International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Preventing Cardiac Arrest Using Novel Clustering Technique With Bio-Inspired Optimization Algorithm

[1]Jayachandran A, [2]Mohammad Asif Iqbal, [3]Anu Sharma, [4]Kumod Kumar Gupta

Submitted: 23/04/2023 **Revised**: 20/06/2023 **Accepted**: 28/06/2023

Abstract*:* Cardiac illness is the most infectious disease in the world currently for individuals of all ages. An essential necessity to anticipate heart illness correctly in a short period. Problems that are both complex and persistent are best tackled using optimization methods. The majority of applications of machine learning and clustering techniques are in the field of cardiovascular disease prediction. When making predictions, clustering makes heavy use of classification algorithms. For data preparation and cleaning, the hamming distance feature selection approach is suggested in this article for use across various heart illness datasets. In order to provide a reliable forecast of heart illness, a bio-inspired clustering model like Bilinear Fuzzy K-means Clustering (BFKC) is used with the Chaotic Drift Cuckoo Search Optimization Algorithm (CDCSOA). The findings show that BFKC-trained CDCSOA performs well, with an accuracy of 95 percent.

Keywords: *Cardiac arrest prediction, Bio-inspired algorithm, Optimization and Clustering.*

1. Introduction

Among the many deadly illnesses that have plagued humanity for centuries, heart disease has attracted much attention from scientists. It's one of the most apparent disorders people have in their middle life or later, and it may cause fatal complications in certain situations. Men are disproportionately affected by heart disease compared to women. Prediction relies heavily on the discovery of cardiac disease [1]. Reduced health risks and heart failure may be prevented with an accurate diagnosis of heart disease. Cigarette smoking, becoming older, being of a different ethnic background, having a family background of heart disease, having hypertension or high cholesterol levels, having diabetes, being overweight or obese, not getting enough physical activity, being stressed out, and having inflamed blood vessels are all of the potential causes of cardiovascular illness. Heart disease kills 17.9 million individuals a year, making it the top cause of mortality globally, according to the World Health Organisation (WHO). The provision of valuable services at affordable prices is a serious challenge faced by healthcare providers such as healthcare facilities [2]. Among those who exhibit COVID-19 signs, heart disease is now the leading cause of death. Also, CHD, the most prevalent form of cardiovascular disease, accounts for around 2% of deaths in India out of a global total of 17.9 million. Published studies on heart disease during the last 20 years show that CHD instances have surged by 100%, creating a complicated scenario or perhaps an epidemic over the next decade worldwide. Cardiovascular diseases are a leading cause of death globally, accounting for an estimated 30% of all deaths. “Coronary artery disease, heart attack, heart failure, cardiac disease, heartburn, a condition of the heart pleural disease, strokes, and arterial disease” are all associated with heart ailments [3]. Predicting the onset of cardiac conditions like hypertrophic cardiomyopathy will be one area of study. The process of giving or receiving blood is known as a blood donation. It's likely to contribute to a heart patient's future health problems significantly. Bleeding may occur during cardiopulmonary bypass due to factors including the individual's age, the surgeon's experience, the patient's prior count of platelets, and the depth of the circulation stoppage [4]. Machine learning entered the scene with many models focusing on different people's health-related issues, mainly as a response to the shortage of healthcare services provided by physicians. Clustering innovations, bolstered by bio-inspired algorithms often used for tackling optimization issues, have lately emerged as an alternative to classic machine-learning approaches for improving computing power when identifying heart disease categorization types [5]. Modern architecture, materials, and techniques are flourishing thanks to bio-inspiration. The biological advancements throughout millions of years that have resulted in adaption, processes, and salvation boost this. The idea is to improve

---

[1]*Professor & HOD, Department of Computer Science and Engineering, Presidency University, Bangalore, India, Email Id: ajayachandran@presidencyuniversity.in*

[2]*Assistant Professor, Department of Electrical Engineeing, Vivekananda Global University, Jaipur, India, Email Id: mohammad.asif.iqbal@vgu.ac.in*

[3]*Assistant Professor, College of Computing Science and Information Technology, Teerthanker Mahaveer University, Moradabad, Uttar Pradesh, India, Email id: er.anusharma18@gmail.com*

[4]*Assistant Professor, Department of Artificial Intelligence (AI), Noida Institute of Engineering and Technology, Greater Noida, Uttar Pradesh, India, Email id: kumodkumargupta@gmail.com*

the contouring and reproduction of the organismal structure to grasp better the environment's scientific constructional elements, such as a supplement, for use in bioinspired design. While many branches of study into natural phenomena had little funding in the 1980s and 2010s, the bio-inspired area saw explosive development, and a wide range of unique algorithms was effectively developed by 2020. Adopting bio-inspired approaches allows clustering optimization to address complex optimization issues, including topology design, tuning, and extreme parameter acquisition. Classification problems include making diagnoses for large numbers of patients using data obtained under several labels for the same ailment. Since cardiac disease categorization data is expected to be large, using clustering approaches embedded with bio-inspired algorithms might help improve diagnostic precision [6].

