Journal of Information & Optimization Sciences
ISSN 0252-2667 (Print), ISSN 2169-0103 (Online)
Vol. 44 (2023), No. 5, pp. 865–894
DOI : 10.47974/JIOS-1175

CDIPEM: Cardiovascular disease incidence probability estimation CNN model

Sunny Bodiwala *
Department of Computer Engineering
Gujarat Technological University
Ahmedabad 382424
Gujarat
India

Nirali Nanavati
Department of Computer Engineering
Sarvajanik College of Engineering and Technology
Surat 395001
Gujarat
India

Abstract

Heart disease affects the majority of the world's population, with a high rate of morbidity and mortality. The prediction of heart disease at an early stage is viewed as one of the imperative issues in clinical trials. Since there is a huge amount of medical data and it keeps on increasing, it becomes very difficult to analyze and process such data. Hence, machine learning algorithms become an obvious choice to handle such data. This paper presents a classification framework for clinical decision-making that uses an optimized convolutional neural network (CNN). The pre-processed clinical dataset is used for the training and testing of the classifier. A publicly available UCI dataset is used to examine how the system performs. The statistical features are extracted at the beginning and then subsequent data minimization is carried out. Therein-after, the prediction of heart diseases is done with the help of CNN. The identification of optimal weights is done to enhance the performance of CNN which in turn gives accurate disease prediction results. In this paper, we propose a hybrid Particle Swarm Optimization (PSO) and Gray wolf Optimization (GWO) technique namely PS-GW technique to identify the optimal weight parameters. Finally, performance analysis is done and the experimental results are compared with the existing approaches which show improved classification accuracy over conventional optimization

* *E-mail:* **sunny.bodiwala@gmail.com**

techniques and some well-known classifiers such as k-nearest neighbour, decision tree, logistic regression, naive bayes and random forest.

Subject Classification: *68Txx, 11Kxx.*

Keywords: *Convolutional neural network, Heart disease prediction, Meta-heuristics, Medical diagnoses.*

1. Introduction

Individuals' lives have been significantly impacted by disease prediction and diagnosis models which has been viewed as a major concern, as early disease prediction is very important for individuals to lead an all around healthy and settled life [1, 2]. Prediction of disease consequently stays much significant for the health care category for managing better clinical consideration for patients [3, 4]. The new advancement of data mining has originated various diseases prediction frameworks [5, 6] to come out [7], which are utilized for various applications. Hence, data mining holds a significant job in disease prediction during medical management [8]. For disease prediction, huge analyses must be done [9, 10]. Convolutional neural networks (CNN) have also played a major role in image encoding with several deep learning method to examine various image captioning methods and to employ CNN to generate captions for images which can in turn give accurate prediction results[11]. Notwithstanding, by applying data mining techniques, the analysis count could be significantly decreased. This helps enhancing time and performance required for rigorous analysis of data [10]. Several data mining techniques like regression, clustering, association rule and classification techniques such as FeedForward Back Propagation (FBNN), Radial Basis Neural Network (RBNN), Random Forest (RF) and K-Nearest Neighbor (KNN) are used to classify various disease attributes in predicting diseases [12, 13, 14, 15].

Cardiovascular failure, also called congestive heart failure, occurs at the point when the heart can't pump adequate blood to satisfy the requirements of the human body [16]. Several risk factors for coronary illness involves respiratory failure, hypertension, stoutness, smoking, liquor fanatic, nutrient deficiency, substantial metal poisonings, rest apnea, getting dormant, and an irregular eating routine (along with unknown fats and salts) [17, 18]. Accordingly, the clinical experts perceive the harms that had happened in the heart of a patient and check if the blood is siphoned well in the patient's heart or not [19]. Such testing methods give enormous sequential data. Further, it is still a major thing to do precise analysis with such significant information, fundamentally in the early phases. Indeed, a strategy for cardiovascular failure prediction methods, which has a low error rate, is basically needed for clinical trials and treatments [20]. By looking at these datasets.Hence, a chance exists to provide early medicines, diagnoses and treatments for people who are plausible to get afflicted by heath illness and furthermore to provide them a better life ahead [21]. One way to deal with the accurate diagnosis problem and its treatments is the significant goal of a physical assessment [22]. Be that as it may, issue emerges while doing heart disease prediction with vast data, which still has research gaps. Despite the fact that the processed data contains a defined feature set for detecting cardiac issues, sufficient commitments have not been accounted for to change or concentrate highlights to recognize the class labels. A large number of researchers have focused mainly on records or data gathering where the features can be extracted [23, 24]. Consequently, now features in the records that are processed do not differentiate well between the class labels. This requires supervised learning procedure. However, training algorithm performance is dependant on supervised learning, which requires sufficient arrangement to deal with the non-linearities of the removed features. Motivated by hybrid algorithms, which have been effectively applied in many challenging applications [25] and a probabilistic Decision Support system (PDSS) to detect cardiovascular disease with AI techniques[26], this study strives to present a novel hybrid calculation [25]. Subsequently, an effective heart disease prediction technique is vital, and with its accurate predictions people would be benefited.

The main contributions of our proposed work are given below.

(1) A framework for classification of disease with formulation and implementation of improved training procedure of supervised learning and feature extraction.

(2) In this paper, we propose a novel hybrid optimization algorithms i.e. Particle Swarm Optimization (PSO) and Grey Wolf Optimization (GWO). The proposed techniques trains the CNN to acquire the knowledge of optimized data from Principle Component Analysis (PCA).

(3) Based on a robust feature selection technique using shallow convolutional layers, the introduced method improves heart disease prediction rates in imbalanced medical data.

The remainder of this paper is structured as follows. Section 2 presents the related work. Section 3 proposes the classification, feature extraction, and disease prediction system with the optimization process of feature dimensionality reduction. Further, Section 4 shows the experiment and results. Finally, conclusion and future work are given in Section 5.

2. Related Work

A method was introduced by the authors in [27] for extracting the important features for the treatment of heart diseases. The Differential Evolution (DE) model was used for selecting features for coronary illnesses. Among 10 existing DE algorithm principles, the seventh rule was taken in the introduced work for the corresponding examination. The results of the proposed methodology were considered and its accuracy was discovered to be outperforming several conventional approaches. Further, the prediction time of health disease was examined by the adopted approach.

The author in [26] presented a Probabilistic decision support system for medical diagnostics that employs Min-Max techniques for normalisation of data, following which the author used artificial intelligence methods to add Kullback-Leibler Divergence, which is required in application parameters. The author examined output in terms of F1-score, recall, accuracy, and precision using a range of computational intelligence methodologies.

A novel strategy utilizing data mining techniques for perceiving the important features have been setup in [2], and consequently the heart diseases prediction accuracy was improved. Prediction strategies were presented using an assorted combination of features, and several classification approaches, like Decision Tree, NN (Neural Network), SVM (support vector machine) and NB(Naive Bayes). Experimentation results have exhibited that the heart disease prediction was better by identifying major features and data mining models, as depicted by the proposed approach.

Authors in [1] have set up a methodology that concentrated on the features improvement and deletion of problems brought in by the prediction framework. The issues that happened attributable to underfitting and overfitting were removed by the introduced model, and consequently, it can give better results on training and testing data. Unsuitable features and inappropriate setup of the model regularly cause training data to overfit. For eliminating the improper features, χ^2 arithmetical models were set up alongside CNN. In addition, the examination of the proposed model uncovered that the proposed prediction scheme can be used by the medical staff for the accurate prediction of heart disease.

A novel approach is introduced in [28] that estimated the logical values of physiological information which was accumulated by the prior HF events identification. The given approach has two significant stages; the first one was pattern similarity analysis and the subsequent was a prediction strategy. The underlying framework consolidated the HWD (Haar Wavelet Decomposition) and it allowed to pick the minimum group of bases that attract the activities of fundamental time series. Hence, the accomplished results proposed the viability of the embraced model in managing the prior recognition of heart failure.

Authors in [3] have introduced a powerful system for heart failure prediction. The significant work here was to distinguish the heart failure through a CNN model. Specifically, one-hot encoding and word vectors were utilized for demonstrating the heart disease diagnoses by using the short and long term network with its key standards. Computations relying upon real dataset have uncovered the competent utility and efficacy of the set up model in heart failure prediction.

