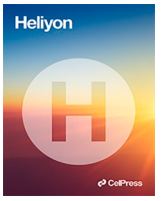
Heliyon 10 (2024) e26799

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

An enhanced and efficient approach for feature selection for chronic human disease prediction: A breast cancer study

Munish khanna [a,*], Law Kumar Singh [b], Kapil Shrivastava [b], Rekha singh [c]

[a] School of Computing Science and Engineering, Galgotias University, Greater Noida, Gautam Buddh Nagar, India
[b] Department of Computer Engineering and Applications, GLA University, Mathura, India
[c] Department of Physics, Uttar Pradesh Rajarshi Tandon Open University, Prayagraj, Uttar Pradesh, India

ARTICLE INFO

Keywords:
Feature selection
Breast cancer prediction
Machine learning
Soft-computing
Teaching learning based optimization
Elephant herding optimization
Hybrid approach

ABSTRACT

Computer-aided diagnosis (CAD) systems play a vital role in modern research by effectively minimizing both time and costs. These systems support healthcare professionals like radiologists in their decision-making process by efficiently detecting abnormalities as well as offering accurate and dependable information. These systems heavily depend on the efficient selection of features to accurately categorize high-dimensional biological data. These features can subsequently assist in the diagnosis of related medical conditions. The task of identifying patterns in biomedical data can be quite challenging due to the presence of numerous irrelevant or redundant features. Therefore, it is crucial to propose and then utilize a feature selection (FS) process in order to eliminate these features. The primary goal of FS approaches is to improve the accuracy of classification by eliminating features that are irrelevant or less informative. The FS phase plays a critical role in attaining optimal results in machine learning (ML)-driven CAD systems. The effectiveness of ML models can be significantly enhanced by incorporating efficient features during the training phase. This empirical study presents a methodology for the classification of biomedical data using the FS technique. The proposed approach incorporates three soft computing-based optimization algorithms, namely Teaching Learning-Based Optimization (TLBO), Elephant Herding Optimization (EHO), and a proposed hybrid algorithm of these two. These algorithms were previously employed; however, their effectiveness in addressing FS issues in predicting human diseases has not been investigated. The following evaluation focuses on the categorization of benign and malignant tumours using the publicly available Wisconsin Diagnostic Breast Cancer (WDBC) benchmark dataset. The five-fold cross-validation technique is employed to mitigate the risk of over-fitting. The evaluation of the proposed approach's proficiency is determined based on several metrics, including sensitivity, specificity, precision, accuracy, area under the receiver-operating characteristic curve (AUC), and F1-score. The best value of accuracy computed through the suggested approach is 97.96%. The proposed clinical decision support system demonstrates a highly favourable classification performance outcome, making it a valuable tool for medical practitioners to utilize as a secondary opinion and reducing the overburden of expert medical practitioners.

* Corresponding author.
E-mail addresses: munishkhanna.official@rocketmail.com (M. khanna), lawkumarcs1@gmail.com (L.K. Singh), kapil1411@gmail.com (K. Shrivastava), singh.rekha70@gmail.com (R. singh).

https://doi.org/10.1016/j.heliyon.2024.e26799
Received 9 July 2023; Received in revised form 15 January 2024; Accepted 20 February 2024
Available online 28 February 2024

1. Introduction

Currently, there is an unprecedented generation of data in various fields due to the rapid accumulation of vast amounts of information [1]. Data with high dimensions generally comprises a substantial quantity of information. Nevertheless, it is common to encounter numerous features that are either unrelated or weakly correlated, which can have an impact on data processing [2]. Therefore, it is imperative to devise effective data mining methodologies that can effectively reduce the dimensionality of high-dimensional data across various domains such as medicine, bioinformatics, text mining, and the internet of drones [3,4]. The feature selection (FS) method is a proficient technique for reducing the dimensions of data. The application of this technology has been extensive and carries substantial significance in the domains of machine learning and pattern recognition [5]. The process of reducing the dimensionality of the dataset can enhance the computational efficiency of the model. As per [6], the problem of FS is approached as a combinatorial optimization problem during the preprocessing stage. The objective of FS is to eradicate extraneous and duplicative attributes from specific datasets [7]. Different FS techniques can be selected for data processing depending on specific learning algorithms [8]. highlights the significance of the search strategy in the FS method, especially when dealing with high-dimensional data. There has been a notable surge in the utilisation of meta-heuristics employing varied search strategies to address FS problems [9]. Adding global search to Swarm Intelligence (SI)-based search methods could make the field of FS much easier to understand and use while also lowering the cost and time needed for computing. This has been evidenced in prior research [10]. The optimization algorithms utilized in SI are founded on the behavioural patterns of natural animal communities, particularly their foraging and predation behaviours, which are continuously simulated [11].

At present, various SI algorithms are employed to tackle FS issues. The algorithms under consideration are the Genetic Algorithm [12], Particle Swarm Optimisation [13], and Grey Wolf Optimisation [14]. Numerous algorithms have been proposed in recent times, including the Whale Optimisation Algorithm [15], the Butterfly Optimisation Algorithm [16], and the Grasshopper Optimisation Algorithm [17]. During each iteration, optimization algorithms based on population produce a set of potential solutions. This has the potential to result in improved outcomes in the field of financial services. Abualigah et al. [18] have introduced a novel meta-heuristic optimization technique named Reptile Search Algorithm, which draws inspiration from the hunting patterns of crocodiles. The Dwarf Mongoose Optimization algorithm has been developed to replicate the foraging patterns of the dwarf mongoose. The algorithm exhibits bionic characteristics in its search capabilities [19]. A novel meta-heuristic algorithm, named the Ebola Optimization Search Algorithm, has been introduced. The strategy employed involves the utilisation of a bionic disease transmission mechanism that is based on the propagation mechanism of the Ebola virus disease [20]. This methodology endeavours to enhance the efficacy of the algorithm. Furthermore, the distribution of primary arithmetic operators can be leveraged to devise a meta-heuristic optimization technique called the Arithmetic Optimisation Algorithm [21]. Furthermore, a method for optimization based on population, known as the Aquila Optimizer, has been proposed. This approach is derived from the predatory habits of the aquila [22]. Evolutionary algorithms have shown considerable promise in feature selection, owing to their straightforwardness and ability to conduct global searches. This inspiration has led us to integrate these algorithms into our feature selection work.

The incidence of cancer has progressively transformed into a persistent ailment that has a significant impact on human well-being. Based on the data provided by the World Cancer Research Fund International [23], breast cancer (BC) is the predominant form of cancer among women globally, with approximately 1.7 million new cases detected in 2012, making it the second most prevalent cancer type overall. BC accounts for roughly 12% of newly diagnosed cancer cases and 25% of all cancers in the female population. This cancer ranks fifth in terms of frequency as a cause of mortality among women due to cancer. Furthermore, there has been a rise in the incidence rate coupled with a reduction in the age of onset [24]. According to research, BC is a prevalent form of cancer and a significant contributor to mortality rates among women in Jordan [25]. According to the Jordan Cancer Registry's 2011 report, BC accounts for 20.1% of all cancer cases in Jordan, encompassing both males and females, and 37.3% of all cancer cases in females [26]. BC has a prevalence of 12% among women in the United States throughout their lifespan. In 2017, the United States recorded over 250, 000 new cases of this disease [27].

As per the GLOBOCAN 2012 report, the global count of newly diagnosed cases of breast cancer was 1.7 million [28–30]. According to Ref. [31], there has been a notable increase of approximately 20% in the incidence of breast cancer and a corresponding rise of approximately 14% in the mortality rate, as compared to the statistics from 2008. Moreover, it has been identified as a significant factor that contributes to cancer-related mortality in developing countries [31]. This is a heterogeneous disease, with distinct biological subtypes that exhibit diverse clinicopathological and molecular features. These features have prognostic implications, such as tumour size, grade, and the presence of axillary lymph node metastasis, as well as predictive implications, such as the expression of oestrogen receptor and progesterone receptor and overexpression of human epidermal growth factor receptor 2 proteins. The aforementioned findings have been documented in prior research [32–35]. Triple-negative breast cancer (TNBC), a subtype of breast cancer, is recognised to have a higher incidence rate among younger women and is regarded as the most aggressive manifestation of the ailment. According to the results of the analysis, the incidence of TNBC exhibits variability across diverse nations, with percentages ranging from 6.7% to 27.9%. It is noteworthy that India has the highest incidence of TNBC, followed by Indonesia, Algeria, and Pakistan. The prevalence of triple-negative breast cancer in various countries, including the Netherlands, Italy, London, and Germany, is lower than the average rate of 15%. The incidence of TNBC in the Indian population is associated with several risk factors, including lifestyle, socioeconomic status, obesity, family history, high mitotic indices, and BRCA1 mutations [36].

The preceding discourse serves as motivation for utilizing a soft-computing-based feature selection approach to predict the rapid spread of the aforementioned infection among women worldwide. For this particular infection prediction, we have adopted two algorithms that have a proven track record and have not been previously used to predict this disease to the best of our knowledge. We propose a third algorithm that combines the aforementioned two algorithms. Four well-performing and widely accepted models are

utilized for the performance evaluation of the feature subset returned by the soft-computing algorithms. The performance of these employed ML models is evaluated on several well-defined metrics. This paper suggests an automatic and useful clinical decision-classification system for finding BC. The public UCI repository features set (WDBC) of 32 features is used to feed this classification system.

The main contribution of this work can be considered as.

- The aim is to determine the most effective features for predicting BC infection. The outcome of this empirical study also showcases the most relevant features, among researchers' communities, required for this infection prediction.
- To implement TLBO, EHO, and their hybrid algorithm for the first time for chronic human disease identification.
- To come up with an effective system by revealing the predictive factors of BC patients and comparing the model's robustness using different standard performance measures. A novel hybrid intelligent clinical system for better prediction of BC by using soft computing algorithms and ML applications is presented.
- To show a more thorough comparison (and analysis) of how soft computing algorithms and ML applications have been used to find BC and make it more visible, which will help to validate the model.

The structure of the paper is as follows: Section two is reviewing a literature survey of prior studies. The third section gives a brief overview of feature selection and outlines the recommended methods (nature-inspired computing) for classifying BC. The fourth section showcases the resulting outcomes. The work is concluded in the fifth section.

2. Literature review

There are many medical studies available in the literature that focus on the diagnosis of Breast Cancer. The majority of these algorithms utilize either the WBCD dataset ([37–48]), the WDBC dataset ([49–52]), or a combination of both ([52–59]). The datasets comprise of classifications for both benign and malignant cases. In their study, the authors suggest the use of a genetically optimized neural network (GONN) for breast cancer classification, which is detailed in Ref. [60]. The neural network architecture is optimized using the genetic algorithm (GA). Benteng et al. [40] presented a GA named Tribe-Competition-Based GA (TCbGA) that is used for feature selection. A method for designing rule-based fuzzy BC classifiers was presented by Marco Pota et al. [41]. The design procedures make use of the naïve Bayes approximation, which optimizes the required parameters independently, resulting in fast computation. In Ref. [42], the FAEMODE algorithm is employed for feature selection. This algorithm is based on elitism and multi-objective differential evolution [43]. provides information on how to implement the graph-based skill acquisition method (GSL). This tool captures both the dynamics of the environment and the experiences of the agent. Liangjun et al. [44] introduced a technique known as full-correntropy-based multilayer-extreme-learning-machine (FC-MELM) for classifying a corrupted BC dataset that has been impacted by impulsive noise. In their study [40], the authors introduce a method called sparse-pseudoinverse incremental-ELM (SPI-ELM) for detecting BC. Compared to ELM, it has a lower run-time complexity. Marco Pota et al. [46] proposed an approach called likelihood-fuzzy analysis (LFA) in their study. This approach is used for extracting statistical information from labeled data. The data is subsequently utilized as input for the fuzzy classification system. Ed-Daoudy et al. [47] proposed a two-stage approach to reduce features and classify BC using Association Rules (AR) in their study's first stage. The SVM classifier has been equipped with the reduced features for the BC classification task, utilizing the 3 CV approach. Zhongliang et al. [49] proposed a new Adaboost algorithm called AdaBoost.FT, which includes a floating threshold. The principle of maximum likelihood is utilized to improve the stability of classification. Yamuna Prasad et al. [50] employed a hybrid SVM classifier that integrated particle swarm optimization (PSO), GA, or Ant colony optimization (ACO) to examine the WDBC dataset in their research study. In their study, Zheng et al. [51] developed a hybrid system of SVM and K-means, known as K-SVM, to identify hidden patterns of benign and malignant tumours in WDBC. This approach resulted in improved accuracy. The PSO algorithm underwent testing on 13 distinct datasets, including WDBC and Wisconsin Breast Cancer Database (WBCD) [52]. The PSO algorithm was used in Ref. [53] to optimize feature subset and kernel bandwidth for BC detection, utilizing a non-parametric kernel density estimation (KDE). The use of the KDE-PSO algorithm results in better performance for processing the WDBC dataset in comparison to other algorithms. Peng et al. [54], introduced the artificial immune semi-supervised learning algorithm as a means to enhance the precision of BC detection. The researchers in Ref. [61] employed a Gauss-Newton-Representation-Based Algorithm to calculate the most suitable weights for training samples for the purpose of classifying BC. The WBCD dataset attained an accuracy rate of 98.54%, whereas the WDBC dataset achieved an accuracy rate of 79.54%. The article [55], introduces ST-ONCODIAG, a diagnostic system specifically developed for the detection of breast cancer. The system begins by reading input from datasets, followed by filtering and cleaning the data. A coordinated set of rules was developed utilizing ontologies and rule languages. Afterwards, a framework for representing knowledge is established and modifications are implemented to the ST system. The sensitivity values obtained for WBCD and WDBC were 0.81 and 1, respectively, while the specificity values were 0.89 and 0.706. Ruholla et al. [56] presented the LS-SOED algorithm in their study. This algorithm improves the efficiency of artificial neural networks by integrating life-sensitivity, self-organization, and error-driven mechanisms. The process involves leveraging the advantages of unsupervised and supervised artificial neural network learning techniques, while simultaneously reducing the costs associated with misclassification. Feng Li et al. [57] have proposed a technique known as smooth group L 1/2 (GLSGL1/2) regularization. This method is used to detect and remove unnecessary input nodes in feed-forward neural networks (FFNN). The WBCD and WCBC models attained accuracies of 92.94% and 91.04%, respectively. Khandezamin et al. [58] employed the logistic regression (LR) method for feature selection in the initial stage of their study. The GMDH neural network is used to diagnose breast cancer. The method resulted in a success rate of 99.4% for WBCD and 99.6% for WDBC. A new Gauss-Newton representation-based algorithm (GNRBA) for

breast cancer classification is presented in Ref. [59]. Practitioners employ sparse representation and training sample selection. So far, sparse representation has only worked for pattern recognition. The innovative GaussNewton-based algorithm finds optimal training sample weights for classification. Compared to the l1-norm approach, it evaluates sparsity computationally efficiently. The GNRBA is tested on the UCI Machine Learning repository's WBCD and WDBC.

