Study of classification and prediction of heart disease using ANN

Dr.Ankur Jain, Akash kant, Giya ambasta, Bhaavan goyal, Tina Chelwani, Anshul pandey

VIT BHOPAL University, Bhopal, India

Abstract

Numerous different cardiac conditions fall under the general label of "heart disease." Cardiovascular disease is yet another name for it, which refers to both heart and blood vessel illness. So, due to the abundance of health data available today, computers are already assisting with diagnosis utilising machine learning (ML) techniques. We studied several datasets for the algorithms recurrent neural network (RNN), convolutional neural network (CNN), k-nearest neighbours (KNN), artificial neural network (ANN), and random forest while reading roughly 45 papers on these topics. In order to get the best accuracy, we used the ANN algorithm for model training. To classify and predict heart disease, our paper uses Artificial Neural Network (ANN) trained through back-propagation. ANN recognises data trends by iterating through layers of functions to create models. To achieve the optimum outcome, we combined many datasets. The results showed a high accuracy of 86.75%. We may thus draw the conclusion that ANN can be employed for heart disease categorization and prediction.

INTRODUCTION

a) Application: It needs a bit of model training which will be used in the nursing phase of heart patients in future.

b) Focus and purpose: The main focus is to use ANN to analyse the medical Cleveland dataset which was published on UCI repository to predict the likelihood of developing heart diseases and take necessary preventative measures.

d) Broad Problem: The broad problem is to develop an accurate and efficient method for early diagnosis and prevention of heart disease.

f) The Literature review /previous work

DATASET

The heart disease prediction system, the dataset in all the methodologies discussed in this paper is Cleveland hospital heart disease dataset available on UCI machine repository. The features considered from this dataset are mentioned below, As this dataset contains 303 instances and 76 attributes, but only 14 are ref erred to by other publishers.

1.	age	Age in years
2.	sex	sex
3.	cp	Chest pain type
4.	trestbps	Resting blood pressure on admission to hospital in mm/Hg
5.	chol	Serum cholesterol in mg/dl

6.	fbs	Fasting blood sugar is greater than 120 mg/dl or not
7.	restecg	Resting electrocardiographic results
8.	thalach	Maximum heart rate achieved
9.	exang	Exercise induced angina
10.	oldpeak	ST depression induced by exercise relative to rest
11.	slope	The slope of the peak exercise ST segment
12.	ca	Number of major vessels (0-3) colored by fluoroscopy
13.	thal	The heart status
14.	num	Diagnosis of heart disease

1. Title/Authors: Classification of Heart disease using Artificial Neural Network. / Voon Khai Tick, Ng Yung Meeng, Nur farahiyah Mohammad, Nor Hazlyna Harun, Hiam Alquran and Mohamad Mohsin.

Objectives: To improve the reliability of diagnosis and reduce the dependence on human experts by doing the test to investigate whether there are any heart attacks that they snapped in order for the physicians to further identify them.

Work done: The accuracy of prediction is compared within different ML algorithms. After this they have focused over ANN and collected the data after it training it by Back propagation to classify heart disease.

Tools/Method/Algorithms used: Back-propagation with ANN

Future works: The future work as mentioned will be through investigations for real-world datasets instead of just theoretical methods and simulations.

2.Title/Authors: Heart Disease Prediction Using Artificial Neural Network. / Rachana Deshmukh, Omeshwari Bhomle, Apurva Chimote, Shubhada Lunge, Shruti Dekate, Payal Gourkhede, Sonali Rangari

Objectives: Prediction of heart disease by an automated medical diagnosis system based on Machine Learning to provide accurate diagnosis than traditional way and reduce the cost of treatment.

Work done: They have worked over Cleveland heart disease dataset and used 13 out of 75 raw features. ANN along with back-propagation then used to develop the system. The model gives realistic results over experimental results with 100% accuracy.

Tools/Method/Algorithms used: Data Mining, Back-propagation with ANN

Future works: This paper contains the experimental prediction results as it can be used for diagnosis over major prediction of heart disease.

3.Title/Authors: Heart Disease Prediction Using Artificial Neural Network and Image Processing. / Mamta Rani, Aditya Bakshi, Kundan Munjal, Anurag Singh Tomar.

Objectives: To remove the features in numbers by image processing so that ANN may give the results with minimum featured elements.

