Asian Journal of Convergence in Technology **Volume V Issue I**
ISSN NO: 2350-1146 I.F-5.11

Deep learning based heart disease prediction

Ashwini S. Phalke,*Student,IT Department,Vivekanand Education Society's Institute of Technology* , Dr Shanta Sondur,*Professor, IT Department,Vivekanand Education Society's Institute of Technology*

***Abstract*—Data mining is the process of data analyzing from various perspectives and combining it into useful information. This technique is used for finding heart disease. Based on risk factor the heart diseases can be defined very easily. The main aim of this work is to evaluate different classification techniques in heart diagnosis. First, the ECG numeric dataset is extracted and preprocess them. After that using extract the features that is condition to be find to be classified by Convolution Neural Network (NN).Compared to existing; Convolution Neural Network provides better performance. After classification, performance criteria including accuracy, precision, F-measure is to be calculated. Compared to KNN, Convolution Neural Network provides better performance. The comparison measure expose that Convolution Neural Network is the best classifier for the diagnosis of heart disease on the existing dataset.**

Keywords— Data mining, Heart diagnosis, Convolution Neural Networks (CNN).

I. INTRODUCTION

The main reason for death worldwide, including South Africa is heart attack diseases and possible detection at an earlier stage will prevent these attacks. Medical practitioners generate data with a wealth of concealed information present, and it's not used effectively for predictions. For this reason, the research converts the unused data into a dataset for shaping using different data mining techniques. People die having encountered symptoms that were not taken into considerations. There is a requirement for medical practitioners to defined heart disease before they occur in their patients.

The features that increase the chances of heart attacks are smoking, lack of physical exercises, high blood pressure, high cholesterol, unhealthy diet, detrimental use of alcohol, and high sugar levels. Cardio Vascular Disease (CVD) constitutes coronary heart, cerebro-vascular or Stroke, hypertensive heart disease, congenital heart, peripheral artery, rheumatic heart disease, and inflammatory heart disease. Data mining is a knowledge discovery technique to examine data and encapsulate it into useful information. The current research intends to forecast the probability of getting heart disease given patient data set. Prophecies' and descriptions are principalgoalsofdatamining;inpracticepredictionindata mining involves attributes or variables in the data set to locate unknown or future state values of other attributes. Description emphasize on discovering patterns that describes the data to be interpretedbyhumans. Feed forward neural network multiple hidden layers in artificial neural networks is usually known as Deep Neural Networks (DNNs). Convolutional Neural Networks (CNN) is one kind of feed forward neural network. In 1960s, when Hubel and Wiesel researched the neurons used for local sensitive orientation-selective in the cat's visual system, they found the special network structure can effectively reduce the complicatedness of Feedback Neural Networks and then proposed Convolution Neural Network. CNN is an competent recognition algorithm which is extensively used in pattern recognition and image processing. It has many appearance such as simple structure, less training parameters andadaptability.

II. . RELATEDWORK

Nearest neighbor (KNN) is very simple, most popular, highly efficient and effective technique for pattern recognition. KNN is a straight forward classifier, where parts are classified based on the class of their nearest neighbor. Medical data bases are big volume in nature. If the data set contains excessive and irrelevant attributes, classification may create less accurate result. Heart disease is the best cause of death in INDIA. In Andhra Pradesh heart disease was the best cause of mortality accounting for 32%of all deaths, a rate as high as Canada (35%) and USA. Hence there is a need to define a decision support system that helps clinicians to take precautionary steps. In this work proposed a new technique which combines KNN with genetic technique for effective classification. Genetic technique perform global search in complex large and multimodal landscapes and provide optimal solution[1].

Image de-noising includes the manipulation of the image data to produce a visually high quality image. The Non Local means filter is originally designed for Gaussian noise removal and the filter is changed to adapt for speckle noise reduction. Speckle noise is a firstly source of medical ultrasound imaging noise and it should be filtered out. This work reviews the existing Non-Local Means based filters for image denoising [2].