The stagnation point flow of unstable compressible Casson hybrid nanofluid over a vertical stretched sheet was studied. Xue, Yamada Ota, and Tiwari Das hybrid nanofluid models were compared [60]. Lorentz force affected normal flow. Authors examined nonlinear radiation. Signal and multi-wall carbon nanotubes were water-tested. Boundary layer approximations in partial differential equations under flow suppositions solved a mathematical model. Lie symmetry created the right transformation. From partial differential equations, suitable transformations created ODEs. Dimensionless system revealed by bvp4c numerical method. Tables and graphs showed how flow parameters affected skin friction, Nusselt number, temperature, and velocity distributions. This study found that Yamada Ota outperformed Tiwari Das and Xue hybrid nanofluid models in heat transmission. Increased solid nanoparticle concentration increased skin friction and decreased temperature gradient. Higher solid nanoparticle values increased skin friction by resisting fluid motion. Two-dimensional incompressibility Stable Sutterby fluid flow over a nonlinear stretching cylinder [61]. Under the Buongiorno nanofluid model, fluid properties vary. An induced magnetic field affects sutterby fluid using heat and chemistry. Nonlinear radiative thermal and concentration slip study. A governing equation-based mathematical model meets expectations. These equations become partial differential equations with boundary layer approximation. Ordinary differential equations are transformed from partial differential equations. Dimensionless system numerical solution. Physical factor regulation is shown visually and tabulated. High variable thermal conductivity raised fluid temperature. High liquid thermal conductivity raised fluid temperature. Due to Brownian motion, fluid temperature rose. Different Brownian motion levels increase kinematic energy and fluid temperature. Nanofluid's second-order micropolar stagnation point flow over an exponentially permeable stretched sheet is calculated. This study examines thermal slip-affected freestream velocity. This model is simplified to partial differential equations using flow assumptions before boundary layer approximations [62]. Transformations simplify governing equation math. Solve the differential system with bvp4c. Data is shown in graphs and tables. For mild and strong concentrations, skin friction and Nusselt number are shown. Light concentration causes less skin friction than strong concentration. Under pair stress, weak concentration skin friction exceeds strong concentration. Micropolar profile depends on micropolar and micro-gyration parameters. Sherwood and Nusselt numbers are higher in strong concentration. Micropolar Casson nanofluid flow on a vertical nonlinear Riga stretching sheet was compared. Brownian movements and thermophoresis study thermal and velocity slip. Create coupled nonlinear ODEs by modifying nonlinear PDEs. The numerical method and Runge–Kutta 4th-order scheme solve nonlinear coupled ODEs. Both tabular and graphical representations show how flow parameters affect skin friction, Nusselt number, temperature, and velocity [63]. Brownian motion, Sherwood number magnitude, and Nusselt number are opposite. High Sherwood and Nusselt numbers worked. The Casson fluid parameter increment decreased with fluid velocity, reducing thickness. Increases in micropolar parameter increase fluid velocity distribution curves.

The experimental inquiry utilizes three feature selection strategies, namely Correlation-based selection, Information Gain-based selection, and Sequential feature selection, to select a set of features [64]. The feature subsets undergo evaluation by several machine learning classifiers, and the selection of the ideal feature subset is determined based on its performance. Finally, it is recommended that the Max Voting Classifier, which is constructed using ensembles, be considered as the optimal model among the three leading options. This work is specifically implementing and dependent on a statistical-based approach for the feature selection process; however, the feature selection phase using soft-computing (and their hybrid)-based approaches is missing in this work, which is the core of our work. By integrating the Principal Component Analysis (PCA) technique with ReliefF Feature Ranking, this work offers a novel way to feature selection [65]. When paired with the recommended feature selection process, the use of hybridization approaches has been proven to considerably increase the classifier's precision, namely the k-nearest neighbor strategy. This improvement is shown in the datasets for both chronic illnesses. When compared to five other methods—Correlation Based Feature Selection, Fast Correlation Based Feature Selection, Mutual Information Based Feature Selection, MODTree Filtering Approach, and ReliefF Feature Selection—the hybrid framework described in this study performs better. This study focuses on the implementation and reliance on a statistical-based strategy for the feature selection process. However, it is important to note that the feature selection phase utilizing soft-computing and hybrid-based techniques, which is the fundamental aspect of our research, is not included in this work. Moreover, very limited machine learning classifiers are used for classification. In this study, the bioinformatic tool known as the Oscillating Search Algorithm for Feature Selection was utilized to iteratively enhance the selection of features for the training of Support Vector Machines (SVM) to enhance the accuracy of breast cancer prediction [66]. Only one feature selection approach and a small number of machine learning classifiers are used for the classification assessment in this study.

Recently, various meta-heuristic optimization techniques have been utilized to create hybrid machine learning solutions that can effectively tackle a wide range of challenges in different applications. Several optimization techniques have been developed to improve the performance of various algorithms. These techniques include the Biogeography Based Optimization [67,68], Particle Swarm Optimization [69], Chimp Optimization Algorithm [70,71], sine-cosine Algorithm [72], salp swarm algorithm [73,74], Autonomous Groups Particle Swarm Optimization, whale optimization algorithm [75,76], and Emperor Penguin algorithm [77]. According to the principle of No Free Lunch (NFL), it is not possible to deem any meta-heuristic algorithm as superior to another meta-heuristic approach. Metaheuristic algorithms are better suited for specific optimization problems when compared to other meta-heuristic approaches [76]. We were motivated by this theorem to use these 3 algorithms (TLBO, EHO and Hybrid of these two) in combination with different classifiers to create an efficient and fast BC classification system. Furthermore, to the best of our knowledge, these algorithms have not been utilized in a similar manner as ours for the purpose of feature selection in predicting chronic human disease infections. The aim of this study is to accurately and automatically classify BC as either malignant or benign. This can be achieved by integrating the different classifier with our utilized three algorithms (TLBO, EHO and Hybrid of these two).

3. Approach implemented

In this empirical study, three soft-computing algorithms have been selected and implemented for feature selection: TLBO, EHO, and a hybrid of these two. Fig. 1 illustrates the comprehensive layout of the entire work.

3.1. Elephant herding optimization

The elephant herding optimization approach is one of the more recent developments in the field of swarm intelligence algorithms. In 2016, Wang, Deb, Gao, and Coelho published a paper [78] in which they proposed the notion. In spite of the fact that it is a relatively new algorithm for optimization, it has already been implemented in a number of different scenarios. For community discovery in intricate social networks, an elephant herding optimization technique was presented [79]. Elephant herding behavior served as an inspiration for EHO. Swarm intelligence algorithms were developed by dissecting the complicated natural behavior of elephants. The elephant population is made up of many different lineages. The matriarch of a specific elephant clan is in command. Every generation, a small number of male elephants depart their herd and relocate [80].

$$e_{new,li,k} = e_{li,k} + \beta.\left(e_{best,li} - e_{li,k}\right).n \tag{1}$$

$e_{new,li,k}$ is representing the new position of elephant k in the clan l_i and $e_{li,k}$ shows the old position. $e_{best,li}$ is the best of clan l_i, $\beta \in [0,1]$ is algorithm's parameter indicating the influence of the matriarch, while $n \in [0,1]$ is random number used to improve diversity of the population in the later stages of the algorithms.

$$e_{new,li} = \eta.e_{center}, l_i \tag{2}$$

$$e_{center.l_i,dia} = \frac{1}{N_{li}} \sum_{s=1}^{N_{li}} e_{li},s, dia \tag{3}$$

where dia is representing the dimension (Equation (3)). N_{li} is the number of clan in l_i,where $\eta \in [0,1]$ is the second parameter of the algorithm.

$$e_{wors}, l_i = e_{\min} + (e_{\max} - e_{\min} + 1).rad \tag{4}$$

where $e_{\min}$ and $e_{\max}$ are lower and upper bound of search space. $rad \in [0,1]$ is representing the random number. It is chosen by random

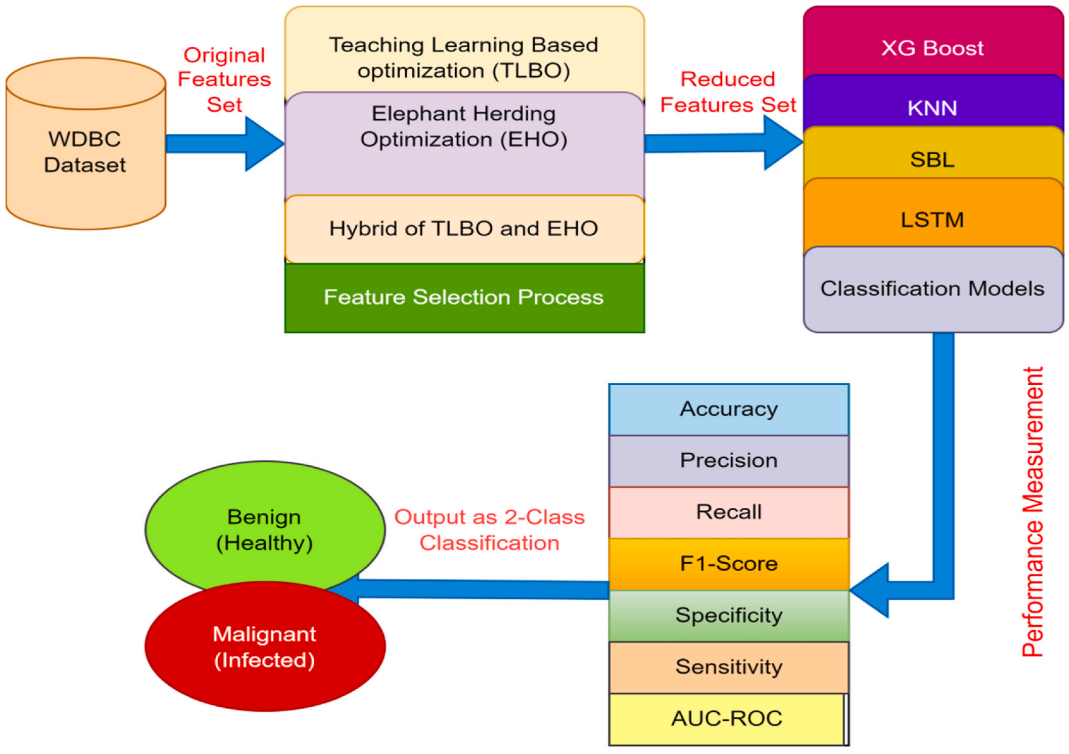


Fig. 1. Framework of our proposed approach.

distribution.

Algorithm 1. The following is the pseudo code for the EHO algorithm.

1.	Initialization	
2.	Set generation counter p,set the maximum generation *MaximGene*	
3.	Initialize the population	
4.	Repeat	
5.		Sort all the elephant according to their fitness
6.		for all clans l_i in the population do
7.		For all elephant k in the clan l_i do
8.		Update $e_{li,k}$ and generate $e_{new}, l_{i,k}$ by Eq. (1)
9.		if $e_{li,k} = e_{li,k} = e_{best,li}$ then
10.		Update $e_{li,k}$ and generate $e_{new}, l_{i,k}$ by Eq. (2)
11.		end if
12.		end for
13.	end for	
14.	for all clan l_i in the population do	
15.	Replace the worst elephant in $clanl_i$ by Eq. (4)	
16.	end for	
17.	Evaluate population by newly updated positions	
18.	until $p < MaximGene$	
19.	Return the best found solution	

3.2. Teaching learning based optimization

The TLBO is like a school, where the teacher and the learners (the students) work together to teach and learn. By using the students' grades as the output, the teacher can find the best mix among the students' work. The teacher gives a test to see how well the students did on their work. Since a teacher is usually seen as someone who knows a lot and shares that information with their students, their effectiveness has an effect on how well their students do. A good teacher, of course, helps students learn so they can get better grades or test scores [81].

TLBO simulates the knowledge transfer and collaboration that occurs between a teacher and students in a classroom setting. The algorithm maintains a population of candidate solutions called "students" and uses a teacher to guide their learning process [82]. The teacher represents the best solution found so far in the population [83].

Algorithm 2. The key steps involved in the TLBO algorithm:

Step 01	Initialization phase
Step 02	Performance Evaluation -Using fitness function calculate values for each individual
Step 03	Select best solution among all
Step 04	Calculate mean
Step 05	Teacher phase $y^{new} = y^{old} + rand(0,1) \bullet (y^{teacher} - EF \cdot y^{m})$
Step 06	Evaluate new solutions
Step 07	Compare new solution with old solutions and update
Step 08	Learner Phase $y^{new} = \begin{cases} y^{old} + rand(0,1) \bullet (y^{rand} - y^{old}) & if\ f(y^{r}) < f(y^{old}) \\ y^{old} + rand(0,1) \bullet (y^{old} - y^{rand}) & otherwise \end{cases}$
Step 09	Update current best solution
Step 10	Check stopping criteria
Step 11	Display best solution

3.3. Hybrid (HEHO) of elephant herd optimization (EHO) and Teaching Learning Based Optimization (TLBO)

A hybrid algorithm combining Hybrid Elephant Herd Optimization (HEHO) and Teaching Learning Based Optimization (TLBO) can leverage the strengths of both algorithms to enhance the optimization process. Here's a general outline of how a hybrid algorithm could be constructed:

Step 1: Initialization: Generate an initial population of candidate solutions (students) randomly, representing the elephant herds.

