RESEARCH ARTICLE OPEN ACCESS

A Neural Network Approach For Heart Disease Prediction

Tanu Verma[1], R. K. Srivastava[2]
[1] Mewar University
[2] Dept. of Computer Science, Bareilly College Bareilly, UP, India

Abstract
Artificial Neural Network for intelligent medical diagnoses has been shown to be an interesting topic. In this paper we have used a neural network technique to predict a system, which can detect heart disease from their physical symptoms. We have used a prediction method with the help of the back propagation algorithm to train the networks. The actual procedure of medical diagnosis that usually is employed by physicians was analyzed and converted to a computer program implemented format. After selecting some symptoms a data set contains the information of fifty two cases was configured and applied to a M L P neural network.
Keywords: artificial neural network, back propagation algorithm

I. Introduction

Neural Network have been used in the field of artificial intelligence, preferred as these use the relation of dependency, a generation of function. The model and language of neural networks are more mathematically formulated hence most of the doctors avoid to use neural networks for prediction of disease. A prediction of heart disease system usually starts with patients complaints and the physician learn more about the patients situations interactively during the diagnosis as well as by measuring some metrics such as blood pressure, hemoglobin, s. urea, s. creating, FBS(mg/d), PPb(mg/d),RBS etc. The quantity of examples is playing an important role for training purpose but examples need to be selected carefully for the reliability and efficient system. I have considered the indirect relation of various parameters for data and presume that their relations are time invariant in view of the patient pathological test reports. The proposed method is more efficient and provide better forecast. The forecasted values obtained through the back propagation algorithm process have been compared with the observed productivity and their robustness has been examined.

II. Artificial Neural Network

Back Propagation through time is a powerful tool of artificial neural network with application to many areas as pattern recognition, dynamic modeling and nonlinear systems. Back propagation algorithm(BPA) provides an efficient way to calculate the gradient of the error function using chain rule of differentiation. The error after initial computation in the forward pass is propagated backward from the output units, layer by layer. BPA, a generalized Delta rule is commonly used algorithm for supervised training of multi layer feed forward artificial neural network. In supervised learning, we try to adapt an artificial network so that the actual outputs ($\overline{Y}$) come close to some target outputs(Y) for a training set, which contains T patterns. The goal is to adapt the parameters of network so that it performs well for pattern from outside the training set.

2.1 Back propagation Algorithm : We have proposed a neural network that will combine the features of multi perceptron concept of both feed forward part of back propagation algorithm and Let the training set be $\{x(k),d(k)\}_{k=1}^{N}$,Where x(k) is the input pattern vector to the network and d(k) is the desired output vector for the input pattern x(k).The output of the jth output unit is denoted by yj , connections weights from the ith unit in one layer to the jth unit in the layer above are denoted by wij. If m be the no. of output units and dj(k) is the desired output from the jth output unit whose actual output in response to the kth input exemplar x(k) is yj ,for j=1,2,3,……..,m. The sum of squares of the error over all the output unit for this kth exemplar by

$$E(k)=(1/2)\sum_{j=1}^{m}[y_j(k)-dj(k)]^2$$

Error E(k) is affected by the output from unit j at the output layer and is determined by

$$\frac{\partial E(k)}{\partial y_j}=y_j-d_j$$

The net input to output layer is of the form

$$Sj=\sum_i y_i^{(1)} w_{ij}-□_j$$

Where $y_i^{(1)}$ *is* the output from the i^{th} unit in the first layer below the output layer, w_{ij} is connection weight multiplying $y_i^{(1)}$ and $□_j$ is the threshold of unit j. The negative of threshold is defined to be the bias.