This work has analyzed prediction systems for Heart disease using more number of input attributes. The work uses medical terms such as sex, blood pressure, cholesterol like 13

attributes to predict the likelihood of patient getting a Heart disease. Until now, 13 attributes are used for prediction. This research work added two more attributes i.e. obesity and smoking. The data mining classification algorithms, namely Decision Trees, Naive Bayes, and Neural Networks are analyzed on Heart disease database [3].

Medical Diagnosis Systems play important role in medical practice and are used by medical practitioners for diagnosis and treatment. In this work, a medical diagnosis system is defined for predicting the risk of cardiovascular disease. This system is built by combining the relative advantages of genetic technique and neural network. Multilayered feed forward neural networks are particularly adapted to complex classification problems. The weights of the neural network are determined using genetic technique because it finds acceptably good set of weights in less number of iterations[4].

A wide range of heart condition is defined by thorough examination of the features of the ECG report. Automatic extraction of time plane features is valuable for identification of vital cardiac diseases. This work presents a multi-resolution wavelet transform based system for detection 'P', 'Q', 'R', 'S', 'T' peaks complex from original ECG signal. 'R-R' time lapse is an important minutia of the ECG signal that corresponds to the heartbeat of the related person. sudden increase in height of the 'R' wave or changes in the measurement of the 'R-R' stand for various abnormality of human heart. Similarly 'P-P', 'Q-Q', 'S-S', 'T-T' also corresponds to various anomalies of heart and their peak amplitude also envisages other cardiac diseases. In this proposed method the 'PQRST' peaks are considered and reserved over the entire signal and the time interval between two ensuing 'R' peaks and other peaks interval are measured to find anomalies in behavior of heart, if any[5].

The ECG signal is well known for its nonlinear changing behavior and a key characteristic that is utilized in this research; the nonlinear component of its dynamics changes more automatically between normal and abnormal conditions than does the linear one. As the higher-order statistics (HOS) maintain phase information, this work makes use of one-dimensional slices from the higher-order spectral region of normal and ischemic subjects. A feed forward multilayer neural network (NN) with error back propagation (BP) learning technique was used as an automated ECG classifier to find the possibility of recognizing ischemic heart disease from normal ECG signals[6].

are certain flaws in existing work such as,1.Poor validation of methodology, 2. Feature selection is not accurate.3.Lack in Accuracy, 4.Less No of records had been used. i.e. An MIT-BIH arrhythmia database consisting of 48 recordings from 1975 and 1979 were used. Whereas in this work dataset that will be used contains 710,369 evaluable

Automatic ECG classification is a showing tool for the cardiologists in medical diagnosis for effective treatments. In this work, propose efficient techniques to automatically classify the ECG signals into normal and arrhythmia affected (abnormal) parts. For these categories morphological features are extracted to illustrate the ECG signal. Probabilistic neural network (PNN) is the modeling technique added to capture the distribution of the feature vectors for classification and the performance is calculated. ECG time series signals in this work are bind from MIT-BIH arrhythmia database[7].

The heart diseases are the most extensive induce for human dying. Every year, 7.4 million deaths are attributed to heart diseases (cardiac arrhythmia) including 52% of deaths due to strokes and 47% deaths due to coronary heart diseases. Hence identification of different heart diseases in the primary stages becomes very important for the protection of cardiac related deaths. The existing conventional ECG analysis methods like, RR interval, Wavelet transform with classification algorithms, such as, Support Vector machine K-Nearest Neighbor and Levenberg Marquardt Neural Network are used for detection of cardiac arrhythmia Using these techniques large number of features are extracted but it will not identify exactly the problem[8].

According to Gupta and Chatur (2012), ECG analysis is highly beneficial for diagnosing the heart diseases. The researchers have applied supervised artificial neural networks and data mining techniques for ECG classification. The application of both the techniques is then compared, where the use of ANN is rendered as more effective[7].

