JCSDA, Vol. 01, No. 2, 1–19
DOI: 10.69660/jcsda.01022401
ISSN 2959-6912

Predictive Machine Learning Model for Low Birth Weight in Newborns

Serkalem Negusse Kassaye
Department of Information Technology and Computer Science,
Kotebe University of Education, Addis Ababa, Ethiopia.
Corresponding author: serkicho@gmail.com

Abebe Alemu Balcha
Department of Computer Science and Engineering,
Adama Science and Technology University, Adama, Ethiopia.

Kula Kakeba
HPC and Big Data Analytics Center of Excellence,
Department of Software Engineering College of Engineering,
Addis Ababa Science and Technology University, Ethiopia.

Samuel Alemu
Department of Information Technology and Computer Science,
Kotebe University of Education, Addis Ababa, Ethiopia.

Low birth weight (LBW) occurs when a newborn weighs less than 2500 grams regardless of the gestational age of the infant. LBW is one of the indicators of complex public health problems of an infant, which is 25 times more likely to die than those at expected birth weight. The neonatal mortality rate measures the number of neonates dying before reaching 28 days of age per 1,000 live births within a given year. It affects one out of every seven newborns, about 14.6 percent of babies born worldwide. The prevalence is 7.2% in developed regions and 13.7% in Africa. The neonatal mortality rate in Ethiopia is 29.524 deaths from 1000 live births in 2023. Thus, building accurate LBW prediction models and finding the related risk factors is critical. The early identification and prediction of such disease would reduce premature death rate caused by LBW. The exponentially increasing data availability on the web has a significant role in extracting better insight, and it is vital to develop machine learning models. The current technology facilitates and supports professionals in health care by providing services to societies. Classification techniques help to classify the case according to a certain feature from data and to predict the probabilities of LBW in infants. We have used a public dataset obtained from the Ethiopia Demographic Health Survey to build models. In this work, artificial neural networks and support vector machines are used for training and model building. Moreover, data preparation techniques like scaling and transformation, hyperparameter tuning approaches like GridSearch CV, and feature engineering techniques have been employed and tested. Based on performance evaluation results, the proposed classifiers were capable of predicting of LBW. The neural network with hyperparameter tuning techniques scored an accuracy of over 97.2% with 98.0% of sensitivity, 97.0% of precision, 96.0% of specificity, and an area under the curve (AUC) of 99%.

Keywords: Feature engineering, Hyperparameter tuning, Low birth weight, Machine Learning, Neural Networks, Prediction model, Support Vector Machine.

1. Introduction

Birth weight is a critical health status indicator of an infant besides being one of the principal factors that determines the infant's physical, survival, and mental growth. It also reflects the maternal health status. A UNICEF report from March 2023 highlighted a 25% increase in pregnant and breastfeeding women suffering from acute malnutrition across 12 countries in Africa and Asia, including Ethiopia, since 2020 [1]. Low birth weight (LBW) is defined as infants born weighing less than 2500 grams, regardless of gestational age, and poses a significant public health challenge in developing nations, particularly in sub-Saharan Africa. Around 20 million infants worldwide are born with LBW each year, representing 17% of all births in developing countries, with 6% in industrialized nations and 21% in developing regions globally [2], [3] , [4]. LBW serves as a critical indicator of maternal and fetal health, predicting mortality, stunting, and chronic conditions in adulthood. LBW infants are 25 times more likely to die than those born at expected weights. The neonatal mortality rate, measuring the number of neonates dying before 28 days per 1,000 live births within a year, impacts one in seven newborns globally, accounting for approximately 14.6% of all births. The prevalence of LBW varies substantially by region, with 7.2% in developed regions and 13.7% in Africa. Despite a gradual decline in Ethiopia's neonatal mortality rate from 47.8 deaths per 1,000 live births in 2001 to 27 deaths in 2020, the country still bears a significant burden of LBW, ranking among the highest in Africa. In 2023, Ethiopia's infant mortality rate decreased by 4.82% compared to 2022, with 29.524 deaths per 1,000 live births. It is estimated that an infant dies every 10 seconds due to LBW-related complications [3], [4], [5], [6].

Although developments in neonatal care and maternal health, LBW remains a significant public health concern across the globe, particularly in low and middle-income countries. Various literature consistently identifies LBW as a major predictor of infant mortality. However, there is a critical gap in understanding the interaction of socio-economic, environmental, and healthcare access factors contributing to LBW and its subsequent impact on infant mortality rates. In light of machine learning-based approaches for LWB, current studies often lack comparative analysis, generalization concerns, evaluation metrics, data quality and size, and technical implementation on model architecture and training process. This gap hinders the development of LBW interventions aimed to reduce and eradicate LBW and improve survival outcomes for newborn infants [7], [8], [9]. This study aims to develop a predictive model using machine learning (ML) techniques to identify factors influencing LBW in newborns. Specifically, we employed artificial neural networks and support vector classifiers to predict LBW among infants. Some researchers have achieved promising results by applying ML techniques to predict LBW in infants [4], [8], [9], [10], [11]. This work highlights the application of artificial intelligence in creating effective predictive models in this area.

The main objective of this study is to develop a predictive model using machine learning techniques to evaluate the factors influencing low birth weight (LBW). This model aims to enhance the identification and management of LBW infants, ultimately reducing neonatal mortality rates and improving overall healthcare outcomes in the field of maternal and child health. In general, some key contributions of the work included:

- The proposed ML model architecture for LBW prediction.
- Preprocessing Techniques to employ a comprehensive preprocessing pipeline to ensure high-quality input data for our model.
- Hyperparameter tuning and optimization techniques to fine-tune the proposed LBW model and maximize its predictive accuracy.
- Enhanced prediction accuracy.
- LBW prediction model
- Enhanced LBW prediction accuracy

The subsequent sections of the study were organized as follows: Section II presented related works previously conducted in the area, Section III discussed the methods and materials used in the study, Section IV covered the evaluation and results, and Section V concluded with conclusions and future directions of the research.

