumption about the distribution of the data or the form of interaction between factors. It can be easily updated with new data. The output of the Neural Network algorithm does not produce an explicit set of rules and it is lacking in classical statistical properties (confidence interval and testing of hypothesis) Rabindra et al (2014).

Model specification for Neural Networks

A basic feedforward neural network consists of an input layer, one or more hidden layers, and an output layer. Each layer comprises units (neurons) that are connected to the units in the previous and next layers.

Neuron Activation

The output of a neuron is computed as:

$$a_j = \phi\left(\sum_{i=1}^{n} w_{ji}\, x_i + b_j\right)$$

where:

- x_i are the input features.
- w_{ji} are the weights associated with the connections.
- b_j is the bias term.
- ϕ is the activation function (e.g., sigmoid, tanh, ReLU).

Activation Functions

Common activation functions include:

- Sigmoid: $\phi(z) = \frac{1}{1+e^{-z}}$
- $Tanh: \phi(z) = tanh\,(z)$
- $ReLU: \phi(z) = max\,(0, z)$

Forward Propagation

The forward propagation step computes the outputs of the neurons

https://dx.doi.org/10.4314/swj.v19i4.9

layer by layer from the input layer to the output layer. For a network with LL layers, the output of layer ll is given by:

$$a^{(l)} = \phi(W^{(l)}a^{(l-1)} + b^{(l)})$$

Cost Function

For classification problems, a common cost function is the cross-entropy loss:

$$\boldsymbol{J}(\boldsymbol{w}, \boldsymbol{b}) = -\frac{1}{m}\sum_{i=1}^{m}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

where $\hat{y}_i$ is the predicted output for the i-th training example, and y_i is the true label.

Backpropagation

Backpropagation is used to compute the gradients of the cost function with respect to the weights and biases, which are then used to update the parameters using gradient descent. The gradients are computed as follows:

1. Compute the error at the output layer:
$$\delta^{(L)} = \boldsymbol{a}^{(L)} - y$$
2. Compute the error at layer l:
$$\delta^{(L)} = \left(W^{(\boldsymbol{l}+\boldsymbol{1})}\right)^T \delta^{(l+1)} \odot \phi'\left(z^{(l)}\right)$$

where $\odot$ denotes element-wise multiplication, and ϕ' is the derivative of the activation function.

3. Compute the gradients:
$$\frac{\partial J}{\partial W^{(l)}} = \frac{1}{m}\delta^{(l)}\left(a^{(l-1)}\right)^T$$
$$\frac{\partial J}{\partial b^{(l)}} = \frac{1}{m}\sum_{i=1}^{m}\delta^{(l)}$$

By applying these mathematical principles, Neural Networks can be used to analyze the socioeconomic factors affecting BMI in babies, providing insights and predictions based on complex relationships between the input features and Birth weight outcomes.

Random Forest

Random forest is an ensemble classifier, like decision trees, can be used to solve classification and regression problems. It uses the concept of generating multiple random trees with, bootstrap of training dataset, bagging on samples, voting scheme and the features are randomly selected in each decision split, which improves the predictive power and results in higher efficiency. It achieves better results most of the time compared to decision trees. Selection of a random subset of features is an example of the random subspace method. It was founded in 2001 and used in the different number of applications, includes medical research, image processing, etc. Hitesh et al (2013), Jehad et al (2012), Ned (2010). The advantages of random forest are; it does not depend on the data, appropriate for high dimensional data modelling, overcoming the problem of over fitting, eliminates prune the trees. It will generate the most important variable used for classification. It runs efficiently on large databases produce high prediction accuracy. It is good with dealing missing values, outlier and maintain accuracy when a large proportion of the data are missing. The model interprets ability and prediction accuracy provided by Random Forest is very unique among popular machine learning methods. It also supports a method for detecting interaction between variables. The main disadvantage is observed to over fit for some datasets with noisy classification/regression tasks.

Model specification for Random Forest

Decision Tree

A Decision Tree is the basic building block of a Random Forest. For a single tree, the goal is to partition the feature space into distinct regions and assign a class label to each region for classification tasks. The tree is constructed by recursively splitting the data based on feature values to maximize a certain criterion, such as information gain or Gini impurity.

