# Feature Optimization using Teaching Learning Based Optimization for Breast Disease Diagnosis

**Mohan Allam, M. Nandhini**

**Abstract**: *Disease diagnosis is a significant challenge in the field of medical science because most of the medical datasets contain irrelevant and redundant attributes which are not mandatory to obtain an accurate estimate of the disease diagnosis. In this work, we have used Teaching Learning Based Optimization (TLBO) algorithm for feature Optimization in automatic breast disease diagnosis. We have used a naive Bayes classifier for finding the fitness of individual and Multilayer Perceptron (MLP), J48, random forest, logistic regression algorithms for estimating the effectiveness of the proposed system. The results confirmed that the expected scheme produced higher accuracy on Wisconsin diagnosis breast cancer (WDBC) data set to classify the malignant and benign tumors. In short, the proposed TLBO variant presents an efficient technique to optimize the features for sustaining data-based decision making systems.*

*Index Terms*: *Feature Optimization, Teaching Learning based Optimization, Breast Cancer.*

## I. INTRODUCTION

Breast cancer is the most commonly spotted disease and the primary reason for cancer death in women around the world. 25% of all new cancer diagnoses as breast cancer in women as stated by American Cancer Society (ACS). One woman every minute and more than 1400 woman every day die from breast cancer. A tumor is an accumulation of unusual tissue. Breast tumors are grouped into cancerous and non-cancerous, which are called as benign and malignant respectively. Improvement in the availability of facilities and advancement in the testing modalities caused a large data to accumulate in the medical data centers. The diagnosis may be hard even for a medical specialist because of vagueness in the data. Improvements in early detection and effective treatment can significantly decline breast cancer deaths and improve the probability of long-term endurance. Automation of diagnosis is one of the best solutions for analyzing medical data records in a faster way in case of fatal diseases like cancer. Shuihua Wang et al. [1] implemented a novel CAD system for identifying abnormal breasts in mammogram imagery. They used principal component analysis (PCA) to reduce the spectrum, feed-forward neural network (FNN) to spawn the classifier, Jaya algorithm to train the classifier and achieved an accuracy of 92.27%. Jaya algorithm [39] is more efficient in training FNN than others. Sasikala et al. [2] gave a Shapely Value Embedded Genetic Algorithm to improve the breast disease identification accuracy by selecting the best subset of features. The classification precision is checked with four classifiers namely, Support vector machine (SVM), K-Nearest Neighbor (KNN), Naïve Bayes (NB), and J48 on the Kent ridge biomedical database.

Different Data mining techniques are used for getting improved results from those datasets which can facilitate doctors in taking better decisions regarding time and accuracy. Emina et al. [3] presented a data mining technique for analysis of breast cancer. GA is used for extraction of essential attributes and performance is evaluated using the Rotation Forest model and achieved 99.48 % classification accuracy on Wisconsin Breast Cancer dataset. Hasan et al. [4] developed a feature selection algorithm based on PCA with the help of Artificial Neural Network as a classifier to make a distinction between benign and malignant tumors. Noel et al. [5] developed a model for breast cancer classification in segmented mammographic lesions using Correlate-based Feature Selector (CBF) and classified with Feed Forward Back Propagation Neural Network (FFBP), SVM, J48 models for a relative performance assessment.In addition to that, the massive dimension of data increases the computation complexity and diminishes the performance of the classifier. One of the best solutions tosolve this problemis feature optimization.Thefeatureoptimization techniques are typically intended to improve the effectiveness of classification algorithms by eliminating the irrelevant and redundant attributes. ShokoufehAalaei et al.

[6] addressed GA supported attribute selection method for breast cancer diagnosis with the help of Artificial Neural Network (ANN), GA, particle swarm optimization (PSO) algorithm supported classifier (PS-classifier) on Wisconsin diagnosis and prognosis breast cancer repository.

The main initiative of this paper is to employ the TLBO for optimizing the values of features in breast cancer diagnosis. The idea is arranged as follows. In the second section, we presented the details of the past work on Breast Cancer disease diagnosis using soft computing techniques. Section III provides the proposed scheme and following subsections discusses TLBO for feature optimization to attain better accuracy results. In Section IV, we discussed the performance of various classifiers by comparing the evaluated results for the diagnosis of breast cancer. Finally, we concluded the paper with the summary.

**Revised Manuscript Received on 30 September 2018.**
**\*** Correspondence Author
   **Mohan Allam\***, Research Scholar, Pondicherry University, Kalapet, Puducherry, India.
   **M. Nandhini**, Department of Computer Science, Pondicherry University, Kalapet, Puducherry, India.

## II. RELATED WORK

We have studied different optimization methods for selecting the best features to improve the performance of various classifiers. Li-Yeh Chuang et al. [7] developed catfish binary particle swarm optimization (CatfishBPSO) algorithm in which few features are selected by replacing new catfish particles into the search space and achieved high classification accuracy withthe k-NN algorithm. Zheng et al. [8] have proposed a hybrid K-means for pattern recognition and classified the WDBC dataset with 97.38 % accuracy using SVM (K-SVM) algorithm.

