

BIO-MOTIVATED NOVEL PREDICTION MODEL SHUFFLED FROG LEAPING ALGORITHM AND EXTREME LEARNING MACHINES (SFLA ELM) FOR BREAST CANCER PREDICTION AND CLASSIFICATION

¹P.LAURA JULIET, ²DR. P.R.TAMILSELVI

¹Assistant Professor of Computer Applications, Vellalar College for Women(Autonomous),Erode, India.

²Assistant Professor of Computer Science, Government Arts & Science College, Komarapalayam, India.

DOI: 10.47750/pnr.2022.13.S10.135

Abstract

Breast Cancer is driving reason for death among women's. As indicated by Cancer Report Breast cancer is appears to be continually expanding all over worldwide in past years and it's a most ghastly sickness for women's. Clinical data mining manages classification of data as well as with the prediction of clinical data. In this stage, a bio-motivated novel prediction model called SFLAELM has been proposed by consolidating the Shuffled Frog Leaping Algorithm (SFLA) and Extreme Learning Machines (ELM), which is first of its sort in the investigation of non picture breast cancer data examination. In this stage, the goal was to foresee the hour of repeat and the outcome got for this case is discovered to be 93:75% exact with a RMSE of 0:30. In both, the cases tanh work performed better.

KEYWORDS: Breast Cancer, SFLA algorithm, Extreme Learning Machines, Wisconsin Breast Cancer Prognostic, Root Mean Square Error.

1. INTRODUCTION

Breast cancer begins in the cells of the breast. A cancerous tumour is a gathering of cancer cells that can develop into and annihilate close by tissue. It can likewise spread to different pieces of the body. Cells in the breast some of the time change and at this point don't develop or carry on regularly. These progressions may prompt non-cancerous breast conditions like abnormal hyperplasia and cysts. They can likewise prompt non-cancerous tumors, for example, intraductal papillomas. However, sometimes, changes to breast cells can cause breast cancer. Frequently, breast cancer begins in cells that line the conduits, which are the cylinders that convey milk from the glands to the nipple. This sort of breast cancer is called ductal carcinoma. Cancer can likewise begin in the cells of the lobules, which are the gatherings of glands that make milk. This kind of cancer is called lobular carcinoma. Both ductal carcinoma and lobular carcinoma can be in situ, which implies that the cancer is still where it began and has not developed into encompassing tissues. They can likewise be intrusive, which implies they have developed into encompassing tissues.

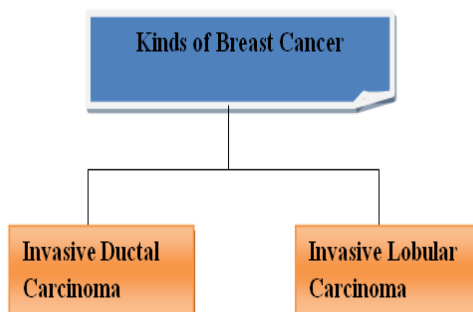


Figure 1.Major Kinds of Breast Cancer

1.1 Medical Data Mining

Health associations today are equipped for creating and gathering a lot of data. This expansion in data volume naturally requires the data to be recovered when required. With the utilization of data mining techniques is feasible to remove the knowledge and decide fascinating and helpful examples. The knowledge acquired in this manner can be utilized in the legitimate request to improve work proficiency and upgrade the nature of dynamic. Knowledge Discovery (KDD) is an interaction that permits programmed examining of high volume data to discover helpful examples that can be viewed as knowledge about the data. Data mining errands can be separated into descriptive and predictive.

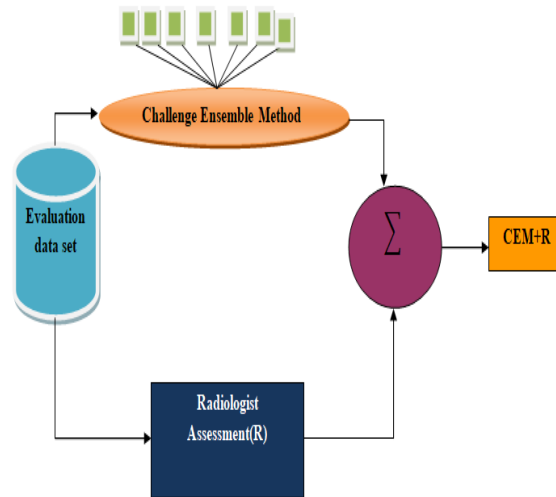


Figure 2. Medical Data Mining

The regular classification problems depend on the plan of classifiers chipping away at the sort of structural parameters chose. For a fuzzy classifier, the guidelines, forerunner, subsequent, and so on, goes about as the structural parameters. On account of K-Nearest Neighbor (KNN) classifier, distance metric goes about as a structural parameter. While, for an Artificial Neural Network (ANN), the quantity of stowed away layers, weights and predispositions fill in as the structural parameters. Tweaking of the parameters requires aptitude. Then again, meta-heuristic algorithms work with the planning of non-parametric classifiers, which can straightforwardly characterize the data dependent on activity taken by choice function otherwise called a cost function, or sometimes it depends on rules and conditions. The meta-heuristic algorithms give an optimal solution even in a perplexing inquiry space. Aside from this, they are likewise equipped for getting away from the issue of local minima or maxima (Yang, 2010a). These two qualities of meta-heuristic algorithms make them fit for delivering highly exact and hearty solutions in the briefest time.