Step 2: Evaluation: Assess the fitness of each student (elephant herd) based on the objective function being optimized.

Step 3: Leader Selection: Identify the leader elephant herd with the highest fitness value.

Step 4: Movement and Interaction (HEHO component): Each elephant herd updates its position based on the position of the leader herd and interacts with neighboring herds to share information, as performed in HEHO. This step allows the herds to explore the solution space effectively and exchange valuable information.

Step 5: Learning Phase (TLBO component): Each student learns from the teacher by adjusting its position based on a learning equation, similar to the TLBO algorithm. This phase emphasizes cooperation and collaboration among the students to improve their solutions and explore the solution space.

Step 6: Local Search (HEHO component): Perform local search within each herd to further refine the solutions in their respective regions.

Step 7: Update Phase: Evaluate the fitness of the new positions of students (elephant herds) and update the population accordingly.

Step 8: Termination: Repeat the movement, interaction, learning, local search, and update phases until a termination condition is met.

By combining the movement and interaction mechanisms from HEHO and the learning and collaboration aspects of TLBO, the hybrid algorithm can leverage the advantages of both algorithms. The herds can benefit from the collective intelligence and exploration capabilities of HEHO, while also incorporating the knowledge transfer and learning from the teacher-student interaction in TLBO.

The time complexity of TLBO is analyzed in terms of the number of iterations required to reach convergence or find an acceptable solution within a given problem space. It is commonly considered to have a time complexity of O(GNP), where G is the number of iterations or generations; N is the population size and P is the problem dimensionality (number of decision variables).

The time complexity of the Elephant Herding Optimization (EHO) algorithm is O(T $\times$ NP $\times$ D), where T is the number of iterations, NP is the population size, and D is the dimension of the problem 1. This means that the time required to solve a problem using EHO is proportional to the number of iterations, population size, and dimension of the problem.

The overall computational complexity of Hybrid of TLBO algorithm and EHO is O(Max Iteration $\times$ N^3 + computational overhead). The complexity depends on maximum iteration, elimination and dispersion, reproduction, fitness function computation and elimination, if any, additionally including computational overhead like synchronization, data exchange, or decision-making mechanisms between the algorithms.

3.4. Real life applicability of the proposed approach

Cancer ranks number one or second in global mortality, according to 2019 World Health Organization statistics. In 2020, 10 million people died from cancer, and female breast cancer (11.7%) outnumbered lung cancer (11.4%). According to reports, breast cancer kills the most women. Nearly 25% of breast cancer fatalities occur in women aged 40–49. Although infrequent before 25–30, this malignancy has been reported in young women. The World Health Organization reports that 1 in 8–10 women have breast cancer. One in 10–15 Iranian women get this cancer. Researchers say late diagnosis causes breast cancer's high fatality rate. We have selected the breast cancer as the subject of investigation as breast cancer is the most common cancer among women such that the existence of a precise and reliable system for the diagnosis of benign or malignant tumours is critical. Since breast cancer is the most common malignancy in women, a precise and reliable technique for diagnosing benign or malignant tumours is essential. It is the most frequent cancer in women, is varied and caused by inherited and environmental risk factors. It arises when breast tissues grow abnormally. Breast tissue cells proliferate irregularly, forming benign or malignant tumours. DNA mutations induce breast cell proliferation abnormalities. This categorization (benign or malignant) is used to analyze breast tumours, masses, and other abnormalities. Benign cancer is usually not life-threatening and has a better survival rate than malignant cancer. Malignant tumours can quickly invade the lymph system and surrounding healthy tissues, having devastating repercussions. Benign tumours cannot grow beyond a certain size and stay inside their mass. Early cancer detection ensures treatment success and survival. Since there are few risk factors for breast cancer, scientists have tried to find its cause. Age, genetics, family history, obesity, gene variation, smoking, and alcohol intake increase breast cancer risk. The exterior of breast cancer cells is so tiny that early detection is difficult.

Since there are no breast cancer prevention methods, scientists have focused on developing better treatments. Screening for breast cancer in women has also been prioritized to save more lives through treatment. At least 30 years, researchers have studied mammography, clinical breast examination, and biopsy for breast cancer screening. Mammography is one of the best ways to screen for breast cancer, although radiologists' judgements differ. Surgical biopsy is more accurate than mammography but expensive and intrusive. Thus, better breast cancer detection systems are needed. These approaches help identify patients as benign or noncancerous or malignant or cancerous. Breast cancer patients survive better with early detection. Clinicians will need highly predictive and reliable diagnostic methods to distinguish benign and malignant breast tumours. Mammography, ultrasound, dynamic magnetic resonance imaging, and elastography are the only early cancer detection methods. Clinicians must often read a lot of imaging data, which reduces accuracy. This approach takes a long time and may misdiagnose cancer. Medical personnel keep making this diagnosis to determine which one has the biggest impact. Precision in cancer detection and early diagnosis is now possible with Fine Needle Aspiration cytology and machine learning.

Several breast cancer prediction studies use machine learning and data mining. Some improve learning models, others pre-process data. Others work on feature selection to find relevant dataset features for a better classification system. Feature selection methods include filters, wrappers, and embedding. Wrapper-based feature selection was used in this investigation. Subset combinations about the prediction model's accuracy are repeatedly evaluated to discover optimal features via the wrapper technique. Though computationally expensive, it is accurate. Recently, machine learning, deep learning, and bio-inspired computers have been used in medical diagnostics. Several studies have examined machine learning for breast cancer detection. Multiple research use UCI repository datasets to clinically forecast this condition. Examples include WDBC dataset. No matter the dataset, the research aims to improve prediction accuracy to appropriately diagnose cancer.

Data storage has expanded due to the rapid growth of internet, IOT, and RFID technology. Given the expanding amount of data

processed by application systems integrated inside devices that are internet-accessible, saving the data is essential. Clearing and extracting appropriate information and feature selection approaches are becoming more critical. Feature selection decreases running time by eliminating unnecessary and redundant information, boosting classification accuracy, and simplifying learnt classifiers or models. With several features, feature selection is difficult. Complex categorization problems involve several features. Thus, the classifier classifies observations across time. Features are selected to minimize dataset dimensionality, maximize classification accuracy, and prevent overfitting to increase wireless sensor network efficiency and energy consumption and extend network lifetime. The main challenge in feature selection difficulties is discarding some of the preprocessed data without affecting quality. Many methods have been developed for feature selection. These algorithms had a hefty computing cost when they were introduced 30 years ago. This challenge was solved by fast computers and large storage resources, but creating a fast solution to deliver this function is still relevant due to new challenges' enormous data sets. Classification is a key machine learning approach. In certain problems, data sets are so high-dimensional that all aspects are unimportant, reducing algorithm accuracy and performance. In this case, feature selection will be crucial, and removing unrelated features will boost algorithm efficiency. This paper proposes a hybrid feature subset selection strategy using evolutionary-based EHO and TLBO. The given method seeks to simplify calculation and search time for optimal solutions to high-dimensional datasets feature selection issue.

One of the most popular machine learning strategies is feature selection(FS). Features are reduced to improve learning algorithm performance by removing extraneous features and choosing the best primary features. In machine learning, feature construction and selection are key characteristics. Since both factors require manual fabrication, they are time-consuming and complicated. Raw data attributes are aggregated, combined, or separated to create characteristics. A computationally cost-effective all-around search for the most features is impossible. Thus, FS is a fundamental pattern identification and machine learning difficulty. This strategy is important in categorization and regression, where numerous features are worthless or diminish learning accuracy. Eliminating these attributes reduces computational complexity and improves accuracy. Feature selection strategies choose a subset of primary features to decrease data dimensionality. These strategies aim to extract a subset with the smallest dimensions for the application. Classification analysis performs better on the reduced space than the main space in most circumstances. Feature selection methods determine the optimal subset of N features and 2^N subsets. A subset is selected to optimize an Evaluation Function in all of these approaches, depending on the application and definition. Each method seeks the most important attributes, however in big and medium datasets, finding the optimized response is challenging and expensive due to the high number of alternative responses.

A FS procedure begins with an exhaustive search through the subset of features to find the best feature among the primary probable subclasses based on a given assessment criterion. If the feature set has n features, the best subset must be selected using optimum feature selection. Since evolutionary computation approaches provide global search, they are used as a strong solution and alternative to standard searching methods to handle these challenges. Particle swarm optimization, genetic algorithms, genetic programming, and ant colony optimization are popular feature selection methods. Heuristic models use diverse tactics to find the tradeoff between exploration and exploitation. Exploration funds clearcut search spaces, whereas exploitation scans the local search space to maintain improved solutions. Some meta-heuristic search methods use exploration, while others use exploitation for superior results. Using hybrid approaches can improve search algorithm performance. The yield of each approach increases when hybridization combines positive qualities of at least two procedures. This study uses EHO and TLBO, two novel and effective meta-heuristics, to construct a hybrid strategy to improve general categorization tasks. A hybrid model for feature selection was proposed in this research. This methodology uses information to simplify feature selection with high-dimensional datasets to reduce computational complexity. Selecting dataset features is easier with a novel hybrid method. To extract dataset-influencing characteristics, an EHO, TLBO, and their combination was applied. EHO and TLBO are integrated to balance exploration and exploitation to build an algorithm without the previous deficiency by employing rapid convergence speed and exploration ability. Only features with a selection probability matching the final subset are expected to align with classification, and the feature selection procedure is only applied to these. Using prediction algorithms can yield good outcomes, according to prior authors. Thus, the algorithm's computational cost is greatly lowered and a subset with fewer features is chosen. The suggested approach accelerates feature selection for high-dimensional datasets, improves classification accuracy, and reduces feature selection. To our knowledge, no other study has used these three nature-inspired methods to reduce unproductive features. Effective feature selection is achieved with these cost-effective techniques. By exhibiting a very positive categorization performance outcome, the suggested clinical decision support system can be used as a secondary opinion, relieving the pressure on more experienced medical practitioners. Furthermore, this technique can be utilized as an early and effective prediction tool in rural areas or underdeveloped nations where there is a shortage of skilled medical professionals.

3.5. About dataset

The WDBC dataset includes data on 569 patients about 30 characteristics of cell nuclei taken from a digital image of a breast cancer fine needle aspirate (FNA). At diagnosis, each patient's cancer was categorized as benign or malignant. There are 569 data points in the dataset overall, of which 212 are categorized as malignant and 357 as benign. These are the features of the dataset: The ten qualities are as follows: Radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, fractal dimension, and so on are the first ten. Three pieces of information are included in each feature: (1) the mean; (2) the standard error; and (3) the "worst" or largest value (the mean of the three largest values). having a total of 30 attributes in the dataset as a result. A total of thirty features were obtained by calculating the mean, standard error, and maximum value (mean of the three highest values) of these features for each image. These characteristics outline the characteristics of the cell nuclei seen in the picture. Both the benign and malignant cases obtained through FNA were confirmed with surgical biopsy in order to ensure the excellent quality of the collected data. Additionally, the effectiveness of the Xcyt software was verified in a sample of 131 newly diagnosed cases (94 benign, 37 malignant), with an

impeccable accuracy rate of 100%.

4. Results and discussion

4.1. Results computed using TLBO

11 features, the matrix given below, have been retrieved using TLBO from the initial features set of 32 features WDBC dataset [84]. Table 1 depicts the performance of four implemented ML classifiers in terms of six vital efficiency measuring metrics. Figs. 2–5 represents the 4 results in graphical format (Accuracy, Confusion Matrix, AUC-ROC curve and Accuracy vs iteration graph) for XGBoost, KNN, LSTM and SBL classifiers respectively. 5-fold cross validation approach has been implemented in this work to overcome over-fitting issue.

['concave point mean','radius_mean', 'texture_mean', 'area_mean', 'concavity_mean', 'radius_worst','symmetry mean', 'perimeter_worst', 'compactness_worst', 'concavity_worst','perimeter mean']

4.2. Results computed using EHO

A total of 18 features, matrix given below, were obtained through the implementation of the EHO algorithm. These features were extracted from an initial set of 32 features in the WDBC dataset. Table 2 presents the performance of four machine learning classifiers that have been implemented, as measured by six important efficiency metrics. Figs. 6–9 depict the graphical representations of the four results, namely Accuracy, Confusion Matrix, AUC-ROC curve, and Accuracy vs iteration graph, for the XGBoost, KNN, LSTM, and SBL classifiers, respectively.

Features Extracted

["concavity_mean","concave_points_mean","symmetry_mean","fractal_dimension_mean","radius_se","texture_se","perimeter_se","area_se","area_worst","smoothness_worst","compactness_worst","concavity_worst","concave_points_worst","symmetry_worst","fractal_dimension_worst","perimeter_mean","area_mean","smoothness_mean"]

4.3. Results computed using hybrid algorithm (TLBO-EHO)

A total of 11 features have been extracted from the initial set of 32 features of the WDBC dataset using the TLBO algorithm. These features are presented in the matrix provided below. Table 3 illustrates the performance of four implemented machine learning classifiers in terms of six essential efficiency measuring metrics. Figs. 10–13 depict the graphical representation of four results, namely Accuracy, Confusion Matrix, AUC-ROC curve, and Accuracy vs iteration graph. These results correspond to the XGBoost, KNN, LSTM, and SBL classifiers, respectively. The implementation of a 5-fold cross-validation approach has been utilized in this study to address the problem of over-fitting.

Fetched features out: 11 ["concave point mean", "texture_mean", "perimeter_mean", "area_mean", "smoothness_mean", "compactness_mean","concavity_mean","perimeter_worst","compactness_worst","concavity_worst", "perimeter mean"].

