Heart Disease Prediction using Neuro-Genetic Algorithm and CNN-MDRP Classifier

O. Bhaskaru[1] and M.Sreedevi[2]

[1]Research Scholar, CSE Department, KoneruLakshmaiah Education Foundation, Green Fields, Guntur District, Vaddeswaram,
Andhra Pradesh, 522502.
[2]Professor, CSE Department, KoneruLakshmaiah Education Foundation, Green Fields, Guntur District, Vaddeswaram,
Andhra Pradesh,522502.
Corresponding author mail: obhaskaru26@gmail.com

Abstract

Prediction of heart disease is vital in healthcare sector due to high risk factor related to the disease. Data analysis plays a vital role in prediction based on patient history. Each factor has to be taken intoconsideration for the prediction to be accurate. Conventional methods involve enormous data rather than accurate prediction. Data has to be chosen correctly for attaining earlier predicting process. If the data collected is partial it's a setback for analysis. The previous work designed an Improved Step Adjustment based Glowworm Swarm Optimization Algorithm with Weighted Feature based Support Vector Machine (ISAGSO-WFSVM) for Heart diseasediagnosis. However, the WFSVM is does not suitable for large data sets and it has long training time. To solve this problem, the proposed system designed a Neuro-Genetic with CNN-MDRP approach for heart disease prediction. In this work, initially perform cleaning of data by converting the incomplete data to informational data from the dataset. The work combines CNN-MDRP classifier for earlier accurate prediction (convolutional neural network (CNN)-based efficient multimodal disease risk prediction)along with neuro genetic approach.The dataset is processed by Naïve Bayes algorithm using structured data. The proposed neuro-genetic approach finds a feasible solution for optimal network configuration. The results prove that combining both effective algorithm and classifier acquires 96.25% of accuracy. The metrics proposed will provide a clear insight on reliable factors.

Key words: CNN-MDRP, Neuro-Genetic approach, Heart disease prediction

1. Introduction

A rapid increase in death rate due to heart disease has been focused for providing a solution in earlier detection through effective methods. The disease occurs due to n number of reasons including stress factors, high cholesterol food and even genetic factors [1]. The existing prediction methods involve automatic diagnosis system which fails to achieve high percentage of accuracy since they include irrelevant features from the dataset. Many researchesfocuson preprocessing of features and feature selection.Earlier prediction is essential because a person is slowly affected by it and he is not aware of the symptoms progressing it to dangerous level[2]. Temperature and heart beat monitoring system was developed which alerts the patient if they are predicted with critical symptoms [3]. Machine learning has made enormous contribution to medical field through perfection in prediction, detection and diagnosing a disease. It contributes its best for earlier detection but still certain pitfalls are observed which leads to irrelevant diagnosis. The automatic prediction is done through structured data where the data is well formed through data available in dataset. For heart disease prediction the data included are ECG output, laboratory records, medical records from medical experts. This particular work handles the data that is really essential for prediction and eliminates irrelevant contents through CNN- MDRP and accuracy of selected features through neuro-genetic algorithm. For obtaining the patterns that are essential from multidimensional database an algorithm was developed for identifying the frequency and maximum transaction length of a prescribed pattern through Apriori algorithm [4]. Previous research implemented neuro genetic approach for selecting the exact features for prediction. Investigations prove that applying genetic algorithm initializes weights in artificial neural network for optimization and overcoming shortcomings of slow convergence [5]. Through unsupervised learning feature selection was implemented by multi-objective genetic algorithm [6]. The results obtained by feature selection through genetic algorithm and optimizing the parameters have never failed to improve the accuracy in classification process [7]. The limitation observed in predictive accuracy of earlier methods contributes very less to cardio disease from multidimensional database. Hence it has motivated us to improve the classification through CNN and applying neuro genetic approach through large dataset.

2. Related Works

This section analyzes and describes the approaches built for chronic heart disease detection and presents the impacts of various genetic algorithms which contribute their best in data mining.

Top K High-Utility pattern mining pattern is used for identifying the threshold regularity of features that are used often. This is implemented by tree-based data structure named as RP-Tree [8].

A fuzzy logic based Extreme Machine Learning was developed for analyzing the risk factor in heart disease [9]. Here the feature selection for risk analysis is performed by Particle Swarm optimization algorithm. The results prove that the diagnosis method attains better detection accuracy. It was stimulated in python and the data set was gathered from Cleveland Clinic Foundation Database.