For a given dataset $\{(x_i, y_i)\}_{i=1}^{N}$ where x_i are the input features and y_i are the class labels, the decision tree algorithm selects the best feature j and threshold t that maximize the information gain:

$$\boldsymbol{IG}(\boldsymbol{D}, \boldsymbol{j}, \boldsymbol{t}) = \boldsymbol{H}(\boldsymbol{D}) - \left(\frac{|\boldsymbol{D}_l|}{|\boldsymbol{D}|}\boldsymbol{H}(\boldsymbol{D}_l) + \frac{|\boldsymbol{D}_R|}{|\boldsymbol{D}|}\boldsymbol{H}(\boldsymbol{D}_R)\right)$$

where:

- D is the dataset.
- $\boldsymbol{D}_l$ and $\boldsymbol{D}_R$ are the left and right subsets created by splitting on feature j at threshold t.
- $H(D)$ is the entropy or Gini impurity of dataset D.

Random Forest

A Random Forest constructs multiple decision trees and aggregates their results. Each tree is trained on a bootstrap sample of the training data, and at each split, a random subset of features is considered.

1. **Bootstrap Sampling**: For each tree, create a bootstrap sample D^* by randomly sampling with replacement from the training dataset D.
2. **Feature Selection**: At each node in the tree, randomly select a subset of features F' from the total set of features F. Choose the best split from this subset.
3. **Tree Growing**: Grow each tree to its maximum depth without pruning.
4. **Aggregation**: For classification, aggregate the results using majority voting. For regression, use the average of the predictions.

Prediction

For classification, the final prediction is made by aggregating the votes from all the individual trees:

$$\hat{y} = mode\left(\{T_b(X)\}_{b=1}^{B}\right)$$

For regression, the final prediction is the average of the predictions from all trees:

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B}T_b(X)$$

where T_b is the $b-th$ tree in the forest, B is the total number of trees, and X is the input feature vector.

Decision Tree

A decision tree is a classifier used in statistics, data mining and machine learning for modeling classification and prediction. In decision analysis, a decision tree can be used to visually and explicitly represent decisions and decision making. There are two varieties in decision tree used data mining are classification tree or regression tree. Decision tree inducers are algorithms that automatically construct a decision tree from a given dataset. Typically the goal is to find the optimal decision tree by minimizing the generalization error, Demsar (2010) Rohach et al(2013), Parashar et al(2012),Ozer et al (2008) Its representation is easy to understand and interpret by non-professional user. It is capable of handling; both nominal and numerical input, requires little data preparation, data sets that may have errors and missing values, efficiency in processing with large datasets. It does not require any domain knowledge or parameter setting, and therefore appropriate for exploratory knowledge discovery. Reliability of the model can be validated using statistical tests. It has some disadvantages such

https://dx.doi.org/10.4314/swj.v19i4.9

as: most of the algorithms support only discrete values as the target attribute. Perform well if a few highly relevant attributes exist, but less if many complex interactions are present. Classification tree analysis when the predicted outcome is the class to which the data belongs and used only for classifying discrete category (the class).

Model specification for Decision Trees

Splitting Criteria

A Decision Tree works by recursively splitting the data into subsets based on an attribute value test. This process is done to create branches until a certain criterion is met (e.g., all data points in a node belong to the same class). The most common criteria for splitting are Information Gain and Gini Impurity.

Information Gain

Information Gain measures the reduction in entropy achieved by partitioning the data based on an attribute. For a given dataset D, the entropy $H(D)$ is defined as:

$$H(D) = -\sum_{k=1}^{K} p_k \, log_2(p_k)$$

where p_k is the proportion of samples belonging to class k.

The Information Gain $IG(D, A)$ of an attribute A is defined as:

$$IG(D, A) = H(D) - \sum_{v \epsilon Values(A)} \frac{|D_v|}{|D|} H(D_v)$$

where D_v is the subset of D where attribute A has value v.

