SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V10I5P115

Volume 10 Issue 5, 165-173, May 2023

Original Article

Heart Disease Prediction using Jellyfish Search Optimization and Deep Residual Networks

Rohan Shinde[1], V. Lavanya[2]

[1, 2]*Department of Networking and Communication, SRM Institute of Science and Technology, Tamil Nadu, India*

[1]*Corresponding Author : shinderohan115@gmail.com*

Received: 12 March 2023 Revised: 22 April 2023 Accepted: 16 May 2023 Published: 31 May 2023

Abstract *- Early heart Disease prediction can reduce heart illness, which is a leading source of worldwide death. Unhealthy eating, mental stress, genetics, a sedentary lifestyle, and other factors have all contributed to the development of heart disease. Angina pectoris, dilated cardiomyopathy, stroke, and heart disease are most commonly caused by congestive heart failure. Moreover, a precise heart disease prognosis is necessary for effective cardiac treatment. In clinical machine learning, it is dangerously challenging to predict heart disease. This paper proposes a method based on Jellyfish Search Optimization (JSO) and enabled Deep Residual Networks (DRN) to improve heart disease prediction. The input dataset is subjected to the pre-processing step, which utilizes missing data imputation and Z- score Normalization. Consequently, the pre-processed output is fed to the feature selection stage, wherein features are selected by Kumar- Hassebrook similarity. At last, in the heart disease diagnosis phase, where DRN makes the heart disease identification, DRN is trained using the proposed JSO. Moreover, the proposed DRN_JSO has effectively delivered better performance parameters with high accuracy of 84.65% F1_score of 83. 65%, Matthews's correlation coefficient (MCC) of 85.65% and a True Positive Rate (TPR) of 82.65%.*

Keywords *- Z- score normalization, Missing data imputation, Kumar- hassebrook similarity, Deep Residual Network, Jellyfish Search Optimization.*

1. Introduction

Heart disease (HD) has been the primary source of most deaths worldwide over the past few years. The coronary arteries, which carry blood to the heart, are thought to be the most common cause of heart failure. HD is also known as cardiovascular disease (CVD) [2]. Family genetics, high blood pressure indications, cholesterol levels, sex(Male or Female), age of the subject, diet followed, calcium rate, a stretched level of blood vessels, and general lifestyle are the main risk factors for heart disease [12, 20, 21, 22]. The majority of risk factors leading to heart disease fall into two categories: changeable factors and other which are non-changeable factors [1]. Cholesterol level, blood pressure, etc., are variables that can be changed, whereas a patient's age, gender, and medical history are variables that cannot be changed. The characteristics of early diagnosis, treatment, and recovery are present in heart disease [31]. Therefore, the key to treatment is early diagnosis of this disease [5, 17], where early detection can significantly improve patient survival rates. Unintentional biases, errors, and high medical costs are expected outcomes of most HD prediction methods, negatively impacting patient care quality.

On the other hand, clinical decision support based on computer-based patient records can reduce medical errors, increase patient safety, reduce unintentional practice variation, and improve patient outcomes [10, 11]. The World Health Organization (WHO) recognized that data mining has the potential to diagnose heart disease and predict its early stages accurately. The discovery of knowledge from a large amount of raw data is the essence of data mining [6].

Recently, numerous researchers have developed various models and methods for creating an expert method that can detect HD in its earliest stages. In deep learning, based on a Neural Network (NN), the data's micro-analytics can generate a desired result that enables the utilization of multiple neurons and layers [14, 15]. In order to accurately and meticulously predict heart disease and provide patients with a means of early detection, deep learning is necessary. More accuracy can be achieved by deep learning models, sometimes surpassing human performance.

A machine learning algorithm uses the input samples to apply a flexible classification (or prediction) model. NN, support vector machines (SVM), decision trees, and other machine learning techniques that use computational methods to diagnose heart disease have emerged over the past ten years [23, 24]. Particularly, there has been a resurgence of interest in using NN to diagnose Heart Disease in

identification tasks. In heart disease detection, neural networks have demonstrated their superior classification capabilities [4, 13]. The problem of the dataset fitting in the classification being underfitting and overfitting is eliminated by the Deep Neural Network (DNN) with two statistical models. Better testing and training data results eliminate these issues [10, 16, 17]. Heart disease prediction using two statistical models and DNN [10, 18, 19] focuses on eliminating underfitting and overfitting issues by fine-tuning the features.

Using the proposed JSO_DRN, the primary goal of this study is to design and develop a method for predicting HD. The input dataset is pre-processed by missing data imputation and Z- score Normalization. After that, the pre-processed output is subjected to a feature extraction stage. Here, features are selected based on similarity like Kumar-Hassebrook. Then, the HD prediction is achieved based on DRN, which has been trained using the proposed JSO.

The following is a description of the paper's contribution:

Developed JSO_DRN for heart disease detection: An efficient approach for heart disease diagnosis is developed using JSO_DRN. Here, DRN is trained to utilize JSO for attaining better performance measures.

The residual section of this research paper is arranged as shown: A review of the paper's relevant literature, as described in section two. Section 3 reveals the proposed method for predicting HD. The proposed approach's outcomes are discussed in section four, and Section five explains the conclusion.

2. Motivation

Pre-processing the data using traditional feature selection methods allows for accurate HD identification; however, this method has significant drawbacks and requires a lot of time, effort, and high computational costs. Identifying and predicting such heart diseases at their earlier stages is much more critical, so researchers should introduce a novel JSO_DRN.

