# MEDICAL DATA CLASSIFICATION USING BIO-INSPIRED ALGORITHM

**Balasaheb Tarle[1]***

**Abstract**
In recent years, the clinical decision support system (DSS) has emerged as an important area in medical sciences to assist clinicians in medical diagnosis. Health records classification is based on learning from various health datasets to improve the better quality of DSS in health care. The main objective of this investigation is to establish a system for the successful classification of health data. Orthogonal Local Preservation Projection (OLPP) has been used to obtain promising outcomes in medical data classification. This is a high-dimensional data input package. A feature-reduction tool is then used to reduce the functionality space without compromising the calculation accuracy. The Artificial Neural Network shall be used as a classifier. We used an optimization algorithm to boost efficiency. The "artificial bee colony algorithm" is a bio-based optimization algorithm a neural network uses. The medical datasets represent the average improvement in the proposed system classification quality with the current form, around 15%.

**Keywords:** Artificial Neural Network, Artificial Bee Colony, Medical Data Classification, Orthogonal Local Preserving Projection, Training Data, Testing Data

---

[1]*Associate Professor, Computer Engineering Department, MVP Samaj's Karmaveer Adv. Baburao Ganpatrao Thakare College of Engineering Nashik, India-422013, E-mail:- tarlebs123@gmail.com

***Corresponding Author: -** Balasaheb Tarle
*Associate Professor, Computer Engineering Department, MVP Samaj's Karmaveer Adv. Baburao Ganpatrao Thakare College of Engineering Nashik, India-422013, E-mail:- tarlebs123@gmail.com

**DOI**: 10.48047/EC/2023.12.si5a.0543

# I. INTRODUCTION

The real-world data contains irrelevant or meaningless data termed noise which can significantly affect various data analysis tasks of data mining. The erroneous training data results in a classifier's low classification performance, increasing the algorithm's time complexity. Several researchers have proposed various techniques for data cleaning. Those techniques include neural networks, and filters, the approaches are discussed below.

The mechanism of assistance for clinical decision-making has become the medical profession's core atmosphere in recent years for doctors to offer help for medical diagnosis (D.V. Patil, 2012). Health background classification relies on studying a better level of DSS in medical care from a variety of health sources (Han et al., 2011). Data mining's key point is to detect emerging market habits and reveal data dynamics to submit relevant and essential data to consumers. (Anindita et al., 2017).One of the issues addressed in this study is the detection of positions in which essential features are detected and irrelevant or redundant. The proposed OLPP method algorithm is used to collect the feature subset. The network is a computational model that imitates the human structure. The best-known feature of the Artificial Neural Network is its capability to learn from empirical data sets and to maximize their performance from learning. ANN can coordinate, adapt, and evolve in real time (Abdullah et al., 2014). The vector of weight for the network is crucial since it contributes to the better fitness function value. This parameter is among the critical optimization problem occurring in ANN. Meta-Heuristic algorithms have been used to optimize the ANNs parameters. Here we have used the ABC algorithm for the optimization.

This paper is organized as follows: In Section 2, related works. In Section 3, the proposed methodology is the ANN-ABC algorithm, In Section 4, the experimental setup, performance evaluation, and results are discussed. Section 5 presents the comparative analysis, and Section 6 presents the conclusion of this research.

# II. LITERATURE REVIEW

In this paper, we discussed some of the more current papers on medical data classification using soft computing techniques of medical trends.

Class-wise work on the methods for cardiovascular prediction was conducted (Chitra and Seenivasagam, 2013, Tarle et al., 2020). The Neural Network was used here to make assumptions regarding medical details. The use of the NB classifier has been implemented in medical applications. A system for classifying NB techniques has been developed. Results help in detecting cardiac failure. A weighted medical decision-making approach has been introduced based on a weighted law (CDSS) for cardiac diagnosis. Dennis and Muthukrishnan (Dennis and Muthukrishnan, 2014) have introduced an effective clustering system for medical data focused on the AGFS. Data is used in the suggested technique to determine the main rule, and optimized rules for genetic algorithms are chosen. A simple way has been developed to coordinate engagement, discretion, and exercise.

The authors (Subbulakshmi and Deepa, 2015) proposed a hybrid methodology focusing on the paradigm of machine learning. This paradigm integrates the successful exploration mechanism called the self-regulated learning capability of the particle swarm optimization (PSO) algorithm with the extreme learning machine (ELM) classifier. As a new off-line learning method, ELM is a single hidden layer feed-forward neural network (FFNN) (Madhiarasan and Deepa, 2017), which proved to be an excellent classifier with many hidden layer neurons (Quang and Jeng, 2013, Gorzałczany et al., 2018). PSO is used to determine the optimum set of parameters for the ELM, thus reducing the number of hidden layer neurons and improving the network generalization performance. The accuracy of algorithms increases with the training dataset's size (Catlett, 1991). Learning from small datasets frequently decreases the accuracy of algorithms due to overfitting. It has been shown that popular algorithms for ANNs cannot deal with massive datasets (Abdullah et al., 2013).

