

RESEARCH

Open Access



A novel oppositional binary crow search algorithm with optimal machine learning based postpartum hemorrhage prediction model

Sujatha Krishnamoorthy^{1*}, Yihang Liu¹ and Kun Liu²

Abstract

Postpartum hemorrhage (PPH) is an obstetric emergency instigated by excessive blood loss which occurs frequently after the delivery. The PPH can result in volume depletion, hypovolemic shock, and anemia. This is particular condition is considered a major cause of maternal deaths around the globe. Presently, physicians utilize visual examination for calculating blood and fluid loss during delivery. Since the classical methods depend on expert knowledge and are inaccurate, automated machine learning based PPH diagnosis models are essential. In regard to this aspect, this study introduces an efficient oppositional binary crow search algorithm (OBCSA) with an optimal stacked auto encoder (OSAE) model, called OBCSA-OSAE for PPH prediction. The goal of the proposed OBCSA-OSAE technique is to detect and classify the presence or absence of PPH. The OBCSA-OSAE technique involves the design of OBCSA based feature selection (FS) methods to elect an optimum feature subset. Additionally, the OSAE based classification model is developed to include an effective parameter adjustment process utilizing Equilibrium Optimizer (EO). The performance validation of the OBCSA-OSAE technique is performed using the benchmark dataset. The experimental values pointed out the benefits of the OBCSA-OSAE approach in recent methods.

Keywords: Postpartum hemorrhage, Predictive model, Machine learning, Metaheuristics, Feature Selection, Classification

Introduction

Postpartum hemorrhage (PPH) is determined as blood loss more than >500 mL within 24 hours after vaginal birth, and is a major reason for pregnancy morbidity, both in the US and world-wide [1]. The severity and incidence of PPH and the morbidity connected to blood product transfusion are increased [2]. The higher occurrences of PPH in the developing world is because of the lack of diagnosis approaches and medication utilized in the active management of the third stage. The implemen-

tation and development of quality initiatives have made aggressive supervision and early recognition of PPH with a reduction in disseminated intravascular coagulopathy, blood product transfusion, renal failure, and respiratory distress syndrome [3]. Most priorities in the PPH security bundles are identification of patients at increased risk of hemorrhage based on intrapartum and antenatal risk and the assessment of hemorrhage risk factors [4].

Existing approaches for forecasting postpartum hemorrhage depend on a risk stratification method. Enhanced prediction capacity is attained using conventional machine learning (ML) and statistical approaches [5]. Most recent developments in ML method applies modern computer-driven algorithms intended to identify pattern

*Correspondence: krishnsu@kean.edu; Sujathaks@ieee.org

¹Department of Computer Science, Wenzhou-Kean University, WENZHOU KEAN UNIVERSITY, 88 UNIVE, Wenzhou 325006, Zhejiang, China
Full list of author information is available at the end of the article



© The Author(s). 2022 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

from data have received more interest due to their better prediction capability. Mainly though defining hospital readmission and intensive care unit admission than statistical methods. While these advanced methodologies have not been broadly tested in the obstetric field. The risk related factors attributed to PPH have been widely studied on the basis of traditional statistical models [6]. Generally, the Lasso regression model, or Logistic regression model has been utilized for predicting the risk of PPH.

Particularly, logistic or lasso regression model shows better discrimination capability [7]. According to maternal medical history and clinical characteristics, a risk score is utilized for PPH prediction. Even though the traditional predictive models are shown to be efficient, the predictive accuracy isn't satisfied: only 60% of women having higher PPH risks are recognized, while another 40% of women who had PPH weren't recognized at an earlier stage [8]. The benefits of ML method involve the capacity of processing non-additive relationships and integrating complex relations among other aspects which don't require pre-determination [9]. For this reason, it is possible that ML methods are capable of precisely identifying women at the maximum risk of postpartum hemorrhage, and enhancing medical results and obstetric decision making [10].

This study develops an oppositional binary crow search algorithm (OBCSA) with an optimal stacked auto encoder (OSAE) model known as OBCSA-OSAE for PPH prediction. The OBCSA-OSAE technique performs the detection and classification process of PPH. Moreover, the OBCSA-OSAE technique involves the design of OBCSA based feature selection (FS) technique for optimum feature subsets. Additionally, the OSAE based classification model is developed to include an effective parameter adjustment process using Equilibrium Optimizer (EO). The experimental outcomes highlighted that the OBCSA-OSAE technique has depicted the other techniques in terms of different evaluation parameters. The method helps to classify the PPH with classification and with good feature selection techniques. Hence, the CSA helps in the optimization algorithm. This classification result will help the doctor to give the patient an early alert and also the counseling session about them to get them prepared mentally and physically if that just happened. The whole paper is structured with an introduction and the related work in the beginning and followed by the proposed OBCSA opposing binary crow search Algorithm, process, and validation step of the model. Results are compared with the existing model, and finally, the discussion and conclusion are presented at the end of the paper with the focused results.

Related works

Venkatesh et al. [11] employed LR method with and without lasso regularization (lasso regression) as the 2 statistical approaches, and XGBoost and RF as the 2 ML methods for predicting postpartum hemorrhage. Model accuracy was evaluated using calibration, decision curves, and C statistics (viz., concordance index). Kumar et al. [12] developed an automation method with wearable devices for preventing PPH in pregnant women. The e device evaluate parameters such as perspiration rate, temperature, pulse rate, and blood pressure. Fuzzy neural method-based rules are utilized for each parameter in predicting the risks of PPH, and for measuring the accuracy of method to reduce morbidity and mortality rates. Wu et al. [13] aim to construct a nomogram integrating clinical and radiomic features of a placenta to forecast the risks of PPH occurring in a caesarian delivery (CD). Radiomic features are selected according to their correlation with EBL. clinico-radiomic, Radiomic, radiological, clinical, and clinico-radiological methods are constructed for predicting the risks of PPH for all patients. The method with a better predictive accuracy was authenticated with its clinical application, discrimination ability, and calibration curve. Betts et al. [14] intended to forecast the risks of general maternal postpartum complication that requires inpatient care. A gradient boosting tree is utilized by 5-fold cross-validation for comparing method accuracy. The better performing methods for all the outcomes are measured later in the independent data validation with the AUC-ROC method.

Man [15] intended to apply the ML classifier methods for better PPH risk prediction. Real datasets were integrated and extracted from EHRS using twelve parameters that are considered to be very appropriate to PPH. The 6 ML methods involving LR, DT, RF, KNN, SVM, and ANN have been compared and tested based on their prediction performance and another matrix such as recall and precision. The RF method is believed to be the optimal method with 89% accuracy. Kumar et al. [16] focused on the detection and symptoms of postpartum hypothermia and hemorrhage. These wearable devices provide a life-saving product which is comfortable, affordable, and easy to use. Also, they help to alleviate other healthcare problems faced by women during and after child birth. At first, a lower cost prototype method was constructed that has sensor nodes that measure and record blood pressure, body temperature, perspiration, heart, and pulse rate. The overall purpose of these devices serve to provide efficient and accurate results to assist in the decline of morbidity rate and maternal issues.