2.1. Literature Survey

Arslan ali et al. [1] introduced the Optimally Configured and Improved Deep Belief Network named OCI-DBN to improve the prediction accuracy for heart disease. Heart disease prediction utilized the Ruzzo-Tompa technique to get rid of features that did not make enough of an impact on system performance and which applied a stacked genetic technique that stacks two genetic methods to produce an optimally designed DBN in order to discover the best configuration for the network. This technique helped medical professionals make better decisions and improved accuracy, but it failed to investigate the time complexity.

Domor Mienye et al. [2] developed a stacked sparse auto-encoder network (SSAE) for efficient heart disease prediction. The SSAE's feature learning and classification capabilities were enhanced by the Particle Swarm Optimizer (PSO). As a result, batch normalization was implemented to prevent this issue. This algorithm achieved more convergence speed but failed to consider stacking autoencoders to enhance classification performance.

Xiaoqing Gu et al. [3] devised the back-propagation (BP) neural network to enhance heart disease prediction. However, the BP neural network's performance declines when dealing with complex medical diagnostic tasks. This technique improved classification performance by maximizing the geometric margin and enhancing the data compactness within the class. Moreover, this technique does not test the noisy data.

Pooja Rani et al. [4] developed the hybrid decision support for early heart disease prediction depending on the patient's clinical data. For data pre-processing, SMOTE (Synthetic Minority Oversampling Technique) and standard scalar methods were also utilized. In the final phase of this method's development, support vector machines, naive bayes, logistic regression, random forest, and adaboost classifiers were utilized by prediction of HD. This method iterated over numerous generations to create the best solution, but the severity of heart disease cannot be diagnosed with this system.

2.2. Major Challenges

Various problems faced by traditional techniques for heart disease prediction are discussed below:

- SSAE was developed for accurate heart disease detection [2]. This method guarantees that the network is learned better representations, but it failed to include feature learning on various classification approaches to increase the efficacy.
- A hybrid decision support system was developed for early heart disease detection [4]. Using this strategy, the dataset's remaining attributes were utilized to calculate the range of missing attributes. In addition, this method does not permit the real-time collection of clinical parameters like body temperature, ECG, oxygen level, and pulse rate.
- In [3], a DNN was developed for heart disease detection. Increasing the class's data compactness and maximizing the geometric margin effectively enhances classification performance, but the speed of this technique is deficient.
- The dataset's overfitting issue is the most prevalent obstacle in heart disease prediction. Further, the availability of a balanced dataset affects the classification performance, a significant challenge the prevailing methods face.

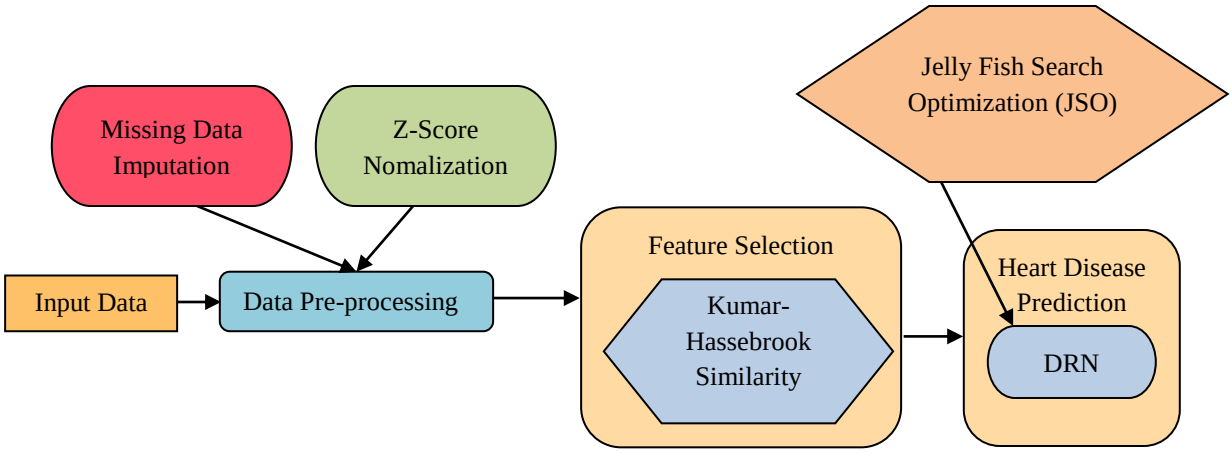

Fig. 1 Block diagram for heart disease prediction using JSO_DRN

3. Proposed Heart Disease Prediction using JSO-DRN

The primary intention of the paper is to plan and devise a new approach for HD identification by JSO_DRN. Initially, the input data from the dataset specified in [29] is fed into the data pre-processing phase, which is critical as the data may have errors, missing data, noise, and many other problems that make the data unsuitable for direct usage. Here, pre-processing is performed using Z- score Normalization [28] and missing data imputation. After that, the pre-processed outcome is transferred to the feature selection stage, wherein features are selected using Kumar-Hassebrook similarity. Finally, the feature vector is subjected to the heart disease prediction phase, wherein heart disease detection is accomplished by DRN [26, 27]. Here, DRN is trained by JSO [25]. In addition, the block diagram for heart disease prediction using JSO_DRN is depicted in Figure 1.

