Cardiac Arrest Prediction to Prevent Code Blue Situation

Mrs. Vidya Zope[1], Anuj Chanchlani[2], Hitesh Vaswani[3], Shubham Gaikwad[4], Kamal Teckchandani[5]

[1]Assistant Professor, Department of Computer Engineering, Vivekanand Education Society's Institute of Technology, Mumbai, India

[2,3,4,5]Student, Department of Computer Engineering, Vivekanand Education Society's Institute of Technology, Mumbai, India

---

Abstract - *Cardiac arrest is defined as the abrupt loss of functioning of the heart in a person who may or may not have been diagnosed with heart disease. It can occur suddenly, or in the wake of other symptoms. If appropriate steps are not taken, cardiac arrest can prove to be fatal. The main cause of cardiac arrest may be irregular heart rhythms. These irregular heart rhythms are called arrhythmias. Ventricular fibrillation is a common arrhythmia associated with cardiac arrest. Ventricular fibrillation occurs when the heart's lower chambers suddenly start beating frantically and do not pump blood. Since cardiac arrest is a major emergency situation, hence Hospitals use code names to alert their staff about an emergency. Code blue is a universally recognized emergency code. Code blue refers to a medical emergency occurring within the hospital.*

Key words: **Decision Trees, Neural Networks, Random Forests J48, Learning Vector Quantization, Fuzzy, Arrhythmia, Code Blue.**

1. INTRODUCTION

The death rate in India due to cardiac diseases is continuously rising whereas in the United States of America this rate has significantly declined. The main reason for the deaths in India is cardiovascular diseases. This underlines the need for efficient cardiac arrest prediction system that taking into account the patient's health record determines the possibility of suffering from cardiac arrest. In this project, we planned to develop a predictive model using machine learning algorithm namely neural networks. This is a supervised machine learning technique. The important purpose of this predictive model was to reduce the computation cost and obtain more accurate results efficiently. Back Propagation Algorithm which is an Artificial Neural Network methodology, was employed.

2. LITERATURE SURVEY

A plethora of work has been done in the field of medical science to develop an efficient system to predict cardiac arrest.

Meghna Sharma and Ankita Dewan in their paper have compared various techniques that can be implied to predict heart diseases. On comparison, they found that the Back propagation algorithm is the best methodology that can be used to develop the required system. In order to solve the drawback of the back propagation algorithm i.e. local minima, they have suggested using Genetic Algorithm as an optimizer.[1]

Varun Kathuria and Prakhar Thapiyal have taken into account the ECG readings of the patients and accordingly have developed a model using supervised learning technique which divides the patients into various cardiac arrhythmic classes. Their main aim is to detect the presence of cardiac arrhythmia and accordingly distinguish it into 13 different classes. [2]

Sheetal Singh, DK Sharma, Sanjeev Bhoi, Sapna Ramani Sardana, in their research paper on code blue situation provide information about general principles of Code Blue. Along with it, they have also employed Naïve Bias Algorithm, J48 Algorithm to predict the possibility of a cardiac arrest. They have conducted their experiment first by taking into account all the attributes available in the patient records and then after performing the attribute selection process.[3]

SY Huang, CH Cheng, PS Hong, AH Chen, and EJ Lin proposed Learning Vector Quantization Algorithm for predicting the presence of heart disease. Learning Vector Quantization Algorithm is a commonly used Artificial Neural Network learning Technique. They have into account 13 unique features from the patient's record and accordingly have trained their model. Hence using these 13 parameters for computation they have developed the heart disease prediction system. The accuracy of the model with which it predicts heart diseases was nearly 80%. [4]

Thendral Puyalnithi and V.Madhu Vishwanathan have used methods like k-fold Cross Validation and Leave One-out. They have utilized 75%of the dataset for training and 25%of the dataset for testing. The methods prescribed above were used for all four algorithms viz. Support Vector Machine, Classification Trees, Random Forest and Naive Bayes.[5]

3. METHODOLOGY

Firstly, the feature selection process was performed. The Dataset selected was analyzed for dirty fuzzy values. From a set of large number parameters, the parameters which were instrumental in determining the output were selected. Also, the attributes which contained mostly null values were discarded.