4.4. Discussion

A well-structured diagnostic methodology has the ability to accurately distinguish between individuals who have a specific ailment and those who do not. Values that surpass the predetermined threshold for a thorough examination consistently indicate the presence of the ailment, whereas values that fall below the threshold consistently exclude the possibility of the ailment. At present, there is no available test that can effectively distinguish between individuals with a disease and those without. As a result, the diagnostic procedures provide only a limited distinction between the mentioned groups of individuals. The occurrence of values exceeding the established threshold does not always indicate the presence of a pathological condition, as individuals without any disease can occasionally display higher values as well. Values that exceed predetermined thresholds for a specific parameter of interest are commonly referred to as false positive values (FP). On the other hand, values that fall below the established threshold are primarily observed in individuals who do not display the disease. However, individuals suffering from the condition may also experience these symptoms. The values correspond to the false negative (FN) values. The implementation of a cut-off value is utilized to categorize the population of individuals under study into four distinct subgroups based on the presence or absence of the disease while taking into account the

Table 1
Results generated though features retrieved from TLBO.

Classifier Vs Performance Metrics	Accuracy	Precision	F1-Score	Specificity	Sensitivity
XGBoost	0.9752	0.9615	0.9591	0.9664	0.9732
KNN	0.9358	0.9456	0.9318	0.9553	0.9039
LSTM	0.9763	0.9544	0.9410	0.9825	0.9455
SBL	0.9396	0.9499	0.9308	0.9565	0.9206

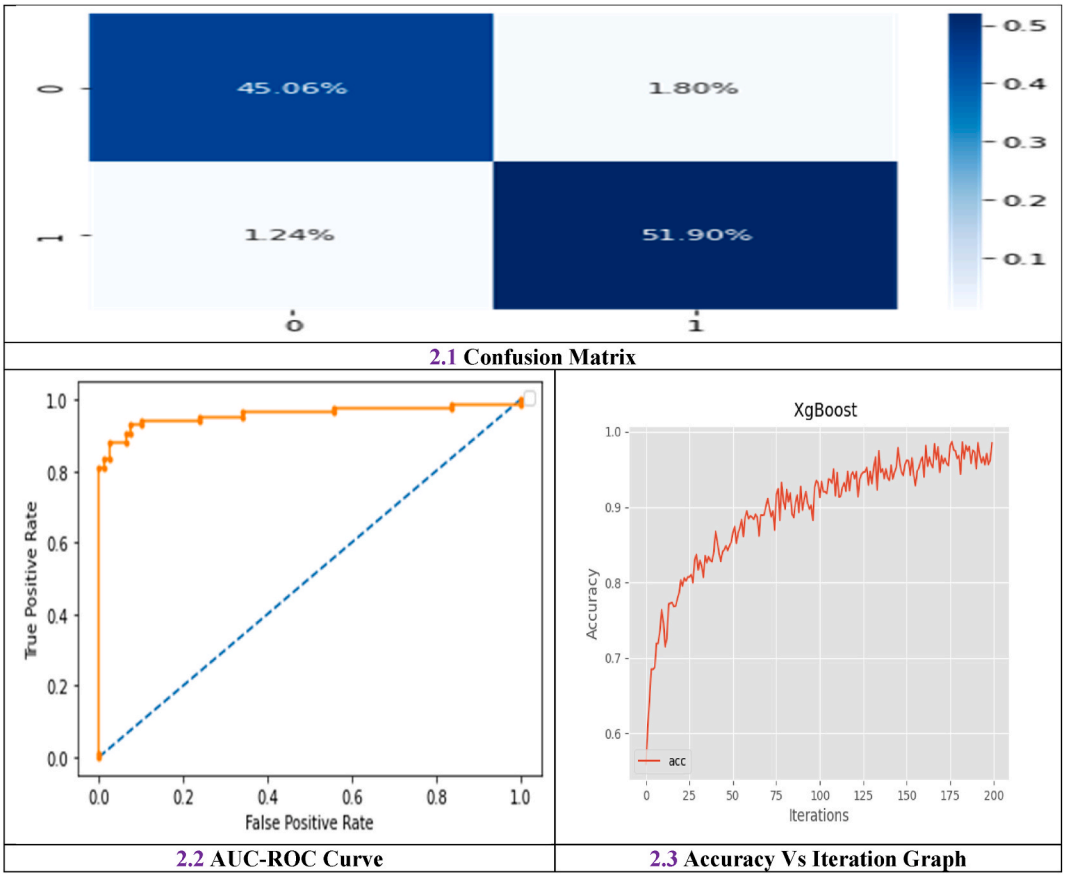


Fig. 2. Results generated through XGBoost classifier on TLBO selected features.

relevant parameter values of interest.

A true positive (TP) refers to individuals who have been diagnosed with the disease and have a parameter of interest that exceeds the specified threshold. The term "false positive" (FP) is employed to describe instances where individuals do not have the disease but display a parameter value that exceeds the specified threshold of significance. A true negative (TN) refers to individuals who do not have the disease and show a parameter value of interest that is lower than the specified cut-off threshold. A false negative (FN) occurs when individuals who have the disease show a parameter value of interest that falls below the specified threshold.

Accuracy (Equation (5)), a major key performance measuring metric, refers to the proportion of correct predictions made by the classification model. Accurate predictions are comprised of two components: true positives (TP) and true negatives (TN). Comprehensive consideration of both positive (P) and negative (N) examples is necessary when formulating predictions. Set P is comprised of true positives (TP) and false positives (FP), whereas set N comprises true negatives (TN) and false negatives (FN). Mathematically, it is represented as the number of correct predictions divided by the total number of predictions across all classes [86]. Our proposed approach is able to generate an auspicious accuracy of 97.96% (in combination with the EHO and KNN ML classifier). The F1-score is determined by computing the weighted mean of precision and recall. To provide greater precision, it can be asserted that the value in question denotes the harmonic mean of precision and recall. The analysis of the F1 score can prove to be a valuable strategy for tackling the tradeoff between precision and recall. Our proposed approach is able to generate a promising F-score accuracy of 0.9627 (in combination with our proposed hybrid algorithm and KNN ML classifier).

Sensitivity (Equation (6)) is a metric expressed as a percentage that denotes the proportion of true positive cases with a particular ailment within a cohort of subjects with the same ailment. Sensitivity refers to the probability of obtaining a positive test result among individuals who have a particular medical condition. This concerns the test's capacity to detect individuals who have the illness. By merging a hybrid algorithm with the SBL, the study's technique reaches a sensitivity level of 0.9800.

The measure of specificity (Equation (7)) serves as a complement to sensitivity in the context of diagnostic test accuracy. The term is defined as the proportion of individuals who test negative and do not have the disease, relative to the overall number of individuals who do not have the disease. The term "specificity" pertains to the probability of obtaining a negative test outcome in an individual who is not afflicted with the disease [85]. In the realm of medical testing, specificity pertains to the precision of a diagnostic examination in identifying individuals who are free of a particular ailment, thereby ruling out the targeted medical condition. The metric assesses the classifier's capacity to recognize individuals who do not manifest the medical condition. Specificity is related to the rate of true

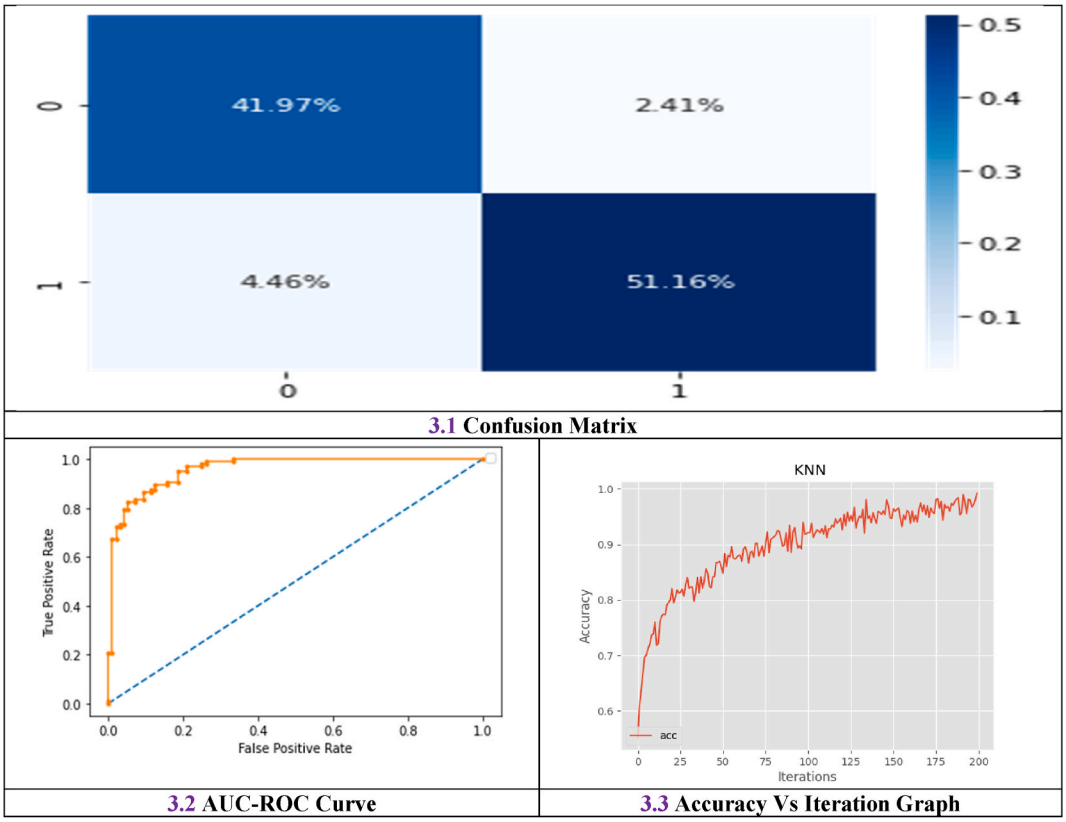


Fig. 3. Results generated through KNN classifier on TLBO selected features.

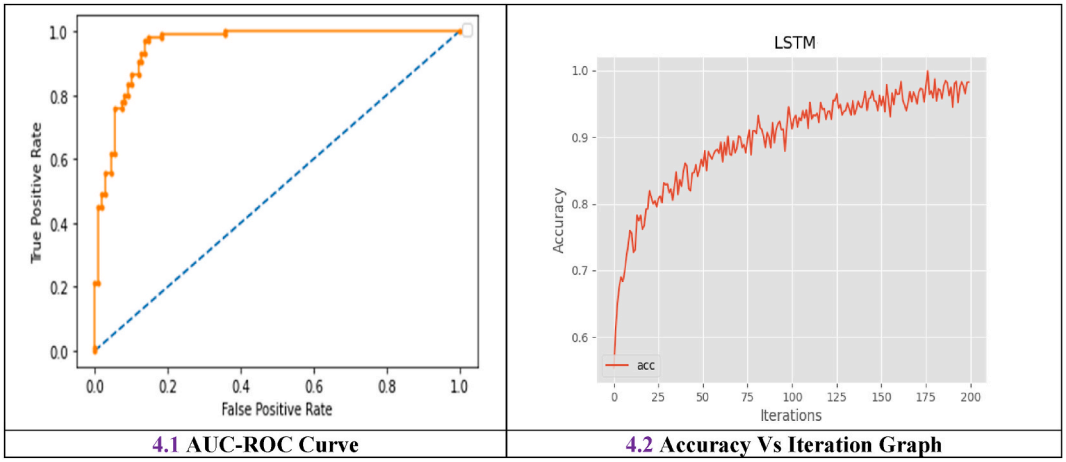
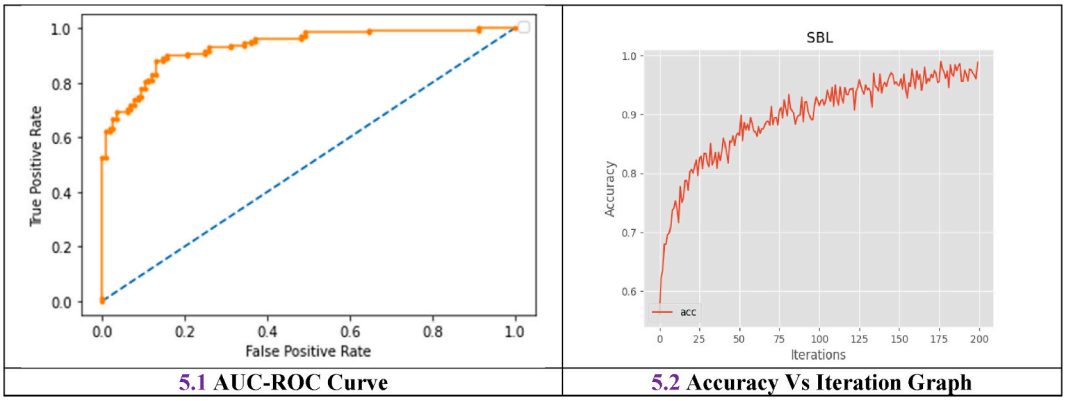


Fig. 4. Results generated through LSTM classifier on TLBO selected features.

negatives. Specificity measures the model's capacity to accurately identify negative instances. Within the realm of probability, the aforementioned statement pertains to the probability of a classifier accurately identifying a healthy patient as such. Within the realm of probabilities, the term "probability of accurate diagnosis by a classifier for a healthy patient" is commonly used. To calculate sensitivity and specificity, create a 2x2 table with subject groups divided by a gold standard or reference method in columns, and test categories in rows (refer to Table 4). The study's methodology achieves a specificity level of 0.9873 by combining a hybrid algorithm with the XGBoost ML classifier.