For this reason, it can be helpful to use pre-processing methods to reduce input space and boost accuracy. Four of the most popular ANN training algorithms have been selected for the scalability of machine learning techniques to verify their effect on different selection approaches. These two algorithms reflect an O (n) and an adaptive learning rate (Bishop, 2006) of a downward gradient (GD). The scaled conjugate gradient (SCG) is a different algorithm, and the scaled conjugate gradient Leven berg-Marquardt (LM), all O (n2) and O (n3) inputs (Moller, 1993).

# III PROPOSED METHOD

## *A. ANN-ABC Algorithm*

Here, there is a decent method of organization for medical data classification in this paper. An orthogonal local preservation prediction and an ideal classification are used as part of the suggested approach to obtain a more robust result as an essential evaluation target. Initially, the knowledge index is selected from the medicinal database. Pre-processing is included at that stage in the data medical knowledge index. We need to derive valuable knowledge from the impressive, beneficial data collection pre-processing task. The attribute's subset collection technique can be helpful to minimize the function without sacrificing the designation's importance. The OLPP function dimensions are diminished here. The classification would be based on the optimal classifier after the process has been reduced. The approach suggested is used to maximize weights in the artificial neural network using an artificial colony algorithm. Fig.1 demonstrates the layout of the proposed architecture.

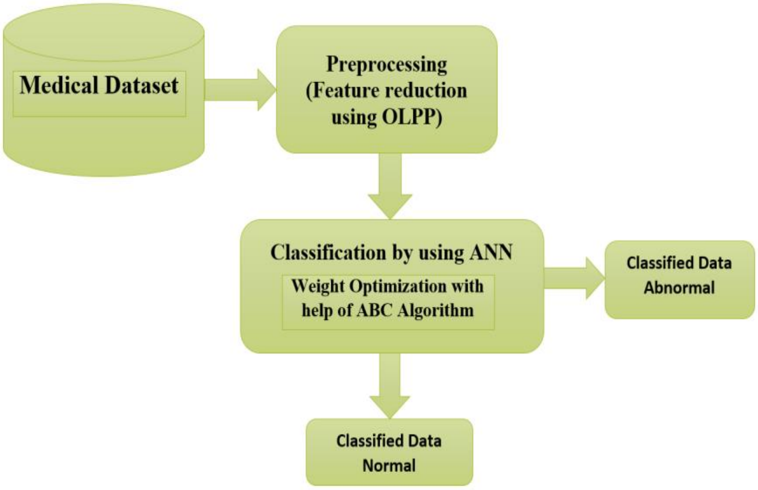


**Fig. 1** The Architecture of the Proposed Method

### *A. Classification Methods*

The ANN training process involves adjusting the weight of neurons iteratively to reduce the error. Hence, NN learning needs more powerful optimization techniques (Hyun et al., 2017). A set of datasets known as the training set is generally presented to inputs to begin the network training to determine the correct outputs. After finishing this process, unseen data, known as the testing set, is given to the network's input to test the classifier's general ability.

There are three primary learning paradigms for ANN. These paradigms include supervised, unsupervised, and reinforced learning. The most common learning algorithm is Back-Propagation (BP), a supervised learning algorithm. BP is a gradient descent-based algorithm. This algorithm calculates the network's output and reduces the mean square error (MSE) between the actual production and the desired outcome by adjusting weights accordingly (Liu, 2009; Kattan and Abdullah, 2013; Abdirashid et al., 2014). However, many researchers have found that BPNN is exposed to the problem associated with local minima in their learning instances. Furthermore, the performance of BP is also attributed to the selection of appropriate learning parameters.

Vector-based and adjacency matrix-based representations are two common strategies used for feed-forward artificial neural network weight representation. The vector-based representation is more appropriate for the HS algorithm. However, the method must accept the typical ranges specified by vectors [X, X] as NN weights have specific value ranges and are not discrete variables for optimization problems. In such a case, the HM representation can be regarded as different musicians exploiting similar musical instruments. Doing so; means the musical instrument will indicate that it will have a typical pitch range. Hence, each [X X] component value is identical to all decision variables. Each harmony vector in HM is represented using a vector representation, as shown in fig. 2. the vector comprises a complete set of NN weights and biases. In this work, HSA is only used for optimizing the weights of ANNs whenever BP's error is unchanged. For example, the error remains the same after six consecutive iterations. We suspect BP has been trapped in the local minimum or overfitting problems. This condition is used as an indicator of BP's unsuccessful learning process. Hence, HSA generates fresh near-optimal weights to continue the training process.

