ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

DEVELOPED A HYBRID IMPROVED WEIGHED CUCKOO SEARCH WITH A DEEP MASK CONVOLUTIONAL NEURAL NETWORK TO PREDICT THE HEART DISEASE AT AN EARLY STAGE

Dr.S.Sathish kumar
Assistant Professor, School of Computer Science, VET Institute of Arts and Science (Co-education), Thindal, Erode, Tamil Nadu, India

Dr.S.Prasath
Assistant Professor, School of Computer Science, VET Institute of Arts and Science (Co-education), Thindal, Erode, Tamil Nadu, India

Abstract:

The biggest medical issue or obstacle that contemporary medicine faces worldwide is cardiovascular disease. It is now a major contributing element to the rising death rate. If heart illness is not detected early on, its severity is much greater and may have dangerous repercussions. Techniques include electronic records of health, ongoing body surveillance via a computer system, consumer health status diagnosis using wearable devices, and healthcare device projections on the bodies of humans.A heart attack diagnosis is discovered to be a major problem, so to take the necessary steps, a diagnosis must be made remotely and frequently. Finding the incidence of coronary artery disease has emerged as a major field of inquiry for scientists in contemporary society and multiple approaches have recently been developed. An important factor in the very accurate detection of cardiac issues is the optimization method. A hybrid Improved Weighed Cuckoo Search with a Deep Mask Convolutional Neural Network (IWCS-DMCNN) approach has been presented to identify heart disease at an early stage.The proposed method is more efficient than other existing methods with a 98.59%. In comparison to an existing methods, the evaluation of the proposed approach improves the performance measures.

Keywords: Deep Mask Convolutional Neural Network, Improved Weighed Cuckoo Search, Heart Disease Prediction, Performance Measures.

1. Introduction

According to a World Health Organisation (WHO) study, heart diseaseor more specificallycardiovascular illness is an important contributor to the elevated mortality rate worldwide. One of the system's components, the heart, pumps and circulates blood to each part of the human physique, including the cerebral cortex, playing an essential role in all other sections [1].The heart can cease pumping circulation to the cerebral cortex and other bodily nerve endings, which can lead to the death of the nervous system itself, which means that all of the tissues in the body and neurons are going to cease functioning and eventually die. Consequently, a living thing's heart is its only source of energy [2].

Therefore, everyone needs their heart to work properly to live a healthy life. To lower the death rate, it's critical to diagnose the illness at its earliest stages and administer the proper care when needed.Heart illness, also known as cardiovascular disease, is categorized into several categories of disorders that must be anticipated early on. This is one of the latest illnesses that is killing people more often than any other in the entire world. This sickness has claimed the lives

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

of the majority of individuals [3]. Numerous risk factors for this illness must be prevented, and if an individual has previous heart problems, precautions must be taken.To lower the incidence of infections from cardiovascular diseases, individuals with heart disease or other cardiovascular disorders should adhere to security precautions and follow their physician's instructions [4].

Heart disease, sometimes referred to as Cardiovascular Disease (CVD) is a group of illnesses affecting the heart and blood vessels. Heart illness can be classified into several forms, with myocardial infarctionalso referred to as angina or a heart attackbeing one of them. Another cardiovascular condition is coronary heart disease, which is brought on by the build-up of plaque, a waxy material that accumulates in the inner region of the arteries in the heart.These coronary artery vessels are going to be utilized to feed the muscles of the heart with circulation that is high in oxygenation [5]. Atherosclerosis is the term for a disease that develops in these arteries when the waxy material known as plaque begins to build up. The artery's internal plaque build-up will continue for several decades. The ventricular septum may rupture, break open, or harden if the plaque's development is not detected in its early stages. The cardiovascular artery will begin to narrow due to plaque that has a tendency to be harder, which will lessen the amount of circulatory blood that is rich in oxygen reaching the coronary arteries [6].

There is going to be a lot of blood on the plaque's surfaces if it becomes harder. Most of the time, blood clots this big would stop the heart's arteries' ability to carry blood. Rapid retention of blood flow is required. If not, the heart muscle begins to die in that area. When heart illness manifests, prompt medical attention must be provided; otherwise, it might end in grave conditions that might result in mortality.The causes of the rising obstruction are associated with dangerous factors [7]. These danger variables can be divided into two categories: risk variables, which can be changed, and risk components, which cannot. Age, sex, and genetics are the criteria for non-modifiable risk factors. These unchangeable risk factors are going to be the primary cause of heart disease onset. Risk factors that we might alter on our own are known as risk variables that can be altered [8]. Modifiable hazards include (1) miscellaneous and biochemical risk factors; (2) stress-related risk factors; (3) habit-related risk factors; and (4) food-related risk factors. Heart illnesses come in different forms: cardiovascular, inflammatory, inheritable, and myocardial in arrhythmia, cardiac arrest, and cholesterol.

Furthermore, there's a potential that these risk factors could make the illness that's already present more severe. Figure 1 shows the risk factors that impact the heart in terms of visual style. Among the many cardiovascular disease indicators are: 1) tobacco; 2) a family history of cardiovascular illness; 3) hypertension; 4) hypercholesterolemia; 5) insulin; 6) overweight; 7) inactivity; and 8) tension. A broader category of disorders attacking different parts of the heart is called cardiovascular illness. 'Cardio' is the word in medicine for the meaning of heartbeat. Therefore, cardiovascular illness is the umbrella term for all ailments [9].

Seventy-five percent of the population dies from cardiovascular disease, which primarily affects those with middle- and lower-class incomes.Furthermore, strokes and heart attacks account for 80% of deaths caused by cardiovascular illnesses. India is among the nations where the number of people with coronary artery disease is rising year over year, according to WHO report. Every year, 2 lakh coronary artery bypass procedures are performed due to the rise in the number of patients affected by heart disease [10].

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

To safeguard against cardiac injuries and preserve the life of the individual in the event of an attack on the heart, medical professionals must respond as soon as possible. Medical professionals employ new technologies to continuously monitor patient information related to heart disease and to give ongoing guidance to individuals on how to recover from their condition [11]. These information mining methods combine elements of computational intelligence, technology for databases, and machine learning. Applications for mining information abound, and one of its most significant uses is its contribution to the early detection of illnesses [12].

The overall process for forecasting heart disease is similar to using an attribute extractor approach to separate important information from a large amount of information. Subsequently, the information obtained is going to be subjected to training using the chosen dataset before being provided to the assessment procedure [13]. Numerous methods are employed to forecast cardiac disease. Decision Trees (DT); Support Vector Machines (SVMs); Neural Networks (NN); and K-Nearest Neighbours (KNN) are the main categories of information mining methods. The development of heart disease and cardiovascular disorders has been predicted using various machine learning algorithms [14].

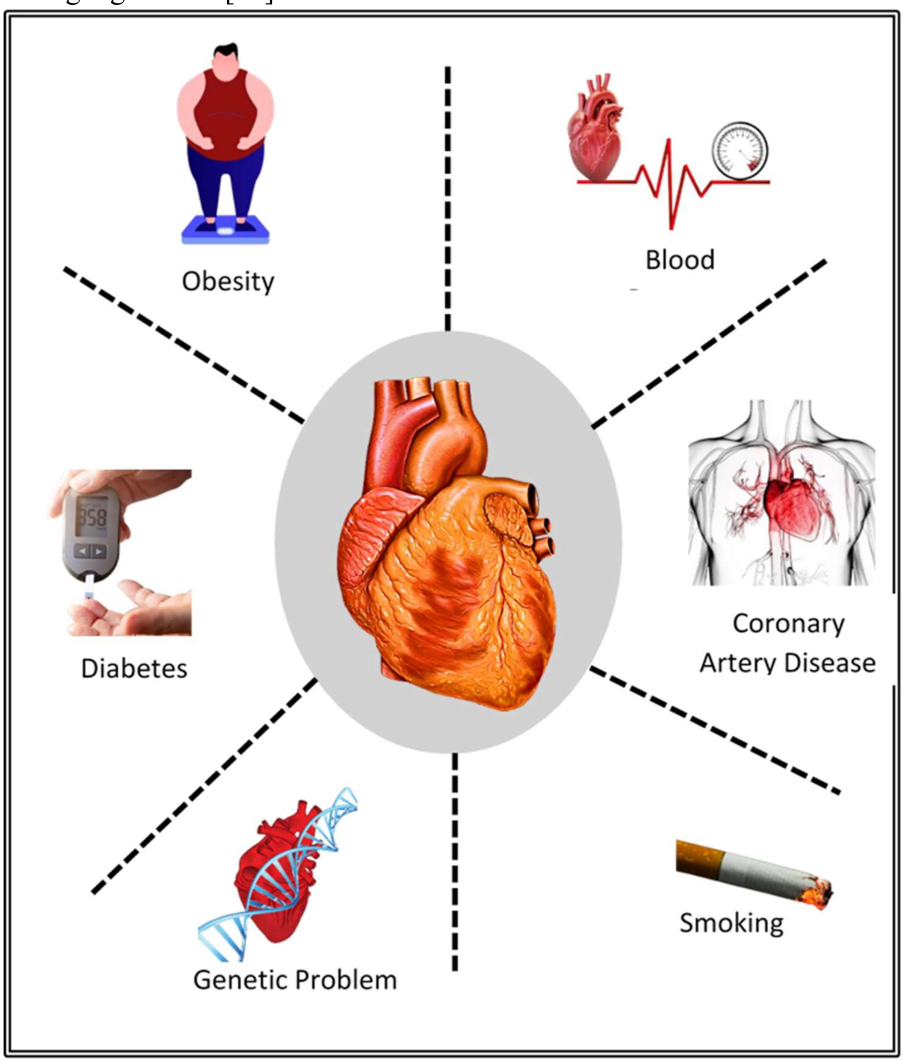


Figure 1: Risk factor of heart disease

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

Convolutional Neural Network (CNN) is useful for the analysis of images because of its capacity to recognize similarities and interpret them. CNN may be employed to quickly recognize and forecast cardiac illnesses by minimizing the number of variables and learning quantity, along with getting more accurate predictions using huge databases.CNN is composed of multiple sections, each of which processes input data according to its purpose [15]. Figure 2 describes the CNN topology. Convolution, maximal pooling, and ReLu constitute a few of the layers that make up the CNN structure. The convolution stage works with output volume ID filtration systems, machines, and image adjustments to create the dot products.The mechanism of the activation layer, which is the additional layer, applies the function of activation to the output that the convolution component produces. The layer with full connectivity retrieves the input from its predecessor and computes the result in the form of 1-D array category values. The next layer up is the pooling level, where storage effectiveness is carried out depending on the results from the preceding layer to minimize the costs of computation [16].