Sridevi et al. [9] developed a feature selection technique for medical analysis of Breast cancer and compared several classification algorithms. They reduced feature vectors 222 for both diagnosis and prognosis wisconsin data sets using rough sets and correlation techniques. Vartika Agrawal et al. [10] proposed a feature selection system for classification of cervical cancer CT images using Artificial Bee Colony algorithm and k-NN classifier. The authors in [11] mentioned about the importance of automatic medical disease diagnosis to handle the problems efficiently in the early stages and also discussed various imaging modalities for capturing the images, feature extraction methods for collecting the required attributes and feature selection techniques for necessary features like texture, and color. Zhi Chen et al. [12] proposed a coarse-grained parallel genetic algorithm (CGPGA) to optimize the features in the dataset and constraints for SVM. They also proposed a new fitness function which is composed of classification accuracy, number of selected features, and the number of support vectors to optimize generalization error. Saleh Shahbeig et al. [13] proposed a mutated fuzzy adaptive particle swarm optimization (PSO) in combination with teaching learning-based optimization (TLBO) algorithm to find the most relevant and smallest set of genes in breast cancer microarray data and achieved the accuracy of 91.88% with SVM classifier. We also studied different classifications algorithms and their efficiencies in different domains. Jung et al. [14] outlined a method to obtain additional numerical parameters from the breast cancer image dataanalysis using a neural network algorithm to find the best classification between benign and malignant. HtetThazin et al. [15] summarized a system using island-based model for training ANN to distinguishing between different classes of breast cancer with better precision and less training time on Wisconsin Diagnostic and Prognostic Breast Cancer. Paul R. Harper [16] addressed different classification techniques and compared their performances in terms of accuracy and computational time over health care data.

In this paper, we developed an optimization model using modified TLBO to determine a good number of significant features from the data set related to breast cancer. The goal is to accomplish the best classification accuracy with the given dataset. The classification performance is estimated based on the classification accuracy, kappa statistic, Precision, Recall, F-Measure, TP and FP Rates. In the proposed model, the important features of valuable information are optimized by using a metaheuristic method. Metaheuristics can find better solutions over a large set of possible solutions with the less computational attempt.

## III. MATERIAL AND METHODS

The most important job for designing an efficient decision-making model is preparing good features from the dataset and makes them more relevant to achieve best accuracy results by the classifier. A good algorithm is required to build an efficient dataset of with more helpful features. Our proposed method has two phases. In the initial phase, an optimization algorithm is used to make the features more reliable for classification. In the subsequent phase, different classification models are applied to calculate the performance of the intended system on WDBC dataset.

In the present paper, we have adopted a novel algorithm called teaching-learning based Optimization algorithm [17] for optimizing the values of the selected dataset. Rao et al. [18] addressed widespread applications of TLBO in various areas for optimization problems. Agrawal et al. [19] developed an iris recognition system in which feature extraction is performed by the Gabor wavelet transform on CASIA dataset. TLBO is applied to these features which generate a template for matching process to enhance the detection rate of iris system. The efficiency of the face detection system advanced with the development of different models [20-21] and medical image segmentation using TLBO [22]-[24].This paper mainly emphasis on the application of TLBO [40] in generating optimal features automatically. The proposed system is as shown in Fig. 1.

The proposed method consists of two modules which include feature optimization using TLBO and validation of the model using the test dataset.The primary steps in our method are as follows,

Step 1: Collect the dataset for diagnosis.

Step2:Optimization of selected features using improved TLBO.

Step 3: Training the classification models with the optimized dataset.

Step 4: Check the classification accuracy of models with the test dataset.

### 3.1. Feature Optimization using TLBO Algorithm.

Different techniques are used for feature optimization such as GA, Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO) and Neural Network by many researchers. A method has been used by panda [25] based on elephant search optimization in combination with the deep neural network for analyzing microarray data. TLBO is one of the Swarm Intelligence (SI) based algorithm which is accurate and efficient. We used TLBO to perform the job of feature optimization for disease diagnosis. TLBO mimics the procedure of educating the students in a classroom.
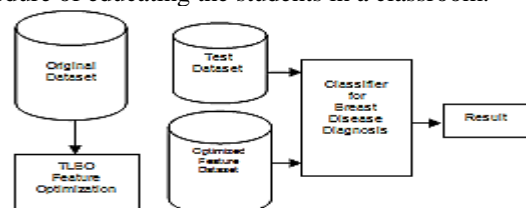


**Figure 1: Proposed Model for Feature Optimization using TLBO.**

First, teacher edifies the students with his knowledge, and later students are trained with the fellow students. TLBO does not demand any explicit factors and only use general factors like the number of instances in the population along with maximum iterations. The authors of papers [26]-[28] used TLBO for optimal feature selection in various applications. TLBO algorithm is based on the random searchmethod consisting of two stages namely, Teaching Phase and Learning Phase.

