Heart Disease Prediction Using Glowworm Swarm Optimization and Support Vector Machine Classifier

¹O. Bhaskaru, ²M. Sreedevi

¹CSE Department, Koneru Lakshmaiah Education Foundation, Green Fields, Guntur District, Vaddeswaram, Andhra, Pradesh, India obhaskaru@gmail.com
²CSE Department, Koneru Lakshmaiah Education Foundation, Green Fields, Guntur District, Vaddeswaram, Andhra, Pradesh, India

msreedevi 27@kluniversity.in

Received: June 18, 2021. Revised: November 17, 2021. Accepted: December 10, 2021. Published: January 3, 2022.

Abstract: At present, health disorder is growing day by way of the day due to existence lifestyle, hereditary. Particularly, heart disease has ended up greater Heart disorder prognosis frequent these days. technique is very quintessential and integral trouble for the patient's health. Besides, it will help out to limit disorder to a larger distinctive level. The role of using strategy like machine learning and algorithm such as heart disease diagnosis using Data Mining(DM) techniques is very significant. In the previous system, the Fuzzy Extreme Learning Machine (FELM) was proposed to predict heart disease, ensuring an accurate and timely diagnosis. However, it only achieves 87.14 % of accuracy. To improve the classification accuracy, the proposed system designed an Improved Step Adjustment based Glowworm Swarm Optimization Algorithm with Weighted Feature based Support Vector Machine (ISAGSO-

I. INTRODUCTION

Heart disorder is a disorder that impacts the operation of the coronary heart [1-2]. There are mixtures of essentials that enlarge the possibility of Heart syndrome. These days, heart disease became the major cause for the occurrence of deaths worldwide [3-6]. The record of World Health Organization (WHO) has anticipated that, for every 12 months there takes place 12 million deaths due to coronary heart diseases. During the period 2008, 17.3 million human being die due to Heart syndrome. More than 80% of deaths in the globe are owing to the fact of Heart syndrome. Nearly 23.6 million human beings would die due to heart failure by 2030, according to WHO reports.

Heart disorders consist of a number of diseases such as chronic blood vessels, such as diseases associated with coronary arteries; issues concerning heart pulses arising when problems increase in regulating heartbeat impulses (heart arrhythmias) and a sort of birth-related coronary heart defects (congenital coronary heart defects), including irregularities. The WFSVM) for Heart disease diagnosis. This proposed venture utilizes the dataset of heart disease for input. Using the Improved Step Adjustment based Glowworm Swarm Optimization Algorithm (ISAGSO) to enhance the true positive rate, optimal features are then selected. Finally, with the aid of the Weighted Feature based Support Vector Machine (WFSVM) classifier, classification is carried out relying selected features. In the proposed method, better performance obtained and that is validated through the experimental results in terms of precision, accuracy, recall and f-measures

Keywords: Heart disease, feature selection, Glowworm Swarm Optimization (GSO) and Weighted Feature based Support Vector Machine (WFSVM) classifier.

time period of coronary heart sickness is occasionally used and interchangeably it can be termed as cardiovascular disease. Cardiovascular Disease (CVD) typically refers to blocked blood vessels that increases the risk and build-up the deposits of fatty acid inside arteries. This puts the life to danger that it can lead to a coronary heart attacks caused by sudden block of heart muscle (Myocardial infarctions), chest ache due to blood flow restriction (angina) or stroke due to improper blood pumping. Recently computer technology and Machine learning technology getting into strategies to strengthen software programs to aid doctors in deciding coronary heart disorder in the early phase [7-9]. The analysis of coronary heart sickness relies upon scientific data. Heart sickness diagnosis machines can aid clinical specialists in detecting coronary heart sickness. It objectives to take advantage of quite a number of records mining strategies on scientific statistics set to assisting heart disease prediction.

Researchers rely on several statistics mining methods to extract the necessary features that are reachable to assist the experts or medical practitioners and provide scientific decision making for diagnosis of coronary heart disease. Decision Tree, k-nearest and Naïve Bayes[10] are regularly often used approaches. Bagging algorithm, kernel density estimation for differentiation of cardiac controls, sequential minimal optimization and neural networks, straight kernel self-organizing map and SVM (Support Vector Machine) are other exceptional classificationbased methods usedOrganization of the paper is as follows: Section II enlists the survey of methods available in forecasting heart disease. Section three offers a description of the Improved Step Adjustment primarily based Glowworm Swarm Optimization Algorithm with Weighted Feature-based Support Vector Machine (ISAGSO-WFSVM) for coronary heart disorder classification. Section 4 deals with the comparison of performance attainment of the proposed method and are evaluated with the prevailing methods.

