Scientia Iranica D (2022) 29(2), 613–631

Sharif University of Technology
Scientia Iranica
Transactions D: Computer Science & Engineering and Electrical Engineering
http://scientiairanica.sharif.edu

A modified metaheuristic algorithm-integrated ELM model of cancer classification

P. Paramita Debata* **and P. Mohapatra**

Department of Computer Science and Engineering, International Institute of Information Technology Bhubaneswar, Odisha, India, 751019.

Received 25 June 2020; received in revised form 23 September 2021; accepted 3 January 2022

KEYWORDS
Self-adaptive multi-population-based Elite Jaya algorithm;
Functional link artificial neural network;
Extreme learning machine;
Classification model.

Abstract. In a speedily degrading environment, cancer is acknowledged as the most menacing disease whose death rate is higher than others. For this reason, a number of researchers have analyzed the cancer-inducing genes and designed an efficient classification model to diagnose cancer effectively and quickly. In this study, random parameters of Extreme Learning Machine (ELM) were optimized through Self-Adaptive Multi-Population-based Elite strategy Jaya (SAMPEJ) approach. This strategy constructs a robust classifier called SAMPEJ-ELM model. This model was tested on datasets of breast, cervical, and lung cancers. To this end, a comparative analysis of the proposed model with ELM, Jaya optimized ELM (Jaya-ELM), SAMPEJ optimized Neural Network (SAMPEJ-NN), Teaching Learning Based Optimization (TLBO) hybridized ELM (TLBO-ELM), and SAMPEJ optimized Functional Link Artificial Neural Network (SAMPEJ-FLANN) models was conducted. Numerous performance metrices namely the accuracy, specificity, Gmean, sensitivity, and F-score with Receiver Operating Characteristic (ROC) were employed to evaluate the proposed approach. Moreover, this model was compared with 11 existing models. Of note, SAMPEJ-ELM approach had the highest degree of accuracy, sensitivity, and specificity in the datasets of breast (0.9895, 1, 0.9853), cervical (0.9822, 0.9948, 0.9828), and lung cancers (0.9787, 1, 1). The experimental outcomes revealed that SAMPEJ-ELM approach could classify the benign and malignant samples of cancer datasets significantly better than others.

1. Introduction

According to a statistical report of GLOBOCAN database [1] and World Health Organization (WHO) [2], cancer is currently the most threatening disease against humans. Early diagnosis of some cancers can definitely ensure rehabilitation for one's entire lifespan. Although some cancer types cannot be cured by surgery, others can still be treated by medicines for lifetime. In this respect, an automatic diagnosis gains considerable significance in cancer research. To handle the growing complexity of the cancer data, especially machine-learning-area-based data, researchers devote considerable effort to design an efficient classifier to classify cancer data with low computational overhead. The sole objective here is to achieve high classification accuracy in an effective time period. Many researches have been conducted on designing a better classifier for cancer data, yet it is still challenging due to its growing complexity.

*. *Corresponding author.*
E-mail addresses: c117007@iiit-bh.ac.in (P. Paramita Debata); puspanjali@iiit-bh.ac.in (P. Mohapatra)

doi: 10.24200/sci.2022.56265.4630

Different types of statistical models and machine-learning-based classifiers have been successfully explored in terms of classification including Bayesian decision theory [3], Euclidean Minimum Distance (EMD), Artificial Neural Network (ANN) [4], Functional Link Artificial Neural Network (FLANN) [5], K-Nearest Neighbor (KNN) [6], Multilayer Perceptron (MLP) [7], Radial Basis Function Neural Networks (RBFNNs) [7], fuzzy rule-based systems [8], Back-Propagation (BP) [9], and Support Vector Machine (SVM) [10,11]. These learning approaches are prone to numerous complications such as slow learning rate, random variables adjustment, and cantrap at local optimum [12]. To deal with these complications, Huang et al. suggested a powerful machine-learning-based classification model known as Extreme Learning Machine (ELM) [13]. This classification model is a generalized form of Single-hidden Layer Feed forward Networks (SLFNs). Dissimilar to the traditional SLFN, ELM classifier does not need to adjust its variables at every iteration. Therefore, its learning speed and generalization ability are better than those of SLFN. Numerous upgraded versions of ELM have been explored so far to interpret different binary and multiclass data classification problems [14–18]. However, they still face some problems. To choose random variables, the best activation function remains still substandard in the ELM classifier. Mostly, the random parameters of ELM cause instability in resolving the classification and regression problems. The outcome of ELM is calculated based on the arbitrarily chosen hidden biases and input weights. In this process, the training error should be minimized. While selecting the random biases and input weights, the output matrix does not reflect full column rank [19] which leads to instability and produces suboptimal solutions. In addition, a number of evolutionary algorithms [20–26] were employed to optimize the random variables of ELM. Among these evolutionary algorithms, Genetic Algorithm (GA) optimized ELM [27], Differential Evolution (DE) optimized ELM [25], Particle Swarm Optimization (PSO) optimized ELM [21,22,24], Ant Colony Optimization (ACO) optimized ELM [28], Cuckoo Search Optimization (CS) [29], and Cat Swarm Optimization (CSO) algorithms optimized kernel ridge regression [30] that are mostly used for classifying cancer data.

In this study, an efficient Jaya algorithm was hybridized with ELM classifier to classify cancer data. Rao [31] developed the applied Jaya algorithm to resolve both constraint- and nonconstraint-based issues mapped on the conception of achieving the best result and discarding the worst one. Similar to Teaching Learning Based Optimization (TLBO) algorithm [32], Jaya does not need any algorithm-specific variable. According to the literature, Jaya algorithm has been favorably used in different fields such as dimension optimization of a micro-channel heat sink [33], optimization of power flow [34], recognition of facial emotion [35], image processing [36], and stock market prediction [37]. To improve the search process of Jaya, the subpopulation search process [38] was applied by separating the entire population into sub-groups assigned all over the search area. This mechanism concentrates on the diversification of search rather than single search space. In this technique, each subgroup or subpopulation is related to the search processes of the algorithm by either exploring or exploiting. According to the related literature, compared to single population-based algorithms, multipopulation-based ones could produce more effective global solutions [38]. In 2000, Branke et al. presented a multi-populational evolutionary algorithm to solve dynamic optimization problems [39]. Du and Li [40] and Yang and Li [41] employed multiswarm PSO and clustering-based PSO approaches to solve the optimization problems. Rao and Patel [42] presented multiple teacher-based TLBO to optimize heat exchanger.

Although multipopulational algorithms are beneficial to maintaining the overall diversity of the population, their performance is still influenced by how the number of divisions of the whole population is determined. This number changes repeatedly during the search process. There is a high possibility that the solutions found in the subgroups may not be optimal for ample diversity. In this regard, Rao and Saroj (2019) [43] proposed Self-Adaptive Multi-Population-based Elite strategy Jaya (SAMPEJ) algorithm to solve these issues. Of note, the SAMPEJ algorithm can adaptively modify the number of subgroups according to the modified quality of the solution. Rao et al. [44] also used SAMPEJ algorithm to optimize the heat pipe design. The current study puts its main focus on the following subjects:

(a) Optimization of the random parameters of ELM classifier through a newly designed SAMPEJ algorithm, especially for cancer classification;

(b) Evaluation of the suggested approach using three cancer datasets including breast, lung, and cervical cancers;

(c) Comparative analysis of the SAMPEJ-ELM model and basic ELM, Jaya-ELM, TLBO-ELM, SAMPEJ-NN, and SAMPEJ-FLANN models.

The rest of this paper is organized as follows: Section 2 covers the description of the suggested model. Section 3 describes all the supported methodologies. Section 4 presents all the experimental discussion and result validation parts. Section 5 gives the concluding remarks.

2. Presented model discussion

Figure 1 depicts the overall flow of the suggested classification model. Three standard cancer datasets including those of breast cancer Wisconsin (original), cervical cancer, and lung cancer were collected from UCI repository to assess this model [45]. At the beginning phase, the datasets ranged between −1 and 1 based on Eq. (1):

$$y_n = m + (m - n) * \frac{(y - X_{\min})}{(X_{\max} - X_{\min})}, \quad (1)$$

where y_n is the scaled form of y (main dataset); m and n variables take the values of 1 and 1, respectively. $X_{\min}$ and $X_{\max}$ are considered as the minimum and maximum values of the dataset, respectively.

Then, each dataset was shuffled using randperm function. Followed by shuffling, each dataset was split into training set (70%) and testing set (30%). Then, the presented model was trained by basic ELM, TLBO-ELM, Jaya-ELM, SAMPEJ-NN, SAMPEJ-FLANN, and SAMPEJ-ELM. The performances of these models were estimated by different performance metrices, namely classification accuracy%, specificity%, F-score, sensitivity%, G-mean, and Receiver Operating Characteristic (ROC) graphs.

Figure 2 presents the suggested SAMPEJ optimized ELM classification model for cancer data classification. It can be trained using basic ELM based models where the output is the classification accuracy.

3. Methodologies supported

3.1. Basic ELM model

The methodology of the basic ELM is summarized below:

1. Randomly take the hidden layer biases (bi) and input weights (wi);
2. Compute output matrix (H) of the hidden layer;
3. Determine the output weight ($\hat{\beta}$) using Eq. (2):

$$\hat{\beta} = H^{\dagger}T, \quad (2)$$

here, T is the target variable.

3.2. Neural network models

In case the number of nodes increases in the hidden layer of NN models, their computational complexity also grows; hence, much time is needed for training. However, compared to other classifiers, NN needs less formal statistical training. This study considered a three-hidden-layer NN model where five nodes were used in each hidden layer. Figure 3 shows the basic structure of the NN model.

FLANN is another type of the NN model that generates non-linear decision boundaries [46] to solve complex problems. The input vector of the FLANN model is expanded in the functional expansion block using any linearly independent function to enhance the dimensionality of the input vector. The basic layout of the FLANN model is depicted in Figure 4.

