International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 1s
DOI: https://doi.org/10.17762/ijritcc.v10i1s.5827
***Article Received**: 30 August 2022 **Revised**: 05 November 2022 **Accepted**: 29 November 2022*

Classification Model for Meticulous Presaging of Heart Disease Detection through SDA and NCA using Machine learning :CMSDANCA

Ritu Aggarwal[1], Suneet Kumar[2]
[1]Deparment of Computer Science and Engineering
Maharishi Markandeshwar Engineering College and Maharishi Markandeshwar institute of computer technology and Business management ,Maharishi Markandeshwar Deemed to be University
Mullana, Ambala
errituaggarwal@gmail.com
[2]Deparment of Computer Science and Engineering
Maharishi Markandeshwar Engineering College,Maharishi Markandeshwar Deemed to be University
Mullana, Ambala
Suneecit81@gmail.com

Abstract— For the design and implementation of CDSS, computation time and prognostic accuracy are very important. To analyze the large collection of a dataset for detecting and diagnosis disease ML techniques are used. According to the reports of World Health Organizations, HD is a major cause of death and killer in urban and rural areas or worldwide. The main reason for this is a shortage of doctors and delay in the diagnosis. In this research work, heart disease is a diagnosis by the data mining techniques and used the clinical parameters of patients for early stages diagnosis. The intend of this learning to develop a representation that relies on the prediction method for coronary heart disease. This proposed work used the approach of self-diagnosis Algorithm, Fuzzy Artificial neural network, and NCA & PCA and imputation methods. By the use of this technique computation time for prediction of Coronary HD can be reduced. For the implementation of this the two datasets are using such as Cleveland and Statlog datasets that is collected from the UCI kaggle the ML repository. The datasets for the disease prediction measure are used to accurately calculate the difference between variables and to determine whether they are correlated or not. For this classification model, the performance measure is calculated in requisites of their accuracy, precision, recall, and specificity. This approach is evaluated on the heart disease datasets for improving the accuracy performance results obtained. The outcome for KNN+SDA+NCA+FuzzyANN for Cleveland dataset accuracy achieved 98.56 %.and for Statlog dataset 98.66 %..

Keywords- PCA, NCA, RF, Classification, Fuzzy, Artificial neural network, Cleveland, Statlog

I. INTRODUCTION

Coronary HD is a kind of main eradicator in a silent form disease which is the main reason of demise globally. According to CDSS analyze the impact of clinical decisions about the individual patients are made. With the use of CDSS, the machine learning classifiers played a vital role to enhance the quality and efficiency of the system model in healthcare in the prognosis of heart disease (Sridevi et al. 2019; Jain and Singh 2018). The use of machine learning techniques has been used in data manipulation, removal, efficiency, estimation, and calculation to achieve the best results. Heart disease is also named coronary artery disease where the CA to supply the blood to the heart muscle are blocked or narrowed this disturbs the normal functioning of the heart may lead to a heart attack. .Blood flows through blood vessels (Baccouche et al. 2020). These blood vessels deliver the blood to the heart. In the heart, the veins are that a blood vessel which delivers that delivers blood from the heart to the whole body. The symptoms of heart failure or heart attack are chest pain, jaw pain, discomfort in shoulders and back, unstable heartbeat, fatigue, heartburn that finally cause the heart attack or angina. The reason for blockage is LDL cholesterol and fatty acid deposits in the arteries that reduce the supply of oxygen and blood in the blood vessel. Angiography or echocardiography of diagnosing heart disease is costly and risky (Widiyaningtyas and Zaeni). Heart disease is a very major disease of causing death if it is early detected it could be procured otherwise caused heart attack or failure. Coronary HD caused million of death every year. The different machine learning techniques are used for diagnosing heart disease by non- protruding tests at early stages. The results of current study are evaluating by create decision on the recent test results of the patients. The same conditions are examined by the previous decisions taken by other patients. In this study to fill up the gap in previous existing works, implement a new method by using SDA,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 1s
DOI: https://doi.org/10.17762/ijritcc.v10i1s.5827
***Article Received**: 30 August 2022 **Revised**: 05 November 2022 **Accepted**: 29 November 2022*

