

PREDICTING THE MODE OF CHILD DELIVERY THROUGH DATA MINING***Akinsola Adeniyi F., Sokunbi Michael A., Okikiola Folasade M., and Ogundele Israel O.**

Yaba College of Technology, Computer Tech. Dept., Yaba, Lagos Nigeria.

***Corresponding author:** akinsolafun@gmail.com **Phone:** 08033818385

Manuscript received on 20/09/22 and accepted on 15/12/22

ABSTRACT

In maternity care, the mode of child delivery is crucial to prepare the patient mind considering her health condition before her due date. It helps obstetricians in decision-making to predict the most accurate childbirth mode in advance to avoid unnecessary emergencies and increase the survival of the mother and her unborn child. The type of delivery currently depends on the attending physician's judgment. Humans are prone to error; in several cases, doctors have recommended cesarean section birth when it is medically unnecessary. A wrong decision might also imperil the life of the mother and her unborn child. A computer-aided decision-making model could be a great solution to this problem. With this in mind, we used five data mining methods (Naive Bayes, Logistic model tree (LMT), Decision stump, Logistic Regression, and Random Forest) to create a model that predicted the best mode of infant birth. After training and analyzing the performance of each classifier using WEKA (Waikato Environment for Knowledge Analysis), our results showed that the Naives Bayes classifier has the highest accuracy, precision, and recall scores (89% each) and is the most stable model. This algorithm (Naives Bayes) is embedded to create a predictor that predicts the method of infant birth using a fresh data set that follows the trend of the previous data set. Using this model to decide the mode of childbirth will dramatically reduce maternal and newborn health risks.

Keywords: Mode of delivery, Data mining, Maternity care, Prediction**1. INTRODUCTION**

Examining extensive pre-existing datasets to discover new information is a technique known as "data mining." It implies utilizing one or more software programs to analyze data patterns in big batches of data. The term "data mining" is ambiguous because the objective is not the extraction (or "mining") of data but rather the discovery of patterns and knowledge from vast amounts of data. Data mining techniques continue to be a simple tool for identifying patterns in complex and large amounts of data and turning them into information that can be used for decision-making. Pregnancy traits can be used to predict the type of birth using data mining classification algorithms. A training set of attributes is used by the classification algorithm to identify correlations between the qualities that could be used to predict the outcome. Two forms of data analysis can be used for extracting models they are Classification and Prediction. Prediction models predict

continuous valued function, while classification models predict categorical class labels.

Pregnancy-related issues can include the following: Placenta abruption, Placenta previa, Uterine rupture, Breech position, etc. Pregnancy can be a time of great excitement to the patient, but it can also be a time of danger, and there are certain serious illnesses of pregnancy to be aware of (Sonia et al., 2015). In the UK, pregnant women attend a booking appointment with a midwife between 8- and 12 weeks gestation (Sonia et al., 2015).

The United Nations International Children's Emergency Fund (UNICEF) estimates that Nigeria presently accounts for 10% of all maternal fatalities worldwide. The fourth highest on Earth, according to the most recent statistics, the maternal mortality rate is 576 per 100,000 live births. About 262,000 newborns die at birth each year, which is the

second-highest national total in the world. (2016). Also, the federal government reported last year that Nigeria records 830 daily maternal deaths during childbirth, amounting to roughly 302,950 annually (Onyebuchi, E 2021). By choosing the right birthing method, this high maternal mortality rate can be greatly decreased.

There are several methods used to deliver babies in hospitals; however, the following are the most typical:

- 1) Vaginal delivery
- 2) Cesarean section.

The most typical method of giving birth is vaginal delivery. Methods of assisted delivery are required when necessary. The most natural way to give birth is through the vaginal canal. In this procedure, the mother's vagina and the birth canal are used to deliver the child. The mother will go through three separate stages of labour and delivery during this procedure, beginning with the contractions and concluding with the delivery of the placenta. If a vaginal delivery is not feasible, a cesarean section is a major operation that is carried out. Through surgical incisions created in the abdomen and uterus during this surgery, the baby is born. When normal delivery is not more secure for the baby, the mother, or both, it is advised. While some cesarean deliveries take place as a result of complications during labour, others may be planned and scheduled.

Nigeria's health care system is stratified in such a way that the majority of births take place in facilities where a cesarean section cannot be done and there are no emergency services. This explains why referral clinics receive a significant number of unbooked patients, who are much more likely than booked parturient to undergo an emergency cesarean section with negative obstetric outcomes. However, many of these patients can not have access to sufficient health services in time. Early risk factor identification enables nurses to provide prenatal counseling and hospital referrals for pregnant women before they go into labour, which is essential to the provision and adoption of cesarean sections. Even though the risk factors for cesarean sections have

been extensively defined, it was deemed reasonable to assess these factors locally and identify the primary causes and predictors of cesarean sections in this particular group.

This study uses data mining algorithms to identify obstetric and pregnancy characteristics that can be utilized to predict the best delivery method using actual data acquired at the obstetrics and gynecology section of a Public Hospital in Lagos.

Many data mining models have been developed for the prediction of the type of delivery. The performance of those models is measured based on the accuracy, Area Under the ROC Curve (AUC) or Receiver Operating Characteristic (ROC). Below are various techniques used by researchers:

Kathryn et al (2004) discovered that women are more likely to have cesarean deliveries when the theory that weariness and sleep disturbance in late pregnancy is associated with the type of delivery and the length of labour was examined.

Hussam et al (2016) used the WEKA software to determine which classification algorithm provides the highest accuracy and to build a model that can assist in accurately predicting the possible delivery complication to prepare for and to reduce the risk of the life of both the mother and the baby.

Alisha et al (2015) created a system that acts like a predictor tool allowing users to enter values of specified parameters and giving the most likely mode of delivery as output. They used a classification algorithm to predict the mode of delivery. The classification algorithm used is Naïve Bayes classification and Iterative dichotomiser 3(ID3). This study explains the mathematical representation of the algorithm. According to Alisha Kamat et al; Naïve Bayes theorem gives a way to estimate the posterior probability while Iterative Dichotomiser 3(ID3) is an algorithm that generates a decision tree for a given data set based on the attributes of the data. The system created cannot process a query having a single missing value. Apart from the shortlisted parameters, the system fails to consider other factors that might contribute to determining the mode of delivery.

Kavitha and Balasubramanian (2018) were able to determine the kind of birth by looking at prenatal traits using the J48 decision tree algorithm. It was difficult for the system to use real data from GPP. The accuracy of the system was 100% and the sensitivity is 99.5%

Sonia et al (2015) created a system that was able to predict the mode of delivery using registered obstetrics characteristics using Cross Industry Standard Process for Data Mining (CRISP-DM). The system was unable to identify a model that could predict the best delivery method individually.