Hochman et al. [17] validated and developed an ML based PPD predictive method using EHR data, and recognized new PPD predictor. PPD was determined by the novel diagnoses of any antidepressant prescription/depressive episode within the first postpartum year. A gradient boosting DT method was employed for clinical, obstetric features and EHR derived socio-demographic. The ML based methods that integrate HER derived predictors can increase symptoms-based screening practices by recognizing the higher risk population as a fundamental necessity for preventive interference before the onset of PPD. Yang et al. [18] adapted a univariate LR method for selecting the important features ($P < 0.01$). Then, they trained numerous ANN and binary LR methods for predicting postpartum hemorrhaging the NN method includes RBF, MLP, and BP. To identify and compare the precise networks, they utilized the ROC curve and the confusion matrix.

Zheutlin et al. [19] proposed a new risk assessment method and relate its performance to those employed in present practice. With a large set of potential identified and risk factor removed from EMR that was accessible previous to delivery, they trained a gradient boosting method from the subset of cohorts. In held-out test samples, they related performances of this method to three medical risk tools and one formerly published method. The result suggests that this method is an outstanding candidate for prospective assessment and can eventually decrease PPH mortality and morbidity due to earlier detection alongside prevention measures.

The proposed OBSCSA-OSAE model

In this study, an effective OBSCSA-OSAE technique is established for the detection and classification of PPH. The proposed OBSCSA-OSAE technique incorporated two major stages, choice of features and classification.

Principally, the OBSCSA technique is presented for optimal selection of a subset of features. Followed by the EO algorithm to derive the optimally choose parameters involved in the SAE model. Lastly, the SAE model is used for PPH classification. Figure 1 illustrates the general process of the OBSCSA-OSAE model.

Algorithmic design of OBSCSA-FS technique

In the initial stage, the OBSCSA-FS technique is designed to choose an optimal subset of features. The OBSCSA algorithm is designed for the integration of oppositional based learning (OBL) concept alongside BCSA. The Crow Search Algorithm is one of the latest evolutionary algorithms developed by Askarzadeh [20], which was stimulated by the social habits of crow through a search procedure that mimics their behavior in the wild. The concept of CSA was inspired by the way that these animals hide food in a given place and retrieve it at a later time. Mathematically, a flock of crows is represented as n_c , and in the search space, the position of crow i at iteration t is x_i^t . In CSA, the hidden location of food can be remembered by the crow i . The procedure of upgrading the location of the crow thieves (crows who want to steal other crow food) was performed by

$$x_i^{t+1} = x_i^t + \tau \times fl \times (M_j^t - x_i^t), \quad i = 1, 2, \dots, n_c, \quad (1)$$

Whereas fl denotes the flight length and τ signifies an arbitrary value in the range of zero and one, M is a fitness function is used to evaluate each crow, and its value is put as an initial memory value for the updation.

The next state of the problem is that crow j , the owner of the food, knows that the crow i is observing him and following him, hence the crew owner would deceive crow i by going to any other location in the search space. In CSA, the location of crow i is upgraded by an arbitrary location, and the accurate state can be defined as:

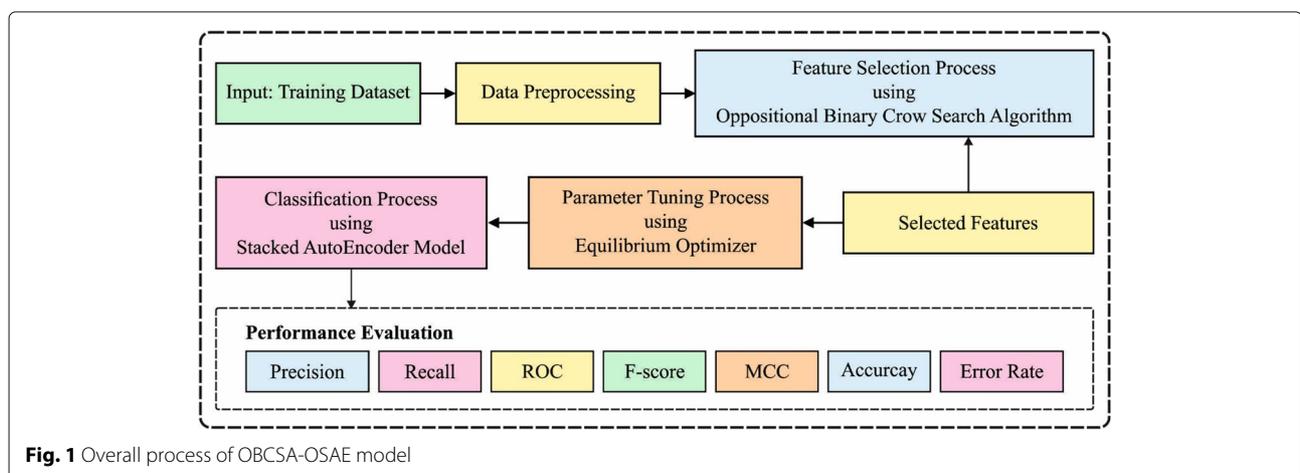


Fig. 1 Overall process of OBSCSA-OSAE model

$$x_i^{t+1} = \begin{cases} x_i^t + \tau \times fl \times (M_j^t - x_i^t), & \text{if } \theta \geq AP \\ \text{random position,} & \text{otherwise} \end{cases} \quad (2)$$

Whereas θ represents the arbitrary value from the range of $[0, 1]$ and AP implies the probability of awareness. The CSA was adapted and is utilized in FS by suggesting a binary search model [21]. In the BCSA, the search space was developed based on the Boolean lattice of n dimension and the essential solution is upgraded through the corners of a hypercube, different from the typical CSA where solution is upgraded from the continuous spaces. The binary vector has been utilized for FS, whereas one equivalent to either the feature that chosen for creating the novel datasets, and or else 0. The idea of OBL has been employed the fl variable from the BCSA to prevent trapping from the local optimal and for enhancing the quality of resultant solution by attaining a balance among exploration as well as exploitation and obtain effective solution. The variable fl is initiated in the BCSA according to the OBL instead of utilizing an arbitrary initiation which might be far from the optimum global solutions and is made by:

$$x = a + b - x, \quad (3)$$

Whereas \bar{x} represent the opposite number and $x \in R$ denotes a real number determined on range of $x \in [a, b]$. While $a = 0$ and $b = 1$ Eq. (3) becomes

$$x = 1 - x, \quad (4)$$

While there is a point $P(x_1, x_2, \dots, x_n)$ in n dimension coordinate and $x_1, x_2, \dots, x_n \in R$ later, the opposite point \bar{P} is determined as its coordinates $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n$:

$$\bar{x}_i = a_i + b_i - x_i \quad i = 1, \dots, n \quad (5)$$

In such cases, have 2 values, x represent initial arbitrary value in $[a, b]$ and \bar{x} denotes the opposite values of x . They calculate $f(x)$ & $f(\bar{x})$ in all the iterations of OBCSA, later, employ on the evaluation function g if $g(f(x)) \geq g(f(\bar{x}))$ select x or else select \bar{x} . Consequently, the fl would be in range: $fl \in [fl_{min}, fl_{max}]$. The opposite number \bar{fl} can be determined by:

$$\bar{fl} = fl_{min} + fl_{max} - x, \quad (6)$$

Later, evaluate the fitness for the first fl value and the fitness for \bar{fl} in all the iterations. When $fitness(fl) \geq fitness(\bar{fl})$, they select fl , or else \bar{fl} would be selected. The stages of presented method can be given in the following.

