Applied Mathematics and Computational Intelligence
Volume 13, No. 3, 2024 [66-86]

An Ensemble Learning Algorithm with Hyperparameter Optimization (ELAHO) Model for Early Detection of Maternal Health Risk (MHR) Level Using Machine Learning

Samaila Abdullahi[1], Saratha Sathasivam[2] *

[1,2]School of Mathematical Sciences, Universiti Sains Malaysia (USM), 11800 USM, Penang Malaysia

* Corresponding author: saratha@usm.my

ABSTRACT

Despite the advancement in the healthcare system, maternal health risk remains high, which is the most challenging aspect nowadays. There is a need to develop an effective model for early detection, monitoring, and prediction of maternal health risk levels during pregnancy. The machine learning intelligence model has proven its effectiveness and robustness in providing accurate and reliable prediction, analysis, and interpretation of medical data, reducing several risk factors for early diagnosis in healthcare. In this research work, we proposed an ensemble learning algorithm with hyperparameter optimization (ELAHO) model using machine learning algorithms to improve its robustness, effectiveness, and model performance. The proposed method uses a hybrid model of logistic regression and support vector machines (LG-SVM) to predict maternal health risk levels during pregnancy. The method utilized Python software for training, testing, and validation. We evaluated the performance of our proposed model using accuracy, precision, sensitivity, f1-score, and ROC-AUC score. The proposed models outperformed the conventional models and achieved 100% predictive accuracy. The proposed approach has the potential to be adapted as an intelligence-monitoring system for early medical diagnosis during pregnancy. The proposed techniques will help medical professionals make quick decisions accurately and enhance monitoring to improve the level of care offered to pregnant mothers and their unborn children.

Keywords: Maternal Health Risk, Logistic Regression, Support Vector Machine, Hyperparameter Optimization, Ensemble Learning Algorithm.

1 INTRODUCTION

Maternal health risk (MHR) refers to possible health issues during the stages of embryonic development, delivery, and postpartum care, which include mental, physical, and social health. The World Health Organization (WHO) estimates that pregnancy problems result in nearly 280,000 female deaths per year [1]. Maternal health is a critical component of public health due to its direct influence on the overall welfare of the mother and child. Notwithstanding the progress made in the field of medicine, maternal mortality continues to be a substantial concern in numerous countries, with a special emphasis on developing nations [2]. Continuous monitoring of each stage of pregnancy is necessary to secure the delivery of a normal child and the healthy growth of the baby [3]. Notwithstanding recent technological advancements, maternal mortality is declining, posing challenges to ensuring the safety of both the expectant mother and her unborn child. Ensuring timely identification and providing appropriate medical help are essential concerns for preserving optimal health during pregnancy. Traditional risk assessment methods often rely on a limited set of factors and may not accurately predict adverse outcomes. Recent years have seen the emergence of artificial intelligence (AI) as a potent tool in healthcare, providing novel approaches to the interpretation and analysis of complex medical data. In several clinical applications, including disease diagnosis, therapy planning, and diagnostic models, it has been demonstrated to have proven outcomes. AI-based models can perform massive amounts of structured and unstructured health analysis to identify individuals who may be at high risk for undesirable consequences. The previous predictive analysis used a single algorithm for modelling the risk factor of maternal health risk and the results need to improve. To improve the previous models, we utilized a hybrid learning algorithm with hyperparameter optimization to predict maternal health risk due to its potential to improve early identification of high-risk pregnancies, thereby enabling timely interventions and reducing maternal and fetal mortality and morbidity. It has been proven that hybrid models are better at attaining predictive accuracy and model performance. According to [4] the proposed machine learning approach was successful in predicting the risk level during pregnancy, achieving 97% accuracy. According to [5] the Internet of Things (IoT) and machine learning models were proposed to early detect risk levels during pregnancy, and the approach has effectively predicted all risk levels and achieved better model performance.

Logistic regression (LR) is a widely used expanded linear model for classification and regression due to its simplicity of use and implementation [6]. Furthermore, it has a simple probabilistic interpretation of its model. Apart from providing the user with exact probabilities of classification, the advantage of logistic regression is that it yields class labels as well as being easily generalizable to the problem of classifying data points into different categories [7]. Additionally, the logistic regression technique predicts the likelihood of a categorical variable using a sigmoid function [8].

Support vector machines (SVM) are robust and reliable classifiers that have been effectively used to solve both classification and regression problems. Also, it has been proven to be effective in the fields of handwriting and pattern recognition [9]. Support Vector machines (SVMs) are extensively used to solve classification and regression problems in fields such as computational biology, economics, and text categorization. This can be attributed to multiple factors, which include the utilization of computational learning theory, the modelling of non-linear distributions with kernel functions that separate all the data points, and the capacity to train and predict on large datasets using innovative mathematical optimization approaches [10].

hyperparameter optimization process of finding a model's optimal parameter using the current dataset. In this study, grid search is taken into consideration to select the optimal parameters that will enhance the model's functionality and lessen overfitting. In [11] Machine learning algorithms can automatically learn from datasets and adjust their own parameters.

The ensemble learning technique refers to predicting the final model evaluation by hybridizing the performance of multiple base models. Enhancing prediction accuracy, stability, and robustness is the main purpose of the ensemble learning algorithm. Different methodologies exist for hybridizing learning models to enhance conventional models. According to [12] Combining a variety of projected models into an effective model that generates dependable and accurate prediction results is the primary goal of the ensemble learning technique. Utilizing the effectiveness and robustness of machine learning (ML) algorithms to predict maternal health risk is the main objective of this research work. We proposed a hybrid model and compared its performance with that of conventional models. To determine the "risk level," which is the desired output in multiclass classification, we used a data preprocessing approach on the obtained dataset, which included 1014 instances and six (6) feature attributes, of which 40% belong to the low-risk category, 33.1% belong to the mid-risk category, and 26.8% belong to the high-risk category of the dataset. The Grid Search CV was utilized in the hyper-parameter tuning technique to identify the optimal estimator values for each parameter combination. The proposed model has improved in performance, robustness, and ability to predict more accurately compared with conventional models.

The findings indicate that the proposed model had better performance in terms of accuracy, precision, and the ROC-AUC curve when compared to the conventional models. In this study, the combination of stratified k-fold cross-validation and the conceptual method of hyperparameter optimization improved the performance of predictive analysis. The proposed model has improved in performance, robustness, and ability to predict more accurately compared to conventional models. The results obtained demonstrate that the proposed model has outperformed the conventional models based on accuracy, precision, and the ROC-AUC curve. This study utilized the conceptual method of hyperparameter optimization combined with stratified k-fold cross validation to improve the performance of predictive analysis. The main contribution of our proposed method includes enhancing the performance of the existing model using the hybrid method, training all data points to achieve low bias and variance using k-fold cross validation, evaluating the model's performance and accuracy using hyperparameter optimization compared with conventional methods, and enhancing the model's robustness, accuracy, and efficiency using an ensemble learning algorithm.

2 MATERIAL AND METHODS

This section describes the research methodology and proposed method used in this paperwork.

