International Journal of Engineering Trends and Technology
ISSN: 2231 – 5381 / https://doi.org/10.14445/22315381/IJETT-V70I4P206
Volume 70 Issue 4, 59-81, April 2022

Original Article

Heart Disease Prediction based on Ensemble Classification Model with Tuned Training Weights

Parvathaneni Rajendra Kumar[1], Suban Ravichandran[2], S. Narayana[3]

[1,2]*Annamalai University, Annamalainagar.*
[3]*Gudlavalleru Engineering College, Andhra Pradesh, India.*

prajendrakumar@sircrrengg.ac.in

Received: 19 February 2022 Revised: 28 March 2022 Accepted: 28 March 2022 Published: 26 April 2022

Abstract *- Heart disease (HD) is the most serious human disease, causing havoc on people's health. Heart disease detection must be accurate and timely to prevent and cure heart failure. In many instances, the diagnosis of HD based on standard medical history is seen as unreliable. Therefore, this paper introduces a novel HD prediction system that includes five major phases such as (a) Preprocessing, (b) Imbalance processing, (c)Feature extraction, (d) Feature Selection and (e) Classification. Originally, the input data is given to the preprocessing phase. Subsequently, the imbalance processing phase is carried out, where an improved strategy for the class imbalance process is performed. The features, including raw features, improved mutual information, higher-order statistical features, entropy, correlation, and statistical features, are extracted in the feature extraction phase. Moreover, appropriate features will be selected from the extracted features in the feature selection phase, for which an improved ReliefF process will be carried out. These selected features is then subjected to the classification phase, where the ensemble classifiers include Neural Network (NN), Recurrent Neural Network (RNN), Random Forest (RF), and K-Nearest Neighbour (k-NN) model. Here, the output of NN, RNN, and RF is given as the input of k-NN. To make the system more precise in disease prediction, the weights of NN and RNN are optimally tuned by a Self-improved Shark Smell Optimization with Gaussmap Estimation and Cycle crossover Operation (SISSGECO) model. Then, the final output is obtained effectively in a precise manner. Finally, the outcomes of the adopted scheme are computed to the other extant schemes in terms of various measures like precision, sensitivity, accuracy, specificity, NPV, MCC, FPR, F1-score, and FNR, respectively.*

Keywords *- Heart Disease Prediction, Imbalance Processing, Improved ReliefF, Ensemble Classifiers, Optimization.*

Nomenclature

Abbreviation	Description
OCFS	Optimality Criterion Feature Selection
SMOTE	Synthetic Minority Oversampling Technique
EMRs	Electronic Medical Records
RNN	Recurrent Neural Network
SSO	Shark Smell Optimization
WHO	World Health Organization
NN	Neural Network
UCI	University Of California, Irvine
RF	Random Forest
DT	Decision Tree
HPO	Hyper Parameter Optimization
K-NN	K-Nearest Neighbour
GA	Genetic Algorithm
LR	Logistic Regression
NB	Naïve Bayes
SVM	Support Vector Machine
RFS-IE	Rough Feature Selection based on Information Entropy
ANN	Artificial Neural Networks
PM-LU	PSO Merged LA Update
LA	Lion Algorithm
PSO	Particle Swarm Optimization
HD	Heart Disease
WOA	Whale Optimization Algorithm
XGBoost	Extreme Gradient Boosting
OH	One-Hot
AUC	Area Under the Curve
FPR	False Positive Rate
RMSE	Root Mean Square Error
AI	Artificial Intelligence
DBN	Deep Belief Network
ROC	Receiver Operating Characteristic
LM	Levenberg–Marquardt
MCC	Matthews Correlation Coefficient
FF	Firefly
CHD	Cleveland Heart Disease
TPOT	Tree-Based Pipeline Optimization Tool
PCA	Principle Component Analysis
DBNKELM	Deep Belief Network and Extreme Learning Machine
CMBO	Cat Mouse Based Optimizer
NPV	Net Present Value
BOA	Butterfly Optimization Algorithm
FNR	False Negative Rate
PRO	Poor and Rich Optimization
SISSGECO	Self-Improved Shark Smell Optimization with Gaussmap Estimation and Cycle Crossover Operation
Social SO	Social Spider Optimization

1. Introduction

One of the most important and challenging health concerns is the automatic prediction of HD in the world [9]. HD [10] [11] produces coronary artery infections and impaired blood vessel function that damages the body of patients, mostly adults and the elderly. According to the WHO, cardiovascular illnesses cause more than 18 million deaths worldwide [12]. In addition, the United States of America invests $1 billion every day in HD treatments. HDs, including stroke, heart attack, and hypertension, are the leading causes of mortality in the United States. Consequently, early HD prediction [13] [14] is more helpful for treating heart patients efficiently earlier than a stroke or heart attack occurs.

Medical testing, as well as wearable sensors, are used to detect cardiovascular diseases [15]. Nevertheless, as clinicians aim to identify the patients promptly and precisely, collecting valuable risk indicators for HD from computerized medical testing is challenging. Due to the frequent medical testing, these EMRs are unstructured and growing in size. Wearable sensors are currently being used to identify cardiac problems by continually monitoring the patient's body both inside and outside. On the other hand, Signal abnormalities such as missing values and noise contaminate wearable sensor data produce an error in the accurate prediction of HD [16] [17].

Combining wearable sensors with EMRs for monitoring cardiac patients is a substantial and demanding challenge. Furthermore, obtaining meaningful and relevant characteristics from data to predict HD is a difficult task [18] [19]. Consequently, an intelligent scheme is essential to automatically merge information derived from EMRs and sensor data and evaluate the extracted data to anticipate heart illness [20] earlier than a heart attack happens and to find hidden indications of heart issues.

Numerous approaches are suggested using data mining approaches and hybrid models to diagnose and predict HD [21]. Risk factors are extracted from unstructured textual data using a data mining approach. Furthermore, a hybrid model [52] [53] [56] is the combination of two or more techniques that perform better collectively than they do the individual method. Moreover, the conventional algorithms [55] [54] for HD diagnosis [22] are based on feature weighting approaches. For all classes, these techniques assign the same weight to each feature. Disease prediction systems [23] play a significant part in human lives, and it's been regarded as a crucial issue, as disease prediction is necessary for people to live a peaceful life [60]. Disease detection [24] is critical for health care organizations to provide the best medical care to patients.

Moreover, the latest advancements in data mining systems resulted in several other illness prediction models that may be utilized for more purposes. Consequently, data mining plays an important role in illness prediction [25] throughout the healthcare system. For a successful prediction, the data prediction method incorporates classification methodologies such as DT, k-NN, SVM, NB, LR, and NN [50] [51] [58] [59], as well as additional clustering methods. The main contribution of this work is given below:

- Determines higher-order statistical features, statistical, raw features, improved mutual information, entropy, and correlation features and improved ReliefF based feature selection takes place.
- Implements a novel Self-Improved Shark Smell Optimization with Gaussmap Estimation and Cycle Crossover Operation (SISSGECO) model for the training of ensemble model via the optimal weights selection.

The rest of this paper is ordered as follows: Section II determines the review of the HD prediction model. Section III describes the overall framework of adopted HD prediction. Section IV portrays the preprocessing, imbalance processing, feature extraction and feature selection based on HD prediction. Section V describes classification via ensemble classifiers: RF, NN, RNN AND K-NN. Section VI specifies the weight optimization of NN and RNN via self-improved shark smell optimization with gauss map estimation and cycle crossover operation. Section VII portrays the results of the presented scheme.

2. Literature Review

2.1 Related Works

In 2020, Farman *et al.* [1] presented a smart healthcare system to predict HD employing ensemble DL and feature fusion methodologies. First, the derived features from sensor data integrated with the feature fusion approach provide useful health information. Furthermore, the information gain approach eliminates unnecessary and redundant features while focusing on the most essential, reducing computing load and improving outcomes. The suggested scheme was 98.5 % accuracy greater than that of conventional systems. In contrast to other traditional methodologies, this finding suggests that the system was more successful at predicting HD.

In 2021, Valarmathi *et al.* [2] had explored a technology that can forecast HD. To improve the performance of the RF classifier and XG Boost classifier, 3 HPO approaches have been suggested: Grid Search, Genetic programming, and Randomized Search. RF and TPOT Classifiers attained the maximum accuracy of 97.52 % for the CHD Dataset. With the ZAlizadeh Sani Dataset, RF with Randomized Search had the greatest accuracy for detecting stenos in 3 vessels: LCX, RCA, and LAD.

In 2021, Rani *et al.* [3] had established a decision support model which would help the diagnosis of HD depending on the patient's clinical factors. The authors employed a multivariate imputation approach based on chained equations to deal with missing data. To select the acceptable features in the provided dataset, a recursive feature elimination and GA were combined to form the hybridized feature selection technique was utilized. SMOTE and normal scalar techniques were also employed for data preprocessing. It was sorely tested using the CHD dataset, which may be found at UCI. The adopted model has shown higher accuracy than other HD prediction methods currently available in the literature.

In 2021, Harika *et al.* [4] adopted a unique ensemble framework for quick HD detection depending on classifiers such as ANN, SVM, and NB. The current study also validates that the three algorithms have been the most effective. The information was collected from UCI. The ensemble methodology takes the findings of individual classifiers and utilizes the majority voting method to get a solution. The ensemble model was shown to have an accuracy of 87.05 % in predicting HD than the SVM, ANN, and NB. The presented ensemble classifier was suggested to be used to forecast the cardiac status to improve accuracy and reduce misclassification.

In 2019, Prakash *et al.* [5] had suggested a novel technique that extracts characteristics from the HD dataset and builds a decision table using those features. They reported the accepted technique for illness detection and prediction using OCFS. Using indiscernible optimality criteria extends the RFS-IE approach to feature selection using the optimality criterion. It was determined how long it took to forecast cardiac disease compared to RFS-IE and MRPS. The OCFS approach required the least amount of time to execute compared to other techniques.

In 2020, Renji *et al.* [6] examined a novel HD prediction that incorporates Attribute Minimization, Feature Extraction, Classification, and Record procedures. Furthermore, the NN model that considers the dimensionally reduced information performs the prediction procedure. This study offers the PM-LU method, a new hybrid approach for NN weight optimization that combines the concepts of LA and PSO. The outcomes provided by the adopted scheme were more accurate than the traditional algorithms based on accuracy.

In 2020, Shiny *et al.* [7] had created a hybridized approach that combines DBNKELM and FKMAW based ensemble methods researchers to enhance the diagnosis. Moreover, the input qualities were first weighted with the FKMAW model. A regression analysis of the HD detection method was presented using weighted attributes with DBNKELM. For all 6 datasets, the findings show that FKMAW + DBNKELM performed well in resolving the issues in medical data categorization.

In 2020, Kartik *et al.* [8] had offered a diagnostic model which uses an enhanced XGBoost classifier to detect cardiac disease in this research. Any classifier's application would not be effective without proper hyperparameter optimization. They employed Bayesian optimization, which was a very effective approach for hyper-parameter optimization, to improve the hyper-parameters of XGBoost. They applied the OH encoding approach to encoding definite information in the dataset to boost prediction accuracy. For performance evaluation, 5 separate measures were used: accuracy, F1-score, sensitivity, AUC of ROC charts, and specificity. The study's findings demonstrated its validity and usefulness in predicting cardiac disease. Furthermore, the new model performed better compared to the previously mentioned models.

2.2 Review

Table 1 illustrates the reviews on the prediction of the HD model. Originally, the ensemble deep learning model was determined in [1] that offers higher accuracy, improved precision, better recall, and lower RMSE; however, the more sophisticated method was not determined to remove the irrelevant features. HPO techniques were exploited in [2] that offer the highest accuracy, precision, and increased F_1-score, but the HD predictions were not performed in real-time. The GA model [3] offers better accuracy, sensitivity, specificity, and precision. Nevertheless, experiment with more feature selection methods like ACO and PSO was not performed. Likewise, the AI-based Ensemble approach was used in [4], which provides larger accuracy, high sensitivity, improved specificity, and higher AUC. Still, the proposed work does not explore Python to make HD detections. OCFS scheme was portrayed in [5] as having better computational time, minimal execution time, and low error rate; however, the attribute's existence would not offer any information regarding the objects. In addition, the PM-LU algorithm was introduced in [6], which offers high accuracy, maximum specificity, improved sensitivity, better precision, and increased MCC. However, need to predict HD based on other hybrid classifier models or with other hybridization algorithms. FKMAW + DBNKELM model was suggested in [7] that offers larger precision, better accuracy, improved recall, high F-measure, and maximum AUC. However, the decision support system was not supported via DBN. Lastly, the XGBoost classifier introduced in [8] offers maximum accuracy, high specificity, larger sensitivity, and better F_1-score but needs to test other related data sets or similar tasks to attain similar accuracy. These challenges were considered effective for the heart HD prediction model in the present work.

Table 1. Reviews on Conventional Hd Prediction Models: Features And Challenges

Author [citation]	Adopted scheme	Features	Challenges
Farman *et al.* [1]	Ensemble deep learning model	❖ Higher accuracy ❖ Better recall ❖ Lower RMSE	❖ A more sophisticated method was not determined to remove the irrelevant features.
Valarmathi *et al.* [2]	HPO techniques	❖ Highest accuracy ❖ Maximum specificity ❖ Better sensitivity ❖ Improved precision ❖ Increased F_1-score	❖ The HD predictions were not performed in real-time.

Rani *et al.* [3]	GA model	❖ Best accuracy ❖ Higher sensitivity ❖ Better specificity ❖ Improved precision ❖ Increased F-Measure	❖ The experiment was carried out with more feature selection methods like ACO; PSO was not performed in this work.
Harika *et al.* [4]	AI-based Ensemble Model	❖ Larger accuracy ❖ High sensitivity ❖ Better specificity ❖ Higher AUC	❖ The proposed work does not explore Python to make HD detections.
Prakash *et al.* [5]	OCFS model	❖ Better computational time ❖ Minimal execution time ❖ Low error rate	❖ The attribute's existence would not offer any information regarding the objects.
Renji *et al.* [6]	PM-LU algorithm	❖ High accuracy ❖ Maximum specificity ❖ Improved sensitivity ❖ Better precision ❖ Increased MCC	❖ Need to predict HD based on other hybrid classifier models or other hybridization algorithms.
Shiny *et al.* [7]	FKMAW + DBNKELM model	❖ Larger precision ❖ Better accuracy ❖ Improved recall ❖ High F-measure ❖ Maximum AUC	❖ The decision support system was not supported via DBN.
Kartik *et al.* [8]	XGBoost classifier	❖ Maximum accuracy ❖ High specificity ❖ Larger sensitivity ❖ Better F1-score	❖ Need to test other related data sets or similar tasks to attain similar accuracy.

3. The Overall Framework of Adopted Hd Prediction

This work intends to introduce a novel HD prediction system that includes five major phases, such as

- Preprocessing,
- Imbalance processing,
- Feature extraction,
- Feature Selection, and
- Classification.

✓ Originally, the input data is given to the preprocessing phase, in which the data normalization process takes place.

✓ Subsequently, the imbalance processing phase is carried out, where an improved strategy on the class imbalance process is performed.

✓ Once the imbalanced problem is solved, the feature extraction is carried out, where the raw feature, improved mutual information, entropy, correlation, statistical and higher-order statistical features are extracted.