In order to assess the precision (Equation (8)) of a given test, it is necessary to compute the ratio of correctly identified positive and negative cases to the total number of cases evaluated. Precision, which is also referred to as positive predicted value, is a metric that

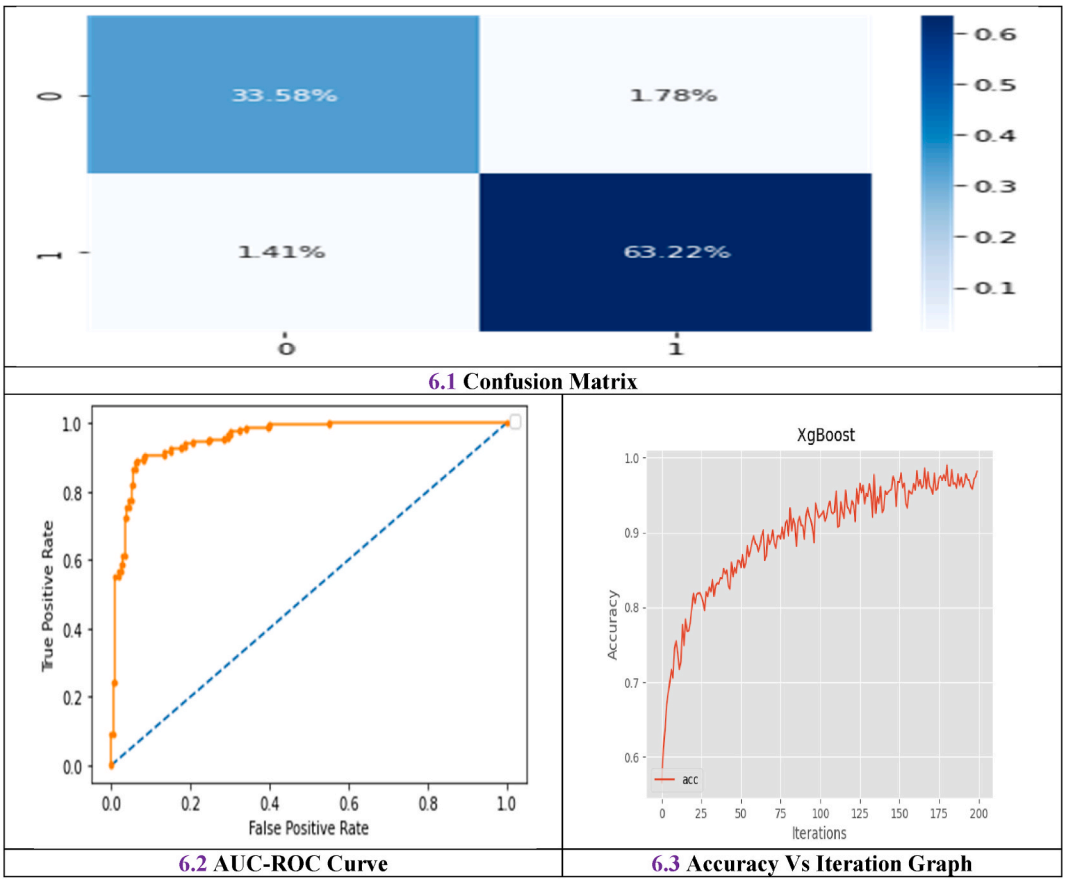


5.1 AUC-ROC Curve | 5.2 Accuracy Vs Iteration Graph

Fig. 5. Results generated through SBL on TLBO selected features.

Table 2
Results generated though features retrieved from EHO.

Classifier Vs Performance Metrics	Accuracy	Precision	F1-Score	Specificity	Sensitivity
XGBoost	0.9680	0.9496	0.9549	0.9726	0.9594
KNN	0.9796	0.9841	0.9601	0.9852	0.9302
LSTM	0.9403	0.8598	0.9206	0.9539	0.8998
SBL	0.9347	0.9168	0.9294	0.9307	0.9397



6.1 Confusion Matrix

6.2 AUC-ROC Curve | 6.3 Accuracy Vs Iteration Graph

Fig. 6. Results generated through XGBoost on EHO selected features.

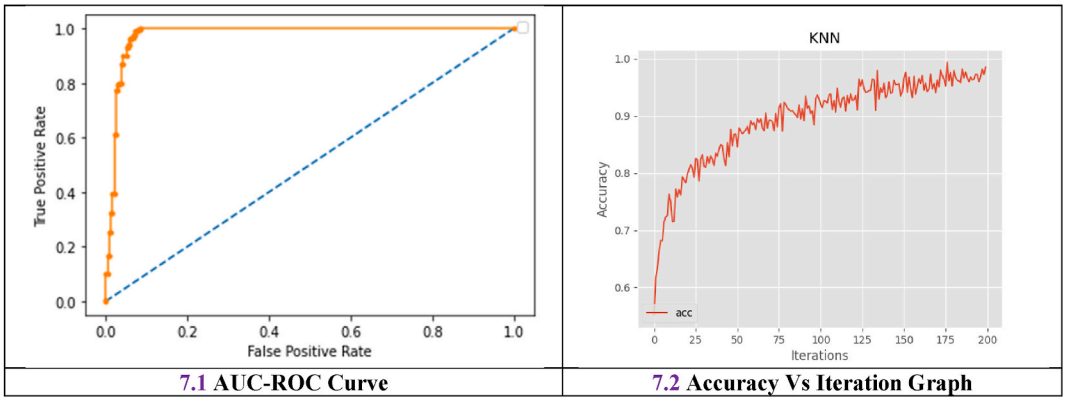


Fig. 7. Results generated through KNN on EHO selected features.

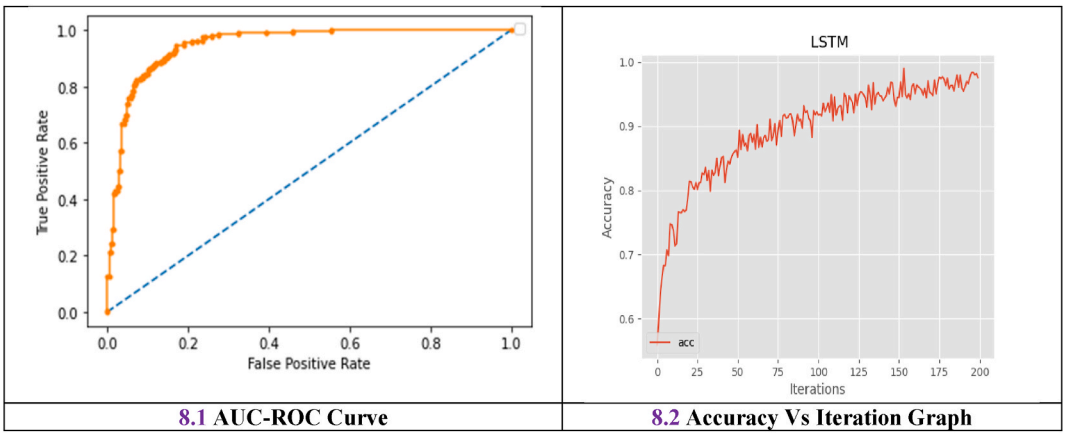


Fig. 8. Results generated through LSTM on EHO selected features.

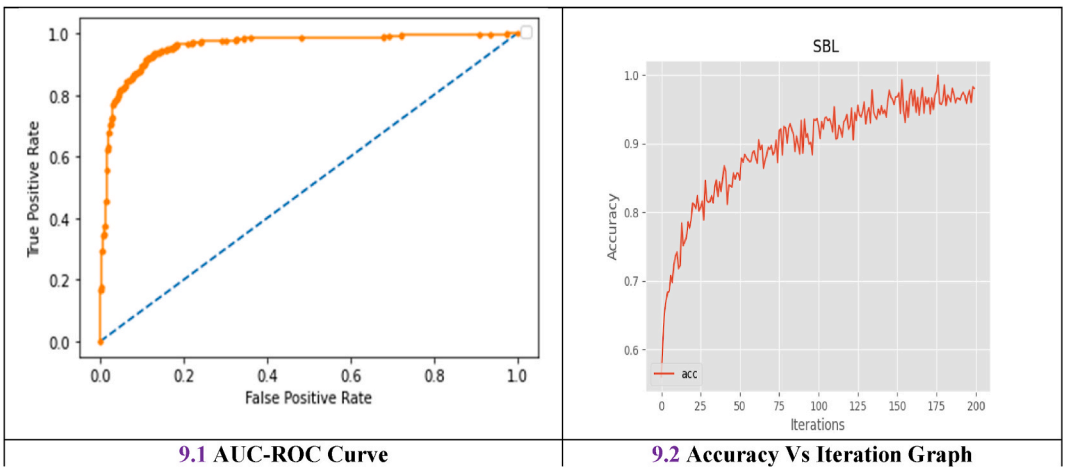
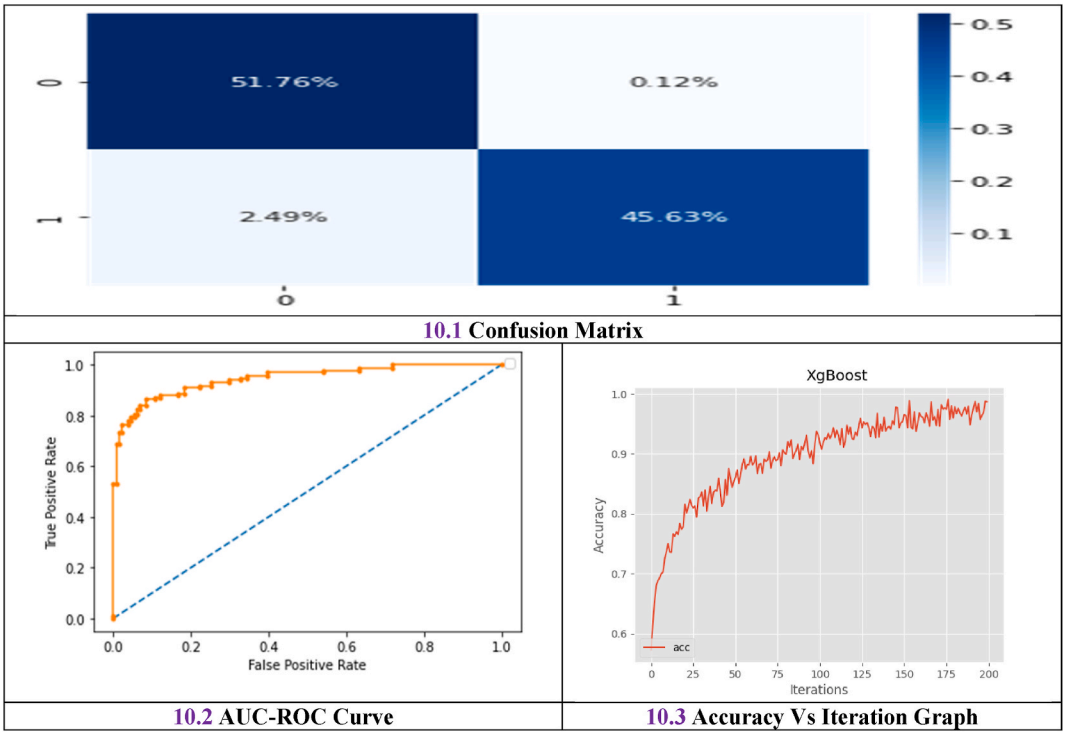


Fig. 9. Results generated through SBL on EHO selected features.

measures the proportion of patients correctly identified as having the disease by the classifier relative to the total number of patients identified by the classifier as having the disease. Expressed probabilistically, it can be stated as follows: "The probability of a patient being classified as sick by the classifier, given that the patient is truly sick." If precision = 1, it means that all patients diagnosed by the classifier really had the disease. The presented methodology exhibits a notable degree of precision, attaining a rate of 0.9876 through

Table 3
Results generated though features retrieved from Hybrid algorithm.

Classifier Vs Performance Metrics	Accuracy	Precision	F1-Score	Specificity	Sensitivity
XGBoost	0.9739	0.9876	0.9615	0.9873	0.9541
KNN	0.9492	0.9842	0.9626	0.9855	0.8956
LSTM	0.9682	0.9536	0.9590	0.9616	0.9762
SBL	0.9779	0.9783	0.9617	0.9754	0.9800

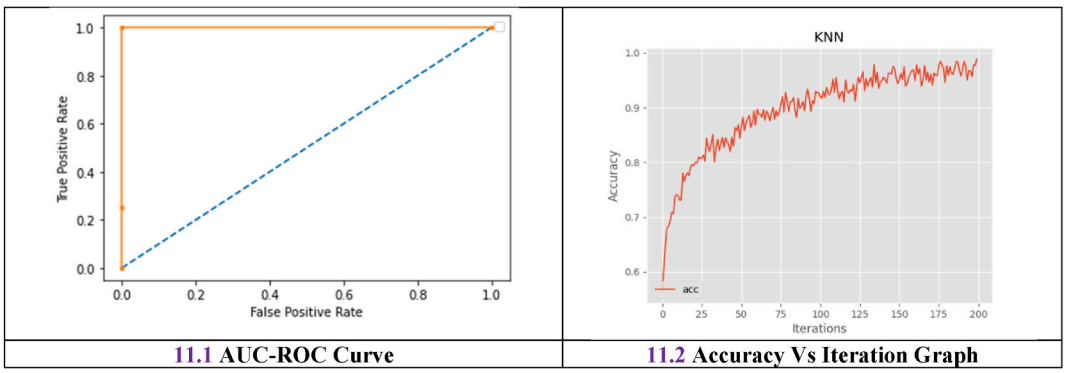


10.1 Confusion Matrix

10.2 AUC-ROC Curve

10.3 Accuracy Vs Iteration Graph

Fig. 10. Results generated through XGBoost on Hybrid-algorithm selected features.



