Cardiovascular Disease Detection Using Sand Cat Swarm-driven AdaBoost Approach

Ish Kapila[1], Dr. Raj Kumar[2], Dr. Suhas Ballal[3], Alok Kumar Moharana[4], Girija Shankar Sahoo[5], Sidhant Das[6]

[1]Centre of Research Impact and Outcome, Chitkara University, Rajpura- 140417, Punjab, India, ish.kapila.orp@chitkara.edu.in https://orcid.org/0009-0001-0976-6585
[2]Assistant Professor, Department of uGDX, ATLAS SkillTech University, Mumbai, Maharashta, India, Email Id- raj.kumar@atlasuniversity.edu.in, Orcid Id- 0000 0001 8878 7407
[3]Assistant Professor, Department of Biochemistry, School of Sciences, JAIN (Deemed-to-be University), Karnataka, India, Email Id- b.suhas@jainuniversity.ac.in, Orcid Id- 0000-0002-6041-8332
[4]Associate Professor, Department of Pharmacy, ARKA JAIN University, Jamshedpur, Jharkhand, India, Email Id- alok.m@arkajainuniversity.ac.in, Orcid Id- 0000-0002-6169-6000
[5]Assistant Professor, Maharishi School of Engineering & Technology, Maharishi University of Information Technology, Uttar Pradesh, India, Email Id-gssahoo07@gmail.com, Orcid Id- 0000-0001-5512-1441
[6]Chitkara Centre for Research and Development, Chitkara University, Himachal Pradesh-174103 India, sidhant.das.orp@chitkara.edu.in, https://orcid.org/0009-0003-3540-5817

Cardiovascular diseases (CVDs) constitute a major cause of death and disability worldwide and represent a major public health problem. Timely intervention and better patient outcomes depend on early identification of CVD. A novel Sand Cat Swarm-driven AdaBoost (SCS-AB) technique is presented for the diagnosis of CVD. The suggested strategy chooses useful features from intricate cardiovascular (CV) datasets by utilizing the shared knowledge of the SCS optimization technique. The SCS method, which addresses feature selection issues in CVD diagnosis, is motivated by the social behaviour of sand cats. It demonstrates a cooperative and self-organizing decision-making process. The AdaBoost incorporates the chosen features, improving its ability to distinguish between patients. The gathered dataset is used in this study's analysis of the suggested SCS-AB approach. The SCS-AB driven CVD prognosis is tested on the Python platform in this study. When contrasted with other standard approaches, the experimental results show that the suggested methodology is beneficial with regard to accuracy, sensitivity and specificity. The method has demonstrated its potential for practical clinical applications as well, offering a useful instrument for rapid CVD identification and risk evaluation.

Keywords: Cardiovascular diseases (CVDs), patient outcomes, early identification, Sand Cat Swarm-driven AdaBoost (SCS-AB).

1. Introduction

Due to their enormous contribution to global morbidity and mortality, CVD continues to be a significant global health concern. Due to changing lifestyle patterns and an aging population, the incidence of CVD is increasing, calling for a more sophisticated and all-encompassing approach to diagnosis and treatment. Advances in diagnostic tools, advanced analytics and new technologies have brought a paradigm change in CVD diagnosis in recent years [1]. In the past, gender, age, high blood pressure, cholesterol levels, tobacco use and family history were among the traditional risk factors that were mainly used to detect CVD. These parameters provide a partial understanding of the complex relationship between CV health and other variables [2]. Advanced diagnostic methods enabled by Machine Learning (ML) and Artificial Intelligence (AI) have transformed the domain of CVD detection.

In this field, developing non-invasive techniques is one of the most significant achievements. Considerable progress has been made in developing advanced visual technology, including Computed Tomography (CT), Magnetic Resonance Imaging (MRI) and cardiac ultrasound. With the ability to quantify heart performance, evaluate blood flow and identify anatomical anomalies early, these tools offer previously unattainable insights into both the structure and operation of the heart. Recent technologies like cardiac MRI and three-dimensional echocardiography have emerged [3, 4]. Molecular tests have become vital tools in the identification of CVD. Heart troponins, Brain Natriuretic Peptides (BNP) and C-reactive protein (CRP) are some markers that provide essential insights into inflammation, heart failure and heart function. These markers provide a thorough profile of CVD and aid in early identification and personalized therapy when paired with patient-specific medical data analysis [5]. In recent years, wearable technology has emerged as a trailblazer in the monitoring and diagnosing of CVD. Fitness trackers, smart-watches and other wearable gadgets with robust sensors can track essential indicators, such as saturation in oxygen and heart rate [6]. These devices enable continuous monitoring and, by employing AI algorithms to detect minor variations from baseline, notify medical professionals of any CVD. This proactive strategy allows people to take charge of their CV health and facilitates timely medical intervention. Due to artificial intelligence (AI) and ML approaches, CV risk analysis has improved. To build prediction models, these techniques could examine large databases, such as genetic information, radiological findings, medical records and biomarker profiling. Afterward, by classifying patients into risk categories using these prediction models, clinicians can create individualized treatment plans and focused treatments [7, 8]. The paradigm of forecasting in CVD detection has been transformed by the combination of ML and big data analytics techniques. ML models can recognize intricate patterns and mild correlations among various risk factors as well as disease outcomes by utilizing large repositories of clinical data along with empirical data. This allows doctors to implement specific therapies and targeted interventions. Precision medicine is made possible by the combination of data-driven understanding and clinical experience [9]. In this paper a novel Sand Cat Swarm-driven Adaboost (SCS-AB) for diagnosing cardiovascular disease (CVD) is introduced.

