Healthcare Informatics Research

Machine Learning–based Classifiers for the Prediction of Low Birth Weight

Mahya Arayeshgari[1], Somayeh Najafi-Ghobadi[2], Hosein Tarhsaz[1], Sharareh Parami[1], Leili Tapak[1,3]
[1]Department of Biostatistics, School of Public Health, Hamadan University of Medical Sciences, Hamadan, Iran
[2]Department of Industrial Engineering, Faculty of Engineering, Kermanshah Branch, Islamic Azad University, Kermanshah, Iran
[3]Modeling of Noncommunicable Diseases Research Center, Hamadan University of Medical Sciences, Hamadan, Iran

Objectives: Low birth weight (LBW) is a global concern associated with fetal and neonatal mortality as well as adverse consequences such as intellectual disability, impaired cognitive development, and chronic diseases in adulthood. Numerous factors contribute to LBW and vary based on the region. The main objectives of this study were to compare four machine learning classifiers in the prediction of LBW and to determine the most important factors related to this phenomenon in Hamadan, Iran. **Methods:** We carried out a retrospective cross-sectional study on a dataset collected from Fatemieh Hospital in 2017 that included 741 mother-newborn pairs and 13 potential factors. Decision tree, random forest, artificial neural network, support vector machine, and logistic regression (LR) methods were used to predict LBW, with five evaluation criteria utilized to compare performance. **Results:** Our findings revealed a 7% prevalence of LBW. The average accuracy of all models was 87% or higher. The LR method provided a sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, and accuracy of 74%, 89%, 7.04%, 29%, and 88%, respectively. Using LR, gestational age, number of abortions, gravida, consanguinity, maternal age at delivery, and neonatal sex were determined to be the six most important variables associated with LBW. **Conclusions:** Our findings underscore the importance of facilitating timely diagnosis of causes of abortion, providing genetic counseling to consanguineous couples, and strengthening care before and during pregnancy (particularly for young mothers) to reduce LBW.

Keywords: Infant, Low Birth Weight, Gestational Age, Abortion, Induced, Logistic Models, Machine Learning

I. Introduction

Infant survival, physical and mental growth, and maternal health status and history are associated with a vital health indicator: birth weight [1]. The World Health Organization classifies low birth weight (LBW) as a weight of below 2,500 g obtained after birth for a live-born infant. Annually, more than 20 million infants (15%–20%) are born with LBW worldwide. An objective of the World Health Organization is to reduce the number of LBW infants by 30% by 2025 [2]. Rates of LBW have been reported at approximately 7%, 16.5%, and 18.6% of births in developed countries, less developed or developing countries, and least developed countries, respectively [3]. In Iran, an LBW prevalence of

8.5% was reported in a systematic review in 2020. According to that study, the highest percentage of LBW was found in Hamadan at 19.1% in 2007 [4]. A major public health problem, LBW has a variety of short- and long-term consequences. Infants with LBW are about 20 times more likely to die than heavier infants [5]. In addition to fetal and neonatal mortality, LBW is associated with higher risk of several adverse outcomes, including intellectual disability, impaired cognitive development, and future chronic health problems such as diabetes and cardiovascular diseases [6]. Specific maternal characteristics before and during pregnancy may provide a basis for predicting LBW. Although many researchers have endeavored to identify factors contributing to LBW, the causes differ by region, with poor fetal growth (due to poor maternal nutrition before and during pregnancy) a major cause in less developed regions and prematurity (due to high maternal age, multiparity, cesarean section, and smoking) in more developed ones [7]. LBW is also influenced by many other factors discussed in previous studies, such as maternal educational level, residence (urban or rural), family income, maternal occupation and health status, birth order, miscarriage, interpregnancy interval, and multiple pregnancies.

Recently, machine learning (ML), an important branch of artificial intelligence, has been widely used in many fields. In particular, breakthroughs have been made with ML methods in medical diagnosis and outcome prediction [8,9]. These methods are generally categorized into supervised or unsupervised learning. In a supervised ML method, a model is first trained on a variety of features related to a known outcome. The model can then make outcome predictions based on new data. When studying a discrete outcome (such as normal birth weight [NBW] versus LBW), the fitted model is termed a classification algorithm. Various ML methods have been proposed to improve the precision of data classification. Unlike traditional parametric statistical methods, ML techniques require no distributional assumptions about the dataset and are excellently suited for large datasets [10]. Nevertheless, each proposed ML method has specific features for outcome classification and estimation, and its performance may vary across conditions and datasets. Therefore, the purpose of this study was to determine the best technique for LBW prediction by comparing the predictive performance of five popular supervised ML methods—decision tree (DT), random forest (RF), artificial neural network (ANN), support vector machine (SVM), and logistic regression (LR)—using a dataset of neonates born at Fatemieh Hospital in Hamadan, Iran. Moreover, we aimed to determine the most important factors associated with LBW.

