Victor et al. BMC Pregnancy and Childbirth (2025) 25:320
https://doi.org/10.1186/s12884-025-07351-3

BMC Pregnancy and Childbirth

RESEARCH

Open Access

Predicting low birth weight risks in pregnant women in Brazil using machine learning algorithms: data from the Araraquara cohort study

Audêncio Victor[1,4*], Francielly Almeida[2], Sancho Pedro Xavier[3] and Patrícia H.C. Rondó[1]

Abstract

Background Low birth weight (LBW) is a critical factor linked to neonatal morbidity and mortality. Early prediction is essential for timely interventions. This study aimed to develop and evaluate predictive models for LBW using machine learning algorithms, including Random Forest, XGBoost, Catboost, and LightGBM.

Methods We analyzed data from 1,579 pregnant women enrolled in the Araraquara Cohort, a population-based longitudinal study. Predictor variables included maternal sociodemographic, clinical, and behavioral factors. Four ML algorithms Random Forest, XGBoost, CatBoost, and LightGBM, were trained using an 80/20 train-test split and 10-fold cross-validation. To address class imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. Model performance was assessed using metrics such as area under the receiver operating characteristic curve (AUROC), F1-score, and precision-recall. Variable importance was evaluated using Shapley values.

Results XGBoost demonstrated the best performance, achieving an AUROC of 0.94, followed by CatBoost (0.94), Random Forest (0.94), and LightGBM (0.94). Maternal gestational age was the most influential predictor, followed by marital status and prenatal care frequency. Behavioral factors, such as physical activity, also contributed to LBW risk. Shapley analysis provided interpretable insights into variable contributions, supporting the clinical applicability of the models.

Conclusion Machine learning, combined with SMOTE, proved to be an effective approach for predicting LBW. XGBoost stood out as the most accurate model, but Catboost and Random Forest also provided solid results. These models can be applied to identify high-risk pregnancies, improving perinatal outcomes through early interventions.

Keywords Low birth weight, Machine learning, XGBoost, Random forest, Araraquara cohort

*Correspondence:
Audêncio Victor
audenciovictor@gmail.com
[1]School of Public Health, University of São Paulo (USP), Faculdade de Saúde Pública- USP Avenida Doutor Arnaldo, 715 – São Paulo, São Paulo 01246904, Brazil
[2]Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, FEA-RP/USP, Ribeirão Preto, São Paulo, Brazil
[3]Institute of Collective Health, Federal University of Mato Grosso. Cuiabá, Mato Grosso, Brazil
[4]Department of Nutrition, Ministry of Health of Mozambique, Maputo, Mozambique

Introduction

Birth weight is one of the primary determinants of newborn survival chances [1, 2]. Low birth weight (LBW), defined as the birth of a baby weighing less than 2,500 g, is a significant public health challenge associated with increased risk of infant mortality, neonatal morbidity, and chronic diseases in childhood and adulthood [3]. Several factors are considered in evaluating birth quality, including maternal nutritional status, prenatal care, and sociodemographic characteristics [4].

Recently, there has been an increase in LBW prevalence, predisposing children to respiratory diseases, growth retardation, heart conditions, and diabetes mellitus [5], LBW can result from preterm birth or inadequate fetal growth, often described as small for gestational age (SGA) to distinguish between immature infants and those with insufficient prenatal growth [6]. Intrauterine growth composition can influence the risk of cardiometabolic diseases in newborns, affecting both small and large for gestational age infants. The compensation of intrauterine growth through postnatal recovery or reduction may result in adverse outcomes [7]. Various factors are identified as predictors of fetal growth, including intrauterine growth retardation, unfavorable socioeconomic conditions, inadequate prenatal care, low maternal education, maternal nutritional status, marital status, ethnicity/race, maternal weight, adolescent or advanced maternal age, urinary infections, and complications such as preeclampsia and bleeding during pregnancy [4, 5].

In maternal and child health, artificial intelligence (AI), including machine learning (ML) algorithms, has been applied for outcome prediction and monitoring in perinatal health, offering new approaches for predictive modeling, diagnosis, early detection, and monitoring in perinatal health [8]. Machine learning is a subfield of AI aimed at extracting knowledge from large amounts of data, where algorithms are trained from previous examples [9]. This field has been on the rise in recent years due to the exponential increase in structured and unstructured data, also known as Big Data (BD), making ML approaches increasingly important in data analysis as traditional methods often rely on unrealistic assumptions [10, 11].

