



# OPEN Optimized breast cancer diagnosis using self-adaptive quantum metaheuristic feature selection

Alok Kumar Shukla<sup>1</sup>, Shubhra Dwivedi<sup>1</sup>, Deepak Singh<sup>2</sup>, Sunil Kumar Singh<sup>3✉</sup>,  
Diwakar Tripathi<sup>4</sup> & Ram Kishan Dewangan<sup>5</sup>

Breast cancer is a leading cause of mortality among women and is increasing rapidly around the world. For early diagnosis of breast cancer, precise classification, and finding the best subset for cancer identification, evolutionary-based feature selection methods play a vital role in effective treatment. Previous studies have shown that existing evolutionary methods are complicated in correctly differentiating BC disease subtypes with high consistency, which seriously affects the performance of classification methods. To prevent diagnostic errors with hostile implications for patient health, in this study, we develop a new evolutionary method called SeQTLBOGA that incorporates the learner quantization before the search capability of the feature space to prevent premature falls into the local optima. In the SeQTLBOGA algorithm, quantum theory and a self-adaptive mechanism are employed to update the Teaching Learning-based Optimization (TLBO) rule to enhance convergence search capabilities. Most importantly, a self-adaptive genetic algorithm (GA) is also incorporated into TLBO to tradeoff between exploration and exploitation to handle slow convergence and exploitation competence, and simultaneously optimizing parameters of support vector machines (SVM) and the best features subset is our primary objective. Comparative results based on optimal computing time and performance are also offered to empirically analyze the traditional algorithms. Therefore, this paper aims to evaluate the most recent quantum-inspired metaheuristic algorithms in WBCD and WDBC databases, emphasizing their advantages and disadvantages.

**Keywords** Teaching Learning-based optimization, Breast Cancer, Feature selection, Genetic algorithm, Quantum, Support vector machine

Cancer is not brought on by external factors; rather, it is produced by specific genetic abnormalities in cells<sup>1</sup>. One of the biggest worries in the world, especially for women, is the rising number of cancer-related deaths<sup>2</sup>. This keeps up the pattern that breast cancer (BC) is the second most common cause of cancer-related deaths among women, and the cancer that is diagnosed in women more often than skin cancer. As of 2020<sup>3</sup>, lung cancer has been surpassed by breast cancer as the most frequent cancer worldwide. By 2040, there will likely be 3.2 million female cases of breast cancer (compared to the present 2.3 million instances), and 1 million projected deaths from the disease. The good news is that, if found early enough to prevent metastatic cancer, the 5-year survival rate for breast cancer can reach 90%<sup>4,5</sup>.

Therefore, figuring out how to efficiently extract critical features for breast cancer diagnosis and reduce dimensionality to ease computational and space complexity while enhancing model performance is crucial. Many automated procedures are available to differentiate between benign and malignant tumors<sup>6,7</sup>. Quantum computing (QC), one of the best ways to solve the dimensionality ruin problem, has been applied to enhance the efficiency of the automated process for early cancer diagnosis<sup>8</sup>. The concept of quantum computing is based on the principles of the quantum bit, promotes population variety because of the combination of states, and swift convergence<sup>9</sup>. Generally speaking, feature selection (FS) works by framing the task of finding a subset of characteristics as an optimization problem. For such an approach, a cost function that gauges a feature subset's quality, usually via a classifier's accuracy, must be defined<sup>10</sup>.

Recently, the principles of quantum computing have been applied to address the time and space complexity by combining natural methods such as artificial bee colonies, teaching learning-based optimization, particle

<sup>1</sup>Thapar Institute of Engineering and Technology, Patiala, Punjab, India. <sup>2</sup>National Institute of Technology, Raipur, India. <sup>3</sup>Department of Computer Science and Engineering, Manipal Institute of Technology Bengaluru, Manipal Academy of Higher Education, Manipal, India. <sup>4</sup>National Institute of Technology, Jamshedpur, India. <sup>5</sup>SCSET, Bennett University, Greater Noida, India. ✉email: sunil.singh@manipal.edu

swarm optimization (PSO), grasshopper optimization algorithm (GOA), GAs, and differential evolution<sup>11</sup>. Quantum evolutionary algorithms (QEAs) outperformed standard EAs for combinatorial optimization while avoiding premature convergence<sup>12</sup>. Furthermore<sup>13</sup>, frequently experience problems, including poor convergence accuracy and a tendency to converge towards the Local Optima (LO). The study presented an integrated strategy that integrates the techniques for Differential Evolution, Slime Mold Algorithm, and Salp Swarm Algorithm to address these issues presented by the SDSSA—this cutting-edge hybrid MIS model combined components from the SSA, SMA, and DE algorithms. The SDSSA model primarily depended on 2D Kapur's entropy and a non-local means 2D histogram. It was first contrasted with analogous algorithms to assess the suggested approach efficiently.

The approach to decision-making in many aspects of life has been revolutionized by the ongoing development of artificial intelligence (AI). A growing area of artificial intelligence called Explainable Artificial Intelligence focuses on creating a wide range of devices and methods for opening up BlackBox AI solutions by producing transparent, perceptive, and understandable explanations of AI judgments<sup>14</sup>. Moreover, medical practitioners' confidence in AI systems in the future may be strengthened by explainability in AI. It is not so much a product of AI as it is a challenge that has existed for at least as long. In<sup>15</sup>, the author presented a novel approach that merges a lightweight parallel depth-wise separable convolutional neural network (LPDCNN) with a ridge regression extreme learning machine to precisely classify three lung cancer types alongside normal lung tissue using CT images. The proposed methodology combined contrast-limited adaptive histogram equalization and Gaussian blur to enhance image quality, reduce noise, and improve visual clarity. LPDCNN extracted discriminant features while minimizing computational complexity (0.53 million parameters and 9 layers). The ridge-ELM model was developed to enhance classification performance, replacing the traditional pseudoinverse in the ELM approach. Notably, integrating the SHAP (Shapley Additive Explanations) in the proposed framework enhances Explainability, providing insights into decision-making and boosting confidence in real-world lung cancer diagnoses.

When tackling feature selection optimization issues, metaheuristic algorithms may typically meet two fundamental preconditions: exploration and exploitation, although they employ various operators and techniques<sup>16</sup>. The algorithm can further refine the quality of the result by searching inside a favorable region through exploitation, a local search capability. The best solution will be of higher quality if the two capabilities are balanced appropriately. Rao et al. proposed TLBO robust and relatively recent metaheuristic algorithm inspired by the dynamical behavior of teaching and learning events, to illustrate better the benefits and contributions of metaheuristics<sup>17,18</sup>. The drawbacks above can be addressed by genetic algorithms with self-adaptive parameters comprising crossover and mutation operators compared to other EAs. The quantity of a solution that can be obtained in terms of convergence speed is determined mainly by genetic algorithms' crossover and mutation probabilities. Rather than relying on constant  $P_c$  and  $P_m$  values, SA-GA uses population data from every generation to modify  $P_c$  and  $P_m$  in an adaptive manner that preserves population variety and convergence ability.

Improving prediction skills mostly depends on choosing the most crucial variables or input parameters. Adopting an adaptive EA method that incorporates the advantages of multiple techniques has significantly improved the learning model performance; in most cases, adaptive algorithms outperform standalone algorithms. In real-world engine building, as opposed to other disciplines like feature engineering, there is a trade-off between feature dimension and the necessary computing cost, and the available feature space, or feature candidates, is typically constrained but noisy. The primary contributions of this paper are:

- This paper introduces quantum self-adaptive TLBO and GA algorithms, SeQTLBOGA, to determine the most useful features for the hybrid engine using SVM of BC in order to handle the stated technical obstacles methodically. In addition, the proposed method is applied to enhance the penalty factor and kernel parameter of the SVM.
- In order to improve exploration and exploitation search capability, the proposed algorithm first uses a quantum technique to generate solutions. After that, novel crossover and mutation operators are used to identify nearly optimal solutions and generate a population with incredible diversity.
- To provide interpretability, the prediction of the model can be interpreted through SHAP, which provides interpretable explanations.
- SeQTLBOGA aims to create a parsimonious predictive model that can distinguish benign from malignant breast tumors with the best features of high quality. In addition, this model can help the analyst provide diagnostic recommendations to general practitioners while achieving the desired prediction accuracy level.

The remainder of this study is organized as follows: Section “Related work” discovers related work on human intelligence-based algorithms and nature-inspired techniques. Section “Existing works” describes the data preparation and existing methods adopted for the study. Sections “Proposed methodology for classification system” and “Experimental study and analysis” describe the proposed work, and the results of the experiment are discussed in detail, followed by the concluding remarks in Section “Conclusion”.

## Related work

Healthcare professionals are currently gathering a wide range of patient data, especially in BC. In this regard, several EA-based decision-support systems have already received approval for use in clinical settings<sup>19</sup>. Wrapper techniques are still not widely used in clinical practice due to their interpretability, the ability to grasp how the algorithms make decisions, despite all the efforts to integrate them into physicians' workflows. Therefore, given the previously indicated increase in BC data, rigorous examination of stochastic EAs should reveal pertinent insights on the impact of these techniques on the breast cancer subtyping problem. Since variations

of EAs can learn embedded representations, we concentrate our study on self-adaptive and quantum EAs and their conditional equivalents<sup>20</sup>. Accurately identifying and describing the morphological traits is crucial to recognizing and diagnosing breast cancer disorders. The complicated morphological features restrict the success of human specialists in the field. Subjectivity and other human factors can lead to differences in diagnosis and treatment plans, which can harm patient outcomes. Practitioners could not be exposed to a variety of collections of breast cancer tissue samples, which can result in an incorrect diagnosis, such as false-positive or false-negative. Additionally, the manual procedure of assessing slides can take a long time, delaying the detection and management of the disease<sup>21,22</sup>.

Various sensing modalities can yield distinct tissue behavior for cancer and adjacent normal tissues, serving as a basis for cancer diagnosis. A novel multimodal diagnostic tool that can concurrently assess the optical, electrical, and mechanical bulk tissue properties can substantially augment the clinical findings, such as histopathology, potentially aiding the clinician in establishing an accurate and rapid cancer diagnosis. This review aims to discuss the clinical and engineering aspects and the unmet challenges of these physical sensing modalities, primarily in optical, electrical, and mechanical fields. The challenges of combining two or more of these sensing modalities that can significantly enhance the effectiveness of the clinical diagnostic tools are further investigated<sup>23</sup>.

A hybrid AlexNet-Extreme Learning Machine (ELM) approach for breast cancer diagnosis using mammography images was proposed in<sup>24</sup>. Batch normalization was applied to improve AlexNet's performance, and the chimp optimization algorithm (ChOA) was utilized to avoid sub-optimal solutions in ELM. The Nelder-Mead simplex (NEMS) technique was then employed to enhance the convergence behavior of ChOA. The study's main contributions were the proposed hybrid model and the application of ChOA and NEMS techniques to improve the performance of ELM. The proposed model was evaluated using the CBIS-DDSM dataset with Wiener and CALHE filters used as preprocessors. The effectiveness of the classification was examined using five optimization algorithms and several metrics. The outcomes demonstrated that the CALHE filter offered the best performance overall, and AlexNet-BN-ELM-CHOA-NEMS was the most accurate of the five models, with a sensitivity of 96.03%, specificity of 94.60%, and overall accuracy of 95.32%. The findings demonstrate the effectiveness of the proposed model in breast cancer diagnosis.

Breast cancer diagnosis involves the assessment of histopathological slide images by skilled pathologists. In<sup>25</sup>, the author was prone to human subjectivity, which can lead to diagnostic errors with adverse implications for patient health and welfare. Artificial intelligence-based models have yielded promising results in other medical tasks and offer tools for potentially addressing the shortcomings of traditional medical image analysis. The BreakHis breast cancer dataset suffers from insufficient data for the minority class with an imbalance ratio of more than 0.40, which poses challenges for deep learning models. To avoid performance degradation, researchers have explored a variety of data augmentation schemes to generate adequate samples for analysis. The author studied a Deep Convolutional Neural Network (DCGAN) with a specific generator and discriminator architecture to mitigate model instability and generate high-quality synthetic data for the minority class. The balanced dataset was passed to the fine-tuned ResNet50 model for breast tumor detection. The study produced high accuracy in diagnosing benign/malignant at 40X magnification, outperforming the state-of-the-art. The results demonstrated that deep learning methods can potentially support effective screening in clinical practice.

Malignant cells predominantly affect women, and researchers in the field of clinical sciences are increasingly interested in leveraging Artificial Intelligence techniques to predict and diagnose breast cancer. In<sup>26</sup>, the study utilized the widely-used Wisconsin Breast Cancer Dataset (WBCD) from the University of California, Irvine (UCI) machine learning repository. The dataset contains 30 features, including mean, standard error, and worst values, providing insights into different aspects of breast cancer. A stacked ensemble classifier was proposed to evaluate the effectiveness of the proposed model, incorporating multiple machine learning algorithms such as the Decision Tree Classifier, AdaBoost Classifier, Gaussian Naive Bayes, and Multi-layer Perceptron Classifier. Various performance measures, including Receiver Operating Characteristic Curve, Area Under the Curve, specificity, F1-score, sensitivity, and accuracy, were used to assess the model performance. The results demonstrated that the proposed ensemble technique achieved an accuracy of 97.66%, surpassing the performance reported in the existing literature. The study underscores the potential of AI techniques and machine learning algorithms in advancing breast cancer prediction and diagnosis, offering promising avenues for improving clinical decision-making and patient outcomes.