Different learning rules exist which can be used to a train the model so as to appropriately obtain the desired output. The various learning rules are supervised, unsupervised and reinforced learning. Since we had defined inputs as the parameters from the patient records and ECG readings, therefore we settled for a supervised learning technique. By supervised, we mean well-defined inputs after computation generates well-defined outputs. Since the output will show whether the patient has a chance of suffering from a heart attack or not, the output variables are in category type. Thus it is a classification type of "supervised learning".

Initially, random weights were assigned to the neurons along with the predetermined constant value known as bias. After the first iteration output generated was compared with the desired output, and accordingly after computing the error, weights were adjusted. The process is as shown in Figure 1

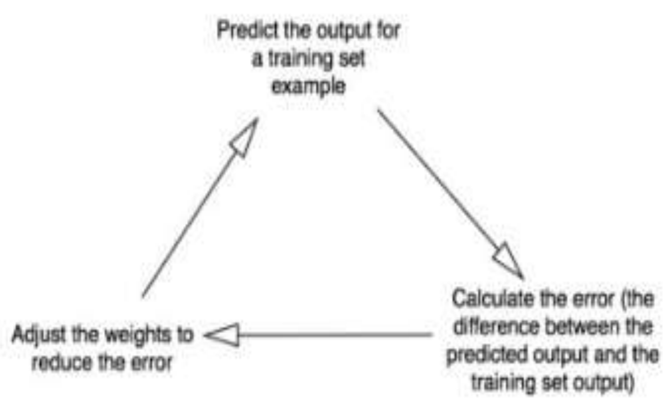

Fig -1: Backpropagation Network Flow

Neural networks are self-learning models. But in order to achieve the desired accuracy, it needs to be trained on some test cases so that when the model encounters a new scenario it accordingly predicts the output based on the past experiences. To reduce training time, back propagation was used.

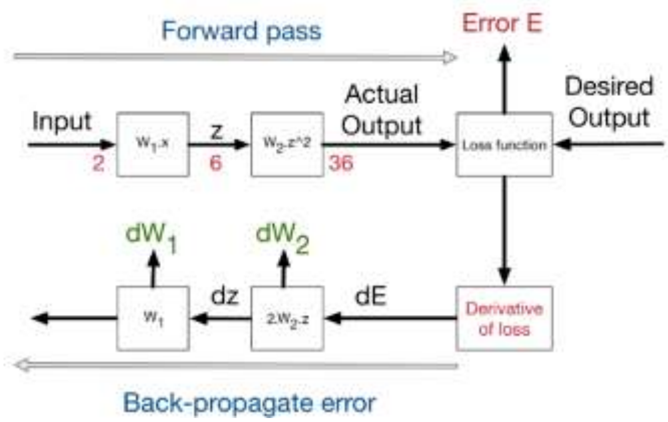

Fig -2: Backpropagation Network

3.1 ARTIFICIAL NEURAL NETWORK (ANN)

The human brain consists of more than 100 billion cells known as neurons. They are connected together by synapses. If inputs to neuron fire, the neuron fires and this process is known as thinking. The ANN developed was of feed-forward type (input layer accepts input, hidden layer after performing computation sends the results to the output layer.) The ANN developed for predicting the presence of heart disease consisting of three layers namely input, hidden, output.

The following figure shows working of the multilayer layer of the perceptron neural network.