II. RELATED WORK

By exploring DM techniques and Machine Learning (ML) algorithms in the infirmary, various studies related to diagnostic systems were identified. In his paper, Liu et al (2017) guided the determination of coronary disease through a method fixated on the RelieF and Rough Set (RFRS) strategy which considered as the hybrid approach. Two subsystems are included in the researcher's proposed approach, example. RFRS feature selection for and classification methods in consort with an ensemble classifier. Three phases were integrated in initial system. This performs (i) Discretization of data for transformation, (ii) the ReliefF algorithm used for feature extraction, and (iii) the heuristic Rough Set reduction algorithm for feature reduction. Upon following framework, the classifier named ensemble classifier relying C4.5 model is implied. The dataset incorporated in the experiments are Statlog (Heart) dataset and it is collected from the UCI database. Endowing to a jackknife cross-validation module, a maximum of 92.59 percent classification precision can be obtained. The designed framework has improved comparable performance to the other classification techniques previously documented and is presented in the reports of this paper[11].

Akgül et al (2019) suggests a model named Hybrid model-based ANN – GA approach for forcasting Heart Disease in his paper. In this examine, the parameters of Artificial Neural Network (ANN) are optimized by means of GA to diagnose heart disease initially. The Artificial Neural Network (ANN) is then offered in hybrid combination with the Genetic Algorithm (GA) to achieve precise of classification. In the end, to illustrate the feasibility of the conceptual framework, the 'Cleveland' dataset is chosen for this medical experiment taken from the UCI (University of California Irvine) repository. The proposed hybrid ANN - GA method, which outflanks other algorithms such as Naive Bayes, C4.5 and K-Nearest Neighbor algorithm[12], reports good performance in terms of accuracy rate, precision, F-measure and recall.

Hasan et al (2018) structured a diagnosis system to make a diagnosis of the heart condition by changing the medical parameters into numerical representation of the sufferers. This system is equipment actualized utilizing FPGA stage pack (Spartan-3A DSP 3400A) used by the doctors. The proposed framework uses two categories of database for the classification of heart conditions as ordinary and strange i.e. abnormal. The first category of database is the web source utilized from dataset repository of machine learning of UCI. The second sort are the collected real database (CD). There are 13 clinical components in the database gathered and that components are effective in heart illness prediction. Using the obtained results, it was reported that the propounded method namely Numeral Virtual Generating random Access Memory - Weighted Neural classifier (NVG-RAM) shows highest percent of accuracy. The presentation of this classifier using a CD was examined. Besides, the designed algorithm is viewed as a decent one by the researcher in [13] to diagnosis multiclass sickness of heart that shows up 88% of accuracy.

Tarle and Jena (2017) proposes an Artificial Neural Network (ANN) heart disease diagnosis . The neural biological functions associated in the brain can be compared with the work done by ANN. That is, ANN is assumed to be the simulation of human brain. ANNs are trained in solving the mathematical problems which found to be un-quantifiable or easily calculated by the primitive techniques. The most important process to select relevant attributes from the huge data set is Feature selection. The extracted attributes form the principal components which are assigned as input given to ANN. Neural net training is used by ANN for best results using back propagation algorithm. Normally, ANN utilizes the human brain observation, which makes it more beneficial for producing results with high clarity. Hence the researcher designed this paper as an aid for diagnosing epidemic and artery diseases [14].

Mai Shouman, et al (2012) recommends k-means clustering integration with the decision tree model for optimizing disease predictions. Different centroid selection techniques for k-means clustering are suggested in this work to improve reliability. It receives data sets from the Cleveland Clinic Foundation for Heart Disease. 13 input attributes were extracted from that data set. Different centroids selections methods and cluster approach are used to calculate sensitivity, specificity, and accuracy. Ten iterations were sought using random attribute and random row methods using in estimating the average and best results. Comparatively incorporating Kmeans clustering with decision tree delivers better with the collected data set greater accuracy in diagnosing heart disease than conventional decision tree approaches. The proposed method also yields higher accuracy than paging algorithm. 83.9% of accuracy was achieved by enabler method with clusters adopted [15].