NCA, PCA, and FANN, for the classification of heart disease each stage evaluate the patient's dataset for better accuracy and improve the efficiency by the clinical data of patients. This method is used to improve the dimensionality reduction and noise removal and missing value in the dataset (Gujare 2020). This combination of NCA and PCA with FANN could be improved. In this proposed work used NCA, which is a non parametric method to select the relevant features from the dataset after that self diagnosis algorithms is used to detect the patients have a disease or not. The findings and results get about in terms of accuracy by hybridizing the different approaches. Correlation between the different attributes and their accuracy computed by the Cleveland dataset and Statlog dataset. To compute and find the missing values in the dataset KNN with hotdeck and cold deck imputation method is employed, after new data is collected and apply NCA and PCA to obtained results. The purpose of this study is to get more accurate results that a patient has a disease or not. This paper is prepared is as follows: in Section 1 introduction discussed CHD, in Section 2 Related work is that reviews the various existing research, Section 3 Background Material and methods in Section 4 Proposed research methodology discussions about the algorithms and datasets used in this proposed work, Section 5 Results and Performance evaluation with their discussions, Section 6 Summarized Conclusion and results Nilashi et al. (2020) proposed a method of using a machine learning classifier such as SVM, various imputation methods and noise removal techniques for HD. It analyzed the Fuzzy SVM to improve the accuracy in their proposed work.T. Karthikeyan et.a l(Karthikeyan & Kanimozhi, 2017). In this study proposed a method using deep learning network classification and CNN .in this work using the DNN algorithms achieve an accuracy of 82%. Rutuja Gujare et.al(Gujare, 2020) In this study proposed work to analyzing to find the missing values using the machine learning classifiers such as RF, DT, SVM, KNN, and LR for the prognosis of heart disease. The obtained accuracy is 86.81%.it used other ensemble techniques like boosting and bagging by which the accuracy obtained by using the voting classifier that is 96.55%. Siddeswara Mayura Guru et.al(Guru et al., 2005) Proposed a method GSOM using PSO for cluster information which is the very fastest algorithm to achieved accuracy.R. Sridevi et.al(Sridevi et al., 2019). In this study researchers proposed a classification model using the Kohonen network .It analyze the feature selection methods by the two different classes .one is a relevant feature class and the other is irrelevant.SOM method is used for the implementation of this classification method. Baccouche A et.al(Baccouche et al., 2020) proposed an approach to develop the show classification model by using BiLSTM or BiGRU model with a CNN model. It achieved accuracy and F1score 91 % to 96% for HD.the concluded work as shown by table 1 comparative work.

Table 1: Comparative Work

Ref No	Proposed method	Datset	Results
[1].	Kohen method and SOM method with feature selection	Cleveland and Statlog	Analyzes feature by two classes
[2].	SVM and noise imputation methods	Cleveland and Statlog	Fuzzy SVM to improve the accuracy
[3].	Deep learning with CNN	Cleveland and Statlog	DBN achieve accuracy 82 %
[4].	RF,SVM,DT,KNN find the missing values (ensemble method)	Cleveland and Statlog	Boosting and Bagging with vote classifier achieved accuracy 96.55 %
[5].	GSOM using PSO	Cleveland and Statlog	Higher accuracy
[6].	BiLSTM or BiGRU with CNN model	Cleveland and Statlog	Accuracy and F1score 91% to 96%
[7]	**Our Work**	**Cleveland and Statlog**	**KNN+SDA+NCA+FuzzyANN for Cleveland dataset accuracy achieved 98.56 %.and for Statlog dataset 98.66 %..**

2. Materials and Methods

The self-diagnosis algorithms as intriguing artificial neural network methods and proposed as machine learning techniques. By SDA clustering the data based on the identical patterns in the form of high dimensional input space. It has two layers. One is input and the other is the Kohonen output layer (Widiyaningtyas& Zaeni). The neurons have a vector space for each input layer, for the output layer by their weights is connecting with neurons of Kohonen layer neurons. In SDA have training and testing, group resolve .this is a normalization process. Each layer contains specific neuron values for each input of a cluster. In an SDA network the following stages to be check and on the performance of these models should be analyzed (Mohan et

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 1s
DOI: https://doi.org/10.17762/ijritcc.v10i1s.5827
***Article Received**: 30 August 2022 **Revised**: 05 November 2022 **Accepted**: 29 November 2022*

al., 2019). In the first step normalized the input data and number of class, in the second step the input weight (wij) and take the random value x. third step minimum learning rate set iteration as =0.4. If the learning rate is less than to minimum least rate of value x, select data randomly with value x, data is j-x. Calculate the weight –i(Dj). It is used for solving the problem according to different samples[33].

2.1 Fuzzy Cognitive Map

It is a form of neural networks based on the dynamic state machine with the fuzzy states i.e. just for 1 and 0. the pair of a fuzzy set is considered by the ordered pair that is
X={a, MX(a).

2.2 Imputation for using incremental Methods

In these, two imputation techniques i.e. hot-deck and cold deck are used with the features of KNN. In the hot deck select the sample based on similar variables as others and as a process to fill the missing values randomly (Mohan et al. 2019). As compared to this, in cold deck the imputation is based on choosing the missing value systematically as in individually. Imputation is performed on the corresponding values of randomly selected values (Chen et al., 2017).