The shuffling frog leaping algorithm was combined with an innovative and effective minimum attribute reduction approach based on quantum-inspired self-adaptive cooperative co-evolution<sup>27</sup>. Initially, multi-state quantum bits were utilized to represent evolutionary frog individuals. Then, the self-adaptive quantum rotation angle and quantum mutation probability technique were implemented to update the quantum revolving door's function. Second, a self-adaptive cooperative co-evolutionary model for minimal attribute reduction was created to partition the evolutionary attribute sets into manageable subsets. Based on their past performance records, the subsets were given the self-adaptive mechanism, and each one was evolved using the quantum-inspired shuffling frog leaping algorithm. Ultimately, the suggested algorithm's global convergence was demonstrated theoretically, and its effectiveness was examined using UCI datasets, magnetic resonance, and a few other global optimization functions.

The Butterfly optimisation algorithm and the Ant Lion optimiser served as the foundation for the hybrid BOAALO approach, a feature selection technique that was developed in<sup>28</sup>. Three classifiers—an artificial neural network, a support vector machine, and an adaptive neuro-fuzzy inference system—were fed the optimal set of features selected by BOAALO in order to determine if breast tissue is benign or malignant. 651 mammography pictures were used to evaluate the suggested technique's efficacy. The findings demonstrate that the first BOA and ALO are outperformed by the receiver operating characteristics curve, Type-I and Type-II error, kappa value,

and BOAALO. Additionally, using a benchmark dataset, the suggested strategy's robustness was evaluated and contrasted with five established methods.

Author<sup>29</sup> developed a noninvasive artificial intelligence approach to characterize the tumor microenvironment by fusing deep learning's prediction capabilities with the explainability of human-interpretable imaging phenotypes (IMPs). IMPs were created using 342 breast tumors' 4D dynamic imaging linked to clinical and genomic data. This allows for the genetic relationship between cancer phenotypes and genotypes. A deep graph clustering technique with unsupervised dual attention was developed to divide bulk tumors into many geographically separated and phenotypically consistent subclusters. Intratumorous subcluster architecture, interaction, and proximity were captured using IMPs varying in spatial and kinetic heterogeneity. IMPs positively correlate with established TME markers and may accurately, consistently, and reliably predict specific molecular fingerprints, such as the expression of hormone receptors, epithelial growth factor receptors, and immunological checkpoint proteins.

### Existing works

High-dimensional datasets may require extensive computations to explore the space of possible search strategies fully<sup>30,31</sup>. For example, GA simultaneously handled feature and instance selection. The primary drive behind the development of quantum self-adaptive EAs is their efficacy in resolving medical issues. It is to be expected that the behavior of nature is always at its best. Probabilistic algorithms underpin all EA-based algorithms, and standard parameters like population size, tuning parameters, and maximum number of iterations are necessary. Many algorithms need their algorithm-specific control parameters and standard control parameters<sup>32</sup>. Utilizing the competition technique, a more recent study aims to efficiently search for less expensive features by enhancing the global search capability of an EA-based filter method for cost-sensitive feature selection.

### Quantum concept

The quantum concept denotes that specific physical attributes, such as energy, charge, and angular momentum, exist in discrete, indivisible units known as quanta, rather than in a continuous form. This essential concept underlies the entire domain of quantum mechanics, which elucidates the behaviour of matter and energy at atomic and subatomic levels. Quantum systems can simultaneously occupy several states until observation occurs. This underpins phenomena such as Schrödinger's cat thought experiment, when a system exists in a superposition of potential outcomes until an observation occurs<sup>33</sup>. Each solution (feature subset) is represented as a qubit in superposition, enabling parallel evaluation of multiple states.

$$|\psi\rangle = \alpha |0\rangle + \beta |1\rangle, \text{ where } |\alpha|^2 + |\beta|^2 = 1$$

A rotation gate  $U(\theta)$  adjusts qubit amplitudes to steer solutions toward high-fitness regions. For the BC problem, the quantum rotation gate is described as follows:

$$U(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Applied to a qubit  $|\psi\rangle$ , the updated state becomes

$$|\psi'\rangle = U(\theta) \cdot |\psi\rangle = \begin{bmatrix} \alpha \cos\theta - \beta \sin\theta \\ \alpha \sin\theta + \beta \cos\theta \end{bmatrix}.$$

The rotation angle  $\theta$  is dynamically tuned using fitness feedback. Large  $\theta$  value promotes exploration by amplifying low-probability states. And, small  $\theta$  value enhances exploitation near high-fitness solutions. Generate  $Q(t)$  quantum population with qubits in equal superposition. Collapse qubits into classical solutions via probabilistic sampling.

$$x_{ij} = \begin{cases} 1 & \text{if } |\beta_{ij}|^2 \geq \text{rand}(0,1) \\ 0 & \text{otherwise} \end{cases}$$

Compute  $f(x_i)$  for all solutions. Adjust  $\alpha_{ij}, \beta_{ij}$  using  $U(\theta_{ij})$  to steer solutions toward better candidates. For an  $n$ -dimensional feature selection problem, quantum population  $Q(t) \{q_1^t, q_2^t, \dots, q_m^t\}$  is maintained,

$$\text{where each } q_i^t \text{ is a matrix of qubits: } q_i^t = \begin{bmatrix} \alpha_{i1} & \beta_{i1} \\ \alpha_{i2} & \beta_{i2} \\ \vdots & \vdots \\ \alpha_{in} & \beta_{in} \end{bmatrix}$$

Here,  $Q(t) = \{q_1^t, q_2^t, q_3^t, \dots, q_N^t\}$  is a population of  $N$  qubit learners at generation  $t$ ;  $Q(t)$  is  $i$ th 1, 2, ...,  $N$ , individual defined as above. Each qubit's superposition allows simultaneous exploration of  $2^n$  Feature combinations exponentially increase search space coverage compared to classical methods.

Quantum superposition and adaptive rotation angles reduce the risk of stagnation in local optima. Gate entangle qubits to ensure correlated feature updates, enhancing global search coherence. In this way, the  $n$ -qubits learner can simultaneously represent the information of  $2^n$  binary states. QEAs have better characteristics of diversity than classical approaches since they can represent a linear superposition of many states, thanks to the

qubit representation. For the update process of QEAs, a suitable quantum gate  $U(\theta)$ . it is usually adopted in compliance with the BC problem.

$$q_i^t = \begin{bmatrix} \alpha_{11}^t & \dots & \alpha_{im}^t \\ \beta_{11}^t & \dots & \beta_{im}^t \end{bmatrix}$$

and  $P(t) = \{X_1^t, X_2^t, X_3^t, \dots, X_N^t\}$  is a set of binary solutions of observation states of  $Q(t)$ , where  $X_i^t$  is the binary solution by observing  $q_i^t (1, 2, \dots, n)$ . In the initialize  $Q(t)$  step, each pair of qubit probability amplitudes, at  $\alpha_{1r}^t$  and  $\beta_{1r}^t; r=1, 2, \dots, m$ , are initialized with  $\frac{1}{\sqrt{2}}, \forall q_i^t \in Q(t)$ . The next step generates a set of binary solutions  $P(t) = \{X_1^t, X_2^t, X_3^t, \dots, X_N^t\}$  where each bit of  $X_i^t = 1, 2, \dots, N$ , is formed by determining the explicit state of each qubit of  $q_i^t |0\rangle$  state or  $|1\rangle$  state, according to either  $\alpha_{1r}^t$  or  $\beta_{1r}^t$  of  $q_i^t$ . For example, to form an explicit state of  $r$ th bit of  $X_i^t = 1, 2, \dots, r$ , the number  $r$  between 0 and 1 is generated randomly. The initial best solution  $X_i^t$  best is then selected and stored among the binary solutions  $P(t)$ . In the while loop, the quantum gate  $U(\theta)$  is used to update  $Q(t-1)$  so that fitter states of the qubit chromosomes are generated. The  $j$ th qubit value  $\alpha_j^t; \beta_j^t$  of  $q_i^t$  is updated as below Eq. (1).

$$\begin{pmatrix} \alpha_j^{t+1} \\ \beta_j^{t+1} \end{pmatrix} = U(\theta_j^t) \begin{pmatrix} \alpha_j^t \\ \beta_j^t \end{pmatrix} = \begin{bmatrix} \cos \theta_j^t & -\sin \theta_j^t \\ \sin \theta_j^t & \cos \theta_j^t \end{bmatrix} \begin{pmatrix} \alpha_j^t \\ \beta_j^t \end{pmatrix}. \quad (1)$$

Here,  $\theta_j^t$  is the rotation angle of  $q_j^t$ . The best solution among  $P(t)$  is then selected in the next step, and if the current solution is fitter than the best-stored solution, the best-stored solution  $X_t^{best}$  will be replaced by this current solution.

### Teaching learning based optimization

Metaheuristic techniques incorporate procedures and strategies to create novel solutions and promote a proper balance of search diversification, intensification, and standard regulating parameters such as population size and maximum number of iterations<sup>34</sup>. It can take the form of two primary methods: population-based algorithms that evolve collections of solutions like TLBO and GA and single-solution-based algorithms that modify a single solution as SA. Because of the convergence rate and the reduced number of changing parameters, EAs are powerful. The algorithm known as Teaching Learning Based Optimization (TLBO) was created lately and is based on the chasing mechanism of the knowledge-acquisition phase, inspired by the classroom teaching and learning process. Owing to its distinct structure, TLBO possesses virtuous global search capabilities. The low convergence rate of the individual TLBO algorithm is one of its shortcomings, which restricts its practical use. When tackling complex global optimization issues, SA-GA smoothly enters the local optimal region while exhibiting a rapid rate of convergence, provided that sufficient fine-tuning of preliminary constraints is not required. An active SeQTLBOGA algorithm is suggested to solve the feature selection challenge given the features of TLBO and SA-GA.

This algorithm aims to identify the most significant learner by having students work together and share information. It is divided into two phases: the teacher and student phases. Students gain knowledge from their classmates and a capable teacher during both. For the TLBO algorithm to work, there can be very skilled students who can get excellent marks or test scores. Together, these students make up what is known as a class. In this class, information is shared through the neighbor learning phenomenon, which facilitates the sharing of perspectives and knowledge:

$$X_{i,k} = \{X_{i,1}, X_{i,2}, \dots, X_{i,Fe}\}$$

Where,  $L_b$  is lower bound and  $U_b$  is upper bound of the  $Fe$  dimension in the search space  $X_{i,Fe} \in [L_b, U_b]$ . The learner is randomly initialized within the search space capabilities. The evolution of  $X_{i,k}$  is randomly generated by the following Eq. (2).

$$X_{i,k} = L_k + r_1 * (U_k - L_k). \quad (2)$$

Where,  $i=1, 2, 3, \dots, \text{nop}$ ,  $k=1, 2, 3, \dots, Fe$ ,  $r_1$  shows a random variable,  $L_k$  shows the lower bound and  $U_k$  shows the upper bound value. The simulation of the classical learning process for all learners is categorized into two significant phases of the TLBO algorithm, namely the Teacher phase and the Learner Phase.

#### Teacher phase

This stage involves the teacher teaching the students. The instructor improves the mean grade for the class as a whole by keeping the class informed and interacting. The teacher is a gauge of how near the best outcomes from optimization problems have been achieved thus far. A skilled teacher refreshes their expertise based on the students' understanding. The highest point to which a teacher can boost the class mean score will depend on the skill of the entire class. Assume that  $M_{i,k} = (1/\text{nop}) (\sum X_{i,k})$  represents the average value for the given subject, with  $k=1, 2, \dots, K$ . The process's update equation, as outlined in Eq. 3.

$$X_{i,k}^{new} = X_{i,k}^{old} + r_2 * (X_{teacher,k} - T_f * M_{i,k}) \& T_f = \text{round} [1 + \text{rand}(0, 1)] \quad (3)$$

Here,  $X_{teacher,k}$  is the best learner of the adopted population at the current iteration of the algorithm, 2 is the random number lies between 0 and 1, respectively;  $T_f$  is a teaching factor that decides the value of the mean



to be changed. In every iteration,  $X_{i,k}^{new}$  is updated from the value of  $X_{i,k}^{old}$ .  $X_{i,k}^{new}$  and  $X_{i,k}^{old}$  denote the  $k$ th learners select after or before learning from the teacher<sup>35</sup>.

#### Learner phase

The second part of this process is called the learner phase, and it attempts to increase students' knowledge in two ways: first, by having teachers provide feedback, and second, by having students engage with one another. To develop their communication abilities, each student's goal is to interact with other students randomly<sup>17</sup>. To put it another way, every learner engages in interpersonal contact during the student phase. As a result, the surrounding solution may get trapped in a local optimum, leading to the current solution getting stuck in a local optimum close to the optimal solution.  $X_p$  and  $X_q$  where through mutual interaction, to choose the  $i$ th learner where ( $p \neq q$ ). In the learning phase, the updating equation of the  $i$ th learner  $X_p$  is as follows:

```

For  $i = 1$ : nPop.
  Randomly choose two learners  $X_p$  and  $X_q$ 
  if ( $f(X_p) < f(X_q)$ )
     $newX_i = oldX_i + r * (X_p - X_q)$ .
  Else
     $newX_i = oldX_i + r * (X_q - X_p)$ .
  End if.
End for.

```

Where,  $r$  is a uniformly distributed random number between 0 and 1,  $f(X_p)$  and  $f(X_q)$  are the best solution of the learners  $X_p$  and  $X_q$ , respectively.

#### Cuckoo search

The Cuckoo Search (CS) algorithm was introduced in 2009 by Xin-She Yang and Suash Deb<sup>36</sup> as an optimization method inspired by nature. The way that cuckoo birds breed served as the model for this optimization problem-solving technique. The strategy is predicated on the fact that cuckoos are known to deposit their eggs in other bird species' nests. The CS algorithm's main ideas and procedures are to set the initial population of potential solutions, which are referred to as nests in an algorithmic context. Since each nest indicates a possible fix for the optimization issue, Levy Flight: A Levy flight is when a flock of cuckoos walks randomly through the search space. Step sizes in Levy flights, a random walk, are determined by a probability distribution.