Step1: The count of crows is $nc = 25, fl_{min} = 0.1, fl_{max} = 1.8, AP = 0.3$, and the maximal number of iterations is $t_{max} = 100$.

Step2: The position that represent the features are made by $U(0, 1)$.

Step3: The fitness function (FF) can be determined by

$$Fitness = C + W \times \left(1 - \frac{F_{all}}{F_{sub}}\right), \quad (7)$$

Whereas C represent the classification performance, W represent the weighted factors in the range of zero and one, F_{all} represent the overall amount of features and F_{sub} signifies the length of elected feature.

Step4: The position of the crows are upgraded as Eq. (2)

Step5: Steps 3 & 4 are repetitive till a t_{max} is attained.

Process involved in OSAE based classification model

At the time of classification process, the chosen features are passed to the OSAE model. AEs have been unsupervised ANN utilized to representation learning. The AE structure has been planned for imposing a bottleneck from the network which forces a compressed knowledge illustration of the input. So, the correlation among the input features is learned and recreated. AEs are encoding-decoding frameworks. The encoding map the original input x to hidden layer that has been regarded as latent space representation. The decoding then regenerates this latent representation as to \hat{x} . The encoder as well as decoder models are determined in Eqs. (8) and (9) correspondingly

$$h = \sigma(Wx + b), \quad (8)$$

$$X = \sigma(W'h + b'), \quad (9)$$

where $x = (x_1, x_2, \dots, x_n)$ signifies the input data vector, $h = (h_1, h_2, \dots, h_n)$ refers the low dimension vector reached in the hidden layer, and $\hat{x} = (\hat{x}_1, \hat{x}_2, \dots, \hat{x}_n)$ represents the recreated input. W and W' defines the weight matrices, b and b' demonstrated the bias vectors, and σ denotes the sigmoid activation function, for instance, $\sigma = \frac{1}{1+e^{-x}}$. It can utilize the MSE function as the recreated error function amongst h and X :

$$E = \frac{1}{N} \sum_{i=1}^N \|\hat{x}_i - x_i\|^2 \quad (10)$$

Overfitting has been common challenge which occurs if trained AE network. An effectual manner for solving this issue is by implementing a weight penalty to cost function:

$$E = \frac{1}{N} \sum_{i=1}^N \frac{1}{2} \|\hat{x}_i - x_i\|^2 + \frac{\lambda}{2} \left(\|W\|^2 + \|W'\|^2 \right) \quad (11)$$

where λ refers the weight attenuation coefficients. Moreover, a sparse penalty term has been presented from the AE hidden layer for achieving optimum feature learning in sparse constraint and keep a condition in which the AE copies the input data to outcome [22]. Considering

\hat{p}_j represents the average activation of hidden layer neuron, it can be determined as $\hat{p}_j = \frac{1}{N} \sum_{i=1}^N h_j(x_i)$, and ρ refers the sparsity proportion, frequently a small positive value nearby 0. For achieving sparsity, it can be restricted to $\hat{p}_j = \rho$, and the Kullback-Leibler (KL) divergence was established for the loss function as regularization term:

$$KL(\hat{\rho} \parallel \rho) = \sum_{j=1}^K \rho \log \left(\frac{\rho}{\hat{\rho}_j} \right) + (1 - \rho) \log \left(\frac{1 - \rho}{1 - \hat{\rho}_j} \right), \tag{12}$$

where K stands for the number of hidden neurons. Therefore, the loss function of sparse AE now has 3 parts: the MSE, weight attenuation, and sparsity regularization parts:

$$E = \frac{1}{N} \sum_{i=1}^N \frac{1}{2} \|x_i - x_i\|^2 + \frac{\lambda}{2} \left(\|W\|^2 + \|W'\|^2 \right) + \beta KL(\hat{\rho} \parallel \rho), \tag{13}$$

where β refers the sparsity regularization parameter. Also, it can stack various sparse AEs for achieving improved feature learning. The framework entails linking the encoder to input layers of the next sparse AE, so make sure the network gains optimum representation learning. Figure 2 depicts the framework of SAE.

During this analysis, the SSAE network is presented. In the SSAE network, the hidden layer of the earlier sparse AE serves as input to the next sparse AE. The last hidden layer is then linked to the softmax classifier that carries out the classification. So, the presented SSAE network contains train sparse AEs and a Softmax classifier. The BP was implemented for fine-tuning the parameter of the total network with the trained samples and their labels. The fine-tuning stage assumes the many layers of the network

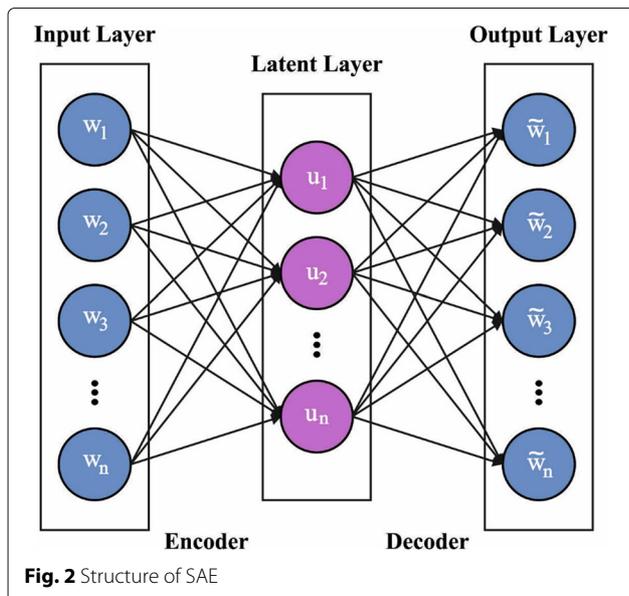


Table 1 Best cost analysis of OBSCSA-FS model

Methods	No. of Selected Features	Best Cost
OBSCSA-FS	15	0.07462
BCSA-FS	24	0.08412
EHO-FS	36	0.09321
GSO-FS	40	0.09468
ACO-FS	48	0.09741

as one model. Considering $\{y^1, y^2, \dots, y^m\}$ implies the target variable of the trained data, the cost function of the whole network is determined as:

$$E = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^N 1\{y^i = j\} \log \frac{e^{\theta_i^T x^i}}{\sum_{i=1}^N e^{\theta_i^T x^i}} \right], \tag{14}$$

where $1\{\bullet\}$ indicates the indicator function, for instance, $1\{y^i = j\} = 1$ if $y = 1$, and $1\{y^i = j\} = 0$ if $y \neq j$, N implies the number classes, and θ_i represents the weight matrix associating the i th output unit. In addition, it can employ the EO technique for optimizing the SSAE parameter, for instance, an optimum weight as well as bias values. The selection of weights as well as bias are vital from trained robust NNs.

