Cardiovascular Disease Prediction using Deep learning Algorithm

M.Kavitha[1], A.Ramya[2], V.Mohanraj[3], V.Muthukoodalingam[4]
[1, 2, 3, 4] Dept of Computer Science and Engineering
[1, 2, 3, 4] Kathir college of engineering

Abstract- *In recent days, many populations are affected by the Heart disease which is mainly caused as a result of poor diet and life routine of the people. The improper forecasting done by the doctors and the individual people with their own body are considered to be the main reasons of mortality in the world today. This forecasting can be made accurate by implementing recurrent neural networks (RNN), which can be used to make predictive decisions based upon the past medical records of a particular patient. Deep learning has shown to be effective for making prediction from the large quantity of data produced by the health care industry. This proposed method for heart disease prediction, will help to reduce diagnosis time and improve diagnostic accuracy*

Keywords- Heart disease, forecasting, Disease diagnosis, RNN- algorithm

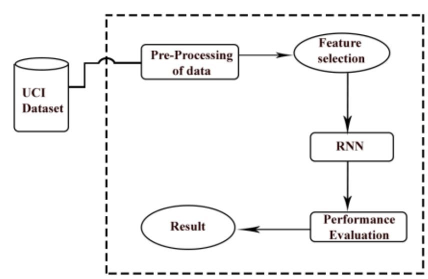
I. INTRODUCTION

In this fast-moving world, people want to live a very luxurious life,so they work like a machine to earn a lot of money and live a comfortable life so in this competition they forget to take care of themselves , they are more stressed in this kind of lifestyle, and they could have blood pressure, sugar at a very young age. This would lead to a major risk of heart disease. Identifying heart disease is difficult due to several contributing factors such as diabetes, high blood pressure, high cholesterol, irregular pulse rate and many other factors. The essence of the heart disease is complex and therefore the condition needs to be treated carefully. Failure to do so may affect the heart or cause premature death. Earlier time decision support system is based on predicate logic and the first expert system called Dendril is used to identify blood-related diseases in humans. Earlier, with the advent of computer processing resources, machine learning concentrated on dataset and manipulation Every approach has its own merits and its own merits. It has been noted that there is a trend in the art of research with regard to many aspects such as lack of accuracy, lack of data independence, a more generic reach and an automated pre-processing step. All major algorithms such as SVM, Naïve Bayes, Decision Tree Neural Network are used to build a support network to predict the risk of a heart attack in humans.

Neural networks are widely seen as the best tool for detecting heart disease and brain disease. The proposed method, which we use, has 13 attributes for predicting heart disease. The restorative field has an extraordinary enthusiasm for a deep learning technique that can make use of every single, accessible credit from a patient to predict illnesses later on. The expectation of illness turns out to be progressively unpredictable due to the relationship of dependence between the different factors. Repetitive neural systems are the ideal answer to this case. The RNN dependent clinical preference strategy has a larger number of points of concern than the normal definition of coronary disease using machine learning. The unique system offers a further path for clinicians to predict heart infections.

In this paper, we propose another protection to safeguard a prescient RNN-dependent clinical choice plan.

II. RELATED WORKS



"Sarath Babu, Vivek EM, Famina KP, Fida K, Aswathi P, Shanid M, Hena M". Heart Disease Diagnosis Using Data Mining TechniqueIn this paper, using only 14 attributes example age, gender,bmi,sugar , down sloping and fat. Less accuracy.

"Jyoti Soni" Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction, this paper using Naive Bayes, Decision List and KNN. Here the Classification based on clustering is not performing well.

S.Deepa and B.Aruna Devi, “Artificial Neural Networks design for Classification of Brain Tumour”.In this system, we manipulate the ability of Back Propagation Neural Networks (BPN) and Radial Basis Function Neural network (RBFN) to categorize mind MRI pictures to either cancer or noncancerous growth instantly. The results revealed outperformance of RBFN criteria in comparison to BPN with category precision of 85.71% which performs as appealing device for category and needs expansion in mind growth analysis.

III. PROPOSED SYSTEM

In this study, we have used an deep learning technique which is the Recurrent Neural Network for effective heart disease prediction. The RNN process starts from pre-processing data phase followed by feature selection which reduces the amount of data for processing and analysis, that makes the prediction performance even more accurate. This system extracts hidden knowledge from the historical database of heart disease suspects, it also predicts the likelihood of patients getting heart disease.

Fig 1. Proposed System

A. DATA PRE-PROCESSING

Heart disease data is pre-processed after collection of varied records. The dataset contains a total of 303 patient records, where 6 records are with some missing values. Those 6 records have been removed from the dataset and the remaining 297 patient records are used in pre-processing .The multi-class variable and binary classification are introduced for the attributes of the given dataset. The multi-class variable is used to check the presence and absence of heart disease. In the instance of the patient having heart disease, the value is set to 1, else the value is set to 0 indicating the absence of heart disease in patient.

