Cardiac disease detection using cuckoo search enabled deep belief network

Nandakumar P, Subhashini Narayan[*]

School of Information Technology and Engineering, Vellore Institute of Technology

ABSTRACT

Cardiac disease is the most infected disease in the world nowadays for all ages of people. An emergency need arises to predict cardiac disease accurately in a short time. In this article, hamming distance feature selection method is proposed for the data preprocessing and data cleaning process in different cardiac disease datasets. Deep learning model such as deep belief networks is used with cuckoo search bio-inspired algorithm for finding the accurate prediction of cardiac disease. The results demonstrate that deep belief networks with the cuckoo search algorithm have achieved good performance with an accuracy of 89.2% from Cleveland, 89.5% from South Africa, and 89.7% from Z-Alizadeh Sani, 90.2% from Framingham, and 91.2% from Statlog cardiac disease datasets.

1. Introduction

Over the decades, people in the world have faced various life-threatening diseases, among that heart disease has reached a great deal of observance in medical research. It is one of the noticeable illnesses that influence numerous individuals during their center or advanced age and in a few cases, it, at last, prompts deadly difficulties. Heart disorders are more routine in men than in ladies. The detection of heart illness is vital in the prediction process. An ideal finding of heart disorder is significant in diminishing well-being chances and forestalling heart failures. Some of the risk factors faced by the people that will rapidly increase heart diseases are smoking, age, ethnicity, family history of the disease, high blood pressure, high blood cholesterol, diabetes, poor diet, lack of exercise, obesity, stress, and vessel inflammation. The World Health Organization (WHO) (Li, Hu & Gu, 2020) lists cardiovascular diseases as the dominant reason for their demise globally with 17.9 million people departing every year. A pertinent threat endured by healthcare organizations, such as hospitals and medical centers, is the allocation of worth amenities at moderate costs.

Heart disease is a predominant reason for demise today, with COVID-19 symptoms. In addition to it, coronary heart disease (CHD) is the most well-known type of cardiovascular sickness that records around 2% of demise in India compared with the world count of 17.9 million. In the last 20 years, the research papers (Saliba et al., 2020) published on heart diseases showed that cases of CHD are increased by 100% which will lead to a complex situation or even an epidemic for the next 10 years in the world. Cardiovascular disorders endure crucial stress worldwide and

are considered for 30% of fatalities in the world. There are different heart disease names related to CVD and CHD are coronary artery disease (CAD), heart attack, heart failure, heart muscle disease (cardiomyopathy), heart inflammation (endocarditis), heart valve disease, pericardial disease, stroke, and vascular disease (blood vessel disease) and its types are shown in Fig. 1 and the clinical representation of cardiac disease types is shown in Fig. 2. Among them, our research will be focused on the prediction of heart diseases (Almustafa, 2020) such as hypertrophic cardiomyopathy (HCM). Blood transfusion is a term used for transferring blood from one person to the other. It is one of the major causes that will affect a person with heart disease in the future. It is used in cardiopulmonary bypass (CPB) which results in bleeding by the type of procedure the physicians follow such as the urgency of surgery, age of the patient, preoperative platelet count, and deep hypothermic circulatory arrest. Due to the lack of healthcare service from the physician, the need for an automatic diagnosis system (Tuli et al., 2020) emerged with artificial intelligence as a root concept, where then machine learning came into the picture with a lot of models focused on various human health-related problems. For more computational performance in diagnosing cardiac disease classification type, in contrast to traditional machine learning methods the deep learning technologies have recently emerged with support from bio-inspired algorithms that are mainly used for solving optimization problems.

Deep Learning is a sub-field of machine learning that emerged as a noticeable model in the world specifically for the disease classification field. The deep learning concept was first initiated by Walter Pitts and Warren McCulloch in the year 1943 through a digital pattern that is
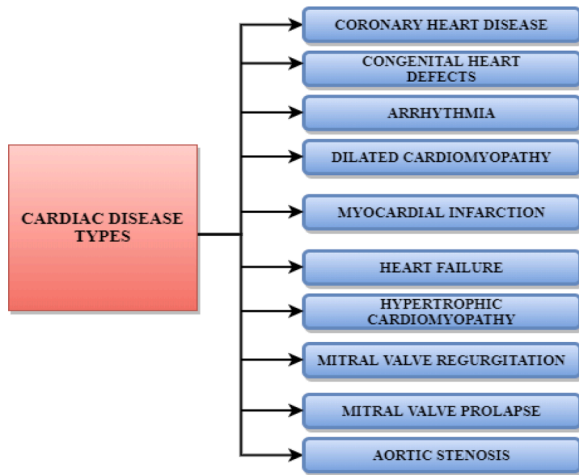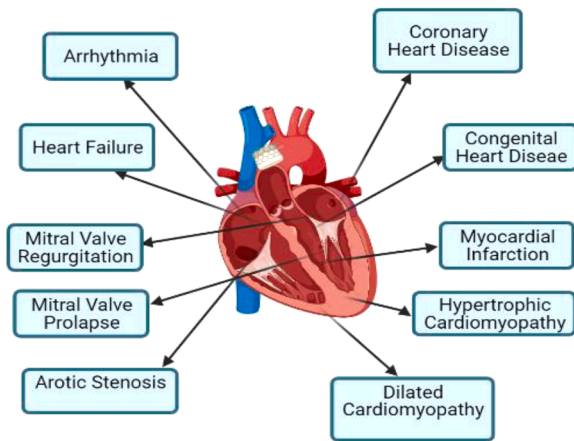
Fig. 1. Types of Cardiac Disease.

Fig. 2. Clinical representation of Cardiac Disease.
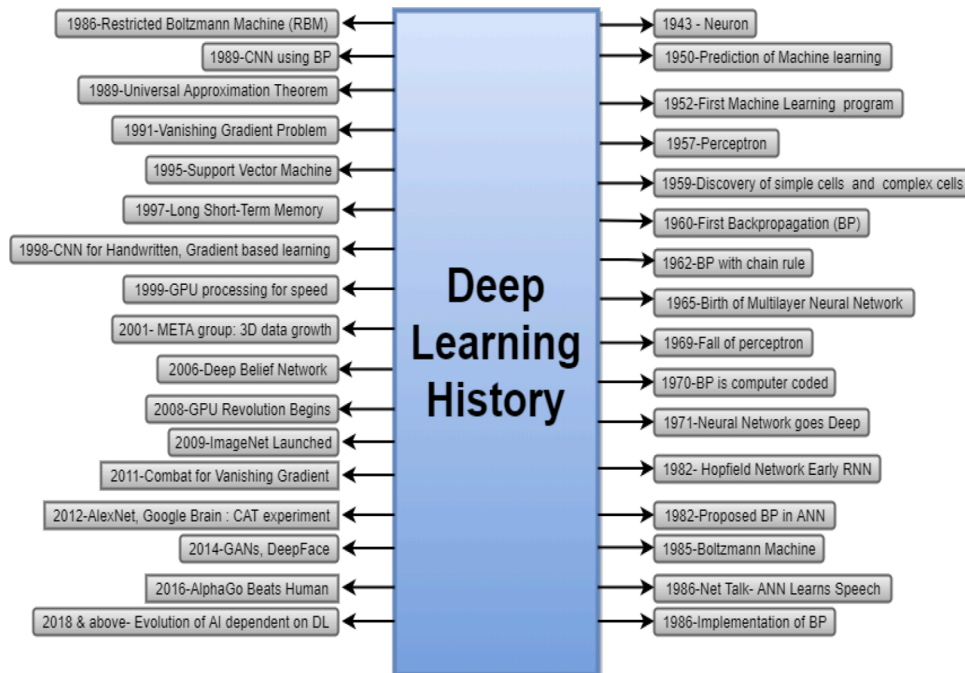
relying on the neural networks of the human brain. With the continuation of the existing model, Henry J Kelley developed Back Propagation Model in 1960 and the clearer kind of it was developed by Stuart Dreyfus in 1962 is based on the chain rule. Both of these models were inefficient and not be utilized in the research until 1985. An effort has been taken for developing deep learning algorithms by Ivakhnenko and Lapa in 1965 with models of concept from polynomial activation functions. Later, in 1979 the first convolutional neural networks were used by Kunihiko Fukushima and developed an artificial neural network that is named Neocognitron and which is used to recognize visual patterns. Neural networks and deep learning-based research emerged in the year 1985 to 1990, with the invention of the Support Vector Machine by Dana Cortes and Vladimir Vapnik in the year 1995 and long short-term memory in 1997, by Sepp Hochreiter and Juergen Schmidhuber.

With the adoption of Graphics Processing Units, there was a significant advancement in deep learning for the process of better and faster results. The influence of deep learning in the industry began in early 2000, with the vanishing gradient problem and it was solved by layer-by-layer pre-training and LSTM. In 2009, Fei-Fei Li, established ImageNet a free database that is used for storing more than 14 million labeled images that are useful for training neural nets. Deep learning significantly holds high efficiency and speed in the year 2011 and 2012, to the speed of GPUs and with the invention of AlexNet a CNN-based architecture developed by Krishevsky et al. The history of Deep learning evolution is shown in Fig. 3.

Bio-inspiration is the blooming of modern structures, resources, and methods. It is stimulated by results found in adaptation, procedures, and salvation, which have biologically progressed over millions of years. The concept is to enhance the sculpting and replication of the biological system to better understand nature's analytical constructional features, such as an extension, for use in bioinspired design advancement. The bio-inspired field starts in the early 1980s and 2010s with many types of research on the natural phenomenon in limited support but in 2020 it's been a huge growth in the variety of novel algorithms that have been defined successfully. Complex optimization problems (Martinez et al., 2021) such as designing of topology, tuning, and learning of hyperparameters can be solved efficiently using deep learning optimization with the adoption of bio-inspired algorithms. Diagnosing diseases from

Fig. 3. Deep Learning History.

many patients through the data collected with a different category of a specific disease is a classification problem. For the Cardiac disease classification type, the data will be big, so with the support of deep learning methods embed with bio-inspired algorithms it will be able to classify it more accurately in diagnosing. Most of the early machine learning techniques predict various diseases type with accuracy suitable for small datasets (Gárate-Escamila, Hajjam El Hassani & Andrès, 2020), for large datasets deep learning optimization with a bio-inspired algorithm, can predict it with more accuracy, which leads to the idea of embedding deep learning with bio-inspired algorithms for cardiac disease classification.