Prototype analysis of various clustering and classification approaches of data mining techniques where provided which helped in extracting only essential features from large dataset [10]. The extraction process involved is obtaining only exact features from the dataset and converting them into meaningful data for further retrieval of data. The main data mining approaches are classification and clustering where data retrieval is efficient through neural networks, decision trees and Bayesian networks. A survey on supervised and semi- supervised detection techniques for both category based and unlabeled datasets in real time synthetic data.

Data clustering in uncertain data streams is a challenging concept in relevant to relational data aspects. An Adaptive Connection based Clustering approach (ACCA) was proposed where conventional matrix formation was defined which highlighted the matrix formations based on the similarity between various clusters using various attributes [11]. A connection-based algorithm was implemented for finding out the assessment which is similar from categorical data for generating final clusters from similar attributes.

With respect to prediction of heart disease knowledge discovery and data mining are explored in various ways. This is because each patient has a unique set of features to be considered for prediction. The attributes has to be arranged in systematic manner for data which are uncertain in data indexing approach. A novel Fuzzy based Partitioned Genetic algorithm (NFPGA) was developed especially for categories of data that are uncertain [12]. Initially data set is partitioned with maximum number clusters are combined. The process is repeated until the clusters are equal to pre-defined clusters relevant to dataset. UCI repository dataset was utilized for obtaining optimal fitness function; cross-over and mutation operations that are evaluated based on parallel partitioning procedure.

3. Proposed system

Heart Disease detection using CNN-MDRP

This section introduces CNN based uni-modal disease risk detection algorithm. The method identifies whether an individual undergoes critical symptoms of heart disease based on the aspects in medical record. The input holds essential factors like gender, indication factor and other information as (f1,f2,f3,…,fn). Another attribute measuring the intensity of risk factor is R0- high risk and R1- minimal risk. The following section covers dataset and its characteristics. The data can be either organized in a certain way or textual information. Textual information seems to be unstructured set of data. Here identification of essential features for prediction is complicated. Organized and unorganized information are considered as multidimensional database and obtains risk factor of an individual presented in the dataset. The unimodel disease risk prediction is based on textual data of the patient. In general, both organized and unorganized data is not sufficient for examining the disease prediction. There is time consumption while performing above mentioned process, to overcome this issue we use CNN-MDRP algorithm.

CNN-MDRP based data processing

The identification of risk factors with relevance to heart disease can be obtained by neural network algorithm using both structured and unstructured textual data. The usage of convolutional neural network (CNN) automatically extracts the necessary information from the data available. CNN has the potential to extract the information that in indirect and lacking from huge medical records. The statistical knowledge is taken into consideration for regulating general illness. Medical experts handling those patients provide an organized data in a particular format for reusable properties. These steps filter out the data that is not needed and provides the exact information for predicting the criticality [13, 14]. The phases of CNN are described below

(i) Data Acquisition

Data Acquisition is the process of obtaining information that is indirectly available in textual data. When data is replaced its termed as “attribute unit”, if component of data is replaced its named as “attribute assignment”. Prior to data acquisition we must remove the unsure or irrelevant information from patient data. Through this process

of data segregation occurs attaining uniqueness in prediction. Aggregate index is acquired and the dataset is considered to be multidimensional. Let Rx_y be dimension of data where x represents the overall data available and y be the unique characteristics of every patient.

(ii) Embedding trained words

The commonly used terms relevant to chronic heart disease has to be embedded in data repository to avoid various terms for a single word. This makes our process simple in identification and prediction of symptoms.

(iii) Steps of CNN-MDRP algorithm

The steps involved in algorithm are as follows:

(i) Creating textual information
(ii) CNN based text conversion
(iii) Convolution layer
(iv) Grouping of text in CNN
(v) CNN classifier

(i)Creating textual information:

Every information related to textual data of patient dataset utilizes a common word embedded in processing the terms present.

(ii)CNN Based text conversion:

At every scenario word are evaluated, where the preferred words are evaluated in beginning and at the end.

(iii)CNN Convolution layer:

The output generated by the convolution layer is considered as an input for terminology grouping operation. The grouping process is essential because every word is unique, differentiation among them is needed to make precise prediction. Maximum number of combinations takes place in convolutional layer and final element is obtained through that process.

(iv)Grouping of text in CNN:

Grouping region is connected via neural network involving real artifacts and deviations. Below diagram depicts the process involved in CNN-MDRP.