The approach proposed in [4] that gauges the "multi-fractal autonomic elements" that were acquired from the heartbeat signals in the non-stop time. Likewise, it was featured that the registered markers from the introduced model were moreover adequate for death prediction. Here, the multi-fractal investigation was converged inside a procedural structure contingent upon the pulses. Wavelet pioneers and coefficients, in particular, were evaluated for predicting the time until the next heartbeat using probability density functions.

A strategy proposed in [29] to manage the problems of restricted risk prediction availability schemes that represented both clinical and biomarkers boundaries. Likewise, an occurrence coronary disease prediction system was set up for handling the previously described problems. Here, in the given method, a multi-changed hazard element scheme was presented for the prediction of the underlying occurrence of hospitalization or heart casualty. Also, a risk parameter was extricated that have gathered the subjects into five risk classes that ranges between 5% & $\geq 20\%$.

Authors in [5] have proposed a method based on mortality for foreseeing the pace of death in heart failure patients. Proper death rates prediction has upgraded the clinical results and allotment of medical supplies, consequently preventing the insufficient treatment and over-treatment of patients having high and low mortality. This strategy has incorporated three goals for death rates prediction, (1) predicting death rates in hospitals, (2) one year and thirty days death prediction. The relative investigations results have revealed a superior death rates predictions utilizing the given model with higher accuracy.

An approach proposed in [30] for predicting significant issues during hospitalization named as multiple tasks deep neural network (MT-CNN). The introduced algorithm was dissected on a "Chinese PLA General Hospital" assembled dataset. Via completing renal dysfunction, the AUC (Area under the ROC Curve) of the used model was 0.94, which was an imperative improvement (p under 0.02) over conventional approaches. The experimental results exhibited that the presented MT-CNN model accomplished better prediction in heart failure patients than conventional methodologies.

Ali et al. [1] have used χ^2 mathematical approach and CNN is likewise ideally chosen through exhaustive search strategy. The intensity of the presented combination strategy χ^2 - NN was measured by assessing its outcomes over conventional CNN and NN models for heart disease prediction. The presented model accuracy was 93.34%. The accomplished results were better compared to existing procedures. The conclusion of the technique has proposed that the used analytic framework could be modified by experts to accurately predict heart disease.

The analysis of different classifiers, along with solo and ensemble classifier (EC) have been given in [31]. Besides, different assessment measures were considered for measuring the performance of given classifier models with standard datasets. Finally, statistical analysis was completed for assessing the distinctions in execution among the three classifiers.

In [32], a unique technique for cardiac disease prediction guided by adaptable multi-layer networks (AMLN) is described, which employs the hierarchical neighbourhood component-based-learning (HNCL) model. In these, HNCL component was first utilized for the interrelation learning among the HFR parameters and risk elements for developing a bunch of important features, assumed as main weight vector, that mirrors each risk element.From the analysis, the introduced procedure was discovered to have accomplished steady and better predictions accuracy with an improvement of around 11.2% compared to most common strategy.

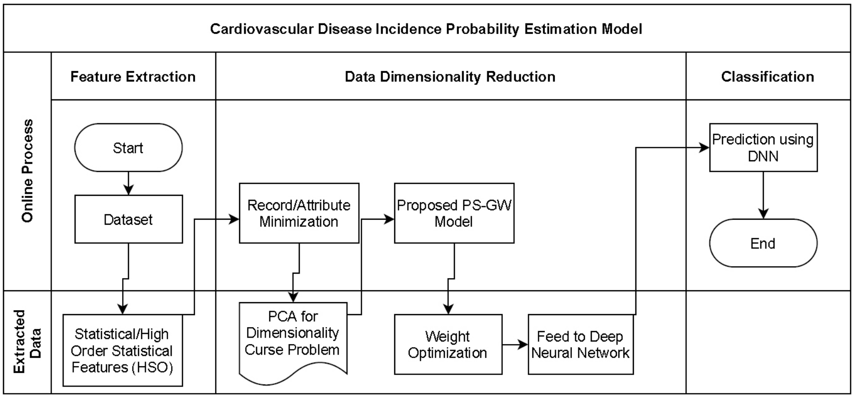


Figure 1

Diagrammatic depiction of the Cardiovascular Disease Incidence Probability Estimation Model

3. Proposed Cardiovascular Disease Incidence Probability Estimation Model

Figure 1 depicts the proposed model for predicting cardiac disease. Feature Extraction (FE), Classification, and Attributes Optimization are all included in the proposed model. At first, the input image is processed to get the higher order statistical (HOS) features and normal features. The whole feature extraction process is given in the next section. Now the raw data is available in huge quantity, the features extracted from such data will also be huge and taking care of those in an effective manner is also difficult. Consequently, it is necessary to reduce the dimensions of features through a standard methodology. In this paper, our focus is on attribute and record based shrinking measure. The detailed description of the process is stated in the next part. More significantly, the adopted technique tackles the issue of dimensionality curse.

Additionally, the shrunken features set are used for classification, and to perform such task a notable classifier such as CNN is utilized in this paper. The CNN model determines whether or not the patient has a coronary artery disease. The proposed approach includes weight optimization and subsequent prediction utilising a CNN model with other operations. The main contribution of the given approach is optimization of CNN weights which takes the proposed novel hybrid PS-GWO CNN algorithm and gives accurate prediction results.

3.1 *Feature Extracting Technique*

In the introduced work, the HSO and statistical features for every record, are settled in the process. The dataset used here incorporates 14 features like age, sex, resting blood pessure, chest pain location (4 types), fasting blood sugar > 120 mg/dl, serum cholesterol in mg/dl, resting electrocardiograph results (values 0, 1, 2), greatest pulse accomplished, work out induced angina, the slope of the peak exercise ST segment, oldpeak = ST exercise-induced ST depression, number of significant vessels (0–3) hued by fluorosopy & thal: 3 = typical; 6 = fixed deformity; 7 = reversable defect, for which the mean, median, standard deviation, minimum and maximum like statistical estimates are given. The minimum estimate is accomplished according to Eq. 1, although the greatest metrics is accomplished according to Eq. 2, in which NF_T represent the no. of features.

$$\text{Optimum Value} = \frac{\text{Smallest Position Features}}{NF_T} \tag{1}$$

$$\text{Optimum Value} = \frac{\text{Largest Position Features}}{NF_T} \tag{2}$$

Further, HSO features given are dependent on the likelihood of features occurrence. Think about the already existing features $HSOF_T$ = f_{t1}, f_{t2}, f_{t3} .... $f_{t_{NF}}$, if f_{t1} occurs for n_1 intervals in F_T, it is calculated by Eq. 3, in which *P* gives the likelihood of $f_t 1$ in F_t.

$$P = \frac{n_1}{NF_T} \tag{3}$$

Similarly, $P_2, P_3 .... P_N F_T$ is determined and the entropy function is given by Eq. 4. for all the probabilities.

$$\text{Cross Entropy} = -P_i \cdot logP_i \tag{4}$$

At this point, the statistical features measure like mean, median, standard deviation, smallest and largest optimum values are computed according to the previously mentioned equation for the accomplished entropy features. At last, all the extracted features are given below, where 5 denotes the quantity of features obtained from the given measures, NF_T denotes number of existing features, NF_E denotes number of entropy features. Hence, an aggregate of 36 such features are obtained from the

