RESEARCH

Machine learning-based approach for predicting low birth weight

Amene Ranjbar[1], Farideh Montazeri[2], Mohammadsadegh Vahidi Farashah[3], Vahid Mehrnoush[2], Fatemeh Darsareh[2*] and Nasibeh Roozbeh[2]

Abstract

Background Low birth weight (LBW) has been linked to infant mortality. Predicting LBW is a valuable preventative tool and predictor of newborn health risks. The current study employed a machine learning model to predict LBW.

Methods This study implemented predictive LBW models based on the data obtained from the "Iranian Maternal and Neonatal Network (IMaN Net)" from January 2020 to January 2022. Women with singleton pregnancies above the gestational age of 24 weeks were included. Exclusion criteria included multiple pregnancies and fetal anomalies. A predictive model was built using eight statistical learning models (logistic regression, decision tree classification, random forest classification, deep learning feedforward, extreme gradient boost model, light gradient boost model, support vector machine, and permutation feature classification with k-nearest neighbors). Expert opinion and prior observational cohorts were used to select candidate LBW predictors for all models. The area under the receiver operating characteristic curve (AUROC), accuracy, precision, recall, and F1 score were measured to evaluate their diagnostic performance.

Results We found 1280 women with a recorded LBW out of 8853 deliveries, for a frequency of 14.5%. Deep learning (AUROC: 0.86), random forest classification (AUROC: 0.79), and extreme gradient boost classification (AUROC: 0.79) all have higher AUROC and perform better than others. When the other performance parameters of the models mentioned above with higher AUROC were compared, the extreme gradient boost model was the best model to predict LBW with an accuracy of 0.79, precision of 0.87, recall of 0.69, and F1 score of 0.77. According to the feature importance rank, gestational age and prior history of LBW were the top critical predictors.

Conclusions Although this study found that the extreme gradient boost model performed well in predicting LBW, more research is needed to make a better conclusion on the performance of ML models in predicting LBW.

Keywords Low birth weight, Fetal weight, Birth weight, Machine learning, X gradient boost model

*Correspondence:
Fatemeh Darsareh
famadarsareh@yahoo.com
[1] Fertility and Infertility Research Center, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
[2] Mother and Child Welfare Research Center, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
[3] Amirkabir University of Technology, Tehran, Iran

Background

Birth weights less than 2500 g are called low birth weight (LBW). LBW has been linked to infant mortality and its consequences [1]. Predicting LBW is thus a valuable preventative tool and predictor of newborn health risks. Previous research has found that maternal demographics, preexisting health conditions, and prenatal care level are all linked to LBW [2, 3]. Thus, pinpointing which pregnant patients are most likely to have a baby with LBW during the preconception or early pregnancy stages is

Ranjbar *et al. BMC Pregnancy and Childbirth*   (2023) 23:803

Page 2 of 7

critical for saving neonatal lives and reducing potentially avoidable medical costs through direct clinical and health policy interventions. There are some documented studies on using ML in perinatal care and maternal health. Previous LBW prediction studies achieved good performance in predicting LBW; however, all previous studies recommended more studies due to study limitations such as small sample size or limited feature selection [4–7]. In this study, we aimed to evaluate the performance of eight different ML algorithms in predicting LBW.

Methods

The findings of this retrospective cohort study are based on birth records obtained from the "Iranian Maternal and Neonatal Network (IMaN Net)," a legitimate national system, from January 2020 to January 2022. IMaN Net is a comprehensive system for registering maternal and newborn information on the outcomes of each delivery, which is completed daily by midwives in all birth centers and hospitals throughout Iran in an integrated manner. All patients' personal information was deidentified and not disclosed.

Women with singleton pregnancies above the gestational age of 24 weeks who gave birth during a study period were included. The target population in this study was divided into LBW (≤ 2499 g) and not LBW (≥ 2500 g), which is the national standard definition [8]. Exclusion criteria included multiple pregnancies and fetal anomalies.