Kowsher et al.(2020) has demonstrated good accuracy in predicting the most suitable birthing mode using machine learning-based recommendation systems.

Prema and Pushpalatha (2019) made a survey on related works on preterm birth prediction using data mining techniques. The study compared different data mining techniques such as Naive Bayes, decision tree, Support Vector Machine(SVM), etc. The researchers were able to balance the dataset using data oversampling techniques. In this study, it was difficult to identify preterm birth in first-time mothers and the risk factors considered are limited and the data sample size is small.

Nafiz et al. (2020) using information mining, determine whether a cesarean section is necessary, and then increase the mother's and the baby's safety throughout and after labor by avoiding an unnecessary cesarean segment.

Francisca et al (2016) induced the k nearest neighbor's algorithm with oversampling and testing all the data and they were able to achieve 94% sensitivity, 84% specificity, and 97% accuracy. There was inequality in the number of cases existing on the dataset.

Yavar et al used Naive Bayes, decision trees, SVM, logistic regression, and associative classifier on a Fraser of historical and maternal newborn records to predict preterm birth. The prediction performance of all algorithms is very poor, even doing feature selection by contrast sets does not improve the results.

Kalu et al (2012) used ultrasonographic scan measurement of cervical length and concluded that cervical length in mid-pregnancy can be of value in predicting the mode of delivery in early pregnancy. In this work, the dynamics of labor were examined using a composite analysis of both primigravidae and multigravida.

Saima et al (2017) The researchers explain how numerous ailments are recognized and diagnosed for human health using data mining techniques for medical data. They went ahead to explain the advantages and disadvantages of various data mining techniques. It was also revealed in the study that data mining benefits depend on how clean the data is. They struggled to choose the best strategy for a categorization challenge. It was unable to investigate data mining applications and their effects on society, because data mining technique, user interface, performance, and scalability were all inadequate.

In our study, we will be using classification techniques for the prediction of the type of delivery. The main risk factors considered are age, number of times of pregnancy, weight, number of the fetus, fetal presentation, fetal lie, placenta location, and estimated fetal weight. The class to be considered is the type of delivery. The noise and ambiguity in the dataset will be removed by preprocessing. We assure to achieve high accuracy in our study

3.0 METHODOLOGY

This study utilized data from the department of gynecology and obstetrics of a Public Hospital in Lagos. Five supervised learning algorithms (Naives Bayes, Logistic model tree (LMT), Decision stump, Logistic Regression, Random forest) were employed using the WEKA toolkit. The simulated result was analyzed and compared.

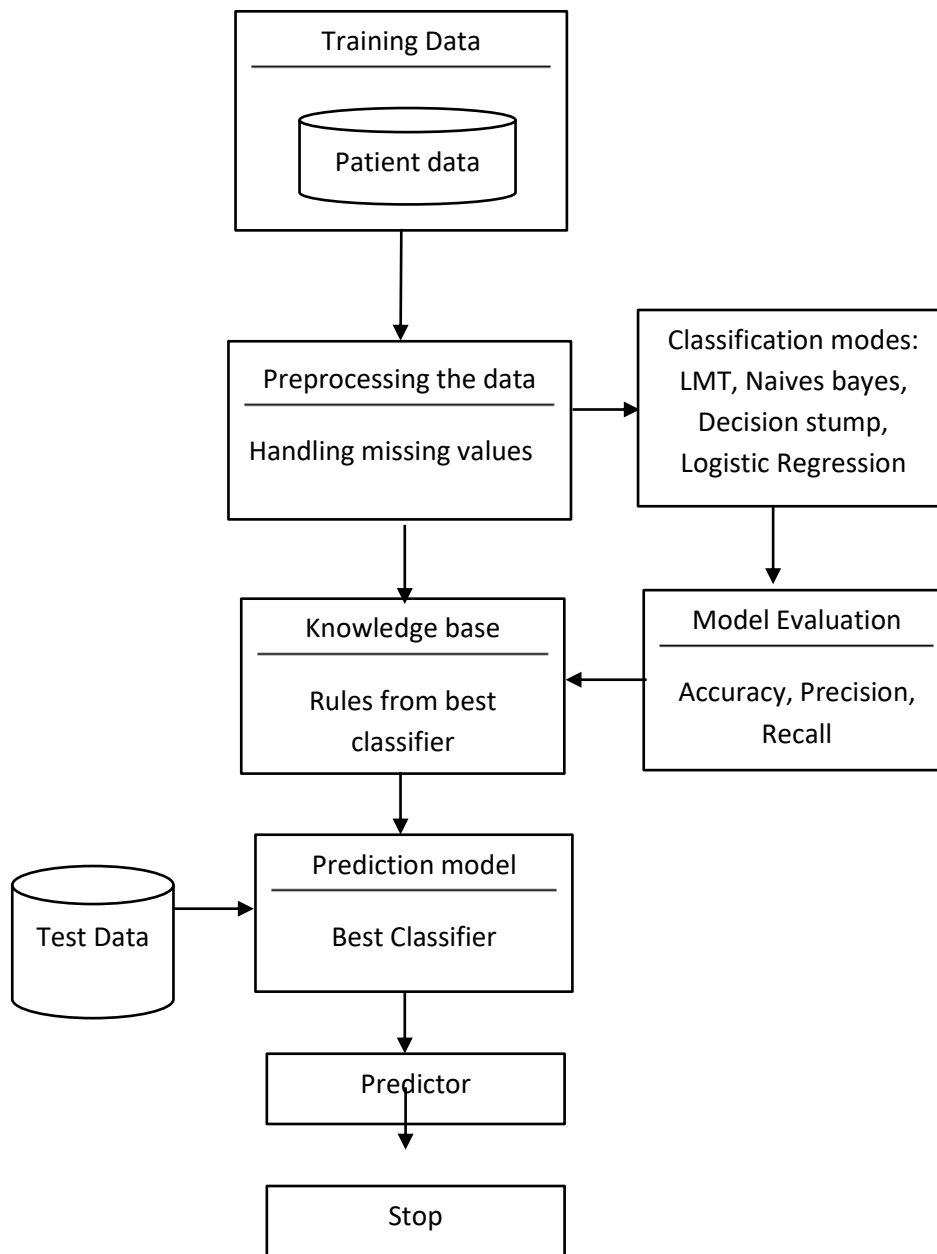


Figure 1. Proposed Model workflow

3.1 Dataset used

The dataset used was real tested data obtained from the department of gynecology and obstetrics of a public hospital in Lagos. The dataset includes 761 records with nine

attributes. Data transformation and pre-processing were performed on the data to make them suitable for mining activities.