III. METHODOLOGY OF PROPOSED RESEARCH WORK

The proposed framework structured an Improved Adjustment Step based Glowworm Swarm Optimization Algorithm with Weighted Feature based Support Vector Machine (ISAGSO-WFSVM) approach for coronary illness forecast. At first the heart disease dataset is taken as input and optimal feature selection is finished by utilizing Improved Glowworm Swarm Adjustment based Step Optimization Algorithm (ISAGSO). As indicated by the selected features heart disease characterization is performed with the assistance of Weighted Feature based Support Vector Machine (WFSVM) classifier. The proposed methodology can be seen as a flow diagram as depicted below (figure 1).



Figure 1: Proposed Work Flow

A. DATASET

The Z-Alizadeh Sani dataset is taken as an input. There are 53 attributes and 1 class attribute for the data set obtained for the heart disease prediction. The class attribute implies the nature of a person's heart disease.

FEATURE SELECTION USING IMPROVED STEP ADJUSTMENT BASED GLOWWORM SWARM OPTIMIZATION ALGORITHM:

Feature selection methods provide a way to minimize computational time, boost predictive efficiency, and deeper knowledge of machine learning data [16-18]. In this research proposal, Improved Phase Adjustment based Glowworm Swarm Optimization Algorithm (ISAGSO) is implemented for the selection of features[19]. A relatively novel algorithm is the Glowworm Swarm Optimization (GSO), an optimized and nature-inspired algorithm. The flashing performance of glow worms inspired this algorithm. The glowworm carries its own luminescence amount called luciferin. Initially glowworms are deployed randomly in the solution space for GSO. All the glowworms are allotted random position inside the given search space and they signify the solution of objective function. The agent's present position is related to with luciferin level. The glowworm which has improved position is brighter, this means the better solution. All glowworms can be attracted towards the neighborhood set who is brighter that its own within its local decision domain and probably make a move towards it using a probabilistic mechanism [20]. Thus, the position is adjusted based on the neighborhood set selected. The local-decision domain will be enlarged when the neighbor density becomes low. That depicts that the density of neighbors of glowworms disturbs the decision radius. The density change determines the size of localdecision domain that it enlarges and finds more neighbors when it is low or else the splitting of swarm into smaller groups will be reduced. The algorithm is repeated until it satisfies the end condition. The majority of the agents gather around the glowworms with brighter luminance. In brief, the GSO involves six phases as discussed below.

Glowworms' initialization: The glowworms are considered as a feature (or) attributes in this phase. These are randomly distributed in the prescribed fitness function space initially. The equivalent measure of luciferin is contained in glowworms. In addition, it is set 1 for the current iteration. In this proposed work, classification accuracy is considered as fitness value.

Volume 16, 2022

Luciferin-update phase: Fitness value which is also termed as accuracy and previous luciferin value is the major terms that are utilized by luciferin update. Rule of this phase is given by

$$l_i(t+1) = (1-\rho)l_i(t) + \gamma Fitness x_i(t+1) \quad (1)$$

The functions in the rules are mentioned as follows; $l_i(t)$ is the luminance amount of glowworm termed as luciferin (feature) i at time t, luciferin decay constant is given by ρ that ranges between $0 < \rho < 1$, γ represents the luciferin enhancement constant, position of glowworm is given by $x_i(t+1)\epsilon R^M$ and fitness at the location i at the time t + 1 is given by *Fitness* $x_i(t+1)$.

Neighborhood-Select Phase: The brighter one within the neighborhood set of glowworm (feature) $N_i(t)$ with i at t iteration illustrated as below.

$$N_{i}(t) = \{j: \left| \left| d_{ij}(t) < r_{d}^{i}(t) \right| \right|; \ l_{i}(t) < l_{j}(t)$$
(2)

 $r_d^i(t)$ is the dynamic decision domain radius or also termed as local-decision domain, and in time t, $d_{i,j}(t)$ implies the Euclidean interval between the characteristics i and j.

Moving Probability-Computer Phase: A probability rule is used by the glowworm to move on the way to other glowworms possessing greater luciferin level. This can be represented by $p_{ij}(t)$ of glowworm (feature) given as

$$p_{ij}(t) = \frac{l_j(t) - l_i(t)}{\sum_{k \in N_i(t)} l_k(t) - l_i(t)}$$
(3)

The above-mentioned formula represents the probability of glowworm that move towards neighbor, i.e. i to j.

