Indonesian Journal of Electrical Engineering and Computer Science
Vol. 25, No. 1, January 2022, pp. 265~272
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v25.i1.pp265-272

A novel salp swarm clustering algorithm for prediction of the heart diseases

Nitesh Sureja[1], Bharat Chawda[2], Avani Vasant[1]
[1]Department of Computer Science and Engineering, Krishna School of Emerging Technology and Applied Research, KPGU, Vadodara, India
[2]Department of Computer Engineering, Bhailalbhai and Bhikhabhai Institute of Technology, Vallabh Vidyanagar, India

Article Info

Article history:

Received Jul 12, 2021
Revised Oct 18, 2021
Accepted Nov 21, 2021

Keywords:

Classification
Clustering
Nature-inspired algorithms
Salp swarm optimisation
SVM classifiers

ABSTRACT

Heart diseases have a severe impact on human life and health. Cardiovascular deaths and diseases have increased at a fast rate worldwide. The early prediction of these diseases is necessary to prevent deaths. Now a day; a considerable amount of medical information is available and collected as databases. An efficient technique is required to analyse this data and predict the disease. Clustering can help medical practitioners in diagnosis by classifying the patient's data collected for a disease. Clustering techniques can analyse such data based on each patient-generated and predict disease. A new prediction model based on salp swarm algorithm and support vector machine is proposed in this research for predicting heart diseases. Salp swarm algorithm is used to select the useful features from the database. Support vector machine classifier is used to predict heart diseases. Results obtained are compared with the other algorithms available in the literature. It is observed that the proposed approach produces better results with accuracy 98.75% and 98.46% with the dataset 1 and 2, respectively. In addition to this, the algorithm converges in significantly less time in comparison to other algorithms. This algorithm might become a perfect supporting tool for medical practitioners.

Corresponding Author:

Nitesh Sureja
Department of Computer Science and Engineering
Krishna School of Emerging Technology and Applied Research, KPGU
Vadodara, Gujarat, India
Email: nmsureja@gmail.com

1. INTRODUCTION

The world health organisation (WHO) says that around 18 million deaths were reported due to heart diseases across the world in 2015 [1]. These diseases can be dangerous if a patient does not take care of his/her symptoms. In most cases, this ignorance leads to the death of the patient. One out of five persons develops the risk of heart disease over a lifetime. Despite numerous techniques in the health care field, predicting heart disease remains a difficult job, as health data contains helpful and not valuable features (variables). An accurate prediction in good time depends on extracting the useful features (variables) from the health data. On the other hand, removing not useful features reduces the computation time, risk of over-fitting, risk of under-fitting and improves the predictability.

Various traditional clustering algorithms have been applied to predict heart diseases [2]-[12]. All of them have their pros and cons. The traditional clustering algorithms used for the diagnosis of heart diseases often produce a delayed and inaccurate diagnoses. They suffer from extensive computational efforts, high convergence time, more parameters, stuck in local optima and stagnation. Also, they are not easy to implement due to their complex structure. Therefore, it is necessary to have techniques to overcome the problems with state of the algorithms to produce good accurate predictions. Recently, various nature-inspired algorithms have been applied to predict heart diseases. A brief review of nature-inspired algorithms proposed for predicting heart diseases is given in the next paragraphs.

A kernel extreme learning machines (KELM) approach with enhanced grey wolf optimiser (IGWO) is proposed in [13] for heart disease diagnosis. The IGWO algorithm is used to extract an optimal subset of features from the heart dataset. In [14], a model for reducing features is proposed. This model combines fuzzy c-means clustering with parameter optimisation by firefly and genetic algorithms. Comparing with other algorithms such as simple bayesian, support vector machine and neural network, it is found that the model performs better. A differential evolution (DE) approach for selecting the more important attributes for heart disease is proposed in [15]. The approach is tested with a heart dataset with hypertension, congenital heart disease, coronary heart disease, chronic pulmonary and rheumatic valvular heart disease. In [16], two algorithms inspired by the cuckoo search are proposed for feature selection on the heart disease dataset. Both the approaches used a filtering technique during the generation of subsets. The algorithms performed good by generating fewer features and good accuracy. A genetic algorithm-based fuzzy decision support system is proposed in [17]. A genetic approach is used to generate fuzzy rules in this approach. In [18], a binary particle swarm optimisation algorithm is proposed to select the significant heart dataset features. The selected subset is then given to support vector machine to classify and predict heart disease. The algorithm produced good accuracy maintaining the integrity of the specifications. In [19], a hybrid model of an opposition firefly with BAT algorithm (OFBAT) and rule-based fuzzy logic (RBFL) is introduced. A locality preserving projection (LPP) algorithm is used to select the features. After that, the rules for the fuzzy logic system are created. During the process of creating the rules, some important rules are generated using OFBAT algorithm. Furthermore, in the last, the fuzzy system is designed to perform classification inside it. In [20], an algorithm that uses BAT algorithm for feature selection and support vector machine (SVM) for classification is introduced. This model is proposed to improve the prediction of Alzheimer disease. In [21], a radial basis function network (RBFN) approach is proposed, which combines quantum computing algorithm and cloning operator to overcome local optima and global searching. This algorithm tried to improve learning for achieving excellent prediction accuracy. In [22], three butterfly approaches are presented to diagnose pneumonia disease. Three approaches: Basic, modified using levy flights and hybridisation of BFO with fuzzy membership function are proposed in this research.