In this figure, $D = [d_1, d_2, ..., d_I]^T$ is the original pattern of the input that can be enhanced using trigonometric functions as shown below:

$$Y = \left[d_1, \cos\left(\prod d_1\right), \sin\left(\prod d_1\right), ..., \cos\left(n\prod d_1\right), \sin\left(n\prod d_1\right), ..., \sin\left(\prod d_1\right), ..., d_2, \cos\left(\prod d_2\right), \sin\left(\prod d_2\right), ..., \cos\left(n\prod d_1\right), \sin\left(n\prod d_1\right)\right]^T. \quad (3)$$

In Eq. (3), n values are considered as 1, 2, and 3 for different numbers of function expansions. Cross-validation was performed to obtain the best structure

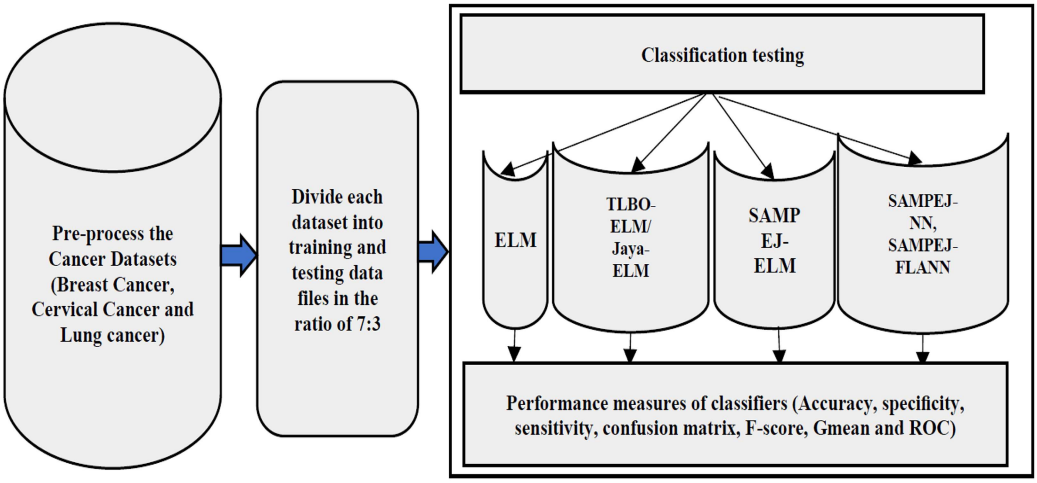

Figure 1. Flow of the suggested classification model.

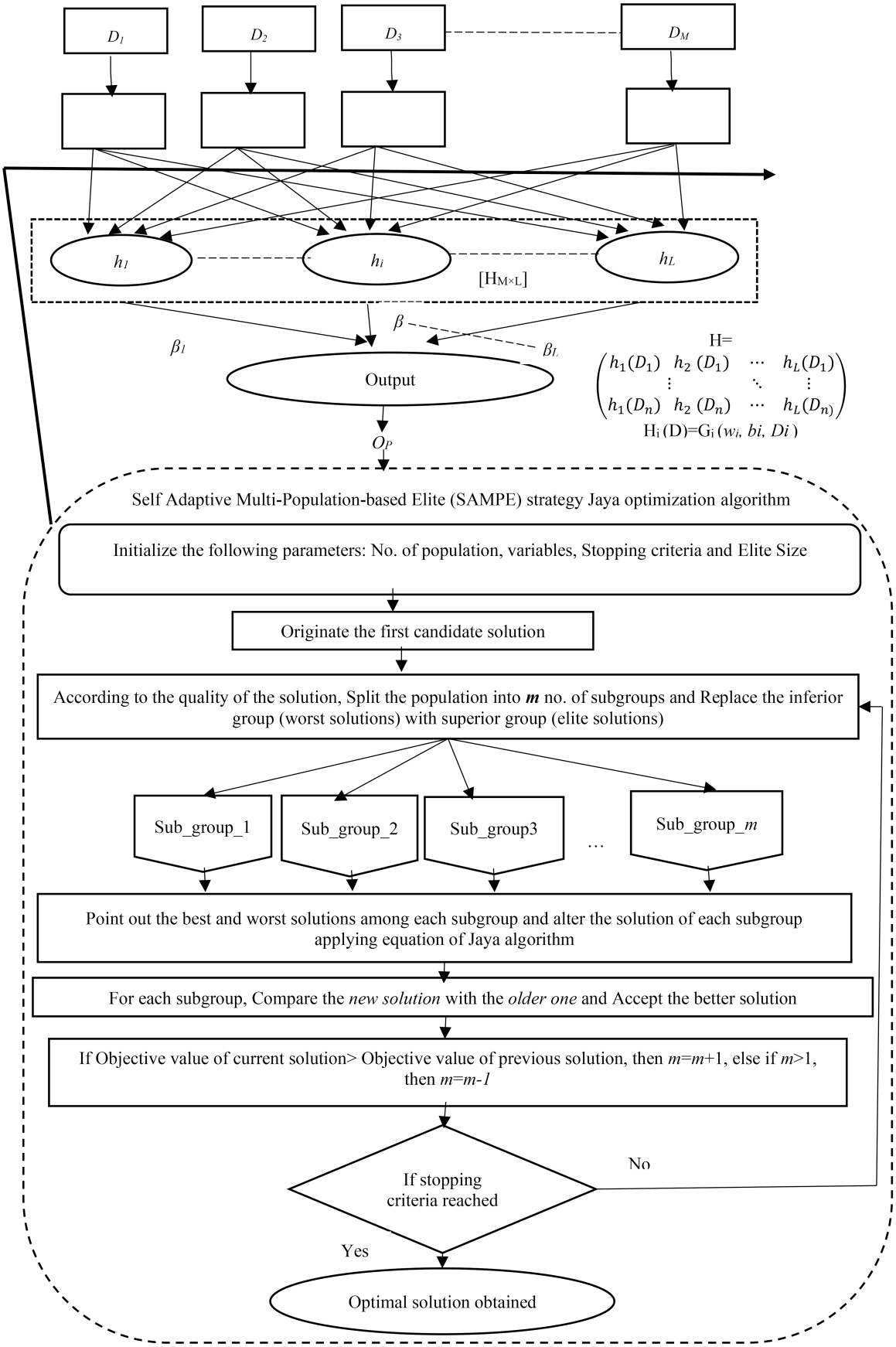

Figure 2. SAMPEJ algorithm optimized ELM classification model.

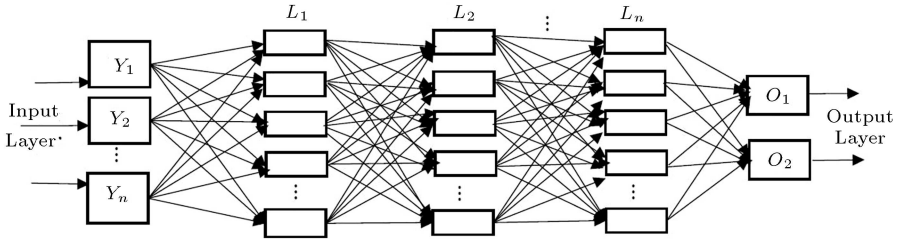

Figure 3. Abstract view of the NN technique with three hidden layers.

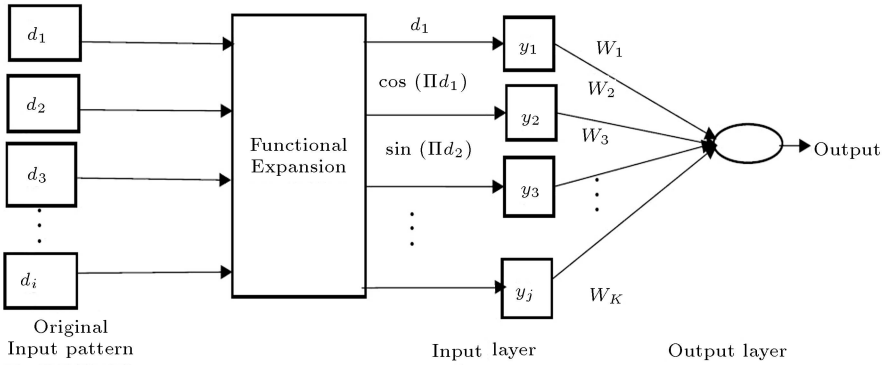

Figure 4. The basic layout of FLANN.

of the FLANN model. In addition, sigmoidal function was used in the output layer as the activation function to produce the outputs ranging from 0 to 1 in this model. Compared to MLP, FLANN is subject to less computational complexity due to its single-layer structure. NN model is more helpful for classification in non-linear problems than other models. However, the FLANN model uses trigonometric expansion that makes it better than the other NN models.

3.3. Jaya algorithm

Like TLBO [47], Jaya [31] needs only two parameters: the number of population and maximum iteration size. The steps of Jaya optimization algorithm are summarized below:

1. Set the numbers of design variables (dv) and population as well as the maximum iteration size. The best and worst candidate solutions are derived among the whole populations.
2. Then, modify the results based on Eq. (4) and the best and worst solutions:

$$K'_{y,z,x} = K_{y,z,x} + r_{1,y,x}\left[\left(K_{y,best,x}\right) + \left|\left(K_{y,z,x}\right)\right|\right] - r_{2,y,x}\left[\left(K_{y,worst,x}\right) + \left|\left(K_{y,z,x}\right)\right|\right]. \quad (4)$$

In Eq. (4), $K_{y,z,x}$ is the value of the xth variable for the zth candidate, k the number of population, i the maximum iteration size, and y the number of design variables

3.4. SAMPEJ algorithm

In the proposed approach, Jaya is modified by SAMPEJ algorithm to gain the optimum result. The flow of the SAMPEJ algorithm is summarized below:

1. Set the population size (P), Elite Size (ES), design variables (d), and stopping criteria;
2. Create the first candidate solution;
3. Split the population into m numbers of sub-groups (set $m = 2$ initially) according to the quality of the solution and substitute the worst solutions of the inferior group (equal to ES) with those of the superior group (elite solutions);
4. Then, use Jaya algorithm for modifying the solutions of the individual subgroup, compare a *new candidate solution* with the *older solution*, and accept a *better solution* in each subgroup;
5. Merge all the sub-populations and check whether the previous best result of the whole population (Z_best_before) is better than the current best one in the whole population (Z_best_after). If the value of Z_best_after is greater than that of Z_best_before, the m value is incremented by one (i.e., $m = m + 1$) due to the growth of exploration feature in the search process. Otherwise, the value m will be decremented by 1 (i.e., $m = m - 1$);
6. In the case of reaching the stopping criterion, exit the loop and produce the best optimal result. Otherwise, the identical solutions should be replaced with new randomly created solutions. Then, continue the procedures again from step 3 for resplitting the population.

3.5. The Jaya-ELM and proposed SAMPEJ-ELM algorithm

The current study proposes a SAMPEJ algorithm to optimize the weight and bias of ELM. In the proposed algorithm, ELM is considered as the objective function. The miss-classification (error) rate obtained from comparing the target value with the estimated value observed by ELM is also used as the fitness function for all approaches. To minimize the misclassification rate, ELM is trained using randomly generated weights and bias that leads to sub-optimal solutions. The randomly taken weight and bias values in this model are optimized using SAMPEJ algorithm. According to SAMPEJ algorithm [43], only the best one is forwarded to the next generation. This strategy upgrades the searching procedures of Jaya. Here, the sub-population size is determined adaptively based on the solution of the problem.

According to the quality of the solution, split the population into M numbers of sub-groups (set $M = 2$ initially) and exchange the worst solutions (equals to ES) of the inferior group with those of the superior group (elite solutions). Then, use Jaya algorithm to modify the solutions of any individual subgroup, compare the *new solution* with the *older one*, and accept the *better one* in each subgroup. Merge all the sub populations and check whether the previous best result of the whole population (Z_best_before) is better than the current best one in the entire population (Z_best_after). If the value of Z_best_after is greater than that of Z_best_before, the value of M is incremented by one (i.e., $M = M + 1$) due to an increase in the exploration feature in the search process. Otherwise, the m value will be decremented by 1 (i.e., $M = M - 1$). When reaching the stopping criterion, exit the loop and produce the best optimal result. Otherwise, replace the identical solutions with the new randomly created solutions and continue the procedures of re-splitting the population again.