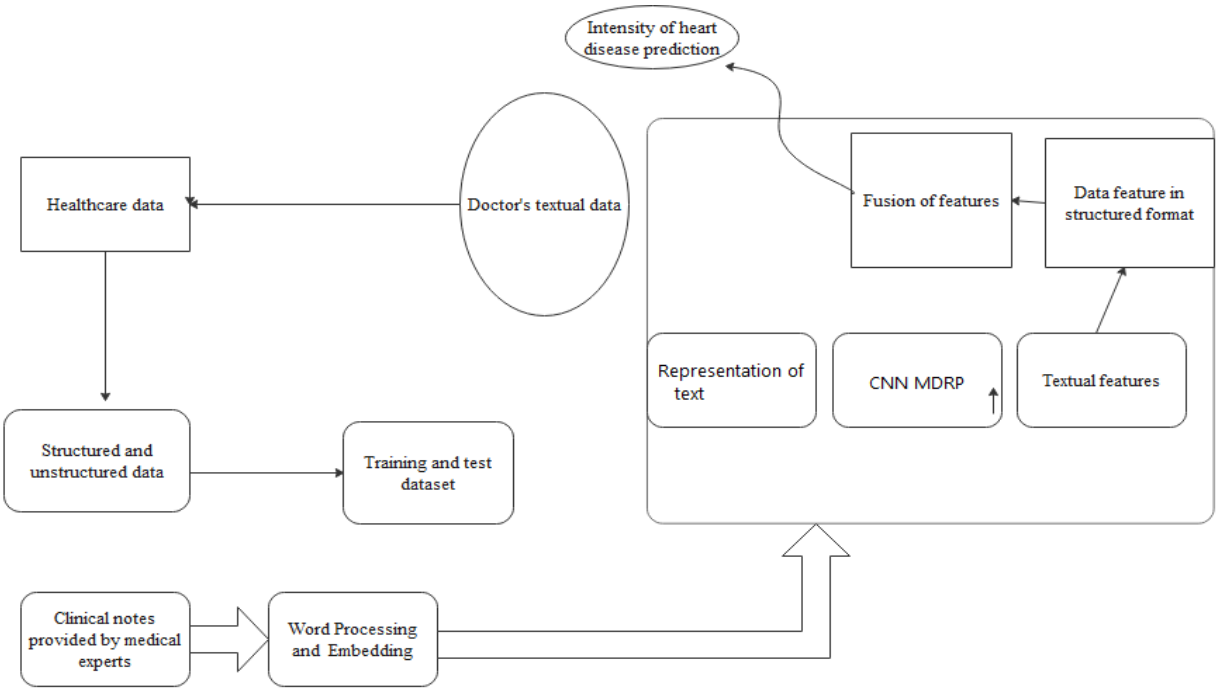

Fig 1: CNN-MDRP architecture

4. Neuro-Genetic approach

The section presents the fitness evaluation in multi-objective function for selection process of feature subset by efficient GA. The depicted subsets make use of strings that are binary where we have assumed value 1 as inclusion of a feature while training the convolutional neural network and binary value 0 represents the feature masking while the network undergo training.

(i) Optimization Function

While deriving best feature subset a CNN model is developed for each subset for evaluating the fitness value. The CNN-MDRP is assumed to have same number of inputs as the number of 1's in corresponding string that is in binary format while representing subset of features. During the process of CNN training with each relevant subset concludes with an error value F(b) in classification of heart disease, b represents the binary string and $0 \leq F(b) \leq 1$. For formulating optimization function, we consider the cost of training the features present in the subset. For acquiring the cost function, we assume training feature cost as 1 and it is directly proportional to number of neurons available a input. C(b) the cost function in binary string is obtained by dividing number of 1s in string to number of 1s in the string to the number of features. The range of cost function lies between 0 to 1. In evaluating the cost function, error value and optimization function can be formulated as,

$$O(b) = (2 - F(b)) - \left(\frac{c(b)}{2 - F(b)} \right) \quad (1)$$

This optimization minimizes the error rate and training cost of convolutional neural network. Insignificant solution is avoided and solutions are provided with high level of accuracy and cost factor that is acceptable. For reducing the nature of complexity in CNN-MDRP we integrate Neuro-genetic approach for optimization of features. The next section provides the steps involved in the algorithm.