$X_i^j=X_k^j, k=\arg\min$……..(1)

2.3 KNN

By employing this method the attributes which have non-missing values used for this KNN are used. Consider the two vectors Xi and Xj in the dataset, find the k nearest neighbor value K neighbors (Jain & Singh, 2018) are an essential grouping and relapse technique. Given a preparation dataset, for the new information test, view as the K tests nearest to the example in the preparation dataset. The majority of the k examples have a place with which class, and these examples are delegated this class. For a given preparation dataset.

$T = \{(x1, y1), (x2, y2), \ldots (xn, yn)\}$……(2)

where xi is the element vector of the example, $yi \in \{c1, c2, \cdots ck\}$ is an example classification The classification, i = 1, 2, ... , N; the example highlight vector is x; yield y is the class to which the example has a place:

$y = \text{argmax}_{cj} \sum_x I(yi = cj)$……(3)

i, = 1,2, ⋯ ,K; $i \in Nk(x)$ (4) Where I is the marker work, ie I is 1 when $yi = cj$, in any case, I is 0 (Vijayashree and Iyengar, 2016).

2.4 Random Forest

Random forest is classifiers by which the records are taken from a dataset that is N number of records by selecting that decision trees are build which can be used for classification and regression. The output of the class is computed by individual trees (Methaila et al., 2014). Random forests (Nikookar and Naderi, 2018) utilize an irregular way to deal with synthesize numerous decision trees into a RF, and every decision tree votes to decide the last classification of the test at the hour of arrangement. To begin with, the bootstrap strategy is utilized to create m preparation sets, then, at that point, for each preparing set, a decision tree is developed. At the point when the hubs track down highlights to part, a piece of the elements are arbitrarily separated from all elements, then view as the ideal arrangement by the removed elements, applied to the hubs, split, lastly accomplish the impact of multi-order.

2.5 NB

Naïve Bayes is a ML classifier and its results are based on the naïve Bayes theorem that has conditional independence. It is a statistical classifier and all the value and their attribute consider as independently. It is very simple and easy to use and the performance provides more accurate or large datasets (Kumar et al.). Naïve Bayes is a characterization strategy in view of Bayes hypothesis and the free presumption of component conditions. For a given arrangement of preparing information, the joint likelihood circulation (X, Y) in the dataset is learned (Jain and Singh 2018). Guileless Bayes made a restrictive autonomy theory for the conditional probability distribution explicitly:

$P(X = x|Y = ck) = \prod P\,n\,j = 1\,(X\,(j) = x\,(j)\,|Y = ck)$…….. (3)

The joint likelihood dissemination (X, Y) can be acquired.

While involving Naive Bayes for arrangement, for the info, the back likelihood conveyance $P(X = x|Y = ck)$ is determined by the order model, and the class with the biggest back likelihood is yield as the classification of x

2.6 Dimensionality Reduction

The numbers of variables are reduced by the process of dimensionality reduction. It used the raw datasets to extract the latest features and maintain the structure by reducing the data. Figure 1 A is showing a model for Machine learning using the Decision Support System. In this proposed work the different kinds of methods are used for selecting of best feature datasets (Nilashi et al. 2020).

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 1s
DOI: https://doi.org/10.17762/ijritcc.v10i1s.5827
***Article Received**: 30 August 2022 **Revised**: 05 November 2022 **Accepted**: 29 November 2022*

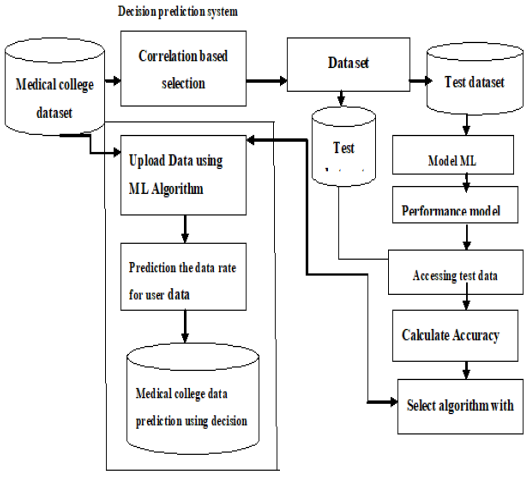

Figure 1: A Model for Machine learning using in the Decision Support System

For the Coronary HD analysis the ML techniques. In this proposed work the clustering approach, dimensionality reduction, SDA, and the classification a method was used (Figure 2).