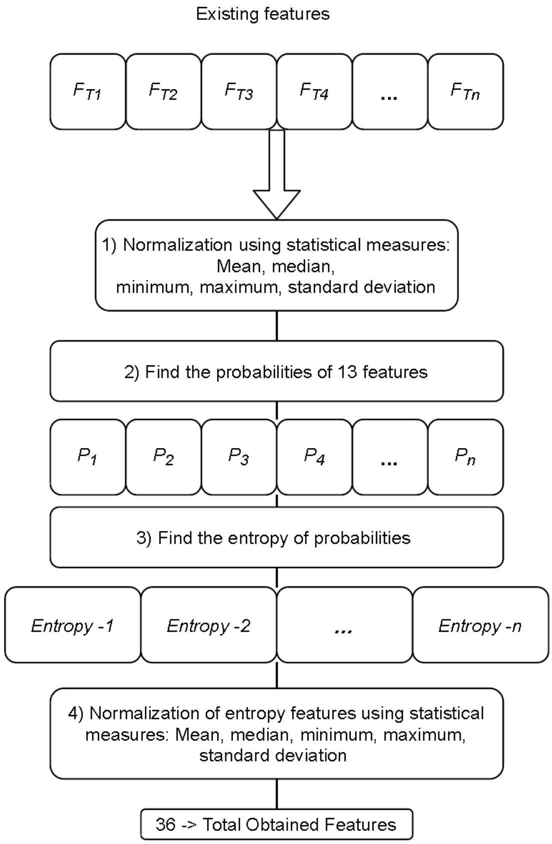


Figure 2

Features Extraction Process

whole process. The representation of the adopted method is shown in Figure 2.

$$\begin{aligned} 5 + NF_T + NF_E + 5 &= NF_T + NF_E + 10 \\ &= 13 + 13 + 10 \\ &= 36 \end{aligned}$$

3.2 *Features Shrinking Process*

To consider the appropriate features for further stages, it is essential to follow certain processes. In this paper, the features dimensionality is narrowed by data and attributes reduction.

3.2.1 Data Minimization

This method extracts the records of class with same records under two class labels i.e. 1 (disease exist)) and 2(absence of disease) and from the extracted class data, the relationship coefficients are given. By records minimization, alike records are combined and accordingly it makes it ready for more straightforward activity of the framework. Additionally, a fixed threshold of 0.97 is set, by which the comparative information could be obtained furthermore, the obtained coefficients are subjected to threshold. According to the comparative results if it is above the threshold, the similar records are then merged. By these the quantity of data can be reduced, whereas the attributes quantity still remains same. Further, for reducing the attributes volume, we adopt PCA technique [33], which satisfies the problem of dimensionality curse. The features dimensionality are just lowered using PCA. However, this doesn't necessarily loose important features. The whole process is shown in Figure 3. Finally, the

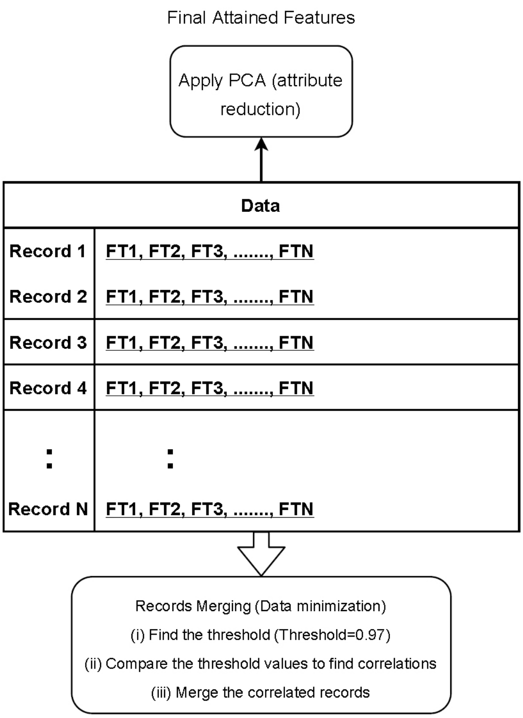


Figure 3

Data Minimization Process

classification of heart disease is done, utilizing DNN based predictions [34]. The F_T denotes the extracted features which are then feed to DNN architecture for disease prediction. The features can be given as in Eq.5.

$$F_T = \{F_{T1}, F_{T2}, F_{T3}, \ldots F_{TN}\} \tag{5}$$

3.2.2 Convolutional neural network

This section gives a basic idea of how the process is further carried out using CNN. CNNs are a type of deep neural network with spectral layers that can learn lower and higher level features. CNNs are useful models for predicting statistics, modelling, and other tasks. Simply put, it outperforms CNNs thanks to three additional concepts: local filters, weight sharing and max-pooling. Figure 4 shows the architecture of CNN and model summary is given in Table 1, which is used to predict cardiac disease. Convolutional layer is always followed by the pooling layer. Pooling is commonly used in the frequency domain. For the problem of variability, max-pooling produces good results. The maximum filter activation from different points within a specific window is collected by a max-pooling layer. It creates a low spatial variant of the convolution features in this stage. The max-pooling layer allows the model to tolerate tiny changes in the placements of the object's pieces, resulting in faster convergence. Fully linked layers, on the other hand, aggregate inputs from all points into a one dimensional feature space. The overall inputs are then classified using the softmax activation function layer.

Several different kernels are used to create the entire feature maps. The feature value at location (i, j) in the k^{th} feature map of the l^{th} layer is calculated mathematically. $z^l_{i,j,k}$ can be given as,

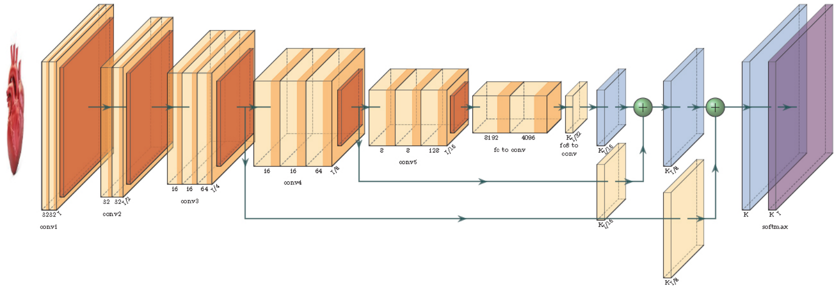


Figure 4

Convolutional neural network

$$z^l_{i,j,k} = w^{i^T}_k \times x^l_{i,j} + b^i_k \tag{6}$$

Where w^k_i and b^k_i are the weight and bias vectors of the l^{th} layer's k^{th} filter, respectively, and $k^l_{i,j}$ is the l^{th} layer's input patch centred at (i, j). The activation value $a^l_{i,j,k}$ of a convolutional feature can be calculated using the following formula:

$$a^l_{i,j,k} = a(z^l_{i,j,k}) \tag{7}$$

Sigmoid, tanh, and ReLU are examples of activation functions. By lowering the resolution of the feature maps, the pooling layer seeks to achieve shift-invariance. We have the following for each feature map $a^l_{i,j,k}$ if we refer to the pooling function as pool:

$$y^l_{i,j,k} = pool(a^l_{m,n,k}), \forall (m,n) \in R_{ij} \tag{8}$$

Where $R_i j$ is a small neighbourhood in the vicinity of the area (i, j). Average and maximum pooling are two common pooling operations. The output layer is the final layer of CNNs. The softmax operator is often used for classification problems . SVM, which may be used with CNN features to handle various classification tasks, is another widely utilised method.

Moreover, the difference between the predicted and actual values is determined by Eq. 9 and that is to be reduced. Now, O_g refers to the output neuron quantity, $\hat{O}_G$ and $\hat{O}_{\hat{G}}$ indicates the actual and the predicted output.

$$O_g = |\hat{O}_G - \hat{O}_{\hat{G}}| \leq \frac{1}{2L} \tag{9}$$

3.2.3 Maximum function and Entities Encoding

This paper intends to improve the overall accuracy of heart disease prediction, the CNN is trained utilizing the new PS-GW algorithm through the selection of the ideal weights (W) of CNN training. The proposed CNN model takes the optimal number of weights where the best prediction accuracy is tried to achieve. The solution as shown in Eq. 10 is chosen to get the maximum accuracy where A refers to the accuracy and x indicates the entities.

