International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V73I4P136

Volume 73 Issue 4, 440-452, April 2025

Original Article

Prediction of Birth Weight Using Machine Learning Models

Keerthana P[1], Sasidhar Babu Suvanam[2]

[1]*School of Computer Science Engineering, REVA University, Bangalore, India.*
[2]*Department of Computer Science Engineering, Presidency University, Bangalore, India.*

[1]*Corresponding Authors: keerthiikeerthanayadav@gmail.com*

Received: 18 November 2024 Revised: 03 April 2025 Accepted: 17 April 2025 Published: 26 April 2025

Abstract - *The study focuses on the fetal birthweight using machine learning (ML) based prediction models. The software used for this research is Python, which is used to design and implement the birthweight forecasting model. The model developed uses six different regression models, using Linear, Ridge, ElasticNet, SVM, XGB and Ensemble (XGB+GB) models, respectively. The performances of each regression model are evaluated using R2 estimation. The results show that the ensemble model achieved a higher accuracy level of 90% (.9039392149835611), whereas the other five models obtained less than 80%. The accuracy achieved by the elastic-net model is 34%, ridge and linear models are 39%, the SVM mode gained 42.9%, and the XGB model achieved 77.7%. From the findings, it is proven that the XGB+GB ensemble model obtained higher accuracy with better performance. Thus, from the results, it can be concluded that the ensemble model is more efficient in accurately predicting the birth weight of fetuses than other traditional models.*

Keywords - *Birthweight, Ensemble model, Fetal birthweight prediction, Fetuses birthweight prediction, Regression models, R squared estimation.*

1. Introduction

Predicting a baby's birth weight is an important factor of prenatal care as it enables early intervention and specific medical care for expectant mothers and their unborn children. Birth weight has the greatest impact on its potential for survival. Low birth weight (LBW) and high birth weight (HBW) are becoming more of a problem, specifically in developing countries. LBW, in general, continues to be a main global concern for general well-being, increasing the risk of death and impairment for young children and babies. Birth weight (BW) also has a major impact on unborn babies' health and endurance; thus, predicting the exact BW can help healthcare professionals make the best decisions [1].

The linear regression and logistic regression methods are two different kinds of regression analysis techniques that can be used to solve regression problems using machine learning. In the context of predictive modelling, regression analysis looks at the relationship between an independent variable and a dependent variable in a dataset. Regression analysis employs several techniques depending on whether there is a linear or non-linear association between the target and independent variables. The major applications of the regression technique are time series analysis, forecast trend analysis, and cause and effect analysis [2]. Machine learning uses regression analysis as its main technique for resolving regression problems. The gradient boosting approach, in general, can be implemented effectively with the help of the open-source software known as Extreme Gradient Boosting (XGBoost). Because of its efficacy as a tree-based ensemble learning technique, data scientists consider it a strong and effective instrument. After its development and initial release, XGBoost has been the standard technique and frequently the critical element when it comes to dealing with a variety of issues in predictive analysis in machine learning contests [3]. Both mother and newborn physical conditions may suffer largely because of LBW, including infant death and various long-term health issues when a baby survives.

In this context, various machine learning algorithms, including XGBoost, Naïve Bayes, Support Vector Machine (SVM), etc., play an influential role in predicting low birth and normal birth weight babies. The work in [4] carried out a study to predict LBW using machine learning algorithms, and among the other machine learning models considered, XGBoost is proven to be effective and provide increased accuracy in birth weight prediction. The recognition and acceptance of XGBoost as a machine learning algorithm and its scalability have exhibited the technique's efficacy in various machine learning applications. Machine learning is becoming increasingly popular in optimizing the diagnosis of infant-related diseases due to several methodological advantages [5]. The problem of predicting birth weight is well-defined; however, discussing the long-term health impacts of birth weight could provide a more comprehensive understanding of its significance. LBW is associated with increased risks of infant mortality, developmental delays, and chronic health conditions such as diabetes, cardiovascular diseases, and obesity later in life. Additionally, LBW infants may experience cognitive impairments, reduced academic performance, and lower socioeconomic outcomes in adulthood. On the other hand, macrosomia (HBW) can lead to complications during delivery, higher chances of obesity, and metabolic disorders. Understanding these long-term health consequences emphasizes the importance of accurate birth weight prediction models, as they enable early interventions and preventive healthcare strategies to improve maternal and child health outcomes.

The predictive model, in general, can identify risk and protective factors linked with infant mortality rate, which could help to guide future interventions to lessen risk factors and encourage various protective factors. XGBoost can particularly be used to resolve demanding problems in the real world with the minimum amount of resource consumption. To define whether a model delivers an increasing level of accuracy and associations between real-time variables, multi-layer perceptron models, along with XGBoost models can be used in prediction analysis. Also, in [6], it has been emphasized that, among other machine learning methods, random forest and XGBoost performed best in predicting neonatal death. Among the different machine learning models used, XGBoost provides the best performance and has proven to be effective on the original imbalanced dataset.

In general, Low Birth Weight (LBW) is said to have a strong relationship with newborn babies' mortality rate. Accordingly, as a useful prophylactic measure and indicator of newborn health vulnerabilities, LBW prediction is considered useful. Using direct medical and health policy interventions, it is significant to figure out which pregnant patients are most likely to have a baby with LBW during the preconception or early stages of pregnancy to save newborn lives and lessen possibly avoidable healthcare conditions. Previous studies on LBW prediction have been proven to be effective concerning LBW prediction. Also, performance metrics exhibit that the XGBoost classification has outperformed the other machine learning models [7].

XGBoost is one of the machine learning techniques that use gradient boosting to carry out various problems relating to ranking, regression, and classification. It makes use of a gradient descent optimization method to train decision trees iteratively to reduce a loss function. Machine learning plays a significant role in the medical field and prediction analysis relating to the healthcare field. Many machine learning models have been developed to predict medical complaints before they become a life-threatening process. Unrecognized Small for Gestational Age (SGA) prior to delivery is considered one of the primary risk factors for miscarriage. However, antenatal prediction of SGA considers closer monitoring and timely delivery to prevent premature fetal outcomes. When the syndrome is identified before delivery, the risk could be significantly lessened. The ultimate prediction model is determined to be the XGBoost algorithm, as it is the highest-performing model. The XGBoost model is found to be best suited for predicting SGA babies, indicating that machine learning is ultimately a potential method for this kind of prediction [8].

1.1. Research Gap

Existing studies on fetal birthweight prediction primarily rely on traditional statistical methods or single machine learning models, often yielding suboptimal accuracy due to limitations in capturing complex nonlinear relationships. Many prior works overlook the potential of ensemble learning, which can enhance predictive performance by combining multiple models. This research addresses the gap by implementing and evaluating six regression models, demonstrating that traditional models like Linear, Ridge, and ElasticNet regressions perform poorly, while the XGB+GB ensemble model achieves significantly higher accuracy. The novelty of this research lies in the integration of ensemble learning for birthweight prediction, proving its effectiveness in improving predictive accuracy compared to standalone models, thereby offering a more reliable approach for fetal health assessment.

In this context, machine learning algorithms have already become a standard choice for efficient clinical applications, i.e. birth weight categorization and assessment [9]. The current study has evaluated the effectiveness of different machine-learning algorithms in predicting birth weight.