### 3.2. Modified TLBO:

Our contribution in this paper is the application of TLBO algorithm for binary classification of the breast disease. The two phases are repeated until a stopping criterion has met. In general, we can take stopping criteria as a fixed number of iterations. The Best learners are the best solutions in a particular iteration and process continuous with the number of generations.The improved process of TLBO for a single iteration is as follows,

### 3.3. Teacher Phase

The teacher phase for a given t number of instances (i.e., population, individuals, {i= 1, 2…t}) is as follows,

Step 1: Initialize a number of instances and separate the population into two groups (g) based on the classes (malignant and benign) in the dataset. Define the number of featuresas $X_{f,i,g}$.

Step 2: Compute the mean of every feature for both groups as $M_{f,m}$ and $M_{f,b}$

Step 3: Find out the finest individual in each group based on their fitness value using the (1).

$$\text{Fitness}(X_{f,i,g}) = \text{Accuracy}(X_{f,i,g}) \quad (1)$$

The fitness of each individual will be calculated with respect to every individual in the other group and select the best combination as teachers from each group. We have used a naivebayes classifier to find the accuracy of each instance and selected the best result with the highest fitness value.Usually, the teacher is considered as an extremely learned individual whoeducates learnersto improve theresults.

Step 4: Evaluate the difference mean for all the attributesby means of thefinest individual as shown in (2).

$$\text{Diff\_Mean}_{f,i,g} = r(X_{f,ibest,g} - T_F M_{f,g}) \quad (2)$$

Difference means can be evaluated as the subtraction of attribute mean from the finest individual attribute given in (2). Where, $X_{f,ibest,g}$ is the best individual in the subject f.TF is the teaching factor with the value 1 or 2 for best practice and $r_k$ is the random number range from 0 to 1. Best individual acts as a teacher and trains the remaining individuals by considering them as learners to improve their knowledge based on his ability as shown in (3).

$$X'_{f,i,g} = X_{f,i,g} + \text{Diff\_Mean}_{f,g} \quad (3)$$

Where, $X'_{f,i,g}$is the trained value of $X_{f,i,g}$

Step 5: If result related to $X'_{f,i,g}$is better than the outcome related to$X_{f,i,g}$ ,

Continue the prior value

Otherwise,
 Substitute the prior value.

*Learner Phase*

In the second section,the learner's skill can be improved by interacting among themselves in the population. Learners with more knowledge can enhance the familiarity of the weak learner with the arbitrary communication. Learner phase is presented in further steps.

Step 6: Choose two instances U and V with the condition$X'_{total-U,g} \neq X'_{total-V,g}$ at random.

Where, $X'_{total-U,g}$, $X'_{total-V,g}$are restructured attributes of $X_{total-U,g}$, $X_{total-V,g}$of U and V respectively.

Step 7: If the result $X'_{total-U,g}$ is better than the result$X'_{total-V,g}$

$$X''_{f,U,g} = X'_{f,U,g} + r(X'_{f,V,g} - X'_{f,U,g}) \quad (4)$$

Otherwise,
$$X''_{f,U,g} = X'_{f,U,g} + r(X'_{f,U,g} - X'_{f,V,g}) \quad (5)$$

Step 8: If result related to $X''_{f,U,g}$is better than the result related to $X'_{f,U,g}$

 Then continue the prior value
 Otherwise,
 Substitute the prior value.

Step 9: If the termination condition fulfilled,
 Thenreport the result
 Otherwise,
 Go to Step 2

The population contains only the top individuals with best fitness values and a set of features after multiple iterations of teacher and learner phases. The dataset with the new individuals can be used to train the classifier models for better performance.

### 3.4. Classification Models

Classification is the problem of forecasting a definite objective using the training dataset. We have adopted four classifiers, namely, Multilayer Perceptron classifier (MLPC), J48, Random Forest, and Logistic Regression algorithms suitable for this work [41].MLPC is a feed forward artificial neural network model which usesbackpropagation for learning the model. It maps input data sets onto a set of suitable outputs and used in supervised learning problems with a pair of the input-output training set. J48 is an implementation of the C4.5 algorithm can be used for classification. Random forest (RF) is a leading supervised machine learning classifier which avoids the over fitting problem if there are enough trees in the forest and will achieve high classification precision. Logistic regression is a statistical model in which one or more independent variables will determine the outcome for analyzing the data set.

# Feature Optimization using Teaching Learning Based Optimization for Breast Disease Diagnosis

In machine learning, classification is a supervised learning method in which the computer algorithm learns from the labeled input data given to it and then uses this learning to classify fresh observation. This data set may basically be bi-class or it may be multi-class. The application of classification algorithms on bi-class problems givesonly binary predictions. Some examples of classification problems areabnormal/normal or yes/no or present/absence, etc. The performance of a binary classifier is usually summarized in a confusion matrix or error matrix. There are different measures for evaluating the performance of classification algorithms over the datasets. Classification accuracy or correctly classified Rate (CCR) is a preferable measure for specifying the classification accuracy. It is frequently illustrated in terms of percentage as shown in (6).