A predictive model was built using eight statistical learning models, including logistic regression, decision tree classification, random forest classification, deep learning feedforward, extreme gradient boost classification (XGBoost), light gradient boost (LGB), support vector machine (SVM), and permutation feature classification with k-nearest neighbors (KNN). Expert opinion and prior observational cohorts were used to select candidate LBW predictors for all models [9, 10]. Predictor factors included maternal age, educational level, maternal occupation, place of residence, inadequate prenatal care (less than three prenatal care visits), smoking, drug addiction, maternal anemia, cardiovascular disease, chronic hypertension, hepatitis, COVID-19, overt diabetes, gestational diabetes and thyroid dysfunction, parity, preeclampsia, fetal gender, method of childbirth, previous history of LBW, supplementary and vitamins intake were obtained from patient medical records. We used Chi-square test to evaluate the association between predicting factors mentioned above and LBW. Then we performed ML analysis approach. We followed the Guidelines for Developing and Reporting Machine Learning Predictive Models in Biomedical Research: A Multidisciplinary View to report our findings. The programming language Python was chosen to create the machine learning model. Scikit-learn was used to implement the ML algorithm. Scikit-learn is a machine-learning library written in Python. It includes an extensive collection of cutting-edge machine-learning algorithms for both supervised (including the multi-output classification and regression algorithm) and unsupervised problems [11].

Internal validation was carried out with the help of k-fold cross-validation. The cases were randomly assigned to either the "training set" (70%) or the "test set" (30%) using a random number generator. The original dataset kept the rate of LBW and non-LBW groups in the training and test sets constant. Using the training set, we arranged the parameters of the prediction models and evaluated their performance using the "test set". The average performance was calculated by repeating these ten times.

Metrics, including area under the receiver operating characteristic curve (AUROC), accuracy, precision, recall, and F1 score, were used to assess the predictive power of the models. The accuracy metric calculates how often a model is correctly predicted across the entire dataset. Precision measures how many of the model's "positive" predictions were correct. The model's recall estimates how many positive class samples in the dataset were correctly identified. The F1 score combines precision and recall by using their harmonic mean, and maximizing the F1 score implies maximizing both precision and recall simultaneously. As a result, researchers have chosen the F1 score to evaluate their models in conjunction with accuracy. We used AUROC as the primary performance metric because it is a widely used index to describe the ML model's ability to predict outcomes. The metrics were scaled from 0 to 1, with higher values indicating a better model [12].

Results

Of 8850 eligible cases, we found 1280 women with a recorded LBW, for a frequency of 14.5%. The demographic and clinical characteristics of study population is given in Table 1. As it shown, maternal age, living residency, gestational age, parity, access to prenatal care, maternal anemia, chronic hypertension, preeclampsia, drug addiction, COVID-19, previous LBW, and newborn gender was linked to LBW.

In this study, we attempt to evaluate parameters and feature selection based on performance parameters using various ML algorithms. A plot ROC chart, as shown in Fig. 1, and calculate AUROC as a plot that allows the user to visualize the tradeoff between the classifier's sensitivity. Deep learning (AUROC: 0.86), random forest classification (AUROC: 0.79), and XGBoost classification

Ranjbar *et al. BMC Pregnancy and Childbirth*   (2023) 23:803

Page 3 of 7

Table 1 Demographic and clinical factors associated with low birth weight

Outcome	Non-LBW (*n* = 7570)	LBW (*n* = 1280)	*P*-value
Maternal age			< 0.001
13–19	137 (1.8)	36 (2.8)	
20–35	6247 (82.5)	995 (77.7)	
Above 35	1186 (15.7)	249 (19.5)	
Education			0.348
Illiterate	480 (6.3)	73 (5.7)	
Primary	2344 (31.0)	378 (29.5)	
High-school/ Diploma	3463 (45.8)	609 (47.6)	
Advanced	1283 (16.9)	220 (17.2)	
Occupation			0.299
Housewife	6798 (89.8)	1146 (89.5)	
Worker/employee	772 (10.2)	134 (10.5)	
Living residency			0.045
Urban	5056 (66.8)	823 (64.3)	
Rural	2517 (33.2)	457 (35.7)	
Gestational age			< 0.001
24–36^{+6}	429 (5.7)	801 (62.6)	
37–41	7141 (94.3)	479 (37.4)	
Parity			< 0.001
Primiparous	2056 (27.1)	439 (34.3)	
Multiparous	5517 (72.9)	841 (65.7)	
Access to prenatal care			0.030
Yes	7343 (97.0)	1255 (98.0)	
No	230 (3.0)	25 (0.2)	
Maternal anemia			0.047
No	7364 (97.2)	1233 (96.3)	
Yes	209 (2.8)	47 (3.7)	
Chronic hypertension			0.005
No	7501 (99.0)	1256 (98.1)	
Yes	72 (1.0)	24 (1.9)	
Cardiovascular disease			0.803
No	7492 (98.9)	1267 (99.0)	
Yes	81 (1.1)	13 (1.0)	
Diabetes			0.276
No	6420 (84.8)	1094 (85.5)	
Yes	1153 (15.2)	186 (14.5)	
Preeclampsia			< 0.001
No	7196 (95.0)	1083 (84.6)	
Yes	377 (5.0)	197 (15.4)	
Drug addiction			< 0.001
No	7530 (99.4)	1251 (97.7)	
Yes	42 (0.6)	29 (2.3)	
Previous low birth weight			< 0.001
No	7479 (98.8)	1089 (85.1)	
Yes	94 (1.2)	191 (14.9)	