The proposed idea is mainly on predicting heart diseases using the deep belief network model which is the type of deep learning method. This investigation means to distinguish the key patterns among various sorts of Deep Learning algorithm, and their performance and utilization for heart disease risk prediction. A bio-inspired algorithm is additionally used for anticipating coronary illness with the help of classification algorithms. In this study, popular datasets have been listed and analyzed in the prediction process of Cardiac disease with their performance using various deep learning techniques.

In this paper, the main contributions are as follows,

- Heart disease is one of the important illnesses that need to be addressed by detecting it earlier.
- A deep learning model such as Deep Belief Network is used for the detection of cardiac disease.
- The bio-inspired algorithm is further used for the optimization purpose that enhances accuracy in the detection of cardiac disease.
- So, with the help of the Cuckoo search algorithm with a deep belief network for cardiac disease detection, we had achieved an accuracy of 89.2% from Cleveland, 89.5% from South Africa, 89.7% from Z-Alizadeh Sani, 90.2% from Framingham, and 91.2% from Statlog heart disease datasets.
- The results obtained by the proposed model show better when compared with other machine learning and deep learning models.

The proposed research work is planned as follows: Section 2 contributes a detailed description of the literature survey of various results and methods developed by several authors according to cardiac disease prognosis. The existing methodology is explained in Section 3. In Section 4, for data pre-processing work, hamming distance is proposed. In Sections 5 and 6, the proposed ensemble classification of deep learning algorithms is utilized and compared with other algorithms for better accuracy of cardiac disease prediction. In Section 7, the obtained results are described with suitable simulations, and tables with the collected dataset from a well-known data sets repository, and finally, in Section 8, the research work is concluded with the future work suggestions for implementing it in real-time applications.

2. Literature survey

The following survey is based on the various machine learning and deep learning models that are used by the researchers for the prediction of cardiac disease which comes among our proposed topic is analyzed with their dataset details as follows:

Nasarian et al. (2020) have developed a novel hybrid feature selection algorithm called heterogeneous hybrid feature selection (2HFS) that is used for mainly feature extraction. For implementation purposes, they have taken the dataset of CAD from Nasarian, and some popular algorithms are used such as decision tree (DT), Gaussian Naive Bayes (GNB), random forest (RF), and XGBoost Classifiers. They used the dataset which contains business-related highlights notwithstanding other clinical characteristics. The model they had proposed is additionally applied to three notable UCI CAD datasets to be specific Hungarian, long_beach_va, and Z-Alizadeh Sani datasets. In the future, the testing process can be enhanced by immense datasets with more aspects.

Additionally, they had planned to apply different evolutionary algorithms and ensemble learning with other new features.

Abdar (2017) have used the Cleveland dataset, which includes 303 records, and found a better decision tree algorithm that is used for extracting rules in predicting heart disease or CAD or CVD. They had proposed a C5.0 algorithm with an accuracy provided of about 85.33%. The main acclaimed reasons for predicting heart disease are a combination of attributes such as trestbps, restecg, thalach, slope, oldpeak, and cp. In the future, this model may be used for identifying the factors influencing heart patients.

Ann (2019) have designed an effective model using several machine learning and data mining methods to detect CAD. The method used for that is the diagnosis support system (DSS). Also, they had implemented two bio-inspired algorithms for optimization purposes, one is the genetic algorithm and the other is particle swarm optimization. They also suggested using other pre-processing approaches in the future rather than data normalization, such as standardization, variance scaling, an encoding of categorical features, and non-linear transformation.

Alizadehsani et al. (2013) have implemented several algorithms on the Z-Alizadeh Sani datasets. They are sequential minimal optimization (SMO), Naive Bayes, bagging with SMO, and neural networks. It can be used for analyzing the datasets. In the future, with larger datasets, more features, and some broader data mining approaches utilized to achieve better and very valuable results in predicting or diagnosis CAD.

Alizadehsani, Hossein and Javad (2016) have explained the detection of CAD by support vector machine (SVM) on the Z-Alizadeh Sani datasets. The implementation part is carried out with two approaches in the selection of features. They are by selecting distinct features for each artery separately and by selecting the same features for all the arteries. The two feature selection methods used are average information gain and combined information gain. In the future, the extension of datasets which accumulate more patient details for predicting it more accurately. New research areas are also suggested such as LAD (Left Anterior Descending), LCX (Left Circumflex Artery), and RCA (Right Coronary Artery) in the CAD detection process.

Arabasadi, Alizadehsani, Roshanzamir and Moosaei (2017) have developed a detection model based on artificial neural networks and genetic algorithms for predicting CAD. Datasets are taken from Z-Alizadeh Sani which contains information on 303 patients and 216 of them are suffering from CAD. In the future, instead of a genetic algorithm, they suggested some other bio-inspired algorithm to enhance the performance of the proposed method by comparing it with different types of neural networks for testing.

Kivimäki and Steptoe (2017) have identified the reason behind the progression of CVD related to the stress effect of adulthood. Many reasons they had identified such as hypertension, high serum cholesterol levels, diabetes, smoking, and obesity in the people at a young age itself. They had gone through street studies for the identification of evidence on the role of stress in disease progression in individuals at critical risk or with CVD.

Kolukisa et al. (2019) have evaluated distinctive classification algorithms and proposed a novel hybrid feature selection method for the diagnosis of CAD. For testing, they had taken resources from UCI (ML) repository and Z-Alizadeh Sani datasets.

Yamamoto et al. (2020) had clarified the relationship between physical performance and prognosis among patients with heart failure (HF). They had changed the assessment parameters with the aged patients with heart failure for better performance using meta-analysis.

Gong et al. (2020) have identified complicated medicinal syndromes that lead to ventricular filling or impaired ejection and characteristics of heart failure defined by the treatment difficulties, poor prognosis, and high mortality. From the data of the Shanxi academy of medical sciences, they had found the reason for heart failure patients by building a 30-day mortality prognosis method. For this, they had used the recurrent attention model and the doctors can easily stipulate drugs conforming to the symptoms.

D'Alto et al. (2019) have compared several reasons for CHD with various medical terms of complex situations that arise in gender differences which will affect the patients a lot. Also, they stated that pregnancy and sex-related evidence for increased risk factors that lead to CHD.

Al-Makhadmeh and Tolba (2019) had proposed an IoT medical device named (HOBDBNN) Higher-Order Boltzmann Deep Belief Neural Network for heart disease prediction. This system or model is also used in the process of finding big data and recognizing abnormal heart patterns in the minimum time. They also suggested that in the future, developing the IoT-based process using optimized techniques with an effective feature selection process.

Mohan, Thirumalai and Srivastava (2019) had developed a novel model for predicting heart disease called (HEFLM) Hybrid Random Forest with a Linear Model. The authors also suggested enhancing this research with different mixtures of ML techniques to better prediction methods for improving the performance of heart disease prediction.

Jin et al. (2018) have used (EHRs) electronic health records for analyzing and predicting heart diseases. They have used a neural network concept for predicting the possibilities of cardiac disease depending on the patient's electronic medical data. They also used LSTM methods and compared them with several other popular methods such as LR, RF, and AdaBoost. They also suggested expanding this model in the future for health care applications.

Mufudza and Erol (2016) have used the model that is based on clustering techniques that can predict and identify heart disorder via Poisson mixture regression models. Based on the sex, exercise, and level of chest pair factor they had shown the risk factors for the prediction process for heart disease. The work done previously in heart disease prediction with dataset details is shown in Table 1.

3. Existing methodology

From the above literature survey, we have found that the various machine learning and deep learning models such as decision trees, ensemble model, Gaussian Naive Bayes, logistic regression, random forest, support vector machine, long short-term memory, and artificial neural network were used for the prediction of cardiac-related diseases. With these models collectively, they have given the accuracy at an average of 85% only for the datasets utilized from the UCI machine learning repository and also from the clinical open-source database. The proposed work used the method of deep learning models such as deep belief networks embedded with bio-inspired algorithms and produces an accuracy greater than the existing models such as 90% and above. The obtained results are further enhanced for more accuracy in an optimized way by applying feature selection concepts that are not used in existing types. The pertinent problem in machine learning techniques is the high dimensionality of the datasets that leads to the anatomy of many features which results in a huge amount of storage requirements. For this reason, the proposed model is used to store a huge amount of dataset features and execute them in a fast manner. The final results obtained are compared with the early machine learning models in terms of performance and accuracy. The various research concepts used in this work assure that it will produce highly accurate results.

4. Proposed methodology

The proposed architecture is shown in Fig. 4. The proposed methodology work is explained with the feature selection methodology, followed by the data cleaning process by hamming distance, and finally, the selection of important features is done. Cardiac disease datasets have been collected and analyzed which are then fed into our cleaning model for pre-processing. In the data pre-processing step, missing values are handled and data is cleaned with duplicates and removed unnecessary attributes with hamming distance feature selection method. Hamming distance is utilized in our proposed work for the estimation of the distance of heart disease dataset features such as age, trestbps, chol, thalach, oldpeak, cp, restecg, slant, ca, thal. The next step is to cross-validate the dataset with training and testing data. The training set is processed with machine learning classification models such as decision

Table 1

Various machine learning and deep learning methodologies accuracy for heart disease prediction with dataset details.