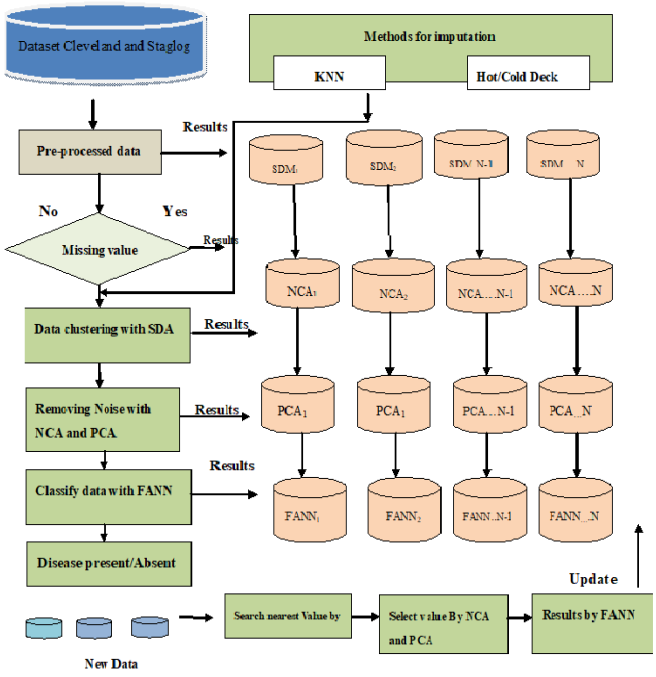

Figure 2: Methodology

Step 1: firstly take the dataset and preprocessed the data to fill the null values. In the next phase data analysis. The preprocessed data is used for classification and regression. To handle the null apply the SDA approach to improve the patient's records
by clustering the data.
Step 3 and 4: The noise removal techniques are applied such as PCA and NCA along with FANN. The classification method
Is used to classify patient's records for heart disease .for this SDA approach is used.

3. Results and Performance Evaluation

It is calculated on basis of testing and training the dataset. The assumption behind the SDA, KNN, Hot deck/Cold Deck was computed for the missing value and null values in the dataset. This all work depends on the clustering techniques that follow in the projected work. The significance k is set to get out non-missing data in the datasets. To accurately find the results applying NCA and PCA to get more accuracy as compared to existing work. For final calculation and results are estimated by fuzzy artificial neural networks.

Specificity= $T_{neg}\_Rate = \frac{True\_NEG}{True\_NEG+False\_POS}$ ..... (4)

Precision= $\frac{True\_POS}{True\_POS+False\_POS}$ ..... (5)

Recall = $\frac{True\_POS}{True\_POS+False\ \_NEG}$ ..... (6)

Accuracy= $\frac{True_{\_POS}+True\_NEG}{True_{\_POS}+False_{\_POS}+True_{NEG}+False\_NEG}$ ..... (7)

3.1 Datasets Classification

In this proposed work Statlog and Cleveland datasets and the chi-square method was used for heart disease in these experiments. In this, the Statlog contains the 270 records, and 13 attributes, while the Cleveland datasets contain 303 features, and 14 attributes. Results presented in Table 1 revealed the dataset for Cleveland and Statlog describes the Coronary angiography and 17 independent features. For Coronary angiography, the NUM is use to define the identification of HD by value yes or no. It shows the presence or absence of HD .by non-invasive tests like exercise electrocardiogram, exercise thallium .the datasets have 17 features. The cardiologist estimates and interprets the patient's coronary angiography results without knowing their non-invasive tests.

3.1.1 Statlog and Cleveland Dataset

In Statlog and Cleveland have different dataset for heart disease prediction in which statlog have 270 instances and 13 attributes were used, as in Cleveland dataset 303 instances and 14 attributes were used. Basically the main difference only the target value is missing in statlog dataset. Both these dataset is used to check the correlation between the attributes according to different approaches and ML classifiers.

3.1.2 Chi-Square Test

A chi-square test is a measurable test used to contrast noticed outcomes and prediction outcomes. The motivation behind

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 1s
DOI: https://doi.org/10.17762/ijritcc.v10i1s.5827
***Article Received**: 30 August 2022 **Revised**: 05 November 2022 **Accepted**: 29 November 2022*

this test is to decide whether a contrast between observed information and expected that information is expected should risk, or then again assuming it is because of a connection between the factors