Work done: This Paper contains a detailed explanation of heart disease and in this investigation paper they have joined image processing procedures to ANN to envision the coronary illness in patients. Image processing is helping to remove more features that will help to get an exactness in results.

Tools/Method/Algorithms used: Image processing method , ANN algorithms

Future works: In future this investigation of heart disease can help us with anticipating the guts illness proficiently. Feature extraction will help to dispose of gathering that enhances the affirmation rate with the smallest amount of element.

4.Title/Authors: Analysis of heart disease prediction system using Artificial Neural Network/ Akruti Dave, Prof. Gayatri Pandi.

Objectives: To find out the best approach over other neural networks approaches and also using ann for better accuracy of results.

Work done: Detail explanation and comparison over various research papers, Ann various uses are explained and worked over with the combined neural approach with back-propagation to train the data.

Tools/Method/Algorithms used: Matlab tool, backpropagation approach, Combined approach neural network.

Future works: The future research may focus on hybrid approaches as well as ensemble neural network approaches to improve the prediction capability with better accuracy.

5.Title/Authors: Diagnosis of heart failure from imbalance datasets using multilevel classification / Dengao Li,Chao Zheng

Objectives: This study establishes an automatic heart failure diagnosis system based on deep learning from imbalance datasets.

Work done: On the testing set, this method improves the average accuracy by 3,78% compared to the traditional random under-sampling method, and the accuracy of our end-to-end multi-class classification experiment is 84.44%.

Tools/Method/Algorithms used: under-sampling method combined with instance Selection, multi-level classification model

Future works: To have a method that can address data imbalance problems and assist overburdened healthcare systems

6.Title/Authors: Automatic diagnosis of cardiovascular diseases using wavelet feature extraction and convolutional capsule network/ Imane El Boujnouni

Objectives: A diagnosis system to classify-eight different diseases using a capsule network with wavelet decomposed images of short ECG segments.

Work done: A computer aided diagnosis system based on the combination of wavelet transform features and capsule network is proposed for the automated classification of eight classes of cardiovascular diseases.

Tools/Method/Algorithms used: the focal loss is adopted to solve the class imbalance issue and 5-fold cross-validation and train-test split techniques

Future works: use the superiority of our approach over state of the art techniques using the same dataset.

7.Title/Authors: An efficient honey badger based Faster region CNN for chronic heart Failure prediction/ S,Irin Sherly

Objectives: To extract the crucial information present in the ECG signal by visual analysis and classify the different abnormalities exhibited in the ECG signal.

Work done: this paper develops an automatic CHF prediction system to assist cardiologists to achieve this target, HBA-FRCNN is proposed for CHF prediction with high accuracy. The HBA-FRCNN classifies the abnormal ECG signals from BIDMC and MIT-BIH database.

Tools/Method/Algorithms used: Honey Badger Algorithm

Future works: Effectiveness of the FRCNN technique in predicting arrhythmias and CHF in patients by taking the increased number of features in the ECG signal

8.Title/Authors: Prediction of single-stranded and double-stranded DNA-binding protein using convolutional neural networks/ Farnoush Manavi,Alok Sharma ,Ronesh Sharma

Objectives: To accurately predict DSB and SSB

Work done: CNN-Pred can enhance the DSB and SSB prediction accuracies by 4.28% on the independent test.

Tools/Method/Algorithms used: SVM, RF,KNN

Future works: Extract evolutionary-based features in the form of mono-gram and bi-gram profiles using position specific scoring matrix.

9.Title/Authors: Prediction of Heart Diseases using Random Forest / Madhumita Pal and Smita Parija

Objectives: To predicts the occurrence of heart disease of a patient using random forest algorithm

Work done: Using random forest algorithm,we obtained accuracy of 86.9% for prediction of heart disease with sensitivity value 90.6% and specificity value 82.7%.From the receiver operating characteristics, we obtained the diagnosis rate for prediction of heart disease

Tools/Method/Algorithms used: Data mining technology, random forest algorithm

Future works: Health care is a major application of data mining. Heart Disease is the most dangerous life-threatening chronic disease globally.This algorithm is the most efficient algorithm for classification of heart disease

10.Title/Authors: Effective Heart Disease Prediction / Senthilkumar Mohan, Chandrasegar Thirumalai,Gautam Srivastava

Objectives: To predict heart disease using the hybrid random forest method.

Work done: In this paper, we propose a novel method that aims at finding significant features resulting in improving the accuracy in the prediction of cardiovascular disease.