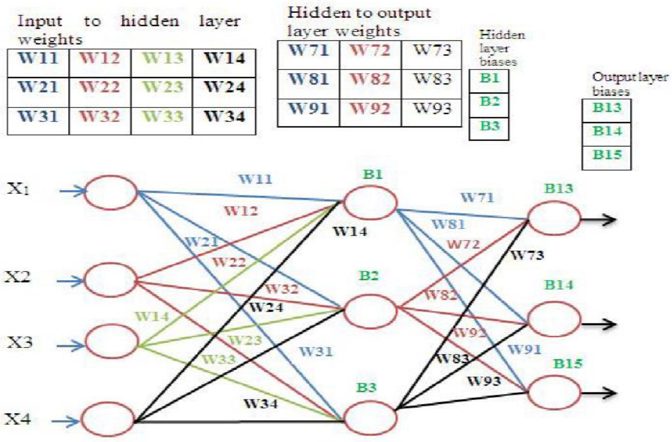


**Fig. 2** FFANN Sample for Weight Vector Representation

In our planned strategy, health data classification is composed of ANN. Presently specific components are isolated into two for training and testing reasons. In the first place, a few elements are used in the training stage, and then the unacquainted components are classified in the testing stage. An ANN involves an input layer, a hidden layer, and an output layer. The first favorable position of NN systems is that around no compelling reason to know information connections. It makes them learn skills and self-tuning capabilities. The following steps are to choose the parameter in the NN system. It implies selecting the input, hidden, and output layers. Each segment includes registering components known as neurons in hidden layers and output layers, which compute a weighted whole of the data sources. After that, it performs a nonlinear change on the entirety. Neurons having a place with different layers are connected via adaptive weights. The number of hidden layers and the number of neurons in every segment rely upon the utilization. We have utilized 19 input, ten hidden, and the unit output layers. We have run our work on the ABC technique in the present review. This technique is anticipated to find optimal weights to convey the NN system structure in the testing stage.

The following is the preparation of the neural system. Essential features in various categories (normal or abnormal) are acquired and assisted for training and testing in the neural network.

### Training Stage

In the training stage, the system is trained by giving it input and output designs. Through this stage, the neural system can change the association weights to partner with the ideal output in an iterative procedure until a fancied outcome is obtained. In the training method, exertion is made by altering the weights. Keeping in mind the end goal to modify the weights of the NN system, the anticipated strategy utilizes the optimization method to choose the ideal weights. Here for selecting the ideal weights, ABC techniques are being used. The well-ordered system of the ABC algorithm is appeared underneath the segment

### *A. Artificial Bee Colony Algorithm*

Karaboga initially published the Artificial Bee Colony algorithm (ABC) in 2005 as a technical report for numerical optimization problems. ABC is a new swarm intelligence algorithm proposed by Karaboga in 2005, inspired by honey bees' behavior. This dance is known as the waggle dance. Fig. 4 shows the waggle dance. The employed bees use this waggle dance to communicate to other bees in the hive to report three main types of information. It is regarding the availability of flower patches, which are the direction of food sources' location, quality, quantity, and distances from these food sources.

This awareness allows other bees in the hive to go further, often without the aid of other bees, towards the sensed patches. After wagging dances, the scout bees fly back to the flower spots with the supporters or the working bees.

An artificial bee colony algorithm is one of the three fundamental components of its operation (Karaboga, 2005). It applies to working bees, unpaid beekeepers, and food supplies.

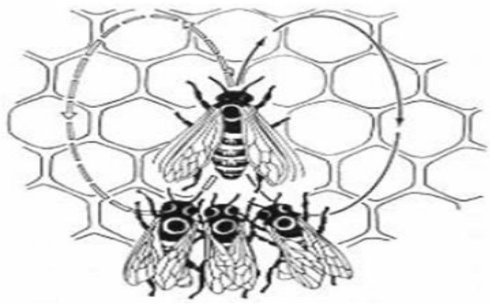

**Fig. 4** The Bees Waggle Dance

Employed bees: An employed artificial bee is often employed at one particular food source at a time which she exploits. She carries all important information about this specific food source and shares it with the rest of the bees waiting in the hive. Other information she shares includes the distance of the food source from the hive, its direction, and its profitability.
Unemployed bees: The group of forager bees looking for food sources to exploit is called unemployed bees. They can be scout bees that search around the environment randomly or onlooker bees who try to find food sources by using the information given by the employed bees. The mean number of scouts is about a percentage.
Food Sources: An artificial bee analyses several factors concerning a given food source before selecting that food source. These factors include the food source's closeness to the hive, richness, and quality of the energy, taste of its nectar, and the ease or difficulty of extracting this food from the source.

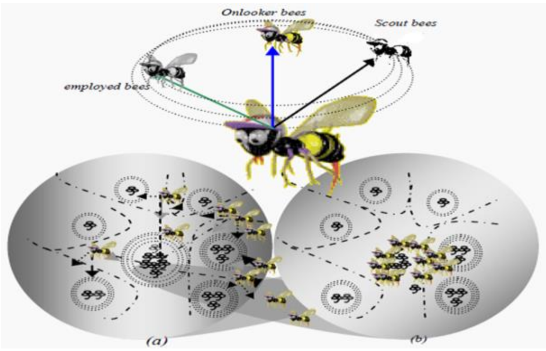


**Fig. 5** The Basic Mechanism Search of ABC (Mogaka et al., 2015)

In short, the artificial foraging bees consist of a group of employed bees, onlookers, and scout bees. Half of this colony comprises employed bees, which form the majority. Every food source has an employed bee associated with it. Once a food source is depleted, the employed bee automatically becomes a scout. The amount of nectar in a batch of flowers determines the fitness value of that solution, in this case, the food position. The primary mechanism search of ABC is well presented in fig. 5, where a) Initial situation, b) Final situation.

