
**ISSN Print:** 2664-8350
**ISSN Online:** 2664-8369
**Impact Factor:** RJIF 5.26
IJPN 2025; 7(1): 14-19
www.pediatricsjournal.net
Received: 15-10-2024
Accepted: 21-11-2024

**Archana Maju**
MSc Nursing, Faculty, Rajkumari Amrit Kaur college of Nursing, DGHS, Ministry of Health and Family Welfare, Lajpat Nagar, New Delhi, Delhi, India

**Sarita Shokandha**
MSc Nursing, Associate Professor, Rajkumari Amrit Kaur college of Nursing DGHS, Ministry of Health and Family Welfare, Lajpat Nagar, New Delhi, Delhi, India

**Dr. Sugandha Arya**
MD (Pediatrics) Professor, Paediatrics, DGHS, Ministry of Health and Family Welfare Safdarjung Hospital, New Delhi, Delhi, India

**Corresponding Author:**
**Archana Maju**
MSc Nursing, Faculty, Rajkumari Amrit Kaur college of Nursing, DGHS, Ministry of Health and Family Welfare, Lajpat Nagar, New Delhi, Delhi, India

# AI driven predictive model for low birth weight among neonates

**Archana Maju, Sarita Shokandha and Sugandha Arya**

**DOI:** https://doi.org/10.33545/26648350.2025.v7.i1a.115

## Abstract
**Background and Objectives:** Low birth weight (LBW) is a critical indicator of maternal health and prenatal care quality worldwide. This study aimed to develop a predictive model using artificial intelligence (AI) to identify risk factors contributing to LBW.
**Methods:** A case-control study design was adopted, comparing 100 postnatal mothers with LBW neonates (cases) to 200 postnatal mothers with normal-weight neonates (controls). Logistic regression, enhanced by AI, was employed to develop the predictive model.
**Results:** The logistic regression model identified significant risk factors for LBW, including inadequate weight gain during pregnancy (<9 kg, $p < 0.001$), fetal complications during pregnancy ($p < 0.001$), maternal height <145 cm (p = 0.001), gestational age <37 weeks (p = 0.034), multiple pregnancies (p = 0.044), and maternal weight <45 kg (p = 0.048). The model demonstrated a high accuracy of 90%, with an AUC of 0.91, indicating excellent discriminatory power to distinguish between LBW and normal-weight neonates.
**Conclusion:** Most identified risk factors are modifiable through regular prenatal care and targeted interventions. Integrating this AI-driven predictive model into hospital information systems and public health programs enables early identification and proactive management of risk factors, significantly reducing LBW incidence and improving neonatal outcomes.

**Keywords:** Risk factors, low birth weight, predictive model, logistic regression model

## Introduction
Low birth weight (LBW), defined as a birth weight of less than 2,500 grams, this condition is classified further into Very Low Birth Weight (VLBW) for infants weighing less than 1,500 grams and Extremely Low Birth Weight (ELBW) for those weighing less than 1,000 grams [1]. It serves as an important indicator of a newborn's nutrition, health, and chances of survival. LBW acts as a crucial biomarker, that results potential negative health and developmental outcomes both in early life and later on, such as cognitive and behavioural impairments, growth delays, neurological issues, and an increased risk of chronic diseases [2]. Additionally, infants with LBW face a significantly higher risk of mortality compared to those with normal birth weights. Neonatal mortality is a critical public health issue worldwide, largely due to preventable factors affecting vulnerable newborns, such as premature birth and birth complications like asphyxia and trauma. These issues are exacerbated in infants with low birth weight (LBW), who are particularly at risk due to their underdeveloped organs and smaller size. Many factors contributing to low birth weight are preventable [3]. By implementing targeted healthcare interventions and providing comprehensive prenatal care, it is possible to significantly reduce the incidence of low birth weight and improve neonatal outcomes globally. Newborns with LBW encounter immediate challenges such as respiratory distress syndrome, difficulties in regulating body temperature, feeding issues, and infections. These infants are also at a greater risk of experiencing developmental delays, learning disabilities, and chronic health issues later in life. The care for LBW infants requires specialized attention in neonatal intensive care units (NICU), with a focus on essential functions like breathing, feeding, and temperature regulation. Nutritional support, including breast milk supplements or specialized formulas, is crucial for their growth and development. Preventative measures for LBW involve comprehensive prenatal care and broader public health initiatives that tackle social determinants like poverty and healthcare accessibility [4-7].