II. Methods

1. Data Collection
In this retrospective cross-sectional study, we selected a random sample of 800 infants born at Fatemieh Hospital in the city of Hamadan. For data collection, we used a researcher-designed questionnaire based on case records available on the Iranian Maternal and Neonatal Network (IMaN Net) in 2017. The IMaN Net was designed by Iran's Ministry of Health and Medical Education to evaluate the maternal and neonatal health status in Iran. After collecting the data, we excluded multiple pregnancies, stillbirths, and infants who died for any reason before discharge from the hospital or who had at least one abnormality. After this exclusion, 741 infants were included in the study, and the associated information was extracted as follows: (1) maternal data included place of residence (urban or rural); maternity insurance (insured or uninsured); delivery type (cesarean section or vaginal); maternal age at delivery (< 18, 18–35, or > 35 years); gestational age in weeks; preterm delivery (yes [< 37 weeks] or no [≥ 37 weeks of gestation]); consanguinity (yes or no); pregnancy risk factors such as chronic blood pressure, hepatitis, thyroid disease, cardiovascular disease, and preeclampsia/eclampsia (yes or no); gravida; parity (i.e., number of previous live and non-live births); number of abortions; and number of previous live births. (2) Neonatal data included sex (male or female) and birth weight in grams.

Infants were classified using a binary outcome (1 for LBW and 0 for NBW), with a birth weight of 2,500 g as a threshold.

2. Data Preprocessing and Missing Values
Before the analysis, the data were evaluated to ensure that no outliers were present. However, some missing values were found for seven variables, ranging from 0.13% to 5% of the dataset. The mean and median were used to impute quantitative and qualitative variables, respectively.

3. Machine Learning Classifiers
In classification, class imbalance and a bias toward the majority class may lead to misclassification. Thus, the data were first evaluated for imbalance, and the Synthetic Minority Oversampling TEchnique (SMOTE) was then used for the ML methods as an efficient algorithm for data balancing [11]. In this technique, the minority class is oversampled by creating synthesized samples according to the similarities between pairs of the existing minority instances.

1) Decision Tree (DT)

A simple ML technique, DT learning generates a tree-like structure by repeatedly splitting the dataset based on a criterion that maximizes the separation of the data [12]. This technique is first executed at distant parts of the tree and then returns to its beginning according to a method termed retrograde return [13]. DTs play an important role in medical diagnosis. We used recursive partitioning and regression trees with the tuning parameter of "cp" (the complexity parameter).

2) Random forest (RF)

The RF algorithm is based on an ensemble of large, correlated decision trees and combines the decisions of individual trees to produce accurate, stable results [14]. We used Breiman and Cutler's RF method with the tuning parameter of "mtry" (randomly selected predictors at each split).

3) Artificial neural network (ANN)

An ANN is a mathematical model designed to simulate the structure and function of biological neural networks in the brain [15]. Each ANN consists of a set of specially arranged neurons acting in coordination to solve a problem. This method is among the best for medical assessment and diagnosis due to its minimal error and maximal confidence. We used a single-hidden-layer neural network with tuning parameters of "size" (hidden units) and "decay" (weight decay).

4) Support vector machine (SVM)

SVMs use a decision boundary termed the hyperplane to separate classes. The hyperplane is located at a maximum distance from the closest data points of each class. These points are known as support vectors [16]. Two uses of SVMs in clinical medical research are prediction models for disease diagnosis and prognosis based on a specific diagnosis. We used a vector machine with polynomial kernel and tuning parameters of "degree" (polynomial degree), "scale" (scale), and "C" (cost).

5) Logistic regression (LR)

LR, a special case of generalized linear modeling, is extensively used for binary outcomes in epidemiology and medicine. By fitting data to a logistic function, the probability of an occurrence may be predicted [17]. We used binary LR with the enter method.

