Computers in Biology and Medicine 147 (2022) 105671

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

HGSORF: Henry Gas Solubility Optimization-based Random Forest for C-Section prediction and XAI-based cause analysis

Md Saiful Islam [a], Md. Abdul Awal [b,*], Jinnaton Nessa Laboni [b], Farhana Tazmim Pinki [c], Shatu Karmokar [b], Khondoker Mirazul Mumenin [b], Saad Al-Ahmadi [a], Md. Ashfikur Rahman [d], Md. Shahadat Hossain [e], Seyedali Mirjalili [f,g]

[a] Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
[b] Electronics and Communication Engineering Discipline, Khulna University, Khulna 9208, Bangladesh
[c] Computer Science and Engineering Discipline, Khulna University, Khulna 9208, Bangladesh
[d] Development Studies Discipline, Khulna University, Khulna 9208, Bangladesh
[e] Department of Quantitative Sciences, International University of Business Agriculture and Technology, Dhaka 1230, Bangladesh
[f] Centre for Artificial Intelligence Research and Optimization, Torrens University, Australia
[g] Yonsei Frontier Lab, Yonsei University, Seoul, South Korea

ARTICLE INFO

Dataset link: https://github.com/MIrazul29/HGSORF_CSection

Keywords:
Cesarean section
Machine learning
Hyperparameter optimization
ADASYN
HGSORF
XAI
SHAP
LIME

ABSTRACT

A stable predictive model is essential for forecasting the chances of cesarean or C-section (CS) delivery, as unnecessary CS delivery can adversely affect neonatal, maternal, and pediatric morbidity and mortality, and can incur significant financial burdens. Limited state-of-the-art machine learning models have been applied in this area in recent years, and the current models are insufficient to correctly predict the probability of CS delivery. To alleviate this drawback, we have proposed a Henry gas solubility optimization (HGSO)-based random forest (RF), with an improved objective function, called HGSORF, for the classification of CS and non-CS classes. Real-world CS datasets can be noisy, such as the Pakistan Demographic and Health Survey (PDHS) dataset used in this study. The HGSO can provide fine-tuned hyperparameters of RF by avoiding local minima points. To compare performance, Gaussian Naive Bayes (GNB), linear discriminant analysis (LDA), K-nearest neighbors (KNN), gradient boosting classifier (GBC), and logistic regression (LR) have been considered in this research. The ADAptive SYNthetic (ADASYN) algorithm has been used to balance the model, and the proposed HGSORF has been compared with other classifiers as well as with other studies. The superior performance was achieved by HGSORF with an accuracy of 98.33% for the PDHS dataset. The hyperparameters of RF have also been optimized by using commonly used hyperparameter-optimization algorithms, and the proposed HGSORF provided comparatively better performance. Additionally, to analyze the causes of CS and their significance, the HGSORF is explained locally and globally using eXplainable artificial intelligence (XAI)-based tools such as SHapely Additive exPlanation (SHAP) and Local Interpretable Model-Agnostic Explanations (LIME). A decision support system has been developed as a potential application to support clinical staffs. All pre-trained models and relevant codes are available on: https://github.com/MIrazul29/HGSORF_CSection.

1. Introduction

The importance of C-Section (CS) delivery to reduce fetal and maternal complications and mortality is recognized worldwide [1]. This procedure delivers babies through surgical incisions made in the mother's uterus and abdomen [2]. However, over the last decade, the growing rate of CS delivery has turned into a severe cause of anxiety for public health professionals worldwide [3].

The CS has been rising in low, medium, and high-income countries throughout the last decade [4,5]. The World Health Organization (WHO) reported that the CS rate has upsurged in developed and developing countries [6]. Unnecessary CS delivery can adversely affect neonatal, maternal, and pediatric morbidity and mortality. In addition, the tremendous cost of CS can be a significant financial burden for households and put additional strain on the health care system,

* Corresponding author.
E-mail addresses: saislam@ksu.edu.sa (M.S. Islam), m.awal@ece.ku.ac.bd (M.A. Awal), labonihasan456@gmail.com (J.N. Laboni), farhana@ku.ac.bd (F.T. Pinki), shatunatore98@gmail.com (S. Karmokar), k.mirazulmumenin@gmail.com (K.M. Mumenin), salahmadi@ksu.edu.sa (S. Al-Ahmadi), ashfikur@ku.ac.bd (M.A. Rahman), shahadat_qs@iubat.edu (M.S. Hossain), ali.mirjalili@torrens.edu.au (S. Mirjalili).

https://doi.org/10.1016/j.compbiomed.2022.105671
Received 12 November 2021; Received in revised form 24 May 2022; Accepted 24 May 2022
Available online 30 May 2022
0010-4825/© 2022 Elsevier Ltd. All rights reserved.

especially in low-and-medium-income countries (LMICs) [7–9]. To put an end to the negative consequences of unnecessary CS and save both mother and infant (in an emergency), the WHO has recommended a standard CS rate of 5%–15% [10]. The CS rate has also increased in south and southeast Asian countries, and the rise is more in urban than rural areas [11]. For example, in the wealth quantile of Pakistan, CS rates among educated women and average women are 39% and 34%, respectively (PDHS, (2012–13)). Additionally, the CS rate in Pakistan has drastically increased from 2.7% in 1990–1991 to 15.8% in 2012–2013, a rise of almost six times in two decades [12].

The Sustainable Development Goals (SDGs) aim to decrease the maternal mortality rate (MMR) per 100,000 live babies to below 70 and ensure a healthy life for all ages by 2030.[1] According to the latest Pakistan Maternal Mortality Survey (PMMS 2019), the Maternal mortality ratio (MMR) in Pakistan has declined from 276 deaths per 100,000 live births as per PDHS of 2006–7 to 186 [13]; however, the country failed to achieve a target of 75% reduction in maternal deaths by 2015 [14]. In addition, South Asian and sub-Saharan African countries have reported more CS delivery rates for deep socio-economic disparities than other parts of the world [15].

1.1. Related works

Many studies have shown that the rate of CS delivery relies on multiple demographic and socio-economic factors such as maternal age at birth, increased institutional delivery, prenatal care sites (private or public), number of previous deliveries, socio-economic condition of the family, and access to antenatal care [16–18]. Although the application of Machine Learning (ML) algorithms is increasing substantially in the medical diagnosis concerning prognosticative analysis, there is a limited number of studies in which researchers used ML-based models for predicting CS delivery with high accuracy on large datasets.

Here, we review the contemporary literature relevant to this study. Abbas et al. [19] used supervised ML algorithms to classify CS and normal delivery. They collected data from two hospitals in Muzaffarabad for classification purposes, which contained 23 factors with 488 subjects. The authors observed that the CS rates of women aged 17–20 and over 30 were much higher than the middle-aged group (22–28 years). The researchers used random forest (RF), linear discriminant analysis (LDA), K-nearest neighbors (KNN), support vector machine (SVM), Naïve Bayes (NB), and AdaBoost classifiers for making predictions. Among all of these classifiers, RF provided the maximum accuracy of 91.8%.

Dulitzki et al. [20] used logistic regression (LR) analysis to determine outcomes of pregnancies in women of 20–44 years old and also to identify the factors related to CS delivery. The researchers also demonstrated the correlation between maternal age and CS. They also observed that women of 44 years and above had more medical complications like diabetes, premature delivery, premature contractions, hypertension, etc. Some researchers predicted the results of deliveries based on ultrasound measurement values. For example, Khazardoost et al. [21] collated the Bishop score with a predictive value measured from translabial ultrasound to ascertain the appropriateness of labor induction. They took 100 women's data with a mean age of 25.1 ± 4.4. Maternal body mass index(BMI) of the CS patients (P : 0.02) was much higher than the normal vaginal delivery (NVD) patients (P : 0.004). The researchers also considered BMI to be an essential factor that affects the progress of the delivery. Cervical length and fetal head symphysis distance were measured by translabial ultrasound with a test sensitivity of 90% and 88%, respectively. Robu and Holban [22] applied the NB, J48, KNN, RF, AdaBoost, SVM, JRipp, LogitBoost, REP Tree, and simple CART algorithms to historical data consisting of 2086 recorded births, where the highest accuracy was obtained from LogitBoost (80%). Sodsee [23] applied a modified nearest neighbor analysis named as CPD-NN algorithm to predict CS, and the accuracy significantly increased. With 100 training and 702 test cases, Euclidean distance, cosine similarity, and correlation provided 90%–100% of prediction accuracy and less than 90% accuracy by using Manhattan distance with $D_{max} < 0.7$. Sana et al. [24] developed a birth prediction classification (BPC) model and selected important factors responsible for the increasing rate of CS. They collected data from 15 hospitals in Sargodha, Pakistan, and identified 50 factors that influence the type of birth. Hasan et al. [25] applied stepwise LR to a dataset of 4422 married Bangladeshi women who have at least one child (age $\leq$ 5 years). They observed that location of residence, education of husband and his family, age of first birth, husband's occupation, wealth index, baby's weight, and the total number of children were the most significant reasons for CS delivery among the Bangladeshi women.

Several limitations can be identified in the existing studies [19–26]. Most of the research works on ML in mother's healthcare used small sample sizes and was mainly based on LR. These methods were not very successful in solving the optimization problem, and as a consequence, they yielded low prediction accuracy. Studies show that tree-based classifiers such as RF provide better accuracy, and the model behaviors can easily be explained using explainable artificial intelligence (XAI). RF has several hyperparameters that influence the learning process and, consequently, determine the prediction accuracy. Another limitation of several existing methods in the literature is that most of the studies mainly focused on predicting delivery types rather than the essential features affecting the delivery modes. These constraints inspired us to predict delivery modes and find the important factors associated with CS delivery to help care providers and mothers.

1.2. Our contributions

We proposed a novel Henry Gas Solubility Optimization (HGSO)-based random forest (RF), namely HGSORF, to mitigate the hyperparameter optimization problem and applied XAI to analyze the cause of CS. This study has utilized PDHS datasets to build the CS model. The major contributions of this study are given as follows:

- A novel optimized ML classifier called HGSORF has been proposed and applied for the first time to CS birth prediction balanced by the adaptive synthetic (ADASYN) algorithm.
- The performance of the proposed HGSORF has been compared with other state-of-the-art classifiers and contemporary studies in terms of statistical measures and also other frequently used optimization algorithms.
- XAI has been used to interpret the model performance using Shapely additive explanations (SHAP) and local interpretable model-agnostic explanations (LIME), and a feature importance curve has been presented to understand the causes and influencing factors of CS delivery.

In this study, we have observed the relationship between the CS rate in a particular population over time to inform policy-makers/caregivers and emphasized strategies for improving maternal mortality.