## Materials and Methods

The research setting selected for the present study was a tertiary hospital in India. Post-natal wards, nurseries, and KMC wards were selected as the settings for the study to evaluate risk factors associated with low birth weight. The study was conducted from December 2023 to July 2024.The institutional ethics committee approved the study (IEC/VMMC/SJH/Project/2023-11/400).

The sample was selected through purposive sampling. The inclusion criteria for selecting sample subjects were the post natal mothers admitted in hospital after labour and their newborns and mothers who delivered through normal, LSCS and instrumental labour. Exclusion criteria were mothers who had delivered stillborn babies while assessing the risk factors of low birth weight among neonates.

An unmatched case-control study design was employed for this research to investigate various risk factors for Low birth weight thereby identifying various predictors of Low birth weight among neonates. Data was gathered from 300 postnatal mothers and from their records, out of which 100 were postnatal mothers and their neonates with low birth weight as cases, and 200 were postnatal mothers and their neonates with normal birth weight as controls. A structured risk assessment Performa was used for assessing risk factors which includes socio demographic, maternal factors. A record review Performa was also used to assess maternal, health care and fetal factors that contributes to Low birth weight.

The Independent variable was the risk factors for Low birth weight and they were categorized as Socio demographic factors, Maternal factors, Fetal factors and Health care factors. Socio demographic factors include Age of mother, Religion, Current state of residence, Education, Occupation, Type of family, Area of dwelling, Monthly income and Dietary pattern. Maternal factors include categories related to their present pregnancy, mental well-being, personal habits, obstetrical history, physical and physiological parameters in various trimesters during pregnancy that include height, weight, weight gain during pregnancy, various investigations performed during pregnancy etc. Health care factors include registration of the present pregnancy, first visit of antenatal mother in health care setting, total number of antenatal visits performed by mother, visit performed in different trimesters in antenatal clinics, ASHA assistance during pregnancy etc. The Dependent variable was the birth weight among neonates.

The nature and purpose of the study were communicated to the participants. Consent was obtained in writing from the participants, ensuring their confidentiality and anonymity throughout and after the study. A structured interview schedule and a record review Performa were employed to collect data from the participants. Antenatal records, including investigation reports and postnatal records were utilized to support the data collection, with the help of a structured record review Performa.

The predictive model development process began with data collection, where information on various risk factors influencing birth weight—such as socio-demographic, maternal, fetal, and healthcare factors—was gathered directly from reliable sources such as post natal mothers and health records. This was followed by data preprocessing, which involved converting categorical variables into numerical formats (e.g., one-hot encoding), normalizing numerical variables for uniform contribution, and addressing issues like multicollinearity, statistical insignificance through feature selection. The prepared dataset was subsequently divided into training and testing subsets, adhering to the conventional 80/20 split ratio. Specifically, 240 samples (80%) were allocated for model training, while the remaining 60 samples (20%) were reserved for testing. This stratification enables the evaluation of the model's generalizability and performance on unseen data, thereby ensuring the reliability and validity of the predictive outcomes.

The logistic regression model was trained using the processed training data, targeting low birth weight as a binary outcome (1 for low birth weight, 0 for normal birth weight). The model's performance was evaluated on the test dataset using metrics like accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC). These metrics assessed its ability to differentiate between low and normal birth weight cases. Model interpretation focused on understanding the significance of each predictor variable by analysing logistic regression coefficients and identifying the most influential features. Finally, the findings were summarized, highlighting the model's strengths, limitations, and potential applications, along with recommendations for future research or improvements.

