International Journal of Education and Science Research Review
P-ISSN 2349-1817, E- ISSN 2348-6457
www.ijesrr.org December- 2017, Volume-4, Issue-6 Email- editor@ijesrr.org

MEDICAL DIAGNOSIS OF CHRONIC KIDNEY DISEASE USING ANN

Ajay Lakhnotra
Department of Computer Science & I.T.
Kathua Campus
University of Jammu

Amit Sharma
Department of Computer Science & I.T.
Kathua Campus
University of Jammu

Vinod Sharm
Department of Computer Science & I.T.
Kathua Campus
University of Jammu

ABSTRACT :
In this paper our efforts are to indicate the chronic kidney disease by various sorts and separating facts of the disease to the greatest extent by Generalized Regression Neural Network (GRNN). The data required to study has been chosen by more complex neural network structure before its inspection on the clinical data. In order to ensure the adeptness of the physician, diagnosing the kidney disease and prescribing the absence of sensational ways is very energetic assignment. To make the process more meaningful and scientific a data of about 400 patients, who were undergoing treatment of the doctors in various hospitals, is collected. Since the study includes the detailed information of the patient, so pre-processing was done. The GRNN have been applied over the patient data. The results of these valuation show that GRNN can be applied successfully for advising the anesthetic for kidney disease patient.
Keywords: Artificial Intelligence, Data Mining, Machine Learning, Chronic Kidney Disease, Generalized Regression Neural Network (GRNN), Medical Diagnosis.

INTRODUCTION:

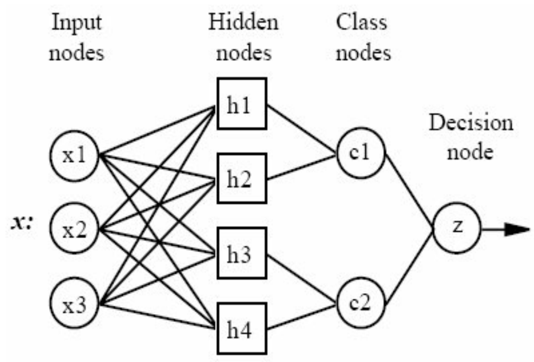
As per latest survey, about 14% of the world population is affected by Chronic Kidney Disease (CKD), and millions die each year because of income resources to undergo the costly treatment. According to the 2010 Global Burden of Disease study, chronic kidney disease was ranked27th in 1990 in the list of worldwide deaths but rose to 18th in 2010. Over 2 million people worldwide currently receive treatment with dialysis/kidney transplant to stay alive which is only 10% of total kidney patients. More than 80% of all patients who receive treatment for kidney failures are from elderly population, usually come from wealthy countries.

Chronic kidney disease damages kidneys and decreases the ability to work. Wastes can be built to high levels in your blood makes you feel sick because of Chronic Kidney disease. Complications like high blood pressure, anemia, weakness, poor nutritional health and nerve damage can be developed. Over a long period of time the kidney disease increases your risk of having cardiac and blood vessel disease. Chronic kidney disease may be caused by diabetes, high blood pressure and other disorders. Early detection and treatment keeps kidney disease from getting worse. When the disease progresses, it may lead to kidney failure which requires dialysis or kidney transplant.

The primary concern of AI in medicine is the creation of AI programs that can enable medical doctors in performing expert diagnosis. Such programs which have been derived as a result of various calculations with the help of sciences such as statistics succeed in finding out the hidden patterns from the training data and with the help of these patterns they classify the test data into one of the possible categories. The firmness of these AI programs are these various data sets prepared from different clinical cases and play a vital role in training the system. The decision and recommendation as an outcome of these systems can be illustrated with experience of human experts.

GENERAL REGRESSION NEURAL NETWORK (GRNN):

The authors selected the GRNN (general regression neural network) for the classification because of its good performance and accuracy. The basic GRNN was published in 1991 by Donald F.Spechtand reinvented by Schiùler and Hartmann in 1992. It is based on established statistical principles and converges with an increasing number of samples asymptotically to the optimal regression surface. GRNN has advantages of instant training and easy tuning. A GRNN would be formed instantly with just 1-pass training with the development data. GRNN is considered to be one of the most efficient and effective inductive learning algorithms for machine learning and data mining. GRNN is a memory-based network that provides estimates of continuous variables and converges to the underlying (linear or nonlinear) regression surface is described. GRNN is a one-pass learning algorithm with a highly parallel structure. It is shown that, even with sparse data in a multidimensional measurement space, the algorithm provides smooth transitions from one observed value to another. The algorithmic form can be used for any regression problem in which an assumption of linearity is not justified.