$$Max\,Function = Max(A, x) \tag{10}$$

Table 1

Detailed Model Architecture

Layers	Filters	Filters Size	Output Size
Input			32 × 32 × 32
Convolution Layer (Conv1)	32	2 × 2	32 × 32 × 32
Batch Normalization 1			32 × 32 × 32
Convolution Layer (Conv2)	32	2 × 2	32 × 32 × 32
Batch Normalization 2			32 × 32 × 32
Convolution Layer (Conv3)	32	2 × 2	32 × 32 × 32
Batch Normalization 3			32 × 32 × 32
Max Pooling 1		2 × 2	16 × 16 × 32
Convolution Layer (Conv4)	64	2 × 2	16 × 16 × 64
Batch Normalization 4			16 × 16 × 64
Max Pooling 2		2 × 2	8 × 8 × 64
Convolution Layer (Conv5)	128	2 × 2	8 × 8 × 128
Batch Normalization 5			8 × 8 × 128
Flatten			8192, None
Dense 1			128
Dense 2			2, None

3.3 *Grey Wolf Optimizer*

This paper focus on accurate heart disease prediction by training CNN model using optimized weight. Here, the combination of GWO and PSO concept is used to offer better prediction results. The grey wolves behavior has given the best way to converge the weight of the CNN. Upon individual implementation, the convergence results were not that effective, and thus we have decided to update the grey wolf position by the PSO algorithm's velocity update process. The PSO algorithm incorporates a basic idea, simple execution, robust operation, and it is computational effective when compared with most common heuristic optimization algorithms. Nonetheless, it is not difficult to get into local optima in higher-dimension space and doesn't go well in the recurring cycle. Hence, a hybrid approach is presented, that combines GWO and PSO, that makes the convergence better as individually PSO does not guarantee a better convergence rate.

The GWO was proposed by authors in [35], who depicted a metaheuristic algorithm which is based on swarm. The algorithm imitates the social leadership and hunting strategy adapted by the grey wolves. The mathematical aspects are utilized by researchers for describing the significant steps of the hunting strategy by the wolves, and utilized it for optimization issues. The GWO algorithm's mathematical model for grey wolf communal behaviour divides the population into four categories, such as alpha α, beta β, delta δ, & omega ω. The first three groups i.e. α, β, and δ has the fittest wolves and they guide the other wolves ω towards the ideal regions inside an search space. During the optimization, the wolves encircle their prey, and the equations for the same can be as shown below:

$$\vec{\mathrm{D}} = |\vec{\mathrm{C}}\vec{\mathrm{X}}p(t) - \vec{\mathrm{X}}(t)| \tag{11}$$

$$\vec{\mathrm{X}}(t+1) = |\vec{\mathrm{X}}p(t) - \vec{\mathrm{A}}\vec{\mathrm{D}}| \tag{12}$$

where $\vec{\mathrm{C}}$ and $\vec{\mathrm{A}}$ denotes the co-efficient vectors, $\vec{\mathrm{X}}p(t)$ is the position vector of the prey, and $\vec{\mathrm{X}}$ gives the grey wolf position vector. $\vec{\mathrm{C}}$ $\vec{\mathrm{A}}$ vectors are given as below.

$$\vec{\mathrm{A}} = 2\vec{\mathrm{a}}r_1 - \vec{\mathrm{a}} \tag{13}$$

$$\vec{\mathrm{C}} = 2\vec{\mathrm{a}}r_2 \tag{14}$$

where the $\vec{\mathrm{a}}$ constituents are sequentially reduced to 2 to 0 using few steps and r_1, r_2 are some vectors in the range $[0,1]$. Utilizing the above eqs, the grey wolf available in position (X, Y) can refresh self position dependent on the location of the prey (X^*, Y^*). Various position encompassing best agents could be attained concerning the current location after changing the $\vec{\mathrm{A}}$ and $\vec{\mathrm{C}}$ values of vector. It's worth noting that the arbitrary vectors $r1$ and $r2$ allow grey wolves to gain access to a point in the middle of these 2 vectors. Using the above equation, the wolf could then arbitrarily change its current position within the search space that includes the prey [35].

The algorithm considers *alpha*, *beta*, and *delta* as potential prey positions. While optimising the underlying three most effective solutions are considered individually to be *alpha*, *beta*, and *delta*. Following that, other *omega* wolves may shift their positions based on the positions of the *alpha*, *beta*, and *delta* wolves. The mathematical representation of the rearranging of the *omega* wolves' position is shown below [36]:

$$\vec{\mathrm{D}}\alpha = |\overrightarrow{\mathrm{C1}}\vec{\mathrm{X}}\alpha - \vec{\mathrm{X}}| \tag{15}$$

$$\vec{D}\beta = |\,\overrightarrow{C2}\vec{X}\beta - \vec{X}\,| \tag{16}$$

$$\vec{D}\delta = |\,\overrightarrow{C3}\vec{X}\delta - \vec{X}\,| \tag{17}$$

where $\vec{X}\alpha$ refers to the α position, $\vec{X}\beta$ refers to the β position, and $\vec{X}\delta$ represents the δ position. $\overrightarrow{C1}$, $\overrightarrow{C2}$, $\overrightarrow{C3}$ are the arbitrary vectors, and $\vec{X}$ shows position taken in this analysis.

Eqs. 15, 16, and 17, respectively. can be used to specify the step size of the *omega* wolf towards the *alpha*, *beta*, and *delta* wolves. Following is an approximation of the wolves' final placements based on the present solution once this distance has been determined:

$$\overrightarrow{X1} = \overrightarrow{X1}\alpha - \overrightarrow{A1}\overrightarrow{D\alpha} \tag{18}$$

$$\overrightarrow{X2} = \overrightarrow{X1}\beta - \overrightarrow{A2}\overrightarrow{D\beta} \tag{19}$$

$$\overrightarrow{X3} = \overrightarrow{X1}\delta - \overrightarrow{A3}\overrightarrow{D\delta} \tag{20}$$

$$\vec{X}(t+1) = \frac{\overrightarrow{X1} + \overrightarrow{X2} + \overrightarrow{X3}}{3} \tag{21}$$

where the position of α, β and δ is shown by $\vec{X}\alpha, \vec{X}\beta$ and $\vec{X}\delta$.

As given in Eqs. 18, 19, 20, and 21, the final position of the ω wolves is set. Also, $\vec{A}$ and $\vec{C}$ are arbitrary and adaptive, that assist in exploiting and exploring the grey wolf algorithm. The same has been simulated in target programming language. The arbitrary components impact is controlled by the velocity update of PSO . Now the adjustment is performed by combining the PSO velocity updating in GWO, the introduced model is named as PS-GW model. Algorithm 1 contains the pseudo code for the proposed PS-GW model. Figure 5 depicts the flowchart of the entire proposed model.