Neuro Genetic Algorithm

In general GA is meant to be an iterative process where chromosomes are involved in representing a genomes string for obtaining best preferable solution in a space where the problem is defined. The genetic algorithm is a complicated process of searching and exploring large space for obtaining the best optimal solution. The steps involved in the algorithm are described below:

(i) Chromosome of GA consists of encoded candidate solutions.
(ii) Fitness function helps in evaluating the quality of every candidate.
(iii) The operators in GA are iteratively applied until optimal solution is obtained.

The Neuro-Genetic algorithm combined with convolutional neural network is depicted in below flowchart. The optimized feature subsets is used for predicting the chronic heart disease by training the CNN. The accuracy is not compensated by this hybrid methodology.

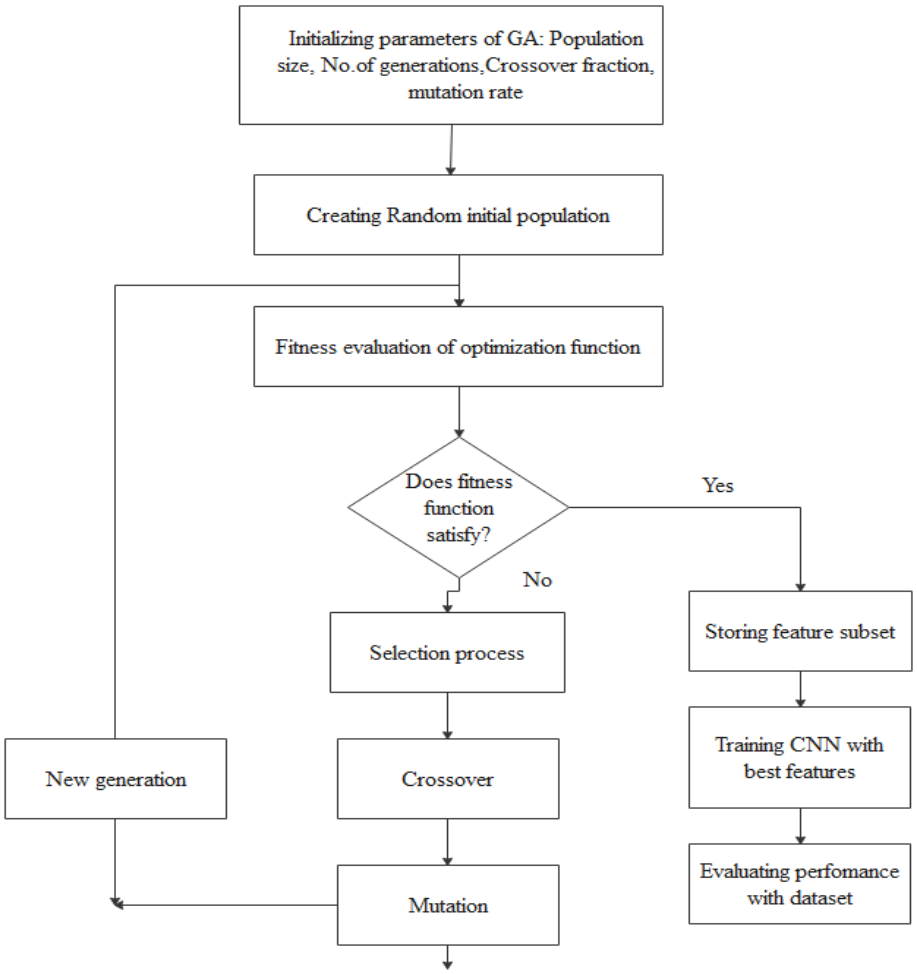

Fig 2 : Flowchart of CNN trained Neuro-Genetic Approach

The model proposed begins with initializing population from random spaces including GA operators like size of the population, Number of generations involved, Fraction of crossover and mutation rate. Fitness of optimization function is obtained by calculating the training cost of the convolutional neural network and error percentage in classification. If fitness function seems to be non-optimal the algorithm is performed with new set of parameters for obtaining new generation. The methodology undergoes iteration until optimal subset of features is obtained. CNN is trained with minimal set and its performance is evaluated through dataset present. The CNN model used machine learning algorithm through various hidden layers for accurate prediction.

5. Results and Discussions

A. Analysing the data of the chronic heart disease

The data has been collected from 400 patients relevant to the symptoms of heart disease. Initially the data is analysed for clearing the values that are missing and irrelevant type values. The input parameters for predicting the chronic heart disease is listed below in the table. The data is collected from Cleveland Clinic Foundation Database which is available in instant format.