#### Levy flight

A Levy flight behavior is incorporated into Cuckoo Search to replicate the haphazard steps cuckoos take when looking for nests. A heavy-tailed probability density function, or Levy distribution, can be used to predict the step sizes in the Levy flight. The following represents the Levy flight in Eq. (4).

$$l(t) = \lambda \text{Levy}(\alpha) \quad (4)$$

Where  $l(t)$  is the step size at time  $t$ ,  $\lambda$  is a scaling factor, and  $\text{Levy}(\alpha)$  represents random samples from a Levy distribution with a specific parameter  $\alpha$ .

---

**Input:** Initialize the parameters nPop,

*Dimension,  $\varepsilon$  (objective function),  $t_{max}$  number of iterations,  $t = 1$ ;*

**Output:** Best learner

1. Set nPop, *Dimension,  $t_{max}$*
  2. **for**  $i = 1$ : nPop // Each search agent
  3.  $X_i = \text{initialize}(\text{nPop})$ ; //Generate a random population with D-dimensional vector
  4. **end for**
  5.  $Z = \text{fitness}(X_i, \varepsilon)$  //Estimate the fitness of each learner (X)
  6.  $X^* = \text{the best search learner}$ ; // called Teacher
  7. **while** ( $t < t_{max}$ ) //maximum iterations are not meet
  8.  $t = t + 1$ ;
  9. **for**  $i = 1$ : nPop
  10. Update position vector by teacher phase
  11. **end for**
  12. **for**  $i = 1$ : nPop
  13. Calculate the fitness of all updated learners
  14. **end for**
  15. Update learner  $X_i$  and calculate learner value by learner phase
  16. **end while**
  17. Store the optimal feature subset in all generations with highest fittest value;
- 

**Algorithm 1.** Basic Teaching Learning-based Optimization algorithm.

#### Egg laying and nest replacement

When one cuckoo's egg is superior to what is currently in the nest, it replaces the previous one. Cuckoos lay their eggs in nests. The likelihood that a cuckoo egg will be superior to the contents of the nest is expressed in Eq. (5).

$$P_a = \exp(-\alpha * \text{fitness\_cuckoo}). \quad (5)$$

where fitness\_cuckoo is the cuckoo egg's fitness,  $P_a$  is the acceptance probability, and  $\alpha$  is a parameter governing the acceptance probability.

#### Particle swarm optimization

Particle swarm optimization (PSO) is a kind of optimization that tackles a range of optimization problems by drawing inspiration from the natural world. PSO was first introduced by James Kennedy and Russell Eberhart in 1995<sup>37</sup>. PSO traverses the search space with a population of potential solutions, represented as particles, in order to find the optimal solution to an optimization issue. Every particle's position in the search space represents a possible solution, and the particles travel based on their unique experiences as well as the collective behavior of the swarm. The PSO algorithm's main ideas and procedures are as follows: Initialize the population of particles in the search space with random positions and velocities using the following update methods. Equations (6) and (7).

##### Position update

Using the particle current position, velocity, and best solution discovered thus far, update each particle's position within the population. Equation (6) is the position update Eq.

$$x_i(t+1) = x_i(t) + v_i(t+1). \quad (6)$$

Where,  $x_i(t+1)$  is the updated position of particle  $i$  at time  $t+1$ ,  $x_i(t)$  is the current position of particle  $i$  at time  $t$ ,  $v_i(t+1)$  is the velocity update of particle  $i$  at time  $t+1$ .

##### Velocity update

The population best solution (gbest) and the best solution it has discovered thus far (pbest) Using the present velocity of each particle, determine its velocity. A possible solution to the optimization problem is represented by each particle. Determine the fitness (objective function value) of each particle's current position to ascertain how effectively it solves the problem, Individual Best (best). Equation (7) is the velocity update Eq.

$$v_i(t+1) = w*v_i(t) + c1*r*(pbest_i - x_i(t)) + c2*rand()*(gbest - x_i(t)). \quad (7)$$

Where,  $v_i(t+1)$  is the updated velocity of particle  $i$  at time  $t+1$ ,  $v_i(t)$  is the current velocity of particle  $i$  at time  $t$ ,  $w$  is inertia weight.  $c1$  acceleration coefficients and  $c2$  are control the influence of pbest and gbest,  $r$  is a random number between 0 and 1.

Based on each particle best-known position, update its individual best position. The global best position is updated based on identifying the particle with the best fitness among the total population. The pbest position is where the particle obtained the best fitness. Based on the particle's current velocity, the difference between its current location and best, and the difference between its current position and best, update each particle's velocity and position. According to the previously mentioned equations, the velocity update enables particles to move toward promising regions while exploring the search space. Fittest, gbest, velocity update, position update, and pbest iterations should all be repeated for a predetermined number of times, or until a stopping criterion is satisfied.

#### Genetic algorithm

Genetic algorithm (GA) is the name given to an EA technique developed in 1975 that is modeled after natural selection<sup>38</sup>. It evolves a population of potential solutions across multiple generations to determine the optimal answer to a problem. The following provides a detailed description of each step in the genetic algorithm pseudo code of Algorithm 2 based on Darwin's theory: At the start of the process, a population of individuals selected at random is initialized. Each chromosome represents a distinct strategy for solving the problem. The code in question consists of characters from a predefined character set into random phrases. Every fitness level of a chromosome is calculated. Here, fitness refers to a chromosome's proximity to the desired result. Fitness is evaluated with the EAs by counting the characters that match between the desired sentence and the problem. Chromosomes are selected from the existing population to act as parents for the next generation. Fitness plays a part in selection; fitter people are more likely to be chosen. Candidates are selected with probability based on various selection techniques and fitness levels. To create offspring, selected parent pairs undergo crossover.

**Input:** Fitness function, maximum number of generations, current generation

**Output:** Best population

**Start**

1. Initialize Population //Generate an initial population of candidate solutions.
2. Evaluate Fitness value // Calculate the fitness of each individual in the population.
3. **While** Termination Condition Met ? // Check if the termination condition (e.g., finding the optimal solution or reaching a maximum number of generations) is met.  
     No: Proceed to the next steps.  
     Yes: The algorithm terminates.
4. Select Parents: Choose individuals from the current population to act as parents for the next generation. Individuals with higher fitness are more likely to be selected.
5. Crossover and Mutation: Apply crossover and mutation operators to create offspring individuals.
6. Create Next Generation: Combine parents and offspring to form the next generation.
7. **End while**
8. Fittest individual.

**End**

**Algorithm 2.** Pseudo code of Genetic algorithm.

#### *Self-adaptive genetic algorithm*

Genetic algorithms are robust optimization techniques, but usually, canonical ones concentrate on a limited number of parameters such as population size, number of crossovers, and mutation rate for each population. To improve the usability of GAs and increase efficiency in a dynamic environment, self-adaptation is an effective technique with several merits compared to canonical genetic algorithms, a difficult and time-consuming process<sup>39</sup>. This approach self-adapts to crossover and mutation-related parameters in other systems. We evaluate self-adaptive genetic algorithm performance on the BC type prediction issue against that of SVM, and the canonical genetic algorithm can be seen in Eqs. (8–9).

$$P_c = \begin{cases} \left[ 1 - c_1 \frac{(f_{max} - fb)}{(f_{max} - f_{avg})} \right] - c_2, & fb \geq f_{avg} \\ c_3, & fb < f_{avg} \end{cases} \quad (8)$$

$$P_m = \begin{cases} \left[ 1 - c_4 \frac{(fb)}{(f_{max} - f_{min} + f_{avg})} \right] - c_5, & fb \geq f_{avg} \\ c_6, & fb < f_{avg} \end{cases} \quad (9)$$

Here  $f_{max}$ ,  $f_{min}$ ,  $f_{avg}$ ,  $fb$  represent the maximum value of the learner, the minimum value of the learner, the median value of the learner, and the top value after crossover and mutation in learners by the selection operation, respectively.  $c_1$  to  $c_6$  represent the constant value between 0 and 1. GA eliminates the requirement for domain experts to comprehend and adjust any operator-related parameter settings by self-adapting. We study how well it works to let EA make its parameter adjustments during a run, reducing the need for human parameter adjustment and improving the usability of EA for non-experts. Because of its self-adaptive methodology, the EA evolves many parameter values that specify the algorithm's operation and potential solutions to a target problem. To cover a range of genetic operator behaviors, we incorporate several crossover and mutation operators in the list of possible operators. Algorithm 3 displays the SA-GA pseudocode.



**Input:** Fitness function, maximum number of generations, current generation

**Output:** Best population

**Start**

1: Initialize population of candidate solutions

2: Evaluate population (fitness)

3: **while** condition do

4: **for** probabilistically selecting parents for the next generation (selection), for each set of parents do

6: probabilistically select the crossover operator according to encoded weights (see Eq. 8)

7: Apply crossover to parents to generate children

8: **end for**

9: **For** each child, do

10: Probabilistically select the mutation operator according to encoded weights (see Eq. 9)

11: Apply a mutation to the child

12: **end for**

13: Replace population with children

14: **end while**

15. Fittest individual.

**Algorithm 3.** Pseudo-code of self-adaptive genetic algorithm (SA-GA).

### Proposed methodology for classification system

Previous attempts at developing self-adaptive-based metaheuristics have not successfully retained the appropriate feature sets to satisfy the various demands of practical medical systems. As a result, we recently formed a quantum self-adaptive TLBO with SA-GA, which is influenced by students' intrinsic social intelligence and utilized to identify novel disease kinds. It also contains fewer control parameters, which reduces the cost of solving the FS issue in terms of convergence rate and computing time. Due to the population effect divergence, some individuals can quickly escape from the local optimum. However, most search learners descend into local optima, especially when the optimal solution also falls, so SeQTLBOGA, with SVM, can deteriorate.

### Quantum representation for feature selection

Feature selection aims to choose an optimal feature subset from all available features. Quantum representation for feature selection utilizes quantum computing techniques to improve the efficiency and efficacy of discovering pertinent features in BC datasets. It can simultaneously explore numerous feature combinations through quantum parallelism by encoding conventional data into quantum states, such as qubits or quantum superposition states. This facilitates a more rapid assessment of feature subsets in contrast to traditional methods, which frequently depend on iterative or greedy techniques.

Furthermore, quantum feature selection can leverage entanglement and interference to reveal correlations between features that classical methods may neglect. As a result, we reasoned for the representation of the features using qubits, with one qubit standing in for each feature. A series of qubits is used to represent every feature. If the state value of a qubit is  $|0\rangle$ , it indicates that the feature is selected, and if it is  $|1\rangle$ , it indicates that the feature is not selected. A string of qubits of length  $n$ , where  $n$  is the number of features, represents a population  $Q_{gi}$ . Following is a representation of  $Q_{gi}$  as follows:

$$Q_{gi} = \begin{vmatrix} \alpha_1 & \dots & \alpha_n \\ \beta_1 & \dots & \beta_n \end{vmatrix}$$

$|\alpha|^2 + |\beta|^2 = 1$ , where  $|\alpha|^2$  and  $|\beta|^2$  are the probabilities of states 0 and 1, respectively. Quantum population represents the superposition of the population positions, which contains  $2^n$  possibilities. The diversity of the population can be achieved with fewer individuals.

---

```

Input :  $Q_{gi}$  = i-th quantum student
Output :  $B_{gi}$  = i-th quantum student
For each qubit j in  $Q_{gi}$  do
  Generate a random number in between 0 and 1.
  If  $r > \alpha^2$  then
     $B_{gi,j} = 1$ 
  else
     $B_{gi,j} = 0$ 
  End if
End for

```

---

**Algorithm 4.** Pseudo code of quantum observation.

---

#### Quantum observation

The quantum student can through from quantum superposition to a definite state through quantum observation. A uniformly distributed random number  $r$  in the range  $[0, 1]$  is generated for each qubit. The relevant bit is assigned a value of 1 if  $r > \sin^2(\theta)$ , else a value of 0. Algorithm 3 displays an observation algorithm.

#### Distance between two binary learners

The SQTlBOGA method makes use of a Hamming distance version. In general, one can measure the difference between two symbol sequences using the Hamming distance. The number of unique bits between two binary answers in the BC problem is represented by the utilized distance. The proposed distance provides the ratio of the number of items not selected in the two binary students' solutions to the Hamming distance.

If  $d(B_{gi,k}, B_{gj,k})$ , where  $B_{gi,k} \neq B_{gj,k}$  then  
 $B_{gi,k} = 1$   
 else  
 $B_{gj,k} = 0$

$$\delta(B_{gi}, B_{gj}) = \sum_{k=1}^m d(B_{gi,k}, B_{gj,k})$$

$$d(B_{gi}, B_{gj}) = \frac{\delta(B_{gi}, B_{gj})}{2m - (\sum_{k=1}^m B_{gi,k} + \sum_{k=1}^m B_{gj,k})}$$

where  $m$  is the size of a binary learner solution.

#### Quantum movement according to learner strategy

The probability amplitude state of quantum students is updated using a rotation gate; the update formula is presented below in Eq. (10).

$$\begin{pmatrix} \alpha^d(t+1) \\ \beta^d(t+1) \end{pmatrix} = \begin{bmatrix} \cos(\Delta\theta^t) & -\sin(\Delta\theta^t) \\ \sin(\Delta\theta^t) & \cos(\Delta\theta^t) \end{bmatrix} \begin{pmatrix} \alpha^d(t) \\ \beta^d(t) \end{pmatrix}. \quad (10)$$

$\theta_i^t = \Delta_d^i s(\alpha_d^i, \beta_d^i)$ , represents the qubit rotation angle, where  $\alpha_d^i$  is the size of rotation and  $s(\alpha_d^i, \beta_d^i)$ , is the direction of rotation which we can get from Table [56]. This strategy aids in maintaining a balance between early-stage exploration and late-stage exploitation. By leveraging historical information from the search process, the proposed approach can detect potential populations where the population becomes trapped in local optima. This enables the population to break free from the current local optima, enhancing its exploration capability.