✓ Moreover, appropriate features are selected from the extracted features in the feature selection phase, for which an improved relief process is carried out.

✓ These selected features are given to the classification phase, where the ensemble classifiers include NN, RNN, RF, and k-NN models.

✓ Here, the output of NN, RNN, and RF is given as the input of k-NN.

✓ The proposed SISSGECO model optimally tunes the weights of NN and RNN to make the system more precise in disease prediction.

✓ Then, the final output is obtained effectively in a precise manner.

Fig. 1 represents the architecture of the adopted framework.

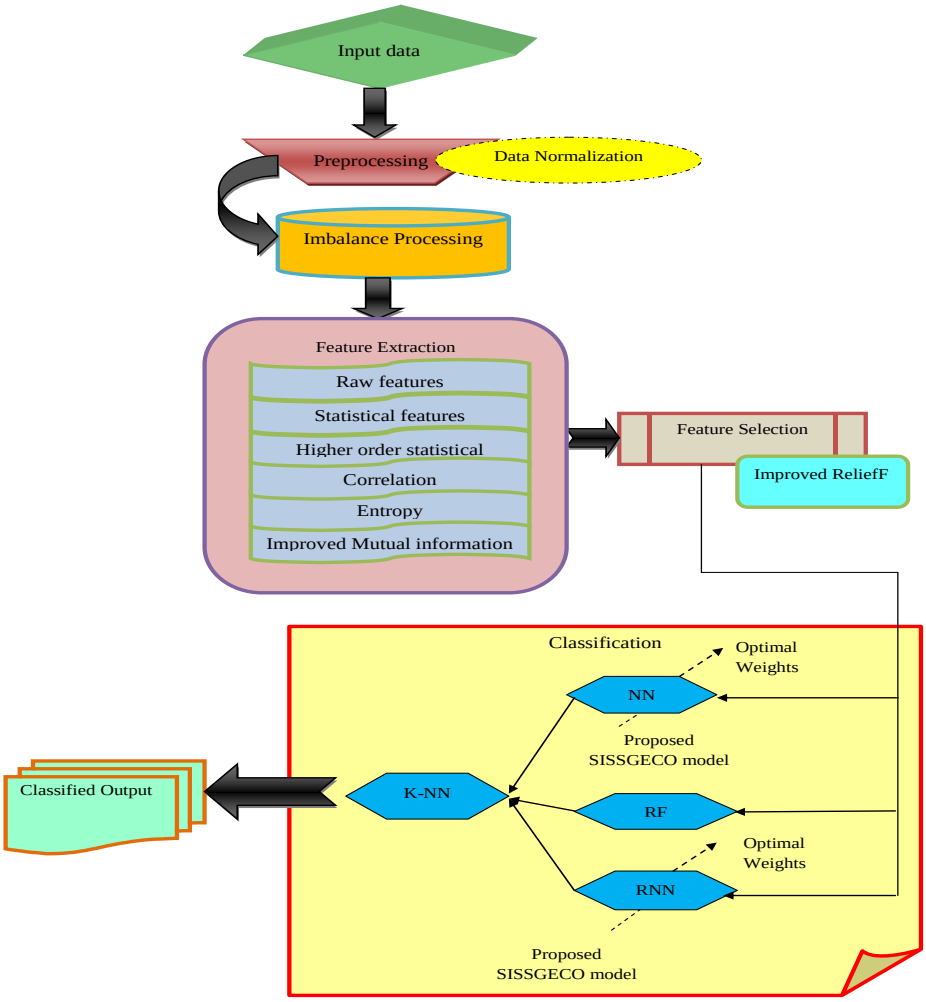


Fig. 1 The architecture of the adopted framework

4. Preprocessing, Imbalance Processing, Feature Extraction and Feature Selection Process based on Hd Prediction

4.1 Preprocessing

The preprocessing is performed via the data normalization process. Data Normalization adjusts the scales of the features to have a standard scale of measure.

4.2 Imbalance processing

The number of abnormal and normal samples must be equal. The imbalance processing is given in Algorithm 1.

Algorithm 1: Imbalance processing

Input: Training set $D=\{D_l, l=1,2,.....,C\}$; C =number of classes, and $|D|=S$; S =Total number of samples

Output: Balanced training set D' ;

Steps: $I_{resam}=Int\left(\frac{S}{C}\right)$

For $l=1$ to C do

If $|D_l|<I_{resam}$ then

$D_l'=Smote(D_l, I_{resam})$

$|D_l'|=I_{resam}$

End if

If $|D_l|>I_{resam}$ then

$V_k=MKFCM(D_l, C)$//use multi-kernel FCM to cluster D_l into clusters C

For $k=1$ to C do

$V_k'=\text{Re}sample\left(V_k, \frac{I_{resam}}{C}\right)$

End for

$D_l'=Concatenate(V_k')$

End if

$D'=Concatenate(D_l')$

End for

Return D'

The process of MKFCM is as follows

The multi-kernel will be used in FCM.

1. Polynomial kernal
2. Gaussian

$$K_1(y_l, y_i) = (y_l, y_i + d)^2 \quad (1)$$

$$K_2(y_l, y_i) = \exp\left(-|y_l - y_i|^2 / r^2\right) \quad (2)$$

$$K_{com} = K_1 + \alpha K_2 \quad (3)$$

The objective function of MKFCM is given in Eq. (4), which P_{com} indicates the kernel.

$$P = \sum_{l=1}^{C}\sum_{i=1}^{n} \mu_{li}^{m} \left\| P_{com}(y_i - O_l) \right\|^2 \quad (4)$$

4.3 Feature Extraction

The obtained preprocessed data is subjected to extracting the features, including

- ❖ Raw features
- ❖ Statistical features
- ❖ Higher-order statistical features
- ❖ Entropy
- ❖ Improved mutual information
- ❖ Correlation

4.3.1 Raw Features

Here, the original input data is considered the raw features. These features are indicated as RS .

4.3.2 Statistical features

These features are determined as follows.

- ✓ Mean
- ✓ Median
- ✓ SD

Mean (Average) [40]:

The process in which the sum of all values divided through the sum of the count of values is known to be the mean value.

$$\overline{G} = \frac{1}{b}\sum_{q=1}^{b} G_q \quad (5)$$

Eq. (5) G indicates the observed value, b represents the number of values and $\overline{G}$ refers to the symbol of the sample mean.

Median [40]:

It is the process in which the middle value in a dataset are organized in ascending order. If the dataset contains 2 values in the middle, then the mean of 2 middle values is regarded as the median of the data.

$$Median = \begin{cases} G\left(\frac{b}{2}\right) & if\ b\ is odd \\ \frac{G\left(\frac{b-1}{2}\right) + G\left(\frac{b+1}{2}\right)}{2} & if\ b is\ even \end{cases} \quad (6)$$

SD

It is a measure of the set of dispersion values or amount of variation. The lower SD [41] denotes the values that tend to be nearer to the mean value, whereas a larger SD denotes the extended values over a larger range. The SD is given in Eq. (7). Here, σ refers to the symbol of SD.

$$\sigma = \sqrt{\frac{1}{b-1}\sum_{q=1}^{b}\left(G_q - \overline{G}\right)^2} \quad (7)$$

The statistical features are indicated SF , and it is defined in Eq. (8).

$$SF = \overline{G} + Median + \sigma \quad (8)$$

4.3.3 Higher-Order Statistical Features

These features are given as follows.

- ✓ Skewness
- ✓ Kurtosis

Skewness [39]

It is a symmetry measure or the lack of symmetry exactly. A data set or distribution is symmetric only if it is similar to the right and left of the centre point. The mathematical expression of skewness is given in Eq. (9).

$$Skewness = \frac{\sum_{q=1}^{b}\left(G_q - \overline{G}\right)^3 / b}{\sigma^3} \quad (9)$$

In Eq. (9), $G_q = G_1\ G_2 ...,\ G_b$ $\overline{G}$ indicates the mean value, σ refers to the SD and b denotes the number of data points. Further, σ indicates the SD, and it is calculated with b the present in the denominator rather than $b-1$ while computing the skewness.

Kurtosis [39]

This is the measure that identifies if the data are light-tailed or heavy-tailed related to the normal distribution. Datasets with less kurtosis tend to provide the lack of outliers or lower tails. Moreover, the datasets with larger kurtosis tend to provide outliers or heavy tails. The formula of kurtosis for univariate data such as $G_1\ G_2\ G_b ...$ is expressed in Eq. (10). The SD is examined through the b value present in the denominator rather than $b-1$ while computing the kurtosis.

$$Kurtosis = \frac{\sum_{q=1}^{b}\left(G_q - \overline{G}\right)^4 / b}{\sigma^4} \quad (10)$$

The higher-order statistical features are indicated HF , and it is given in Eq. (11).

$$HF = Skewness + Kurtosis \quad (11)$$

4.3.4 Entropy

Entropy is known as the average level of surprise, uncertainty, or information inbuilt in the variable's feasible resultant of the data theory. The conception of information entropy is sometimes known as the Shannon entropy. Further, the entropy is given in Eq. (12).

$$Entropy = \frac{-\sum_{q=1}^{b}\left[G_q\right]\log\left[G_q\right]}{length\left(G_q\right)} \quad (12)$$

4.3.5 Improved Mutual Information

An improved mutual information is in Eq. (13).

$$IMI = \frac{J(u,v)}{1/2\left[H(u)+H(v)\right]} \tag{13}$$

In Eq. (13), $H(u)$ and $H(v)$ specifies the conditional entropies and $J(u,v)$ denotes the mutual information.

$$H(v) = -\sum_{uv} Q(u,v)\log\frac{Q(u,v)}{Q(u)} \tag{14}$$

$$H(u) = -\sum_{uv} Q(u,v)\log\frac{Q(v,u)}{Q(v)} \tag{15}$$

4.3.6 Correlation

Correlation is the measure of similarity between the 2 features. The correlation approach is used to determine the relationship among the features. To determine the correlation between 2 random variables, there are 2 basic groups to examine. For a set of variables (Y,Z), the linear correlation coefficient ' $\hat{r}$ ' is determined in Eq. (16).

$$\hat{r} = \frac{\sum\left(Y_j-\bar{Y}_j\right)\left(Z_j-\bar{Z}_j\right)}{\sqrt{\sum\left(Y_j-\bar{Y}_j\right)^2}\sqrt{\sum\left(Z_j-\bar{Z}_j\right)^2}} \tag{16}$$

The overall extracted features are indicated *FE* , and it is given in Eq. (17).

$$FE = RS + SF + HF + Entropy + IMI + \hat{r} \tag{17}$$

4.4 Feature Selection

After extracting the huge set of features *FE* , the feature selection gets carried out, for which an improved Relief algorithm is used. This determines the dimensionality reduction process.

ReliefF determines the features scores on the basis of the feature value difference and class values among the nearby instances. The neighbouring instances group has dissimilar feature values; however, the similar class value would then reduce the score of relief value.

An improved ReliefF: The step of an improved ReliefF is given as follows.

Set all weights $\hat{w}\left[\hat{k}\right]=0$;

For $h=1$ to number of runs

Select 2 features randomly

Cluster with K-means on selected features

For $t=1$ to T

Randomly select a tuple x from dataset

For $p=1$ to $\hat{n}$

$$\hat{w}_p = \hat{w}_p - diff\left(x_p, nearest\,Hit(x)_p\right)^2 + diff\left(x_p, nearest\,Miss(x)_p\right)^2$$

End

End

Calculate the Harmonic mean $\hat{w}_p$ by the number of runs

end

The selected Relief feature is denoted as *FS* .

5. Classification Via Ensemble Classifies: RF, NN, RNN AND K-NN

These selected features are given to the classification phase, in which the ensemble model includes NN, RNN, RF, and K-NN models. Here, the output of NN, RNN, and RF is given as the input of k-NN. For making the system more precise in HD prediction, the weights of NN and RNN are optimally tuned by an improved SSO Algorithm.

5.1 RF

The selected features *FS* are provided to RF classifier as its input. Further, the RF [32] is an approach that combines multiple tree predictors and each tree depending on the value of a randomly selected vector between all trees in the forest. The quality of characterization is improved by decreasing the overriding problems as the RF is associated with one DT. The characterization of the parameter assessment is more important in the RF technique. The RF algorithm consists of certain primary constraints like the parting basis, the extreme tree profundity, the number of attributes to be regarded during the best split searching, and a number of forest trees as well. The Gini debasement is used as the parting method, and it is determined in Eq. (18).

$$Gini_{\hat{M}} = 1 - \sum_{\hat{h}=1}^{\hat{H}} \hat{U}^2\left(\hat{G}_{\hat{p}}\right) \tag{18}$$

Eq. (18) $\hat{U}\left(\hat{G}_{\hat{p}}\right)$ denotes the subset of the element $\hat{G}_{\hat{p}}$ class present in the tree node $\hat{M}$. Moreover, the splitting criterion in the case of binary classification is expressed in Eq. (19).

$$Gini_{\hat{M}}^{split} = \hat{C}_1/\hat{C}\; Gini_{\hat{M}1} + \hat{C}_2/\hat{C}\; Gini_{\hat{M}2} \rightarrow \min \tag{19}$$

In Eq. (19), $\hat{C}_1$ and $\hat{C}_2$ denotes the object count in the nodes to both descendants-nodes of binary tree $\hat{M}_1$ and $\hat{M}_2$ $\hat{C}$ refers to the objects count in the present node. The classified outcomes of the RF classifier are denoted as CL_{RF} .

5.2 Optimized NN

NN [33] is a set of algorithms that endeavoured for recognizing the interaction in a data group via a process that decreases the technique the human brain works. The selected features *FS* are subjected to the NN as its input, and it is given in Eq. (20). Here, $\hat{z}$ denotes the total number of extracted features.

$$FS^2 = \left\{FS_1^2, FS_2^2, ...FS_{\hat{z}}^2\right\} \tag{20}$$

The NN framework consists of hidden, output, as well as input layers. Moreover, the hidden layer output $\hat{q}^{(\hat{x})}$ is determined in Eq. (21). Here, $\bar{c}$ and $\hat{E}$ represents the neurons in the input layer as well as a hidden layer, respectively, $\overline{A}$ determines the activation function, $\hat{g}_{\hat{v}}$ refers to the input neurons count, $w_{\left(\hat{d}\hat{E}\right)}^{(\hat{y})}$ denotes the bias weight with $\hat{v}^{th}$ hidden neuron, and $w_{\left(\bar{c}\hat{E}\right)}^{(\hat{y})}$ denotes the weight between the $\bar{c}^{th}$ input neuron to $\hat{E}^{th}$ hidden neuron. Moreover, the network output $\hat{R}_{\hat{j}}$ is determined in Eq. (22), which $w_{\left(\hat{E}\hat{j}\right)}^{\left(\hat{R}\right)}$ specifies the weight of the $\hat{E}^{th}$ hidden layers to $\hat{j}^{th}$ the output layer, *Out* specifies the output neurons, $w_{\left(\hat{d}\hat{j}\right)}^{\left(\hat{R}\right)}$ represents the output bias weight of $\hat{j}^{th}$ the output layer and n portrays the hidden neurons count. Therefore, the error (Er^*) attained in both actual and predicted values should be low, as given in Eq. (23). Where $\hat{l}_{\hat{c}}$ specifies the output neuron count $\tilde{R}_{\hat{j}}$ and $\hat{R}_{\hat{j}}$ determines the predicted and actual output, respectively.

$$\hat{q}^{(\hat{x})} = \overline{A}\left(w_{\left(\hat{d}\hat{E}\right)}^{(\hat{y})} + \sum_{\bar{c}=1}^{\hat{l}_{\hat{E}}} w_{\left(\bar{c}\hat{E}\right)}^{(\hat{y})} IN \right) \tag{21}$$

$$\hat{R}_{\hat{j}} = \overline{A}\left(w_{\left(\hat{d}\hat{j}\right)}^{\left(\hat{R}\right)} + \sum_{\hat{E}=1}^{\hat{l}_{\hat{v}}} w_{\left(\hat{E}\hat{j}\right)}^{\left(\hat{R}\right)} \hat{q}^{(\hat{x})} \right) \tag{22}$$

$$Er^* = \underset{\left\{w_{\left(\hat{d}\hat{E}\right)}^{(\hat{y})}, w_{\left(\bar{c}\hat{E}\right)}^{(\hat{y})}, w_{\left(\hat{d}\hat{j}\right)}^{\left(\hat{R}\right)}, w_{\left(\hat{E}\hat{j}\right)}^{\left(\hat{R}\right)}\right\}}{\arg\min} \sum_{\hat{j}=1}^{\hat{l}_{\hat{c}}} \left| \tilde{R}_{\hat{j}} - \hat{R}_{\hat{j}} \right| \tag{23}$$

The classified outcome of the optimized NN classifier is denoted as CL_{NN}.