The steps of the proposed SAMPEJ-ELM algorithm are given in Algorithm 1.

3.6. Description of performance measure

The present study employed several performance measures based on confusion matrix [24] including True Positive (TP), False Negative (FN), True Negative (TN), False Positive (FP) with accuracy percentage, specificity, sensitivity, F-score, Gmean, ROC (graphical representation of sensitivity versus (1-specificity)), and AUC. These performance measures are described in Eqs. (5)–(9).

$$S_n = \frac{TP}{(TP + FN)}, \tag{5}$$

$$S_p = \frac{TN}{(TN + FP)}, \tag{6}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \tag{7}$$

$$Gm = \sqrt{Sp * Sn}, \tag{8}$$

$$Fs = \frac{(2 * Sp * Sn)}{(Sp + Sn)}. \tag{9}$$

4. Experimentation and result validation

4.1. Simulation environment

Processing unit: Intel(R) Core (TM) i5-7200U with 2.5 GHz processing speed, operating system: Windows 10, RAM capacity: 8 GB, and Programming Language platform: R2015b MATLAB.

4.2. Dataset description

In this experimentation, three types of standard cancer namely breast Cancer Wisconsin (Original), cervical cancer, and lung cancer were taken into account. Breast Cancer Wisconsin (Original) is a binary class (Benign and Malignant) dataset that consists of 699 samples, 10 attributes, and one target variable. Cervical Cancer is a binary class dataset containing 858 samples and 36 attributes, and four target variables namely Cytology, Hinselmann, Schiller, and Biopsy. In this study, Biopsy target variable is considered as the class label for experiment. Lung cancer has three dataset classes with 32 samples, 56 attributes, and one target variable.

4.3. Parameter set up

This study empirically compare the performances of Jaya-NN, SAMPEJ-NN, Jaya-NN, Jaya-FLANNELM, TLBO-ELM, ELM-Jaya, and ELM-SAMPEJ. The initial values of the parameters used in these algorithms are as follows: population size (Jaya, SAMPEJ, TLBO): 100; number of iterations (Jaya, SAMPEJ, TLBO): 100; number of hidden layers (ELM):10 to 100 (increment of 10 neurons in each run); number of hidden layer (NN): 3; number of nodes in each hidden layer (NN): 5; size of expansion (FLANN): 10 (for breast cancer), 36 (for cervical cancer), and 17 (for primary tumor); activation function (NN, FLANN, ELM): sigmoidal function.

4.4. Experiment with datasets

In this experimentation, three aforementioned cancer datasets were taken into account to evaluate the proposed technique. The testing accuracy, training accuracy, training time, specificity, sensitivity, F-score, and Gmean were calculated considering four ELM-based models for each dataset with 10–100 hidden nodes (with the addition of 10 neurons in every run).

The proposed algorithm also increases the convergence rate and leads to more investigations into the search area with no confinement to a specific regional

Input: Cancer dataset; Hidden layer size (H_c); Population size (PS); Maximum no. of iteration (K), Elite size (ES)

Output: Accuracy%

Description $newP_i$ is the new population, Obj_i is the objective value of the new population, W is the weight vector, M is the population size (Consider M=2 initially), *train_acc%* and *test_acc%* are taken for *train_data* and *test_data* accuracy percentages respectively

1. Set the 70% dataset as *train_data* and 30% of data as *test_data*
2. Set the no. of random weight population *(H_c),* each population having size of 1× H_c P_i= $\{W_1^i, W_2^i, W_3^i, \ldots, W_{H_c}^i\}$ for i=1,2,3, …PS
3. Generate the initial candidate solution for entire population P by using the equation of ELM algorithm and find *g_best* and *g_worst*
4. While $i < K$
5. The population PS is divided into M subpopulation $\{P_1, P_2, P_3, \ldots, P_M\}$ and the worst solutions (equals to ES=2) of the inferior group with superior group (elite solutions) is replaced
6. For each sub population P_j where j= *1, 2 ,3, ..., M*
7. Get *g_best* P_j where $arg_{min}(Obj_j)$
8. Get *g_worst* P_j where $arg_{max}(Obj_j)$
   by estimating the error value in ELM from Step 9 to Step 13
9. For every *training_input* derive
10. H = *training_input* × P_i
11. β_i = *pseudo_inverse* × *training_output*
12. *obtained_output* = (*testing_input* × P_i) × β_i
13. Compare *obtained_output* and *ts_output* and *miss_classification_rate* (Obj_i) is calculated
14. For each sub-population P_j follow Equation 5.
$$newP_i = P_i + rand(1, H_c) \times (g\_best - |P_j|) - rand(1, H_c) \times (g\_worst - |P_j|) \quad (10)$$
15. Find $nObj_j$ by Step 9 to Step 13
16. If $Obj_j > nObj_j$ (Here, $nObj_i$ and Obj_i are considered as the objective value of the current best solution and previous best solution for the whole population respectively)
17. Exchange P_j with $newP_j$
18. All the sub-populations $\{P_1, P_2, P_3, \ldots, P_M\}$ are merged into P
    If $nObj_i > Obj_i$
    $M = M+1$;
    else if $M>1$
    $M = M-1$;
    End if
19. Repeat from Step 4 -18 till stopping condition reaches
20. *g_best* is taken as the final weight W_{g_best} and hold on the corresponding β_{g_best}
21. *Calculated_output* = (*testing_input* × W_{g_best}) × β_{g_best} and determine the actual label of the class
22. Compare the actual label with expected label of the class and obtain the *miss_classification_rate.*
23. *train_acc_percentage = 1- miss_classification_rate_train_data / sizeof(train_data)*
24. *test_acc_percentage = 1- miss_classification_rate_test_data / sizeof(test_data)*

Algorithm 1. SHMDEJ-ELM algorithm.

Table 1. Comparison of training accuracy and testing accuracy with respect to the number of hidden neurons of all the considered approaches in breast cancer.

	ELM		TLBO-ELM		Jaya-ELM		SAMPEJ-ELM	
HNs	**TRAcc**	**TSAcc**	**TRAcc**	**TSAcc**	**TRAcc**	**TSAcc**	**TRAcc**	**TSAcc**
10	0.9569	0.9308	0.9628	0.9341	0.9631	0.9371	0.9662	0.9491
20	0.9526	0.9384	0.9557	0.9417	0.9588	0.9371	0.9619	0.9467
30	0.9532	0.9416	0.9557	0.9449	0.9594	0.9479	0.9619	0.9499
40	0.9562	0.9416	0.9593	0.9465	0.9624	0.9495	0.9655	0.9515
50	0.9592	0.9459	0.9623	0.9492	0.9624	0.9522	0.9685	0.9642
60	0.9653	0.9459	0.9684	0.9492	0.9715	0.9539	0.9746	0.9659
70	0.9683	0.9483	0.9714	0.9465	0.9715	0.9546	0.9746	0.9566
80	0.9694	0.9486	0.9725	0.9519	0.9737	0.9498	0.9747	0.9739
90	0.9634	0.9516	**0.9736**	0.9549	**0.9796**	0.9579	0.9827	0.9799
100	**0.9675**	**0.9552**	0.9725	**0.9585**	0.9756	**0.9615**	**0.9968**	**0.9895**

Note: TRAcc: Training Accuracy; TSAcc: Testing Accuracy; HN: Number of Hidden Neurons.

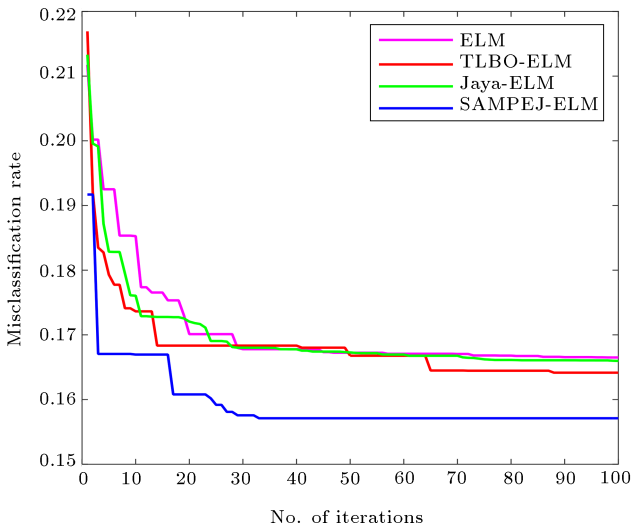

Figure 5. Error convergence graph (breast cancer).

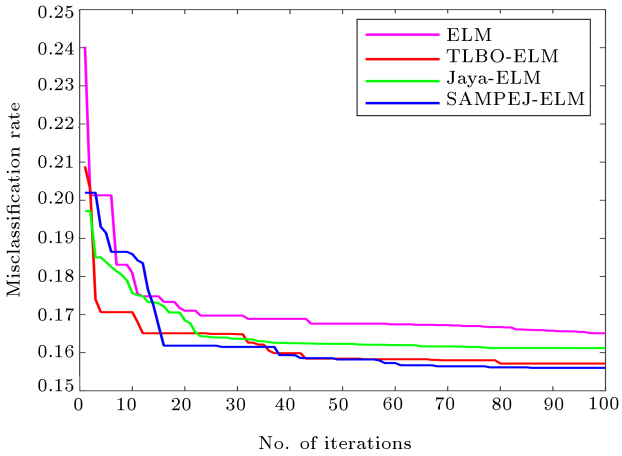

Figure 6. Error convergence graph (cervical cancer).

optimum solution. In this experiment, ELM was used as an objective function. The current study offered the global best candidate solution, determining which solution has the minimum objective value, and the global worst candidate solution, determining which solution has the maximum objective value. The error convergence graphs of all ELM-based approaches are presented in Figures 5–7 in three datasets. Based on Figures 5–7, it can be clearly concluded that the converging speed of the suggested SAMPEJ-ELM approach outperforms others.

4.4.1. Experiment I (breast cancer dataset)

Breast cancer dataset was taken into account to evaluate the proposed model in Experiment I. Table 1

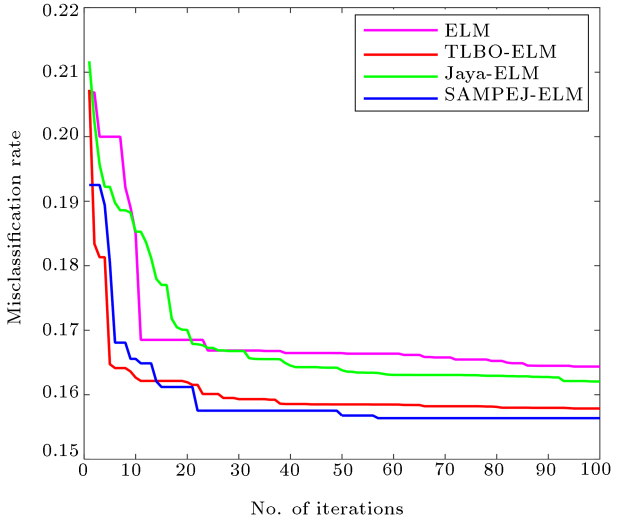

Figure 7. Error convergence graph (lung cancer).