The CNN operates in conjunction with the classifications to generate the characteristics. The benefit of employing this CNN as an extractor of features is that it is the most straightforward way to apply a difference function to convert the volume of input to the quantity of output. This method's drawback is that it doesn't incorporate the thing's direction and placement into its forecasts. Furthermore, the learning procedure for the deep network is going to take more time [17].

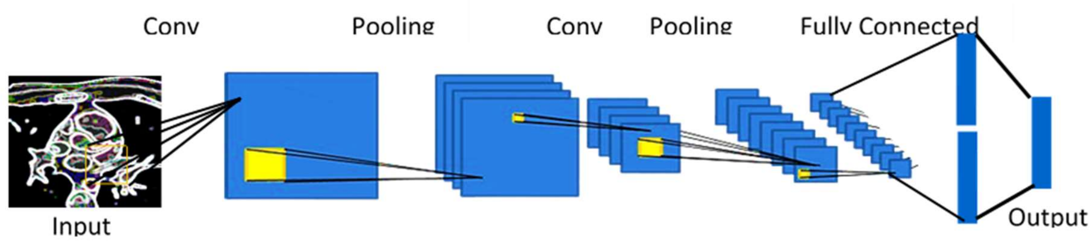


Figure 2: Typical architecture of CNN

Several factors, including recognition of patterns, ML, NN, and data are combined to create DM. It assumes that a method of sifting through vast databases to find concealed data exists in models and trends. The method of DM in the healthcare sector is depicted in Figure 3.A significant health problem that has impacted millions of individuals globally is Heart Disease (HD). The typical signs of HD include swollen feet, weakness of the muscles, and difficulty breathing. Since existing diagnosis and treatment approaches are poor in early identification for many causes, involving efficiency and speed of implementation, specialists are trying to create a successful method for early identification of HD [18].

Heart disease is very hard to identify and manage when advances in technology and healthcare providers are not accessible. Deep Convolutional Neural Networks (DCNNs) are referred to as heterogeneous if their classifications have various kinds and homogeneous if the classifiers are of a single type [19]. Every DCNN gets trained using its feature field, yet on occasion, the characteristics might include noise made up of duplicate and undesirable information. Both the training duration and the rate of false positives increase in these situations. The choice of features method is utilized to get around this issue. The categorization ensembles are the primary application for the characteristic identification approach. With the optimized characteristics, the DCNN performs better when a feature is chosen [20]. The Swarm method,

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

along with other techniques, is used to obtain the optimized characteristic collection. A lot of feature methods, including Particle Swarm Optimization (PSO), NN, SVM and deep learning computations, have been used to carry out the optimization [21].

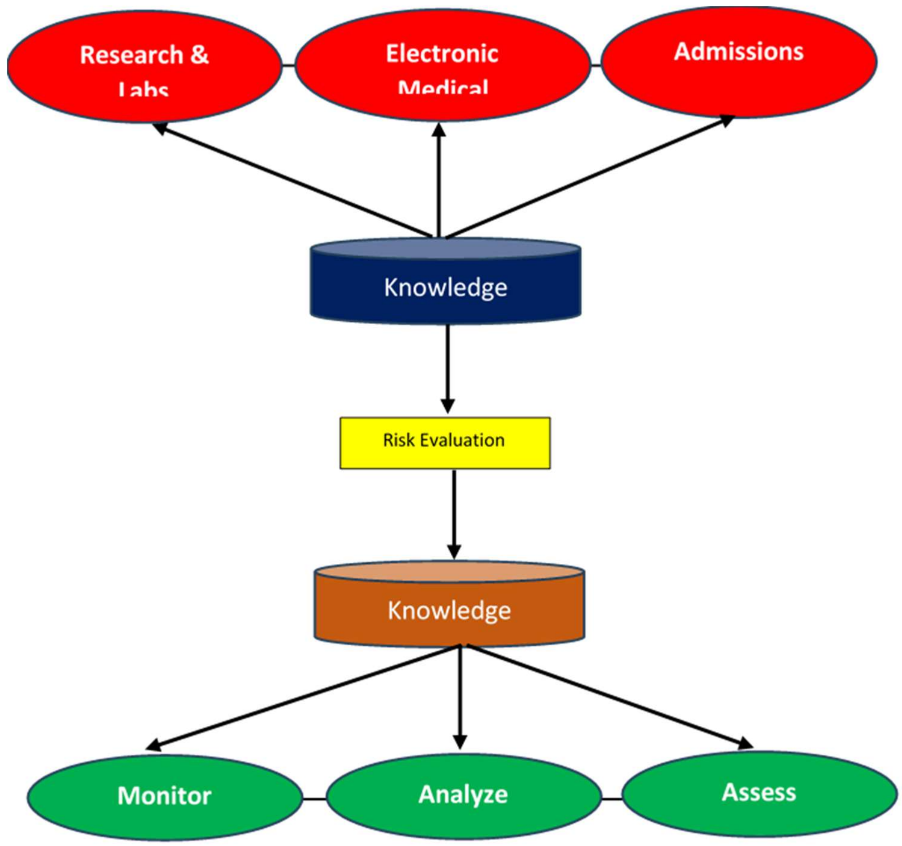


Figure 3: Flow of DM process in Healthcare

1.1 Problem Statement

In general, mining information is a rapidly developing technology that can be used in many different industries that are headed for success. Retail, e-business, and numerous other industries use information mining techniques. Methods based on data mining have recently been introduced into the field of medicine to predict or detect a variety of illnesses at their primary or initial stages. The use of preliminary data processing, characteristic extraction, and grouping approaches, which are the main components of data mining methods, can be used to predict heart disease [22].

Apart from the extraction of features, the categorization technique is a crucial component of artificial intelligence techniques. There are various sorts of classification methods, and everyone includes advantages and disadvantages of its own. In recent years, entropy-based randomly generated forest classifiers, MLPs, and ANNs with weights have been successfully employed.To create effective detection and forecasting of heart conditions, the method known as deep learning may be combined with additional methods of machine learning. One of the issues with applying machine-learning approaches is the use of natural categorization, whereas the use of just one classification algorithm could be deemed static [23].

1.2 Motivation

According to research published in the World Heart Federation Report, early detection of coronary artery disease may reduce the mortality incidence by one-third. Pain in the chest, difficulty breathing, headaches, discomfort in the jaw, oesophageal reflux, upper pain in the stomach, or back pain are common signs of heart disease. In addition, a few cardiovascular risk factors that help reduce the likelihood of heart disease include regularly exercising, low cholesterol levels, regulated blood pressure and quitting drinking [24].

Most of the time, a heart attack or stroke happens before the heart disease is discovered. Thus, monitoring cardiac indicators and consulting with physicians are essential. The medical field can now collect and store ongoing health information, which helps with important medical decisions, thanks to technical advancements in computing and information.To make critical medical choices such as diagnosis, treatment strategy, prognosis, and analysis of images, the information that has been saved may be analyzed. Unfortunately, the health care system's accessible information is rich but lacks expertise. In recent decades, Machine Learning (ML) approaches have played a significant role in handling complicated issues, including extremely nonlinear forecasting and categorizing. As a result, it is likely to develop a framework that may determine if heart disease is present or absent based on various heart-related indications [25].

A robust and open medical system is the primary objective and the primary requirement of the majority of the world's nations. The lack of this system causes misunderstandings and a communication gap between individuals and clinicians, which delays or prevents treatment. For better care and faster diagnosis, the medical professionals in these nations can view each patient's record in its entirety. This approach will be especially useful in situations where the patient suddenly loses consciousness and cannot provide an explanation. Thus, a centralized approach may prove more beneficial in minimizing the use of the wrong drug and other similar errors, improving the likelihood of recovery.

1.3 Objectives

The primary goal of this research is to use a method based on machine learning to accurately determine whether a patient has had cardiac disease. With the aid of stored information as the input value, healthcare experts are gathering information about patients. The machine-learning algorithm receives the input value and uses it to forecast the likelihood of a heart attack or stroke. The following are the research's primary goals:

- Efficient extraction and categorization of features are necessary to improve the forecast algorithm's success percentage, and this should be taken into account when diagnosing coronary artery disease at the beginning of the process.

- The Deep Neural Network with convolution was utilized to gather the characteristics and perform the level fusion function of the feature that was removed to identify the important and pertinent characteristics.

- The classifier's efficiency can also be enhanced by applying a dimensionality method. Both the level of detail and the amount of information kept are decreased as a result of the dimensionality reduction.

- To divide patient information from the information being entered into categories such as

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

good and questionable information.

2. Related Works

Modern computer technology and communication techniques have made it possible for medical professionals to compile and store patient data on an ongoing basis, which helps in decision-making. To determine the necessary medical choices, such as an estimate, evaluation, picture evaluation, and course of therapy, the saved health data may be examined. Currently, a range of machine learning approaches is frequently used to categorize and forecast illnesses [26].