ABC weight optimization technique is proposed. Three types of honey bees are unique to the artificial bee colony algorithm: employed, onlooker, and scout. Initially, it was viewed and defined as starting food-forward relationship weights that the customer described as the size of employees' bees. Health is recorded for all working bees. Set up a bee viewer and evaluate the well-being of each bee viewer.
The absence of the opportunity to validate honeybees' execution during user time is not clarified. Then the honeybee scout is expelled and restored by the bee used, inserted arbitrarily. This process repeats the number of cycles specified by the consumer. Better honey bees characterize the final weights of the artificial neural system. The top and bottom of the ABC are noticeable in the bottom segment. As seen in fig. 6, the ABC algorithm's measurements are generally summarised (Mogaka et al., 2015).

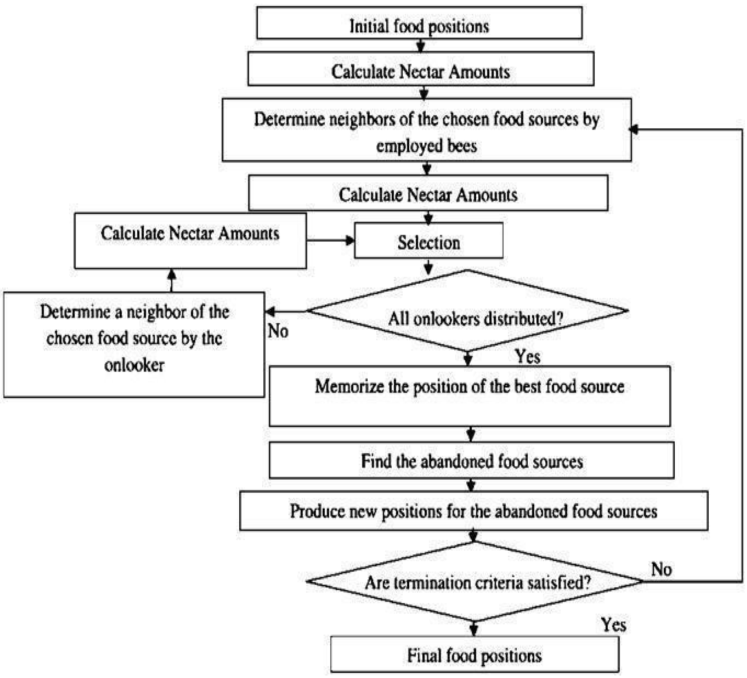


**Fig. 6** the Flow Chart of the Artificial Bee Colony Algorithm

### *A. Steps of ABC Algorithm*

**Algorithm 1 The ABC algorithm**

**Input:** Random weights

**Output:** Optimal weights

1. Begin
2. Initialize the population of weights,
3. Evaluate the fitness,
4. Repeat
5. Produce Employed bee to select the new food source

*For i=1, 2….N do*

$W_i^{New} = W_{ij} + \varphi(W_{ij} - W_{kj})$

$\forall k \in (1,2...N); k \neq j$ *// $\varphi$ is the random value [-1, 1].*

*Calculate* $Fit_i^{New}$

*If (* $Fit_i^{old} \leq Fit_i^{New}$ *) then*

$W_i = W_i^{New}$

*End if*

*End for*

6. Produce Onlooker bee to select the new food source

*For i=1, 2…., N do*

$W_i^{New} = W_i^{\max} + rand(W_i^{\max} - W_i^{\min})$

*Calculate* $Fit_i^{New}$

*If (* $Fit_i^{old} \leq Fit_i^{New}$ *) then*

$W_i = W_i^{New}$

*Calculate the probability value,*

$$prob = \frac{Fit_i}{\sum_{i=1}^{N} Fit_i}$$

*End if*

End for

7. Produce Scout bee to select the new food source

*For i=1, 2….N do*

*If (scout (i) =limit) then*

*Generate the weights randomly*

*End if*

*End for*

8. *End*

Finally, the optimized weight is used in artificial neural systems for classification of medical data.

### *B. The General Phases of ANN*

**Phase 1:** Set the mass parameter without the input layers for each neuron first.

**Phase 2:** Build a neural structure with {f1, f2, f3, ...,fn} components, showing Hn as a hidden parameter and O as output as input values.

**Phase 3:** The input layer calculates the input layer used by the output task.

$$Input(F) = \sum_{n=0}^{N} w_{(n1)}\, f_1(n) + w_{(n2)}\, f_2(n) + w_{(n3)}\, f_3(n) + .... + w_{(nN)}\, f_N(n)$$

(1)

For the output layer, the activation function is evaluated as,

$$Activation\ fun\ (F) = \frac{1}{1+e^{-F}} \qquad (2)$$

**Phase 4:** Detection of fault

The training results are checked according to the mean absolute error (MAE)

*MAE = 1/n Σ|Target-output|2* (3)

Based on the ' faultiness' of our concept, the neural system is used for categorizing health data.

**Testing Stage**

The ANN method is generated using inputs in the training process but less solutions. It is used to identify and detect values for each function. Various output measurements, such as sensitivity, precision, and consistency, are noted when classifying the test results using the proposed ANN. The success of our plan is measured and clarified in section 5.2 below.