Tools/Method/Algorithms used: Hybrid random forest with a linear model .We produce an enhanced performance level with an accuracy level of 88.7% through the prediction model for heart disease with the hybrid random forest with a linear model.

Future works: For prediction of heart disease as it is here visible that hybrid random forest with a linear model algorithm is the most efficient algorithm for classification of heart disease.

11. Title/Authors: Machine learning for real time prediction of complications in critical care: a retrospective study/ Alexander Meyer, Dina Zverinski, Boris Pfahringer, Jörg Kempfert, Titus Kuehne, Simon H Sündermann, Christof Stamm, Thomas Hofmann, Volkmar Falk, Carsten Eickhoff

Objectives: The goal of this work was to use deep machine learning approaches to predict serious problems during critical care following cardiothoracic surgery in real time.

Work done: This RNN model outperformed the clinical tools in terms of accuracy, resulting in a distinct separation of graphs. In addition, hemorrhage and renal failure were predicted with the highest level of accuracy.

Tools/Method/Algorithms used: Deep Learning method - RNN

Future works: The final model will soon be incorporated into EHR systems, guaranteeing that all relevant data is available in real time and that no extra data entry is required.

12.Title/Authors: ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network / Zhaohan Xiong, Martyn P Nash, Elizabeth Cheng , Vadim V. Fedorov, Martin K Stiles, Jichao Zhao

Objectives: Cardiovascular illness is one of the major causes of mortality globally, and irregular heart rhythms, such as Atrial Fibrillation (AF), are the most prevalent kind of cardiac arrhythmia, thus they intended to develop an automated method for detecting AF from ECGs.

Work done: RhythmNet was developed and tested. The major challenge they encountered in this effort was over-fitting on the cross validation set, demonstrating the difficulties of training large neural networks like the one presented in this paper with minimal data. This was improved by increasing dropout or training the network for less time during cross-validation.

Tools/Method/Algorithms used : Convolutional Neural Networks, Deep Learning, Recurrent Neural Networks

Future works: The current detection accuracy for AF is 87%, which may be increased by including more databases during training and testing.

13.Title/Authors: Full Left Ventricle Quantification via Deep Multitask Relationships Learning/ Wufeng Xue, Gary Brahm, Sachin Pandey, Stephanie Leung, Shuo Li

Objectives: As cardiac left ventricular measurement is one of the most significant clinical tasks for identifying and diagnosing heart disorders. As a result, they endeavoured to create an end-to-end integrated framework of the cardiac left ventricle for full quantification.

Work done: DMTRL integrated the benefits of CNN-based cardiac MR representation with RNN-based dynamic modelling for cardiac sequences, achieving accurate assessment for all three types of left ventricle indices and cardiac.

Tools/Method/Algorithms used : RNN, CNN and deep multitask relationship learning (DMTRL).

Future works: Since it is capable of collecting varied patterns during the cardiac cycle, it can be used in clinical practice.

14.Title/Authors: Hybrid deep learning model using recurrent neural network and gated recurrent unit for heart disease prediction / Surenthiran Krishnan, Pritheega Magalingam, Roslina Ibrahim

Objective: Infusing AI and ML, deep learning technologies in the healthcare system to increase accuracy of prediction of heart disease.

Work done: During model training, the RNN with GRU selects features for categorization and predicts diseases based on priority. The RNN has the advantage of processing data instances that are unrelated to previous occurrences.

Tools/Method/Algorithms used : RNN, Long short term memory (LSTM), Gated Recurrent Unit (GRU)

Future works: This predictive deep learning model might be one of the AI and ML technologies that can potentially enhance the quality of life for billions of people worldwide who suffer from heart disease.

h) Current state of the problem: The current state of our problem is that we have framed a prediction model using ANN which can provide the accurate results whether a person is having heart disease or not.

j) Motivation: Heart disease is the major health problem and as per data available on internet 17% of all deaths in south Asian countries is due to heart disease. The limitations of traditional methods used for predicting heart diseases like lifestyle habits and medical history do not always provide accurate results. After all the comparisons between different prediction algorithms we found that by using ANN we can provide the best accurate prediction results which can be used for early diagnosis of heart disease.

l) Novelty: This paper is dealing with the full study of different classification of algorithms and after all comparisons we got to know that ANN is giving better accuracy with different attributes. We are using back propagation so that error will not occur while prediction. So, we are giving a full prediction model with better accuracy over predicting heart disease.

p) Key Contents: Backward propagation, Artificial Neural Network, Cleveland dataset

PROBLEM FORMULATION

Modelling of the problem: In this paper, we look at the application of synthetic facts generation to a coronary heart sickness prediction problem. Heart ailment prediction is a nicely-studied type hassle and earlier analyses function as adequate baselines for our evaluation.As mentioned in all the above papers, the researchers are predicting heart disease using machine learning algorithms.