Movement Phase: The discrete-time model of the movement of glowworm (feature) i activity that selects the glowworm $\in N_i(t)$ with $p_{ii}(t)$ is given as

$$x_{i}(t+1) = x_{i}(t) + s(t) \left(\frac{x_{j}(t) - x_{i}(t)}{||x_{j}(t) - x_{i}(t)||} \right)$$
(4)

Here, step size denoted by S and $\|.\|$ is the Euclidean norm operator

Decision Radius Update Phase: Decision radius of glowworm i changes for each update and it is given as follows:

$$r_{d}^{j}(t+1) = \min \{r_{s}, \max\{0, r_{d}^{j}(t) + \beta(n_{t} - |N_{i}(t)|)\}\}$$
(5)

Here, β is considered a constant, r_s signifies the sensory radius of glowworm i, and n_t is a limit that in which neighbor number is controlled

Improved Step Adjustment based GSO (**ISAGSO**): The test results can be restricted by two aspects with the constant step size in the glowworm swarm optimization algorithm. Smaller step size influences the iterations proceeded in the algorithm. Larger step size skips the global optimum in searching very easily. An improved step size adjustment was designed and the formula is given as below

$$c = \frac{1}{t_{max}} \ln \left(\frac{S_{min}}{S_{max}}\right) \quad (6)$$
$$s(t) = S_{max} e^{c.g_t} \quad (7)$$

In the above formula, current iteration number is represented as g_t and the maximum number of converged iterations is given by t_{max} . S_{min} , S_{max} expresses the minimum value and maximum value of s respectively. On substitution of equation (7) in equation (4).

$$\begin{aligned} x_i(t+1) &= x_i(t) + \\ S_{max} e^{c.g_t} \left(\frac{x_j(t) - x_i(t)}{||x_i(t) - x_i(t)||} \right) \end{aligned} \tag{8}$$

Algorithm 1: Improved Step Adjustment based Glowworm Swarm Optimization Algorithm (ISAGSO)

- 1. Initialize Dimensions = m Dimensions
- 2. Declare glowworms number (Features)
- 3. Assign G=1 where G is generation
- 4. Let the step size be s
- 5. Glowworms (features) are presented randomly standardisation in the search space
- 6. While (G < max generation) do
- 7. for every glowworm (features) i do
- 8. Update according to Luciferin (equation 1)
- 9. Set of neighbors should be confirmed according to (2)
- 10. Find the movement choice according to (3)
- 11. Glowworm i shifts toward j based on the Equation (8);
- 12. Find succeeding position x(t+1) of each glowworm and the succeeding decision radius $r_d(t+1)$
- 13. Neighborhood range update based on Equation (5)
- 14. Apprise position
- 15. Get selected features
- 16. End for
- 17. End while

C.3.3 Classification Using Weighted Feature Based Support Vector Machines

In this proposed research work, Weighted Feature based Support Vector Machines (WFSVM) utilized for classification of heart disease using the selected features. Support Vector Machine (SVM) imprinted its best performance in classification approaches [21-24] is one of the popular supervised machine learning. It is far more sophisticated and it may be utilized as a discriminative classifier which demonstrated to be more appropriate and accurate than any other classification models [25]. The Structural Risk Minimization (SRM) principle is good for controlling the generalization ability and so used in SVM that involves finding optimal separating hyper-plane. This makes it more accurate classifier which is adopted by most of the applications.

The finding of a discriminant function f(x) such that $y_i = f(x_i)$ has been the two-class classification problem in general for the N data samples (attributes) (x_1, y_1) ... (x_i, y_i) ... (x_N, y_N) . Moreover f(x) =sgn($w \cdot x - b$) is given as linear discriminant function were $w \cdot x - b = 0$ is assigned for separating hyperplane in the considered data space. The maximum separating margin possessing hyperplane can be chose by discriminant function in terms of two classes. The composition of the linear discriminant is given as

$$f(\mathbf{x}) = \operatorname{sgn}(\sum_{i=1}^{l} \alpha_i y_i (x_i \cdot \mathbf{x} - \mathbf{b}) \quad (9)$$

in which, training records is counted as l, the training data associated label is given as $y_i \in \{-1, +1\}$ with the range $0 \le \alpha_i \le C$ (constant C > 0), and x_i represents the support vectors. Considering two classes separated by the surface goes not linear, then the data points are brought into linear separable by transforming data points to higher dimensional space. The nonlinear discriminant function of SVM is:

$$f(x) = \text{sgn} \left(\sum_{i=1}^{l} \alpha_i y_i \ K(x_i \cdot x) + b \right) - \dots - (10)$$