Every algorithm reviewed above attempted to achieve high predictability and accuracy traditionally. Many have achieved good accuracy but fails in optimal feature selection and vice versa. The prediction time is also a concern while dealing with the big-sized datasets, which is an issue still for investigation.

The research in this work tends to develop a technique that accurately predicts heart diseases with optimised features in acceptable convergence time. A novel salp swarm algorithm (SSA) is proposed to predict heart diseases in this research. The SSA algorithm [23] is selected based on its merits, such as simplicity, efficiency, flexibility, ease of implementation, and requirement of fewer parameters to initialisation. The salp swarm algorithm is very stochastic in nature. The SSA has already proved its ability to solve any size of problems. SSA is used to select the useful features from the original heart dataset. It is found from the comparison that the proposed algorithm produces results with great accuracy in a very less amount of time. In addition to this, the algorithm also proved itself best in feature selection.

The manuscript is organised as: Section 2 introduces research methodology, which includes traditional and proposed SSA and the fitness function used. Section 3 discusses the results obtained by the proposed algorithm. A comparison of the results with other existing algorithms is also given in section 3. And, section 4 concludes paper with the future improvements in the algorithms. Acknowledgement and references follow section 4.

2. RESEARCH METHOD

No-free-lunch theorem [24] says that there is no single algorithm capable of solving all kinds of optimisation problems. One algorithm can solve some problems very well, but not all. Therefore, many opportunities are still there for developing and improving algorithms. It is possible that new developed algorithm can surpass the available algorithms. The no-free-lunch theorem (NFL) has inspired us to develop a feature selection tool by improving salp swarm algorithm (SSA). Our main goal is to introduce a very straightforward, uncomplicated and reasonable algorithm. The highlights of the developed algorithm include:

- This work introduces a nature-inspired algorithm known as the binary salp swarm algorithm for selecting an optimal subset of features. The introduced algorithm is applied to detect heart disease.
- The algorithm is evaluated using different metrics such as G-mean, F-measure, specificity, precision, sensitivity and accuracy.

– The new algorithm has been evaluated on two datasets related to heart disease.

2.1. Salp swarm algorithm

Mirjalili *et al.* have introduced the salp swarm algorithm in 2017 [23]. This algorithm mimics the food foraging behaviour of salps in an ocean. SSA mimics swarming and navigation behaviours of salps. Salps have a bottle-shaped transparent body. They form chains of the salps in the ocean for the direction-finding and foraging process. The chain of salps has followers (salps) and a leader salp. The leading salp leads the followers during direction-finding for searching a good food source in a multidimensional search space. The algorithm starts with some random solutions and works iteratively. Every salp explores and exploits the search space in iterations. The best-fit salp with its fitness is found at the end of every iteration. The position of the leader salp is changed using (1). It is based on the distance between a food source and the salp. Pseudocode of salp swarm algorithm (SSA) is given in Figure 1.

$$A_i^1 = \begin{cases} FD_i + r_1\big((u_i - l_i) \times r_2 + l_i\big) & r_3 \geq 0 \\ FD_i - r_1\big((u_i - l_i) \times r_2 + l_i\big) & r_3 < 0 \end{cases} \tag{1}$$