11.1 AUC-ROC Curve

11.2 Accuracy Vs Iteration Graph

Fig. 11. Results generated through KNN on Hybrid-algorithm selected features.

the integration of hybrid with the XGBoost ML classifier.

$$A_{cc} = \frac{T_p + T_N}{T_p + T_N + F_P + F_N} \tag{5}$$

Accuracy is given by.
Sensitivity is given by

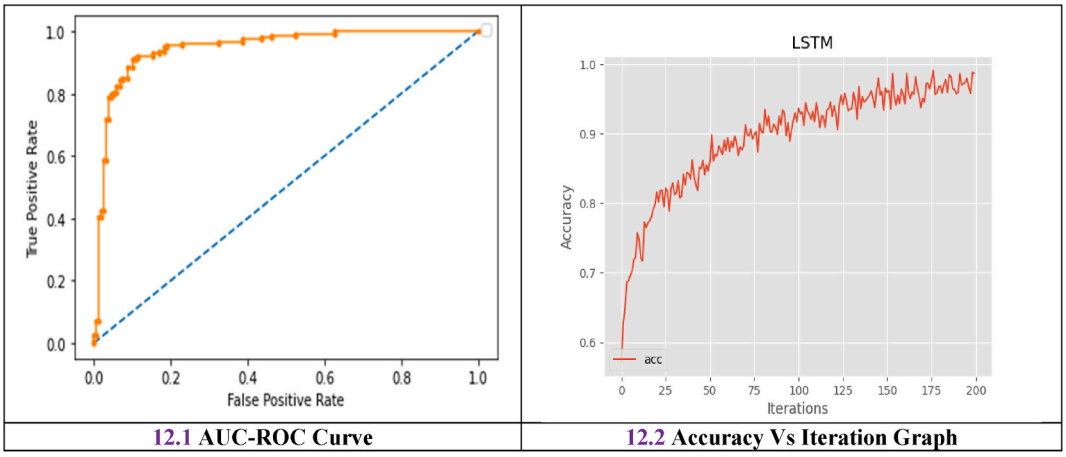


Fig. 12. Results generated through LSTM on Hybrid-algorithm selected features.

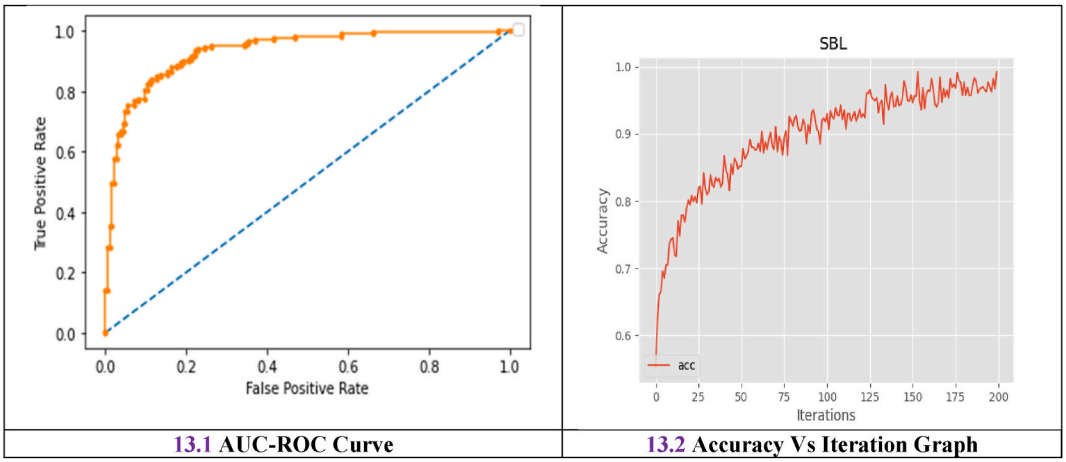


Fig. 13. Results generated through SBL on Hybrid-algorithm selected features.

Table 4
Table for computing parameters.

	Subjects with the disease	Subjects without the disease
positive	TP	FP
negative	FN	TN

$$S_{en} = \frac{T_P}{T_P + F_N} \tag{6}$$

Specificity is given by

$$S_{pe} = \frac{T_N}{T_N + F_P} \tag{7}$$

Precision is given by

$$\mathrm{P}_{re} = \frac{T_p}{T_P + F_P} \tag{8}$$

Every cut-off point is associated with a specific set of diagnostic sensitivity and specificity values. To construct a ROC curve, the values are plotted on a graph where the x-axis represents 1-specificity and the y-axis represents sensitivity, as illustrated in Fig. 14. The configuration of the ROC curve and the value of the AUC serve as an approximation of the discriminatory capability of a given test [87]. Tests that exhibit curves positioned nearer to the upper-left quadrant and greater areas beneath the curve demonstrate superior

discriminatory capacity in distinguishing between individuals with and without disease. The definite integral of the curve over the interval [0,1] is a dependable indicator of the effectiveness of the test. A diagnostic assessment that exhibits an AUC value of 1.0 is deemed ideal, whereas a test that lacks discriminatory power is characterized by an area of 0.5. The Area under the Curve (AUC) is a widely accepted and standardized measure utilized for evaluating the accuracy of diagnostic tests. By merging a hybrid algorithm with the KNN, the suggested approach has an AUC of 0.9901.

In the burgeoning fields of big data and artificial intelligence, researchers are progressively pursuing uncomplicated and effective models. There exist several practical applications that entail high-dimensional data, such as text mining [88], genomics [89], and image retrieval [90]. It is imperative to acknowledge that not all data holds relevance in problem-solving. The inclusion of irrelevant features may lead to a decrease in the performance of a model. Utilizing feature selection methods to eliminate irrelevant and redundant features from the data is a waste of time. The primary concept underlying feature selection is to streamline the acquired model by removing extraneous or duplicative features. The reduction of features results in a decrease in the inference time of the model and enhances the model's generalization ability [91].

Several research studies have demonstrated that feature selection is a viable technique for improving model performance and simplifying it [92]. The categorization of feature selection can be based on the solution methods employed, which include wrapper methods [93], filter methods [94], and embedded methods [95]. Despite the extensive research on feature selection, it continues to be a challenging task. Wrapper methods employ the performance of a classifier as the evaluation metric for subsets of features. The aforementioned evaluation method is characterized by its simplicity and precision, yielding subsets that are more refined compared to filtering methods. The process of evaluating each feature subset can be computationally expensive, especially for high-dimensional data, as it necessitates retraining a classifier. Therefore, this approach does not possess scalability. It is also possible for the training set to become too well-fitted when subsets are evaluated using classifier performance metrics [96]. Filter methods employ the correlation between the one-dimensional features of the data and the target variable for evaluating the significance of the features. This methodology is frequently beneficial in terms of optimizing time efficiency. However, the task of determining the most suitable number of features to select can be quite challenging in the absence of domain knowledge or extensive experimentation. Furthermore, this approach may lack the ability to detect interactions among multiple features [97]. Embedded methods are deemed to be more efficient than wrapper methods as they incorporate the feature ranking process within the classifier model training. However, ascertaining the ideal number of features to retain poses a challenge. It is apparent from the information provided that different FS algorithms exhibit unique strengths and weaknesses. The integration of diverse feature selection algorithms presents a viable solution to tackle the challenge of FS.

Evolutionary computation techniques are widely used in feature selection due to their exceptional global search capabilities, as demonstrated by their significant attention in previous studies [98–101]. The present paper proposes a strategy for expeditious evaluation that substantially reduces evaluation expenses and improves the efficacy of the feature selection procedure. In this particular instance, the original feature set consisting of 32 features was successfully reduced to 11 features, resulting in a reduction of 65% in the case of two algorithms. Additionally, a reduction of 75% was achieved through the use of the last algorithm, resulting in a final feature set of 8. Hence, our methodology has effectively accomplished the implementation of the feature selection process, resulting in a 75% reduction in the number of features.

4.5. Comparison with prior stare-of-the-art studies

The proposed system must be compared to other accepted models in order to be thoroughly examined. The results of 12 efficient approaches are shown in Table 5 and have been published in reputable academic journals. In the process of choosing comparative

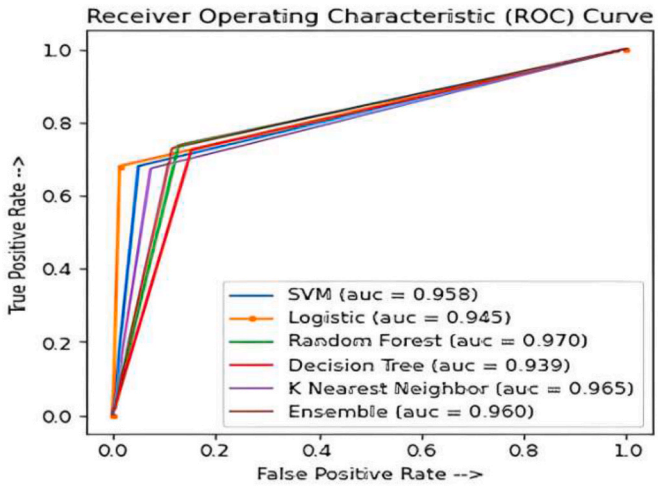


Fig. 14. Sample ROC curve (Relationship between the area under the ROC curve and diagnostic accuracy).

studies, high-quality studies that reflect the state of the art were given priority. The abbreviation "**UD**" stands for the authors' have not disclosed this information in their research study.

When reducing the original feature set, the authors in Ref. [102] used a statistically based FS approach. In this process, the most pertinent features are identified, and the machine learning models are trained and evaluated based on these features. 95.1691% was the highest accuracy recorded. We have computed some of the other key performance measuring parameters that were not computed in this study, so our results compare somewhat better to the results of this study. In Ref. [103], a novel intelligent diagnosis approach has been proposed that employed an information gain-directed simulated annealing genetic algorithm wrapper for FS. In this process, authors perform the ranking of features according to the information gain algorithm, and extracting the top-most optimal feature utilized the cost-sensitive support vector machine learning algorithm. The highest accuracy recorded was 95.8%. We have calculated additional key performance metrics that were not included in this study, which has resulted in a more favourable comparison of our results to those of the study.

The paradigm under consideration [104] employs a cooperative coevolutionary method wherein FS and instance selection (IS) are treated as distinct and separate subproblems. FS and IS techniques are employed to enhance the performance of a system by eliminating less pertinent features and instances, respectively. In this study, the wrapper strategy is employed for both feature and instance selection. The wrapper approach utilizes a combination of cooperative co-evolution and a random forest classifier. The reduced dataset was utilized for training a random forest classifier, and the resultant model played a crucial role in informing clinical decision-making. Our results in terms of accuracy, precision, specificity, and sensitivity are better than this approach.

Researchers in [105] describe a novel feature selection method that combines gradient-boosting decision trees and bee colony optimization. This strategy aims to address issues with the effectiveness and informativeness of the chosen characteristics. To find the most informative features, the suggested method performs global optimization of the decision tree's input variables using the bee colony algorithm. The feature space that the dataset encompasses must be initialized using the technique. According to how much each feature contributes to decision-making, less important features are suppressed using the artificial bee colony algorithm. The highest accuracy recorded was 97.9%. We have conducted an analysis of supplementary key performance metrics that were not originally incorporated in this study. As a result, our findings now present a more favourable comparison between our results and those of the study.

In this work [106], practitioners have combined a multivariate statistical approach with an artificial intelligence-based learning technique to create a prediction model. This paper suggests a hybrid feature selection technique that can be applied alongside artificial neural networks and PCA. The most important features are extracted from the data and preprocessed using principal component analysis. Our results in terms of accuracy, specificity, and sensitivity are better than this approach. This paper [107] proposes a hybrid optimization approach for the feature selection issue that combines particle swarm optimization and the slap swarm algorithm. By combining the two methods, a hybrid algorithm is produced that enhances the effectiveness of both the exploration and exploitation phases. The highest accuracy recorded was 97.9%; however, the best value of our accuracy is 97.96%. Precision, sensitivity, and specificity are not discussed in the study.

In the next study [108], research has led to the creation of a novel knowledge-based system for classifying breast cancer cases. To improve the precision and dependability of the system, our method combines techniques for clustering, noise reduction, and classification. As a clustering technique, the Expectation Maximization algorithm is frequently used to organize data into clusters that show similarities. The fuzzy rule-based reasoning technique uses classification and regression trees to build fuzzy rules for the classification of breast cancer disease in the knowledge-based system. The knowledge-based solution that has been suggested incorporates principal component analysis to solve the multi-collinearity issue. The highest accuracy recorded was 94.1%; however, the best value of our accuracy is 97.96%. Other key metrics like precision, sensitivity, and specificity are not discussed in the study. The purpose of [109] was to examine the impact of combining the FS algorithm with the classification algorithms on the prognosis of breast cancer. The authors suggested that by employing FS strategies to decrease the number of characteristics, they could enhance most classification systems. Compared to other aspects, some features are more significant and have a greater impact on the classification algorithms'

Table 5
Brief evaluation among the suggested modality and the state-of-the-art studies.

S.No.	Literature	Highest Accuracy	Highest F1-Score	Highest Precision	Highest Specificity	Highest Sensitivity
1.	Chaurasia & Pal [102]2020	95.1691	UD	UD	UD	UD
2.	Liu et al., [103] 2019	95.8	UD	UD	UD	UD
3.	Christo et al., [104] 2022	97.1	UD	0.967	0.975	0.967
4.	Rao et al., [105] 2019	97.9 (Highest) 92.8 (Average)	UD	UD	UD	UD
5.	Sahu et al., [106] 2019	97.0	UD	UD	0.98	0.95
6.	Ibrahim et al., [107] 2019	97.8	0.9834	UD	UD	UD
7.	Nilashi et al., [108] 2017	94.1	UD	UD	UD	UD
8.	Sakri et al., [109] 2018	81.3	0.876	0.883	0.632	0.869
9.	Dheeba et al., [110] 2014	93.67	UD	UD	0.92105	0.94167
10.	Ramadevi et al., [111] 2015	97.89	UD	UD	UD	UD
11.	Idris & Ismail [112] 2021	94.534	0.94137	0.9429	UD	0.9400
12.	Rajaguru & SR [113] 2019	95.61	UD	0.9726	0.9500	0.9595
13.	Ours (Highest)	97.96	0.9627	0.9876	0.9873	0.9800

outcomes. The outcomes of this research on three well-known classification algorithms—naïve Bayes, IBK, and REPTree—with and without the FS approach, particle swarm optimization, are shown. In conclusion, naive Bayes outperformed the other two approaches when employed with particle swarm optimization, while the former gave superior results both with and without particle swarm optimization. The result of our suggested approach performs very well in terms of all performance-measuring parameters.