In this paper, the idea of meta-heuristics Shuffled Frog Leaping Algorithm (SFLA) and advanced neural network called Extreme Learning Machines for the forecast of breast cancer data. For this reason, a novel SFLAELM model is proposed which is a mix of Shuffled Frog Leaping Algorithm (SFLA) and simple Extreme Learning Machines (ELM).

This paper presents as Shuffled Frog Leaping Algorithm (SFLA) for include determination issues utilizing standard microarray cancer informational indexes. This addresses a novel methodology, which will diminish the arrangement of accessible features. Each frog is addressing a subset of genes. The number of inhabitants in frogs is divided into subsets called image lexes. During the process of advancement, in each memplex most noticeably terrible frogs jump to the best frogs in the meme lexes. For jump most exceedingly awful frog to the best frog, low positioned genes will be eliminates and high positioned genes will be added into the most noticeably terrible frog. The local updating of frog makes the unessential features less attractive as well as assists Frogs with choosing applicable features. In addition, in light of the Shuffled of frogs at each stage, frogs select that subset of features, which has never recently been investigated.

2. Literature Survey

Kapil and Chhavi proposed suggested an IWDT methodology for BC expectation. At the preliminary stage, the element positioning was embraced on singular attributes using a χ^2 test. Those positioned highlights were by and large processed by Cost and Entropy measures to preset the hub level principles. Ultimately, the DT was utilized on those related element loads to inspect the accessibility of BC. This methodology encompasses disadvantages like i) χ^2 test was muddled to achieve the accurate highlights and ii) devours time. In another paper has developed a compelling LFOA-SVM methods which can well diagnose the breast disease in clinical diagnosis and furnish specialists with the significant clinical choice. In this examination, the improved FOA method, LFOA, was utilized to upgrade the two key parameters pair including punishment factor and width of the portion work in the SVM method and acquire the optimal model (LFOA-SVM). Moreover, this model will be explored to diagnose breast malignant growth on an undeniable level highlights dataset. Apparently, this paper is quick to tackle the boundary optimization problem of SVM with LFOA.

S.Kharya dealt with breast cancer expectation and expressed that artificial neural networks are broadly utilized. The paper highlighted about the benefits and weaknesses of utilizing machine learning methods like SVM, Naive Bayes, Neural network and Decision trees. Naresh Khuriwal, Nidhi Mishra took information from Wisconsin Breast Cancer database and chipped away at breast cancer diagnosis. The consequences of their analyses demonstrated that ANN and Logistic Algorithm worked better and gave a decent arrangement. It accomplished an exactness of 98.50%.

Breast Cancer Prediction Using Genetic Algorithm Based Ensemble Approach composed by Pragya Chauhan and Amit Swami proposed a framework where they found that Breast cancer prediction is an open zone of exploration. In this paper diverse machine learning algorithms are utilized for identification of Breast Cancer Prediction. Decision tree, random forest, support vector machine, neural network, linear model, adabost, Naive Bayes methods are utilized for prediction. A troupe method is utilized to build the prediction precision of breast cancer. New procedure is executed which is GA based weighted average outfit method of characterization dataset which over-came the restrictions of the old style weighted average method. Genetic algorithm based weighted average method is utilized for the prediction of various models.

Dongdong Sun et al. have proposed a deep learning (DL) method named D-SVM for the expectation of human breast cancer prognosis. The calculation viably took in various levelled and dynamic portrayal from crude information and effectively incorporated conventional classification method.

D. Selvathi et al. have proposed an automated system for accomplishing blunder free identification of breast cancer utilizing a Sparse Auto encoder (SAE) which takes in include portrayals from the mammogram and a classifier which is fell with the SAE plays out the order dependent on these learned highlights.

3. Proposed Methodology

In this paper, a bio-inspired novel prediction model called SFLAELM has been proposed by joining the Shuffled Frog Leaping Algorithm (SFLA) and Extreme Learning Machines (ELM), which is first of its sort in the investigation of non picture breast cancer data analysis.