2. Related works

Numerous studies have explored the use of ML classifiers for predicting birth weight, with varying approaches in preprocessing and hyperparameter tuning. The study [4] applied the machine learning Gaussian Naïve Bayes and Random Forest classifier models on a dataset of 445 instances only, and through data analysis the classifiers model Gaussian Naïve Bayes scored 86% and the Random Forest scored perfect accuracy. The study utilized a preprocessing technique to handle missing values through mean values and it lacked applying vast amount of data, and an advanced preprocessing techniques like data normalization, hyperparameter tuning.

The researcher [7] attempted to construct a predictive model using an artificial neural network for birth weight prediction based on maternal features and pregnancy-related factors, the ANN model scored 100% accuracy. However, the study lacked clear presentation of preprocessing techniques that utilized in the research. Additionally, the study is not explicitly mentioned the dataset size. It's uncertain to generalize the model reported a genuine predictive capability because perfect accuracy may score due to model overfitting and or small dataset also lead to it. A study conducted by [10], introduced a novel deep learning model, deep Preterm Birth Survival Risk monitoring by utilizing long short-term memory – LSTM. The study achieved an accuracy of 0.88, recall of 0.78, and AUC of 0.897 using 285 infant datasets from the neonatal intensive care unit at St. Louis Children's Hospital, Washington University. Therefore, these researchers have not considered the possible features used for predicting LBW.

The study conducted by [12] presented a study comparing the performance of random forest and binary logistic regression algorithms in predicting low birth weight using data from the 2012 Indonesian Demographic and Health Survey. The results indicated that the random forest classification approach outperformed the binary logistic regression model in predicting low birth weight. It is very important to implement advanced machine learning approaches. The very recent study [13] utilized 1863 instances of dataset sourced from BDHS 2017-2018. Considerably larger dataset utilized in this study compared to some other studies. The study utilized multiple ML classifiers including Random Forest, Support Vector Machine, and XGBoost, to build effective model for predicting low birth weight. To identify the significant maternal and demographic predictors, the study utilized a feature selection techniques Boruta and Wrapper. Despite of the model performance evaluation, the Random Forest classifier scored 85.86% model accuracy. However, the study lacked applying hyperparameter tuning like GridSearchCV and optimization techniques.

The conference paper by [14], proposes a regressor Machine learning model. The researchers utilized gestational, perinatal factors to investigate LBW. A comparison was done with previous works from the results, using the Dichotomized classification LR, RF, and SVC. The performance of the regression and classification improved significantly, from R^2 0.23 to 0.74 and from the area under the roc curve (AUC) 0.85 to 0.94 respectively. However, the process had limitations due to the utilization of numerous features. A study by [15] analyzed multiple ML algorithms including random forest (RF) to predict the probability of LBW in Ethiopia by using the 2016 EDHS dataset. The algorithm RF scored 91.6% model accuracy, 96.8% ROC-AUC, and 91.6% recall. The researcher also identified the key predictors such as child's gender, marriage-to-birth interval, mother's occupation, and age. However, limitations included a small sample size (14% of surveyed births) and class imbalance. Finally, the study recommends using larger, updated datasets and exploring additional socio-economic factors for improved generalizability.

In a study [16] five machine learning models such as decision trees, random forest, artificial neural networks, support vector machines, and logistic regression were evaluated for predicting LBW using maternal and neonatal data. The Logistic regression performed better and scored 88% model accuracy. The study emphasized the importance of improved prenatal care and genetic counseling to mitigate LBW. However, the study is limited by the use of a dataset from a single region, emphasizing the need for validation with larger, multi-regional data. A recent study [17] evaluated eight ML models for LBW prediction and found the XGBoost model performance became best with an accuracy of 79% and 87% precision. The key predictors identified were gestational age and prior LBW history. The study underscores the ML utility in maternal care, it highlights the necessity for broader datasets to improve generalizability. Another study [18] used multiple ML algorithms for classifying the probability of LBW in newborns. Algorithm Logistic regression with SMOTE for LBW classification, achieving 90% accuracy, 87.6% precision, and 90.2% recall but noted limitations such as class imbalance and small sample data size, recommends larger, diverse datasets and enhanced data quality measures.

Therefore, existing research literature has explored comprehensive predictive models that can accurately identify infants at risk of low birth weight, comprehensive models with high accuracy and broad applicability remain limited. This study aimed to bridge this gap by developing a machine-learning model that can predict low birth weight with high accuracy, thereby enabling early intervention and improved health outcomes for newborns.

3. Material and methods

3.1. *Dataset collection and preprocessing*

For this study, we utilized the LBW dataset from the publicly available Ethiopia Demographic and Health Survey repository. This dataset comprises 10,641 instances and encompasses both newborn and maternal features. Relevant features were selected and utilized during experimentation.

The large volume of data used in our study necessitated extensive preprocessing techniques to address potential inconsistencies, missing values, and noise. Even insignificant errors in the dataset could significantly impact model performance. One prevalent issue we encountered was missing values, which required careful handling before further analysis. Before addressing missing values, it was important to normalize the data using the MinMAx Scalar. This transformation technique scaled features to a specified range of values, thereby ensuring uniformity and facilitating more effective modelling.

The MinMaxScaler is particularly appropriate for the Ethiopia Demographic and Health Survey (EDHS) dataset due to its ability to preserve the original range of values, which is vital when dealing with health-related metrics that have specific clinical significance. Moreover, the MinMax Scaler helps in handling the diverse scales of various features, ensuring that all variables contribute equally to the predictive model's performance. This normalization process played a critical role in preparing the dataset for subsequent modelling farm duties. By standardizing the data, we alleviated potential biases and inconsistencies, thus enhancing the reliability and accuracy of our analyses. Therefore, by systematically cleaning and transforming the data, we uncover valuable insights and inform evidence-based decision-making in LBW prediction in healthcare, and other domains.