III. PROPOSEDSYSTEM

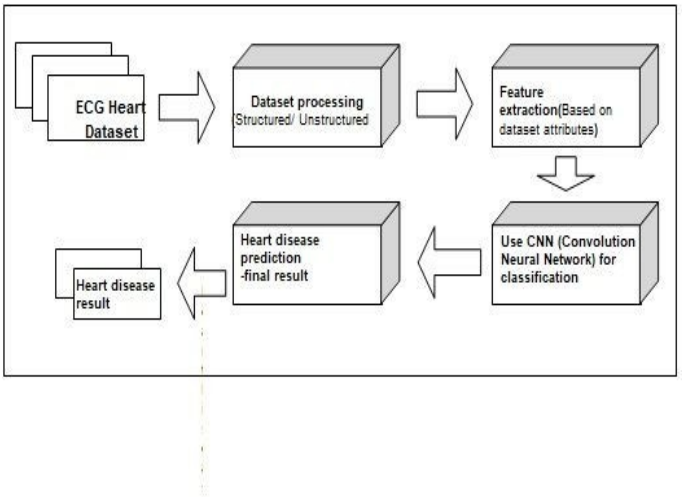
In recent times, different methods have been applied for efficient classification of arrhythmia and heart diseases such as Genetic algorithms, Fuzzy Logic, SelfOrganizing Map, Bayesian, Hidden Markov Models and SVMs. to help health care professionals in the diagnosis of heart disease.There are many machine learning algorithms with some appropriate choice of features which deal with time series data. But as it stated in introduction the work presents use of feature extractors – convolutional neural networks (CNN) in special case. This feature extraction should free the developers from use of expert knowledge and handcrafted features. It is not remove expert knowledge at all from development process but decrease the time to show the first model of ECG classification algorithm.Considering the study of ECG signal patterns in diagnosis, it is very comprehensive time-consuming, and gradual thus requiring the computerized means for effective diagnosis of cardiac diseases. Neural networks are used for classification once the pre-processingis done and feature extraction from the ECG signal. Existing work has developed a diagnostic system by using an artificial neural network classifier for heart disease prediction.There

ECG records over a 17-year study period. Validation can be further improve with this.The proposed system offers a predication accuracy of 90%, which has further scope for improvement. This work is based on detection of heart diseases using deep learning(Convolution neural network)Which is an advanced technology to AI. This work is

used for finding various heart diseases. Based on risk factor the heart diseases can be predicted very easily. The main aim of this paper is to predict the heart diagnosis.First,

the ECG numeric dataset is extracted and preprocess them. After that using extract the features that is condition to be find to be classified by Convolution Neural Network (NN).Compared to existing; Convolution Neural Network provides better performance. After classification, performance criteria including accuracy, precision, F-measure is to be calculated. The comparison measure reveals that Convolution Neural Network is the best classifier for the diagnosis of heart disease on the existing data.

III. PROPOSED SYSTEMARCHITECTURE



IV.ALGORITHM Convolution NeuralNetwork

Generally, the structure of CNN includes two layers one is feature extraction layer, the input of each neuron is connected to the local receptive fields of the earlier layer, and extracts the local feature. Once the local features are extracted, the positional relationship between it and other features also will be determined. The other is feature map layer; eachcomputing layer of the network is composed of a plurality of feature map. Every feature map is a plane, the weight of the neurons in the plane are equal. The structure of feature map uses the sigmoid function as activation function of the convolution network, which makes the feature map have shift invariance. Besides, since the neurons in the same mapping plane share weight, the number of free parameters of the network is reduced. Each convolution layer in the convolution neural network is followed by a computing layer which is used to calculate the local average and the second extract, this unique two feature extraction structure reduces the resolution.