Table 1: Dataset for Cleveland and Statlog

Sr. No.	Group	Patient Feature Group	Feature Code	Description
1	Patient data	Age in years	AG	29 and 77.
2	Patient data	Sex in numeric	SEX	M: 1, F: 0
3	Patient data	Rest blood sugar in body	RBPR	94 mm Hg-200 mm Hg
4	Exercise electrocardiogram	Exercise based induced angina	EXIA	Yes: 1, No : 0
5	Patient record data	Depression ST induced	STDIE	1,2,3 by levels
6	Patient Record data	ST Slope of the peak Segment.	SPE	0 -6.2
7	Blood Vessel colored by fluoroscopy	No. of major vessels (0-3)	NMV	0 -3
8	Exercise electrocardiogram	Thal	THM	N- 3, FD : 6,RD : 7
9	Blood Vessel	Serum Chol	SCL	126 mg/dl -564 mg/dl
10	Blood Vessel	Fasting sugar in blood	FBSR	FBSR > 120 Mg
11	Exercise electrocardiogram	Electrocardiographic Results in Resting	RER	N- 0, ST-T: 1, HT: 2)
12	Exercise electrocardiogram	Heart rate maximum	MHRA	71 -202
13	PR Data	Maximum Smoking	CIGS	Per day Cigarettes Consumption
14	PR Data	Diagnose HD	Num	Yes or No
15	Exercise electrocardiogram	Beta-blocker used during exercise ECG	Prop	1 = yes; 0 = no)
16	Exercise electrocardiogram	Target	TAR	Heart disease absent =0 ,or Present=1
17	Exercise thallium scintigraphy	Peak exercise systolic blood pressure Exercise Thallium heart	THAL	Exercise Thallium heart scan: 3 ¼ normals; 6 ¼ fixed defects; 7 ¼ reversible defect

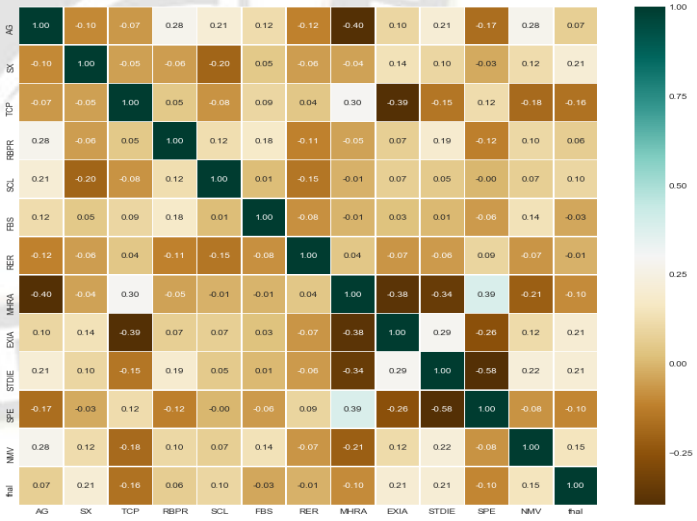

Figure 3: Correlation between Features in Statlog dataset

By the above figure 3 correlation between features in Statlog 13 attributes are to be selected such as AG, SX, TCP, RBPR, SCL, FBS, RER, MHRA, EXIA, STDIE, SPE ,NMV, THAL. The significance of each correlated attribute compares about the connection among's features and eliminate one of two features that have a relationship higher than 0.9.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 1s
DOI: https://doi.org/10.17762/ijritcc.v10i1s.5827
***Article Received**: 30 August 2022 **Revised**: 05 November 2022 **Accepted**: 29 November 2022*

Table2: Results for Cleveland dataset

Approach	Precision	call	ccuracy	Specificity
KNN+SDA +NCA +FuzzyANN	0.9545	0702	0.9856	0.9345
Hot-Deck+Cold Deck+SDA +PCA +NCA FuzzyANN	0.9425	0654	0.9669	0.9123
SDA+NCA +ANN	0.9263	0623	0.9779	0.9023
PCA+ANN	0.8873	0235	0.9012	0.8903
SDA+ FuzzyANN	0.8338	0963	0.8978	0.789
ANN	0.8652	0979	0.8999	0.6789
NB	0.7906	0989	0.7995	0.8926
RF	0.6949	0959	0.7982	0.7919

In the below figure 5 the results for KNN+SDA+ NCA +FuzzyANN in terms of accuracy is 0.9856, for Recall 0.9702, Precision 0.9545.