In the context of maternal and child health, mobile health (mHealth) emerges as a promising AI application, particularly useful in prenatal care in low-resource settings [12]. The application of these algorithms can enhance the precision and reliability of predictions, contributing to early prevention and intervention strategies for fetal growth issues.

Despite these advancements, most studies on LBW prediction using ML have been conducted in high-income countries, which limits their applicability to low- and middle-income settings like Brazil [13, 14]. Furthermore, many existing models lack interpretability, a crucial factor for clinical adoption, and fail to address class imbalance issues, which are common in LBW datasets [15–17]. The Araraquara cohort offers a unique opportunity to overcome these limitations, providing rich data on maternal, clinical, and behavioral predictors in a low-resource setting [18, 19].

The aim of this study is to develop a predictive model for LBW in pregnant women using machine learning algorithms. By incorporating advanced techniques such as SMOTE and Shapley values, we aim to improve the accuracy and interpretability of predictions, thereby enabling more effective strategies to address fetal growth problems early and improve neonatal outcomes.

Methods

This is a cohort study using data from the original longitudinal population-based cohort study conducted in Araraquara, São Paulo, Brazil, titled "Araraquara Cohort" [18]. The sample included women with a gestational age of ≤ 19 weeks who received prenatal care at Basic Health Units in Araraquara, São Paulo, Brazil. The pregnant women were followed quarterly throughout their pregnancy until the birth of their children between 2017 and 2022. Women with twin pregnancies and those who experienced miscarriage were excluded. In cases of fetal and stillbirth, only data from the pregnancy were considered. The data used in this study were accessed between May 2024 and July 2024.

The outcome of low birth weight was analyzed based on the dichotomous classification of birth weight, defined as low birth weight: < 2500 g and normal weight: ≥ 2500 g. The predictor variables are illustrated in Table 1.

The study was approved by the Research Ethics Committee with Human Subjects at the School of Public Health, University of São Paulo, prior to data collection, under protocol number CAEE: 59787216.2.0000.5421, opinion number 1.885.874. All participants signed the Informed Consent Form before participating. The participants were informed about the objectives of the study, the associated risks, and benefits, ensuring that their participation was entirely voluntary. All participants provided informed consent, consistent with the principles outlined in the Helsinki Declaration.

Model design

As shown in Fig. 1, regarding the fetal growth outcome LBW, different model specifications were tested to evaluate whether changes in strategy improved model performance. Models for the outcome were estimated independently, without information sharing between them. Before discussing the models, quantitative variables were normalized using the z-score separately for the training and test sets. All qualitative variables were

Table 1 Description of model variables

Variable	Description	Category
Maternal Age	Age of the mother at the time of pregnancy.	In years
Marital Status	Mother's marital status.	Married; Stable union; Single, widow, or separated
Maternal Education (in years)	Mother's completed years of education.	None; 1–7; 8–11; 12 or more
Race/Color	Mother's ethnicity or color.	White; Black; Asian; Mixed; Indigenous
Nutritional Status	Physical evaluation of the pregnant woman.	Anthropometry: BMI (kg/m^2): Underweight, Normal, Overweight, Obesity; Arm circumference (cm): Underweight (< 23 cm), Normal (25–28 cm), Overweight or Obese (≥ 28 cm); Body fat percentage
Household Data	Information about the gestational home.	Number of household members per room (terciles)
Previous Pregnancies	Number of previous pregnancies.	0; 1; ≥ 2
Lifestyle	Pregnant woman's lifestyle habits.	Physical activity; Smoking; Alcohol consumption
Morbidity	Pre-existing medical conditions.	Diabetes; Hypertension; Urinary tract infection; Cervicitis/Vaginitis
Other Relevant Predictors	Other relevant information.	Gestational age at birth; Glycemic profile: Fasting glucose (mg/dL), Insulin (μIU/mL), HOMA (μIU/mL), Glycated hemoglobin (%); C-reactive protein (ng/mL); Hemoglobin (g/dL); Lipid profile: Total cholesterol, LDL-c, HDL-c, and triglycerides (mg/dL)