Table 1 (continued)

Outcome	Non-LBW (*n* = 7570)	LBW (*n* = 1280)	*P*-value
COVID-19			0.020
No	7465 (98.6)	1250 (97.7)	
Yes	108 (1.4)	30 (2.3)	
Thyroid dysfunction			0.999
No	6778 (89.5)	1146 (89.5)	
Yes	795 (10.5)	134 (10.5)	
Hepatitis			0.079
No	7543 (99.6)	1279 (99.1)	
Yes	30 (0.4)	1 (0.1)	
Newborn gender			< 0.001
Male	3942 (52.1)	599 (46.8)	
Female	3631 (47.9)	681 (53.2)	
Supplementary intake			0.078
No	4 (0.1)	5 (0.4)	
Yes	7569 (99.9)	1275 (99.6)	

Data are presented as n (%)

(AUROC: 0.79) all have higher ROC_AUC and perform better than others, as shown in Fig. 1.

Other performance parameters for each algorithm are shown in Table 2. Other performance parameters indicate that the XGBoost classification performs more than all. Random forest classification and deep learning feedforward are also very close. When the accuracy, precision, recall, and F1 score of the models mentioned above with higher AUROC were compared, the XGBoost model was the best model to predict LBW with an accuracy of 0.79, precision of 0.87, recall of 0.69, and F1 score of 0.77.

The confusion matrix of the XGBoost classification model is shown in Fig. 2.

Figure 3 presents an analysis of the importance of variables in the XGBoost algorithm. As the feature importance rank was identified, gestational age and previous history of LBW were the top critical predictors.

Discussion

With the exponential growth in the quantity and dimension of healthcare data in recent years, ML approaches for dealing with complex and high-dimensional data have been introduced [13–15]. In this study, we aimed to evaluate the performance of eight different ML algorithms in predicting LBW. According to our findings, the XGBoost classification model had a more significant diagnostic performance parameter with an AUROC of 0.79, accuracy of 0.79, precision of 0.87, recall of 0.69, and F1 score of 0.77. XGBoost classification is a supervised machine learning algorithm based on a distributed gradient-boosted decision tree [16]. It can produce consistent
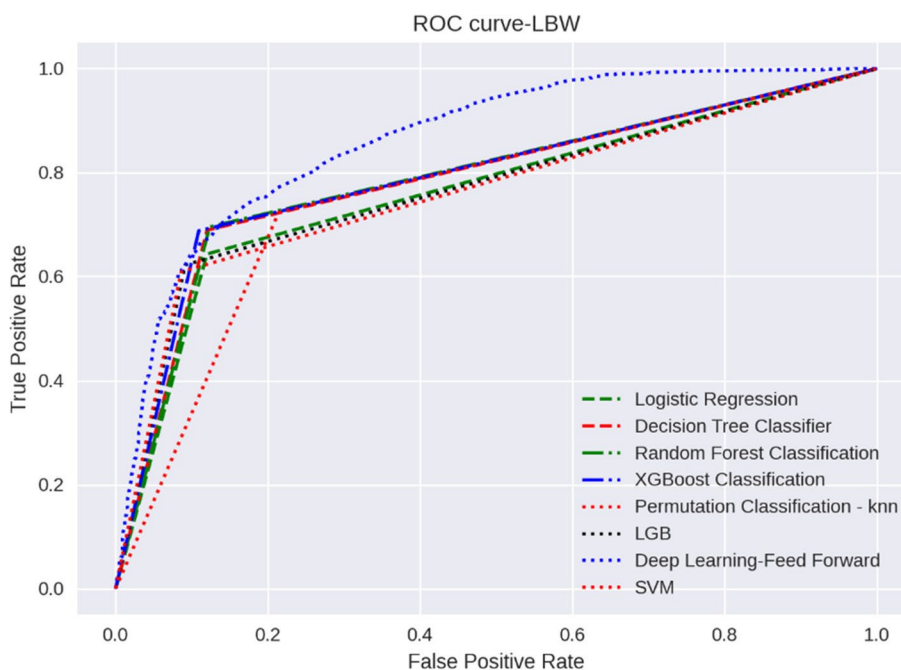
Fig. 1 AUROC of ML models