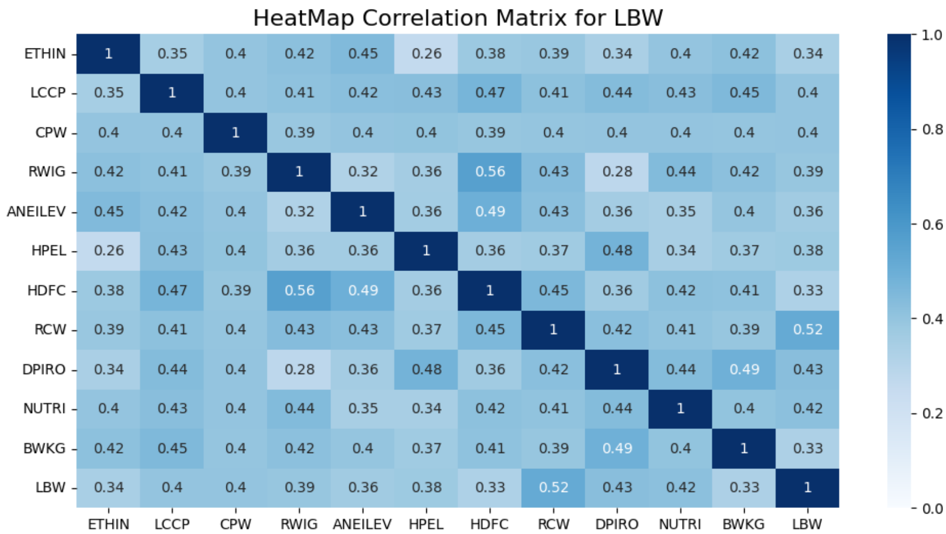

Fig. 1. Heat-Map Correlation matrix for LBW

3.2. *Feature selection*

LBW is defined as the birth weight of a new infant below 2500 grams. Based on the threshold we labeled each infant indicating the presence or absence of LBW. Specifically

- Infants with a birth weight $<$ 2500 grams were labelled as '1', indicating the presence of LBW;
- Infants with a birth weight $\geq$ 2500 grams were labelled as '0', indicating the absence of LBW.

After labeling the data, feature selection is important to improve the model's performance. The heatmap correlation analysis (Fig 1) was used to refine the feature selection process. By visualizing the correlation between independent features and the target variable LBW, we can identify the most significant features. This allows us to prioritize features with high correlation values, providing valuable insights into which variables are most relevant for predicting LBW in newborn infants.

3.3. *Classifiers*

SVMs and ANN algorithms are well-suited for healthcare applications due to their ability to handle complex data, and high-dimensional features, learn complicated patterns, and provide accurate and valuable predictions. The state-of-the-art performance in healthcare tasks like disease diagnosis makes them well-suited for predicting LBW among newborns using the EDHS dataset. Their flexibility with kernel functions enables them to capture feature relationships in newborn and maternal

Table 1. Proposed ML algorithms

Algorithm	**Description**
Support Vector Machine (SVM)	The SVM classifies the radial and linear kernel of the support vector classifier with randomly assigned cost and gamma parameters for tuning the hyperparameter. Hyperparameter tuning for the RBF kernel classifier involved defining the C: 1, 10, 100 and Gamma: 0.01, 0.001, 0.0001 to find the optimum hyperplane. The GridSearch Cross Validation (GridSearchCV) was employed to search for the optimum hyperparameter from the grid through model fitting. The dataset was initially split into training and test sets, and the training set was divided into actual train and validation sets with a (70+10)%, 20% ratio respectively. The training set comprised 70% of the total samples and was used for model building, while the validation set constituted 10% and was utilized for tuning the hyperparameters and validating the model during training. The remaining 20% formed the test set for final evaluation.
Artificial Neural Network (ANN)	In the artificial neural network used in LBW prediction, the model was trained using 16 neural networks and a cross-validation technique. Hyperparameters were tuned by varying the number of neurons, activation function, optimizer, learning rate, batch size, and epochs. Specifically, the model was tuned using various optimizer parameters such as **rmsprop and Adam**. The best-performing multilayer ANN model utilized the **Adam** optimizer with 50 epochs and a batch size of 10.

characteristics effectively. Like SVM, ANN specializes in recognizing complicated patterns within the EDHS dataset, including demographic features and maternal health indicators. Proposed machine learning algorithms are given in Table 1

3.4. *Training and test set*

Well-defined training and testing datasets are crucial for accurate predictions. In this study, the training set was used to train and build the models, while the test set evaluated model performance. The training and testing dataset samples were created using the 10-fold cross-validation technique. According to studies, 10 folds seem to be the ideal amount to maximize the time needed to complete the test as well as bias and variation related to validation [19], [20], [21].

3.5. *Hyperparameter tuning*

Hyperparameters, set before model training, play a crucial role in model optimization. Manual tuning involves iterative changes, training, and evaluation to enhance performance. The GridSearchCV evaluates model performance using cross-validation techniques and selects optimized hyperparameters from a grid. We select

Table 2. Confusion Matrix representation

	Predicted healthy infant Negative (0) case	Predicted LBW infant Positive (1) case
Actual – Negative (0) case	TN (True Negative)	FP (False Positive)
Actual – Positive (1) case	FN (False Negative)	TP (True Positive)

GridSearchCV for its ability to systematically search through a predefined hyperparameter space to find the best-performing combination. The optimum hyperparameter values were determined using a GridSearchCV strategy.

3.6. *Performance evaluation metrics*

In evaluating the performance of a machine learning classifier, it's important to consider a variety of metrics including Accuracy, Sensitivity, Specificity, Precision, Misclassification rate and ROC-AUC Curve [22], [23], [24]. These metrics provide insights into various aspects of the classifier's performance and help to assess its predictive accuracy and effectiveness. Python's sklearn.metrics library is a powerful tool for evaluating classifier predictive performance. Particularly, the confusion-matrix package within this library simplifies the calculation of classifier performance by providing a structured representation of true positive, false positive, true negative, and false negative predictions. The confusion matrix representation is given in Table 2

The key performance metrics for the machine learning model include accuracy, precision, sensitivity and specificity. In this analysis, to evaluate the proposed ML models, these metrics are considered. Accuracy refers to the proportion of correctly predicted LBW and non-LBW cases among all predictions. Precision measures the proportion of correctly predicted cases among all cases predicted as LBW infants. Sensitivity evaluates the proportion of correctly identified infants with LBW among all infants who have LBW. Specificity assesses the proportion of correctly identified infants without LBW among all infants who do not have LBW. Therefore, using the confusion matrix the study evaluates the correctly predicted values of the selected classifiers model.