The subsequent sections of the paper have been structured in the following order. In section 2, an examination of the related works was conducted. Section 3 provides an in-depth explanation of the proposed approach methodology, while section 4 shows findings obtained from the study. Section 5 discusses the conclusions derived from the study.

2. Related works

The study [10] developed a CVD prediction model based on ML techniques based on the health screening datasets. 4699 individuals aged 45 and over, who had their diagnoses made using the international classification of diseases (I20-I25) made comprised the CVD group. In addition, 4699 arbitrary patients who did not have a CVD diagnosis were added to the non-CVD group. According to their research, the medical screening dataset-based CVD prediction system that used machine learning (ML) algorithms was more accurate than previous CVD models in predicting outcomes, was easily adjustable, and produced confirmed results. The University of California Irvine repository (UCI) Lab collected data on CVD patients, and the author of [11] applied to find patterns methods like Decision Trees (DT), Neural Networks, Rough Sets, Support Vector Machines (SVM), and Naive Bayes, evaluated their accuracy and predictions. A hybrid strategy was proposed to increase the accuracy of these approaches. Study [12], six ML algorithms were employed and their respective performances were evaluated as well as compared. DT performed better than the other classifiers in terms of testing accuracy. They carried out the six distinct ML algorithms to forecast the illness and shared the outcomes. To examined that the SVM, logistic regressions (LR) and random forests (RF) methods [13] were used to predict CVD as issues of classification utilizing ML-based methodologies. These ML-based methods of CVD prediction were capable of dealing with non-linear classification due to strong mathematical concepts support. SVM performed better than LR and RF. The adoption of ML-based technologies could help with disease prediction.

To discuss two reputable ML algorithms, multi-layer perceptrons (MLP) and K-nearest neighbors (K-NN) that were utilized to identify CVD using publicly available data from the UCI [14]. According to the comparison of findings, the MLP model had greater prediction accuracy. To investigate those data-driven strategies used supervised ML algorithms to detect CVD patients. Multiple ML models (LR, SVM, RF and gradient boosting) [15] were tested on the accuracy of classification using various data sets and time-frames. The models were combined to generate a weighted ensemble model. That study examined into those models' applicability in diagnosing people with CVD. The examined at the Kaggle dataset, including characteristics related to heart disease such as age, sexual orientation, cholesterol and hypertension. They evaluated the precision of several ML techniques. Using a Kaggle heart attack dataset, they compared their accuracy scores in an attempt to employ artificial neural networks (ANN) [16] to improve the accuracy of predictions. A list of the many prediction parameters was created by them. The study [17] developed a heart illness predictive system that evaluated the possibility of receiving a heart problem diagnosis based on their past medical history. They used many ML techniques, including LR, KNN, to forecast and identify people with CVD illness. Their objective was to ascertain that the design could be employed to enhance the accuracy that a CVD can be predicted in any individual.

The author of, [18] presented a novel strategy that used ML techniques to identify important components in an effort to improve the accuracy of CVD diagnosis. Well-known methods for classification and multiple feature combinations were used to analyze the predictions model. The results demonstrated that each attribute selected and ML technique used was effective in predicting cardiac disease compared to conventional models. Study [19] examined the accuracy of ML approaches in assessing overall dangers and forecasting cardiac diseases. The proposed experiment used a combination of standard ML approaches, including RF, K-NN, DT, SVM and LR, to predict CVD with high accuracy. A forecasting algorithm is discussed that may be used to identify or make people aware of cardiac problems. They accomplished this by comparing the precision of applying rules to each of the individual results of SVM [20], Gradient Boosting, RF, and Naive Bayes. The outcomes were shown, and the accuracy and results generated were evaluated. The algorithms yielded the most accurate results in this instance.

3. Methodology

In this paper, Figure 1 depicts the proposed method architecture for detecting CVD using a patient dataset. The CVD data is preprocessed using min max normalization to eliminate duplicate values. We suggest that a Sand Cat Swarm-driven Adaboost (SCS-AB) offers a unique and potentially more effective approach for improved CVD detection.