2.1 Logistic Regression algorithm

Logistic regression is a statistical technique employed to model the connection between one or more independent variables and a dependent variable. Logistic regression is a widely used statistical technique that can provide good models. Its type of supervised learning algorithm is used to solve classification and regression problems by predicting the probability that an observation will belong to a particular class [13]. Multinomial logistic regression, which combines the probabilities of other two or more class combinations of algorithms using one-vs-one techniques, allows for the direct implementation of the logistic more class in a model with categorical responses. In this case, it is considered that the distribution of the dependent variable is multinomial [14]. Multinomial allows two or more feature vectors $y^{(i)} \in \left(1, 2, \ldots, k\right)$ where K is the number of class labels and $X \in \mathbb{R}^n$ is a vector with a variable probability distribution p_X that consists of linearly independent features with the backing $X \in \mathbb{R}^n$ and $\sum_{j=1}^{L} p_j(x) = 1$ for any $x \in X$. while we look at a multinomial logistic regression model, we assume that assumed that

$$In\frac{pk(x)}{pK(x)}=\eta_K^T x,\quad k=1,\ldots,K-1 \tag{1}$$

and $\eta_k \in \mathbb{R}^n$ are the feature vectors of regression coefficients, then

$$pk(x)=\frac{\exp\eta_k^T x}{1+\sum_{k=1}^{K-1}\exp(\eta_k^T x)}, \text{ and } pK(x)=\frac{1}{1+\sum_{k=1}^{K-1}\exp(\eta_k^T x)},\ \forall k=1,\ldots,K-1$$

the generalized form of the multinomial logistic regression is given as:

$$pK(x)=\frac{\exp(\eta_k^T x)}{1+\sum_{k=1}^{K}\exp(\eta_k^T x)}, k=1,\ldots,K \text{ with } \eta_K=0 \tag{2}$$

The class with the highest likelihood function is assigned to x in accordance with the classification criteria [15].

The generalized form for cost function is given as:

$$j(\eta)=-\left[\sum_{i=1}^{n}\sum_{k=1}^{K}1\{p^{(i)}=k\}\log\frac{\exp(\eta^{(k)T}x^{(i)})}{\sum_{j=1}^{K}\exp(\eta^{(k)T}x^{(i)})}\right] \tag{3}$$

For each possible feature of $y\in\{1,2,\ldots,K\}$ assigned the feature vector $\mu\in\{0,1,2\}^K$ with $\mu k=I\{y=k\}$. let $(x_1,y_1),\ldots,(x_n,y_n)\sim f\eta(x,y)$, be the matrix of regression coefficient in (1) with columns $\eta_1,\ldots,\eta_K$ and let $f_\eta(x,y)$ be the equal shared distribution of (x,y) that is,

$$df_\eta(x,y)=\prod_{k=1}^{K}pk(x)^{\xi k}dPx(x) \text{ where } pk(x)=\frac{\exp\eta_k^T x}{1+\sum_{k=1}^{K-1}\exp(\eta_k^T x)}.$$

Given a random feature $(x_1,y_1),\ldots,(x_n,y_n)\sim f\eta(x,y)$, the log likelihood function is given as:

$$l(\eta)=\sum_{i=1}^{n}\left\{X_i^T\eta\xi_i-\ln\sum_{i=1}^{k}\exp(\eta_i^T X_i)\right\} \tag{4}$$

2.2 Support Vector Machines

Support vector machines (SVMs) is the most used supervised machine learning models for predictive analysis, both for classification and regression. SVMs operate by determining which hyperplane in the input data best divides the various classes. The closest data points from each class are separated from the hyperplane by a distance known as the margin, which is maximized to define the hyperplane. By maximizing this margin, SVMs aim to achieve better generalization and robustness in classifying new, unseen data points [16].

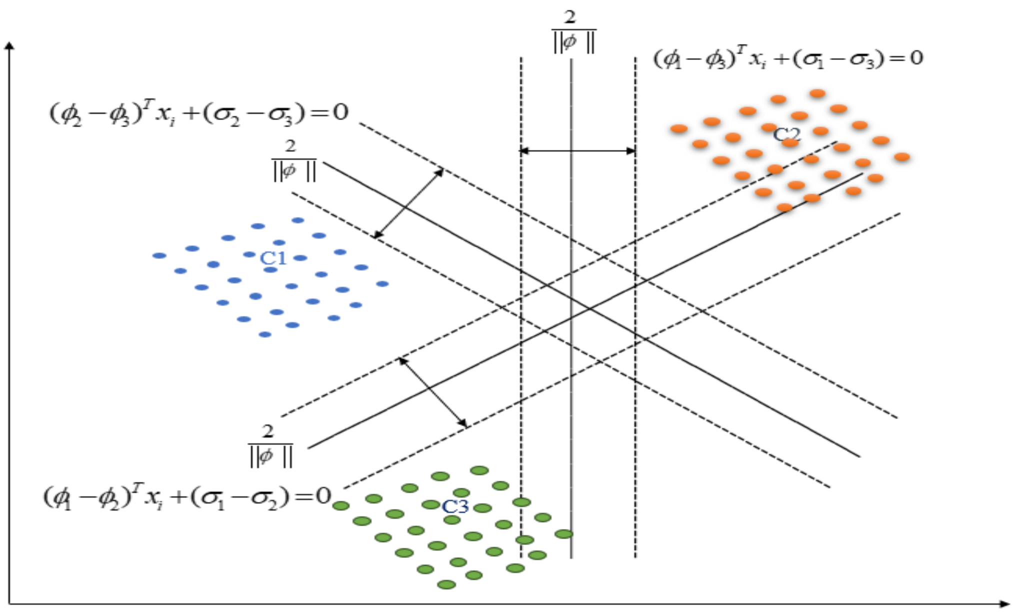

Figure 1: Multiclass Support vector machine (MCSVM)

The multi-class classification problem can now be handled with SVM enhancements using one-against-rest (OAR) and one-against-one (OAO) methods, which involve building and combining several binary classifiers [29]. The first attempt to solve a multi-class problem using SVM is using one-versus-all (OVA) or one-versus-rest (OVR) techniques. Suppose there are N training features in the form of (x_i, y_i) and k number of class label, we want to develop k binary classification using the SVM model, then we train w^{th} support vector model and define the class w^{th} as a positive class label and rest of the classes as negative label. If there are an equal number of training data in each class, this decision problem is an unbalanced classification method, as shown in the following representation:

$$\min_{\phi_w,\sigma_w} \frac{1}{2}\|\phi_w\|_2^2 + c\sum_{i=1}^{N} \xi_i^w$$

Subject to $\phi_w{}^T x_i + \sigma_w \geq 1 - \xi_i^w$, if $y_i = w$

$$\phi_w{}^T x_i + \sigma_w \le -1 + \xi_i^w, \text{ if } y_i \ne w$$

$$\xi_i^w \ge 0 \tag{5}$$

The new sample features x_i belong to the class w with the largest decision function value

$$\bar{y}_i = \arg\max_w \phi_i{}^T x_i + \sigma_w \tag{6}$$

The primary defect in this approach, as we have indicated, is that every binary classification is out of balance. This attribute may influence how one strategy performs in comparison to the others on a multi-class classification problem. To solve such a multiclass problem, we utilized one-versus-one or one-versus-rest by training more binary SVM models. With these methods, we train a single binary SVM model for every two classes, for a total of $\frac{k(k-1)}{k}$ models. For class w and v, the decision boundary or margin between them is $\phi_{wv} x_i + \sigma_{wv} = 0$, and The following can be learned from the problem.