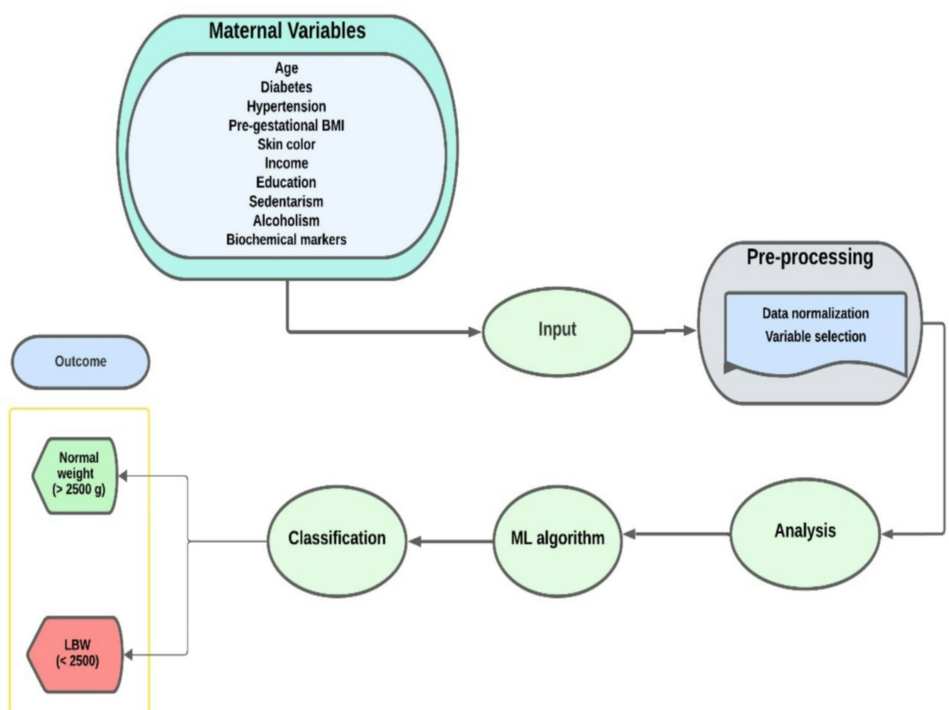

Fig. 1 An example of the workflow diagram for classifying birth weight adequacy

treated through one-hot encoding, where each category was separately considered for this procedure. Additionally, pregnant women were excluded due to missing information, and variables with missing data below 20% were imputed by the mean.

In this study, four different machine learning algorithms were tested: Catboost [20], Xgboost [21], Lightgbm [22] and Random Forest. For Catboost, XGBoost, and LightGBM, were used [23]. The data analysis was conducted using Python (version 3.9), with key libraries including pandas (1.3.5) and numpy (1.21.4) for data preprocessing, and scikit-learn (1.0.2). Additionally, the official libraries for CatBoost (1.0.4), XGBoost (1.5.1), and LightGBM (3.3.2) were used for these respective algorithms.

The key hyperparameters for CatBoost, XGBoost, and LightGBM included a learning rate of 0.1, a maximum depth of 10, and 500 estimators. These parameters were fine-tuned using Bayesian optimization, which was applied across 50 iterations. For Random Forest, the number of trees (n_estimators) tested ranged from 100 to 1000 at evenly spaced intervals, with maximum depths varying from 5 to 20 and features considered for splits (max_features) set to either 'log2' or 'sqrt'. Hyperparameter optimization for Random Forest was performed using RandomizedSearchCV with 30 iterations and a random seed of 42 for reproducibility [24]. In cases of significant class imbalance, where the minority class represents less than 25% of the total outcomes, the Synthetic Minority Over-sampling Technique (SMOTE) was applied. Additionally, the Boruta method was employed for feature selection [25]. The performing models in the training set (80% of the data) were selected for evaluation in the test set (20%). The evaluation of ML algorithms was conducted in the test set using metrics such as area under the ROC curve (AUC-ROC), Matthew's correlation coefficient (MCC), precision, recall, positive predictive value, negative predictive value, and F1-score. Additionally, the performance of the algorithms in the top 20% of high-risk patients (20% k-tops) was assessed using metrics such as true positive, false positive, precision, and recall. All analyses followed the Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD) guidelines [26].

Results

Maternal characteristics

According to Table 2, the maternal characteristics of 1,579 pregnant women from the Araraquara cohort were evaluated. The women had an average age of 28.4 years, a height of 162 cm, a pre-pregnancy body mass index (BMI) of 24.7 kg/m^2, and a gestational age of 39.3 weeks. Most women (88.4%) had an education level equal to or greater than 8 years, 53.7% were non-white, and 87.7% were married or in a stable relationship, with most having a family income of R$563 and being non-smokers. Only 5.0% of the women had diabetes, and 7.0% had hypertension. A total of 11.3% of the women had a urinary tract infection during pregnancy. CRP (C-reactive protein) levels were 3.3 mg/L (interquartile range: 1.4–7.8), and HOMA (homeostasis model assessment) values were 2.9 units (interquartile range: 1.3–6.1), shows the distribution of low birth weight in the Araraquara cohort, where 1,309 (91.2%) had normal birth weight, and 126 (8.8%) had LBW.