5.2.1 Optimized RNN

The *FS* is given to RNN classifier as its input. Moreover, the RNN is a special type of NN that is used for the purpose of recognition, prediction, and classification. The RNN [43] includes the hidden layer, input layer, and output layer. In RNN, each layer includes neurons. The RNN is trained with the BPTT model using Bayesian regulation that depends upon the forward and backward pass.

The input units $\hat{I}$ at a time $\hat{T}$ in the input layer with the set of vector as $\left\{..., \hat{J}_{\hat{T}-1}, \hat{J}_{\hat{T}}, \hat{J}_{\hat{T}+1}, ...\right\}$, in which $\hat{J}_{\hat{T}} = \left(\hat{J}_1, \hat{J}_2, ...\hat{J}_{\hat{I}}\right)$. Each input unit in the fully convolutional RNN is associated in the hidden layer with each hidden unit, and it is determined as the weight matrix W_{Ih}. In addition, the hidden layer $\tilde{H}$ consists of hidden units $\hat{D}_{\hat{T}} = \left(\hat{D}_1, \hat{D}_2, ...\hat{D}_{\tilde{H}}\right)$ that are associated with each other via the recurrent links in the matrix W_{hh}. Moreover, the hidden layer of RNN is determined in Eq. (24).

$$\hat{V}_{\hat{L}} = \hat{A}_{\hat{b}}\left(W_{Ih}\hat{J}_{\hat{L}} + W_{hh}\hat{V}_{\hat{L}-1} + \hat{B}_{\hat{b}}\right) \tag{24}$$

In Eq. (24), $\hat{B}_{\hat{b}}$ and $\hat{A}_{\hat{b}}(.)$ specifies the hidden units bias vector and the activation function, correspondingly. Moreover, the hidden units are associated via the weight matrix W_{ho} to the output layer. Here, the output layer consists of $\hat{K}$ units as $\hat{X}_{\hat{L}} = \left(\hat{X}_1, \hat{X}_2, ...\hat{X}_{\tilde{C}}\right)$, and it is given in Eq. (25).

$$\hat{X}_{\hat{L}} = \hat{A}_o\left(W_{ho}\hat{V}_{\hat{L}} + \hat{B}_o\right) \tag{25}$$

In Eq. (25), $\hat{B}_o$ and $\hat{A}_o(.)$ indicates the output units bias vector as well as the activation function, correspondingly. In RNN, the weight matrix W_3 connects among the input and hidden layer, W_1 with hidden-to-hidden recurrent connections that is it exists in the hidden layer nodes, as well as W_2 connected to the hidden layer and the output layer. All these parameters (W_3, W_1, W_2) are jointed across time. Further, the stages of the RNN are given below:

- ✓ The weight matrices are given as W_3, W_1, W_2, and the bias function $\hat{P}_1, \hat{P}_2$ is initialized with 0.
- ✓ The RNN forward pass is determined in Eq. (26) to Eq. (29), correspondingly.

$$\tilde{Q}(\hat{L}) = \hat{P}_1 + W_3.\hat{F}(\hat{L}) + W_1.\hat{V}(\hat{L}-1) \tag{26}$$

$$\hat{V}(\hat{L}) = \tanh.(\tilde{Q}(\hat{L})) \tag{27}$$

$$\tilde{K}(\hat{L}) = \hat{P}_2 + \hat{V}(\hat{L}).W_2 \tag{28}$$

$$\hat{Z}(\hat{L}) = soft\max(\tilde{K}(\hat{L})) \tag{29}$$

- ✓ The loss function of RNN is determined in Eq. (30). Where $\tilde{N}$ indicates the number of class labels, $\tilde{E}$ specifies the binary indicator which finds the class label *cls* s split correctly for the observation *obs*.

$$Loss = -\sum_{\hat{P}_2=1}^{\tilde{N}} \tilde{E}_{obs,cls}.\log(\tilde{M}_{obs,cls}) \tag{30}$$

- ✓ The gradients are computed via the back-propagation. The RNN output is denoted as CL_{RNN}.

5.3 KNN

The classified outcome of RF, NN, and NN (CL_{RF} CL_{NN}, and CL_{RNN}) *FS* are subjected to K-NN classifier as its input. The K-NN [34] classifiers depend on learning

through its similarity as it compares the specified test tuples with training tuples that are similar to them. Moreover, each tuple specifies the point in a $\tilde{n}$ -dimensional space. In addition, the k-NN classifier explored the k training tuples on providing unfamiliar tuples that reside near to the unfamiliar one in pattern space. Further, the k-NN of the unfamiliar tuples is known to be k training tuples[35]. Closeness is defined with respect to distance metrics like ED. The ED between the points $KN_1 = (\tilde{r}_{11}, \tilde{r}_{12}...\tilde{r}_{1\tilde{n}})$ $KN_2 = (\tilde{r}_{21}, \tilde{r}_{22}...\tilde{r}_{2\tilde{n}})$ and or the two tuples is computed in Eq. (31).

$$dis(KN_1, KN_2) = \sqrt{\frac{\sum_{\tilde{y}=1}^{\tilde{Y}} (\tilde{r}_{1\tilde{y}} - \tilde{r}_{2\tilde{y}})^2}{\tilde{Y}}} \quad (31)$$

The classified output of k-NN is indicated by CL_{K-NN}.

6. Weight Optimization of NN and RNN Via Self Improved Shark Smell Optimization With Gaussmap Estimation and Cycle Crossover Operation

6.1 Objective Function and Solution Encoding

The weights of NN and RNN are tuned optimally via an adopted SISSOOL method. Fig. 3 illustrates the input solution to the adopted SISSOOL model. Here, the entire number of weights in RNN is indicated as N , and the entire number of weights in NN is denoted as $\tilde{J}$. The objective function Obj of the implemented scheme is determined in Eq. (32). Here $Loss$ is depicted the K-NN loss function.

$$Obj = \min(Loss) \quad (32)$$

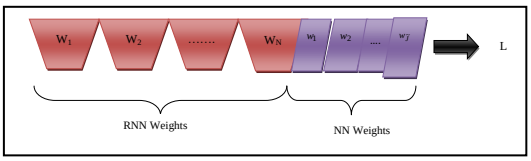


Fig. 2 Solution Encoding

6.2 Proposed SISSGECO model

Although, SSO [26] is implemented on the basis of the shark's ability for hunting with a high smell sense for solving real-world engineering problems. However, it suffers in preserving the convergence speed and convergence rate. To overcome this, the SISSGECO model is proposed. Generally, the self-enhancement is proved to be capable in the existing optimization models [27] [28] [29] [30] [31]. SSO includes 4 primary phases such as initialization, forward movement, rotational movement, and position update.

6.2.1 Initialization

For modelling SSO, the initial solution population are arbitrarily created in the searching space. Each solution denotes a particle of odour which is a feasible shark position at the start of the searching procedure. The initial solution vector is determined as in Eq. (33) and (34), in which the $U_g^1 = g^{th}$ initial populace vector position and n signifies the size of the populace.

$$U^1 = \lfloor U_1^1, U_2^1, ...U_s^1 \rfloor \quad (33)$$

The associated optimization issue is determined in Eq. (34), wherein the $U_{g,a}^1 = a^{th}$ dimension of g^{th} the position of the shark and c signify the decision variable count.

$$U_g^1 = \lfloor U_{g,1}^1, U_{g,2}^1, ...U_{g,c}^1 \rfloor \quad (34)$$

6.2.2 Forward Movement

When blood is mixed with water, the Shark in each position produces strong odour particles with a "velocity B " to come nearer to the target (prey). Therefore, the initial velocity vector A is expressed based on its positions as per Eq. (35), and each B comprises a dimensional element as specified in Eq. (36).

$$B_g^1 = \lfloor B_1^1, B_2^1, ...B_s^1 \rfloor \quad (35)$$

$$B_g^1 = \lfloor B_{g,1}^1, B_{g,2}^1, ...B_{g,c}^1 \rfloor \quad (36)$$

Therefore, the velocity in every dimension is evaluated as in Eq. (37), where $z = 1,2,....X_{\max}$, $\left.\frac{\partial(OB)}{\partial \chi a}\right|_{\chi_{g,a}^z}$ point out derivative OB at the position $\chi_{g,a}^z$, z symbolizes stage count, $X_{\max}$ indicates stage count for shark's forwarding movement, and $\Re 1$ symbolizes arbitrary integer among (0, 1).

$$B_{g,a}^z = \eta_z . \Re 1 . \left.\frac{\partial(OB)}{\partial \chi a}\right|_{\chi_{g,a}^z} \quad (37)$$

The raise i B s depicted by raise in the intensity of the odour. In each phase $B_{g,a}^z$, the velocity limiter is employed as exposed in Eq. (38), wherein ψ_z denotes the inertia coefficient in (0, 1) and $\Re 2$ signifies arbitrary integer amongst (0, 1), β which denotes the velocity limit ratio for z .

$$\left|B_{g,a}^z\right| = \left[\left|\eta_z . \Re 1 . \left.\frac{\partial(obj)}{\partial \chi a}\right|_{\chi_{g,a}^z} + \psi_z . \Re 2 . A_{g,a}^{z-1}\right|, \left|\beta_z . A_{g,a}^{z-1}\right|\right] \quad (38)$$

The new shark position is depicted owing to its preceding velocity and position as per Eq. (39), which $\Delta \hat{t}_z$ points out the time interval of the stage z .

$$L_g^{z+1} = M_g^z + B_g^z . \Delta \hat{t}_z \quad (39)$$

6.2.3 Rotational Movement

The shark makes a rotational movement for discovering the stronger odour particle. This process is called a local search as defined in (40), where $e = 1,2...E$ and $\Re 3$ points out arbitrary integers amongst (0, 1).

$$F_g^{z+1,e} = L_g^{z+1} + \Re 3.L_g^{z+1} \quad (40)$$

Conventionally, the parameters $\Re 1$ $\Re 2$ $\Re 3$ are randomly generated. As per the proposed SISSGECO model, the gauss map is used for this parameter estimation as per Eq. (41).

$$R_{f+1} = \begin{cases} 0; & R_f = 0 \\ \dfrac{1}{R_f \bmod(1)}; & Otherwise \end{cases} \quad (41)$$

6.2.4 Particle Position Update

The searching path of the shark continues with rotational movement as it comes nearer to the strong odour particle that is exposed in Eq. (42), which here F_g^{z+1} represents the subsequent position of the shark with a higher OB value.

$$M_g^{z+1} = \arg\max\left\{OB\left(L_g^{z+1}\right), OB\left(F_g^{z+1,g}\right), ..., OB\left(F_g^{z+1,E}\right)\right\} \quad (42)$$

In addition, as per the proposed logic, and arithmetic crossover operation is performed.

As per the proposed model, the cyclic crossover operation is performed.

6.2.5 Cycle Crossover

A gene from one parent is replicated into a child in cycle crossover, but the offspring must acquire the location of the other parent.

Steps for Cycle Crossover

- As indicated in Fig. 5a, begin a cycle from the 1st gene of the 1st parent to the 1st gene of the 2nd parent.
- Choose the gene in the 1st position of the 2nd parent and go to the 1st parent's corresponding gene.
- Vertically travel from the 1st parent's present gene to the 2nd parent's gene.
- Examine if the 2nd parent's gene is identical to the 1st parent's 1st gene. If true, proceed to step 6; otherwise, proceed to step v.
- Go to step 3 and find the gene in the 1st parent that corresponds to the present gene in the 2nd parent.
- To get the 2nd offspring, repeat the previous processes.
- Move the genes from the 1st parent's cycle to the 1st offspring's equivalent places, as indicated in Fig. 5b.
- Copy the genes from the 2nd parent's cycle to the 2nd offspring's equivalent places, as indicated in Fig. 5b.
- Copy the remaining genes from the 2nd parent to the first offspring's appropriate places, as indicated in Fig. 5c.
- As illustrated in Fig. 5c, copy the remaining genes of the 1st parent to their respective places in the 2nd offspring.
- The final matching offspring is formed by the present gene sequence in each of the children.

Fig. 3 illustrates the Cycle crossover (a) formation of the cycle, (b) formation initial stage (c) completion of offspring.

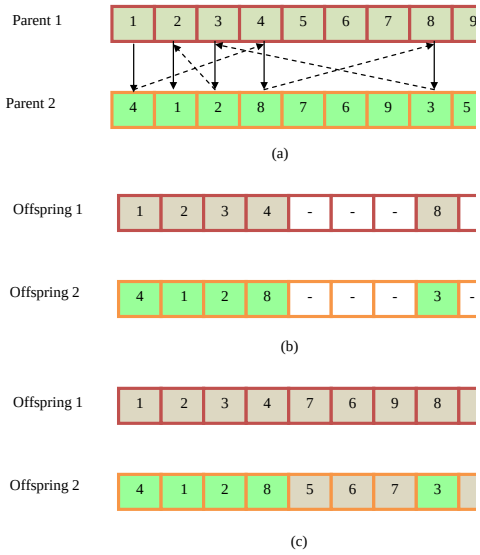


Fig. 3 Cycle crossover (a) formation of cycle(b) formation initial stage (c) completion of offsprings

The Pseudocode of the presented SISSGECO scheme is specified in Algorithm 2.