Input Layer Hidden Layer Summation Layer Output Layer

Figure 1 A GRNN Architecture

A GRNN is a special kind of the radial basis neural networks, which is based on kernel regression networks. A GRNN does not require an iterative training procedure as back propagation networks. It approximates any arbitrary function between input and output vectors, drawing the function estimate directly from the training data.

A GRNN consists of four layers: input layer, pattern layer, summation layer and output layer as shown in Fig. 1. The number of input units in input layer depends on the total number of the observation parameters. The first layer is connected to the pattern layer and in this layer each neuron presents a training pattern and its output. The pattern layer is connected to the summation layer. The summation layer has two different types of summation, which are a single division unit and summation units. The summation and output layer together perform a normalization of output set. In training of network, radial basis and linear activation functions are used in hidden and output layers. Each pattern layer unit is connected to the two neurons in the summation layer, S and D summation neurons. S summation neuron computes the sum of weighted responses of the pattern layer. On the other hand, D summation neuron is used to calculate un-weighted outputs of pattern neurons. The output layer merely divides the output of each S-summation neuron by that of each D-summation neuron, yielding the predicted value Y_i to anunknown input vector x .

$$Y_i = \frac{\sum_{i=1}^{n} \quad y_i e^{-D(x,\ x_i)}}{\sum_{i=1}^{n} \quad e^{-D(x,\ x_i)}}$$

Where

$$D(x,\ x_i) = \sum_{k=1}^{m} \quad (\frac{x - x_{ik}}{\sigma})^2$$

y_i is the weight connection between the i^{th} neuron in the pattern layer and the S-summation neuron, n is the number of the training patterns, D is the Gaussian function, m is the number of elements of an input vector, x_k and x_{ik} are the j^{th} element of x and x_i, respectively, σ is the spread parameter, whose optimal value is determined experimentally.

RADIAL BASIS FUNCTION NEURAL NETWORK:

RBFN is an alternative technique to the more widely used MLP network which uses less computer time for network training. RBFN consists of three layers: an input layer, a hidden (kernel) layer, and an output layer. The nodes within each layer are fully connected to the previous layer. The input variables are each assigned to the nodes in the input layer and they pass directly to the hidden layer without weights. The transfer functions of the hidden nodes are RBF. An RBF is symmetrical about a given mean or center point in a multidimensional space. In the RBFN, a number of hidden nodes with RBF activation functions are connected in a feed forward parallel architecture. The parameters associated with the RBFs are optimized during the network training. These parameter values are not necessarily the same throughout the network nor are they directly related to or constrained by the actual training vectors. When the training vectors are presumed to be accurate, i.e. no stochastic, and it is desirable to perform a smooth interpolation between them, then linear combinations of RBFs can be found which give no error at the training vectors. The method of fitting RBFs to data, for function approximation, is closely related to distance weighted regression. The RBF expansion for one hidden layer and a Gaussian RBF is represented by

$$Y_k(X) = \sum_{i=1}^{H} W_{ki}\, e^{\left(-\frac{\|X-u_i\|}{\sigma_i^2}\right)}$$

An interpolation RBFN is characterized by equal number of basic functions with training points. However, each input training point serves as a center for the basis function. In order to ensure a smooth fit of the desired outputs, the width of each kernel has to incorporate the training points.

METHODOLOGY;

This data set was obtained from the UCI Repository of data sets. The data set was collected and created by Dr. P. Soundarapandian and L.Jerlin Rubini. The data set have 400 cases and each case has 25 clinical parameters. Out of 400 cases 250 are of class one and 150 of class two. The various clinical parameters are shown in the table below.