2. Materials and methods

The workflow of our proposed HGSORF model is shown in Fig. 1. It comprises several sequential steps.

Initially, the CS dataset is collected, scaled, and imputed, followed by a class balancing step with the help of the ADASYN algorithm. Next, the dataset is split into train and test sets, which are used for learning and classification by implementing several state-of-the-art classification algorithms, such as gradient boosting classifier (GBC), Gaussian Naive Bayes (GNB), KNN, LDA, LR, and the proposed method HGSORF. Our

[1] https://sustainabledevelopment.un.org/post2015/transformingourworld (Accessed 05 August 2021).

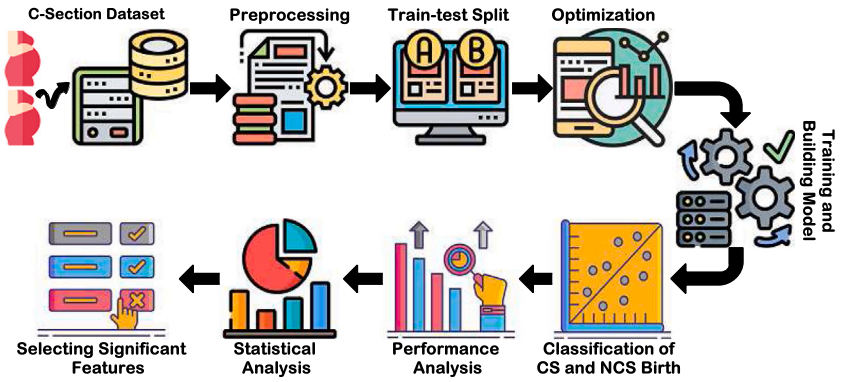

Fig. 1. The workflow of the proposed study.

Table 1
Regional cesarean rates of respective participants.

Sl. no.	Region	Cesarean rate in 2017–2018	Cesarean rate in 2012–2013
1	Islamabad (ICT)	31.9%	25.5%
2	Federally Administered TribalAreas (FATA)	2.6%	No data available
3	Balochistan	5.1%	1.9%
4	Khyber Pakhtunkhwa (KPK)	9.1%	6.1%
5	Sindh	24.3%	18.1%
6	Punjab	27.8%%	19.1%
7	Azad Kashmir and Jammu(AKJ)	21.6%	No data available
8	Gilgit Baltistan	12.3%	4.2%

recently proposed HGSO [27] has been adapted with an improved objective function to optimize the hyperparameters of the classifiers. A model is established and trained, followed by the model's assessment. Several statistical analyses, such as confusion matrix, violin plot, ANOVA test, etc., have been used before selecting the most dominant features with LIME and SHAP analysis.

2.1. Data source

The Demographic and Health Surveys (DHS) program provides the most representative and widely available data for CS births based on country populations worldwide. The dataset used for the present study was collected from the PDHS, and we used PDHS'17-18 and '12-13 datasets for the current study. These were all cross-sectional studies directed by the National Institute of Population Studies (NIPS) of the Ministry of National Health Services, Regulations and Coordination of Pakistan.[2] The PDHS provides a comprehensive guideline of the population study including child and maternal health status, including other related areas of health care and women's empowerment in family planning. The details of the survey's strategy, procedure, and questionnaires can be found in the final country report of Pakistan.[3] We used secondary datasets of PDHS, which were collected from DHS and are available online.[4] The PDHS'17-18 and '12-13 datasets contain a total of 15,409 married women aged between 15 and 49 years who had at least one delivery in the last five years. The region-based CS rates are presented for Pakistan in Table 1 and the CS rate of PDHS'17-18 is exhibited in Fig. 2.

The DHS datasets are publicly available for research and academic purposes. All the DHS survey protocols received ethical consent from both the Institutional and country-specific review boards [28]. We used secondary data of the PDHS, which obtained ethical approval for the primary studies; thus, there are no issues of ethical approval for the present study. The organization that collected the data is responsible for protecting appropriate ethical permissions before collecting the data. Therefore, further ethical consent is not required for the present study to predict CS delivery. The PDHS'17–18 and '12–13 datasets contained a total of 875 features, where some features were related to household information, contraceptive knowledge and practice, post-delivery, children's health care, nutritional information, and migration

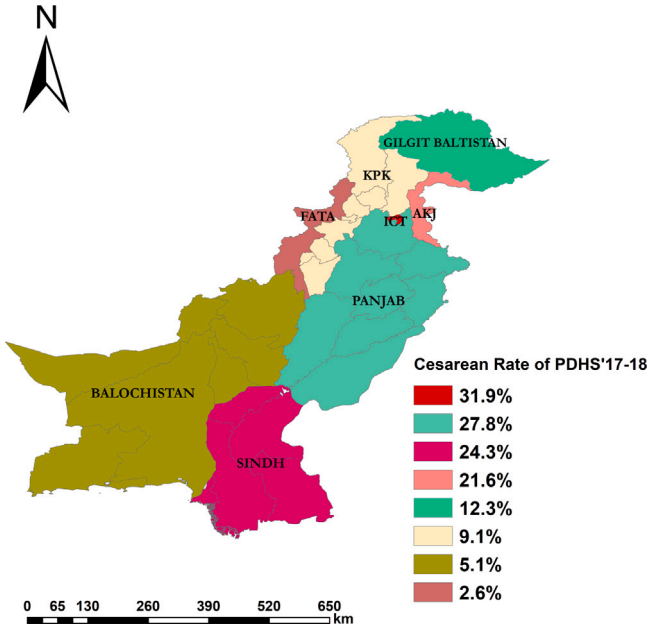

Fig. 2. Division-wise CS birth rate. The ArcGIS 10.8 software was used to plot this Figure.

[2] http://dhsprogram.com/pubs/pdf/FR354/FR354.pdf (Accessed 05 April 2021).
[3] http://dhsprogram.com/pubs/pdf/FR354/FR354.pdf (Accessed 05 April 2021).
[4] https://www.dhsprogram.com/ (Accessed 06 April 2021).

Table 2
Description of the variables.

Features name	Features short descriptions	
Numerical Variables		Mean ± STD
Age of Mother	Current age of mother at last birth ranged 15–49 years.	29.192 ± 5.818
Number of household members	How many members have in respondent's family.	9.122 ± 4.748
Total children ever born	The number of pregnancy occurred before the last birth.	3.308 ± 2.081
Number of living children	The number of children respondents have at present.	3.103 ± 1.922
Mother's BMI	Mother's Body Mass Index ranged 12.93–52.48	25.676 ± 5.458
Number of ANC Visits	The number of antenatal care visits during pregnancy ranged 0–20.	4.045 ± 2.722
Categorical Variables		χ^2-test (P-value)
Type of Place of Residence	The place (Urban or Rural) where respondent lives.	827.450(5.807e − 182)
Mother's Education	The educational level (No Education, Primary, Secondary, Higher) of respondent .	476.459(6.025e − 103)
Births in last five years	The number of children respondent gave birth within last 5 years.	195.827(2.964e − 41)
Ever had a Terminated Pregnancy	Whether the mother has had an abortion.	182.038(1.739e − 41)
Sources of Drinking Water	The type of water (Improved or Non-improved) mother's family drink.	6.172(0.012)
Household Toilet Facility	The type of toilet (Hygienic or Unhygienic) mother's family used.	30.210(3.876e − 08)
Reading Newspaper or Magazine	Frequency of reading newspaper or magazine in a week.	12.205(0.0004)
Listening to Radio	Frequency listening radio in a week.	70.214(5.320e − 17)
Watching-TV	Frequency watching TV in a week.	174.986(6.028e − 40)
Household Wealth Quantile	The economic status of that family.	682.118(2.593e − 146)
Smoker	Frequency of smoking in a week.	72.840(1.405e − 17)
Husband's Education	The educational level (No Education, Primary, Secondary, Higher) of respondent's husband.	827.450(1.133e − 37)
Husband's occupations	The working field (No Work, Agricultural, Professional/Services, Business, Others) of respondent's husband.	588.089(5.858e − 126)
Mother's Occupation	The working status (Yes or No) of respondent.	116.977(2.903e − 27)
Decision Making Power on Delivery Place	The person who takes decision on health care in a family.	553.519(1.199e − 119)
Suffered by Domestic Violence	Frequency of domestic violence suffer by mother.	987.086(1.151e − 216)
Had Previous C-Section	The history of previous C-Section of mother.	4272.328(0.0)
Size of child at birth	The growth of infant during labor.	199.357(5.164e − 42)
Last Birth a Caesarean Section		

patterns.[5] These features were unnecessary for prediction purposes, so they were excluded from the features set based on our experience and the literature. An extensive review was performed on the prevalence of CS delivery in Bangladesh [25,29–32], Pakistan [33,34], Nepal [35–39], India [18,40,41], Ghana [42], Iran [43], Tanzania [28], and East African countries [44,45] to identify independent variables. In the refined dataset, 24 independent and correlated variables were used for predicting the target variable (CS birth).

For life-threatening and high-risk CS operations, the dependent variable, i.e., usage status variable, was built from the combined response to whether a respondent underwent any CS operation in her last delivery. The target variable "Last birth a CS" is binary in the dataset, where non-CS delivery is denoted by '0' and CS delivery is characterized by '1'. Independent features included socio-demographic characteristics of the mother, household members, community, and variables related to pregnancy. In the selected 24 independent variables, there were six numerical and 18 categorical variables, which have been provided in Table 2. For example, $Mean \pm STD$ has been enumerated to describe the numerical variables, whereas the meaningful relationship between categorical and dependent variables has been depicted using χ^2-test.

2.2. Data preprocessing

In this section, the dataset containing the selected independent variables and the target variable (Last birth a CS) was prepared to load into the prediction model. We eliminated all noisy data from the converted CSV. There were also some missing values of independent variables. We have used the KNN imputer by Scikit-learn library to impute these missing values, where the nearest neighbors are enumerated with the Euclidean distance metric. The value estimated by $k = 5$-nearest neighbors replaced missing values in the dataset. However, there was a class imbalance problem in the dataset: 2131 CS birth instances compared to 13278 NCS birth instances. To develop a balanced model, we require the same amount of data for positive and negative classes [46]. There were two resampling methods to handle this problem: (i) over-sampling and (ii) undersampling. The over-sampling method is preferable because it increases samples of minority classes for balancing with the majority class. Hence, the ADASYN [47] technique was used to solve the problem by generating new instances adaptively for the minority class. The generated synthetic data can reduce the learning bias of actual imbalanced data distribution and shift the decision boundary adaptively. After applying ADASYN, the number of instances became equal for both CS birth and NCS birth classes (15,409 cases). The train-test-split was performed randomly to split the dataset with 70% of the data for the training and 30% for testing purposes. Then, the training dataset was ready to be fed into the ML model to predict CS birth.