Gini Impurity

Gini Impurity is another criterion used for splitting. For a dataset D, the Gini Impurity $G(D)$ is defined as: $\mathrm{G(D)} = 1 - \sum_{k=1}^{K} p_k^2$

The reduction in Gini Impurity when splitting on an attribute A is:

$$\Delta G(D, A) = G(D) - \sum_{v \epsilon Values(A)} \frac{|D_v|}{|D|} G(D_v)$$

K - Nearest Neighbors Algorithms

K-Nearest Neighbor is one of the simplest Machine Learning algorithms based on Supervised Learning technique. It is an algorithm assumes the similarity between the new case/data and available cases and put the new case into the category that is most similar to the available categories. K-NN algorithm stores all the available data and classifies a new data point based on the similarity. This means when new data appears then it can be easily classified into a well suite category by using K- NN algorithm. K-NN algorithm can be used for Regression as well as for Classification but mostly it is used for the Classification problems. K-NN is a **non-parametric algorithm**, which means it does not make any assumption on underlying data. It is also called a **lazy learner algorithm** because it does not learn from the training set immediately instead it stores the dataset and at the time of classification, it performs an action on the dataset. KNN algorithm at the training phase just stores the dataset and when it gets new data, then it classifies that data into a category that is much similar to the new data.

Confusion Matrix

The confusion matrix is a contingency table which compares actual class to the model predictions. It is divided into true positive, false positive, true negative and false negative values:

True positive (TP): this is a case where actual positive values are predicted as positive. For instance, the number of cases correctly classified as birth weight.

False positive (FP): this is a case where actual negative values are predicted as positive. For instance, the number of cases falsely classified as birth weight.

True negative (TN): this is a case where actual negative values are predicted as negative. For instance, the number of cases correctly classified as not birth weight.

False negative (FN): this is a case where actual positive values are predicted as negative. For instance, the number of cases falsely classified as not birth weight.

Sensitivity (True Positive Rate): The proportion of actual positive cases that are correctly identified as positive. Ist measure how well the model detects true positives.

Sensitivity = TP/ (TP+FN)

Where TP = True Positives, FN = False Negatives

Specificity (True Negative Rate): The proportion of actual negative cases that are correctly identified as negative. It measures how well the model detects true negatives.

Specificity = T$A = \pi r^2$N/ (TN+FP)

Where TN = True Negatives, FP = False Positives

Kappa (Cohen's kappa): A measure of the agreement between the predicted and actual classifications, adjusted for chance. It takes into account the possibility of agreement occurring by chance.

Kappa = (p_o-p_e) / (1-p_e)

Where p_o is the observed agreement and p_e is the expected agreement by chance.

Positive Predicted value (PPV): The proportion of positive prediction that are actually true positive . It measures the accuracy of positive predictions.

PPV = TP/ (TP + FP)

Negative Predicted value (NPV): The proportion of negative predictions that are actually negatives. It measures the accuracy of negative predictions.

NPV = TN / (TN+FN)

Where;

- TP = True Positive
- TN = True Negative
- FP = False Positive
- FN = False Negative

Interpretation:

1) High sensitivity and specificity indicate good performance.
2) High kappa value (close to 1) indicates strong agreement between predicted and value classification
3) High PPV (Positive Predicted Value) and NPV (Negative Predicted Value) Indicates accurate prediction

RESULTS AND DISCUSSION

The main objective of this research was to build a machine learning model for predicting birth weight. The first of these three specific objective was to comprehensively access each algorithm's ability to accurately predict birth weight and identify the best performing model based on a combination of these metrics. To achieve these objectives, data was collected from Obstetric department of Gwagwalada teaching hospital Abuja. The data was subjected to pre-processing in order to clean it and prepare for the modelling tasks. Six machine learning algorithms were trained and tested namely; Random Forest, Linear regression, Neural Network, support Vector Machine, Gradient Boosting, Decision Tree and K-Nearest Neighbors and logistic regression. After training and testing them, they were evaluated based on performance metrics specifically used for regression and classification problems as presented in tables 4.1 and 4.2 below:

Science World Journal Vol. 19(No 4) 2024 https://dx.doi.org/10.4314/swj.v19i4.9
www.scienceworldjournal.org
ISSN: 1597-6343 (Online), ISSN: 2756-391X (Print)
Published by Faculty of Science, Kaduna State University

Table 1: Predictive performance of various Regression methods.
The table below shows the summary of the regression method used and their performance metric

Algorithm	RMSE	R-squared	MAE	MSLE	Explained Variance	Max Error
Random Forest	0.410	0.804	0.148	0.00878	0.804	2.403
Linear Regression	0.578	0.612	0.390	0.01826	0.612	2.032
Neural Network	0.602	0.579	0.402	0.01903	0.588	2.293
Support Vector Machine	0.682	0.458	0.390	0.02267	0.504	2.255
Gradient Boosting	0.473	0.740	0.282	0.01136	0.740	2.162
Decision Tree	0.519	0.686	0.153	0.01412	0.686	2.700
K-Nearest Neighbors	0.570	0.622	0.303	0.01749	0.623	2.700

Table 2: Predictive performance of various classification methods
The table below shows the summary of the classification used and the performance metrics.