In which,

– A_i^1 - represents the position of the leader at *ith* position

– FD_i - represents food source position at *i*th position

– u_i *and* l_i - represents the upper and lower bounds at *i*th position

– r_1, r_2 *and* r_3 - represents some random numbers

– r_1 is used to control exploration and exploitation

$$r_1 = 2e\left(\frac{4i}{I}\right)^2 \tag{2}$$

In which,

– i - represents current iteration and

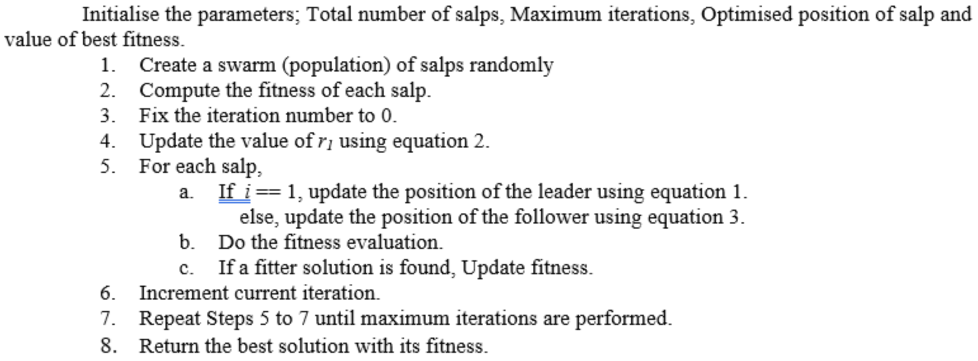
– I - represents total iterations

In (3) is used to update the follower’s position.

$$A_i^j = \frac{1}{2}\left(A_i^j + A_i^{j-1}\right) \tag{3}$$

In which,

– A_i^j - represents the current position of *j*th follower in *i*th dimension.



```
Initialise the parameters; Total number of salps, Maximum iterations, Optimised position of salp and
value of best fitness.
    1. Create a swarm (population) of salps randomly
    2. Compute the fitness of each salp.
    3. Fix the iteration number to 0.
    4. Update the value of $r_1$ using equation 2.
    5. For each salp,
        a. If $i$ == 1, update the position of the leader using equation 1.
           else, update the position of the follower using equation 3.
        b. Do the fitness evaluation.
        c. If a fitter solution is found, Update fitness.
    6. Increment current iteration.
    7. Repeat Steps 5 to 7 until maximum iterations are performed.
    8. Return the best solution with its fitness.
```

Figure 1. Pseudo code of the salp swarm algorithm (SSA)

2.2. Proposed salp swarm algorithm

2.2.1. Solution construction

In clustering problems, solution values are represented in binary [0 and 1]. We need to have a binary version of the algorithm to be used. So, a binary SSA is developed and used in this research. We use a vector to define a solution. The number of features in the real dataset is used to define the length of the solution vector. Each unit in the vector will have the value of either 1 or 0. The value "1" specifies the selection of the feature, and the value “0” says that the feature is not chosen. Mapping of continuous values into binary ones is done using (4). An example solution subset for eight attributes dataset is shown in the Figure 2. Here all "1" represents the chosen features, and "0" represents not chosen features.

$$D_{mn} = \begin{cases} 1, & if\ E_{mn} > 0.5 \\ 0, & otherwise \end{cases} \tag{4}$$

In which,
- m – represents the search agent
- n – represents the dimension
- D_{mn} – represents a discrete form of solution vector E.
- E_{mn} – represents the continuous position of the search agents.

0	1	1	0		0	0	1	1

Figure 2. A sample subset solution

2.2.2. Fitness function

The research problem considered in this research has two objectives. That is why it is called multi-objective. Two objectives of the problem are maximising accuracy and minimising features. So, the algorithm developed has to deal with the two objectives. The final output must have significantly fewer numbers of very accurate features. The SVM classifier is used for the classification of the features selected by the proposed salp swarm algorithm. Flowchart of new proposed salp swarm algorithm is given in Figure 3. So, SVM classifier accuracy is used as a fitness function to assess the performance in (5).

$$Fitness\ function = \alpha * err - rate(E) + \beta \left(\frac{|FS|}{|TF|}\right) \tag{5}$$

In which,
- E – represents error rate
- α & β - represents controlling constants, [$\alpha = [0, 1]$ & $\beta = 0.8$]
- $|FS|$ - represents the reduction in features
- $|TF|$ - represents the number of total features