2.3. Data classification using RF

The decision tree algorithm is relatively easy to understand and explain. However, a single decision tree is often insufficient for massive statistical analysis. RF is a multiple-decision-tree-based ML algorithm that attempts to solve the problems. The ensemble learning algorithm is a general term for an RF-type classifier [48]. The RF classifier generates a number of trees like CART, and trained on a bootstrapped sample of the original training data. The output of the classifier is decided by a majority vote. The RF algorithms search randomly selected subsets of input variables to define a node. Bagging is a method of creating multiple random subsamples by replacing the original training dataset with the same value multiple times. It helps to adjust each selected feature. The default value is set to the square root of the number

[5] https://www.dhsprogram.com/ (Accessed 06 April 2020).

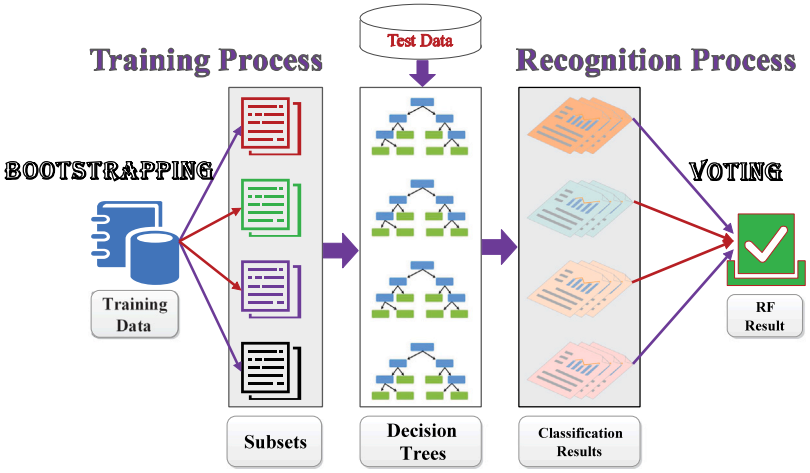

Fig. 3. Conceptual framework of random forest classifier.

of inputs. The algorithm's computational complexity depends on the correlation between trees and can be reduced by limiting the number of variables used for division. Finally, the pruning of the trees is done, further reducing the load. As a result, high dimensional data can be handled by this algorithm with a considerable number of trees in this ensemble algorithm. The entire process of classification based on the RF is shown in Fig. 3.

The main steps in creating an RF classifier are as follows:

Step 1: Set an exact value for the variable named M, the number of elements of a feature subset.
Step 2: Depending on the value of M, a new feature subset (θ_k) is selected randomly from the exclusive feature.
Step 3: Training dataset with feature subsets is used to create decision trees for each group of the training set. The expression of each classifier is $h(Y, \theta_k)$, where Y denotes the inputs.
Step 4: Choose a new θ_k and repeat the process above until all feature subsets have traveled.
Step 5: Input the test dataset and determine the class labels based on the voting results.

In the RF classifier, *n_estimators*, *max_features*, *min_samples_split*, *min_samples_leaf*, and *max_samples* hyperparameters play an important role during the model training and learning process. Here, *n_estimators* represents the total number of trees in the forest, *max_features* denotes the number of features used during tree split, *min_samples_split* indicates the most undersized samples needed to separate the internal node, *min_samples_leaf* represents the most miniature samples needed to be present in a leaf node, and *max_samples* denotes the number of samples required to train a base estimator. These important hyperparameters are required to be optimized in order to obtain an improved classification result.

2.4. Optimization problem formulation

It has already been mentioned that most of the classifiers are hyperparameter-sensitive, which implies that the classification performance is likely to be accelerated by optimizing the hyperparameters of the classifiers [49]. In our proposed method, five important hyperparameters of RF, namely *n_estimators*, *max_features*, *min_samples_split*, *min_samples_leaf*, and *max_samples*, have been optimized using HGSO together with an improved objective function. The general framework of the proposed optimization problem is defined as follows:

$$\arg\min_{d \in D} J\,(Clf(d); D) \tag{1}$$

where $J(.)$ represents the objective function, d symbolizes the hyperparameters $d_1, d_2, d_3, \cdots, d_n \in D$, Clf denotes the RF classifier. This optimization problem can be interpreted as *"select the optimal hyperparameters of RF (D) for which $J(\cdot)$ is the minimum"*. The HGSO will be adapted to solve this optimization problem.

In this study, we have proposed to use the mean of $N_{CVT} = 5$ fold kappa score loss calculated from the training dataset as an objective function $J(\cdot)$. The kappa score indicates the reliability of a classifier, and the objective function can be expressed by Eq. (2).

$$J = 1 - \left[\frac{1}{N_{CVT}} \times \sum_{CVT=1}^{N_{CVT}} \left[\frac{2\,(TP \times TN - FP \times FN)}{(TP+FP)(FP+TN)+(TP+FN)(FN+TN)}\right]_{CVT}\right] \tag{2}$$

2.5. Hyperparameter tuning using HGSO algorithm

Hyperparameter optimization is a major problem in ML applications. Since ML algorithms are susceptible to many hyperparameters, the performance of these algorithms is magnified by determining an appropriate set of hyperparameters [50]. There are many optimization techniques, such as exhaustive search, gradient descent, genetic algorithms [51], whale optimizer [52], Bayesian optimization [53], etc. However, we have used the novel HGSO, a physics-based optimization technique that follows Henry's gas law, to balance the exploitation and exploration phases in the search space and avoid local optima [27]. We have adapted this state-of-the-art HGSO algorithm for tuning the hyperparameters of RF. For this reason, we need to initialize the position of hyperparameters with population size, N, using Eq. (3).

$$X_i(m+1) = X_{lb} + r \times (X_{ub} - X_{lb}) \tag{3}$$

where $r \epsilon [0, 1]$, X_{ub} and X_{lb} are the lower and upper bounds of X, respectively, and $i = 1, 2, 3, \ldots, N$. In addition, it is necessary to initialize several parameters such as $H_j(m)$, $P_{i,j}$, C_j, which denote Henry's coefficient of j cluster, partial pressure, and enthalpy of dissolution, respectively. After that, divide the initial hyperparameters of RF into clusters having the same Henry's constant H_j. Then evaluate each cluster (j) to identify and rank the optimal hyperparameters, and update Henry's coefficient of each cluster using Eq. (4).

$$H_j(m+1) = H_j(m) \times e^{-C_j\left(\frac{1}{T(m)} - \frac{1}{298.15}\right)}, T(m) = e^{\left(-\frac{m}{M}\right)} \tag{4}$$

where M is the total number of iterations. Then update the solubility $(S_{i,j})$ of the hyperparameter of RF using Eq. (5).

$$S_{i,j}(m) = K \times H_j(m+1) \times P_{i,j}(m) \tag{5}$$

where K is constant. Afterward, update the position of hyperparameters using Eq. (6),

$$X_{i,j}(m+1) = X_{i,j}(m) + rF\left[\gamma\left(X_{i,best}(m) - X_{i,j}(m)\right) + \alpha\left(S_{i,j}(m) \times X_{best}(m) - X_{i,j}(m)\right)\right] \quad (6)$$

where $\gamma = \beta \times e^{\left[-\frac{(F_{best}(m)+0.05)}{(F_{i,j}(m)+0.05)}\right]}$, $\alpha = 1$ and β is constant. $F_{i,j}$ is the fitness of hyperparameter i in j cluster, F_{best} is the fitness of best hyperparameter, $X_{i,best}$ denotes best hyperparameter i in j cluster, and X_{best} is the best hyperparameter agent. Next rank and select worst hyperparameter agent using Eq. (7),

$$N_w = N \times \left(r(C_2 - C_1) + C_1\right); C_1 = 0.1; C_2 = 0.2 \quad (7)$$

Then update the position of the worst hyperparameter agents using Eq. (8).

$$G_{i,j} = G^{lb}_{i,j} + r \times (G^{ub}_{i,j} - G^{lb}_{i,j}) \quad (8)$$

where $G_{i,j}$ represents the position of gas i in cluster j, r is random number, and $G^{ub}_{i,j}$ and $G^{lb}_{i,j}$ are the boundary of the problem. Now, update the best hyperparameter $X_{i,best}$ and best hyperparameter agent X_{best}.

Finally, it iterates until it meets the maximum number of iterations M. The pseudo-code regarding the HGSO has also been adapted to optimize the hyperparameters of the RF classifier; hence the algorithm is named as HGSORF and portrayed in **Algorithm 1**.

Algorithm 1: The hyperparameter optimization of RF using the HGSO algorithm

Input: Training dataset, D-dimensional hyperparameters of RF, N, M, X_{ub}, X_{lb}, i, H_j, $P_{i,j}$, C_j.
Output: Optimized hyperparameters of RF and the HGSORF model.

```
Initialization of RF hyperparameters position X_i between (X_ub) and (X_lb) where X_i is (N × D) dimensional matrix.
Divide the RF hyperparameters into several clusters with the same H_j.
Evaluate each cluster (j).
Get X_{i,best} and X_best.
while m < M, do
    for each search agent, do
        Update the positions of D dimensional RF hyperparameters using Eq. (6).
    end for
    Update Henry's coefficient for each RF hyperparameter using Eq. (4).
    Update the solubility (S_{i,j}) of the hyperparameter of RF using Eq. (5).
    Rank and select out the worst RF hyperparameter agents using Eq. (7).
    Update the position of the worst RF hyperparameter agents using Eq. (8).
    Update X_{i,best} and X_best.
    m ← m + 1
end while
return the best hyperparameters of RF and the best fitness value.
Build the HGSORF model using the best hyperparameters of RF from the training dataset.
```

2.6. Model evaluation

We have used some model evaluation metrics, namely accuracy score, F1 score, kappa score, Matthews correlation coefficient (MCC) score, and sensitivity score, and computed their values from the confusion matrix to evaluate the performance of the proposed model. The higher value of these scores represents that the model provides better performance than other ML models. From the confusion matrix, we can obtain true positive (TP), false positive (FP), true negative (TN), and False Negative (FN), which are necessary for evaluating performance metrics. Here, TP, FP, TN, and FN represent a correct prediction of cesarean birth, incorrect prediction of cesarean birth, correct prediction of NCS birth, and incorrect prediction of NCS birth, respectively. The evaluation metrics are calculated using Eq. (9) to (13).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (9)$$

$$F1score = 2 \times \frac{precision \times recall}{precision + recall} \quad (10)$$

where $precision = \frac{TP}{TP+FP}$ and $recall = \frac{TP}{TP+FN}$.

$$Kappa = \frac{2(TP \times TN - FP \times FN)}{(TP + FP)(FP + TN) + (TP + FN)(FN + TN)} \quad (11)$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (12)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (13)$$

For a classification model, we have also used other performance evaluation techniques described in the following paragraphs.