| Author Name	Reference	Year	Models	Dataset	Accuracy
Hossain M. E et al.,	(Hossain, Uddin & Khan, 2021)	2021	Logistic Regression, SVM, DT, RF, NB, and KNN	CBHS health fund in Australia	79% to 88%
Beunza J et al.,	(Beunza et al., 2019)	2019	Decision Tree, Random Forest, Support Vector Machines, Neural Networks, and Logistic Regression	Framingham Heart Study	85%
Dutta A et al.,	(Dutta, Batabyal, Basu & Acton, 2020)	2020	Efficient Neural Network with Convolutional Layers	National Health and Nutritional Examination Survey	79.5%
Spencer R et al.,	(Spencer, Thabtah, Abdelhamid & Thompson, 2020)	2020	Chi-Squared Feature Selection with The Bayes net Algorithm	The Cleveland-Dataset	85%
Jin B et al.,	(Jin et al., 2018)	2018	Long Short-Term Memory (LSTM) Neural Network Model.	Electronic Health Record (EHR) Data From Real-World Datasets	Receiver operator characteristics ROC & Area under the curve AUC =68.27
Amin Ul Haq et al.,	(Haq et al., 2018)	2018	K-NN, ANN, SVM, DT, and NB	Cleveland Heart Disease Dataset 2016	83%
Mufudza C, Erol H	(Mufudza & Erol, 2016)	2016	Poisson Mixture Regression Model	Cleveland Clinic Foundation Heart Disease Data Set	86.7%
Dwivedi A	(Dwivedi, 2018)	2018	Artificial Neural Network (ANN), Support Vector Machine (SVM), Logistic Regression, K-Nearest Neighbor (KNN), Classification Tree, and Naïve Bayes	Statlog Heart Disease Dataset	85%
Orphanou K et al.,	(Orphanou, Stassopoulou & Keravnou, 2016)	2015	Extended Dynamic Bayesian Network (DBN) Model	Stulong	85.5%
Tay D et al.,	(Tay, Poh, Van Reeth & Kitney, 2015)	2015	Support Vector Machine (SVM)	Cardiovascular Health Study (CHS) Dataset	84.8%
Henriques J et al.,	(Henriques et al., 2015)	2015	Telemonitoring Data model	Historical Dataset - My heart Telemonitoring Study	78% and 80%
Dogan M et al.,	(Dogan, Grumbach, Michaelson & Philibert, 2018)	2018	Random Forest Classification Model for Symptomatic CHD And Integrated Genetic-Epigenetic Algorithms	Framingham Heart Study	78%
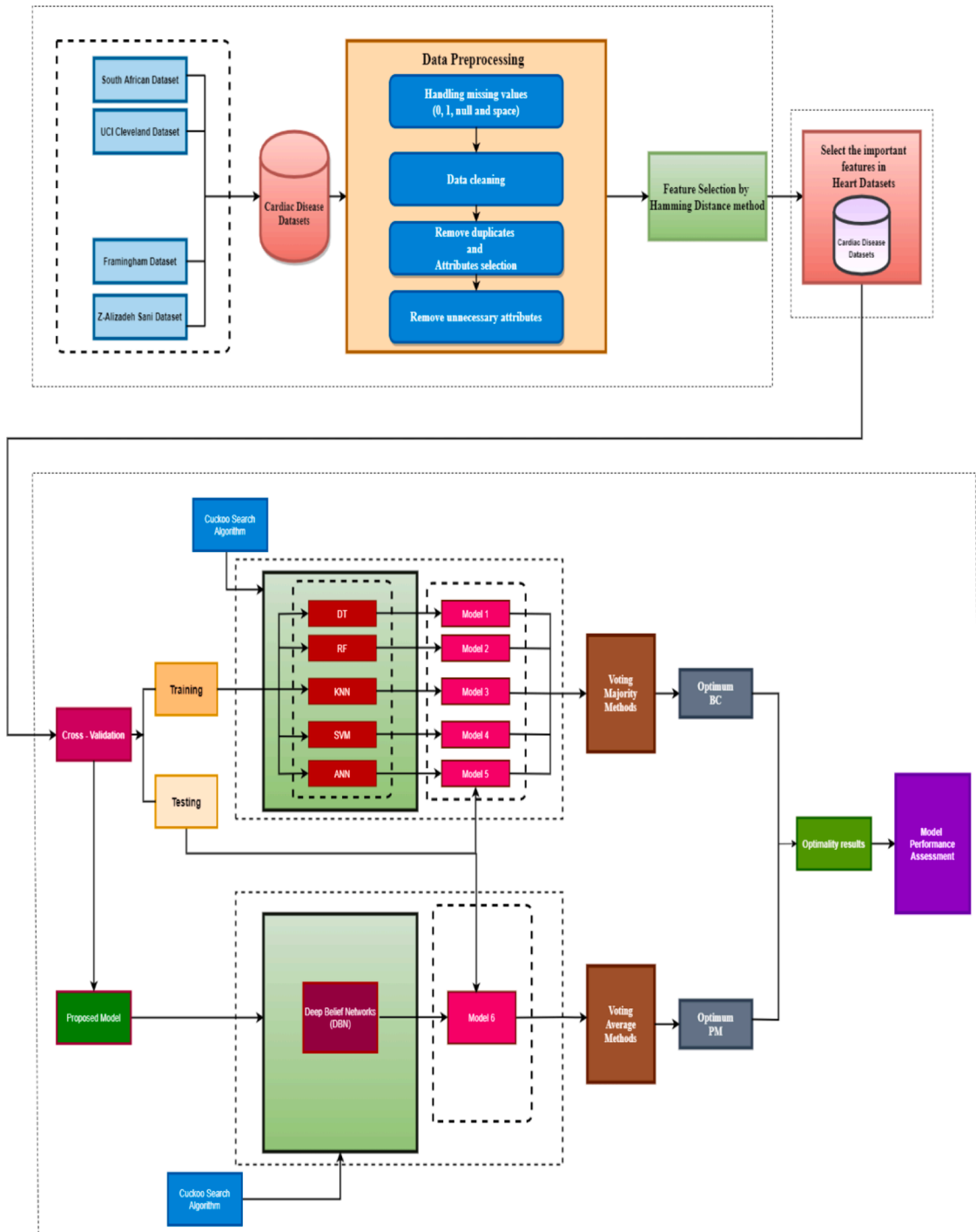
Fig. 4. The Architecture of Proposed Work.

trees, Naïve Bayes, logistic regression, and support vector machine. The testing set is processed with an optimized classification of deep learning models such as deep belief networks combined with a cuckoo search bio-inspired algorithm. After comparing the prediction results of both training and testing data with accuracy, the results entered in Tables 6–10 show that deep belief networks with cuckoo search provide the best accuracy for the prediction of cardiac disease.

4.1. Dataset feature selection methodology

The proposed work is initiated with the dataset feature selection process with a collection of different cardiac datasets that are collected

from a reputed source. It has a greater number of features with missing values present in it. So, to reduce the features which are not well organized for our selection process, we have identified some missing values and then the feature selection method is widely used to filter them. The feature selection by hamming distance is shown in Fig. 5. The collected cardiac datasets are used for our feature selection and analysis process. The most accepted cardiac disease-associated dataset is the Cleveland dataset. Even though the entirety of the datasets has various kinds of highlights, a few of them are generally more significant for one or the other forecast or grouping. Henceforth, the general significance of highlights for the discovery of cardiac disease is examined. For instance, sex and age are examined independently because the effect of these 2
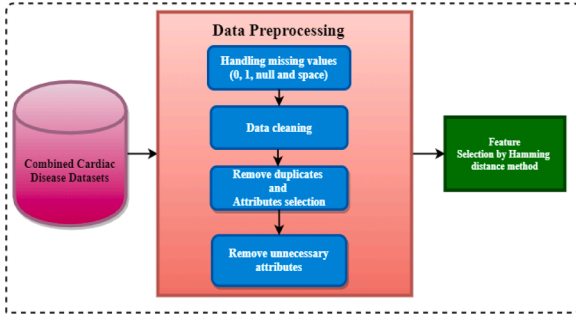
Fig. 5. Feature Selection by Hamming Distance Method.

elements is major for research objectives in most sickness states.

4.2. Data cleaning process by hamming distance method

The collected datasets can be combined and cleaned using the method of hamming distance (Djatna, Hardhienata & Masruriyah, 2018) by the calculation of two binary vectors with the sum of the difference between the two vectors. The hamming distance can be used for cleaning the dataset to incorporate the non-linearity property of the data. Hamming distance is also referred to as binary strings or bitstrings. A capacity on expressions of fixed length over a letter set portraying the number of changes to the images of a single word needed to diminish it to another. The data preprocessing task is shown in Fig. 6.

Let A be an alphabet of symbols and B be a subset of A^n, where the set of words of length is n over A.

Let $a = (a_1, ..., a_n)$ and $b = (b_1, ..., b_n)$ be words in B.

The Hamming distance $d(a, b)$ is defined as the number of places in which a and b can differ by, $i : a_i \neq b_i$, with $i = 1, ..., n$. The Hamming distance is satisfied by the following mathematical model,

$$d(a, b) \geq 0 \text{ and } d(a, b) = 0 \text{ if and only if } a = b \qquad (1)$$

$$d(a, b) = d(b, a); \qquad (2)$$

$$d(a, b) \leq d(a, w) + d(w, b) \qquad (3)$$

where w is the hamming weight and also it is represented as a metric in B. In the concept of error-correcting codes, the words from B are transferred down to a noisy channel with the assumption of changing the symbols. If the nominal hamming distance among words of B is δ, then this task is skilled in noticing up to $\delta - 1$ errors and ordered up to $[\delta - 12]$ errors. It is also defined for strings type having the same length. For example, the two strings s_1 and s_2 is evaluated in the hamming distance as $D(s_1, s_2)$ that can differ with the number of places that consists of distinct characters. Then the distance between two strings s_1 and s_2 is $\sum |s_{i1} - s_{i2}|$. The Hamming weight $w(A)$ of a word, A over an alphabet incorporating a notable symbol 0 is the distance $d(0, b)$ where zero is all zero words. For the linear type of code that is represented in a vector space over a finite field, then the hamming distance is determined by the weight:

$$d(A, B) = w(A - B) \qquad (4)$$

5. Classification algorithm

The mathematical reason for the neural network's prowess at classifying is the universal approximation theorem which states that a neural network can approximate any continuous real-valued function on a compact subset. Therefore, in our work, we have used the deep belief networks (DBN) model. The concept of ensemble classification is also used in our work to attain a set of classifiers but not just an individual one can be called an ensemble of classifiers (Latha & Jeeva, 2019), and it will combine their predictions for the different classification problems. After the completion of dataset pre-processing work with the identification of important features for the analysis of cardiac disease, the next step is to apply ensemble classification algorithms for finding better accuracy in the prediction process. Ensemble classification has the procedure of collecting the training dataset and transferring it to a classifier and then ensemble it with the test dataset. Then the combined results can be calculated by averaging and voting them, which is the basic method for predicting the final output with accuracy. The task done by the deep belief network classifier is explained in the following section.

5.1. Deep belief networks

Deep Belief Networks (Bengio, Lamblin, Popovici & Larochelle, 2007) were introduced in the year 2007 by Larochelle et al. as probabilistic generative models collected by stacked modules of Restricted Boltzmann Machines (RBMs) and provide an alternative to the discriminative nature of traditional neural nets. The primary attributes of DBN are quick induction and the capacity to encode higher-request network structures. Deep belief networks are models that utilized the two probabilities and unaided figuring out how to deliver yields. They are made out of double inert factors, and they contain both undirected layers and coordinated layers. In contrast to different models, each layer in deep belief networks learns the whole information. It can be used for identification, clustering, image processing, video sequels, and signal-capture data and also for training nonlinear autoencoders.