Researchers reviewed the current machine learning algorithms for predicting heart disease in this section. In addition, an investigation of Outlier Detection (OD)-dependent heart disease models for prediction is conducted in conjunction with an evaluation of CDSS. In addition, a thorough examination of comparisons is performed to determine the traits of the evaluated models for forecasting [27].Long periods are needed, which are very hard to come by in many real-world situations involving sizable and biological datasets. People are becoming more concerned about their health due to the constant rise in the population, the prevalence of chronic diseases, and the rapid development of Information and Communication Technology (ICT) and the Internet of Things (IoT). The system of healthcare has evolved from conventional e-health to m-health, telemedicine, and currently ubiquitous healthcare (u-health) [28].

Each patient's behavior is monitored by the U-Healthcare structure, which notifies medical staff or clients for investigation or trend analysis. The proposed approach for heart disease diagnostics may be integrated with the U-Healthcare surveillance framework, which may notify medical personnel or patients while supporting doctors and cardiologists in making decisions.In a study, different DM techniques were used to analyze the forecasting of heart disease [29]. One of the DM tools used is Weka 3.6.6. The results of the simulation showed that NN produced 100% accurate results, while Naïve Bayes and DT produced 90% and 95.62% precision, respectively. Fuzzy Logic and Genetic Algorithms (GA) combine fuzzy-trained systems for experiments and GA for choosing characteristics in a mat lab by using a fuzzy tool [30].

In a study, different DM techniques were used to analyse the forecasting of heart disease. One of the DM tools used is Weka 3.6.6. The results of the experiment showed that NN produced 100% accurate results, while Naïve Bayes and DT produced 90% and 95.62% precision, respectively. Fuzzy Logic and Genetic Algorithms (GA) combine fuzzy expertise systems for testing and GA for choosing characteristics in a mat lab by using a fuzzy tool. Improved the HDPS, and it combines two input characteristics, including being overweight or smoking, to fully increase the estimated accuracy. In contrast to NB and DT, NN provides results that are 99.25 percent precise, whereas DT and NB achieve 96.66% and 94.44% precision, respectively [31]. DM for medical evaluation via a combination of techniques including connection and grouping standards, ANN, time series analysis, soft computing techniques, and so on. While DT yields superior results and Bayesian classification shares the same effectiveness as DT, NN, KNN, and clustering-based categorization are not doing well. The use of GA improves DT precision and Bayesian categorization [32].

Selecting features, also known as searching for future distances, is the process of identifying the best characteristic within a subset. Each characteristic of the set of data has an assigned weight value, which is then computed and contrasted to the initial collection [33].When a subset of

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

characteristics is created using the wrapper approach by adding and removing characteristics, the preciseness of each feature subset is determined to determine the effectiveness of the subset's creation. When contrasted with the filtering approach, the wrapper approach yields superior outcomes, according to the investigator's different outcomes [34].

Solutions to assist doctors in eliminating medical mistakes have been developed as a result of the problem's extent and growth. The current computer tools for watching over patients are mostly knowledge-based. The ability to track data using pre-existing information from medical specialists and reflected computing expertise. The construction of these frameworks takes time, and health care is covertly limited.It provides a thorough explanation of the methodology along with the real results. It is assumed that detecting abnormal patient movements or unforeseen behaviors can help identify medical malpractice. The outlier-based surveillance and alerting strategy is used in conjunction with knowledge economy reporting systems to improve the overall clinical alerting frequency. This innovative approach has the potential to serve as an established paradigm in medical subspecialties where knowledge-based warning is still lacking [35].

Patient warning and surveillance technologies are often used by hospitals using Electronic Health Records (EHRs) to improve patient care. These devices swiftly review information about patient examples that require clinical expertise to identify diseases or occurrences. Announcements are sent out as alerts or notifications when an event is present. One important alert module involves identifying possible errors in handling patients. Examples include notifications that are programmed to identify important medication mistakes, prescription inconsistencies, and important results of laboratory tests related to the patient's healthcare and status [36].

Most current computer systems that identify problems and raise alarms when they occur are based on knowledge. Healthcare specialists' expertise is codified and provided as rules that are applied to consumer information to identify dangerous illnesses and occurrences.Here, DM is utilized to examine linkages and concealed trends in heart disease across the medical information at hand. Machine learning methods have become more important in recent years for resolving complex, nonlinear categorization and prediction issues [37].

As a result, it is likely to create an algorithm for forecasting that determines the presence of heart disease based on several heart-related complaints (features). An examination of the current machine learning techniques for predicting heart disease has been provided, along with an extensive comparison aimed at defining the features of the algorithms under evaluation.Each characteristic of the set of data has an assigned weight value, which is then computed to be compared to the original dataset. When the feature subset is created using the wrapper approach by adding and removing features, the precision of the subset of features is determined to determine the effectiveness of the subset. When contrasted with the filter approach, the wrapper approach yields greater outcomes, according to the investigators various findings [38].

3. Proposed System

The tedious, demanding, and stressful process of transferring information from an expert in the field to a computer program mostly relies on the judgment of the medical specialist. IWCS-DMCNNwas employed as an example of deep learning for predictions to effectively tackle this situation. The proposed methodology's structure is depicted in Figure 4. Information about patients is gathered at the beginning of the procedure. After collection, the information can move

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

on to the pre-processing stage. The method described in the data initial processing portion is used to either fill in or eliminate the values that are not present.

Several prepossessing methods have been employed on the information to enhance it. The test and training datasets are the two subsets created from the cleansed information. Using the training information from the IWCS-DMCNNmethod, the model was trained using the initial information. The proposedIWCS-DMCNN model efficacy was assessed using the evaluation information. The outcome is the output of the proposed method following model testing and training processes. The individual in question is classified either well or poorly by the proposed method.

3.1 Feature Selection

The choice of features is a crucial preliminary processing phase that is applied to several techniques for data mining, such as pattern identification. By eliminating the unnecessary characteristics in the information space, the method for choosing characteristics only chooses the features that are required. If not, the undesired information adds redundant and undesired information to the computation, lengthening the computation and making it more complicated.

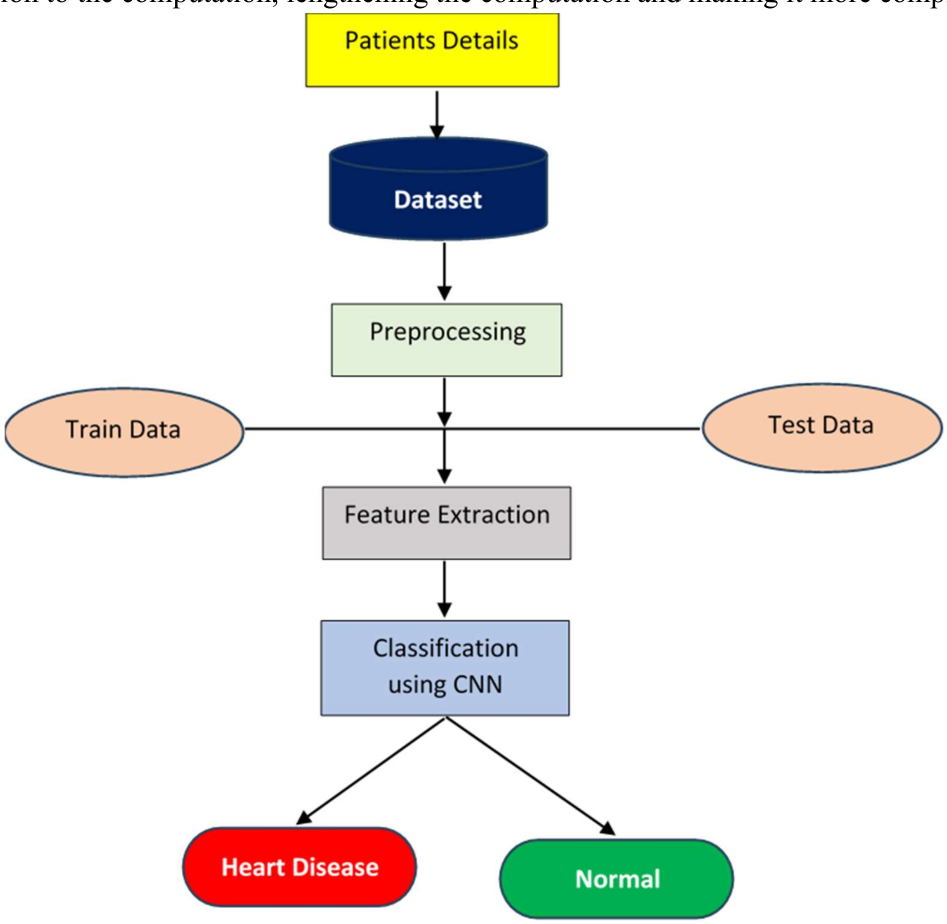


Figure 4: Proposed methodology

Significant features are chosen in choosing characteristics based on how significant they are, making it simple to classify features without modifying the initial subset. Numerous studies have demonstrated that characteristic selection-derived characteristics have a higher success rate in categorization than characteristics selected at random. Numerous algorithms have been created

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

to choose the attributes. The two techniques for choosing characteristics and pattern identification are filter methodology and the wrapper approach.

It is referred to as a filter-based strategy and depends on the information's characteristics if the method for choosing features is autonomous. It is referred to as a wrapper strategy if the classifier is employed; the characteristics derived from the wrapping approach are solely dependent on the classification algorithm employed. Two distinct subsets of attributes are offered by two distinct two different classification methods offer two distinct subsets of characteristics. Classifier methods. Although the wrapper technique requires more time to complete, the feature subset it yields is more efficient than the filter approach.

3.2 Dataset

There were 1050 patients in the dataset that was gathered via Kaggle, and it had 76 attributes. Fourteen of the seventy-six traits were utilized in predicting cardiac illness.This is because these characteristics have a greater impact on the illness than the other characteristics. Before being used for categorization, the dataset is organized and processed to remove any duplicate or absent variables. 80% of the specimens in the one used for training information and 20% in the one used for testing databases were chosen at random. 205 samples have been used to conduct tests, and 820 pieces of information from a total of 1025 medical records were used for learning. TheIWCS-DMCNN technique was used to train the model using information for training, and its effectiveness was verified using information from validation.