The Receiver Operating Characteristic (ROC) curve, a two-dimensional graphical representation, visualizes the classification performance of an ML model at all possible thresholds. While plotting the ROC curve, the true positive rate (TPR) or sensitivity is represented along the y-axis, and the false positive rate (FPR) or specificity is on the x-axis. Moreover, the area under the ROC curve (AUC) is calculated from the area covered by the ROC curve, where a higher value indicates a better performance.

A precision–recall curve (PRC) is a convenient evaluation measure for imbalanced classes classified by a binary classification model. The PRC is plotted based on two evaluation measures-recall and precision at all possible thresholds, where recall and precision are presented on the x and y-axis, respectively. The recall is identical to sensitivity, and precision is identical to positive predictive value. A pair of x and y values (x, y) in the PRC is called a precision–recall point.

Bootstrap is a resampling method where many small samples of the same size from existing data are repeatedly taken with replacement to estimate a population. It is used to determine the various hyperparameters of a population and evaluate the efficiency of the ML model for predicting data that are not included in the training data. The bootstrap technique obtains the sampling distributions of outputs for the positive class and negative class. The Bootstrap ROC curve is plotted by all the (FPR, TPR) pairs by varying the decision threshold over the whole range of the bootstrap sampling distribution.

2.7. XAI

An array of tools used to interpret and comprehend any predictive model using ML algorithms thoroughly is defined as XAI. The XAI is a key part of broader human-centric responsible AI practices, which illustrates the potential biases, expected impact, explainability, and accountability of the model. The interpretability focuses on model understanding techniques, translating these explanations into understandable human terms. Thus, the XAI bridges the gap between ML models and human systems. The enormous growth of AI-based techniques can play a pivotal role in explaining the ML models and attaining the faithfulness of the models to the ultimate users. Numerous XAI techniques (e.g., SHAP, DeepSHAP, DeepLIFT, CXplain, and LIME) have been introduced. All the XAI techniques, except SHAP and LIME, requires a long time to explain the model (as well as the significance of the features) and, thus, could not gain acceptability to the users. Because of this, SHAP (for global explanation) and LIME (for local explanation) have been utilized in our research to explain the feature

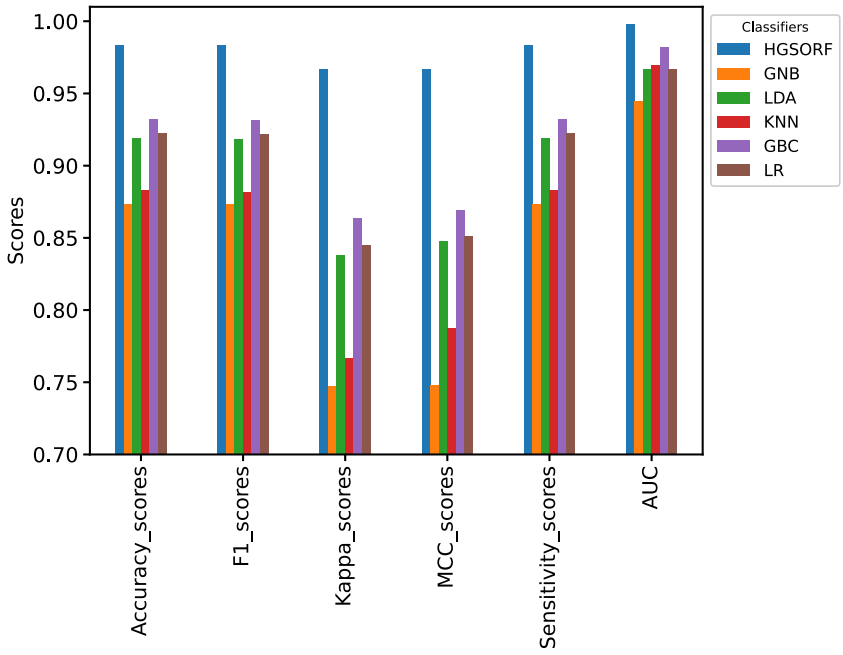

Fig. 4. Comparative performance of different classifiers.

importance and, most importantly, the abstraction behind it. The use of both SHAP and LIME will help the readers and users grasp the latent meaning of our proposed framework. The mathematical formulation for ascertaining the SHAP value, ϕ_i, was initially introduced by Lundberg and Lee [54], which is given as follows:

$$\phi_i = \sum_{S \subseteq Q \setminus \{i\}} \frac{|S|!(M-|S|-1)!}{M!}[f_x(S \cup i) - f_x(S)] \tag{14}$$

where M, Q, and S symbolize the total input features, set of all input features, and a subset of the input features, respectively.

On the other hand, to provide a mathematical explanation of LIME, Ribeiro et al. [55] introduced a local surrogate model, where the training process was executed on the perturbations of the data chosen for investigation. A formal framework shown in Eq. (15) was proposed to illustrate the trade-off between interpretability and fidelity.

$$\psi(x) = \arg\min_{h \in H} \mathcal{L}\left(f, H, \pi_x'\right) + \chi(h) \tag{15}$$

where π_x' is the local kernel to enumerate the loss; $\psi(x)$ denotes the explanation; the unfaithfulness of h while approximating f is measured using the inverse value of the local fidelity, $\mathcal{L}\left(f, H, \pi_x\right)$; $\chi(h)$ represents the complexity of the local model h (also called penalizing constraint to mitigate the time complexity). Therefore, it is essential to maximize the local fidelity, i.e., to minimize the unfaithfulness of the proposed ML model.

3. Results and discussion

The experimental analyses of our proposed HGSORF model are described in Section 3.1. The feature importance calculated by XAI and cause analysis of features are discussed in Section 3.2. Also, the proposed model is compared with other existing methods and some other optimization algorithms, which are discussed in Sections 3.3 and 3.4, respectively.

3.1. Performance evaluation

In this study, an optimized model has been created, where the hyperparameters have been controlled using HGSO. The performance of the model has been enumerated for different classifiers and compared. Fig. 4 presents the comparative study of various statistical measures for the classifiers used here. It is clear from the graphical visualization that the proposed HGSORF achieved the uppermost point in classification measures, whereas GNB occupied the lowermost point. It could be observed that HGSORF achieved classification performance (accuracy, F1 score, and AUC) above 98%, while its sensitivity is approximately 98.94%.

The confusion matrix, a two-dimensional visualization tool, illustrates the correct classification of the target variable of any ML model. For instance, Fig. 5(a) shows the percentage of correct classification of births through CS and NCS using GBC, where the percentage of CS is 43.56 with a total accuracy of 93.20%. Fig. 5(b) represents the confusion matrix while using GNB and obtained an accuracy of 87.36%. Similarly, with KNN, LDA, and LR, the obtained accuracy was 88.32%, 91.90%, and 92.24%, respectively. On the other hand, the proposed HGSORF provided us with the topmost accuracy of 98.34%.

Since the ROC curve represents the performance of the classification algorithms at the probable thresholds, the ROC curve has been plotted to evaluate our proposed model's performance. For instance, the ROC curve (Fig. 6) shows the results of the successfully classified CS and NCS classes for the ADASYN-balanced CS dataset. This figure illustrates that HGSORF provides the topmost accuracy.

Besides ROC analysis, the PRC is employed to visualize the overall performance of the classification task while evaluating and comparing the test results. The perfectness of the test is determined by the closeness of PRC to the top right corner of the graph. As Fig. 7 delineates, the PRC obtained for HGSORF is the closest to the top right corner, implying the superiority of the model.

The sensitivity of the optimized model to the training dataset is determined by using bootstrapping on the HGSORF-based model. A slightly different discriminating ability is obtained through providing N_{boot} to the HGSORF. For instance, in Fig. 8, $N = 100$ bootstraps were supplied with "mean" ROC and 95% confidence interval using optimized HGSORF. The obtained result suggests that an HGSO optimized and ADASYN-balanced model can achieve very high performance.

The comparative study regarding the cross-validation (CV) accuracy of different classifiers has been visualized in Fig. 9 using violin plot representation. It has been observed that the average accuracy obtained from KNN, LDA, GBC, and LR is approximately 92%. The most exciting

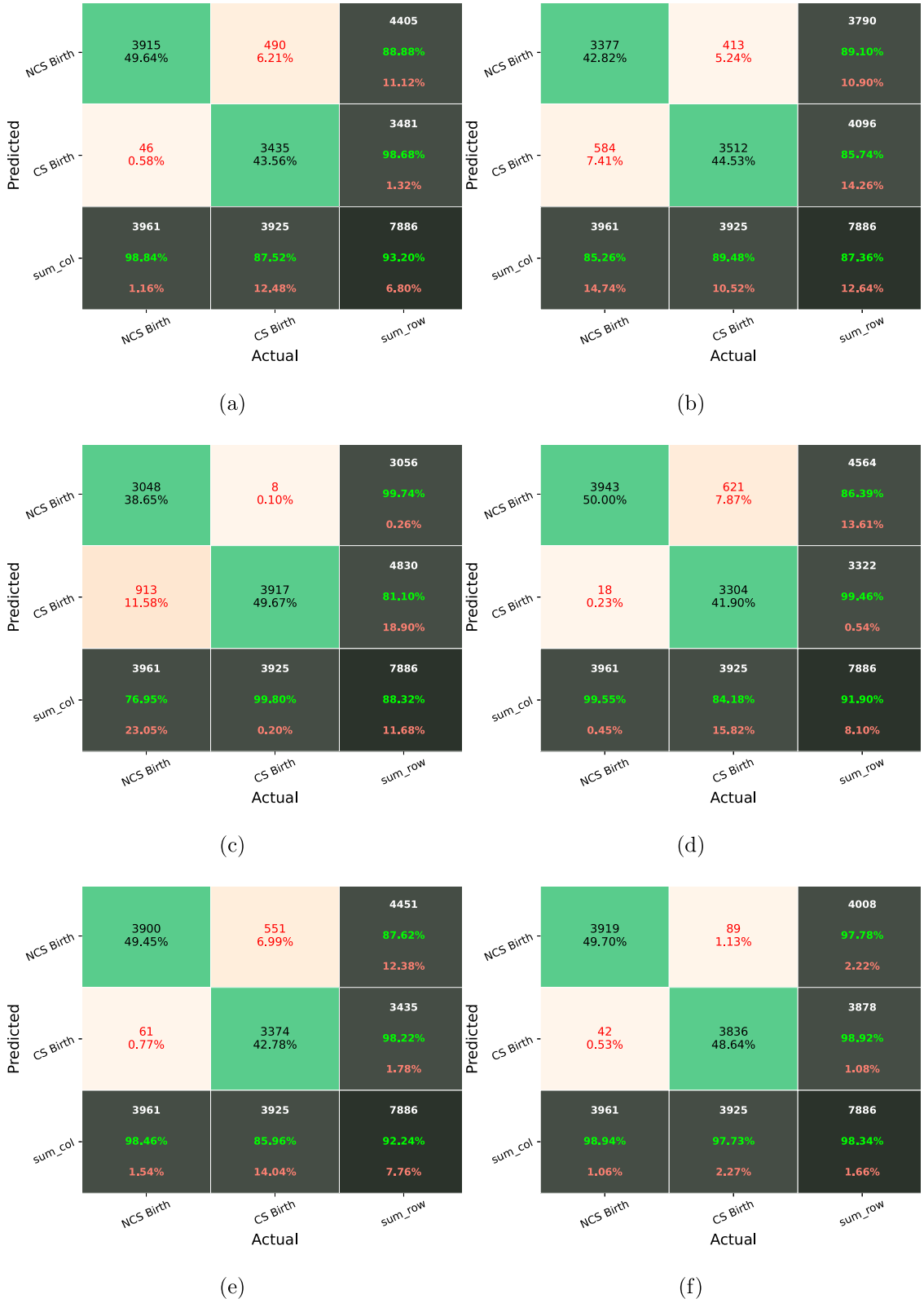

Fig. 5. Confusion matrix for (a) GBC, (b) GNB, (c) KNN, (d) LDA, (e) LR, and (f) HGSORF.