Table 1: Dataset used

Variable	Description	Possible values
CARD NUMBER	Patient Identification No	Nominal
PRIMP/P	Is the patient carrying her first pregnancy	Yes or No
MULTI/<3 MULTI/>=3	The patient has children already.	Yes or No Yes or No
AGE (16 -27) AGE (28-50)	Age of the patient	Yes or No Yes or No
GESTATION	Number of foetus	Single or Multiple
FETAL LIE	It describes how the spines of your baby and the mother align. Between your infant's head and tailbone is his spine.	Longitudinal or Oblique
FETAL PRESENTATION	It describes how the infant is positioned up to enter the birth canal and be delivered. It describes the way the baby is situated in the uterus	Cephalic or Breech presentation
TYPE OF DELIVERY	The type of delivery the patient used to give birth	Caesarean section or Vaginal delivery
PLACENTA LOCATION	Where the placenta is located either front or side	Anterior, posterior, fundal

i. **Longitudinal lie:** The term "longitudinal lie" refers to the position of your baby's spine when it is parallel to your spine. Almost all infants are lying longitudinally.

ii. **Oblique lie:** This is when your baby's head is just to the side of the pelvic inlet, that is it is lying diagonally across the uterus. A baby seldom maintains this position up to labor. Only 1% of newborns will have transverse or oblique limbs. Early in the pregnancy, many babies lie sideways; during the final trimester, they shift their heads downward.

iii. **Cephalic presentation:** Head down is the best position for your baby within your uterus during delivery. This is when your baby is present head first. Sometimes between weeks 37 and 40 of your pregnancy, your baby will drop into a cephalic position on its own.

iv. **Breech presentation:** In this case, the baby will come out with feet or buttocks rather than head first. Because the delivery canal must widen for your baby to emerge bottom first, a breech baby might provide challenges for both mother and child.

v. **Posterior placenta:** A fertilized egg typically implants on the uterine wall's back. In this instance, the placenta also forms or grows on the uterus's posterior wall. The placenta is referred to as the posterior placenta when it is in this position.

vi. **Anterior Placenta:** The placenta forms on the front wall of the womb when the fertilized egg attaches to the front of the uterus, and the baby develops behind it. Anterior placenta refers to the placenta when it is on the front side of the uterus.

vii. **Fundal placenta:** The placenta is referred to as a fundal placenta when it lies in the fundus or on top of the uterus. The placenta may occasionally be in either a fundal posterior or anterior position. The top of the womb and the front of the womb are often where a fundal anterior placenta is found. At the top of the womb, a fundal-posterior placenta can be found, but it extends to the rear.

viii. **Vaginal delivery:** It is the birth of offspring through the vagina in mammals.

ix. **Caesarean section:** It involves delivering a baby surgically through an incision created in the mother's abdomen and uterus.

3.2 SUPERVISED LEARNING ALGORITHM

Five supervised learning algorithms: Naives Bayes, Logistic model tree (LMT), Decision stump, Logistic Regression, and Random forest were used to classify the data.

3.2.1 NAÏVE BAYES

Naive Bayes is a simple technique for constructing classifiers: models that assign class labels to problem instances, represented as vectors of feature values, class labels are drawn from some finite set. There is no single algorithm for training such classifiers but a family of algorithms based on shared principles. All Naive Bayes classifiers assume that the value of a particular feature is independent of any other features, given the class variable. For some probability models, Naive Bayes classifiers can be relatively efficient in a supervised learning setting. In many practical applications, parameter estimation for the Naive Bayes model without accepting Bayesian probability or using any Bayesian methods. (Debjani, C et al. 2020)

3.2.2 LMT (Logistic Model Tree)

LMT is a classification model with a supervised approach that combines decision tree learning and logistic regression (LR). The algorithm can handle numeric and nominal properties, missing values, and binary and multiclass target variable. (Niels, L. et al. 2005)

3.2.3 LOGISTIC REGRESSION

Logistic regression is a statistical analysis technique that predicts a data value based on previous observations from a data set. This method enables a machine learning algorithm to classify incoming data based on historical data. A logistic regression model forecasts a dependent data variable by examining the connection between one or more independent data variables. Logistic regression is a quick and easy technique that can be helpful in some challenges. (Jason, B. 2016)

3.2.4 RANDOM FOREST

The random forest classification technique is composed of several decision trees. It is employed in the solution of classification and regression issues. It contains some decision trees on different subsets of the supplied dataset and takes the average to increase the dataset's predicting accuracy. Instead of relying on a single decision tree, the random forest uses the majority vote of projection from each tree and predict the final result. The larger the number of trees in a forest, the better the accuracy and the prevention of overfitting. (Sruthi, E. R (2021)

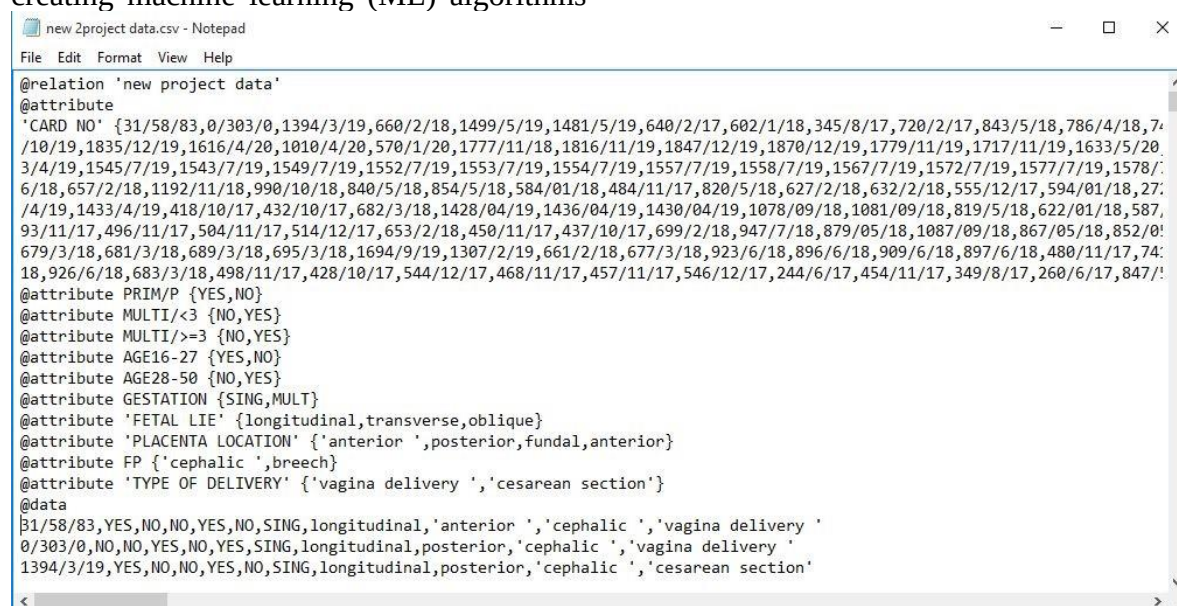
3.2.5 DECISION STUMP

A decision is a one-novel decision tree-based machine learning model. It is a decision tree with a single internal node (the root) connected to the terminal nodes (its leaves). A decision stump predicts something based only on the worth of one particular input feature. They are occasionally referred to as "1 rules". (Sammur, C. et. al.2010)