The focus of the suggested solution is on the use of a bilinear fuzzy k-means clustering model for the prediction of cardiovascular illnesses. Objectives of this study include identifying commonalities in algorithm types, Efficiency, and application in predicting cardiovascular disease risk. Predicting coronary sickness using systems for classification also uses a bio-inspired method.

Contributions:

- Early diagnosis is crucial for tackling severe diseases like heart disease.
- Researchers use cluster models like bilinear fuzzy k-means clustering to diagnose heart disease.
- Furthermore, the bio-inspired approach is optimized for better identification of heart illness.
- We found that using a Cuckoo search algorithm with a BFKC for detecting heart problems increased accuracy to 95%.
- The suggested model produces superior outcomes compared to the alternatives.

The following is a schedule for the proposed study: The literature review of diverse outcomes and methodologies established by numerous writers in accordance with heart illness prediction is described at length in Section 2. Section 3 explains the approach used. In Section 4, appropriate simulations and tables illustrate the findings. Section 5 provides the study's conclusions and recommendations for its use in practical settings.

2. Related Work

There are a number of issues with the current approach to computer-aided diagnosis (CAD) for liver cancer, which is founded on the conventional design of features techniques. Modern deep learning algorithms have solved these issues by implicitly collecting complex patterns from massive medical picture data. However, they are still influenced by the variables and topology of the underlying network. Therefore, the current state of the art may be improved by incorporating bio-inspired notions into models trained with deep learning. They present a unique bio-inspired deep learning strategy for improving liver cancer prediction in that study. That method adds to the canon in two distinct ways. First, the SegNet network, the Unsupervised Neural Network, and the Artificial Bee Colony Optimisation are combined into a unique hybrid delineation technique called SegNet-UNet-ABC for extracting liver lesions from CT images. The SegNet is used to segment the liver from the stomach CT image, and then the UNet is utilized to identify and isolate liver lesions [7]. The goal of that study is to implement and assess the efficacy of a bio-inspired artificial intelligence-based student separation approach for detecting students with elevated levels of efficiency anxiety. In order to categorize students using their feelings and degrees of anxiety about performance throughout the event, a Mayfly-based clustering optimization technique is deployed to a dataset including 774 occurrences of students. The suggested technique is differentiated by higher classification skills, as shown by contrast to genetic algorithms and particle swarm optimization [8]. One of the leading causes of cancer mortality in women is breast cancer. Computer-assisted diagnosis (CAD) aids radiologists in swiftly spotting anomalies. A new CAD system is suggested in that paper for analyzing mammography images using grey wolf optimizer and general theory. Mass segmentation mammograms may have texture, magnitude, and shape-based characteristics retrieved from them. A new reducing dimensionality approach combining GWO and rough set theory is provided for deriving the proper features from the retrieved feature set. Inspired by the grey wolf's hunting methods and social structure, GWO is a cutting-edge bio-inspired optimization algorithm. In that study, they combine GWO and Rough Set techniques to extract useful characteristics from mammography scans [9]. To improve the heart illness diagnostic accurateness of classification, that chapter proposes a bio-inspired optimized approach called the Chimp optimization algorithm (ChOA) to pick characteristics. In that method, an average filter is first used to eliminate distracting noise in the cardiac picture. The cardiac picture is then processed to isolate GLCM characteristics. The ChOA technique selects the best features from the set of features retrieved. The classifier receives as input the specified characteristics. In that method, a support vector neural network classifier is employed [10]. With cardiovascular disease now the primary cause of death worldwide, there is a pressing need to develop reliable early detection methods so that individuals can take preventative measures or receive therapy before the disease progresses to a fatal stage. To that end, they propose a Modified Artificial Plant

Optimisation (MAPO) algorithm that, combined with other ML methods, can accurately predict a person's heart rate from their fingertip video dataset and, thus, their current risk of developing coronary heart disease. After the video footage has been cleaned and the noise removed, MAPO is used to accurately predict the heart rate with a 0.9541 correlation coefficient and a 2.418 average error Estimate [11]. Artificial intelligence (AI) provides the software technologies essential for analyzing the available data and making the most accurate predictions possible. Several strategies for handling data are available in an organization model for classifying heart disease. The kind of data is of special importance to our study. A model that predicts may be derived from the training and test information with the help of the categorization. Using a mix of computational techniques and computational approaches, an approach to classification screens this data and generates an entirely novel model able to perform detailed data. We may examine the existing data using the optimization method and make accurate predictions. Utilizing several classifier methods such as “Naive Bayes, Bayesian Optimised Support Vector Machine, K-Nearest Neighbours, and Slap Swarm Optimised Neural Network”, this study intends to provide a structure for predicting cardiovascular disease utilizing the main risk variables. Using the heart illness dataset from the UCI Machine Repository, that study aims to improve the accuracy of cardiac disease detection [12]. Significant advancements in health care have emerged from the use of ML methods, which have allowed for earlier identification and evaluation of conditions like heart disease (HD). Accurate medical diagnoses are essential, and ML's ability to detect HD may help doctors do just that. HD is currently the most common illness globally, and early detection greatly improves patients' prognosis. Signs of HD may be mitigated and understood using ML techniques. Therefore, a new strategy, distinct from traditional guided ML techniques, is proposed in that paper. The study compared Independent Factor Analysis, a methodology for reducing dimensions, with the ensemble method and an ANN. The research used data from the Heart Disease archive at UCI ML Source [13].