#### Weighted teaching learning-based optimization

TLBO method uses inertia weights assigned to distinct populations based on the rank of the population<sup>44</sup> as demonstrated in the inertia weight factor, which observes the scenario to search and adjust the inertia weight value based on multiple feedback values in Eq. (11).

$$\omega_i = \frac{Te_i^*}{T_{Pop}} (\omega_{max} - \omega_{min}) + \omega_{min}. \quad (11)$$

Where  $Te_i^*$  is the rank of the  $i^{\text{th}}$  population when the population is ordered based on its population's best fitness,  $\omega_{min}$  and  $\omega_{max}$  is constant and  $T_{Pop}$  denotes as total number of populations.

### Teacher updating phase

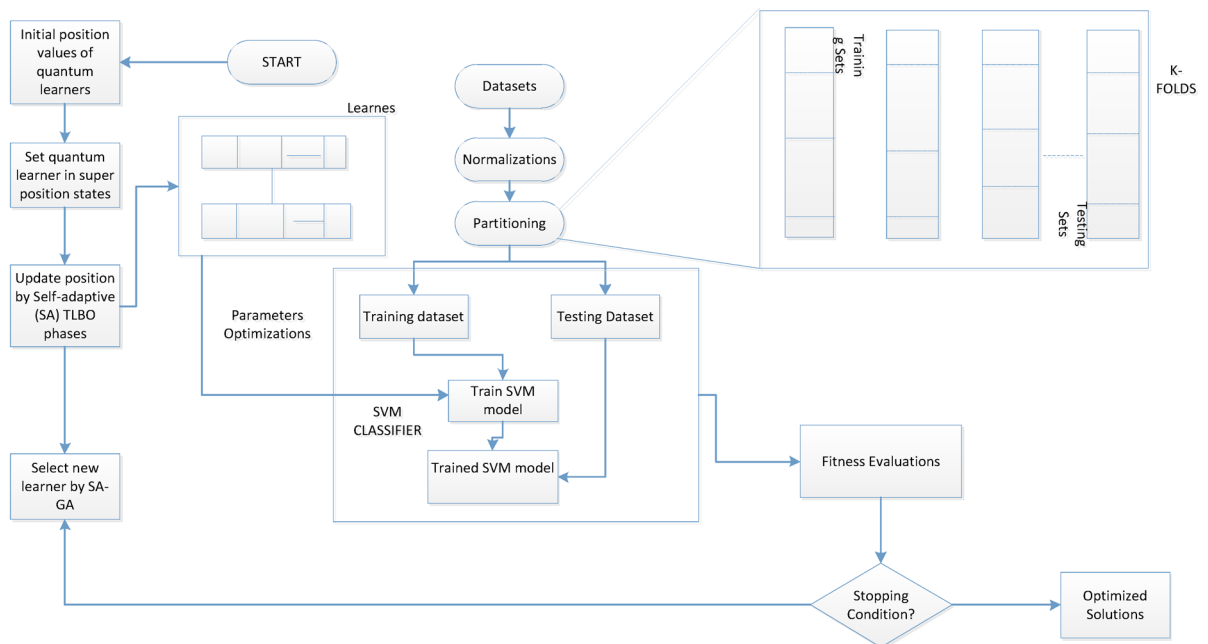
The metaheuristic approach known as teaching-learning-based optimization (TLBO) was motivated by the teaching and learning process in the classroom. The population is considered a group of pupils, and a fitness evaluation is used to choose the best student to be the teacher. The Teacher and Learner stages are the two distinct phases of the TLBO procedure. To avoid premature convergence, we use the weighted technique in TLBO to resolve problems of this kind. We seek to improve the learning represented in Eq. (12) using the wTLBO approach. This stage involves analyzing Eq. (2), knowing the precise solution (\*) and the person utilized in the subsequent iterations. Using the fundamental TLBO algorithm ( $Te^*$ ). We generate the learners collectively from the previous position and best position; here, we introduced a new position-updating rule through an inertial weight strategy on basic TLBO, whereby the position of every class member is updated by the following Eq. (12). First, we randomly select learners for the full class. Second, we assess the best values of these two learners and get the best one.

$$X_{i,k}^{new} = \omega_i * X_{i,k}^{old} + r * (X_{Te^*,k} - T_f * M_{i,k}) \& T_f = \text{round} [1 + \text{rand} (0, 1)] \quad (12)$$

Where,  $X_{i,k}^{new}$  and  $X_{i,k}^{old}$  consist of the learner's new value and old values,  $r$  is a random number in the range  $[0, 1]$ ,  $T_f$  is the teacher factor, with a range of  $1, 2$ , random  $M_{i,k}$  is created from the learner's desired mean values, and  $Te^*$  represents the top learner in the current iterations. For learners  $p$  and  $q$ :  $|\psi_i\rangle \otimes |\psi_q\rangle \rightarrow |\psi_p\rangle \otimes |\psi_q \oplus \psi_p\rangle$ .

This ensures feature subset changes in one learner directly influence others, mimicking group learning. Instead of fixed  $r$ , crossover probabilities derive from qubit amplitudes ( $p_{\text{cross}} = |\beta|^2$ ) promoting diversity. Mutation rates scale with qubit phase differences:  $p_m$  proposanal to  $\arccos(|\langle\psi_p|\psi_q\rangle|)$ . This prevents stagnation in local optima, reducing feature redundancy. Quantum-evolved solutions undergo elitist selection and mutation:  $p_c, p_m$  over generations. This refines feature subsets while preserving quantum-enhanced diversity. By merging quantum superposition with the TLBO teaching paradigm, SeQTLBOGA achieves faster convergence and higher precision in BC classification, addressing TLBO's historical limitations in feature selection stability.

The following advancements are incorporated into our process. To enable the system to use previous data or information from other patients, primary data must first be added to the records. Second, each action necessitates the algorithm to examine patient data, and a learner's estimated knowledge can only be updated following the action. Third, our technique uses self-adaptive quantum TLBO and GA to minimise unnecessary characteristics from BC datasets. Creating a framework as shown in Fig. 1 is one possible solution for the aforementioned issues. The Boolean operator XOR ( $\otimes$ ) connects the best learner as the teacher and the learner. This is the first kind of work that uses self-adaptive (SA) processing with quantum TLBO and GA, which calls for many important advancements in comparison to the regular EAs learning works. Our framework uses SVM learning algorithms and feature selection that is inspired by nature to identify a patient's illness early.



**Fig. 1.** Optimizing parameters of SVM by the new optimizer SeQTLBOGA for breast cancer types detection.

*Self-adaptive teaching learning-based optimization*

Studies conducted in the past have mainly concentrated on the self-adaptive (SA) processing of teaching elements. These factors are crucial for teachers because they indicate how much they impact students' average performance. For instance, in 2020<sup>40</sup>, the author presented a hybrid adaptive teaching-learning-based optimization. Unlike in the original text, teaching factors are not a fixed value in this article. A strategy of adaptive selection can change its value.

$$[T_f = \begin{cases} 2, & \text{if } RP < \epsilon \\ 1, & RP \text{ otherwise} \end{cases}$$

In the above formula,  $RP$  represents the learner's ranking probability. This article establishes a threshold condition to determine the interim method for computing teaching factors.  $\epsilon$  is 0.5 as the threshold. Using the first technique, the teaching factor equals 2 when  $RP$  is greater than 0.5. If the second option is selected, the teaching factor remains at 1. In contrast to the ones mentioned before, the adaptive strategy in this study intends to process the instructional components adaptively by applying a new formula. The precise equation is as follows:

$$T_f = \cos\left(\frac{i\omega * it}{Maxit}\right). \quad (13)$$

The new teacher phase is defined by the following Eq. (14);

$$X_{i,k}^{new} = \omega_i * X_{i,k}^{old} + r * (X_{Te^*,k} - T_f * M_{i,k}). \quad (14)$$

In the above Eqs. (13) and (14),  $i\omega$  and  $it$  refer to the inertia weight and current iteration, respectively. The  $Maxit$  refers to the maximum iteration. By this formula of Eq. (14), the teaching factor ( $T_f$ ) adaptively changes. Other steps of self-adaptive quantum on TLBO with GA are the same as those of the original algorithm. The teacher and student levels updating process is still based on the original formulas (16). The SeQTLBOGA algorithm pseudo-code is given in Algorithm 5.

**Begin**

**Input:** Initialization population size 'Pop',  $\epsilon$ ,  $\varphi$ , inertia weight as  $iw$ , Max number of iteration:  $t_{max}$ , and termination criterion; D: training set, F: all features`

**Output:** Best Optimal Solution

---

```

for i= 1: nPop                                     //Each search learner
Y_i = initialize ( Population B_g );                 //Generate a random population with D
Get binary student population B_g by using Algorithm 4,
end for
Estimate  $X_{mean} = \sum_{i=1}^{i=NP} X_i / nPop$ 
Z=fitness (Y_i,  $\epsilon$ )                                // Estimate the fitness of each learner (Y)
Te* = the best search learner;
 $T_f = \cos(\frac{i\omega * it}{Maxit})$                       // self-adaptive  $T_f$ .
While (t < t_max )
t=t+1;
Diff_Vector = (Te* -  $T_f * X_{mean}$  )
For i= 1 : nPop
For k=1 : D
 $X_{i,k}^{new} = iw * (X_{i,k}^{old}) + Diff\_Vector(k)$ 
 $\varphi_{i,k} = self\_crossover(X_{i,k}^{new} \otimes Te^*)$           //Apply self-adaptive crossover operator of GA
Estimate fitness value ( $\epsilon$ ) of each updated learner  $\varphi_{i,k}$ 
End for
if (f ( $X_{i,k}^{new}$ ) < f ( $\varphi_{i,k}$ ))
 $\varphi_{i,k} = X_{i,k}^{new}$ 
End if
End for
For i= 1 : nPop
Randomly select the other learner Xp where {p  $\neq$  i}
 $\mathfrak{M}_i = self\_crossover(\varphi_{i,k} \otimes Xp)$               //Apply self-adaptive crossover operator of GA
 $\mathfrak{B}_i = \mathfrak{M}_i$                                        // select best population
For k=1: nPop
 $\eta_i = self\_mutation(\mathfrak{B}_k)$                         //Apply self-adaptive mutation operator of GA
End for
Estimate fitness value of each updated learner  $\eta_i$ 
if (f( $\eta_i$ ) < f(  $X_i$  ))
 $X_i = \eta_i$ 
End if
End for
Increase fitness value and reduce features;
End while
End

```

---

**Algorithm 5.** Pseudo-code SeQTLBOGA for feature selection.

In addition, evaluate the performance of SeQTLBOGA, we have used SVM as fitness function in the self – adaptive and quantum TLBO (see Fig. 1) and find the best agents with optimal number features.

The proposed processes for the optimized SVM model are as follows, as seen in Fig. 1. Set up the fundamental TLBO and GA settings. Give the students  $n + 2$  dimensions to encode. Finally, there are two dimensions: C and

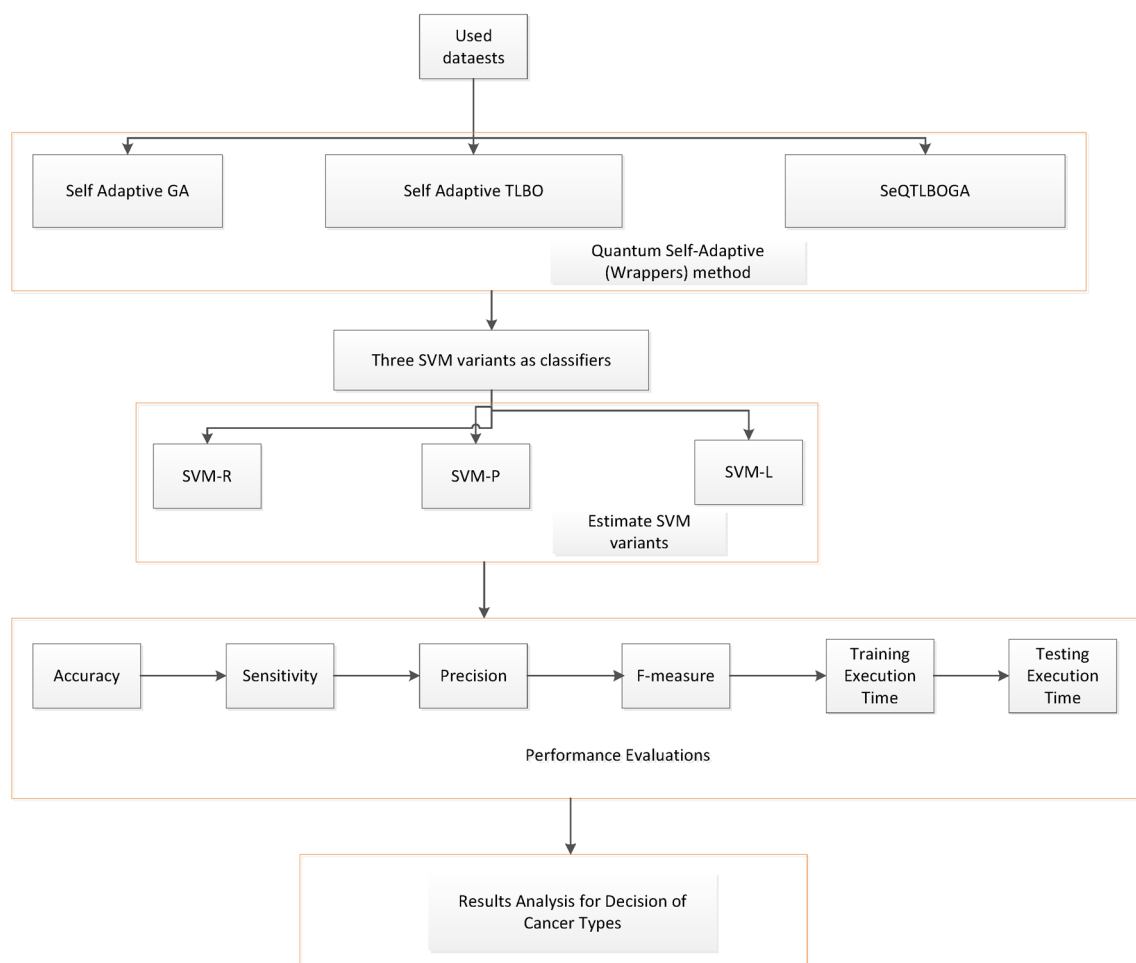


$\sigma$ . The values of the  $n$  dimensions are features, where 1 denotes that a feature is chosen and 0 denotes that a feature is discarded. Update the self-adaptive notion and apply the quantum concept to the teacher phase. Train the SVM and assess the learner fitness (FA) for every search. The ideal SVM parameters are obtained, and the learner fitness value is maintained if it approaches the best FA. A learner can get a high FA if they have a modest number of carefully chosen characteristics and a high classification accuracy (CA).