3.1. Data Acquisition

Let us take into consideration the input data from the dataset specified in [29], which is expressed as follows:

$$A=\{c_1, c_2,\ldots\ldots..c_a,\ldots..c_b\} \quad (1)$$

Where, A denotes input dataset, a^{th} the dataset is specified as c_a an entire amount of datasets represented by b

3.2. Data Pre-Processing

In the pre-processing phase, input data is subjected to Z-score normalization [28] and missing data imputation is used to remove errors, missing data, and noise.

3.2.1. Z-score Normalization

Putting various variables on the same scale is known as Z-score Normalization, which divides a score's deviation by a data set's standard deviation to create uniform scores on the same scale. This idea lets you compare scores for different kinds of variables.

$$d` = \frac{d_e - B_e}{std(B)} \quad (2)$$

Where d' denotes the output's normalization value, d considered a range of the normalized attribute, B_e represents the mean value in attributes. $std(B)$ is derived from the standard deviation attributes B .

3.2.2. Missing Data Imputation

When dealing with a single attribute value that is missing, this approach is used, which is applied where the data is missing. In other words, the imputation process for missing data substitutes the value calculated based on the details already available for the missed value. In other words, the process of replacing missed data with substituted values, and finally, the pre-processed output H.

3.3. Feature Selection

After pre-processing, the data H is subjected to a feature selection step, where the procedure of choosing relevant features is known as feature selection. The evaluation criteria are used to select the best feature subset. Here, the feature selection is utilized by Kumar-hasse Brook [30] similarity

3.3.1. Kumar-Hassebrook Similarity

Kumar-hassebrook similarity [30] is employed for measuring the Peak-to-Correlation Energy (PCE). The input of feature selection is $S_{b\times a}$, afterwards, $S_{b\times i}$ depicted as output, where $a > i$. Moreover, Kumar-hassebrook similarity is expressed as,

$$C_{jac} = \frac{\sum DE}{\sum D^2 + \sum E^2 + \sum DE} \quad (3)$$

I_t signifies the candidate feature, C_{jac} denotes the Jaccard function and E represents the highest target value. Here, the features selected based on the top feature with the

highest value, and the outcome of a feature selection is considered as I.

3.4. Heart Disease Prediction

The selected feature I is transferred to the heart disease prediction phase. HD describes a condition that disrupts normal heart function. Heart disease prediction is essential to provide treatment plans and diagnose as early as possible. The heart disease prediction output is denoted as R , and it is done by DRN, which is carried out using JSO.

3.4.1. Architecture of DRN

The DRN [26] architecture includes two types of layers: convolutional and pooling layers in a residual structure. This method automatically yields significantly superior results compared to learning in which this approving layer is not utilized. It shows a positive effect on lowering the error rate of classification tests.

Convolution Layer: The 2 Dimension convolution layer is used in training to reimburse weight sharing and reduce accessible attributes. The conv layer helps process the input image by utilizing a local connection and a sequence of filters known as the kernel. The convolution layer uses a mathematical approach to slide the filter with the input matrix and calculates the kernel's dot product.

Pooling Layer: The purpose of the layer, which is linked to the convolution layer, is to decrease the size of the spatial feature map. In order to apply the average pooling technique to each feature map slice and depth, this decision is made.

Residual Structure: This structure indicates that shortcut associations exist in convolution layers. Direct connections between the input and output only occur when they are of equal size. For a variety of sizes, the dimension matching factor is altered to match input and output, where the DRN output is P_e

3.4.2. Training of DRN using JSO

The JSO is used to train the DRN for heart disease prediction. Here, JSO behaves as jellyfish in the ocean. Jellyfish's following of ocean current, their movements within a JSO (active and passive), and simulation of their search behaviour includes their convergence into a jellyfish bloom and a mechanism to control time for switching among these movements.

Initialization of Population: Jellyfish populations are typically initial with randomness, which is considered the following equation,

$$F_{e+1} = \beta F_e(1\text{-}F_e),\ \ 0 \le F_0 \le 1 \tag{4}$$

Where F_e represents the logistic chaotic location value of e^{th} jellyfish, F_0 denotes to utilize for jellyfish initial population generating, $F_0 \in (0, 1)$, $F_0 \in \{0.0, 0.25, 0.75, 0.5, 1.0\}$ and β represents the 4.0.

Fitness Function: Here, the fitness function is achieved based on a Mean Square Error (MSE), which is expressed as,

$$\mathrm{MSE} = \frac{1}{a}\sum_{e=1}^{a}(P_e^* - P_e)^2 \tag{5}$$

Here, a it represents the number of samples, P_e^* denotes the expected output and P_e explains the DRN output

Ocean Current: Jellyfish are drawn to ocean current because it is rich in nutrients. The ocean currents direction $\overrightarrow{trend}$ is computed by averaging all the vectors from each jellyfish in the ocean (overall space) to the jellyfish currently in the best (optimal) location.

$$F_e(g+1) = F_e(g) + \mathrm{rand}\,(0,1)\,x\,\overrightarrow{\mathrm{trend}} \tag{6}$$

Where $F_e(g)$ is derived as e^{th} jellyfish have located at the time, represents the current ocean direction and denotes the random number with a value between 0 and 1.