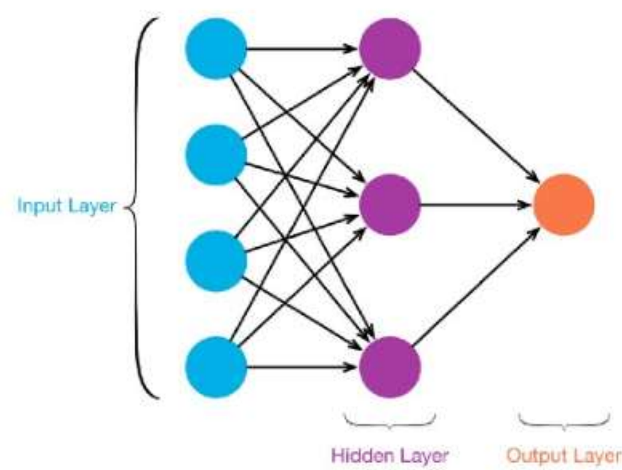

Fig -3: Layers of Neural Network

Referring to [Fig-3], we calculated the dot product as shown below.

$$W \cdot X = w_1x_1 + w_2x_2 + ... + w_mx_m = \sum_{i=1}^{m} w_ix_i$$

Here, w1, w2, w3 are the weights associated with each neuron. With hidden layer, we perform n dot products to get hidden output h: (h1,h2, h3 and so on). We performed dot product for input data set X: X1, X2, and X3. Afterwards, we added a small constant value called as the bias. So the final equation for determining output becomes

$$z = \sum_{i=1}^{m} w_ix_i + bias$$

After computing z, we feed it into the activation function so as to get output for the first neuron of the first hidden layer.

$$h_1^1 = f(z), \ (note\ that\ f()\ needs\ not\ be\ step\ function)$$

The same procedure gets repeated for each neuron of each hidden layer. After calculating for each and every neuron of the hidden layer, we calculated the final output. For the final output, we performed dot product for hidden layer outputs and hidden layer weights. Following is the final equation to calculate the actual output.

$$z = \sum_{i=1}^{n} w_i^{h_1} h_i^1 + bias$$

The above-mentioned procedure is for one hidden layer. Based on the complexity of the network, the number of hidden layers varies.

3.2 BACKPROPAGATION NEURAL NETWORK

Now each neuron was initialized with some random value. To generate output sigmoid function was used as the activation function. The formula is as follows

$$o = \sigma(\vec{w} \cdot \vec{x}) \qquad \sigma(y) = \frac{1}{1+e^{-y}}$$

Where x is the vector of network input values and w is the vector of unit weight values.

Afterwards error δ_k is calculated. Afterwards error signal is propagated to again to all the neurons. Then the error is calculated for the output layer with the equation as shown below.

$$\delta_k \leftarrow o_k(1 - o_k)(t_k - o_k)$$

Here, o_k is the output of the network for the o/p unit k.

Similarly, for each hidden neuron, the hidden error is computed. The equation is as follows

$$\delta_h \leftarrow o_h(1 - o_h) \sum_{k \in outputs} w_{kh}\delta_k$$

Here Wkh is the network weight from hidden unit to output unit k.

Now, we need to update the network weights. The formula for computing it are as follows

$$w_{ji} \leftarrow w_{ji} + \Delta w_{ji} \quad \text{where } \Delta w_{ji} = \eta \delta_j x_{ji}$$

Here δ_j is the learning rate.

3.3 EXPERIMENTAL RESULTS

The prediction of cardiac arrhythmia was done using python with Tensor flow along with Keras for training the neural network.

3.4 DATABASE

The dataset used was taken from the site https://archive.ics.uci.edu/ml/datasets/Arrhythmia. The dataset consists of 452 patient records and 279 attributes. The description of few attributes is provided below.