2. Literature Review

This section highlights the analysis of various ML models used for the prediction of birth weight. The work in [8] adopted a model of machine learning to predict LBW. The predictive model was developed based on statistical learning such as random forest classification, extreme gradient boost, decision tree classification, logistic regression, deep learning feedforward, support vector machine, light gradient boost, and k-nearest neighbors-based permutation feature. It was identified that extreme gradient boosts performed better in predicting LBW. Machine learning concentrates on analyzing the data using different learning processes and statistical tools.

The main intention of this study was to adopt techniques of machine learning on LBW data with specific reference to Indonesia. This study carries out machine learning tasks that entail prediction and classification. A model of binary logistic regression (BLR) was adopted for training and testing data. A random approach was adopted to set the data. It was found that BLR had the best performance in terms of prediction; on the other hand, it showed a poor approach in terms of classification. The random forest had better performance for the classification and prediction of LBW [9].

Linear regression can be considered a popular regression learning procedure for predictive analysis. Francis Galton (1894) originally proposed the idea of linear regression. When used on a dataset, linear regression is a statistical test that contributes to figuring out and measuring the link between the variables under consideration. Chi-square, Fisher's exact test, t-test, and ANOVA are some examples of univariate statistical tests that do not allow accounting for the effect of additional factors or confounders during experiments [10].

In contrast, regression and partial correlation provide the researcher with more control over the effect of confounders on the interpretation of the link between two variables. Linear regression is largely used in biological and medical fields for prediction analysis. It is a statistical method of defining relationships between two or more variables in biological or medical data using estimation. For example, it is possible to use linear regression to determine whether age and weight have any effect on the healthcare risk. Age, weight, and sex are some of the important examples of response factors that describe the dependent variable, which is the variable that should be explained, i.e., SBP.

The linear regression method in [11] was used to examine the link between a dependent variable and an independent variable. A popular statistical method for establishing a relationship model between two different variables is known as regression analysis. The two variables are the predictor variable and the response variable. Of these, values of the predictor variable can be observed through experiments, whereas the value of the response variable can be derived from the predictor variable.

The study in [12] carried out a study for predicting infant weight using the linear regression model. Multivariate linear regressions have been employed to define which variables play a significant role in significant for prediction. The birth weight of babies has been predicted using the linear regression model. The findings emphasized that the resulting model can give 65% accuracy in data weight variation. According to the findings, the suggested simplified regression model can accurately predict the term birth weight of low-birth-weight babies.

High dimensionality leads to experiential non-identifiability, which makes it difficult to fit the linear regression model to a high-dimensional dataset. By expanding the loss function with a penalty, penalized regression avoids this non-identifiability. Ridge regression is the effect of combining the squared regression coefficients or the ridge penalty. As high-dimensional datasets became available, there has been a resurrection of interest in the ridge regression estimator, which was primarily developed to deal with collinearity. An analysis using ridge regression could be compared with some kind of averaging, which is comparatively strong and reproducible. It was identified in [13] that the "ridge predictor" can perform better in terms of prediction than the lasso equivalent.

Any data which has a high degree of multicollinearity could be assessed using ridge regression. Ridge regression is also known as L2 regularization. Predicted values differ significantly from actual values when multicollinearity is present, least-squares are impartial, and differences are high to a certain extent. Also, when there is multicollinearity, the ridge estimator acts incredibly well at improving the least-squares estimate. It is evident from the study in [14] that ridge regression of machine learning techniques is proven to be effective in predicting mechanical properties and resources, making them suitable for materials scientists and engineers.

Ridge regression is one of the significant forms of linear regression where it is possible to improve long-term predictions by integrating a modest bit of bias. One regularization technique for reducing the model's complexity is ridge regression. This technique modifies the cost function by integrating the penalty term. Also, the term "ridge regression penalty" denotes the extent of bias in the model. When there is considerable collinearity between the independent variables, a general linear regression will probably not work, and in this context, ridge regression can be utilized to deal with these issues.

When there are more parameters than samples, the issues are comparatively simpler to solve. Ridge regression integrates all the model's features and is generally used to reduce over-fitting issues [15]. A reduction in coefficients is likely to reduce the complexity of the model. Regression is a statistical method for predicting costs since it is a dependent variable and an independent variable. For huge sample sizes, ridge regression should not be used. The variance reduction attained through ridge estimation in these situations is not sufficient to offset the bias of the approach because the variance of the ML estimate is comparatively tiny. The variation of the ML estimator increases significantly as the sample size decreases for smaller sample sizes; however, the variance of the ridge estimation barely changes. Therefore, the ridge method proved to be more effective than ML estimation for smaller data.

The study in [16] developed a method for predicting the infants' weight range using an approach of machine learning. The authors adopted five models of machine learning such as ANN (artificial neural network), support vector machine, decision tree, Naïve Bayes and logistic regression. The accuracy of naïve bayes and ANN is 70 percent and 60 percent for logistic regression and support vector machine. On the other hand, the decision tree was below 60 percent. Research was carried out in [17] to develop a predictive model using machine learning techniques for predicting the birth weight range of fetal into high, low, or normal. It was found that developed machine learning systems can give reliable predictions. Numerous authors adopted neural networks for predicting fetal weight prediction.

The work in [18] predicted fetal weight using ANN with an accuracy of 100 percent in Palestine. The work in [19] adopted the ANN method to predict fetal weight with an error rate below 15 percent in the USA. The work in [20] used genetic algorithm-back propagation NN to predict fetal weight with an accuracy of 76.3 percent. The work in [21] used pre-trained CNN (convolution neural network) to predict fetal features with an accuracy of 97.05 percent.

Authors followed various machine learning algorithms for predicting fetal weight, namely the XG boost method [22]-[26], support vector machine [27] [28] and random forest method [29] [30]. The work in [31] studied algorithms of machine learning to predict low birth weight (LBW) in Ethiopia. This research adopted extreme gradient boosting, decision tree, logistic regression, random forest, gradient boosting, and K-nearest neighbour to compare the results. Classifier categories based on LBW and normal weight. Random forest was the best classifier which predicted accuracy, F1 score and recall (91.6 percent), Jaccard score (81.86) and hamming loss (1.05). It was clear that random forest predicts the LBW rate effectively and correctly when compared with other classifiers. Child gender, mother's age and occupation, and interval from marriage to birth were the main predictors of LBW in Ethiopia.

The work in [32] conducted retrospective cross-sectional research to predict LBW using machine learning algorithms. ANN, random forest, decision tree, logistic regression, and support vector machine were adopted to predict LBW. The average accuracy of all machine learning models was 87 percent or higher. The logistic regression method gave an accuracy, negative likelihood ratio, positive likelihood ratio, specificity, and sensitivity of 88 percent, 29 percent, 7.04 percent, 89 percent, and 74 percent. Using the best classifiers to predict the main LBW-related factors could allow healthcare providers to take preventive steps to minimize LBW.

To facilitate timely diagnosis for abortion causes, enhancing care before pregnancy and during pregnancy (particularly for young mothers), genetic counselling to needed couples.