Furthermore, as the number of feature subsets is related to the SVM generalization error, the learner that scans fewer features obtains a more enormous FA. Thus, we consider each of them in this study and build the Fitness function as Eq. (15). To reach the FA, update learners through the self-adaptive genetic algorithm phase. Using the training dataset and the SVM optimal parameters, obtain the learning engine. To determine the best FA, use the model to predict test results. Obtain the feature subset, optimal C, and  $\sigma$  from the best learner (as Teacher).

### Optimizing parameters using proposed method

SeQTLBOGA, the most recent search technique used to identify the best answers to various real-world problems, is shown in Fig. 2 for predicting BC illnesses. It is a global search heuristic that mimics socially inspired behavior that humans influence. The first learner in TLBO is created arbitrarily using a set of candidate solutions, and it is then quantized using the idea of quantum computing<sup>41</sup>. In the instructor phase of the TF, an adaptive teaching factor takes the place of the original TLBO teaching factor. The algorithm's overall performance is improved because the adaptive teaching factor can successfully balance search efficiency and convergence speed. After all potential learners have been assessed using fitness values, the best learner among the current learners is selected as the instructor. Other students begin to move in that direction as the group leader draws attention to themselves by drawing another person to the vicinity. GA searches the solution space by simulating survival of the fittest and using that evolution. To address the BC problem, they explore every region of the state space and exploit potential locations by applying adaptive behavior of crossover and mutation to individuals within the population. SeQTLBOGA finds kernel function parameters and feature vectors simultaneously. Learners must, therefore, possess two of these elements. In these experiments, we used the RBF kernel function for two



**Fig. 2.** SeQTLBOGA implementation approach for prediction of BC disease.

reasons. First, the kernel can handle scenarios with a nonlinear relationship between the features and class labels because it places samples into a higher-dimensional space. The second explanation relates to the number of hyperparameters  $C$  and  $\sigma$ , which influence how complex the model is. The RBF kernel requires a maximum of two parameters.

Like other EAs, basic TLBO becomes locked in local optima when utilized with varying optimization challenges. As a result, the self-adaptive nature is combined with GA in TLBO, allowing the top learners to depart from local optima. Parameters  $\sigma$  and  $C$  of the SVM are optimized utilizing TLBO amalgamation of GA to combine the quantum search capability notion with the investigation of SeQTLBOGA. The process for SVM parameter optimization using SeQTLBOGA is divided into two parts, as shown in Fig. 1. a, SVM parameter encoding and population formation; b, fitting value computation for each learner based on Eq. (15). The calculation process involves the use of the testing dataset to classify whether or not diseases are present using an expert detection engine, the cross-validation of SVM model prediction, the application of the SeQTLBOGA model to train the SVM model for generating new solutions, and the selection of the best solution obtained by optimizing the SVM tuning parameter.

### Fitness function

The quality of each produced individual is assessed using predefined parameters. For this use, a fitness function should be defined as Eq. (15). Our model fitness function is determined by averaging the classification accuracy of SVM models trained using user-supplied data and parameters. The function of fitness, which is expressible, can be utilized to assess the quality of each individual  $I_i^t$  at iteration  $t$ , where  $\text{lab}(x_j)$  is the classification result of the  $j^{\text{th}}$  instance in the testing data set,  $x_j$  is the actual class label of the  $j^{\text{th}}$  instance,  $\theta$  is the relation between  $c(x_j)$ , and  $y_j$  that is if  $\text{lab}(x_j) = y_j$  then  $\theta = 1$  otherwise  $\theta = 0$ .  $N$  is the number of instances in the testing data set. Note that the fitness access is measured based on  $k$ -cross-validation, where  $K$  denotes the number of folds. The latter step is essential to avoid the problem of overfitting and, consequently, to obtain more robust results. It can be referred as cross-validation to as an internal cross-validation

$$\text{Fitness}_{\text{value}} = \frac{1}{K} \sum_{k=1}^K \frac{1}{N} \sum_{j=1}^N \theta(\text{lab}(x_j), y_j). \quad (15)$$

### Computational complexity

The computational complexity of SeQTLBOGA is expressed as  $O(m \cdot n \cdot T \cdot \log k) + O(T * n) + (n * m + \Psi)$ .

Several key factors, including population size, dimensionality, chaos-inspired developments, self-adaptive strategies, and genetic algorithms, influence the computational complexity of the self-adaptive quantum Teaching Learning-based Optimization Algorithm, SeQTLBOGA. Understanding the complexity of SeQTLBOGA requires breaking down its significant computational steps and analyzing their impact on overall efficiency. Where,  $m$ : Population size,  $n$ : Number of features,  $T$ : Generations until convergence,  $k$ : Reduced feature subset size ( $\log k$  term from quantum mechanism),  $\Psi$ : Computation Training. SeQTLBOGA reduces time complexity compared to basic methods like GA, PSO, DE through wrapper-inspired and adaptive learning. For large-scale datasets, our quantum approach may complement its strengths.

The final computational complexity of our technique is expressed:  $O(m \cdot n \cdot T \cdot \log k) + O(T * n) + (n * m + \Psi)$ . Suppose that  $m$ : Population size,  $n$ : Number of features,  $T$ : Generations until convergence,  $k$ : Reduced feature subset size ( $\log k$  term from quantum mechanism),  $\Psi$ : Computation Training. According to our algorithm, population initialization for computation training is  $\Psi$  times. It needs  $O(n * \log k)$  time complexity to transform qubit learners to binary learners. The position of each learner needs to be updated, which takes  $O(\log k)$  time complexity. An algorithm uses a Self-adaptive strategy to find a better solution, which takes time complexity  $O(n)$ . Therefore, the overall time complexity of our algorithm is  $O(m \cdot n \cdot T \cdot \log k) + O(T * n) + (n * m + \Psi)$ .

### Model interpretation

#### Explainable AI models

This section addresses the strategies and tactics that can be applied to improve the intuitiveness of machine-learning models for human comprehension. In the context of our method, explainability is the capacity to comprehend and analyze the choices or forecasts made by a model. This entails bringing the evolutionary system decision-making process into clear and human-understandable terms. Furthermore, it is essential to foster systemic trust, particularly in BC applications, where poor decisions can have grave repercussions. Several strategies for attaining explainability have been examined in previous research, including using interpretable models, producing explanations for intricate models, and creating models with naturally interpretable structures<sup>42,43</sup>. Developers can guarantee that the model is accurate but also dependable and trustworthy by integrating Explainable AI (XAI) approaches. Shapley Additive Explanations (SHAP) in the XAI technique evaluate the impact of individual attributes on machine learning model predictions<sup>15</sup>. The model baseline performance is assessed by comparing the data. Consequently, each feature is evaluated by permuting its value randomly, reassessing the model performance, and determining how it deviates from the baseline. A negative score suggests a potential dependency on other variables, whereas a more considerable performance decline denotes more significant elements. Additionally, the ability of our model to continue operating and producing accurate forecasts in the face of unanticipated or hostile conditions is referred to as the proposed method of resilience in this context.

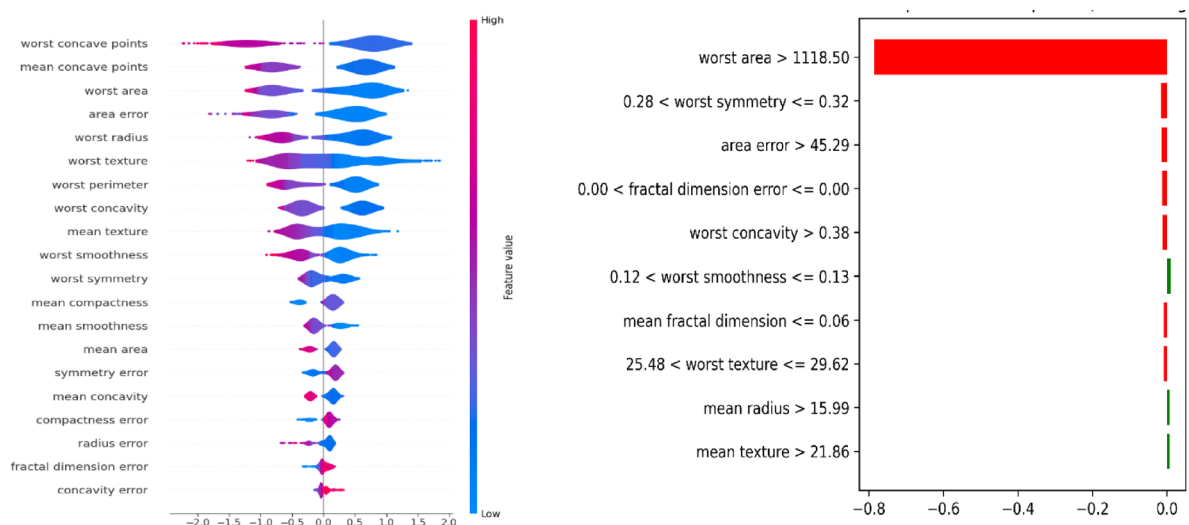
Additionally, the well-liked method for providing a local and interpretable explanation for any machine learning model's predictions is called Local Interpretable Model-agnostic Explanations (LIME)<sup>14</sup>. This is especially helpful when complicated models such as ensemble methods are involved, and it is necessary to comprehend

how a model came to a specific conclusion. This model is mainly used in applications related to breast cancer, where the input data may be dynamic, diversified, and confront various difficulties. guarantees that this model is not unduly sensitive to small input changes and generalizes well. We employ data augmentation, adversarial training, and hybrid techniques to address input uncertainties to enhance resilience. The total dependability and credibility of our built system are influenced by both robustness and explainability. In terms of explainability, our model can aid in problem diagnosis and comprehension of the various ways the model operates under multiple circumstances, assisting in detecting and mitigating potential vulnerabilities. However, this model is more likely to instill trust due to its robustness and ability to function well under various circumstances. In essence, a new AI method is proposed with proof of both robustness and explainability. This comprises assessments using reference BC datasets to gauge the model performance in various scenarios and data metrics to reveal how our model makes decisions. Our model architecture and performance algorithms are shown in Fig. 3. The decision-making process helps clarify the model architecture.

### Future directions and limitations

In addition to the above contributions, many other aspects of this study have the potential to be advanced in some exciting ways. Therefore, this section suggests and discusses some aspects of future research. Although SeQTLBOGA has several benefits for diagnosing breast cancer, there are also some possible disadvantages, especially with regard to feature subset variability, dataset size, and overfitting concerns. Large, varied datasets are essential for the model's success; smaller or unbalanced datasets may result in less-than-ideal feature selection and less generalizability. Another issue is overfitting, which occurs when high-dimensional feature spaces are used because the quantum-inspired optimization process may over-specialize to training data noise. Inconsistent clinical interpretations may also be introduced by variations in chosen feature subsets among runs or datasets, which could erode confidence in the model's predictions. Future studies could investigate hybrid AI models that combine deep learning with quantum evolutionary algorithms to address these problems. These models would combine the representational strength of neural networks with the reliable search capabilities of SeQTLBOGA. Examining multi-objective optimization techniques that strike a compromise between feature selection effectiveness and classification accuracy is another exciting avenue to pursue. This will guarantee that the model retains good performance while reducing redundancy. Additionally, by utilizing pre-trained feature selection paradigms, transfer learning frameworks can expand the applicability of SeQTLBOGA to other medical classification issues. These developments could improve the model's therapeutic usefulness, scalability, and dependability in a range of healthcare situations. In the updated manuscript, we also referenced a study on medical AI applications that used quantum evolutionary optimization. Overfitting, in which models learn noise rather than generalizable patterns, is made worse by small datasets. This results in inflated training accuracy but subpar test performance. Rotation gate learning rates are an example of a quantum metaheuristic with a high parameter complexity that may overfit to training data. Rankings of features' importance are inconsistent. irrelevant feature selection, like giving noise precedence over biomarkers with clinical validation. Due to their extreme sensitivity to initialization, quantum algorithms generate feature subsets that vary from run to run. Clinical adoption is made more difficult by this instability since inconsistent biomarkers erode diagnostic confidence. For automated feature extraction, integrate deep learning architectures with quantum feature selection in the future.