The mathematical model of deep belief networks is as follows:

A DBN with l hidden layers contains l weight matrices: $W^{(1)}, ..., W^{(l)}$

It also contains $l + 1$ bias vectors: $b^{(0)}, ..., b^{(l)}$ where $b^{(0)}$ provide the biases for the visible layer.

The probability distribution for DBN is given by,

$$P\left(h^{(l)}, h^{(l-1)}\right) \propto exp\left(b^{(l)^T} h^{(l)} + b^{(l-1)^T} h^{(l-1)} + h^{(l-1)^T} W^{(l)} h^{(l)}\right) \qquad (5)$$

$$P\left(h_i^{(k)} = 1 \middle| h^{(k+1)}\right) = \sigma\left(b_i^{(k)} + W_{:,i}^{(k+1)^T} h^{(k+1)}\right) \qquad (6)$$

where $\forall i, \forall k \in 1, ..., l - 2$

$$P\left(v_i = 1 \mid h^{(1)}\right) = \sigma\left(b_i^{(0)} + W_{:,i}^{(1)^T} h^{(1)}\right) \forall i. \qquad (7)$$

In the case of real-valued visible units, substitute

$$v \sim N\left(b^{(0)} + W^{(1)^T} h^{(l)} \beta^{-1}\right), \qquad (8)$$

with β diagonal for tractability $\sigma(x) = 1/(1 + exp(-x))$

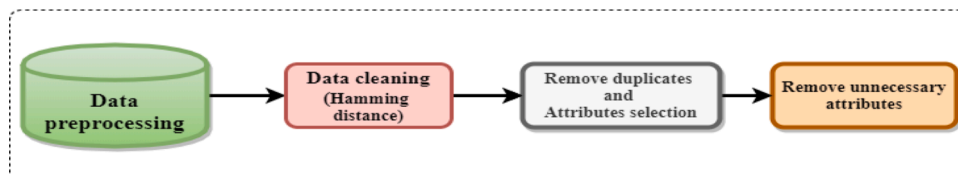The weights from the trained DBN can be used as the initialized

Fig. 6. Data pre-processing.

weights of a DNN

$$h^{(1)} = \sigma\big(b^{(1)} + v^T W^{(1)}\big), \tag{9}$$

$$h^{(l)} = \sigma\Big(b_i^{(l)} + h^{(l-1)T} W^{(l)}\Big), \ \forall l \in 2, ..., m \tag{10}$$

also, then, at that point, the entirety of the loads are tweaked by applying backpropagation or other discriminative models to enhance the efficiency of the entire network. The current trends and applications of DBN is shown in Table 2.

6. Bio-Inspired optimization algorithms

Optimization is the most noticed mathematical problem in all engineering disciplines. It is prominently used to find the best feasible solution. Nowadays, an active research topic is an optimization problem that is widely spread all over the world. It can be both deterministic and stochastic and has the models to solve optimization problems that require immense computational efforts. Bio-inspiration is the blooming of modern structures, resources, and methods. It is stimulated by results found in adaptation, procedures, and salvation, which has biologically progressed over millions of years. The bio-inspired field starts in the early 1980s and 2010s with many types of research on the natural phenomenon in limited support but in 2020 it's been a huge growth in the variety of novel algorithms that have been defined successfully. The notions of nature and biological activities have motivated the burgeoning of many enlightened algorithms for problem-solving. The critical thinking feature of an artificial evolvement is incorporated into the choice interaction, which comprises two essential ways. 1) an assessment of the phenotype that gives a quantitative score, otherwise called a fitness value, and 2) a multiplication operator that makes an enormous number of duplicates of genotypes compared to phenotypes with high fitness values. These algorithms are commonly classified as evolutionary computation and swarm intelligence algorithms. Evolutionary computation is an idiom used to embellish algorithms that were enthused by the essence of the fittest. Swarm intelligence is an idiom used to embellish algorithms for disseminated problems-solvers that were enthused by the concerted team intellect of swarm or collective habits of bug enclaves and other species inhabitants. This leads to adding a bio-inspired optimization algorithm in the proposed work for the prediction of cardiac disease to produce more efficacy.

During the last few decades, bio-inspired computation (Kar, 2016) has become one of the most researched disciplines of AI. Bio-inspired algorithms have been widely used to speed up the learning process of deep learning models, with a particular focus on deep belief networks. Similarly, bio-inspired heuristics (Del Ser et al., 2019) are frequently

used to prescribe near-optimal actions based on DL model predictions, completing what is now known as actionable data science. Future possibilities are expected to grow rapidly due to the enormous number of parameters featured by the Deep Learning family of models, which is recently prospecting the adoption of bio-inspired techniques as an alternative.

In the existing literature, most of the researchers have performed other traditional methods for the feature selection process such as PCA, Chi-Square test, Correlation coefficient, etc. In our proposed work, when compared with other methods, hamming distance method is used which can produce more enhancement for the feature selection process. The cuckoo search algorithm is used for the class inequity problem for the classification of data. In (Mohsin, Li & Abdalla, 2020), the authors have proposed an Adam-Cuckoo search-based Deep Belief Network (Adam-CS-based DBN) for the classification of data. First, the classification is performed by the DBN classifier and trained using an Adam-based cuckoo search algorithm and achieved better performance using the metrics such as accuracy, sensitivity, and specificity. In (Li et al., 2021), the authors have used an adaptive hybrid Simulated Annealing Cuckoo Search algorithm (SA-ACS) model, and then it was applied to the Deep Belief Network (DBN). They have also stated that the SA-ACS-DBN algorithm is utilized for the improvement of the training speed and convergence accuracy. In (Gampala et al., 2022), the authors have used the deep belief network method with an enhancement of hosted cuckoo optimization approach for getting optimum hyper tuning parameters in the diagnosis of COVID-19. The proposed model has achieved better performance when compared with other models such as Convolutional Neural Network with Social Mimic Optimization (CNN-SMO) and Support Vector Machine classifier using Bayesian Optimization algorithm (SVM-BOA).

Based on the above literature, for the classification of data, cuckoo search with deep belief networks has done better performance in their problem statements. So, we have utilized cuckoo search enabled deep belief networks in our work for the detection of cardiac disease in an accurate manner.

6.1. Cuckoo search algorithm

Cuckoo search (Yang & Deb, 2014) is an important population-based optimization algorithm enhanced by Xin-she Yang et.al., in the year 2009. The working principle of it is by smearing their ova in the shelters of different feeding birds. Cuckoo search is inspired by the help of breeding behavior and also applied for several optimization challenges. From the nature-inspired area, a Cuckoo search is also one of the techniques, which is used mainly to solve several optimization problems in different domains of engineering. It can able to maintain the balance between local and global random walks in solving global optimization using the switching parameter. Due to less research on the switching parameter, it was fixed at 25%, so the impact of dynamic in nature was not able assessed by the method. It is broadly exploited in the applications of speech reorganization, job scheduling, neural computing, and global optimization.

The mathematical models of the cuckoo search algorithm are as follows: The CS method is explained with the following three idealized rules:

1. Each cuckoo lays each egg in turn and dumps it in a haphazardly picked home.
2. The best homes with top-caliber of eggs will extend to the following ages.
3. The quantity of accessible host homes is fixed, and a host can find an outsider egg with a likelihood $P_a \in [0, 1]$. For this situation, the host bird can either discard the egg or leave the home to assemble a new home (nest) n in another area.

Based on these three principles, the essential strides of the CS can be

Table 2

Proposed deep learning model used in current trends and applications.

Deep learning models	Methods used	Applications	Advantages	Accuracy
DBN	Knowledge-based deep belief network (KBDBN) and RBM (Yu & Liu, 2020)	Machining Roughness Prediction and Knowledge Discovery	It keeps the physics knowledge of the process and provides good predictions	70-80%
	Deep Belief Networks, SVM, KNN, and HMM (Movahedi, Coyle & Sejdic, 2018)	Electroencephalography signals	It provides better data modeling and task classification for EEG data	91.31%

summed up as follows:

When generating new solutions $x^{(t+1)}$ for, say, a cuckoo i, a Lévy flight is defined by Eq. (11) as follows,

$$x_i^{(t+1)} = x_i^{(t)} + \alpha \oplus Levy(\lambda) \tag{11}$$

where $x_i^{(t)}$ is the current location of the cuckoo, α is a step size and positive constant tuned according to the dimensions of the search space, $\oplus$ is the entry wise multiplication and λ is the levy exponent.

The above equation is essentially the stochastic equation for a random walk. The random step length is drawn from a Lévy distribution defined in Eq. (12) as follows,

$$Levy(\lambda) \sim u = t^{-\lambda}, \ (1 < \lambda \le 3) \tag{12}$$

where λ is the levy exponent that defines the decay of the probability

density function (PDF) with t. In most cases, $\alpha = 1$ and $\lambda = 1.5$.