TABLE: UCI dataset range and datatype

AGE	Numeric [29 to 77;unique=41;mean=54.4;median=56]
SEX	Numeric [0 to 1;unique=2;mean=0.68;median=1]
CP	Numeric [1 to 4;unique=4;mean=3.16;median=3]
TESTBPS	Numeric [94 to 200;unique=50;mean=131.69;median=130]
CHOL	Numeric[126to564;unique=152;mean=246.49;median=24]
FBS	Numeric [0 to 1;unique=2;mean=0.15;median=0]
RESTECG	Numeric [0 to 2;unique=3;mean=0.99;median=1]
THALACH	Numeric [71 to 202;unique=91;mean=149.61;median=153]
EXANG	Numeric [0 to 1;unique=2;mean=0.33;median=0.80]
OLPEAK	Numeric [0 to 6.20;unique=40;mean=1.04;median=0.80]
SLOPE	Numeric [1 to 3;unique=3;mean=1.60;median=2]
CA	Categorical [5 levels]
THAL	Categorical [4 levels]
TARGET	Numeric [0.00 to 4.00;unique=5;mean=0.94;median=0.00]

TABLE : UCI dataset attribute details

Attribute	Describtion	Type
Age	Patient's age is completed years	Numeric
Sex	Patient's gender (Male represent as 1 and fe-male as 0)	Nominal
Cp	The type of chest pain categorized into 4 values : 1. typical angina, 2. atypical angina, 3. non-angina pain and 4. asymptomatic	Nominal
Trestbps	Level of blood pressure at resting mode (in mm/Hg at the time of admitting in the hospital)	Numeric
Chol	Serum cholesterol in mg/dl	Numeric
FBS	Bloodsugar levels on fasting>120mg/dl:represented as 1 in case of true, and 0 in case of false	Nominal
Resting	Results of electrocardiogram while at rest are represented in 3 distinct values:Normal state represent vale as 0, Abnormality in ST-T wave as value 1, (which may include inversions of T-wave and/or depression or elevation of ST of >0.05mV) and any probability or certainty of LV hypertrophy by Estes' criteria as value 2	Nominal
Thali	The accomplishment of the maximum rate of heart	Numeric
Exang	Angina induced by exercise. (0 depicting 'no' and 1 depicting 'yes')	Nominal
Oldpeak	Exercise-induced St depression in comparison with the state of rest	Numeric
Slope	ST segment measured in terms of the slope during peak exercise depicted in three values : 1. unsloping, 2. flat and 3. downsloping	Nominal
Ca	Fluoroscopy coloured major vessels numbered from 0 to 3	Numeric
Thal	Status of the heart illustrated through three distinctly numbered values. Normal numbered as 3, fixed defect as 6 and reversible defect as 7.	Nominal
Num	Heart disease diagnosis represented in 5 values, with 0 indiacting total absence and 1 to 4 representing the presence in different degrees.	Nominal

B. FEATURE SELECTION AND REDUCTION

From among the 13 attributes of the dataset, two attributes pertaining to age and sex are used to identify the personal information of the patient. The remaining 11 attributes are considered important as they contain vital clinical records. Clinical records are vital to diagnosis and learning the severity of heart condition.

C. CLASSIFICATION MODELLING

1). Recurrent Neural Network (RNN): A Recurrent Neural Network (RNN) is a subclass inside the fake neural system which goes under the expansive field of Deep Learning idea where associations between hubs structure a coordinated chart along with a fleeting grouping. This enables it to show fleeting powerful conduct. In contrast, to encourage forward neural systems, RNNs can utilize their inner state (memory) so as to process groupings of sources of info and yields. This influences them to be utilized in zones were errands, for example, unsegmented, penmanship acknowledgment or discourse acknowledgment or content to discourse transformation and in certain fields were bringing the last happened verifiable information is much criticized for making forecasts. The expression "recurrent neural system" is utilized explicitly to allude to two wide classes of systems, where one is limited motivation and other is an interminable drive. The two classes of systems display transient powerful conduct. A limited drive repetitive system is a coordinated non-cyclic chart that can be unrolled and supplanted with a carefully feed-forward neural system, while an endless motivation intermittent system is a coordinated cyclic diagram that cannot be unrolled. Both limited motivation and boundless drive repetitive systems can have an extra put away state, and the capacity can be under direct control by the neural system. The capacity can likewise be supplanted by another system or chart if that fuses time delays or has input circles.

2). Long-Short Term Memory (LSTM): Long Short-Term Memory (LSTM) systems are an expansion for recurrent neural systems, which fundamentally expands their memory. it is appropriate to gain from imperative encounters that have an extremely prolonged stretch of time slacks in the middle.

IV. ACCURACY TABLE

In this table used to compare the accuracy between genetic algorithm and RNN algorithm. RNN algorithm consists more accuracy then other algorithm.