Table 2. Comparison of sensitivity (Sn), TP, FN, FP, and TN values of all the considered approaches in breast cancer.

HNs	ELM					TLBO-ELM					Jaya-ELM					SAMPEJ-ELM				
	Sn	TP	FN	FP	TN	Sn	TP	FN	FP	TN	Sn	TP	FN	FP	TN	Sn	TP	FN	FP	TN
10	.9310	135	10	10	54	.9448	137	8	8	56	.9655	140	5	6	58	.9931	143	1	3	62
20	**.9328**	125	9	9	66	.9478	127	7	7	68	.9701	130	4	5	70	**1**	141	0	2	66
30	.9241	134	11	8	56	.9379	136	9	6	58	.9586	139	6	4	60	.9861	142	2	1	64
40	.9254	124	10	9	66	.9403	126	8	7	68	.9627	129	5	5	70	.9929	140	1	2	66
50	.9310	135	10	7	57	.9448	137	8	5	59	.9655	140	5	3	61	.9931	143	1	1	64
60	.9328	125	9	10	65	**.9478**	127	7	8	67	**.9701**	130	4	6	69	**1**	141	0	3	65
70	.9241	134	11	9	55	.9379	136	9	7	57	.9586	139	6	5	59	.9861	142	2	2	63
80	.9179	123	11	8	67	.9328	125	9	6	69	.9552	128	6	4	71	.9858	139	2	1	67
90	.9310	135	10	9	55	.9448	137	8	7	57	.9655	140	5	5	59	.9931	143	1	2	63
100	.9030	121	13	9	66	.9179	123	11	7	68	.9403	126	8	5	70	.9716	137	4	2	66

Table 3. Comparison of specificity (Sp), F-score (Fs), and Gmean (Gm) values using all the considered approaches in breast cancer dataset.

	ELM			TLBO-ELM			Jaya-ELM			SAMPEJ-ELM		
HNs	**Sp**	**Fs**	**Gm**	**Sp**	**Fs**	**Gm**	**Sp**	**Fs**	**Gm**	**Sp**	**Fs**	**Gm**
10	.8438	.8852	.8863	.8750	.9086	.9092	.9063	.9349	.9354	.9538	.9731	.9733
20	.8800	.9056	.9060	.9067	.9268	.9270	.9333	.9514	.9516	.9706	.9851	.9852
30	.8750	.8989	.8992	.9063	.9218	.9220	.9375	.9479	.9480	.9846	.9854	.9854
40	.8800	.9021	.9024	.9067	.9232	.9233	.9333	.9478	.9479	.9706	.9816	.9817
50	.8906	.9104	.9106	**.9219**	.9332	.9333	**.9531**	.9593	.9593	.9846	.9965	.9965
60	.8667	.8985	.8991	.8933	.9197	.9201	.9200	.9444	.9447	.9559	.9774	.9777
70	.8594	.8906	.8912	.8906	.9137	.9140	.9219	.9399	.9401	.9692	.9776	.9776
80	**.8933**	.9055	.9055	.9200	.9264	.9264	.9467	.9509	.9509	**.9853**	.9856	.9856
90	.8594	.8938	.8945	.8906	.9169	.9173	.9219	.9432	.9434	.9692	.9810	.9811
100	.880	.8913	.8914	.9067	.9123	.9123	.9333	.9368	.9368	.9706	.9711	.9711

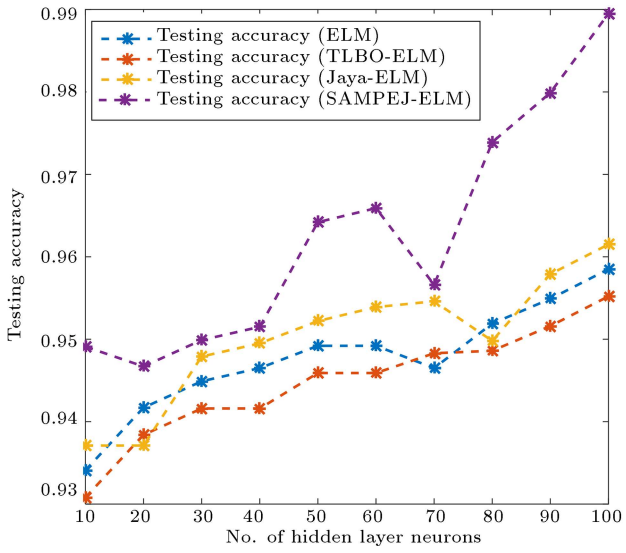

Figure 8. Number of hidden layers versus testing accuracy of all the considered approaches.

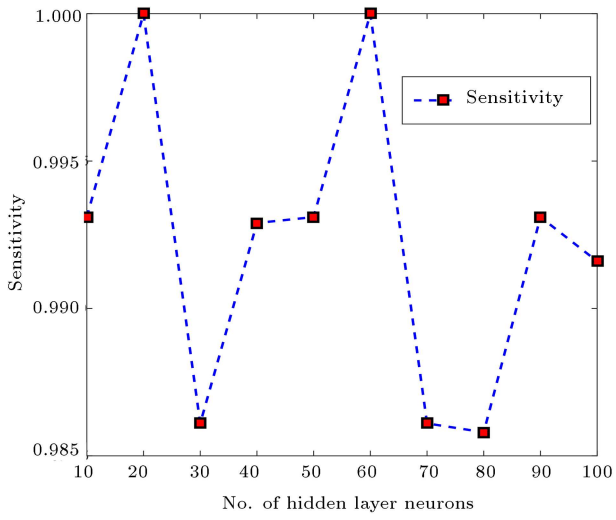

Figure 9. Number of hidden layers versus sensitivity (SAMPEJ-ELM).

shows the training with the testing accuracy of four ELM-based models. Based on the results of Table 1, the graph in Figure 8 presents a comparison among the testing accuracy values of the ELM classified approaches. Table 2 shows a comparison of TP, FN, FP, TN, and sensitivity values. Table 3 presents a comparison between specificity, F-score, and Gmean values of all approaches. Based on the results from Table 2, Figure 9 determines the number of Hidden

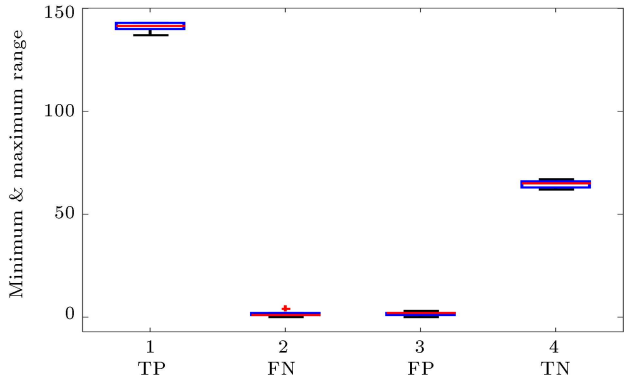

Figure 10. Box plot (SAMPEJ-ELM).

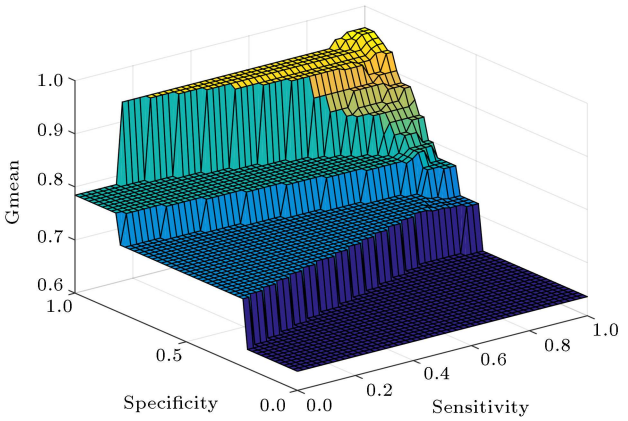

Figure 11. Sensitivity versus specificity versus Gmean of SAMPEJ-ELM.

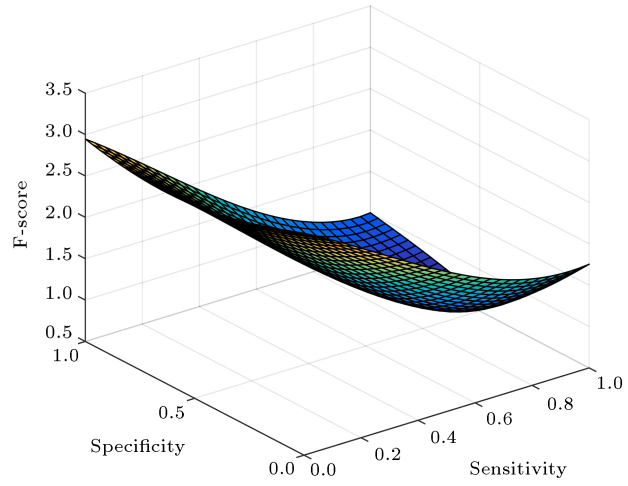

Figure 12. Sensitivity versus specificity versus F-score of SAMPEJ-ELM.

Neurons (HNs) versus sensitivity of all ELM-based approaches. According to the results of Table 2, Figure 10 shows the box plot of SAMPEJ-ELM model. Figure 11 presents sensitivity versus specificity versus Gmean graph. Figure 12 displays sensitivity versus specificity versus F-score graph of the SAMPEJ-ELM model. Figure 13 illustrates the ROC plots of all considered approaches with their AUC values, respectively.

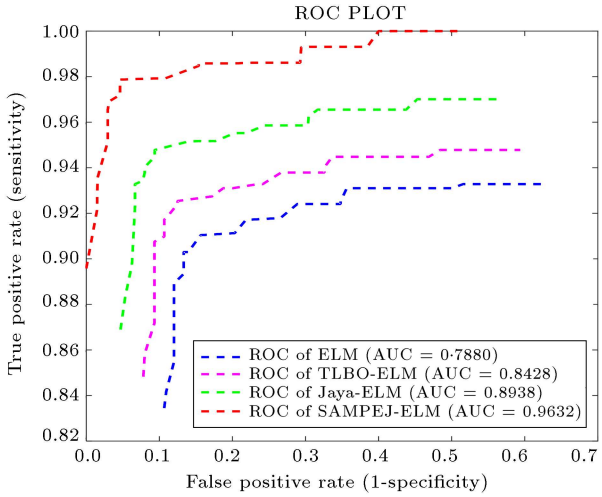

Figure 13. ROC (TPR versus FPR) of ELM, TLBO-ELM, Jaya-ELM, and SAMPEJ-ELM with AUC.