Algorithm 2: Adopted SISSGECO scheme
Start
Initialization
Apply OBL and generate opposite solutions
Assign constraints, s ψ_z $z_{\max}$ η_z and $z = 1,2,....X_{\max}$
Generate primary populace with all individuals
initializing z =1
For z =1: $X_{\max}$
Forward movement
Compute every element of $B_{g,a}$
The shark position is given as per forwarding movement in Eq. (38).
Rotational movement
Attain novel shark position as per the rotational movement $F_g^{z+1,e}$
The gauss map is used for this parameter estimation as per the proposed logic in Eq. (41).
Choose subsequent shark position depending upon 2 movements
Carry out cycle crossover
End for z
Fix $z = z + 1$
Choose the best shark position with a higher OB value
End

7. Results and Discussions

7.1 Simulation Procedure

The adopted ensemble classifier+ SISSGECO based HD prediction scheme was implemented in PYTHON, and their results were confirmed. Furthermore, the outcomes of the

presented ensemble classifier+ SISSGECO scheme was computed over the conventional schemes such as ensemble classifier +SSO [26], ensemble classifier + BOA [45], ensemble classifier +CMBO [37], ensemble classifier +PRO [36], and ensemble classifier +social SO [35], correspondingly. The dataset was collected from [57], and the respective 4 datasets were "dataset-1 (Cleveland), dataset-2 (Hungary), dataset-3 (Switzerland), and dataset-4 (VA Long Beach)," respectively. "This database contains 76 attributes, but all published experiments refer to using a subset of 14 of them. In particular, the Cleveland database is the only one that has been used by ML researchers to this date. The goal field refers to the presence of heart disease in the patient. It is integer-valued from 0 (no presence) to 4. Experiments with the Cleveland database have concentrated on simply attempting to distinguish presence (values 1, 2, 3, 4) from absence (value 0)". In addition, the performance was computed by altering the learning percentage from 40%, 50%, 60%, 70%, and 80% for different metrics, including precision, sensitivity, accuracy, specificity, FNR, F-measure, MCC, FPR, and NPV, respectively.

7.2 Performance Analysis

The performance analysis of the adopted ensemble classifier+ SISSGECO scheme is computed over the existing schemes like ensemble classifier +SSO, ensemble classifier + BOA, ensemble classifier +CMBO, ensemble classifier +PRO, and ensemble classifier +social SO, correspondingly in terms of certain metrics and it is given in Fig. 4 to Fig. 15 for datasets 1,2, 3, and 4. Moreover, the positive measures such as precision, sensitivity, accuracy, and specificity are illustrated in Fig. 4 to Fig. 7. Similarly, the adopted ensemble classifier+ SISSGECO scheme holds maximum accuracy (~0.9) for a learning percentage of 50% than the other existing schemes in Fig. 4(b) for dataset 1. Likewise, the scheme attains higher specificity (~0.97) for a learning percentage of 70% than the other existing schemes for dataset 3 in Fig. 6(d). Further, the proposed ensemble classifier+ SISSGECO scheme has shown a higher precision value with better performance at a learning percentage of 40% than at a learning percentage of 60% for dataset 4 in Fig. 7(c). This analysis outcome has proven the impact of an ensemble classifier that gets trained with the appropriate features. Further, as both the NN and RNN weights are tuned optimally, the proposed ensemble classifier+ SISSGECO technique paved the way for better results in the HD prediction model with lower error.

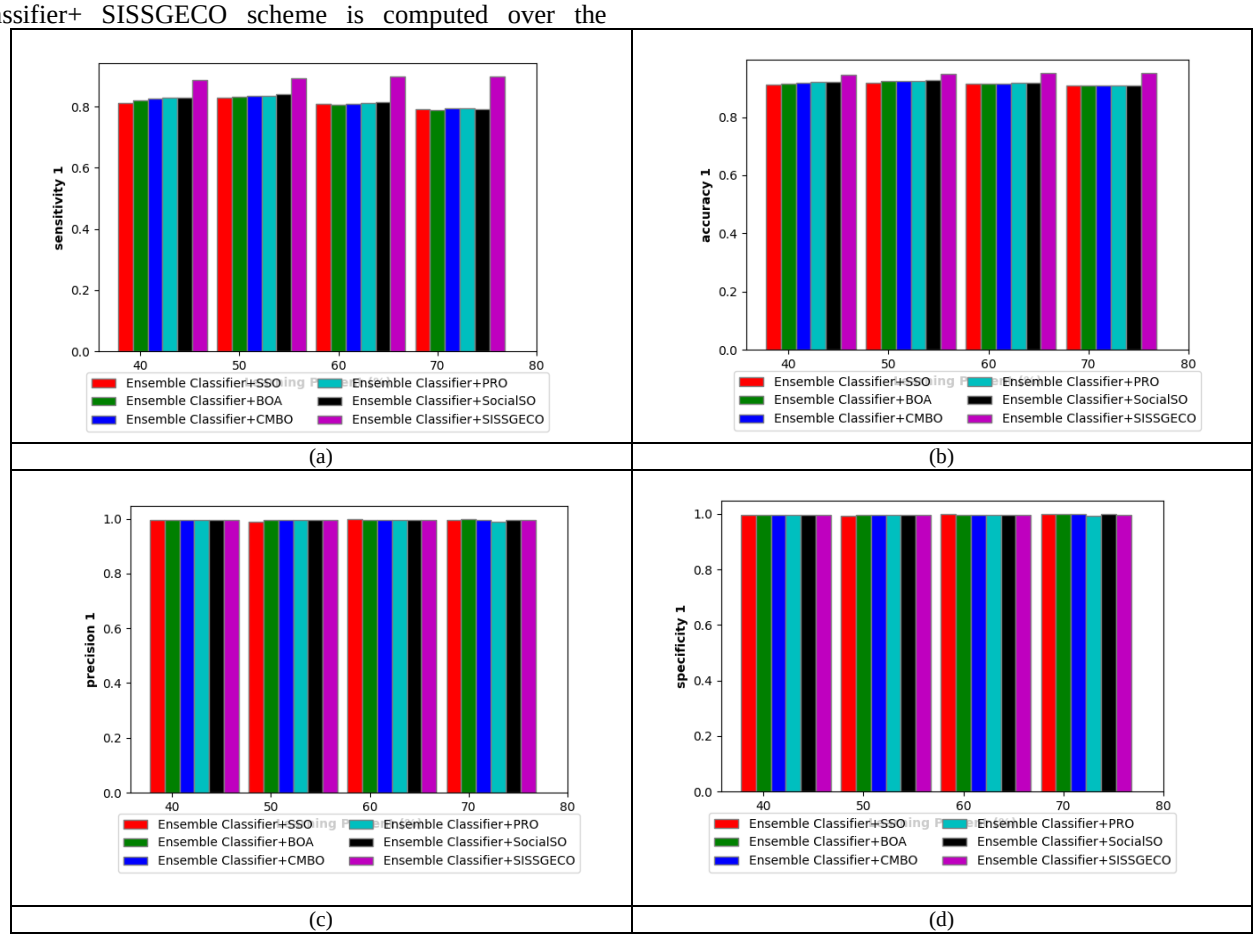


Fig. 4 Performance analysis of the developed scheme to the traditional approaches for (a) sensitivity, (b) accuracy, (c) precision (d) specificity for dataset 1

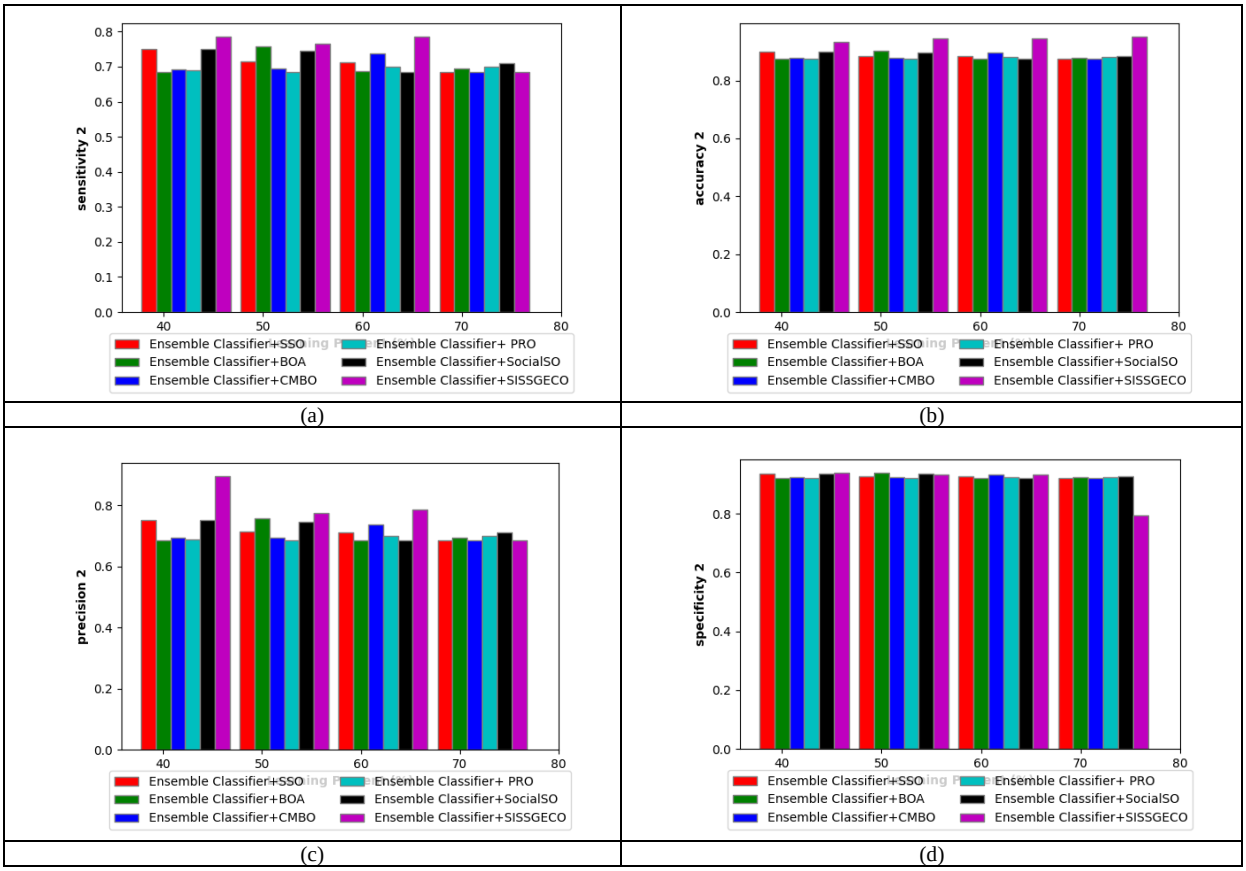


Fig. 5 Performance analysis of the developed scheme to the traditional approaches for (a) sensitivity, (b) accuracy, (c) precision (d) specificity for dataset 2

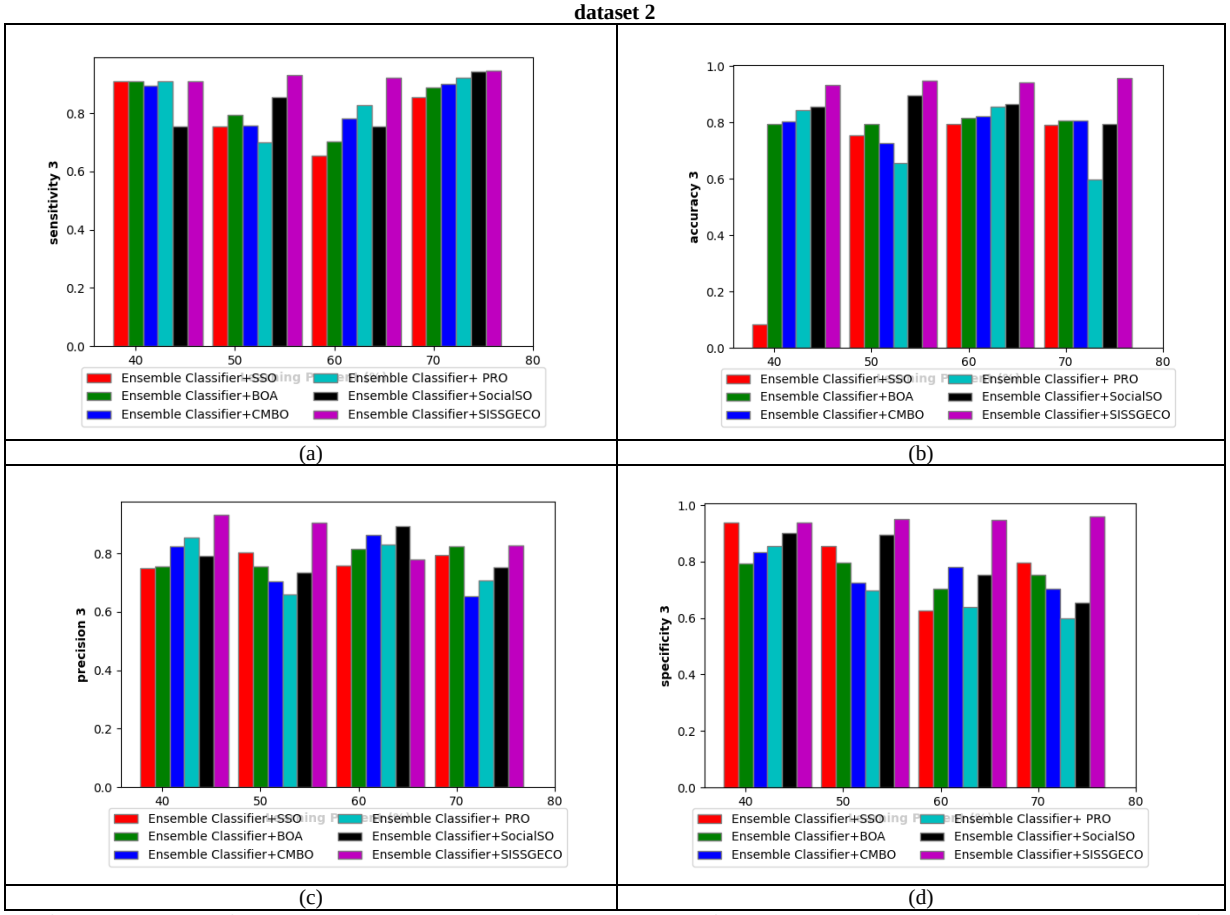


Fig. 6 Performance analysis of the developed scheme to the traditional approaches for (a) sensitivity, (b) accuracy, (c) precision (d) specificity for dataset 3