3. Proposed Methodology

The suggested layout is displayed in Figure 1. The work of the recommended approach is broken down into four stages: data collection, data pre-processing by hamming distance, feature selection, and classification. Our cleaning model receives cardiac illness datasets that have been gathered and analyzed. The data pre-processing phase accounts for missing values eliminates duplicates, and sorts out extraneous characteristics using the hamming distance selection of features technique. Our proposed study uses Hamming distance to calculate the metric between heart disease dataset characteristics.

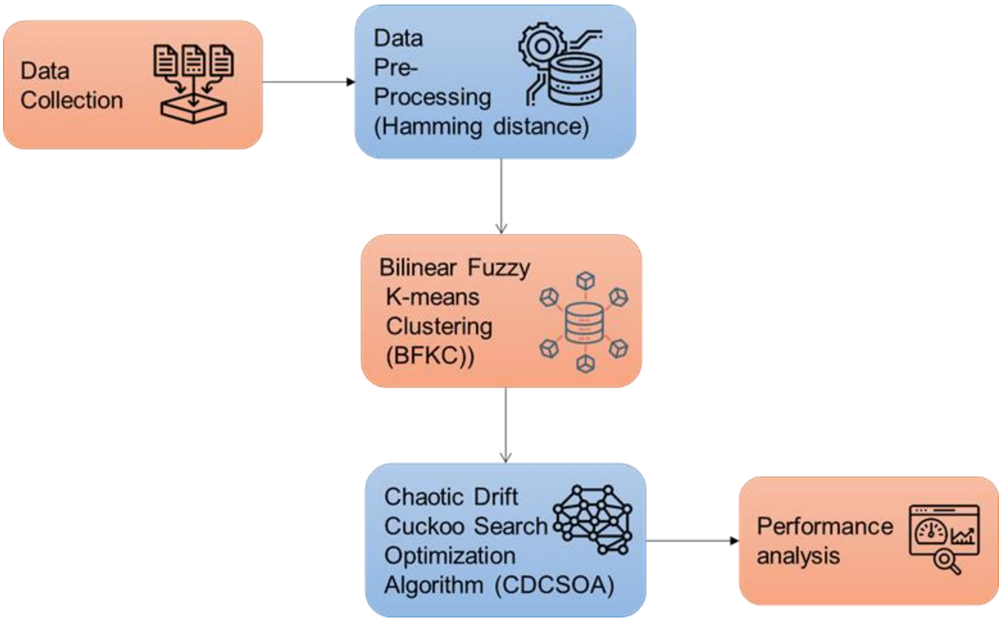

Fig.1. Proposed Framework

3.1 Data acquisition

Starting with various high-quality cardiac information sets, the suggested study begins with the database feature selection procedure. There are more blanks in its array of characteristics. We have found some unavailable values, and the feature extraction approach is generally utilized to filter out the elements that could be more effectively organized for our selection procedure. Our selecting characteristics and analysis method is applied to the gathered cardiac datasets. The Cleveland dataset is the gold standard for studying heart disease. While each dataset contains interesting outliers of varying types, some are more useful for making specific predictions or classes than others. The importance of accents in general for diagnosing heart disease is discussed. For instance, the effects of gender and age are studied separately.

3.2 Data preprocessing

Using the hamming distance technique, in which the difference among two binary vectors is calculated, the gathered datasets may be aggregated and cleaned. Incorporating the data's non-linearity into the cleaning process is facilitated by the hamming distance. Binary strings, often called bit strings, are another name for the Hamming distance. An ability on statements of a particular length over a character set depicting the amount of transformations to a word's pictures required to shorten it to another.

The set of all possible words of length n over the alphabet A is denoted by B.

Let b = (b1... bm) and a = (a1... am) be words in a.