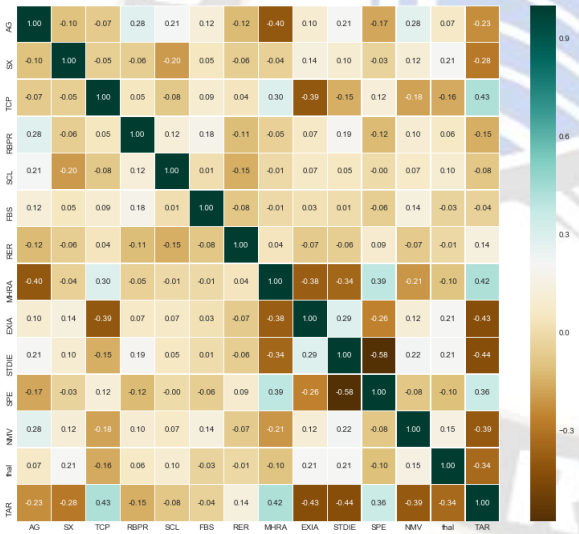

Figure 4: Correlation between features in Cleveland dataset

By the above Figure 4 correlation between features in Cleveland Dataset 14 attributes are to be selected like AG, SX,TCP, RBPR ,SCL,FBS,RER,MHRA,EXIA,STDIE,SPE,NMV,THAL,TAR.

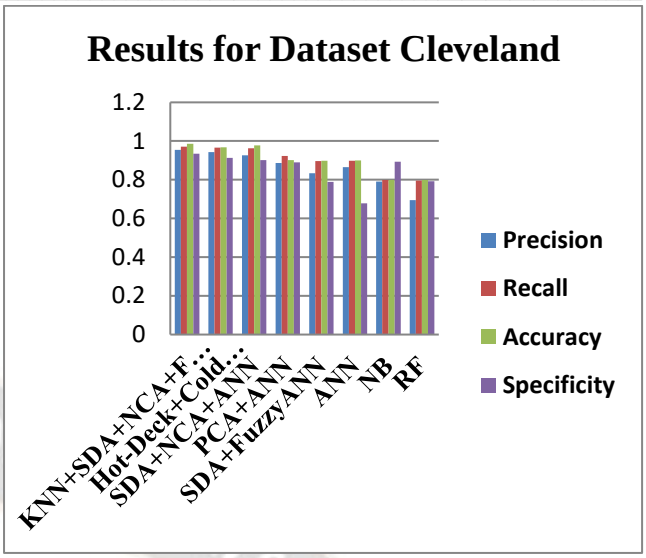

Figure 5: Results for Dataset Cleveland

By the above graph the accuracy achieved 0.9856 for better outcomes for Cleveland.

Table 3: Results for Statlog dataset

Approach	Precision	Recall	Accuracy	Specificity
KNN+SDA +NCA +FuzzyANN	0.9534	0.9822	0.9866	0.9355
Hot-Deck+ Cold Deck+ SDA +PCA +NCAFuzzy ANN	0.9415	0.9752	0.9679	0.9126
SDA+NCA + ANN	0.9253	0.9674	0.9789	0.9523
PCA+ANN	0.8863	0.9255	0.9015	0.8923
SDA+FuzzyANN	0.8238	0.8967	0.8988	0.7819
ANN	0.8551	0.8949	0.8999	0.6889
NB	0.7905	0.7679	0.7979	0.8916
RF	0.6939	0.7969	0.7989	0.782

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 1s
DOI: https://doi.org/10.17762/ijritcc.v10i1s.5827
***Article Received**: 30 August 2022 **Revised**: 05 November 2022 **Accepted**: 29 November 2022*

In the below figure 6 the results for KNN+SDA+NCA+FuzzyANN in terms of accuracy is 0.9866, for Recall 0.9822, Precision 0.9534.

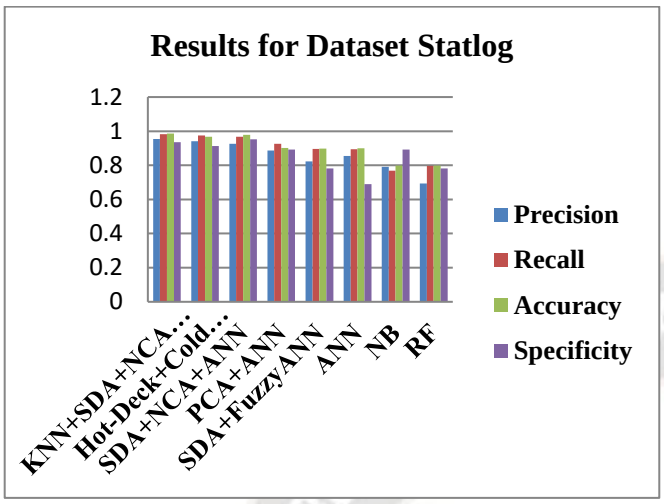

Figure 6: Results for Dataset Statlog

The accuracy achieved 0.9866 that gives best results for connection among's features and eliminate one of two features that have a relationship higher than 0.9.