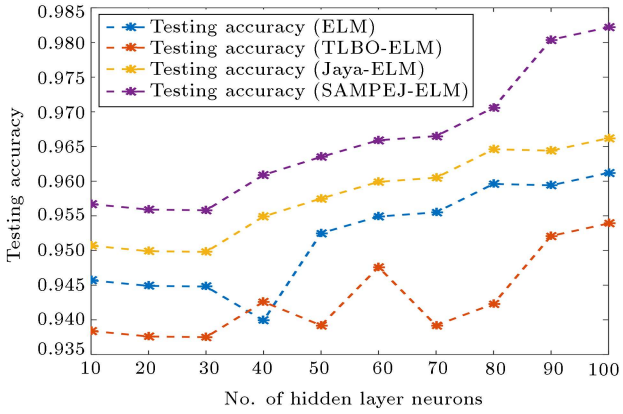

Figure 14. Number of hidden layers versus testing accuracy of all the considered approaches.

4.4.2. Experiment II (cervical cancer dataset)

Cervical cancer dataset was used for assessing the performance of the proposed model in Experiment II. Table 4 discusses the training with the testing accuracy values of ELM, TLBO-ELM, Jaya-ELM, and SAMPEJ-ELM. Based on the result of Table 4, Figure 14 presents a comparison of the testing accuracy values of all ELM-based models. Table 5 shows a comparison among FP, TN, TP, FN, and sensitivity values, and Table 6 presents a comparison among specificity, F-score, and Gmean scores of four ELM wrapped approaches. Based on the result of Table 5, Figure 15 compares the NHs with the sensitivity of all ELM hybridised models. According to the result of Table 5, Figure 16 illustrates the boxplots of all the proposed approaches. Figure 17 presents sensitivity versus specificity versus Gmean graph, and Figure 18 displays sensitivity versus specificity versus F-score graph SAMPEJ-ELM model. Figure 19 illustrates the ROC plots of all the considered approaches with their AUC values, respectively.

Table 4. Comparison of training accuracy and testing accuracy with respect to number of hidden neurons of all considered approaches in cervical cancer dataset.

	ELM		TLBO-ELM		Jaya-ELM		SAMPEJ-ELM	
HNs	**TRAcc**	**TSAcc**	**TRAcc**	**TSAcc**	**TRAcc**	**TSAcc**	**TRAcc**	**TSAcc**
10	0.9432	0.9384	0.9508	0.9457	0.9558	0.9507	0.9618	0.9567
20	0.9438	0.9376	0.9514	0.9449	0.9558	0.9499	0.9624	0.9559
30	0.9473	0.9375	0.9519	0.9448	0.9599	0.9498	0.9653	0.9558
40	0.9473	0.9426	0.9559	0.9399	0.9609	0.9549	0.9669	0.9609
50	0.9468	0.9392	0.9538	0.9525	0.9619	0.9575	0.9669	0.9635
60	0.9514	0.9476	0.9590	0.9549	0.9598	0.9599	0.9700	0.9659
0	0.9515	0.9392	0.9611	0.9555	0.9661	0.9605	0.9691	0.9665
80	0.9578	0.9423	0.9614	0.9594	0.9704	0.9646	0.9782	0.9706
90	0.9566	0.9521	**0.9701**	0.9596	0.9700	0.9644	0.9832	0.9804
100	**0.9625**	**0.9539**	0.9592	**0.9612**	**0.9751**	**0.9662**	**0.9911**	**0.9822**

Note: TRAcc: Training Accuracy; TSAcc: Testing Accuracy; HN: No. of Hidden Neuron.

Table 5. Comparison of sensitivity, TP, FN, FP, and TN values of all the considered approaches in cervical cancer dataset.

HNs	**ELM**					**TLBO-ELM**					**Jaya-ELM**					**SAMPEJ-ELM**				
	Sn	TP	FN	FP	TN	Sn	TP	FN	FP	TN	Sn	TP	FN	FP	TN	Sn	TP	FN	FP	TN
10	.9312	176	13	13	55	.9365	177	12	17	51	.9733	182	5	15	55	.9312	196	3	1	57
20	.9421	179	11	11	56	.9474	180	10	13	54	.9840	185	3	11	58	.9895	188	2	5	62
30	.9430	182	11	11	53	.9482	183	10	14	50	.9843	188	3	12	54	.9896	191	2	6	58
40	.9474	180	10	10	57	**.9536**	181	9	12	55	.9894	186	2	10	5 9	.9947	189	1	4	63
50	**.9485**	184	10	10	53	.9526	185	9	10	53	**.9896**	190	2	8	5 7	**.9948**	193	1	2	61
60	.9309	175	13	13	56	.9362	176	12	21	48	.9731	181	5	19	52	.9787	184	4	13	56
70	.9211	175	15	15	52	.9263	176	14	20	47	.9628	181	7	18	51	.9684	184	6	12	55
80	.9371	164	17	17	59	.9429	165	10	22	60	.9827	170	3	20	64	.9886	173	2	14	68
90	.9397	187	12	10	48	.9447	188	11	9	49	.9797	193	4	7	53	.9812	185	4	9	59
100	.9250	152	18	16	71	.9313	149	11	35	62	.9747	154	4	33	66	.9849	157	3	27	70

Table 6. Comparison of specificity, F-score, Gmean values using all the considered approaches in cervical cancer dataset.

HNs	**ELM**			**TLBO-ELM**			**Jaya-ELM**			**SAMPEJ-ELM**		
	Sp	F-score	Gmean	Sp	F-score	Gmean	Sp	F-score	Gmean	Sp	F-score	Gmean
10	.8254	.8827	.8848	.8413	.8939	.8957	.8769	.9299	.9316	**.9828**	.9838	.9838
20	.7910	.8600	.8633	.8060	.8710	.8738	.8406	.9067	.9095	.9254	.9564	.9569
30	.7656	.8451	.8497	.7813	.8567	.8607	.8182	.8936	.8974	.9063	.9461	.9470
40	.8060	.8710	.8738	.8209	.8819	.8843	.8551	.9173	.9198	.9403	.9668	.9671
50	**.8276**	.8801	.8819	**.8448**	.8920	.8934	**.8833**	.9290	.9303	.9683	.9814	.9815
60	.6812	.7867	.7963	.6957	.7982	.8070	.7324	.8358	.8442	.8116	.8874	.8912
70	.6866	.7867	.7952	.7015	.7984	.8061	.7391	.8363	.8436	.8209	.8886	.8916
80	.7195	.8140	.8211	.7317	.8240	.8306	.7619	.8583	.8653	.8293	.9019	.9054
90	.7353	.8217	.8275	.7500	.8329	.8381	.7857	.8695	.8745	.8676	.9199	.9216
100	.6289	.7487	.7627	.6392	.7581	.7715	.6667	.7918	.8061	.7216	.8317	.8415

4.4.3. Experiment III (lung cancer dataset)

Lung cancer dataset was employed to evaluate the proposed model in Experiment III. Table 7 shows the training with the testing accuracy of four ELM hybridized models. Based on the results of Table 7, Figure 20 presents a comparison of the training and testing accuracy values of all the considered approaches. Table 8 also presents a comparison of the sensitivity, TP, FN, FP, and TN values. Table 9 presents a comparison of specificity, F-score, and Gmean values

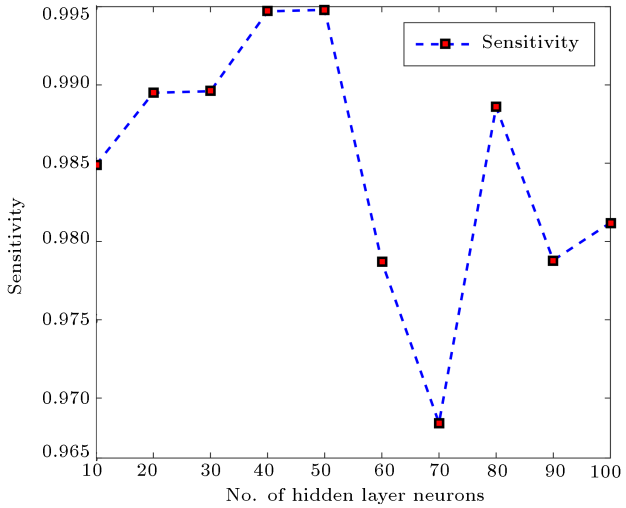

Figure 15. Number of hidden layers versus sensitivity score (SAMPEJ-ELM).

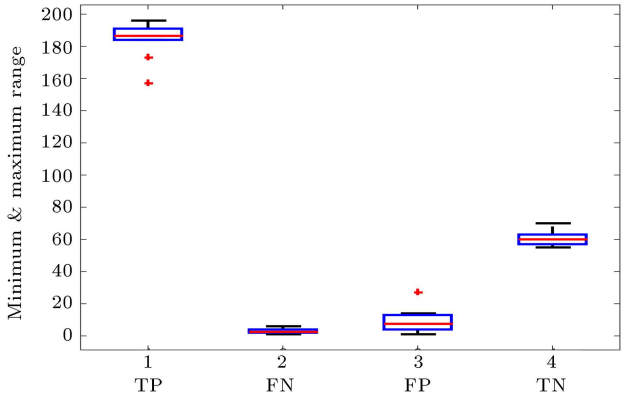

Figure 16. Box plot of SAMPEJ-ELM.

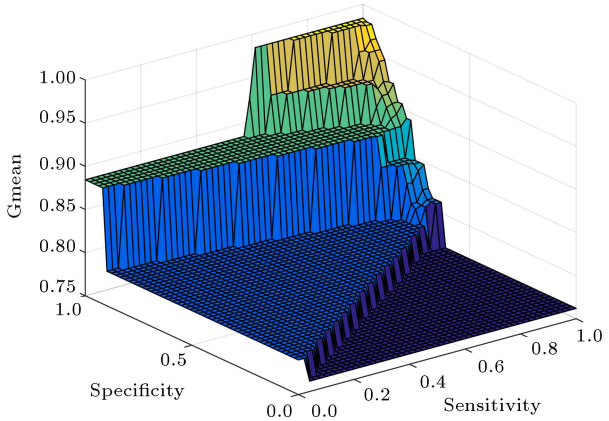

Figure 17. Sensitivity versus specificity versus Gmean of SAMPEJ-ELM.

of all the considered approaches. Based on the results given in Table 8, Figure 21 displays the sensitivity with respect to the number of HNs of ELM-based models. According to the result in Table 8, Figure 22 illustrates the box plots of all the considered approaches. Figure 23 presents sensitivity versus specificity versus Gmean graph. Figure 24 shows the sensitivity versus

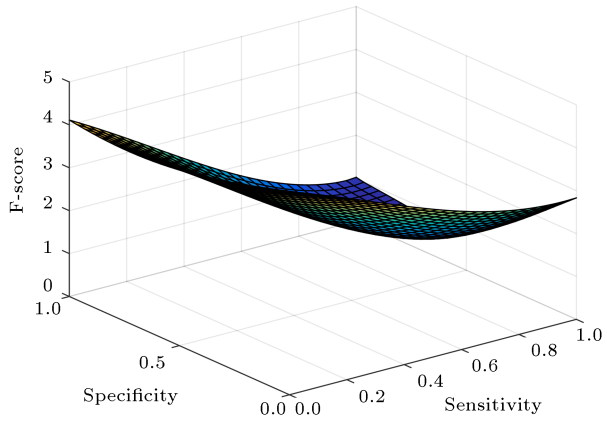

Figure 18. Sensitivity versus specificity versus F-score of SAMPEJ-ELM.