Algorithms	**Performance Metrics**			
	Accuracy	**Kappa**	**Sensitivity**	**Specificity**
Logistic Regression	0.9806	0.9592	0.9741	0.9908
Support Vector Machine	0.9806	0.9592	0.9941	0.9600
Decision Tree	0.9823	0.9631	0.9853	0.9778
Neural Network	0.9806	0.9592	0.9941	0.9600
Gradient Boosting Machine	0.9806	0.9595	0.9824	0.9778
K-Nearest Neighbors	0.9452	0.8859	0.9501	0.9378

The metrics used for regression method were Root Mean Squared Error (RMSE), R -Squared, Mean Absolute Error (MAE), Mean Square Logarithmic Error (MSLE), Explained Variance and Max Error. Among these algorithms, Random Forest exhibits the lowest Root Mean Squared (RMSE), highest R -Squared, and lowest Mean Absolute Error (MAE) indicating superior overall performance. It achieves an RMSE of (0.41), an R -Squared (0.80), and an MAE of (0.15). Gradient boosting also performs well the competitive metrics, demonstrating the second - lowest RMSE and the highest Explained Variance. In Addition, Support Vector Machine displays the highest RMSE and lowest R -Squared suggesting less effective predictive capability for birth weight. The decision-making process for selecting the best model involves considering multiple metrics, Where Random Forest emerges as the most favorable algorithms for regression method.
The Random Forest model stands out as the best performer in predicting birth weight, as it attains the lowest Root Mean Squared Error (RMSE) of 0.41, indicating accurate predictions with minimal deviation from the actual values. The high R-squared value of 0.80 reflects the model's ability to explain a significant portion of the variance in the target variable. The Mean Absolute Error (MAE) of 0.15 signifies the average magnitude of errors in the predictions, further supporting the model's accuracy. Additionally, the low Mean Squared Logarithmic Error (MSLE) of 0.0087 indicates the model's proficiency in handling logarithmic differences between predicted and actual values. The Explained Variance of 0.80 underscores the model's capacity to capture the variability in the data. The maximum error of 2.40 suggests that, in the worst case, the model's predictions deviate by approximately 2.40 units from the true birth weight values. Therefore, these comprehensive metrics affirm the Random Forest as the most reliable model for birth weight prediction in this context. The forecasting results using the Random Forest model demonstrated its capability to provide accurate predictions for birth weights using regression method.
Furthermore, the ROC-AUC was used to test the classification ability of the models to differentiate between the low-birth-weight cases and the cases without low birth weight. In terms of accuracy, the best machine learning model was the Decision tree with an accuracy of 0.9823. The other five models produced an accuracy that ranged between 0.9806 to 0.9822. Based on the kappa, decision tree again emerged to be the best with a value of 09631. The rest of the models had a kappa that ranged from 0.8859 to 0.9592. Sensitivity was also evaluated and Neural Network and support vector machine had the same sensitivity value of 0.9941 whereas the other models managed a recall score ranging from 0.9501 to 0.9853. Moreover, Specificity was also examined. Logistic Regression model had the best specificity value of 0.9908. The rest of the models ranged from 0.9378 to 0.9778. Furthermore, the ROC curves of all the tested models were plotted and the area under the curved evaluated. The decision tree had the highest area under the curve of 0.9896. The AUC of the other models ranged from 0.9440 to 0.9816. Therefore, from these results based on the performance metrics and ROC-AUC, decision tree emerged to be the most robust model for classification method. Previous study conducted by Sharon J. Sawe used machine learning models to perform prediction with six machine learning algorithms were trained and tested namely; the random forest, decision tree, gradient boosting, XGBoost, SVM and logistic regression. After training and testing them, they were evaluated based on performance metrics specifically used for classification problems. The metrics used were accuracy, precision score, recall score and F1 score. He found out that random forest is the best model for predicting low birth weight since it had the highest accuracy and effectiveness in terms of recall and precision. Moreover, random forest yielded the best AUC therefore the best classification model. It was also important to identify the variables that contributed most to the robustness of the model. This technique known as feature importance was performed using the random forest technique. It was ascertained that mother's weight, height, age and number of antenatal visits attended during pregnancy are the most important variables that contributed most to the model's accuracy. Furthermore, the ROC-AUC was used to test the classification

ability of the models to differentiate between the low-birth-weight cases and the cases without low birth weight. In terms of accuracy, the best machine learning model was the random forest with an accuracy of 0.956679. The other five models produced an accuracy that ranged between 0.666667 to 0.939832. Based on the precision score, random forest again emerged to be the best with a value of 0.956831.