The significant advantage of our algorithm is that it selects the best relevant features and also removes inconsistent features that cause the degradation of the accuracy of the classification algorithms. This algorithm plays a significant role in the accurate classification of BC data sets that work as a single objective. The disadvantage of our method is that it does not take advantage of parallel concurrent processing of high-dimensional data. Moreover, current studies should also inquire about patients who have completed treatment on what they



**Fig. 3.** Impact of Model in SHAP and LIME.

believe to be the most crucial aspect in determining treatment release. We believe that the proposed system would be beneficial to doctors in making their final decisions on their patients. Therefore, our technique has a significant advantage in correcting breast cancer classification to achieve accurate decisions compared to existing algorithms, and on-time and accurate treatment may be provided to the patients. In future studies, we will use high-dimensional multi-class classification data to study breast cancer diseases that affect the performance of multi-classification methods.

## Experimental study and analysis

### Datasets description

Breast cancer is brought on by breast tumors, or abnormally developing cells in the breast tissue. The two BC dataset types, WBCD (Wisconsin Breast Cancer Database) and Wisconsin Diagnostic Breast Cancer (WDBC)<sup>44</sup>, obtained from human breast tissues, are categorization datasets that document breast cancer case measurements. Digital pictures of the breast tissues of the patients were used to compile 569 cases of the WDBC cancer dataset. The features mean, standard error, and worst or largest (mean of the three most significant values) were computed for each image, resulting in thirty features excluding id and label. It is split up into two categories. The WDBC dataset<sup>45</sup> has 212 (37.2%) cases classified as malignant tumors and 357 (62.8%) cases classified as benign tumors. Only 699 suitable samples remained, and they were the only ones evaluated. The collection contains 65.5% (458) benign samples and 34.5% (241) cancer samples. Nine attributes defining the BC features are collected for each case sample. Data from 699 patients' digital breast images are stored in WBCD. In each patient sample, 9 tumor traits are displayed<sup>46</sup>. 699 instances total from patient needle aspirates are included in the collection; 458 of these cases belong to the benign class, and the remaining 241 cases to the malignant class. Nine qualities are linked to every record in the database.

### Experimental setup

We have implemented our SeQTLBOGA approach, which employs EAs directly, in MATLAB R2016a on Windows 10 with an Intel®Core™ i7 processor and 16 GB of RAM. We used the Wisconsin Breast Cancer Databases utilizing this approach. The usual parameter settings were used in this paper to ensure a fair comparison. The following is the configuration for the SeQTLBOGA detail parameter.

#### Parameter configurations

For the EAs approached, primary parameters were established. These are the number of learners (population size) and the number of iterations is assigned the values 30 and 100, respectively. As a result, all EA algorithms, kept the population size and number of iterations constant. In this investigation, the number of iterations, population size, crossover probability, and mutation probability were the four primary parameters that is established for the GA technique corresponding values are 100, 30, 0.7, and 0.4. Quantum Rotation Angle ( $\theta$ ) is  $0.01\pi$ . Learning Rate set as 0.01, tournament size set as 3. Inertia Weight ( $w$ )=0.4, Cognitive Coefficient ( $c_1$ ), and Social Coefficient ( $c_2$ ) set as 1.4; in DE the value of differential weight ( $F$ ) set as 0.5.

### Performance measures

Depending on the particular application at hand, a variety of measures can be used to assess the EA algorithms. Researchers frequently use performance indicators like sensitivity and specificity to evaluate systems. In this part, we establish a set of performance indicators that we use to assess how well our technique is working. The performance metrics, such as accuracy, sensitivity, specificity, F-measure, and area under the curve (AUC), are shown as follows:

$$Accuracy = \frac{TN + TP}{TP + TN + FP + FN} \quad Sensitivity = Recall (Re) = \frac{TP}{TP + FN} \quad Precision (Pr) = \frac{TP}{TP + FP}$$

$$F_{score} = \frac{2 * Pr * Re}{Pr + Re} \quad AUC = \frac{(Sensitivity + Specificity)}{2} \quad Specificity = \frac{TN}{FP + TN}$$

Additionally, True Positive Rate ( $TPR = \frac{TP}{P}$ ), True Negative Rate ( $TNR = \frac{TN}{N}$ ), False Positive Rate ( $FPR = \frac{FP}{FP + FN}$ ), and False Negative Rate ( $FNR = \frac{FN}{TP + FN}$ ) are used to examine the proposed approach. Here, TP is the number of true positives; FN, the number of false negatives; TN, the number of true negatives; FP, the number of false positives.

### Results analysis and discussion

Our method is evaluated on WBCD and WDBC in this section. The proposed method applied with SVM training is compared to different up-to-date existing methodologies for detecting cancers by the expert system, including GA, DE, CS, PSO, and GOA. Results show that the proposed approach outperforms other EAs that are selected from up-to-date, relevant research found on used WBCD data using various modeling tools in previously referred literature<sup>47</sup>. All of the results presented are the averages of ten runs.

#### Comparison of feature selection methods using SVM

In situations where a complete evaluation appears unfeasible, the advantage of EAs like GA, DE, CS, PSO, and GOA is that they make it easier to select the best feature sets that promise improved performance. It chooses from two datasets the appropriate features. SeQTLBOGA, in conjunction with various evolutionary approaches and SVM learning variations, helps with classification optimization problems.

This section displays the results obtained from running our system on the WBCD and WDBC datasets. As a result, Table 1 displays the classification results of our system for WBCD and WDBC datasets, taking into account different setups of SVM classifiers and EAs-based feature selection techniques. This section presents the results of the evaluation of our method using five traditional EA-based studies as a comparison. Based on average metrics from ten runs of 10-fold cross-validation, these results are presented.

SVM-R classifier outperformed the other SVM classifier, based on Table 1 results. Additionally, proposed approach utilizing SVM-R performed better than the other methods, yielding outstanding outcomes when taking into consideration percentages for accuracy, sensitivity, F-measure, and specificity. Table 1 illustrates the efficacy of several wrapper techniques such as Genetic Algorithm, Particle Swarm Optimisation, Cuckoo Search, Grasshopper Optimisation Algorithm, and proposed method when utilised with SVM classifiers employing various kernels (RBF), Polynomial, and Linear on the WBCD and WDBC datasets. Our technique consistently surpasses GA, PSO, CS, and GOA across all SVM kernels. The SVM-RBF attains the best accuracy as 96.11%, sensitivity as 92.62%, specificity as 93.54%, and F-measure as 92.11% in WBCD, indicating exceptional discriminative capability in feature selection. Likewise, for SVM-Polynomial, our technique achieves an accuracy of 89.62%, markedly surpassing the next-best GA at 82.38%, while preserving balanced sensitivity as 82.62% and specificity as 80.65%. In the instance of SVM-Linear, the suggested methodology achieves an accuracy of 82.67%, although with a narrower margin, indicating that non-linear kernels (RBF, Polynomial) derive greater advantages from sophisticated feature selection.

The suggested method demonstrates superior performance, especially with SVM-RBF, with 92.91% accuracy, 92.12% sensitivity, and 93.65% specificity, signifying its resilience in managing WDBC data. The SVM-Polynomial achieves an accuracy of 90.87%, much above that of conventional optimisers GA as 77.63%, PSO as 74.88%, whilst the SVM-Linear demonstrates an accuracy of 88.72%, further affirming its generalisation capacity. The F-measure consistently exhibits a high value, reflecting a robust balance between precision and recall. Classical optimizers such as GA, PSO, CS, and GOA demonstrate modest efficacy, with GA typically outperforming the others, indicating that evolutionary algorithms are more consistent than swarm-based techniques like PSO, GOA. Non-linear SVM kernels (RBF, Polynomial) derive the greatest advantage from the suggested feature selection, presumably due to their superior management of intricate feature interactions.

Table 2 presents a comparison between the proposed method and the current SVM-R technique with respect to training and testing accuracy and estimated run time. Based on ten rounds of tests using both datasets, these conclusions have been drawn. Based on the statistics, it can be concluded that the proposed method consistently yields more reliable and consistent results than its five rivals.

Comparative studies in terms of performance

In this section, our method outcomes and debates are shown alongside five alternative wrapper strategies that use an SVM-R classifier. Table 2 shows that the training accuracy rate for the proposed system is 98.95, and the WBCD execution time is 877.98 (in seconds). Additional analysis has been done on the WDBC dataset 97.43, as Table 2 displays the training accuracy rate of 544.89 (in seconds). It provides an overview of the BC datasets-focused performance analysis. The results of our strategy and additional methods are displayed in the table, accounting for factors. As seen by Fig. 4, the F-measure rate, we can demonstrate that our approach outperforms existing state-of-the-art EAs in terms of performance; it achieves an optimal sensitivity rate on the dataset. When compared to previous strategies, the results show that our technique produces better operational results in terms of classification performance. As demonstrated in Tables 1 and 2. Furthermore, as compared to the previous strategies used, the proposed strategy greatly reduces training and testing times.

The performance of the proposed method is compared over time to different wrapping techniques for different types of malignancies using simulations. Table 3 presents an analysis of the results of comparing the suggested method with five conventional studies, as well as a comparison with several state-of-the-art methodologies. The comparison was initially done for datasets from UCI. Table 3 compares the suggested model for classification on BC datasets with the comparative analysis of EAs such as GA, DE, CS, PSO, and GOA.

WBCD						WDBC					
Classifiers	GA	PSO	CS	GOA	Proposed	Measure	GA	PSO	CS	GOA	Proposed
SVM-R	88.47	85.54	83.14	86.66	96.11	Accuracy	83.29	81.91	81.36	84.94	92.91
	80.62	81.69	84.98	84.45	92.62	Sensitivity	81.00	78.24	79.73	80.01	92.12
	83.55	82.04	81.68	83.89	93.54	Specificity	80.38	82.53	80.19	81.43	93.65
	81.23	80.64	83.44	82.81	92.11	F-measure	83.96	85.34	82.45	82.97	91.02
SVM-P	82.38	79.75	78.68	81.15	89.62	Accuracy	77.63	74.88	74.83	74.88	90.87
	80.22	78.01	79.37	79.01	82.62	Sensitivity	73.47	77.06	77.04	77.04	87.04
	77.81	76.05	77.16	75.11	80.65	Specificity	76.62	77.39	73.38	75.57	79.39
	79.65	76.72	77.88	78.45	81.06	F-measure	79.38	81.02	80.91	77.87	82.79
SVM-L	80.65	78.57	77.57	76.57	82.67	Accuracy	80.43	75.95	80.45	83.31	88.72
	73.04	77.68	76.53	72.4	83.21	Sensitivity	79.44	76.16	77.14	74.85	79.95
	70.84	72.49	75.45	71.02	81.81	Specificity	74.44	75.77	78.53	77.43	78.03
	71.07	72.91	75.62	77.98	85.35	F-measure	74.23	72.95	75.15	74.03	78.15

Table 1. Average performance in the wrapper methods of the two datasets.



Data	Methods	Parameters		Train Accuracy	Test Accuracy	Runtime(secs)
		$\sigma$	c			
WBCD	Single SVM	0.993	5.342	98.87	97.65	1022.43
	DE-SVM	0.892	4.987	98.32	98.01	995.45
	GA-SVM	0.873	4.556	98.23	98.12	890.32
	CS-SVM	0.786	3.654	96.34	95.99	1011.34
	GOA-SVM	0.432	3.112	92.54	91.57	889.04
	SeQTLBOGA	0.065	1.443	98.95	98.34	877.98
WDBC	Single SVM	0.889	5.453	89.32	88.11	1088.43
	DE-SVM	0.765	3.554	87.45	85.43	1011.43
	GA-SVM	0.544	3.654	88.11	86.98	899.90
	CS-SVM	0.435	3.987	89.23	87.22	880.54
	GOA-SVM	0.323	1.342	89.32	88.42	911.32
	SeQTLBOGA	0.034	1.123	97.43	94.19	544.89

Table 2. Evaluation of BC databases on different eas.

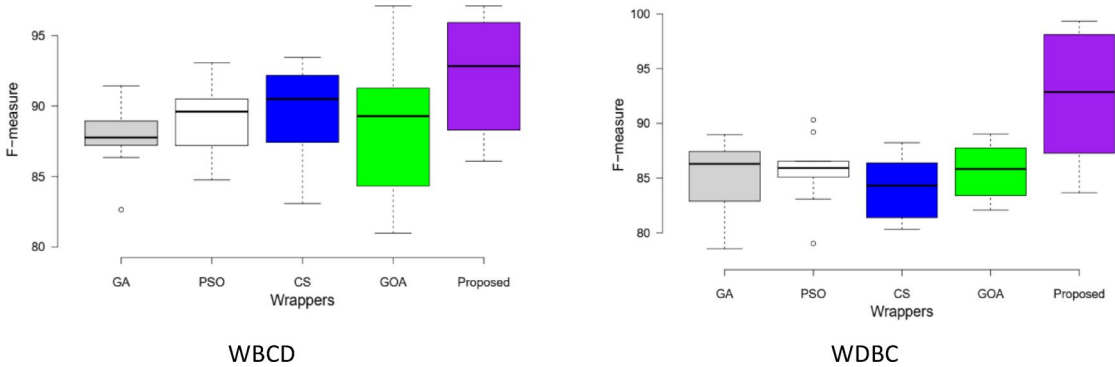


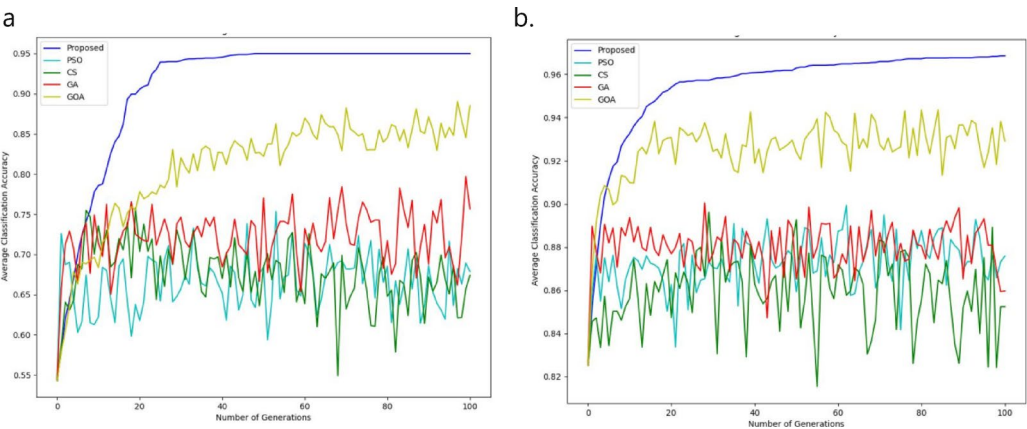
Fig. 4. Box-plot of wrappers of F-measure rate on both datasets.