3.7. *PROPOSED MACHINE LEARNING MODEL*

The proposed machine learning predictive model for LBW in newborn infants is depicted in Fig. 2, outlining the proposed prediction model development framework.

4. Results and discussion

4.1. *Dataset analysis*

From the EDHS dataset exploratory analysis, the preprocessed data classified instances as follows: 1 for infants who were LBW and 0 for infants who were not LBW.

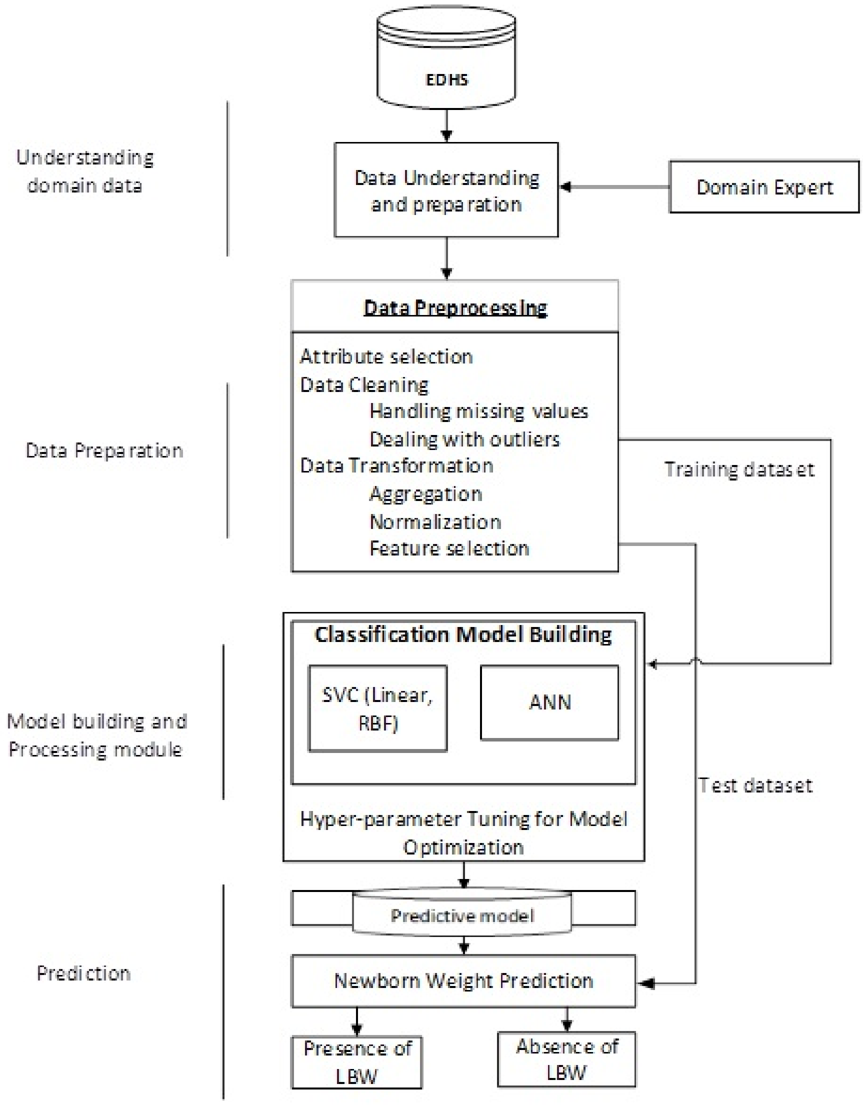

Fig. 2. Proposed system architecture for LBW prediction model building

Out of these instances, 7717 (72.5%) were classified as low birth weight, while 2924 (27.5%) were not. This distribution shows that the majority, three-fourths of the instances, tested positive for LBW, while the remaining one-fourth showed no signs of low birth weight. These features contribute to the existence of low birth weight in newborn infants.

4.2. *Results of preprocessing techniques*

The preprocessing steps included data cleaning, normalization, and feature selection, which are required to ensure the EDHS dataset is suitable for training the machine learning models. Initial data cleaning techniques involved in handling missing values, removing duplicates, and correcting inconsistencies within the EDHS dataset. This step ensured that the data was accurate and complete, providing a solution for the following analysis. Normalization techniques standardize the feature values, the Min Max normalization was applied by using the MinMaxScaler() function. This technique transformed feature values to a range between 0 and 1. Doing this is important in mitigating the risk of the model overfitting or underfitting by ensuring that all features contribute equally to the model training process. Before normalization, feature values were unbounded, leading to potential model overfitting or underfitting issues. After normalization, feature values were transformed to a range between 0 and 1 using the MinMaxScaler() function [19], [20], [21], [25], [26], [27].

To address the class imbalance in the EDHS dataset, we employed SMOTE as an additional preprocessing techniques and step. SMOTE generates synthetic samples for the minority class called non-LBW infants to balance the dataset, improving the candidate model's ability to correctly classify both LBW and non-LBW instances. This technique ensures that the model is not biased towards the majority class – LBW and enhances its predictive performance for both classes. After preparing the data, the model was trained using 80% of the EDHS dataset. The dataset was split into training–validation and testing sets to evaluate the model's performance. The classifiers ANN and SVM were implemented to assess their prediction capabilities. The performance of the predictive models was evaluated both before and after scaling and transformation. The normalization techniques significantly improved the model performance in predicting LBW.