Where $K(x_i.x)$ is used in transformation of data points which is also called as the kernel function.All the features extracted from the training data sets will be treated alike and it can be found rom the Kernel function format $K(x_i.x)$ of SVM which has an impact on accuracy. The importance of diverse features can be considered by adding weightages to the Kernel function. In generic, new kernel function can be formulated as $K(\omega x_i.\omega x)$, here ω represents the weights of extracted features of training dataset. The formulation of nonlinear discriminant function is given as,

$$f(\mathbf{x}) = \operatorname{sgn}(\sum_{i=1}^{l} \alpha_i y_i \operatorname{K}(\omega x_i . \omega x) + b) - \dots - (11)$$

This is the enhanced kernel function which is influenced by the weights of the dataset and it is liberated from particular kernel functions. The suitable kernel functions from training data sets are calculated by the system through rough set theory and it varies for different applications. The proposed system works one the basis of the algorithm mentioned below. This algorithm adopts rough set theory for feature ranking and their weight calculation respectively. After the completion of ranking process, the features possessing null will be considered as of less importance and is neglected.

Algorithm 2: Measurement of individual feature weights

Input : *D*, *Heart disease dataset Output*: ω, weight vectortput: Find out all the reducts of D using rough sets; $N_{feature} \leftarrow number of features in D;$ $N_{reduct} \leftarrow number of reducts of D;$ for $(i \leftarrow 0 \text{ to } N_{feature})$ do $\omega i \leftarrow 0;$ end for for $(i \leftarrow 0 \text{ to } N_{feature}) do$ for $(j \leftarrow 0 \text{ to } N_{reduct}) do$ if (feature i in the jthreduct R_i) then $m \leftarrow number of features in R_i$; $\omega i \leftarrow \omega i + 1/m$; end if end for end for *Scale the values of feature weights into the interval* [0, 100]

IV. EXPERIMENTAL RESULTS

The experimental assessment of the proposed system is executed in the MATLAB. Z-Alizadeh Sani dataset is taken from the UCI AI archive as benchmark issues to assess the execution of proposed and existing strategy. Table 1 demonstrates the dataset attributes of the Z-Alizadeh Sani. It consists of 54 features. The performance of the proposed Improved Step Adjustment based Glowworm Swarm Optimization Algorithm with Weighted Feature based Support Vector Machine (ISAGSO-WFSVM) and existing methods such as J48, Naïve Bayes (NB), Hybrid Differential Evolution based Fuzzy Neural Network (HDEFNN) [26] and Fuzzy Extreme Learning Machine (FELM) [27] is compared accuracy, precision, recall and f-measure.

Name of the Dataset	Z-Alizadeh Sani dataset		
Instances count	303		
Feature count	54		
Group count	4 (Demographic, Symptom and examination ECG, Echo feature)		
Classification	Each patient could be in two possible categories yes or no		

Table 1: Data Set Characteristics

Performance metrics

A. Accuracy

In image processing, accuracy is the most intuitive measure of success that occurs in the correctly classified instance of occurrence and is simply the dividend of expected accurate values to the total observed measurements.

Accuracy =
$$\frac{TP+TN}{TP+FP+FN+TN}$$
 (12)

Where,

TP - True Positive

FN - False Negative

FP - False Positive

TN- True Negative

B. Precision

Precision defines the relevance of the results and is given by the ratio positive observations predicted correctly to the positive observations predicted in total.

Precision
$$= \frac{\text{TP}}{\text{TP}+\text{FP}}$$
 (13)

C. Recall

It is the percentage of positive measures receive adequate to entire observations comes under the "YES" class

Recall =
$$\frac{TP}{TP+FN}$$
 (14)

D. F-measure

The F1 Score is also a measure of test accuracy and is characterized by the weighted average of accuracy

and recall. This score thus takes both false positives and false negatives into account.

F-measure =
$$2*\frac{(Recall*Precision)}{(Recall+Precision)}$$
 (15)

The proposed approach ISAGSO-WFSVM is compared with four algorithms. The J48, Naïve Bayes (NB), HDEFNN and FELM algorithms are taken into consideration. Table 2 shows the performance comparison for accuracy, precision, recall and f-measure.