The parameter optimization of SAE model takes place using the EO algorithm [23]. It follows dynamic mass balance system that runs on control volume. An arithmetical expression is utilized for representing a mass balance to determine the concentration of a non-reactive constituent under the dynamic environments of control volume. This expression is a function with their several processes under the different kinds of source and sink. The entire description of the EO approach can be explained by the following: The haphazard population (primary concentration) is ini-

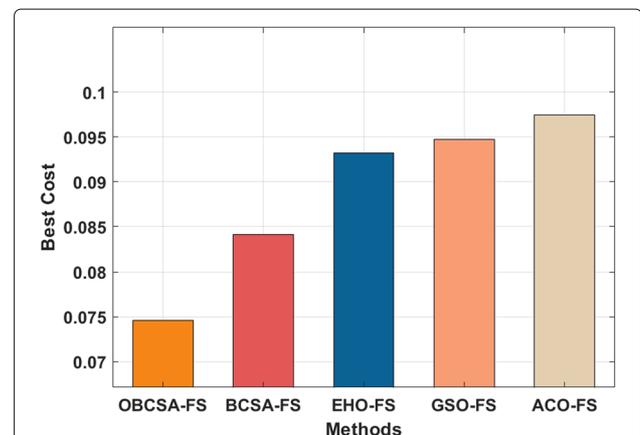


Fig. 3 Best cost analysis of OBSCSA-FS model

tiated by standard distribution depending on amount of particles and dimensional in provided search region:

$$C_i^{initial} = C_{min} + rand_i (C_{max} - C_{min}) \quad i = 1, 2, \dots, n \quad (15)$$

where as $C_i^{initial}$ represent vector of early concentration of i th particles, C_{min} and C_{max} represents lower and upper bounds, $rand_i$ indicates uniform arbitrary number from the range of zero and one and n denotes the size of population.

In order to define the equilibrium state (global optimal), a pool of 4 optimum to this point candidates should be found along with other particles using a concentration equivalent to arithmetical mean of these 4 particles. Thus particles together procedure a pool vector as follows

$$\vec{C}_{eq.pool} = \left(\vec{C}_{(1)}, \vec{C}_{(2)}, \vec{C}_{(3)}, \vec{C}_{(4)}, \vec{C}_{eq(ave)} \right) \quad (16)$$

In the evolution phase, initial particle updates its concentration in initial generation according to on $\vec{C}_{eq(1)}$ and in next generation, then upgrading might takes place

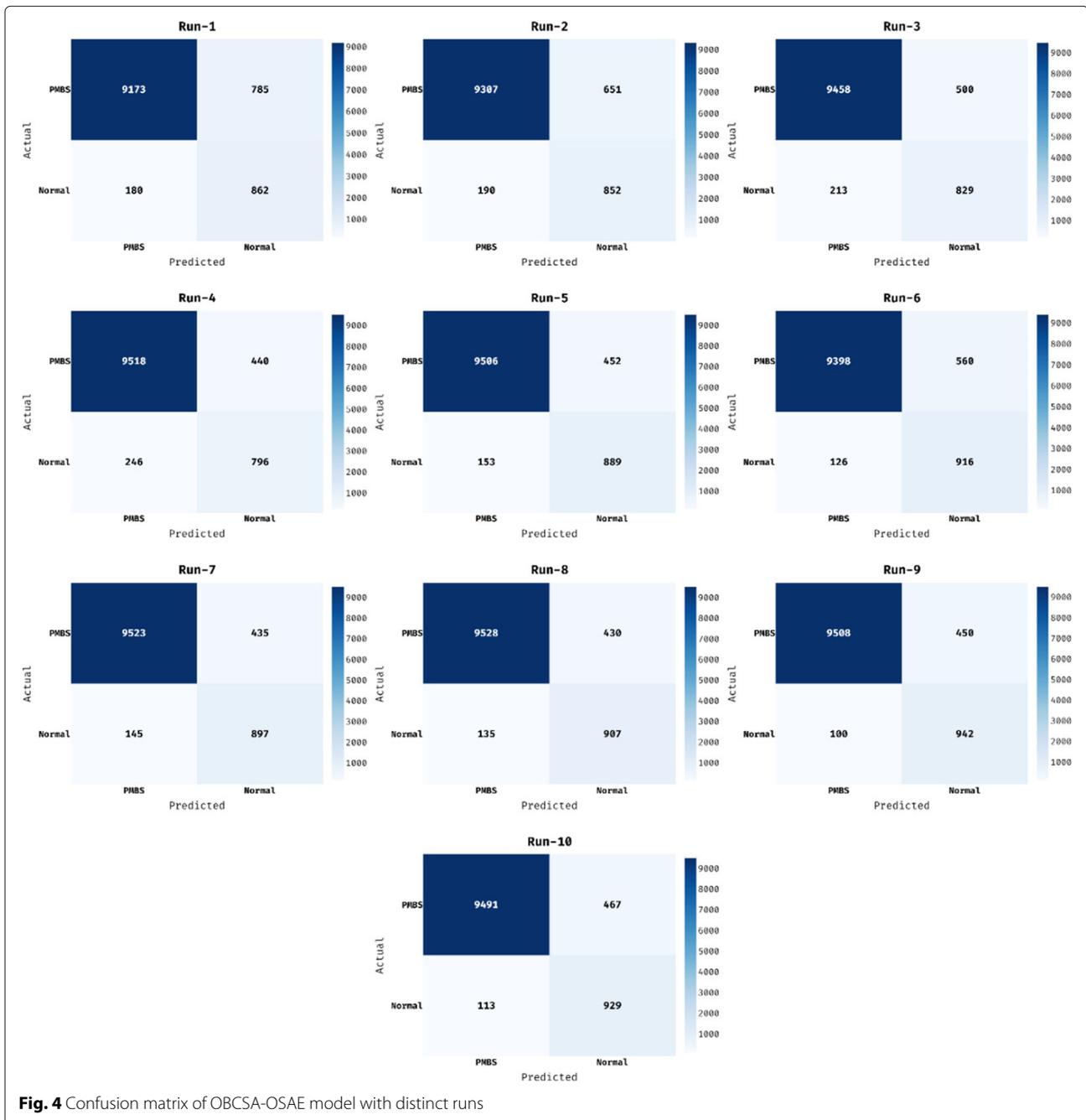


Table 2 Result analysis of OBCSA-OSAE model with different measures

No. of Runs	Precision	Recall	Accuracy	F-Score	MCC	Error Rate
Run-1	0.9808	0.9212	0.9123	0.9500	0.6143	0.0877
Run-2	0.9800	0.9346	0.9235	0.9568	0.6414	0.0765
Run-3	0.9780	0.9498	0.9352	0.9637	0.6697	0.0648
Run-4	0.9748	0.9558	0.9376	0.9652	0.6674	0.0624
Run-5	0.9842	0.9546	0.9450	0.9692	0.7230	0.0550
Run-6	0.9868	0.9438	0.9376	0.9648	0.7069	0.0624
Run-7	0.9850	0.9563	0.9473	0.9704	0.7335	0.0527
Run-8	0.9860	0.9568	0.9486	0.9712	0.7414	0.0514
Run-9	0.9896	0.9548	0.9500	0.9719	0.7565	0.0500
Run-10	0.9882	0.9531	0.9473	0.9704	0.7431	0.0527
Average	0.9833	0.9481	0.9384	0.9654	0.6997	0.0616

on $\vec{C}_{eq(ave)}$. Then, all the particles with each candidate solution are upgraded until the completion of evolution process.