The following mathematical model satisfies the Hamming distance:

$$\mathrm{c(b,a)} \geq 0 \text{ and } \mathrm{c(b,a)} = 0 \text{if and only if } \mathrm{b = a} \quad (1)$$

$$\mathrm{c(b,a)} = c(a,b); \quad (2)$$

$$\mathrm{c(b,a)} \leq c(b,u) + c(u,a) \quad (3)$$

Whereas u is the hammering weight, which is likewise a B measurement. The idea behind correction codes is to change the symbols used to convey the words from point B down a noisy channel. This assignment is proficient in detecting up to $[\delta - 12]$ mistakes and ordering up to $[\delta - 1]$ errors if the minimal hamming distance amongst words of B is. The same is true for string types of the same length. The hamming distance between two strings say s1 and s2, is calculated as E (s1, s2), and this value may change depending on the amount of elements that separate the two strings. If s1 and s2 are two strings, their separation is denoted by $\sum| si1 - si2|$. For an alphabet with the unique character 0, the distance $e\ (0, b)$ between all the zero-words in the sequence and the word A is the Hamming weight $u(B)$. The hamming duration for a linear code, expressed in a vector structure over a limited field, is proportional to its weight.

$$c(B,A) = u(B - A)$$
(4)

3.3 Bilinear fuzzy k-means clustering

Clustering, often known as cluster analysis, is a method used in machine learning for classifying unlabelled data. It is "a method of organizing data by forming groups of records that share common characteristics," as one definition puts it. The most similar items are kept together, while those with less or no characteristics are separated.

Despite its apparent lack of complexity, this technique serves as a building block for a wide variety of more complex clustering methods. According to this algorithm is both flat and exclusive. There are a variety of forms that this program might take. However, they all use iterative algorithms to estimate the following for a given number of clusters.

(1) Collecting cluster centers as points, these values represent a weighted average of the cluster values.

(2) To cluster the data, find the shortest distance from every specimen to the cluster's center, and put the sample there.

The interpretation of fundamental patterns in matrix profiling of genes has seen extensive usage of clustering, for which many different clustering methods have been developed. Due to its ease of use and high computational effectiveness, K - mean has become one of the most often used methods for gene clustering. However, the K-means algorithm must be more picky about where to put the first cluster nodes. The algorithm may quickly get trapped in a desirable state if the important centers are chosen randomly.

Data is divided into k clusters $Sl\ (l = 1,2...k)$ using the fuzzy k-means clustering method, with each cluster Sl having a representative (cluster center) Cl. A data point's association with a cluster's typical member is hazy at best. In other words, the degree of closeness between data point w_j^n and cluster cen is represented by an affiliation $w_j^n\ [0,1]$. Let's call this group of numbers S = Xi. The FKM method is predicated on reducing the following kind of errors:

$$i = \sum_{i=1}^{l} \sum_{j=1}^{M} w_j^n\ , c_{ji} \quad (5)$$

N is the number of information points, m is the fuzzy-factor parameter, k is the amount of clusters, and JCL is the symmetrical Euclidean radius among data point Yi and clusters indicative c_{ji}. The following restriction on $w_{j,i}$ must be observed.

$$\sum_{i=1}^{l} w_{j,i} = 1, for\ j = 1\ to\ M. \quad (6)$$

By dividing data points, FKM maps a set of typical vectors onto a better group of vectors. It starts with a seed set of cluster centers and iteratively maps them until some stopping requirement is met. No two clusters are assumed to have the same democratic cluster. It is recommended to disrupt a cluster center if two cluster centers overlap throughout the iterative phase. If $w_{j,i}$ is less than a small positive integer, then c_{ji}= 1 and c_{ji}= 0 for l > j. Here we introduce the fuzzy k-means clustering algorithm.

Algorithm 1: Bilinear Fuzzy K-means Clustering

Step 1: To begin, provide a value ε for the starting cluster centers $TD_0 = \{D_i(0)\}$. Set$o = 1$.

Step 2: Determine $c_{hi}$for $j = 1$ to M and $i = 1$ to, given the group of clustering centers 〚_〛. The following calculation should be used to update w_{ji} memberships:

$$w_{j,i} = \left[\left(c_{ji}\right)^{\frac{1}{n}-1} \sum_{k=1}^{l} \left[\frac{1}{c_{jl}}\right]^{\frac{1}{n}-1} \right]^{-1}$$

(7)

Step 3: To generate a new set of cluster delegates TD_{o+1},, we first need to calculate the center of each cluster using Eq.(8).

$$D_{i(o)} = \frac{\sum_{j=1}^{M} e_{ji}^{n} v_j}{\sum_{j=1}^{M} e_{ji}^{n}}$$

(8)

Step 4: Where >0 is a tiny positive value, the program will terminate if $\|D_i(o) - D_i(o-1)\| < \varepsilon$) holds for $i = 1\ to\ l$ if $p + 1 \to o$, go to the second step.

FKM is primarily hard to figure out because of steps 2 and 3. Step 3, on the other hand, is a lot easier to figure out than Step 2. So, concerning the number of distance computations, FKM has a computing difficulty of $O\ (Nkt)$, where t represents the number of rounds.