Performance of machine learning models

In this study, we evaluated the predictive capacity of various ML models, including Random Forest, XGBoost, LightGBM, and Catboost, to predict LBW in neonates. The analysis focused on multiple performance metrics such as AUROC, accuracy, precision, recall, F1-score, and the MCC to provide a comprehensive evaluation of the ability of these models to classify cases of low birth weight.

The results of the ML models evaluation for LBW prediction are summarized in Table 3; Fig. 2. The XGBoost model showed the best performance with an AUROC of 0.941, demonstrating excellent discrimination between neonates with normal and low birth weight. Catboost followed closely with an AUROC of 0.939, while Random Forest and LightGBM achieved AUROCs of 0.938 and 0.937, respectively.

In addition to AUROC, other performance metrics were considered. Random Forest exhibited the highest overall accuracy (0.94), while Catboost provided the best balance between precision (0.80) and recall (0.78), resulting in an F1-score of 0.79 (Table 3). The ROC curves in Fig. 3 visually illustrate the similar performance between the models, with all curves approaching the top-left corner, indicating high discriminatory capacity.

Variable importance for LBW prediction

Figure 3, shows the importance of the predictors for LBW using Shapley values for the best-performing model, XGBoost. The most important variable identified was gestational age, standing out as the factor with the most influence on predicting low birth weight. Following this, maternal marital status and the absence of regular physical activity during pregnancy were significant predictors. Other factors contributing substantially included maternal race, parity, and fewer prenatal visits, underscoring the importance of socioeconomic and behavioral variables. Variables such as smoking and alcohol consumption during pregnancy, although important predictors, appeared with less relevance compared to the previously mentioned factors. S.1, illustrates the strength of variable contributions to the prediction of LBW using Shapley values in the XGBoost model. The most influential variables, shown in the graphs, provide a detailed view of how each factor contributes to increasing or reducing the risk of LBW.

Discussion

The findings from this study confirm and expand the evidence that machine learning (ML) models, such as Random Forest, XGBoost, LightGBM, and Catboost, are effective for predicting low birth weight (LBW) in neonates. Among the models tested, XGBoost exhibited the best performance with an AUROC of 0.941, which places it as the most effective algorithm. This outcome aligns with previous studies that have explored ML approaches in predicting LBW [27–29]. The use of ML in LBW prediction is gaining traction, as these models demonstrate significant potential across various clinical and public health applications.

Table 2 Maternal characteristics of pregnant women in the Araraquara cohort, Brazil (2021–2023)

Variable	*N* (%)
Age (years)	
< 20	154 (9.7)
20 to 35	1,185 (74.9)
> 35	244 (15.4)
Height (cm)	
< 153	534 (33.7)
153–165	566 (35.8)
> 165	483 (30.5)
Pre-pregnancy BMI (kg/m^2)	
Underweight	64 (4.0)
Normal weight	1,084 (68.5)
Overweight	450 (28.4)
Obesity	48 (3.2)
Arm circumference (cm)	
< 23	674 (42.6)
≥ 23	919 (57.4)
Gestational age (weeks)	39.3 (38.5–40.3)
Maternal Education (years)	
≤ 8	184 (11.6)
> 8	1,395 (88.4)
Number of people per room	
≤ 1	536 (33.9)
> 1	1,043 (66.1)
Family income (R$)	
< 937	311 (32.5)
937–1,866	563 (34.6)
> 1,866	666 (33.0)
Race	
Non-white	724 (46.3)
White	833 (53.7)
Marital Status	
Single	135 (9.3)
Married or in a stable relationship	1,359 (87.7)
Separated, divorced, or widowed	198 (12.7)
Smoking	
No	1,434 (91.7)
Yes	123 (7.9)
Alcohol Consumption	
No	1,235 (79.5)
Yes	319 (20.5)
Diabetes	
No	1,479 (95.0)
Yes	78 (5.0)
Hypertension	
No	1,445 (93.7)
Yes	108 (7.0)
Urinary Tract Infection	
No	1,378 (88.7)
Yes	175 (11.3)
Vigorous Physical activity	
Adequate	319 (31.2)
Inadequate	924 (68.8)
Candidiasis/Vaginitis	

Table 2 (continued)