In this paper [110], a particle swarm-optimized wavelet neural network classification technique for the detection of breast anomalies in digital mammograms is investigated. The suggested abnormality identification approach is predicated on taking Laws Texture Energy Measures out of the mammograms and using a pattern classifier to categorize the areas that are questionable. The highest accuracy recorded was 93.67%; however, the best value of our accuracy is 97.96%. Our sensitivity and specificity were also higher than those of their reported values. F1-score and precision values were not discussed in the study. PCA, a well-liked feature extraction technique, is tested in this work using several classifiers and a variety of breast cancer datasets [111]. The outcomes are totaled and examined. PCA was applied, and it was found that the classifiers' performance improved on some data sets, decreased on others, and stayed unchanged. Although the highest accuracy that was ever recorded was 97.89%, the best value that our accuracy has ever had is 97.96%. The study does not go into detail about other significant parameters such as accuracy, sensitivity, and specificity.

Next study presents the Fuzzy-ID3 (FID3) algorithm, a classification method for breast cancer screening that makes use of a fuzzy decision tree [112]. This work aims to alleviate the limitations of the ID3 algorithm; furthermore, it aims to improve decision trees' classification accuracy. Using the ID3 algorithm for decision tree learning, the FID3 approach combines the concepts of fuzzy systems and decision tree methods. The fuzzy database that is used for data fuzzification in the FID3 algorithm is designed using the FuzzyDBD methodology, an automated method for constructing fuzzy databases. It was used to build an existing fuzzy database before the fuzzy rule base was created. The fuzzified dataset was used in the FID3 approach, a fuzziness-infused version of the ID3 technique. The FID3 technique uses a simple inference process whereby new input instances are classified by directly pulling rules from the tree that has been generated. Although the highest accuracy that was ever recorded was 94.534%, the best value that our accuracy has ever had is 97.96%. The other key parameters were computed, but the performance of our suggested approach is better than that of the results presented in Ref. [112]. Following feature selection by PCA, the two machine learning methods are validated on the WDBC dataset [113]. The decision tree and the K-Nearest Neighbor method are two machine learning algorithms that are compared using common performance criteria. Although the highest accuracy that was ever recorded was 95.61%, the best value that our accuracy has ever had is 97.96%. The remaining key parameters have been calculated, and our proposed approach demonstrates superior performance compared to the results presented in Ref. [113].

5. Conclusion

Over the past decade, breast cancer (BC) has become a serious medical condition with a significant mortality rate. Among women, BC is the most commonly diagnosed cancer. This research recommends a new clinical support classification system for the intelligent diagnosis of BC. To evaluate the effectiveness of the proposed method, this paper performs experiments on the BC benchmark dataset (WDBC). The results indicate that the proposed approach computes highly auspicious results in terms of many statistically significant measuring parameters. Although classification performance and fitness have improved significantly, there is still room for enhancement when working with datasets that have fewer features, especially in real-world scenarios. The proposed method has the potential to be applied to a broader range of practical applications, including disease diagnosis, prediction, and engineering optimization problems. This could help to tackle real-life challenges.

Funding

None.

Ethical approval

Not required.

Data availability statement

Also same data is also freely available and easily downloadable from the internet repository (UCI).

This work is not funded by any external, internal or any Government agency.

CRediT authorship contribution statement

Munish khanna: Writing – review & editing, Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Law Kumar Singh:** Visualization, Validation, Supervision, Software. **Kapil Shrivastava:** Visualization, Validation, Supervision, Software. **Rekha singh:** Supervision, Software, Resources, Project administration.

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.

References

[1] L. Abualigah, A. Diabat, Chaotic binary group search optimizer for feature selection, Expert Syst. Appl. 192 (2022) 116368.
[2] X.F. Song, Y. Zhang, Y.N. Guo, X.Y. Sun, Y.L. Wang, Variable-size cooperative coevolutionary particle swarm optimization for feature selection on high-dimensional data, IEEE Trans. Evol. Comput. 24 (5) (2020) 882–895.
[3] R. Abu Khurma, I. Aljarah, A. Sharieh, M. Abd Elaziz, R. Damaševičius, T. Krilavičius, A review of the modification strategies of the nature inspired algorithms for feature selection problem, Mathematics 10 (3) (2022) 464.
[4] L. Abualigah, A. Diabat, P. Sumari, A.H. Gandomi, Applications, deployments, and integration of internet of drones (iod): a review, IEEE Sensor. J. 21 (22) (2021) 25532–25546.
[5] M.A. Tawhid, A.M. Ibrahim, Feature selection based on rough set approach, wrapper approach, and binary whale optimization algorithm, International Journal of Machine Learning and Cybernetics 11 (3) (2020) 573–602.
[6] P. Agrawal, T. Ganesh, A.W. Mohamed, A novel binary gaining–sharing knowledge-based optimization algorithm for feature selection, Neural Comput. Appl. 33 (11) (2021) 5989–6008.
[7] M. Rostami, K. Berahmand, S. Forouzandeh, A novel community detection based genetic algorithm for feature selection, Journal of Big Data 8 (1) (2021) 1–27.
[8] X.L. Deng, Y.Q. Li, J. Weng, J.L. Zhang, Feature selection for text classification: a review, Multimed. Tool. Appl. 78 (3) (2019) 3797–3816.
[9] M. Sharma, P. Kaur, A comprehensive analysis of nature-inspired meta-heuristic techniques for feature selection problem, Arch. Comput. Methods Eng. 28 (3) (2021) 1103–1127.
[10] W.P. Ma, X.B. Zhou, H. Zhu, L.W. Li, L.C. Jiao, A two-stage hybrid ant colony optimization for high-dimensional feature selection, Pattern Recogn. 116 (1) (2021) 107933.
[11] L. Abualigah, Group search optimizer: a nature-inspired meta-heuristic optimization algorithm with its results, variants, and applications, Neural Comput. Appl. 33 (7) (2021) 2949–2972.
[12] J. Yang, V. Honavar, Feature subset selection using a genetic algorithm, in: H. Liu, H. Motoda (Eds.), Feature Extraction, Construction and Selection, Springer, 1998, pp. 117–136.
[13] J. Kennedy, R. Eberhart, Particle swarm optimization, Proceedings of ICNN'95-international conference on neural networks 4 (1995) 1942–1948.
[14] S. Mirjalili, S.M. Mirjalili, A. Lewis, Grey wolf optimizer, Adv. Eng. Software 69 (2014) 46–61.
[15] S. Mirjalili, A. Lewis, The whale optimization algorithm, Adv. Eng. Software 95 (2016) 51–67.
[16] S. Arora, S. Singh, Butterfly optimization algorithm: a novel approach for global optimization, Soft Comput. 23 (3) (2019) 715–734.
[17] S. Saremi, S. Mirjalili, A. Lewis, Grasshopper optimisation algorithm: theory and application, Adv. Eng. Software 105 (2017) 30–47.
[18] L. Abualigah, M. Abd Elaziz, P. Sumari, Z.W. Geem, A.H. Gandomi, Reptile search algorithm (RSA): a nature-inspired meta-heuristic optimizer, Expert Syst. Appl. 191 (2022) 116158.
[19] J.O. Agushaka, A.E. Ezugwu, L. Abualigah, Dwarf mongoose optimization algorithm, Comput. Methods Appl. Mech. Eng. 391 (2022) 114570.
[20] O.N. Oyelade, A.E.S. Ezugwu, T.I. Mohamed, L. Abualigah, Ebola optimization search algorithm: a new nature-inspired metaheuristic optimization algorithm, IEEE Access 10 (2022) 16150–16177.
[21] L. Abualigah, A. Diabat, S. Mirjalili, M. Abd Elaziz, A.H. Gandomi, The arithmetic optimization algorithm, Comput. Methods Appl. Mech. Eng. 376 (2021) 113609.
[22] L. Abualigah, D. Yousri, M. Abd Elaziz, A.A. Ewees, M.A. Al-Qaness, A.H. Gandomi, Aquila optimizer: a novel meta-heuristic optimization algorithm, Comput. Ind. Eng. 157 (2021) 107250.
[23] C.E. DeSantis, F. Bray, J. Ferlay, J. Lortet-Tieulent, B.O. Anderson, A. Jemal, International variation in female breast cancer incidence and mortality rates, Cancer epidemiology, biomarkers & prevention 24 (10) (2015) 1495–1506. https://www.wcrf.org/int/cancer-facts-figures/data-specific-cancers/breast-cancer-statistics.Accessed16January2015.
[24] C. Yu, H. Chen, Y. Li, Y. Peng, J. Li, F. Yang, Breast cancer classification in pathological images based on hybrid features, Multimed. Tool. Appl. 78 (2019) 21325–21345.
[25] H. Abdel-Razeq, F. Attiga, A. Mansour, Cancer care in Jordan, Hematology/oncology and stem cell therapy 8 (2) (2015) 64–70.
[26] A. Alzu'bi, H. Najadat, W. Doulat, O. Al-Shari, L. Zhou, Predicting the recurrence of breast cancer using machine learning algorithms, Multimed. Tool. Appl. 80 (2021) 13787–13800.
[27] A.G. Waks, E.P. Winer, Breast cancer treatment: a review, JAMA 321 (3) (2019) 288–300.
[28] P. Boyle, Triple-negative breast cancer: epidemiological considerations and recommendations, Ann. Oncol. 23 (2012) vi7–vi12.
[29] P. Kumar, R. Aggarwal, An overview of triple-negative breast cancer, Arch. Gynecol. Obstet. 293 (2016) 247–269.
[30] M. Sharma, J.D. Sharma, A. Sarma, S. Ahmed, A.C. Kataki, R. Saxena, D. Sharma, Triple negative breast cancer in people of North East India: critical insights gained at a regional cancer centre, Asian Pac. J. Cancer Prev. APJCP 15 (11) (2014) 4507–4511.
[31] J. Ferlay, I. Soerjomataram, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, F. Bray, Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012, Int. J. Cancer 136 (5) (2015) E359–E386.
[32] M. Akhtar, S. Dasgupta, M. Rangwala, Triple negative breast cancer: an Indian perspective, Breast Cancer 7 (2015) 239.
[33] D.C. Doval, A. Sharma, R. Sinha, K. Kumar, A.K. Dewan, H. Chaturvedi, A. Mehta, Immunohistochemical profile of breast cancer patients at a tertiary care hospital in New Delhi, India, Asian Pac. J. Cancer Prev. APJCP 16 (12) (2015) 4959–4964.
[34] R. Königsberg, G. Pfeiler, T. Klement, N. Hammerschmid, A. Brunner, R. Zeillinger, C. Dittrich, Tumor characteristics and recurrence patterns in triple negative breast cancer: a comparison between younger (< 65) and elderly (≥ 65) patients, Eur. J. Cancer 48 (16) (2012) 2962–2968.
[35] J.A. Sparano, R.J. Gray, D.F. Makower, K.I. Pritchard, K.S. Albain, D.F. Hayes, G.W. Sledge, Prospective validation of a 21-gene expression assay in breast cancer, N. Engl. J. Med. 373 (21) (2015) 2005–2014.
[36] K.K. Thakur, D. Bordoloi, A.B. Kunnumakkara, Alarming burden of triple-negative breast cancer in India, Clin. Breast Cancer 18 (3) (2018) e393–e399.
[37] A. Bhardwaj, A. Tiwari, Breast cancer diagnosis using genetically optimized neural network model, Expert Syst. Appl. 42 (10) (2015) 4611–4620.
[38] V.S. Kompalli, U.R. Kuruba, Combined effect of soft computing methods in classification, in: Proceedings of the First International Conference on Computational Intelligence and Informatics: ICCII 2016, Springer Singapore, 2017, pp. 501–509.
[39] H. Lu, H. Wang, S.W. Yoon, A dynamic gradient boosting machine using genetic optimizer for practical breast cancer prognosis, Expert Syst. Appl. 116 (2019) 340–350.
[40] B. Ma, Y. Xia, A tribe competition-based genetic algorithm for feature selection in pattern classification, Appl. Soft Comput. 58 (2017) 328–338.
[41] M. Pota, M. Esposito, G. De Pietro, Designing rule-based fuzzy systems for classification in medicine, Knowl. Base Syst. 124 (2017) 105–132.
[42] S.K. Nayak, P.K. Rout, A.K. Jagadev, T. Swarnkar, Elitism based multi-objective differential evolution for feature selection: a filter approach with an efficient redundancy measure, Journal of King Saud University-Computer and Information Sciences 32 (2) (2020) 174–187.
[43] F. Shoeleh, M. Asadpour, Graph based skill acquisition and transfer learning for continuous reinforcement learning domains, Pattern Recogn. Lett. 87 (2017) 104–116.
[44] C.H.E.N. Liangjun, P. Honeine, Q.U. Hua, Z. Jihong, S. Xia, Correntropy-based robust multilayer extreme learning machines, Pattern Recogn. 84 (2018) 357–370.
[45] P.H. Kassani, A.B.J. Teoh, E. Kim, Sparse pseudoinverse incremental extreme learning machine, Neurocomputing 287 (2018) 128–142.
[46] M. Pota, M. Esposito, G. De Pietro, Likelihood-fuzzy analysis: from data, through statistics, to interpretable fuzzy classifiers, Int. J. Approx. Reason. 93 (2018) 88–102.