Jellyfish Swarm: Jellyfish have passive (type A) and active (type B) motions in a swarm. When the swarm is still in its infancy, most jellyfish move in a type A manner, gradually exhibiting a type B motion over time. Type A motion is when jellyfish move independently, with each jellyfish providing its most recent location represented.

$$F_e(g+1) = F_e(g) + \beta\ x\ \mathrm{rand}\,(0,1)\ x\ (G_f - K_f) \tag{7}$$

$$X_i(t+1) = X_i(t) + \overrightarrow{\mathrm{step}} \tag{8}$$

$F_e(g)$ is derived as e^{th} jellyfish is positioned at the time g, β is the motion coefficient, G_f and G_f it portrays the upper and lower bound search spaces and $\overrightarrow{step}$ refers to the step size of the movement.

Time control Mechanism: This scenario is simulated by introducing the time control mechanism. A time control function is inbuilt into the time control mechanism, and a constant to regulate the jellyfish's movement between moving inside the swarm and following the ocean current.

$$h(g) = \left|\left(1 - \frac{g}{Max_{iter}}\right) x\,(2^{*}rand\ (0,1) - 1)\right| \tag{9}$$

Where the amount of iteration is represented as g . An initialized parameter, the maximum allowed iterations count is Max_{iter}

Boundary Condition: There are oceans worldwide, and as the earth is roughly spherical, a jellyfish returns to the opposite bound whenever it leaves the bound of its search area.

$$\begin{cases} \grave{F_{e,c}} = (F_{e,c} - G_{f,c}) + K_f(c) & if\ F_{e,c} > G_{f,c} \\ \grave{F_{e,c}} = (F_{e,c} - K_{f,c}) + G_f(c) & if\ F_{e,c} > K_{f,c} \end{cases} \quad (10)$$

$F_{e,c}$ represents the location of e^{th} jellyfish in a c^{th} measurement. Then, $F`_{e,c}$ restructured location after the boundary condition has been checked $G_{f,c}$ and $K_{f,c}$ denotes the upper and lower bound e^{th} dimensions in the search spaces.

Conceptual Representation of Artificial JSO: In the artificial JSO, a time control mechanism moves between exploration and exploitation by moving toward a current and within a jellyfish swarm. At first, exploration is more likely than exploitation to uncover areas with promising optimal positions. As the likelihood of exploitation increases significantly, the jellyfish choose the most advantageous location within the designated areas.

Termination: Finally, the procedure above is repeated until the maximum number of iterations has been attained.

4. Results and Discussion

This work represents an overview of the experimental results of the newly devised JSO-DRN approach for heart disease prediction.

4.1. Experimental Set-Up

The developed method is implemented with the help of the Python tool.

4.2. Dataset Description

The heart disease prediction input data is taken from the Heart Disease Data Set [29], which contains four databases: Cleveland, Hungary, the VA Long Beach and Switzerland. Here, heart disease prediction is utilized in Cleveland and Hungary datasets.

4.3. Assessment Metrics

The four assessment parameters the developed method uses to predict heart disease are accuracy, F1_score, TPR, and MCC.

Accuracy: One of the most common metrics for prediction performance is prediction accuracy, defined as the proportion of samples correctly classified to total samples. It is expressed as

$$Acc^* = \frac{PQ+MQ}{PQ+PO+MQ+MO} \quad (11)$$

Where PQ shows the true positive rate, MQ denotes the false positive rate, PO represents the actual negative rate, MO consider the false negative rate.

TPR: The following equation represents TPR, where TPR compares the number of positive samples to the total number correctly classified.

$$PQS = \frac{PQ}{PQ+MO} \quad (12)$$

Matthews Correlation Coefficient (MCC): MCC is calculated directly from the confusion matrix and represents the correlation between the predicted and observed classifications based on a true positive, true negative, false positive and false negative. Here, MCC is expressed as,

$$MCC` = \frac{PQ^*PO - MQ^*MO}{\sqrt{(PQ+MQ)(PQ+MO)(PO+MQ)(PO+MO)}} \quad (13)$$

F1_Score Measurement: F1_Score measure is calculated using the test's precision and recall, where recall is the number of results that are deemed true positive divided by all the samples that should have been identified as positive, and accuracy is the number of results that are positive and true divided by all positive results, also considering those that were not correctly identified. F_Score measure is expressed as,

$$F_{Score} = \frac{2PQ}{2PQ+MQ+MO} \quad (14)$$

4.4. Comparative Techniques

In comparative evaluation, various traditional techniques, such as Generative Adversarial Network One-Dimensional Convolutional neural Network (GNA-1D-CN) [7], Recurrent Neural Network + Gated Recurrent Units (RNN + GRU) [8, 9], Heart Disease Prediction Model (HDPM) [10], and Hybrid decision support system [4], are considered.

4.5. Comparative Evaluation

The performance parameters varied with esteem to training data based on the Cleveland and Hungarian data are used to compare the introduced JSO_DRN, as shown below.

4.6. Comparative Assessment Considering the Cleveland Dataset

Here, figure 2 portrays the comparative discussion of evaluation parameters and training data of JSO_DRN. The accuracy evaluation of the modelled method is depicted in Figure 2a. With 90% data to propose the JSO_DRN approach attained an accuracy rate of 0.846, and the additional existing technique, GNA-1D-CN of 0.665, RNN+GRU of 0.7265, HDPM of 0.746 and Hybrid decision

support system of 0.816. The proposed method has a better accuracy rate of 11.82% than the HDPM method. The F1-score-based analysis of JSO_DRN is depicted in Figure 2b.