Variable	*N* (%)
No	1,445 (91.9)
Yes	108 (6.8)
Number of previous births	
0	764 (53.7)
≥ 1	655 (43.9)
CRP (mg/L)	3.3 (1.4–7.8)
HOMA (units)	2.9 (1.3–6.1)
Glycated Hemoglobin (%)	5.1 (4.7–5.6)
Fasting Glucose (mg/dL)	74.3 (67.3–81.6)
Cholesterol (mg/dL)	195 (168–224)
LDL (mg/dL)	104 (86–123)
HDL (mg/dL)	57 (47–68)
Triglycerides (mg/dL)	94 (65–143)

Note: Data are presented as mean and interquartile range (25th percentile – 75th percentile) or number (percentage). Abbreviations: BMI, Body Mass Index; LDL, Low-Density Lipoprotein; HDL, High-Density Lipoprotein; CRP, C-reactive Protein; HOMA, Homeostasis Model Assessment of Insulin Resistance

Table 3 Test set performance comparison for LBW weight classification models

Model	Accuracy	Precision	Recall	F1-Score	MCC	AUC
XGBoost	0.92	0.76	0.77	0.77	0.54	0.94
LightGBM	0.92	0.76	0.77	0.77	0.54	0.94
CatBoost	0.93	0.80	0.78	0.79	0.58	0.94
RandomForest	0.94	0.87	0.72	0.77	0.57	0.94

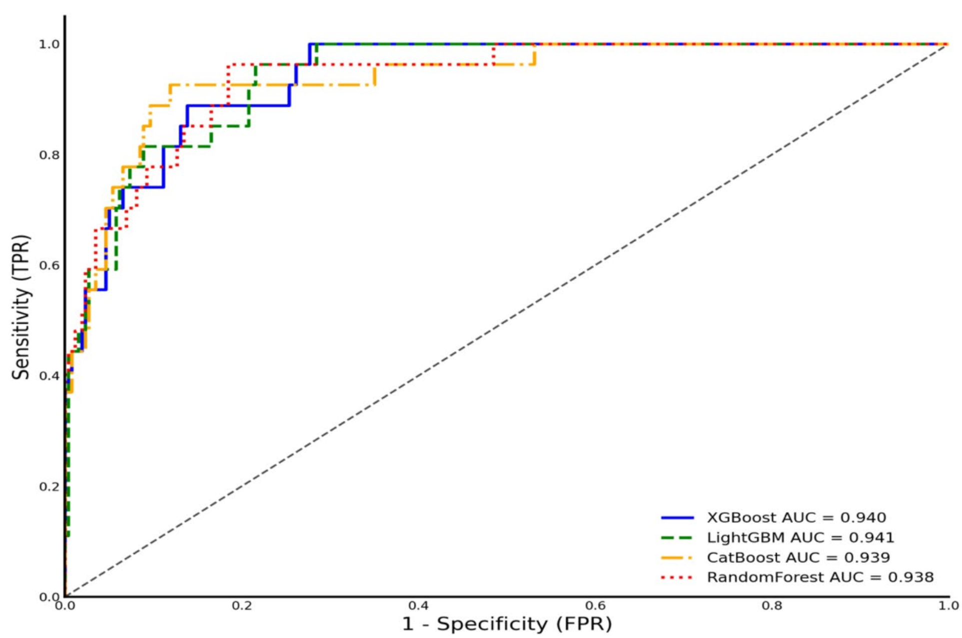

Fig. 2 Performance by ROC curve of ML (Random Forest, XGBoost, LightGBM, and Catboost) for LBW prediction

In agreement with the current literature, our study corroborates the efficacy of ML algorithms in maternal-fetal health. Previous studies have underscored the importance of selecting the most suitable ML algorithm, given the variations in performance and reliability among different models [30]. For example, a study published in BMC Pregnancy and Childbirth evaluated the performance of eight different ML algorithms for predicting LBW. It found that deep learning (AUROC: 0.86), random forest classification (AUROC: 0.79), and extreme gradient boost classification (AUROC: 0.79) performed best in distinguishing LBW from normal birth weight across a cohort of 8,853 births, where 1,280 resulted in LBW [16].