3.4 Bio-inspired optimization

In all of engineering, optimization is the most prominent mathematical issue. The most effective remedy may be found with its help. An optimization issue with widespread application is an ongoing research area. It has the capability to handle very computationally intensive optimization problems and may be either mechanistic or unpredictable. The proliferation of cutting-edge tools, materials, and processes is what we call "bio-inspiration." The biological advancement made throughout millions of years, including findings related to adaptability, method, and redemption, serves as inspiration. While many branches of study into natural phenomena had little funding in the 1980s and 2010s, the bio-inspired area saw explosive development, and a plethora of unique programs was effectively developed by 2020. Common names for these methods include swarm intelligence and computational evolution. The term "evolutionary computation" describes using the principle of survival of the fittest to improve programs. Inspired by the collective group brainpower of swarms or the cooperative behaviors of beetle enclaves and other species' occupants, "swarm intelligence" is an idiom used to decorate techniques for distributed problem-solvers. In order to increase the efficient of the recommended work for predicting the likelihood of heart illness, a bioinspired optimization method has been included.

3.5 Cuckoo search optimization

Cuckoo search is an effective method for optimizing populations. The females spread their eggs throughout the nesting sites of many bird species. Several optimization problems may benefit from the cuckoo search algorithm's inspiration from breeding behavior. Cuckoo search is one of the strategies drawn from the field of natural computation that is often used to find optimal solutions to engineering challenges. Using a shifting parameter may preserve a healthy equilibrium among local and global fluctuations while tackling global optimization problems. The approach could not evaluate the influence of a dynamic since the switching value was set at 25% due to a lack of study in this area. Numerous fields, including speech reorganization, work organizing, neural computing, and global optimization, use this technique extensively.

The cuckoo search method may be described mathematically as follows: Three hypothetical guidelines are presented to illustrate the CS approach:

- Each cuckoo takes turns laying one egg, which she then abandons in a randomly selected nesting box.

- Top-quality egg-laying hens will ensure that the healthiest households survive into the future.

- There is a limited supply of host dwellings, and the odds that a host will come upon an alien egg are Pa [0, 1]. In this case, the parent bird has two options: abandon the egg or relocate to start a new nest elsewhere.

The key developments of the CS may be distilled down to the following three principles:

When coming up with new answers $v_j^{(s+1)}$for, say, a cuckoo, it defines a L'evyflight as follows:

$$v_j^{(s+1)} = v_j^{(s)} + \alpha \oplus Levy(\lambda)$$

(9)

Where $v_j^{(s)}$i is the cuckoo's initial position, is a stride size and beneficial integer tuned to the searched space parameters, is the entering wise multiplying, and is the levy coefficient.

The preceding stochastic calculation describes a random walk. A L'evy distribution is used to generate an arbitrary step size.

$$Levy(\lambda) \sim w = s^{-\lambda}, (1 < \lambda \leq 3)$$

(10)

Where λ characterizes the rate of decrease of the PDF as a function of time.

L'evy-flight is unique in that it successively increases population diversity, which helps the algorithm escape the local minimum.

$$Levy(\lambda) \sim \frac{\phi \times w}{|x|^{1/\lambda}}$$
(11)

It defines in the case when w and x have normal distributions:

$$\phi = \left[\frac{\gamma(1+\lambda) \times sin(\pi \times \lambda/2)}{\gamma\left(((1+\lambda)/2) \times \lambda \times 2^{(\lambda-1)/2}\right)}\right]^{1/\lambda}$$
(12)

Where γ is the conventional gamma distribution, and ϕ is the convergence ratio.

In conclusion, the suggested method avoids the local optimum using random walk with Levy-flight, significantly improving the algorithm's global search capabilities. It's also been seen to provide more fruitful outcomes, particularly concerning single-modal and several-modal reference services.

The qualities of L'evy flying, in which a large followed PDF usurps the advancement lengths, explain why many animals and insects search for sustenance in a seemingly random or erratic fashion. Due to the infinite resources and versatility of L'evy flights, some new configurations will be centered on the most successful ones at the present moment. In contrast, others will be placed in a different relative position. This feature enables CS to explore the inquiry space efficiently, unlike other algorithms that use regular Gaussian interaction, which tend to get stuck on local minima. Figure 2 illustrates the connection between DBNs and the cuckoo search method. Algorithm 2 details the stages of the cuckoo search algorithm.