3.3 WEKA (Tools used)

To carry out this analysis, we used the WEKA toolkit. WEKA is an open-source data mining tool widely used in data mining applications. It is a contemporary tool for creating machine learning (ML) algorithms

and applying them to data mining issues (Consolata 2019). WEKA Explorer was used to loading the data file, and using the classify panel, we ran classification algorithms on the generated dataset.



```

new 2project data.csv - Notepad
File Edit Format View Help
@relation 'new project data'
@attribute
'CARD NO' {31/58/83,0/303/0,1394/3/19,660/2/18,1499/5/19,1481/5/19,640/2/17,602/1/18,345/8/17,720/2/17,843/5/18,786/4/18,74/10/19,1835/12/19,1616/4/20,1010/4/20,570/1/20,1777/11/18,1816/11/19,1847/12/19,1870/12/19,1779/11/19,1717/11/19,1633/5/20,3/4/19,1545/7/19,1543/7/19,1549/7/19,1552/7/19,1553/7/19,1554/7/19,1557/7/19,1558/7/19,1567/7/19,1572/7/19,1577/7/19,1578/7/19,1581/7/19,1584/7/19,1587/7/19,1590/7/19,1593/7/19,1596/7/19,1599/7/19,1602/7/19,1605/7/19,1608/7/19,1611/7/19,1614/7/19,1617/7/19,1620/7/19,1623/7/19,1626/7/19,1629/7/19,1632/7/19,1635/7/19,1638/7/19,1641/7/19,1644/7/19,1647/7/19,1650/7/19,1653/7/19,1656/7/19,1659/7/19,1662/7/19,1665/7/19,1668/7/19,1671/7/19,1674/7/19,1677/7/19,1680/7/19,1683/7/19,1686/7/19,1689/7/19,1692/7/19,1695/7/19,1698/7/19,1701/7/19,1704/7/19,1707/7/19,1710/7/19,1713/7/19,1716/7/19,1719/7/19,1722/7/19,1725/7/19,1728/7/19,1731/7/19,1734/7/19,1737/7/19,1740/7/19,1743/7/19,1746/7/19,1749/7/19,1752/7/19,1755/7/19,1758/7/19,1761/7/19,1764/7/19,1767/7/19,1770/7/19,1773/7/19,1776/7/19,1779/7/19,1782/7/19,1785/7/19,1788/7/19,1791/7/19,1794/7/19,1797/7/19,1800/7/19,1803/7/19,1806/7/19,1809/7/19,1812/7/19,1815/7/19,1818/7/19,1821/7/19,1824/7/19,1827/7/19,1830/7/19,1833/7/19,1836/7/19,1839/7/19,1842/7/19,1845/7/19,1848/7/19,1851/7/19,1854/7/19,1857/7/19,1860/7/19,1863/7/19,1866/7/19,1869/7/19,1872/7/19,1875/7/19,1878/7/19,1881/7/19,1884/7/19,1887/7/19,1890/7/19,1893/7/19,1896/7/19,1899/7/19,1902/7/19,1905/7/19,1908/7/19,1911/7/19,1914/7/19,1917/7/19,1920/7/19,1923/7/19,1926/7/19,1929/7/19,1932/7/19,1935/7/19,1938/7/19,1941/7/19,1944/7/19,1947/7/19,1950/7/19,1953/7/19,1956/7/19,1959/7/19,1962/7/19,1965/7/19,1968/7/19,1971/7/19,1974/7/19,1977/7/19,1980/7/19,1983/7/19,1986/7/19,1989/7/19,1992/7/19,1995/7/19,1998/7/19,2001/7/19,2004/7/19,2007/7/19,2010/7/19,2013/7/19,2016/7/19,2019/7/19,2022/7/19,2025/7/19,2028/7/19,2031/7/19,2034/7/19,2037/7/19,2040/7/19,2043/7/19,2046/7/19,2049/7/19,2052/7/19,2055/7/19,2058/7/19,2061/7/19,2064/7/19,2067/7/19,2070/7/19,2073/7/19,2076/7/19,2079/7/19,2082/7/19,2085/7/19,2088/7/19,2091/7/19,2094/7/19,2097/7/19,2100/7/19,2103/7/19,2106/7/19,2109/7/19,2112/7/19,2115/7/19,2118/7/19,2121/7/19,2124/7/19,2127/7/19,2130/7/19,2133/7/19,2136/7/19,2139/7/19,2142/7/19,2145/7/19,2148/7/19,2151/7/19,2154/7/19,2157/7/19,2160/7/19,2163/7/19,2166/7/19,2169/7/19,2172/7/19,2175/7/19,2178/7/19,2181/7/19,2184/7/19,2187/7/19,2190/7/19,2193/7/19,2196/7/19,2199/7/19,2202/7/19,2205/7/19,2208/7/19,2211/7/19,2214/7/19,2217/7/19,2220/7/19,2223/7/19,2226/7/19,2229/7/19,2232/7/19,2235/7/19,2238/7/19,2241/7/19,2244/7/19,2247/7/19,2250/7/19,2253/7/19,2256/7/19,2259/7/19,2262/7/19,2265/7/19,2268/7/19,2271/7/19,2274/7/19,2277/7/19,2280/7/19,2283/7/19,2286/7/19,2289/7/19,2292/7/19,2295/7/19,2298/7/19,2301/7/19,2304/7/19,2307/7/19,2310/7/19,2313/7/19,2316/7/19,2319/7/19,2322/7/19,2325/7/19,2328/7/19,2331/7/19,2334/7/19,2337/7/19,2340/7/19,2343/7/19,2346/7/19,2349/7/19,2352/7/19,2355/7/19,2358/7/19,2361/7/19,2364/7/19,2367/7/19,2370/7/19,2373/7/19,2376/7/19,2379/7/19,2382/7/19,2385/7/19,2388/7/19,2391/7/19,2394/7/19,2397/7/19,2400/7/19,2403/7/19,2406/7/19,2409/7/19,2412/7/19,2415/7/19,2418/7/19,2421/7/19,2424/7/19,2427/7/19,2430/7/19,2433/7/19,2436/7/19,2439/7/19,2442/7/19,2445/7/19,2448/7/19,2451/7/19,2454/7/19,2457/7/19,2460/7/19,2463/7/19,2466/7/19,2469/7/19,2472/7/19,2475/7/19,2478/7/19,2481/7/19,2484/7/19,2487/7/19,2490/7/19,2493/7/19,2496/7/19,2499/7/19,2502/7/19,2505/7/19,2508/7/19,2511/7/19,2514/7/19,2517/7/19,2520/7/19,2523/7/19,2526/7/19,2529/7/19,2532/7/19,2535/7/19,2538/7/19,2541/7/19,2544/7/19,2547/7/19,2550/7/19,2553/7/19,2556/7/19,2559