## III. EXPERIMENTAL SETUP AND PERFORMANCE EVALUATION

### *A. Experimental Setup*

We have proposed ANN with the ABC algorithm. The four standard medical datasets from the UCI repository (Dua and Graff, 2019) were used for conducting the experiments. The datasets are Cleveland, Hungarian, Switzerland, and Chronic kidney disease datasets. We have used MATLAB 2015b, an implementation for the proposed system, and the performance measures were done using i5 CPU with 4 GB Memory. The 5-fold cross-validation method is used to test the performance of the proposed classifier on the said datasets.

### *B. Results and Performance Evaluation*

The performance of the proposed classifier (ANN with ABC) is evaluated using performance measures, namely: testing accuracy, sensitivity, specificity, and F-score. These parameters are explained in detail below. Four possible results are expected when any test is done using a classifier. If the instance is positive and classified as positive, it is considered True Positive (TP). If the example is positive and classified as unfavourable, it is considered False Negative (FN).

- Overall Accuracy: Accuracy is the ratio of the True predicted values to the Total predicted values.

$$OverallAcaracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (4)$$

- Specificity: Specificity is considered as the TNR (True Negative Rate). Specificity measures the ability of the proposed method to identify typical cases.

$$Specificity = \frac{TN}{FP + TN} \quad (5)$$

- Sensitivity: Recall, also known as sensitivity, is the fraction of examples classified as positive among the total number of positive examples. In other words, the number of true positives is divided by the number of true positives plus false negatives.

$$Sensitivity = \frac{TP}{TP + FN} \quad (6)$$

- Precision: Precision is defined as the ratio of the total number of correctly classified positive classes divided by the total number of predicted positive classes.

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

**Table 1** The table presents the classification performance of the proposed ANN with ABC classifier obtained using performance metrics TN, TP, FP, FN, FNR, and FPR on four medical datasets.

| Dataset | TP | TN | FP | FN | FNR | FPR |
|---|---|---|---|---|---|---|
| Cleveland | 148 | 128 | 11 | 16 | 0.0791 | 0.0976 |
| Hungarian | 187 | 86 | 20 | 01 | 0.1887 | 0.0053 |
| Switzerland | 04 | 114 | 01 | 04 | 0.0087 | 0.50 |
| Chronic kidney | 235 | 145 | 07 | 15 | 0.0461 | 0.06 |

- The F-score: It is also called the F1-score, which measures a model's accuracy on a dataset. The F-score combines the precision and recall of the model, and it is defined as the harmonic mean of the model's precision and recall. The F-score is the geometric mean of precision and recall.

$$F1 = \frac{2TP}{(2TP + FP + FN)} \quad (8)$$

**Table 2** The table presents the classification performance of the proposed ANN with ABC classifier obtained using performance measures Accuracy, Sensitivity, Specificity, Precision, and F-score on four medical datasets.

| Dataset | Accuracy | Sensitivity | Specificity | Precision | F-Score |
|---|---|---|---|---|---|
| Cleveland | 91.09 | 90.24 | 92.09 | 93.08 | 91.66 |
| Hungarian | 92.86 | 99.46 | 81.13 | 90.34 | 94.68 |
| Switzerland | 95.93 | 50.00 | 99.13 | 80.00 | 61.54 |
| Chronic kidney | 94.53 | 94.00 | 95.39 | 97.11 | 95.53 |

Table 1 presents the classification performance of the proposed ANN with ABC classifier obtained using performance metrics TN, TP, FP, FN, FNR, and FPR on four medical datasets. The efficiency of the proposed system is presented in Table 2. The table shows the classification performance of the proposed ANN with ABC classifier obtained using performance measures Accuracy, Sensitivity, Specificity, Precision, and F-score on four medical datasets.

The data in Table 2 is graphically represented in Fig. 7; See diagram graphically presents results obtained on performance measures Accuracy, Sensitivity, Specificity, Precision, and F-score using the proposed technique on four medical datasets.

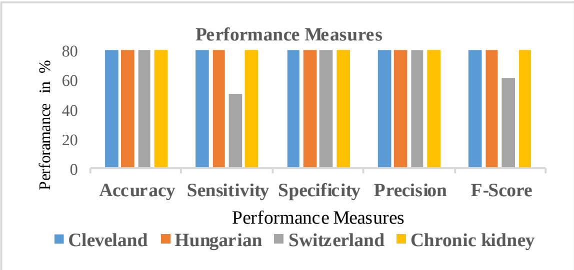


**Fig.7** The graphical representation of performance measures Accuracy, Sensitivity, Specificity, Precision, and F-score on the ANN-ABC Method using four medical datasets

## I. Comparative Analysis

In this section, the performance of the proposed ANN with the ABC algorithm is compared with the existing ANN algorithm using various performance measures. The comparison is made in terms of sensitivity, specificity, and accuracy. The Comparison outcomes are presented in the tables given below.

From fig. 8, it can be observed that the performance measure sensitivity (using the proposed method) archives better values on all three datasets, except for the Switzerland dataset. In the case of the Cleveland dataset, the sensitivity of ANN with ABC is 90.2%; whereas, the traditional neural network ANN achieves 81.09% sensitivity. For the Hungarian dataset and the chronic kidney dataset, the sensitivity value for the proposed algorithm is 99.46% and 94.00% respectively, contrary to this, the existing algorithm achieves 87.2% and 10.00% for both the Hungarian and the chronic kidney datasets. The Switzerland dataset attains 50.00% sensitivity for the proposed technique compared to 12.50 % for the existing technique.