Table 2 Performance parameters of models with the highest AUROC

Row	Algorithms	Accuracy	Precision	Recall	F_1 Score
1	Random Forest Classification	0.78	0.85	0.70	0.77
2	XGBoost Classification	0.79	0.87	0.69	0.77
3	Deep Learning- Feed Forward	0.78	0.84	0.70	0.76

results while minimizing overfitting by employing a parallel tree-boosting strategy. Furthermore, XGBoost can use the importance score to determine the importance of each feature. Previous studies evaluating different ML machines for predicting LBW will also have promising results. According to Ahmadi et al., the random forest model performed well in terms of diagnostic performance, with an accuracy of 0.95, recall of 0.72, and AUROC of 0.89 [5]. Another study by Desiani et al. found that naive Bayes had excellent diagnostic performance, with an accuracy of 0.85 and a recall of 0.72 [17]. However, both studies were limited by a small sample size (less than 1000 participants).

Recent studies with larger sample sizes also demonstrated good performance. For example, in a survey by Eliyati et al., with a sample size of 12,500 study participants, SVM showed high diagnostic performance in predicting LBW with an accuracy of 0.93 [18]. Ren et al. used a more extensive study in this field, with a sample size of 266,687 birth records over six years. According to their findings, the XGBoost classification model had the highest recall score of 0.85, but the AUROC score was only 0.61 [19].

Although our study did not have the largest sample size of any study in this field, we believe that using hospital records made our feature selection rich enough to make a reasonable conclusion on identifying LBW risk factors. In our study, we surveyed maternal age, educational level, place of residence, inadequate prenatal care (fewer than three prenatal care visits), drug addiction, maternal anemia, cardiovascular disease, chronic hypertension, pyelonephritis, hepatitis, COVID-19, overt diabetes, gestational diabetes and thyroid dysfunction, parity, preeclampsia, and history of LBW. Among all the potential predisposing factors of LBW, gestational age and previous history of LBW were the top critical predictors. In line with previous findings [20, 21], gestational age is the highest predictor of LBW. Being born too soon (premature birth) is the most common cause of LBW. The prior history of LBW was another weighted factor in predicting LBW. It has been reported that the risk of LBW recurs between pregnancies. Women with a
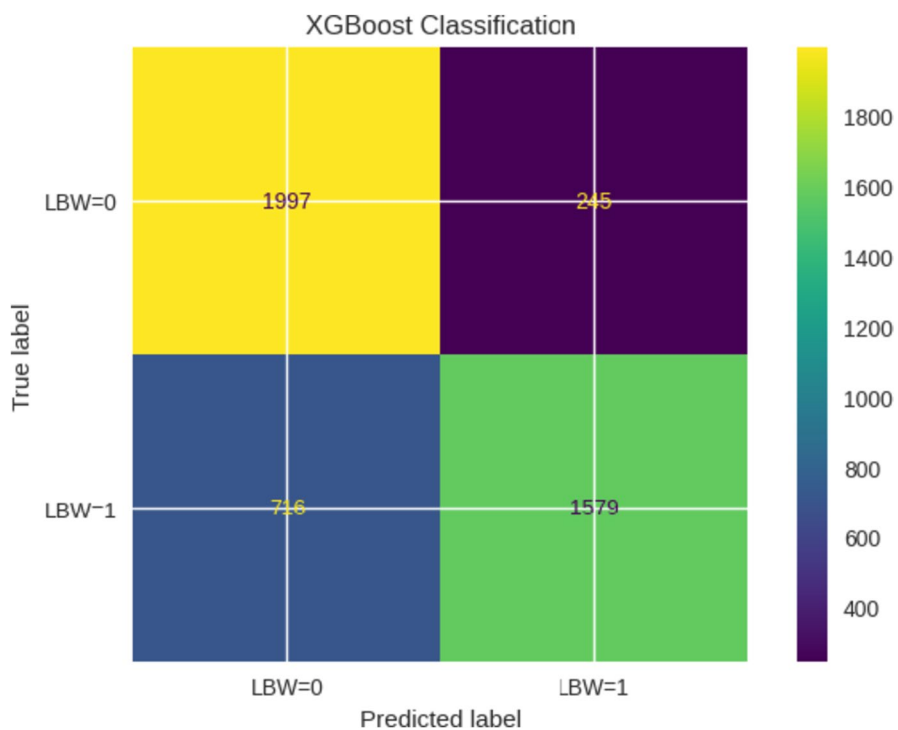
Ranjbar *et al. BMC Pregnancy and Childbirth*   (2023) 23:803