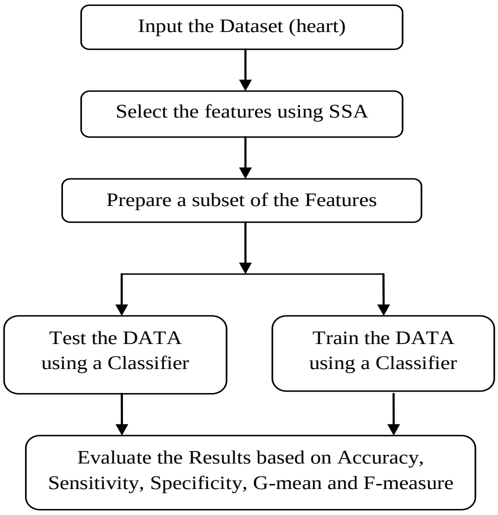


Figure 3. Flowchart of proposed salp swarm algorithm (SSA)

3. RESULTS AND DISCUSSION

3.1. Datasets

The proposed approach evaluated using two datasets that are given below [25], [26].

Dataset 1: This dataset is known as the heart dataset and is available on UCI and KABBLE. This dataset includes test results of 303 people with 76 features. The researchers used only 14 features out of 76 for their research.

Dataset 2: This dataset is known as heart-failure-clinical-records. This dataset contains records of 299 heart patients of faislabad institute cardiology, Pakistan. It contains the data of 194 male and 105 female patients with 12 different physiological measures. This dataset is available on UCI as well as Kaggle. The specifications of the datasets are given in Table 1.

3.2. Evaluation measures

Different metrics such as specificity, sensitivity, accuracy, precision and F-Measure given by Karalolis *et al.* in 2010 [27] are used for the evaluation. They use the following defined rules.

- True Positive (TP): The number of Heart Disease Patients classified correctly.
- True Negative (TN): The Number of patients not having heart disease correctly classified.
- False Positive (FP): The Number of healthy patients wrongly classified as heart disease patient.
- False Negative (FN): The number of healthy patients classified as heart disease patients.

We can define all measures as per the rules given above. Accuracy shows how the algorithm performs altogether. Accuracy is computed as a ratio of the correct predictions to total predictions. It is computed using (6).

$$Accuracy = \frac{TP+TN}{TP+FN+FP+TN} \times 100\% \tag{6}$$

Sensitivity measures the number of actual instances rightly predicted. It is computed using (7).

$$Sensitivity = \frac{TP}{TP+FN} * 100\% \tag{7}$$

Table 1. The specifications of dataset 1 and 2

Dataset 1		Dataset 2			
Feature	Description	Feature	Description	Measurement	Range
Age	Age in years	Age	Age of the patient	Years	[40,..., 95]
Sex	1 = male 0 = female	Anaemia	Decrease of hemoglobin	Boolean	0, 1
Cp	Chest pain type	High BP	If a patient has hypertension	Boolean	0, 1
Trestbps	Resting blood pressure	Creatinine phosphokinase	Level of the CPK enzyme in the blood	mcg/L	[23,..., 7861]
Chol	Serum cholesterol	Diabetes	If the patient has diabetes	Boolean	0, 1
Fbs	Fasting blood sugar	Ejection fraction	Percentage of blood leaving the heart at each contraction	Percentage	[14,..., 80]
Restecg	Resting cardio results	Sex	Woman or man	Binary	0, 1
Thalach	Maximum heart rate	Platelets	Platelets in the blood	kplatelets/mL	[25.01,. 850.00]
Exang	Exercise-induced angina	Ser creatinine	Level of creatinine in blood	mg/dL	[0.50,.., 9.40]
Oldpeak	ST depression-induced	Serum sodium	Level of sodium in the blood	mEq/L	[114,..., 148]
Slope	slope of the peak exercise	Smoking	If the patient smokes	Boolean	0, 1
Ca	Number of major vessels colored by fluoroscopy	Time	Follow-up period	Days	[4,...,285]
Thal	Heart condition	DEATH EVENT	If patient died in follow-up	Boolean	0, 1

Specificity measures number of unreal instances rightly predicted. It is computed using (8).

$$Specificity = \frac{TN}{TN+FP} * 100\% \tag{8}$$

The proportion of the rightly predicted actual instances is known as precision. It is computed using (9).

$$Precision = \frac{TP}{TP+FP} * 100\% \tag{9}$$

The accuracy is not a good measure when the number of false instances is increased. Other measures like G-mean and F-measure proposed by Kubat and Matwin [28], and Lewis and Gale [29], respectively, are used to evaluate the algorithm's performance. G-mean and F-measure are computed using (10) and (11).

$$G - mean = \sqrt{Sensitivity * Specificity} \tag{10}$$

$$F - measure = \frac{(\beta 2 + 1) * Precision * Sensitivity}{\beta 2 * Precision + Sensitivity} \tag{11}$$

3.3. Discussion

The experimental results obtained with both the datasets using the proposed algorithm are compared with those obtained with other stochastic clustering algorithms [30]-[40]. The algorithm is executed several times using a random population for testing the effectiveness. Experiments are performed on an Intel Core i3 seventh generation, 8-GB RAM system and coded with Matlab 12.