Algorith m	Performance metric in (%)				
	Accura cy	Precisio n	Reca ll	F- measu re	
J48	81.42	58.5	86.5	71	
Naive Bayes	82.52	64.1	79.8	78.4	
HDEFN N	84.46	65	87.1	88.8	
FELM	87.14	71	87.8	89.9	
ISAGSO - WFSVM	92	79	91	94	

 Table 2: Performance Analysis

The proposed ISAGSO-WFSVM approach is compared with the existing HDEFNN and FELM approaches in relationship of accuracy and precision. The accuracy and precision performance are shown in figure 2. Adaptive features are selected in this proposed research work by using the Glowworm Swarm Optimization Algorithm based on Improved Step Adjustment (ISAGSO). This strengthens the reliability of precision and accuracy. From the analysis, it can be inferred that 92 percent accuracy is achieved by the suggested technique, whereas other techniques such as J48, Naïve Bayes (NB), HDEFNN and FELM achieves 81.42 %, 82.52%, 84.46% and 87.14% respectively. And also precision of the proposed system is 79 % when existing method such as J48, NB, HDEFNN and FELM achieves 58.5 %, 64.1 %, 65% and 71% respectively. The proposed ISAGSO-WFSVM approach is compared with the existing HDEFNN and FELM approaches in relationship of accuracy and precision. The accuracy and precision performance are shown in figure 2. Adaptive features are selected in this proposed research work by using the Glowworm Swarm

Optimization Algorithm based on Improved Step Adjustment (ISAGSO).



Figure 2: Accuracy and precision comparison

This strengthens the reliability of precision and accuracy. From the analysis, it can be inferred that 92 percent accuracy is achieved by the suggested technique, whereas other techniques such as J48, Naïve Bayes (NB), HDEFNN and FELM achieves 81.42 %, 82.52%, 84.46% and 87.14% respectively. And also precision of the proposed system is 79 % when existing method such as J48, NB, HDEFNN and FELM achieves 58.5 %, 64.1 %, 65% and 71% respectively.



Figure 3: Recall and f-measure comparison

Performance of Recall and f-measure of the proposed ISAGSO-WFSVM and existing HDEFNN and FELM approaches are shown in figure 3. Based on the selected characteristics, the heart disease is categorized using Weighted Feature-based Support Vector Machines in this proposed research work (WFSVM). It correctly predicted positive observations. From the graph, it tends to be inferred that 91% of recall recorded for proposed system, whereas other methods such as J48, NB, HDEFNN and FELM attains 86.5 %, 79.8%, 87.1% and 87.8% respectively. And also, f-measure of the proposed

system is 94 % when existing method such as J48, NB, HDEFNN and FELM achieves 71 %, 78.4 %, 88.8% and 89.9% respectively.

V. CONCLUSION

One of the key causes of loss of life globally is heart disease and early prediction of heart disease is significant. The Improved Step Adjustment-based Glowworm Swarm Optimization Algorithm with Weighted Feature-based Support Vector Machine (ISAGSO-WFSVM) is designed for the prediction of heart disease in this proposed approach. For optimal feature selection, the Improved Step Adjustment based Glowworm Swarm Optimization Algorithm is presented to improve predictive performance. The classification is carried out according to the features extracted by means of the Weighted Feature-based Support Vector Machine (WFSVM) classifier. In terms of precision, accuracy, recall and fmeasurement, the outcomes of the experiment revealed that higher performance was achieved as comparison to the present system.

REFERENCES

- [1] Soni, J., Ansari, U., Sharma, D., & Soni, S. (2011). Predictive data mining for medical diagnosis: An overview of heart disease prediction. *International Journal of Computer* Applications, 17(8), 43-48.
- [2] Bharti, S., & Singh, S. N. (2015, May). Analytical study of heart disease prediction comparing with different algorithms. In *International Conference on Computing, Communication & Automation* (pp. 78-82). IEEE.
- [3] Dbritto, R., Srinivasaraghavan, A., & Joseph, V. (2016). Comparative analysis of accuracy on heart disease prediction using classification methods. *International Journal* of *Applied Information Systems*, 11(2), 22-25.
- [4] Dangare, C. S., & Apte, S. S. (2012). Improved study of heart disease prediction system using data mining classification techniques. *International Journal of Computer Applications*, 47(10), 44-48.
- [5] Saini, M., Baliyan, N., & Bassi, V. (2017, August). Prediction of heart disease severity with hybrid data mining. In 2017 2nd International Conference on Telecommunication and Networks (TEL-NET) (pp. 1-6). IEEE.
- [6] Kumar, G. V., Bharadwaja, A., & Sai, N. N. (2017, May). Temperature and heart beat monitoring system using IOT. In 2017 international conference on trends in electronics and informatics (ICEI) (pp. 692-695). IEEE.