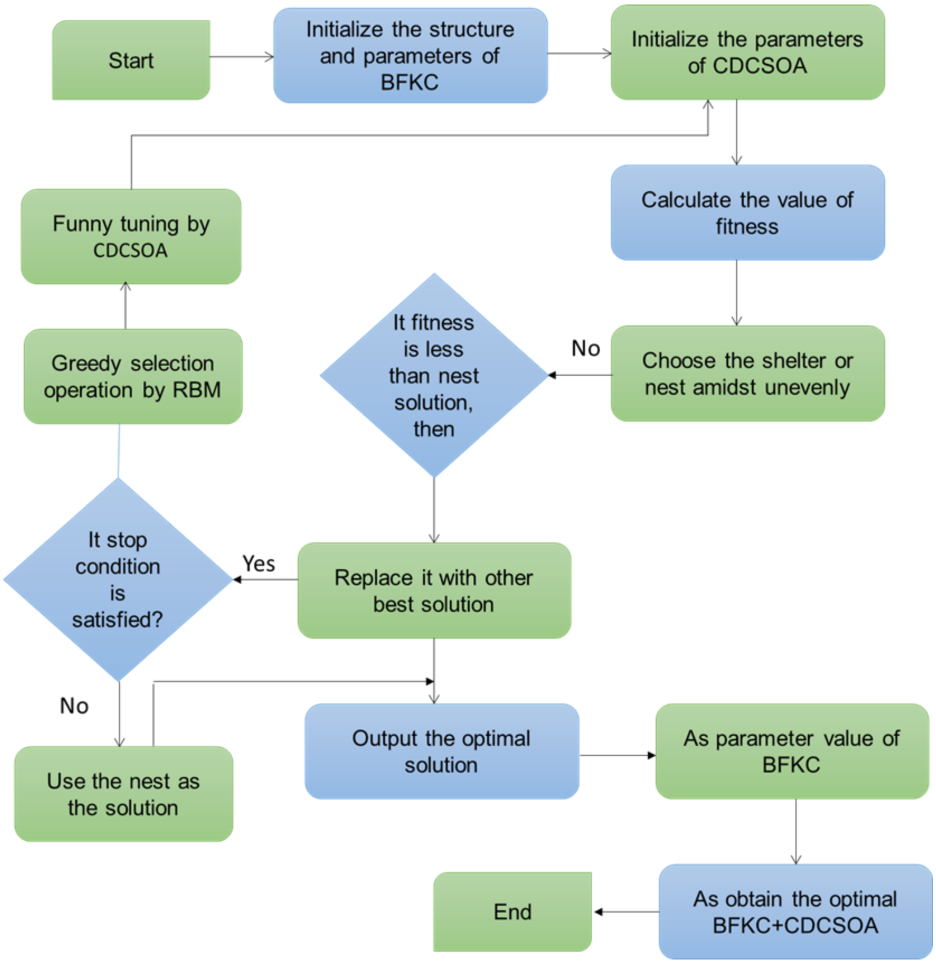

Fig. 2. Flowchart of the suggested method (BFKC-CDCSOA)

Algorithm 2: Chaotic Drift Cuckoo Search Optimization Algorithm

Step 1: Initialize the process

Step 2: Establish an artificial population

Step 3: Heavy-tailed cuckoo-bird obtaining.

Step 4: Determine the monetary benefit of fitness.

Step 5: Pick a hiding place in this world of uneven

Step 6: Assuming a nest resolution exists, if Efficiency is less than it,

Step 7: Find an alternative that works better

Step 8: Take refuge in a nest

Step 9: Discard the fractionally worse nest and build a new one using heavy-tailed spiders for stages 7 and 8.

Step 10: Keep track of the present moment's optimal solution

Step 11: If nests are fewer than or equivalent to the highest number of variations, go to Step 3.

Step 12: Putting a stop to it

CSO is a population-level algorithm, much like GA and PSO. The first benefit is that any size step is possible when the amount of steps is heavy-tailed, making randomness more effective. The second perk is that, unlike GA and PSO, it requires fewer tuning parameters, allowing it to be used for various optimization problems. Additionally, each nest may stand in for a group of solutions, making CS a meta-population technique. The original CDCSOA may be classified into three subtypes according to the kind of understanding used to solve the issue. Both local search strategies and productive heuristics are examples. The third kind is a combination of elements from multiple algorithms. Identifying the presence or absence of cardiac illness is the intended goal of this study, and the third category is carried out using an adaptation of BFKC to classify, with CDCSOA employed for the final tuning of the variables.

4. Result and Discussion

This section provides the findings and explains the different heart disease forecast methods tried using the suggested model. Using the confusion matrix, it is possible to derive many performance indicators. Accuracy, precision, specificity, and the F1 score are these. The findings favor the BFKC over competing models.

4.1 Accuracy

A measurement's, an analysis's, or an outcome's "accuracy" refers to its precision or correctness. This concept is crucial to several disciplines, including science, engineering, Medicine, economics, and more. Accuracy is sometimes used as a proxy for quality, appropriateness, or dependability when assessing the effectiveness of structures, designs, or data. The correct categorization of various cases in the dataset relies heavily on accuracy. The following mathematical formula expresses it:

$$Accuracy = TP + TN / TP + FN + FP + TN \quad (13)$$

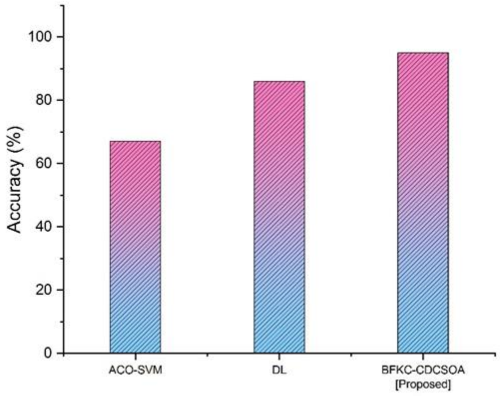

Fig.3. Accuracy of existing and proposed method

Table 1. Accuracy comparison

Methods	Accuracy (%)
ACO-SVM	67
DL	86
BFKC-CDCSOA [Proposed]	95

The statistics imply that the proposed approach may perform better than the current standard research methodologies, as shown in Figure 3 and Table 1. The proposed technique outperforms state-of-the-art methods like ACO-SVM (67% accuracy) and DL (86% accuracy). This demonstrates that our approaches (CDCSOA 95%) are superior to the gold standard.