Table 2 presents results attained by all algorithms in terms of all performance measures using datasets 1 and 2. Dataset 1 and 2 are represented as set 1 and set 2 in the results table. For dataset 1, it is observed that, among all algorithms, the SSA algorithm performs well with the values of 98.75% accuracy, 98.15% sensitivity, 97.37% specificity, 99.39% precision, 97.76% G-mean, and 98.77% F-measure. For dataset 2, it can be seen that, among all algorithms, the SSA algorithm performs the best with the highest values of 98.46% accuracy, 98.32% sensitivity, 97.65% specificity, 99.14% precision, 97.98% G-mean, and 98.73% F-measure. The BGWO algorithm follows SSA with 97.45% accuracy, 98.08% sensitivity, 96.67% specificity, 98.29% precision, 97.37% G-mean, and 98.19% F-measure with dataset 1. For dataset 2, the BGWO algorithm produces 97.85% accuracy, 98.24% sensitivity, 96.86% specificity, 98.65% precision, 97.55% G-mean, and 98.44% F-measure. The plots in Figures 4 and 5 graphically depict a comparison among SSA and other algorithms in all measures. SSA performs best among all methods.

Table 2. The results obtained with dataset 1 and 2

Algorithm	Accuracy (%)		Sensitivity (%)		Specificity (%)		Precision (%)		G-mean (%)		F-measure (%)	
	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2	Set 1	Set 2
PSO [30]	95.7	97.2	96.8	96.8	95.6	96.6	96.0	96.3	96.2	96.2	96.4	96.5
GA [31]	93.9	94.2	94.3	94.4	94.6	94.7	95.9	95.8	94.4	94.4	95.1	95.1
FFA [32]	96.4	96.9	97.1	97.2	96.6	96.5	97.0	96.9	96.9	96.9	96.9	97.1
BGWO [33]	97.5	97.9	98.1	98.2	96.7	96.9	98.3	98.7	97.4	97.4	98.2	98.4
SSA	98.8	98.5	98.2	98.3	97.4	97.7	99.4	99.1	97.8	97.8	98.8	98.8
BCS [34]	68.0	71.3	69.3	72.7	70.2	73.9	70.6	84.2	69.8	69.8	69.9	77.9
BMVO [35]	96.1	97.3	95.9	97.5	96.8	97.9	96.9	97.9	96.3	96.3	96.4	97.7
SSPA [36]	80.9	83.9	81.9	84.4	82.3	84.9	83.2	85.9	82.1	82.1	82.6	85.1
RBMF [37]	89.5	92.3	91.2	93.9	90.9	92.9	91.9	93.9	91.1	91.1	91.5	93.9
WOA [38]	93.1	95.3	94.2	96.5	94.7	96.9	94.9	97.3	94.5	94.5	94.6	96.9
CALO [39]	82.4	85.8	83.7	86.9	83.9	87.8	84.3	89.5	83.8	83.8	83.9	88.2
MAHA [40]	78.0	81.2	79.2	82.3	79.9	83.2	79.9	83.9	79.5	79.5	79.6	83.2