/7/19,2562/7/19,2565/7/19,2568/7/19,2571/7/19,2574/7/19,2577/7/19,2580/7/19,2583/7/19,2586/7/19,2589/7/19,2592/7/19,2595/7/19,2598/7/19,2601/7/19,2604/7/19,2607/7/19,2610/7/19,2613/7/19,2616/7/19,2619/7/19,2622/7/19,2625/7/19,2628/7/19,2631/7/19,2634/7/19,2637/7/19,2640/7/19,2643/7/19,2646/7/19,2649/7/19,2652/7/19,2655/7/19,2658/7/19,2661/7/19,2664/7/19,2667/7/19,2670/7/19,2673/7/19,2676/7/19,2679/7/19,2682/7/19,2685/7/19,2688/7/19,2691/7/19,2694/7/19,2697/7/19,2700/7/19,2703/7/19,2706/7/19,2709/7/19,2712/7/19,2715/7/19,2718/7/19,2721/7/19,2724/7/19,2727/7/19,2730/7/19,2733/7/19,2736/7/19,2739/7/19,2742/7/19,2745/7/19,2748/7/19,2751/7/19,2754/7/19,2757/7/19,2760/7/19,2763/7/19,2766/7/19,2769/7/19,2772/7/19,2775/7/19,2778/7/19,2781/7/19,2784/7/19,2787/7/19,2790/7/19,2793/7/19,2796/7/19,2799/7/19,2802/7/19,2805/7/19,2808/7/19,2811/7/19,2814/7/19,2817/7/19,2820/7/19,2823/7/19,2826/7/19,2829/7/19,2832/7/19,2835/7/19,2838/7/19,2841/7/19,2844/7/19,2847/7/19,2850/7/19,2853/7/19,2856/7/19,2859/7/19,2862/7/19,2865/7/19,2868/7/19,2871/7/19,2874/7/19,2877/7/19,2880/7/19,2883/7/19,2886/7/19,2889/7/19,2892/7/19,2895/7/19,2898/7/19,2901/7/19,2904/7/19,2907/7/19,2910/7/19,2913/7/19,2916/7/19,2919/7/19,2922/7/19,2925/7/19,2928/7/19,2931/7/19,2934/7/19,2937/7/19,2940/7/19,2943/7/19,2946/7/19,2949/7/19,2952/7/19,2955/7/19,2958/7/19,2961/7/19,2964/7/19,2967/7/19,2970/7/19,2973/7/19,2976/7/19,2979/7/19,2982/7/19,2985/7/19,2988/7/19,2991/7/19,2994/7/19,2997/7/19,3000/7/19,3003/7/19,3006/7/19,3009/7/19,3012/7/19,3015/7/19,3018/7/19,3021/7/19,3024/7/19,3027/7/19,3030/7/19,3033/7/19,3036/7/19,3039/7/19,3042/7/19,3045/7/19,3048/7/19,3051/7/19,3054/7/19,3057/7/19,3060/7/19,3063/7/19,3066/7/19,3069/7/19,3072/7/19,3075/7/19,3078/7/19,3081/7/19,3084/7/19,3087/7/19,3090/7/19,3093/7/19,3096/7/19,3099/7/19,3102/7/19,3105/7/19,3108/7/19,3111/7/19,3114/7/19,3117/7/19,3120/7/19,3123/7/19,3126/7/19,3129/7/19,3132/7/19,3135/7/19,3138/7/19,3141/7/19,3144/7/19,3147/7/19,3150/7/19,3153/7/19,3156/7/19,3159/7/19,3162/7/19,3165/7/19,3168/7/19,3171/7/19,3174/7/19,3177/7/19,3180/7/19,3183/7/19,3186/7/19,3189/7/19,3192/7/19,3195/7/19,3198/7/19,3201/7/19,3204/7/19,3207/7/19,3210/7/19,3213/7/19,3216/7/19,3219/7/19,3222/7/19,3225/7/19,3228/7/19,3231/7/19,3234/7/19,3237/7/19,3240/7/19,3243/7/19,3246/7/19,3249/7/19,3252/7/19,3255/7/19,3258/7/19,3261/7/19,3264/7/19,3267/7/19,3270/7/19,3273/7/19,3276/7/19,3279/7/19,3282/7/19,3285/7/19,3288/7/19,3291/7/19,3294/7/19,3297/7/19,3300/7/19,3303/7/19,3306/7/19,3309/7/19,3312/7/19,3315/7/19,3318/7/19,3321/7/19,3324/7/19,3327/7/19,3330/7/19,3333/7/19,3336/7/19,3339/7/19,3342/7/19,3345/7/19,3348/7/19,3351/7/19,3354/7/19,3357/7/19,3360/7/19,3363/7/19,3366/7/19,3369/7/19,3372/7/19,3375/7/19,3378/7/19,3381/7/19,3384/7/19,3387/7/19,3390/7/19,3393/7/19,3396/7/19,3399/7/19,3402/7/19,3405/7/19,3408/7/19,3411/7/19,3414/7/19,3417/7/19,3420/7/19,3423/7/19,3426/7/19,3429/7/19,3432/7/19,3435/7/19,3438/7/19,3441/7/19,3444/7/19,3447/7/19,3450/7/19,3453/7/19,3456/7/19,3459/7/19,3462/7/19,3465/7/19,3468/7/19,3471/7/19,3474/7/19,3477/7/19,3480/7/19,3483/7/19,3486/7/19,3489/7/19,3492/7/19,3495/7/19,3498/7/19,3501/7/19,3504/7/19,3507/7/19,3510/7/19,3513/7/19,3516/7/19,3519/7/19,3522/7/19,3525/7/19,3528/7/19,3531/7/19,3534/7/19,3537/7/19,3540/7/19,3543/7/19,3546/7/19,3549/7/19,3552/7/19,3555/7/19,3558/7/19,3561/7/19,3564/7/19,3567/7/19,3570/7/19,3573/7/19,3576/7/19,3579/7/19,3582/7/19,3585/7/19,3588/7/19,3591/7/19,3594/7/19,3597/7/19,3600/7/19,3603/7/19,3606/7/19,3609/7/19,3612/7/19,3615/7/19,3618/7/19,3621/7/19,3624/7/19,3627/7/19,3630/7/19,3633/7/19,3636/7/19,3639/7/19,3642/7/19,3645/7/19,3648/7/19,3651/7/19,3654/7/19,3657/7/19,3660/7/19,3663/7/19,3666/7/19,3669/7/19,3672/7/19,3675/7/19,3678/7/19,3681/7/19,3684/7/19,3687/7/19,3690/7/19,3693/7/19,3696/7/19,3699/7/19,3702/7/19,3705/7/19,3708/7/19,3711/7/19,3714/7/19,3717/7/19,3720/7/19,3723/7/19,3726/7/19,3729/7/19,3732/7/19,3735/7/19,3738/7/19,3741/7/19,3744/7/19,3747/7/19,3750/7/19,3753/7/19,3756/7/19,3759/7/19,3762/7/19,3765/7/19,3768/7/19,3771/7/19,3774/7/19,3777/7/19,3780/7/19,3783/7/19,3786/7/19,3789/7/19,3792/7/19,3795/7/19,3798/7/19,3801/7/19,3804/7/19,3807/7/19,3810