4. Evaluation Criteria

To compare the performance of the utilized classifiers, we divided the data into training (70% of the data) and test (30%) sets and repeated this process 10 times. Then, the performance of the trained models was evaluated using the test set based on criteria of sensitivity (or recall), specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and accuracy as follows:

$$Sensitivity = \frac{TP}{TP + FN},$$

$$Specificity = \frac{TN}{TN + FP},$$

$$PLR = \frac{Sensitivity}{1 - Specificity},$$

$$NLR = \frac{1 - Sensitivity}{Specificity},$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}.$$

A false positive (FP) indicates NBW neonates that were incorrectly identified as LBW, a true positive (TP) indicates LBW neonates that were correctly diagnosed as LBW, a true negative (TN) indicates NBW neonates correctly identified as NBW, and a false negative (FN) indicates LBW neonates incorrectly identified as NBW.

5. Hyperparameter Tuning

To find the optimum values of the hyperparameters for the methods (DT, RF, ANN, and SVM), we applied a 10-fold cross-validation strategy to the training set. Hyperparameters were chosen when the maximum value of the receiver operating characteristic was observed. We repeated this process 10 times with different training and test partitions.

6. Variable Importance

In this study, depending on the ML model, different methods were used to compute variable importance on a numerical scale from 0 to 100. For DT, it has been stated that "an overall measure of variable importance is the sum of the goodness of split measures for each split for which it was the primary variable, plus goodness (adjusted agreement) for all splits in which it was a surrogate" [18]. For RF, the increase in the percentage of instances in which a case was out-of-bag and misclassified when the variable was permuted was considered to indicate variable importance. For ANN, the variable importance was computed based on the weights method [19]. For SVM, variable importance was calculated using the area under the receiver operating characteristic curve. For

Table 1. Maternal and neonatal demographic characteristics

Variable	NBW (n = 690)	LBW (n = 51)	Total (n = 741)	p-value[a]
Neonatal sex				0.558
Male	354 (51.30)	24 (47.06)	378 (51.01)	
Female	336 (48.70)	27 (52.94)	363 (48.99)	
Place of residence				0.406
Urban	378 (54.78)	31 (60.78)	409 (55.20)	
Rural	312 (45.22)	20 (39.22)	332 (44.80)	
Maternity insurance				0.276
Uninsured	29 (4.20)	4 (7.84)	33 (4.45)	
Insured	661 (95.80)	47 (92.16)	708 (95.55)	
Delivery type				0.538
Vaginal	514 (74.49)	36 (70.59)	550 (74.22)	
Cesarean section	176 (25.51)	15 (29.41)	191 (25.78)	
Maternal age at delivery (yr)				0.470
< 18	29 (4.20)	4 (7.84)	33 (4.45)	
18–35	556 (80.58)	40 (78.43)	596 (80.43)	
> 35	105 (15.22)	7 (13.73)	112 (15.12)	
Preterm delivery				<0.001
No (≥ 37 wk)	651 (94.35)	15 (29.41)	666 (89.88)	
Yes (< 37 wk)	39 (5.65)	36 (70.59)	75 (10.12)	
Consanguinity				0.133
No	594 (86.09)	40 (78.43)	634 (85.56)	
Yes	96 (13.91)	11 (21.57)	107 (14.44)	
Pregnancy risk factors				0.122
No	573 (83.04)	38 (74.51)	611 (82.46)	
Yes	117 (16.96)	13 (25.49)	130 (17.54)	
Gravidity				0.354
1	256 (37.10)	21 (41.18)	277 (37.38)	
2	206 (29.85)	10 (19.61)	216 (29.15)	
3	138 (20.00)	11 (21.57)	149 (20.11)	
4	64 (9.28)	5 (9.80)	69 (9.31)	
≥ 5	26 (3.77)	4 (7.84)	30 (4.05)	
Parity				0.230
0	291 (42.17)	24 (47.06)	315 (42.51)	
1	225 (32.61)	10 (19.61)	235 (31.71)	
2	132 (19.13)	12 (23.53)	144 (19.44)	
≥ 3	42 (6.09)	5 (9.80)	47 (6.34)	
Number of abortions				0.095
0	555 (80.44)	39 (76.47)	594 (80.16)	
1	110 (15.94)	7 (13.73)	117 (15.79)	
2	19 (2.75)	3 (5.88)	22 (2.97)	
≥ 3	6 (0.87)	2 (3.92)	8 (1.08)	

Table 1. Continued

Variable	NBW (n = 690)	LBW (n = 51)	Total (n = 741)	*p*-value[a]
Number of previous live births				0.320
0	299 (43.33)	25 (49.02)	324 (43.72)	
1	227 (32.90)	11 (21.57)	238 (32.12)	
2	133 (19.28)	11 (21.57)	144 (19.44)	
≥ 3	31 (4.49)	4 (7.84)	35 (4.72)	
Gestational age (wk)	38.90 ± 1.39	34.17 ± 3.87	38.57 ± 2.06	<0.001
Neonatal birth weight (g)	3,226.36 ± 386.48	1,943.33 ± 503.77	3,138.05 ± 511.73	-

Values are presented as number (%) or mean ± standard deviation.