	WBCD							WDBC					
K-Fold	Proposed	GA	PSO	DE	CS	GOA	Measure	Proposed	GA	PSO	DE	CS	GOA
2	99.43	89.34	87.42	88.67	89.21	87.02	Sensitivity	93.45	92.98	89.03	88.56	87.23	88.34
4	99.00	90.22	84.43	86.34	82.02	81.76		92.99	90.76	89.09	81.89	83.76	85.23
6	98.31	92.43	95.03	94.77	96.02	95.95		90.12	87.43	86.46	85.23	82.45	82.98
8	98.62	95.54	86.78	91.43	90.65	92.65		91.78	90.34	88.21	83.04	80.92	85.23
10	99.11	94.54	95.23	94.79	95.98	96.03		94.23	90.78	89.23	85.43	87.23	82.98
2	98.54	97.45	96.34	95.65	94.55	97.23	Specificity	88.67	88.32	84.5	82.56	80.67	85.98
4	96.47	95.33	94.02	94.00	94.67	95.43		86.32	80.12	81.23	80.43	82.43	83.64
6	97.89	96.98	95.88	96.03	98.45	98.01		87.12	81.32	86.02	82.32	81.32	82.54
8	96.66	91.43	90.87	89.03	93.43	88.04		81.32	80.27	81.87	82.87	81.09	78.67
10	97.21	96.03	95.55	94.78	93.09	94.67		77.23	78.54	79.76	75.27	76.87	78.99
2	344.34	388.45	390.76	498.43	789.34	1000.4	Etr(sec)	123.54	166.54	187.76	200.43	189.34	212.56
4	423.87	546.23	672.54	657.55	711.45	690.54		156.76	154.66	298.22	156.67	187.21	211.05
6	576.54	723.45	833.65	743.37	900.23	892.45		232.44	254.23	322.49	233.44	298.23	311.43
8	981.32	1111.45	1453.54	1004.54	966.45	897.56		433.54	456.44	533.65	675.33	644.34	634.43
10	865.56	980.34	879.34	1023.87	1117.8	1000.4		511.34	522.23	799.12	583.66	546.71	567.11
2	1189.04	1345.22	1211.56	1342.65	1411.23	1244.65	Ete(sec)	709.02	1008.4	988.43	899.34	923.45	834.6
4	1243.43	1232.43	1304.5	1299.45	1344.5	1444.45		699.34	934.54	899.34	922.45	899.34	902.54
6	1356.45	1511.43	1623.56	1322.65	1228.56	1235.65		733.45	965.45	877.65	875.45	788.45	899.34
8	1427.80	1511.34	1345.23	1123.43	1228.45	1311.45		945.56	1324.43	1022.5	933.5	1056.45	1077.98
10	1132.98	1332.57	1244.47	1542.25	1622.31	1445.65		1070.5	1277.5	1321.65	1225.6	1355.4	1233.6

Table 3. Comparative analysis of the experimental performance on two datasets.

Method	Dataset	Accuracy
48	WBCD	96.56
49	WBCD	91.15
50	WDBC	95.17
51	WDBC	96.49
51	WDBC	94.74
our	WDBC	97.43

**Table 4.** Comparative analysis of proposed in datasets.



**Fig. 5.** Convergence graph on (a) WBCD. (b) WDBC.

Datasets	Measures	GA	PSO	DE	CS	GOA	Proposed
WDBC	AUROC	96.99	97.43	89.34	98.32	94.23	99.06
	AUPR	97.66	98.22	89.98	98.03	96.47	99.06
WBCD	AUROC	98.23	97.44	95.03	97.34	98.90	99.11
	AUPR	97.34	95.22	94.34	94.88	98.91	99.12

**Table 5.** AUROC and AUPR values dataset.

Furthermore, it is evident that our strategy performs better than the feature subset in both datasets based on the actual results obtained by testing the suggested technique on the dataset. Based on these data, our approach performs better than other state-of-the-art methods, as shown by all the evaluation indicators listed in Table 4.

As shown in Table 4, the proposed approach has achieved the highest accuracy compared to other recent techniques, such as<sup>47–49</sup> In addition to its superior accuracy, the proposed technique also significantly outperforms other prevailing state-of-the-art techniques in terms of accuracy.

Figure 5 shows the convergence graph on two datasets based on WBCD and WDBC, respectively. From graphs, we can see that our proposed method performs best regarding classification accuracy compared to other nature-inspired algorithms.

Classifier performance is measured using AUPR (Area under Precision-Recall) Curve values ranging from 0 to 1, with a perfect classifier reaching an AUPR score of 1. Table 5 shows the AUROC and AUPR values, suggesting that the proposed technique outperforms other current models in identifying disease across the dataset used. In contrast, all of the techniques used have a very good fit. In Table 6, Central Nervous System (CNS)<sup>52</sup> Data shows the comparative study by our wrapper method and another wrapper method in terms of four measures, like accuracy, F-measure, sensitivity, and specificity.

*Statistical results*

One of the most popular statistical techniques for ranking algorithm performance is the Friedman test<sup>53,54</sup>. Whether looks for any discernible differences between the outcomes of several algorithms. Its foundation is the null hypothesis, which states that there are no differences in how the algorithms are presented. Based on the following factors, the algorithm that performs the best gets the lowest rank, while the algorithm that performs the worst gets the highest rank (see Table 7).

- (1) The observed accuracy value for each algorithm and dataset pair.

EAs	Performance	SVM-L	SVM-P	SVM-R
GA	Accuracy	83.52	81.25	86.32
	Sensitivity	82.25	80.25	84.63
	Specificity	78.63	79.36	85.6
	F-measure	84.25	78.31	84.87
PSO	Accuracy	85.67	83.54	81.63
	Sensitivity	82.63	81.74	80.52
	Specificity	79.34	79.86	79.38
	F-measure	80.52	76.58	77.94
CS	Accuracy	75.36	74.22	76.32
	Sensitivity	74.25	71.02	70.05
	Specificity	77.89	75.23	71.95
	F-measure	72.84	73.65	74.25
GOA	Accuracy	78.25	73.56	74.85
	Sensitivity	77.84	73.89	73.45
	Specificity	76.52	75.89	79.65
	F-measure	77.25	76.84	77.06
Our	Accuracy	88.25	87.62	85.32
	Sensitivity	87.67	85.66	84.79
	Specificity	84.63	86.07	85.97
	F-measure	84.25	83.96	82.78

**Table 6.** Comparative analysis of four measures on the CNS dataset.

Algorithms	Ranking
GA	2.25
PSO	3.875
CS	5
GOA	2.875
Proposed	1

**Table 7.** Summary of statistical results.

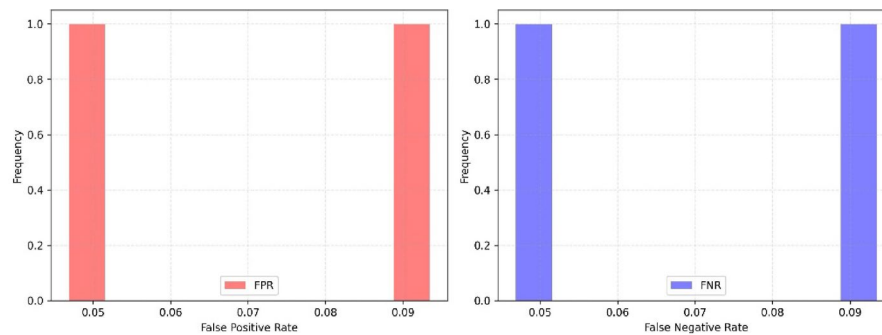
i	Algorithms	$z = (R_0 - R_i)/SE$	$p$	Holm Hochberg	Li
4	GA	3.577709	0.000347	0.0125	0.03876
3	PSO	2.571478	0.010127	0.016667	0.03876
2	CS	1.677051	0.093533	0.025	0.03876
1	GOA	1.118034	0.263552	0.05	0.05

**Table 8.** Post hoc comparison table for  $\alpha = 0.05$ .

- (2) For each BC dataset, rank values from 1 (best result) to 5 (worst result).
- (3) For each method, average the ranks obtained in the BC datasets to achieve the final rank.

The tests are applied to the accuracy obtained from the BC data sets. Table 7 displays Friedman's statistic, which determines the level of responsibility based on the average rank that each method achieved on two datasets. Based on the computed results, the calculated Friedman statistic is 14.95, and the test p-value was 0.004806 at the 5% significance level. As a result, the null hypothesis is disproved, indicating variations in the methods' performances. Using Li and Holm's process, the post-hoc method for pairwise comparisons of metaheuristic methods further validates the suggested method. P-values that were determined by utilizing post hoc techniques on the Friedman procedure findings are displayed in Table 8.

In order to comprehend model performance and potential biases, error analysis in breast cancer datasets entails assessing false positive (FP) and false negative (FN) rates. False negatives happen when a cancer case is overlooked, possibly postponing treatment, whereas false positives happen when a benign case is mistakenly identified as malignant, resulting in needless operations. The goal of bias detection is to determine if these mistakes are consistently linked to particular characteristics, as seen in Fig. 6.



**Fig. 6.** Histograms of false positive and false negative rates.

## Conclusion

Breast cancer is a prevalent disease that affects women globally. However, the number of deaths from breast cancer has decreased by the early detection of the disease through the application of medical evolution techniques. Particularly in breast cancer prediction, our novel development based on a self-adaptive quantum evolutionary algorithm known as SeQTLBOGA is vibrant for enhancing the quality of life for millions of patients. Our method has two main benefits. First, using fewer features can build a diagnostic model that is easy to understand. The other is that optimizing the SVM model parameters can produce the best prediction model. Specifically, using a series of experiments with EAs on both BC databases, we show that our method selects the most valuable features and maximizes speed. The experimental results show that our method produces more accurate results than regular EAs when EAs are combined with SVM. Our method performs best in BC datasets regarding ROC area, F-measure, accuracy, sensitivity, and specificity. These findings demonstrate that our technology can be used as a practical replacement for the existing techniques for diagnosing breast cancer. We may apply our SeQTLBOGA with the SVM kernel to predict the prognosis of additional deadly illnesses. It will consequently be a lifeline for patients in terms of the clinical management of cancer and other life-threatening conditions and their prognosis. This algorithm can be utilized as a diagnostic tool in BC research.

## Data availability

The datasets used during the current study are publically available in the Wisconsin Diagnostic Breast Cancer datasets. The data generated during the current study available from the corresponding author on reasonable request.