result is that the proposed HGSORF provided us with the utmost accuracy (around 99%). The one-way ANOVA was also tested, and it was found that HGSORF provides statistically more significant ($p < 0.05$) results than other classifiers [46].

3.2. XAI based cause analysis using SHAP and LIME and validation

In analyzing the significance of the features, the most dominating features are categorized in a plummeting mode, which implies that the most important features occupy the top position, and the less

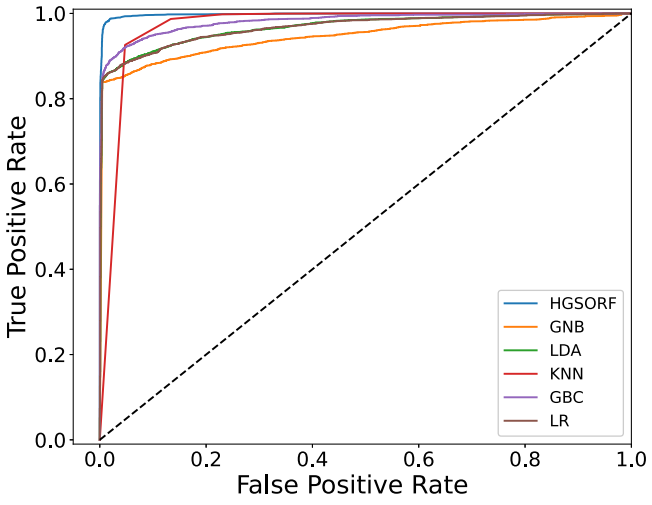

Fig. 6. ROC analysis.

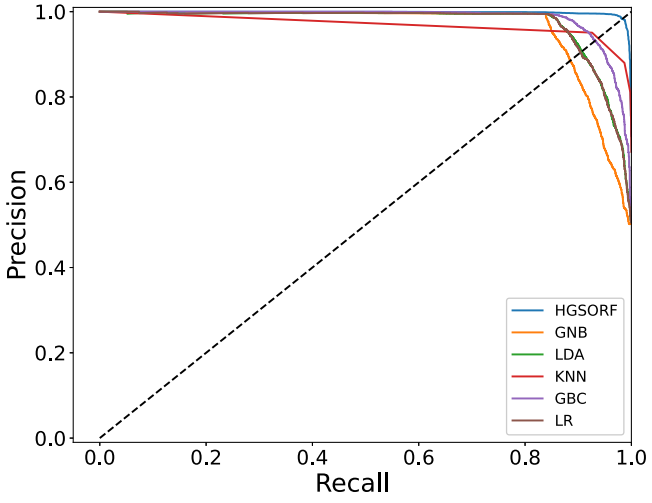

Fig. 7. Precision vs. Recall curve.

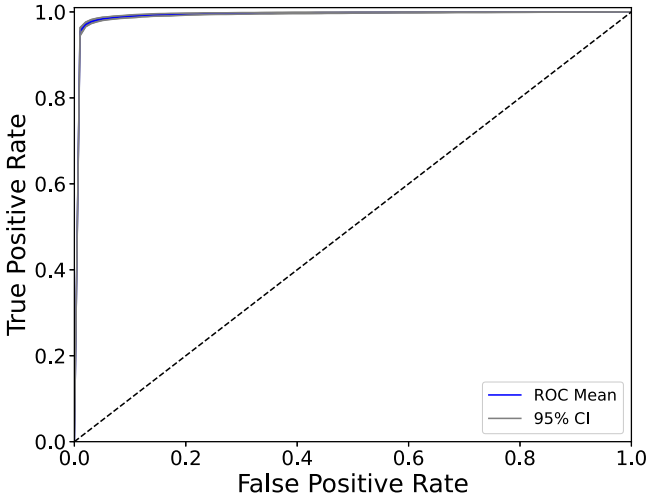

Fig. 8. Bootstrap ROC.

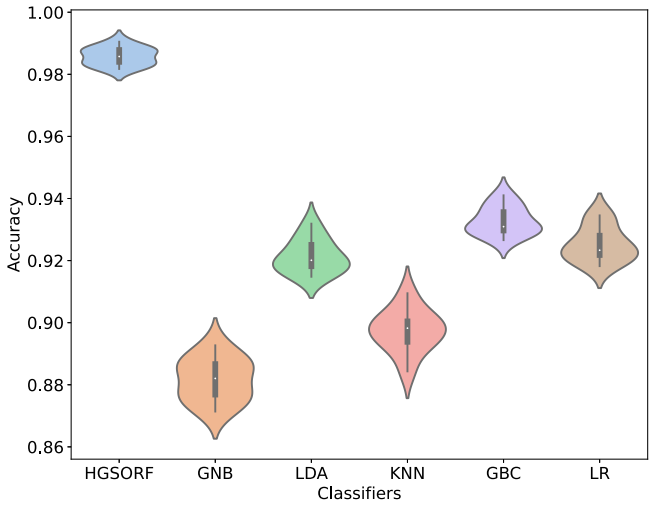

Fig. 9. Violin plot representation of comparative CV accuracy for different classifiers.

essential features take a nosedive. Fig. 10 illustrates the stacked bars for binary outputs visualizing the mean absolute SHAP value of the twenty dominating features concerning the CS and NCS classes, where the CS class is denoted using "1", and NCS class is using "0". Among the twenty visualized features, a couple of features have been marked as the most important, such as (i) *Had_Previous_C-Section*, (ii) *Suffered_by_Domestic_Violence*, (ii) *Husbands_Occupations*, (iv) *Mothers_BMI*, and (v) *Number_of_household_members*, where *Had_Previous_C-Section* occupies the incredible priority. In contrast, *Reading_Newspaper_or_Magazine* remains underneath compared to the rest of the features displayed in the given diagram.

Sometimes, the global interpretation, such as the SHAP analysis mentioned above, is insufficient to interpret the importance of local features because globally important features may not be locally significant. To investigate this issue, we need to build numerous surrogate models to interpret the local features (to explain individual predictions of the black-box model), which is a convenient tool to get the model's robustness. In addition, the LIME tweaks the feature values to modify a single data sample, and observes the resulting impact on the output (a list of explanations, which reflects how much each feature affects the ultimate prediction). As shown in Fig. 11(a), the right bar plot describes each feature's contribution to the prediction of the Yes-class, and the left bar plot in Fig. 11(b) delineates the faithfulness of the proposed approach to prognosticate the No-class. The orange color signifies the positive impact, and the blue color symbolizes the negative effect of that feature on the target. For example, *Had_Previous_C-Section* has the paramount contribution to predicting the CS and NCS classes correctly.

The features we have found from SHAP analysis are associated with CS, and now their importance is also explained. Women who had previous CS prefer CS delivery to normal delivery [56] due to many medical complications, as well as their gynecologists' suggestion to choose CS. Women who have suffered from domestic violence are also likely to use CS delivery [57] as domestic violence can cause severe physical and mental illness for a mother, which may result in pregnancy-related complications similar to the complications resulting from previous CS delivery. Husband's higher education is also a significant factor associated with CS delivery [58]. Generally, the procedure requires financial support; it is believed that higher education may ensure better-fitted jobs and secure a family's financial solvency, which might enhance the utilization of CS delivery. Women who reside in urban areas are likely to use CS delivery [58] as they are more educated, may have an excellent financial position, and may not feel comfortable enduring labor pain, which may increase their chances of undergoing CS. Mother's BMI is another significant factor in the prevalence of

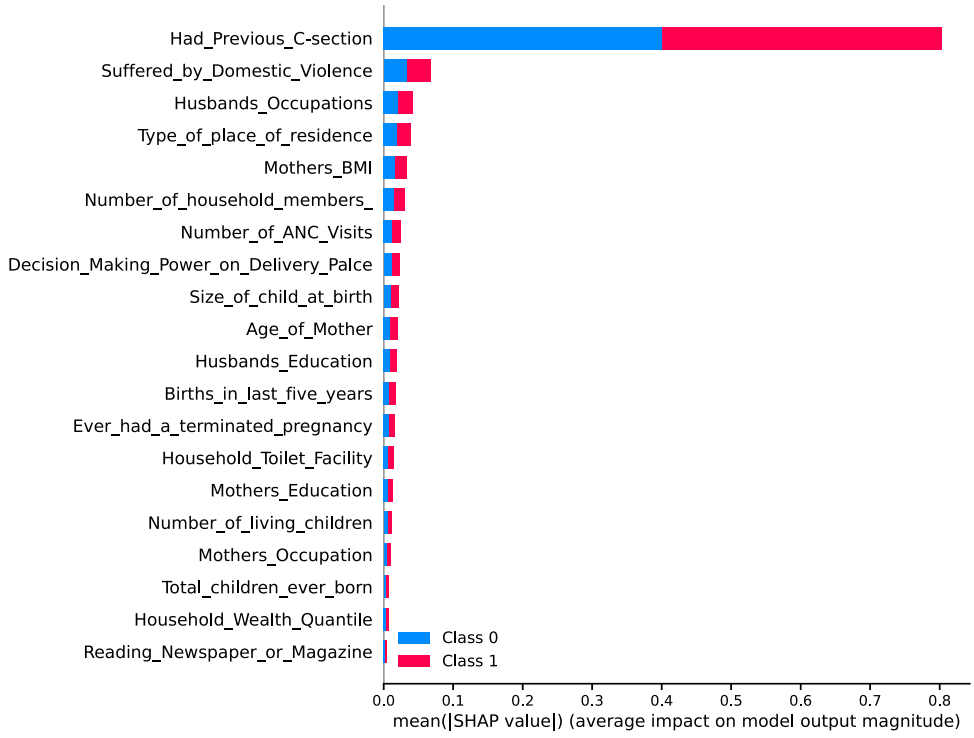

Fig. 10. Summary plot of mean absolute SHAP value on target of the model.