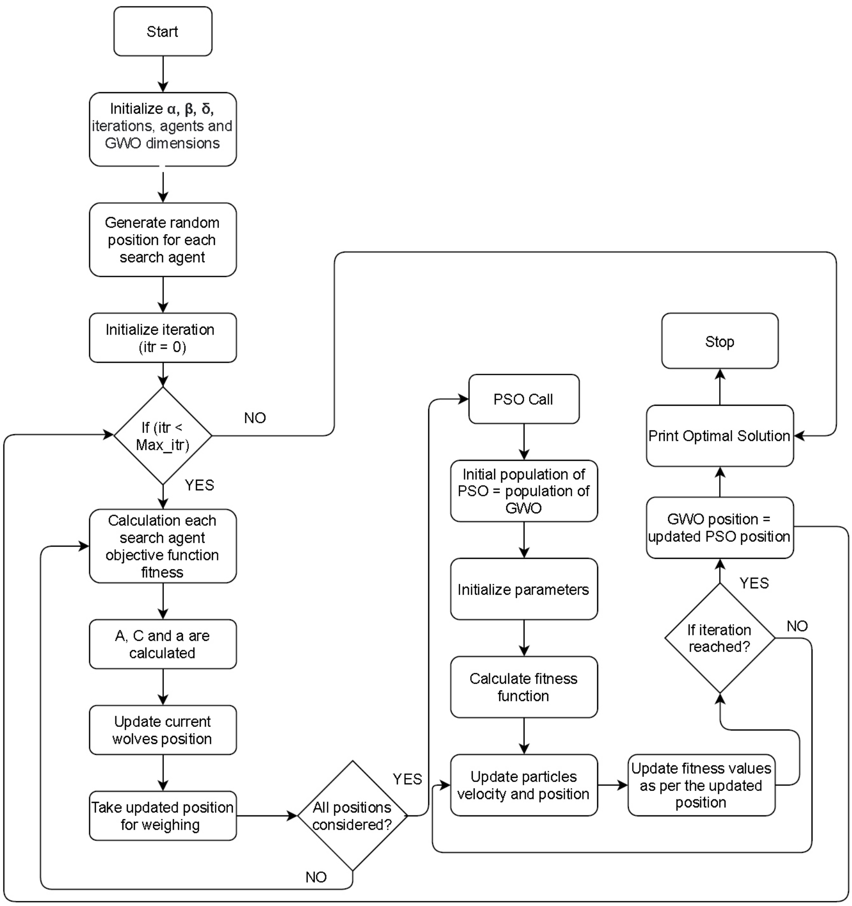


Figure 5

Flowchart of the proposed model

Algorithm 1 : PS-GW Model

Input : Bias Scaling Factor $s_{bias} \in R$

Input scaling factor $s_{dot} \in R$

Output: Input features coefficients $r_{bias}, r_{dot} \in R^m$

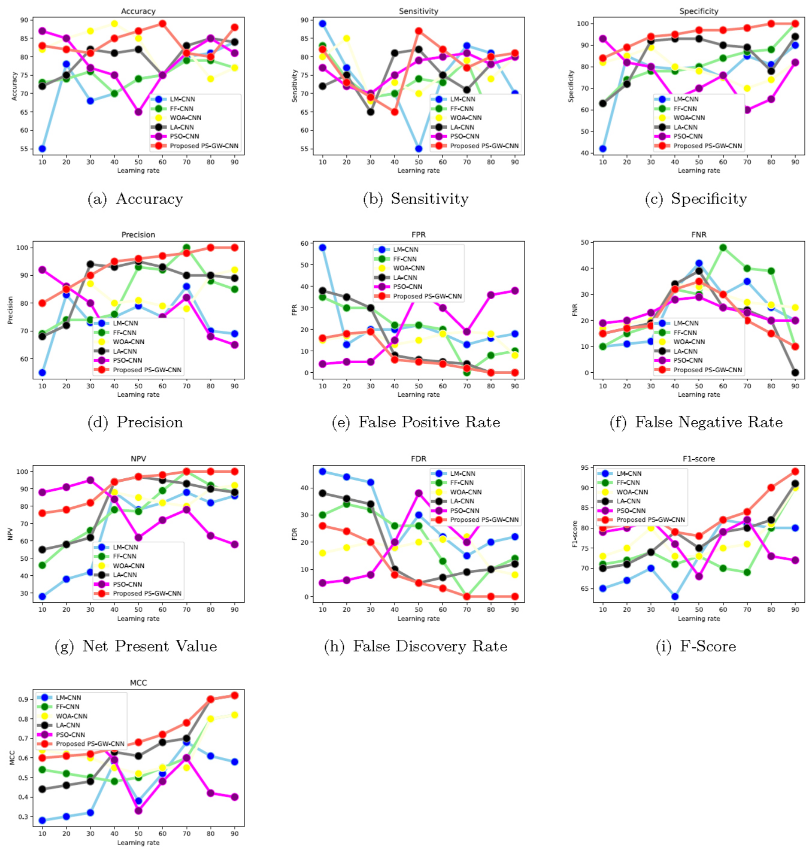
Output $\vec{X}_\alpha \in R$

1. Initialize grey wolf population $X_i (i = 1, 2, ..., n)$

2. Initialize $\vec{C}, \vec{A}$ and $\vec{d}$
3. Every agent or wolf fitness value is calculated
4. $\vec{X}_\alpha$ - First fine agent
5. $\vec{X}_\beta$ - Second fine agent
6. $\vec{X}_\delta$ - Third fine agent
7. While $(t_i > max.\ no.\ of\ iterations)$
8. **for** *each search agent* **do**
9. change the current search agent position
10. **end**
11. Adjust A, α and C
12. Compute all search agent fitness
13. Perform agent computation, i.e. PSO velocity measure is combined with updated agent fitness value
14. Modify $\vec{X}_\alpha$, $\vec{X}_\beta$ and $\vec{X}_\delta$
15. $t_i = t_i + 1$
16. end While
17. return $\vec{X}_\alpha$

4. Experiments and Results

The proposed PS-GW model implementation for predicting heart disease was carried out in python with Google Colab and the comparable outputs were achieved. The dataset used for the experiments was "https://archive.ics.uci.edu/ml/datasets/heart+disease". The dataset has Multivariate characteristics and 303 number of Instances. The attributes characteristics are Integer, Real and Categorical. The number of main attributes are 14 which incorporates age, sex, resting blood pressure, chest pain location (four types), fasting blood sugar > 120 mg/dl, serum cholesterol in mg/dl, resting electrocardiograph results (values 0, 1, 2), greatest pulse accomplished, work out induced angina, the slope of the peak exercise ST segment, oldpeak = ST depression induced by exercise relative to rest, number of significant vessels (0–3) hued by flourosopy and thal: 3 = typical; 6 = fixed deformity; 7 = reversable deformity. Subsequently, the associated tasks are classification. Further, the absent heart disease quantity is 164, while, the quantity of information with the presence of coronary illness is 139. The insights about the presence & absence of disease is also given in the dataset. The proposed approach is compared with existing techniques like Levenberg–Marquardt (LM-CNN), Whale Optimization Algorithm (WOA-CNN), Firefly (FF-CNN), Particle Swarm Optimization (PSO-CNN) and Lion Algorithm (LA-CNN) with performance measures like F Score, FPR (False Positive Rate), FDR (False Discovery Rate), FNR (False Negative Rate), MCC (Matthews Correlation Coefficient), sensitivity, precision, specificity and accuracy. For fair



(a) Accuracy (b) Sensitivity (c) Specificity

(d) Precision (e) False Positive Rate (f) False Negative Rate

(g) Net Present Value (h) False Discovery Rate (i) F-Score

(j) Matthews Correlation Coefficient

Figure 6

Performance Analysis: Proposed VS. Existing Models

comparison outcomes, we try the technique by changing the learning rates, that refers to the number of inputs utilized for training the CNN model. The information considered for training the model are given by learning rates. Expansion in learning rate enhances the accuracy percentage and henceforth, the learning rates are traversed from 10%, 20%, 30%, 40%, 50 60%, 70%, 80% and 90%. In this paper for demonstrating the improvement of the proposed model, 70% learning rate represents 70% heart patient records & ordinary records, thus the training data contains {164 * 0.3 = 115 and 139 * 0.7 = 97} i.e. 212 data combined. Finally, the feature analysis was additionally done to show the presented work performance.

4.1 *Performance of Proposed Model*

The proposed PS-GW CNN model performance analysis is shown in Figure 6 for different learning rates. For all the measures, overall better results are achieved by the given model. The test is performed by changing the performance metrics and learning rate of models. Both the positive and negative measures are considered. The positive matrices includes F score, MCC, NPV, precision, specificity, sensitivity and accuracy. Additionally, the negative metrics incorporates FNR, FPR & NPV. Here, the performance distinction over different approaches are determined utilizing equation $\frac{proposed\ model - existing\ model}{existing\ model} \times 100.$

Moreover, from Figure 6(a), the presented PS-GW CNN model accuracy is more at 90^{th} learning rate, that is 9%, 16.5%, 16.5%, and 9.6% better compared to LM-CNN, FF-CNN, WOA-CNN and PSO-CNN algorithms. Now, Sensitivity as shown in Figure 6(b), the presented strategy is 11.2% better compared to LA-CNN and LM-CNN algorithm at 90^{th} learning rate. At the 90^{th} learning rate, the sensitivity proportion of lion CNN in Figure 6(b) is discovered 10.5 percent higher than the proposed approach. hence, the wide range of various metrics like precision, accuracy, etc are discovered to be better at a higher rate for the proposed model. Subsequently, these minor things can be neglected.