4.3. *Results of hyperparameter tuning*

Hyperparameter tuning plays a pivotal role in identifying the optimal parameters from the given dataset. GridSearchCV stands out as one of the most effective techniques for hyperparameter tuning, as it systematically explores the defined search space of hyperparameter values to pinpoint the optimal configuration [7], [26]. In this study, we leveraged GridSearchCV for hyperparameter optimization. We particularly defined the parameter ranges, initiated the tuning process, and applied a 10-fold cross-validation strategy to the training set. Through iterations involving different training and validation sets, we discerned the best-performing parameters based on maximizing the Area Under the Receiver Operating Characteristic curve. Consequently, the Support Vector Machine (SVM) and Artificial Neural Network (ANN) classifiers were fine-tuned to their optimal configurations. The optimized ANN model train test accuracy is presented in Fig. 3.

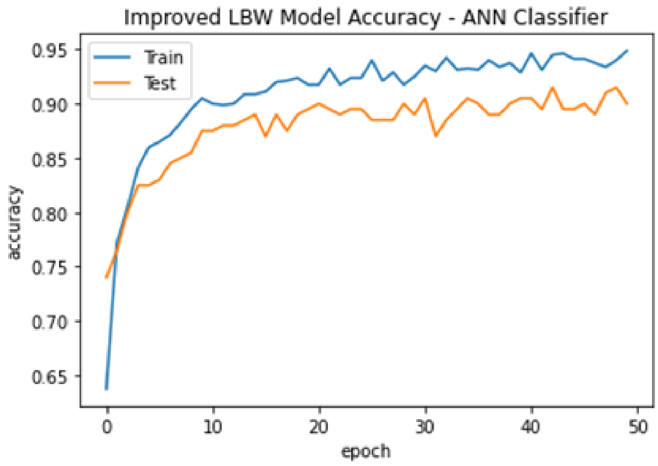

Fig. 3. Optimized ANN model train test accuracy

The model was systematically fine-tuned by adjusting parameters such as the learning rate, the number of nodes in the dense layers, the activation functions, and the optimizers. Using the Adam optimizer and Softmax activation instead of the rmsprop optimizer enhanced prediction accuracy by about 3%, a significant improvement (see Figure 3). For the support vector classifier, we explored various combinations of cost and gamma values for both linear and RBF kernels. GridSearchCV was used to fine-tune the model, with a 10-fold cross-validation applied on the training set. The optimal parameters were identified as Cost: 1 and gamma: 0.01 for the linear kernel, and Cost: 100 and gamma: 0.01 for the RBF kernel.

The classifier models exhibited robust performance in classifying a dataset comprising 10,641 rows and 12 features. The dataset was partitioned into 80% for training and validation and 20% for testing. During model training, 10-fold cross-validation ensured model robustness and prevented model overfitting. GridSearchCV was employed to screen the best hyperparameters for the candidate classifiers. Our analysis revealed that hyperparameter tuning has shown a significant impact on model performance. For example, the Adam optimizer and softmax activation improved neural network learning speed and stability. Tuning the cost and gamma values for the SVC enhanced the model generalization and decision-making performance. This systematic investigation of hyperparameters helps us to optimize the candidate model performance and prevent overfitting, ultimately attaining the highest model accuracy in predicting LBW in newborns.

The above Fig. 4 shows that the final model performance was evaluated using a test set, and the results were remain close to those of the training set. Overall, the models demonstrated high accuracy and performance in predicting low birth weight.

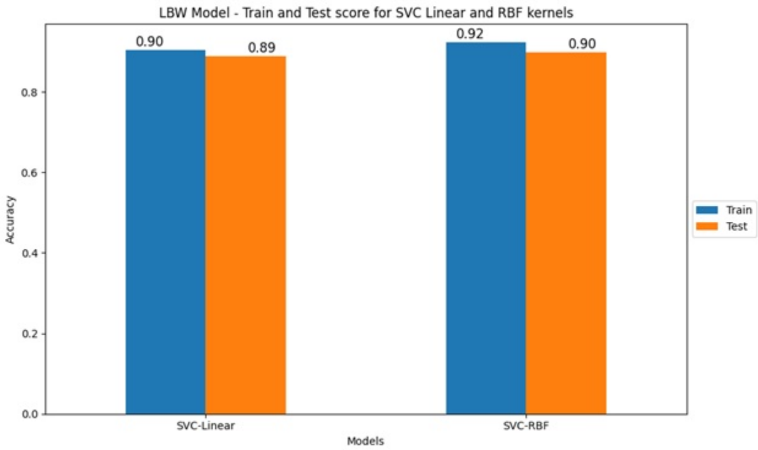

Fig. 4. Train-test set scores for the SVC classifier model

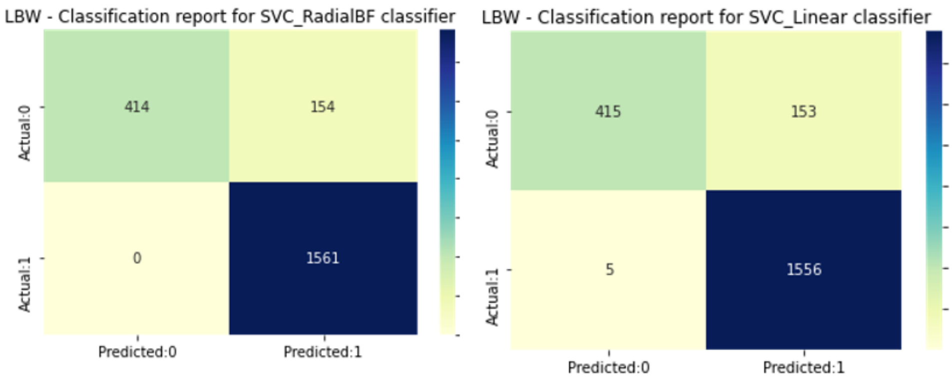

Fig. 5. SVM Classifiers confusion matrix result

4.4. *Classifiers performance*

The candidate classifiers' models measured their performance in predicting the probabilities of LBW or non-LBW in newborn infants. The study evaluated the correctly predicted values of the selected classifiers' models using the confusion matrix.