Thus, it was summarized that logistic regression performed better than other machine learning classifiers. The work in [33] developed and evaluated a model of machine learning for predicting failure in postnatal growth among very LBW infants. Models of machine learning were built using random forest, CNN, support vector machine, and extreme gradient boosting (XGB) for comparison against conventional MLR (multiple logistic regression) models. It was found that XGB indicated the best performance.

The work in [34] predicted the risks associated with LBW using machine learning with specific reference to Bangladesh. The average percentage of LBW was 16.2 percent. Respondent's education, region, wealth index, height, alive child, and twin child were the main risk factors for LBW babies. A classifier based on logistic regression showed 87.6 percent accuracy; on the other hand, the decision tree showed 85.4 percent accuracy. A classifier based on logistic regression had shown the best accuracy in classifying LBW babies.

It was suggested that there is a need for an integrated, cost-effective, and efficient complementary approach for reducing and correctly predicting LBW babies with specific reference to Bangladesh. Birth weight is the main factor during the process of fetal development to protect infant and maternal safety. This research predicted LBW infants using a classifier based on LSTM (long short-term memory).

Classification accuracies are evaluated based on various classifiers for SGA (small-for-gestational-age), AGA (appropriate-for-gestational-age) and LGA (large-for-gestational-age) groups. The findings of the research indicate the accuracy rate for the model of prediction using random forest, CNN, backpropagation NN, support vector machine, linear regression, and the proposed hybrid model. The proposed hybrid model maximizes the convergence rate and enhances the accuracy in the prediction of birth weight [35]. Table 1 highlights the advantages and limitations of the reviewed current models in this review.

Table 1. Critical analysis of current models

Ref	Models	Advantages	Disadvantages
[8-9]	Random Forest, XGBoost, Decision Tree, Logistic Regression, Deep Learning Feedforward, SVM, LightGBM, KNN	XGBoost performed the best for LBW prediction and Random Forest for classification.	Logistic Regression had poor classification performance.
[10]	Linear Regression	Measures relationship between variables and controls confounders.	Poor performance in high-dimensional data lacks predictive power for complex relationships.
[11]	Linear Regression	Simple, widely used for regression analysis.	Limited to linear relationships, high bias.

[12]	Multivariate Linear Regression	65% accuracy in weight variation prediction.	High dimensionality leads to non-identifiability.
[13]	Ridge Regression	Better performance than Lasso for prediction.	It is not suitable for very large sample sizes.
[14, 15]	Ridge Regression	Effective in reducing overfitting and multicollinearity issues.	Poor performance when large sample sizes are available.
[16]	ANN, SVM, Decision Tree, Naïve Bayes, Logistic Regression	Naïve Bayes (70%) and ANN (60%) performed best.	The decision Tree had the lowest accuracy.
[17]	Neural Network	Reliable predictions using ML.	Lacks specific performance comparisons.
[18]	ANN	Achieved 100% accuracy.	Lacks generalizability outside the dataset.
[19]	ANN	Low error rate (<15%).	Potential overfitting issues.
[20]	Genetic Algorithm+Backpropagation NN	Higher accuracy.	Complexity in model tuning.
[21]	Pre-trained CNN	97.05% accuracy.	Requires large labelled dataset.
[22-26]	XGBoost	High predictive power.	Computationally expensive.
[27, 28]	SVM	Works well with small datasets.	Poor scalability with large datasets.
[29, 30]	Random Forest	Robust, handles feature importance well.	It can be computationally expensive.
[31]	XGBoost, Decision Tree, Logistic Regression, Random Forest, Gradient Boosting, KNN	Random Forest had the highest accuracy (91.6%).	Some classifiers had poor performance.
[32]	ANN, Random Forest, Decision Tree, Logistic Regression, SVM	Logistic Regression (88%) performed best.	Requires careful selection of features.
[33]	Random Forest, CNN, SVM, XGBoost	XGBoost outperformed MLR.	High computational cost.
[34]	Logistic Regression, Decision Tree	Logistic Regression (87.6%) outperformed Decision Tree.	Limited dataset for generalization.
[35]	Random Forest, CNN, Backpropagation NN, SVM, Hybrid Model	The hybrid model maximized convergence rate and accuracy.	High computational complexity.

3. Proposed Methodology

The research method in this research uses a hybrid approach where two regression algorithms are used in the machine learning model to detect the birth weight of the fetuses. Generally, the birth weight of a fetus is determined using the clinical examination outcomes and ultrasound (ultrasonic examination) datasets.

In this research, the data obtained from the patients (mothers) are used as the inputs, and the predictions of the fetuses as “birth weight” using the data are obtained by the hybrid ML model developed. Here, weight, age, height, medical history, habits, ethnicity, and more attributes are used as datasets to determine the birth weight of fetuses.

3.1. Proposed Design

The current research aims at finding the birth weight of fetuses using different algorithms in one ensemble model. The developed ensemble model uses the Voting Regressor (VR) technique that combines two or more regression methods into one regression model for predictions.

This method is a hybrid approach where both Gradient-Boosting (GB) and XGBoost (XGB) algorithms are utilized to achieve better performance than the existing prediction models.

The ensemble model overcomes the over-fitting issues, and thus, using large datasets is not a complication. The ensemble model developed here initially gains the pre-processed inputs (datasets) and passes them to the model for learning the features. Once the model is trained to extract the features, it is then passed onto the testing phase to predict the weight accurately. Once the model achieves a higher score, the model is retained. If not, it is re-trained by adjusting the parameters for better accuracy and minimal data loss. Thus, the design of the birth weight prediction model (refer to figure 1) is planned such as:

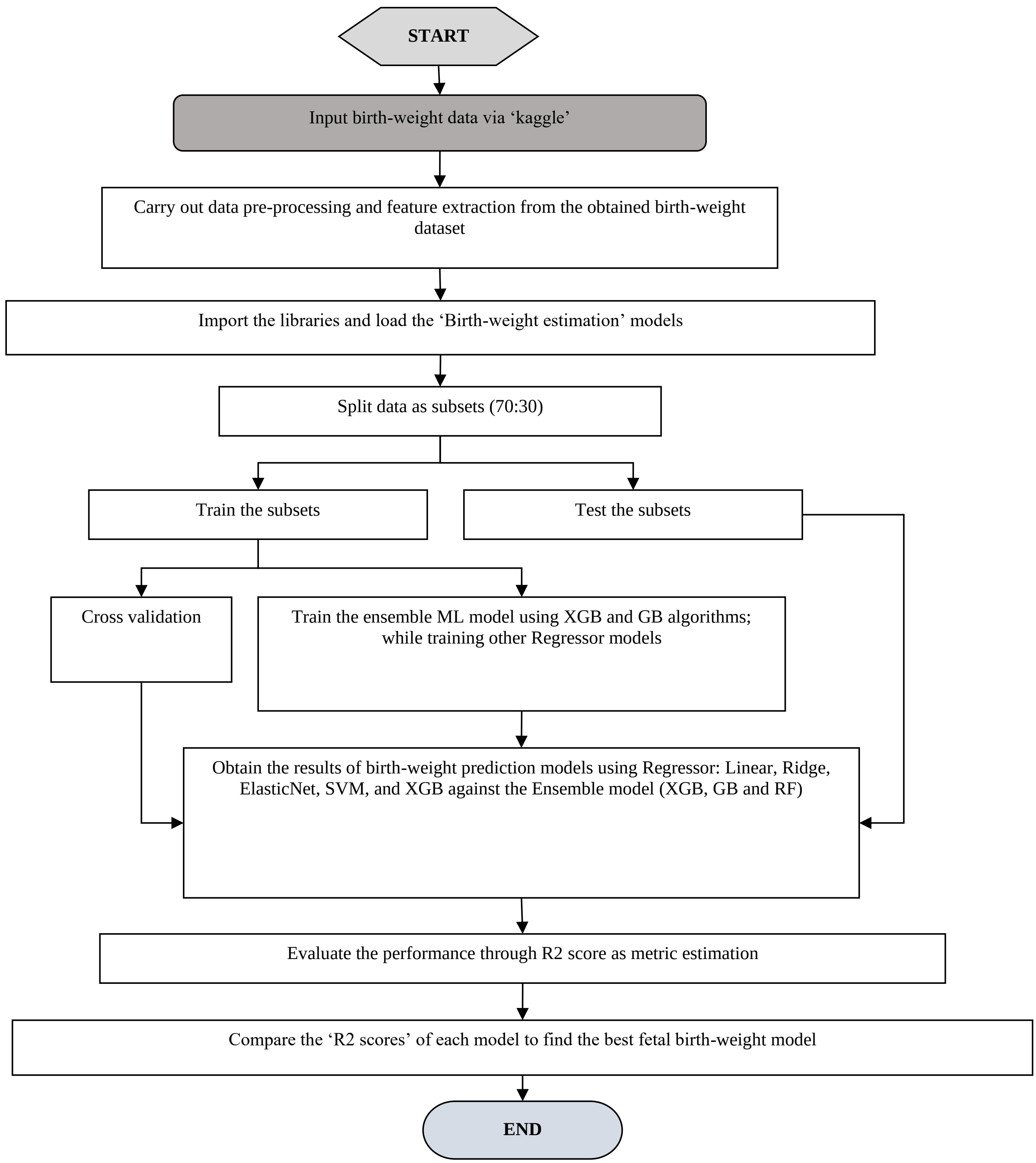

Fig. 1 Workflow of research design

3.2. Architecture

The proposed deep learning-based prediction research focuses on comparing birthweight prediction-based regressor models against the developed model. The model designed here includes two regressor algorithms, namely XGB and GB, where the random forest (RF) is used as the classifier algorithm. The inputs as texts are obtained from "Kaggle" and are passed through the "ensemble model" as a hybrid approach to gain the expected outcome with higher accuracy than other models (refer to figure 2).

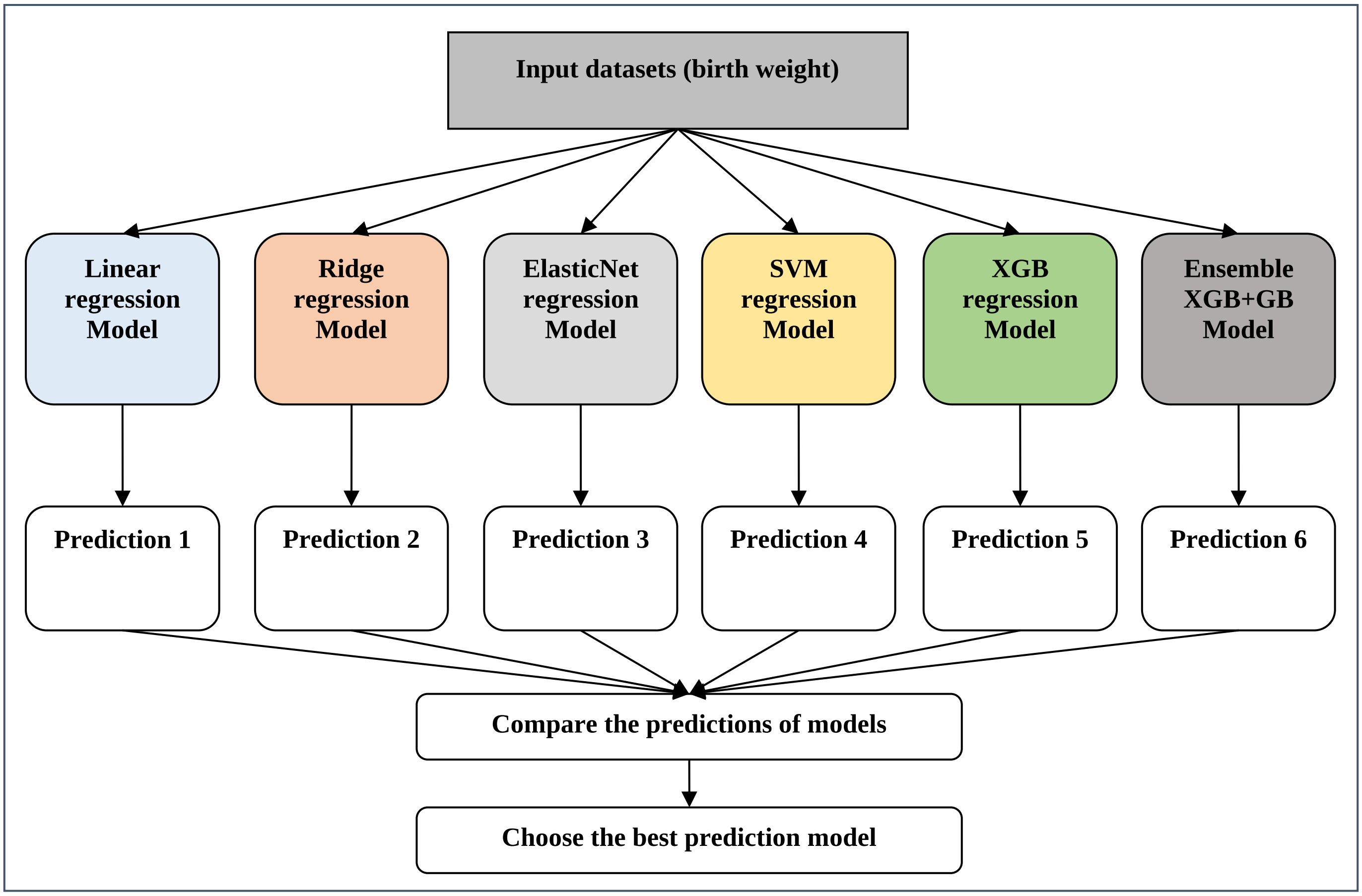

Fig. 2 Block diagram of research model

The block diagram depicts the current study process. The architecture of the ensemble model developed explains the model process and how the data is processed. The model developed is an ensemble model that uses the voting regressor architecture with the neural networking (NN) of sequential layers in deep learning. Majorly, most used deep learning models with sequential layers adopt the CNN (convolutional), RNN (recurrent NN), MLP (multi-layer perceptron) and LSTM (long-short-term memory) based architectures. The current model was developed in an ensemble model (hybrid) that uses the decision trees-based gradient boosting (GB) and extreme gradient boosting (XGB) algorithms as voting regressors for predictions (refer to Figure 3). The predictions from the ensemble model are examined, and the voting regressor analyses the predictions using the algorithms. These predictions are then finalized as one result, which is later identified by the model as the 'final prediction' and considered the model's output. Thus, the outputs are gained, and the model's performance is evaluated using the metric evaluation technique. The results obtained are then compared against the existing regressor model to measure the model's performance.