Proposed novel SFLAELM model (SFLAELM)

The proposed novel SFLAELM model is a novel model which is a mix of Shuffled Frog Leaping Algorithm (SFLA) and Extreme Learning Machines (ELM). The idea of SFLA and ELM which has numerous benefits when contrasted with the current algorithms of their classification has inspired the creators to assemble a novel model that can anticipate the clinical data with high exactness and negligible blunder. Here, SFLA algorithm is utilized to improve the boundaries of ELM with the goal that the prediction task is done productively. The primary point of ELM is to foresee the data with least mistake. For achieving a negligible blunder, benchmark Wisconsin Breast Cancer Prognostic (WBCP) data set has been tried upon three distinctive learning capacities (sigmoid, sine and tanh) and the capacity which delivers the best outcome has been considered as the last. This eliminate has been conveyed to help novel SFLAELM model.

SFLA algorithm

The Shuffled Frog Leaping Algorithm (SFLA) is perhaps the most imaginative streamlining algorithms propelled by the social behavior of frogs in nature, and regarding classification, among behavioral algorithms or Mimetic Algorithms are incorporated. Different names for the frog bounce enhancement algorithms incorporate the frog hop algorithm, frog leaping algorithm, and SFLA algorithm. This algorithm was first proposed by Eusuff and Lansey in 2003, albeit numerous papers have since been submitted to improve this algorithm. At the lower part of the download, the interface is the fundamental article on the frog optimization algorithm.

This method consolidates the advantages of two categories of genetics-based algorithms, (for example, meme tics) and social conduct based algorithms, (for example, the PSO bird calculation). It tries to find some kind of harmony between broad scrutinizes in the space of potential answers. In this populace calculation, a populace of frogs (answers) comprises, each frog will have a chromosome-like structure in the hereditary calculation. The entire populace of frogs is isolated into more modest groups, each group addressing various sorts of frogs that are scattered in better places of the answer space. Each group of frogs at that point starts a precise neighborhood search around their environment.

Each frog in each category is influenced by other members of its group as well as other groups. After a few steps, the mixing takes place, and the information is spread among all the groups to establish the condition of convergence and reaching the answer. How to find the best solution in this algorithm consists of two stages of global and local search.

Shuffled Frog Leaping Algorithm (SFLA)

The SFLA is a mimetic meta-heuristic technique that is gotten from a virtual population of frogs wherein each frog addresses a bunch of doable solutions. Each frog is into a subset of the population saw as meme lexes. A neighborhood search is performed in each memplex. To guarantee worldwide exploration, rearranged information trade will happen between meme lexes after a characterized number of advancement steps.

In this paper, have sfla_p frog. Each frog is representing a subset of genes. Each frog has its own most extreme length. The population of frogs is partitioned into subsets called image partitioned. Sfla_m is the quantity of meme lexes. Thusly, there are sfla_nfrogs in each memplex. The distinctive image lexes are considered as various societies of frogs, each performing a neighborhood search. Inside each memplex, there is a submemplex. In each submemplex there are sfla_q frogs arbitrarily chose by the accompanying probability function.

$$P_j = \frac{2(sfl_n+1-j)}{sfla_n(sfla_n+1)}, j = 1, 2, \dots, sfla_n \quad (1)$$

Submemplex causes the algorithm stalls out in a local Optimum.

Where P_j is the probability of selecting j -th frog after a number of mimetic evolution steps, genes are passed among frogs of meme lexes in a shuffling process. The local search and the shuffling processes continue until some convergence criteria are satisfied iteration, within each submemplex of memplexes, the frog with the best fitness and the frog with the worst fitness are identified as P_b and P_w , respectively. The frog with the global best fitness is identified as P_g each iteration only the worst fitness frog will be modified. Therefore, the position of the frog with the worst fitness is adjusted as:

$$\begin{cases} \min \{ \text{int}(\text{rand} \cdot [P_g - P_w]), S_{max} & \text{for positive leap} \\ \max \{ \text{int}(\text{rand} \cdot [P_g - P_w]), -S_{max} & \text{for negative leap} \end{cases} \quad (2)$$
$$P'_w = P_w + S_B \quad (3)$$

Where S_{max} The maximum length took into consideration the leap Note for positive leaping in size S_B qualities which have the most elevated position in the t-test filter method will be added to the worst frog. Additionally for negative leaping in size S_B qualities that have a low position in the t-test filter method will be killed from the worst frog. In the event that the new frog (P'_w) is superior to the worst frog (P_w) it will be supplanted by the worst frog. Something else, the position of the worst frog is altered by the position of the frog with the global best fitness as:

$$S_G = \begin{cases} \min \{ \text{int}(\text{rand} \cdot [P_g - P_w]), S_{max} & \text{for positive leap} \\ \max \{ \text{int}(\text{rand} \cdot [P_g - P_w]), -S_{max} & \text{for negative leap} \end{cases} \quad (4)$$

$$P'_w = P_w + S_G \quad (5)$$

Nota for positive leaping in size SG genes which have the most elevated position in the t-test filter method will be added to the most noticeably terrible frog. Likewise for the negative leaping in size SG genes that have a low position in the t-test filter method will be wiped out from the most exceedingly awful frog. A similar state before the new frog (P'_w) was better than the most exceedingly terrible frog (P_w) it will supplant the most exceedingly terrible frog. On the off chance that no improvement gets possible for this situation an arbitrary frog is produced which replaces the most noticeably awful frog in submemeplex. These means are rehashed a few times (ITmem), in another word again all frog rearranging together and again be separated into sfla_m memeplex. This activity will continue until the termination conditions are fulfilled.