/7/19,3813/7/19,3816/7/19,3819/7/19,3822/7/19,3825/7/19,3828/7/19,3831/7/19,3834/7/19,3837/7/19,3840/7/19,3843/7/19,3846/7/19,3849/7/19,3852/7/19,3855/7/19,3858/7/19,3861/7/19,3864/7/19,3867/7/19,3870/7/19,3873/7/19,3876/7/19,3879/7/19,3882/7/19,3885/7/19,3888/7/19,3891/7/19,3894/7/19,3897/7/19,3900/7/19,3903/7/19,3906/7/19,3909/7/19,3912/7/19,3915/7/19,3918/7/19,3921/7/19,3924/7/19,3927/7/19,3930/7/19,3933/7/19,3936/7/19,3939/7/19,3942/7/19,3945/7/19,3948/7/19,3951/7/19,3954/7/19,3957/7/19,3960/7/19,3963/7/19,3966/7/19,3969/7/19,3972/7/19,3975/7/19,3978/7/19,3981/7/19,3984/7/19,3987/7/19,3990/7/19,3993/7/19,3996/7/19,3999/7/19,4002/7/19,4005/7/19,4008/7/19,4011/7/19,4014/7/19,4017/7/19,4020/7/19,4023/7/19,4026/7/19,4029/7/19,4032/7/19,4035/7/19,4038/7/19,4041/7/19,4044/7/19,4047/7/19,4050/7/19,4053/7/19,4056/7/19,4059/7/19,4062/7/19,4065/7/19,4068/7/19,4071/7/19,4074/7/19,4077/7/19,4080/7/19,4083/7/19,4086/7/19,4089/7/19,4092/7/19,4095/7/19,4098/7/19,4101/7/19,4104/7/19,4107/7/19,4110/7/19,4113/7/19,4116/7/19,4119/7/19,4122/7/19,4125/7/19,4128/7/19,4131/7/19,4134/7/19,4137/7/19,4140/7/19,4143/7/19,4146/7/19,4149/7/19,4152/7/19,4155/7/19,4158/7/19,4161/7/19,4164/7/19,4167/7/19,4170/7/19,4173/7/19,4176/7/19,4179/7/19,4182/7/19,4185/7/19,4188/7/19,4191/7/19,4194/7/19,4197/7/19,4200/7/19,4203/7/19,4206/7/19,4209/7/19,4212/7/19,4215/7/19,4218/7/19,4221/7/19,4224/7/19,4227/7/19,4230/7/19,4233/7/19,4236/7/19,4239/7/19,4242/7/19,4245/7/19,4248/7/19,4251/7/19,4254/7/19,4257/7/19,4260/7/19,4263/7/19,4266/7/19,4269/7/19,4272/7/19,4275/7/19,4278/7/19,4281/7/19,4284/7/19,4287/7/19,4290/7/19,4293/7/19,4296/7/19,4299/7/19,4302/7/19,4305/7/19,4308/7/19,4311/7/19,4314/7/19,4317/7/19,4320/7/19,4323/7/19,4326/7/19,4329/7/19,4332/7/19,4335/7/19,4338/7/19,4341/7/19,4344/7/19,4347/7/19,4350/7/19,4353/7/19,4356/7/19,4359/7/19,4362/7/19,4365/7/19,4368/7/19,4371/7/19,4374/7/19,4377/7/19,4380/7/19,4383/7/19,4386/7/19,4389/7/19,4392/7/19,4395/7/19,4398/7/19,4401/7/19,4404/7/19,4407/7/19,4410/7/19,4413/7/19,4416/7/19,4419/7/19,4422/7/19,4425/7/19,4428/7/19,4431/7/19,4434/7/19,4437/7/19,4440/7/19,4443/7/19,4446/7/19,4449/7/19,4452/7/19,4455/7/19,4458/7/19,4461/7/19,4464/7/19,4467/7/19,4470/7/19,4473/7/19,4476/7/19,4479/7/19,4482/7/19,4485/7/19,4488/7/19,4491/7/19,4494/7/19,4497/7/19,4500/7/19,4503/7/19,4506/7/19,4509/7/19,4512/7/19,4515/7/19,4518/7/19,4521/7/19,4524/7/19,4527/7/19,4530/7/19,4533/7/19,4536/7/19,4539/7/19,4542/7/19,4545/7/19,4548/7/19,4551/7/19,4554/7/19,4557/7/19,4560/7/19,4563/7/19,4566/7/19,4569/7/19,4572/7/19,4575/7/19,4578/7/19,4581/7/19,4584/7/19,4587/7/19,4590/7/19,4593/7/19,4596/7/19,4599/7/19,4602/7/19,4605/7/19,4608/7/19,4611/7/19,4614/7/19,4617/7/19,4620/7/19,4623/7/19,4626/7/19,4629/7/19,4632/7/19,4635/7/19,4638/7/19,4641/7/19,4644/7/19,4647/7/19,4650/7/19,4653/7/19,4656/7/19,4659/7/19,4662/7/19,4665/7/19,4668/7/19,4671/7/19,4674/7/19,4677/7/19,4680/7/19,4683/7/19,4686/7/19,4689/7/19,4692/7/19,4695/7/19,4698/7/19,4701/7/19,4704/7/19,4707/7/19,4710/7/19,4713/7/19,4716/7/19,4719/7/19,4722/7/19,4725/7/19,4728/7/19,4731/7/19,4734/7/19,4737/7/19,4740/7/19,4743/7/19,4746/7/19,4749/7/19,4752/7/19,4755/7/19,4758/7/19,4761/7/19,4764/7/19,4767/7/19,4770/7/19,4773/7/19,4776/7/19,4779/7/19,4782/7/19,4785/7/19,4788/7/19,4791/7/19,4794/7/19,4797/7/19,4800/7/19,4803/7/19,4806/7/19,4809/7/19,4812/7/19,4815/7/19,4818/7/19,4821/7/19,4824/7/19,4827/7/19,4830/7/19,4833/7/19,4836/7/19,4839/7/19,4842/7/19,4845/7/19,4848/7/19,4851/7/19,4854/7/19,4857/7/19,4860/7/19,4863/7/19,4866/7/19,4869/7/19,4872/7/19,4875/7/19,4878/7/19,4881/7/19,4884/7/19,4887/7/19,4890/7/19,4893/7/19,4896/7/19,4899/7/19,4902/7/19,4905/7/19,4908/7/19,4911/7/19,4914/7/19,4917/7/19,4920/7/19,4923/7/19,4926/7/19,4929/7/19,4932/7/19,4935/7/19,4938/7/19,4941/7/19,4944/7/19,4947/7/19,4950/7/19,4953/7/19,4956/7/19,4959/7/19,4962/7/19,4965/7/19,4968/7/19,4971/7/19,4974/7/19,4977/7/19,4980/7/19,4983/7/19,4986/7/19,4989/7/19,4992/7/19,4995/7/19,4998/7/19,5001/7/19,5004/7/19,5007/7/19,5010/7/19,5013/7/19,5016/7/19,5019/7/19,5022/7/19,5025/7/19,5028/7/19,5031/7/19,5034/7/19,5037/7/19,5040/7/19,5043/7/19,5046/7/19,5049/7/19,5052/7/19,5055/7/19,5058/
```