Assumptions: Heart disease is a prime killer in India and for the duration of the sector, the utility of promising technology like device studying to the initial prediction of coronary heart sicknesses will have a profound effect on society. The early prognosis of heart sickness can be a useful resource in making decisions on way of life adjustments in high-threat sufferers and in turn reduce the complications, which may be a fantastic milestone within the discipline of medicine. The number of humans facing coronary heart sicknesses is on a boost each 12 months. This prompts for its early prognosis and remedy. The utilisation of appropriate era guides in this regard can prove to be extraordinarily useful to the scientific fraternity and sufferers.

Constraints: Such small datasets are traditional in healthcare in which they are okay for human comprehension but insufficient in quantity for device mastering models . One purpose for the small length of clinical information units is because of the shortage of centralised clinical databases. Hospitals, coverage groups, clinics and studies establishments all keep and defend their own patient databases. Patient privacy legal guidelines are the principal purpose for this fragmentation . These statistics sets and features are categorised as small-statistics and often are available with much less than 500 observations and may be as small as 10 observations.

Objectives: The foremost assignment in heart ailment is its detection. There are units available that could predict coronary heart disease but either they're expensive or are not efficient to calculate the hazard of heart sickness in people. Early detection of cardiac

diseases can decrease the mortality rate and average complications. However, it is not possible to display sufferers normal in all cases correctly and sessions of an affected person for twenty-four hours by way of a medical doctor isn't always to be had since it requires greater patience, time and know-how. On account that we have a terrific quantity of facts in today's world, we are able to use numerous machine learning algorithms to analyse the statistics for hidden styles. The hidden patterns can be used for health diagnosis in medicinal records.

METHODOLOGY

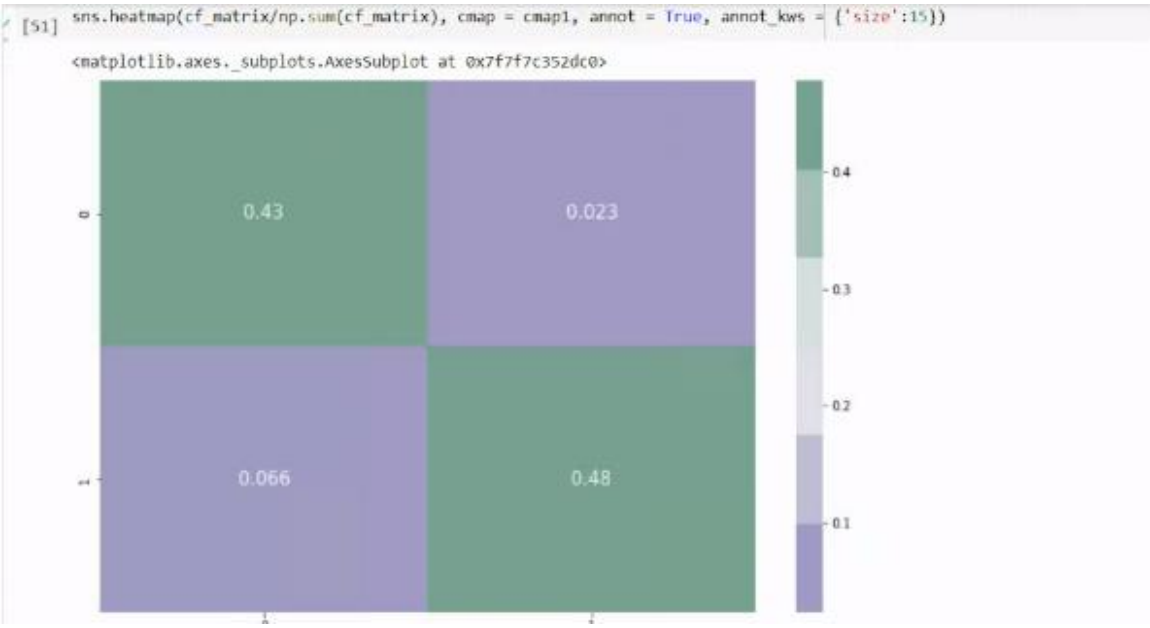
Design:A major challenge in modern healthcare is providing the highest quality services and effective and accurate diagnoses. Heart disease has been considered the leading cause of death worldwide in recent years, but it can also be effectively controlled and managed. All accuracy in heart prediction depends on the right time to detect this disease. The proposed work attempts to identify these heart defects at an early stage to avoid catastrophic outcomes. Analyse massive medical data records generated by healthcare professionals and extract valuable insights. Machine learning algorithms are the means of extracting valuable and hidden information from the large amount of data available.