The pseudo-code of SFLA is given in Table (1). In view of this algorithm, the most noticeably terrible frog can jump to a superior frog. By rehashing this activity, the mean wellness of the population increases in the mimetic development steps. The best arrangement found during the inquiry interaction can be considered as the output of the algorithm. Truth is told, during the interaction of advancement, the most noticeably awful frogs leap to the best frogs in the memeplex or best frog in the population. The frog will be refreshed utilizing first-stage data and low ranked qualities will be taken out and high ranked qualities will be added into the most noticeably terrible frog. In view of the development of data between frogs, the likelihood of tracking down the best subset of quality increases.

PSEUDO-CODE OF SFLA

1. rankgenebase1-testfiltermethod
2. Create an initial population of SFLA_P frog's generated randomly.
3. Divide the frogs inform afa_m memplexeseach holdings sfla_n frogs.
 - 3.2 i=0
 - 3.3.wtllleI < ITmem
 - 3.3.J.createasubmemeplexforeachmemep/ex
 - 3.3.2.the position of the worst frog P_w' for the memplex is adjusted such as (3)
 - 3.3... 3.if (fitness (P_w)< fitness (P11))
 - position of the worst frog P_w' for the memplex is adjusted such as (5)
 - 3.3.4. If {fitness (P_w) < fitness (P_w)}
 - ndom frog is generated which replaces the worst frog.
 - 3.3.5. otherwise $P_w = P_w$
 - 3.3.6. otherwise $P_w = P_w$
 - 3.3.7.i=i+1
 - 3.4. Frog shuffling together
4. Check the convergence. If the convergence criteria are satisfied stop, otherwise return to the step3.
5. finish

ELM

The extreme learning machine (ELM) is broadly utilized in cluster learning, consecutive learning, and gradual learning due to its quick and proficient learning speed, quick combination, great speculation capacity, and simplicity of execution. ELM depends on experimental risk minimization theory and its learning cycle needs just a solitary emphasis. The calculation stays away from numerous cycles and neighborhood minimization. It has been utilized in different fields and applications on account of better speculation capacity, vigor, and controllability and quick learning rate.

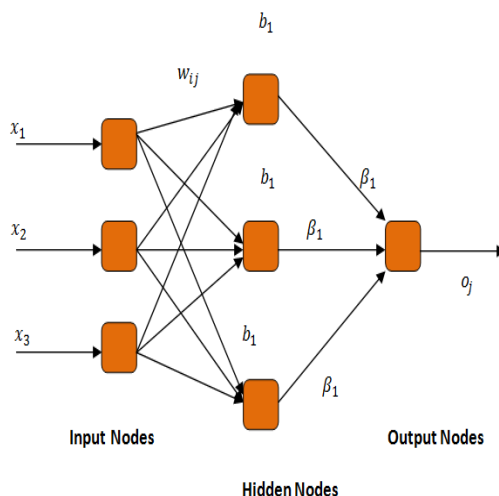


Figure 7.3 General structure of ELM

Consider a preparation set of data with the N number of tests. Considering an illustration of a single output node, the ELM with L secret nodes can be addressed mathematically as in Equation (6).

$$f_L(x) = \sum_{i=1}^L \beta_i h_i(x) = h(x)\beta \quad (6)$$

Here, x indicates the input; β is the yield weight vector and is addressed as in Equation (7).

$$\beta = [\beta_1, \beta_2, \dots, \beta_L]^T \quad (7)$$

$h(x)$ is the hidden layer output vector addressed as in Equation (8).

$$h(x) = [h_1(x), h_2(x), \dots, h_L(x)] \quad (8)$$

The main aim of ELM is to minimize $\|H\beta - O\|^2$ and $\|\beta\|$ where represented, o is the output vector also called as target vector and H is the hidden layer output matrix as in equation (9) and (10).

$$\beta = [t_1, t_2, \dots, t_L]^T \quad (9)$$

$$H = \begin{bmatrix} h(x_1) \\ h(x_2) \\ \vdots \\ h(x_N) \end{bmatrix} \quad (10)$$

The essential implementation of ELM utilizes the minimal standard least square method addressed as in Equation (11).

$$\beta' = H * O = H^T (HH^T)^{-1} O \quad (11)$$

Here, H^* is the special form of inverse of H called as Moore-Penrose generalized inverse.