4. RESULT AND DISCUSSION

The simulated results obtained after the various algorithms were tested on the 761 dataset are partitioned into various tables for

easier analysis and evaluation. The algorithms are compared by analyzing their accuracy, precision, recall, and F-measure.

Table: 2. The accuracy of the classifiers

Algorithm (total instances 761)	accurately classified instance % (value)	Inaccurately classified instance (values)%	Absolute mean error	Time Taken (seconds)
LMT (Logistic Model Tree)	75.43	24.57	0.36	0.08
Naïve Bayes	88.57	11.43	0.28	0.06
Logistic Regression	100	0	0	0.11
Decision Stump	69.25	30.73	0.42	0.02
Random forest	100	0	0.16	0.11

Table 3: The classifier accuracy evaluation measures by class

Algorithm	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
LMT (Logistic Model Tree)	0.943	0.625	0.752	0.943	0.837	0.406	0.782	0.853	Vaginal delivery
	0.375	0.057	0.766	0.375	0.504	0.406	0.782	0.639	Cesarean section
	0.754	0.436	0.757	0.754	0.726	0.406	0.782	0.782	Weighted Average
Naïve Bayes	0.959	0.261	0.881	0.959	0.918	0.738	0.967	0.987	Vaginal delivery
	0.739	0.041	0.899	0.739	0.811	0.738	0.967	0.908	Cesarean section
	0.886	0.188	0.887	0.886	0.883	0.738	0.967	0.960	Weighted Average
Logistic Regression	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Vaginal delivery
	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Cesarean section
	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Weighted Average
Decision Stump	0.913	0.751	0.709	0.913	0.799	0.220	0.581	0.706	Vaginal delivery
	0.249	0.087	0.589	0.249	0.350	0.220	0.581	0.396	Cesarean section
	0.693	0.530	0.669	0.693	0.649	0.220	0.581	0.603	Weighted Average
Random Forest	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Vaginal delivery
	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Cesarean section
	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	Weighted Average

From the comparison result of evaluation metric in table 2 and 3

Naïve Bayes with an accuracy of **88.57%** and the total time taken to build the model is 0.03 seconds. It has a weighted average of 0.886 TP rate, 0.188 FP rate, 0.887 precision, 0.886 recall, 0.883 F-measure, 0.738 MCC, 0.967 ROC area, and 0.960 PRC area appeared to be most stable and generalizable; LMT has an accuracy of 75.43%, Logistic regression has accuracy of 100%, Decision stump has an accuracy of 69.25% and Random forest has an accuracy of 100%. Logistic regression and Random forest gave 100% accuracy. When a classifier generates 100% accuracy, it means it has overfitting problem. Data scientists refer to this as overfitting, which happens when a statistical model matches its training data exactly. This defeats the algorithm's aim and makes it impossible for

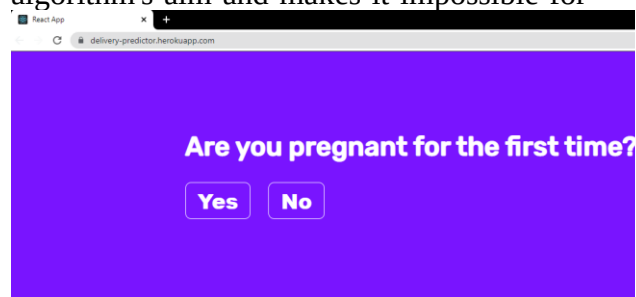


Figure 3 The Predictor interface

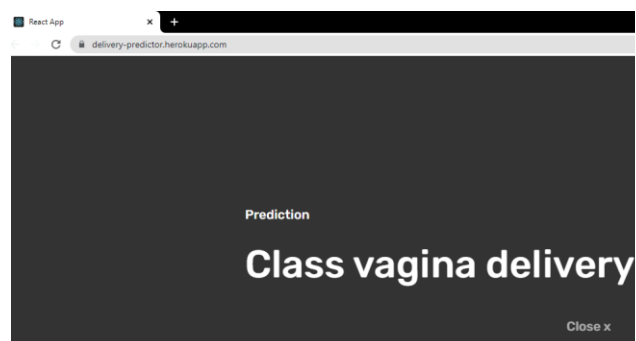


Figure 4. Prediction output for vagina delivery

6. CONCLUSION

A family may be able to determine if a woman genuinely need a caesarean section with the help of an early prediction, which would reduce unnecessary anxiety, procedures, and greatly increase medical

it to generalize successfully to new data. When this occurs, the algorithm cannot perform accurately against unseen data. A model won't be able to carry out the classification or prediction task it was designed for if it can't generalize successfully to new data.

5. PREDICTOR

The Naïve Bayes algorithm, which is the most stable, was used to build the rules for the predictor, which was created using JavaScript. An obstetrics scan must have been completed by the user before filling out the prediction form. In less than 5 seconds, the predictor makes a single prediction. Once the patient has completed the form and provided information regarding her obstetrics scan, the predictor will predict the patient's potential birth type.

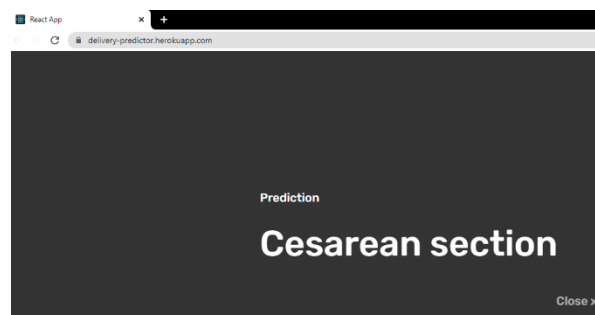


Figure 5. Prediction output for caesarean section

benefits. This research work has been an attempt to carry out study and comparative analysis on LMT, Naive Bayes, Logistic Regression, Decision Stump and Random Forest.