Table1:Attributes involved for predicting the chronic heart disease

S.No	Attributes taken into consideration
1.	Age factor of the patient(in yrs)
2.	Sex(1=Female,0=Male)
3.	Categories of chest pain: 1=atypical angina, 2=typical angina, 3=Non-angina pain, 4= asymptomatic
4.	Blood Pressure(BP)
5.	Cholesterol Serum
6.	Fasting blood sugar>120mg/dl 1=True,0=false
7.	ECG results: 0=Normal; 1=ST elevation; 2= left ventricular Hypertrophy
8.	Maximum heart rate
9.	Whether angina is induced (Yes/No)
10.	Depression level of ST induction by exercise
11.	Slope of peak exercise ST segment, 1= Sloping up, 2= flat,3=down sloping
12.	Number of vessels ranging from 0-3
13.	4=Normal, 6=fixed defect, 7= reversible defect
14.	Heart disease diagnosis Value 0: <50% diameter narrowing Value 1: >50% diameter narrowing

B. Performance Indices

The proposed model is evaluated through its performance metrics obtained bycomputing the Sensitivity percentage (SE), Specificity Percentage(SP) and. Accuracy (AC). These validation parameters are defined as:

$$Sensitivit\quad y = \frac{True\quad Positive}{True\quad positive\quad + False\quad negative} \tag{2}$$

$$Specificit\quad y = \frac{True\quad Negative}{True\quad negative\quad + False\quad positive} \tag{3}$$

$$Accuracy\quad = \frac{True\quad Positive}{True\quad positive\quad + False\quad negative\quad + True\quad positive\quad + False\quad negative} \tag{4}$$

Where True Positive is the number of attributes correctly classified as healthy, True Negatives is the number of subjects as abnormal; False Negatives is the number of subjects misclassified as abnormal when actually normal, and FP (False Positives) is the number of subjects misclassified as normal when actually abnormal.

C. Experimental Results

CNN MDRP experimentation

The factors involved are height and weight of the patient and bias range of BMI(Body Mass Index). These factors together contribute a major role in risk assessment of chronic heart disease. The values of the complete layer deal with connections and deviations of the network layer. The layer is directly connected to the classifier for further processing. For processing structured and unstructured data, we are experimenting CNN-MDRP for efficient handling and prediction. They extract certain features from text data-set and undergo function level analysis. GA based selection of feature subset considers the parameters mentioned in Table 2. The data set errors corresponding to training, validation and test set for CNN model is depicted in Table 3. The achieved accuracy factor is shown in Table 4.

Table 2 : Parameters considered for the applied Neuro-Genetic approach

Size of the population	60
Total number of generations	120
Crossover function	0.5
Rate of mutation	0.04
Category of mutation and selection	Uniform and Rank based
Generation Limit	12
Time limit	Infinite

Table 3: Classification of errors from the dataset

Attribute set	Weights	Fitness	Training error	Validation error	Test error
Full feature subset(13)	**500**	**1.84**	**0.65**	**0.389**	**0.625**
Optimal subset(8)	**8**	**1.56**	**0.48**	**0.51**	**0.50**

Table 4: Accuracy levels

Attribute	Training Accuracy	Validation Accuracy	Test Accuracy
Feature Subset	88%	89.35%	74%
Optimal Subset	88%	87.75%	96.25%

TABLE 5: Comparison table with accuracy rate based on other methodologies in literature

Novel Pruning method[15]	68%
Kernel Difference weighted[16]	71%
SVM with Gaussian Kernel[17]	76%
Learning Vector Optimization[18]	77%
Fuzzy Weighted AIRS[19]	81%
NN with fuzzy membership function[20]	81%
MLP with two hidden layers[21]	86%
Neuro-Genetic approach[22]	90%
ISAGSO-WFSVM	92 %
Neuro-Genetic approach with CNN-MDRP	96.25%

The above table produces a clear vision on reduced complexity indicating highest accuracy rate compared to other methods in literature.