The equation (6.6) can also be alternatively represented as Equation (12).

$$\beta' = H^T \left(\frac{1}{c} + HH^T \right)^{-1} \mathbf{O} \quad (12)$$

By using equation (7), the output function in equation (1) can be rewritten as Equation (13)

$$f_L(x) = h(x)\beta = H(x) H^T \left(\frac{1}{c} + HH^T \right)^{-1} \mathbf{O} \quad (13)$$

ELM makes use of Kernels function to produce fine output thus, the final output Function ELM can be represented as in equation (14)

$$f_L(x) = [k(x, x_1), k(x, x_2) \dots k(x, x_N)] \left(\frac{1}{c} + HH^T \right)^{-1} \mathbf{O} \quad (14)$$

By default Gaussian function is used as kernel function and is represented as in Equation (15)

$$K_{gau}(x, x_i) = \exp\left(\frac{-\|x-x_i\|^2}{\gamma}\right) \quad (15)$$

Here, γ is the local Gaussian kernel parameter.

$$\gamma = \sigma^2 \quad (16)$$

And σ is the functionality of the hidden layer containing output weight and biases and is represented as in Equation (17)

$$\sigma = \sum_1^N \omega_i x_i + b_i \quad (17)$$

Pseudo code for ELM

Given a training set of N instances, with an activation function $g(x)$, and n hidden neurons

Step - 1: Initiate arbitrary input weights w_i and biases b_i , where $i = 1$ to n .

Step - 2: Calculate hidden layer output matrix H.

Step - 3: Calculate the output weight matrix.

Step - 4: Calculate the output matrix.

Proposed Algorithm:

The proposed novel SFLAELM model is quite simple and works in four simple steps as follows:

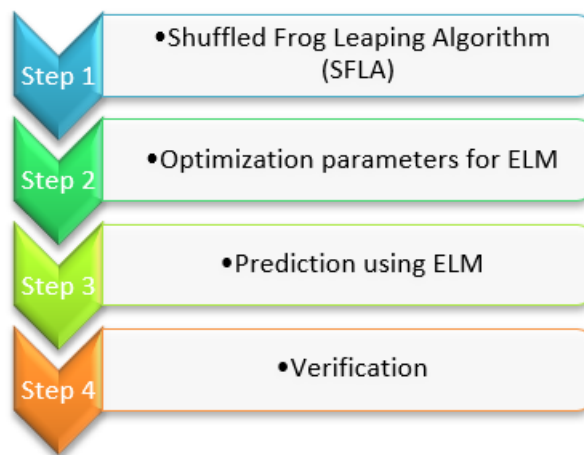


Figure 4. Proposed Algorithm SFLAELM model steps

Step 1: Selection of appropriate information is the first and most significant advance in data mining and machine learning task. The yield of any undertaking relies on the subjective info gave. Rather than going for experimentation based element choice or arbitrary choice, connection investigation is picked to choose the element. Since the principle point is to foresee the breast cancer data, Pearson connection coefficient approach is followed which can distinguish how solid or how feeble the objective variable is related with the supporting variable. Coefficients esteem lies between - 1 to +1. Negative worth shows negative relationship, positive qualities demonstrate positive connection, and zero methods no relationship. On the off chance that the coefficient esteem is 1 or close to 1, they are firmly related. In the event that the co-proficient worth is under 0.5, they are viewed as pitifully related. The Pearson Correlation coefficient condition is given by Equation (18)

$$P(X, Y) = \frac{\text{cov}(X,Y)}{\sigma(X)\sigma(Y)} \quad (18)$$

Here, $P(X, Y)$ is the Pearson coefficient, $\text{Cov}(X, Y)$ is the covariance of X with respect to Y , $\sigma(X)$ and $\sigma(Y)$ are the standard deviation of X and Y . Where X is the predicting variable and Y is the supporting variable.

Step 2: Optimization of parameters for ELM: Shuffled Frog Leaping Algorithm (SFLA) is utilized to advance the parameters of ELM, for example, regression coefficient C and kernel parameter λ . The intrinsic limit of Bat algorithm helps in discovering the global optimum, along these lines, maintaining a strategic distance from the model to get caught into local minima.

Step 3: Prediction using ELM: Subsequent to streamlining the parameters, right arrangement of sources of info, loads, and inclinations, threesome is taken care of to the ELM to anticipate the information.