Accuracy ( Correctly classified observations Total classified observations ) * 100(

Accuracy alone will be unclear if thedata set contains three or more classes. We can overcome this limitation of classification accuracy with the confusion matrix.The performance of a classification model can be evaluated to know the correctness and erroneous using this matrix.It gives the visualization of the performance of the model. Right and wrong guesses are computed relating to each class. In bi-class problems, the confusion matrix can berepresented with a table of two columns and two rows which gives true positive (TP), true negative (TN) along the diagonal and false positives (FP) and false negatives (FN) in opposite sides as shown in TableI. One class represents positive and another negative.

In this study, we defined the confusion matrix with the ratio of a number of actual malignant records predicted as benign and anumber of actual benign records predicted as malignant.The other measures which we focused on for better understanding are Kappa statistic, F-Measure,Precision, Recall and true positive, false positive rates.

## IV. RESULTS & DISCUSSION

We have used various data mining methods namely Multilayer Perceptron, J48, Random Forest and Logistic Regression for binary classification of the test dataset. We performed classification in 2 cases: With original features of the datasetand TLBO optimized dataset. The proposed algorithm is evaluated by measuring the accuracy of classifiers on standard Wisconsin Breast Cancer medical dataset.

**Table I. Confusion Matrix**

| | | Predicted | |
|---|---|---|---|
| | | "+" ve | "-" ve |
| Actual | "+" ve | TP | FN |
| | "-" ve | FP | TN |

### 4.1. Experimental Setup and Dataset

The performance of TLBO is evaluated with Breast Cancer Wisconsin (Diagnostic) Dataset from UCI machine learning datasets and can be accessed from HTTP://archive.ics.uci.edu/ml. Dataset has total 569 records belong to either Malignant (M) or Benign (B) tumors. There are 357 benign records and 212 malignant records in the dataset. The features are derived from a digitized picture of a fine-needle aspirate (FNA) of patient s' breast matter. Each record contains 30 features which are made from ten real-valued features, along with their mean, standard error and the ''worst'' (mean of 3 highest values) values [29].

Two different cases were set up for training the classification models with WBC diagnostic dataset. In the first case, we trained the classifiers with the original features of the dataset.In the second case, we applied TLBO algorithm on the dataset to get optimized values for the features. Most of the researchers used k-fold cross-validationto achieve excellent results for breast cancer diagnosis. We used tenfold cross-validation in training and testing the above three selected set of features for measuring the classification accuracy of the proposed model. In 10-fold cross-validation procedure, the dataset is partitioned into ten subsets randomly. Nine subsets are utilized for training the models and remaining one set of features are used for testing the trained models for getting the classification accuracy. Finally, we evaluated the performance of the proposed method against the remaining cases.

### 4.2. Comparison of Accuracy Results

Our recommended system is an optimization process to improve the identification rate and minimize the computation time. To assess the performance of the suggested model over standard diagnosis systems with original features and optimized feature set, we have used themeasures like kappa statistic, precision, recall, FP rate, TP rate, F-Measure and accuracy rate as the most significant point for evaluation. Theoutcomes are presented in Table II.Classification accuracy (CA) can also be defined as shown in (7).

$$CA = \frac{TP+TN}{TP+TN+FP+FN}$$

(7)

Classification accuracy of the proposed model for the WBC (Diagnostic) dataset was approximated using 10-fold cross-validation. We got the classification accuracy of 96.66 % with original dataset featuresand 97.89% with TLBO optimized features for multilayer perceptron classifier. Accuracy measures achieved with the J48 classifier are 93.15% with an original set of features and 95.8% with TLBOoptimizedfeatures. Random Forest gave the values of 95.78% with the

**Table II. Classification Accuracy**

| | With Original Features | With TLBO Optimized Features |
|---|---|---|
| Multilayer Perceptron | 96.66 | 97.89 |
| J48 | 93.15 | 95.78 |
| Random Forest | 95.78 | 96.66 |
| Logistic Regression | 93.49 | 97.36 |

Original dataset and TLBO selected features. Finally, Logistic Regression classifier provided 93.49% and97.36% accuracy values with the two feature sets respectively. TLBO based optimized features achieved abettor result except with the Random Forest classifier.

This result proves that multilayer perceptron is better as compared to other classifiers. The graphical illustration of the classification accuracy is represented in Fig. 2.