LMT has an accuracy of 75.43%, Naïve Bayes has an accuracy of 88.57%, Logistic regression has accuracy of 100%, Decision stump has an accuracy of 69.25% and Random forest has an accuracy of 100%

The most accurate algorithm using the patient data is Nave Bayes, which takes 0.03 seconds to create. Its accuracy is 88.57%. It has a weighted average of 0.886 TP rate, 0.188 FP rate, 0.887 precision, 0.886 recall, 0.883 F-measure, 0.738 MCC, 0.967 ROC area, and 0.960 PRC area. With the help of this model, the doctor will be able to decide more accurately and quickly. Based on the performance generated by Naïve Bayes a predictor was developed. The predictor provides a clickable form that helps the patients provide information of their obstetrics scan result. Pregnant women would benefit immensely from this predictor by knowing their expected delivery method, which will save unnecessary anxiety.

7. ACKNOWLEDGEMENT

The dataset was gathered from the gynecology and obstetrics department of a public hospital in Lagos. We are unable to reveal the name in this context due to data sensitivity and confidentiality concerns.

REFERENCES

- Alisha, K. Veenal, O. & Manalee, D. (2015). Implementation of classification algorithms to predict mode of delivery. *International journal of computer science and information technologies*, 6(5),4531-4534. Retrieved May 15, 2022, from <https://ijcsit.com/docs/Volume%206/vol6issue05/ijcsit2015060579.pdf>
- Consolata, G. & Jeniffer, J (2019). A classification model for water quality analysis using decision tree. *European Journal of Computer Science and Information Technology*, 7, No.3, pp.1-8, June. Retrieved. May 15, 2022. From-<https://www.coursehero.com/file/131523736/A-classification-model-for-water-quality-analysis-using-decision-treepdf/>
- Debjani, C. Ahona, G. Sriparna, S. (2020). A survey on Internet-of-Thing applications using electroencephalogram. In *Emergence of Pharmaceutical Industry Growth with Industrial IoT Approach*, Pages 21-47. Retrieved May 15, 2022 <https://doi.org/10.1016/C2019-0-00219-1>
- Francisca, F., Hugo, P., Filipe, M., Jose, M. & Antonio, A. (2017). Step towards prediction of Perineal tear. *International Workshop on Healthcare Interoperability and Pervasive Intelligent Systems. Procedia Computer science* 113:565-570. <https://doi.org/10.1016/j.procs.2017.08.284>
- Hussam, A. Amer, K. & Mwaffaq, O. (2016) Mode of delivery prognosis through data mining. 3rd International Conference on Soft Computing & Machine Intelligence (ISCMI), 1,63-66. DOI: 10.1109/ISCMI.2016.10
- Jason, B. (2016) Logistic Regression for Machine Learning. *Machine Learning Algorithms*. Retrieve May 15, 2022 from <https://machinelearningmastery.com/logistic-regression-for-machine-learning/>
- Kalu, C.A. Umeora, O.U. Egwuatu, E.V. & Okwor, A. (2012), Predicting mode of delivery using mid-pregnancy ultrasonographic measurement of cervical length” *Nigerian Journal of Clinical Practice*, 15,3, Pg 338-343, Retrieved May 17, 2022. From <https://www.njconline.com/article.asp?issn=1119-3077;year=2012;volume=15;issue=3;page=338;epage=343;aulast=Kalu>
- Kavitha, D. & Balasubramanian, T. (2018). Predicting the mode of delivery and the risk factors associated with cesarean delivery using decision tree model. *International Journal of Engineering Sciences & Research Technology*, 7(8), 116-124. DOI: 10.5281/zenodo.1336707

- Kathryn, A. & Caryl, L. (2004) Sleep in late pregnancy predicts length of labor and type of delivery". PMID: 15592289, DOI: 10.1016/j.ajog.2004.05.086
- Kowsher, M. Nusrat, J.P. Anik, T. A. & Babul, I.M. (2020). Machine Learning Based Recommendation Systems for the Mode of Childbirth. International Conference on Cyber Security and Computer Science. Retrieved May 17, 2022. from <https://www.springerprofessional.de/en/machine-learning-based-recommendation-systems-for-the-mode-of-ch/18230538>
- Nafiz, I.K. Tahasin, M. Muhammad, N. I. Sumaiya, N.M. (2020) "Prediction of Cesarean Childbirth using Ensemble Machine Learning Method". 22nd International Conference on Information Integration and Web-based Applications & Services. DOI:10.1145/3428757.3429138
- Niels, L., Mark, H., & Eibe, F. (2005) Logistic Model Trees. Machine Learning, 59, pages 161–205. DOI:10.1007/s10994-005-0466-3
- Onyebuchi, E. (2021) “302,950 Pregnant Women Die Yearly in Nigeria”, This Day New Paper. Retrieved September 12, 2022. from <https://www.thisdaylive.com/index.php/2021/05/20/302950>
- Prema, N.S. & Pushpalatha, M.P. (2019). Prediction of preterm birth using data mining. IIOAB Journal, 10, |2| |13-17|. Retrieved September 12, 2022. from https://www.iioab.org/IIOABJ_10.2_13-17.pdf
- Saima, A.L. Rosziati, I. Norhalina, S. & Taujuddin, N.S.A.M. (2018). Application of data mining techniques for medical data classification. *MATEC Web of Conferences* **150**, 06003; doi.org/10.1051/mateconf/201815006003
- Sammut, C., Webb, G.I (2010) Encyclopedia of Machine Learning. Springer, Boston, MA DOI: https://doi.org/10.1007/978-0-387-30164-8_202
- Sónia, P. Filipe, P. Manuel, F. José, M. & António, A. (2015). Predicting Type of Delivery by Identification of Obstetric Risk Factors through Data Mining. *Procedia Computer Science* 64 :601 – 609. <https://doi.org/10.1016/j.procs.2015.08.573>
- Sruthi, E. R (2021) Understanding Random Forest. Retrieved from- May 15, 2022 <https://www.analyticsvidhya.com/blog/2021/06/understanding-random-forest/>
- Yavar, N. Mojdeh, J.H. & Amit, S. (2008). Predicting preterm birth based on maternal and fetal data. *International Journal of Engineering & Technology* 7(2):5-8; Retrieve May 17, 2022. From- <http://yavar.naddaf.name/downloads/Predicting%20Preterm%20Birth%20Based%20on%20Maternal%20and%20Fetal%20Data.pdf>