3.3 Data Pre-processing

By eliminating any insufficient information from the gathered data set, pre-processing the data lowers the forecast rate and reliability. Before the model is trained, any information that is lost must be eliminated due to the possibility of human error. The dataset is pre-processed once it has been collected to eliminate any data that is lacking or has duplicate values. The collection of data is subsequently utilized to train the algorithm once superfluous variables are eliminated.When numerous entries that are absent occur in a single row, the row is removed from the information set. Researchers knew the widely used median and average approaches for numeric and category characteristics and, accordingly, provided data that was absent.

3.4 Proposed method

A hybrid IWCS-DMCNN yields superior results for solving optimization challenges. The DT, SVM, RF, and NB methods are the four methods that make up the IWCS-DMCNN. The artificial cuckoo colonies are employed in the proposed model's IWCS method to find characteristics and a particular portion of the characteristic generation. DMCNNis then used to evaluate the characteristic subset that is developed.

The precision of the ensembles is increased by the employment of genetic and colony algorithms in the IWCS, an intelligent approach to characteristic optimization. The proposed IWCS-DMCNN allows for the calculation of every characteristic's capability inside a group, or in the characteristic subsets.To determine the precision of the accessible characteristic in the portion, a 10-fold approach is applied. Binary strings 0 or 1 denote every used cuckoo. The total number of characteristics is equal to the dataset's duration, which signifies the characteristic choice process carried out by the cuckoo search. One indicates that something has been chosen in the binary string, whereas zero indicates that a characteristic has not been picked.

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

Eight thick layers preceded the two layers based on convolution in the algorithm. The completely interconnected dense structure included the 14 chosen properties. To develop the IWCS-DMCNN structure, an overall of eight entirely linked thick layers was employed shown in Figure 5. To prevent the model from excessively fitting, the percentage of participants who dropped out was set at 3%.A cuckoo search is used once the breeding procedure for the initial group of birds has been completed. The process of extracting a characteristic from the data set is called initializing cuckoos. In the first place, C_x is a characteristic determined by the cuckoo in the Equation (1) and placed at random locations in the search area:

$$C_x = C - 1, C_2, \dots, C_n, \quad where, x = 1,2,3, \dots, n \quad (1)$$

For the meta-heuristics optimization designs, several first choices are utilized in the beginning to track the disparity between answers concurrently and improve. The meaning of an opposing point may be expressed using the Equation (2):

$$\hat{C}_x = g_y + h_y - c_x \quad (2)$$

Equation (3) may be used to derive the function of fitness for the OC_x, depending on the assumed function.

$$OC_x = MAX(Accuracy) \ (3)$$

To create a new position, one of the flocks of birds is randomly selected and the innovative position of the bird is obtained.

Algorithm: IWCS-DMCNN

Step 1: Start
Step 2: Random population initialization
Step 3: Cuckoo unevenly tailed (heavy)
Step 4: Compute the fitness value
Step 5: DMCNN detect the food available place
Step 6: Shelter or nest unevenly chosen using DMCNN Choose the shelter or nest amidst unevenly
Step 7: If Fitness < the best solution
  Step 7.1: Update the position
  else
  Step 7.2: Use the current nest as solution
Step 8: From steps 7.1 to 7.2, it perform in finding the position (worst case nest) or create a new nest
Step 9: Update and store the best solution using DMCNN
Step 10 If number of nest <= iterations (maximum) then find the best solution or else goto step 2.
Step 11: End the process

Using the following Equation (4):

$$PC^{x+iter+1} = \begin{cases} PC_{x+iter+1} + a_x \times cl^{x,iter} \times \left(M^{y,iter} - PC^{x,iter}\right) \, if \; a_y \geq APC^{y,iter} \\ \quad ran \; PC \; otherwise \end{cases} \quad (4)$$

Update the position

$$M^{x,iter+} = \begin{cases} PC^{x,iter} \; c\left(PC^{y.iter+}\right) > c\left(M^{y,iter}\right) \\ M^{x,iter} \; otherwisee \end{cases} \quad (5)$$

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

In the maintained posture, the cuckoo's current location's health value is regulated to be at its maximum. Cuckoo frequently adds additional places to storage spaces. The optimal location of memory that corresponds to the goal shall be regarded as the ideal filtering set of feature resolution if multiple rounds are carried out. The purpose of every cuckoo is to determine the role's appropriateness and health evaluation:

$$fitness_1(S) = \frac{\sum_{y=1}^{m} accuracy(S)}{m} \quad (6)$$

$$fitness_2(S) = consensus(S) \quad (7)$$

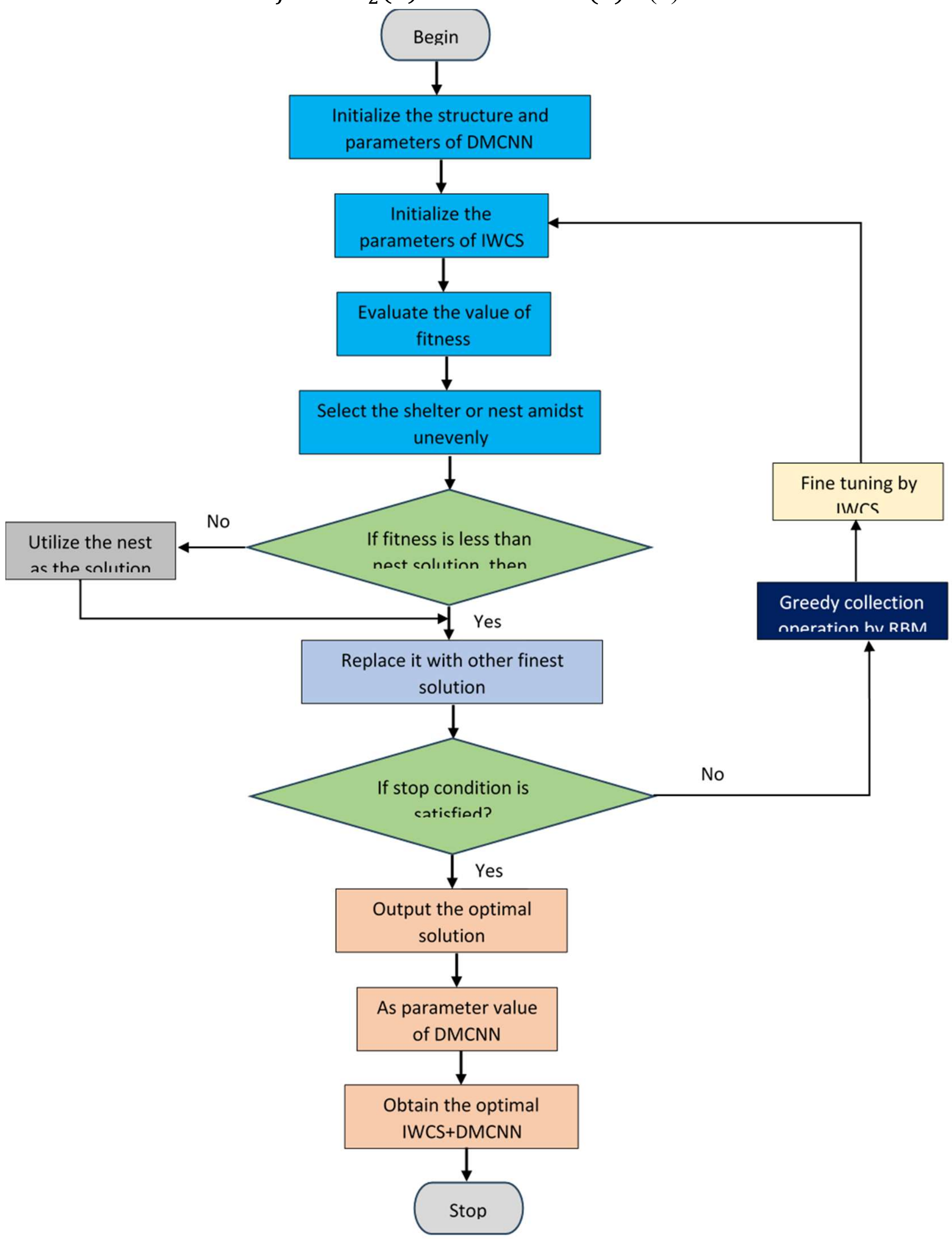


Figure 5: Proposed model of heart disease prediction in detail

$$fitness = \frac{fitness_1(S)+fitness_2(S)}{2} \quad (8)$$

The anticipated result in the group of classifiers is called efficiency (S), while the precision of classification in the S characteristic group is denoted by consensus (S). Mean precision and accuracy may be used to assess fitness; average precision determines whether characteristics have the capacity for precise categorization, whileIWCS-DMCNNattempts to optimize characteristics. The feature subset generalization capability is enhanced by its median precision.Consensus assessment, the subsequent stage of fitness assessment, determines if a feature collection is best suited for achieving a high degree of accuracy in categorization. The data is passed to the onlooker cuckoo by the cuckoo searches, which use Equation (8) to determine the likelihood of choosing a feature shown in Figure 6.