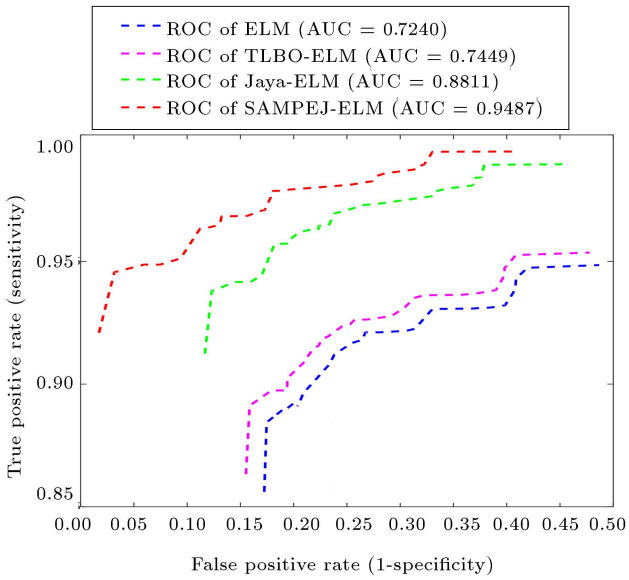

Figure 19. ROC (TPR vs. FPR) of ELM, TLBO-ELM, Jaya-ELM, and SAMPEJ-ELM with AUC.

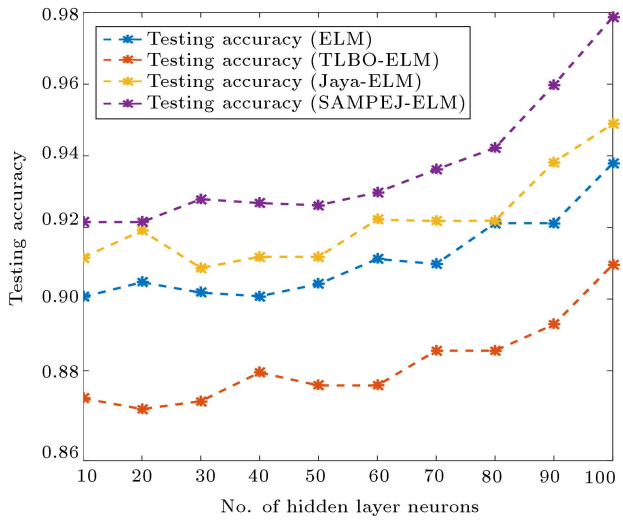

Figure 20. Number of HNs vs. testing accuracy of ELM, TLBO-ELM, Jaya-ELM, and SAMPEJ-ELM.

Table 7. Comparison of training accuracy and testing accuracy with respect to number of hidden neurons of all the considered approaches in lung cancer dataset.

	ELM		TLBO-ELM		Jaya-ELM		SAMPEJ-ELM	
HNs	**TRAcc**	**TSAcc**	**TRAcc**	**TSAcc**	**TRAcc**	**TSAcc**	**TRAcc**	**TSAcc**
10	.8894	.8723	.9180	.9006	.9292	.9116	.9389	.9215
20	.8876	.8693	.9162	.9047	.9274	.9191	.9388	.9215
30	.8776	.8715	.9062	.9018	.9174	.9087	.9295	.9279
40	.8927	.8795	.9213	.9008	.9325	.9118	.9428	.9268
50	.8734	.8759	.9020	.9042	.9132	.9118	.9335	.9262
60	.8860	.8759	.9146	.9112	.9258	.9222	.9355	.9298
70	.9282	.8856	.9568	.9098	.9680	.9218	.9783	.9362
80	.9240	.8856	.9590	.9212	.9638	.9219	.9739	.9422
90	**.9304**	.8929	.9584	.9212	**.9702**	.9382	.9893	.9598
100	.9298	**.9487**	.9526	**.9379**	.9696	**.949**	**.9918**	**.9787**

Note: TRAcc: Training Accuracy; TSAcc: Testing Accuracy; HN: Number of Hidden Neuron.

Table 8. Comparison of TP, FN, FP, TN, and sensitivity score of four ELM-based models in lung cancer dataset.

HNs	**ELM**					**TLBO-ELM**					**Jaya-ELM**					**SAMPEJ-ELM**				
	Sn	**TP**	**FN**	**FP**	**TN**	**Sn**	**TP**	**FN**	**FP**	**TN**	**Sn**	**TP**	**FN**	**FP**	**TN**	**Sn**	**TP**	**FN**	**FP**	**TN**
10	.5	2	2	3	3	.5	3	3	3	1	.6	3	2	3	2	.6	3	2	2	3
20	.5	3	3	2	2	.6	3	2	2	3	.6	3	2	2	3	.8	4	1	2	3
30	.6	3	2	3	2	.6	3	2	2	3	.6	3	2	2	3	.6667	4	2	1	3
40	.4	2	3	3	1	.5	3	3	3	2	.4	2	3	3	2	.75	3	1	2	4
50	.6667	4	2	3	1	.5	3	3	2	2	.6667	4	2	1	3	.6667	4	2	1	4
60	.5714	4	3	1	2	.5	2	2	2	4	.6667	4	2	1	3	1	4	0	1	5
70	.6	3	2	2	3	.5	3	3	2	2	.75	3	1	2	4	.8333	5	1	1	3
80	.75	3	1	3	3	.8	4	1	1	4	.75	3	1	3	3	1	5	0	1	4
90	**.8333**	5	1	1	3	**.8333**	5	1	2	2	**.8333**	5	1	1	3	1	5	0	0	5
100	.7143	5	2	1	4	.75	3	1	1	4	.8	4	1	1	4	.8333	5	1	0	4

Table 9. Comparison of specificity, F-score, and Gmean values using all the considered approaches in lung cancer.

HNs	**ELM**			**TLBO-ELM**			**Jaya-ELM**			**SAMPEJ-ELM**		
	Sp	Gmean	F-score	Sp	Gmean	F-score	Sp	Gmean	F-score	Sp	Gmean	F-score
10	.5	.5	.5	.25	.3536	.3333	.4	.4899	.48	.6	.6	.6
20	.5	.5	.5	.6	.6	.6	.6	.6	.6	.6	.6928	.6857
30	.4	.4899	.48	.6	.6	.6	.6	.6	.6	.75	.7071	.7059
40	.25	.3162	.3077	.4	.4472	.4444	.4	.4	.4	.6667	.7071	.7059
50	.25	.4083	.3636	.5	.5	.5	.75	.7071	.7059	.8	.7303	.7273
60	.6667	.6172	.6154	.6667	.5774	.5714	.75	.7071	.7059	.8333	.9129	.9091
70	.6	.6	.6	.5	.5	.5	.6667	.7071	.7059	.75	.7906	.7895
80	.5	.6124	.6	.8	.8	.8	.5	.6124	.6	.8	.8944	.8889
90	.75	.7906	.7895	.5	.6455	.625	.75	.7906	.7895	1	1	1
100	**.8**	.7559	.7547	**.8**	.7746	.7742	**.8**	.8	.8	1	.9129	.9091

specificity versus F-score graph of the SAMPEJ-ELM model. Figure 25 shows the ROC of all the considered approaches with their AUC values, respectively.

4.5. Result analysis

The present study employed a metaheuristic algorithm-optimized ELM model called SAMPEJ-ELM to classify three types of cancer datasets namely breast, cervical, and lung cancers. The calculated results were then compared with those from ELM, TLBO-ELM, Jaya-ELM, Jaya-NN, SAMPEJ-NN, Jaya-FLANN, and SAMPEJ-FLANN. Here, testing with training accuracy, sensitivity, Gmean, specificity, F-score, and ROC with AUC values were taken as performance evaluation measures. From Tables 1, 4 and 7, it is clearly observed that SAMPEJ-ELM needs fewer HNs than all

Table 10. Comparison of the maximum testing accuracy of all models in three cancer datasets.

Dataset	Jaya-NN	SAMPEJ-NN	Jaya-FLANN	SAMPEJ-FLANN	ELM	Jaya-ELM	SAMPEJ-ELM
Breast cancer	.8974	.9012	.9274	.9552	.9312	.9615	**.9895**
Cervical cancer	.8102	.8365	.8936	.9539	.9182	.9662	**.9822**
Lung cancer	.7432	.7567	.7967	.9096	.9487	.949	**.9787**

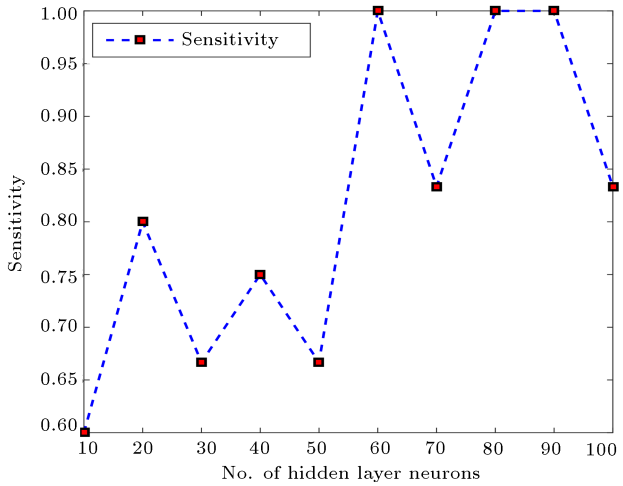

Figure 21. Number of hidden layers vs. sensitivity (SAMPEJ-ELM).

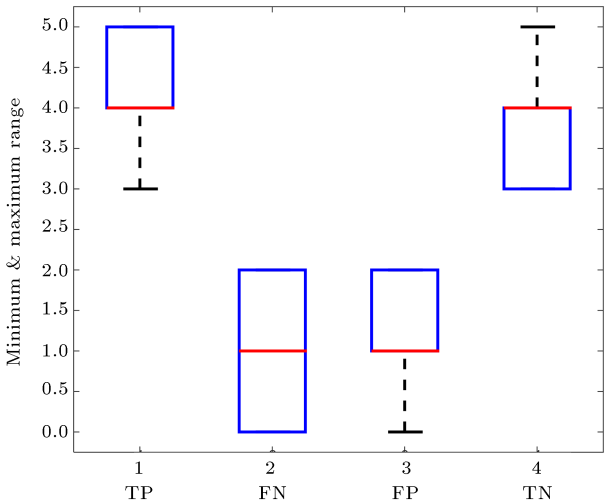

Figure 22. Box plot (SAMPEJ-ELM).

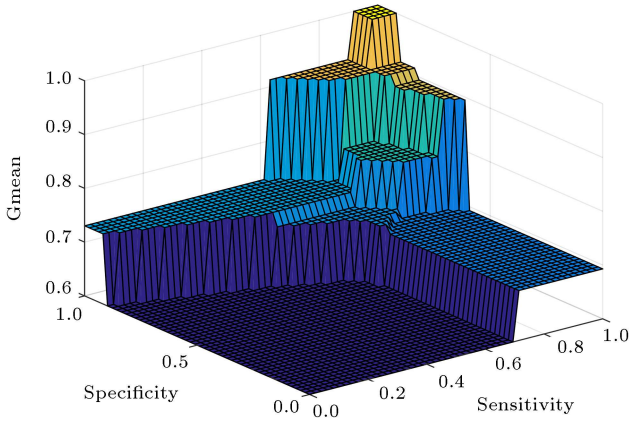

Figure 23. Sensitivity vs. specificity vs. Gmean of SAMPEJ-ELM.