Lévy-flight has the distinct property of increasing population diversity sequentially, allowing the algorithm to effectively jump out of the local optimum. The following Eq. (13) is calculated as Levy random numbers,

$$Levy\ (\lambda) \sim \frac{\phi \times u}{|\nu|^{1/\lambda}} \tag{13}$$

Where u and ν are both standard normal distributions, then ϕ is defined in Eq. (14) as follows,

$$\phi = \left[\frac{\gamma\ (1+\lambda)\ \times\ sin\ (\pi\ \times\ \lambda/2)}{\gamma\ (((1+\lambda)/2)\ \times\ \lambda\ \times\ 2^{(\lambda-1)/2})} \right]^{1/\lambda} \tag{14}$$

where γ is a standard gamma function and ϕ is a random angle used for
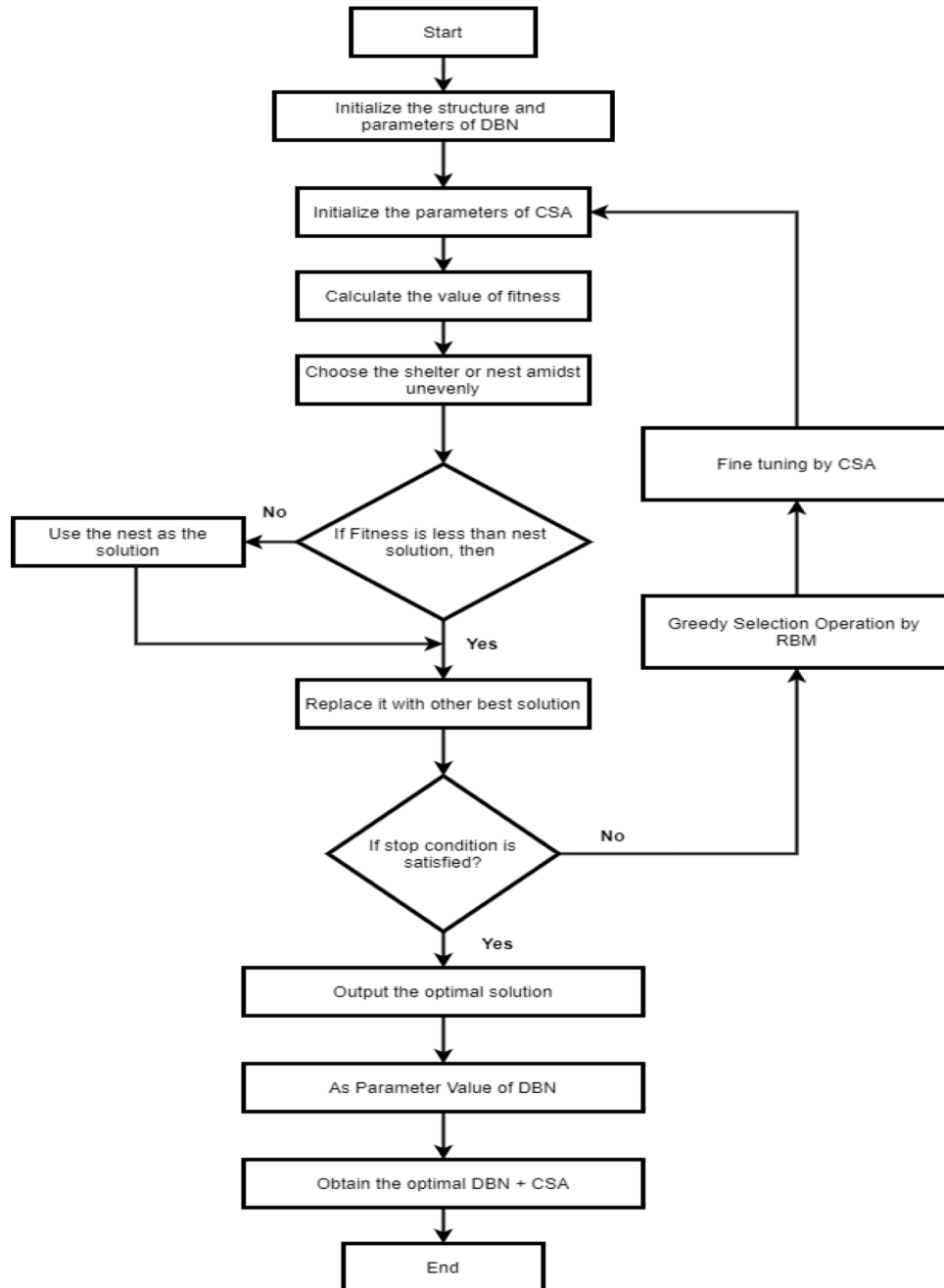

Fig. 7. Flow chart representation of the proposed model (DBN+CSA).

convergence.

To summarise, the proposed algorithm's global search ability is improved by utilizing random walk with Levy-flight and its local minimum is avoided. It is also seen to produce more successful results, especially for unimodal and multimodal benchmark functions.

In nature, numerous creatures and creepy crawlies look for food in an irregular or semi-arbitrary way following the properties of Lévy flight, in which the progression lengths are appropriated by a substantial followed PDF. As Lévy flights have boundless means and change, some new arrangements will be situated around the best arrangements got a time t, and others will be set away from the current most popular arrangements. This component permits CS to investigate the inquiry space more productively than different calculations that utilize standard gaussian interaction, trying not to be caught around nearby optima. The link between deep belief networks and the cuckoo search algorithm is shown in Fig. 7.

In a similar fashion to GA and PSO, CS is a population-based algorithm. The first advantage is that randomization is more efficient because the step length is heavy-tailed, allowing for any huge step. The second advantage is that it has fewer tuning factors than GA and PSO, making it potentially more flexible in application to a broader range of optimization issues. Furthermore, each nest can represent a set of solutions, allowing CS to be used as a meta-population method. The original CSA, by incorporating problem-specific knowledge can be split into three kinds. One is the constructive heuristics and the second one is the local search methods. The third one is the hybridization with structures borrowed from other algorithms. In this work, the third kind is implemented by adapting DBN to classify and CSA is used for the final tuning of the parameters with the proposed aim is to detecting whether the cardiac disease is present or absent. The cuckoo search algorithm is shown in Algorithm 1 with its steps.

7. Results and discussion

This segment exhibits the results and its discussion of the various cardiac disease datasets that have been experimented with by the proposed model. The cardiac datasets used in the proposed model are the South African heart disease dataset, UCI Cleveland, Hungary, Switzerland, and the VA long beach, Statlog, Z-Alizadeh Sani, Framingham heart, and heart failure dataset. First, the datasets are loaded with various features into the weka tool for data pre-processing. Feature selection methodology is used for identifying and classifying the right features for the analysis process. Some of the features commonly used in the above cardiac datasets are age, sex, maximum heart rate, thal, type A, family history, tobacco, obesity, resting blood sugar, and serum cholesterol. Some of them are not used for selection because it reduces the performance by increasing the time complexity that indicates the identification of errors. Hamming distance is used for the feature selection process, which minimizes the data by the cleaning process and is

also used on the categorical data, and selection of the categorical values is forwarded so that the cluster is most likely to take on for the data pre-processing work. The classification model used in our proposed work is the deep belief network which is the type of deep learning model. The training samples are classified with traditional classifiers whereas testing samples are done by optimized classifiers such as deep belief networks embedded with the cuckoo search algorithm. The results show that the deep belief network outperforms all the other models. Tables 4 and 5 shows the performance evaluation metrics results from the deep belief networks model and Tables 6–10 shows the comparison of proposed model accuracy with existing models. Figs. 8–14 represents the graphical representation of performance with the deep belief networks model.

7.1. Confusion matrix

The confusion matrix is a kind of table representation that is used for the description of the performance of a classification type of model which contains either true or false. It is simple to understand, but complex to analyze with a range of features. For the cardiac disease prediction process, there are two possibilities of prediction either yes or no that represent whether the disease is present or not. The basic terms used for the analysis of the confusion matrix are true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Table 3 shows the performance metrics used in the proposed work using the confusion matrix.

True positive (TP): It is used for identifying the predicted result correctly.
False-positive (FP): It is used for identifying the predicted result incorrectly that represents in Type I error.
False-negative (FN): It is used for rejecting the result incorrectly that represents in Type II error.
True negative (TN): It is opposite to FN that correctly rejects the result.

There are some different performance metrics available that are calculated by using the confusion matrix. They are Accuracy, Precision, Recall, and F1 score.

Accuracy is decisive for the proper classification of several instances present in the dataset. The mathematical formula for calculating accuracy was represented by,

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{15}$$

Precision is used for predicting the positive values from the given dataset and it is widely used in information retrieval. The mathematical formula for calculating precision was represented by,

Algorithm 1
Cuckoo Search Algorithm.

Steps:
A1. Start the process
A2. Initialize the random populace
A3. Obtain a cuckoo unevenly by heavy-tailed
A4. Calculate the value of fitness
A5. Choose the shelter or nest amidst unevenly
A6. If Fitness is less than nest solution, then
A7. Replace it with other best solution else
A8. Use the nest as the solution
A9. For both the steps, 7 and 8, then perform abandon a fraction worst nest and create a new nest by heavy-tailed
A10. Store the current best solution
A11. If no. of the nest is less than or equal to maximum iterations, then find the best nest, else go to step 3
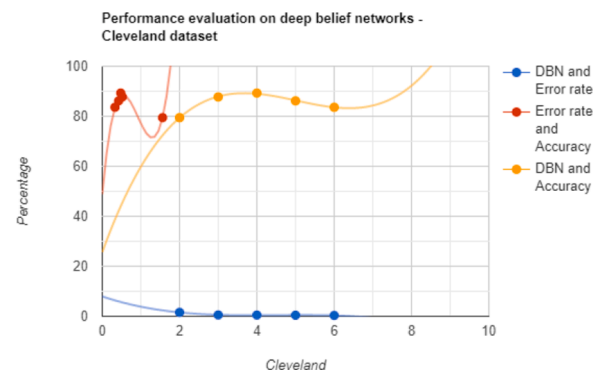*A12. End the process*

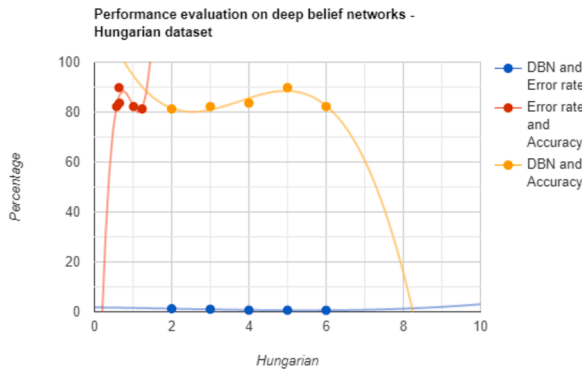**Fig. 8.** Graphical representation of performance evaluation result - Cleveland dataset.

Fig. 9. Graphical representation of performance evaluation result - Hungarian dataset.
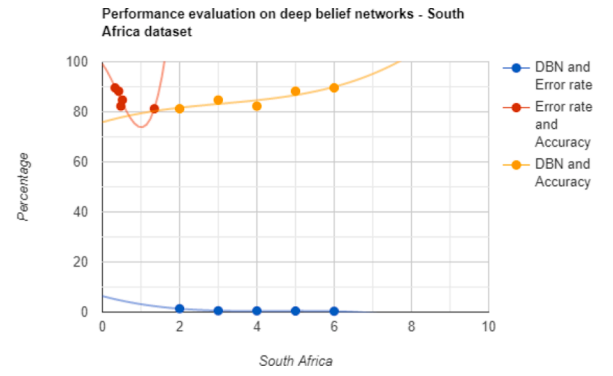

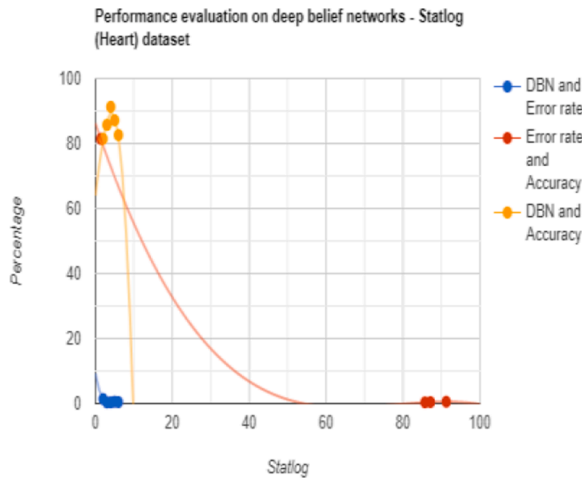**Fig. 12.** Graphical representation of performance evaluation result - South Africa dataset.