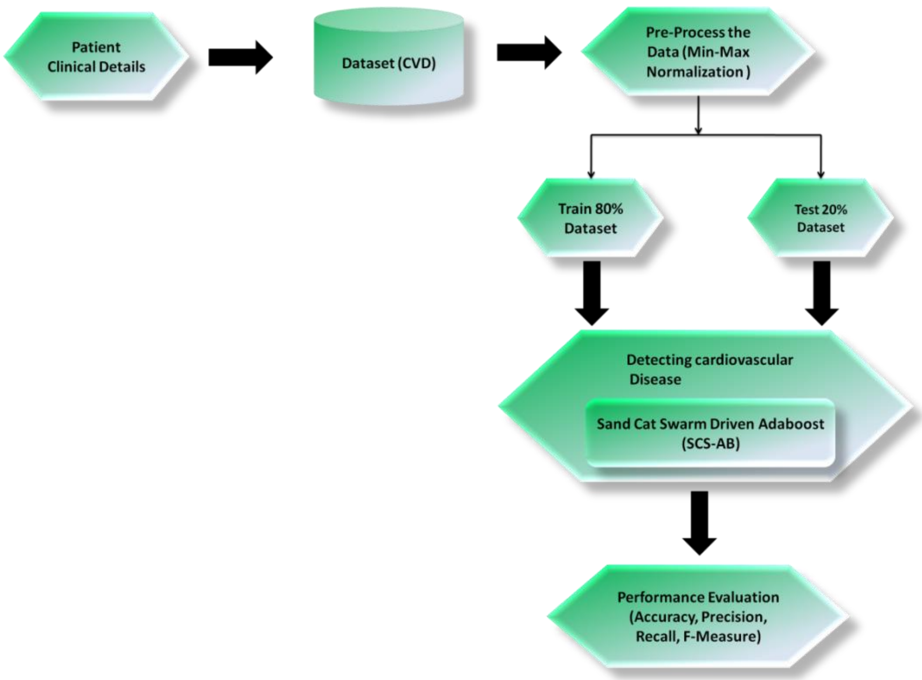

Figure 1: Proposed method architecture (Source: Author)

3.1 CVD Dataset

There is a lot of patient data and related health records in the dataset that is used to diagnose and identify CVD. This dataset was gathered by Kaggle from three distinct sources. Factual information on CV disorders is given by the objective resource, test findings are included in the examination and patient-generated material is included in the subjective resource. The format of the CVD dataset is (68783, 12) and it's the latest version of the CVD dataset as shown in Table 1 [21]. Two categories of data were created: training data (80%) and testing data (20%).

Table 1: CVD dataset

Feature	Information
Age	The individual years of age
Sexuality	Dichotomous value: 1-male, 2 - female.
Altitude	Depicting a patient's height
Load	Indicating the strength of the individual's
AP_HIGHT	Systolic hypertension
AP_LOW	High blood pressure diastolic
Chlorosterolic	The blood's cholesterol level (1: Typical, 2: Exceeding the Norm, 3: Significantly Surpassing the Norm)
SUGARSE	A value of the blood sugar categories (1: Healthy, 2: Elevated, 3: High)
USES TOBACCO	Tobacco Consuming (0: No, 1: Yes)
Liquor	Use of alcohol (1: agree 0: None)
Physical movement	Type of physical activity
Heart disease	Targeted value indicating whether CVD is present or absent.

3.2 Min Max Normalization

We apply techniques to preprocess data for CVD detection and standardize features such as blood pressure or cholesterol levels. We offer the Min-Max Normalization approach that can be stated as follows in Equation (1):

$$\mathrm{X} = \frac{\mathrm{x} - \mathrm{Min}}{\mathrm{Max} - \mathrm{Min}} \tag{1}$$

Each dimension's maximum and minimum values are represented by the terms Min and Max, respectively. The Min-Max Normalization can map data between 0 and 1 without disrupting the original data's linear connection, improving the model's accuracy and convergence speed.

3.3 Predicting CVD

In this section we integrate sand cat swarm optimization (SCSO) and AdaBoost algorithm to predict CVD efficiently. AdaBoost Algorithm increases classification accuracy by improving weaker classifiers and SCSO algorithm improves the classification algorithm. This system improves pattern detection in CV data, allowing for more precise diagnosis and treatment techniques.

3.4. SCSO

SCSO is utilized to forecast CV illness. Through simulating the collective behavior of sand cats, SCSO improves our ability to recognize and manipulate CVD via improving the precision and efficacy of illness detection through intricate optimization methods. Every sand cat is an $1 \times \mathrm{dim}$ array in the dim dimensions optimizing issue. It reflects the issue's solution, as demonstrated in Figure 2. Each Opt in a collection of values for variables $(\mathrm{Opt}_1, \mathrm{Opt}_2, \dots\dots, \mathrm{Opt}_{\mathrm{dim}})$ Opt must have the limits at the bottom and upper. The initializing procedure produces a starting point depending on the size of the issue matrix$(\mathrm{M} \times \mathrm{dim})$. Additionally, iteration, the pertinent response will be displayed. If the output that follows values is superior, the previous answers have to be substituted. If no alternative occurs in the following iteration, the current iteration's result cannot be saved.

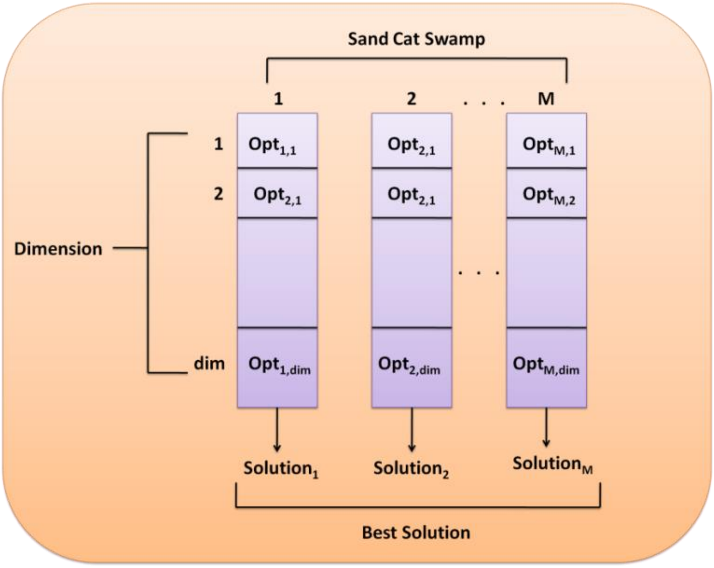

Figure 2: Group initiation (Source: https://www.mdpi.com/2079-9292/12/9/2042)