$$\min_{\phi_{wv}, \sigma_{wv}} \frac{1}{2} \|\phi_{wv}\|_2^2 + c \sum_{i=1}^{N} \xi_i^{wv}$$

Subject to $\phi_{wv}{}^T x_i + \sigma_{wv} \ge 1 - \xi_i^{wv}$, if $y_i = w$

$$\phi_{wv}{}^T x_i + \sigma_{wv} \le -1 + \xi_i^{wv}, \text{if } y_i = v$$

$$\xi_i^{wv} \ge 0 \tag{7}$$

The voting procedure is used as a test, if $sign(\phi_{wv} x_i + \sigma_{wv})$ says x_i is in class w, then the vote for class w is added by one; if not, the vote for class v is added by one. The final predicted class is the one with the most vote. To address this issue, we propose to use k the feature vectors required to model each of these binary classifiers, for example, a classifier between class w and v, $\phi_{wv} = \phi_w - \phi_v$. This avoids the unbalanced training data subproblem while trying to bring our space usage back to that of the one-versus-rest strategy. We want to compute k decision function in testing without using a voting method like one-versus-one approach. The soft margin objective function for this problem is to maximize the margin between class w and v datapoints by solving the binary classification problem between them.

$$\min_{\phi_w, \phi_v \in \mathbb{R}^d} \frac{1}{2} \|\phi_w - \phi_v\|_2^2 + c \sum_{y_i \in \{w,v\}} \xi_i^{wv}$$

Subject to $y_i^{wv} f_{wv}(x_i) \ge 1 - \xi_i^{wv}, \forall y_i \in \{w, v\}$

$$\xi_i^{wv} \ge 0 \tag{8}$$

Where $x_i \in \mathbb{R}^d$, non-negative slack variable ξ_i^{wv} was introduced to separate all datapoints linearly and $\sum_{y_i \in \{w,v\}} \xi_i^{wv}$ is the penalty term aid to reduce error during training. The parameter c is to balance between regularization term $\|\phi_w - \phi_v\|_2^2$ and training error. where the decision function is $f_{wv}(x_i) = (\phi_w - \phi_v)^T x_i + (\sigma_w - \sigma_v)$ and $y_i^{wv} = \begin{cases} +1 & if \quad y_i = w \\ -1 & if \quad y_i = v \end{cases}$. There are $\frac{k(k-1)}{2}$ binary classifiers for all class labels. However, we now develop a multiclass support vector machine (MCSVM) as:

$$\min_{\phi \in \mathbb{R}^{d*k}, \sigma \in \mathbb{R}^k} \frac{1}{2} \sum_{w=1}^{k-1} \sum_{v=w+1}^{k} \|\phi_w - \phi_v\|_2^2 + c \sum_{w=1}^{k-1} \sum_{v=w+1}^{k} \sum_{y_i \in \{w,v\}} \xi_i^{wv}$$

Subject to $y_i^{wv} f_{wv}(x_i) \geq 1 - \xi_i^{wv}, \forall y_i \in \{w,v\}$

$$\xi_i^{wv} \geq 0 \tag{9}$$

The equation (9) will only have one optimal solution, As can be seen, if $\phi \in \mathbb{R}^{d*c}$ and $\sigma \in \mathbb{R}^d$ is the optimum solution. Since there are multiple solution. Therefore, the two minimizing constrain will be imposed in the equation (9) and the final optimization function is given as:

$$\min_{\phi \in \mathbb{R}^{d*k}, \sigma \in \mathbb{R}^k} \frac{1}{2} \sum_{w=1}^{k-1} \sum_{v=w+1}^{k} \|\phi_w - \phi_v\|_2^2 + \frac{1}{2} \sum_{w=1}^{k} \|\phi_w\|_2^2 + \frac{1}{2} \sum_{w=1}^{k} \sigma_w^2 + c \sum_{w=1}^{k-1} \sum_{v=w+1}^{k} \sum_{y_i \in \{w,v\}} \xi_i^{wv}$$

Subject to $y_i^{wv} f_{wv}(x_i) \geq 1 - \xi_i^{wv}, \forall y_i \in \{w,v\}$

$$\xi_i^{wv} \geq 0 \tag{10}$$

Where the decision function $f_{wv}(x_i) = (\phi_w - \phi_v)^T x_i + (\sigma_w - \sigma_v)$ and $y_i^{wv} = \begin{cases} +1 & if \quad y_i = w \\ -1 & if \quad y_i = v \end{cases}$.

Now that we use this model to predict a new feature, we apply the same voting techniques as the one-versus-one approach for multiclass support vector machines. If $sign((\phi_w - \phi_v)^T x_i + \sigma_w - \sigma_v) = 1$. The vote for class label w is added by one. Notwithstanding, we can find more k that satisfies $\phi_w^T x_i + \sigma_w > \phi_v^T x_i + \sigma_v$, then w will have votes. The predicted class label should have maximum decision function value,

$$\bar{y}_i = \arg\max_{w} \phi_w^T x_i + b_i . \tag{11}$$

We just need to compute decision function k times and identify the class with the maximum value.

2.3 Hyperparameter optimization

Hyperparameter optimization in machine learning is the process of controlling parameter values during training. Most of the previous work using hyperparameter optimization has focused solely on a grid search, random search, with some other researchers comparing two techniques to search for the best parameter combination to achieve an optimal model. During parameter optimization phase, there is a need to look at cross validation to train all data points to reduce overfitting. According to [17] hyperparameter tuning was used to find the best parameter setting or combination for random forest, artificial neural networks, and Bayesian optimization. Optimizing hyperparameters is the process of finding such parameter values that will improve model performance on a particular dataset [18]. Hyperparameter tuning is a vital stage in machine learning for achieving a robust and effective model during the training phase. Following the optimization of the default parameter value to get more accurate and consistent results. In this paperwork, we consider grid search techniques to search for the best parameter combination for a hybrid model to improve performance, achieve better accuracy, and reduce overfitting. To utilize and implement a grid search, hyperparameter parameter values of interest are defined within a network.

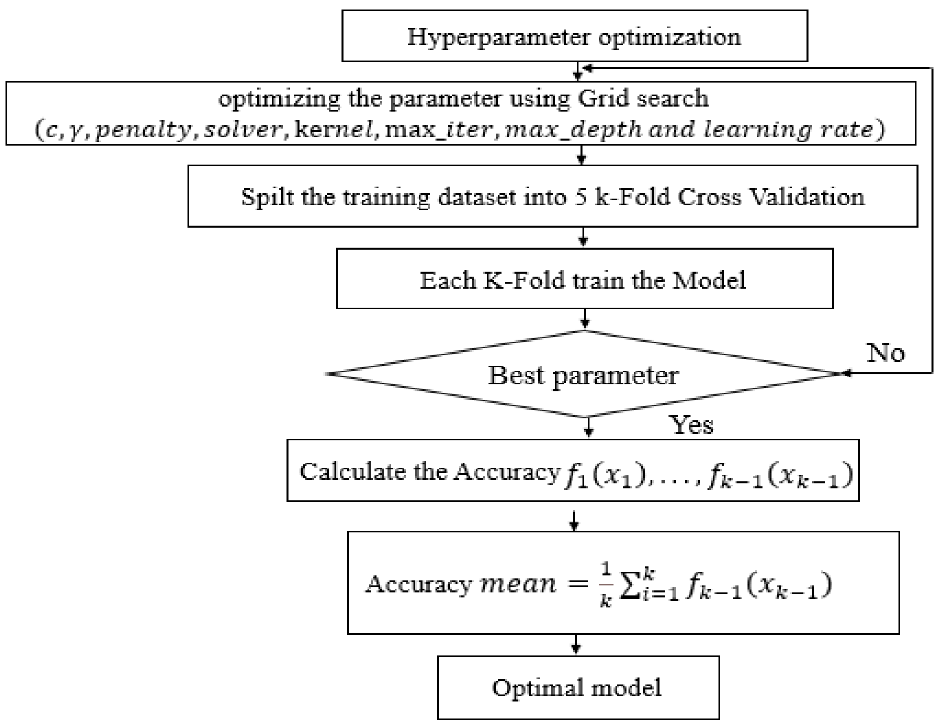

Figure 2: Flowchart of hyperparameter optimization.