## Results

The logistic regression model identified eight significant predictors of low birth weight among neonates (Table 1). Factors such as Inadequate weight Gain during pregnancy <9kg (p=<0.001), Fetal complications reported during pregnancy (p=<0.001), Height of the mother <145cm (p=<0.001), Gestational age <37 weeks (p=0.028), multiple pregnancy(p=0.044) and Weight of the mother <45Kg (p=0.048) have p-values less than 0.05, showed positive co-efficient with Low birth weight indicating these factors might contribute to Low birth weight. On the other hand, joint family and urban dwellings showed negative co-efficient with Low birth weight, suggesting that these factors associated with a lower risk for Low birth weight among neonates (Table 1).

The model demonstrates a high overall accuracy of 90%. For the control group, precision is 0.88, indicating that when it predicts an instance to be a control, it is correct 88% of the time. The recall of 0.97 for the control group shows that it correctly identifies 97% of all actual controls. For the case group, precision is higher at 0.95, meaning the model is correct 95% of the time when it predicts a case. However, the recall is only 0.79, which means it only correctly identifies 79% of all actual cases. The F1-score, which is the harmonic mean of precision and recall, is 0.92 for controls and 0.86 for cases, suggesting that the model is better at identifying control instances than case instances overall. The AUC for this ROC curve is 0.91 (Figure 1), which is quite good and suggests that the classifier has a high ability to distinguish between the positive (low birth weight) and negative classes (Normal weight) (Table 2).

**Table 1:** Predictive Model for Low-Birth-Weight N=240

| S. No | Features | Co-efficient | St. d error | P Value |
|---|---|---|---|---|
| 1 | Inadequate weight gain during pregnancy <9 kg | 2.8181 | 0.67 | <0.001 |
| 2 | Fetal complications reported during pregnancy | 2.5026 | 0.668 | <0.001 |
| 3 | Height of the mother <=145 cm | 3.8529 | 1.204 | 0.001 |
| 4 | Joint family | -1.5907 | 0.724 | 0.028 |

| | | | | |
|---|---|---|---|---|
| 5 | Gestational age <37 weeks | 1.2518 | 0.59 | 0.034 |
| 6 | Residing in urban set up | -1.1017 | 0.546 | 0.043 |
| 7 | Multiple Pregnancy | 3.7175 | 1.849 | 0.044 |
| 8 | Weight of the mother <45 kg | 1.2634 | 0.64 | 0.048 |
| 9 | Obstetrical complications reported during pregnancy | 0.9835 | 0.598 | 0.1 |
| 10 | Co- morbidities of mother | -1.7207 | 1.054 | 0.103 |
| 11 | High blood pressure (>140/90 mm of Hg) | 2.6182 | 1.706 | 0.125 |
| 12 | Consanguineous marriage | -1.5458 | 1.012 | 0.127 |
| 13 | History of preterm labour | 1.8194 | 1.201 | 0.13 |
| 14 | Unplanned Pregnancy | 0.734 | 0.591 | 0.215 |
| 15 | Access of health services (>5km) | -0.9203 | 0.849 | 0.278 |
| 16 | First birth order of the Newborn | 0.7303 | 0.69 | 0.29 |
| 17 | History of miscarriages | -0.6386 | 0.688 | 0.353 |
| 18 | Lack of adequate sleep during pregnancy <8 hrs | 0.607 | 0.67 | 0.365 |
| 19 | Feel anxious, fearful, or stressed because of the emotional threat by partner | 0.5478 | 0.702 | 0.435 |
| 20 | Dietary pattern: Non vegetarian | -0.6111 | 0.805 | 0.448 |
| 21 | Income of the family <Rs 15000 | 0.3766 | 0.614 | 0.54 |
| 22 | Exposure to passive smoking | 0.4184 | 0.691 | 0.545 |
| 23 | Medications used in pregnancy other than supplements | -0.4592 | 0.78 | 0.556 |
| 24 | Consumption of extra protein supplements during pregnancy | -0.4505 | 0.889 | 0.612 |
| 25 | Total number of Antenatal visits performed | 0.3948 | 0.937 | 0.673 |
| 26 | Age of the mother <20 years or >35 years | -0.295 | 0.747 | 0.693 |
| 27 | Sex of the newborn | -0.1952 | 0.508 | 0.701 |
| 28 | Lack of health care services and assistance by ASHA | 0.1998 | 0.609 | 0.743 |
| 29 | Low Blood pressure (<110/65 mm of Hg) | 0.1901 | 0.81 | 0.814 |
| 30 | Haemoglobin level <11gm/dl | 0.1022 | 0.574 | 0.859 |
| 31 | Separate room for kitchen | -0.0684 | 0.598 | 0.909 |
| 32 | Non consumption of Iron and calcium supplements (<3Months) | -0.0421 | 0.676 | 0.95 |