The exponential word F displayed in Eq. (3) helps EO method by attaining an appropriate balance among intensification and diversification. λ represent an arbitrary number from the range zero and one for controlling the turn-over rate in actual control volume.

$$\vec{F} = e^{-\vec{\lambda}(t-t_0)} \tag{17}$$

Whereas t represents the amount of iteration (*Iter*) as follows:

$$t = \left(1 - \frac{Iter}{Max_iter}\right) \left(a_2 \frac{Iter}{Max_iter}\right), \tag{18}$$

In which *Iter* = current iteration, *Max_iter* = maximum iteration and parameters a_2 is used for controlling exploitation capability of EO [24]. For ensuring the convergence when improving local and global search capability of the method:

$$t_0 = \frac{1}{\lambda} \ln \left(-a_1 \text{sign}(\vec{r} - 0.5) \left(1 - e^{-\vec{\lambda}t}\right) \right) + t, \tag{19}$$

In which a_1 & a_2 is utilized for controlling global as well as local search capability of EO method. The $\text{sign}(\vec{r} - 0.5)$, is accountable for the way of exploitation and exploration. In EO, the a_1 & a_2 values are selected to be 2 and 1 correspondingly.

By replacing Eq. (19) in Eq. (17), the equation would become:

$$\vec{F} = a_1 \text{sign}(\vec{r} - 0.5) \left[e^{-\vec{\lambda}t} - 1 \right], \tag{20}$$

The generation rate in EO is employed for improving exploitation as a function of time. The initial order expo-

ponential decay procedure in the method of generation rate of multi-purpose models:

$$\vec{G} = \vec{G}_0 e^{-\vec{k}(t-t_0)}, \tag{21}$$

Whereas G_0 = first value, k = decay variable.

Lastly, the generation rate expression assumes $k = \lambda$:

$$\vec{G} = \vec{G}_0 e^{-\vec{\lambda}(t-t_0)} = \vec{G}_0 \vec{F}_0, \tag{22}$$

$$\vec{G}_0 = GCP \left(\vec{C}_{eq} - \vec{\lambda} \vec{C} \right), \tag{23}$$

$$GCP = \begin{pmatrix} 0.5r_1, & r_2 \geq 0 \\ 0, & r_2 < 0 \end{pmatrix}, \tag{24}$$

Let r_1, r_2 be the 2 arbitrary numbers from the range of [0, 1] and *GCP* variable is for controlling generation rate.

Based on the above equation, the last upgrading equation of concentration (particles) is determined by:

$$\vec{C} = \vec{C}_{eq} + \left(\vec{C} - \vec{C}_{eq} \right) \vec{F} + \frac{\vec{G}}{\lambda V} \left(1 - \vec{F} \right), \tag{25}$$

The upgrading equation has 3 terms: initial term is an equilibrium concentration; the next term is used to global search and last term has been accountable to local search for attaining solution more precisely. In order to optimally adjust the parameter of the SAE algorithm, the EO model is utilized and the thorough functioning can be given as follow. The training method of the SAE algorithm is performed via a FF. Additionally, ten-fold cross-validation procedure is used for evaluating the FF. In ten-fold CV, the training databases are arbitrarily segmented to a set of ten mutually exclusive subsets of almost equivalent size whereas 9 subsets are utilized for training the information and the residual one is employed for testing the information. This processes are repetitive for the collection of ten