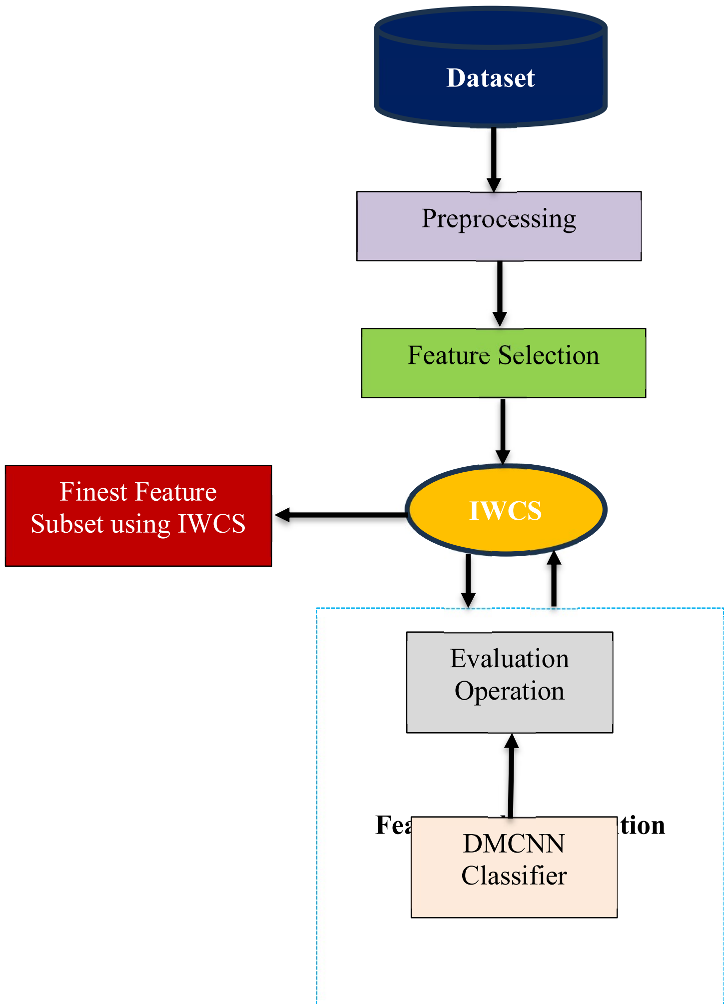


Figure 6: Proposed Architecture of IWCS-DMCNN

$$P_x = \frac{fitness}{\sum_{x=1}^{m} fitness} \quad (9)$$

$$v_x = i_x + \mu_x(I_x - I_y) \quad (10)$$

Where I_x is the precision of the item that was chosen, and I_y is the reliability of the component that the cuckoo observer chosen. A number chosen at random in the spectrum of (0, 1) is represented by µi. As a result, when a fresh function is added to the cuckoo look, the birds

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

that are watching make complete use of it, establishing a fresh setup for the segment. Following this technique, all of the characteristics have been utilized to create a fresh characteristic subgroup, and the data that is now accessible in the characteristics attempts to advance towards an improved feature portion setting up. The staff member becomes a scout cuckoo if the cuckoo searching is not improved. Next, the scout cuckoo, which is shown as follows, is given the recently introduced segment:

$$I_{xy} = I_y^{max} + rand(0,1)\left(I_y^{max} - I_y^{min}\right) \quad (11)$$

Where the minimum value is represented by I_y^{min} and the maximum value by I_y^{max}. To obtain the best characteristics, both the lower and upper bounds and the same procedure are continued until the end requirements are met. The algorithm for genetics uses the best possible outcome as input, evaluating every characteristic's health. This allows the IWCS-DMCNN to choose the characteristics according to their classification, which reduces the amount of time that the noisy and unfavorable characteristics take to select the key feature from the characteristic group. Since there are numerous characteristics to manage in a large set of data, the accuracy of the classifier suffers. The DMCNN method improves classifier computational efficiency by prioritizing features according to their significance as shown in Figure 7.

The spatial frames were provided into a box extrapolation layer and a categorization layer after their mapping to a low dimensionality characteristic. In this paper, IWCS-DMCNN where the RPN was removed, the final layer of the third MLPCONV framework was used to map the features of the neural network proposed. For minute details, this particular layer has a substantial level of abstraction.

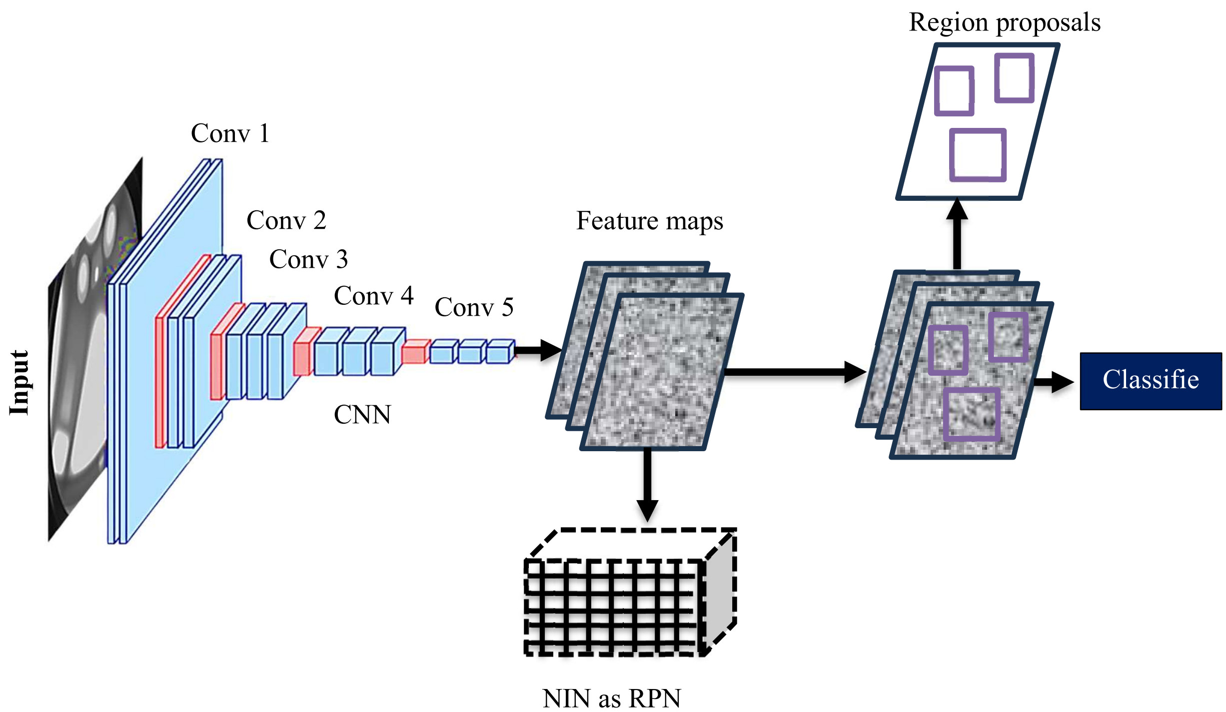


Figure 7: DMCNN Architecture

4. Results and Discussions

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

Among the 303 examples with 14 features in the Cleveland heart disease collection of data, 138 have heart disease and 165 are healthy. The Iraqi dataset on Erbil heart illness has 333 samples with 20 attributes; 118 of these samples have heart disease, whereas the remaining 215 do not. In this phase, information processing is used to recognize null values, filter corrupt, gone, inconsiderate, and wrong standards, and eliminate the repetition of some features.Subsequently, splitting, feature scaling, and normalization are used to determine the typical information format. The dataset is divided into a collection for training, which comprises 80% of the information, and an assessment set, which includes 20% of the information after the data has been prepared.Segmentation of sufferers into groups with and without cardiac disease: There are 138 individuals with heart disease and 165 patients without heart disease as shown in Figure 8 and the model assessment.

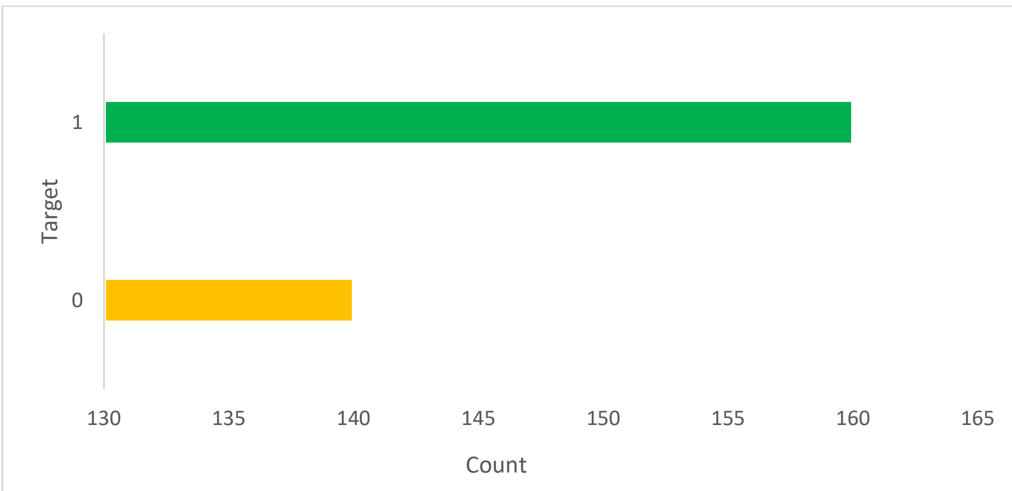
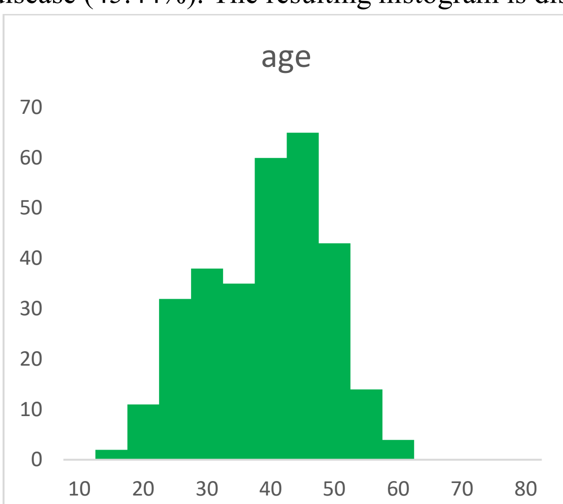
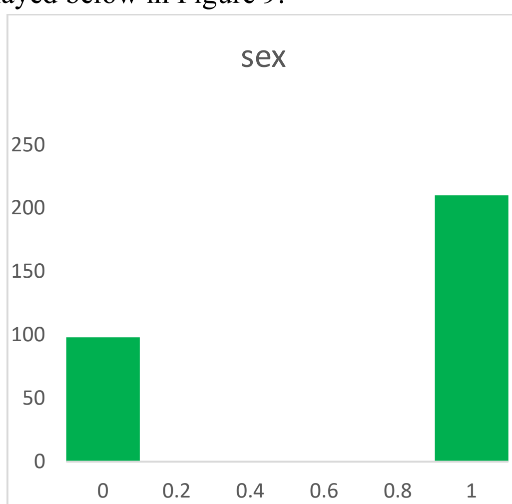