3.3. Algorithm Model

Ensemble model using two different algorithms named Voting Regressor where the XGB-Regressor and GB-Regressor are combined as one model and used to predict the birth weight of a fetus.

Algorithm 1: Voting regressor (XGB+GB)

Initialize: Constant value for the prediction model
r1 = XGB regressor,
r2 = RF regressor,
r3 = GB regressor,
Voting regressor:
Train XGB:
Step 1: Initialize the model by loading the libraries;
Step 2: Generate the estimator with a size of 1000;
Step 3: Set time as 0.1 with column sample by-tree as 0.8 and sub-sample as 0.7;
Step 4: The maximum depth value is set to 10, and
Step 5: Set the enable categorical parameter to false.
Train GB:
Step 1: Generate the estimator with a size of 500 with a maximum depth size of 10 and
Step 2: Receive the predicted outcomes and vote for the most accurate outcome.

By using the hybrid approach, the model is presumed to outperform the existing birthweight prediction models.

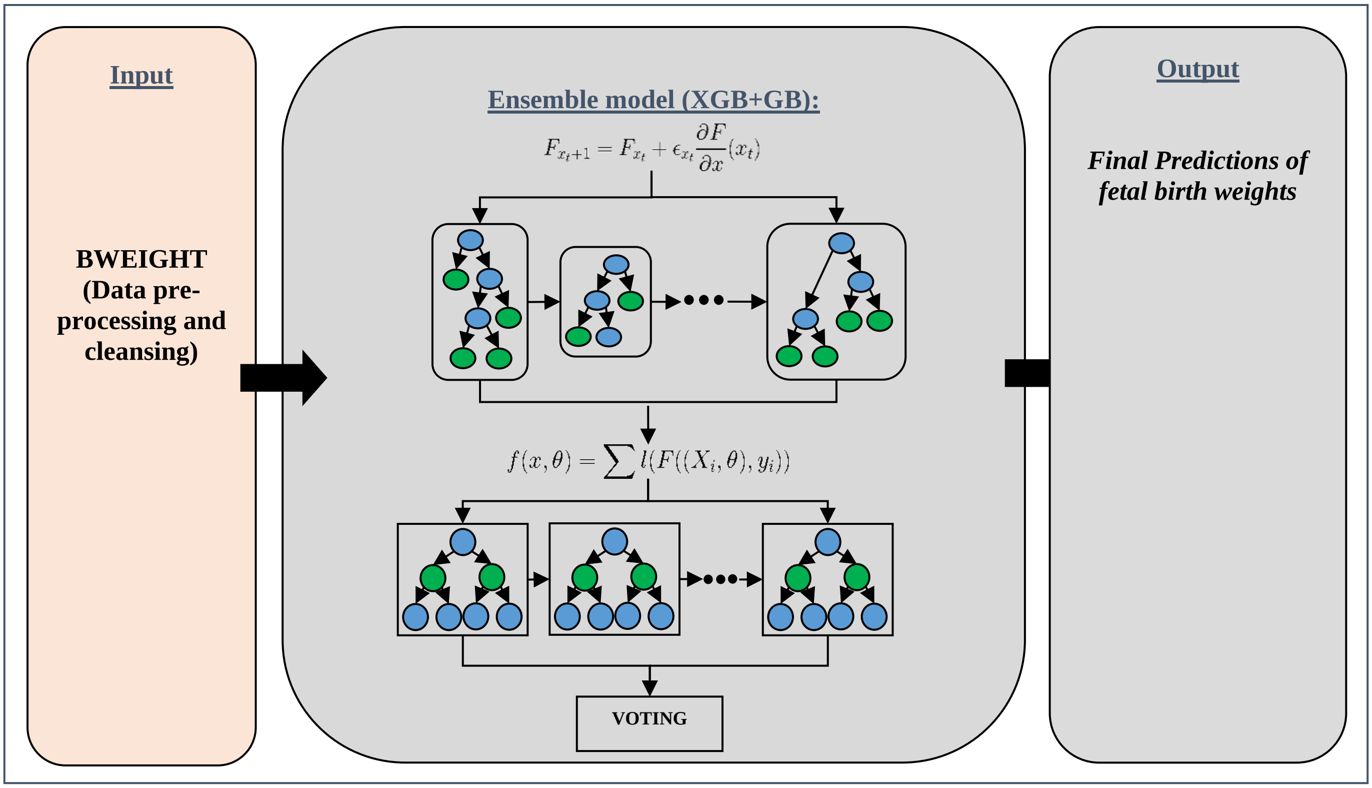

Fig. 3 Architecture of research model

3.4. Dataset

The dataset has been accessed from "Kaggle" and gathered and pre-processed by the author [36]. The database has one .txt and two .csv files as extensions where the .csv includes two folders, namely: "baby-weights-dataset" and "judge-without-labels". In this study, "baby-weights-dataset" with the label "BWEIGHT" is used, but another dataset is not used. This file has thirty-seven variables that have been used here for birthweight determination.

Data split: The data used here is split into 70:30 (train: test) with the random state value set as 99. The dataset includes a total of 101400, whereas per the data split, training includes 70976, while the testing includes 30424, respectively.

Units: The study examines the weight as a metric value, and hence, it utilizes the "lbs" as a unit here. The same can be converted into kilo-weights (kg) where 1lbs equals 0.4535kgs. The unit conversion is applicable for studies that use "kg" units.

Parameters: The parameters used here are thirty-seven individual entries (variables). They are: ID, MARITAL, SEX, MAGE, FAGE, VISITS, GAINED, MEDUC, FEDUC, TOTALP, TERMS, BDEAD, LOUTCOME, RACEDAD, HISPDAD, RACEMOM, HISPMOM, CIGNUM, DRINKNUM, ANEMIA, CARDIAC, ACLUNG, DIABETES, HERPES, HYDRAM, HEMOGLOB, WEEKS, HYPERPR, HYPERCH, CERVIX, ECLAMP, PINFANT, RHSEN, RENAL, PRETERM, UTERINE, and BWEIGHT. From these 37 variables, the collective data like mother's health condition, diseases, abortion, children alive, children dead, habits, race/ ethnicity, fathers' characteristics, race, and more to determine the fetus's weight where Hispanic race has an impact on the fetal weight. The non-numeric data (ordinal and nominal) values are transformed into numeric data for easy examination.

4. Experimental Results

4.1. Implementation

The implementation of the designed ensemble model includes two major categories, namely the adaptation of the software and hardware necessity and the implementation of the model and testing the predictions.

4.1.1. Software and Hardware Requirements

The model designed here uses "Python" as the software, which is an open language that is free to use. The software adopted is user-friendly, easy to understand, and works well on many OSs. To design the ensemble, model the following libraries are used in Python, namely: Panda (for extraction and manipulation of vectorized data operations), Numpy (core library for rapid scientific computations), Seaborn (used for visualizing the statistical graphs) from Matplotlib and Pytorch (for computer vision-based operations). Table 2 presents the specification requirements for the experiment.