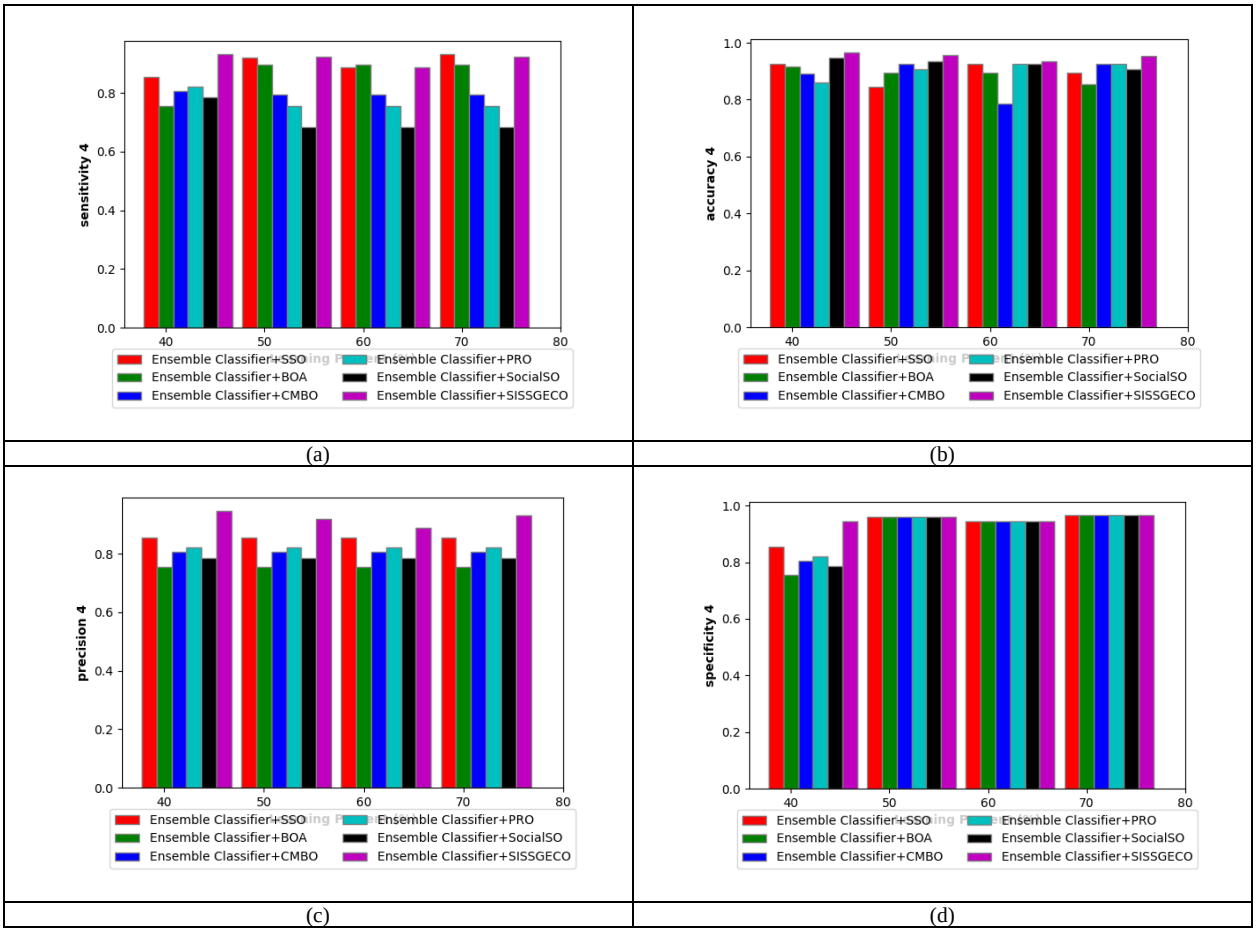


Fig. 7 Performance analysis of the developed scheme to the traditional approaches for (a) sensitivity, (b) accuracy, (c) precision (d) specificity for dataset 4

The negative metrics like FPR, and FNR of the adopted ensemble classifier+ SISSGECO scheme to the traditional schemes like ensemble classifier +SSO, ensemble classifier + BOA, ensemble classifier +CMBO, ensemble classifier +PRO, and ensemble classifier +social SO, respectively for dataset 1, 2, 3, and 4 are represented in Fig. 8 to Fig. 11. In addition, the variations in learning percentage demonstrate the difference in the outcomes. This performance has proven that the adopted work has converged with the objective (minimization of error). In addition, the proposed ensemble classifier+ SISSGECO model proves the less FPR value (~0.15) as a better performance than the conventional models at a learning percentage of 60% for dataset 3 in Fig. 1(b). Less FNR value of the proposed ensemble classifier+ SISSGECO model has proven that the model is less prone to error that direct to precise outcomes at a learning percentage of 40% in Fig. 8(a).

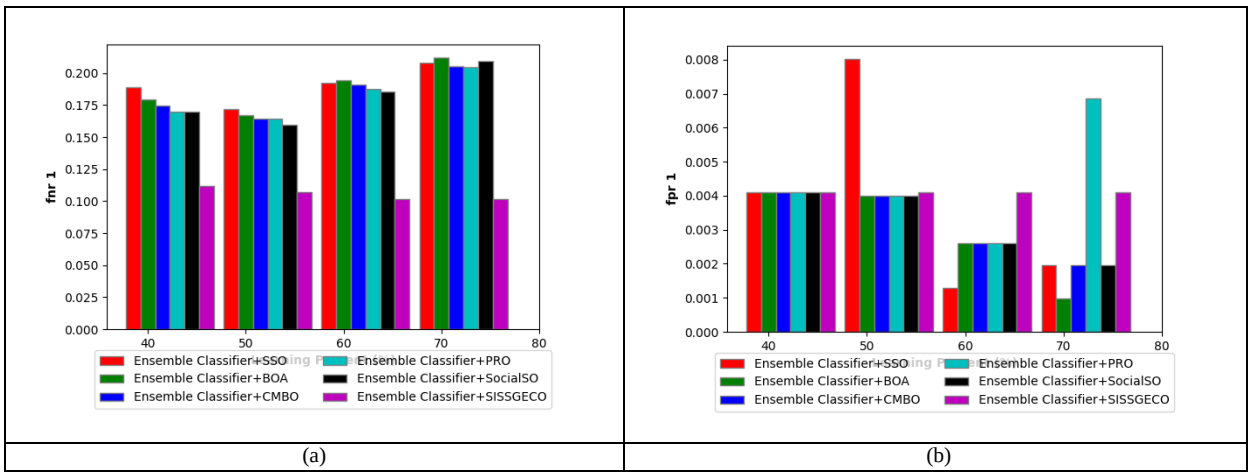
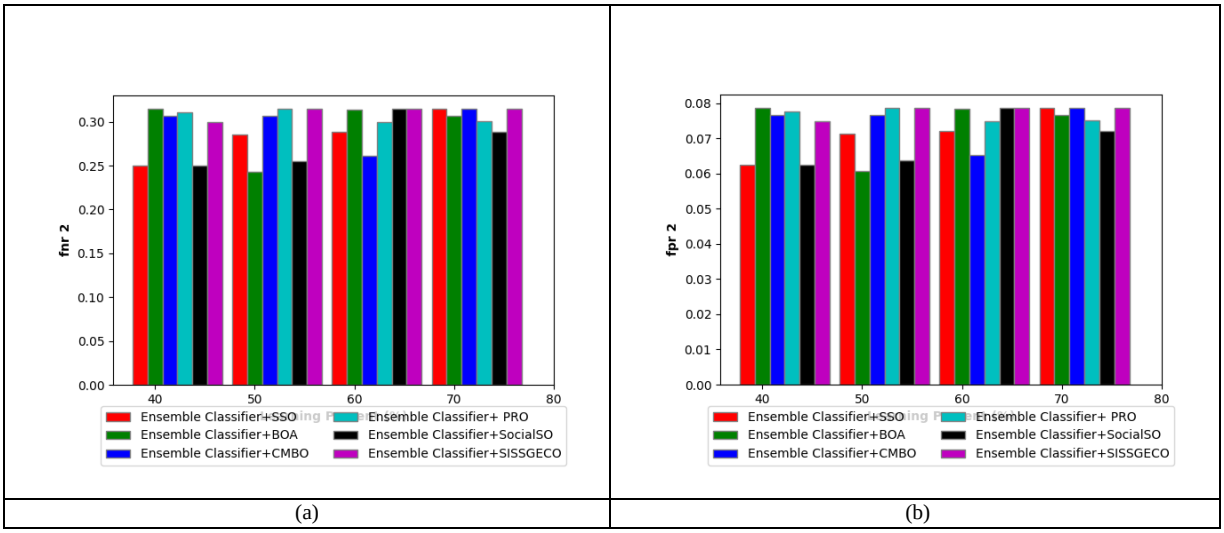
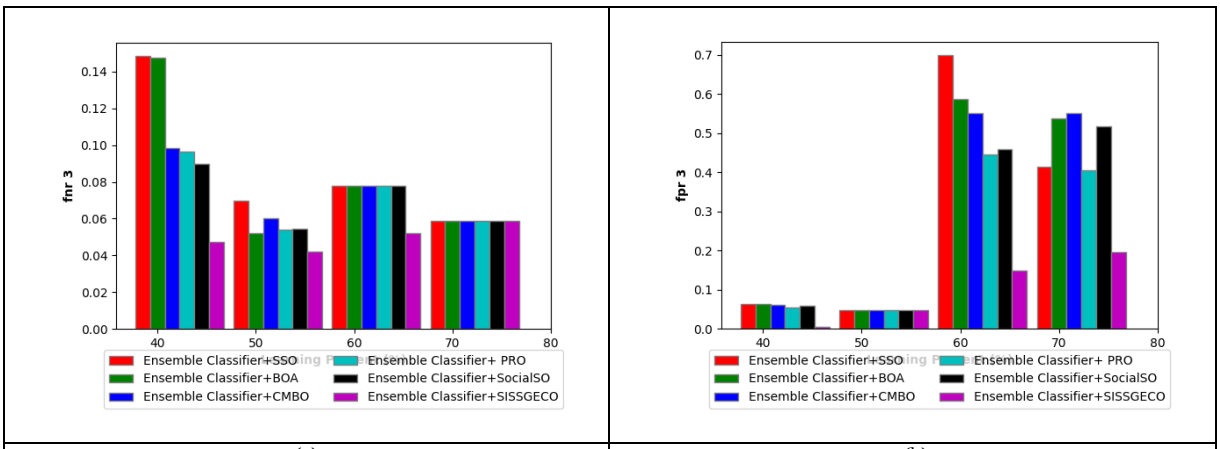


Fig. 8 Performance analysis of the developed scheme to the traditional approaches for (a) FNR (b) FPR for dataset 1



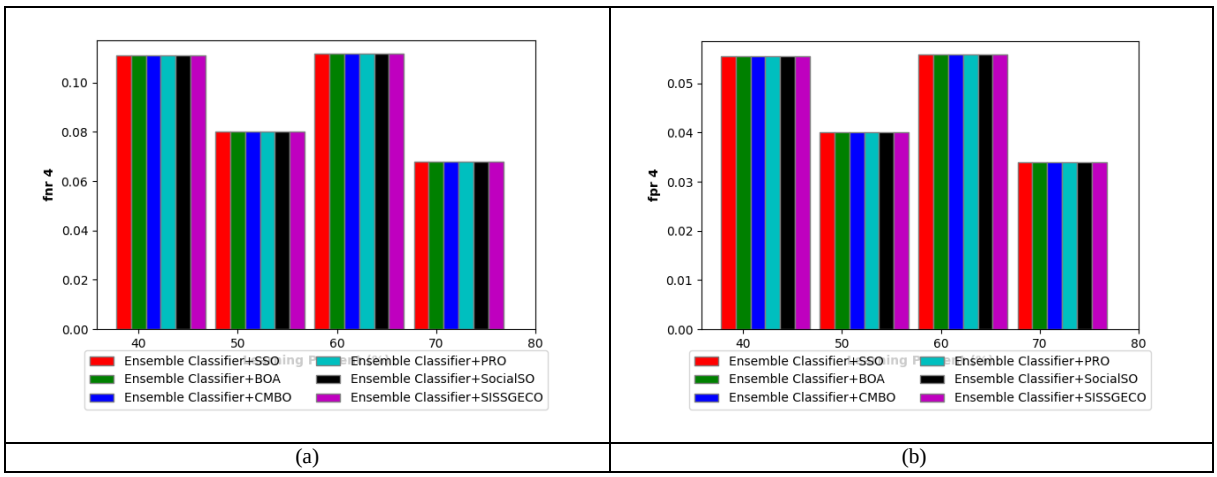
(a) (b)

Fig. 9 Performance analysis of the developed scheme to the traditional approaches for (a) FNR (b) FPR for dataset 2



(a) (b)

Fig. 10 Performance analysis of the developed scheme to the traditional approaches for (a) FNR (b) FPR for dataset 3



(a) (b)

Fig. 11 Performance analysis of the developed scheme to the traditional approaches for (a) FNR (b) FPR for dataset 4

Fig. 5 represents the other metrics analysis like MCC, NPV, and F-measure of the proposed ensemble classifier+ SISSGECO model over other conventional schemes. Similarly, the F-measure of the adopted ensemble classifier+ SISSGECO model for a learning percentage of 60% in Fig. 12(b) is superior to other traditional for dataset 1. Likewise, the adopted ensemble classifier+ SISSGECO model attains maximum NPV (~0.92) for a learning percentage of 50% than other extant schemes for dataset 4 in Fig. 15(a). From the graph, it is clearly shown that the MCC of the adopted ensemble classifier+ SISSGECO model attains a higher value for learning percentage of 50%; however, the compared existing models attain lower values for dataset 2 as per Fig. 13(c). Therefore, the performance of the presented ensemble classifier+ SISSGECO model has shown its improvement over other traditional models.

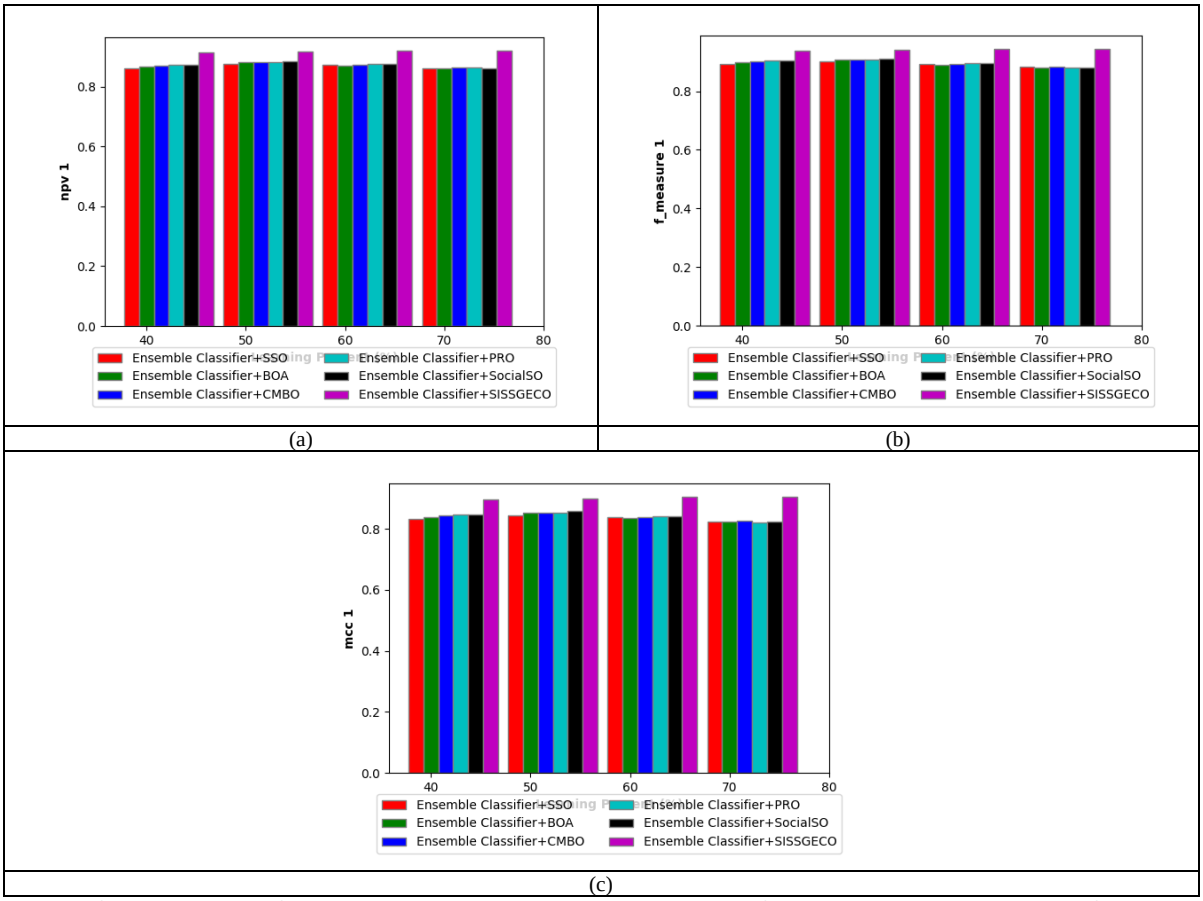


Fig. 12 Performance analysis of the developed scheme to the traditional approaches for (a) NPV (b) F-measure (c) MCC for dataset 1