The goal of the hyperparameter optimization technique is to use an automated search to determine the optimal set of parameter values for a given dataset to maximize accuracy. We specify the model's parameters before training it.

2.4 Ensemble Learning Algorithm

Ensemble learning algorithms refer to learning methods that hybridize two or more base learning models to improve performance and robustness. The meta model was used to predict the final model's performance by combining base learners' predictive probabilities. By leveraging the heterogeneity and combination of predictive performance of several models, these ensemble learning techniques can improve the predictive performance and generalization of optimum models. In this paperwork, we utilize ensemble learning algorithms via stacking techniques to hybridize logistic regression and support vector machines to model maternal health risk (MHR) to predict the risk level during pregnancy among women. The meta model is defined as given below, which will determine the final model evaluation.

$$\hat{y}(x) = q_i(x) + q_{i+1}(x) + q_{i+2}(x) +, \ldots, + q_{n(x)} \quad \text{where } i = 1, 2, 3, \ldots, n \tag{12}$$

where $\hat{y}(x)$ is the final prediction and $q_i(x)$ is the predicted probability for the base learners. The predictive probability features $q_i(x)$ for meta-model maps each data point to its corresponding target features. Sequential ensemble and parallel ensemble methods are two categories into which ensemble methods fall. According to [19] when it comes to several prediction and classification problems, combining the results from multiple predictors gives maximum accuracy.

2.5 Data description

The maternal health risk dataset employed in this research was obtained from the machine learning repository at the University of California, Irvine (UCI). The dataset has seven (7) attributes for predicting risk levels during pregnancy, which include age, systolic blood pressure, diastolic blood pressure, blood sugar, body temperature, and heart rate. Each variable was represented by a numeric value. predicted target variable for risk level, was present in three different class labels: high risk, mid risk, and low risk. There were 1,014 instances in the dataset; 406 samples belonged to the low-risk category, 336 samples to the medium-risk category, and 272 samples to the high-risk category. The percentage distribution of the dataset is shown in Figure 3.

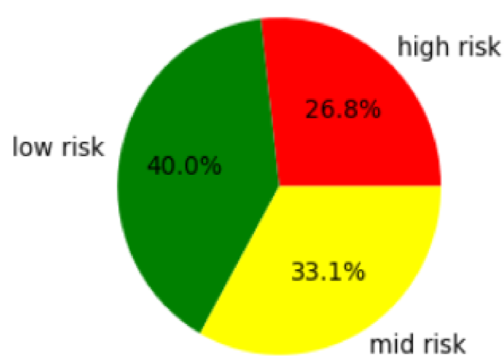

Figure 3: The distribution of Maternal Health Risk (MHR) Dataset based on Risk Level

There are no missing values in the dataset. We use 80% for the training using k-fold cross-validation and 20% for the test to check model performance. Pre-processing and standardization procedures were executed at this level as the most fundamental data preprocessing steps in machine learning prior to the classification of the dataset class labels. Some of the previous research used the same dataset and split the dataset 80 by 20% for training, validation, and testing. According to [20] the proposed Light GBM, Cat Boost, Random Forest, Gradient Boosting Machines, and KNN models were applied to model the maternal health risk (MHR) dataset, using 80% for the training sample and 20% for testing, and the model's hyperparameters were optimized. Gradient Boosting Classifier, Extra Trees Classifier, Decision Tree Classifier, Random Forest Classifier, and Snap Boosting Machine Classifier are proposed for predicting risk level with different split portions, with 80% reserved for training and 20% for testing [21]. Most of the previous work used individual algorithms without stacking different algorithms to improve prediction performance and attain better accuracy.

Table 1: Sample of Maternal Health Risk (MHR) Dataset

Age	Systolic BP	Diastolic BP	Blood Sugar	Body TP	Heart Rate	Risk Level
25	130	80	15	98	86	high risk
35	140	90	13	98	70	high risk
29	90	70	8	100	80	high risk
30	140	85	7	98	70	high risk
35	120	60	6.1	98	76	low risk
23	140	80	7.01	98	70	high risk
23	130	70	7.01	98	78	mid risk
35	85	60	11	102	86	high risk
32	120	90	6.9	98	70	mid risk
42	130	80	18	98	70	high risk
23	90	60	7.01	98	76	low risk
19	120	80	7	98	70	mid risk

The attributes in the dataset are as follows: (1) Age: age of the pregnant woman; (2) Systolic BP: upper value of blood pressure in mmHg; (3) Diastolic BP: low value of blood pressure in mmHg; (4) Body Temp: The measured body temperature of the patient in degrees Fahrenheit; (5) BS: Blood glucose levels are in terms of molar concentration, mmol/L; (6) Heart Rate: A normal resting heart rate in beats per minute; and (7) Risk Level: A predicted risk intensity level during pregnancy.

2.6 Proposed Model

It has been proven that ensemble learning performs significantly better in terms of generalization than any single learning algorithm. In this paper, stacked generalization techniques are applied to hybridized logistic regression and support vector machines to enhance predictive performance and boost model accuracy by reducing bias and variance. By methodically combining the predictive probabilities of two or more base learning algorithms, the ensemble learning algorithm improved predictive performance.