[47] A. Ed-daoudy, K. Maalmi, Breast cancer classification with reduced feature set using association rules and support vector machine, Network Modeling Analysis in Health Informatics and Bioinformatics 9 (2020) 1–10.
[48] S.A. Mohammed, S. Darrab, S.A. Noaman, G. Saake, Analysis of breast cancer detection using different machine learning techniques, in: Data Mining and Big Data: 5th International Conference, DMBD 2020, Belgrade, Serbia, July 14–20, 2020, Proceedings, Springer Singapore, 2020, pp. 108–117, 5.
[49] Z. Fu, D. Zhang, X. Zhao, X. Li, Adaboost algorithm with floating threshold, in: International Conference on Automatic Control and Artificial Intelligence (ACAI 2012), IET, 2012, March, pp. 349–354.
[50] Y. Prasad, K.K. Biswas, C.K. Jain, SVM classifier based feature selection using GA, ACO and PSO for siRNA design, in: Advances in Swarm Intelligence: First International Conference, ICSI 2010, Springer Berlin Heidelberg, Beijing, China, 2010, pp. 307–314. June 12-15, 2010, Proceedings, Part II 1.
[51] B. Zheng, S.W. Yoon, S.S. Lam, Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms, Expert Syst. Appl. 41 (4) (2014) 1476–1482.
[52] I. De Falco, A. Della Cioppa, E. Tarantino, Facing classification problems with particle swarm optimization, Appl. Soft Comput. 7 (3) (2007) 652–658.
[53] R. Sheikhpour, M.A. Sarram, R. Sheikhpour, Particle swarm optimization for bandwidth determination and feature selection of kernel density estimation based classifiers in diagnosis of breast cancer, Appl. Soft Comput. 40 (2016) 113–131.
[54] L. Peng, W. Chen, W. Zhou, F. Li, J. Yang, J. Zhang, An immune-inspired semi-supervised algorithm for breast cancer diagnosis, Comput. Methods Progr. Biomed. 134 (2016) 259–265.
[55] O.N. Oyelade, A.A. Obiniyi, S.B. Junaidu, S.A. Adewuyi, ST-ONCODIAG: a semantic rule-base approach to diagnosing breast cancer base on Wisconsin datasets, Inform. Med. Unlocked 10 (2018) 117–125.
[56] R. Jafari-Marandi, S. Davarzani, M.S. Gharibdousti, B.K. Smith, An optimum ANN-based breast cancer diagnosis: Bridging gaps between ANN learning and decision-making goals, Appl. Soft Comput. 72 (2018) 108–120.
[57] F. Li, J.M. Zurada, W. Wu, Smooth group L1/2 regularization for input layer of feedforward neural networks, Neurocomputing 314 (2018) 109–119.
[58] Z. Khandezamin, M. Naderan, M.J. Rashti, Detection and classification of breast cancer using logistic regression feature selection and GMDH classifier, J. Biomed. Inf. 111 (2020) 103591.
[59] L. Dora, S. Agrawal, R. Panda, A. Abraham, Optimal breast cancer classification using Gauss–Newton representation based algorithm, Expert Syst. Appl. 85 (2017) 134–145.
[60] N. Abbas, W. Shatanawi, K. Abodayeh, Computational analysis of MHD nonlinear radiation casson hybrid nanofluid flow at vertical stretching sheet, Symmetry 14 (7) (2022) 1494.
[61] N. Abbas, W. Shatanawi, A.M. Taqi, Thermodynamic study of radiative chemically reactive flow of induced MHD sutterby nanofluid over a nonlinear stretching cylinder, Alex. Eng. J. 70 (2023) 179–189.
[62] N. Abbas, K.U. Rehman, W. Shatanawi, A.A. Al-Eid, Theoretical study of non-Newtonian micropolar nanofluid flow over an exponentially stretching surface with free stream velocity, Adv. Mech. Eng. 14 (7) (2022), 16878132221107790.
[63] N. Abbas, W. Shatanawi, Heat and mass transfer of micropolar-casson nanofluid over vertical variable stretching riga sheet, Energies 15 (14) (2022) 4945.
[64] A. Sharma, P.K. Mishra, Performance analysis of machine learning based optimized feature selection approaches for breast cancer diagnosis, Int. J. Inf. Technol. (2022) 1–12.
[65] D. Jain, V. Singh, Diagnosis of breast cancer and diabetes using hybrid feature selection method, in: 2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC), IEEE, 2018, December, pp. 64–69.
[66] C. Henneges, D. Bullinger, R. Fux, N. Friese, H. Seeger, H. Neubauer, B. Kammerer, Prediction of breast cancer by profiling of urinary RNA metabolites using Support Vector Machine-based feature selection, BMC Cancer 9 (2009) 1–11.
[67] M.R. Mosavi, M. Khishe, M. Akbarisani, Neural network trained by biogeography-based optimizer with chaos for sonar data set classification, Wireless Pers. Commun. 95 (2017) 4623–4642.
[68] M. Kaveh, M. Khishe, M.R. Mosavi, Design and implementation of a neighborhood search biogeography-based optimization trainer for classifying sonar dataset using multi-layer perceptron neural network, Analog Integr. Circuits Signal Process. 100 (2019) 405–428.
[69] D. Biswas, S.K. Das, S. Roy, Dependence of the individual growth process upon allometric scaling exponents and other parameters, J. Biol. Syst. 16 (1) (2008) 151–163.
[70] M. Khishe, M.R. Mosavi, Classification of underwater acoustical dataset using neural network trained by Chimp Optimization Algorithm, Appl. Acoust. 157 (2020) 107005.
[71] T. Hu, M. Khishe, M. Mohammadi, G.R. Parvizi, S.H.T. Karim, T.A. Rashid, Real-time COVID-19 diagnosis from X-Ray images using deep CNN and extreme learning machines stabilized by chimp optimization algorithm, Biomed. Signal Process Control 68 (2021) 102764.
[72] C. Wu, M. Khishe, M. Mohammadi, S.H. Taher Karim, T.A. Rashid, Evolving deep convolutional neutral network by hybrid sine–cosine and extreme learning machine for real-time COVID19 diagnosis from X-ray images, Soft Comput. 27 (6) (2023) 3307–3326.
[73] A.S. Elkorany, Z.F. Elsharkawy, Automated optimized classification techniques for magnetic resonance brain images, Multimed. Tool. Appl. 79 (37–38) (2020) 27791–27814.
[74] M. Khishe, H. Mohammadi, Passive sonar target classification using multi-layer perceptron trained by salp swarm algorithm, Ocean Eng. 181 (2019) 98–108.
[75] M. Khishe, M.R. Mosavi, Improved whale trainer for sonar datasets classification using neural network, Appl. Acoust. 154 (2019) 176–192.
[76] W. Qiao, M. Khishe, S. Ravakhah, Underwater targets classification using local wavelet acoustic pattern and Multi-Layer Perceptron neural network optimized by modified Whale Optimization Algorithm, Ocean Eng. 219 (2021) 108415.
[77] S. Harifi, M. Khalilian, J. Mohammadzadeh, S. Ebrahimnejad, Emperor Penguins Colony: a new metaheuristic algorithm for optimization, Evolutionary intelligence 12 (2019) 211–226.
[78] G.G. Wang, S. Deb, X.Z. Gao, L.D.S. Coelho, A new metaheuristic optimisation algorithm motivated by elephant herding behaviour, Int. J. Bio-Inspired Comput. 8 (6) (2016) 394–409.
[79] K. Ahmed, A.E. Hassanien, E. Ezzat, An efficient approach for community detection in complex social networks based on elephant swarm optimization algorithm, in: Handbook of Research on Machine Learning Innovations and Trends, IGI Global, 2017, pp. 1062–1075.
[80] W. Li, G.G. Wang, A.H. Alavi, Learning-based elephant herding optimization algorithm for solving numerical optimization problems, Knowl. Base Syst. 195 (2020) 105675.
[81] R.V. Rao, V.J. Savsani, D.P. Vakharia, Teaching–learning-based optimization: a novel method for constrained mechanical design optimization problems, Comput. Aided Des. 43 (3) (2011) 303–315.
[82] M. Črepinšek, S.H. Liu, L. Mernik, A note on teaching–learning-based optimization algorithm, Inf. Sci. 212 (2012) 79–93.
[83] R.V. Rao, V.J. Savsani, D.P. Vakharia, Teaching–learning-based optimization: a novel method for constrained mechanical design optimization problems, Comput. Aided Des. 43 (3) (2011) 303–315.
[84] UCI Machine learning Repository. Accessed: June. 1, 2023. [Online]. Available: http://archive.ics.uci.edu/ml/.
[85] D. Painuli, S. Bhardwaj, Recent advancement in cancer diagnosis using machine learning and deep learning techniques: a comprehensive review, Comput. Biol. Med. 146 (2022) 105580.
[86] R.R. Rajammal, S. Mirjalili, G. Ekambaram, N. Palanisamy, Binary grey wolf optimizer with mutation and adaptive k-nearest neighbour for feature selection in Parkinson's disease diagnosis, Knowl. Base Syst. 246 (2022) 108701.
[87] D. Deepika, N. Balaji, Effective heart disease prediction using novel MLP-EBMDA approach, Biomed. Signal Process Control 72 (2022) 103318.
[88] X. Bai, X. Gao, B. Xue, Particle swarm optimization based two-stage feature selection in text mining, in: 2018 IEEE Congress on Evolutionary Computation (CEC), 2018, pp. 1–8.
[89] Spiliopoulou B M P W R, A Hayward C Rudan I Campbell H Wright AF Wilson JF Agakov F Navarro P Haley CS, Application of high-dimensional feature selection: evaluation for genomic prediction in man Sci Rep 5 (10312) (2015), 10-1038.

[90] Y. Rui, T.S. Huang, S.F. Chang, Image retrieval: Current techniques, promising directions, and open issues, J. Vis. Commun. Image Represent. 10 (1) (1999) 39–62.
[91] M. Dash, Feature selection via set cover, in: Proceedings 1997 IEEE Knowledge and Data Engineering Exchange Workshop, 1997, pp. 165–171.
[92] F. Yang, K.Z. Mao, Robust feature selection for microarray data based on multicriterion fusion, IEEE/ACM Trans Comput Biol Bioinform 8 (2011) 1080–1092, https://doi.org/10.1109/TCBB.2010.103.
[93] I. Guyon, A. Elisseeff, An introduction to variable and feature selection, J. Mach. Learn. Res. 3 (2003) 1157–1182.
[94] L.C. Molina, L. Belanche, A. Nebot, Feature selection algorithms: a survey and experimental evaluation, in: 2002 IEEE International Conference on Data Mining, 2002, pp. 306–313, 2002. Proceedings.
[95] H. Liu, Yu Lei, Toward integrating feature selection algorithms for classification and clustering, IEEE Trans. Knowl. Data Eng. 17 (2005) 491–502, https://doi.org/10.1109/TKDE.2005.66.
[96] J. Loughrey, P. Cunningham, Overfitting in wrapper-based feature subset selection: the harder you try the worse it gets, in: M. Bramer, F. Coenen, T. Allen (Eds.), Research and Development in Intelligent Systems XXI, Springer, London, 2005, pp. 33–43.
[97] A. Jakulin, I. Bratko, Testing the significance of attribute interactions, in: Proceedings of the Twenty-First International Conference on Machine Learning, association for computing machinery, New York, NY, USA, 2004, p. 52.
[98] J. Kennedy, R.C. Eberhart, A discrete binary version of the particle swarm algorithm, in: 1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation, vol. 5, IEEE, 1997, October, pp. 4104–4108.
[99] R. Cheng, Y. Jin, A competitive swarm optimizer for large scale optimization, IEEE Trans. Cybern. 45 (2015) 191–204, https://doi.org/10.1109/TCYB.2014.2322602.
[100] S. Gu, R. Cheng, Y. Jin, Feature selection for high-dimensional classification using a competitive swarm optimizer, Soft Comput. 22 (2018) 811–822, https://doi.org/10.1007/s00500-016-2385-6.
[101] L.K. Singh, M. Khanna, H. Garg, R. Singh, Emperor penguin optimization algorithm-and bacterial foraging optimization algorithm-based novel feature selection approach for glaucoma classification from fundus images, Soft Comput. (2023) 1–37.
[102] V. Chaurasia, S. Pal, Applications of ML techniques to predict diagnostic breast cancer, SN Computer Science 1 (5) (2020) 1–11.
[103] N. Liu, E.S. Qi, M. Xu, B. Gao, G.Q. Liu, A novel intelligent classification model for breast cancer diagnosis, Inf. Process. Manag. 56 (3) (2019) 609–623.
[104] V.E. Christo, H.K. Nehemiah, J. Brighty, A. Kannan, Feature Selection and instance selection from clinical datasets using co-operative co-evolution and classification using random forest, IETE J. Res. (2020) 1–14.
[105] H. Rao, X. Shi, A.K. Rodrigue, J. Feng, Y. Xia, M. Elhoseny, L. Gu, Feature Selection based on artificial bee colony and gradient boosting decision tree, Appl. Soft Comput. 74 (2019) 634–642.
[106] B. Sahu, S. Mohanty, S. Rout, A hybrid approach for breast cancer classification and diagnosis, EAI Endorsed Transactions on Scalable Information Systems 6 (20) (2019).
[107] R.A. Ibrahim, A.A. Ewees, D. Oliva, M. Abd Elaziz, S. Lu, Improved salp swarm algorithm based on particle swarm optimization for Feature Selection, J. Ambient Intell. Hum. Comput. 10 (8) (2019) 3155–3169.
[108] M. Nilashi, O. Ibrahim, H. Ahmadi, L. Shahmoradi, A knowledge-based system for breast cancer classification using fuzzy logic method, Telematics Inf. 34 (4) (2017) 133–144.
[109] S.B. Sakri, N.B.A. Rashid, Z.M. Zain, Particle swarm optimization Feature Selection for breast cancer recurrence prediction, IEEE Access 6 (2018) 29637–29647.
[110] J. Dheeba, N.A. Singh, S.T. Selvi, Computer-aided detection of breast cancer on mammograms: a swarm intelligence optimized wavelet neural network approach, J. Biomed. Inf. 49 (2014) 45–52.
[111] G.N. Ramadevi, K.U. Rani, D. Lavanya, Importance of feature extraction for classification of bc datasets, a study, International Journal of Scientific and Innovative Mathematical Research 3 (2) (2015), 763-368.
[112] N.F. Idris, M.A. Ismail, Breast cancer disease classification using fuzzy-ID3 algorithm with FUZZYDBD method: automatic fuzzy database definition, PeerJ Computer Science 7 (2021) e427.
[113] H. Rajaguru, S.C. SR, Analysis of decision tree and k-nearest neighbor algorithm in the classification of breast cancer, Asian Pac. J. Cancer Prev. APJCP: Asian Pac. J. Cancer Prev. APJCP 20 (12) (2019) 3777.