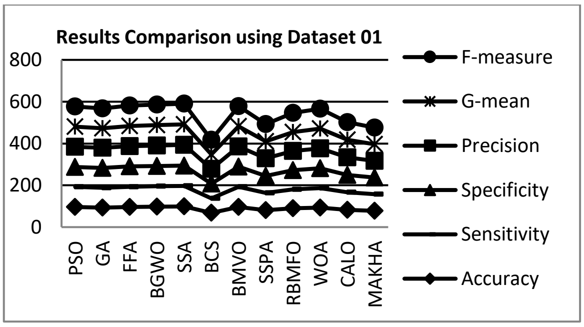


Figure 4. Comparison of results using dataset 1

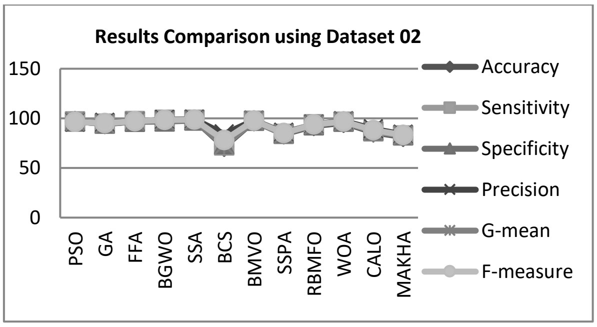


Figure 5. Comparison of results using dataset 2

4. CONCLUSION

In this paper, a novel salp swarm algorithm is proposed to diagnose the heart diseases. Firstly, the proposed SSA is applied for selecting the most informative and useful features from the two different heart datasets. Then, the SVM classifier is applied to predict the heart disease on the subset derived in the first step. The results attained are compared with the results attained from other nature inspired methods used for detecting heart diseases using a set of assessment criteria. The results attained have proved that the proposed SSA algorithm converges very quickly and produces accurate solutions. The algorithm also demonstrates high quality classification performance with less number of selected features. In the future, this algorithm will be applied to more real-life problems. Also, this algorithm will be implemented using a parallel approach to achieve the high performance.

ACKNOWLEDGEMENTS

We thank all WHO have supported us directly or indirectly in this research.

REFERENCES

[1] World Health Organization, (2020) *Cardiovascular Diseases*, WHO, Geneva, Switzerland, https://www.who.int/health-opics/cardiovascular-diseases/#tab=tab_1. Accessed 9 May 2021.

[2] D. Shah, S. Patel, and S. K. Bharti, "Heart Disease Prediction using Machine Learning Techniques," *SN Computer Science,* vol. 1, no. 6, 2020, doi: 10.1007/s42979-020-00365-y.

[3] H. F. Kareem, M. S. A.-Husieny, F. Y. Mohsen, E. A. Khalil, and Z. S. Hassan, "Evaluation of SVM performance in the detection of lung cancer in marked CT scan dataset," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 3, pp. 1731-1738, 2021, doi: 10.11591/ijeecs.v21.i3.pp1731-1738.

[4] K. M. Almustafa, "Prediction of heart disease and classifiers' sensitivity analysis," *BMC Bioinformatics*, vol. 21, no. 1, doi: 10.1186/s12859-020-03626-y.

[5] H. Jindal, S. Agrawal, R. Khera, R. Jain, and P. Nagrath, "Heart disease prediction using machine learning algorithms," *IOP Conference Series: Materials Science and Engineering*, vol. 1022, 2021, p. 012072, doi: 10.1088/1757-899x/1022/1/012072.

[6] G. Saranya and A. Pravin, "A comprehensive study on disease risk predictions in machine learning," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 4, pp. 4217-4225, 2020, doi: 10.11591/ijece.v10i4.pp4217-4225.

[7] H. David, Benjamin, and S. Belcy, "Heart disease prediction using data mining techniques," 09. *ictact journal on soft computing*, October 2018, vol. 09, no. 01, 2018, doi: 10.21917/ijsc.2018.0253.

[8] H. A. Lafta, Z. F. Hasan, and N. K. Ayoob, "Classification of medical datasets using back propagation neural network powered by genetic-based features elector," *International Journal of Electrical and Computer Engineering*, vol. 9, no. 2, pp. 1379-1384, 2020, doi: 10.11591/ijece.v9i2.pp.1379-1384.

[9] P. K. Sahoo and P. Jeripothula, "Heart Failure Prediction Using Machine Learning Techniques," *SSRN Electronic Journal*, 2020, doi: 10.2139/ssrn.3759562.

[10] L. Yang *et al.*, "Study of cardiovascular disease prediction model based on random forest in eastern China," *Scientific Reports*, vol. 10, no. 1, 2020, doi: 10.1038/s41598-020-62133-5.

[11] A. Ibrahem, R. A. Ahmed, M. A. Mohialden, and Y. M. Ali, "Efficient method for breast cancer classification based on ensemble offending tree and naïve Bayes," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 18, no. 2, pp. 1074-1080, 2020, doi: 10.11591/ijeecs.v18.i2.pp1074-1080.