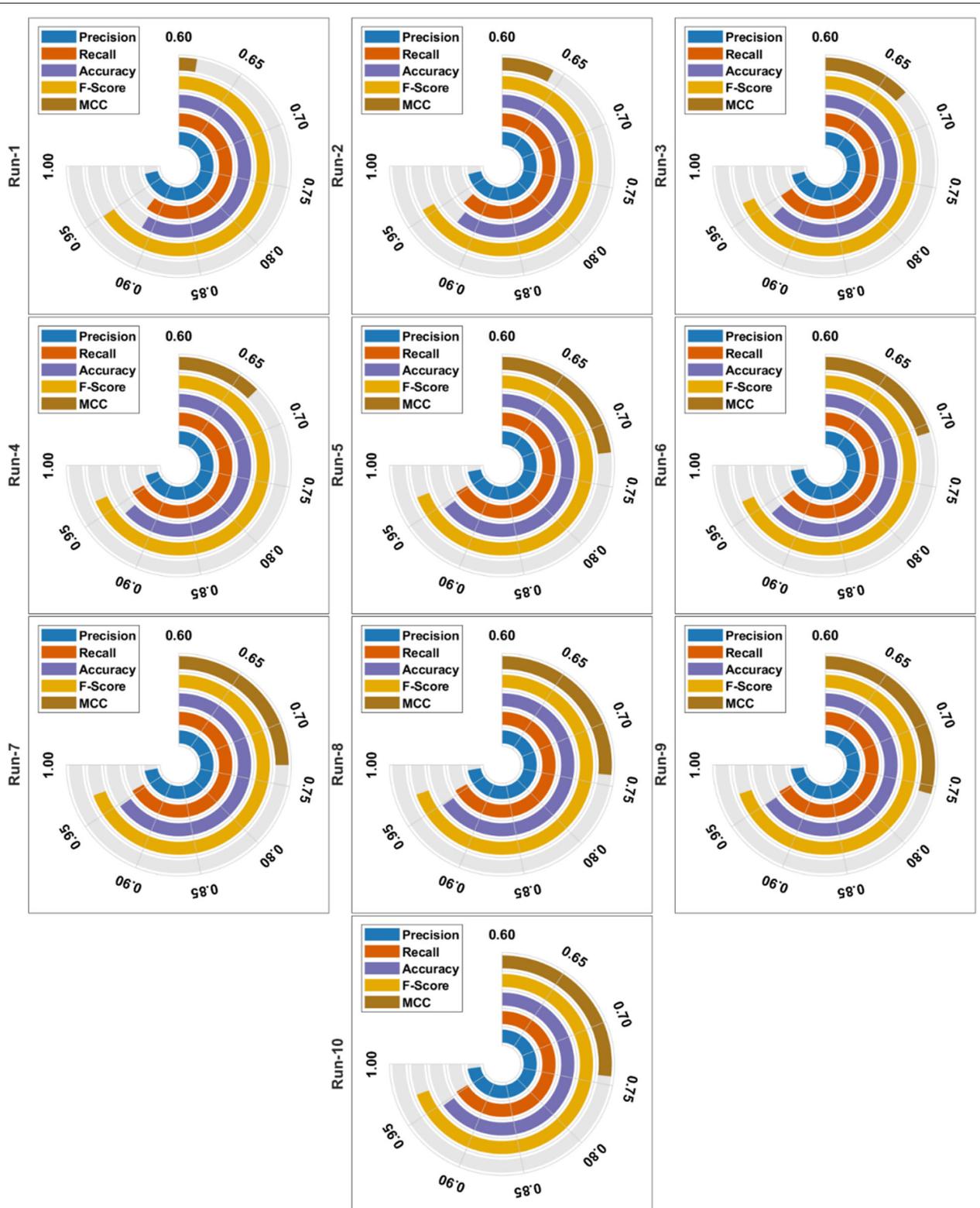
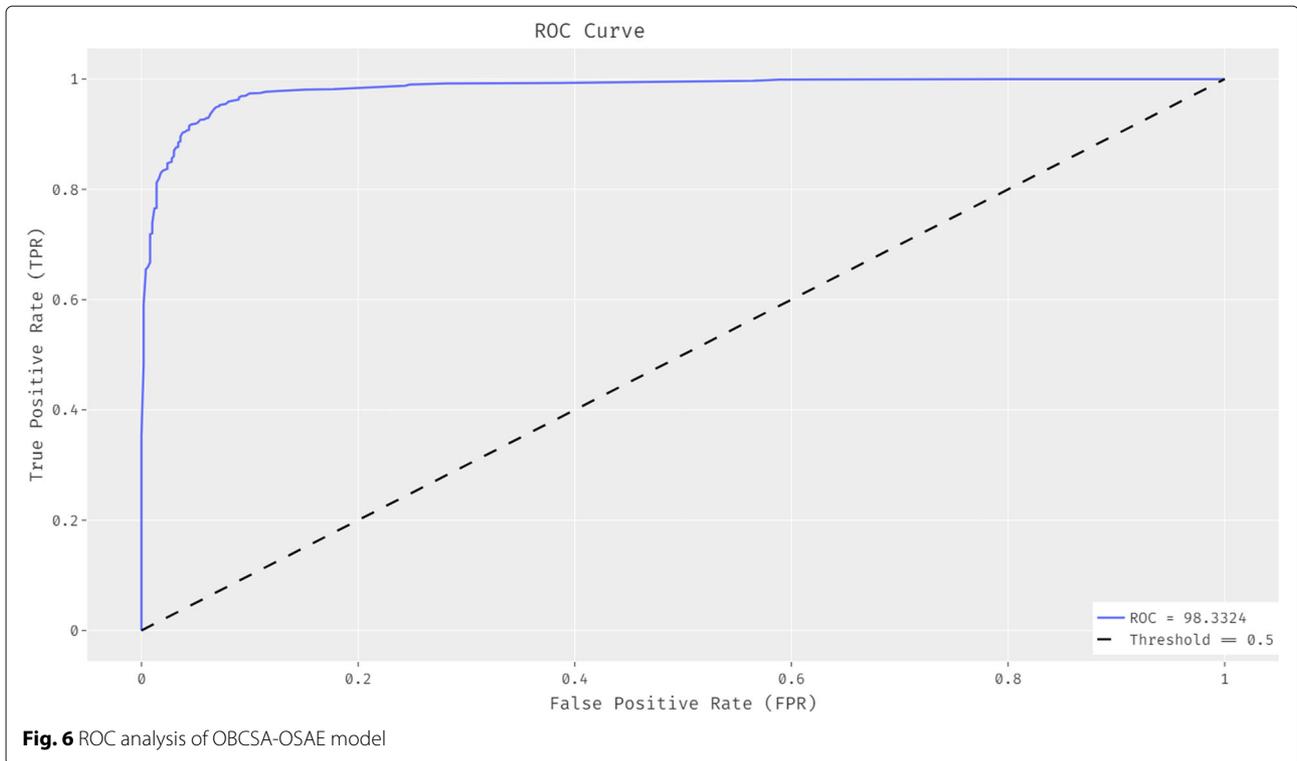


Fig. 5 Result analysis of OBSCA-OSAE model with different runs



iterations that the subsets utilize for testing the method. The FF is represented as $1 - CA_{validation}$ of the ten-fold CV method in the training data, as determined in Eq. (26). As well, a solution with maximal $CA_{validation}$ results in minimum fitness value.

$$Fitness = 1 - CA_{validation} \tag{26}$$

$$CA_{validation} = 1 - \frac{1}{10} \sum_{i=1}^{10} \left| \frac{y_c}{y_c + y_f} \right| \times 100, \tag{27}$$

Whereas, y_c & y_f indicates the amount of true and false classification. Lastly, the hyperparameter included in the SAE algorithm is optimally picked up by the EO method in which the classification performances get enhanced.

Experimental validation

The performance validation of the OBCSA-OSAE model takes place using a data set collected from an obstetrics and gynecology hospital in Hebei province during 2020-2021. It comprises 11000 instances with a total of 149 features. Among the available instances, a set of 1042 samples comes under the presence of PPH. As the trial and error we performed the test and train with variable size to data starting from 50 percent of test and same amount of training data. Generally machine learning to need a huge dat to train the algorithm m results are displayed with The best cost analysis of the OBCSA-FS with existing techniques are demonstrated in Table 1 and Fig. 3. The results demonstrated that the GSO-FS and ACO-FS techniques

have showcased ineffective outcomes with the maximum best cost of 0.09468 and 0.09741 respectively. In line with, the EHO-FS and BCSA-FS techniques have obtained a moderate best cost of 0.09321 and 0.08412 respectively. But the OBCSA-FS technique has accomplished better outcomes with the least best cost of 0.07462.

The confusion matrix produced by the OBCSA-OSAE technique on the execution of ten dissimilar runs is illustrated in Fig. 4. On the test run-1, the OBCSA-OSAE technique has classified 9713 and 862 into PMBS and Normal classes respectively. Also, on the test run-3, the OBCSA-OSAE method has classified 9458 and 829 into PMBS and Normal classes correspondingly. Besides, on the test run-6, the OBCSA-OSAE algorithm has classified 9398 and 916 into PMBS and Normal classes respectively.

Table 3 Comparative analysis of OBCSA-OSAE model with existing methods in terms of different measures

Methods	Accuracy	Recall	Precision	F-Score	MCC
RF	0.912	0.747	0.937	0.830	0.630
GBDT	0.919	0.655	0.736	0.690	0.670
XGB	0.916	0.739	0.878	0.800	0.660
SVM	0.923	0.748	0.937	0.830	0.640
EL-HC	0.907	0.656	0.953	0.780	0.650
EL-SC	0.932	0.697	0.847	0.760	0.690
OBCSA-OSAE	0.938	0.948	0.983	0.965	0.700

At the same time, on the test run-8, the OBCSA-OSAE approach has classified 9508 and 942 into PMBS and Normal classes respectively. Eventually, on the test run-10, the OBCSA-OSAE system has classified 9491 and 929 into PMBS and Normal classes correspondingly.