In this analysis, the JSO_DRN technique was used to measure the F1_score of 0.865 for 90% of the training data, while RNN+GRU, GNA_ID_CN, HDPM, and the hybrid decision support system is 0.755, 0.725, 0.806, and 0.825 respectively. The proposed JSO_DRN method has a better F1 score of 12.71% than the RNN+GRU method.

The MCC obtained for the training data of 90% by GNA-1D-CN is 0.856, RNN+GRU is 0.736, HDPM is 0.786, the hybrid decision support system is 0.785 and proposed JSO_DRN is 0.856 depicted in figure 2c. The proposed method's MCC rate is 8.17% higher than the HDPM method's. Figure 2d portrays the evaluation of JSO_DRN regarding TPR, here 90% training percentage of the JSO_DRN approach TPR is 0.826 and a TPR attained by the other methods, like RNN+GRU is 0.726, GNA-1D-CN is 0.685, the hybrid decision support system is 0.786, and HDPM is 0.746. The TPR rate of the introduced method is better by 12.10% than the rate of the RNN+GRU method.

4.7. Comparative Assessment with Hungarian Dataset

Figure 3 depicts the evaluation of the proposed JSO_DRN by training data value for various evaluation indicators using the Hungarian dataset. Figure 3a signifies the evaluation of various techniques regarding accuracy.

If a training percentage is 90%, the accuracy rate achieved by JSO_DRN of 0.846 and other conventional methods, like, GNA-1D-CN, RNN+GRU, HDPM, Hybrid decision support system attained an accuracy of 0.698,0.726, 0.795 and 0.806. The improvement of the performance of 14.18% is achieved compared to the RNN+GRU method.

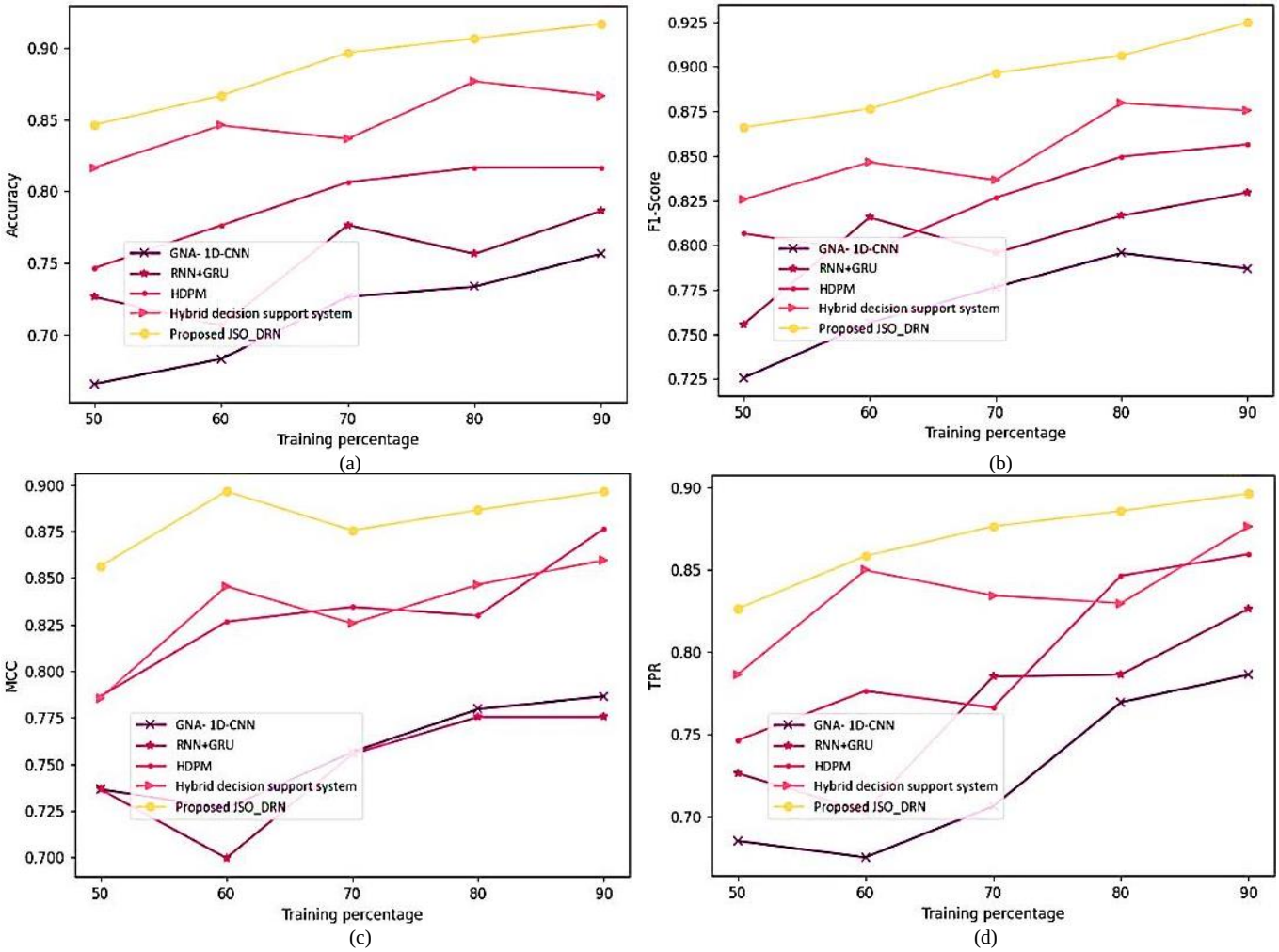

Fig. 2 Comparative evaluation with cleveland dataset of developed JSO_DRN based on a) Accuracy, b) F1_Score, c) MCC, and d) TPR