[12] M. F. Darmawan, A. F. Z. Abidin, S. Kasim, T. Sutikno, and R. Budiarto, "Random forest age estimation model based on length of left hand bone for asian population," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 1, pp. 549-558, 2020, doi: 10.11591/ijece.v10i1.pp549-558.

[13] Q. Li *et al.*, "An Enhanced Grey Wolf Optimization Based Feature Selection Wrapped Kernel Extreme Learning Machine for Medical Diagnosis," *Computational and Mathematical Methods in Medicine*, pp. 1-15, 2017, doi: 10.1155/2017/9512741.

[14] N. C. Long, P. Meesad, and H. Unger, "A highly accurate firefly based algorithm for heart disease prediction," *Expert Systems with Applications*, vol. 42, no. 21, pp. 8221-8231, 2015, doi: 1016/j.eswa.2015.06.024.

[15] T. Vivekanandan, and N. C. S. N. Iyengar, "Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease," *Computers in Biology and Medicine*, vol. 90, pp. 125-136, 2017, doi: 10.1016/j.compbiomed.2017.09.011.

[16] A. M. Usman, U. K. Yusof, and S. Naim, "Cuckoo inspired algorithms for feature selection in heart disease prediction," *International Journal of Advances in Intelligent Informatics*, vol. 4, no. 2, p. 95, 2018, doi: 10.26555/ijain.v4i2.245.

[17] A. K. Paul, P. C. Shill, M. R. I. Rabin, and M. A. H. Akhand, "Genetic algorithm-based fuzzy decision support system for the diagnosis of heart disease," in *Informatics, Electronics and Vision (ICIEV), 5th International Conference*, IEEE, 2016, pp. 145-150.

[18] M. N. Elbedwehy, H. M. Zawbaa, N. Ghali, and A. E. Hassanien, "Detection of Heart Disease using Binary Particle Swarm Optimization," *Federated Conference on Computer Science and Information Systems*, FedCSIS 2012, pp. 177-182.

[19] G. T. Reddy and N. Khare, "An Efficient System for Heart Disease Prediction Using Hybrid OFBAT with Rule-Based Fuzzy Logic Model*," Journal of Circuits, Systems and Computers,* vol. 26, no. 04, p. 1750061, 2016, doi: 10.1142/s021812661750061x.

[20] A. T. Shereen and G. Wafaa, "A new model for early diagnosis of Alzheimer's disease based on BAT-SVM classifier," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 2, pp. 759-766, April 2021, doi: 10.11591/eei.v10i2.2714.

[21] L. E. Aik, T. W. Hong, and A. K. Junoh, "An improved radial basis function networks based on quantum evolutionary algorithm for training nonlinear datasets," *IAES International Journal of Artificial Intelligence*, vol. 8 no. 2, pp. 120-131, 2019, doi: 10.11591/ijai.v8.i2.pp120-131.

[22] I. K. Baydaa and Y. A. Manar, "Pneumonia detection using butterfly optimisation and hybrid butterfly optimisation algorithm," *Bulletin of Electrical Engineering and Informatics*, vol. 10, no. 4, pp. 1262-1270, August 2021, doi: 10.11591/eei.v10i4.2872.

[23] S. Mirjalili, A. H. Gandomi, S. Z. Mirjalili, S. Saremi, H. Faris, and S. M. Mirjalili, "Salp Swarm Algorithm: A bio-inspired optimiser for engineering design problems," *Advances in Engineering Software*, vol. 114, pp. 163-191, 2017, doi: 10.1016/j.advengsoft.2017.07.002.

[24] D. H. Wolpert, and, W. G. Macready, "No free lunch theorems for optimization," *IEEE Transactions on Evolutionary Computation*, vol. 1, no. 1, pp. 67-82, 1997, doi: 10.1109/4235.585893.

[25] https://www.kaggle.com/andrewmvd/heart-failure-clinical-data. Accessed 09 January 2019.

[26] https://archive.ics.uci.edu/ml/machine-learning-databases/heart-disease/. Accessed 9 January 2019.

[27] M. A. Karaolis, J. A. Moutiris, D. Hadjipanayi, and C. S. Pattichis, "Assessment of the Risk Factors of Coronary Heart Events Based on Data Mining With Decision Trees," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 3, pp. 559-566, doi: 10.1109/titb.2009.2038906.

[28] M. Kubat and S. Matwin, "Addressing the curse of imbalanced training sets: one-sided selection," *Proceedings of the 14th International Conference on Machine Learning*, 2000.