Classification accuracy is a good choice for performance analysis but, it has its own constraints and runs into troubles in practice. The main difficulty with classification accuracy is that it conceals the information to better understand the performance of the classification model. The first condition to encounter this problem is when the data has more than 2 classes. If the classification accuracy is 80% with 3 or more classes, then we don't know if that is for the reason that all classes are being predicted uniformly well or whether one or two classes are being ignored by the model. It is not good if the achieved accuracy is 90% with the data that does not have an even number of classes. We can achieve a good score if 80 records for every 100 belong to one class by always predicting the most common class value. Confusion matrix can better overcome this problem to measure the performance of the model.

The confusion matrix illustrates misunderstandings made by the classifier during the predictions. The count of correct and incorrect guesses for every class is organized into a table or matrix as shown in TableIII. Every row of the matrix represents actual positive and negative classes in the dataset. Every column of the matrix represents predicted positive and negative classes.The sum of the predictions in each row will give the overall count of the instances related to each class. Diagnosing the cancer is a bi-class categorization problem in which the tumor will be identified as Malignant (M) or Benign (B). Multilayer Perceptron Classifier correctly predicted 201 as malignant tumors out of 212 and 349 out of 357 as benign tumors. It wrongly predicted 8 as malignant and 11 as benign with the original features of dataset. The exact predicted values (201 + 349) are positioned at top left and bottom right locations in the matrix table. The majorityof the mistakes occurred while guessingmalignant rather than benign.The classifiers improved the prediction rate for both classes with optimized features.Logistic Regression predicted 22 benign tumors as malignant and reduced to 10 after feature selection. But the count has decreased to 6 after TLBO based feature optimization.
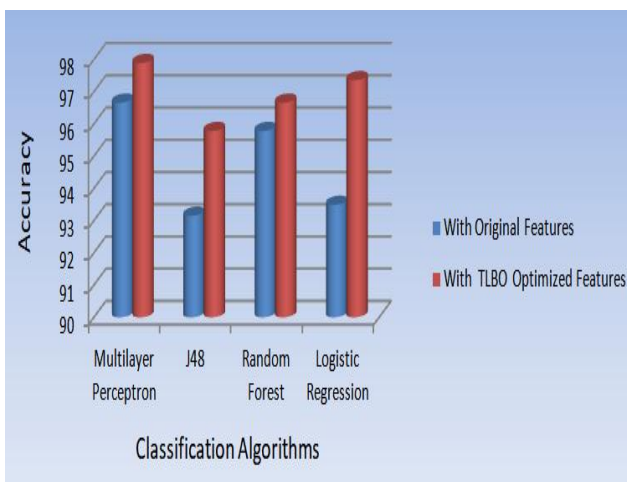


**Figure 2: Classification Accuracy with WDBC Dataset**

**Table III.  Confusion Matrix**

| Classifier | With Original Features | | With TLBO Optimized Features | | Actual Result / Classification |
|---|---|---|---|---|---|
|  | M | B | M | B |  |
| Multilayer Perceptron | 201 | 11 | 203 | 9 | M |
|  | 8 | 349 | 3 | 354 | B |
| J48 | 196 | 16 | 200 | 12 | M |
|  | 23 | 334 | 12 | 345 | B |
| Random Forest | 197 | 15 | 200 | 12 | M |
|  | 9 | 348 | 7 | 350 | B |
| Logistic Regression | 197 | 15 | 203 | 09 | M |
|  | 22 | 335 | 06 | 351 | B |

*(Header: Predictive Result / Classification)*

We have compared the performance of different feature selection techniques over the WDBC dataset for automatic diagnosis of breast cancer as shown in Table IV. Chee Kau Lim et al. [30] developed a Bandler Kohout (BK) sub-product-based inference systems to classify WDBC dataset and achieved the accuracy of 95.26% with BK-IVFS (weighted). Carlos J. Mantas et al. [31] proposed an improved model of the C4.5 algorithm based on a mathematical assumption of inaccurate probabilities and ambiguity measures on credal sets to classify WDBC dataset and achieved the accuracy of 95.01% with C4.5 and 95.12% with new Credal-C4.5. Joaquín Pacheco et al. [32] developed a technique for expanding classification trees based on a metaheuristic approach known as Greedy Randomized Adaptive Search Procedure for optimization problems and achieved 94.8% classification accuracy with WDBC dataset. Michel Ballings et al. [33] proposed a method for kernel machines to split the train data into kernel matrix for binaryclassification and achieved the accuracy of 95.61% with the random forest classifier. Ying Huang et al. [34] composed a hybrid predictive model based on k-means clustering and a classic inductive rule learning method (FOIL) and evaluated the classification accuracy of 94.6% with the WDBC dataset.