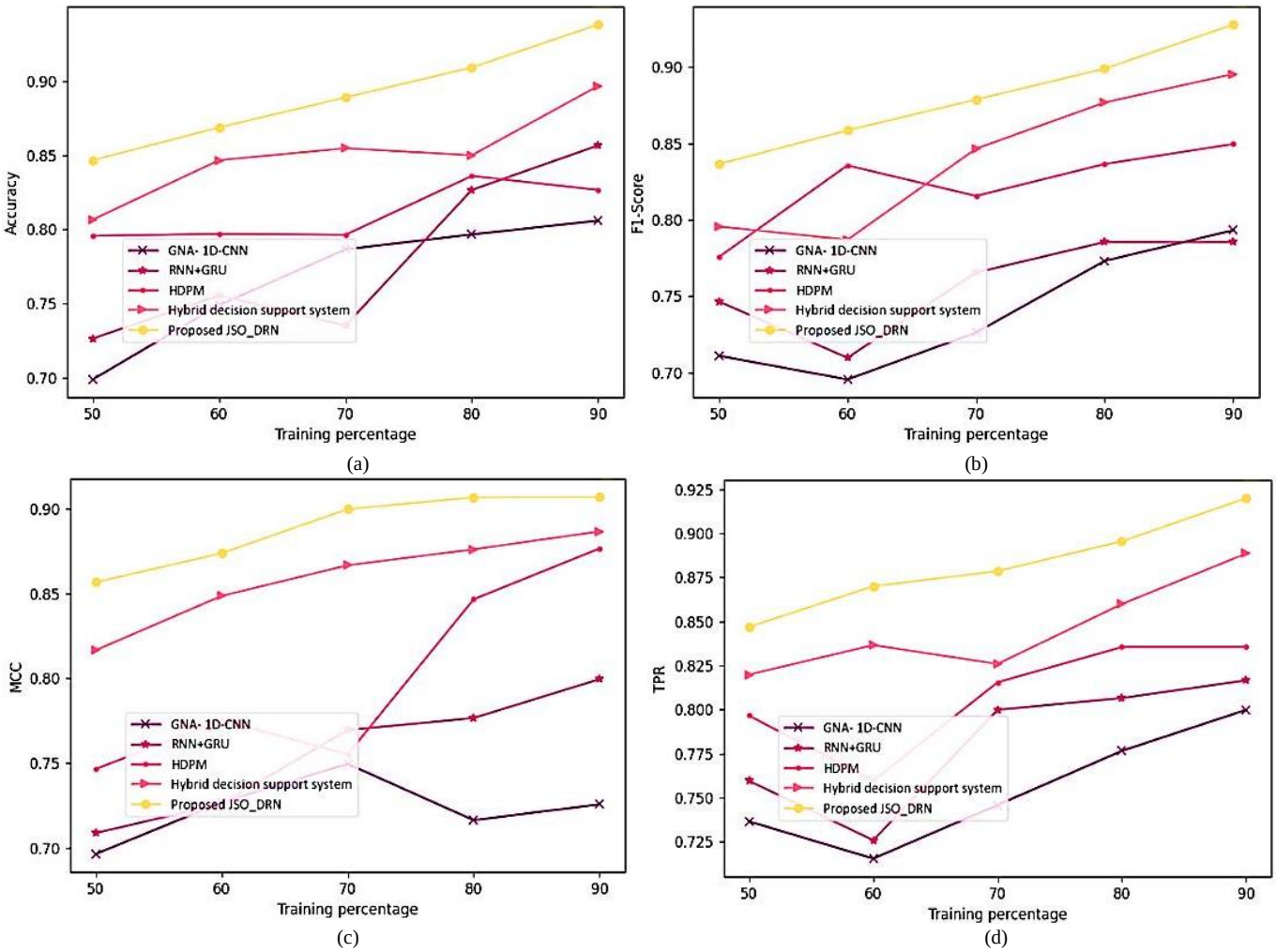

Fig. 3 Comparative evaluation with hungarian dataset of developed JSO_DRN considering a) Accuracy, b) F1_Score, c) MCC, and d) TPR

The evaluation of JSO_DRN concerning the F1_score is depicted in Figure 3 b. With training data 90%, the F1_score gained by the developed method of 0.836 exposes a performance improvement when compared with the existing techniques like 0.746 of GNA-1D-CN, 0.711 of RNN+GRU, 0.775 of HDPM, 0.795 of Hybrid decision support system. The proposed method JSO_DRN has a better score of 10.76% than the GNA_ID_CN method.

Figure 3c portrays the evaluation of JSO_DRN in regard to MCC. If a training percentage is 90%, the accuracy achieved by GNA-1D-CN is 0.696, RNN+GRU is 0.708, the hybrid decision support system is 0.816, HDPM is 0.746, and the proposed JSO_DRN is 0.856. The proposed method has a better MCC rate of 4.77% than the hybrid decision support system method rate.