Our study further supports the finding that boosting algorithms such as XGBoost and Catboost are

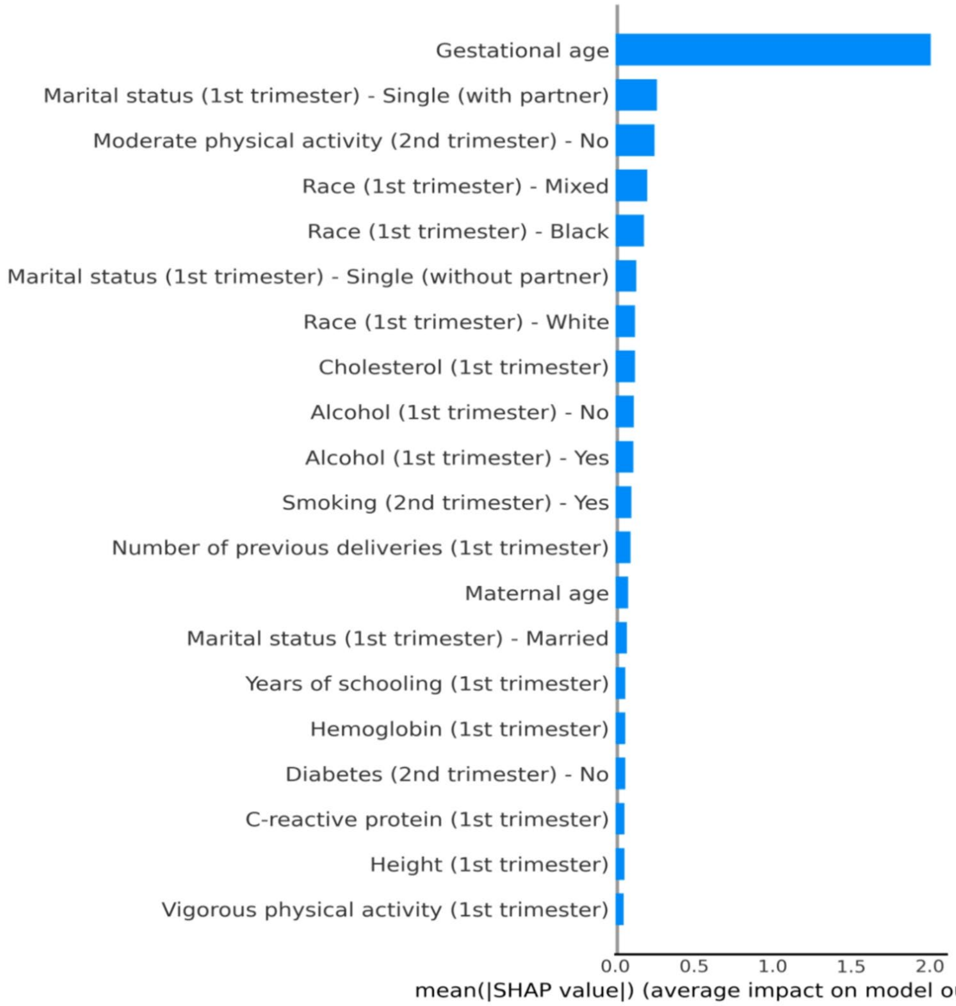

Fig. 3 Predictors of LBW (Shapley Variable Importance) for the best model - XGBoost

particularly powerful for tabular datasets, as seen in our cohort. These models outperform more traditional methods like logistic regression and are advantageous in dealing with complex non-linear relationships among predictors [31–33]. Recent studies indicate that boosting algorithms represent the state-of-the-art for tabular data. They have shown high performance across a wide range of tasks, including classification [31, 32].

Studies like that of Pollob et al. (2022), which used ML to predict LBW in Bangladesh, similarly demonstrated that ensemble and boosting models provide more accurate predictions than classical statistical models, with LBW rates around 16.2% in their study population. Key risk factors included region, education, wealth index, and height, consistent with our findings on the importance of gestational weight gain, race, and socioeconomic status in predicting LBW [34].

The use of Shapley values in our study provided additional insights into the interpretability of the models, allowing for a more granular understanding of how each variable contributes to the prediction of LBW. This aspect is particularly important in clinical applications where transparency and explainability of the models are crucial for their adoption by healthcare professionals. For example, the Shapley analysis in our study indicated that gestational weight gain had the strongest influence on LBW predictions, followed by maternal race and prenatal care visits. These findings mirror global trends in LBW prediction, where maternal health, nutrition, and prenatal care are recognized as key determinants of birth outcomes.

Our findings also align with the work of Patterson et al. (2023), who developed a predictive model for LBW in low- and middle-income countries [13]. Their study found that socioeconomic factors, including maternal education and access to prenatal care, significantly contributed to the risk of LBW, similar to our results from the Araraquara cohort. Both studies emphasize the importance of early identification of high-risk pregnancies through predictive models, especially in resource-limited settings where timely interventions can significantly improve neonatal outcomes.