**Table 2:** Validity of Predictive Model N=60

| Accuracy: 0.9 | | | | **60** |
|---|---|---|---|---|
| **Confusion Matrix:** | | | | |
| Control | 35 | 1 | | 36 |
| Case | 5 | 19 | | 24 |
| **Classification Report:** | | | | |
| | Precision | Recall | F1-Score | Support |
| Control | 0.88 | 0.97 | 0.92 | 36 |
| Case | 0.95 | 0.79 | 0.86 | 24 |

### Roc Curve

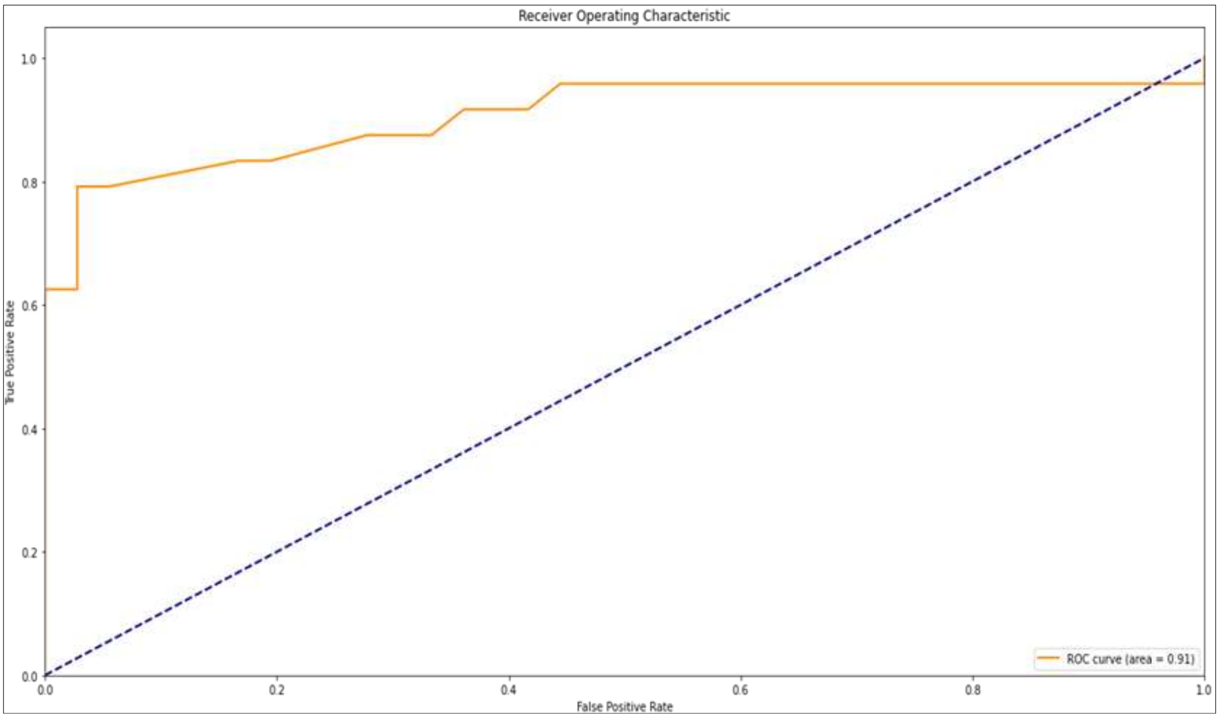


**Fig 1:** ROC curve showing the area of developed predictive model

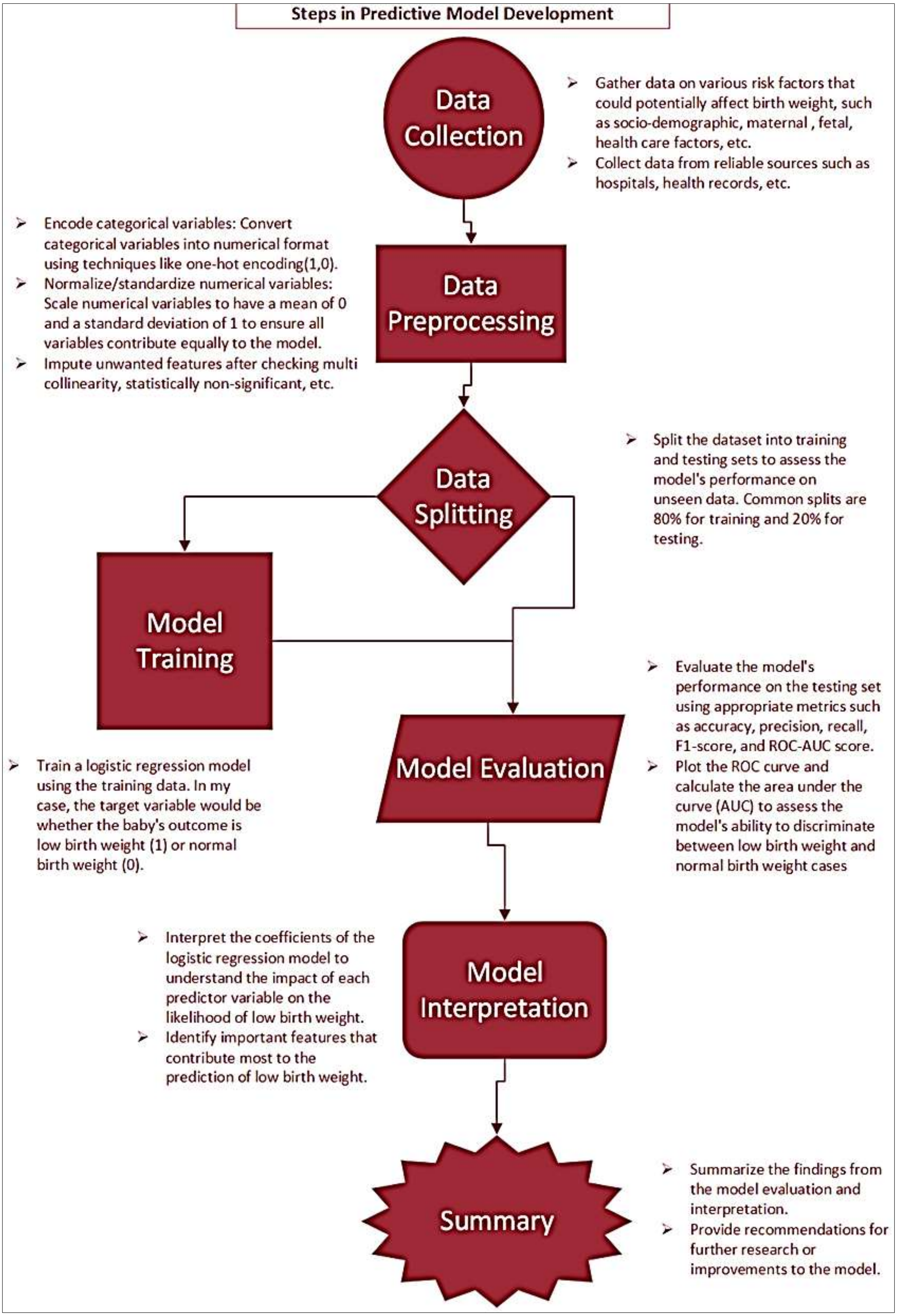


**Fig 2:** Process Of Predictive Model development using Artificial Intelligence

## Discussion

In this study, logistic regression was employed to develop a predictive model for the early identification of low birth weight during the antenatal period. Patterson *et al*. (2023) reported that logistic regression performed best among various predictive models, achieving an area under the curve (AUC) of 0.72 [8]. In comparison, the AUC of the present study was 0.91, indicating a high discriminative ability to differentiate between low birth weight and normal weight neonates. This superior performance suggests that the developed model could be a valuable tool for improving early detection and management of low-birth-weight cases.