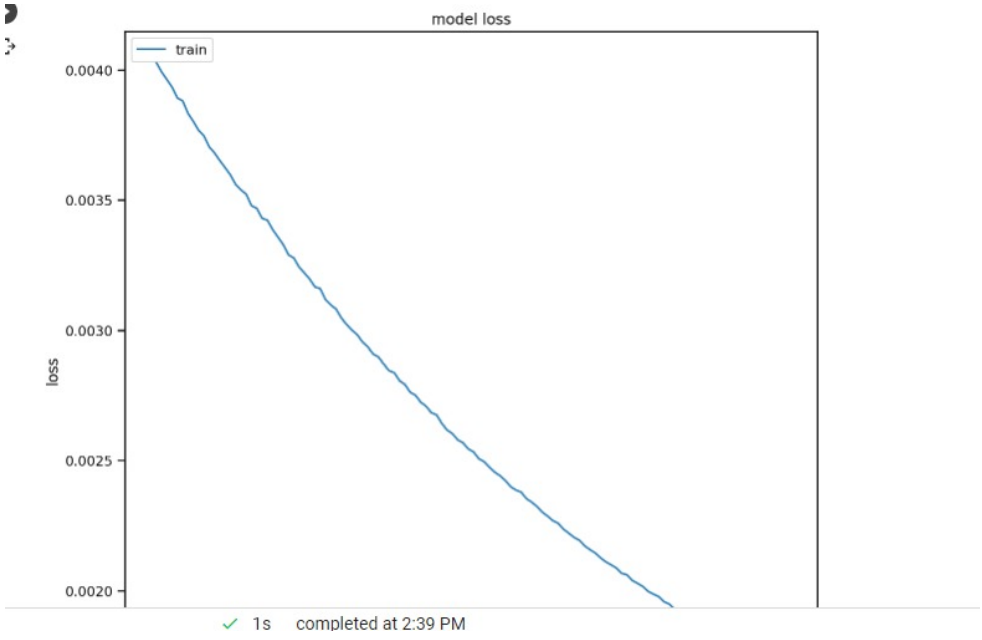

Fig 3: Training model

The training model of the proposed Neuro-Genetic with CNN-MDRP approach is shown in Fig 3. In order to reducing the nature of complexity in CNN-MDRP , integrate Neuro-genetic approach for optimization of features.

```
print(Data.head(20))
```

```
    age  sex  cp  trestbps  chol  fbs  ...  exang  oldpeak  slope  ca  thal  target
0    63    1   3       145   233    1  ...      0      2.3      0   0     1       1
1    37    1   2       130   250    0  ...      0      3.5      0   0     2       1
2    41    0   1       130   204    0  ...      0      1.4      2   0     2       1
3    56    1   1       120   236    0  ...      0      0.8      2   0     2       1
4    57    0   0       120   354    0  ...      1      0.6      2   0     2       1
5    57    1   0       140   192    0  ...      0      0.4      1   0     1       1
6    56    0   1       140   294    0  ...      0      1.3      1   0     2       1
7    44    1   1       120   263    0  ...      0      0.0      2   0     3       1
8    52    1   2       172   199    1  ...      0      0.5      2   0     3       1
9    57    1   2       150   168    0  ...      0      1.6      2   0     2       1
10   54    1   0       140   239    0  ...      0      1.2      2   0     2       1
11   48    0   2       130   275    0  ...      0      0.2      2   0     2       1
12   49    1   1       130   266    0  ...      0      0.6      2   0     2       1
13   64    1   3       110   211    0  ...      1      1.8      1   0     2       1
14   58    0   3       150   283    1  ...      0      1.0      2   0     2       1
15   50    0   2       120   219    0  ...      0      1.6      1   0     2       1
16   58    0   2       120   340    0  ...      0      0.0      2   0     2       1
17   66    0   3       150   226    0  ...      0      2.6      0   0     2       1
18   43    1   0       150   247    0  ...      0      1.5      2   0     2       1
19   69    0   3       140   239    0  ...      0      1.8      2   2     2       1

[20 rows x 14 columns]
```

Fig 4: input model

The row and column wise input models are shown in fig 4.