The clinical implications of our findings are significant. By accurately identifying pregnancies at high risk for LBW, these ML models could enable healthcare providers to implement early interventions, such as nutritional supplementation, increased prenatal visits, or targeted counseling on lifestyle modifications. Such interventions could mitigate the risks associated with LBW, including neonatal mortality, morbidity, and long-term health consequences like developmental delays and chronic conditions. This is particularly relevant in low- and middle-income countries, where LBW rates are higher, and healthcare resources are often limited [13].

While the results of this study are promising, there are several limitations to consider. First, the cohort used in this study was from a specific region in Brazil, which may limit the generalizability of the findings to other populations. Future research should aim to validate these models across different regions and populations to ensure their broader applicability. Additionally, while our models demonstrated high predictive accuracy, further research is needed to assess their integration into clinical workflows and their potential impact on perinatal care. For example, future studies could explore the use of these models in combination with mobile health (mHealth) technologies to improve prenatal care in low-resource settings.

Moreover, although we employed robust techniques such as SMOTE to handle class imbalance and Shapley values to interpret model predictions, the models still require validation in real-world clinical environments. The practical deployment of ML models in healthcare settings involves challenges related to data privacy, model fairness, and bias mitigation, which should be thoroughly addressed before these models can be widely adopted.

Conclusion

This study successfully developed and evaluated machine learning models, including XGBoost, Catboost, LightGBM, and Random Forest, for predicting low birth weight in neonates. The XGBoost model demonstrated the highest predictive performance, with excellent discrimination between neonates at risk of LBW and those with normal birth weight. The application of these models in clinical practice has the potential to improve early detection of high-risk pregnancies, enabling timely and personalized interventions that could significantly improve neonatal outcomes.

Given the increasing global focus on maternal and neonatal health, these findings hold important implications for both clinical practice and public health policy. The integration of machine learning models into prenatal care systems could offer a transformative approach to preventing adverse birth outcomes, particularly in low-resource settings where LBW remains a critical challenge.

Supplementary information

The online version contains supplementary material available at https://doi.org/10.1186/s12884-025-07351-3.

Supplementary Material 1

Acknowledgements

The authors gratefully acknowledge the professionals, undergraduate, and graduate students who collaborated in the data collection for the Araraquara cohort.

Author contributions

AV, FA and PHCR: conceptualization, methodology. AV, SPX and PHCR: investigation. AV: Data analysis, AV, SPX: visualization and writing - original draft. AV, FA, and PHCR: supervision, writing, review, and editing. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the São Paulo Research Foundation (FAPESP) (grant number 2015/03333–6). Audêncio Victor has received a scholarship from São Paulo Research Foundation (FAPESP) (grant number 2023/07936-3).

Data availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. The code developed for constructing the algorithms along, is available on Github (https://github.com/Audency/Predictors-of-Low-Birth-Weight-using-Machine-Laerning-.git.)

Declarations

Ethics approval and consent to participate

The research received ethical approval from the Research Ethics Committee with Human Subjects at the School of Public Health, University of São Paulo, before the commencement of data collection, as per protocol CAEE: 59787216.2.0000.5421 and opinion number 1.885.874. All participants provided informed consent, consistent with the principles outlined in the Helsinki Declaration.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 17 November 2024 / Accepted: 19 February 2025
Published online: 19 March 2025

References

1. Wachamo TM, Yimer NB, Bizuneh AD. Risk factors for low birth weight in hospitals of North Wello zone, Ethiopia: A case-control study. PLoS ONE. 2019;14:1–15.
2. Sangi R, Ahsan AK, Khan AT, Aziz SN, Afroze M, Jamro S, et al. Evaluating association of maternal nutritional status with neonatal birth weight in term pregnancies: A Cross-Sectional study with unexpected outcomes. Cureus. 2021;13:1–6.
3. Lopes AAT, Tani G, Maia JAR. Desempenho neuromotor, prematuridade e Baixo peso à nascençe. Revista Brasileira De Cineantropometria E Desempenho Humano. 2011;13:73–81.
4. Christine B, de GM, Oliveira Pessoa TA, Aguiar Lima FC, Munhoz Gaíva MA. O Crescimento e desenvolvimento Frente à prematuridade e Baixo peso Ao nascer. Av En Enfermería. 2016;33:401–11.
5. Moreira AIM, de Sousa PRM, Sarno F. Low birth weight and its associated factors. Einstein (Sao Paulo). 2018;16:eAO4251.