Step 4: Verification: In each preliminary, the yield is confirmed by contrasting it and the threshold. On the off chance that Root Mean Square Error (RMSE) is not exactly the threshold, the yield esteems are viewed as substantial, else they are disposed of and the model is diverted to stage 2 where the streamlining happens once more. The RMSE is considered as the wellness capacity of SFLA. Lower the wellness esteem more dependable is the model. The wellness capacity of SFLA is given by Equation (19).

$$\text{fitness} = \sqrt{\sum_{i=1}^N (y_i^2 - y_i)^2} \quad (19)$$

The procedure of working of the proposed novel SFLAELM model is clearly represented using a flow chart as shown in Figure 7.5.



Figure 5. Flow Chart of Proposed Novel SFLAELM Model

4. Experiment Result

- True Positives: Observations where the genuine and anticipated Class were immaterial
- True Negatives: Observations where the genuine and anticipated class were immaterial
- False Positives: Observations where the true class were immaterial however anticipated to be pertinent
- False Negatives: Observations where the genuine class were immaterial however weren't anticipated to be pertinent

Table 1. Experiment result definition

Parameters	Definition
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Sensitivity	$\frac{TP}{TP + FN}$
Specificity	$\frac{TN}{TN + FP}$

Precision	$\frac{TP}{TP + FP}$
------------------	----------------------

4.1 Accuracy

Table 2. Comparison Table of Accuracy

No of Data	LFOA-SVM	IWDT	Proposed SFLAELM
100	73	65	79
200	89	85	93
300	85	81	91
400	82	76	89
500	72	69	80

The Comparison table 7.2 of Accuracy values explains the different values of existing algorithms (ANN, GA) and proposed SFLAELM. While comparing the Existing algorithm (ANN, GA) and proposed SFLAELM, provides the better results. The existing algorithm values start from 73 to 89, 65 to 85 and proposed SFLAELM values starts from 79 to 93.

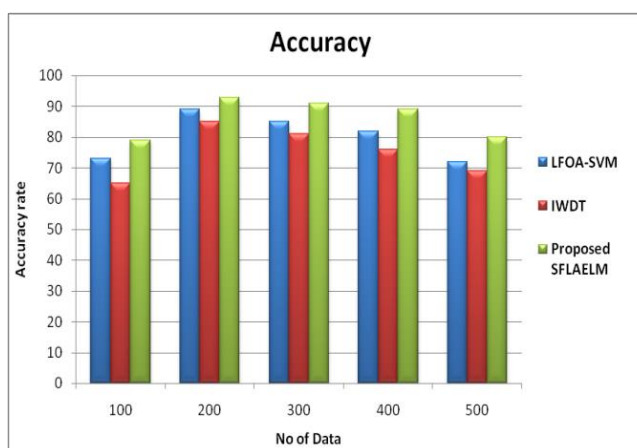


Figure 6. Comparison Chart of Accuracy

The Figure 6 Shows the comparison chart of Accuracy demonstrates the existing1, existing 2 (LFOA-SVM, IWDT) and proposed SFLAELM values. X axis denote the validation and y axis denotes the performance values in Accuracy. The proposed SFLAELM values are better than the existing algorithm. The existing algorithm values start from 73 to 89, 65 to 85 and proposed SFLAELM values starts from 79 to 93.

4.2 Sensitivity

Table 3. Comparison Table of Sensitivity

No of Data	LFOA-SVM	IWDT	Proposed SFLAELM
100	83.48	81.22	86.87
200	84.74	83.52	88.74
300	88.21	84.01	90.55

400	90.48	87.35	92.46
500	93.66	90.65	96.91

The Comparison table 3 of Sensitivity values explains the different values of existing algorithms (LFOA-SVM, IWDT) and proposed SFLAELM. While comparing the Existing algorithm (LFOA-SVM, IWDT) and proposed SFLAELM, provides the better results. The existing algorithm values start from 83.48 to 93.66, 81.22 to 90.65 and proposed SFLAELM values starts from 86.87 to 96.91.

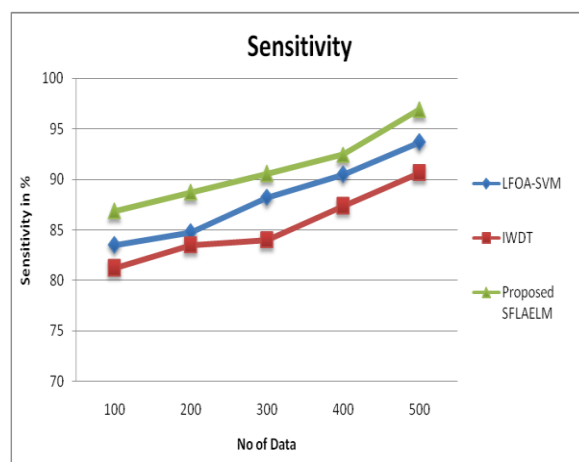


Figure 7. Comparison Chart of Sensitivity

The Figure 7 Shows the comparison chart of Sensitivity demonstrates the existing1, existing 2 (LFOA-SVM, IWDT) and proposed SFLAELM values. X axis denote the validation and y axis denotes the performance values in sensitivity. The proposed SFLAELM values are better than the existing algorithm. The existing algorithm values start from 83.48 to 93.66, 81.22 to 90.65 and proposed SFLAELM values starts from 86.87 to 96.91.