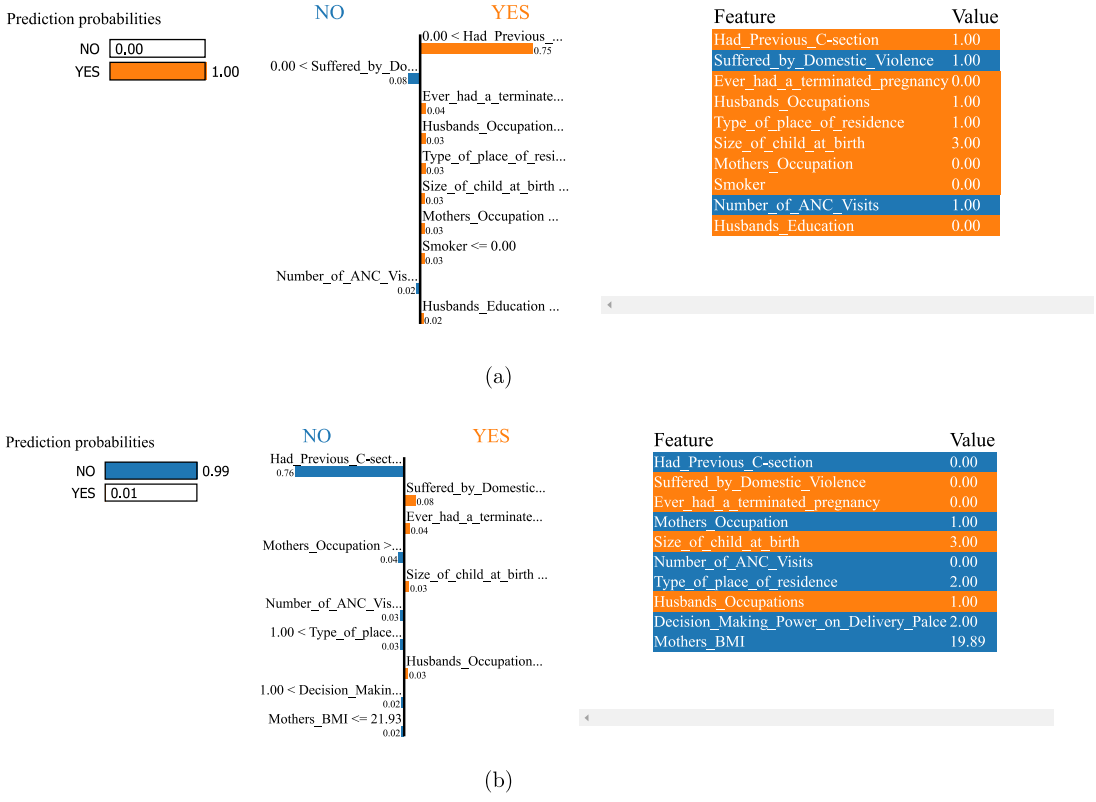

Fig. 11. LIME analysis for (a) Yes-class and (b) No-class.

CS [59]. Women suffering from malnutrition and high obesity may face complications during pregnancy and delivery period. Thus, CS is recommended by their doctors, and they also choose CS of their own volition to prevent any risk to themselves and their child(ren). The women who have more ANC visits that are from private hospitals and clinics are more likely to have CS because they are more aware of maternal service care. In addition, they have been through different medical treatments; if any complications are addressed during ANC, they are recommended to choose CS [60]. Various medical conditions such as maternal age, children's size at birth, birth order, and previous

Table 3
Comparative analysis.

References	Classification algorithms	Dataset used	No. of sample	Performance indices			
				ACC	SE	SP	Precision
Abbas et al. [19]	Supervised machine learning method	Clinical Data	488	91.8%	94.5%	–	95.5%
Khazardoost et al. [21]		Clinical Data	ultrasound measurements	100	–	90%	–
Robu and Holban [22]	Naive Bayes, J48, KNN, RF, SVM, AdaBoost, LogitBoost, JRipp, REPTree.	Historical Data	2086	80.24%	–	–	–
Sodsee [23]	Modified nearest neighbor analysis	Medical Data	802	92.29%	100%	83%	–
Sana et al. [24]	Decision Tree, ANN, J48	Clinical Data	Data (40% normal birth and 60% cesarean)	80%	71.8%	–	77.8%
Abbas et al. [62]	Bagging function (BagFda)	Historical, maternal, social, and physical data of Kashmir	488	93.44%	91.5%	–	94.5%
Proposed	**HGSORF**	**PDHS'17–18 & PDHS '12–13 data**	**15409**	**98.33%**	**98.33%**	**98.33%**	**98.34%**

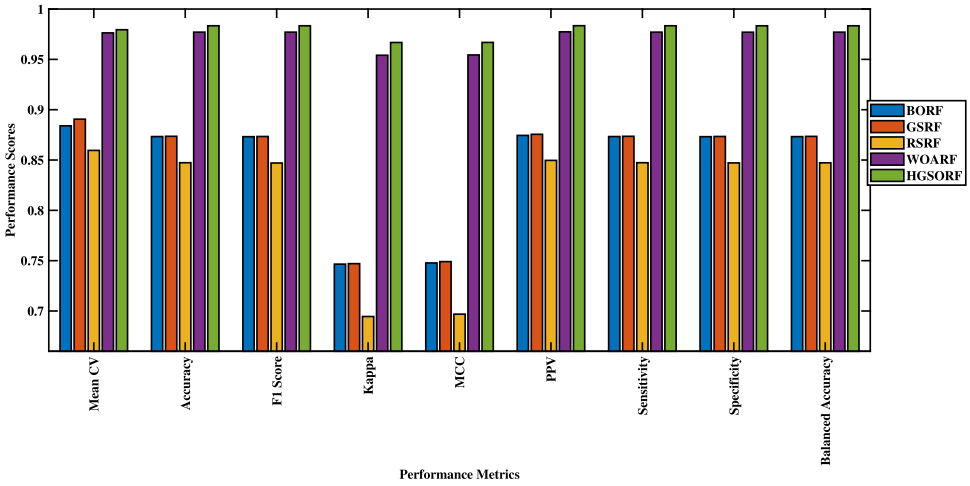

Fig. 12. Performance comparison among various hyperparameter optimization algorithms.

terminated pregnancy also influence women to opt for CS delivery [61]. Some hygienic issues also affect the prevalence of CS delivery, such as toilet facilities and drinking water. Women who use the unhygienic lavatories and drink unimproved water suffer from many diseases and parasitical infections that cause complications during pregnancy and delivery; these factors also influence women to opt for CS delivery.

3.3. Comparisons with other studies

The proposed system is compared with some of the existing works, and this comparative analysis is shown in Table 3.

From the comparison with other works, it is observed that the ACC of our proposed model is greater than that of all other studies. Using only 25 features comprised of socio-economic condition, maternal characteristics, health care, and features related to mother and family information, our model achieved an ACC of 98.33%. The other performance metrics, such as SE, SP, and precision, are also higher than other methods. Note that the nature of the datasets used in Table 3 are closely related to the dataset used in this study and, therefore, presented for a better comparison. Hence, the above discussion suggests that our developed HGSORF model is able to superiorly classify CS and NCS births, providing the best results with the least number of features.

3.4. Comparison with other optimization algorithms

The performance of the proposed HGSORF model relies on the hyperparameter optimization using an HGSO algorithm that balances exploration and exploitation behavior in the search space and tries to avoid local optima to find global optima. We can also logically justify its use by comparing our proposed HGSORF with common optimization techniques [46,50,63], such as Bayesian optimization-based RF (BORF), grid search-based RF (GSRF), random search-based RF (RSRF), and whale optimization algorithm-based RF (WOARF) using some statistical measures. This performance comparison is shown in Fig. 12, from which it can be seen that the mean cross-validation, accuracy, F1 score, kappa score, MCC score, positive predicted value, sensitivity, specificity, and balanced accuracy are superior for the proposed HGSORF than other optimization techniques using RF. In addition, we have compared HGSORF with BORF, GSRF, RSRF, and WOARF and calculated the run time for each case for the purpose of demonstrating the run-time complexity. For a fair justification, all optimization algorithms were run on a Core i9 processor with 64 GB RAM and Windows 10 operating system. It is evidenced that BORF, GSRF, RSRF, WOARF, and HGSORF required 136.47269 s, 518.4401 s, 155.83131 s, 4673.717163 s, 3183.8153 s, respectively. Although HGSORF required more time to complete the optimization process than BORF, GSRF, and RSRF, efficient implementation using parallel computing could decrease computational time, which is one of our intended future works.