NBW: normal birth weight, LBW: low birth weight.

[a]*p*-values were obtained from chi-square testing or two-sample independent *t*-testing.

LR, the absolute value of the Wald statistic corresponding to the model was used to compute variable importance.

7. Software

SPSS version 25 (IBM Corp., Armonk, NY, USA) was used to calculate descriptive statistics, after which R software (version 4.2.1; R Foundation for Statistical Computing, Vienna, Austria; packages: caret, themis, rpart, randomForest, nnet, and kernlab) was used to apply ML classifiers to the dataset and measure the evaluation criteria. *p*-values of less than 0.05 were considered to indicate significance for all statistical inferences.

8. Ethics Approval and Consent to Participate

The data were collected from the IMaN Net. Therefore, a waiver of informed consent was awarded for this study. All methods were carried out in accordance with relevant guidelines and regulations, and the study was approved by the Ethical Committee of the Hamadan University of Medical Sciences (No. IR.UMSHA.REC.1401.779).

III. Results

1. Data Description

Among the 741 neonates, 51 (7%) had LBW (birth weight < 2,500 g). The mean ± standard deviation of the birth weight was approximately 3,138 ± 511 g. Overall, 51% of the neonates were male, and 10% were delivered before 37 weeks of gestation. Approximately 55% of mothers lived in urban areas, 95% were insured, and 74% had a vaginal delivery. The maternal age at delivery was between 18 and 35 years in most cases (80%). The mean ± standard deviation of the mother's gestational age was approximately 38 ± 2 weeks. Other maternal and neonatal characteristics are presented in Table 1.

2. ML Classifier Results

The importance levels of variables calculated using each ML model are shown in Figure 1. In addition, the results of the LR model are presented in Table 2 (reference category for the outcome, NBW). The performance of the five classifiers in the prediction of LBW is shown in Table 3 in terms of sensitivity, specificity, PLR, NLR, and accuracy on the test set with 10 repetitions. The means and standard deviations of these criteria for each classifier are reported in Table 3. The average accuracy of all models was 87% or higher. Furthermore, all models had high mean specificity (≥ 88%), with the lowest specificity observed with ANN and the highest with RF (97%). However, the mean sensitivity ranged between 44% (for RF) and 74% (for LR). The mean PLR was higher for RF than for the other models, at 15.27. In contrast, the mean NLR varied between 29% (for LR) and 58% (for RF). As a result, LR was found to be the best approach for the prediction of LBW in this study. As shown in Table 2 and Figure 1, LR analysis indicated that the six most important variables were gestational age, number of abortions, gravida, consanguinity, maternal age at delivery, and neonatal sex.