Figure 3d signifies the evaluation of different technique regard to TPR. If the training data is 90%, TPR reached by the JSO_DRN method is 0.826, and the TPR attained by the other methods, RNN+GRU of 0.759, GNA-1D-CN of 0.736, hybrid decision support system 0.819, and HDPM of 0.796. The proposed method's TPR rate is superior to the HDPM method's by 4.67%.

4.8. Comparative Result

This section compares the designed model to various evaluation parameters, demonstrating that the developed JSO_DRN strategy performed better. Because here, we utilized the Z-score Normalization and missing data imputation for pre-processing phase, the superior value under the parameters with an accuracy of 84.63%, F1_score of 86.59%, MCC of 85.65%, and TPR of 84.68%. Moreover, the traditional GNA-1D-CN technique obtains 66.58% accuracy, 72.56% F1_score,85.9% MCC, and 68.54% TPR.

Existing RNN+GRU technique measured accuracy, F1_score, MCC, and TPR of 72.65%,75.56%,73.65%, and 72.65%, HDPM measured 74.65%,80.65%,78.65%, and 74.65%.

Table 1. Discussion of the developed method concerning various traditional methods

Dataset	Evaluation parameters	GNA-1D-CN	RNN+ GRU	HDPM	Hybrid decision support system	Proposed JSO_DRN
Cleveland	Accuracy (%)	66.58	72.65	74.65	81.65	84.63
	F1_score(%)	72.56	75.56	80.65	82.56	86.59
	MCC (%)	73.65	73.65	78.65	78.56	85.64
	TPR (%)	68.54	72.65	74.65	78.65	82.65
Hungarian	Accuracy (%)	69.89	72.65	79.56	80.65	84.65
	F1_score (%)	71.10	74.65	77.56	79.56	83.65
	MCC (%)	69.65	70.88	74.65	81.65	85.65
	TPR (%)	73.65	75.95	79.65	81.98	84.68

The hybrid decision support system reported an accuracy of 81.65%, a F1 score of 82.56%, a MCC of 78.65%, and TPR of 82.65%. As a result, the above discussion reveals that the JSO approach improves the DRN's performance for effective heart disease prediction with increased evaluation metrics. Table 1 compares and discusses the proposed method for predicting heart disease.

5. Conclusion

Because it is required to pump blood rich in oxygen, the heart is the most important part of the human body. Arrhythmia, coronary artery disease, congenital heart disease, and others are all examples of heart diseases. Therefore, an approach that can predict heart disease at less cost is necessary. This study used the JSO-enabled DRN approach to predict heart disease. The HD data set serves as the input data source, which is then pre-processed with Z-Score Normalization and imputation of missing data to improve data quality.

The feature selection phase uses the similarity algorithm - kumar-hassebrook to select features from the pre-processed data. Consequently, HD detection depends on DRN, which JSO trains. Proposed DRN_JSO received better outcomes as compared with several methods, such as GNA-1D-CN, RNN+GRU, HEART DISEASEPM and Hybrid decision support system techniques with 84.65% accuracy, 83.65% F1_score, 85.65% of MCC and 82.65% TPR. Further, the devised approach will investigate the effects of feature extraction for enhancing heart prediction in future studies.

Reference

[1] Syed Arslan Ali et al., "An Optimally Configured and Improved Deep Belief Network (OCI-DBN) Approach for Heart Disease Prediction Based on Ruzzo–Tompa and Stacked Genetic Algorithm," *IEEE Access*, vol. 8, pp. 65947-65958, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[2] Ibomoiye Domor Mienye, and Yanxia Sun, "Improved Heart Disease Prediction using Particle Swarm Optimization Based Stacked Sparse Auto Encoder," *Electronics*, vol. 10, no. 19, pp. 1-15, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[3] Xiaoqing Gu, Yizhang Jiang, and Tongguang Ni, "Discriminative Neural Network for Coronary Heart Disease Detection," *Journal of Medical Imaging and Health Informatics*, vol. 10, no. 2, pp. 463-468, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[4] Pooja Rani et al., "A Decision Support System for Heart Disease Prediction Based Upon Machine Learning," *Journal of Reliable Intelligent Environments*, vol. 7, no. 3, pp. 263-275, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[5] Liaqat Ali et al., "An Automated Diagnostic System for Heart Disease Prediction Based on X^2 Statistical Model and Optimally Configured Deep Neural Network," *IEEE Access*, vol. 7, pp. 34938-34945, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[6] Raniya R. Sarra et al., "Enhanced Heart Disease Prediction Based on Machine Learning and X^2 Statistical Optimal Feature Selection Model," *Designs*, vol. 6 no. 5, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[7] Raniya R. Sarra et al., "A Robust Framework for Data Generative and Heart Disease Prediction Based on Efficient Deep Learning Models," *Diagnostics*, vol. 12, no. 12, pp. 1-19, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Surenthiran Krishnan, Pritheega Magalingam, and Roslina Ibrahim, "Hybrid Deep Learning Model using Recurrent Neural Network and Gated Recurrent Unit for Heart Disease Prediction," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 6, pp. 5467-5476, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Vaibhav Gupta, and Pallavi Murghai Goel, "Heart Disease Prediction using ML," *SSRG International Journal of Computer Science and Engineering*, vol. 7, no. 6, pp. 17-19, 2020. [CrossRef] [Publisher Link]