4.3 Specificity

Table 4. Comparison Table of Specificity

No of Data	LFOA-SVM	IWDT	Proposed SFLAELM
100	83.154	83.121	84.745
200	85.649	84.024	87.999
300	85.425	85.021	86.452
400	90.267	90.135	91.406
500	92.623	92.365	93.91

The Comparison table 4 of Specificity values explains the different values of existing algorithms (LFOA-SVM, IWDT) and proposed SFLAELM. While comparing the Existing algorithm (LFOA-SVM, IWDT) and proposed SFLAELM, provides the better results. The existing algorithm values start from 83.154 to 92.623, 83.121 to 92.365 and proposed SFLAELM values starts from 84.745 to 93.91.

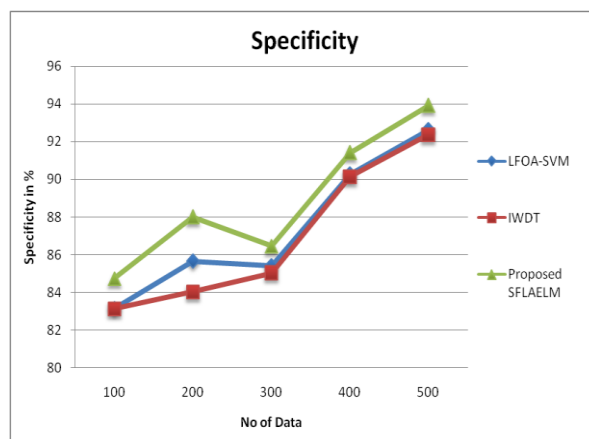


Figure 8. Comparison Chart of Sensitivity

The Figure 8 Shows the comparison chart of Sensitivity demonstrates the existing1, existing 2 (LFOA-SVM, IWDT) and proposed SFLAELM values. X axis denote the validation and y axis denotes the performance values in sensitivity value. The proposed SFLAELM values are better than the existing algorithm. The existing algorithm values start from 83.154 to 92.623, 83.121 to 92.365 and proposed SFLAELM values starts from 84.745 to 93.91.

4.4 Precision

Table 5. Comparison Table of Precision

No of Data	LFOA-SVM	IWDT	Proposed SFLAELM
100	70.74	68.02	72.26
200	73.09	70.26	75.23
300	72.89	71.18	75.02
400	75.04	72.11	77.14
500	76.85	74.22	78.99

The Comparison table 5 of Precision values explains the different values of existing algorithms (ANN, GA) and proposed SFLAELM. While comparing the Existing algorithm (ANN, GA) and proposed SFLAELM, provides the better results. The existing algorithm values start from 0.221 to 0.601, 0.129 to 0.511 and proposed SFLAELM values starts from 0.399 to 0.996.

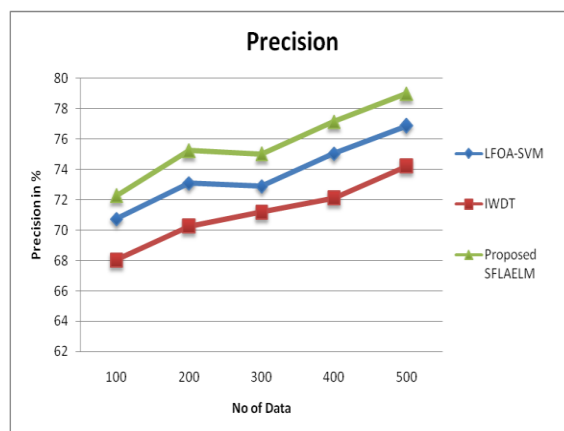


Figure 9. Comparison Chart of Precision

The Figure 9 Shows the comparison chart of Precision demonstrates the existing1, existing 2 (ANN, GA) and proposed SFLAELM values. X axis denote the validation and y axis denotes the performance values in Precision. The proposed SFLAELM values are better than the existing algorithm. The existing algorithm values start from 0.221 to 0.601, 0.129 to 0.511 and proposed SFLAELM values starts from 0.399 to 0.996.