IV. Discussion

Different classification approaches for LBW have been utilized in several studies. In the United Arab Emirates, Khan et al. [20] assessed the performance of several ML algorithms. Through 5-fold cross-validation, they showed that the RF approach was superior to alternatives in birth weight estimation with regard to mean absolute error (294.53 g). However, the best classification performance was achieved using LR with SMOTE regarding accuracy (90.24%), precision (87.6%), recall (90.2%), and F1-score (0.89). Maternal diabetes, hypertension, and gestational age were found to be vital
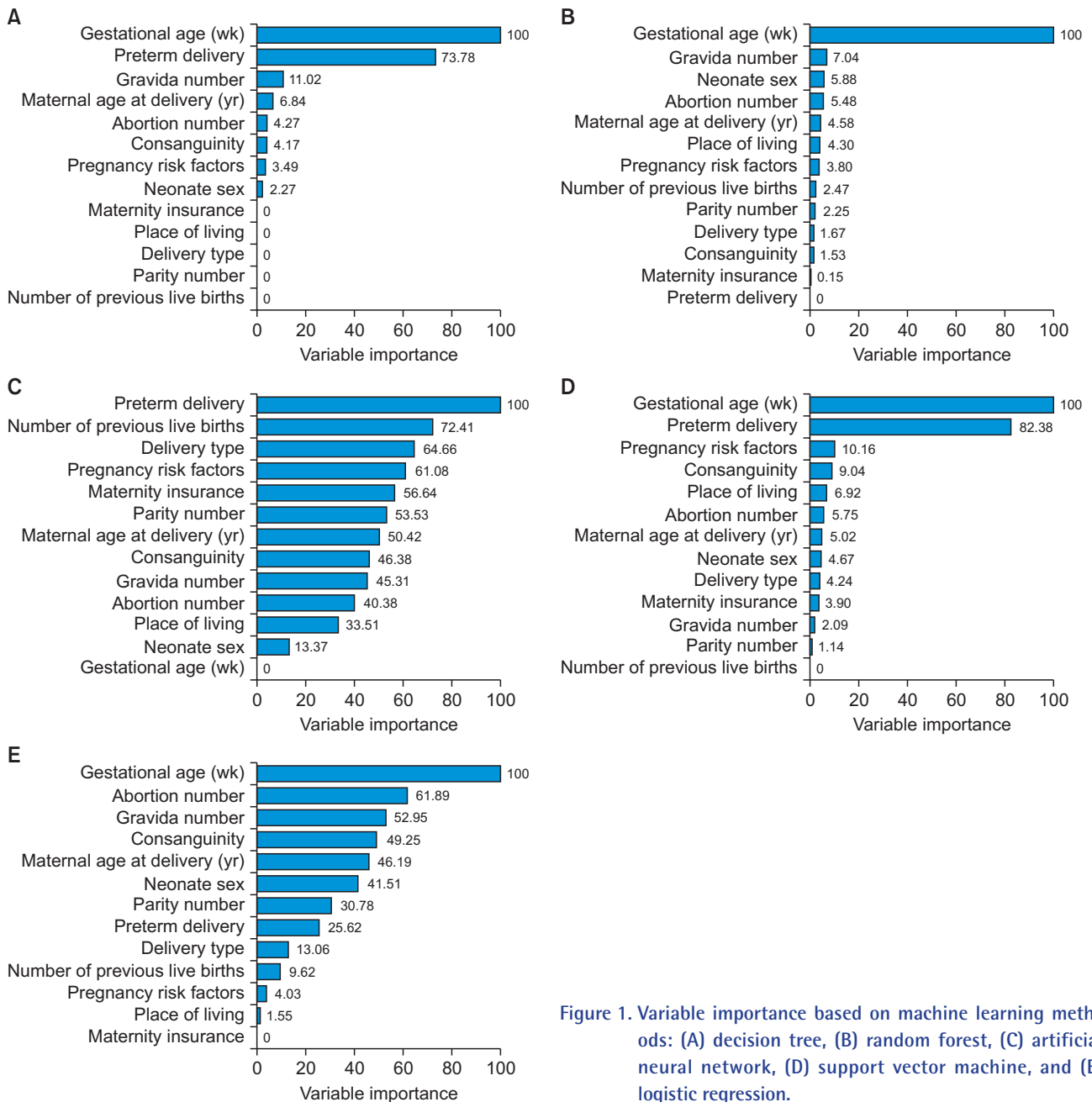
Figure 1. Variable importance based on machine learning methods: (A) decision tree, (B) random forest, (C) artificial neural network, (D) support vector machine, and (E) logistic regression.

factors in the classification of LBW. In a study by Zahirzada and Lavangnananda [21], five popular ML techniques (k-nearest neighbor, naive Bayes [NB], ANN, RF, and SVM) were applied to data obtained from the Afghanistan Demographic and Health Survey to determine the most effective strategy for predicting LBW. The data were divided into a training set (80% of the data) and a test set (20%). For both rural and urban areas, RF was the best method in terms of all four-evaluation metrics (accuracy, area under the curve, precision, and recall). In a study by Borson et al. [22] in Bangladesh, six classification techniques (LR, NB, RF, SVM, k-nearest neighbor, and multilayer perceptron artificial neural

network [MLP-ANN]) were used to predict infant LBW. According to 10-fold cross-validation, LR and SVM exhibited the greatest accuracy, at 80.3%. The highest precision (0.80) and F-measure (0.89) were obtained using SVM, while MLP-ANN was associated with the greatest recall (0.81). A train-test split (75:25) analysis also showed that the MLP-ANN, SVM, and LR methods showed almost identical performance, yielding the highest accuracy (81.6%), precision (0.81), recall (0.81), and F-measure (0.89) compared to the other classifiers. Senthilkumar and Paulraj [23] employed a cross-validation technique to predict the performance of six data mining algorithms (LR, NB, RF, SVM, ANN, and classi-

Table 2. Factors associated with low birth weight in infants based on logistic regression

