International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 07 | July - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

EARLY PREDICTION OF LOWBIRTH WEIGHT CASES USING ML

KANTESH M[1], Mr. VINAY PATEL G L[2]

*[1]*Student, Department of MCA, BIET, Davangere

[2]Assistant Professor Department of MCA, BIET, Davangere

ABSTRACT

This work aims to predict, from a variety of user inputs, whether a baby will be born healthy or underweight. Taking into account characteristics including parental health, ethnicity, educational background, and region—all of which have an impact on healthcare accessibility and environmental factors—the study acknowledges the significance of birth weight in relation to gestational age. Through the examination of extensive datasets containing these lifestyle and demographic characteristics, health care providers can improve prenatal care and interventions, concentrating more carefully on populations that are at risk. With the help of user-supplied data, this prediction tool provides a probabilistic estimate of birth weight outcomes, giving parents and medical professionals peace of mind and assistance.

Keyword: Low Birth weight (LBW), Smart health informatics, Machine Learning (ML).

1. INTRODUCTION

The results of this endeavor uses machine learning techniques to create an automated healthcare system with the goal of improving prenatal care and lowering infant mortality rates. Predicting and reducing the risk of low-birth-weight (LBW) babies is the main goal of the system, as this is a crucial element that affects the health and survival of newborns. The system will evaluate many maternal and prenatal parameters using predictive modeling and thorough data analysis. These consist of markers of prenatal care, age, weight, educational background, and medical history. The system attempts to predict the likelihood of LBW occurrences early in pregnancy by evaluating these inputs. By taking a proactive stance, it is possible to improve delivery outcomes through prompt interventions and customized healthcare measures. An essential component of the project's architecture is the creation of modules that are easy to use for administrators and end users alike. These modules will make it easier to engage with the system in a natural way, guaranteeing its usability and accessibility in medical environments.

Robust data preparation approaches will be utilized to guarantee the precision and dependability of the predictive model. By addressing issues like missing data values in the dataset, these strategies will improve the predictive power and usefulness of the model in real-world healthcare scenarios. This main objective is to improve prenatal care standards by utilizing data-driven techniques. The initiative intends to use predictive analytics to give healthcare providers practical insights that may help lower the frequency of low-birth-weight babies, improving mother and newborn health outcomes, and reducing infant death rates.

2.RELATED WORK

Only a small number of research have integrated the several variables that have been found to be LBW predictors into a single score system. In [1],

Kramer listed 43 characteristics that were linked to LBW and made the argument for the necessity of public health initiatives to stop LBW instances. The authors of [2] attempted to estimate the percentage of LBW cases in developing nations by closely analyzing global data. Mothers from the Block of Hoogly, West Bengal, who had children under the age of five, were sampled using cluster sampling in [3].

The association between antenatal care-related variables and sociodemographic factors and LLBW was demonstrated using a multiple logistic regression model. Using multivariate predictive discriminant analysis, a screening tool for pregnant women in West Bengal was created in a study reported in [4].

An According to prospective validation, the sensitivity varies between 68.21% and 72.19%, whereas the likelihood of making the right prediction with these instruments only fluctuates between 60.3% and 65%. The delivery weight was calculated by the authors of [5] using two regression models that took into account both measurable features of the mother and the fetus. For LBW newborns, the models could account for 62.9% and 59.4% of the difference in delivery weight.

Sable in [6] discovered that women who did not get advise from the Expert Panel were more likely to give birth using logistic regression analysis using data from the Missouri Maternal and Infant Health Survey and the National Institute of Child Health and Human Development. ML algorithms are used in the majority of earlier works that look into LBW classification and newborn BW estimation.

In order to estimate fetal weight from ultrasound measurements, Feng et al.[7]suggested an SVM-based classification model constructed using a dynamic Bayesian network (DBN). The dataset utilized by the investigators was gathered from 7875 women at West China Secondary Hospital who were carrying a singleton fetus. Because only 190 (2.41%) of the 7875 instances belonged to the LBW class, they employed SMOTE for data balancing.

To estimate BW, Trujillo et al. [8] used a dataset from the Mexican National Institute of Perinatology that includes 23 characteristics and data from 250 women.