Table 2. Specification requirements

System: Operating System	Advanced OS such as Windows 10, MacOS, Windows 11, and more are utilized in large datasets-based operations.
Processor: Central Processing Unit	Higher-end CPU is generally adopted in birthweight research due to data analysis complications. Here, the Intel core of Intel-i7-10750H is adopted.
Memory: Random Access Memory	8GB to 16GB is used by researchers for large data analyses. Here, 16GB is used.
Graphic Processing Unit	8GB is necessary for better performance. NVIDIA-2080 with 8GB GPU is used here.
Storage capacity	To reduce errors and system crashes for large data processing, large storage is preferred. 1TB NVME is used here.

4.1.2. Model Implementation

The ensemble model developed uses two algorithms, namely XGB and GB, where the boosting parameters are set as mentioned in the algorithm. The random forest classifier is set with the number-of estimator as 10 and the random state as 1. By using these conditions and limits, the algorithm is developed, and the model is implemented to predict the birth weights of the infants.

Table 3 presents the training dataset samples for this research. During the training phase, the non-numeric columns are transformed into numeric values in Python for accurate estimations. Transforming non-numeric values (M & N) into numeric values (1 & 2), respectively, for the parents' race in this research as samples are given in Table 4. From Table 4, the data are converted from nominal values and the model is trained only with numeric datasets for accurate and reliable results.

Table 3. Training dataset samples

Variables	0	1	2	3	4
ID	2001	2002	2003	2004	2005
SEX	2	2	2	1	1
MARITAL	1	2	1	1	2
FAGE	33	19	33	25	21
GAINED	26.0	40.0	16.0	40.0	60.0
VISITS	10	10	14	15	13
MAGE	34	18	31	28	20
FEDUC	12.0	11.0	16.0	12.0	12.0
MEDUC	4	12	16	12	14
TOTALP	2	1	2	3	2
...	...	...	...	...	...
HYPERCH	0	0	0	0	0
HYPERPR	0	0	0	0	1
ECLAMP	0	0	0	0	0
CERVIX	0	0	0	0	0
PINFANT	0	0	0	0	0
PRETERM	0	0	0	0	0
RENAL	0	0	0	0	0
RHSEN	0	0	0	0	0
UTERINE	0	0	0	0	0
BWEIGHT	4.3750	6.9375	8.5000	8.5000	9.0000

Table 4. Transformation of non-numeric data

	HISPMOM	HISPDAD	Transformed Into	
0	M	M	1	1
1	N	N	2	2
2	N	N	2	2
3	N	N	2	2
4	N	N	2	2
…	…	…	…	…
101395	M	M	1	1
101396	N	N	2	2
101397	N	N	2	2
101398	N	N	2	2
101399	N	N	2	2

4.2. Testing

The batch size is set as 64, with the model's training parameters, namely, Train = false and Shuffle = true. The model loss is estimated using the MSE loss estimation, whereas the accuracy is estimated using the R2 score. The linear regression model is tested with no additional parameters. However, the ridge regressor has an alpha value of 30.0. The ElasticNet's alpha value is set as 0.1 with 0.9 as the l1_ratio value. The SVM regressor model is trained and tested as it is without any altered values. The XGB model is trained and tested with n-estimator as 10000, maximum depth value as 10 with eta (time) 0.l. The subsample value is set as 0.7 with colsample_bytree as 0.8.

As mentioned earlier in model implementation, the ensemble model incorporates the XGB (with the same values as the XGB model) and GB regressors. The hyper-parameters of XGB such as colsample_bytree (0.8), enable_categorical (false), eta (0.1), max_depth (10), missing (nan), n-estimator (1000), and subsample (0.7) are set in the hybrid model with all other hyper-parameters as 'none' (no values). Similarly, the GB algorithm's hyper-parameters, namely n-estimators (500) and max_depth (10), are set in this model for prediction. Based on these hyper-parameters, the model classifies and votes the most accurate result for fetuses' birth weight in their mothers' wombs.

4.3. Data Analysis

The birth weight prediction models examined here are compared with their R2 scores. The R2 is estimated in [39] by employing the following equation (refer to equation 1):

$$R2 - Score\ (r2) = 1 - \frac{SQER}{SQEM} \tag{1}$$

Where:

$$SQER = \sum_a (\hat{x}_a - \bar{x})^2 \tag{2}$$

{SQER= sum-squared total-error; $\sum$ = data-points sums; (x_a) = data-point, $(\bar{x})$ = mean-value and $\widehat{(x)}$ = value predicted for observations for 'a'}

$$SQEM = \sum_a (x_a - \bar{x})^2 \tag{3}$$

{SQEM= sum-squared mean-value; $\sum$ = data-points sums; (x_a) = data-point observed, $(\bar{x})$ = mean-value}. The r2 scores of the birthweight models (refer to Table 5) are estimated using Equation 1 [37].

Table 5. Performance analysis of r2 scores

Model	Training	Testing
Linear Regressor	0.3855	0.3902
Ridge Regressor	0.3855	0.3902
ElasticNet Regressor	0.3391	0.3418
SVM Regressor	0.4757	0.4292
XGB Regressor	0.7999	0.7773
Ensemble (XGB+GB)	0.9039	0.8959