Variable	β	Exp (β)	Wald	p-value
Intercept	27.07	-	18.38	<0.001*
Neonatal sex (ref: male)	0.88	2.42	4.08	0.043*
Place of residence (ref: urban)	−0.14	0.86	0.12	0.723
Maternity insurance (ref: uninsured)	−0.29	0.74	0.08	0.772
Delivery type (ref: vaginal)	0.38	1.47	0.69	0.404
Maternal age at delivery (yr)	−1.17	0.30	4.91	0.027*
Preterm delivery (ref: no)	0.97	2.65	1.84	0.174
Consanguinity (ref: no)	1.13	3.12	5.50	0.019*
Pregnancy risk factors (ref: no)	−0.24	0.78	0.21	0.647
Gravidity	−2.52	0.08	6.24	0.012*
Parity	1.95	7.03	2.47	0.115
Number of abortions	2.35	10.50	8.25	0.004*
Number of previous live births	0.66	1.94	0.47	0.490
Gestational age (wk)	−0.71	0.49	19.91	<0.001*

The reference category for the outcome is normal birth weight.

*$p < 0.05$.

Table 3. Performance comparison of ML classifiers in the prediction of LBW on a test dataset with 10 repetitions

Model	Sensitivity	Specificity	PLR	NLR	Accuracy
Decision tree	0.61 ± 0.16	0.93 ± 0.03	9.91 ± 3.42	0.41 ± 0.16	0.91 ± 0.03
Random forest	0.44 ± 0.13	0.97 ± 0.01	15.27 ± 4.45	0.58 ± 0.13	0.93 ± 0.01
Artificial neural network	0.71 ± 0.13	0.88 ± 0.03	6.37 ± 1.72	0.33 ± 0.14	0.87 ± 0.03
Support vector machine	0.68 ± 0.10	0.94 ± 0.01	11.92 ± 3.11	0.34 ± 0.11	0.92 ± 0.01
Logistic regression	0.74 ± 0.09	0.89 ± 0.03	7.04 ± 2.03	0.29 ± 0.10	0.88 ± 0.02

Values are presented as mean ± standard deviation.

ML: machine learning, LBW: low birth weight, PLR: positive likelihood ratio, NLR: negative likelihood ratio.

fication tree) using data gathered at Baystate Medical Center in Springfield, MA, USA. They demonstrated that compared to other data mining methods, the classification tree method provided superior overall prediction accuracy (89.95%), specificity (72.88%), area under the curve (93.80%), F-value (93.04%), and precision (88.81%). The highest recall was provided by RF, with a value of 99.23%. Last maternal weight (in pounds) before pregnancy and maternal age were the two main factors associated with LBW.

Numerous studies have indicated that low gestational age is one of the most critical risk factors for LBW [3]. In addition, similar to the present study, other research has indicated that poor obstetric history (such as past abortion) is associated with LBW. In a study by Brown et al. [24], women with one, two, and three or more prior abortions were, respectively, 2.8, 4.6, and 9.5 times more likely to have LBW

infants than those who had never had an abortion. Recently, Ghelichkhani et al. [25] assessed the maternal risk factors for preterm delivery (gestational age < 37 weeks) at Hamadan's Fatemieh Hospital. Their study revealed a history of abortion to be one of the most critical factors associated with preterm delivery. A study conducted by Cogendez et al. [26] showed that early detection of congenital and acquired intrauterine causes of abortion is possible with post-abortion hysteroscopy. Therefore, we anticipate that timely diagnosis of the causes of abortion and proper intervention can play a vital role in reducing preterm delivery and LBW. In a study by Demelash et al. [6], the rates of LBW for primigravida, multigravida, and grand multigravida cases were 47.3%, 38.7%, and 14%, respectively. This finding aligned with our study. The other factor related to LBW in the present study was consanguinity. Poorolajal et al. [27] performed a meta-

analysis to explore the effect of consanguinity on LBW. Their findings showed that consanguineous marriage can increase the risk of LBW. In many nations, such as in North America, it is forbidden or even illegal to marry close relatives; however, in other nations, especially those in Asia, the Middle East, and Africa, it may be preferred [28]. Lower maternal age at delivery has also been repeatedly found to be significantly related to adverse neonatal outcomes, specifically a higher prevalence of LBW. This can be explained by the fact that younger mothers are less likely to receive adequate prenatal care than older ones [29]. In the present study, female neonates had a higher risk of LBW than male infants. This may result from the greater lean body mass and lower body fat seen in male neonates relative to female infants or the Y chromosome's influence on the weight of the male fetus [30].