Table -1: Dataset Attributes

No.	Clinical Features	Description
1	Age	Age
2	Ca	Number of major vessels (0-3) colored by fluoroscopy
3	Chol(mg/dl)	Serum Cholesterol
4	Cp	Chest pain type
5	Exang	Exercise induced angina
6	Fbs	Fasting blood sugar
7	Num	Diagnosis of heart disease
8	Oldpeak	ST depression induced by exercise relative to rest
9	Restecg	Resting electrocardiographic results
10	Sex	Gender
11	Slope	The slope of the peak exercise ST segment
12	Thal	3=normal, 6=fixed defect, 7=reversible defect
13	Thalach	Maximum heart rate achieved
14	Trestbps(mmHg)	Resting Blood Pressure

Table -2: Classification Attributes

No.	Features	Description
1	Normal	Normal parameters
2	Ischemic Changes (Coronary artery)	Blood flow rate to heart via coronary artery
3	Old Anterior Myocardial Infarction	Tissue supply by left coronary artery
4	Old Inferior Myocardial Infarction	Ischemia and infarction to the inferior region of the heart
5	Sinus tachycardia	Sinus rhythm with an changed rate of impulses per min
6	Sinus bradycardia	Heart rate lower than normal condition
7	Premature ventricular contractions (PVC)	Premature heartbeats originating from the ventricles of the heart
8	Supraventricular premature contractions	Atrial contractions triggered by ectopic foci rather than the sinoatrial node
9	Left bundle branch block	A delay or blockage of electrical impulses to the left side of the heart
10	Right bundle branch block	The right ventricle is not directly activated by impulses travelling through the right bundle branch
11	Left ventricular hypertrophy	Enlargement and thickening (hypertrophy) of the walls of your heart's main pumping chamber
12	Atrial fibrillation or flutter	An irregular, often rapid heart rate that commonly causes poor blood flow
13	Others	Extra additional data

The dataset was divided into two groups, 70% training set and 30% testing set.

Now the metrics used for measuring the efficiency of the trained neural network are as follows

Precision: It was basically done so as to determine based on the input of how many were the positive test cases (patients suffering having a risk of cardiac arrest). It is calculated using the formula as shown below

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

Clearly, from the denominator, we can observe that it is total predicted positive.

		Predicted	
		Negative	Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

True Positive + False Positive = Total Predicted Positive

Therefore above-mentioned formula becomes as follows

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$= \frac{True\ Positive}{Total\ Predicted\ Positive}$$

Recall: Same logic was applied while calculating recall

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$= \frac{True\ Positive}{Total\ Actual\ Positive}$$

		Predicted	
		Negative	Positive
Actual	Negative	True Negative	False Positive
	Positive	False Negative	True Positive

True Positive + False Negative = Actual Positive

Recall actually helped to determine how many actual positive our model was able to capture through labelling.

Taking into above-mentioned factors, the back-propagation artificial neural network model was providing the accuracy of 58.069%.

4. CONCLUSION

Our proposed heart disease prediction model was able to provide us with moderate accuracy. But there exists scope for improvement since the existing system is taking into account the reading interpreted from ECG ratings. So, better pre-processing techniques will help us to remove more redundant data plus at the same time more patient records will be required so as to train the model for various different scenarios

REFERENCES

[1] A. Dewan and M. Sharma, "Prediction of heart disease using a hybrid technique in data mining classification," 2015 2nd International Conference on Computing for Sustainable

Global Development (INDIACom), New Delhi, 2015, pp. 704-706.

[2] Varun Kathuria and Prakhar Thapiyal "Prediction of Cardiac Arrhythmia" Manipal Institute of technology, Manipal -576 104, Karnataka, India.

[3] Sheetal Singh, DK Sharma, Sanjeev Bhoi, Sapna Ramani Sardana, Sonia Chauhan Code Blue Policy for a Tertiary Care Trauma Hospital in India, JRFHHA.

[4] A. H. Chen, S. Y. Huang, P. S. Hong, C. H. Cheng and E. J. Lin, "HDPS: Heart disease prediction system," 2011 Computing in Cardiology, Hangzhou, 2011, pp. 557-560.

[5]Thendral Puyalnithi and V. Madhu Viswanathan "Preliminary Cardiac Disease Risk Prediction Based on Medical and Behavioural Data Set Using Supervised Machine Learning Techniques" Indian Journal of Science and Technology, Vol 9(31), DOI: 10.17485/ijst/2016/v9i31/96740, August 2016.

.