Page 5 of 7

Fig. 2 XGBoost classification confusion matrix
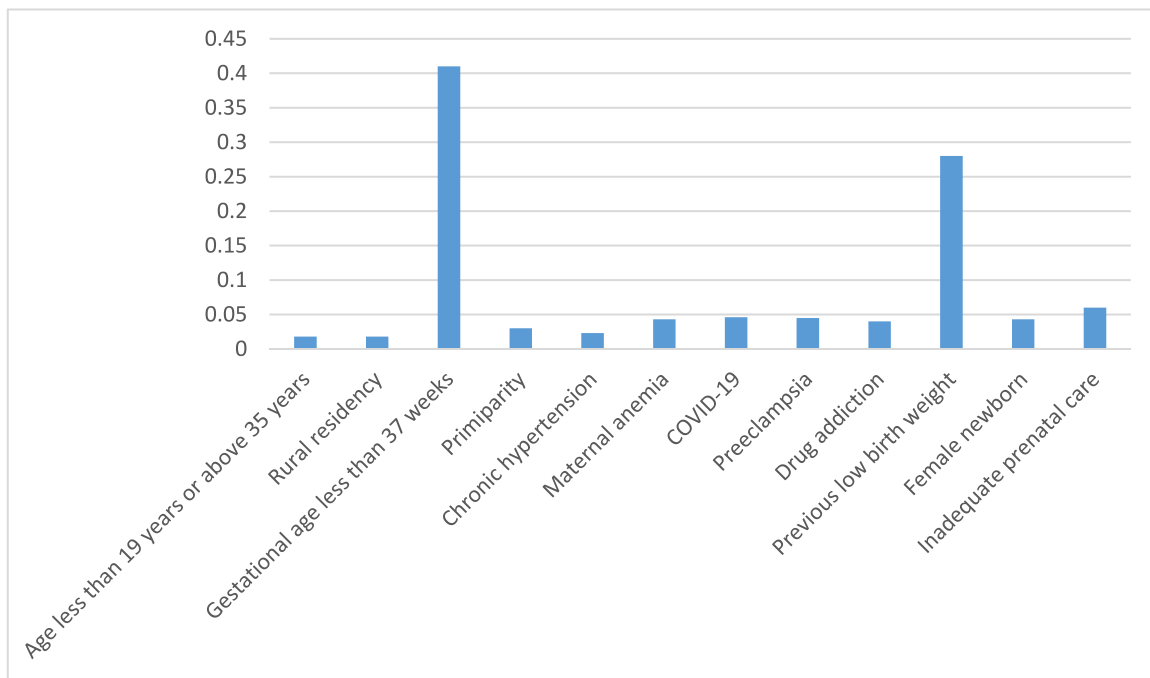
Fig. 3 Feature importance of the XGBoost classification in the prediction of low birth weight