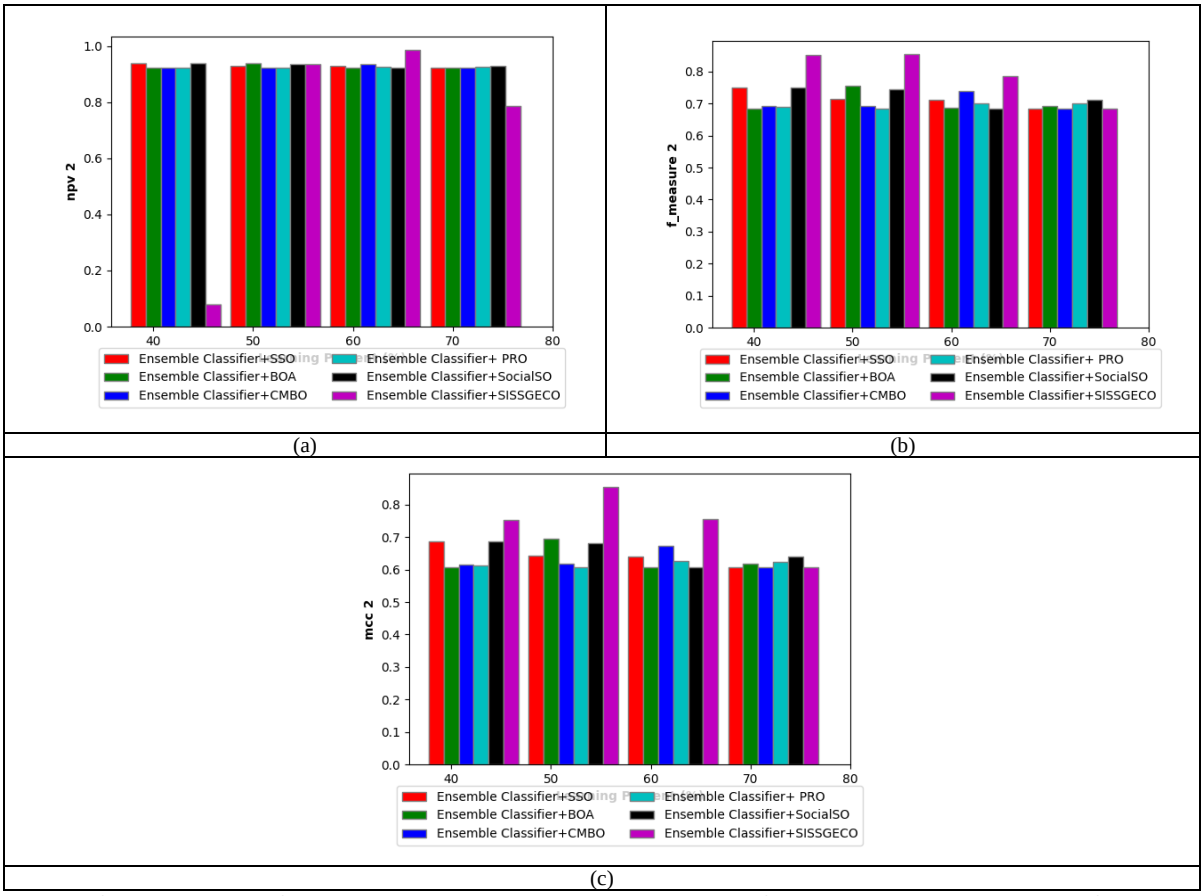


Fig. 13 Performance analysis of the developed scheme to the traditional approaches for (a) NPV (b) F-measure (c) MCC for dataset 2

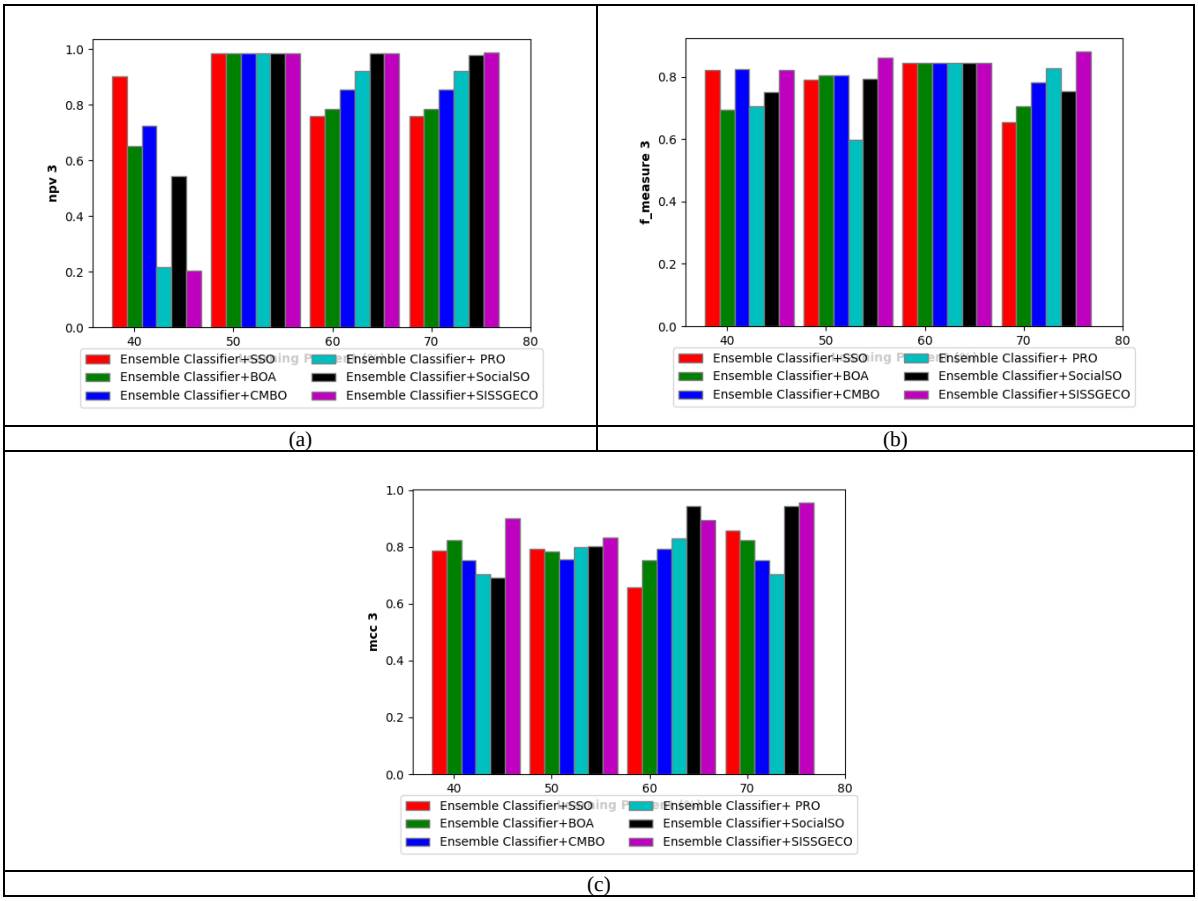


Fig. 14 Performance analysis of the developed scheme to the traditional approaches for (a) NPV (b) F-measure (c) MCC for data

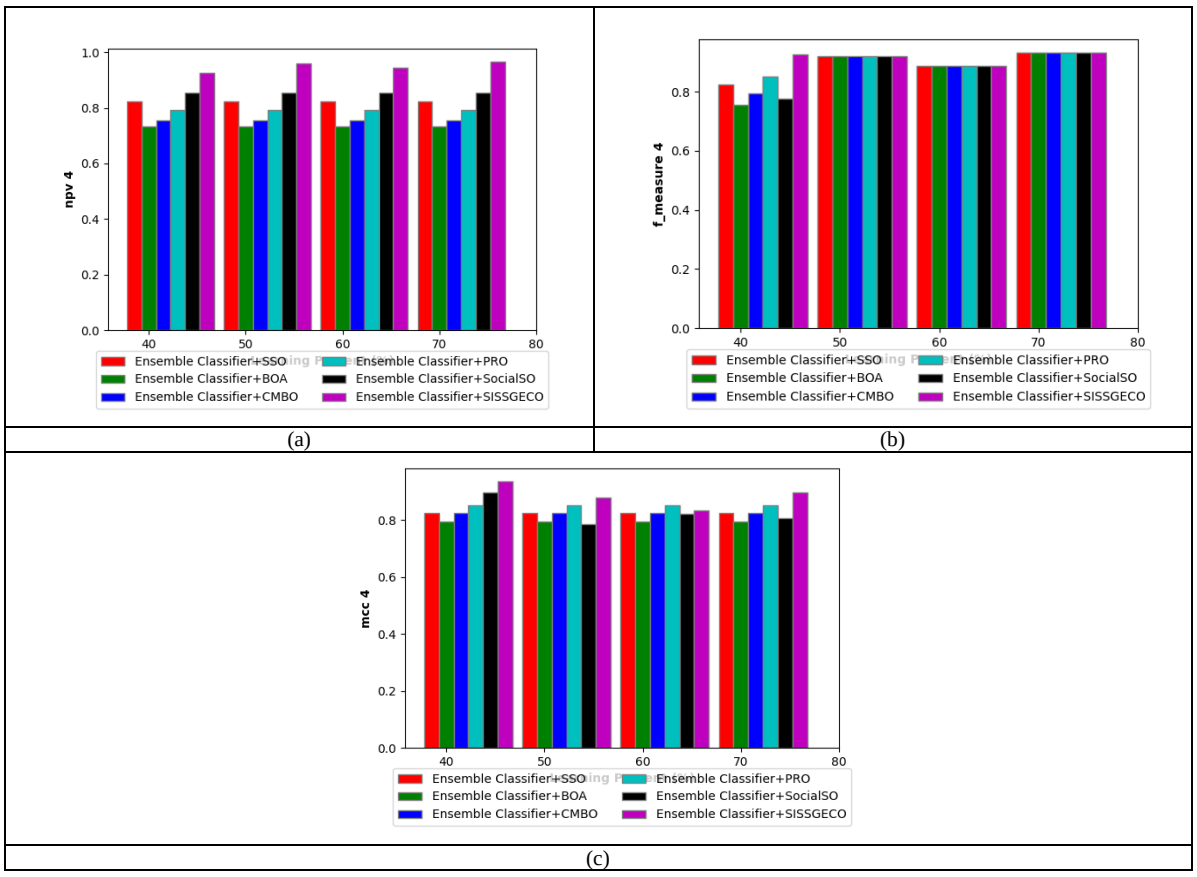


Fig. 15 Performance analysis of the developed scheme to the traditional approaches for (a) NPV (b) F-measure (c) MCC for dataset 4

7.3 Analysis on Classifier

Table II to Table VI describes the performance of classifier analysis of the adopted work to extant models for datasets 1, 2, 3, and 4, respectively. Moreover, the performance of the presented ensemble classifier+ SISSGECO scheme is compared over existing schemes such as DBN, NB, SVM, CNN, and ensemble classifier+ HB-EA model, respectively, for all datasets based on different measures. From the table, the presented ensemble classifier+ SISSGECO scheme has held higher positive values and minimal negative values for all metrics than other extant schemes. In Table II, the adopted scheme provides a larger specificity value (0.995) to extant approaches, including DBN, NB, SVM, CNN, and ensemble classifier+ HB-EA model, correspondingly for dataset 1. Further, the accuracy value of the presented scheme in dataset 4 is higher than in dataset 2. Likewise, the NPV of the suggested approach for dataset 4 is better than the traditional models. The presented scheme has shown minimal FPR values (~0.0042) than the extant approaches for dataset 3. Thus, the proposed work has attained best outcomes than the extant approaches.

Table 2. Analysis of the classifier of adopted and traditional schemes for dataset-1

Metrics	DBN [44]	NB [47]	SVM [46]	CNN [48]	Ensemble classifier + HB-EA [38]	Proposed Ensemble classifier + SISSGECO model
FPR	0.214592	0.184549	0.198856	0.254649	0.195972	0.004098
sensitivity	0.356223	0.446352	0.403433	0.236052	0.941041	0.88835
MCC	0.141631	0.261803	0.204578	-0.0186	0.842673	0.896096
precision	0.356223	0.446352	0.403433	0.236052	0.934367	0.994565
F-Measure	0.356223	0.446352	0.403433	0.236052	0.937692	0.938462
specificity	0.785408	0.815451	0.801144	0.745351	0.804028	0.995902
NPV	0.785408	0.815451	0.801144	0.745351	0.880631	0.913534
Accuracy	0.678112	0.723176	0.701717	0.618026	0.916202	0.946667
FNR	0.643777	0.553648	0.596567	0.763948	0.058959	0.11165

Table 3. Analysis of the classifier of adopted and traditional schemes for dataset-2

Metrics	DBN [44]	NB [47]	SVM [46]	CNN [48]	Ensemble classifier + HB-EA [38]	Proposed Ensemble classifier + SISSGECO model
FPR	0.194444	0.208333	0.263889	0.402778	0.119597	0.074938
sensitivity	0.734177	0.716981	0.650307	0.49711	0.940178	0.78524
MCC	0.456203	0.427113	0.317579	0.075511	0.822647	0.75214
precision	0.453125	0.431818	0.358108	0.228723	0.916945	0.895214
F-Measure	0.560386	0.539007	0.461874	0.313297	0.924281	0.851436
specificity	0.805556	0.791667	0.736111	0.597222	0.880403	0.938851
NPV	0.932476	0.926829	0.902896	0.831721	0.930234	0.078521
Accuracy	0.792711	0.778157	0.720272	0.577828	0.911152	0.932541
FNR	0.265823	0.283019	0.349693	0.50289	0.059822	0.29975

Table 4. Analysis of the classifier of adopted and traditional schemes for dataset-3