Table 2

Performance Analysis: Proposed Model vs. Existing Models

Metrics	Optimized and Sampled CNN					
	WOA	**LM**	**FF**	**LA**	**PSO**	**Proposed**
PS-GW						
Accuracy	0.77	0.84	0.77	0.84	0.81	0.93
Specificity	0.82	0.9	1	0.94	0.82	1.0
Sensitivity	0.8	0.7	0.6	0.8	0.8	0.8
Precision	0.79	0.86	1	0.92	0.8	1.0
False Positive Rate	0.2	0.2	0	0.06	0.2	0
Net Present Value	0.8	0.9	1	0.94	0.82	1.0
False Negative Rate	0.23	0.2	0.47	0.3	0.2	0.27
False Discovery Rate	0.22	0.15	0	0.084	0.2	0
MCC	0.51	0.7	0.61	0.69	0.62	0.77
F-score	0.76	0.83	0.7	0.82	0.89	0.85

Table 3

Proposed output and targets with the LM and FF Techniques

Target	Train PS-GW	Difference	Train LM	Difference	Train FF	Difference
0	0.21658	0.21658	-0.05432	-0.05432	-0.04269	-0.04269
1	0.91254	-0.07944	0.97847	-0.03126	0.94522	-0.04026
1	0.93122	-0.04558	0.97543	-0.02946	0.97638	-0.04152
0	-0.11044	-0.11044	-0.15851	-0.15519	-0.16911	-0.16911
0	0.010297	-0.010297	-0.12092	-0.12092	0.13041	-0.13041
1	**0.12942**	**-0.89716**	**-0.0089**	**-0.9801**	**0.8902**	**-0.8902**
0	0.26678	0.26678	0.18846	0.18864	0.20513	0.20513
0	0.23867	0.23867	0.45656	0.45656	0.55236	0.55236
1	0.95216	-0.08493	1.0007	0.0007	1.0009	0.0009
0	-0.16881	-0.16881	-0.43445	-0.43454	-0.52674	-0.52674
0	0.13818	-0.13818	-0.09119	-0.09281	-0.07271	-0.07271
1	**-0.02745**	**-1.02745**	**0.0438**	**-0.9761**	**-0.0438**	**-0.09925**
1	1.3120	0.3120	1.081	0.018	1.021	-0.0721
1	0.72025	-0.49484	1.0859	0.0886	1.0952	0.0952
0	0.30261	0.30216	0.027146	0.037163	0.048246	0.048246

Contd...

1	0.86424	-0.13758	0.62893	-0.58621	0.73951	-0.69412
1	0.82678	-0.04804	0.8731	-0.2278	0.8721	-0.3268
1	1.0687	0.0687	1.2138	0.1373	1.4283	0.4283
0	0.21206	0.21206	0.1843	0.1844	0.1956	0.1956
1	1.059	0.059	0.04619	-0.07073	0.98861	-0.07052
1	0.98695	-0.00341	1.0209	0.0207	1.8107	0.8107
1	1.2437	0.2437	0.98691	-0.03231	0.996741	-0.03359
1	0.98793	-0.01216	0.97034	-0.04847	0.99077	-0.03147
1	0.9314	-0.1895	0.78612	-0.24797	0.87031	-0.32856

Table 4

Performance of the PS-GW method

Epoch/100	Gradient	Mean Square Error
0	4131.23	31.861
2	97.2469	0.655
4	20.7281	0.0127
6	0.0071	3.0531e-007
8	4.96048e-021	1.2815e-031

Table 5

Feature Analysis: Proposed Model vs. Existing Models

Metrics	Existing feature + Dimensionality Reduction	Proposed Feature + Dimensionality Reduction	Non Dimensionality Reduction	Dimensionality Reduction
Accuracy	0.67	0.88	0.8	0.9
Specificity	0.6	1.0	0.76	1.0
Sensitivity	1	0.8	0.7	0.8
Precision	0.6	1.0	0.71	1.0
False Postive Rate	0.6	0	0.3	0
Net Present Value	0.5	1.0	0.8	1.0
False Negative Rate	0	0.3	0.31	0.27
False Disvovery Rate	0.5	0	0.31	0
MCC	0.6	0.8	0.51	0.78
F-score	0.7	0.85	0.7	0.88

Table 6

Performance Analysis: Proposed VS. Existing Classifiers

Metrics	K-NN	Decision Tree	Logistic Regression	Naive Bayes	Random Forest	Our CNN
Accuracy	0.68	0.81	0.852	0.852	0.86	0.93
Specificity	0.71	0.78	0.91	0.83	0.76	1
Sensitivity	0.29	0.4	0.41	0.87	0.73	0.88
Precision	0.66	0.66	0.76	0.78	1	1
False Positive Rate	0.32	0.4	0.21	0.6	0.41	0.11
Net Present Value	0.8	0.9	0.79	0.85	0.9	1
False Negative Rate	32.35	17.64	11.76	8.52	8.52	0.27
False Discovery Rate	0.23	0.21	0	0.21	0.075	0
MCC	-0.3	-0.8	-0.3	0.71	0.6	0.64
F-Score	0.69	0.83	0.86	0.57	0.95	0.90

Table 7

Proposed Classifier Analysis: Existing Work vs. Adopted Work

Model	Classifier	Accuracy
J.Tele'19 [2]	Data Mining Techniques	86
AISC'19 [37]	Fuzzy Rules	88
RAM'20 [38]	Memetic Classifier	95.24
DMAI'20 [39]	RBF-TSVM	97
ICTCS'19 [40]	Hybrid-DT	78
JCS'17 [41]	Rule Based Fuzzy Logic	78
EvolInt'20 [42]	CDTL-RF	89.30
AIHC'2020 [43]	T-CNN	85.4
CC'19 [44]	GA-SVM	88.34
Proposed	PS-GW CNN	93

The specificity of the presented PS-GW CNN model at 90^{th} learning rate is 22%, 11.2%, 11.2%, 11.2%, and 28.9% better than LM-CNN, FF-

CNN, WOA-CNN, LA-NN, and PSO-CNN models is shown in Figure 6(c). Additionally, from Figure 6(d), the precision of the proposed model at 90^{th} learning rate is high at 25.2%, 11.2%, 11.2%, 51.6%, and 11.2% than LM-CNN, FF-CNN, WOA-CNN LA-CNN, and PSO-CNN algorithms. The FPR ((False Positive Rate)) that is accomplished from Figure 6(e) at 90^{th} learning rate is 10%, 9.89%, 10%, 10%, & 36% better compared to LM-CNN, FF-CNN, WOA-CNN, LA-CNN and PSO-CNN algorithms. As shown in Figure 6(e), FNR (False Negative Rate) of embraced PS-GW CNN plot at 90^{th} learning rate is half better compared to LM-CNN, and PSO-CNN models. In expansion, Figure 6(g) shows, the used approach for NPV (Net Present Value) at 90^{th} learning rate is 22%, 11.2%, 11.2%, 11.2%, and 28.9% better than LM-CNN, FF-CNN, WOA-CNN, LA-CNN, and PSO-CNN models. Likewise, FDR, F-score and MCC are also compared and shown. Consequently the viability of the proposed PS-GW CNN model for accurate prediction of heart disease has been demonstrated by the achieved outcomes. Table 2 shows proposed PS-GW CNN approach overall performance at 90^{th} learning rate for cardiac disease prediction.

Table 3 compares a sample of the network's output to the desired output. We can see that the output of the network is not confined to 0 and 1 merely because we employed the linear transfer function in the output layer. It is possible to assess the severity of the sickness in this manner. A patient with a network output of -0.79, for example, is unlikely to have cardiac disease. A network output of 1.49, on the other hand, confirms the presence of the disease and can be regarded serious. In addition, examples where the doctor can make mistake are shown in bold.

For each strategy, the weight computing technique change has been evaluated in order to reduce the mean square error (MSE) between the actual output and the network output. The training began with a 26.461 MSE and ended after 22 epoch with a minimum of 0.043631 when the robust backpropagation approach was applied to our data. Table 4 shows that the training for the PS-GW algorithm began with a mean square error of 33.861 and ended with a minimum of $1.2815\times10^{-0.3}$ after 8 epochs. We can consider the results to be reliable at these levels of error.