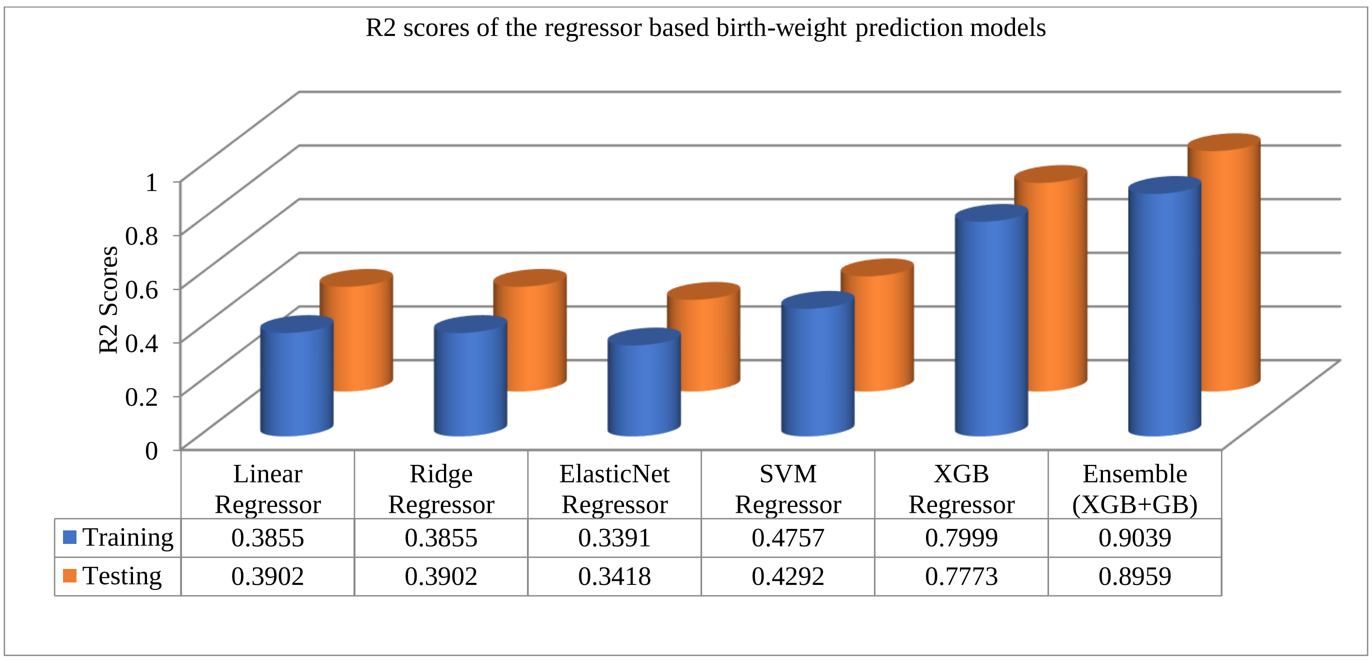

Fig. 4 Birthweight prediction regressor models

The performance of the ensemble model developed during the training phase is 0.9039, whereas during the testing phase is 0.8959. The model thus achieved 90% accuracy, which is higher than the other five models examined to predict the birth weights of the fetuses. From the analysis and the findings of the results obtained, it is evident that the six regressor models used here predicted the fetal birthweight using deep learning algorithms. However, the developed model achieved higher accuracy than existing regressors, which insists that either optimizing a model or combining two or more algorithms gains higher accuracy than original ML models.

Authors in [29] examined the birth weights of fetuses using ML algorithms and found that decision tree methods such as random forest are effective in estimating the birth weights. Authors in [42] used the XGBoosting and GB algorithms to find the accurate models that had more accuracy and speed with different datasets. They found through their analysis that XGB is more accurate than GB with less error. Similarly, the authors in [17] used different ML algorithms based on eight models where XGB and GB are two of the methods that predicted the birth weight of fetuses. The study found that the XGB model with the ET classifier achieved higher accuracy than the other seven models [38].

Authors in [40] compared different gradient boosting algorithms and found GB as the most effective while XGB was the second-most effective algorithm. In the study conducted by authors in [29], it was found that decision trees in predicting the birthweight in infants and fetuses are found to be effective rather than adopting oversampling and feature selection methods since they are not reliable. According to the study by Henseler *et al.*, [41] the R2 score that obtains ≤.40→≤.50 shows that, the model developed is poor. Whereas the R2 score between the values of ≥.50→≤.75 is considered significant, and the model is considered good. The R2 score of ≥.75 shows that the developed prediction model is significantly strong and reliable. As per the standard, the developed model obtained an R2 score of .90, which proves that the model is efficient, effective, and significant in examining and predicting the birthweights of fetuses. Thus, based on these findings and observations, the study proposed a hybrid approach to the decision tree method-based prediction model. The XGB + GB algorithm is used in an ensemble model using a voting regressor approach, and the results are measured based on the model's performance [42].

4.4. Performances of the ML Models are Estimated Using R^2 scores. The R^2 scores of each model used here are

1. Linear regression model with the r2 score of 0.385542939400517 (training) and 0.39022512274705534 (testing);
2. Ridge regressor model with the r2 score of 0.38554287665370224 (training) and 0.39022246506412195 (testing);
3. ElasticNet regressor model with the r2 score of 0.33911152197211514 (training) and 0.3418611611202438 (testing);
4. SVM regressor model with the r2 score of 0.47570916421169196 (training) and 0.42921858238606214 (testing);
5. XGB regressor model with the r2 score of 0.7999997295324917 (training) and 0.7773414286036746 (testing); and
6. Ensemble (XGB+GB) model with the r2 score of 0.9039392149835611 (training) and 0.8959362743855217 (testing).

5. Conclusion

Birthweight prediction using ML models has been increasing rapidly due to its accuracy and the advanced technologies in the medical field. Earlier in the 1800s, using variables such as the mother's weight, weeks of pregnancy, mothers' health attributes and ethnicity, the weights were approximately assumed by the medical practitioners until the medical field implemented the machine learning models for birth weight predictions. The current study focuses on comparing different regressors that are majorly adopted to predict the birth weights of the fetuses using several parameters as inputs. The study obtained data from Kaggle where the inputs of the mother (health attributes, maternity profile: age, preterm, stillbirths, childbirth records and so on), father (age, race, and ethnicity, and more) and more related to the pregnancy are gathered, cleansed, and pre-processed.

Mostly, the birthweight prediction models use the decision tree method by adopting one algorithm per model. Hence, in this study, the researcher attempted a hybrid approach of using XGB + GB in the ensemble prediction model. The results showed that the XGB+GB model achieved 90% accuracy, whereas other models gained ≤.50 (i.e. <50%) except the XGB regressor model (77%). Thus, clearly shows that, for the prediction of birthweights in fetuses, using the decision tree method, especially XGB+GB, produces higher accuracy than other regressor models. Thus, it's also evidently concluded that a hybrid approach is better among the gradient boosting methods in ML than normal boosting methods. The research focuses on a single dataset; further, it can be extended by including multiple diverse datasets from various populations and healthcare systems.

Acknowledgment

The authors would like to thank the Reva University, Bengaluru and Presidency University, Bengaluru, for their constant support throughout the completion of this work.

References

[1] Deping Hu et al., "Inclusion of Machine Learning Kernel Ridge Regression Potential Energy Surfaces in On-the-Fly Nonadiabatic Molecular Dynamics Simulation," *The Journal of Physical Chemistry Letters*, vol. 9, no. 11, pp. 2725-2732, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[2] Tasneem Binte Reza, and Nahid Salma, "Prediction and Feature Selection of Low Birth Weight Using Machine Learning Algorithms," *Journal of Health, Population and Nutrition*, vol. 43, no. 1, pp. 1-13, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[3] Jagjit S. Teji et al., "Neoai 1.0: Machine Learning-Based Paradigm for Prediction of Neonatal and Infant Risk of Death," *Computers in Biology and Medicine*, vol. 147, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[4] M. Poongodi et al., "New York City Taxi Trip Duration Prediction Using MLP and XGBoost," *International Journal of System Assurance Engineering and Management*, vol. 13, no. S1, pp. 16-27, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[5] Alfensi Faruk et al., "Prediction and Classification of Low-Birth-Weight Data Using Machine Learning Techniques," *Indonesian Journal of Science and Technology*, vol. 3, no. 1, pp. 18-28, 2018. [Google Scholar] [Publisher Link]

[6] Amene Ranjbar et al., "Machine Learning-Based Approach for Predicting Low Birth Weight," *BMC Pregnancy and Childbirth*, vol. 23, no. 1, pp. 1-7, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] S. Ramraj et al., "Experimenting Xgboost Algorithm for Prediction and Classification of Different Datasets," *International Journal of Control Theory and Applications*, vol. 9, no. 40, pp. 651-662, 2016. [Google Scholar]

[8] Khushbu Kumari, and Suniti Yadav, "Linear Regression Analysis Study," *Journal of the Practice of Cardiovascular Sciences*, vol. 4, no. 1, pp. 33-36, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[9] Christopher Freyder, "*Using Linear Regression and Mixed Models to Predict Health Care Costs After an Inpatient Event*," Master's Thesis, University of Pittsburgh, pp. 1-43, 2016. [Google Scholar] [Publisher Link]