Metrics	DBN [44]	NB [47]	SVM [46]	CNN [48]	Ensemble classifier + HB-EA [38]	Proposed Ensemble classifier + SISSGECO model
FPR	0.271845	0.213592	0.262136	0.359223	0.192414	0.004215
sensitivity	0.641026	0.710526	0.652361	0.54321	0.93389	0.909091
MCC	0.302656	0.416295	0.320891	0.147955	0.836135	0.901245
precision	0.348837	0.424084	0.36019	0.262948	0.915286	0.932146
F-Measure	0.451807	0.531148	0.464122	0.354362	0.924495	0.821918
specificity	0.728155	0.786408	0.737864	0.640777	0.807586	0.9375
NPV	0.899281	0.924658	0.903686	0.856031	0.913499	0.201921
Accuracy	0.712025	0.772655	0.72209	0.622152	0.899511	0.932642
FNR	0.358974	0.289474	0.347639	0.45679	0.06611	0.047518

Table 5. Analysis of the classifier of adopted and traditional schemes for dataset-4

Metrics	DBN [44]	NB [47]	SVM [46]	CNN [48]	Ensemble classifier + HB-EA [38]	Proposed Ensemble classifier + SISSGECO model
FPR	0.087379	0.082524	0.058252	0.089806	0.235261	0.055556
sensitivity	0.825243	0.834951	0.883495	0.820388	0.94118	0.932146
MCC	0.737864	0.752427	0.825243	0.730583	0.872974	0.935813
precision	0.825243	0.834951	0.883495	0.820388	0.947679	0.945813
F-Measure	0.825243	0.834951	0.883495	0.820388	0.942296	0.925813
specificity	0.912621	0.917476	0.941748	0.910194	0.764739	0.944444
NPV	0.912621	0.917476	0.941748	0.910194	0.885314	0.925813
Accuracy	0.883495	0.889968	0.92233	0.880259	0.937056	0.965512
FNR	0.174757	0.165049	0.116505	0.179612	0.05882	0.111111

7.4 Statistical Analysis

The statistical analysis of the developed ensemble classifier+ SISSGECO approach is computed to the existing scheme based on the accuracy measure is represented in Table VI to Table IX. In Nature, the meta-heuristic algorithms are stochastic; thus, the algorithms are performed several times for determining the achievement of the defined objective. The mean performance of the adopted ensemble classifier+ SISSGECO scheme holds a better mean value than the traditional schemes for dataset 2. The best-case scenario proves an enhancement of the proposed ensemble classifier+ SISSGECO scheme attains (~0.04879) with more accurate results than the other traditional models like ensemble classifier +SSO, ensemble classifier + BOA, ensemble classifier +CMBO, ensemble classifier +PRO, and ensemble classifier +social SO, correspondingly for dataset 2. The proposed ensemble classifier+ SISSGECO scheme has proved its improvement almost in all cases. Therefore, the development of the presented ensemble classifier+ SISSGECO scheme has been validated effectively.

Table 6. Statistical analysis with respect to accuracy: Proposed vs Conventional models for dataset 1

Methods	Best	Worst	Mean	Median	Standard Deviation
Ensemble classifier + SSO [26]	0.074074	0.154333	0.102052	0.089901	0.03274
Ensemble classifier + BOA [45]	0.084074	0.145333	0.109552	0.104401	0.022263
Ensemble classifier + CMBO [37]	0.074333	0.214468	0.118552	0.092704	0.057217
Ensemble classifier + PRO [36]	0.074333	0.140074	0.095802	0.084401	0.026825
Ensemble classifier + Social SO [35]	0.053333	0.094468	0.071651	0.069401	0.015146
Proposed Ensemble classifier + SISSGECO model	0.034488	0.064468	0.046906	0.044333	0.01093

Table 7. Statistical analysis with respect to accuracy: Proposed vs Conventional models for dataset 2

Methods	Best	Worst	Mean	Median	Standard Deviation
Ensemble classifier + SSO [26]	0.099917	0.125895	0.113822	0.114738	0.009237
Ensemble classifier + BOA [45]	0.097252	0.125895	0.117818	0.124063	0.011945
Ensemble classifier + CMBO [37]	0.10458	0.125895	0.118984	0.122731	0.008417
Ensemble classifier + PRO [36]	0.1199	0.125895	0.122565	0.122231	0.002569
Ensemble classifier + Social SO [35]	0.099917	0.125895	0.110824	0.108743	0.010583
Proposed Ensemble classifier + SISSGECO model	0.04879	0.067459	0.056365	0.054605	0.006832

Table 8. Statistical analysis with respect to accuracy: Proposed vs Conventional models for dataset 3

Methods	Best	Worst	Mean	Median	Standard Deviation
Ensemble classifier + SSO [26]	0.205488	0.916779	0.394215	0.227296	0.302119
Ensemble classifier + BOA [45]	0.185488	0.205287	0.197373	0.199359	0.00809
Ensemble classifier + CMBO [37]	0.178543	0.274316	0.210752	0.195074	0.037315
Ensemble classifier + PRO [36]	0.145488	0.401459	0.261856	0.250239	0.113283
Ensemble classifier + Social SO [35]	0.104588	0.205873	0.147809	0.140388	0.036734
Proposed Ensemble classifier + SISSGECO model	0.043189	0.067358	0.054825	0.054377	0.008814

Table 9. Statistical analysis with respect to accuracy: Proposed vs Conventional models for dataset 4

Methods	Best	Worst	Mean	Median	Standard Deviation
Ensemble classifier + SSO [26]	0.081111	0.091111	0.086204	0.086296	0.003984
Ensemble classifier + BOA [45]	0.076667	0.092222	0.08463	0.084815	0.005509
Ensemble classifier + CMBO [37]	0.075556	0.09	0.08287	0.082963	0.005135
Ensemble classifier + PRO [36]	0.075556	0.092222	0.0825	0.081111	0.006105
Ensemble classifier + Social SO [35]	0.073333	0.091667	0.08162	0.080741	0.006562
Proposed Ensemble classifier + SISSGECO model	0.048889	0.053333	0.050556	0.05	0.001843

7.5 Analysis Based on Features

The analysis based on features work for proposed and conventional features in terms of certain metrics are illustrated in Table X to Table XIII. In addition, the proposed ensemble classifier+ SISSGECO model holds better accuracy (~0.946667) than other feature comparisons, including the proposed model with conventional mutual information, the proposed model without imbalance processing, LDA, and PCA, respectively, for dataset 1. Further, the proposed ensemble classifier+ SISSGECO model holds lower FNR (0.111111) with better performance than other feature comparisons, including the proposed model with conventional mutual information, proposed model without imbalance processing, LDA, and PCA, respectively, for dataset 4. This has proved that with proposed ensemble classifier+ SISSGECO model helps to analyze more accurately, whereas other extant approaches show the least performance with the proposed concept. This absolutely evolves that the developed combination is a fit for the HD prediction model.

Table 10. Analysis based on features of adopted and traditional schemes for dataset-1

Metrics	Proposed model with conventional mutual information	Proposed model without imbalance processing	LDA [49]	PCA [42]	Proposed Ensemble Classifier + SISSGECO model
FPR	0.241774	0.16	0.198856	0.23319	0.004098
sensitivity	0.201575	0.419355	0.310744	0.222576	0.88835
MCC	-0.02607	0.179352	0.074672	-0.00692	0.896096
precision	0.070407	0.178082	0.119138	0.079096	0.994565
F- Measure	0.104362	0.25	0.17224	0.116715	0.938462
specificity	0.758226	0.84	0.801144	0.76681	0.995902
NPV	0.912692	0.945946	0.930696	0.916396	0.913534
Accuracy	0.711869	0.807882	0.76208	0.72188	0.946667
FNR	0.798425	0.580645	0.689256	0.777424	0.11165

Table 11. Analysis based on features of adopted and traditional schemes for dataset-2

Metrics	Proposed model with conventional mutual information	Proposed model without imbalance processing	LDA [49]	PCA [42]	Proposed Ensemble Classifier + SISSGECO model
FPR	0.416667	0.1	0.25	0.347222	0.074938
sensitivity	0.482759	0.857143	0.666667	0.556213	0.78524
MCC	0.052861	0.679934	0.344051	0.168368	0.75214
precision	0.21875	0.642857	0.375	0.273256	0.895214
F- Measure	0.301075	0.734694	0.48	0.366472	0.851436
specificity	0.583333	0.9	0.75	0.652778	0.938851
NPV	0.823529	0.967742	0.909091	0.862385	0.078521
Accuracy	0.563758	0.892562	0.734694	0.634421	0.932541
FNR	0.517241	0.142857	0.333333	0.443787	0.29975

Table 12. Analysis based on features of adopted and traditional schemes for dataset-3

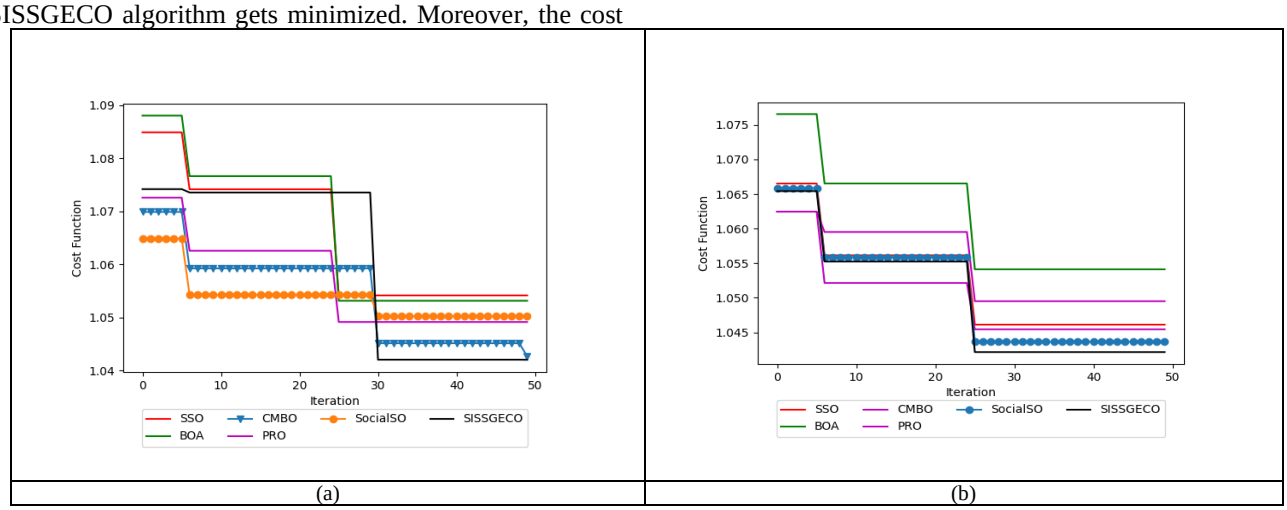
Metrics	Proposed model with conventional mutual information	Proposed model without imbalance processing	LDA [49]	PCA [42]	Proposed Ensemble Classifier + SISSGECO model
FPR	0.417476	0.227273	0.291262	0.514563	0.004215
sensitivity	0.481928	0.693878	0.618644	0.3861	0.909091
MCC	0.051548	0.388627	0.266904	-0.10298	0.901245
precision	0.218182	0.404762	0.327354	0.15873	0.932146
F- Measure	0.300375	0.511278	0.428152	0.224972	0.821918
specificity	0.582524	0.772727	0.708738	0.485437	0.9375
NPV	0.823045	0.918919	0.890244	0.758725	0.201921
Accuracy	0.56294	0.758364	0.691943	0.465477	0.932642
FNR	0.518072	0.306122	0.381356	0.6139	0.047518

Table 13. Analysis based on features of adopted and traditional schemes for dataset-4

Metrics	Proposed model with conventional mutual information	Proposed model without imbalance processing	LDA [49]	PCA [42]	Proposed Ensemble Classifier + SISSGECO model
FPR	0.325243	0.139175	0.101942	0.330097	0.055556
sensitivity	0.263736	0.633484	0.722467	0.255474	0.932146
MCC	-0.04245	0.381997	0.501664	-0.05144	0.935813
precision	0.097035	0.341463	0.438503	0.093333	0.945813
F- Measure	0.141872	0.44374	0.545757	0.136719	0.925813
specificity	0.674757	0.860825	0.898058	0.669903	0.944444
NPV	0.873664	0.953741	0.967067	0.871212	0.925813
Accuracy	0.626661	0.837575	0.88063	0.621251	0.965512
FNR	0.736264	0.366516	0.277533	0.744526	0.111111

7.6 Convergence Analysis

The convergence of the presented SISSGECO approach to the traditional schemes is examined by varying the iteration count from 0, 10, 20, 30, 40, and 50, correspondingly. Fig. 16 represents the convergence analysis of the presented scheme over the traditional schemes. The proposed SISSGECO approach attains the minimum cost function as per the defined objectives in Eq. (32). As the count of iterations rises, the cost function of the SISSGECO algorithm gets minimized. Moreover, the cost function of the proposed SISSGECO model had a fall in between the range 26th to 30th iteration for dataset 1. The cost function of the adopted SISSGECO scheme provides a lower constant value (1.042) from the 25th to 50th iteration than other existing models like SSO, BOA, CMBO, PRO, and Social SO correspondingly for dataset 4. Consequently, it is shown clearly that the adopted SISSGECO approach had attained the lower cost function with superior outcomes.



(a) (b)

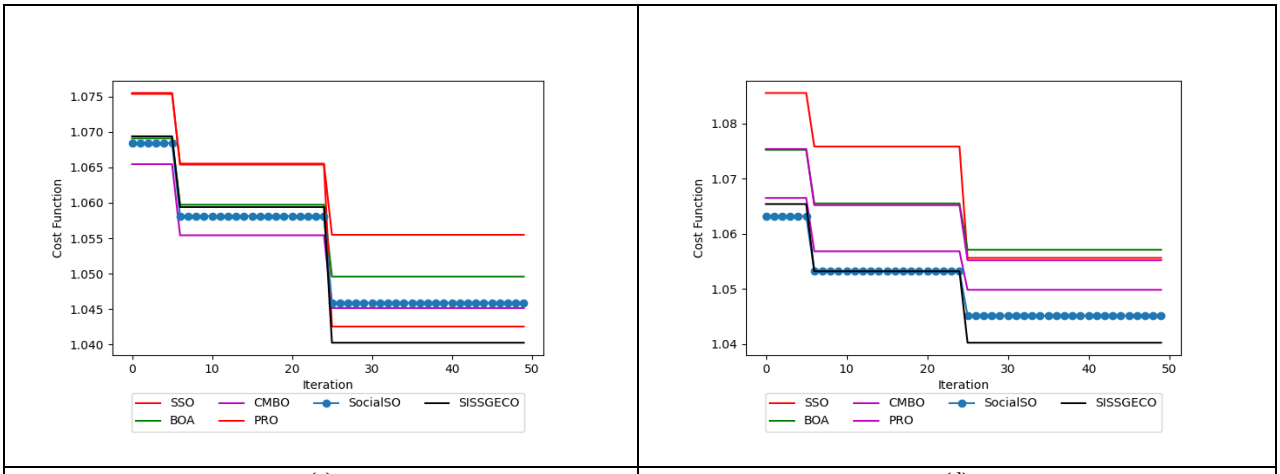


Fig. 16 Performance analysis of the developed scheme to the traditional approaches for (a) dataset 1 (b) dataset 2 (c) dataset 3 (d) dataset 4

7.7 Computational Analysis

The computation complexity for the proposed SISSGECO model with the conventional SSO, BOA, CMBO, PRO, and Social SO, is depicted in Table XIV. From the result, it can be noticed that the outcomes of the proposed SISSGECO model are 3.17%, 4.45%, 1.66%, 0.88%, and 3.26% higher than traditional SSO, BOA, CMBO, PRO, and Social SO methods. From the result, it is evident that the proposed SISSGECO model takes less time to compute when compared to other existing methods.