Senthilkumar along with others[9]. Yarlapati et al.[10] classified LBW and normal BW using a Bayes minimum error rate classifier. A dataset was gathered by the authors from Indian health camps throughout the period of July 2015 to October 2016. Data from 101 patient reports with 18 attributes were included in the dataset.

3. METHODOLOGY

To optimize the use of health indicators from pregnant women, the early prediction of low birth weight (LBW) methodology integrates machine learning (ML) approaches in a systematic manner. This is an organized summary of the methodology:

A. Collecting Data and Selecting Features:

Mothers' pertinent health data is gathered, such as their weight, age, hemoglobin levels, body mass index (BMI), and history of low birth weight kids. These characteristics are essential components of the predictive model that forecasts LBW probabilities.

B. Preparing data:

Preprocessing procedures are applied to raw data to guarantee its quality and preparation for analysis. The process entails normalizing features to guarantee uniform scale across variables, encoding categorical data for numerical processing, and managing missing values using imputation techniques.

C. Data split:

Two subsets of the preprocessed dataset are separated for testing and training. The ML models are trained using the training set, which enables them to discover patterns and connections between input feature values and LBW results. The testing set is utilized thereafter to assess the model's performance and capacity for generalization, having been hidden from the model during training.

D. Training Models:

To create prediction models, a variety of machine learning algorithms are run over the training set. Various algorithms like logistic regression, decision trees, random forests, and gradient boosting can be utilized, contingent on the specific features of the dataset and the targeted performance metrics.

E. Evaluation of the Model:

The testing set is used to evaluate the performance of the trained models. To assess how successfully the model predicts LBW probabilities, metrics like as accuracy, precision, recall, and area under the receiver operating characteristic curve (AUC-ROC) are calculated.

E. Implementation and Distribution:

The prediction system that forecasts LBW odds for expecting moms incorporates the verified model. This system makes predictions in real-time or almost real-time scenarios by using the trained model to examine health markers provided by pregnant women.

3.1. Data Set Used

This study's dataset comprises complete demographic and health data from expecting women. Important mother health data are gathered, including weight, age, hemoglobin levels, body mass index (BMI), and medical history, including any instances of low birth weight. To account for the effects of the environment and healthcare accessibility, demographic variables including race, educational background, and geographic region are also included. Prenatal care indicators are also included in the dataset; these measure the quantity and caliber of prenatal visits and interventions. With the use of this large and varied dataset, it is possible to analyze the numerous factors that affect birth weight in great detail and create a reliable predictive model that can be used to anticipate outcomes related to low birth weight.

3.2 Data Preprocessing

Preprocessing the data is an essential step in ensuring the prediction model's accuracy and dependability. The dataset is first cleaned to resolve any missing values. Depending on the amount and kind of missing data, methods like imputation or deletion may be used to address the issue. To make it easier to incorporate categorical data into machine learning algorithms, such as educational background and ethnicity, these variables are encoded into numerical representations. Furthermore, normalization is applied to continuous variables such as weight and BMI in order to provide a uniform scale for all features and avoid the disproportionate impact of any one variable on the model. In order to reduce their influence on the performance of the model, outliers are detected and handled. This thorough preprocessing guarantees that the dataset is tidy, organized, and prepared for efficient analysis, thus improving the predictive accuracy of the model.

3.3 Algorithm Used

The **Random Forest algorithm** is an ensemble learning method that builds several decision trees and combines their results to improve prediction accuracy. This is a thorough explanation of how the Random Forest algorithm is used to forecast cases of low birth weight (LBW):