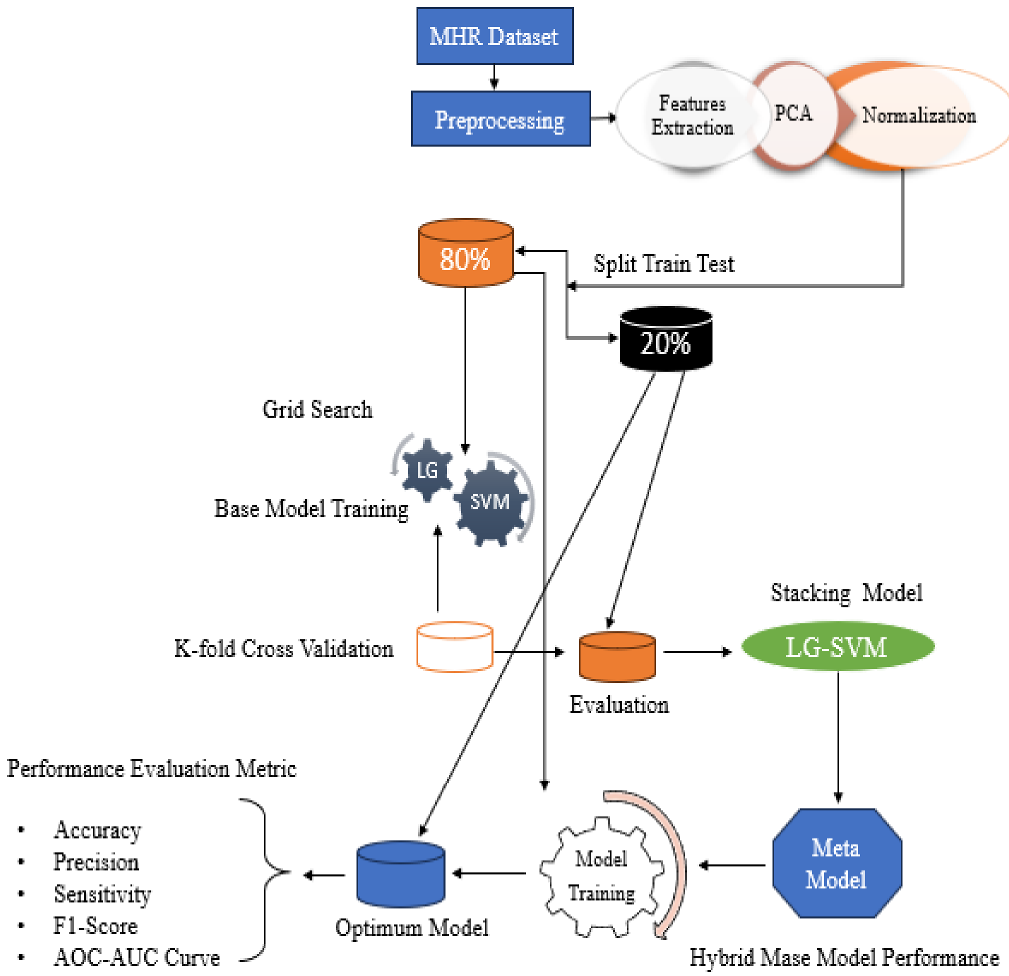

Figure 4: Flow chart of the proposed model.

Figure 4 represents the proposed model that will utilize the maternal health risk (MHR) dataset to predict the risk level during pregnancy. First, we supply the MHR dataset to the algorithm after applying pre-processing techniques, which are vital for achieving the best accuracy. Then, split the dataset into training and testing sets, with 80% reserved for training and validation and 20% reserved for testing. Train base learners by utilizing hyperparameter optimization and cross-validation to train all datapoints to achieve the best model performance and reduce over-fitting. Secondly, we used stacking techniques to combine the performance of each base learner to enhance the model's performance, and we employed a meta-model approach to predict the ultimate or optimal model performance. Finally, we evaluate the model's performance using standard performance evaluation metrics.

Algorithm 1: Pseudocode for both LG and SVM Model Using One-vs-rest (OVR)

Step 1: Input Training data $D=\{x_i, y_i\}^N$ where $x, y \in R^m$ $i=1,2,\ldots,N$

Step 2: Apply hyperparameter optimization using Grid Search.

Step 3: Apply cross validation using K-Fold in training phase for the first and second level base models.

Step 4: Split D into K Equal partition: $(D_1, D_2, \ldots, D_k)$

For $k:1 \leftarrow K$ do

End

Step 5: β, a learner (classifier for binary classification)

Step 6: features labels y where $y_i \in \{1,\ldots,k\}$ is the target labels

For each $k \in \{1,\ldots,K\}$

Construct a new label feature ϕ where $\phi_i = y_i$, if $y_i = k$ and $\phi_i = 0$

Apply β to D, ϕ to get p_k

Step 8: Return $\hat{y} = \underset{k\in\{1,\ldots,K\}}{\arg\max}\, p_k(x)$

2.7 Performance Evaluation Metric

We analyze the performance of our proposed model by executing a series of coding exercises using Python software. The performance evaluation metric is critical in the classification and regression models to gain a better understanding of how well the model performs in predicting each class label. We have employed various measures for performance evaluation to investigate and examine different levels of model performance. In this research work, the ROC-AUC curve, accuracy, precision, sensitivity, F1-score, and sensitivity are the typical performance evaluation measures that we use for classified algorithms. The employed metrics are as follows:

The accuracy is the ratio between the actual number of correctly identified features and the total classified features. the accuracy is constrained to be between 0 and 1 [22]. The model's accuracy is given as:

$$Accuracy = \frac{T_{pp} + T_{np}}{T_{pp} + T_{np} + F_{pp} + F_{np}} \tag{13}$$

Precision can be expressed as the ratio of correctly classified features compared to all features assigned to that specific class. The precision is constrained to be between 0 and 1 [23]. The precision is given as:

$$\Pr ecision = \frac{T_{pp}}{T_{pp} + F_{pp}} \tag{14}$$

The sensitivity, also termed the true positive rate (TPR), is the proportion of correctly categorized features that are correctly identified, and all features belong to the positive class. The sensitivity is constrained to be between 0 and 1 [24]. The sensitivity is given as:

$$Sensitivity = \frac{T_{pp}}{T_{pp} + F_{np}} \tag{15}$$

The F1-score is the value obtained by calculating the harmonic mean of precision and recall. F1-score is constrained to be between 0 and 1 Prediction [25]. The F1-score is given as:

$$F1 - Score = 2 * \frac{precision * recall}{precision + recall} \tag{16}$$

The ROC-AUC curve shows the predicted probability curve made by TPR and FPR using various threshold values to test how well and reliably a learning model can tell the difference between class labels [26]. When the ROC-AUC is greater than 0.5, the learning model has a chance of accurately classifying all class labels. The ROC-AUC is given as:

$$ROC - AUC = \frac{T_{PR}}{F_{PR}} \tag{17}$$

True positive prediction (T_{pp}) is the number of times the model correctly predicted the class labels. True negative prediction (T_{np}) The number of times the model correctly predicted the negative class and true positive rate (T_{PR}) are used to identify the correct classification in the confusion matrix, while false positive (F_{pp}) is the number of times a model predicted a positive class when the actual class label was negative, false negative (F_{np}) is the number of times a model predicted a negative class when the actual class label was positive and false positive rate (F_{PR}) are also used to identify misclassification of the sample features.

3 RESULTS AND DISCUSSION

This research has proposed a hybridized LG-SVM model to predict maternal health risk (MHR) levels during pregnancy. We compared the proposed model's result with conventional models such as random forest (RF), decision tree (DT), multilayer perceptron (MLP), Gaussian naive bayes (GNB), and logistic regression (LG) without the use of hyperparameter optimization and k-fold cross validation for predicting maternal health risk (MHR) levels. We analyze the findings using the performance evaluation metric to compare our proposed model's performance to conventional models in predicting the maternal health risk level during pregnancy. The proposed model has proven its effectiveness and robustness during training, validation, and testing stages. Our proposed model's efficiency in distinguishing between multiclass features on the target output determines its ability to correctly predict all class labels.