[10] Norma Latif Fitriyani et al., "HDPM: An Effective Heart Disease Prediction Model for A Clinical Decision Support System," *IEEE Access*, vol. 8, pp. 133034-133050, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[11] Ravneet Kaur, "Concept of Pippali (Piper Longum) Verses Heart Diseases," *SSRG International Journal of Medical Science*, vol. 6, no. 5, pp. 5-6, 2019. [CrossRef] [Publisher Link]

[12] Jaishri Wankhede, Magesh Kumar, and Palaniappan Sambandam, "Efficient Heart Disease Prediction-Based on Optimal Feature Selection using DFCSS and Classification by Improved Elman-SFO," *IET Systems Biology*, vol. 14, no. 6, pp. 380-390, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[13] T. M. Angelin Monisha Sharean, and G. Johncy, "Deep Learning Models on Heart Disease Estimation-A Review," *Journal of Artificial Intelligence and Capsule Networks*, vol. 4, no. 2, pp. 122-130, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[14] J. Thomas, and R. Theresa Princy, "Human Heart Disease Prediction System using Data Mining Techniques," *In 2016 International Conference on Circuit, Power and Computing Technologies*, pp. 1-5, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[15] Saba Bashir et al., "Improving Heart Disease Prediction using Feature Selection Approaches," *In 2019 16th International Bhurban Conference on Applied Sciences and Technology*, pp. 619-623, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[16] Srabanti Maji, and Srishti Arora, "Decision Tree Algorithms for Prediction of Heart Disease," *Information and Communication Technology for Competitive Strategies*, vol. 40, pp. 447-454, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[17] Mohd Ashraf, M. A. Rizvi, and Himanshu Sharma, "Improved Heart Disease Prediction using Deep Neural Network," *Asian Journal of Computer Science and Technology*, vol. 8, no. 2, pp. 49-54, 2019. [Google Scholar] [Publisher Link]

[18] Yuanyuan Pan et al., "Enhanced Deep Learning Assisted Convolutional Neural Network for Heart Disease Prediction on the Internet of Medical Things Platform," *IEEE Access*, vol. 8, pp. 189503-189512, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[19] Abhijeet Sudhakar, Rajendra B. Patil, and R. Srivaramangai, "Descriptive Data Analysis of Contagious Diseases using Statistical Parameters," *SSRG International Journal of Computer Science and Engineering*, vol. 6, no. 9, pp. 1-6, 2019. [CrossRef] [Publisher Link]

[20] J. Jeyaranjani, T. Dhiliphan Rajkumar, and T. Ananth Kumar, "Coronary Heart Disease Diagnosis using the Efficient ANN Model," *Materials Today: Proceedings*, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[21] Sesham Anand, "Archimedes Optimization Algorithm: Heart Disease Prediction: Archimedes Optimization Algorithm: Heart Disease Prediction," *Multimedia Research*, vol. 4, no. 3, 2021. [Google Scholar] [Publisher Link]

[22] I. Lakshmi, "Prediction Analysis on Heart Disease using HNB and NB Techniques," *SSRG International Journal of Computer Science and Engineering*, vol. 5, no. 10, pp. 11-15, 2018. [CrossRef] [Publisher Link]

[23] Eric J. Chow et al., "Prediction of Ischemic Heart Disease and Stroke in Survivors of Childhood Cancer," *Journal of Clinical Oncology*, vol. 36, no. 1, pp. 44-52, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[24] Lijetha C. Jaffrin, J. Visumathi, and G. Umarani Srikanth, "Role of Exploratory Analytics and Visualization in Heart Disease Prediction," *International Journal of Engineering Trends and Technology*, vol. 70, no. 3, pp. 22-28, 2022. [CrossRef] [Publisher Link]

[25] Jui-Sheng Chou, and Dinh-Nhat Truong, "A Novel Metaheuristic Optimizer Inspired by the Behaviour of Jellyfish in the Ocean," *Applied Mathematics and Computation*, vol. 389, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[26] Zhicong Chen et al., "Deep Residual Network Based Fault Detection and Diagnosis of Photovoltaic Arrays using Current-Voltage Curves and Ambient Conditions," *Energy Conversion and Management*, vol. 198, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[27] M. Arunanand, and T. Viswanathan, "IoT Based Intelligent Outdoor Healthcare Monitoring System for Heart Disease Patients using Smartphone and Smart Clothing," *International Journal of Biotech Trends and Technology*, vol. 7, no. 2, pp. 7-12, 2017. [CrossRef][Publisher Link]

[28] Hestu Aji Prihanditya, "The Implementation of Z-Score Normalization and Boosting Techniques to Increase the Accuracy of C4.5 Algorithm in Diagnosing Chronic Kidney Disease," *Journal of Soft Computing Exploration*, vol. 1, no. 1, pp. 63-69, 2020. [CrossRef][Google Scholar][Publisher Link]

[29] Heart Disease Data Set. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/heart+Disease

[30] Sung-Hyuk Cha, "Comprehensive Survey on Distance/Similarity Measures between Probability Density Functions," *International Journal of Mathematical Models and Methods in Applied Sciences*, vol. 4, no. 1, pp. 300-307, 2007.[Google Scholar][Publisher Link]

[31] Chetan Nimba Aher, and Ajay Kumar Jena, "Soft Computing Based Approaches for Classifying Diseases using Medical Diagnosis Dataset," *2020 International Conference on Emerging Smart Computing and Informatics*, India, pp. 77-81, 2020. [CrossRef][Google Scholar][Publisher Link]