Table 2 and Fig. 5 tabulates the overall classification results analysis of the OBCSA-OSAE technique under ten runs. The OBCSA-OSAE technique has gained effectual outcomes under every run.

Result analysis of OBCSA-OSAE model with different runs

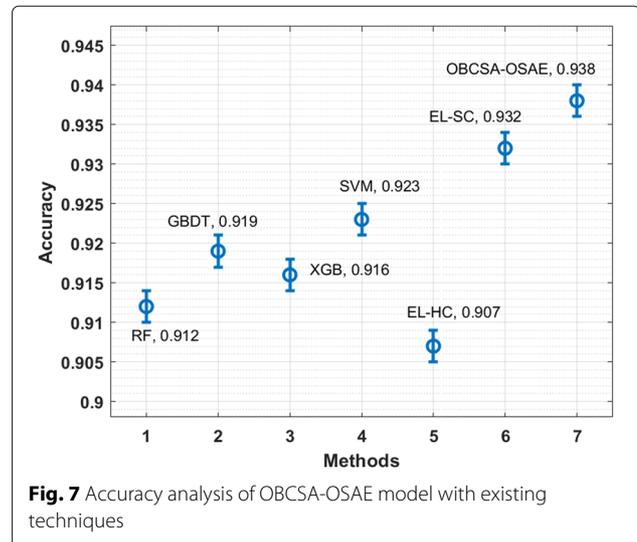
With run-1, the OBCSA-OSAE technique has obtained precision, recall, accuracy, F-score, MCC, and error rate of 0.9808, 0.9212, 0.9123, 0.9500, 0.6143, and 0.0877. Similarly, with run-3, the OBCSA-OSAE methodology has attained precision, recall, accuracy, F-score, MCC, and error rate of 0.9780, 0.9498, 0.9352, 0.9637, 0.6697, and 0.0648. Followed by, in run-6, the OBCSA-OSAE approach has obtained precision, recall, accuracy, F-score, MCC, and error rate of 0.9868, 0.9438, 0.9376, 0.9648, 0.7069, and 0.0624. Finally, with run-10, the OBCSA-OSAE approach has obtained precision, recall, accuracy, F-score, MCC, and error rate of 0.9882, 0.9531, 0.9473, 0.9704, 0.7431, and 0.0527.

The ROC analysis of the OBCSA-OSAE technique is offered in Fig. 6. The figure shows that the OBCSA-OSAE technique has accomplished effectual outcomes with a higher ROC of 98.3324%.

Finally, a brief comparative result analysis of the OBCSA-OSAE technique with existing ones is made in Table 3 [25]. The accuracy analysis of the OBCSA-OSAE technique with compared methods is provided in Fig. 7. The figure displays that the EL-HC, RF, XGB, and GBDT techniques have moderately closer accuracy values. At the same time, the SVM and EL-SC techniques have obtained slightly improved accuracy values. However, the OBCSA-OSAE technique has outperformed the other techniques with a maximum accuracy of 0.938.

The recall, precision, and F-score analysis of the OBCSA-OSAE approach with related techniques are given in Fig. 8. The figure demonstrated that the EL-HC, RF, XGB, and GBDT methods have outperformed moderately closer recall, precision, and F-score values. In addition, the SVM and EL-SC methods have reached somewhat enhanced recall, precision, and F-score values. Finally, the OBCSA-OSAE methodology has exhibited the other systems with higher recall, precision, and F-score of 0.948, 0.983, and 0.965.

The MCC analysis of the OBCSA-OSAE technique with compared methods is provided in Fig. 9. The figure depicts that the EL-HC, RF, XGB, and GBDT manners have moderately closer MCC values. Simultaneously, the SVM and EL-SC approaches have somewhat superior



MCC values. Lastly, the OBCSA-OSAE algorithm show the other techniques with the maximal MCC of 0.700. From these result analysis, the OBCSA-OSAE technique has accomplished maximum PPH prediction outcomes.

Conclusion

In this study, an effective OBCSA-OSAE technique was derived for the detection and classification of PPH. The proposed OBCSA-OSAE technique encompassed two major stages, namely selection of features and classification. Primarily, the OBCSA technique was utilized to optimally select a subset of features. Then, the EO algorithm served to optimally choose the parameters involved in the SAE model. Finally, the SAE model was used for the classification of PPH. For showcasing the improved performance of the OBCSA-OSAE approach was varied out

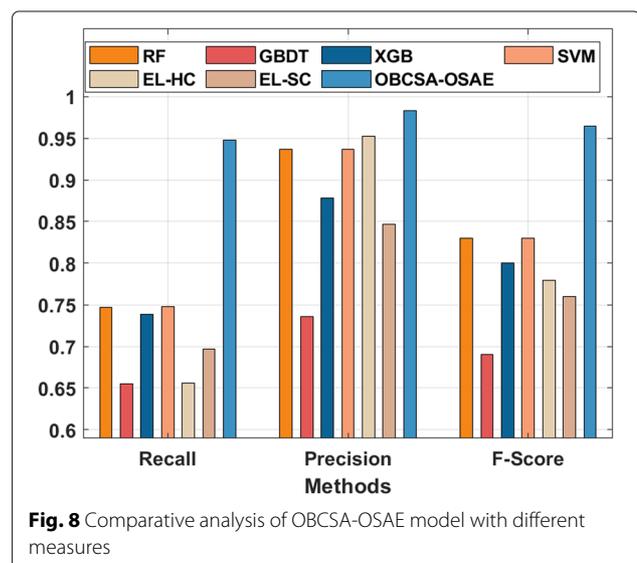
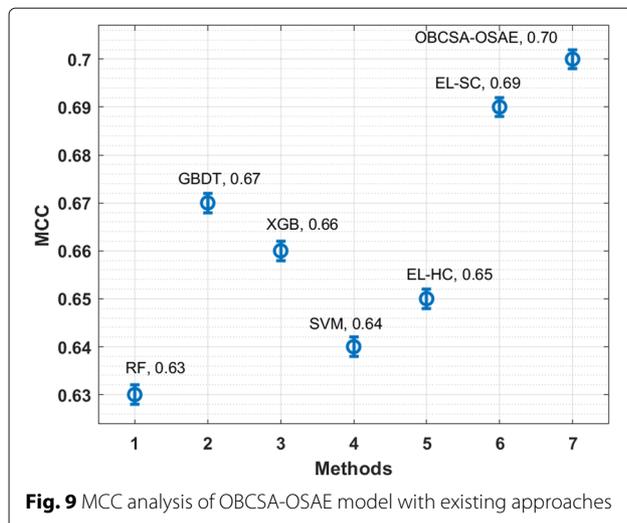


Fig. 8 Comparative analysis of OBCSA-OSAE model with different measures



on the benchmark data set. The experimental outcomes highlighted that the OBCSA-OSAE technique has demonstrated the other techniques featured in terms of different evaluation parameters. In future research, the clustering process could be included to handle massive amounts of PPH data and improve classification outcomes. In the future, we would like to analyze on the time complexity of the algorithm comparing the various parameters.