3.4.1 Explore for Prey (investigation Stage)

$\mathrm{Opt_j}$ represents every sand cat's spot. Sand cats' capacity to sense sound helps the SCSO method with low-frequency identification. Every sand cat is sensitive to frequencies lower than 2 kHz. As a result, in computational modelling, Equation (2) defines a sensitivity value q_H, meaning that the dunes cat's sensitivities range is 2 to 0 kHz. Furthermore, the exploration and utilization capacity of the technique is regulated yet the value of the parameter Q is derived following Equaitons (2-3).

$$q_H = T_N - \left(\frac{T_N \times s}{S}\right) \tag{2}$$

$$Q = 2 \times q_H \times \mathrm{rand}(0,1) - q_H \tag{3}$$

Where, S is the total number of iterations that can be made, s is the number that can be made right now, and T_N is 2.

While looking for prey, every sand cat will move arbitrarily over its sensitive range. It makes it easier to investigate and utilize algorithms. The sensitive range $(q)$of every sand cat varies to prevent them from entering the local optimal range according to Equation (4).

$$q = q_H \times \mathrm{rand}(0,1) \tag{4}$$

Where,q_H is the guiding parameter q.

Every sand cat can search for prey based on the ideal candidate position$(\mathrm{Opt_{ad}})$, current location$(\mathrm{Opt_d}(s))$ and sensitive range(q), as indicated in Equation (5).

$$\mathrm{Opt}(s+1) = q \times \left(\mathrm{Opt_{ad}}(s) - \mathrm{rand}(0,1) \times \mathrm{Opt_d}(s)\right) \qquad (5).$$

3.4.2 Targeted Attacks (Exploitation Stage)

Equation (6) mimics the procedure of the sand cat affecting prey by displaying a distance ($\mathrm{Opt_{qmc}}$) between the cat and the prey. Let us assume that the sand cat has a circle of radius as its sensitive range and its motion path selects a random angle ($\propto$) using the rolling wheels choice mechanism. Given that the chosen randomized angle ranges from 0° to 360°, its value falls between −1 and 1. Every sand cat can move in searching space in a distinct circumference path in this manner. The prey is assaulted in accordance with Equation (7). The sand cat can get to the hunting location more quickly in this manner.

$$\mathrm{Opt_{qmc}} = |\mathrm{rand}(0,1) \times \mathrm{opt_a}(s) - \mathrm{opt_d}(s)| \tag{6}$$

$$\mathrm{opt}(s+1) = \mathrm{Opt_a}(s) - q \times \mathrm{Opt_{qmc}} \times \cos(\propto) \tag{7}$$

3.4.3 SCSO Algorithm's execution

Equation (1) illustrates that the adaptable parameters q_H and Q are controlled by the SCSO method to govern the algorithm's exploration and extraction. q_H falls linearly from 2 to 0 throughout an execution. As a result, a value of $[-4, 4]$for the parameter Q is randomized. If Q equals to or lower than one, the sand cat will assault its victim. Without it, the sand cat will use Equation (8) to find prey.

$$\mathrm{Opt}(s+1) \begin{cases} q \times \left(\mathrm{Opt_{ad}}(s) - \mathrm{rand}(0,1) \times \mathrm{Opt_d}(s)\right) |Q| > 1; \text{exploration} \\ \quad \mathrm{Opt_a}(s) - \mathrm{Opt_{qmc}} \times \cos(\propto) \times q|Q| \leq 1; \text{exploitation} \end{cases} \tag{8}$$

Every sand cat's position updating throughout the exploration and extraction phase is displayed in Equation (8). The sand cat will attack its victim when $Q \leq 1$. If not, the sand cat's job is to search the global region for fresh prey.

3.5 Adaptive Boosting (AdaBoost)

The Adaboost algorithm easily detects CVD by boosting weak classifiers on features such as cholesterol levels, blood pressure and others. Its altering weights are based on classification accuracy. AdaBoost is a ML algorithm. The structure of AdaBoost can be summarized as follows. AdaBoost computed the average weighted classification error per learner s utilizing the subsequent Equation (9):

$$f_s = \sum_{m=1}^{M} c_m^{(s)} \, J\left(z_m \neq g_s(w_m)\right) \tag{9}$$

Where, w_m is the prediction vector value for the observations m, z_m represents the actual class label and g_s represents the hypothesis (learner predictor). In step s, J is the value of the indicator variable and $c_m^{(s)}$ is the measurement weight. AdaBoost instructs individuals consecutively. AdaBoost calculates prediction utilizing the subsequent Equation (10-12) during the conditioning stage:

$$e(w) = \sum_{s=1}^{S} \propto_s \, g_s(w) \tag{10}$$

$$\propto_s = \frac{1}{2} \log \frac{1-\varepsilon_s}{\varepsilon_s} \tag{11}$$

Where $\propto_s$ is the ensemble's weak hypotheses weight, AdaBoost retraining can be thought of as the minimizing of an exponential loss utilizing the equation below.

$$\sum_{m=1}^{M} x_m \exp^{(-z_m e(w_m))} \quad (12)$$

Given that $z_m \epsilon \{-1,1\}$is the real class, x_m is the normalized observed weight and $e(w_m) \epsilon (-\infty, +\infty)$is the projected classification.