**Table IV.  Comparison of Performance**

| S. No | Author | Method | Accuracy |
|---|---|---|---|
| 1 | Chee Kau Lim [30] | BK-IVFS (weighted) | 95.26 |
| 2 | Mantas and Abellan | C4.5 (no pruning) | 94.37 |

| | | | |
|---|---|---|---|
| | [31] | C4.5 (pruning) | 95.01 |
| 3 | Mantas and Abellan [31] | Credal-C4.5 (no pruning) | 95.08 |
| | | Credal-C4.5 (pruning) | 95.12 |
| 4 | Pacheco [32] | GRASP | 94.80 |
| 5 | Ballings [33] | KF | 95.61 |
| 6 | YingHuang [34] | K-means + FOIL | 94.60 |
| 7 | Koloseni [35] | DE | 93.64 |
| 8 | Astudillo [36] | SOM | 93.32 |
| 9 | Sinatabakhi 37] | ACO | 91.76 |
| 10 | Saez [38] | MI | 96.14 |
| 11 | This work | TLBO | 97.89 |

Koloseni et al. [35] proposed differential evolution-based data classification method in which total distance measure is used for the final classification decisions and achieved an average classification accuracy of 93.64 % with WDBC dataset.Astudillo et al. [36] applied the semi-supervised method on WDBC using tree-based topology-oriented SOM to identify malignant and benign breast tumors and got the accuracy of 93.32 % in classification. An unsupervised multivariate filter-based feature selection technique has been proposed by Sinatabakhi et al. [37] using ant colony optimization (ACO) and achieved the accuracies of 90.72 %, 91.76% and 91.29 with respect to SVM, DT and NB classifiers on WDBC dataset. Saez et al. [38] used mutual information (MI) feature weighting algorithmfor NN classifier to classify the WDBC dataset and achieved the accuracy of 96.14%.The proposed TLBO model outperforms other compared methodsin terms of detection rate for diagnosing the breast cancer with a classification accuracy of 97.89 %.

The important accuracy measures like kappa statistic, precision, recall, FP rate, TP rate, and F-Measure are tabulated as shown in TableV. Outcomes proved that the recommended Optimal feature selection scheme amend the performance of all classifiers and achieved higher classification accuracy with multilayerperceptron classifier (97.89%).The results proved that the proposed algorithm can optimize the medical dataset for getting bestresults in breast cancer diagnosis.

Recall is defined as the fraction of the total number of exactly predicted positive samples to the total number of positive samples.It provides the information about when it's really yes, how frequently does it guess yes.Precision is described as the portion of the overall number of correctly predicted positive samples to the total number of classified positive samples. Precision describes when it guesses yes, how frequently it is right. High precision specifies a sample labeled as positive is certainly positive. High recall and low precisionindicate that a good number of the positive samples are correctly identified however there aremanyfalse positives. Low recall and high precisionindicate that there are many false negative samples in the prediction. F-measure characterizes both the precision and recall. It can be calculated with the harmonic mean instead of the arithmetic mean to punish the excessive values and is closer to the smaller value of recall or precision.

**Table V.  Accuracy Measures**

| Accuracy Measures \ Algorithm | Kappa statistic | TP Rate | FP Rate | Precision | Recall | F-Measure | | |
|---|---|---|---|---|---|---|---|---|
| Multilayer Perceptron | 0.9284 | 0.948 | 0.022 | 0.962 | 0.948 | 0.955 | M | With TLBO Optimized Features |
| | | 0.978 | 0.052 | 0.969 | 0.978 | 0.974 | B | |
| J48 | 0.8544 | 0.925 | 0.064 | 0.895 | 0.925 | 0.910 | M | |
| | | 0.936 | 0.075 | 0.954 | 0.936 | 0.945 | B | |
| Random Forest | 0.9093 | 0.929 | 0.025 | 0.956 | 0.929 | 0.943 | M | |
| | | 0.975 | 0.071 | 0.959 | 0.975 | 0.967 | B | |
| Logistic Regression | 0.8618 | 0.929 | 0.062 | 0.900 | 0.929 | 0.914 | M | |

| | | 0.938 | 0.071 | 0.957 | 0.938 | 0.948 | B | |
|---|---|---|---|---|---|---|---|---|
| MultilayerPerceptron | 0.9546 | 0.958 | 0.008 | 0.985 | 0.958 | 0.971 | M | |
| | | 0.992 | 0.042 | 0.975 | 0.992 | 0.983 | B | |
| J48 | 0.9098 | 0.943 | 0.034 | 0.943 | 0.943 | 0.910 | M | With Original Features |
| | | 0.966 | 0.057 | 0.966 | 0.966 | 0.910 | B | |
| Random Forest | 0.9282 | 0.943 | 0.020 | 0.966 | 0.943 | 0.955 | M | |
| | | 0.980 | 0.057 | 0.967 | 0.980 | 0.974 | B | |
| Logistic Regression | 0.9435 | 0.958 | 0.017 | 0.971 | 0.958 | 0.964 | M | |
| | | 0.983 | 0.042 | 0.975 | 0.983 | 0979 | B | |

## V.CONCLUSION

The main theme of our paper is to improve the accuracy of automatic breast cancer diagnosis system using TLBO based optimal feature selection. The performance of the devised technique is verified with the help of MLP, J48, random forest and logistic regression classifiers on Wisconsin diagnosis breast cancer (WDBC) dataset. Logistic Regression achieved the best result in comparison with the initial accuracy values, and MLPC got the best accuracy of 97.89% compared to other classifiers. The investigational results demonstrate that the proposed TLBO based feature optimization system can advance the accuracy of classifiers compared to previous methods. We would like to extend our work with a new hybrid algorithm in the future.