Table 14. Computational time

S.no	**Methods**	Time (sec)
1.	**SSO**	1042
2.	**BOA**	1056
3.	**CMBO**	1026
4.	PRO	1018
5.	Social SO	1043
6.	SISSGECO	1009

8. Conclusion

This paper has proposed an HD prediction system. Here, the output of NN, RNN, and RF is given as the input of k-NN. For making the system more accurate in HD prediction, the weights of NN and RNN were optimally tuned by a SISSGECO model. Then, the final output was obtained effectively in a precise manner. Finally, the outcomes of the developed approach were compared to the other extant schemes based on various measures like precision, sensitivity, accuracy, specificity, NPV, MCC, FPR, F1-score, and FNR, respectively. Further, the sensitivity of the adopted ensemble classifier+ SISSGECO scheme for learning percentage of 60% was 10.25%, 5.12%, 12.82%, 15.38%, and 7.69%, superior to the existing schemes for dataset 2. In addition, the proposed model proves the less FPR value (~0.15) as a better performance than the conventional models at a learning percentage of 60% for dataset 3. The best-case scenario proves an enhancement of the proposed ensemble classifier+ SISSGECO scheme attains (~0.04879) with more accurate results than the other traditional models for dataset 2.

References

[1] Farman AliShaker El-SappaghKyung-Sup Kwak, A Smart Healthcare Monitoring System for Heart Disease Prediction Based on Ensemble Deep Learning and Feature Fusion, Information Fusion. 63 (2020) 208-222.

[2] R. ValarmathiT. Sheela, Heart Disease Prediction Using Hyper Parameter Optimization (HPO) Tuning, Biomedical Signal Processing and Control. 70 (2021) 103033.

[3] Rani, P., Kumar, R., Ahmed, NMOS et al., A Decision Support System for Heart Disease Prediction Based Upon Machine Learning. J Reliable Intell Environ. 7 (2021) 263–275. https://doi.org/10.1007/s40860-021-00133-6

[4] Harika N, Swamy S.R., & Nilima, Artificial Intelligence-Based Ensemble Model for Rapid Prediction of Heart Disease, SN COMPUT. SCI. 2 (2021) 431. https://doi.org/10.1007/s42979-021-00829-9

[5] Prakash S, Sangeetha K, & Ramkumar N, An Optimal Criterion Feature Selection Method for Prediction and Effective Analysis of Heart Disease. Cluster Comput. 22 (2019) 11957–11963. https://doi.org/10.1007/s10586-017-1530-z.

[6] Renji P. CherianNoby ThomasSunder Venkitachalam, Weight Optimized Neural Network for Heart Disease Prediction Using Hybrid Lion Plus Particle Swarm Algorithm, Journal of Biomedical Informatics. 110 (2020) 103543.

[7] D. Shiny IreneT. SethukarasiN. Vadivelan, Heart Disease Prediction Using Hybrid Fuzzy K-Medoids Attribute Weighting Method with DBN-KELM Based Regression Model Medical Hypotheses. 143 (2020) 8.

[8] liyaShailendra Kumar ShrivastavaVivek Sharma, An Optimized Xgboost Based Diagnostic System for Effective Prediction of Heart Disease, Journal of King Saud University - Computer and Information Sciences. (2020).

[9] Ibomoiye Domor MienyeYanxia SunZenghui Wan, An Improved Ensemble Learning Approach for the Prediction of Heart Disease Risk, Informatics in Medicine. 20 (2020) 100402.

[10] C. Beulah Christalin LathaS. Carolin Jeeva, Improving the Accuracy of Prediction of Heart Disease Risk Based on Ensemble Classification Techniques, Informatics in Medicine. 16 (2019) 100203

[11] Md Mamun AliBikash Kumar PaulMohammad Ali Moni, Heart Disease Prediction Using Supervised Machine Learning Algorithms: Performance Analysis and Comparison, Computers in Biology and Medicine.136 (2021) 104672.

[12] Ibomoiye Domor MienyeYanxia SunZenghui Wang, Improved Sparse Autoencoder Based Artificial Neural Network Approach for Prediction of Heart Disease, Informatics in Medicine. 18 (2020) 100307.

[13] Zafer Al-MakhadmehAmr Tolba, Utilizing IoT Wearable Medical Device for Heart Disease Prediction Using Higher-Order Boltzmann Model: A Classification Approach, Measurement. 147 (2019) 106815.

[14] Aniruddha DuttaTamal BatabyalScott T. Acton, An Efficient Convolutional Neural Network for Coronary Heart Disease Prediction Expert Systems with Applications. 159 (2020) 113408.

[15] S. Mohan, C. Thirumalai and G. Srivastava, Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques, IEEE Access. 7 (2019) 81542-81554. doi: 10.1109/ACCESS.2019.2923707.

[16] N. L. Fitriyani, M. Syafrudin, G. Alfian and J. Rhee, HDPM: An Effective Heart Disease Prediction Model for a Clinical Decision Support System, IEEE Access. 8 (2020) 133034-133050. doi: 10.1109/ACCESS.2020.3010511.

[17] Y. Pan, M. Fu, B. Cheng, X. Tao and J. Guo, Enhanced Deep Learning Assisted Convolutional Neural Network for Heart Disease Prediction on the Internet of Medical Things Platform, IEEE Access. 8 (2020) 189503-189512. doi: 10.1109/ACCESS.2020.3026214.

[18] S. A. 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, in IEEE Access. 8 (2020) 65947-65958. doi: 10.1109/ACCESS.2020.2985646.

[19] B. Wang et al., A Multi-Task Neural Network Architecture for Renal Dysfunction Prediction in Heart Failure Patients with Electronic Health Records, IEEE Access. 7 (2019) 178392-178400. doi: 10.1109/ACCESS.2019.2956859.

[20] S. S. Sarmah, An Efficient IoT-Based Patient Monitoring and Heart Disease Prediction System Using Deep Learning Modified Neural Network, IEEE Access. 8 (2020) 135784-135797. doi: 10.1109/ACCESS.2020.3007561.

[21] D. Bertsimas, L. Mingardi and B. Stellato, Machine Learning for Real-Time Heart Disease Prediction, IEEE Journal of Biomedical and Health Informatics. 25(9) (2021) 3627-3637. doi: 10.1109/JBHI.2021.3066347.

[22] M. A. Khan, An IoT Framework for Heart Disease Prediction Based on MDCNN Classifier, IEEE Access. 8 (2020) 34717-34727. doi: 10.1109/ACCESS.2020.2974687.

[23] Javeed, S. Zhou, L. Yongjian, I. Qasim, A. Noor and R. Nour, An Intelligent Learning System Based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection, IEEE Access. 7 (2019) 180235-180243. doi: 10.1109/ACCESS.2019.2952107.

[24] Aggrawal, R., Pal, S. Sequential Feature Selection and Machine Learning Algorithm-Based Patient's Death Events Prediction and Diagnosis in Heart Disease. SN COMPUT. SCI. 1 (2020) 344. https://doi.org/10.1007/s42979-020-00370-1.

[25] Ripan, R.C., Sarker, I.H., Hossain, S.M.M. et al. A Data-Driven Heart Disease Prediction Model Through K-Means Clustering-Based Anomaly Detection. SN COMPUT. SCI. 2 (2021) 112. https://doi.org/10.1007/s42979-021-00518-7

[26] Mohammad-Azari S., Bozorg-Haddad O., Chu X. Shark Smell Optimization (SSO) Algorithm. In: Bozorg-Haddad O. (eds) Advanced Optimization by Nature-Inspired Algorithms. Studies in Computational Intelligence, 720. Springer, Singapore. 720 (2018) https://doi.org/10.1007/978-981-10-5221-7_10.

[27] R. Rajakumar, Impact of Static and Adaptive Mutation Techniques on Genetic Algorithm, International Journal of Hybrid Intelligent Systems. 10(1) (2013) 11-22. doi: 10.3233/HIS-120161.

[28] B. R. Rajakumar, Static and Adaptive Mutation Techniques for Genetic Algorithm: A Systematic Comparative Analysis, International Journal of Computational Science and Engineering. 8(2) (2013) 180-193. doi: 10.1504/IJCSE.2013.053087.

[29] S. M. Swamy, B. R. Rajakumar and I. R. Valarmathi, Design of Hybrid Wind and Photovoltaic Power System using Opposition-based Genetic Algorithm with Cauchy Mutation, IET Chennai Fourth International Conference on Sustainable Energy and Intelligent Systems (SEISCON 2013), Chennai, India. (2013). doi: 10.1049/ic.2013.0361.

[30] Aloysius George and B. R. Rajakumar, APOGA: An Adaptive Population Pool Size based Genetic Algorithm, AASRI Procedia - 2013 AASRI Conference on Intelligent Systems and Control (ISC 2013). 4 (2013) 288-296. doi: https://doi.org/10.1016/j.aasri.2013.10.043.

[31] B. R. Rajakumar and Aloysius George, A New Adaptive Mutation Technique for Genetic Algorithm, In proceedings of IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Coimbatore, India. 18(20) (2012) 1-7. doi: 10.1109/ICCIC.2012.6510293.

[32] Zerina Masetic, Abdulhamit Subasi, Congestive Heart Failure Detection Using Random Forest Classifier, Computer Methods and Programs in Biomedicine. 130 (2016) 54-64.

[33] Yogeswaran Mohan, Sia Seng Chee, Donica Kan Pei Xin and Lee Poh Foong, Artificial Neural Network for Classification of Depressive and Normal in EEG, 2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES). (2016).

[34] M. Akhil Jabbar, B. L. Deekshatulu, Priti Chandra, Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm, Procedia Technology. 10 (2013) 85-94.

[35] Fouad, Ahmed, Social Spider Optimization Algorithm. (2015).

[36] Seyyed Hamid Samareh MoosaviVahid Khatibi Bardsiri, Poor and Rich Optimization Algorithm: A New Human-Based and Multi Populations Algorithm, Engineering Applications of Artificial Intelligence. 86 (2019) 165-181.

[37] Dehghani, Mohammad, ŠtěpánHubálovský, and Pavel Trojovský, Cat and Mouse Based Optimizer: A New Nature-Inspired Optimization Algorithm. 21(15) (2021) 5214. https://doi.org/10.3390/s21155214.

[38] Rajendra Kumar, Identification of Noteworthy Features and Data Mining Techniques for Heart Disease Prediction, In Communication.

[39] [Online]. Available: https://www.itl.nist.gov/div898/handbook/eda/section3/eda35b.htm#:~:text=Skewness%20is%20a%20measure%20of,relative%20to%20a%20normal%20distribution.

[40] [Online]. Availanble: https://en.wikipedia.org/wiki/Statistic.

[41] [Online]. Availanble: https://en.wikipedia.org/wiki/Standard_deviation

[42] Mishra, Sidharth & Sarkar, Uttam & Taraphder, Subhash & Datta, Sanjoy & Swain, Devi & Saikhom, Reshma & Panda, Sasmita & Laishram, Menalsh, Principal Component Analysis. International Journal of Livestock Research. 1 (2017).

[43] Ling-Jing Kao, Chih Chou Chiu, Application of Integrated Recurrent Neural Network with Multivariate Adaptive Regression Splines on SPC-EPC Process, Journal of Manufacturing Systems. 57 (2020) 109–118.

[44] H.Z. Wang, G.B. Wang, G.Q. Li, J.C. Peng, and Y.T. Liu, Deep Belief Network Based Deterministic and Probabilistic Wind Speed Forecasting Approach, Applied Energy. 182 (2016) 80–93.

[45] Arora, Sankalap & Singh, Satvir, Butterfly Optimization Algorithm: A Novel Approach for Global Optimization. Soft Computing. (2019).

[46] E. Avci, A New Intelligent Diagnosis System for the Heart Valve Diseases by Using Genetic-SVM Classifier, Expert Systems with Applications. 36(7) (2009) 10618-10626.
[47] Paraskevas TsangaratosIoanna Ilia, Comparison of a Logistic Regression and Naïve Bayes Classifier in Landslide Susceptibility Assessments: The Influence of Models Complexity and Training Dataset Size, CATENA. 145 (2016) 164-179.
[48] Y. Lecun, K. Kavukvuoglu, and C. Farabet, Convolutional Networks and Applications in Vision, In Circuits and Systems, International Symposium. (2010) 253–256.
[49] Ping DengHongjun WangXinwen Zhu, Linear Discriminant Analysis Guided by Unsupervised Ensemble Learning, Information Sciences. 480 (2018) 211-221.
[50] Malige Gangappa, Kiran Mai C, Sammulal P, Enhanced Crow Search Optimization Algorithm and Hybrid NN-CNN Classifiers for Classification of Land Cover Images, Multimedia Research. 2(3) (2019) 12-22.
[51] G.Gokulkumari, Classification of Brain Tumor using Manta Ray Foraging Optimization-Based DeepCNN classifier, Multimedia Research. 3(4) (2020).
[52] Sesham Anand, Archimedes Optimization Algorithm: Heart Disease Prediction, Multimedia Research. 4(3) (2021).
[53] Yuanhao Liu, Hybrid Shark Smell Optimization Based on World Cup Optimization Algorithm for Minimization of THD, Journal of Computational Mechanics, Power System and Control. 3(3) (2020).
[54] B. Kranthi Kiran, Indian Music Classification using Neural network Based Dragon Fly Algorithm, Journal of Computational Mechanics, Power System and Control. 4(3) (2021).
[55] Vaibhav Ankush Thorat, Cloud Intrusion Detection using Modified Crow Search Optimized Based Neural Network, Journal of Networking and Communication Systems.4(2) (2021).
[56] Sankul Rathod, Hybrid Metaheuristic Algorithm for Cluster Head Selection in WSN, Journal of Networking and Communication Systems. 3(4) (2020).
[57] (2021). [Online]. Availabe: https://archive.ics.uci.edu/ml/datasets/Heart+Disease.
[58] Kalletla Sunitha, Automatically Identifying Wild Animals in Camera-Trap Images with Deep Learning, SSRG International Journal of Computer Science and Engineering. 8(5) (2021) 12-16. https://doi.org/10.14445/23488387/IJCSE-V8I5P102
[59] Nidhi Mongoriya, Vinod Patel, Review the Breast Cancer Detection Technique Using Hybrid Machine Learning, SSRG International Journal of Computer Science and Engineering. 8(6) (2021) 5-8. https://doi.org/10.14445/23488387/IJCSE-V8I6P102
[60] V .P. Amadi, N.D Nwiabu, V. I. E. Anireh, Case-Based Reasoning System for the Diagnosis and Treatment of Breast, Cervical and Prostate Cancer, SSRG International Journal of Computer Science and Engineering. 8(8) (2021) 13-20. https://doi.org/10.14445/23488387/IJCSE-V8I8P103