Acknowledgments

The authors would also like to thank Hebei QINUO Information Technology Co., Ltd for data support. We would like to thank the reviewers for providing such a wonderful comment to structure the work and also some inspiration to words the future work.

Authors' contributions

Dr. SK and Dr. KL are the supervisors of the project. YL analyzed the data and run the code. Dr. SK and Dr. YL helped in contributing the technical ideas and also helped the research scholar in writing. Problem was initially from the research scholar and the machine learning technique was suggested by Dr. SK and the Dr. KL helped in the literature study and coined the optimization Algorithm. After multiple trial of various algorithms worked by all three of us contributed in the writing of the paper. Contributed in section of proposed method and data collection and the algorithm implementation was contributed by KL and YL completed the paper and both SK and YL revised the paper. All authors read and approved the final manuscript.

Funding

This work is funded by the Institute of Societal & Contemporary Computing (ISCC). This research was funded by "Wenzhou-Kean University" in Wenzhou, Zhejiang Province, China (Grant No. ICRP202204).

Availability of data and materials

Data used to support the findings of this study are available from the corresponding author upon request.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of Computer Science, Wenzhou-Kean University, WENZHOU KEAN UNIVERSITY, 88 UNIVE, Wenzhou 325006, Zhejiang, China. ²College of Bioscience and Bioengineering, Hebei University of Economics and Business Shijiazhuang 050061, Hebei, China.

Received: 14 January 2022 Accepted: 18 May 2022

Published online: 13 July 2022

References

- World Health Organization. WHO Recommendations for the Prevention and Treatment of Postpartum Haemorrhage. Geneva: World Health Organization; 2012.
- Callaghan WM, Kuklina EV, Berg CJ. Trends in postpartum hemorrhage: the United States, 1994–2006. *Am J Obstet Gynecol.* 2010;202(4):353–56.
- Einerson BD, Miller ES, Grobman WA. Does a postpartum hemorrhage patient safety program result in sustained changes in management and outcomes? *Am J Obstet Gynecol.* 2015;212(2):140–44.
- Main EK, Goffman D, Scavone BM. National partnership for maternal safety: consensus bundle on obstetric hemorrhage. *Obstet Gynecol.* 2015;126:155–62.
- Chen C, Liu X, Chen D, Huang S, Yan X, Liu H, Chang Q, Liang Z. A risk model to predict severe postpartum hemorrhage in patients with placenta previa: A single-center retrospective study. *Ann Palliat Med.* 2019;8(5):611–21.
- Koopmans CM, Tuuk KVD, Groen H, Doornbos JPR, Graaf IMD. Prediction of postpartum hemorrhage in women with gestational hypertension or mild preeclampsia at term. *Acta Obstet Gynecol Scand.* 2014;93(4):399–407.
- Kramer MS, Berg C, Abenhaim H, Dahhou M, Rouleau J, Mehrabadi A, Joseph KS. Incidence, risk factors, and temporal trends in severe postpartum hemorrhage. *Amer J Obstet Gynecol.* 2013;209(5):449.
- Onan A. On the performance of ensemble learning for automated diagnosis of breast cancer. *Adv Intell Syst Comput.* 2015;347:119–29.
- Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: A practical introduction. *BMC Med Res Methodol.* 2019;19:64.
- Mienye ID, Sun Y, Wang Z. An improved ensemble learning approach for the prediction of heart disease risk. *Informat Med Unlocked.* 2020;20:100402.
- Venkatash KK, Strauss RA, Grotegut C, Heine RP, Chescheir NC, Stringer JS, Stamilio DM, Menard MK, Jelovsek JE. Machine learning and statistical models to predict postpartum hemorrhage. *Obstet Gynecol.* 2020;135(4):935.
- Kumar VA, Sharmila S, Kumar A, Bashir AK, Rashid M, Gupta SK, Alnumay WS. A novel solution for finding postpartum hemorrhage using fuzzy neural techniques. *Neural Computing and Applications.* 2021;33:1–14.
- Wu Q, Yao K, Liu Z, Li L, Zhao X, Wang S, Shang H, Lin Y, Wen Z, Zhang X, Tian J. Radiomics analysis of placenta on T2WI facilitates prediction of postpartum hemorrhage: a multicenter study. *EBioMedicine.* 2019;50:355–65.
- Betts KS, Kisely S, Alati R. Predicting common maternal postpartum complications: Leveraging health administrative data and machine learning. *BJOG Int J Obstet Gynaecol.* 2019;126(6):702–9.
- Man Z. Comparative Study of Machine Learning Models to Predict PPH. Master's Paper. The University of North Carolina at Chapel Hill; 2019.
- Kumar VD, Sharmila S, Kumar A, Subha SS, Singh V, Kannan R. Predictive Analysis of Postpartum Haemorrhage And Hypothermia Using Wearable Device. *Eur J Mol Clin Med.* 2020;7(4):2260–4.
- Hochman E, Feldman B, Weizman A, Krivoy A, Gur S, Barzilay E, Gabay H, Levy J, Levinkron O, Lawrence G. Development and validation of a machine learning-based postpartum depression prediction model: A nationwide cohort study. *Depression Anxiety.* 2021;38(4):400–11.
- Yang J, Guo P, Song Y, Han L, Yang X, Bai H. A Clinical Decision Support System for Prediction of Postpartum Hemorrhage in Vaginal Birth. 2021. <https://doi.org/10.21203/rs.3.rs-618703/v1>.
- Zheutlin AB, Vieira L, Shevcraft RA, Li S, Wang Z, Schadt E, Gross S, Dolan SM, Stone J, Schadt E, Li L. Improving postpartum hemorrhage risk prediction using longitudinal electronic medical records. *Journal of the American Medical Informatics Association: JAMIA.* 2022;29(2):296–305.

20. Askarzadeh A. A novel metaheuristic method for solving constrained engineering optimization problems: crow search algorithm. *Comput Struct.* 2016;169:1–12.
21. Al-Thanoon NA, Algamal ZY, Qasim OS. Feature selection based on a crow search algorithm for big data classification. *Chemometrics and Intelligent Laboratory Systems.* 2021;212:104288.
22. Mienye ID, Sun Y. Improved Heart Disease Prediction Using Particle Swarm Optimization Based Stacked Sparse Autoencoder. *Electronics.* 2021;10(19):2347.
23. Faramarzi A, Heidarinejad M, Stephens B, Mirjalili S. Equilibrium optimizer: A novel optimization algorithm. *Knowl-Based Syst.* 2020;191:105190.
24. Dinkar SK, Deep K, Mirjalili S, Thapliyal S. Opposition-based laplacian equilibrium optimizer with application in image segmentation using multilevel thresholding. *Expert Syst Appl.* 2021;174:114766.
25. Zhang Y, Wang X, Han N, Zhao R. Ensemble Learning Based Postpartum Hemorrhage Diagnosis for 5G Remote Healthcare. *IEEE Access.* 2021;9:18538–48.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