[29] D. D. Lewis and W. A. Gale, "A Sequential Algorithm for Training Text Classifiers," In: *Croft B.W., van Rijsbergen C.J. (eds) SIGIR '94*. Springer, London, 1994, doi: 10.1007/978-1-4471-2099-5_1.

[30] M. N. Elbedwehy, H. M. Zawbaa, N. Ghali, and A. E. Hassanien, "Detection of heart disease using binary particle swarm optimization," *Federated Conference on Computer Science and Information Systems* (FedCSIS), 2012, pp. 177-182.

[31] D. Chicco and G. Jurman, “Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone,” *BMC Medical Informatics and Decision Making*, vol. 20, no. 1, doi: 10.1186/s12911-020-1023-5.

[32] N. Mülayim and A. Alaybeyoğlu, “Designing of an expert system based on firefly algorithm for diagnosis of Heart Disease,” *20th National Biomedical Engineering Meeting* (BIYOMUT), pp. 1-4, 2016, doi: 10.1109/BIYOMUT.2016.7849405.

[33] E. Emary, H. M. Zawbaa, and A. E. Hassanien, “Binary Gray Wolf Optimization Approaches for Feature Selection,” *Neurocomputing*, 2015, doi: 10.1016/j.neucom.2015.06.083.

[34] L. A. M. Pereiraet *et al.*, “A Binary Cuckoo Search and Its Application for Feature Selection,” *Studies in Computational Intelligence*, pp. 141-154, 2013, doi:10.1007/978-3-319-02141-67.

[35] R. H. Hans and H. Kaur, “Binary Multi-Verse Optimization (BMVO) Approaches for Feature Selection,” *International Journal of Interactive Multimedia and Artificial Intelligence*, 2019, doi: 10.9781/ijimai.2019.07.004.

[36] H. M. Zawbaa, E. Emary, A. E. Hassanien, and B. Parv, “A wrapper approach for feature selection based on swarm optimisation algorithm inspired from the behavior of social-spiders,” *7th International Conference of Soft Computing and Pattern Recognition* (SoCPaR), 2015, pp. 25-30, doi: 10.1109/SOCPAR.2015.7492776.

[37] R. A. Khurma, I. Aljarah, and A. Sharieh, “Rank Based Moth Flame optimisation for Feature Selection in the Medical Application,” *IEEE Congress on Evolutionary Computation (CEC)*, pp. 1-8, 2020, doi: 10.1109/CEC48606.2020.9185498.

[38] M. Sharawi, H. M. Zawbaa, and E. Emary, “Feature selection approach based on whale optimisation algorithm,” *Ninth International Conference on Advanced Computational Intelligence (ICACI)*, 2017, pp. 163-168, doi: 10.1109/ICACI.2017.7974502.

[39] H. M. Zawbaa, E. Emary, and C. Grosan, “Feature Selection via Chaotic Antlion Optimization,” *PLoS ONE*, vol. 11, no. 3, p. e0150652, 2016, doi: 10.1371/journal. pone.0150652.

[40] A. I. Hafez, A. E. Hassanien, H. M. Zawbaa, and E. Emary, “Hybrid Monkey Algorithm with Krill Herd Algorithm optimisation for feature selection,” *11th International Computer Engineering Conference (ICENCO)*, 2015, pp. 273-277, doi: 10.1109/ICENCO.2015.7416361.

BIOGRAPHIES OF AUTHORS

Dr. Nitesh Sureja is currently working as a professor at Krishna School of Emerging Technology and Applied Research, KPGU, Vadodara, Gujarat, INDIA. He has about 22 years of experience. He is a life member of ISTE, IEANG and IMETE. He is currently guiding two PhD students. He can be contacted at email: nmsureja@gmail.com.

Dr. Bharat Chawda is currently working as a Lecturer in the Department of Computer Engineering, B. & B. Institute of Technology, Vallabh Vidyanagar. He has published books related to subjects Operating Systems and Database Management System. He has published 10 International Journal papers. He is a life member of ISTE, The Institutions of Engineers (India). He can be contacted at email: bharat.bbit@gmail.com.

Dr. Avani Vasant is working as a Professor and Head of the Department of Computer Science Engineering at Krishna School of Emerging Technology and Applied Research, KPGU, Vadodara. She holds PhD in Computer Engineering. As an academic and researcher, she holds 19+ years of experience, actively participate in research projects, had lectured in different domains. She can be contacted at email: avani.vasant@gmail.com.