Received: 14 February 2025; Accepted: 30 May 2025

Published online: 06 June 2025

## References

- Zhu, J. et al. An integrated approach of feature selection and machine learning for early detection of breast cancer. *Sci. Rep.* **15**(1), 1–12. <https://doi.org/10.1038/s41598-025-97685-x> (2025).
- Zubair, M. et al. Enhanced gastric cancer classification and quantification interpretable framework using digital histopathology images. *Sci. Rep.* **14**(1), 1–17. <https://doi.org/10.1038/s41598-024-73823-9> (2024).
- Arya, N., Mathur, A., Saha, S. & Saha, S. Proposal of SVM utility kernel for breast Cancer survival Estimation. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* **20**(2), 1372–1383. <https://doi.org/10.1109/TCBB.2022.3198879> (2023).
- Song, L., Hsu, W., Xu, J. & Schaar, M. V. D. Using contextual learning to improve diagnostic accuracy: application in breast Cancer screening. *IEEE J. Biomed. Heal. Inf.* **20**(3), 902–914 (2016).
- Almutairi, S., Manimurugan, S., Kim, B. G., Aborokbah, M. M. & Narmatha, C. Breast cancer classification using deep Q learning (DQL) and gorilla troops optimization (GTO). *Appl. Soft Comput.* **142**, 110292. <https://doi.org/10.1016/j.asoc.2023.110292> (2023).
- Meenachi, L. & Ramakrishnan, S. Metaheuristic search based feature selection methods for classification of Cancer. *Pattern Recognit.* **119**, 108079. <https://doi.org/10.1016/j.patcog.2021.108079> (2021).
- Spanhol, F. A dataset for breast Cancer histopathological A dataset for breast Cancer histopathological image classification. *IEEE Trans. Biomed. Eng.* **63**, 1455–1462. <https://doi.org/10.1109/TBME.2015.2496264> (2016).
- Bilal, A. et al. BC-QNet: A quantum-infused ELM model for breast cancer diagnosis. *Comput. Biol. Med.* **175**, 108483. <https://doi.org/10.1016/j.combiomed.2024.108483> (2024).
- Dong, S., Xia, Y. & Kamruzzaman, J. Quantum particle swarm optimization for task offloading in mobile edge computing. *IEEE Trans. Ind. Inf.* **19**(8), 9113–9122. <https://doi.org/10.1109/TII.2022.3225313> (2022).
- Kang, Y. et al. A Hybrid Two-Stage Teaching-Learning-Based Optimization Algorithm for Feature Selection in Bioinformatics. *IEEE/ACM Trans. Comput. Biol. Bioinforma.* **20**(3), 1746–1760. <https://doi.org/10.1109/TCBB.2022.3215129> (2022).
- Agrawal, R. K., Kaur, B. & Agarwal, P. Quantum inspired particle swarm optimization with guided exploration for function optimization. *Appl. Soft Comput.* **102**, 107122. <https://doi.org/10.1016/j.asoc.2021.107122> (2021).
- Han, K. H. & Kim, J. H. Quantum-inspired evolutionary algorithm for a class of combinatorial optimization. *IEEE Trans. Evol. Comput.* **6**(6), 580–593. <https://doi.org/10.1109/TEVC.2002.804320> (2002).
- Guo, H. et al. Multi-threshold Image Segmentation based on an improved Salp Swarm Algorithm: Case study of breast cancer pathology images. *Comput. Biol. Med.* **168** (2024). <https://doi.org/10.1016/j.combiomed.2023.107769> (2023).
- Ding, W., Abdel-basset, M., Hawash, H. & Ali, A. M. Explainability of artificial intelligence methods, applications and challenges: A comprehensive survey. *Inf. Sci. (Nij.)* **615**, 238–292. <https://doi.org/10.1016/j.ins.2022.10.013> (2022).
- Faisal, L. & Arselene, M. A novel framework for lung cancer classification using lightweight convolutional neural networks and ridge extreme learning machine model with SHapley Additive exPlanations (SHAP). *Expert Syst. Appl.* **248**, 123392. <https://doi.org/10.1016/j.eswa.2024.123392> (2023).

16. Mernik, M. Exploration and exploitation in evolutionary algorithms: A survey. *ACM Comput.* **45**(3), 1–33 (2013).
17. Rao, R. V., Savsani, V. J. & Vakharia, D. P. Teaching learning-based optimization: A novel method for constrained mechanical design optimization problems. *Comput. Des.* **43**(3), 303–315. <https://doi.org/10.1016/j.cad.2010.12.015> (2011).
18. Tuo, S. et al. HSTLBO: A hybrid algorithm based on harmony search and Teaching-Learning-based optimization for complex high-dimensional optimization problems. *PLoS One*. **12**(4), 1–23 (2017).
19. Abdelmalak, M. E. S. et al. BER-XGBoost: pothole detection based on feature extraction and optimized XGBoost using BER metaheuristic algorithm. *J. Artif. Intell. Metaheuristics*. **6**(2), 46–55. <https://doi.org/10.54216/jaim.060205> (2023).
20. Gaber, K. S. A review on designing hybrid energy systems for renewable integration, **03**, 02, pp. 21–32, (2025).
21. Özbay, F. A. An enhanced Zebra optimization algorithm with multiple strategies for global optimization and feature selection problems: A hepatocellular carcinoma case study. *IEEE Access*. **13**, no. <https://doi.org/10.1109/ACCESS.2025.3541975> (January, 2025).
22. Özbay, F. A., Özbay, E. & Gharehchopogh, F. S. An improved artificial rabbits optimization algorithm with chaotic local search and Opposition-Based learning for engineering problems and its applications in breast Cancer problem. *C - Comput. Model. Eng. Sci.* **141**(2), 1067–1110. <https://doi.org/10.32604/cmescs.2024.054334> (2024).
23. Kamal, A. M., Sakorikar, T., Pal, U. M. & Pandya, H. J. Engineering approaches for breast Cancer diagnosis: A review. *IEEE Rev. Biomed. Eng.* **16**, 687–705. <https://doi.org/10.1109/RBME.2022.3181700> (2023).
24. Junyue, C., Zeebaree, D. Q., Qingfeng, C. & Zebari, D. A. Breast cancer diagnosis using hybrid AlexNet-ELM and chimpanzee optimization algorithm evolved by Nelder-mead simplex approach. *Biomed. Signal. Process. Control*. **85** <https://doi.org/10.1016/j.bspc.2023.105053> (Aug. 2023).
25. Eshun, R. B., Bikdash, M. & Islam, A. K. M. K. A deep convolutional neural network for the classification of imbalanced breast cancer dataset. *Healthc. Anal.* **5**, 100330. <https://doi.org/10.1016/j.health.2024.100330> (2024).
26. Sharma, A., Goyal, D. & Mohana, R. An ensemble learning-based framework for breast cancer prediction, *Decis. Anal. J.* **10**, 100372, 2024, <https://doi.org/10.1016/j.dajour.2023.100372> (2023).
27. Ding, W. & Wang, J. A novel approach to minimum attribute reduction based on quantum-inspired self-adaptive cooperative co-evolution. *Knowledge-Based Syst.* **50**, 1–13. <https://doi.org/10.1016/j.knsys.2013.03.008> (2013).
28. Li, D., Zhang, L., Zhang, J. & Xie, X. Convolutional feature descriptor selection for mammogram classification. *IEEE J. Biomed. Heal. Inf.* **27**(3), 1467–1476. <https://doi.org/10.1109/JBHI.2022.3235355> (2023).
29. Lv, T. et al. AI-powered interpretable imaging phenotypes noninvasively characterize tumor microenvironment associated with diverse molecular signatures and survival in breast cancer. *Comput. Methods Programs Biomed.* **243**, 107857. <https://doi.org/10.1016/j.cmpb.2023.107857> (2024).
30. El-kenawy, E. S. M. et al. Sunshine duration measurements and predictions in saharan Algeria region: an improved ensemble learning approach. *Theor. Appl. Climatol.* **147**, 3–4. <https://doi.org/10.1007/s00704-021-03843-2> (2022).
31. El-Kenawy, E. S. M. et al. Meta-Heuristic optimization and keystroke dynamics for authentication of smartphone users. *Mathematics* **10**(16), 1–26. <https://doi.org/10.3390/math10162912> (2022).
32. Yang, Z., Li, K. & Niu, Q. A self-learning TLBO based dynamic economic / environmental dispatch considering multiple plug-in electric vehicle loads. *J. Mod. Power Syst. Clean. Energy*. **2**(4), 298–307. <https://doi.org/10.1007/s40565-014-0087-6> (2014).
33. Almutairi, L. et al. Quantum Dwarf Mongoose Optimization with Ensemble Deep Learning Based Intrusion Detection in Cyber-Physical Systems, *IEEE Access* **11** 66828–66837 <https://doi.org/10.1109/ACCESS.2023.3287896> (2023).
34. Agarwal, S., Jain, D., Gupta, S., Sawhney, S. & Mittal, Y. Brain Tumor Detection and Classification Using Deep Learning, *Proc. - IEEE 5th Int. Conf. Adv. Comput. Commun. Control Networking, ICAC3N 2023*, pp. 635–640(2023). <https://doi.org/10.1109/ICAC3N60023.2023.10541584>
35. Nickabadi, A., Ebadzadeh, M. M. & Safabakhsh, R. A novel particle swarm optimization algorithm with adaptive inertia weight. *Appl. Soft Comput.* **11**(4), 3658–3670. <https://doi.org/10.1016/j.asoc.2011.01.037> (2011).
36. Shehab, M., Khader, A. T. & Al-Betar, M. A. A survey on applications and variants of the cuckoo search algorithm. *Appl. Soft Comput.* **61**, 1041–1059. <https://doi.org/10.1016/j.asoc.2017.02.034> (2017).
37. Butler-Yeoman, T., Xue, B. & Zhang, M. Particle swarm optimisation for feature selection: A hybrid filter-wrapper approach, *2015 IEEE Congr. Evol. Comput. CEC 2015 - Proc.*, pp. 2428–2435. <https://doi.org/10.1109/CEC.2015.7257186> (2015).
38. Raymer, M. L., Doom, T. E., Kuhn, L. A. & Punch, W. F. Knowledge discovery in medical and biological datasets using a hybrid Bayes classifier/evolutionary algorithm. *IEEE Trans. Syst. Man. Cybern. Part. B Cybern.* **33**(5), 802–813. <https://doi.org/10.1109/TSMB.2003.816922> (2003).
39. Kirchner-Bossi, N. & Porté-Agel, F. Wind farm power density optimization according to the area size using a novel self-adaptive genetic algorithm, *Renew. Energy* **220**, 119524, (2024). <https://doi.org/10.1016/j.renene.2023.119524>
40. Li, S., Gong, W., Wang, L., Yan, X. & Hu, C. A hybrid adaptive teaching-learning-based optimization and differential evolution for parameter identification of photovoltaic models. *Energy Convers. Manag.* **225**, no. <https://doi.org/10.1016/j.enconman.2020.113474> (June, 2020).
41. Xu, Y., Yang, Z., Li, X., Kang, H. & Yang, X. Dynamic opposite learning enhanced teaching-learning-based optimization. *Knowledge-Based Syst.* **188**, 104966. <https://doi.org/10.1016/j.knsys.2019.104966> (2020).
42. Calzarossa, M. C., Giudici, P. & Zieni, R. Explainable machine learning for phishing feature detection. *Qual. Reliab. Eng. Int.* **362**–373. <https://doi.org/10.1002/qre.3411> (2023).
43. Lu, Q. et al. Building the Future of Responsible AI: A Reference Architecture for Designing Large Language Model based Agents, pp. 1–14, [Online]. (2023). Available: <https://arxiv.org/abs/2311.13148v2>
44. Alshayji, M. H., Ellethy, H., Abed, S. & Gupta, R. Computer-aided detection of breast cancer on the Wisconsin dataset: an artificial neural networks approach. *Biomed. Signal. Process. Control*. **71** <https://doi.org/10.1016/j.bspc.2021.103141> (Jan. 2022).
45. Huang, Q., Ding, H. & Effatparvar, M. M. Breast Cancer diagnosis based on hybrid squeezeNet and improved Chef-based optimizer. *Expert Syst. Appl.* **p. 121470** <https://doi.org/10.1016/j.eswa.2023.121470> (Mar. 2023).
46. Thirumoorthy, K. A two-stage feature selection approach using hybrid quasi-opposition self-adaptive Coati optimization algorithm for breast cancer classification. *Appl. Soft Comput.* **146** <https://doi.org/10.1016/j.asoc.2023.110704> (Oct. 2023).
47. Elmasry, W., Akbulut, A. & Zaim, A. H. Evolving deep learning architectures for network intrusion detection using a double PSO metaheuristic. *Comput. Networks*. **168** <https://doi.org/10.1016/j.comnet.2019.107042> (2020).
48. Hajiabadi, H., Babaiyan, V., Zabihzadeh, D. & Hajiabadi, M. Combination of loss functions for robust breast cancer prediction. *Comput. Electr. Eng.* **84** <https://doi.org/10.1016/j.compeleceng.2020.106624> (Jun. 2020).
49. Bacha, S. & Taouali, O. A novel machine learning approach for breast cancer diagnosis, *Meas. J. Int. Meas. Confed.* **187**, 110233, 2022, (2021). <https://doi.org/10.1016/j.measurement.2021.110233>
50. Xu, C. & Zhang, S. A genetic Algorithm-based sequential instance selection framework for ensemble learning. *Expert Syst. Appl.* **236** <https://doi.org/10.1016/j.eswa.2023.121269> (Feb. 2024).
51. Jonny, A. I. & Arnob, A. K. B. Deep learning paradigms for breast Cancer diagnosis: A comparative study on Wisconsin diagnostic dataset. *Malaysian J. Sci. Adv. Technol.* **4**(2), 109–117. <https://doi.org/10.56532/mjsat.v4i2.245> (2024).
52. Dabba, A., Tari, A., Meftali, S. & Mokhtari, R. Gene selection and classification of microarray data method based on mutual information and moth flame algorithm, *Expert Syst. Appl.*, vol. 166, no. July p. 114012, 2021, (2020). <https://doi.org/10.1016/j.eswa.2020.114012>



53. Iantovics, L. B., Dehmer, M. & Emmert-Streib, F. MetrIntSimil-an accurate and robust metric for comparison of similarity in intelligence of any number of cooperative multiagent systems. *Symmetry (Basel)*. **10**(2), 1–22. <https://doi.org/10.3390/sym10020048> (2018).
54. Wang, Y., Wang, B., Li, H. & Yen, G. G. Incorporating objective function information into the feasibility rule for constrained evolutionary optimization. *IEEE Trans. Cybern.* pp. 1–15, (2015).

### Author contributions

Alok Kumar Shukla, Shubhra Dwivedi wrote the main manuscript proposed method, and analyzed the data, and Deepak Singh and Diwakar Tripath organized manuscript, visualization and Sunil Kumar Singh, Ram Kishan Dewangan design, analysis tools, figure, etc.

### Funding

Open access funding provided by Manipal Academy of Higher Education, Manipal

Open Access charges will be provided by Manipal Institute of Technology, Bengaluru, MAHE Bengaluru Campus, India.

### Declarations

### Competing interests

The authors declare no competing interests.

### Ethical approval

This study was performed using available datasets, as per my compliance with ethical standards, there were no human or animal participants, and therefore, the study did not require ethics approval.

### Human participants and/or animals

This article contains no studies with human participants or animals performed by any of the authors.

### Additional information

**Correspondence** and requests for materials should be addressed to S.K.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

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

© The Author(s) 2025