The present study had some limitations. First, we had no data regarding several important maternal characteristics, such as prenatal care, nutritional status, body mass index, interpregnancy interval, and financial status. Second, we could not include some features, such as maternal education, in the analysis due to a high percentage of missing values. Third, because considering each pregnancy risk factor (such as chronic blood pressure and cardiovascular disease) separately led to extremely unbalanced factor distributions, we considered all pregnancy risk factors as a single combined factor. Fourth, this study may have been vulnerable to potential bias in the evaluation of performance criteria, as the data were obtained from a retrospective registry-based study.

The results of this study showed that LR outperformed the other ML classifiers. Using promising classifiers to identify key LBW-related factors can allow medical practitioners to take preventative steps to minimize LBW. Based on the results, facilitating timely diagnosis of causes of abortion, providing genetic counseling to consanguineous couples, and improving care before and during pregnancy (especially for young mothers) can play an important role in reducing LBW. Additionally, the results of this study could be used to design an online mobile application to predict LBW risk in pregnant women. This would assist healthcare practitioners in the timely detection of mothers at high risk of giving birth to LBW infants and help provide them with appropriate interventions. In addition, researchers should consider the factors noted in the limitations section in further studies.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

ORCID

Mahya Arayeshgari (https://orcid.org/0000-0002-5031-146X)
Somayeh Najafi-Ghobadi (https://orcid.org/0000-0003-0657-8359)
Hosein Tarhsaz (https://orcid.org/0000-0003-2217-168X)
Sharareh Parami (https://orcid.org/0000-0003-4844-2548)
Leili Tapak (https://orcid.org/0000-0002-4378-3143)

References

1. Talie A, Taddele M, Alemayehu M. Magnitude of low birth weight and associated factors among newborns delivered in Dangla Primary Hospital, Amhara Regional State, Northwest Ethiopia, 2017. J Pregnancy 2019;2019:3587239. https://doi.org/10.1155/2019/3587239

2. World Health Organization. Global nutrition targets 2025: low birth weight policy brief [Internet]. Geneva, Switzerland: World Health Organization; 2014 [cited at 2023 Jan 18]. Available from: https://apps.who.int/iris/bitstream/handle/10665/149020/WHO_?sequence=2

3. Ranjbaran M, Jafary-Manesh H, Sajjadi-Hazaneh L, Eisaabadi S, Talkhabi S, Sadat Khoshniyat A. Prevalence of low birth weight and some associated factors in Markazi province, 2013-2014. World J Med Sci 2015;12(3):252-8. https://doi.org/10.5829/idosi.wjms.2015.12.3.93203

4. Sabbaghchi M, Jalali R, Mohammadi M. A systematic review and meta-analysis on the prevalence of low birth weight infants in Iran. J Pregnancy 2020;2020:3686471. https://doi.org/10.1155/2020/3686471

5. World Health Organization. Low birth weight [Internet]. Geneva, Switzerland: World Health Organization; c2023 [cited at 2023 Jan 18]. Available from: https://www.who.int/data/nutrition/nlis/info/low-birth-weight#:%7E:text=Low%20birth%20weight%20has%20been,2500%20grams%20(5.5%20pounds)

6. Demelash H, Motbainor A, Nigatu D, Gashaw K, Melese A. Risk factors for low birth weight in Bale zone hospitals, South-East Ethiopia : a case-control study. BMC

Pregnancy Childbirth 2015;15:264. https://doi.org/10.1186/s12884-015-0677-y

7. World Health Organization. UNICEF-WHO low birth-weight estimates: levels and trends 2000-2015 [Internet]. Geneva, Switzerland: World Health Organization; 2019 [cited at 2023 Jan 18]. Available from: https://www.unicef.org/reports/UNICEF-WHO-low-birthweight-estimates-2019

8. Ahmadi-Jouybari T, Najafi-Ghobadi S, Karami-Matin R, Najafian-Ghobadi S, Najafi-Ghobadi K. Investigating factors affecting the interval between a burn and the start of treatment using data mining methods and logistic regression. BMC Med Res Methodol 2021;21(1):71. https://doi.org/10.1186/s12874-021-01270-5

9. Najafi-Ghobadi S, Najafi-Ghobadi K, Tapak L, Aghaei A. Application of data mining techniques and logistic regression to model drug use transition to injection: a case study in drug use treatment centers in Kermanshah Province, Iran. Subst Abuse Treat Prev Policy 2019;14(1):55. https://doi.org/10.1186/s13011-019-0242-1