SNo.	Parameter	Parameter Description (in Numeric Values)
1	age	age
2	bp	blood pressure
3	sg	specific gravity
4	al	albumin
5	su	sugar
6	rbc	red blood cells
7	pc	pus cell
8	pcc	pus cell clumps
9	ba	bacteria
10	bgr	blood glucose random
11	bu	blood urea
12	sc	serum creatinine
13	sod	sodium
14	pot	potassium
15	hemo	hemoglobin

16	pcv	packed cell volume
17	wc	white blood cell count
18	rc	red blood cell count
19	htn	hypertension
20	dm	diabetes mellitus
21	cad	coronary artery disease
22	appet	appetite
23	pe	pedal edema
24	ane	anemia
25	class	Class

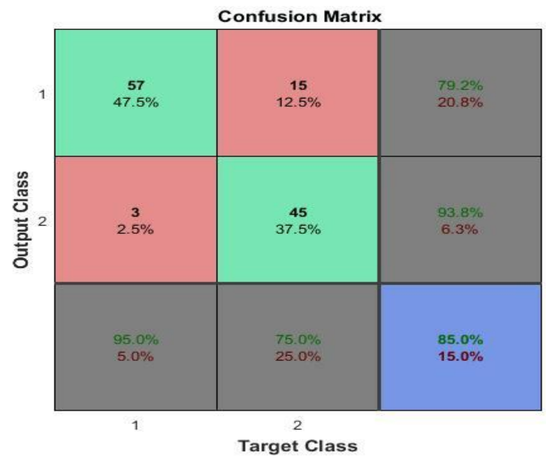


Figure 2 Confusion Matrix

IMPLEMENTATION AND ANALYSIS:

We use 70% of the data as input to the GRNN and rest 30% of data for testing .As evident from confusion matrix shown in the Figure 3, 85% data items are classified correctly by the GRNN. We implement GRNN in MATLAB R2015a.

Also the Receiver Operating Characteristics of the network are plotted in the Figure 3.

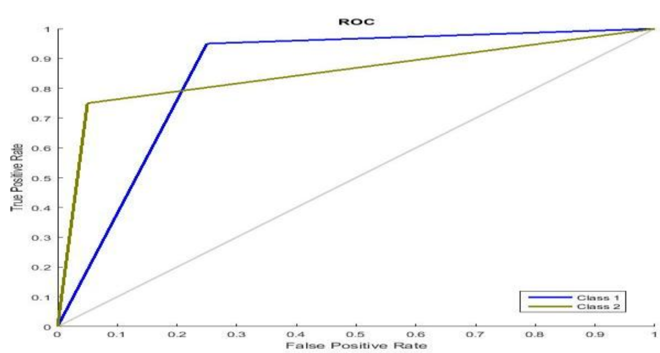


Figure 3 ROC Graph

CONCLUSION

GRNN gives us the better classification results than the classical neural networks. The reliability of the system was evaluated by computing the mean absolute error between the predicted values and exact values the cases. The results suggest that this system can perform good prediction with least error and finally this technique could be an important tool for supplementing the medical doctors in performing expert diagnosis. In this method the efficiency of Forecasting was found to be around 85%. Its performance can be further improved by identifying & incorporating various other parameters and increasing the size of training data.

REFERENCES

1. Abid Sarwar, Vinod Sharma,Intelligent Naïve Bayes Approach to Diagnose Diabetes Type-2,ICNICT 2012, November 2012.
2. Shaikh Abdul Hannan,R. R. Manza,R. J. Ramteke,Generalized Regression Neural Network and Radial Basis Function for Heart Disease Diagnosis ,nternational Journal of Computer Applications (0975 –8887) Volume 7–No.13, October 2010
3. Pandey A.,Bansal K.K.(2014): "Performance Evaluation of TORA Protocol Using Random Waypoint Mobility Model" *International Journal of Education and Science Research Review* Vol.1(2)
4. Tiwari S.P.,Kumar S.,Bansal K.K.(2014): "A Survey of Metaheuristic Algorithms for Travelling Salesman Problem " *International Journal Of Engineering Research & Management Technology* Vol.1(5)
5. Yadav R.K.,Bansal K.K.(2012) : "Analysis of Sliding Window Protocol for Connected Node" *International Journal Of Soft Computing and Engineering(IJSCE)* Vol.2(5) pp.292-294
6. Kamer Kaaer , Tulay Yilirim, Medical Diagnosis on Pima Indian Diabetes Using General Regression Neural Networks.
7. Donald F. Specht, A General Regression Neural Network , IEEE transactions on neural networks. Vol. 2. no. 6. November 1991
8. http://archive.ics.uci.edu/ml/datasets/Chronic_ Kidney_Disease
9. http://in.mathworks.com/help/nnet/ug/generalized-regression-neural-networks.html#bss38j0-2
10. http://www.cdc.gov/diabetes/pubs/pdf/kidney_factsheet.pdf