- [7] Bahrami, B., & Shirvani, M. H. (2015). Prediction and Diagnosis of Heart Disease by Data Mining Techniques. Journal of Multidisciplinary Engineering Science and Technology (JMEST), 2(2), 164-168.
- [8] Amin, S. U., Agarwal, K., & Beg, R. (2013, April). Genetic neural network based data mining in prediction of heart disease using risk factors. In 2013 IEEE Conference on Information & Communication Technologies (pp. 1227-1231). IEEE.
- [9] Bhaskaru, O., Sreedevi, M," Research on classification techniques in data mining", International Journal of Innovative Technology and Exploring Engineering, 2019.
- [10] Ratnam, D., HimaBindu, P., Sai, V. M., Devi, S. R., & Rao, P. R. (2014). Computer-based clinical decision support system for prediction of heart diseases using Naïve Bayes algorithm. *International Journal of Computer Science and Information Technologies*, 5(2), 2384-2388.
- [11] Liu, X., Wang, X., Su, Q., Zhang, M., Zhu, Y., Wang, Q., & Wang, Q. (2017). A hybrid classification system for heart disease diagnosis based on the RFRS method. Computational and mathematical methods in medicine, 2017.
- [12] Akgül, M., Sönmez, Ö. E., & Özcan, T. (2019, July). Diagnosis of Heart Disease Using an Intelligent Method: A Hybrid ANN–GA Approach. In International Conference on Intelligent and Fuzzy Systems (pp. 1250-1257). Springer, Cham.
- [13] Hasan, T. T., Jasim, M. H., & Hashim, I. A. (2018, December). FPGA Design and Hardware Implementation of Heart Disease Diagnosis System Based on NVG-RAM Classifier. In 2018 Third Scientific Conference of Electrical Engineering (SCEE) (pp. 33-38). IEEE.
- [14] Tarle, B., & Jena, S. (2017, August). An artificial neural network based pattern classification algorithm for diagnosis of heart disease. In 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA) (pp. 1-4). IEEE.
- [15] Mai Shouman, Tim Turner and Rob Stocker, "Integrating Decision Tree and K-Means

Clustering with Different Initial Centroid Selection Methods in the Diagnosis of Heart Disease Patients", Proceedings of the International Conference on Data Mining, 2012

- [16] Potharaju, S. P., & Sreedevi, M. (2018). A novel subset feature selection framework for increasing the classification performance of SONAR targets. *Procedia Computer Science*, 125, 902-909.
- [17] Potharaju, S. P., & Sreedevi, M. (2017). A Novel Clustering Based Candidate Feature Selection Framework Using Correlation Coefficient for Improving Classification Performance. *Journal of Engineering Science & Technology Review*, 10(6).
- [18] Potharaju, S. P., & Sreedevi, M. (2017). A novel cluster of feature selection method based on information gain. *IJCTA*, *10*(14), 9-16.
- [19] Potharaju, S. P., & Sreedevi, M. (2017). A Novel M-Cluster of Feature Selection Approach Based on Symmetrical Uncertainty for Increasing Classification Accuracy of Medical Datasets. *Journal of Engineering Science & Technology Review*, 10(6).
- [20] Zhang, J. L., Zhou, G., & Zhou, Y. Q. (2010). A new artificial glowworm swarm optimization algorithm based on chaos method. In *Quantitative Logic and Soft Computing* 2010 (pp. 683-693). Springer, Berlin, Heidelberg.
- [21] Guenther, N., & Schonlau, M. (2016). Support vector machines. *The Stata Journal*, *16*(4), 917-937.
- [22] Ghaddar, B., & Naoum-Sawaya, J. (2018). High dimensional data classification and feature selection using support vector machines. *European Journal of Operational Research*, 265(3), 993-1004.
- [23] Vidyullatha pellakuri, "Performance analysis of machine learning techniques for intrusion detection system", 2019Proceedings - 2019 5th International Conference on Computing, Communication Control and Automation, ICCUBEA 2019.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0 https://creativecommons.org/licenses/by/4.0/deed.en_US