In addition, variable importance was examined. This specific objective was geared to ascertain the variables which are the most important to be considered when predicting birth weight. It was observed that gestation period emerges as the most important variable for predicting birth weight followed by the of the variables like Mother's Age, Birth status, educational status of the mother and Ante natal visit during the pregnancy. It was found out that out of the Nine independent variables used the most important ones for predicting low birth weight were; Gestation period, mother's Age, birth status, and the number of antenatal care visits during pregnancy. Muula et al (2011) conducted a study on determined the significant risk factors for LBW and predicted LBW babies using the critical risk factors with ML in Bangladesh. They implemented the Logistic Regression based method to determine the parity and maternal education that are associated with low birth weight. In their study, they identified poverty and the absence of education (no education) were the most prominent risk factors associated with the prevalence of LBW babies in Bangladesh. LBW is more common in twin babies than in single birth babies. Mostly, second twin babies are at a higher risk of being LBW than single babies.

Conclusion and Recommendations

This study provides a comprehensive discussion and interpretation of the results obtained from the birth weight prediction model using regression and classification algorithms. The aim is to analyze the performance of various machine learning algorithms, examine the correlation matrix of features, and draw meaningful insights from the forecasting results. This chapter synthesizes the findings to derive conclusions, offer recommendations, discuss contributions to knowledge, and suggest potential avenues for further research.

The results obtained from the birth weight prediction model demonstrate that the Random Forest algorithm outperforms other machine learning models, exhibiting the lowest Root Mean Squared Error (RMSE), highest R-squared, and lowest Mean Absolute Error (MAE). This suggests that Random Forest is a robust model for predicting birth weights. Additionally, the correlation matrix provides valuable insights into the relationships between different features, revealing, for instance, the impact of smoking status on birth status and the positive correlation between maternal age and gestation.

Alternatively, using classification method our approach was meticulous, involving a comprehensive evaluation of numerous machine learning models to discern their effectiveness in addressing this critical healthcare challenge. Through rigorous scrutiny of performance metrics such as accuracy, kappa, sensitivity, and specificity, we meticulously assessed the strengths and weaknesses of each model. This exhaustive analysis culminated in the identification of the most robust technique for low-birth-weight prediction, offering a promising avenue for improving maternal and child health outcomes in Nigeria. Furthermore, our study delved deep into the underlying factors contributing to low-birth-weight occurrences, meticulously identifying the pivotal variables driving this phenomenon. By unraveling these key determinants, we provided valuable insights into the complex interplay of social-demographic, environmental, and healthcare-related factors influencing maternal and infant health in Nigeria. This comprehensive understanding serves as a foundation for developing targeted interventions and policy strategies aimed at reducing the prevalence of low birth weight and improving overall maternal and child health outcomes in the country.

In conclusion, the birth weight prediction model, particularly the Random Forest algorithm, proves effective in providing accurate and reliable predictions for regression while Decision tree with highest accuracy kappa was identified as the best algorithm for classification methods. The findings highlight the importance of considering various machine-learning algorithms and feature relationships when predicting birth weights. Additionally, the correlation analysis sheds light on potential influencing factors on birth outcomes. These insights contribute to enhancing our understanding of birth weight prediction, offering practical implications for healthcare professionals and policymakers.

1. Based on the results, it is recommended to consider the following; Consider the adoption of the Random Forest algorithm as the primary model for birth weight prediction due to its superior performance in terms of lower RMSE, higher R-squared, and lower MAE compared to other machine learning models.
2. Consider the adoption of the Random Forest algorithm as the primary model for birth weight prediction due to its superior performance in terms of lower RMSE, higher R-squared, and lower MAE compared to other machine learning models.