According to this, the candidate models can be evaluated using classification performance metrics such as sensitivity, specificity, precision, misclassification rate, and $ROC - AUC$. The classification results have been evaluated using machine learning evaluation metrics, including sensitivity, precision, and specificity. Sensitivity refers to the ability to correctly identify entries that belong to the positive class, while specificity refers to the ability to correctly identify entries that belong

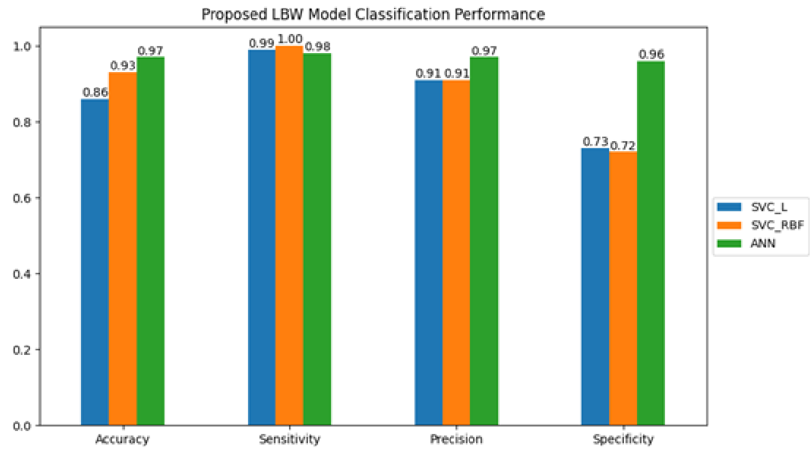

Fig. 6. Proposed LBW prediction model classification performance

to the negative class. This study presents the competitive results of the classifiers. The $SVC - RBF$ classifier model demonstrated better prediction probabilities for low birth weight in newborn infants compared to the $SVC - L$ model. Please see Fig. 5 above for more details.

The study on predicting low birth weight in newborn babies employs a rigorous approach to model evaluation and optimization. With access to ground truth labels, validation methods and metrics serve as essential tools for assessing the effectiveness of classification models. The $ROC - AUC$ curve measure is a common choice for evaluating the discriminatory power of candidate models, providing insights into their ability to distinguish between LBW and normal birth weight cases. Furthermore, the study utilizes the cosine similarity measure to assess the resemblance between training and test set data, aiding in the identification of the optimal hyperplane. To enhance model performance, various hyperparameters such as kernel type, activation function, and optimizer are fine-tuned using GridSearchCV randomly assigned values. This thorough optimization process ensures that the models are finely tuned to predict LBW with minimal error.

The final stage of the analysis involves grid search cross-validation, a robust technique for model comparison that ensures reliable evaluation across multiple train-test splits. This rigorous approach allows for a comprehensive assessment of each model's predictive performance and facilitates the selection of the most effective model for LBW prediction. While the accuracy metrics of candidate models reflect their competency in predicting LBW, it's important to acknowledge that healthcare technologies may exhibit higher sensitivity compared to doctors' perceptions. This sharp sensitivity could lead to more effective service delivery and positive changes in clinical practice. For a detailed overview of the performance of candidate machine learning models in LBW prediction, refer to the summary provided in Table 3, which captures key performance metrics and comparisons.

Table 3. Performance Metrics for proposed Classifiers

Classifiers	**Tuned Parameters and 10-fold CV**	**Accuracy**	**Optimized Model Accuracy**	**Precision**	**Sensitivity**	**Specificity**	**ROC-AUC**
SVM Linear	*C*: {**1.0**, 10.0, 100.0}, *Gamma*: {**0.01**, 0.001, 0.0001}	0.82	0.86	0.91	0.99	0.73	0.83
SVM RBF	*C*: {1.0, **10.0**, 100.0}, *Gamma*: {**0.01**, 0.001, 0.0001}	0.89	0.93	0.91	1.00	0.72	0.97
ANN	*Activation*: {Sigmoid, ReLU, **softmax**}, *Optimizer*: {RMSprop, **Adam**}	0.94	0.97	0.97	0.98	0.96	0.99

The above Table 3 presents that among the proposed classifiers models ANN model outperformed the rest candidate classifiers models across all key metrics. Specifically, the ANN classifier with softmax activation function and Adam optimizer achieved the highest accuracy 0.97 and ROC-AUC 0.99, presenting its significant ability to differentiate the probability of LBW and Non-LBW cases. This classifier model also presented high sensitivity 0.98 and precision rate 0.97. According to the experiment and the result presented by ANN classifier, suggesting that the ANN model is particularly effective in predicting LBW cases with least missing cases.

The SVM classifier with linear and RBF kernel models are performed well. The SVM-RBF model presented an accuracy of 0.98 with perfect sensitivity of 1.0, lower specificity than ANN model, it shows that the model might produce more false positive prediction than ANN model. Whereas the SVM-Linear kernel model demonstrated relatively lower overall performance of the candidate classifiers with an accuracy of 0.86 and also a lower ROC-AUC score 0.83. The SVM-Linear kernel model was less effective than the other candidate classifier models for this particular task. Therefore, these investigation and findings shown that the importance of selecting effective and robust model for healthcare applications is crucial. The candidate classifier ANN model with its overall performance makes it to be the most suitable for identifying the probability of LBW risks in newborns and for other clinical applications.