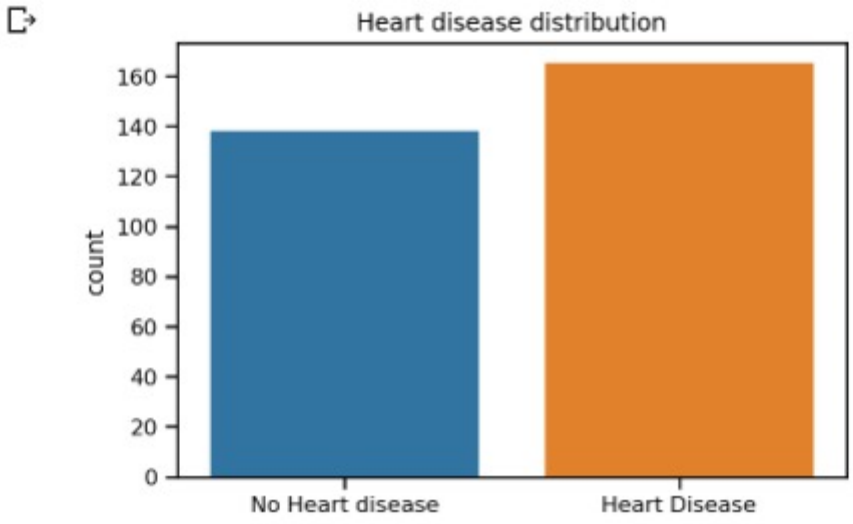

Fig 5: Heart disease prediction results

In this proposed work CNN-MDRP classifier is designed for accurate prediction of heart diseases long with neuro genetic approach. It classifies the input data into two classes such heart disease or no heart disease which is shown in Fig 5. Then gender based classification is represented in Fig 6.

```
plt.legend(['Female', 'Male'])
f.set_title("Heart disease by
f.set_xticklabels(['No Heart d
plt.xlabel("");
plt.rcParams["figure.figsize"]
```

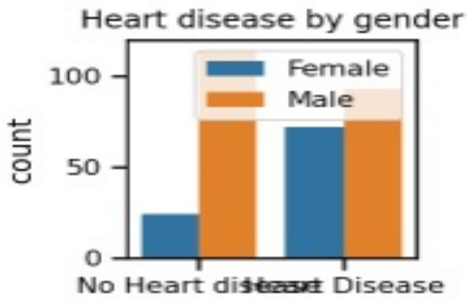

Fig 6 : Gender based classification

```
print('Model Accuracy = ',score*100)
```