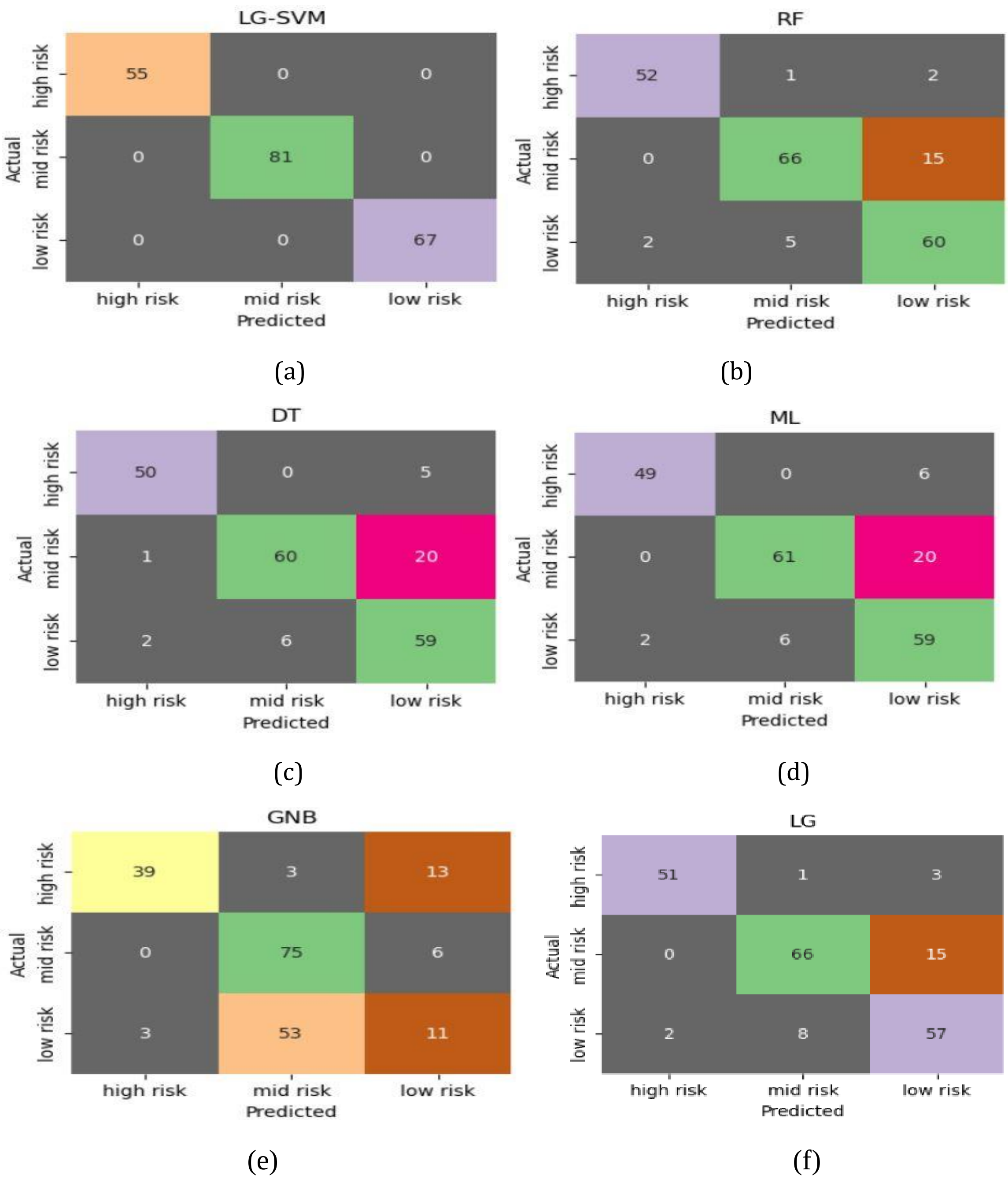

Figure 5: Confusion matrix during the testing phase for the proposed and conventional models for Maternal Health Risk (MHR) Dataset (a) LG-SVM, (b) RF, (c) DT, (d) MLP, and (e) GNB (f) LG Models.

Figure 5 represents the confusion matrix of the five (5) predictive models for the maternal health risk (MHR) dataset. The actual and predicted performance metrics in the confusion matrix are the number of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) for each class predictions. Based on prediction, the proposed model, LG-SVM, outperformed four other conventional models, RF, DT, MLP, GNB, and LG classifies every data point to their corrected classes without misclassification, whereas the conventional model has some misclassification of data points. The proposed model is more effective in predicting maternal health risk (MHR) levels and has proven its effectiveness to achieve better performance compared with the other conventional models since all class labels are predicted correctly.

Table 2: The weighted average performance for predictive maternal health risk (MHR) Dataset of the proposed and conventional models (%).

Models	Precision	Sensitivity	F1-Score	Accuracy
LG-SVM	100	100	100	100
RF	88	88	88	88
DT	85	83	83	83
MLP	85	83	84	83
GNB	60	62	58	62
LG	87	86	86	86

Table 2 shows the weighted average performance of our proposed and conventional model scores on how well the models were able to predict each class label. The results obtained are accuracy, precision, sensitivity, and the F1-score of the LG-KSVM, RF, DT, MLP, GNB and LG models. The proposed model outperformed the other four models for all performance evaluation metrics used to evaluate how well our model performed in predicting maternal health risk (MHR) levels during pregnancy. RF outperforms other conventional models and achieves better accuracy, while DT and MLP perform better compared with GNB and LG which achieves low accuracy. Utilizing hyperparameters for the proposed techniques proves their effectiveness and robustness in predicting all data points in their corrected class labels, unlike the other four models that use default parameters during training. Our proposed model has the potential to be used as an intelligent system for the early detection of risk levels for maternal health during pregnancy.

Looking at the results in Table 3, the accuracy and precision of the LG-SVM outperformed the four conventional models. The results demonstrated that the accuracy scores for RF, DT, MLP, GNB and LG are lower compared to the LG-KSVM model. The proposed model achieved 100% predictive accuracy with 0% misclassification compared to the conventional models that achieved 88%, 83%, 83%, 62%, and 86% predictive accuracy with 12%, 17%, 17%, 38%, and 14% misclassification. The Precision Score LG-SVM also outperformed the four conventional models. The results illustrate that the precision scores for RF, DT, MLP, GNB, and LG are lower compared to the LG-SVM model. This shows that the proposed model is more efficient in predicting and detecting risk levels during pregnancy compared to conventional models.

Table 3: performance analysis of our proposed model with conventional models for Maternal health Risk Level (%)

Models	Classes	Precision	Sensitivity	F1-Score	Accuracy	ROC-AUC
Hybrid LG-SVM	Low Risk	100	100	100	100	100
	Mid Risk	100	100	100		
	Hight Risk	100	100	100		
RF	Low Risk	96	95	95	88	96
	Mid Risk	92	81	86		
	Hight Risk	78	90	83		
DT	Low Risk	94	91	93	83	92
	Mid Risk	91	74	82		
	Hight Risk	70	88	78		
MLP	Low Risk	96	89	92	83	77
	Mid Risk	91	75	82		
	Hight Risk	69	88	78		
GNB	Low Risk	93	71	80	62	79
	Mid Risk	57	93	71		
	Hight Risk	37	16	23		
LG	Low Risk	96	93	94	86	96
	Mid Risk	88	81	85		
	Hight Risk	76	85	80		

The sensitivity and F1-score for LG-SVM achieved better scores than the other conventional models. The results demonstrated that the sensitivity scores for RF, DT, MLP, GNB and Lg are lower compared to the LG-SVM model. The F1-score values for LG-SVM outperformed the four conventional models. The results obtained on F1-score values for RF, DT, MLP, GNB, and LG are lower compared to the LG-SVM model. The receiver operating characteristics area and the area under the curve (ROC-AUC) determine the quality of the predictive models with different threshold values. The ROC-AUC for LG-SVM outperformed the conventional models. The results demonstrated that the ROC-AUC values scored by RF, DT, MLP, GNB, and LG are lower compared to the LG-SVM, which scored the highest ROC-AUC score of 100%. The proposed model's ROC-AUC effectively detects the maternal health risk level during pregnancy. This shows that the proposed model is more effective at predicting risk levels compared to conventional models. The proposed model has proven its effectiveness, robustness, and capability to predict the low, mid, and high-risk level classes correctly compared to conventional models. Our proposed model, LG-SVM, was able to get 100% correct predictive scores, though. This was made possible by using ensemble learning algorithms, hyperparameter optimization, and K-fold cross-validation.