3.5. Summary

A novel HGSO algorithm has been adapted together with a state-of-the-art objective function and then applied in the proposed research

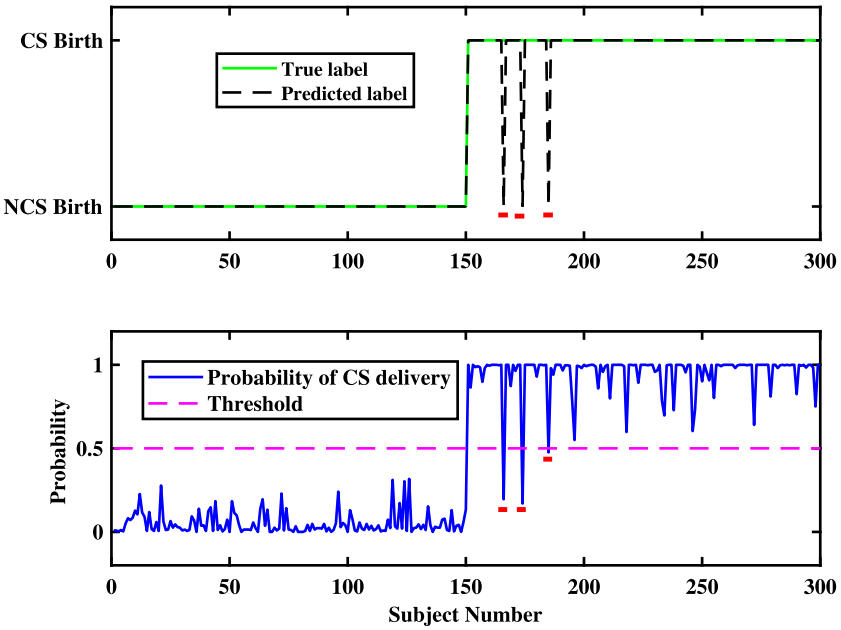

Fig. 13. Development of a DSS using HGSORF.

to tune the hyperparameters of the RF classifier. Some other classifiers with optimized hyperparameters have also been deployed to segregate the CS and NCS classes from the CS dataset, containing 15409 samples. It is obvious that the performance of the proposed HGSORF was superior in terms of all classification metrics considered; for instance, the accuracy was approximately 98.33%, with an AUC of around 99%(as presented in Figs. 4 and 5[f]). The most noteworthy point is that the CV accuracy of the proposed HGSORF also touched the uppermost point, which has been visualized in Fig. 9. Overall, the HGSO-optimized and ADASYN-balanced RF classification algorithm generalized well and performed best on the CS dataset. Therefore, it can be confidently used for further clinical trials and validation. The overall summary is as follows:

- The CS data are real-world and noisy data taken from PDHS. So, RF is an excellent choice to avoid the overfitting problems in a noisy environment compared to other classifiers (GNB, LDA, KNN, LR, and even state-of-the-art GBC). In addition, our PDHS dataset comprises categorical and continuous data, and for such a dataset, RF provides better results than GBC, which ultimately improves the performance. Finally, a globally optimized RF model provides a robust model and, consequently, improves results when applied to testing data.
- The HGSO is able to avoid local minima and, consequently, provide the global optimal point. Therefore, the hyperparameters of RF are highly tuned by using HGSO optimization to produce a globally optimal HGSORF model. As a result, the HGSORF model provided improved performances such as accuracy, sensitivity, specificity, cross-validation, ROC, etc.
- XAI tools such as SHAP and LIME have been utilized to explain the global and local behavior of the HGSORF model, respectively. Fig. 10 calculates the overall feature importance (global behavior) where we can see the importance of different causes that are ultimately responsible for CS delivery. These factors are then justified by various contemporary literature. In addition, two outcomes (one CS Birth and one NCS Birth) have been explained by LIME analysis in Fig. 11.
- The performance of the proposed model has been compared with other studies in Table 3. In addition, our proposed HGSORF has also been compared with other optimization techniques such as BORF, GSRF, RSRF, and WOARF in Fig. 12. It is obvious that the proposed HGSORF provided the best result among these optimization techniques.

4. Potential application of the HGSORF model

A potential application has been demonstrated by developing a decision support system (DSS), as shown in Fig. 13. The DSS is a graphical representation of the probable state of CS delivery and presents the probability of a pregnant mother either going for CS delivery or not.

If the probability calculated by HGSORF is greater than 0.5, it represents that this mother would have a higher chance of CS delivery. This graphical representation will benefit non-specialist users such as clinical staffs and nurses. A possible outcome of CS delivery is presented in Fig. 13 in terms of posterior probability calculated from HGSORF mode. It could be noted that 300 subjects (150 NCS birth and 150 CS birth) have been utilized for demonstration purposes in Fig. 13. It can be seen in the upper figure that our predicted label is a highly accurate to true label except for subjects 165, 167, and 175. Additionally, the corresponding probability of CS delivery is shown in the lower figure. The test data are sorted in ascending order so that the data for NCS birth can appear first, followed by the data for CS birth.

5. Conclusion and future direction

This research proposes an optimized HGSORF classifier and applies it to the ADASYN balanced CS dataset (PDHS'17–18 and '12–13 data) to predict the target. The performance of the proposed model has been ascertained using various classification metrics, such as ACC, SE, kappa index, SP, precision, recall, etc. We obtained 98.33% accuracy with the proposed method, which outperforms other classifiers as well as other studies. In addition, HGSORF provides improved results compared to GSRF, BORF, RSRF, and WOARF in terms of classification metrics. Therefore, it can be concluded that an optimized and ADASYN-balanced HGSORF classifier can reliably predict the CS and NCS classes. One significant implication of this study is that it not only predicts CS delivery using the HGSORF model but also uses XAI-based tools such as SHAP and LIME to interpret the model and extract the major causes of CS delivery. Thus, it is hoped that our proposed method and XAI provide an early prediction procedure to help the government and national policy-makers enforce new policies to stop unnecessary CS delivery. Further development and testing are also required before clinical

trials. One of the limitations of the proposed HGSORF is that it requires a large computing time during the training and optimization process due to the algorithm's evolutionary nature, which can be minimized by an efficient implementation using parallel computing. In the future, we plan to implement HGSORF efficiently, implement an ensemble ML model, and propose a new feature selection algorithm. Furthermore, our proposed approach can be generalized and adapted for classifying hypertension, Alzheimer's disease, asthma patients, diabetes mellitus, and more clinical and socio-demographic data.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data and code availability

Our preprocessed data and pre-trained models are available on Github: https://github.com/MIrazul29/HGSORF_CSection.

Acknowledgments

The authors extend their appreciation to the Deanship of Scientific Research at King Saud University for funding this work through research group no [RG-1441-394].

References

[1] A. Betran, M.R. Torloni, J. Zhang, A. Gülmezoglu, W.W.G. on Caesarean Section, H. Aleem, F. Althabe, T. Bergholt, L. de Bernis, G. Carroli, et al., WHO statement on caesarean section rates, BJOG: Int. J. Obstet. Gynaecol. 123 (5) (2016) 667–670.

[2] T. Morris, Cut It Out: The C-Section Epidemic in America, NYU Press; Reprint edition (November 1, 2016), 2016.

[3] L. Gibbons, J.M. Belizán, J.A. Lauer, A.P. Betrán, M. Merialdi, F. Althabe, et al., The global numbers and costs of additionally needed and unnecessary caesarean sections performed per year: overuse as a barrier to universal coverage, World Health Rep. 30 (1) (2010) 1–31.

[4] A. Dumont, L. De Bernis, M.-H. Bouvier-olle, G. Bréart, M.S. Group, et al., Caesarean section rate for maternal indication in sub-Saharan Africa: a systematic review, Lancet 358 (9290) (2001) 1328–1333.

[5] J.P. Vogel, A.P. Betrán, N. Vindevoghel, J.P. Souza, M.R. Torloni, J. Zhang, O. Tunçalp, R. Mori, N. Morisaki, E. Ortiz-Panozo, et al., Use of the robson classification to assess caesarean section trends in 21 countries: a secondary analysis of two WHO multicountry surveys, Lancet Glob. Health 3 (5) (2015) e260–e270.

[6] X.L. Feng, L. Xu, Y. Guo, C. Ronsmans, Factors influencing rising caesarean section rates in China between 1988 and 2008, Bull. World Health Organ. 90 (2012) 30–39A.

[7] F. Althabe, C. Sosa, J.M. Belizán, L. Gibbons, F. Jacquerioz, E. Bergel, Cesarean section rates and maternal and neonatal mortality in low-, medium-, and high-income countries: an ecological study, Birth 33 (4) (2006) 270–277.

[8] A.P. Betrán, M. Merialdi, J.A. Lauer, W. Bing-Shun, J. Thomas, P. Van Look, M. Wagner, Rates of caesarean section: analysis of global, regional and national estimates, Paediatr. Perinat. Epidemiol. 21 (2) (2007) 98–113.

[9] M.R. Festin, M. Laopaiboon, P. Pattanittum, M.R. Ewens, D.J. Henderson-Smart, C.A. Crowther, Caesarean section in four south east Asian countries: reasons for, rates, associated care practices and health outcomes, BMC Pregnancy Childbirth 9 (1) (2009) 1–11.

[10] W.H. Organization, et al., WHO statement on caesarean section rates, Tech. Rep., World Health Organization, 2015.

[11] V. Verma, R.K. Vishwakarma, D.C. Nath, H.T. Khan, R. Prakash, O. Abid, Prevalence and determinants of caesarean section in south and south-east Asian women, Plos One 15 (3) (2020) e0229906.

[12] S. Mumtaz, J. Bahk, Y.-H. Khang, Rising trends and inequalities in cesarean section rates in Pakistan: Evidence from Pakistan demographic and health surveys, 1990–2013, PLoS One 12 (10) (2017) e0186563.

[13] T.D. Program, Pakistan 2019 Maternal Mortality Survey Summary Report, National Institute of Population Studies (NIPS), Ministry of National Health Services, Regulations and Coordination, Islamabad, Pakistan, 2019, https://dhsprogram.com/publications/publication-SR267-Summary-Reports-Key-Findings.cfm.

[14] L. Alkema, D. Chou, D. Hogan, S. Zhang, A.-B. Moller, A. Gemmill, D.M. Fat, T. Boerma, M. Temmerman, C. Mathers, et al., Global, regional, and national levels and trends in maternal mortality between 1990 and 2015, with scenario-based projections to 2030: a systematic analysis by the UN maternal mortality estimation inter-agency group, Lancet 387 (10017) (2016) 462–474.

[15] Q. Long, T. Kempas, T. Madede, R. Klemetti, E. Hemminki, Caesarean section rates in Mozambique, BMC Pregnancy Childbirth 15 (1) (2015) 1–9.

[16] M. Khawaja, T. Kabakian-Khasholian, R. Jurdi, Determinants of caesarean section in Egypt: evidence from the demographic and health survey, Health Policy 69 (3) (2004) 273–281.

[17] N. Khawaja, T. Yousaf, R. Tayyeb, Analysis of caesarean delivery at a tertiary care hospital in Pakistan, J. Obstet. Gynaecol. 24 (2) (2004) 139–141.

[18] U.S. Mishra, M. Ramanathan, Delivery-related complications and determinants of caesarean section rates in India, Health Policy Plan. 17 (1) (2002) 90–98.

[19] S.A. Abbas, R. Riaz, S.Z.H. Kazmi, S.S. Rizvi, S.J. Kwon, Cause analysis of caesarian sections and application of machine learning methods for classification of birth data, IEEE Access 6 (2018) 67555–67561.

[20] M. Dulitzki, D. Soriano, E. Schiff, A. Chetrit, S. Mashiach, D.S. Seidman, Effect of very advanced maternal age on pregnancy outcome and rate of cesarean delivery, Obstet. Gynecol. 92 (6) (1998) 935–939.

[21] S. Khazardoost, F.G. Vahdani, S. Latifi, S. Borna, M. Tahani, M.A. Rezaei, M. Shafaat, Pre-induction translabial ultrasound measurements in predicting mode of delivery compared to bishop score: a cross-sectional study, BMC Pregnancy Childbirth 16 (1) (2016) 1–7.