Fig. 10. Graphical representation of performance evaluation result - Statlog dataset.
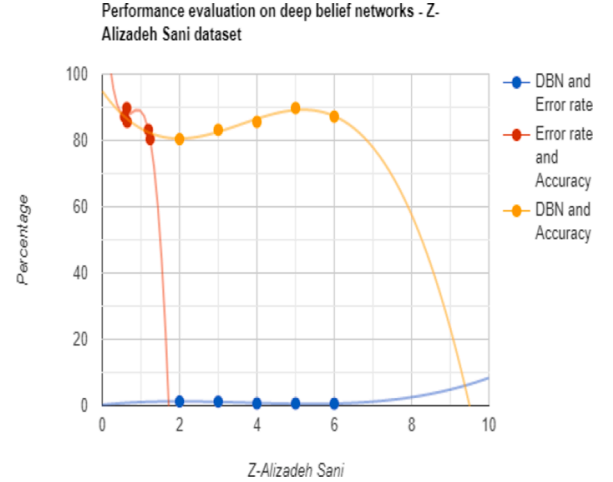
Fig. 13. Graphical representation of performance evaluation result - Z-Alizadeh Sani dataset.
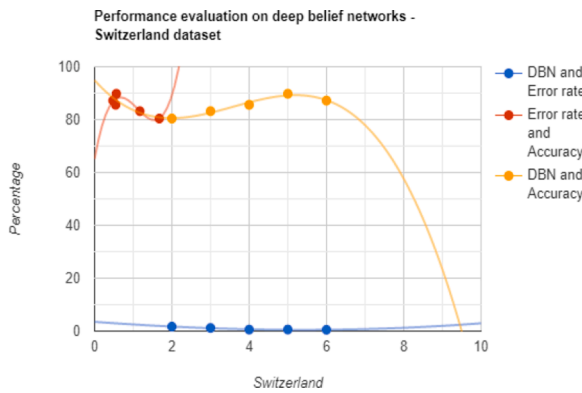
Fig. 11. Graphical representation of performance evaluation result - Switzerland dataset.

$$Precision = \frac{TP}{TP + FP} \tag{16}$$

Recall (Sensitivity) is also widely used in information retrieval with the utilization of proportional data for the prediction of total positive in percentage. The mathematical formula for calculating recall was represented by,

$$Recall\ (Sensitivity) = \frac{TP}{TP + FN} \tag{17}$$
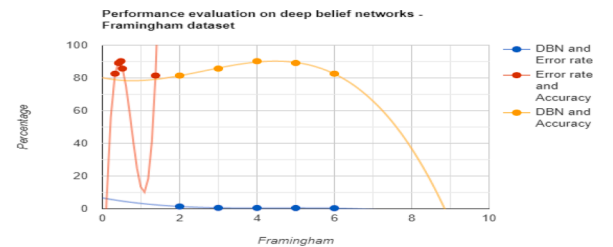
Fig. 14. Graphical representation of performance evaluation result - Framingham dataset.

Table 3
Performance metrics using the confusion matrix.

Class	Yes	No
Yes	True positive (TP)	False-negative (FN)
No	False-positive (FP)	True negative (TN)

Specificity is used for predicting fully true negatives with the proportion of actual negatives that can be formulated by,

$$Specificity = \frac{TN}{TN + FP} \tag{18}$$

F1 score is represented with the harmonic mean of the combination of precision and recall. It handles both false positives and false negatives

that result in a good performance on an imbalanced dataset. The mathematical formula for calculating the F1 score was represented by,

$$F1 \; score = \frac{2 * (Precision * Recall)}{(Precision \; + \; Recall)} \qquad (19)$$

7.2. Discussion of the datasets

Cardiac datasets used in our work consist of several records such as the South African heart dataset with 462 records, UCI Cleveland has 297 records, UCI Statlog has 270 records, UCI Z-Alizadeh Sani dataset has 303 records, UCI Framingham heart dataset has 4240 records, and heart failure dataset has 1000 records. All these records have been utilized with K-fold cross-validation for the implementation part. From the result, we have got the performance value of error rate with different depth values of deep belief networks. Furthermore, the effect of error detection was evaluated by including an error classification stage.

Analysis of Table 4 shows that the increase in network depth value reduces the error rate. The accuracy of the different cardiac datasets is shown with depth value and error rate. Based on the depth value change the error rate is decreased and the accuracy is widely increased. For the Cleveland dataset, the maximum accuracy of 89.2% is attained at the depth value of 4. The same accuracy is attained in the Hungarian dataset, Switzerland heart dataset and z-Alizadeh Sani dataset of 89.7% at the depth value of 5 with the same error rate of 0.63 for Hungarian and z-Alizadeh Sani and it has a different error rate of 0.56 for Switzerland dataset. The Statlog (heart) dataset has attained the maximum accuracy of 91.2% at the depth value of 4. For the South Africa dataset, it has attained an accuracy of 89.5% at the depth value of 6. Finally, for the Framingham dataset, it has attained an accuracy of 90.2% at the depth value of 4. The result shows that the proposed model has attained an average accuracy of 90% with the improved network depth that is detected by deep belief networks proving that the performance is increased for the prediction of cardiac disease. However, the proposed model is less accurate when compared with (Altan, Allahverdi & Kutlu, 2018), where the authors have utilized the wavelet packet decomposition, higher order statistics, morphology, and discrete Fourier transform techniques for the feature extraction with an accuracy of 94.15%. This work is also compared with (Altan, Kutlu & Allahverdi, 2016), where the authors have utilized the ECG waveform and SODP as the feature extraction methods for the detection of arrhythmia with an accuracy of 96.10%. From the above two reviews, our proposed model has differed in feature selection and optimization. We have used hamming distance method for the features selection, with the DBN classifier, and finally, with the utilization of the cuckoo search, the bio-inspired optimization algorithm has achieved an accuracy of 91.26% which is accurate in the prediction of cardiac disease. The other performance metrics such as AUC and ROC are shown in Table 5 with the runtime of the proposed model.

The time complexity of our proposed model slightly increased when using smaller datasets for the detection of cardiac disease. Since the instances are fewer in the datasets, the DBN model structure is modified. Then the new component of the structure must be retrained, resulting in a significant increase in training time due to the algorithm's complexity. The time it takes to train increases as the complexity and number of required iterations grow. However, due to the smaller size of the solution space, the algorithm's time complexity increases slightly. The runtime is slightly increased based on the increase in the depth values of the DBN. In the future, if we used larger datasets, the time complexity may vary based on the size of the data and it also results in either decreasing mode or an increasing mode based on the model we use. The overfitting problem is avoided since we have used hamming distance feature selection method that is cleaned the input data and performed

Table 4
Performance evaluation on deep belief networks.

Dataset	DBN	Error rate	Accuracy %
Cleveland	2	1.56	79.4
	3	0.51	87.7
	4	**0.47**	**89.2**
	5	0.42	86.1
	6	0.32	83.5
Hungarian	2	1.23	81.2
	3	1.01	82.2
	4	0.64	83.6
	5	**0.63**	**89.7**
	6	0.57	82.2
Statlog (Heart)	2	1.38	81.4
	3	0.31	85.7
	4	**0.34**	**91.2**
	5	0.54	87.1
	6	0.43	82.5
Switzerland	2	1.67	80.4
	3	1.17	83.2
	4	0.54	85.6
	5	**0.56**	**89.7**
	6	0.47	87.2
South Africa	2	1.34	81.2
	3	0.51	84.7
	4	0.47	82.2
	5	0.42	88.2
	6	**0.32**	**89.5**
Z-Alizadeh Sani	2	1.25	80.4
	3	1.19	83.2
	4	0.64	85.6
	5	**0.63**	**89.7**
	6	0.57	87.2
Framingham	2	1.37	81.4
	3	0.51	85.7
	4	**0.47**	**90.2**
	5	0.42	89.1
	6	0.32	82.5

Table 5
AUC, ROC, and Runtime performance metrics of the proposed model with an Error rate.

Dataset	Error rate	AUC	ROC	Runtime (s)
Cleveland	1.56	0.85	0.81	16.21
	0.51	0.86	0.83	17.26
	0.47	**0.88**	**0.87**	**18.31**
	0.42	0.82	0.86	21.36
	0.32	0.83	0.88	22.41
Hungarian	1.23	0.83	0.79	16.30
	1.01	0.84	0.81	16.35
	0.64	0.82	0.83	18.42
	0.63	**0.86**	**0.85**	**20.28**
	0.57	0.81	0.87	22.39
Statlog (Heart)	1.38	0.78	0.80	17.30
	0.31	0.83	0.81	19.90
	0.34	**0.87**	**0.86**	**22.56**
	0.54	0.80	0.84	24.80
	0.43	0.86	0.86	28.54
Switzerland	1.67	0.75	0.80	18.28
	1.17	0.79	0.84	19.34
	0.54	0.83	0.82	21.57
	0.56	**0.85**	**0.86**	**24.85**
	0.47	0.88	0.81	28.97
South Africa	1.34	0.76	0.79	17.20
	0.51	0.81	0.82	19.44
	0.47	0.84	0.86	22.17
	0.42	0.86	0.88	25.45
	0.32	**0.88**	**0.87**	**26.15**
Z-Alizadeh Sani	1.25	0.77	0.80	16.10
	1.19	0.81	0.82	20.16
	0.64	0.85	0.84	22.14
	0.63	**0.87**	**0.88**	**24.13**
	0.57	0.86	0.89	27.44
Framingham	1.37	0.82	0.80	16.21
	0.51	0.85	0.84	18.16
	0.47	**0.89**	**0.87**	**20.43**
	0.42	0.86	0.85	22.97
	0.32	0.87	0.88	24.13

well when compared with other correlation coefficients. By the reduction of unnecessary features, we have significantly decreased the training time. The proposed model generates the trained data with different patterns and accurately classified the data samples. So, the performance of the training is better when compared with other deep learning and machine learning models.