5. Conclusion

This study proposed and enhanced a method for heart disease diagnosis using a machine learning classifier and the imputation methods such as FANN, RF NB, ANN, and other Noise removal techniques and also the accuracy improving methods like NCA and PCA. The clustering method SDA is used. The two datasets Cleveland and Statlog dataset are used and provide accuracy, precision, recall, and specificity values for each method comparisons. Evaluated for PCA as compared to NCA is better with the other approaches. The results for KNN+SDA+NCA+FuzzyANN for Cleveland dataset accuracy achieved 98.56 %.and for Statlog dataset 98.66 %.In the future this method could be used for feature selection and used other imputation methods could be used for better results.

Conflicts of Interest

The authors declare that there is no conflict of interest

Funding Information

No

References

[1]. Nilashi, M., Ahmadi, H., Manaf, A.A., Rashid, T. A., Samad, S., Shahmoradi, L., Aljojo, N., &Akbari, E.(2020).Coronary heart diseasediagnosis through self-organizingmap and fuzzysupportvectormachine with incrementalupdates. *International Journal of Fuzzy Systems*, *22*(4), 1376–1388. https://doi.org/10.1007/s40815-020-00828-7

[2]. Karthikeyan, T., &Kanimozhi, V. A.," Deep Learning Approach for Prediction of Heart Disease Using Data mining Classification Algorithm Deep Belief Network", International Journal of Advanced Research in Science. (2017).*EngineeringTechnology*,*4*(1, January).

[3]. Gujare, R. (2020). Heart disease prediction using ensemble learning methods. *International Journal of Advanced Science and Technology*, *29*(6), 76–85.

[4]. Guru, S. M., Hsu, A., &Halgamuge, Saman. (2005).'An Extended Growing Self-Organizing Map for Selection of Clusters in Sensor Networks', al of Distributed Sensor Networks.*International Journal of Distributed Sensor Networks*, *1*(02, June), 0–0

[5]. Sridevi, R., Dinadayalan, P., & Bastin Britto, S.(2019). An appropriate feature classification model using Kohonen network. *International Journal of Computer Engineering & Technology* , *10*(2, March–April), 148–159. https://doi.org/10.34218/IJCET.10.2.2019.016

[6]. Baccouche, A., Garcia-Zapirain, B., Castillo Olea, C., &Elmaghraby, A.(2020).'A Ensemble Deep Learning Models for Heart Disease Classification', A case study from Mexico. *Information*,*11*(4), 207. https://doi.org/10.3390/info11040207

[7]. Jain, D., &Singh, V.(2018).Feature selection and classification systems for chronic disease prediction: A review. *Egyptian Informatics Journal*, *19*(3), 179–189. https://doi.org/10.1016/j.eij.2018.03.002 ISSN, *8665*, 1110.

[8]. Mienye, I.D., Sun, Y., &Wang, Z.(2020).An improved ensemble learning approach for the prediction of heart disease risk. *Informatics in Medicine Unlocked*, 20. https://doi.org/10.1016/j.imu.2020.100402

[9]. Nikookar, E., &Naderi, E.(2018).'Hybrid Ensemble Framework for Heart Disease Detection and Prediction',IJACSA. *International Journal of Advanced Computer Science and Applications*, *9*(5). https://doi.org/10.14569/IJACSA.2018.090533

[10]. Widiyaningtyas, T., &Zaeni, I. A. E.," Self-Organizing Map (SOM) For Diagnosis Coronary Heart Disease", ICITSEE ,2019.

[11]. Mohan, S., Thirumalai, C., & Srivastava, G.(2019).Effective heart disease prediction using hybrid machine learning techniques. *IEEE Access*, *7*(July 3), 81542–81554. https://doi.org/10.1109/ACCESS.2019.2923707

[12]. Gavhane, A., Kokkula, G., Pandya, I., &Prof.Devadkar, K. Prediction of heart disease using machine learning, IEEE Conference Record # 42487; IEEE XploreISBN:978-1-5386-0965-1 (ICECA 2018)

[13]. Karayılan, T., & Kılıç, Ö.(published 2017). *Prediction of heart disease using neural network*. IEEE Publications.

[14]. Dinesh Kumar, G., Arumugaraj, K., Santhosh Kumar, D., &Mareeswari, V.Prediction of cardiovasculardiseaseusingmachinelearningalgorithms. *Publisher: IEEE, proceeding of 2018 IEEE international conference*.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 10 Issue: 1s
DOI: https://doi.org/10.17762/ijritcc.v10i1s.5827
***Article Received**: 30 August 2022 **Revised**: 05 November 2022 **Accepted**: 29 November 2022*

[15]. MrKrishnan, S.(April 26, 2019). J, Dr Geetha. S.*Prediction of Heart Disease Using Machine Learning Algorithms*.