5. Conclusion

Breast cancer whenever found at a beginning paper will help save lives of thousands of ladies or even men. These ventures help this present reality patients and specialists to accumulate as much data as possible. This stage uses the idea of meta-heuristics Shuffled Frog Leaping Algorithm (SFLA) and advanced neural network called Extreme Learning Machines for the forecast of breast cancer information. For this reason, a half and half SFLAELM model is proposed which is a mix of SFLA Algorithm (SFLA) and simple Extreme Learning Machines (ELM). The SFLA is utilized to deliver arbitrary loads and predispositions which are taken care of to the ELM where the forecast task is done. The strength of this expectation model lies in its novelization. The blend of SFLA and ELM transforms into an incredible prescient model delivering higher correctness. The proposed method has achieved a classification accuracy of 93.83%, sensitivity of 91.22%, specificity of 96.53% and MCC of 0.8799 for breast cancer diagnosis based on the high-level features.

REFERENCES

1. Msph LAT, Bray F, Siegel RL, Jacques Ferlay ME, Lortet-Tieulent J, PhD AJD. Global cancer statistics, 2012. *CA Cancer J Clin.* 2015;65(2):69–90.
2. Li Q, Chen H, Huang H, Zhao X, Cai Z, Tong C, Liu W, Tian X. An enhanced Grey wolf optimization based feature selection wrapped kernel extreme learning machine for medical diagnosis. *Comput Math Methods Med.* 2017; 2017:9512741.
3. Ma C, Ouyang J, Chen HL, Zhao XH. An efficient diagnosis system for Parkinson's disease using kernel-based extreme learning machine with subtractive clustering features weighting approach. *Comput Math Methods Med.* 2014;2014(3):985789.
4. Chen H-L, Yang B, Liu J, Liu D-Y. A support vector machine classifier with rough set-based feature selection for breast cancer diagnosis. *Expert Syst Appl.* 2011;38(7):9014–22.
5. Wang M, Chen H, Yang B, Zhao X, Hu L, Cai Z, Huang H, Tong C. Toward an optimal kernel extreme learning machine using a chaotic moth-flame optimization strategy with applications in medical diagnoses. *Neurocomputing.* 2017;267(Supplement C):69–84.
6. Zhao X, Zhang X, Cai Z, Tian X, Wang X, Huang Y, Chen H, Hu L. Chaos enhanced grey wolf optimization wrapped ELM for diagnosis of paraquatpoisoned patients. *Comput Biol Chem.* 2018. <https://doi.org/10.1016/j.compbiolchem.2018.11.017>.
7. Zhu J, Zhao X, Li H, Chen H, Wu G. An effective machine learning approach for identifying the glyphosate poisoning status in rats using blood routine test. *IEEE Access.* 2018;6:15653–62.
8. Zhu J, Zhu F, Huang S, Chen H, Zhao X, Zhang S. A new evolutionary machine learning approach to identify the pyrene induced rat hepatotoxicity and renal dysfunction. *IEEE Access.* 2018. <https://doi.org/10.1109/ACCESS.2018.2889151>.
9. Xu J, Zhang X, Chen H, Li J, Zhang J, Shao L, Wang G. Automatic analysis of microaneurysms turnover to diagnose the progression of diabetic retinopathy. *IEEE Access.* 2018; 6:9632–42.
10. Wang X, Wang Z, Weng J, Wen C, Chen H, Wang X. A new effective machine learning framework for Sepsis diagnosis. *IEEE Access.* 2018; 6:48300–10.
11. Cai Z, Gu J, Wen C, Zhao D, Huang C, Huang H, Tong C, Li J, Chen H. An intelligent Parkinsons' disease diagnostic system based on a chaotic bacterial foraging optimization enhanced fuzzy KNN approach. *Comput Math Methods Med.* 2018;2018:24.
12. Maglogiannis I, Zafiropoulos E, Anagnostopoulos I. An intelligent system for automated breast cancer diagnosis and prognosis using SVM based classifiers. *Appl Intell.* 2009;30(1):24–36.
13. Kaya Y. A new intelligent classifier for breast cancer diagnosis based on rough set and extreme learning machine: RS+ELM. *Turk J Electr Eng Comput*

- Sci. 2014;21(Sup.1):2079–91.
14. Akay MF. Support vector machines combined with feature selection for breast cancer diagnosis. *Expert Syst Appl.* 2009;36(2):3240–7.
 15. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: a review. *IEEE Rev Biomed Eng.* 2009;2:147–71.
 16. Kuse M, Sharma T, Gupta S. A classification scheme for lymphocyte segmentation in H&E stained histology images. Berlin Heidelberg: Springer; 2010.
 17. Dundar MM, Badve S, Bilgin G, Raykar V, Jain R, Sertel O, Gurcan MN. Computerized classification of Intraductal breast lesions using histopathological images. *IEEE Trans Biomed Eng.* 2011; 58(7):1977–84.
 18. Sparks R, Madabhushi A. Content-based image retrieval utilizing explicit shape descriptors: applications to breast MRI and prostate histopathology. *Proc SPIE.* 2011; 7962(8):765–8.