Table 4: Comparison of model's accuracy rate (%) with previous study on maternal health risk (MHR)

XGB Classifier[21]	Light GBM [20]	Multilayer Perceptron [4]	KNN [27]	**Proposed Method**
69	84.24	90.29	96	**100**

Table 4 shows the accuracy rate of the previous works compared with the proposed method. Based on the results attained by each model, the proposed method supports the previous model of maternal health risk (MHR) in classifying different stages of health risk during pregnancy. The proposed method outperforms all four (4) models without misclassification. Based on the accuracy rate, the proposed method shows its robustness and effectiveness in predicting all class labels correctly, unlike the previous work that achieved less with misclassification. XGB attained less accuracy rate while KNN attained higher in the previous models.

Table 5: Accuracy rates of classifiers (%)

Hybrid LG-SVM	RF	DT	MLP	GNB	LG
100	88	83	83	62	86

Table 5 shows the comparison of accuracy scores attained by each model. The proposed approach achieved the highest accuracy and predicted all class labels correctly in the testing phase compared to conventional models. RF predicts 178 data points correctly, followed by DT and MLP, which predict 169 data points correctly, while GNB predicts only 125 data points correctly out of the 203 samples reserved for testing. The accuracy scores for our proposed approach and conventional models are compared in figure 6.

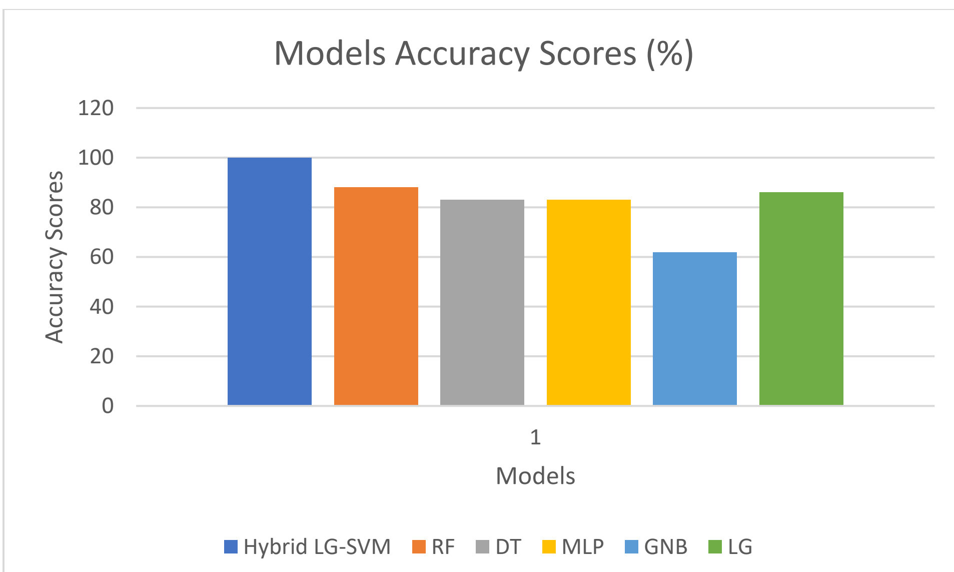

Figure 6: Accuracy Scores for the proposed and conventional models.

Figure 6 represents the accuracy score of each model obtained for modeling three different risk levels from the maternal health risk dataset. Our proposed model achieved the highest accuracy score of 100%, followed by Random Forest (RF) at 88%, Decision Tree (DT) at 83%, Multilayer Perceptron (MLP) at 83%, Gaussian Naive Bayes (GNB) at 62% and Logistic Regression (LG) at 86%. Therefore, the proposed approach achieved better accuracy and ROC-AUC scores compared to other conventional models. The use of ensemble learning algorithms and hyperparameter optimization techniques to hybridize LG-SVM improved performance and accuracy compared to individual models, demonstrating the robustness, effectiveness, and superior performance of hybrid models in prediction problems.

4 CONCLUSION

In this paperwork, we present a workable approach for developing an improved prototype to predict and classify maternal health risk levels during pregnancy. This research used four base models and the proposed hybridized learning models, LG-SVM, RF, DT, MLP, GNB, and LG to efficiently classify all datapoints correctly. The results illustrate that all five (5) models achieved better performance to meet the required objectives. Due to the utilization of hyperparameter optimization and k-fold cross validation during the training stage for our proposed model, it outperformed the other four conventional models. The use of hyperparameters in the training phase demonstrates that the predicting process is more robust and effective at diversifying the input features and searching for the best parameter combination that will achieve better accuracy.

The utilization of k-fold cross-validation is that it is required to train all datapoints to reduce issues of overfitting and underfitting during the training phase to determine the optimal model. We successfully compared the effectiveness and robustness of our proposed model to other techniques. In this paperwork, the result achieved by LG-SVM outperforms the RF, DT, MLP, GNB, and LG models using performance evaluation metrics. To enhance the capability, robustness, and the effectiveness of the machine learning models, we hybridized the outputs of the base models as inputs to the proposed model using an ensemble learning algorithm via the stacking approach. The LG-SVM design's stability and efficiency provide insight into potential dynamics for use in real-world situations. Taking data mining techniques into account, the suggested model identified the best features for predicting the risk level based on the maternal health risk during pregnancy.

Therefore, while most researchers came to different conclusions using different methodologies, the techniques also varied. Nevertheless, our proposed model uses ensemble learning with hyperparameter optimization and stratified K-fold cross validation to quantify and compare performance to standard models. Our model performs well in terms of accuracy and performance. Despite its superior performance, the proposed technique has limitations that must be acknowledged. When the target class has only two outputs, the proposal must be adjusted to meet the predictive analysis. In future investigations, other ensemble learning algorithms, metaheuristic algorithms, and hyperparameter optimization can be employed to perform different types of predictive analysis.

ACKNOWLEDGEMENT

This research was supported by Ministry of Education Malaysia (MOHE) through the Fundamental Research Grant Scheme (FRGS) with Grant Number FRGS/1/2022/STG06/USM/02/11 and Universiti sains Malaysia.

REFERENCES

[1] S. Mondal, A. Nag, A. K. Barman, and M. Karmakar, "Machine Learning-based maternal health risk prediction model for IoMT framework," *International Journal of Experimental Research and Review*, vol. 32, pp. 145–159, Aug. 2023, doi: 10.52756/ijerr.2023.v32.012.