4.2 Precision

Precision is the proportion of True Positives (positive samples that were accurately identified) relative to the overall proportion of identified positive samples. Data retrieval heavily uses it to forecast true values from a given dataset. Accuracy was calculated using the following formula:

$$Precision = TP / TP + FP \quad (14)$$

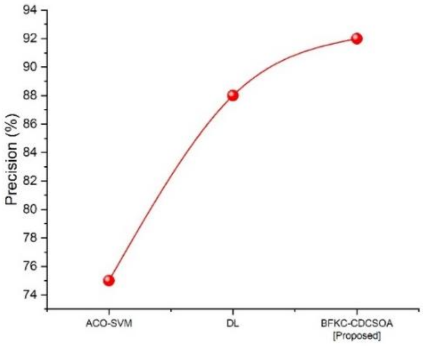

Fig. 4. Precision of existing and suggested method

Table 2. Precision comparison

Methods	Precision (%)
ACO-SVM	75
DL	88
BFKC-CDCSOA [Proposed]	92

The precision measurements provide similar results in Figure 4 and Table 2. The proposed technique has a higher precision of 92% compared to regularly used methodologies such as ACO-SVM (75% precision) and DL (88% precision). The algorithm accurately predicts the false category for both true and false occurrences. A false positive is produced when an approach incorrectly identifies a true class. When the system correctly predicts the negative outcome, this is known as a false negative.

4.3 Specificity

The percentage of false negatives in a binary test is classified as false positives. Predicting 100% true negatives using the fraction of observed negatives is a use of specificity. It is expressed by,

$$Specificity = TN / TN + FP$$
(15)

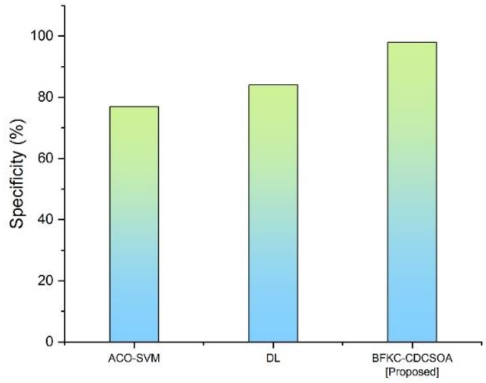

Fig.5. Specificity of existing and suggested method

Table 3. specificity comparison

Methods	Specificity (%)
ACO-SVM	77
DL	84
BFKC-CDCSOA [Proposed]	98

When compared to other techniques, the one we offer achieves more specificity. Figure 5 and Table 3 show the comparison of specificity.

4.4 F1-score

The harmonious average of the memory and accuracy scores shows the F1 score. It deals with false positives and negatives, making it work well on an uneven dataset. The method for figuring out the F1 score looked like this:

$$F1\ score = 2 * (Precision * Recall) / (Precision + Recall)$$
(16)

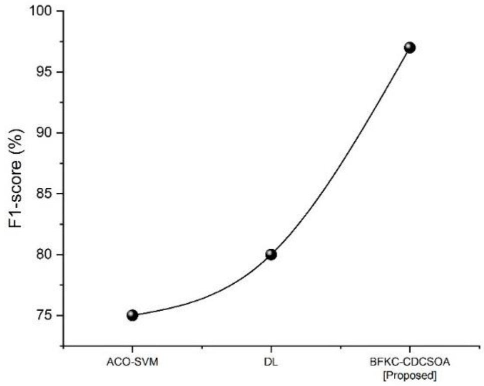

Fig.6. F1-score of existing and proposed method

Table 4. F1-score comparison

Methods	F1-score (%)
ACO-SVM	75
DL	80
BFKC-CDCSOA[Proposed]	97

The F1-score rate of the current method is better than the F1-score rate of the existing process. To make it easier to compare, the effectiveness of the old treatment stays the same over time, while the BFKC-CDCSOA method remains the same within a small range. The BFKC-CDCSOA method gets better over time within its steady state range. Figure 6 and Table 4 show how the results compare to each other.