[16]. Wang, C. W. (2006).New ensemble machine learning method for classification and prediction on geneexpressiondata, Proceeding of 2016 IEEE International Conference. *Conference Proceedings*, 2006, 3478–3481. https://doi.org/10.1109/IEMBS.2006.259893

[17]. Anooj, P. K.(January 2012).Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules. *Journal of King Saud University – Computer and Information Sciences*, *24*(1), 27–40. https://doi.org/10.1016/j.jksuci.2011.09.002

[18]. Abdullah, S., &Rajalaxmi, R. R.(2012).A data mining model for predicting the coronary heart disease using random forest classifier. In*Proceedings of the International Conference Recent Trends Computability Methods, Commun.Controls*,April(pp.22–25).

[19]. Rairikar, A., Kulkarni, V., Sabale, V., Kale, H., &Lamgunde, A.(2017, June). Heart disease prediction using data mining techniques. InInternational Conference on Intelligent Computing and Control (I2C2), *2017* (pp. 1–8). IEEE Publications.

[20]. Vijayashree,J.,&SrimanNarayanaIyengar,N.Ch.(2016).Heart disease prediction system using datamining and hybrid intelligettechniques: A review. *Informatics in Medicine Unlocked* 16 (2019), *8*(4), 139–148 11.

[21]. C. Beulah Christalin Latha, S. Carolin Jeeva ," Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques ",

[22]. Kolukisa, B., Hacilar, H., Goy, Gokhan, Kus, M., Bakir-Gungor, B., Aral, A., &Gungor, V. C.," Evaluation of Classification Algorithms. (2018).Linear discriminantanalysis and a newhybridfeatureselectionmethodology for the diagnosis of coronaryarterydiseaseIEEE International Conference on Big Data (Big Data).

[23]. Chen, M., Hao, Y., Hwang, K., Wang, L., &Wang, L.(2017).Disease prediction by machine learning over big data from healthcare communities. *IEEE Access*, *5*,8869–8879. https://doi.org/10.1109/ACCESS.2017.2694446

[24]. Methaila, A., Kansal, P., Arya, H., &Kumar, P.(2014).Earlyheart disease prediction Using datamining techniques.

[25]. Pouriyeh, S., Vahid, S., Arabnia, H. R., & Sannino, G. (July 2017).*A comprehensive einvestigation on comparison of machine learning techniques on heart diseasedomain*.

[26]. Dinh, A., Miertschin, S., Young, A., & Mohanty, S. D.(2019).A data-driven approach to predicting diabetes and cardiovascular disease with machine learning. *BMC Medical Informatics and Decision Making*, 19(1), 211. https://doi.org/10.1186/s12911-019-0918-5

[27]. Kohli, P. S., &Arora, S.Application of machine learning in disease prediction. *Proceeding of 2018 IEEE international onference*.

[28]. Ghumbre, S. U., &Ghatol, A. A.(January 2012).'Heart Disease Diagnosis Using Machine Learning Algorithm,'advances in intelligent and softcomputing.*Proceedings of the International Conference on Information Systems Design and Intelligent Applications*.

[29]. Dahiwade, D., Prof.Patle, G., &Prof. Meshram, E.(2019).Designing disease prediction mode lusing machine learning approach.*IEEE, ICCMC*.

[30]. Chen, M., Hao, Y., Hwang, K., Wang, L., &Wang, L. (2017 b).Disease prediction by machine learning over big data from healthcare communities.*IEEE Access*, *5*, 8869–8879. https://doi.org/10.1109/ACCESS.2017.2694446

[31]. Aggarwal, R., &Thakral, P.(2022). Meticulous presaging arrhythmia fibrillation for heart disease classification using oversampling method for multiple classifiers based on machine learning. InP.Verma, C.Charan, X. Fernando&S.Ganesan(Eds.), *Advances in data computing, communication and security. Lecture notes on datae ngineering and communicationst echnologies*, *106*. https://doi.org/10.1007/978-981-16-8403-6_9. Springer.

[32]. Aggarwal, R., &Kumar, S.(2022, March). An automated perception and prediction of heart disease based on machine learning. In *AIP Conference Proceedings* 2424 (1), p. 020001). AIP Publishing LLC.

[33]. Aggarwal .,R , Kumar, S.(2022)HRV based feature selection for congestive heart failure and normal sinus rhythm for meticulous presaging of heart disease using machine learning, Measurement: Sensors,2022100573,ISSN 2665-9174,https://doi.org/10.1016/j.measen.2022.100573(https://www.sciencedirect.com/science/article/pii/S2665917422002070)