REFERENCES

Anggraini Dewi, Mali Abdollahian, Kaye Marion (2018) Foetal weight prediction models at a given gestational age in the absence of ultrasound facilities:Application in Indonesia. *BMC Pregnancy Childbirth*. 18(1):1-12. doi:10.1186/s12884-018-2047-z

Agbor VN, Ditah C, Tochie JN, Njim T. (2018) Low birthweight in rural Cameroon: An analysis of a cut-off value. *BMC Pregnancy Childbirth*. 2018;18(1):1-5. doi:10.1186/s12884-018- 1663-y

Amene Ranjbar,Farideh Montazeri, Fatemeh Darsareh, Mohammadsadegh V. Farashah,Nasibeh Roozbeh(2023) Machine learning - based approach for predicting low birth weight:*BMC pregnancy and Childbirth* .2023:23,803. https://doi.org/10.1186/s12884-023

Bankole Oladeinde H, Oladeinde OB, Omoregie R, Onifade AA (2015): Prevalence & determinants of low birth weight: the situation in a traditional birth home in Benin City, Nigeria. *Afri Heal Sci.*2015;15 (4):1123-1132.

Bayingana C, Muvunyi CM, Africa CWJ. (2010) Risk factors of preterm delivery of low birth weight (plbw) in an African population. *J Clin Med Res.* 2(7):114-118. https://www.academicjournals.org/JCMR.

Blencowe H, Krasevec J, de Onis M. (2019) National, regional, and worldwide estimates of low birthweight in 2015, with trends from 2000: a systematic analysis. *Lancet Glob Heal*. 2019;7(7):e849-e860. doi:10.1016/S2214-109X (18)30565-5

Cho, H., Lee, E.H., Lee, KS. *et al (2022).* Machine learning-based risk factor analysis of adverse birth outcomes in very low birth weight infants. *Sci Rep* **12**,12119. https://doi.org/10.1038/s41598-022-16234-y

Daynial, E.Ballot, Tobias F Chirwa & Peter, A. Cooper. (2010) Determinant of survival in very low birth weight neonates in a public sector hospital in Johannesbury; 30; 1471-2431.

Dönmez P(2013)Introduction to Machine Learning, 2nd ed., by Ethem Alpaydın. Cambridge, MA: The MIT Press2010. ISBN: 978-0-262-01243-0. $54/£ 39.95 + 584 pages. *Nat Lang Eng*. 19(2):285-288. doi:10.1017/s1351324912000290

Endalamaw A, Engeda EH, Ekubagewargies DT, Belay GM, Tefera MA. (2018) Birth weight and its associated factors in Ethiopia: A systematic review and meta-analysis. *Ital J Pediatr*. 44(1):1-12. doi:10.1186/s13052-018-586-6

Eliyati N, Faruk A, Kresnawati ES, Arifieni I.(2019) Support vector machines for classification of low birth weight in Indonesia Support vector machines for classification of low birth weight in Indonesia. *IOP J Phys*. doi:10.1088/17426596/1282/1/012010

Faruk A, Cahyono ES, Eliyati N, Arifieni I.(2018) Prediction and classification of low birth weight data using machine learning techniques. *Indones J Sci Technol*.3(1):18-28. doi:10.17509/ijost. v3i1.10799

Feresu SA, Harlow SD, Woelk GB. (2015) Risk factors for low birthweight in Zimbabwean women:A secondary data analysis.*PLoS One*. 10(6):1-17. doi: 10.1371/journal.pone.0129705

Gao H, Wu C, Huang D, Zha D, Zhou C.(2021) Prediction of fetal weight based on back propagation neural network optimized by genetic algorithm.18 (March):44024410.

M. W. Moreira, J. J. Rodrigues, N. Kumar et al. (2017) "Performance assessment of decision tree-based predictive classifiers for risk pregnancy care", *IEEE Global Communications Conference IEEE GLOBECOM,* pp. 1-5, Dec. 4–8.

Senthilkumar, D; Paulraj S.(2015) Prediction of Low Birth Weight Infants and Its Risk Factors Using Data Mining Techniques. *Proc 2015 Int Conf Ind Eng OperManag Dubai, United Arab Emirates*. 2015;3: 186-194.