3.6 Sand Cat Swarm-driven Adaboost (SCS-AB)

Predicting CVD with Sand Cat Swarm-driven Adaboost is a strong and innovative method. Sand cat hunting is modeled by the bio-inspired optimization algorithm SCSO has been utilized for cat selection and modeling optimization of parameters in a range of ML tasks. AdaBoost is an active ensemble learning technique. Through combining SCSO with it, the prediction performance of diagnosing CV illness can be greatly enhanced. Through decreasing complexity and improving the accuracy of the data considering into the AdaBoost ensemble, SCSO improves in identifying the most relevant characteristics from a huge dataset. In addition, AdaBoost changes a number of poor classifiers into a powerful classifier that recognizes complex patterns in the data. With lowering over fitting, this hybrid strategy increases the accuracy of predictions and improves the model's quality. When combined, SCSO and AdaBoost offer an effective as well as efficient approach to CVD early detection, that's essential for enhancing patient outcomes and reducing medical expenses. Algorithm 1 displays the SCS-AB pseudocode.

Algorithm 1: Pseudo code for Sand Cat Swarm-driven Adaboost

```
Initialize Sand Cat Swarm algorithm parameters
Initialize AdaBoost algorithm parameters
while termination condition is not met do
   for each data point do
      Run Sand Cat Swarm algorithm for data point
      Get the result from Sand Cat Swarm
      if result is positive then
         Increase weight of data point using AdaBoost
      else
         Decrease weight of data point using AdaBoost
      end if
   end for
   Normalize weights using AdaBoost
   Train classifier using weighted data points
   Update AdaBoost parameters
```

Update Sand Cat Swarm parameters

end while

4. Result and Discussion

The suggested job is carried out using the open-source Anaconda 2020; Python 3.76 is needed to be installed alongside Python to carry out the procedure. In this section, we assess the proposed approach and gauge its performance with the following criteria: Accuracy (%), Precision (%), Recall (%) and F-Measure (%).We evaluate our proposed technique's performance against other available approaches. The techniques now in use include Naïve Bayes [10], Gradient Boosting [22] and ANN [23].

Accuracy: An indicator of a ML model's for accurate CVD detection. Higher numbers suggest a better forecast. Loss: A measure of a model's prediction error during training and validation that aims to close the gap between expected and actual values. The graphical depiction of the loss rate is shown in Figure 3. With the goal of achieving high accuracy and minimal loss, the training graph will monitor these parameters throughout epochs, enhancing the model's capacity to detect people who are at risk of CVD.

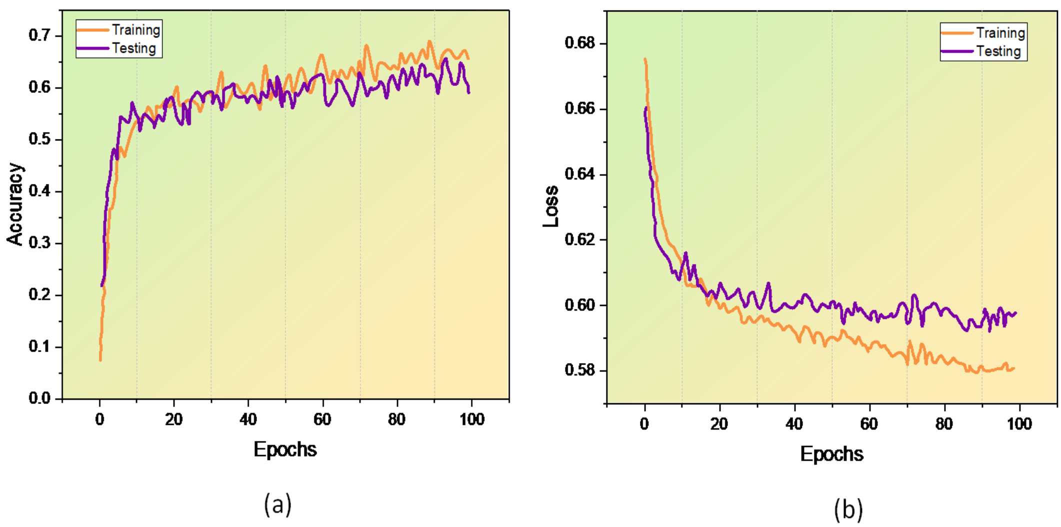

Figure 3: Outcome of Accuracy, Loss (Source: Author)

Predictive performance determines CVD that examine how efficiently a system can recognize CV instances. The correctness of case predictions are measured as a percentage of complete occurrences. Figure 4 and Table 2 depict the comparative evaluation of accuracy in suggested and traditional methods. When compared to the existing methods, such as Naïve Bayes, Gradient Boosting and ANN, which have accuracies of 85.4%, 73% and 81.7%, respectively, the suggested SCS-AB has a 90.8% accuracy rate. Our suggested approach yields better outcomes over the existing methods for CVD detection as shown in Equation (13).

$$\text{Accuracy} = \frac{(\text{TP}+\text{TN})}{(\text{TP}+\text{TN}+\text{FP}+\text{FN})} \tag{13}$$