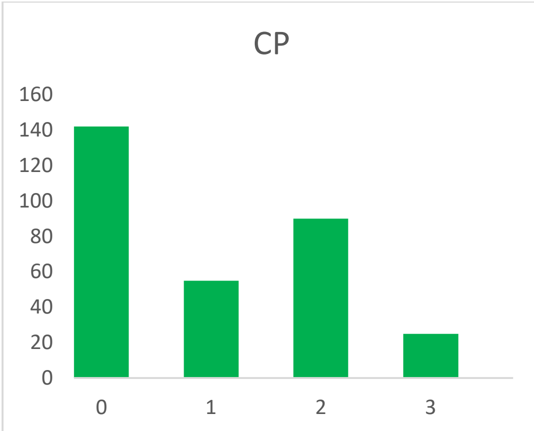
Figure 8: Classification of patients with Heart Disease and without Heart Disease in Dataset

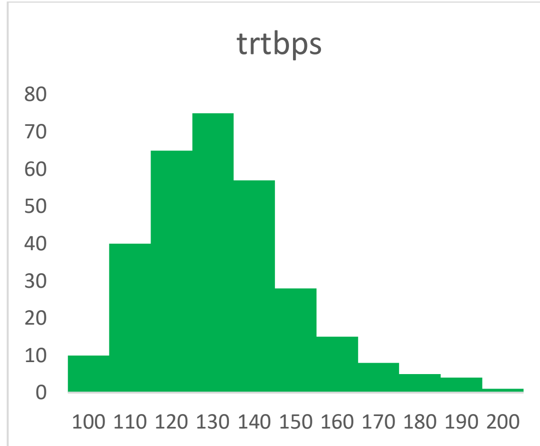
Figure 8 shows that 138 patients have heart disease (54.56%) and 165 patients do not have heart disease (45.44%). The resulting histogram is displayed below in Figure 9.

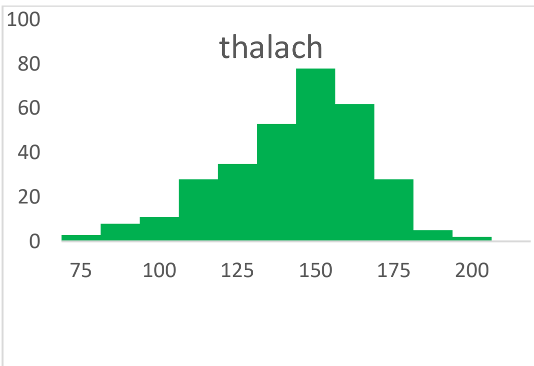


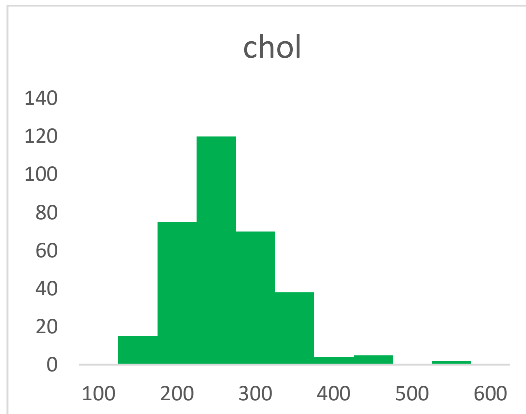


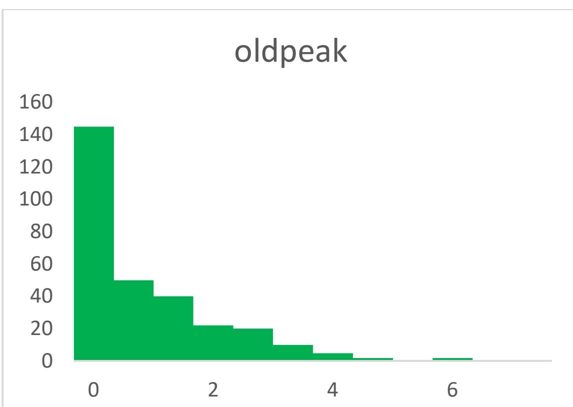
ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

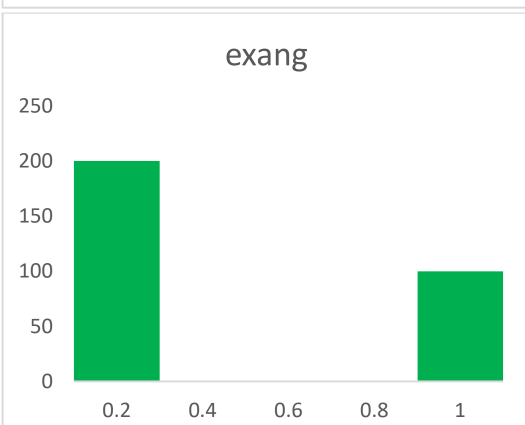












ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

Figure 9: Numerical and categorical attributes histogram in Dataset

There are no values that are missing in the numerical attributes of the Dataset as shown by Figure 9 statistical aspects of the value of the numerical characteristics, which include the minimum (Min), mean (SD), missing (M), and distinct values (D), as shown in Table 1.

Table 1: Numeric attributes of the Dataset

Attributes	Min	Max	Mean	SD	M	D	Proportion
Age	30	78	55.35	9.15	0	42	14.52%
Trtbps	95	202	132.66	18.54	0	50	16.24%
Chol	127	566	247.27	52.95	0	153	51.20%
Thalach	72	204	150.65	23.12	0	92	31.00%
oldpeak	0	6.4	1.040	1.18	0	42	14.20%

Table 2: Categorical attributes of the dataset

Attributes	Label value	Count	Proportion(%)	M	D
Sex	0	97	32.65	0	3
	1	208	69.32		
cp	0	144	48.20		
	1	51	17.50	0	5
	2	88	28.72		
	3	24	0.09		
fbs	0	259	86.18	0	33
	1	46	15.83		

restecg	0	148	49.52		
	1	153	5019	0	4
	2	5	1.45		
exang	0	205	68.35	0	3
	1	100	33.64		
slp	0	22	7.92		
	1	141	47.20	0	4
	2	176	47.90		
	3	176	58.29		
ca	0	66	22.46		
	1	39	13.52	0	6
	2	22	7.62		
	3	6	1.86		
target	0	139	00.68	0	5
	1	166	6.12		
thal	0	3	55.65		
	1	19	39.32	0	3
	2	167	46.54		
	3	118	55.47		

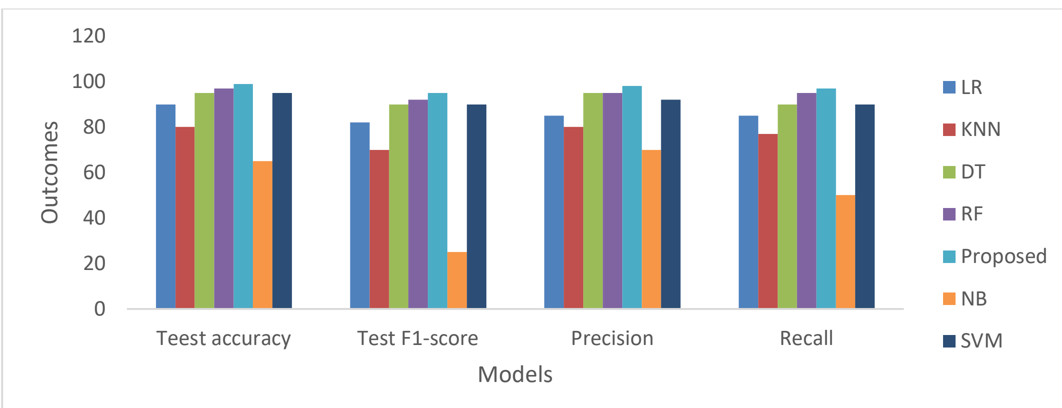


Figure 10:Performance measures of proposed and existing systems

The standard settings are utilized when using artificial intelligence methods in the present study. The results of the method are displayed in Table 2 and Figure 10.In this work, machine learning techniques are applied in normal circumstances. The outcomes of the procedure are shown in Figure 10 and Table 2. The level of engagement also includes the interests of the public, free software, business, and academia. Visual analysis, conversational interpretation, cleanup evidence, and acceptable resolution of issues are used to assess technical efficiency.

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

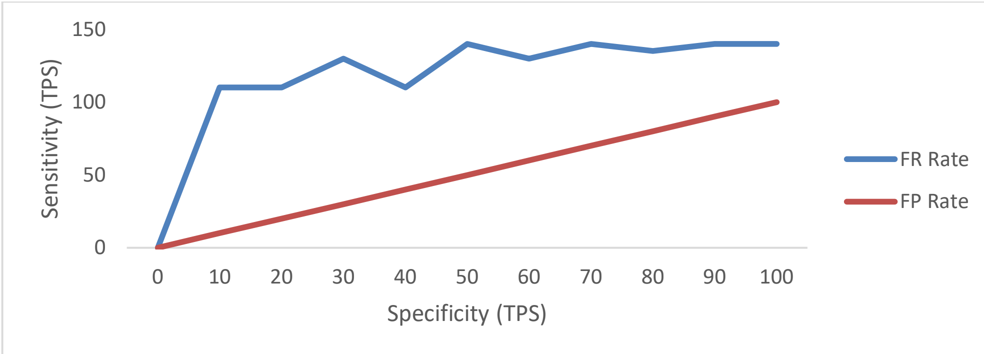


Figure 11: Sensitivity vs. Specificity of proposed system

Additionally, we recognize that the threshold is nearly 140 if we consider that one of the ideas of machine learning using IWCS-DMCNN. The calculation of points could be dependent on a specific collection of training or test groups in certain programs, and a range of trust around the estimated value shown in Figure 11.