Ranjbar *et al. BMC Pregnancy and Childbirth*   (2023) 23:803

Page 6 of 7

previous LBW baby have been identified as potential carriers of the recurrent risk and have a higher recurrence risk of LBW in their subsequent pregnancy than those with a previous normal birth weight baby [22]. Other factors, such as maternal comorbidities, sociodemographic characteristics, and fetal gender, were not among the weighted factors in predicting LBW, in contrast to previous studies. In one study, Bekele et al. found that fetal gender, marriage to birth interval, mother's occupation, and mother's age were all weighted factors in predicting LBW [23]. Another study found that maternal pre-pregnancy weight, mother's age, number of doctor visits during the first trimester, and previous preterm labor were the most significant risk factors for LBW [4]. The differences in findings could be attributed to study design, the type of ML models used, geographical differences, or imbalances in each study's dataset. It should be noted, however, that clinicians can use the key findings of each study to identify high-risk pregnant patients early in their pregnancy using early prenatal care screening tools.

Although we used a large dataset with a lot of maternal and neonatal information, a significant variable, like body mass index, was missing in most of the birth records, so we couldn't use this factor in our selection features, which is a significant limitation of the study.

Conclusion

Using ML approaches to predict LBW yielded promising results. As a result, this study might add to the current perinatal care literature. Although this study found that the XGBoost model performed well in predicting LBW, more research is needed to make a better conclusion on the performance of ML models in predicting LBW.

Abbreviations

LBW	Low birth weight
ML	Machine learning
XGBoost	Extreme gradient boost
LGB	Light gradient boost
SVM	Support vector machine
KNN	K-nearest neighbors

Declarations

Ethics approval and consent to participate
This study complied with the Declaration of Helsinki and was performed according to ethics committee approval. The Ethics and Research Committee of the Hormozgan University of Medical Sciences approved the study. The records of all patients who provided informed consent for using their data for research purposes were analyzed. For those under the age of 18 the informed consent was taken from their gurdian. Statistical analysis was performed with patient anonymity following ethics committee regulations.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

References

1. Chen Y, Li G, Ruan Y, Zou L, Wang X, Zhang W. An epidemiological survey on low birth weight infants in China and analysis of outcomes of full-term low birth weight infants. BMC Pregnancy Childbirth. 2013;13:242. https://doi.org/10.1186/1471-2393-13-242. PMID:24370213;PMCID: PMC3877972.
2. Heaman MI, Martens PJ, Brownell MD, Chartier MJ, Derksen SA, Helewa ME. The Association of inadequate and intensive prenatal care with maternal, fetal, and infant outcomes: a population-based study in Manitoba, Canada. J Obstet Gynaecol Can. 2019;41(7):947–59. https://doi.org/10.1016/j.jogc.2018.09.006. Epub 2019 Jan 11 PMID: 30639165.
3. Cunningham FG, Leveno KJ, Bloom SL, Hauth JC, Rouse DJ, Spong CY. Williams obstetrics. New York: McGraw-Hill; 2010. p. 804- 831. 23.
4. Senthilkumar D, Paulraj S. Prediction of low birth weight infants and its risk factors using data mining techniques. Proceedings of the 2015 International Conference on Industrial Engineering and Operations Management; IMEOM '15. Dubai: IEOM Society; 2015.
5. Ahmadi P, Alavimajd H, Khodakarim S, Tapak L, Kariman N, Amini P, Pazhuheian F. Prediction of low birth weight using random forest: a comparison with logistic regression. Arch Adv Biosci. 2017;8(3):36–43. https://doi.org/10.22037/jps.v8i3.15412.
6. Borson N, Kabir M, Zamal Z, Rahman R. Correlation analysis of demographic factors on low birth weight and prediction modeling using machine learning techniques. Proceedings of the 4th World Conference on Smart Trends in Systems, Security and Sustainability; WorldS4 '20. London: Institute of Electrical and Electronics Engineers; 2020. p. 169–73.
7. Faruk A, Cahyono ES. Prediction and classification of low birth weight data using machine learning techniques. Indones J Sci Technol. 2018;3(1):18–28. https://doi.org/10.17509/ijost.v3i1.10799.
8. International statistical classification of diseases and related health problems, 10th revision. World Health Organization; 2004. Availabe at: https://apps.who.int/iris/bitstream/handle/10665/42980/9241546530_eng.pdf?sequence=1&isAllowed=y.
9. Schimmel MS, Bromiker R, Hammerman C, Chertman L, Ioscovich A, Granovsky-Grisaru S, Samueloff A, Elstein D. The effects of maternal age and parity on maternal and neonatal outcome. Arch Gynecol Obstet. 2015;291(4):793–8. https://doi.org/10.1007/s00404-014-3469-0.
10. Sharifi N, Dolatian M, FathNezhadKazemi A, Pakzad R, Yadegari L. The relationship of the structural and intermediate social determinants of health with low birth weight in Iran: a systematic review and meta-analysis. Sci J Kurdistan Univ Medical Sci. 2018;23(2):21–36. https://doi.org/10.29252/sjku.23.2.21.
11. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É. Scikit-learn: machine learning in Python. J Mach Learn Res. 2011;12:2825–30. https://doi.org/10.1145/2786984.2786995.