From Tables 6–10, the results shown are compared with existing base classifiers on cardiac datasets with the proposed model. The classifiers used for the comparison are decision trees, Naive Bayes, support vector machines, neural networks, logistic regression, and artificial neural networks on the cardiac datasets. A decision tree is a well-known predictive-based algorithm. It shows the output in the form of a tree structure. It uses the concept of splitting a node into two or more sub-nodes. Based on the target variable, the node gets increased and it was constructed by recursively evaluating different features. Through its basic idea of splitting a population of data into smaller segments and it determines the best attribute from the set of attributes. It produces an accuracy of around 73% to 85%. Naïve Bayes algorithm is worked with the concept of Bayes theorem. It is one of the best inductive learning algorithms for machine learning. It utilizes the concepts of structuring a node as a parent node for classification type and learns the parameters for processing and then displays the output in its network. It produces an accuracy of around 71% to 83%. Support vector machine is based on the theory of statistics. It works with the concept of building a set of hyperplanes and then optimizing it for the detection of outliers which classifies the input data points. Based on the observations with hyperplane which separates the given input data into a class, then it predicts the desired result with maximum margin. It produces an accuracy of around 73% to 86%. A neural network is performed based on the neurons in the human brain. It is also called a multi-layer perceptron. It contains three layers, input, hidden, and output layers. Each input with a bias and weights is combined to produce a single output. A combination function or activation function, such as a linear activation function is used with weight and the bias to give the predicted output. It produces an accuracy of around 68% to 86%. Logistic regression is widely used for binary classification problems and also for a prediction of a target variable. It takes input values that are combined linearly using weights or coefficient values. It works with the concept of probability posterior type with a linear function as the input data and estimated the final output by the maximum likelihood method. It produces an accuracy of 82% to 85%. Artificial neural network captures the concepts from the biological neural networks. Through the connection link method, the neurons are connected with another neuron. Each connected neuron has a weight that has information about the input signal. With the help of the activation function, the weighted sum of inputs is passed to the next level. Finally, a binary activation function is used for predicting the output which is resulted in the form of either 0 or 1. It produces an accuracy of around 84% to 87% respectively. Hence, based on a comparison with the above methods, it is articulated that our proposed model is enhanced in heart disease prediction accuracy. In our work, the used symbols with their description are shown in Table 11.

Advantages of our proposed model:

- In this manuscript, the deep belief network method is enhanced by the cuckoo search optimization approach for getting optimum hyper

Table 6

Comparison of performance evaluation accuracy of different techniques on the Statlog dataset.

Classification algorithms	Accuracy (%)
SVM (Chan-Yun, Jui-Jen & Feng-Li, 2013)	85.20
DT (Mantas Carlos & Joaquin, 2014)	80.41
MLP with Genetic Algorithm (Ahmad, Isa, Hussain & Osman, 2013)	86.29
ANN (Iftikhar et al., 2017)	86.04
BPNN (Olaniyi, Oyedotun & Helwan, September 2015)	85.00
Proposed- DBN+CSA	**91.26**

Table 7

Comparison of performance evaluation accuracy of different techniques on the Cleveland dataset.

Classification algorithms	Accuracy (%)
Genetic fuzzy logic system (Nguyen, Khosravi, Creighton & Nahavandi, 2015)	78.7
Type 2 fuzzy logic with firefly algorithm (Long, Meesad & Unger, 2015)	86.0
XGBoost and LR (Nalluri et al., 2020)	85.68
ANFSI and ANN (Abushariah, Alqudah & Adwan, 2014)	87.04
Ensemble learning (Gradient boosted) (Sapra, Sandhu & Goyal, 2021)	84
Proposed- DBN+CSA	**89.26**

Table 8

Comparison of performance evaluation accuracy of different techniques on the South Africa dataset.

Classification algorithms	Accuracy (%)
DT (Babič et al., 2017)	73.87
NB (Babič et al., 2017)	71.17
SVM (Babič et al., 2017)	73.40
NN (Babič et al., 2017)	68.48
Proposed- DBN+CSA	**89.58**

Table 9

Comparison of performance evaluation accuracy of different techniques on the Z-Alizadeh Sani dataset.

Classification algorithms	Accuracy (%)
DT (Babič et al., 2017)	85.38
NB (Babič et al., 2017)	83.33
SVM (Babič et al., 2017)	86.67
NN (Babič et al., 2017)	86.32
EmNNs with PSO (Shahid & Singh, 2020)	88.34
Proposed- DBN+CSA	**89.78**

Table 10

Comparison of performance evaluation accuracy of different techniques on the Framingham dataset.

Classification algorithms	Accuracy (%)
LR (Renganathan, 2019)	82.9
ANN (Renganathan, 2019)	84.4
GBC (Aggrawal & Pal, 2021)	87.61
RF (Latifah, Slamet & Sugiyanto, 2020)	84.4
Proposed- DBN+CSA	**90.26**

tuning parameters in the detection of cardiac disease from the five different benchmark datasets.

- An important advantage of CSA is its simplicity when compared with other metaheuristics algorithms such as PSO, and GA, since it has only a single parameter apart from population size n and it is very easy to implement.
- The local optima problem, inference problem, and learning problem are solved by the proposed methodology.

Disadvantages of our proposed model:

- The dependency on a specific feature selection technique is compared with other methods to produce highly accurate results.
- An adverse effect may happen due to a lot of missing values in the dataset.
- The larger dataset can be incorporated for the improvement of performance using the proposed model.

Table 11
Summary of Symbols and their description used in the article.

Symbol	Description
P_a	Probability
i	Initial cuckoo
$x_i^{(t+1)}$	New solution
$x_i^{(t)}$	Current solution
α	Step size
$\oplus$	Entry wise multiplication
λ	Levy exponent
u	Standard normal distribution
ν	Standard normal distribution
γ	Standard Gamma distribution
ϕ	Random angles in radians
t	Number of iterations
n	Nest
v	Vector of the visible layers
b	Bias vector of the hidden units
W	Weight matrices
h	The state vector of the hidden layers
σ	Sigmoid function
$P(h^{(l)})$	The probability distribution for hidden layers
$P(v)$	The probability distribution for visible layers
$P(v_i = 1\|h^{(1)})$	Joint probability distribution
i	Current iteration
l	Contains hidden layers and weight matrices from 0 *to l*
$\forall i$	For all i^{th} iteration
$\forall l$	For all l^{th} layers

8. Conclusion

In this article, we presented a deep belief network model with a feature selection method for analyzing and predicting cardiac disease with good accuracy. The data preprocessing process and the results are well tabulated with different values of datasets. Different cardiac datasets were taken and analyzed with hamming distance feature selection method for the data cleaning process. With the selection of important features, it is then forwarded to the classification process with deep learning models such as deep belief networks embedded with cuckoo search bio-inspired algorithm with different levels of depth and produced an accurate prediction of cardiac disease. It was additionally seen that the proposed prediction model shows better results over other machine learning models.

In future enhancement, the performance can be improved by combining large datasets with deep learning classification models embedded in bio-inspired algorithms on the ECG data of cardiac disease.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Compliance with Ethical Standards: This article does not contain any studies with human participants performed by any of the authors.