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
dense (Dense)                (None, 30)                420
_________________________________________________________________
dense_1 (Dense)              (None, 20)                620
_________________________________________________________________
dense_2 (Dense)              (None, 1)                 21
=================================================================
Total params: 1,061
Trainable params: 1,061
Non-trainable params: 0
_________________________________________________________________
Model Accuracy =  96.25314712524414
```

Fig 7 : Accuracy results

In this proposed work CNN-MDRP classifier is utilized for accurate prediction of heart diseases. For reducing the nature of complexity in CNN-MDRP integrate Neuro-genetic approach for optimization of features. It improves the accuracy rate.

6. Conclusion

CNN-MDRP along with Neuro-Genetic approach produces higher level of accuracy in chronic heart disease detection. The task performed for preprocessing is through CNN-MDRP classifier which contributes its best in prediction of the disease through various convolutional hidden layers. The features extracted undergo genetic algorithm which is neuro genetic based approach which reduces the complexity of features present in the subset. The genetic algorithm is used for generating an optimized function among the random population. It proves its efficiency by providing accuracy of 96.25%. For a clear picture of accuracy , the percentage of conventional methods are depicted in Table 5. The work has demonstrated improved accuracy value by combining neuro genetic algorithm with CNN-MDRP.

References

[1] SayaliAmbekarandRashmiPhalnikar.(2018).Disease Risk Prediction by Using Convolutional Neural Network,978-1-5386-5257-2/18/$31.00 © IEEE.
[2] U. R. Acharya, H. Fujita, O. S. Lih, M. Adam, J. H. Tan, and C. K. Chua.(2017).``Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network,"Knowl.-Based Syst.,vol. 132, pp. 6271.
[3] Vijay Kumar, G., Bharadwaja, A., Nikhil Sai, N. (2017).“Temperature and heart beat monitoring system using IOT “,Proceedings - International Conference on Trends in Electronics and Informatics, ICEI.
[4] VijayKumarG,krishnachaitanyaT,Pratap.(2016).Mining popular patterns from multidimensional database Parallel and distributed frequent-regular patternmining using vertical format in large databases,Indian Journal of Science and Technology .
[5] D.Shanthi, G.Sahoo and N. Saravanan. (2009).“Evolving connection Weights of ANN using GA with Application to the Prediction of Stroke Disease”,International Journal of Soft computing 4(2): 95-102, ©Medwell Journals.
[6] M.Morita, R,.Sabourin, F.Bortolozzi, and C.Y.Suen.(2003).“Unsupervised feature selection using multi-objective genetic algorithms for handwritten word recognition. In proceedings of the 7th ICDAR, pages666-670, IEEE Computer Society.
[7] C.L.Huang, C.J Wang.(2006).“A GA based feature selection and parameter optimization for SVM”, Expert Systems with Applications, pp.231-240.
[8] VijayKumarG., VishnuSravyaS,SatishG.(2018)."Mining high utility regular patterns in transactional database", International Journal of Engineering and Technology(UAE).
[9] Bhaskaru, O., Sreedevi, M.(2020).”Risk feature aware accurate heart disease prediction system using fuzzy extreme learning machine”, Journal of Advanced Research in Dynamical and Control Systems.
[10] SrinivasKolli M. Sreedevi.(2018).“PROTOTYPE ANALYSIS OF DIFFERENT DATA MINING CLASSIFICATION AND CLUSTERING APPROACHES”, ARPN Journal of Engineering and Applied Sciences.
[11] SrinivasKolli M. Sreedevi.(2018).“Adaptive Clustering Approach to Handle Multi Similarity Index for Uncertain Categorical Data Streams”, Journal of Adv Research in Dynamical & Control Systems, Vol. 10, 04-Special Issue.
[12] SrinivasKolli M. Sreedevi.(2019).”A novel index based procedure to explore similar attribute similarity in uncertain categorical data”, ARPN Journal of Engineering and Applied Sciences.
[13] M. Chen, Y. Hao, K. Hwang, L. Wang and L. Wang.(2017)."Disease Prediction by Machine Learning Over Big Data From Healthcare Communities," in IEEE Access, vol. 5,pp. 8869-8879, 2017.doi: 0.1109/ACCESS.2017.2694446.
[14] SayaliAmbekar and Dr.RashmiPhalnikar.(2018).“Disease prediction by using machine learning”, InternationalJournal of Computer Engineering and Applications,Volume XII, Special Issue.
[15] Ali MirzaMehmood and MrithyunjayaRaoKuppa. (2012).“A novel pruning approach using expert knowledge for data-specific pruning”, Engineering with Computers pp.21-30.
[16] Zuo.W.M, Lu. W.G, Wang K.Q, Zhang.H.(2008).“Diagnosis of cardiac arrhythmia using kernel difference weighted KNN classifier” Computers in Cardiology, pp.253-256.
[17] Uyar A, Gurgen F. (2007).“Arrhythmia Classification Using Serial Fusion of Support Vector Machine and Logistic Regression”, Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, pp.560, 6-8 Sept.
[18] Alaa M. Elsayad.(2010).“Classification of ECG arrhythmia using Learning Vector Quantization Neural Networks” (978-1-4244-5844-8/09/©2009 IEEE), Manuscript received July 30, 2009: revised 1 October 2010.
[19] KemelPolat, SeralSahan, SalihGunes. (2006).“A new method to medical diagnosis; Artificial immune recognition system (AIRS) with fuzzy weighted pre-processing and application to ECG arrhythmia”, Expert systems with applications, Vol.31, Issue 2, pp.264-269.
[20] San-Hong Lee, Jung-Kwon Uhm, Lim J.S.(2007).“Extracting Input Features and Fuzzy Rules for Detecting ECG arrhythmia based on NEWFM”, International Conference on Intelligent and Advanced systems, Division of Software, Kyungwon University, Korea.
[21] M.Meenakshi, H.S.Niranjana Murthy.(2014).“Comparison of ANN based Heart stroke classifiers using Varied folds dataset cross validation”, SPRINGER proceedings of International conference on Intelligent computing, communication & devices .
[22] H.S.Niranjana Murthy, M.Meenakshi.(2014)."Dimensionality Reduction using Neuro-Genetic approach for Early Prediction of coronary heart disease",Proceedings of International Conference on Circuits, Communication, Control and Computing.

Author Details

O. Bhaskaru is presently working as Assistant Professor in Computer Science and Engineering at Sanskrithi School of Engineering, Puttaparthi, Andhra Pradesh. He received Master of Technology in Computer Science & Engineering during 2007-09 from R. G. M College of Engineering and Technology,JNTU Anantpur,Andhra Pradesh. Presently he is pursuing Ph.D. in the field of Data Mining from K L University, Vijayawada, Andhra Pradesh.

Dr. M. Sreedevi is Presently working as Professor in the Department of Computer Science & Engineering, K L University, Vijayawada, Andhra Pradesh, India. She received Ph.D from Acharya Nagarjuna University in 2015. She Published several papers in National and International Journals. Her Research interest includes Data mining and Machine Learning.