## REFERENCES

1. Shuihua Wang, Ravipudi Venkata Rao, Peng Chen, "Abnormal Breast Detection in Mammogram Images by Feed-forward Neural Network Trained by Jaya Algorithm", FundamentaInformaticae, vol. 151, no. 1-4, pp. 191-211, 2017.
2. S.Sasikala, S.Appavu and S.Geetha, "A Novel Feature Selection Technique for ImprovedSurvivabilityDiagnosis of Breast Cancer", 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15), pp. 16-23, 2015.
3. EminaAlicKovic,Abdulhamit Subasi, "Breast cancer diagnosis using GA feature selection and Rotation Forest",Neural Computing and Applications, vol. 28(4), pp. 753–763, 2017.
4. Hasmarina Hasan, Nooritawati Md Tahir, "Feature Selection of Breast Cancer Based on Principal Component Analysis", 6th International Colloquium on Signal Processing & Its Applications (CSPA), IEEE, 2010.
5. Noel Pérez1, Miguel A. Guevara, Augusto Silva, "Evaluation of Features Selection methods for breast cancer classification", 15th International Conference on Experimental Mechanics, 2012.
6. ShokoufehAalaei, Hadi Shahraki, Alireza Rowhanimanesh, and Saeid Eslami, "Feature selection using genetic algorithm for breast cancer diagnosis: experiment on three different datasets", Iranian Journal of Basic Medical Sciences, vol. 19(5), pp. 476-482, 2016.
7. Li-Yeh Chuang a, Sheng-Wei Tsai b, Cheng-Hong Yang, "Improved binary particle swarm optimization using catfish effect for feature selection", Expert Systems with Applications, Vol. 38,pp. 12699–12707, 2011.
8. Bichen ZhengSang WonYoon, Sarah S.Lam, "Breast cancer diagnosis based on feature extraction using a hybrid of K-means and support vector machine algorithms", Expert Systems with Applications, Volume 41, Issue 4, pp. 1476-1482, 2014.
9. T.Sridevi, A.Murugan, "A Novel Feature Selection Method for Effective Breast Cancer Diagnosis and Prognosis", International Journal of Computer Applications, vol. 88(11), pp. 0975 – 8887, 2014.
10. Vartika Agrawal, Satish Chandra, "Feature Selection using Artificial Bee Colony Algorithm for Medical Image Classification", International Conference on Contemporary Computing (IC3), IEEE, pp. 171-176, 2015.
11. Mohan Allam, M. Nandhini, "A Study on Optimization Techniques in Feature Selection for Medical Image Analysis", International Journal on Computer Science and Engineering (IJCSE), vol. 9 (3), pp. 75-82, 2017.
12. Zhi Chen, Tao Lin, Ningjiu Tang, and Xin Xia, "A Parallel Genetic Algorithm Based Feature Selection and Parameter Optimization for Support Vector Machine", Volume 2016 (2016), Article ID 2739621.
13. Saleh Shahbeig, Mohammad SadeghHelfroush, Akbar Rahideh, "A Fuzzy Multi-Objective Hybrid TLBO-PSO Approach to Select the Associated Genes with Breast Cancer", Signal Processing, vol. 131, pp. 58-65, 2016.
14. Jung IS, Thapa D, Wang GN, "Neural Network Based Algorithms for Diagnosis and Classification of Breast Cancer Tumor", Computational Intelligence and Security, Lecture Notes in Computer Science, vol. 380, pp. 107–114, Springer, 2005.
15. HtetThazin Tike Thein, Khin Mo Tun, "An Approach for Breast Cancer Diagnosis Classification using Neural network", Advanced Computing: An International Journal (ACIJ), vol. 6, no.1, 2015.
16. Paul R. Harper, "A review and comparison of classification algorithms for medical decision making Classification", Elsevier, vol. 71, pp. 315–331, 205.
17. R.V.Rao, V.J.Savsani, D.P.Vakharia, "Teaching–learning-based optimization: A novel method for constrained mechanical design optimization problems",Computer-Aided Design, Science Direct, vol. 43(3), pp. 303-315, 2011.
18. R. Venkata Rao, "Review of applications of TLBO algorithm and a tutorial for beginners to solve the unconstrained and constrained optimization problems", Decision Science Letters, vol. 5, pp. 1-30, 2016.
19. Shikha Agrawal, Shraddha Sharma, Sanjay Silakari, "Teaching Learning Based Optimization Based Improved Iris Recognition System", Advances in Intelligent Systems and Computing, vol. 366, pp. 735-740, 2015.
20. GulaveSukdev Madhav, Surendra Vishwakarma,"Improve the Performance of Face Recognition Using Feature Based and Teacher Learning Based Optimization Techniques", International Journal of Master of Engineering Research and Technology vol. 2, no. 10, 2015.
21. AbhilashaWakde, Virendra Shrivastava, "The Improved Performance of Face Detection Based on Partial Feature and TLBO", International Journal of Emerging Technology and Advanced Engineering, vol. 6, no. 9, 2016.