The accuracy of the model developed in this study was 90%, which is notably high. Similarly, Arayeshgari *et al* (2023). reported the development of predictive models with accuracies of 87% or higher [9]. The findings of the present study demonstrate a comparable or superior level of accuracy, highlighting the model's potential utility in early identification of low-birth-weight cases.

AI predictive model represents a novel and emerging approach, the number of published studies in this area remains limited. Despite this, findings from existing research provide valuable insights into the predictors of low birth weight (LBW). For instance, Reza *et al*. (2024) identified maternal weight and multiple pregnancies as significant predictors of LBW, highlighting the importance of maternal health and pregnancy context in influencing neonatal outcomes [14]. Similarly, Ranjbar *et al*. (2023) reported that prematurity is a critical risk factor for LBW, underscoring the role of gestational age in determining birth outcomes [10]. These findings align with the results of the present study, which further validate the utility of AI-based predictive models in identifying key risk factors for LBW. Some studies have highlighted significant factors influencing outcomes, including maternal age, height, body mass index (BMI), the education level of the household head, antenatal care, maternal race, payment source, the number of pre-delivery emergency department visits and inpatient hospitalizations, pre-delivery disease profiles, components of the social vulnerability index, age at first sexual intercourse, and birth order number [11,12,14,18,19]. Conversely, other studies have identified additional factors such as primiparity, maternal education, pregnancy-induced hypertension, chorioamnionitis, and the use of antenatal steroids [15, 16]. This variation in findings underscores the multifactorial nature of maternal and neonatal outcomes, suggesting the need for further research to identify context-specific determinants. However, the limited number of studies emphasizes the need for further research to expand the evidence base and enhance the generalizability of AI-driven insights in this domain.

In conclusion, this study highlights the multifactorial nature of LBW, with significant contributions from socioeconomic, nutritional, obstetric, environmental, and psychosocial factors. Addressing these risk factors through targeted interventions and improved antenatal care can significantly reduce the burden of LBW and improve neonatal outcomes. The study used advanced AI methods for prediction, ensuring accurate and reliable results through careful testing and validation.

## Strength and limitations of the study

The strength of the study lies in its comprehensive approach, combining a thorough literature review with robust statistical analysis to identify significant risk factors for low birth weight. It employs advanced AI techniques for predictive modelling, ensuring high accuracy and reliability through rigorous training, testing, and validation processes. The developed predictive model can be incorporated with the information system of the hospital or health care system and can be used to identify the risk factors for low birth weight early in the antenatal period thereby the health care providers can implement the targeted interventions. This will be helpful for reducing the incidence of Low birth weight. The limitations identified was participants may forget or selectively recall past events, leading to recall bias and misclassification of exposure status. Record reviews were based on data from pre-existing records.

## Conclusion

Many risk factors were found to be modifiable with consistent prenatal care and interventions. The predictive model proved highly accurate in categorizing babies' birth weights as either normal or low.


## Acknowledgment

The authors wish to sincerely thank the authorities of the hospital for permitting to conduct the study and all the participants for their participation in the study and their cooperation. Also, we acknowledge Mr. Maju Alexander for his contribution in providing the technical help for building the Predictive model for Low birth weight using Artificial Intelligence.