Olu, S. Dunart, Maznah, D. Nazar, Azahar, Oche, S.Oche & Norlaili, Abudaziz. (2005) The birth weight of Nigerian babies, J Trop Pediatr ,41: 140 – 141 Risk factors for low birth weight in Nigeria; Evidence from 2013 Demographic and Health Survey (2005); Glob Health Action 2016;9: 10.3402/gha

Raghavendra B.K, S.K. Srivatsa, (2011) "Evaluation of Logistic Regression and Neural Network Model with Sensitivity Analysis on Medical Datasets," International Journal of Computer Science and Security (IJCSS), Volume (5): Issue (5): pp. 503 -511. v9.28822

Olowonyo, T, Oshin, S & Iyabo, S. Bello (2006): Some factor associated with low birthweight. Ogun State, Nigeria;Journal/Nigerian medical practitioner/vol.49 No 6

Oladeinde, H.B; Omoreigie, R and Onifade, A.A (2015): Prevalence and determinants of Low birth weight: the situation in a traditional birth home in Benin City, Nigeria, 15(4):1123-1129.

Takai, I. Usman, Audu, B. Mohammed & Bukar, Mohammed (2014): A prospective study of maternal risk factors for low-birth-weight babies in Maiduguri, Nigeria.; Nigeria journal of basic and clinical sciences 11(2):89

Tshotetsi L, Dzikiti L, Hajison P, Feresu S. (2019) Maternal factors contributing to low-birth-weight deliveries in Tshwane District, South Africa. *PLoS One*.14(3). doi: 10.1371/journal.pone.0213058.

Wachamo TM, Yimer NB, Bizuneh AD. (2019) Risk factors for low birth weight in hospitals of North Wello zone, Ethiopia: A case-control study. *PLoS One*.14(3):1-15. doi: 10.1371/journal.pone.0213054

Wanjaria (2016). a Correlation of Ultrasound and Clinical Fetal Weight. J Med Ultrasound. 24(2004):144-148. 3

Yadav H, Lee N (2013) Maternal factors in predicting low birth weight babies. *Med J Malaysia*. ;68(1):44-47.

Yilgwan, C.S, Abok, II Yinnang, W.D & Vajime, B. A (2009). Prevalence and risk factors of low birth weight in Jos, Nigeria; Journal/ Jos Journal of medicine; vol.4.NO.1 UNICEF and WHO. Low birthweight. *East Afr Med J*. 2004; (2):89-90. WHO Low Birth WeightPolicy Brief. *WHA Glob Nutrition Targets 2025*. 2014; 28(3):1-66. UNICEF and WHO. Low birthweight. *East Afr Med J*.(2004);63(2):89-90. doi:10.1787/9789264183902-17-en

Portia A. Cerny, "Data mining and Neural Networks from a Commercial Perspective,"pp:1-10

Raghavendra B.K, S.K. Srivatsa (2011) "Evaluation of Logistic Regression and Neural Network Model With Sensitivity Analysis on Medical Datasets," International Journal of Computer Science and Security (IJCSS), Volume (5): Issue (5), pp. 503 -511.

Mr. Hitesh H. Parmar, Prof. Glory H. Shah. (2013) "Experimental and Comparative Analysis of Machine Learning Classifiers," International Journal of Advanced Research in Computer Science and Software Engineering 3(10), pp. 955-963.

Ned Horning," Random Forests(2010) : An algorithm for image classification and generation of continuous fields data sets,"International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences.

Jehad Ali, Rehanullah Khan, Nasir Ahmad.(2012) "Imran Maqsood Random Forests and DecisionTrees,"IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 3, pp:272-278.

Hem Jyotsana Parashar, Singh Vijendra, and Nisha Vasudeva. (2012) "An Efficient lassification Approach for Data Mining," International Journal of Machine Learning and Computing, Vol. 2, No. 4, pp.446-448.

Janez Demšar, "Material for lectures on Data mining,", Kyoto University, Dept. of Health Informatics, July 2010, pp.1-53.

Lior Rokach &Oded Maimon.(2013)"Data Mining and Knowledge Discovery Handbook," (Decision Tree Chapter 9), pp.165-198.

Patrick Ozer, Dr. I.G. Sprinkhuizen-Kuyper. (2008) "Data Mining Algorithms for Classification," Thesis Artificial Intelligence, Radboud University Nijmegen, pp.1-27 Classification and Regression Trees,36-350, November 2009, Principles of Data Mining, sections 10.5 and 5.2: Berk, chapter 3.