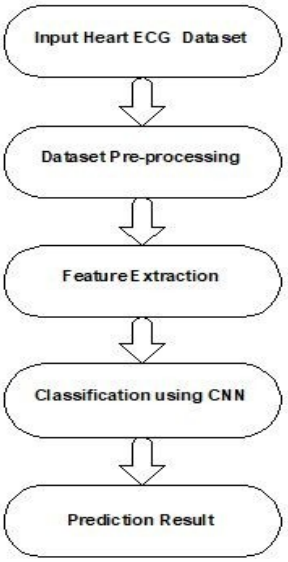


fig.2 Project Flow Diagram

Algorithm Process flow :

Step 1: Select the dataset.

Step 2: Feature selection using information gain and ranking

Step 3: Classification algorithm

Step 4: Each Feature calculate fx value of input layer

Step 5: bias class of each feature calculate

Step 6: Next produce the feature map it goes to forward pass input layer

Step 7: Calculate the convolution cores in a feature pattern

Step 8: Produce sub sample layer and feature value.

Step 9: Back propagation input deviation of the kth neuron in output layer.

Step 10: Finally give the selected feature and classification results.

V. DATASETUSED

Here, we will use ECG numeric dataset and preprocess them for further processing.

VI. TECHNIQUES

HardwareSpecification

1. System: Pentium IV 2.4GHz.
2. Hard Disk: 40GB
3. Floppy Drive: 44Mb.

4. Monitor: 15 VGA Color

SoftwareSpecification

1. Operating system: Ubuntu16.04.
2. Coding Language:Python.
3. IDE: EclipseLuna.

CONCLUSION

The experiment is organized with the dataset of Heart Disease by contemplating the single and multilayer neural network modes. Heart Disease dataset is taken and analyzed to predict the asperity of the heart disease. A convolution neural network approach is used to predict the asperity of the disease. The dataset is preprocessed to make it applicable for classification.

The convolution neural network approach to generate efficient classification rules is proposed. To perform classification task of medical data, the neural network is trained using Convolutions technique. With the help of this system doctors could monitor & diagnose patient's condition continuously & could advise them precautions if any. Also determine possibilities of heart attack based on heart activities. timely diagnosis of heart disease plays a important role in preventing its progress, assess related risk factors and reducing related treatment costs. Also helps in improving treatment policies and prevention of any mistake in hospitals.

References

[1] Algorithm M.Akhil jabbar B.L Deekshatulua Priti Chandra International "Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm" Conference on Computational Intelligence: Modeling Techniques and Applications (CIMTA)2013.

[2] Beshiba Wilson, Dr.Julia Punitha Malar Dhas "A Survey of Non-Local Means based Filters for Image Denoising" International Journal of EngineeringResearch&Technology,Vol.2-Issue10(October–2013.

[3] Chaitrali S Dangare "Improved Study Of Heart Disease Prediction System Using Data Mining Classification Techniques", International Journal Of Computer Applications, Vol.47, No.10 (June2012).

[4] Amma, N.G.B "Cardio Vascular Disease Prediction System using Genetic Algorithm", IEEE International Conference on Computing, Communication and Applications,2012.

[5] SayantanMukhopadhyay1,ShouvikBiswas2,AnamitraBardhanRoy3 , Nilanjan Dey4' Wavelet Based QRS Complex Detection of ECG Signal' International Journal of Engineering Research and Applications (IJERA) Vol. 2, Issue 3, MayJun 2012, pp.2361-2365

[6] Sahar H. El-Khafifand Mohamed A. ElBrawany, "Artificial Neural Network-Based Automated ECG Signal Classifier", 29 May2013.

[7] M.Vijayavanan, V.Rathikarani, Dr. P. Dhanalakshmi, "Automatic Classification of ECG Signal for Heart Disease Diagnosis using morphologicalfeatures".ISSN:22293345Vol.5No.04Apr2014.

[8] I. S. Siva Rao, T. Srinivasa Rao, "Performance Identification of Different Heart Diseases Based On Neural Network Classification". ISSN 0973-4562 Volume 11, Number 6 (2016) pp3859-3864.