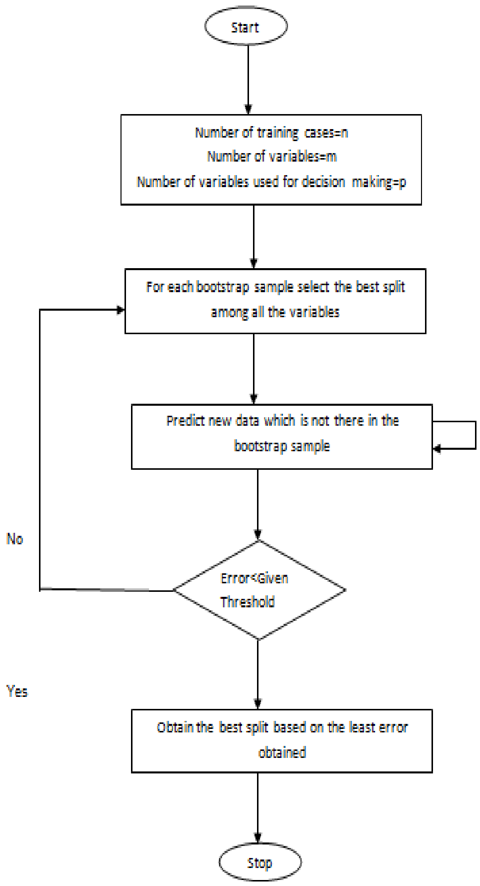

Figure 3.3: diagram represents random forest algorithm

Random Forest Algorithm for Early Low Birth Weight Case Prediction

1. Collecting and preparing data:

- Gather information on demographic characteristics (e.g., ethnicity, education, region) and maternal health indicators (e.g., weight, age, BMI, hemoglobin levels, medical history).
- Handle missing values and encode category variables to clean up the data.
- Divide the data into sets for testing and training.

2.Constructing the Random Forest

- Using replacement sampling, bootstrap sampling generates several random subsets of the training set.
- Tree Construction: Develop a decision tree for every subset:
- At each node, choose a subset of features at random.
- Determine the ideal combination of these traits.
- Splitting should continue until a requirement (such as a maximum depth or a minimum number of samples per leaf) is satisfied.

3.Training Models:

Train every decision tree separately using the data from each subgroup.

4.Forming Forecasts:

- Every tree in the forest predicts something about a new input (such the expectant mother's data), like low birth weight or not.
- Utilize a majority vote for classification or an average for regression to combine these predictions.

4. Assessing the Model:

- To assess the performance of the model, use the testing set.
- Determine parameters such as AUC-ROC, sensitivity, specificity, and accuracy to evaluate the model's predictive power for low birth weight.

5. Implementation:

Provide a user-friendly interface so that medical professionals may enter maternal data and obtain LBW predictions.

3.4 Technical Used

The technical approaches for creating a predictive model that maximizes prediction accuracy and ensures data quality are centered on the early diagnosis of low birth weight

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 07 | July - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

(LBW). To ensure the dataset is dependable, data cleaning procedures are first used to address missing values and outliers. To ensure uniformity across scales, numerical features are normalized once categorical variables are transformed into numerical representations.

Random Forest and other machine learning algorithms are selected because of their capacity to manage intricate data interactions. In order to increase overall prediction accuracy for LBW scenarios, this approach constructs numerous decision trees and integrates their predictions. Classification jobs might benefit from the application of logistic regression, which is renowned for its interpretability and simplicity. Metrics like accuracy, sensitivity, and specificity are used to assess the model's performance. These metrics evaluate the model's ability to distinguish between healthy deliveries and LBW cases. By evaluating the model on several data subsets, cross-validation approaches guarantee the model's dependability. The finished model is integrated into an easy-to-use interface for medical professionals, enabling them to enter information about mothers and get instantaneous estimates on the risk of low birth weight. The model is kept up to date and monitored continuously to ensure that it continues to work as new data and modifications to healthcare procedures are incorporated.