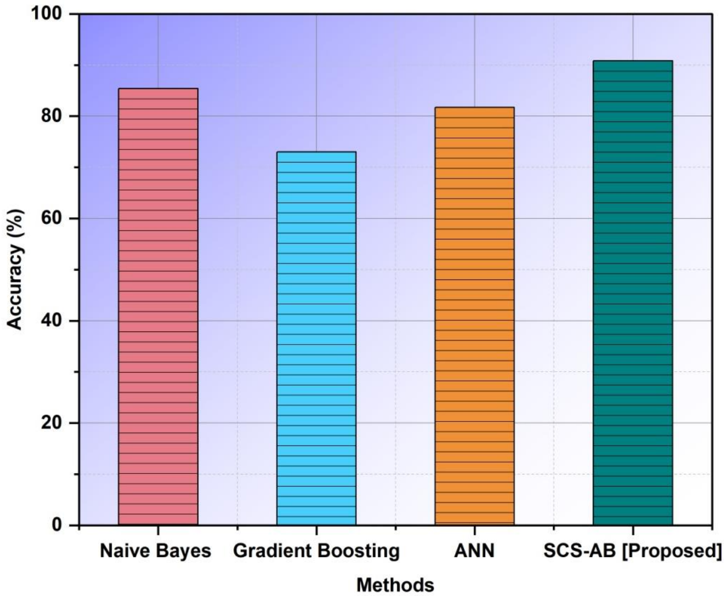

Figure 4: Result of Accuracy (Source: Author)

Table 2: Comparison of Accuracy

Methods	Accuracy (%)
Naïve Bayes	85.4
Gradient Boosting	73
ANN	81.7
SCS-AB [Proposed]	90.8

The efficiency of a CVD prediction model is characterized as the percentage of appropriately identified indications compared to the overall count of optimistic recommendations. The comparative evaluation of precision is shown in Figure 5 and Table 3. When compared to the existing methods, such as Naïve Bayes, Gradient Boosting and ANN, which have a precision of 83.1%, 74% and 83.1%, respectively and the suggested SCS-AB have a precision of 94.8%. The proposed methodology demonstrates superiority over the existing methods for CVD detection as shown in Equaiton (14).

$$\text{Precision} = \frac{\text{TP}}{(\text{TP}+\text{FP})} \tag{14}$$

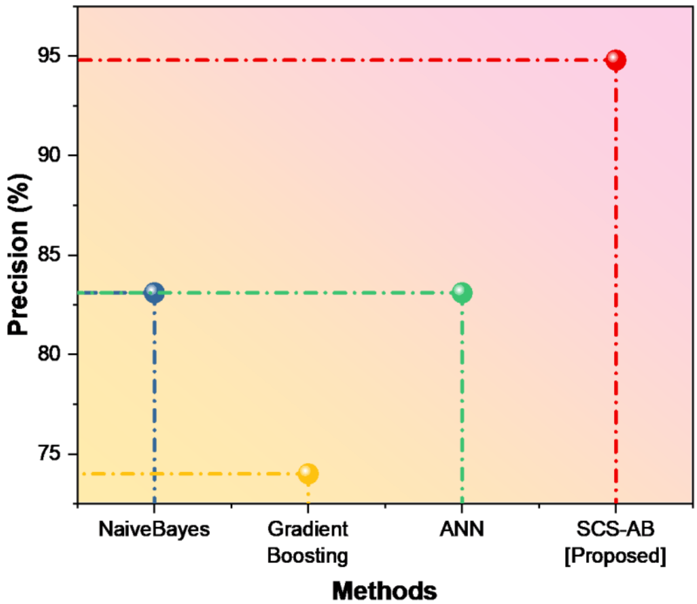

Figure 5: Precision (Source: Author)

Table 3: Comparison of Precision

Methods	Precision (%)
Naïve Bayes	83.1
Gradient Boosting	74
ANN	83.1
SCS-AB [Proposed]	94.8

Figure 6 and Table 4 determine the comparative evaluation of recall of suggested methods along with the other existing methodologies. When compared to the existing methods such as Naïve Bayes, Gradient Boosting and ANN, which have a recall of 84.9%, 73% and 85%, respectively, the suggested SCS-AB have 97.5%. The suggested methodology was superior to the existing method for CVD detection.

$$\text{Recall} = \frac{\text{TP}}{(\text{TP+FN})} \tag{15}$$

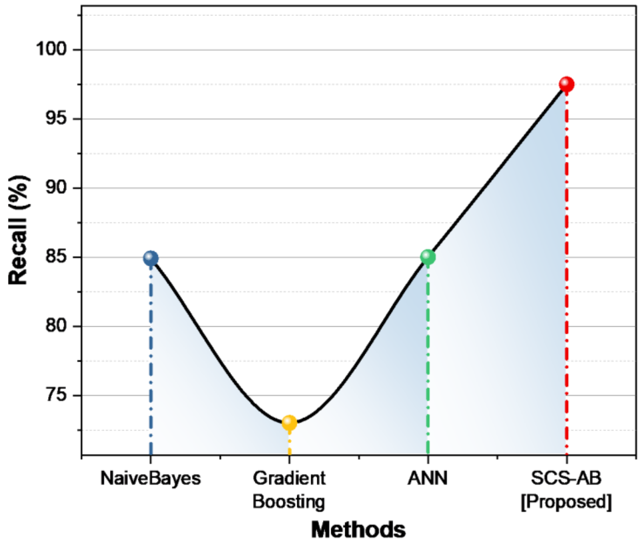

Figure 6: Recall (Source: Author)