## Conflict of Interest

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## References


1. ICD10Volume2_en_2010.pdf. (n.d.). [cited 2024 Feb 21]. Available from: https://icd.who.int/browse10/Content/statichtml/ICD10Volume2_en_2010.pdf
2. Too many babies are born too small [Internet]. [cited 2024 Feb 21]. Available from: https://www.who.int/news/item/16-05-2019-too-many-babies-are-born-too-small
3. 1 in 7 babies worldwide born with a low birth weight—The Lancet Global Health, UNICEF, WHO. (n.d.). [cited 2024 Mar 2]. Available from: https://www.unicef.org/press-releases/1-7-babies-worldwide-born-low-birthweight-lancet-global-health-unicef-who
4. Newborn mortality. (n.d.). [cited 2024 May 1]. Available from: https://www.who.int/news-room/fact-sheets/detail/newborn-mortality
5. Low birth weight. (n.d.). [cited 2024 May 1]. Available from: https://www.who.int/data/nutrition/nlis/info/low-birth-weight
6. Global nutrition targets 2025: Low birth weight policy brief. (n.d.). [cited 2024 Feb 21]. Available from: https://www.who.int/publications-detail-redirect/WHO-NMH-NHD-14.5
7. Low birthweight. (n.d.). UNICEF DATA. [cited 2024 May 23]. Available from: https://data.unicef.org/topic/nutrition/low-birthweight/
8. Patterson JK, Thorsten VR, Eggleston B, Nolen T,

Lokangaka A, Tshefu A, *et al*. Building a predictive model of low birth weight in low- and middle-income countries: A prospective cohort study. BMC Pregnancy Childbirth. 2023;23(1):600. doi:10.1186/s12884-023-05866-1

9. Arayeshgari M, Najafi-Ghobadi S, Tarhsaz H, Parami S, Tapak L. Machine Learning-based Classifiers for the Prediction of Low Birth Weight. Healthc Inform Res. 2023;29(1):54-63. doi:10.4258/hir.2023.29.1.54
10. Ranjbar A, Montazeri F, Farashah MV, Mehrnoush V, Darsareh F, Roozbeh N. Machine learning-based approach for predicting low birth weight. BMC Pregnancy Childbirth. 2023;23(1):803. DOI:10.1186/s12884-023-06128-w
11. Alabbad DA, Ajibi SY, Alotaibi RB, Alsqer NK, Alqahtani RA, Felemban NM, *et al*. Birthweight Range Prediction and Classification: A Machine Learning-Based Sustainable Approach. Mach Learn Knowl Extr. 2024;6(2):2. doi:10.3390/make6020036
12. Rubaiya Mansur M, Alam MM, Rayhan MI. Unraveling birth weight determinants: Integrating machine learning, spatial analysis, and district-level mapping. Heliyon. 2024;10(5):e27341. DOI:10.1016/j.heliyon.2024.e27341
13. Reza T, Salma N. Prediction and Feature selection of Low Birth Weight using Machine Learning Algorithms. [cited 2024 Mar 21]. Available from: https://doi.org/10.21203/rs.3.rs-3972884/v1
14. Ren Y, Wu D, Tong Y, López-DeFede A, Gareau S. Issue of Data Imbalance on Low Birthweight Baby Outcomes Prediction and Associated Risk Factors Identification: Establishment of Benchmarking Key Machine Learning Models With Data Rebalancing Strategies. J Med Internet Res. 2023;25(1):e44081. DOI:10.2196/44081
15. Cho H, Lee EH, Lee KS, Heo JS. Machine learning-based risk factor analysis of adverse birth outcomes in very low birth weight infants. Sci Rep. 2022;12(1):12119. doi:10.1038/s41598-022-16234-y
16. Do HJ, Moon KM, Jin HS. Machine Learning Models for Predicting Mortality in 7472 Very Low Birth Weight Infants Using Data from a Nationwide Neonatal Network. Diagnostics. 2022;12(3):625. DOI:10.3390/diagnostics12030625
17. Bekele WT. Machine learning algorithms for predicting low birth weight in Ethiopia. BMC Med Inform Decis Mak. 2022;22(1):232. DOI:10.1186/s12911-022-01981-9
18. Pollob SMAI, Abedin MM, Islam MT, Islam MM, Maniruzzaman M. Predicting risks of low birth weight in Bangladesh with machine learning. PLoS One. 2022;17(5):e0267190. DOI:10.1371/journal.pone.0267190
19. Prediction and Feature selection of Low Birth Weight using Machine Learning Algorithms. (n.d.). [cited 2024 May 23]. Available from: http://ouci.dntb.gov.ua/en/works/7AKZEv1l/

**How to Cite This Article**
Maju A, Shokandha S, Arya S. AI driven predictive model for low birth weight among neonates. International Journal of Pediatrics and Neonatology. 2025; 7(1): 14-19.