Comprehensive assessments were conducted both before and after optimization to gauge model performance across different stages of fine-tuning. The ROC curve

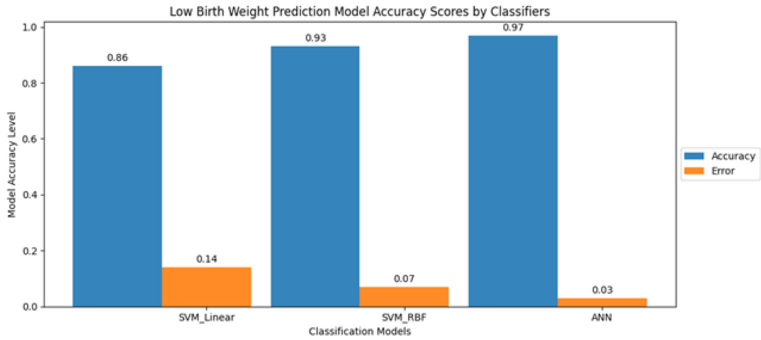

Fig. 7. Classifiers Model accuracy for LBW prediction system

was a crucial tool in assessing how well our models could predict low birth weight cases in newborn babies. It showed us the balance between correctly identifying LBW cases (true positive rate) and incorrectly flagging normal birth weight cases as LBW (false positive rate) across different thresholds. The area under the ROC curve (AUC) acted as a yardstick for our models' ability to distinguish between LBW and normal birth weight instances. A higher AUC value, nearing 1, signaled better predictive power, indicating that our optimized models were adept at distinguishing between LBW and normal birth weight cases with greater accuracy.

In the context of our study on predicting low birth weight cases in newborn babies, we employed ROC-AUC curves to evaluate the performance of the optimized ANN model. These curves plot the true positive rate (sensitivity) against the false positive rate (1-specificity), with the top-left corner representing the ideal point where the true positive rate is one and the false positive rate is zero. Our analysis revealed that the optimized ANN model demonstrated significantly improved classification performance in predicting LBW probabilities compared to previous iterations. For a visual representation of the ROC-AUC curves and further details on the model's performance.

Looking at Figure 8, it presents the candidate classifiers model performance. The optimized ANN classifier model outperforms significantly well in predicting low birth weight than the other two models we studied. Specifically, the neural network model achieves a remarkable accuracy rate of 97.2% in LBW prediction.

In this study, we are compared machine learning classifiers such as ANN, SVM-Linear and SVM-RBF. The candidate classifiers model were selected as a benchmarks due to their robust performance in related tasks and their ability to handle variety of relationships like linear and non-linear. While other researchers in the domain frequently employed ANN, SVM with different kernel, random forest and

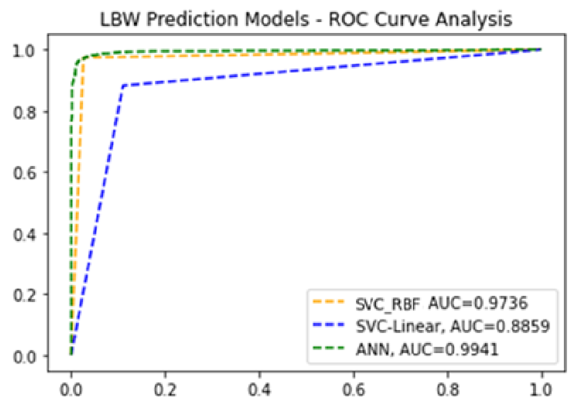

Fig. 8. Receiver Operating Characteristics curve for classifier models

linear regression. These classifiers were chosen for their proven efficiency in low to high dimensional and complex features interactions of data, this also evident in our dataset. This overall comparison demonstrates that the greater model performance of classifier ANN achieving model accuracy of 97.2% compared to 93% for SVM-RBF and 86% for SVM-Linear.

5. Conclusion

Our study employed a classification approach alongside grid search cross-validation to optimize our model's parameters, enhancing its predictive capabilities. We assigned hyperparameters randomly and trained them using various classifiers, including Support Vector Machines (SVM), and Artificial neural network, with a range of settings such as kernel types, learning rates, and activation functions. Utilizing GridSearchCv, we fine-tuned these hyperparameters and ranked them based on their significance in predicting LBW. To ensure the effectiveness of the model, we cross-validated our models using 10-fold cross-validation, splitting our dataset of 12 features of both maternal and new infant and 10,641 instances into 80% for training-validation and 20% for testing. We supervised class sensitivity, specificity, precision, and accuracy, as well as ROC-AUC, to evaluate model performance. Initially, our ANN model achieved approximately 94% accuracy without parameter tuning. However, after fine-tuning the parameters for improved feature extraction, accuracy increased to 97.2%, as indicated in Table 3. This demonstrates the significant impact of parameter tuning on model performance, with significant improvements observed across all machine learning model performance metrics.

Compared to other research, our study stands out for its comprehensive approach to parameter tuning and model evaluation. While many studies overlook optimization techniques, our particular approach highlights their importance in achieving accurate predictions for LBW. Additionally, our study contributes to the field by providing detailed insights into the performance of different classifiers and hyperparameter settings. Despite these contributions, our study is limited by the scope of algorithms compared and the specific dataset used. Future research should explore additional algorithms, such as convolutional neural networks (CNNs) or ensemble methods, and evaluate performance on diverse datasets to ensure generalizability and robustness of the findings.

Acknowledgment
We extend our appreciation to the Vice President's Office of Technology Transfer and Research, Addis Ababa Science and Technology University, as well as the Research Directorate Office, for their support. We are also thankful to the HPC and Big Data Analytics Center of Excellence for providing the resources that made this research possible.

References

1. World Health Organization. Global nutrition targets 2025: low birth weight policy brief geneva. world health organization; 2014., 2014.
2. World Health Organization. Global health estimates for 2020: deaths by cause, age, sex, by country and by region, 2000-2019, 2020.
3. M. Tadese, A. S. Minhaji, C. T. Mengist, F. Kasahun, and G. B. Mulu. Determinants of low birth weight among newborns delivered at tirunesh beijing general hospital, addis ababa, ethiopia: a case-control study. *BMC Pregnancy Childbirth*, 21(1):1–9, 2021. doi: 10.1186/s12884-021-04275-6.
4. Z. Hussain and M. D. Borah. Birth weight prediction of new born baby with application of machine learning techniques on features of mother. *J. Stat. Manag. Syst.*, 23(6):1079–1091, Aug. 2020. doi: 10.1080/09720510.2020.181449 9.
5. D. A. Jember, Z. A. Menji, and Y. A. Yitayew. Low birth weight and associated factors among newborn babies in health institutions in dessie, amhara, ethiopia. *J. Multidiscip. Healthc.*, 2020. doi: 10.2147/JMDH.S285055.
6. A. Alebel et al. Factors associated with low birth weight at debre markos referral hospital, northwest ethiopia: A hospital based cross-sectional study. *BMC Res. Notes*, 12(1):1–6, 2019. doi: 10.1186/s13104-019-4143-1.
7. M. Al-Shawwa. Ann for predicting birth weight. [Online]. Available: www.ijeais.org/ijahmr.
8. H. Y. Hassen, S. H. Gebreyesus, B. S. Endris, M. A. Roro, and J. P. Van Geertruyden. Development and validation of a risk score to predict low birthweight using characteristics of the mother: Analysis from bunmap cohort in ethiopia. *J. Clin. Med.*, 9(5), 2020. doi: 10.3390/jcm9051587.
9. C. Mangold, S. Zoretic, K. Thallapureddy, A. Moreira, K. Chorath, and A. Moreira. Machine learning models for predicting neonatal mortality: A systematic