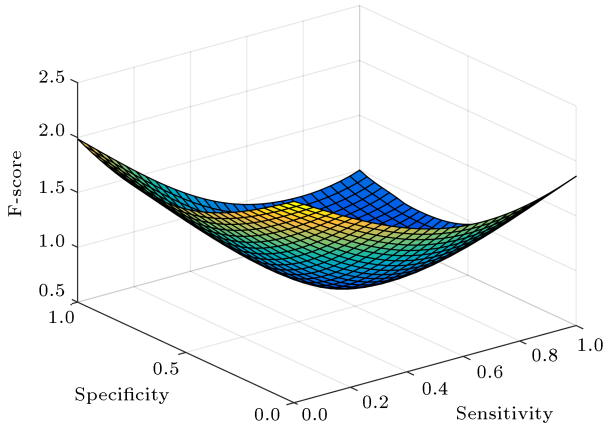

Figure 24. Sensitivity vs. specificity vs. F-score of SAMPEJ-ELM.

the considered approaches to achieve the same testing accuracy. Table 10 presents the comparison between the testing accuracy of various ELM wrapped models and NN wrapped models and proves the superiority of SAMPEJ-ELM model. Table 11 shows AUC scores of all the considered approaches in three cancer datasets.

Likewise, the primacy of SAMPEJ-ELM approach is easily visualized in Figures 11-12, Figures 17-18, and Figures 23-24 graphs of sensitivity versus specificity versus Gmean in three cancer datasets. According to Tables 1, 4, and 7, the sensitivity of SAMPEJ-ELM model is higher than that of other models. Sensitivity implies that positive samples are well classified.

In ROC graphs (Figures 13, 19, and 25), the SAMPEJ-ELM model outperforms the other ELM wrapped approaches. The AUC values of the respective approach are also given in the ROC graphs. It is clearly visualized from the AUC values that the proposed model is superior to others.

In case of the SAMPEJ-ELM model, the box plots (Figures 10, 16, and 22) represent TP, TN rate exceeding FP, and FN rate in three datasets, specifying that negative and positive samples are correctly classified.

Tables 3, 6, and 9 also clearly state that specificity is greater in SAMPEJ-ELM model than other models.

Table 11. AUC values of four ELM hybridized models in three cancer datasets.

Dataset	ELM	TLBO-ELM	Jaya-ELM	SAMPEJ-ELM
Breast cancer	.7880	.8428	.8938	**.9632**
Cervical cancer	.7240	.7449	.8811	**.9487**
Lung cancer	.7440	.7549	.8411	**.9627**

Table 12. A comparative analysis of the suggested approach versus other existing techniques among three cancer datasets (the '–' sign declares the missing of data).

Techniques used	Accuracy % in three cancer datasets		
	Breast cancer	Cervical cancer	Lung cancer
SVM-RBF [48]	96.84	–	–
Random forest [49]	–	89	–
SVM [50]	96.99	–	–
measure based on Yu's norms [51]	96.28	–	**99.99**
Random Forest [52]	-	97.6	-
CART [53]	94.84	–	–
Synthetic Minority Oversampling Technique-PCA [54]	–	–	81
PCA+c4.5 [55]	–	90.70	-
SMO+J48+NB+IBk [56]	97.28	–	–
Genetic algorithm and fuzzy system [57]	-	-	97.5
Random forest [58]	95.78	–	–
SAMPEJ-ELM (Proposed model)	**98.95**	**98.22**	97.87

Note: CART: Classification and Regression Trees;

PCA: Principal Component Analysis; SMO: Sequential Minimal Optimization;

IBK: Instance Based for K-nearest neighbor; J48 and c4.5: Different types of decision tree.

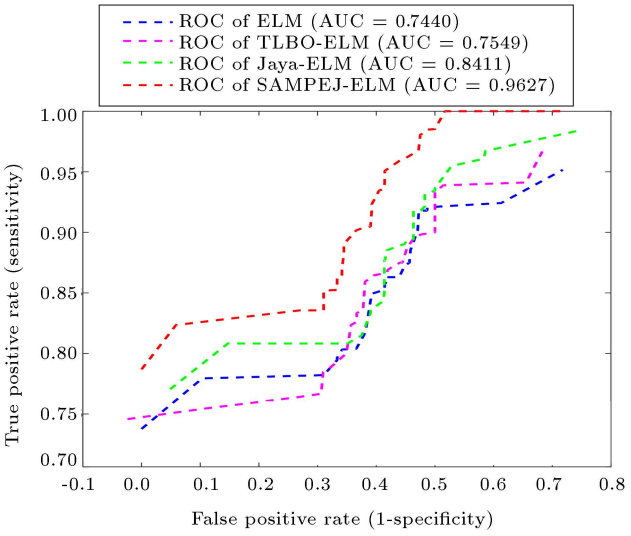

Figure 25. ROC (TPR vs. FPR) of ELM, TLBO-ELM, Jaya-ELM, and SAMPEJ-ELM with AUC.

This work determines that SAMPEJ-ELM approach better classifies the negative samples.

From the basis of the above observations, it is distinctly proved that basic ELM wrapped approaches are better classifiers than NN wrapped approaches. Upon comparing the ELM wrapped approaches with SAMPEJ-ELM, the SAMPEJ-ELM approach is certifiably significantly better.

4.6. Comparison with existing approaches

In this paper, the accuracy% of the presented approach is compared with those of eleven other standard approaches. From this qualitative analysis, a conclusive result may not be drawn as various approaches apply various evaluating measures and various datasets to evaluate their performance. A comparative analysis gives an approximate estimation of the presented technique compared to other standard approaches.

Table 12 displays a qualitative analysis of the presented technique with 11 existing techniques in three cancer datasets. According to this comparison, the proposed model excels in breast and cervical cancer datasets. In the model [51], the lung cancer dataset outperforms the proposed model.

4.7. Statistical results

To examine the means of distinct groups (whether or not they are the same), the most popular statistical approach, i.e., Analysis of variance (ANOVA), is applied in this work. This statistical approach helps

Table 13. ANOVA test with respect to classification accuracy (%).

Dataset	ELM	Jaya-ELM	SAMPEJ- ELM	Total
Number of datasets	3	3	3	9
$\sum X$	285.78	287.67	295.04	868.49
Mean	95.26	95.89	98.34667	96.49889
$\sum X2$	27223.6394	27586.26	19536.14	47122.4
Standard deviation	0.085986918	0.888988	0.551029	0.254334

Table 14. Result details of the ANOVA test.

Source	SS	df	MS	
Between-treatments	15.95962222	2	7.979811111	f-value = 19.74820579
Within-treatments	2.424466667	6	0.404077778	p-value = 0.048974974
Total	18.384088887	8	8.383888889	

estimate the model statistically. Generally, null and alternative hypotheses are incorporated into ANOVA. In this test, F-value is calculated first; then, according to F-value, p-value is determined. The obtained p-value of ANOVA finalizes whether to keep or discard the null hypothesis. The null hypothesis is refused if p-value ≤ 0.05 (taking 5% as the significance level) and it can be decided that the accuracy percentages of all the algorithms are undoubtably different. Moreover, the detailed statistical analysis of ANOVA test is shown in Tables 13 and 14. In this work, the observed p-value is 0.048974974, which is quite lower than the set p-value (i.e., 0.05). Henceforth, the null hypothesis is rejected. Therefore, it is decided that the presented algorithm is statistically better than other ELM-based models.

5. Conclusion

The current study put its main focus on the cancer data classification through the proposed SAMPEJ-ELM model. Here, three datasets of breast, cervical, and lung cancers were taken into account to perform experiments. In this study, three Extreme Learning Machine (ELM) hybridised approaches (ELM, TLBO-ELM, and Jaya-ELM) and four NN-based models (Jaya-NN, SAMPEJ-NN, Jaya-FLANN, and SAMPEJ-FLANN) were matched with the suggested SAMPEJ-ELM model. All of these models were evaluated based on a series of empirical studies. Several performance metrices namely the accuracy%, specificity, sensitivity, Gmean, F-score, and Receiver Operating Characteristic (ROC) graphs were used for an unbiased comparison with other models. Based on different observations, it can be clearly concluded that all the ELM-based classifiers outperformed the NN-based models. However, to develop a more robust and stable classifier, the random parameters of ELM were optimized by the Self-Adaptive Multi-Population-based Elite strategy Jaya (SAMPEJ) algorithm. The findings revealed that the SAMPEJ-ELM could efficiently handle the ill-conditioned problem and achieve better accuracy than other models. Further, the suggested SAMPEJ-ELM approach could efficiently classify cancer data. Therefore, the proposed model can be successfully for microarray data classification.

References

1. Parkin, D.M., Freddie Bray, F., Jacques Ferlay, F., et al. "Estimating the world cancer burden: Globocan 2000", *International Journal of Cancer*, **94**(2), pp. 153–156 (2001).
2. Cooper, J.S., Porter, K., Mallin, K., et al. "National cancer database report on cancer of the head and neck: 10-year update", *Head & Neck: Journal for the Sciences and Specialties of the Head and Neck*, **31**(6), pp. 748–758 (2009).
3. Duda, R.O. and Hart, P.E., *Pattern Classification*, John Wiley & Sons (2006).
4. Zurada, J.M., *Introduction to Artificial Neural Systems*, **8**, St. Paul: West (1992).
5. Naik, B., Nayak, J., Behera, H.S., et al. "A harmony search based gradient descent learning-FLANN (HS-GDL-FLANN) for classification", In *Computational Intelligence in Data Mining*, **2**, pp. 525–539. Springer, New Delhi (2015).
6. Bahrololoum, A., Nezamabadi-Pour, H., Bahrololoum, H., et al. "A prototype classifier based on gravitational search algorithm", *Applied Soft Computing*, **12**(2), pp. 819–825 (2012).
7. Fernández-Navarro, F., Hervás-Martínez, C., Ruiz, R., et al. "Evolutionary generalized radial basis function neural networks for improving prediction accuracy in gene classification using feature selection", *Applied Soft Computing*, **12**(6), pp. 1787–1800 (2012).
8. Aydogan, E.K., Karaoglan, I., and Pardalos, P.M. "hGA: Hybrid genetic algorithm in fuzzy rule-based

classification systems for high-dimensional problems", *Applied Soft Computing*, **12**(2), pp. 800–806 (2012).

9. Heermann, P.D. and Khazenie, N. "Classification of multispectral remote sensing data using a back-propagation neural network", *IEEE Transactions on Geoscience and Remote Sensing*, **30**(1), pp. 81–88 (1992).

10. Malathi, V., Marimuthu, N.S., and Baskar, S. "Intelligent approaches using support vector machine and extreme learning machine for transmission line protection", *Neurocomputing*, **73**(10–12), pp. 2160–2167 (2010).