Table 4: Comparison of Recall

Methods	Recall (%)
Naïve Bayes	84.9
Gradient Boosting	73
ANN	85
SCS-AB [Proposed]	97.5

Figure 7 and Table 5 determine the comparative evaluation of the F-Measure of suggested methods along with the other existing methodologies. When compared to the existing methods such as Naïve Bayes, Gradient Boosting and ANN, which have an F-Measure of 84.6%, 73% and 80%, respectively, the suggested SCS-AB have 94.8%. The suggested methodology was superior to the existing method for CVD detection.

$$\mathrm{F-measure} = 2 \times \frac{\mathrm{Precision} \times \mathrm{Recall}}{\mathrm{Precision} + \mathrm{Recall}} \tag{16}$$

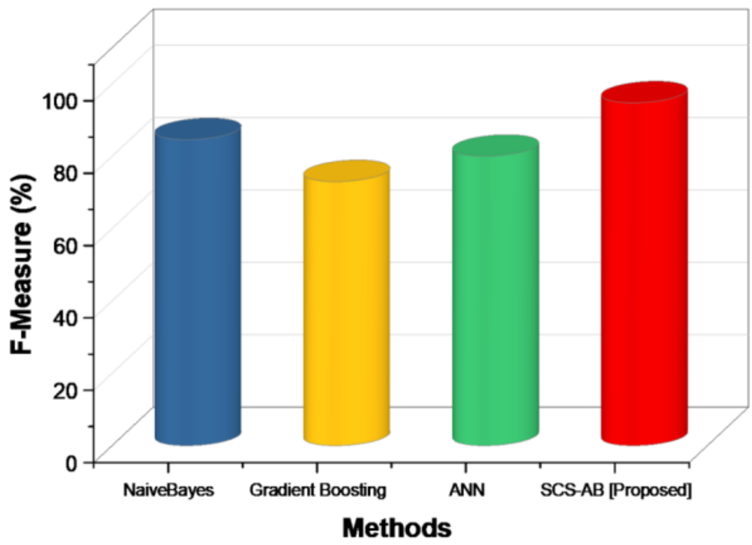

Figure 7: F-Measure (Source: Author)

Table 5: F-Measure

Methods	F-Measure (%)
Naïve Bayes	84.6
Gradient Boosting	73
ANN	80
SCS-AB [Proposed]	94.8

5. Conclusion

In this study, we introduced a novel approach, Sand Cat Swarm-driven Adaboost (SCS-AB), for the early identification of CV disorders, with the detailed examination of many medical factors. This study conducted an early detection and regular monitoring of CVD and reduced the risk of life-threatening complications using the CVD dataset. This dataset was preprocessed using min-max normalization. To evaluate the performance of the proposed method in terms of accuracy (90.8%), precision (94.8%), recall (97.5%) and F-measure (94.8), the proposed method results were contrasted with the other previously used algorithms and the evaluations' results demonstrated that the recommended approach was superior for the identification of CVD. The capacity of available diagnostic techniques estimates an individual's chance for developing CVD without any symptoms or a diagnosis could be restricted. In future research, the combination of big data as well as modern algorithms has the capability to enhance data analytics along with interpretation, hence reducing the possibility of incorrect outcomes, besides promoting more precise and accurate identification of CVD.