[2] T. O. Togunwa, A. O. Babatunde, and K. U. R. Abdullah, "Deep hybrid model for maternal health risk classification in pregnancy: synergy of ANN and random forest," *Front Artif Intell*, vol. 6, 2023, doi: 10.3389/frai.2023.1213436.

[3] R. S. Kambli, "MATERNAL RISK LEVEL PREDICTION USING ENSEMBLE MODEL," 2023. [Online]. Available: www.ijcrt.org

[4] J. J. E. Macrohon, C. N. Villavicencio, X. A. Inbaraj, and J. H. Jeng, "A Semi-Supervised Machine Learning Approach in Predicting High-Risk Pregnancies in the Philippines," *Diagnostics*, vol. 12, no. 11, Nov. 2022, doi: 10.3390/diagnostics12112782.

[5] M. Ahmed, M. A. Kashem, M. Rahman, and S. Khatun, "Review and Analysis of Risk Factor of Maternal Health in Remote Area Using the Internet of Things (IoT)," in *Lecture Notes in Electrical Engineering*, Springer, 2020, pp. 357–365. doi: 10.1007/978-981-15-2317-5_30.

[6] T. Ning, Y. Yang, and Z. Du, "Quantum kernel logistic regression-based Newton method," *Physica A: Statistical Mechanics and its Applications*, vol. 611, Feb. 2023, doi: 10.1016/j.physa.2023.128454.

[7] S. K. Shevade and S. S. Keerthi, "A simple and efficient algorithm for gene selection using sparse logistic regression," *Bioinformatics*, vol. 19, no. 17, pp. 2246–2253, Nov. 2003, doi: 10.1093/bioinformatics/btg308.

[8] I. J. Kakoly, Md. R. Hoque, and N. Hasan, "Data-Driven Diabetes Risk Factor Prediction Using Machine Learning Algorithms with Feature Selection Technique," *Sustainability*, vol. 15, no. 6, p. 4930, Mar. 2023, doi: 10.3390/su15064930.

[9] M. M. Adankon and M. Cheriet, "Model selection for the LS-SVM. Application to handwriting recognition," *Pattern Recognit*, vol. 42, no. 12, pp. 3264–3270, Dec. 2009, doi: 10.1016/j.patcog.2008.10.023.

[10] J. Martínez, C. Iglesias, J. M. Matías, J. Taboada, and M. Araújo, "Solving the slate tile classification problem using a DAGSVM multiclassification algorithm based on SVM binary classifiers with a one-versus-all approach," *Appl Math Comput*, vol. 230, pp. 464–472, Mar. 2014, doi: 10.1016/j.amc.2013.12.087.

[11] E. Elgeldawi, A. Sayed, A. R. Galal, and A. M. Zaki, "Hyperparameter tuning for machine learning algorithms used for arabic sentiment analysis," *Informatics*, vol. 8, no. 4, Dec. 2021, doi: 10.3390/informatics8040079.

[12] M. S. M. Kasihmuddin, M. A. Mansor, and S. Sathasivam, "Hybrid genetic algorithm in the hopfield network for logic satisfiability problem," *Pertanika J Sci Technol*, vol. 25, no. 1, 2017.

[13] V. Nasteski, "An overview of the supervised machine learning methods," *HORIZONS.B*, vol. 4, pp. 51–62, Dec. 2017, doi: 10.20544/horizons.b.04.1.17.p05.

[14] F. Abramovich, V. Grinshtein, and T. Levy, "Multiclass classification by sparse multinomial logistic regression," *IEEE Trans Inf Theory*, vol. 67, no. 7, pp. 4637–4646, Jul. 2021, doi: 10.1109/TIT.2021.3075137.

[15] M. Bourel and A. M. Segura, "Multiclass classification methods in ecology," *Ecol Indic*, vol. 85, pp. 1012–1021, Feb. 2018, doi: 10.1016/j.ecolind.2017.11.031.

[16] E. J. Bredensteiner and K. P. Bennett, "Multicategory Classification by Support Vector Machines," 1999.

[17] J. Wu, X. Y. Chen, H. Zhang, L. D. Xiong, H. Lei, and S. H. Deng, "Hyperparameter optimization for machine learning models based on Bayesian optimization," *Journal of Electronic Science and Technology*, vol. 17, no. 1, pp. 26–40, Mar. 2019, doi: 10.11989/JEST.1674-862X.80904120.

[18] G. Battineni, N. Chintalapudi, and F. Amenta, "Machine learning in medicine: Performance calculation of dementia prediction by support vector machines (SVM)," *Inform Med Unlocked*, vol. 16, Jan. 2019, doi: 10.1016/j.imu.2019.100200.

[19] M. Nilashi, H. Ahmadi, L. Shahmoradi, O. Ibrahim, and E. Akbari, "A predictive method for hepatitis disease diagnosis using ensembles of neuro-fuzzy technique," *J Infect Public Health*, vol. 12, no. 1, pp. 13–20, Jan. 2019, doi: 10.1016/j.jiph.2018.09.009.

[20] H. B. MUTLU, F. DURMAZ, N. YÜCEL, E. CENGİL, and M. YILDIRIM, "Prediction of Maternal Health Risk with Traditional Machine Learning Methods," *NATURENGS MTU Journal of Engineering and Natural Sciences Malatya Turgut Ozal University*, Jun. 2023, doi: 10.46572/naturengs.1293185.

[21] S. Venkatesh, H. Jha, F. Kazmi, and S. Zaidi, "Classification of Maternal Health Risks Using Machine Learning Methods."

[22] M. M. Ali, B. K. Paul, K. Ahmed, F. M. Bui, J. M. W. Quinn, and M. A. Moni, "Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison," *Comput Biol Med*, vol. 136, Sep. 2021, doi: 10.1016/j.compbiomed.2021.104672.

[23] N. Afshan, Z. Mushtaq, F. S. Alamri, M. F. Qureshi, N. A. Khan, and I. Siddique, "Efficient thyroid disorder identification with weighted voting ensemble of super learners by using adaptive synthetic sampling technique," *AIMS Mathematics*, vol. 8, no. 10, pp. 24274–24309, 2023, doi: 10.3934/math.20231238.

[24] M. Atif, F. Anwer, F. Talib, R. Alam, and F. Masood, "Analysis of machine learning classifiers for predicting diabetes mellitus in the preliminary stage," *IAES International Journal of Artificial Intelligence*, vol. 12, no. 3, pp. 1302–1311, Sep. 2023, doi: 10.11591/ijai.v12.i3.pp1302-1311.

[25] D. Sisodia and D. S. Sisodia, “Prediction of Diabetes using Classification Algorithms,” in *Procedia Computer Science*, Elsevier B.V., 2018, pp. 1578–1585. doi: 10.1016/j.procs.2018.05.122.

[26] C. Li *et al.*, “Spread Prediction and Classification of Asian Giant Hornets Based on GM-Logistic and CSRF Models,” *Mathematics*, vol. 11, no. 6, Mar. 2023, doi: 10.3390/math11061332.

[27] A. Raza, H. U. R. Siddiqui, K. Munir, M. Almutairi, F. Rustam, and I. Ashraf, “Ensemble learning-based feature engineering to analyze maternal health during pregnancy and health risk prediction,” *PLoS One*, vol. 17, no. 11 November, Nov. 2022, doi: 10.1371/journal.pone.0276525.