Analysis: This article presents an analysis of various machine learning algorithms. The algorithm used in this article is a Nearest Neighbour (KNN) algorithm that can be useful to medical practitioners or medical analysts in accurately diagnosing heart disease. This paperwork includes a review of journals, published articles, and recent cardiovascular data.

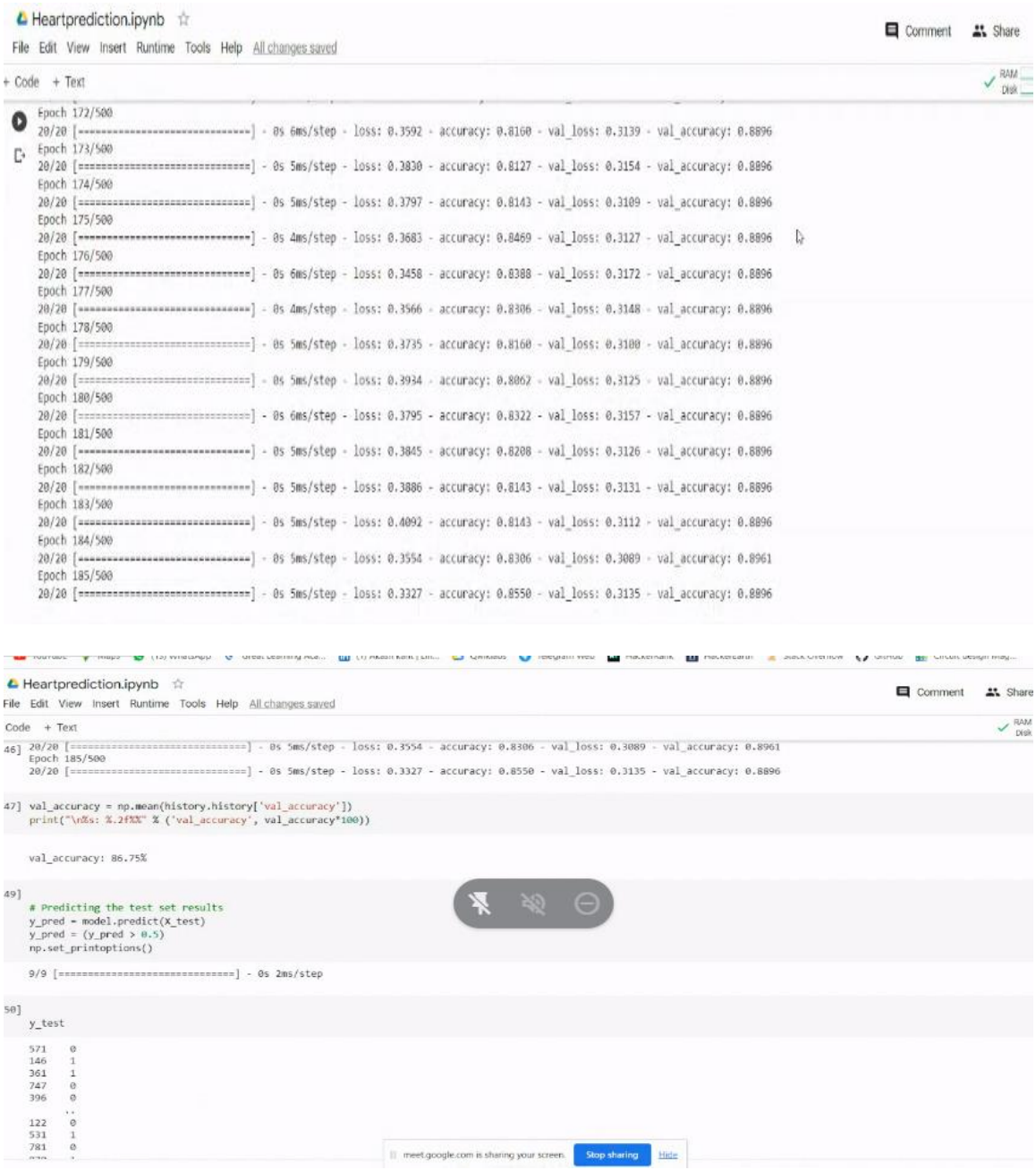
Implementation: Deep learning is a technology that mimics the human brain in that it is composed of multiple layers of nerve cells, just like the human brain. The network thus formed consists of an input layer, an output layer, and one or more hidden layers. With