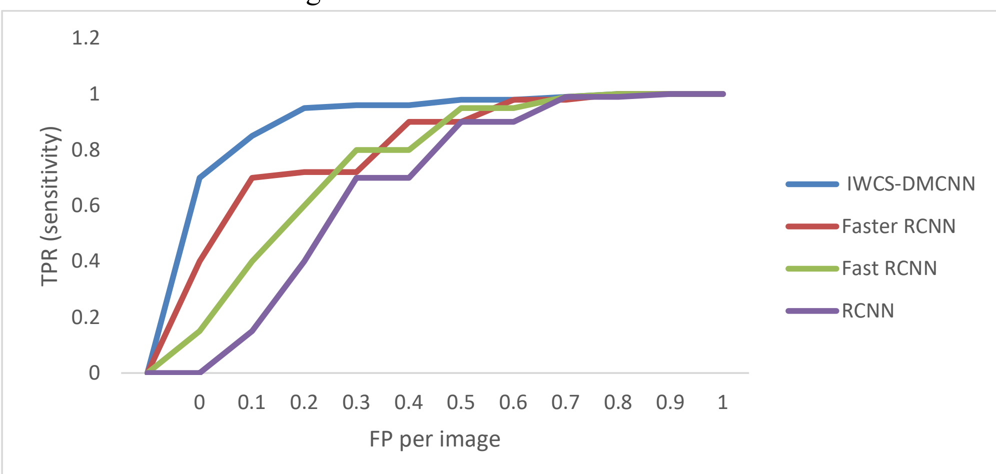


Figure 12: Comparison of IWCS-DMCNNwith existing systems based on FROC curves

Using dataset, IWCS-DMCNN has been contrasted with R-CNN, fast R-CNN, and faster R-CNN. Figure 12 shows the FROC curves and accompanying AUCs, showing that IWCS-DMCNNperforms best among different techniques utilizing both sets of data.An additional method of assessing the proposed approach is to calculate the losses associated with validation and training values. Figure 13 shows a graphical illustration of it. It shows that the verification loss is 17.39% and the learning loss is 24.02. Training and validation loss comparison shown in Figure 14.

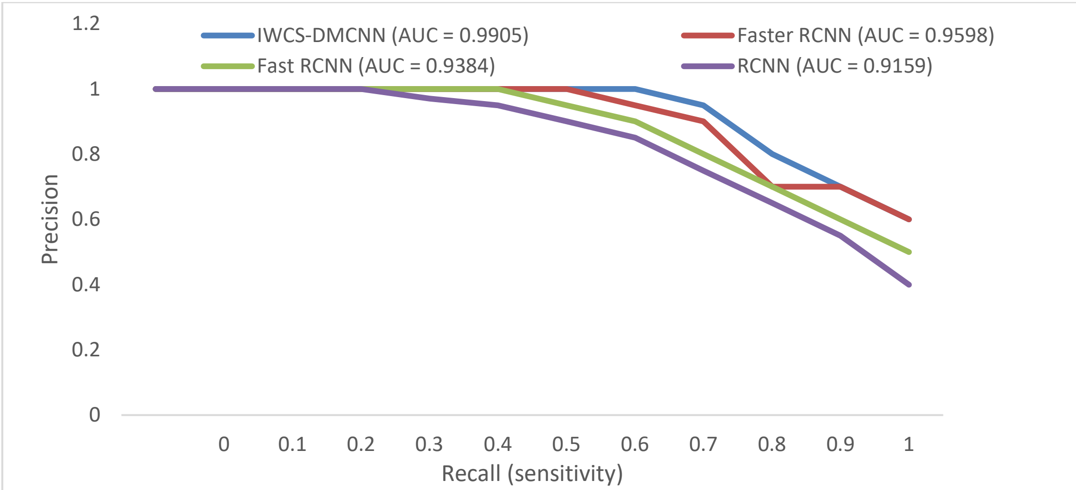


Figure 13: Comparison of IWCS-DMCNNwith existing systems based on Precision recall curves

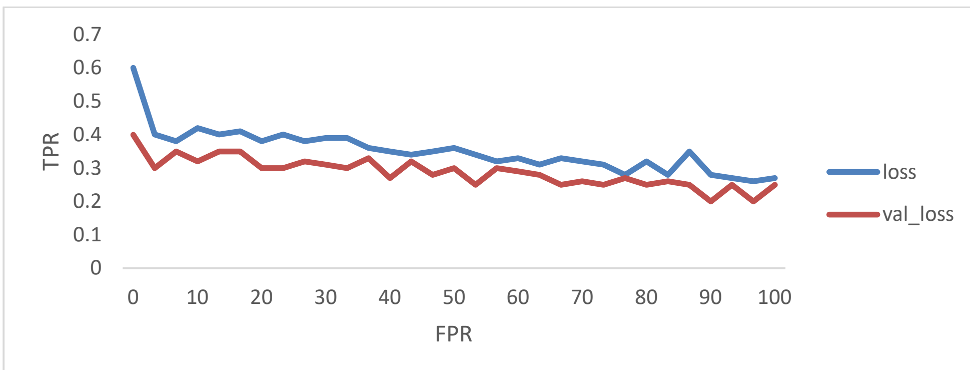


Figure 14:Comparison of training and validation loss

Figure 15 displays a graph that represents the precision of both the training and the validation data. After validation of the information, the training precision was 91.91% and the reliability was 90.12%. The area under the curve (AUC) and categorization algorithm efficiency are inextricably linked as shown in Figure 16. The ROC curve was used to assess the adaptability and durability of the proposed model. Figure 17 displays the receiver operator parameter curve for the proposed model.

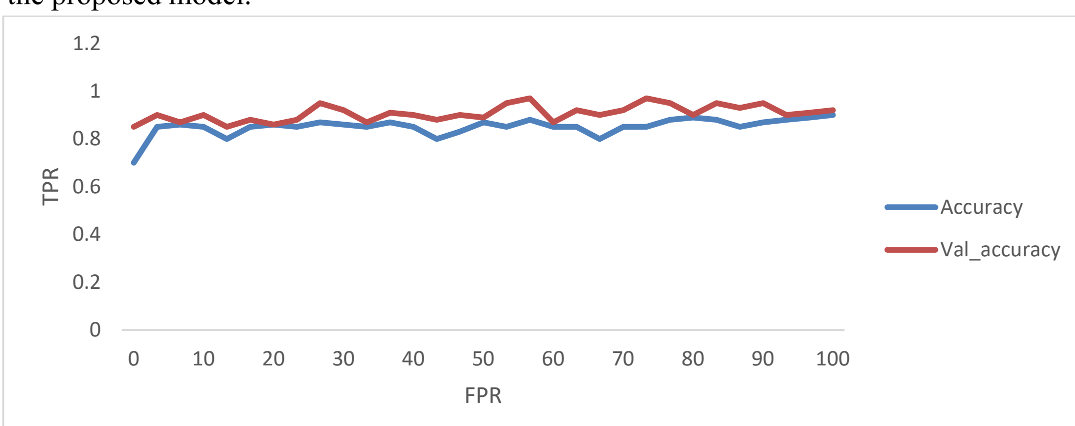


Figure 16:**Comparison of training and validation accuracy**

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

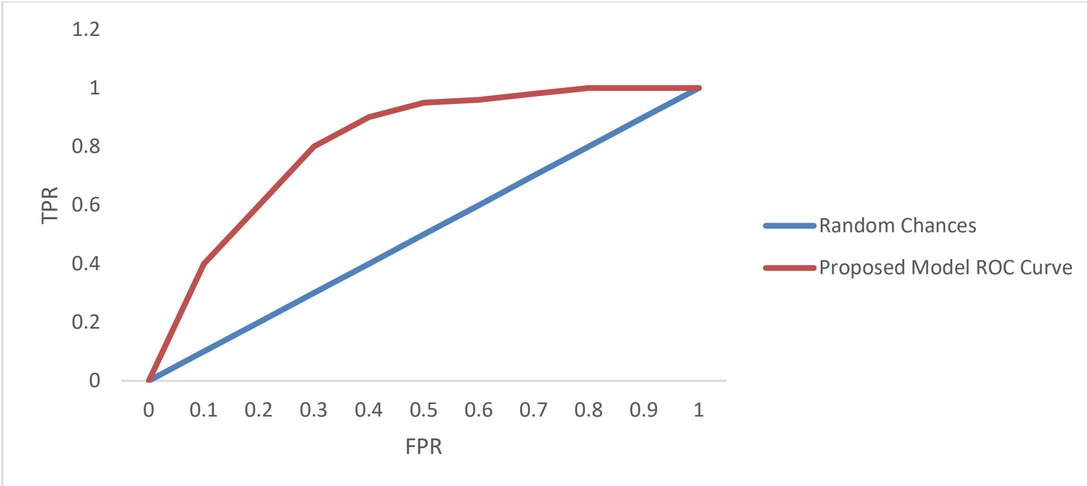


Figure 17: Proposed model ROC curve

5. Conclusions

Numerous accidents can be avoided by early detection of cardiovascular conditions. Physicians can identify potential coronary artery disease before symptoms appear by using an efficient method. This study dealt with the early identification of cardiac illnesses using a cutting-edge UCI library. Fourteen of the 76 instances in the original dataset which was gathered from the UCI repository were utilized in the forecast. The previously processed data was employed before learning the set.Everyone used both quantitative and qualitative assessment indicators to assess our proposed system, and the outcomes were more than satisfactory. A hybrid Improved Weighed Cuckoo Search with a Deep Mask Convolutional Neural Network (IWCS-DMCNN) approach has been presented to identify heart disease at an early stage. In terms of precision, recollection, precision and F1 score, the proposed system performed well as 91.71%, 88.88%, 82.75%, and 85.70%, in that order.Future studies might concentrate on applying this framework to other healthcare object identification domains and integrating cardiac cycle duration information into the identification and localization process.