11. Cristianini, N. and Shawe-Taylor, J. "An introduction to support vector machines and other kernel-based learning methods", Cambridge university press (2000).

12. Huang, G.B., Zhu, Q.Y., and Siew, C.K. "Extreme learning machine: theory and applications", *Neuro Computing*, **70**(1–3), pp. 489–501 (2006).

13. Huang, G.B., Zhu, Q.Y., and Siew, C.K. "Extreme learning machine: a new learning scheme of feed-forward neural networks", *IEEE International Joint Conference on Neural Networks (IEEE Cat. No. 04CH37541)*, **2**, pp. 985–990 (2004).

14. Huang, G.B., Zhou, H., Ding, X., et al. "Extreme learning machine for regression and multiclass classification", *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **42**(2), pp. 513–529 (2011).

15. Huang, G.B., Wang, D.H., and Lan, Y. "Extreme learning machines: a survey", *International Journal of Machine Learning and Cybernetics*, **2**(2), pp. 107–122 (2011).

16. Wang, D. and Alhamdoosh, M. "Evolutionary extreme learning machine ensembles with size control", *Neurocomputing*, **102**, pp. 98–110 (2013).

17. Huang, G.B. and Wang, D. "Advances in extreme learning machines (ELM2010)", *Neurocomputing*, **16**(74), pp. 2411–2412 (2011).

18. Feng, G., Huang, G.B., Lin, Q., et al. "Error minimized extreme learning machine with growth of hidden nodes and incremental learning", *IEEE Transactions on Neural Networks*, **20**(8), pp. 1352–1357 (2009).

19. Zhao, G., Shen, Z., Miao, C., et al. "On improving the conditioning of extreme learning machine: a linear case", In *2009 7th International Conference on Information, Communications and Signal Processing (ICICS)*, pp. 1–5, IEEE (2009).

20. Chen, C.P. and Wan, J.Z. "A rapid learning and dynamic stepwise updating algorithm for flat neural networks and the application to time-series prediction", *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, **29**(1), pp. 62–72 (1999).

21. Pacifico, L.D. and Ludermir, T.B. "Improved evolutionary extreme learning machines based on particle swarm optimization and clustering approaches", *International Journal of Natural Computing Research (IJNCR)*, **3**(3), pp. 1–20 (2012).

22. Han, F., Yao, H.F., and Ling, Q.H. "An improved evolutionary extreme learning machine based on particle swarm optimization", *Neurocomputing*, **116**, pp. 87–93 (2013).

23. Baliarsingh, S.K. and Vipsita, S. "Chaotic emperor penguin optimised extreme learning machine for microarray cancer classification", *IET Syst Biol*, **14**(2), pp. 85–95 (2020).

24. Feng, G., Huang, G.B., Lin, Q., et al. "Error minimized extreme learning machine with growth of hidden nodes and incremental learning", *IEEE Transactions on Neural Networks*, **20**(8), pp. 1352–1357 (2009).

25. Zhu, Q.Y., Qin, A.K., Suganthan, P.N., et al. "Evolutionary extreme learning machine. Pattern recognition", **38**(10), pp. 1759–1763 (2005).

26. Yang, X.S. and Deb, S. "Multiobjective cuckoo search for design optimization", *Computers & Operations Research*, **40**(6), pp. 1616–1624 (2013).

27. Aha, D.W. "Tolerating noisy, irrelevant and novel attributes in instance-based learning algorithms", *International Journal of Man-Machine Studies*, **36**(2), pp. 267–287 (1992).

28. Rashno, A., Nazari, B., Sadri, S., et al. "Effective pixel classification of mars images based on ant colony optimization feature selection and extreme learning machine", *Neurocomputing*, **226**, pp. 66–79 (2017).

29. Mohapatra, P., Chakravarty, S., and Dash, P.K. "An improved cuckoo search based extreme learning machine for medical data classification", *Swarm and Evolutionary Computation*, **24**, pp. 25–49 (2015).

30. Mohapatra, P., Chakravarty, S., and Dash, P.K. "Microarray medical data classification using kernel ridge regression and modified cat swarm optimization based gene selection system", *Swarm and Evolutionary Computation*, **28**, pp. 144–160 (2016).

31. Rao, R. "Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems", *International Journal of Industrial Engineering Computations*, **7**(1), pp. 19–34 (2016).

32. Rao, R.V., Savsani, V.J., and Vakharia, D.P. "Teaching-learning-based optimization: an optimization method for continuous non-linear large scale problems", *Information Sciences*, **183**(1), pp. 1–15 (2012).

33. Rao, R.V., More, K., Taler, J., et al. "Dimensional optimization of a micro-channel heat sink using Jaya

algorithm", *Applied Thermal Engineering*, **103**, pp. 572–582 (2016).

34. Warid, W., Hizam, H., Mariun, N., et al. "Optimal power flow using the Jaya algorithm", *Energies*, **9**(9), p. 678 (2016).
35. Wang, S.H., Phillips, P., Dong, Z.C., et al. "Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya algorithm", *Neurocomputing*, **272**, pp. 668–676 (2018).
36. Wang, S., Rao, R.V., Chen, P., et al. "Abnormal breast detection in mammogram images by feed-forward neural network trained by Jaya algorithm", *Fundamenta Informaticae*, **151**(1–4), pp. 191–211 (2017).
37. Das, S.R., Mishra, D., and Rout, M. "A hybridized ELM-Jaya forecasting model for currency exchange prediction", *Journal of King Saud University-Computer and Information Sciences*, **32**(3), pp. 345–366 (2020).
38. Li, C., Nguyen, T.T., Yang, M., et al. "Multi-population methods in unconstrained continuous dynamic environments: The challenges", *Information Sciences*, **296**, pp. 95–118 (2015).
39. Branke, J., Kauβler, T., Smidt, C., et al. "A multi-population approach to dynamic optimization problems", In *Evolutionary Design and Manufacture*, pp. 299–307, Springer, London (2000).
40. Du, W. and Li, B. "Multi-strategy ensemble particle swarm optimization for dynamic optimization", *Information Sciences*, **178**(15), pp. 3096–3109 (2008).
41. Yang, S. and Li, C. "A clustering particle swarm optimizer for locating and tracking multiple optima in dynamic environments", *IEEE Transactions on Evolutionary Computation*, **14**(6), pp. 959–974 (2010).
42. Rao, R. and Patel, V. "An elitist teaching-learning-based optimization algorithm for solving complex constrained optimization problems", *International Journal of Industrial Engineering Computations*, **3**(4), pp. 535–560 (2012).
43. Rao, R.V. and Saroj, A. "An elitism-based self-adaptive multi-population Jaya algorithm and its applications", *Soft Computing*, **23**(12), pp. 4383–4406 (2019).
44. Venkata Rao, R., Saroj, A., and Bhattacharyya, S. "Design optimization of heat pipes using elitism-based self-adaptive multipopulation Jaya algorithm", *Journal of Thermophysics and Heat Transfer*, **32**(3), pp. 702–712 (2018).
45. https://archive.ics.uci.edu/ml/index.php
46. Dehuri, S., Roy, R., Cho, S.B., et al. "An improved swarm optimized functional link artificial neural network (ISO-FLANN) for classification", *Journal of Systems and Software*, **85**(6), pp. 1333–1345 (2012).
47. Rao, R.V. and Kalyankar, V.D. "Parameters optimization of advanced machining processes using TLBO algorithm", *EPPM, Singapore*, **20**, pp. 21–31 (2011).
48. Aruna, S., Rajagopalan, S.P., and Nandakishore, L.V. "Knowledge based analysis of various statistical tools in detecting breast cancer", *Computer Science & Information Technology*, **2**(2011), pp. 37–45 (2011).
49. Ceylan, Z. and Pekel, E. "Comparison of multi-label classification methods for prediagnosis of cervical cancer", *Graphical Models*, **21**, p. 22 (2017).
50. Christobel, A. and Sivaprakasam, Y. "An empirical comparison of data mining classification methods", *International Journal of Computer Information Systems*, **3**(2), pp. 24–28 (2011).
51. Luukka, P. "Similarity classifier using similarity measure derived from Yu's norms in classification of medical data sets", *Computers in Biology and Medicine*, **37**(8), pp. 1133–1140 (2007).
52. Yang, W., Gou, X., Xu, T., et al. "May. cervical cancer risk prediction model and analysis of risk factors based on machine learning", In *Proceedings of the 2019 11th International Conference on Bioinformatics and Biomedical Technology*, pp. 50–54 (2019).
53. Lavanya, D. and Rani, D.K.U. "Analysis of feature selection with classification: Breast cancer datasets", *Indian Journal of Computer Science and Engineering (IJCSE)*, **2**(5), pp. 756–763 (2011).
54. Naseriparsa, M. and Kashani, M.M.R. "Combination of PCA with SMOTE resampling to boost the prediction rate in lung cancer dataset", arXiv preprint arXiv:1403.1949 (2014).
55. Nasution, M.Z.F., Sitompul, O.S., and Ramli, M. "PCA based feature reduction to improve the accuracy of decision tree c4.5 classification", In *Journal of Physics: Conference Series*, **978**(1), p. 012058 (2018).
56. Salama, G.I., Abdelhalim, M., and Zeid, M.A.E. "Breast cancer diagnosis on three different datasets using multi-classifiers", *Breast Cancer (WDBC)*, **32**(569), p. 2 (2012).
57. Alharbi, A. "An automated computer system based on genetic algorithm and fuzzy systems for lung cancer diagnosis", *International Journal of Nonlinear Sciences and Numerical Simulation*, **19**(6), pp. 583–594 (2018).
58. Prabadevi, B., Deepa, N., Krithika, L.B., et al. "Analysis of machine learning algorithms on cancer dataset", In *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), IEEE*, pp. 1–10 (2020).

Biographies

Prajna Paramita Debata received the MTech degree in Computer Science from Siksha 'O' Anusandhan University, India in 2012. She is currently pursuing her PhD degree at the Department of Computer Science at IIIT Bhubaneswar, India. She has more than ten years of teaching experience. She has published more than ten research articles in peer-reviewed journals such as Journal of King Saud University-Computer

and Information Sciences, Evolutionary Intelligence, Turkish Journal of Electrical Engineering Computer Sciences, and presented various papers in international conferences. Her main research interests are multi-objective optimization, evolutionary computing, medical data mining, and machine learning.

Puspanjali Mohapatra is an Assistant Professor at International Institute of Information Technology, Bhubaneswar. She has more than 18 years of teaching experience, and she holds a PhD in CSE from Siksha O Anusandhan University, India. She completed her MTech Computer Science from Utkal University, Bhubaneswar in 2002 and graduated in BE Electrical Engineering from IGIT Sarang, Odisha in 1999. In 2009, she did her MBA in Systems, IT & HR from BPUT Rourkela, India. She has published more than 35 papers in reputed journals, conferences, and book chapters. She is also associated with different educational and research societies such as IEEE, ACM, IAENG, and OITS. She is currently working on the following research areas: biomedical data mining, time series data analysis, machine learning, deep learning, and evolutionary computing.