the start of our model we have worked over in the Python language. as the configuration of the system on which we have applied the code is a system with WINDOWS 11 operating system and with I5 feature. Inside the model the libraries which have been used are pandas, numpy, sklearn, seaborn, matplotlib, keras, with the following modules we have trained the dataset as with 13 features inside which. We also compare the dataset as to whether it is balanced or not. After that we trained the ANN model with backward propagation. The final results show an accuracy of 88.5%. The full working of the model is done on the google collaborative platform.

RESULTS



The accuracy of our neural network for the prediction of heart disease is 88.5% and it ranges btween 80% to 89%.



CONCLUSIONS:

The back propagation used with ANN algorithm is helping to first train the model and lay down all the results with different features based on the clinical dataset. ANN can learn by processing labelled training dataset. After going through layers of functions, ANN can identify the trends of the data from a model. The model can be used to perform

classification with input testing data. More training data will improve model quality and increase accuracy. As it showing a very accurate research model.

REFERENCES

1. “Classification of Heart disease using Artificial Neural Network.” Voon Khai Tick, Ng Yung Meeng, Nur farahiyah Mohammad, Nor Hazlyna Harun, Hiam Alquran and Mohamad Mohsin. DOI 10.1088/1742-6596/1997/1/012022

2. “ Heart Disease Prediction Using Artificial Neural Network.” Rachana Deshmukh, Omeshwari Bhomle, Apurva Chimote, Shubhada Lunge, Shruti Dekate, Payal Gourkhede, Sonali Rangari. DOI 10.17148/IJARCCE.2019.8119

3. “Heart Disease Prediction Using Artificial Neural Network and Image Processing.” Mamta Rani, Aditya Bakshi, Kundan Munjal, Anurag Singh Tomar. **DOI:** 10.1101/ESCI48226.2020.9167669

4. “Analysis of heart disease prediction system using Artificial Neural Network” Akruti Dave, Prof. Gayatri Pandi.

5. Diagnosis of heart failure from imbalance datasets using multilevel classification / Dengao Li,Chao Zheng

6. Automatic diagnosis of cardiovascular diseases using wavelet feature extraction and convolutional capsule network/ Imane El Boujnouni

7. An efficient honey badger based Faster region CNN for chronic heart Failure prediction/ S,Irin Sherly

8. Prediction of Heart Diseases using Random Forest / Madhumita Pal and Smita Parija

9. Prediction of single-stranded and double-stranded DNA-binding protein using convolutional neural networks/ Farnoush Manavi,Alok Sharma ,Ronesh Sharma

10. Effective Heart Disease Prediction / Senthilkumar Mohan, Chandrasegar Thirumalai,Gautam Srivastava

11. Machine learning for real time prediction of complications in critical care: a retrospective study/ Alexander Meyer, Dina Zverinski, Boris Pfahringer, Jörg Kempfert, Titus Kuehne, Simon H Sündermann, Christof Stamm, Thomas Hofmann, Volkmar Falk, Carsten Eickhoff

12. ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network / Zhaohan Xiong, Martyn P Nash, Elizabeth Cheng , Vadim V. Fedorov, Martin K Stiles, Jichao Zhao

13. Full Left Ventricle Quantification via Deep Multitask Relationships Learning/ Wufeng Xue, Gary Brahm, Sachin Pandey, Stephanie Leung, Shuo Li

14. Hybrid deep learning model using recurrent neural network and gated recurrent unit for heart disease prediction / Surenthiran Krishnan, Pritheega Magalingam, Roslina Ibrahim