References

Abdar, M. (2017). "Using Decision Trees in Data Mining for Predicting Factors Influencing of Heart Disease," no. (June).
Abushariah, M. A., Alqudah, A. A., Adwan, O. Y., et al. (2014). Automatic heart disease diagnosis system based on artificial neural network (ANN) and adaptive neuro-fuzzy inference systems (ANFIS) approaches. *Journal of Software Engineering and Applications, 7*, 1055.
Aggrawal, Ritu, & Pal, Saurabh (2021). Elimination and Backward Selection of Features (P-Value Technique) In Prediction of Heart Disease by Using Machine Learning Algorithms. *Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12*, 2650–2665.
Ahmad, Fadzil, Isa, Nor Ashidi Mat, Hussain, Zakaria, & Osman, Muhammad Khusairi (2013). Intelligent medical disease diagnosis using improved hybrid genetic algorithm-multilayer perceptron network. *Journal of Medical Systems, 37*, 9934.
Alizadehsani, R., et al. (2013). A data mining approach for diagnosis of coronary artery disease. *Comput. Methods Programs Biomed., 111*(1). https://doi.org/10.1016/j.cmpb.2013.03.004
Alizadehsani, R., Hossein, M., & Javad, M. (2016). Knowledge-Based Systems Coronary artery disease detection using computational intelligence methods. *Knowledge-Based Syst, 109*, 187–197. https://doi.org/10.1016/j.knosys.2016.07.004
Al-Makhadmeh, Z., & Tolba, A. (2019). Utilizing IoT wearable medical device for heart disease prediction using higher order Boltzmann model: A classification approach. *Meas. J. Int. Meas. Confed., 147*. https://doi.org/10.1016/j.measurement.2019.07.043
Almustafa, K. M. (2020). Prediction of heart disease and classifiers' sensitivity analysis. *BMC Bioinformatics, 21*(1). https://doi.org/10.1186/s12859-020-03626-y
Altan, G., Allahverdi, N., & Kutlu, Y. (2018). A Multistage Deep Learning Algorithm for Detecting Arrhythmia. In *2018 1st International Conference on Computer Applications & Information Security (ICCAIS)* (pp. 1–5). https://doi.org/10.1109/CAIS.2018.8441942
Altan, Gokhan, Kutlu, Yakup, & Allahverdi, Novruz (2016). A multistage deep belief networks application on arrhythmia classification. *International Journal of Intelligent Systems and Applications in Engineering*, 222–228.
Ann, N. (2019). "A New Machine Learning Technique for an Accurate Diagnosis of Coronary Artery Disease Computer Methods and Programs in Biomedicine A new machine learning technique for an accurate diagnosis of coronary artery disease," no. (August), , doi: 10.1016/j.cmpb.2019.104992.
Arabasadi, Z., Alizadehsani, R., Roshanzamir, M., & Moosaei, H. (2017). Computer Methods and Programs in Biomedicine Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm. *Comput. Methods Programs Biomed., 141*, 19–26. https://doi.org/10.1016/j.cmpb.2017.01.004
Babič, František, et al. (2017). Predictive and descriptive analysis for heart disease diagnosis. In *2017 federated conference on computer science and information systems (fedcsis)*. IEEE.
Bengio, Y., Lamblin, P., Popovici, D., & Larochelle, H. (2007). Greedy layer-wise training of deep networks. In *Advances in Neural Information Processing Systems*. https://doi.org/10.7551/mitpress/7503.003.0024
Beunza, J. J., et al. (2019). Comparison of machine learning algorithms for clinical event prediction (risk of coronary heart disease. *Journal of Biomedical Informatics, 97*. https://doi.org/10.1016/j.jbi.2019.103257
Chan-Yu, Yang, Jui-Jen, Chou, & Feng-Li, Lian (2013). Robust classifier learning with fuzzy class labels for large-margin support vector machines. *Neurocomputin, 99*, 1–14.
D'Alto, M., et al. (2019). Does gender affect the prognosis and risk of complications in patients with congenital heart disease in the modern era? *International Journal of Cardiology, 290*. https://doi.org/10.1016/j.ijcard.2019.05.010
Del Ser, Javier, et al. (2019). Bio-inspired computation: Where we stand and what's next. *Swarm and Evolutionary Computation, 4*, 220–250.
Djatna, T., Hardhienata, M. K. D., & Masruriyah, A. F. N. (2018). An intuitionistic fuzzy diagnosis analytics for stroke disease. *J. Big Data, 5*(1). https://doi.org/10.1186/s40537-018-0142-7
Dogan, M. V., Grumbach, I. M., Michaelson, J. J., & Philibert, R. A. (2018). Integrated genetic and epigenetic prediction of coronary heart disease in the Framingham Heart Study. *PloS one, 13*(1). https://doi.org/10.1371/journal.pone.0190549
Dutta, A., Batabyal, T., Basu, M., & Acton, S. T. (2020). An efficient convolutional neural network for coronary heart disease prediction. *Expert Syst. Appl., 159*, Article 113408. https://doi.org/10.1016/j.eswa.2020.113408
Dwivedi, A. K. (2018). Performance evaluation of different machine learning techniques for prediction of heart disease. *Neural Comput. Appl., 29*(10). https://doi.org/10.1007/s00521-016-2604-1
Gampala, Veerraju, et al. (2022). Diagnosis of COVID-19 patients by adapting hyper parametertuned deep belief network using hosted cuckoo optimization algorithm. *Electromagnetic Biology and Medicin*, 1–15.
Gárate-Escamila, A. K., Hajjam El Hassani, A., & Andrès, E. (2020). Classification models for heart disease prediction using feature selection and PCA. *Informatics Med. Unlocked, 19*. https://doi.org/10.1016/j.imu.2020.100330
Gong, J., Bai, X., Li, D. A., Zhao, J., & Li, X. (2020). Prognosis Analysis of Heart Failure Based on Recurrent Attention Model. *IRBM, 41*(2). https://doi.org/10.1016/j.irbm.2019.08.002
Haq, A. U., Li, J. P., Memon, M. H., Nazir, S., Sun, R., & Garciá-Magarinõ, I. (2018). A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms. *Mob. Inf. Syst., 2018*. https://doi.org/10.1155/2018/3860146
Henriques, J., et al. (2015). Prediction of Heart Failure Decompensation Events by Trend Analysis of Telemonitoring Data. *IEEE J. Biomed. Heal. Informatics, 19*(5). https://doi.org/10.1109/JBHI.2014.2358715
Hossain, M. E., Uddin, S., & Khan, A. (2021). Network analytics and machine learning for predictive risk modelling of cardiovascular disease in patients with type 2 diabetes. *Expert Syst. Appl., 164*. https://doi.org/10.1016/j.eswa.2020.113918
Iftikhar, S., Fatima, K., Rehman, A., Almazyad, A. S., & Saba, T. (2017). An evolution based hybrid approach for heart disease classification and associated risk factor identification. *Biomedical Research, 28*(8), 3451–3455.
Jin, B., Che, C., Liu, Z., Zhang, S., Yin, X., & Wei, X. (2018). Predicting the Risk of Heart Failure with EHR Sequential Data Modeling. *IEEE access : practical innovations, open solutions, 6*. https://doi.org/10.1109/ACCESS.2017.2789324
Kar, Arpan Kumar (2016). Bio inspired computing–a review of algorithms and scope of applications. *Expert Systems with Applications, 5*, 20–32.
Kivimäki, M., & Steptoe, A. (2017). Effects of stress on the development and progression of cardiovascular disease. *Nat. Publ. Gr., 15*(4), 215–229. https://doi.org/10.1038/nrcardio.2017.189

Kolukisa, B. et al., (2019). "Diagnosis of Coronary Heart Disease via Classification Algorithms and a New Feature Selection Methodology," vol. 1, no. (1), pp. 1–8.

Latha, C. B. C., & Jeeva, S. C. (2019). Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques. *Informatics Med. Unlocked, 16.* https://doi.org/10.1016/j.imu.2019.100203

Latifah, Firda Anindita, Slamet, Isnandar, & Sugiyanto. (2020). Comparison of heart disease classification with logistic regression algorithm and random forest algorithm. In *AIP conference proceedings, 2296.* AIP Publishing LLC.

Li, G., Hu, R., & Gu, X. (2020). A close-up on COVID-19 and cardiovascular diseases. *Nutrition, Metabolism and Cardiovascular Diseases, 30*(7). https://doi.org/10.1016/j.numecd.2020.04.001

Li, Honghui, et al. (2021). Fault diagnosis of railway freight car wheelset based on deep belief network and cuckoo search algorithm. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transi,* Article 09544097211029155.

Long, N. C., Meesad, P., & Unger, H. (2015). A highly accurate firefly based algorithm for heart disease prediction. *Expert Syst. Appl, 42,* 8221–8231.

Mantas Carlos, J., & Joaquin, Abellan (2014). Credal-C4. 5: Decision tree based on imprecise probabilities to classify noisy data. *Expert Syst App, 41*(10), 4625–4637.

Martinez, A. D., et al. (2021). Lights and shadows in Evolutionary Deep Learning: Taxonomy, critical methodological analysis, cases of study, learned lessons, recommendations and challenges. *Inf. Fusion, 67.* https://doi.org/10.1016/j.inffus.2020.10.014

Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. *IEEE access : practical innovations, open solutions, 7.* https://doi.org/10.1109/ACCESS.2019.2923707

Mohsin, Mohammed, Li, Hong, & Abdalla, Hemn Barzan (2020). Optimization driven adam-cuckoo search-based deep belief network classifier for data classification. *IEEE access : practical innovations, open solutions,* 105542–105560.

Movahedi, F., Coyle, J. L., & Sejdic, E. (2018). Deep belief networks for electroencephalography: A review of recent contributions and future outlooks. *IEEE J. Biomed. Heal. Informatics, 22*(3). https://doi.org/10.1109/JBHI.2017.2727218

Mufudza, C., & Erol, H. (2016). Poisson Mixture Regression Models for Heart Disease Prediction. *Comput. Math. Methods Med., 2016.* https://doi.org/10.1155/2016/4083089

Nalluri, S., Saraswathi, R. V., Ramasubbareddy, S., Govinda, K., & Swetha, E. (2020). Chronic heart disease prediction using datamining techniques. In *Data Engineering and Communication Technology* (pp. 903–912).

Nasarian, E., et al. (2020). Association between work-related features and coronary artery disease : A heterogeneous hybrid feature selection integrated with balancing approach. *Pattern Recognit. Lett., 133,* 33–40. https://doi.org/10.1016/j.patrec.2020.02.010

Nguyen, T., Khosravi, A., Creighton, D., & Nahavandi, S. (2015). Classification of healthcare data using genetic fuzzy logic system and wavelets. *Expert Syst. Appl, 42,* 2184–2197.

Olaniyi, E. O., Oyedotun, O. K., Helwan, A., et al. (2015, September). Neural network diagnosis of heart disease. In *Proceedings of the International Conference on Advances in Biomedical Engineering* (pp. 21–24).

Orphanou, K., Stassopoulou, A., & Keravnou, E. (2016). DBN-extended: A dynamic Bayesian network model extended with temporal abstractions for coronary heart disease prognosis. *IEEE J. Biomed. Heal. Informatics, 20*(3). https://doi.org/10.1109/JBHI.2015.2420534

Renganathan, V. (2019). Overview of artificial neural network models in the biomedical domain. *Bratislavske lekarske listy, 120,* 536–540.

Saliba, A., et al. (2020). Genetic and genomics in congenital heart disease: A clinical review. *Jornal de Pediatria, 96*(3). https://doi.org/10.1016/j.jped.2019.07.004

Sapra, L., Sandhu, J. K., & Goyal, N. (2021). Intelligent method for detection of coronary artery disease with ensemble approach. *Advances in Communication and Computational Technology,* 1033–1042.

Shahid, Afzal Hussain, & Singh, M. P. (2020). A Novel Approach for Coronary Artery Disease Diagnosis using Hybrid Particle Swarm Optimization based Emotional Neural Network. *Biocybernetics and Biomedical Engineering, 40,* 1568–1585.

Spencer, R., Thabtah, F., Abdelhamid, N., & Thompson, M. (2020). Exploring feature selection and classification methods for predicting heart disease. *Digit. Heal., 6.* https://doi.org/10.1177/2055207620914777

Tay, D., Poh, C. L., Van Reeth, E., & Kitney, R. I. (2015). The effect of sample age and prediction resolution on myocardial infarction risk prediction. *IEEE J. Biomed. Heal. Informatics, 19*(3). https://doi.org/10.1109/JBHI.2014.2330898

Tuli, S., et al. (2020). HealthFog: An ensemble deep learning based Smart Healthcare System for Automatic Diagnosis of Heart Diseases in integrated IoT and fog computing environments. *Futur. Gener. Comput. Syst., 104.* https://doi.org/10.1016/j.future.2019.10.043

Yamamoto, S., et al. (2020). Impact of physical performance on prognosis among patients with heart failure: Systematic review and meta-analysis. *J. Cardiol., 76*(2). https://doi.org/10.1016/j.jjcc.2020.02.022

Yang, Xin-She, & Deb, Suash (2014). Cuckoo search: Recent advances and applications. *Neural Computing and applications, 24,* 169–174.

Yu, J., & Liu, G. (2020). Knowledge-based deep belief network for machining roughness prediction and knowledge discovery. *Comput. Ind., 121.* https://doi.org/10.1016/j.compind.2020.103262