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

11. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 2002;16:321-57. https://doi.org/10.1613/jair.953

12. Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. J Biomed Inform 2002;35(5-6):352-9. https://doi.org/10.1016/s1532-0464(03)00034-0

13. abdulaziz Mohsen A, Alsurori M, Aldobai B, Mohsen GA. New approach to medical diagnosis using artificial neural network and decision tree algorithm: application to dental diseases. Int J Inf Eng Electron Bus 2019;11(4):52-60. https://doi.org/10.5815/ijieeb.2019.04.06

14. Sandhu AK, Batth RS. Software reuse analytics using integrated random forest and gradient boosting machine learning algorithm. Softw Pract Exp 2021;51(4):735-47. https://doi.org/10.1002/spe.2921

15. Suzuki K. Artificial neural networks: methodological advances and biomedical applications. Rijeka, Croatia: InTech; 2011.

16. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of support vector machine (SVM) learning in cancer genomics. Cancer Genomics Proteomics 2018;15(1):41-51. https://doi.org/10.21873/cgp.20063

17. Alanazi HO, Abdullah AH, Qureshi KN. A critical review for developing accurate and dynamic predictive models using machine learning methods in medicine and health care. J Med Syst 2017;41(4):69. https://doi.org/10.1007/s10916-017-0715-6

18. Therneau TM, Atkinson EJ. An introduction to recursive partitioning using the RPART routines. Rochester (MN): Mayo Foundation; 1997.

19. Gevrey M, Dimopoulos I, Lek S. Review and comparison of methods to study the contribution of variables in artificial neural network models. Ecol Model 2003;160(3):249-64. https://doi.org/10.1016/S0304-3800(02)00257-0

20. Khan W, Zaki N, Masud MM, Ahmad A, Ali L, Ali N, et al. 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

21. Zahirzada A, Lavangnananda K. Implementing predictive model for low birth weight in Afghanistan. Proceedings of 2021 13th International Conference on Knowledge and Smart Technology (KST); 2021 Jan 21-24; Bangsaen, Thailand. p. 67-72. https://doi.org/10.1109/KST51265.2021.9415792

22. Borson NS, Kabir MR, Zamal Z, Rahman RM. Correlation analysis of demographic factors on low birth weight and prediction modeling using machine learning techniques. Proceedings of 2020 4th World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4); 2020 Jul 27-28; London, UK. p. 169-73. https://doi.org/10.1109/WorldS450073.2020.9210338

23. 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; 2015 Mar 3-5; Dubai, UAE. p. 186-94.

24. Brown JS, Adera T, Masho SW. Previous abortion and the risk of low birth weight and preterm births. J Epidemiol Community Health 2008;62(1):16-22. https://doi.org/10.1136/jech.2006.050369

25. Ghelichkhani S, Masoumi SZ, Shirzadeh AA, Khazaei S, Shahbazi F. Evaluation of maternal risk factors for preterm delivery in Fatemieh Hospital of Hamadan, Iran, 2019: a case-control study. J Family Med Prim Care 2021;10(10):3832-7. https://doi.org/10.4103/jfmpc.jfmpc_1032_21

26. Cogendez E, Dolgun ZN, Sanverdi I, Turgut A, Eren S. Post-abortion hysteroscopy: a method for early di-

agnosis of congenital and acquired intrauterine causes of abortions. Eur J Obstet Gynecol Reprod Biol 2011; 156(1):101-4. https://doi.org/10.1016/j.ejogrb.2010.12.025

27. Poorolajal J, Ameri P, Soltanian A, Bahrami M. Effect of consanguinity on low birth weight: a meta-analysis. Arch Iran Med 2017;20(3):178-84.

28. Bennett RL, Motulsky AG, Bittles A, Hudgins L, Uhrich S, Doyle DL, et al. Genetic counseling and screening of consanguineous couples and their offspring: recommendations of the National Society of Genetic Coun-selors. J Genet Couns 2002;11(2):97-119. https://doi.org/10.1023/A:1014593404915

29. Friede A, Baldwin W, Rhodes PH, Buehler JW, Strauss LT, Smith JC, et al. Young maternal age and infant mortality: the role of low birth weight. Public Health Rep 1987;102(2):192-9.

30. Kebede A, Kebede A, Belina S, Biratu Y. Trends and determinants of small birth weight in Ethiopia: further analysis of Ethiopian Demographic and Health Surveys. Ethiop J Health Sci 2021;31(2):299-310. https://doi.org/10.4314/ejhs.v31i2.13