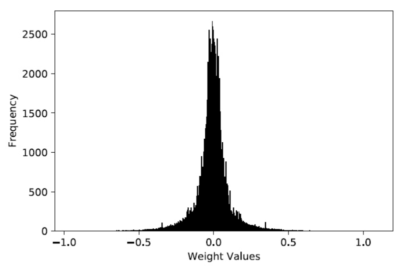
4.2 *Features Analysis*

The significance of feature extraction is given in Table 5 for 70^{th} learning rate. After observation, better performance of proposed scheme is shown as compared to existing schemes. The positive measures such as F-score, MCC, NPV, precision, specificity, and accuracy are assessed. The accuracy intends to be the precise model prediction and subsequently it is important to get high precision than ordinary features. After reducing the dimensionality, the proposed strategy's F-score, specificity, accuracy, precision, MCC and NPV are 31.7%, 27%, 53.2% 40%, 40% and 50% more than the current features. A different proportion of proposed and existing algorithms are analysed. In this manner the advancement of the adopted strategy regarding features analysis has been affirmed at 70^{th} learning rate.

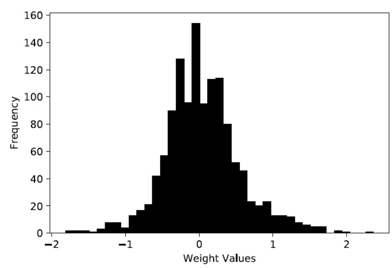
The features of the introduced model with and without reduction of dimension is exhibited by Table 5 at 70^{th} learning rate. Further analysis shows superior achieved performance of the proposed model over the existing approaches by reducing features dimensions. The presented model accuracy with reduced dimensions is 19.6% better than the introduced model without shrinking dimensions. Additionally, the proposed reduced feature dimension model sensitivity is 05.7% better compared to the Non reduction model. In addition, the specificity of the presented approach with reduced dimension is 33.1% better compared to the introduced model with non features dimension reduction. Accordingly, the introduced strategy effectiveness with reduced features dimension is demonstrated for the 70^{th} learning rate.

4.3 *Classifiers Analysis*

The performance of proposed CNN classifier is compared with existing classifiers such as K-NN, Decision Tree, Logistic Regression, Naive Bayes, and Random Forests as shown in Table 6 and the performance of the used CNN system is demonstrated. Through analysis, the accuracy of the proposed work is 25% better compared to K-NN, 12% better



(a)



(b)

Figure 7

Hidden and Output Layer Histogram

compared to decision tree, 7% better compared to logistic regression and so on. In this way, the performance of the proposed classifier is shown viably from the experiments. In addition, the accuracy of the proposed approach compared to existing work is given in Table 7 which shows better classification accuracy than many of the models. Figure 7 shows the proposed CNN weight analysis histogram in hidden layer and output layer.

5. Conclusion

A cardiovascular disease incidence probability estimation model has been introduced in this paper that included stages like feature extraction, data reduction, and classification. At first, both statistical and HSO features were extracted and afterward, the records minimization work was carried out. Here, for settling the dimensionality curse problem, PCA was used. Thereafter, the dimension reduced features were given to CNN, by which the process of disease prediction starts. The main goal of the introduced work is higher prediction accuracy and thus, optimization of CNN weight was planned. Subsequently, the PS-GW model is deployed for optimization, which combines the concept of PS and GWO. Finally, the performance of the presented approach was assessed over other existing methods and the results were obtained. Experimental results shows the introduced PS-GW CNN model accuracy was 9%, 16.5%, 16.5%, 10.1% and 9.6% better compared to LM-CNN, WOA-CNN, FF-CNN, PSO-CNN, and LA-CNN approaches. Likewise, other measures are also given with detailed comparative analysis. In comparison to other classifiers and state-of-the-art algorithms, the PS-GW CNN offers the best result for the UCI dataset in all performance measures. Additionally, the features extraction and dimension reduction significance was assessed, by analyzing the proposed and existing models with different performance metrics. Hence, the performance of the proposed technique was demonstrated. Future work includes the use of different classifiers and hybrid approaches for better heart disease prediction.

References

[1] L. Ali, A. Rahman, A. Khan, M. Zhou, A. Javeed, and J. A. Khan, "An automated diagnostic system for heart disease prediction based on χ^2 statistical model and optimally configured deep neural network," *IEEE Access*, vol. 7, pp. 34 938–34 945, 2019.

[2] M. S. Amin, Y. K. Chiam, and K. D. Varathan, “Identification of significant features and data mining techniques in predicting heart disease,” *Telematics and Informatics*, vol. 36, pp. 82–93, 2019.

[3] B. Jin, C. Che, Z. Liu, S. Zhang, X. Yin, and X. Wei, “Predicting the risk of heart failure with ehr sequential data modeling,“ *IEEE Access*, vol. 6, pp. 9256–9261, 2018.

[4] G. Valenza, H. Wendt, K. Kiyono, J. Hayano, E. Watanabe, Y. Yamamoto, P. Abry, and R. Barbieri, “Mortality prediction in severe congestive heart failure patients with multifractal point-process modeling of heartbeat dynamics,” *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 10, pp. 2345–2354, 2018.

[5] Z.Wang, L. Yao, D. Li, T. Ruan, M. Liu, and J. Gao, “Mortality prediction system for heart failure with orthogonal relief and dynamic radius means,” *International Journal of Medical Informatics*, vol. 115, pp. 10–17, 2018.

[6] R. Ganesan and V. V. Chamundeeswari, “Composite algorithm for pervasive healthcare system{a solution to find optimized route for closest available health care facilities,” *Multimedia Tools and Applications*, vol. 79, no. 7, pp. 5125–5148, 2020.

[7] S. Parisot, S. I. Ktena, E. Ferrante, M. Lee, R. Guerrero, B. Glocker, and D. Rueckert, “Disease prediction using graph convolutional networks: Application to autism spectrum disorder and alzheimer’s disease,” *Medical Image Analysis*, vol. 48, pp. 117–130, 2018.

[8] B. Aramini, P. Geraghty, D. J. Lederer, J. Costa, S. L. DiAngelo, J. Floros, and F. D’Ovidio, “Surfactant protein a and d polymorphisms and methylprednisolone pharmacogenetics in donor lungs,”*The Journal of Thoracic and Cardiovascular Surgery*, vol. 157, no. 5, pp. 2109–2117, 2019.

[9] C. Zhang, L. Zhu, C. Xu, and R. Lu, “Ppdp: An efficient and privacy-preserving disease prediction scheme in cloud-based e-healthcare system,” *Future Generation Computer Systems*, vol. 79, pp. 16–25, 2018.

[10] X. Chen, Y.-W. Niu, G.-H. Wang, and G.-Y. Yan, “Hamda: Hybrid approach for mirna-disease association prediction,” *Journal of Biomedical Informatics*, vol. 76, pp. 50–58, 2017.

[11] S. Kalra and A. Leekha, “Survey of convolutional neural networks for image captioning,” *Journal of Information and Optimization Sciences*, vol. 41, no. 1, pp. 239–260, 2020. [Online]. Available: https://doi.org /10.1080/02522667.2020.1715602

[12] J. Rodríguez, S. Prieto, and L. J. Ramírez López, "A novel heart rate attractor for the prediction of cardiovascular disease," *Informatics in Medicine Unlocked*, vol. 15, p. 100174, 2019.

[13] V. J. Baggen, E. Venema, R. Živná, A. E. van den Bosch, J. A. Eindhoven, M. Witsenburg, J. A. Cuypers, E. Boersma, H. Lingsma, J. R. Popelová, and J. W. Roos-Hesselink, "Development and validation of a risk prediction model in patients with adult congenital heart disease," *International Journal of Cardiology*, vol. 276, pp. 87–92, 2019.

[14] J. Nahar, T. Imam, K. S. Tickle, and Y.-P. P. Chen, "Computational intelligence for heart disease diagnosis: A medical knowledge driven approach," *Expert Systems with Applications*, vol. 40, no. 1, pp. 96–104, 2013.