**Table 3** Comparing sensitivity as a classification performance measure for proposed (ANN with ABC) vs. existing (ANN) method on four medical datasets

| Datasets | Sensitivity (ANN-ABC) | Sensitivity (ANN) |
|---|---|---|
| Cleveland | 90.24 | 81.09 |
| Hungarian | 99.46 | 87.23 |
| Switzerland | 50.00 | 12.50 |
| Chronic kidney | 94.00 | 10.00 |

**Fig. 8** compares the classification performance measure in terms of sensitivity for the proposed vs existing method graphically. Sensitivity is the ability of a test to identify patients with a disease correctly.

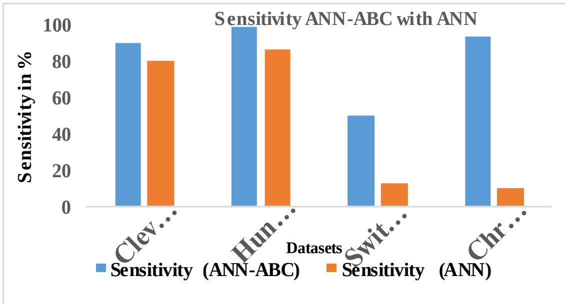


**Fig. 8.**The graphical representation showing the comparison of sensitivity as a classification performance measure for the proposed algorithm (ANN with ABC) vs the existing (ANN) method on four medical datasets

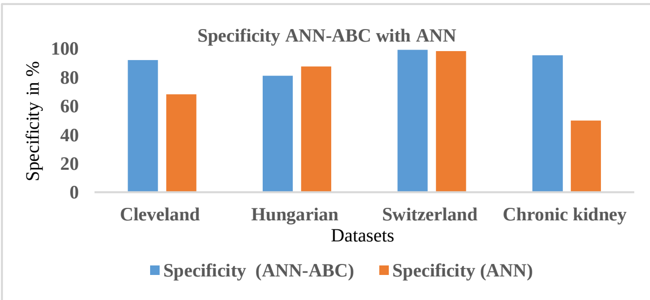


**Fig.9.** The graphical representation showing the comparison of specificity as a classification performance measure for the proposed algorithm (ANN with ABC) vs existing (ANN) method on four medical datasets

Fig. 9 compares the classification performance measure in terms of sensitivity for the proposed vs existing method graphically. Specificity is the ability of a test to identify people without the disease correctly.

**Table 4** The comparison of specificity as a classification performance measure for proposed (ANN with ABC) vs existing (ANN) technique on four medical datasets

| Datasets | Specificity (ANN-ABC) | Specificity (ANN) |
|---|---|---|
| Cleveland | 92.08 | 68.34 |
| Hungarian | 81.13 | 87.73 |
| Switzerland | 99.13 | 98.26 |
| Chronic kidney | 95.39 | 50.00 |

From fig. 9, it can be observed that the performance measure specificity (using the proposed method) archives better values on all four datasets. The specificity value for the Cleveland dataset is 92.08% for the proposed (ANN with ABC algorithm); whereas, the traditional neural network ANN achieves 68.34% specificity. The proposed method archives an 81.13% specificity value on the Hungarian dataset, whereas the conventional neural network ANN achieves a slightly better 87.7% specificity. The proposed method achieves the maximum specificity value of 99.13% for the Switzerland dataset, whereas the traditional neural network ANN achieves 98.26%. The proposed technique achieves a specificity value of 95.39% for the chronic kidney dataset. It is observed that the proposed method achieves a better specificity value than the existing ANN algorithm.

**Table 5** The comparison of classification accuracy achieved using the proposed (ANN with ABC) vs existing (ANN) method with average enhancement**.**

| Datasets | Accuracy (ANN-ABC) | Accuracy (ANN) | Δ1 |
|---|---|---|---|
| Cleveland | 91.09 | 75.25 | 17.39 |
| Hungarian | 92.85 | 87.41 | 5.86 |
| Switzerland | 95.93 | 92.68 | 3.38 |
| Chronic kidney | 94.53 | 62.18 | 34.22 |
| The average enhancement | 93.6 | 79.38 | 15.21 |

In our work, we have proposed the optimization technique ABC algorithm for the weight optimization process in ANN. The results are graphically presented in fig. 10, and the proposed technique ANN with ABC is compared with the existing ANN algorithm.

The classification accuracy is computed as the measure of performance. Let the classification accuracy obtained on the proposed system be represented A. The table presents accuracy on Artificial Neural Network defined as $A_{ANN\text{-}ABC}$. Similarly, the performance obtained on ANN algorithms is defined as $A_{NN.}$ The improvement in classification performance is calculated as $\Delta_1$, and is represented in Eq. (9).