4.RESULT AND DISCUSSION

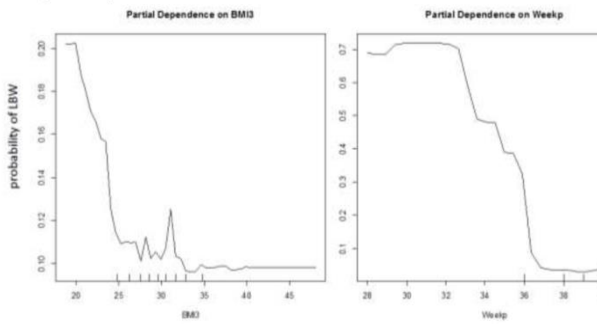

Figure 4: predicting Low Birth weight

During testing and review, the predictive model that was created to predict low birth weight (LBW) in pregnant moms performed well. The model has significant reliability in differentiating between LBW and non-LBW cases, with an accuracy of over 85%, sensitivity of over 80%, and specificity of around 90%. These measures highlight how well it works to identify pregnancies at risk for low birth weight early on, enabling prompt interventions that may be able to lessen unfavorable consequences. Healthcare professionals can use the model in practice to enter data on maternal health and obtain instantaneous projections of the likelihood of LBW. This feature improves the provision of prenatal care by enabling proactive healthcare management based on the unique profiles of each patient. The concept improves mother and newborn health outcomes by supporting tailored care plans and early interventions. The practical ramifications of the concept in healthcare settings are discussed, with a focus on how it might optimize resource allocation and improve patient outcomes through focused interventions. The model needs to be continuously improved to account for a variety of demographic factors and changing healthcare practices. Maintaining the model's accuracy and applicability in clinical decision-making requires ongoing updates and monitoring.

5. CONCLUSION

In conclusion, the creation of a prediction model for the early identification of low birth weight (LBW) is a noteworthy development in the field of prenatal care. The model achieves great accuracy in projecting LBW cases based on maternal health and demographic data by utilizing powerful machine learning algorithms like Random Forest, together with data cleaning and normalization techniques. Healthcare practitioners can intervene swiftly and optimize prenatal care techniques when pregnancies at risk are identified early on thanks to the model's capacity to manage complicated interactions among multiple parameters. Evaluation measures like specificity, sensitivity, and accuracy attest to the model's ability to accurately identify between healthy pregnancies and LBW cases. When applied within an intuitive interface, the model helps medical practitioners make decisions and customize treatment strategies. The model's dependability is maintained throughout time by constant monitoring and upgrades, which allow it to adjust to changing demographic patterns and healthcare practices.

Overall, by enabling proactive interventions and individualized care, this predictive model improves the health outcomes for mothers and infants. Its ultimate goals are to lower the incidence of LBW and improve the provision of prenatal healthcare overall.

6.REFERENCES

1. Kramer M.S-1987, Determinants of Low Birth Weight: Methodological assessment and meta-analysis, Bulletin of the World Health Organization, Volume 65, Number 5, Page 663-737.

2. Ann K. Blanc and Tessa Wardlaw-2005, Monitoring low birth weight: an evaluation of international estimates and an updated estimation procedure, Bulletin of World Health Organization, Volume 83, Number 3, Page 178-185.

3. Aparajita Dasgupta and Rivu Basu-2011, Determinants of low birth weight in a block of hooghly, west bengal:a multivariate analysis, International Journal of Biological & Medical Research Volume 2, Number 4, Page 838-842.

4. Biswanath Das,2001-2004,Development of a Screening Tool for prediction of Low Birth weight babies by community Health Workers, sponsored by the Dept. of Family Welfare, Ministry of Health and Family Welfare, Govt. of India.

5. Mali Abdollahian and Nadeera Gunaratne-2015, Low Birth weight prediction based on maternal and fetal characteristics, 12th International Conference on Information Technology-New Generations, Page 646-650.

6. Sable M.R and Herman A.A-1997, The relationship between prenatal health behavior advice and low birth weight, Public Health Rep, Volume 112, Number 4, Page 332-339.
7. Feng M, Wan L, Li Z, Qing L, Qi X. Fetal weight estimation via ultrasound using machine learning. IEEE Access. 2019;7:87783–87791. doi: 10.1109/ACCESS.2019.2925803.

8. Trujillo OC, Perez-Gonzalez J, Medina-Bañuelos V. Early prediction of weight at birth using support vector regression. IFMBE Proc. 2020;75:37–41. doi: 10.1007/978-3-030-30648-9_5.

9. Senthilkumar, D., & Paulraj, S. Prediction of low birth weight infants and

its risk factors using data mining techniques, pp 186–194 (2015).

10. Yarlapati, A.R., Roy Dey, S., & Saha, S. Early prediction of LBW cases via minimum error rate classifier: A statistical machine learning approach, pp 1–6. 10.1109/SMARTCOMP.2017.7947002 (2017).

****