22. Chereddy Srinivasa Rao, KanadamKarteeka Pavan1, Allam AppaRao, "An Automatic Medical Image Segmentation using Teaching Learning Based Optimization", Proc. of Int. Conf. on Advances in Computer Science, AETACS, 2013.

23. Baljit Singh Khehra, Amar Partap Singh Pharwaha, "Image Segmentation using Teaching-Learning based Optimization Algorithm and Fuzzy Entropy", 15th International Conference on Computational Science and Its Applications, 2015.

24. P. Naga Srinivasu, G. Srinivas,and T. Srinivas Rao, "An automated Brain MRI image segmentation using Generic Algorithm and TLBO", IJCTA, 9(32), pp. 233-241, 2016.

25. Panda M, Elephant search optimization combined with deep neural network for microarray data analysis. Journal of King Saud University – Computer and Information Sciences (2017), https://doi.org/10.1016/j.jksuci.2017.12.002

26. Suresh C. Satapathy, Anima Naik, Parvathi K, "Rough set and teaching learning based optimization technique for optimal features selection", Central European Journal of Computer Science, vol. 3, no. 1, pp. 27-42, 2013.

27. Kapil Verma, Akash Mittal, Yogendra Kumar Jain, "A Hybrid method of face detection based on Feature Extraction using PIFR and Feature Optimization using TLBO", Journal of Engineering Research and Applications, vol. 6, Issue 1, pp.145-150, 2016.

28. Malkhede Seema Krishna, Surendra Vishwakarma," Improve the Performance of Fingerprint Recognition Using Feature Selection and Teacher Learning Based Optimization (TLBO)", International Journal of Master of Engineering Research and Technology, vol. 2, no.10, 2015.

29. UCI Machine Learning Repository: Breast Cancer Wisconsin (Diagnostic) Data Set (2012) Retrieved 15 Mar 2012 http://archive.ics.uci.edu/ml/datasets/Breast+ Cancer+Wisconsin+(Diagnostic)

30. Chee Kau Lim, Chee Seng Chan, "A weighted inference engine based on interval-valued fuzzy relationaltheory", Expert Systemswith Applications, vol. 42, pp. 3410-3419, 2015.

31. Mantas C. J, Abellan. J, "Credal-C4.5: Decision tree based on imprecise probabilities to classify noisy data", ExpertSystems with Applications, vol. 41, no. 10, pp. 4625–4637, 2014.

32. Pacheco J, Alfaro E, Casado S, Gamez M, & Garcia N,"A GRASP method for building classification trees, Expert Systems with Applications", 39(3),3241–3248, 2012.

33. Michel Ballings,Dirk Van den Poel, "Kernel Factory: An ensemble of kernel machines", Expert Systems with Applications, vol. 40,pp. 2904–2913, 2013.

34. Ying Huang, Kechadi T, "An effective hybrid learning system for telecommunication churn prediction", Expert Systems with Applications, 40(14), 5635–5647, 2013.

35. David Koloseni,JouniLampinen, PasiLuukka , "Differential evolution based nearest prototype classifier with optimized distance measures for the features in the data sets", Expert Systems with Applications, vol. 40,pp.4075–4082, 2013.

36. A.Astudillo, B. John Oommen, "On achieving semi-supervised pattern recognition by utilizing tree-based SOMs", Pattern Recognition, Vol (46), pg. 293-304, 2013.

37. SinaTabakhi, Parham Moradi, "Relevance-redundancy feature selection based on ant colony optimization", Pattern Recognition, Volume 48, Issue 9, Pages 2798-2811, 2015.

38. José A.Sáez, JoaquínDerrac , "Statistical computation of feature weighting schemes through data estimation for nearest neighbor classifiers", Pattern Recognition, Volume 47, Issue 12, pp. 3941-3948, 2014.

39. R.V.Rao, "Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems", International Journal of Industrial Engineering Computations, vol. 7, no. 1, pp. 19-34, 2016.

40. R.V.Rao, "Teaching Learning Based Optimization Algorithm And Its Engineering Applications", Springer International Publishing, 2016

41. Danasingh Asir Antony Gnana Singh, Subramanian Appavu Alias Balamurugan, Epiphany Jebamalar Leavline. An Unsupervised Feature Selection Algorithm with Feature Ranking for Maximizing Performance of the Classifiers[J]. International Journal of Automation and Computing, vol. 12, no. 5, pp. 511-517, 2015