METHOD	ACCURACY	SPACIALITY	SENSITIVITY
RNN (RECURRENT NEURAL NETWORK)	0.98	0.9808	0.9847
GENETIC ALGORITHM	0.8366	0.8641	0.8043
K-MEAN	0.81	0.8395	0.777

The Proposed Method using RNN algorithm (0.98% - accuracy).

V. INTENTION OF OUR WORK

Our objective is to achieve accuracy as follows:

1) The system must have a higher accuracy rate than the conventional machine learning algorithm.
2) The system will have the ability to handle large data of the patients since the rate of medical data that is been generated these days are larger.
3) The system must be able to handle delay and noise tolerance very well.
4) The RNN algorithm is good at processing large data which is capable of delay and noise tolerance than any other machine learning algorithms.

VI. EXPERIMENT AND ANALYSIS

We have deployed our Coronary heart disease prediction system using RNN algorithm.

Recurrent Neural Network: By using RNN we are able to get an accuracy more than 90 percent in predicting the coronary heart disease of a patient from a large set of the patient's historical medical data.

A. *RNN* Working

For a neuron, i in the network with action potential y_i, the rate of change of activation is given by:

$$\tau i\, yi = -yi + \sum_{j=1}^{n} \mathrm{Wji}\, \sigma\,(\mathrm{yj} - \Theta \mathrm{j}) + Ii(t)$$

Where

- τ_i: Time constant of postsynaptic node.
- y_i: Rate of change of activation of postsynaptic node.
- y_j: Activation of postsynaptic node
- wji: Weight of connection from pre to postsynaptic node
- σ(x): Sigmoid of x e.g. σ(x) =1/(1+e-x)
- y_j: Activation of presynaptic node
- Θj: Bias of presynaptic node
- Ii(t): Input (if any) to node

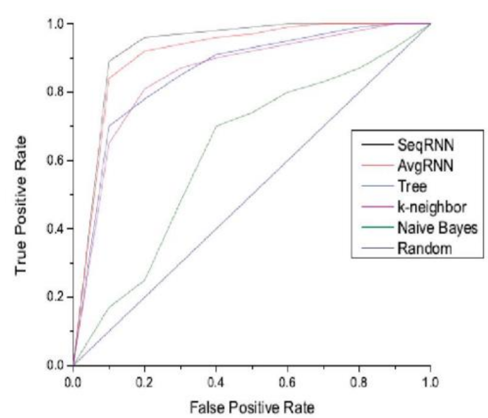


Fig 2. Comparison of RNN With various algorithms

VII. CONCLUSION

The recent statistics from the WHO report shows that 17.7 million people die every year due to cardiovascular diseases, (31% of global deaths). In 2017 nearly 6,16,000 deaths have been encountered due to heart disease. Hence the need for an efficient and accurate prediction of heart disease is on high demand. This paper deals with an deep technique involving the feature selection and classification of heart disease resulting in accurate prediction. The future course of this research can be pre-formed with diverse mixtures of deep learning techniques. Furthermore, new feature selection methods are often developed to urge a broader perception of the many features to extend the performance of heart conditionprediction.

REFERENCES

[1] Sarath Babu, Vivek EM, Famina KP, Fida K, Aswathi P, Shanid M,Hena M" Heart Disease diagnosis Using data mining Technique"-017.

[2] Wei-Pin, Chang, Der-Ming, Liou "Comparison of Three Data Mining Techniques with RNN Classifier Algorithm in the Analysis of Diabetics Data", Journal of Telemedicine and Telecare, 2008.

[3] Z. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei, and A. A. Yarifard, ‘‘Computer aided decision making for heart diseasedetection using hybrid neural network-genetic algorithm,’’Comput.Methods Programs Biomed., vol. 141, pp. 19–26, Apr. 2017

[4] S.B.Patil and Y.Kumaraswamy, ‘‘Intelligent and effective heart attack prediction system using data mining and artificial neural network,’’ Eur.J.Sci.Res., vol.31, no.4, pp.642–656,2000.

[5] A.D.Dolatabadi, S.E.Z.Khadem, and B.M.Asl, ‘‘Automated diagnosis of coronary artery disease (CAD) patients using optimized SVM,’’ Comput. Methods Programs Biomed., vol.138, pp.117–126, Jan.2017.

[6] U.R.Acharya, H.Fujita, O.S.Lih, M.Adam, J.H.Tan, and C.K.Chua, ‘‘Automated detection of coronary artery disease using different durations ofecg segments with convolutional neural network,’’ Knowledge.- Based Syst.,vol.132,pp.62–71, Sep.2017. M.Anbarasi, E.Anupriya,N.Ch.S.N.

[7] Iyengar,"Enhanced Prediction of Heart Disease with Feature Subset Selection using Genetic Algorithm", International Journal of Engineering Science and Technology Vol.2(10), 2010, 5370-5376.

[8] Jyoti Soni, Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction, International Journal of Computer Applications (09758887) Volume 17- No.8, March 2011.