References

1. Bhatt, C. M., Patel, P., Ghetia, T., &Mazzeo, P. L. (2023). Effective heart disease prediction using machine learning techniques. *Algorithms*, *16*(2), 88.

2. Yadav, A. L., Soni, K., &Khare, S. (2023, July). Heart Diseases Prediction using Machine Learning. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)* (pp. 1-7). IEEE.

3. Arumugam, K., Naved, M., Shinde, P. P., Leiva-Chauca, O., Huaman-Osorio, A., & Gonzales-Yanac, T. (2023). Multiple disease prediction using Machine learning algorithms. *Materials Today: Proceedings*, *80*, 3682-3685.

4. Chandrasekhar, N., &Peddakrishna, S. (2023). Enhancing Heart Disease Prediction Accuracy through Machine Learning Techniques and Optimization. *Processes*, *11*(4), 1210.

5. Khan, A., Qureshi, M., Daniyal, M., &Tawiah, K. (2023). A Novel Study on Machine Learning Algorithm-Based Cardiovascular Disease Prediction. *Health & Social Care in the Community*, *2023*.

6. Nandy, S., Adhikari, M., Balasubramanian, V., Menon, V. G., Li, X., &Zakarya, M. (2023). An intelligent heart disease prediction system based on swarm-artificial neural network. *Neural Computing and Applications*, *35*(20), 14723-14737.
7. Kresoja, K. P., Unterhuber, M., Wachter, R., Thiele, H., &Lurz, P. (2023). A cardiologist's guide to machine learning in cardiovascular disease prognosis prediction. *Basic research in cardiology*, *118*(1), 10.
8. Alkayyali, Z. K., Idris, S. A. B., & Abu-Naser, S. S. (2023). A Systematic Literature Review of Deep and Machine Learning Algorithms in Cardiovascular Diseases Diagnosis. *Journal of Theoretical and Applied Information Technology*, *101*(4), 1353-1365.
9. Bakar, W. A. W. A., Josdi, N. L. N. B., Man, M. B., &Zuhairi, M. A. B. (2023, March). A Review: Heart Disease Prediction in Machine Learning & Deep Learning. In *2023 19th IEEE International Colloquium on Signal Processing & Its Applications (CSPA)* (pp. 150-155). IEEE.
10. Radwan, M., Mohamed Abdelrahman, N., Wael Kamal, H., Khaled AbdelmonemElewa, A., &Moataz Mohamed, A. (2023). MLHeartDisPrediction: Heart Disease Prediction using Machine Learning. *Journal of Computing and Communication*, *2*(1), 50-65.
11. Kadhim, M. A., &Radhi, A. M. (2023). Heart disease classification using optimized Machine learning algorithms. *Iraqi Journal For Computer Science and Mathematics*, *4*(2), 31-42.
12. Li, F., Wu, P., Ong, H. H., Peterson, J. F., Wei, W. Q., & Zhao, J. (2023). Evaluating and mitigating bias in machine learning models for cardiovascular disease prediction. *Journal of Biomedical Informatics*, *138*, 104294.
13. Shukur, B. S., &Mijwil, M. M. (2023). Involving machine learning techniques in heart disease diagnosis: a performance analysis. *International Journal of Electrical and Computer Engineering*, *13*(2), 2177.
14. Dalal, S., Goel, P., Onyema, E. M., Alharbi, A., Mahmoud, A., Algarni, M. A., &Awal, H. (2023). Application of Machine Learning for Cardiovascular Disease Risk Prediction. *Computational Intelligence and Neuroscience*, *2023*.
15. Sk, K. B., Roja, D., Priya, S. S., Dalavi, L., Vellela, S. S., & Reddy, V. (2023, March). Coronary Heart Disease Prediction and Classification using Hybrid Machine Learning Algorithms. In *2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA)* (pp. 1-7). IEEE.
16. Trigka, M., & Dritsas, E. (2023). Long-term coronary artery disease risk prediction with machine learning models. *Sensors*, *23*(3), 1193.
17. Bani Hani, S. H., & Ahmad, M. M. (2023). Machine-learning algorithms for ischemic heart disease prediction: a systematic review. *Current Cardiology Reviews*, *19*(1), 87-99.
18. Abdulsalam, G., Meshoul, S., &Shaiba, H. (2023). Explainable Heart Disease Prediction Using Ensemble-Quantum Machine Learning Approach. *Intell. Autom. Soft Comput*, *36*, 761-779.
19. Behera, M. P., Sarangi, A., Mishra, D., &Sarangi, S. K. (2023). A Hybrid Machine Learning algorithm for Heart and Liver Disease Prediction Using Modified Particle Swarm Optimization with Support Vector Machine. Procedia Computer Science, 218, 818-827.

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

20. Dileep, P., Rao, K. N., Bodapati, P., Gokuruboyina, S., Peddi, R., Grover, A., &Sheetal, A. (2023). An automatic heart disease prediction using cluster-based bi-directional LSTM (C-BiLSTM) algorithm. Neural Computing and Applications, 35(10), 7253-7266.

21. Ozcan, M., &Peker, S. (2023). A classification and regression tree algorithm for heart disease modeling and prediction. Healthcare Analytics, 3, 100130.

22. Velmurugan, A. K., Padmanaban, K., Kumar, A. M., Azath, H., &Subbiah, M. (2023, January). Machine learning IoT based framework for analysing heart disease prediction. In AIP Conference Proceedings (Vol. 2523, No. 1). AIP Publishing.

23. Al Ahdal, A., Rakhra, M., Rajendran, R. R., Arslan, F., Khder, M. A., Patel, B., ...& Jain, R. (2023). Monitoring Cardiovascular Problems in Heart Patients Using Machine Learning. Journal of Healthcare Engineering, 2023.

24. Sudha, V. K., & Kumar, D. (2023). Hybrid CNN and LSTM network for heart disease prediction. SN Computer Science, 4(2), 172.

25. Forrest, I. S., Petrazzini, B. O., Duffy, Á., Park, J. K., Marquez-Luna, C., Jordan, D. M., & Do, R. (2023). Machine learning-based marker for coronary artery disease: derivation and validation in two longitudinal cohorts. The Lancet, 401(10372), 215-225.

26. Pant, A., Rasool, A., Wadhvani, R., &Jadhav, A. (2023, January). Heart disease prediction using image segmentation Through the CNN model. In 2023 13th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 374-380). IEEE.

27. Mishra, N., Desai, N. P., Wadhwani, A., &Baluch, M. F. (2023). Visual Analysis of Cardiac Arrest Prediction Using Machine Learning Algorithms: A Health Education Awareness Initiative. In Handbook of Research on Instructional Technologies in Health Education and Allied Disciplines (pp. 331-363). IGI Global.

28. Yaqoob, M. M., Nazir, M., Khan, M. A., Qureshi, S., & Al-Rasheed, A. (2023). Hybrid classifier-based federated learning in health service providers for cardiovascular disease prediction. Applied Sciences, 13(3), 1911.

29. Safa, M., Pandian, A., Gururaj, H. L., Ravi, V., &Krichen, M. (2023). Real time health care big data analytics model for improved QoS in cardiac disease prediction with IoT devices. Health and Technology, 13(3), 473-483.

30. Zhou, C., Hou, A., Dai, P., Li, A., Zhang, Z., Mu, Y., & Liu, L. (2023). Risk factor refinement and ensemble deep learning methods on prediction of heart failure using real healthcare records. Information Sciences, 637, 118932.

31. Hassan, D., Hussein, H. I., & Hassan, M. M. (2023). Heart disease prediction based on pre-trained deep neural networks combined with principal component analysis. Biomedical Signal Processing and Control, 79, 104019.

32. Dritsas, E., &Trigka, M. (2023). Efficient data-driven machine learning models for cardiovascular diseases risk prediction. Sensors, 23(3), 1161.

33. Biswas, N., Ali, M. M., Rahaman, M. A., Islam, M., Mia, M. R., Azam, S., ...&Moni, M. A. (2023). Machine Learning-Based Model to Predict Heart Disease in Early Stage Employing Different Feature Selection Techniques. BioMed Research International, 2023.

34. Gopisetti, L. D., Kummera, S. K. L., Pattamsetti, S. R., Kuna, S., Parsi, N., &Kodali, H. P. (2023, January). Multiple Disease Prediction System using Machine Learning and Streamlit.

ISSN:2153-182X, E-ISSN: 2153-1838
Vol. 17 No. 02 (2023)

In 2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 923-931). IEEE.

35. Gangadhar, M. S., Sai, K. V. S., Kumar, S. H. S., Kumar, K. A., Kavitha, M., &Aravinth, S. S. (2023, February). Machine Learning and Deep Learning Techniques on Accurate Risk Prediction of Coronary Heart Disease. In 2023 7th International Conference on Computing Methodologies and Communication (ICCMC) (pp. 227-232). IEEE.

36. Jain, A., Rao, A. C. S., Jain, P. K., & Hu, Y. C. (2023). Optimized levy flight model for heart disease prediction using CNN framework in big data application. Expert Systems with Applications, 223, 119859.

37. Ahamad, G. N., Shafiullah, Fatima, H., Imdadullah, Zakariya, S. M., Abbas, M., ...& Usman, M. (2023). Influence of Optimal Hyperparameters on the Performance of Machine Learning Algorithms for Predicting Heart Disease. Processes, 11(3), 734.

38. Menshawi, A., Hassan, M. M., Allheeib, N., &Fortino, G. (2023). A Hybrid Generic Framework for Heart Problem diagnosis based on a machine learning paradigm. Sensors, 23(3), 1392.