The Equation is as given below.

$$\Delta_1 = ((A_{ANN-ABC} \; - A_{NN})/A_{ANN-ABC}) * 100 \qquad (9)$$

Observing fig. 10, we can state that the proposed optimization approach outperforms with an accuracy of 91.08% as compared to 75.25% accuracy on conventional ANN for the Cleveland dataset. The proposed ANN with ABC achieves a 92.85% accuracy value for the Hungarian dataset, whereas the exiting ANN achieves 87.41% accuracy. The accuracy value of the Switzerland dataset is 95.93%, whereas the conventional neural network ANN provides a 92.68% accuracy value. The proposed technique attains a 94.50% accuracy value for the chronic kidney dataset, whereas the traditional neural network ANN reaches 62.18%.

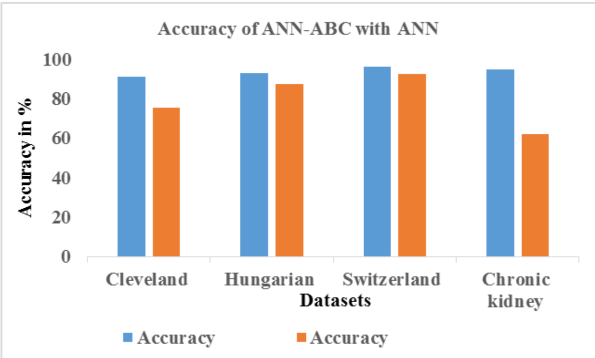


**Fig.10.**The comparison of classification accuracy for proposed (ANN-ABC) vs existing (ANN) method using four medical datasets

It can be observed that the classification performance using the proposed method (on ANN with ABC as compared to ANN) on all the datasets is enhanced by around 3% to 34%. The average improvement in classification accuracy with the proposed method to the existing method is around 15.21% on the medical datasets used. From these results, we can conclude that our method achieves an improved accuracy value compared with the existing methods.

To improve the performance of ANN, the bio-inspired optimization algorithm "Artificial bee colony algorithm" is used to optimize the weights of ANN during the learning process. When ABC is used, the optimization problem is first changed to the problem of finding the best parameter vector that minimizes an objective function. The artificial bees then find a population of initial solution vectors at random and iteratively enhance them using the strategies: migrating toward better solutions via a neighbour search mechanism while deleting poor solutions. Thus, it helps ANN to minimize the objective function. The objective function in ANN is MSE (Mean Square Error), and therefore ABC is the most suitable algorithm to optimize weights in ANN. We can then conclude that, in comparison to the approaches in use, our solution achieves better accuracy efficiency. We have performed two tests in hypothesis testing, t-test and Wilcoxon rank sum test. We have done a Wilcoxon rank sum test on hypothesis pairs (ANN, ANN-ABC). It is observed that the alternative hypothesis is true. We have performed a t-test on hypothesis pairs (ANN, ANN-ABC). It is observed that the alternative hypothesis is true.

## I. CONCUSSION

In recent years, the paradigm for promoting clinical judgments has emerged as a critical medical research area to assist physicians with patient diagnosis. It was effectively expanded to include several classification algorithms and pre-processing data methodologies. To mine valuable data, pre-processing techniques are applied to raw data. Selecting feature subsets is an essential issue in handling high dimensional datasets. Here in this contribution, we have proposed orthogonal local preserving projection for feature selection. To improve the performance of the artificial neural network, we have applied the Artificial Bee Colony Algorithm for optimizing weights in NN. For classification, a neural network classifier is combined with ABC. The implementation is done in the Mat lab. For experimentation, the datasets from the UCI repository are used. Datasets used are Cleveland, Hungarian, Switzerland, and Chronic kidney data. It can be observed that the classification performance of the proposed method (on ANN-ABC as compared to ANN) on all the datasets is enhanced. The average improvement in classification accuracy with the proposed technique to the existing technique is around 15.21% on the medical datasets used. From these results, we can conclude that our method achieves an improved accuracy value compared with the existing methods. As mentioned in detail earlier in this paper, when ABC is used, the optimization problem is first changed to the problem of finding the best parameter vector that minimises an objective function. Thus, it helps ANN to minimize the objective function. The objective function in ANN is MSE (Mean Square Error), and therefore ABC is the most suitable algorithm to optimize weights in ANN. We have performed two tests in hypothesis testing, t-test and Wilcoxon rank sum test. We have achieved a t-test a Wilcoxon rank sum test on hypothesis pairs (ANN, ANN-ABC). It is observed that the alternative hypothesis is true.

## DATA AVAILABILITY STATEMENTS

The datasets analysed during the current study are available in the UCI repository, the web link: https://archive.ics.uci.edu/ml/index.php

## CONFLICT OF INTEREST

The authors whose names are listed immediately below certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

## REFERENCES

1. Jiawei Han, Micheline Kamber, Jian Pei, Data Mining: Concepts and Techniques, 3rd Edition Morgan Kaufmann Publishers Inc. San Francisco, CA, USA, 2011.
2. C. V. Subbulakshmi and S. N. Deepa, Medical Dataset Classification: A Machine Learning Paradigm Integrating Particle Swarm Optimization with Extreme Learning Machine Classifier, The Scientific World Journal March 2015.
3. Patil D.V, Bichkar R..S., "Issues in Optimization of Decision Tree Learning: A Survey", International Journal of Applied Information Systems (IJAIS) Foundation of Computer Science FCS, New York, USA Volume 3 – No.5. 2012.
4. Quang Hung Do and Jeng-Fung Chen, A Neuro-Fuzzy Approach in the Classification of Students", Academic Performance, Hindawi Publishing Corporation Computational