[22] R. Robu, S. Holban, The analysis and classification of birth data, Acta Polytech. Hungarica 12 (4) (2015) 77–96.

[23] S. Sodsee, Predicting caesarean section by applying nearest neighbor analysis, Procedia Comput. Sci. 31 (2014) 5–14.

[24] A. Sana, S. Razzaq, J. Ferzund, Automated diagnosis and cause analysis of cesarean section using machine learning techniques, Int. J. Mach. Learn. Comput. 2 (5) (2012) 677.

[25] F. Hasan, M.M. Alam, M.G. Hossain, Associated factors and their individual contributions to caesarean delivery among married women in Bangladesh: analysis of Bangladesh demographic and health survey data, BMC Pregnancy Childbirth 19 (1) (2019) 1–9.

[26] M.A. Rahman, M.A. Rahman, L.B. Rawal, M. Paudel, M.H. Howlader, B. Khan, T. Siddiquee, A. Rahman, A. Sarkar, M.S. Rahman, et al., Factors influencing place of delivery: Evidence from three south-Asian countries, Plos One 16 (4) (2021) e0250012.

[27] F.A. Hashim, E.H. Houssein, M.S. Mabrouk, W. Al-Atabany, S. Mirjalili, Henry gas solubility optimization: A novel physics-based algorithm, Future Gener. Comput. Syst. 101 (2019) 646–667.

[28] D.R. Bishanga, M. Drake, Y.-M. Kim, A.H. Mwanamsangu, A.M. Makuwani, J. Zoungrana, R. Lemwayi, M.J. Rijken, J. Stekelenburg, Factors associated with institutional delivery: Findings from a cross-sectional study in Mara and Kagera regions in Tanzania, PLoS One 13 (12) (2018) e0209672.

[29] F. Hasan, M. Sabiruzzaman, C.K. Joardar, M. Hossain, et al., Maternal socio-demographic factors and nutritional status as predictors of Caesarean delivery at hospitals in Rajshahi city, Bangladesh, Malays. J. Nutr. 21 (3) (2015).

[30] S.M. Kamal, Preference for institutional delivery and caesarean sections in Bangladesh, J. Health Popul. Nutri. 31 (1) (2013) 96.

[31] F. Karim, N.B. Ali, A.N.S. Khan, A. Hassan, M.M. Hasan, D.M.E. Hoque, S.M. Billah, S. El Arifeen, M.A.K. Chowdhury, Prevalence and factors associated with caesarean section in four hard-to-reach areas of Bangladesh: Findings from a cross-sectional survey, PLoS One 15 (6) (2020) e0234249.

[32] M.N. Khan, M.M. Islam, A.A. Shariff, M.M. Alam, M.M. Rahman, Socio-demographic predictors and average annual rates of caesarean section in Bangladesh between 2004 and 2014, PLoS One 12 (5) (2017) e0177579.

[33] A. Amjad, U. Amjad, R. Zakar, A. Usman, M.Z. Zakar, F. Fischer, Factors associated with caesarean deliveries among child-bearing women in Pakistan: secondary analysis of data from the demographic and health survey, 2012–13, BMC Pregnancy Childbirth 18 (1) (2018) 1–9.

[34] A. Amjad, A. Imran, N. Shahram, R. Zakar, A. Usman, M.Z. Zakar, F. Fischer, Trends of caesarean section deliveries in Pakistan: secondary data analysis from demographic and health surveys, 1990–2018, BMC Pregnancy Childbirth 20 (1) (2020) 1–13.

[35] A.K. Bhandari, B. Dhungel, M. Rahman, Trends and correlates of cesarean section rates over two decades in Nepal, BMC Pregnancy Childbirth 20 (1) (2020) 1–13.

[36] B. Devkota, J. Maskey, A.R. Pandey, D. Karki, P. Godwin, P. Gartoulla, S. Mehata, K.K. Aryal, Determinants of home delivery in Nepal–A disaggregated analysis of marginalised and non-marginalised women from the 2016 Nepal demographic and health survey, Plos One 15 (1) (2020) e0228440.

[37] R. Shah, E.A. Rehfuess, M.K. Maskey, R. Fischer, P.B. Bhandari, M. Delius, Factors affecting institutional delivery in rural chitwan district of Nepal: a community-based cross-sectional study, BMC Pregnancy Childbirth 15 (1) (2015) 1–14.

[38] A. Shahabuddin, V. De Brouwere, R. Adhikari, A. Delamou, A. Bardaj, T. Delvaux, Determinants of institutional delivery among young married women in Nepal: Evidence from the Nepal demographic and health survey, 2011, BMJ Open 7 (4) (2017) e012446.

[39] S.K. Shrestha, B. Banu, K. Khanom, L. Ali, N. Thapa, B. Stray-Pedersen, B. Devkota, Changing trends on the place of delivery: why do Nepali women give birth at home? Reprod. Health 9 (1) (2012) 1–8.
[40] T. Gondwe, K. Betha, G. Kusneniwar, C.H. Bunker, G. Tang, H. Simhan, P. Reddy, C.L. Haggerty, Maternal factors associated with mode of delivery in a population with a high cesarean section rate, J. Epidemiol. Global Health 9 (4) (2019) 252.
[41] S.S. Padmadas, S.B. Nair, A.K. KR, et al., Caesarean section delivery in Kerala, India: evidence from a national family health survey, Soc. Sci. Med. 51 (4) (2000) 511–521.
[42] A.K. Manyeh, A. Amu, D.E. Akpakli, J. Williams, M. Gyapong, Socioeconomic and demographic factors associated with caesarean section delivery in southern ghana: evidence from INDEPTH network member site, BMC Pregnancy Childbirth 18 (1) (2018) 1–9.
[43] L. Miri Farahani, M.J. Abbasi Shavazi, Caesarean section change trends in Iran and some demographic factors associated with them in the past three decades, J. Fasa Univ. Med. Sci. 2 (3) (2012) 127–134.
[44] M. Ochieng Arunda, A. Agardh, B.O. Asamoah, Cesarean delivery and associated socioeconomic factors and neonatal survival outcome in Kenya and Tanzania: analysis of national survey data, Glob. Health Action 13 (1) (2020) 1748403.
[45] E. Yisma, L.G. Smithers, J.W. Lynch, B.W. Mol, Cesarean section in Ethiopia: prevalence and sociodemographic characteristics, J. Matern.-Fetal Neonatal Med. 32 (7) (2019) 1130–1135.
[46] M.A. Awal, M.S. Hossain, K. Debjit, N. Ahmed, R.D. Nath, G.M.M. Habib, M.S. Khan, M.A. Islam, M.A.P. Mahmud, An early detection of asthma using BOMLA detector, IEEE Access 9 (2021) 58403–58420.
[47] H. He, Y. Bai, E.A. Garcia, S. Li, ADASYN: Adaptive synthetic sampling approach for imbalanced learning, in: 2008 IEEE International Joint Conference on Neural Networks (IEEE World Congress on Computational Intelligence), IEEE, 2008, pp. 1322–1328.
[48] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32.
[49] K. Debjit, M.S. Islam, M.A. Rahman, F.T. Pinki, R.D. Nath, S. Al-Ahmadi, M.S. Hossain, K.M. Mumenin, M.A. Awal, An improved machine-learning approach for COVID-19 prediction using Harris Hawks optimization and feature analysis using SHAP, Diagnostics 12 (5) (2022) 1023.
[50] M.A. Awal, M. Masud, M.S. Hossain, A.A.-M. Bulbul, S.H. Mahmud, A.K. Bairagi, A novel bayesian optimization-based machine learning framework for COVID-19 detection from inpatient facility data, IEEE Access 9 (2021) 10263–10281.
[51] S. Mirjalili, Genetic algorithm, in: Evolutionary Algorithms and Neural Networks, Springer, 2019, pp. 43–55.
[52] S. Mirjalili, A. Lewis, The whale optimization algorithm, Adv. Eng. Softw. 95 (2016) 51–67.
[53] M. Pelikan, D.E. Goldberg, E. Cantú-Paz, et al., BOA: The Bayesian optimization algorithm, in: Proceedings of the Genetic and Evolutionary Computation Conference GECCO-99, Vol. 1, Citeseer, 1999, pp. 525–532.
[54] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, pp. 4768–4777.
[55] M.T. Ribeiro, S. Singh, C. Guestrin, "Why should i trust you?" explaining the predictions of any classifier, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, pp. 1135–1144.
[56] T. Begum, A. Rahman, H. Nababan, D.M.E. Hoque, A.F. Khan, T. Ali, I. Anwar, Indications and determinants of caesarean section delivery: evidence from a population-based study in matlab, Bangladesh, PLoS One 12 (11) (2017) e0188074.
[57] N.A. Al Shidhani, A.A. Al Kendi, M.H. Al Kiyumi, Prevalence, risk factors and effects of domestic violence before and during pregnancy on birth outcomes: an observational study of literate omani women, Int. J. Women's Health 12 (2020) 911.
[58] M.M. Rahman, M.R. Haider, M. Moinuddin, A.E. Rahman, S. Ahmed, M.M. Khan, Determinants of caesarean section in Bangladesh: Cross-sectional analysis of Bangladesh demographic and health survey 2014 data, PLoS One 13 (9) (2018) e0202879.
[59] W. Al-Kubaisy, M. Al-Rubaey, R.A. Al-Naggar, B. Karim, N.A.M. Noor, Maternal obesity and its relation with the cesarean section: A hospital based cross sectional study in Iraq, BMC Pregnancy Childbirth 14 (1) (2014) 1–5.
[60] M.A. Rahman, M.S. Rahman, M. Aziz Rahman, E.A. Szymlek-Gay, R. Uddin, S.M.S. Islam, Prevalence of and factors associated with anaemia in women of reproductive age in Bangladesh, Maldives and Nepal: Evidence from nationally-representative survey data, Plos One 16 (1) (2021) e0245335.
[61] R.M. Häger, A.K. Daltveit, D. Hofoss, S.T. Nilsen, T. Kolaas, P. Øian, T. Henriksen, Complications of cesarean deliveries: rates and risk factors, Am. J. Obstet. Gynecol. 190 (2) (2004) 428–434.
[62] S.A. Abbas, A.U. Rehman, F. Majeed, A. Majid, M.S.A. Malik, Z.H. Kazmi, S. Zafar, Performance analysis of classification algorithms on birth dataset, IEEE Access 8 (2020) 102146–102154.
[63] M.K. Hasan, M.T. Jawad, A. Dutta, M.A. Awal, M.A. Islam, M. Masud, J.F. Al-Amri, Associating measles vaccine uptake classification and its underlying factors using an ensemble of machine learning models, IEEE Access 9 (2021) 119613–119628.