6. Wilcox AJ, Cortese M, McConnaughey DR, Moster D, Basso O. The limits of small-for-gestational-age as a high-risk category. Eur J Epidemiol. 2021;36:985–91.
7. Gätjens I, Fedde S, Schmidt SCE, Hasler M, Plachta-Danielzik S, Müller MJ, et al. Relationship between birth weight, early growth rate and body composition in 5 to 7-year-old children. Obes Facts. 2022. https://doi.org/10.1159/000522509
8. Ramakrishnan R, Rao S, He J-R. Perinatal health predictors using artificial intelligence: A review. Womens Health (Lond). 2021;17:17455065211046132.
9. Islam MN, Islam AKMN. A systematic review of the digital interventions for fighting COVID-19: the Bangladesh perspective. Ieee Access. 2020;8:114078–87.
10. Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. Science (1979). 2015;349:255–60.
11. Loring Z, Mehrotra S, Piccini JP. Machine learning in 'big data': handle with care. EP Europace. 2019;21:1284–5.
12. Davidson L, Boland MR. Towards deep phenotyping pregnancy: a systematic review on artificial intelligence and machine learning methods to improve pregnancy outcomes. Brief Bioinform. 2021;22:bbaa369.
13. 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.
14. Sanchez-Martinez S, Marti-Castellote PM, Hoodbhoy Z, Bernardino G, Prats-Valero J, Aguado AM, et al. Prediction of low birth weight from fetal ultrasound and clinical characteristics: a comparative study between a low- and middle-income and a high-income country. BMJ Glob Health. 2024;9:e016088.
15. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority Over-sampling technique. J Artif Intell Res. 2002;16:321–57.
16. 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.
17. 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:e44081.
18. Victor A, de França da Silva Teles L, Aires IO, de Carvalho LF, Luzia LA, Artes R, et al. The impact of gestational weight gain on fetal and neonatal outcomes: the Araraquara cohort study. BMC Pregnancy Childbirth. 2024;24:320.
19. Victor A, Geremias dos Santos H, Silva GFS, Barcellos Filho F, de Fátima Cobre A, Luzia LA, et al. Predictive modeling of gestational weight gain: a machine learning multiclass classification study. BMC Pregnancy Childbirth. 2024;24:733.
20. Prokhorenkova L, Gusev G, Vorobev A, Dorogush AV, Gulin A. CatBoost: unbiased boosting with categorical features. Adv Neural Inf Process Syst. 2018;31.
21. Chen T, Guestrin C, Xgboost. A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016;785–94.
22. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W et al. Lightgbm: A highly efficient gradient boosting decision tree. Adv Neural Inf Process Syst. 2017;30.
23. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: machine learning in Python. J Mach Learn Res. 2011;12:2825–30.
24. Bergstra J, Yamins D, Cox D. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In: International conference on machine learning. PMLR. 2013;115–23.
25. Kursa MB, Rudnicki WR. Feature selection with the Boruta package. J Stat Softw. 2010;36:1–13.
26. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med. 2015;162:W1–73.
27. Schmidt LJ, Rieger O, Neznansky M, Hackelöer M, Dröge LA, Henrich W, et al. A machine-learning–based algorithm improves prediction of preeclampsia-associated adverse outcomes. Am J Obstet Gynecol. 2022;227:e771–7730.
28. Naimi AI, Platt RW, Larkin JC. Machine learning for fetal growth prediction. Epidemiology. 2018;29:290–8.
29. Włodarczyk T, Płotka S, Szczepański T, Rokita P, Sochacki-Wójcicka N, Wójcicki J, et al. Machine learning methods for preterm birth prediction: A review. Electron (Basel). 2021;10:586.
30. Islam MN, Mustafina SN, Mahmud T, Khan NI. Machine learning to predict pregnancy outcomes: a systematic review, synthesizing framework and future research agenda. BMC Pregnancy Childbirth. 2022;22:348.
31. Borisov V, Leemann T, Seßler K, Haug J, Pawelczyk M, Kasneci G. Deep neural networks and tabular data: A survey. IEEE Trans Neural Netw Learn Syst. 2022;1–21.
32. Shwartz-Ziv R, Armon A. Tabular data: deep learning is not all you need. Inform Fusion. 2022;81:84–90.
33. McElfresh D, Khandagale S, Valverde J, Ramakrishnan G, Goldblum M, White C. When Do Neural Nets Outperform Boosted Trees on Tabular Data? arXiv preprint arXiv:230502997. 2023.
34. Ashikul Islam Pollob SM, Abedin MM, Islam MT, Islam MM, Maniruzzaman M. Predicting risks of low birth weight in Bangladesh with machine learning. PLoS ONE. 2022;17:5.

Publisher's note