Intelligence and Neuroscience, Volume 2013, pp 1-7. 2013.
5. Shang H.L., A survey of functional principal component analysis. Advances in Statistical Analysis, a journal of the German Statistical Society, 98 pages 121–142, 2014.
6. R. Alzubi, N. Ramzan, H. Alzoubi and A. Amira, A Hybrid Feature Selection Method for Complex Diseases SNPs, in IEEE Access, vol. 6, pp. 1292-1301, 2018.
7. L. Mogaka, D. K. Murage, M. J. Saulo,. Power Optimization and Prioritization in an Island Supplied by a Rotating Machine Based Distributed Generator Using Artificial Bee Colony Algorithm. International Journal of Energy and Power Engineering. Vol. 5, No. 1, pp. 15-21, 2015.
8. Gorzałczany, Marian B and Filip Rudziński, 2016. Interpretable and accurate medical data Classification-a multi-objective genetic-fuzzy optimization approach, Elsevier on Expert Systems with Applications, pages1-17.
9. Chitra R. and Seenivasagam V., Review of Heart Disease Prediction System Using Data Mining and Hybrid Intelligent Techniques, ICTACT Journal on Soft Computing, Vol. 03, No. 04, pp. 605-609, 2013.
10. Dennis B. and Muthukrishnan S., AGFS: Adaptive Genetic Fuzzy System for medical data classification, Applied Soft computing, Vol. 24, issue December 2014 pp. 242-252, 2014.
11. B. Tarle, C. Sanjay, S. Jena, Integrating multiple methods to enhance medical data classification, Journal Evolving Systems 15 (1), 133-142, 2020.
12. M. Madhiarasan and Deepa S.N., Comparative analysis on hidden neurons estimation in multi-layer perceptron neural networks for wind speed forecasting. Artificial Intelligence Rev 48, pp. 449–471, 2017.
13. Dua D. and Graff C., UCI Machine Learning Repository, http://archive.ics.uci. edu/ml]. Irvine, CA: the University of California, School of Information and Computer Science, 2019.
14. Abdullah, Rosni, Kattan, Ali, Herawan, Tutut, Ghazali, Rozaida, Mustafa Mat, Enhanced MWO Training Algorithm to Improve Classification Accuracy of Artificial Neural Networks, Recent Advances on Soft Computing and Data Mining, Springer International Publishing, pages.183-194, 2014.
15. N. Anindita, H. A. Nugroho and T. B. Adji, A Combination of multiple imputations and principal component analysis to handle missing value with the arbitrary pattern, 7th International Annual Engineering Seminar, Yogyakarta, pages, 1-5, 2017.
16. Kattan, A., Abdullah, R., Training Feed-Forward Artificial Neural Networks for Pattern-Classification Using the Harmony Search Algorithm. In: Proceedings the Second International Conference on Digital Enterprise and Information Systems (DEIS 2013), pp. 84–97, 2013.
17. Catlett J.,On changing continuous attributes into ordered discrete attributes. In Proceedings of the European Working Session on Learning, Springer Verlag, pp. 164– 178, Berlin, 1991.
18. Bishop, C.M. Pattern Recognition and Machine Learning. Springer, Berlin, 2006.
19. M. Moller, A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning, Neural Networks, Vol. 6, No. 4, pp. 525-533, 1993.
20. Liu Q.A, Sung H., Chen Z., Liu J., Huang X., Deng Y., Feature selection and classification of MAQC-II breast cancer and multiple myeloma microarray gene expression data, vol.4 issue 12, pp. 1-24, 2009.
21. Abdirashid Salad Nur, Nor Haizan Mohd Radzi, Ashraf Osman Ibrahimm, Artificial Neural Network Weight Optimization: A Review , TELKOMNIKA Indonesian Journal of Electrical Engineering, Vol. 12, No. 9, September 2014, pp. 6897 -~ 6902, 2014. DOI: 10.11591/telkomnika.v12i9.6264
22. Karaboga, D. An Idea Based on Honey Bee Swarm for Numerical Optimization. Technical Report-TR06, Department of Computer Engineering, Engineering Faculty, Erciyes University, 2005.

**Balasaheb Tarle** received the AMIE degree from The Institution of Engineers (India), Kolkata, in 2000, the BE degree from the K.K. Wagh College of Engineering, Nashik (Pune University) in 2002, and the ME degree from Devi Ahilya Vishwavidyalaya Indore (MP) India in 2007. He has been awarded the Ph.D. degree from GITAM (Deemed to be University) Visakhapatnam July 2022. He is currently working as an Associate Professor in the Department of Computer Engineering, MVPS's KBT College of Engineering Nashik, India. He has more than 21 research publications in reputed and peer-reviewed international journals, conference proceedings, and book chapters. His research interests include Data Mining, Software Engineering, and Machine Learning. He is a Life Member of ISTE (India) and a Member of the Institution of Engineers of India (IEI).