[15] T. Roshini, R. V. Ravi, A. Reema Mathew, A. B. Kadan, and P. S. Subbian, "Automatic diagnosis of diabetic retinopathy with the aid of adaptive average filtering with optimized deep convolutional neural network," *International Journal of Imaging Systems and Technology*, vol. 30, no. 4, pp. 1173–1193, 2020.

[16] T. Honda, D. Yoshida, J. Hata, Y. Hirakawa, Y. Ishida, M. Shibata, S. Sakata, T. Kitazono, and T. Ninomiya, "Development and validation of modified risk prediction models for cardiovascular disease and its subtypes: The hisayama study," *Atherosclerosis*, vol. 279, pp. 38–44, 2018.

[17] Y. Zhang, B. E. Schroeder, P.-L. Jerevall, A. Ly, H. Nolan, C. A. Schnabel, and D. C. Sgroi, "A novel breast cancer index for prediction of distant recurrence in hr+ early-stage breast cancer with one to three positive nodes," *Clinical Cancer Research*, vol. 23, no. 23, pp. 7217–7224, 2017.

[18] A. Menotti and P. E. Puddu, "Lifetime prediction of coronary heart disease and heart disease of uncertain etiology in a 50-year follow-up population study," *International journal of cardiology*, vol. 196, pp. 55–60, 2015.

[19] R. J. P. Princy, S. Parthasarathy, P. S. H. Jose, A. R. Lakshminarayanan, and S. Jeganathan, "Prediction of cardiac disease using supervised machine learning algorithms," in 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS). IEEE, 2020, pp. 570–575.

[20] G. T. Reddy, M. P. K. Reddy, K. Lakshmanna, D. S. Rajput, R. Kaluri, and G. Srivastava, "Hybrid genetic algorithm and a fuzzy logic

classifier for heart disease diagnosis," *Evolutionary Intelligence*, vol. 13, no. 2, pp. 185–196, 2020.

[21] A. Saeed, V. Nambi, W. Sun, S. S. Virani, G. E. Taffet, A. Deswal, E. Selvin, K. Matsushita, L. E. Wagenknecht, R. Hoogeveen, J. Coresh, J. A. de Lemos, and C. M. Ballantyne, "Short-term global cardiovascular disease risk prediction in older adults," *Journal of the American College of Cardiology*, vol. 71, no. 22, pp. 2527–2536, 2018, SPECIAL FOCUS ISSUE: CARDIOVASCULAR HEALTH PROMOTION.

[22] A. Z. Hameed, B. Ramasamy, M. A. Shahzad, and A. A. S. Bakhsh, "Efficient hybrid algorithm based on genetic with weighted fuzzy rule for developing a decision support system in prediction of heart diseases," *The Journal of Supercomputing*, pp. 1–21, 2021.

[23] S. A. Rizwan, A. Jalal, and K. Kim, "An accurate facial expression detector using multi-landmarks selection and local transform features," in 2020 3rd International Conference on Advancements in Computational Sciences (ICACS). IEEE, 2020, pp. 1–6.

[24] M. T. Nguyen, P. Siritanawan, and K. Kotani, "Saliency detection in human crowd images of different density levels using attention mechanism," *Signal Processing: Image Communication*, vol. 88, p. 115976, 2020.

[25] A. N. Jadhav and N. Gomathi, "Digwo: Hybridization of dragonfly algorithm with improved grey wolf optimization algorithm for data clustering," *Multimed Res*, vol. 2, no. 3, pp. 1–11, 2019.

[26] Smita and E. Kumar, "Probabilistic decision support system using machine learning techniques : A case study of cardiovascular diseases," *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 24, no. 5, pp. 1487–1496, 2021. [Online]. Available: https://doi.org/10.1080/09720529.2021.1947452

[27] T. Vivekanandan and N. C. Sriman Narayana Iyengar, "Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease," *Computers in Biology and Medicine*, vol. 90, pp. 125–136, 2017.

[28] J. Henriques, P. Carvalho, S. Paredes, T. Rocha, J. Habetha, M. Antunes, and J. Morais, "Prediction of heart failure decompensation events by trend analysis of telemonitoring data," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 5, pp. 1757–1769, 2015.

[29] A. Driscoll, E. H. Barnes, S. Blankenberg, D. M. Colquhoun, D. Hunt, P. J. Nestel, R. A. Stewart, M. J. West, H. D. White, J. Simes, and A.

Tonkin, "Predictors of incident heart failure in patients after an acute coronary syndrome: The lipid heart failure risk-prediction model," *International Journal of Cardiology*, vol. 248, pp. 361–368, 2017.

[30] B. Wang, Y. Bai, Z. Yao, J. Li, W. Dong, Y. Tu, W. Xue, Y. Tian, Y. Wang, and K. He, "A multi-task neural network architecture for renal dysfunction prediction in heart failure patients with electronic health records," *IEEE Access*, vol. 7, pp. 178 392–178 400, 2019.

[31] C.-H. Weng, T. C.-K. Huang, and R.-P. Han, "Disease prediction with different types of neural network classiffiers," *Telematics and Informatics*, vol. 33, no. 2, pp. 277–292, 2016.

[32] O. W. Samuel, B. Yang, Y. Geng, M. G. Asogbon, S. Pirbhulal, D. Mzurikwao, O. P. Idowu, T. J. Ogundele, X. Li, S. Chen, G. R. Naik, P. Fang, F. Han, and G. Li, "A new technique for the prediction of heart failure risk driven by hierarchical neighborhood component-based learning and adaptive multi-layer networks," *Future Generation Computer Systems*, vol. 110, pp. 781–794, 2020.

[33] Y. Jin, C. Qiu, L. Sun, X. Peng, and J. Zhou, "Anomaly detection in time series via robust pca," in *2017 2nd IEEE International Conference on Intelligent Transportation Engineering (ICITE)*, 2017, pp. 352–355.

[34] Y. Mohan, S. S. Chee, D. K. P. Xin, and L. P. Foong, "Artificial neural network for classiffication of depressive and normal in eeg," in *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 2016, pp. 286–290.

[35] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.

[36] E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary grey wolf optimization approaches for feature selection," *Neurocomputing*, vol. 172, pp. 371–381, 2016.

[37] R. J. P. Princy, S. Parthasarathy, P. S. Hency Jose, A. Raj Lakshminarayanan, and S. Jeganathan, "Prediction of cardiac disease using supervised machine learning algorithms," in 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), 2020, pp. 570–575.

[38] S. P. Siddique Ibrahim and M. Sivabalakrishnan, *An Evolutionary Memetic Weighted Associative Classiffication Algorithm for Heart Disease Prediction*. Singapore: Springer Singapore, 2020, pp. 183–199. [Online]. Available: https://doi.org/10.1007/978-981-15-1362-69

[39] M. Thiyagaraj and G. Suseendran, "Enhanced prediction of heart disease using particle swarm optimization and rough sets with transductive support vector machines classiffier," in *Data Management, Analytics and Innovation*, N. Sharma, A. Chakrabarti, and V. E. Balas, Eds. Singapore: Springer Singapore, 2020, pp. 141–152.

[40] S. Maji and S. Arora, "Decision tree algorithms for prediction of heart disease," in *Information and Communication Technology for Competitive Strategies*, S. Fong, S. Akashe, and P. N. Mahalle, Eds. Singapore: Springer Singapore, 2019, pp. 447–454.

[41] G. T. Reddy and N. Khare, "An efficient system for heart disease prediction using hybrid ofbat with rule-based fuzzy logic model," *Journal of Circuits, Systems and Computers,* vol. 26, no. 04, p. 1750061, 2017.

[42] M. Gopu and P. Swarnalatha, "Optimal feature selection through a cluster-based dt learning (cdtl) in heart disease prediction," *Evolutionary Intelligence*, vol. 14, 06 2021.

[43] S. Sandhiya and U. Palani, "An effective disease prediction system using incremental feature selection and temporal convolutional neural network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, pp. 5547–5560, 2020.

[44] C. Gokulnath and S. Shantharajah, "An optimized feature selection based on genetic approach and support vector machine for heart disease," *Cluster Computing,* pp. 1–11, 2019.

Received November, 2021

Revised May, 2022