Ranjbar *et al. BMC Pregnancy and Childbirth*   (2023) 23:803

Page 7 of 7

12. Yen SJ, Lee YS. Under-sampling approaches for improving prediction of the minority class in an imbalanced dataset. Proceedings of the 2016 International Conference on Intelligent Computing; ICIC '06. Kunming: Springer; 2006. p. 731–40.

13. Boujarzadeh B, Ranjbar A, Banihashemi F, Mehrnoush V, Darsareh F, Saffari M. Machine learning approach to predict postpartum haemorrhage: a systematic review protocol. BMJ Open. 2023;13(1):e067661. https://doi.org/10.1136/bmjopen-2022-067661. PMID:36657750;PMCID: PMC9853215.

14. Mehrnoush V, Ranjbar A, Farashah MV, Darsareh F, Shekari M, Jahromi MS. Prediction of postpartum hemorrhage using traditional statistical analysis and a machine learning approach. AJOG Glob Rep. 2023;3(2):100185. https://doi.org/10.1016/j.xagr.2023.100185. PMID:36935935;PMCID: PMC10020099.

15. Darsareh F, Ranjbar A, Farashah MV, Mehrnoush V, Shekari M, Jahromi MS. Application of machine learning to identify risk factors of birth asphyxia. BMC Pregnancy Childbirth. 2023;23(1):156. https://doi.org/10.1186/s12884-023-05486-9. PMID:36890453;PMCID:PMC9993370.

16. Chen T, He T. xgboost: eXtreme gradient boosting. The Comprehensive R Archive Network; 2017. https://cran.microsoft.com/snapshot/2017-12-11/web/packages/xgboost/vignettes/xgboost.pdf.

17. Desiani A, Primartha R, Arhami M, Orsalan O. Naive Bayes classifier for infant weight prediction of hypertension mother. J Phys Conf Ser. 2019;1282(1):012005. https://doi.org/10.1088/1742-6596/1282/1/012005.

18. Eliyati N, Faruk A, Kresnawati ES, Arifieni I. Support vector machines for classification of low birth weight in Indonesia. J Phys Conf Ser. 2019;1282(1):012010. https://doi.org/10.1088/1742-6596/1282/1/012010.

19. 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. https://doi.org/10.2196/44081. PMID:37256674;PMCID: PMC10267797.

20. Loreto P, Peixoto H, Abelha A, Machado J. Predicting low birth weight babies through data mining. Proceedings of the 2019 World Conference on Information Systems and Technologies; WorldCIST '19; March 27–29, 2018. Naples: Springer; 2019. pp. 568–77.

21. Khan W, Zaki N, Masud MM, Ahmad A, Ali L, Ali N, Ahmed LA. Infant birth weight estimation and low birth weight classification in United Arab Emirates using machine learning algorithms. Sci Rep. 2022;12(1):12110. https://doi.org/10.1038/s41598-022-14393-6.

22. Mvunta MH, Mboya IB, Msuya SE, John B, Obure J, Mahande MJ. Incidence and recurrence risk of low birth weight in Northern Tanzania: a registry based study. PLoS ONE. 2019;14(4):e0215768. https://doi.org/10.1371/journal.pone.0215768.

23. Bekele WT. Machine learning algorithms for predicting low birth weight in Ethiopia. BMC Med Inform Decis Mak. 2022;22(1):232. https://doi.org/10.1186/s12911-022-01981-9.

Publisher's Note