review. 2021. doi: 10.1159/000516891.

10. J. Feng, J. Lee, Z. A. Vesoulis, and F. Li. Predicting mortality risk for preterm infants using deep learning models with time-series vital sign data. *npj Digit. Med.*, 4(1):1–8, 2021. doi: 10.1038/s41746-021-00479-4.
11. R. M. McAdams, R. Kaur, Y. Sun, H. Bindra, S. J. Cho, and H. Singh. Predicting clinical outcomes using artificial intelligence and machine learning in neonatal intensive care units: a systematic review. *npj Digital Medicine*, 5(1): 1–10, 2022. doi: 10.1038/s41372-022-01392-8. URL https://doi.org/10.1038/s41372-022-01392-8.
12. A. Faruk, E. S. Cahyono, N. Eliyati, and I. Arifieni. Prediction and classification of low birth weight data using machine learning techniques. *Indonesian Journal of Science and Technology*, 3(1):18–28, 2018. doi: 10.17509/ijost.v3i1.10799. URL https://doi.org/10.17509/ijost.v3i1.10799.
13. T. B. Reza and N. Salma. Prediction and feature selection of low birth weight using machine learning algorithms. *Journal of Health, Population and Nutrition*, 43(1):157, 2024. doi: 10.1186/s41043-024-00647-8. URL https://doi.org/10.1186/s41043-024-00647-8.
14. A. Seveso, V. Bozzetti, P. Tagliabue, M. L. Ventura, and F. Cabitza. Developing a machine learning model for predicting postnatal growth in very low birth weight infants. In *HEALTHINF 2020 - 13th International Conference on Health Informatics*, pages 490–497, 2020. doi: 10.5220/0008972804900497. URL https://doi.org/10.5220/0008972804900497.
15. W. T. Bekele. Machine learning algorithms for predicting low birth weight in ethiopia. *BMC Medical Informatics and Decision Making*, 2022. doi: 10.1186/s12911-022-01981-9. URL https://doi.org/10.1186/s12911-022-01981-9.
16. M. Arayeshgari, S. Najafi-Ghobadi, H. Tarhsaz, S. Parami, and L. Tapak. Machine learning-based classifiers for the prediction of low birth weight. *Healthcare Informatics Research*, 29(1):54–63, 2023. doi: 10.4258/hir.2023.29.1.54. URL https://doi.org/10.4258/hir.2023.29.1.54.
17. A. Ranjbar, F. Montazeri, M. V. Farashah, V. Mehrnoush, F. Darsareh, and N. Roozbeh. Machine learning-based approach for predicting low birth weight. *BMC Pregnancy and Childbirth*, 23(1):1–7, 2023. doi: 10.1186/s12884-023-06128-w. URL https://doi.org/10.1186/s12884-023-06128-w.
18. W. Khan et al. Infant birth weight estimation and low birth weight classification in united arab emirates using machine learning algorithms. *Scientific Reports*, 12(1):1–12, 2022. doi: 10.1038/s41598-022-14393-6. URL https://doi.org/10.1038/s41598-022-14393-6.
19. A. Panesar. *Machine Learning and AI for Healthcare*. 2021. doi: 10.1007/978-1-4842-6537-6. URL https://doi.org/10.1007/978-1-4842-6537-6.
20. P. Guleria and M. Sood. *Intelligent Learning Analytics in Healthcare Sector Using Machine Learning*. 2020. doi: 10.1007/978-3-030-40850-3_3. URL https://doi.org/10.1007/978-3-030-40850-3_3.

21. K. R. Saraf and P. Malathi. Intelligent learning analytics in the healthcare sector using machine learning and iot. In *Machine Learning, Deep Learning, Big Data, and Internet of Things for Healthcare*. 2022. doi: 10.1201/9781003227595-3. URL https://doi.org/10.1201/9781003227595-3.
22. S. Karthiga, K. Indira, and C. V. N. Angeline. Machine learning model to predict birth weight of newborn using tensorflow. *Computer Science and Information Technology*, 15(9):71–90, 2019. doi: 10.5121/csit.2019.91506. URL https://doi.org/10.5121/csit.2019.91506.
23. M. Supriya and A. Deepa. Machine learning approach on healthcare big data: A review. *Big Data and Information Analytics*, 5(1):58–75, 2020. doi: 10.3934/bdia.2020005. URL https://doi.org/10.3934/bdia.2020005.
24. I. D. Dinov. *Data Science and Predictive Analytics*. 2018. doi: 10.1007/978-3-319-72347-1. URL https://doi.org/10.1007/978-3-319-72347-1.
25. P. Approach and P. Models. *Feature Engineering and Selection*. n.d.
26. S. Senthilkumar and D. Paulraj. Prediction of low birth weight infants and its risk factors using data mining techniques. In *Proceedings of the 2015 International Conference on Industrial Engineering and Operations Management*, Dubai, United Arab Emirates, 2015.
27. B. H. Su, P. H. Hu, C. T. Peng, and C. H. Tsai. Chronic lung disease in extremely low birth weight infants: A two-year retrospective analysis. *Acta Paediatrica Taiwanica*, 41(2):75–79, 2000.