[10] Rahul Pathak et al., "Infant Weight Prediction Using Multivariate Linear Regression," *International Journal of Research and Analytical Reviews*, vol. 5, no. 4, pp. 158-164, 2018. [Google Scholar]

[11] S.S. Pooja, and Gladrene Sheena Basil, "Students Performance Prediction and Classification Using Ridge Regression Algorithm," *International Journal of Advanced Trends in Engineering and Management*, vol. 8, no. 112, pp. 1006-1014, 2023. [Publisher Link]

[12] Saleem Adeeba, Banujan Kuhaneswaran, and B.T.G.S. Kumara "Prediction of Weight Range of Neonate Using Machine Learning Approach," *2022 International Conference on Innovations in Science, Engineering and Technology*, Chittagong, Bangladesh, pp. 427-432, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] Wessel N. van Wieringen, "Lecture Notes on Ridge Regression," *Arxiv*, pp. 1-149, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[14] Gaurav Kumar et al., "Application of Machine Learning in Predicting the Yield Strength of API Steels," *European Chemical Bulletin*, vol. 12, no. 7, pp. 1639-1653, 2023. [Publisher Link]

[15] Dina A. Alabbad et al., "Birthweight Range Prediction and Classification: A Machine Learning-Based Sustainable Approach," *Machine Learning and Knowledge Extraction*, vol. 6, no. 2, pp. 770-788, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[16] Mohammed Al-Shawwa, and Samy S. Abu-Naser, "Predicting Birth Weight Using Artificial Neural Network," *International Journal of Academic Health and Medical Research*, vol. 3, no. 1, pp. 9-12, 2019. [Google Scholar] [Publisher Link]

[17] Chen Bo et al., "A Method for Estimating Fetal Weight Based on Body Composition," *The Journal of Maternal-Fetal & Neonatal Medicine*, vol. 32, no. 19, pp. 3306-3314, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[18] Cheng Gao et al., "Deep Learning Predicts Extreme Preterm Birth from Electronic Health Records," *Journal of Biomedical Informatics*, vol. 100, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[19] Pradeeba Sridar et al., "Decision Fusion-Based Fetal Ultrasound Image Plane Classification Using Convolutional Neural Networks," *Ultrasound in Medicine & Biology*, vol. 45, no. 5, pp. 1259-1273, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[20] Zakir Hussain, and Malaya Dutta Borah, "Birth Weight Prediction of Newborn Baby with Application of Machine Learning Techniques on Features of Mother," *Journal of Statistics and Management Systems*, vol. 23, no. 6, pp. 1079-1091, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[21] Yu Lu et al., "Ensemble Machine Learning for Estimating Fetal Weight at Varying Gestational Age," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 1, pp. 9522-9527, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[22] Emily J. MacKay et al., "Application of Machine Learning Approaches to Administrative Claims Data to Predict Clinical Outcomes in Medical and Surgical Patient Populations," *PLoS One*, vol. 16, no. 6, pp. 1-14, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[23] Zahra Hoodbhoy et al., "Use of Machine Learning Algorithms for Prediction of Fetal Risk Using Cardiotocographic Data," *International Journal of Applied and Basic Medical Research*, vol. 9, no. 4, pp. 226-230, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[24] Nabillah Rahmayanti et al., "Comparison of Machine Learning Algorithms to Classify Fetal Health Using Cardiotocogram Data," *Procedia Computer Science*, vol. 197, pp. 162-171, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[25] Oliver Campos Trujillo, Jorge Perez-Gonzalez, and Verónica Medina-Bañuelos, "Early Prediction of Weight at Birth Using Support Vector Regression," *VIII Latin American Conference on Biomedical Engineering and XLII National Conference on Biomedical Engineering*, Cancún, México, pp. 37-41, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[26] Miao Feng et al., "Fetal Weight Estimation via Ultrasound Using Machine Learning," *IEEE Access*, vol. 7, pp. 87783-87791, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[27] Wasif Khan et al., "Infant Birth Weight Estimation and Low Birth Weight Classification in United Arab Emirates Using Machine Learning Algorithms," *Scientific Reports*, vol. 12, no. 1, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[28] Elena Quatrini et al., "Machine Learning for Anomaly Detection and Process Phase Classification to Improve Safety and Maintenance Activities," *Journal of Manufacturing Systems*, vol. 56, pp.117-132, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[29] Wondesen Teshome Bekele, "Machine Learning Algorithms for Predicting Low Birth Weight in Ethiopia," *BMC Medical Informatics and Decision Making*, vol. 22, no. 1, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[30] Mahya Arayeshgari et al., "Machine Learning-Based Classifiers for the Prediction of Low Birth Weight," *Healthcare Informatics Research*, vol. 29, no. 1, pp. 54-63, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[31] Jung Ho Han et al., "Application of Machine Learning Approaches to Predict Postnatal Growth Failure in Very Low Birth Weight Infants," *Yonsei Medical Journal*, vol. 63, no. 7, pp. 640-647, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[32] Wei-Ting Lin et al., "Predicting in-Hospital Length of Stay for Very-Low-Birth-Weight Preterm Infants Using Machine Learning Techniques," *Journal of the Formosan Medical Association*, vol. 121, no. 6, pp. 1141-1148, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[33] P. Thabrez, and J. Srinivasan, "A Novel Approach Low Birth Weight Prediction Using Machine Learning Techniques," *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, vol. 8, no. 4, pp. 272-276, 2022. [Google Scholar] [Publisher Link]

[34] S.M. Ashikul Islam Pollob et al., "Predicting Risks of Low Birth Weight in Bangladesh with Machine Learning," *PLoS One*, vol. 17, no. 5, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[35] Jing Tao et al., "Fetal Birthweight Prediction with Measured Data by a Temporal Machine Learning Method," *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, pp. 1-10, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[36] Ashis CU Denver, Birth Weight Prediction, *Kaggle*, 2021. [Online]. Available: https://www.kaggle.com/c/birth-weight-prediction

[37] Davide Chicco, Matthijs J. Warrens, and Giuseppe Jurman, "The Coefficient of Determination R-Squared Is More Informative Than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation," *PeerJ Computer Science*, vol. 7, pp. 1-24, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[38] Candice Bentéjac, Anna Csörgő, and Gonzalo Martínez-Muñoz, "A Comparative Analysis of Gradient Boosting Algorithms," *Artificial Intelligence Review*, vol. 54, no. 3, pp. 1937-1967, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[39] Niaz Muhammad Shahani et al., "Developing an Xgboost Regression Model for Predicting Young's Modulus of Intact Sedimentary Rocks for the Stability of Surface and Subsurface Structures," *Frontiers in Earth Science*, vol. 9, pp. 1-13, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[40] Tianqi Chen, and Carlos Guestrin, "Xgboost: A Scalable Tree Boosting System," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, California, USA, pp. 785-794, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[41] Jörg Henseler, Christian M. Ringle, and Rudolf R. Sinkovics, "The Use of Partial Least Squares Path Modeling in International Marketing," *New Challenges to International Marketing*, vol. 20, pp. 277-319, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[42] Hemlata Sharma, Hitesh Harsora, and Bayode Ogunleye, "An Optimal House Price Prediction Algorithm: XGBoost," *Analytics*, vol. 3, no. 1, pp. 30-45, 2024. [CrossRef] [Google Scholar] [Publisher Link]