References

1. Mavrogeni, Koutsogeorgopoulou, Dimitroulas, Markousis-Mavrogenis, Kolovou. "Complementary role of cardiovascular imaging and laboratory indices in early detection of cardiovascular disease in systemic lupus erythematosus." Lupus 26, no. 3 (2017): 227-236. doi:https://doi.org/10.1177/0961203316671810
2. Kumar, Vinod, Jack R,Brent, Munish Shorie, Harmanjit Kaur, Gaganpreet Chadha, Andrew , Thomas, Edward , Lewis "Nanostructured aptamer-functionalized black phosphorus sensing platform for label-free detection of myoglobin, a cardiovascular disease biomarker." ACS applied materials & interfaces 8, no. 35 (2016): 22860-22868. doi:https://doi.org/10.1021/acsami.6b06488
3. Brunese, Luca, Fabio Martinelli, Francesco Mercaldo, Antonella Santone. "Deep learning for heart disease detection through cardiac sounds." Procedia Computer Science 176 (2020): 2202-2211. doi:https://doi.org/10.1016/j.procs.2020.09.257
4. Abubaker, Mohammed , Bilal Babayiğit. "Detection of cardiovascular diseases in ECG images using machine learning and deep learning methods." IEEE Transactions on Artificial Intelligence 4, no. 2 (2022): 373-382. doi:https://doi.org/10.1109/TAI.2022.3159505
5. Sadad, Tariq, Syed Ahmad Chan Bukhari, Asim Munir, Anwar Ghani, Ahmed M. El-Sherbeeny, Hafiz Tayyab Rauf. "Detection of cardiovascular disease based on PPG signals using machine learning with cloud computing." Computational Intelligence and Neuroscience 2022 (2022). doi:https://doi.org/10.1155/2022/1672677
6. Mabbott, Samuel, Syrena , Fernandes, Monika Schechinger, Gerard , Cote, Karen Faulds, Charles R. Mace, and Duncan Graham. "Detection of cardiovascular disease associated miR-29a using paper-based microfluidics and surface enhanced Raman scattering." Analyst 145, no. 3 (2020): 983-991. doi:https://doi.org/10.1039/C9AN01748H
7. Rizzo, Nancy R, Nicole , Hank, Jian Zhang. "Detecting presence of cardiovascular disease through mitochondria respiration as depicted through biophotonic emission." Redox Biology 8 (2016): 11-17. doi:https://doi.org/10.1016/j.redox.2015.11.014
8. Qtaish, Amjad, Dheeb Albashish, Malik Braik, Mohammad , Alshammari, Abdulrahman Alreshidi, Eissa Jaber Alreshidi. "Memory-based Sand Cat Swarm Optimization for Feature Selection in Medical Diagnosis." Electronics 12, no. 9 (2023): 2042. doi:https://doi.org/10.3390/electronics12092042
9. Kim, Joung Ouk, Yong-Suk Jeong, Jin Ho Kim, Jong-Weon Lee, Dougho Park, Hyoung-Seop Kim. "Machine learning-based cardiovascular disease prediction model: A cohort study on the Korean National Health Insurance Service Health Screening Database." Diagnostics 11, no. 6 (2021): 943. doi:https://doi.org/10.3390/diagnostics11060943
10. Esfahani, Hamidreza Ashrafi, and Morteza Ghazanfari. "Cardiovascular disease detection using a new ensemble classifier." In 2017 IEEE 4th international conference on knowledge-based engineering and innovation (KBEI), pp. 1011-1014. IEEE, 2017. doi:https://doi.org/10.1109/KBEI.2017.8324946
11. Baghdadi, Nadiah A., Sally Mohammed Farghaly Abdelaliem, Amer Malki, Ibrahim Gad, Ashraf Ewis, and Elsayed Atlam. "Advanced machine learning techniques for cardiovascular disease early detection and diagnosis." Journal of Big Data 10, no. 1 (2023): 144. https://doi.org/10.1186/s40537-023-00817-1
12. Sun, Weicheng, Ping Zhang, Zilin Wang, and Dongxu Li. "Prediction of cardiovascular diseases based on machine learning." ASP Transactions on Internet of Things 1, no. 1 (2021): 30-35. doi:https://doi.org/10.52810/TIOT.2021.100035
13. Pal, Madhumita, Smita Parija, Ganapati Panda, Kuldeep Dhama, and Ranjan K. Mohapatra. "Risk prediction of cardiovascular disease using machine learning classifiers." Open Medicine 17, no. 1 (2022): 1100-1113. doi:https://doi.org/10.1515/med-2022-0508
14. Dinh, An, Stacey Miertschin, Amber Young, Somya , Mohanty. "A data-driven approach to

predicting diabetes and cardiovascular disease with machine learning." BMC medical informatics and decision making 19, no. 1 (2019): 1-15. doi:https://doi.org/10.1186/s12911-019-0918-5

15. Pasha, Syed Nawaz, Dadi Ramesh, Sallauddin Mohmmad, Harshavardhan. "Cardiovascular disease prediction using deep learning techniques." In IOP conference series: materials science and engineering, vol. 981, no. 2, p. 022006. IOP Publishing, (2020). doi:10.1088/1757-899X/981/2/022006
16. Shah, Devansh, Samir Patel, Santosh Kumar Bharti. "Heart disease prediction using machine learning techniques." SN Computer Science 1 (2020): 1-6. doi:https://doi.org/10.1007/s42979-020-00365-y
17. Jindal, Harshit, Sarthak Agrawal, Rishabh Khera, Rachna Jain, Preeti Nagrath. "Heart disease prediction using machine learning algorithms." In IOP conference series: materials science and engineering, vol. 1022, no. 1, p. 012072. IOP Publishing, (2021). doi:10.1088/1757-899X/1022/1/012072
18. Mohan, Senthilkumar, Chandrasegar Thirumalai, a Gautam Srivastava. "Effective heart disease prediction using hybrid machine learning techniques." IEEE access 7 (2019): 81542-81554. doi:https://doi.org/10.1109/ACCESS.2019.2923707
19. Chauhan, Yash Jayesh. "Cardiovascular disease prediction using classification algorithms of machine learning." International Journal of Science and Research 9, no. 5 (2020): 194-200. doi:10.21275/SR20501193934
20. Dinesh, Kumar , Arumugaraj, Kumar . Santhosh, Mareeswari. "Prediction of cardiovascular disease using machine learning algorithms." In 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), pp. 1-7. IEEE, (2018). doi:https://doi.org/10.1109/ICCTCT.2018.8550857
21. KAGGEL, Cardiovascular Disease dataset. 2020
22. Alalawi, Hana , and Manal , Alsuwat. "Detection of cardiovascular disease using machine learning classification models." International Journal of Engineering Research & Technology 10, no. 7 (2021): 151-7.
23. Nissa, Najmu, Sanjay Jamwal, Shahid Mohammad. "Early detection of cardiovascular disease using machine learning techniques an experimental study." Int. J. Recent Technol. Eng 9, no. 3 (2020): 635-641.