5. Conclusion

This paper introduced a BFKC equipped with a feature selection approach for analyzing and accurately forecasting cardiac illness. The steps to prepare the data and the resulting numbers are neatly summarised. The hamming distance feature selection approach was used for many cardiac datasets during data pre-processing. Once relevant characteristics have been selected, they are sent on to the classification phase, where they are analyzed using clustering approaches like BFC integrated with a cuckoo search bio-inspired algorithm at varying depths to provide a reliable prediction of heart illness. The suggested forecasting system was also shown to outperform competing models. Combining enormous databases with the deep learning categorization models incorporated in bio-inspired algorithms on heart disease ECG data might increase performance in the future.

References

[1] Shobana, R. and Shanthi, R., 2023. Prediction of Cardiovascular Diseases Using Bio-Inspired MKCB Optimization Algorithm. Journal of Survey in Fisheries Sciences, 10(2S), pp.3651-3660.

[2] Metan, J., Prasad, A.Y., Kumar, K.A., Mathapati, M. and Patil, K.K., 2021. Cardiovascular MRI image analysis by using the bio-inspired (sandpiper optimized) fully deep convolutional network (Bio-FDCN) architecture for automated detection of cardiac disorders. Biomedical Signal Processing and Control, 70, p.103002.

[3] Nandakumar, P. and Narayan, S., 2022. Cardiac disease detection using cuckoo search enabled deep belief network—intelligent Systems with Applications, 16, p.200131.

[4] Sureja, N., Chawda, B. and Vasant, A., 2022. A novel salp swarm clustering algorithm for prediction of heart diseases. Indonesian Journal of Electrical Engineering and Computer Science, 25(1), pp.265-272.

[5] Yadav, P., Arora, N. and Sharma, S.C., 2023, March. Performance Analysis of Nature-Inspired Optimization Algorithms for Chronic Disease Prediction. In 2023 International Conference on Device Intelligence, Computing and Communication Technologies,(DICCT) (pp. 1-6). IEEE.

[6] Navin, K.S., Nehemiah, H.K., Nancy Jane, Y. and Veena Saroji, H., A classification framework using filter–wrapper based feature selection approach for diagnosing congenital heart failure. Journal of Intelligent & Fuzzy Systems, (Preprint), pp.1-36.

[7] Ghoniem, R.M., 2020. A novel bio-inspired deep learning approach for liver cancer diagnosis. Information, 11(2), p.80.

[8] Mastrothanasis, K., Zervoudakis, K., Kladaki, M. and Tsafarakis, S., 2023. A bio-inspired computational classifier system for evaluating children's theatrical anxiety at school. Education and Information Technologies, pp.1-24.

[9] Sathiyabhama, B., Kumar, S.U., Jayanthi, J., Sathiya, T., Ilavarasi, A.K., Yuvarajan, V. and Gopikrishna, K., 2021. A novel feature selection framework based on grey wolf optimizer for mammogram image analysis. Neural Computing and Applications, 33(21), pp.14583-14602.

[10] Pradhan, M., 2023. Cardiac Image-Based Heart Disease Diagnosis Using Bio-Inspired Optimized Technique for Feature Selection to Enhance Classification Accuracy. In Machine Learning and AI Techniques in Interactive Medical Image Analysis (pp. 151-166). IGI Global.

[11] Sharma, P., Choudhary, K., Gupta, K., Chawla, R., Gupta, D. and Sharma, A., 2020. Artificial plant optimization algorithm to detect heart rate & presence of heart disease using machine learning. Artificial intelligence in Medicine, 102, p.101752.

[12] Patro, S.P., Nayak, G.S. and Padhy, N., 2021. Heart disease prediction by using novel optimization algorithm: a supervised learning perspective. Informatics in Medicine Unlocked, 26, p.100696.

[13] Mian, T.S., 2022, December. A Comparative Analysis of Intelligent Machine Learning for Cardiovascular Disease Classification. In 2022 International Conference on Intelligent Technology, System and Service for Internet of Everything (ITSS-IoE) (pp. 1-7). IEEE.

[14] Sree Sandhya, N. and Beena Bethel, G.N., 2021. Impact of Bio-inspired Algorithms to Predict Heart Diseases. In Smart Computing Techniques and Applications: Proceedings of the Fourth International Conference on Smart Computing and Informatics, Volume 2 (pp. 121-127). Springer Singapore. Existing-□> ACO-SVM

[15] Pandiyan, N. and Narayan, S., 2023. A Survey on Deep Learning Models Embed Bio-Inspired Algorithms in Cardiac Disease Classification. The Open Biomedical Engineering Journal, 17(1). Existing-□> DL

[16] Mr. Kaustubh Patil. (2013). Optimization of Classified Satellite Images using DWT and Fuzzy Logic. International Journal of New Practices in Management and Engineering, 2(02), 08 - 12. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/15

[17] Patil, R. A. ., & Patil, P. D. . (2023). Skewed Evolving Data Streams Classification with Actionable Knowledge Extraction using Data Approximation and Adaptive Classification Framework. International Journal on Recent and Innovation Trends in Computing and Communication, 11(1), 38–52. https://doi.org/10.17762/ijritcc.v11i1.5985