2.2 Computer code in C Language to trained the Network: Let us consider the components x[0][0] to x[ntmax][ni] so that there are "ntmax * ni" inputs to the network and y[0] to y[ntmax] outputs. The value "ntmax *ni" decides how many neurons in the network, "net" represents the total level of existing a neuron and y[nt][k] represents the intensity of resulting output from the neuron or activation level. We assume the full range of allowed connections, simply for the sake of generality.

```
#include<stdio.h>
#include<math.h>
#include<conio.h>
void main()
{
double
x[52][13],h[13][13],y[13],yd[13],dy[13],dh[13][13],net,whi[13][13],woh[13];
double
xmax[700],xmin[700],ydmax[13],ydmin[13],e,dnet,dwhi[13][13],dwoh[13],diff;
int ni=13,ntmax=52,ncmax,nh,i,j,k,nt,nc;
float eta;
FILE *f,*ee,*o;
clrscr();
printf("type no of iterations");
scanf("%d",&ncmax);
printf("type no. of hidden neurons");
scanf("%d",&nh);
printf("type value of learning rate");
scanf("%f",&eta);
for(i=0;i<=ni;++i)
for(j=0;j<=nh;++j)
whi[i][j]=0;
for(i=0;i<=ni;++i)
woh[i]=0;
f=fopen("data1","r");
o=fopen("out","w");
ee=fopen("error","w");
for(nt=0;nt<=ntmax;++nt)
fscanf(f,"%ld",&yd[nt]);
for(nt=0;nt<=ntmax;++nt)
for(i=0;i<=ni;++i)
fscanf(f,"%ld",&x[nt][i]);
for(nc=0;nc<=ncmax;++nc)
{
for(nt=0;nt<=ntmax;++nt)
{
for(j=0;j<=nh;++j)
{
net=0.0;
for(i=0;i<=ni;++i)
net+=whi[j][i]*x[nt][i];
h[nt][j]=1.0/(1.0+exp(-net));
}
}for(nt=0;nt<=ntmax;++nt)
{net=0;
for(j=1;j<=nh;++j)
net+=woh[j]*h[nt][j];
y[nt]=1.0/(1.0+exp(-net));
}
e=0.0;
for(nt=0;nt<=ntmax;++nt)
{
dnet=y[nt]-yd[nt];
e+=0.5*pow(dnet,2);
dy[nt]=dnet*y[nt]*(1.0-y[nt]);
}
printf("nc=%d,e=......%f",nc,e);
fprintf(ee,"%d%f\n",nc,e);
for(nt=0;nt<=ntmax;++nt)
for(j=0;j<=nh;++j)
{
dnet=0.0;
dnet+=dy[nt]*woh[j];
dh[nt][j]=dnet*h[nt][j]*(1.0-h[nt][j]);
}
for(i=0;i<=nh;++i)
for(j=0;j<=ni;++j)
{
dwhi[i][j]=0.0;
for(nt=0;nt<=ntmax;++nt)
dwhi[i][j]+=x[nt][j]*dh[nt][i];
}
for(j=0;j<=nh;++j)
{
dwoh[j]=0.0;
for(nt=0;nt<=ntmax;++nt)
dwoh[j]+=h[nt][j]*dy[nt];
}
for(i=0;i<=ni;i++)
for(j=0;j<=nh;j++)
whi[j][i]-=eta*dwhi[j][i];
for(j=0;j<=nh;++j)
woh[j]-=eta*dwoh[j];
}
printf("\n observed value \t calculated value \t difference");
fprintf(o,"\n observed value \t calculated value \t difference");
for(nt=0;nt<=ntmax;++nt)
{
diff=0.0;
diff=yd[nt]-y[nt];
printf("\n%f\t%f\t%f",yd[nt],y[nt],diff);
fprintf(o,"\n %f\t %f\t %f",yd[nt],y[nt],diff);
}fclose(f);
getch();
}
```

III. Computation of ANN forecasted value

Considering the patient pathological and other test reports related to heart disease as input x(k) and particular S. No. of patient data set to be predicted as desired output d(k) after applying the BPA, the calculated output is treated as predicted value of cor. Angio. The steps adapted for calculation of predicted cor. Angio. value i.e. output through BPA is as follows:

Step 1: Data set of patient S. No. 1 to 46 as input set and Coronary Angiogram value of patient S. No. 46 as desired output.
Step 2: Data set of patient S. No. 1 to 47 as input set and Coronary Angiogram value of patient S. No. 47 as desired output.
Step 3: Data set of patient S. No. 1 to 48 as input set and Coronary Angiogram value of patient S. No. 48 as desired output.
Step 4: Data set of patient S. No. 1 to 49 as input set and Coronary Angiogram value of patient S. No. 49 as desired output.
Step 5: Data set of patient S. No. 1 to 50 as input set and Coronary Angiogram value of patient S. No. 50 as desired output.
Step 6: Data set of patient S. No. 1 to 51 as input set and Coronary Angiogram value of patient S. No. 51 as desired output.
Step 7: Data set of patient S. No. 1 to 52 as input set and Coronary Angiogram value of patient S. No. 52 as desired output.

The algorithm has been implemented through C programming language, considering two hidden layers and computations have been made by various iterations levels like: 100, 200, 500 & 1000. Out of these, the best suitable forecasted values have been obtained by model with 1000 iterations. The result so obtained has been illustrated in Table 1 as follows:

Table: 1

S sNo.	NAME	AGE (Yrs)	Hb (gm%)	S.UREA (mg/dl)	S.CREATININE (mg/dl)	FBS (mg/dl)	PPBS (mg/dl)	RBS (mg/dl)	TG	S.CHOLES	LDL	VLDL	HDL	FASTING INSULIN	I.R.	COR. ANGIO.	ANN Forecasted COR. ANGIO value
1	H.C.YADAV	63	12.1	33	1.2	93.85	134.5	158	122	206	130	26	50	7.34	1	3	
2	B.M PANDEY	65	12	20	1.07	72.28	118.6	132	117.14	145.5	83.23	23.43	38.9	12.2	2.18	0	
3	K.P YADAV	54	13	24	1.12	76.4	114.8	148	114.6	138	84	22	32	10.2	1.92	3	
4	LALMANI SAROJ	55	12	24	1.23	70.2	116.8	98	117	136	76	24	36	8.13	1.4	3	
5	KAVITA PURVAR	38	12	20	0.98	74.81	136.1	124	115.75	124.97	71.03	23.15	30.79	10.1	1.86	0	
6	KALI PRASAD	62	13.7	23	1.27	92.25	124.6	118	121.34	124.41	76.42	24.27	23.72	16	3.64	0	
7	RAJPAT	60	11.4	24	1.1	68.69	94.7	96	105.45	103.5	48.21	21.09	34.2	2.07	0.35	3	
8	SI KANDEV	60	8.9	33.6	1.37	69.4	126.1	103	73.76	89.76	23.51	14.75	51.5	2.12	0.36	3	
9	RAJDEV	45	13.9	41.3	1.13	78.87	114.2	134	152.1	163.2	82.58	30.42	50.2	13.9	2.7	1	
10	V.D.SINGH	65	12.4	25.4	1.48	88.03	116.7	128	95.68	155.6	34.56	19.14	51.9	22.2	4.8	3	
11	C.B.SINGH	57	13	27.7	1.4	80.2	136.6	150	96	248.5	180.2	19.26	49	39.3	7.78	2	
12	MAKKHAN LAL	61	10.4	31.1	1.24	98.86	126.7	142	86.65	118.26	49.63	17.33	51.3	14.2	3.46	0	
13	CHAMPA DEVI	50	11.5	29	1	78	97.85	132	114	160	104	16	40	2	0.38	0	
14	RAM DEV	50	15	30.5	1.37	84.6	116.6	102	115.9	261.2	191.1	23.18	46.9	13.9	2.9	1	
15	HARISHANKER	57	13.1	28.4	1.23	88.68	121.4	90	125.7	138	73	25	40	4.5	0.98	0	
16	CHHOTAK SINGH	65	13.9	42.5	1.9	72.2	120.7	157	132.3	182.6	112.8	26.46	43.4	10.4	1.85	0	
17	BANKEYLAL YADAV	61	7.6	44.9	1.11	76.6	114.9	72.6	80.6	116.7	54.38	16.12	46.2	11.6	2.19	0	
18	B.M PANDEY	40	12	27	89	79.8	119.8	82	167.9	175.8	117	31	46	13.8	2.6	1	
19	K.P YADAV	64	8.9	31.4	0.93	80.6	134.1	81.9	150.95	138.75	59.59	30.19	42.2	8.6	1.71	1	
20	LALMANI SAROJ	65	6.8	21.9	0.9	74.8	125.8	84.6	99.65	119.87	63.14	19.03	36.14	10.4	1.92	3	

21	KAVITA PURVAR	55	10.5	21.9	0.76	69.9	140.9	33.3	156.86	163.98	92.42	31.37	40.2	14.2	2.45	1	
22	KALI PRASAD	52	11.1	28.9	1.05	78.9	138.5	150	80.15	129.93	64.22	16.03	49.68	6.2	1.2	0	
23	RAJPAT	65	10.5	20.2	1.36	89.5	136.3	100	75.71	109.34	49.94	15.14	44.26	11.6	2.56	2	
24	SI KANDEV	50	13	24.6	0.55	78.9	127.3	66.1	80.6	166	52.1	15.39	44.6	10.6	2.06	1	
25	RAJDEV	57	9	42.7	1.42	68.4	124.9	157	76.96	146.4	86.4	15.34	44.6	11.4	1.92	0	
26	V.D.SINGH	60	8.5	55.3	1.22	85.7	134.6	130	83.81	153.83	90.22	16.76	46.84	13.1	2.77	2	
27	C.B.SINGH	65	8.3	27.1	0.84	76.1	115.9	113	110.56	174.65	118.1	22.11	34.64	9.8	1.84	1	
28	MAKKHAN LAL	55	11.5	37.2	0.94	66.4	123.8	128	131.29	177.96	102	26.26	49.68	5.6	0.91	0	
29	CHAMPA DEVI	63	9.8	64	1.99	78.9	124.6	126	148.18	179.3	100.8	29.64	48.18	9.3	1.81	3	
30	RAM DEV	55	8.5	21.7	0.85	81.4	134.6	177	126.06	176.73	106.8	25.21	44.68	15.1	3.03	1	
31	HARISHANKER	43	9.5	27.8	1.4	69.76	128.9	173	86.1	132.8	69.68	17.22	45.9	7.2	1.24	0	
32	CHHOTAK SINGH	65	11.5	24.7	0.95	72.92	133.9	158	91.9	145.8	80.42	18.32	47.2	11.5	2.07	1	
33	BANKEYLAL YADAV	55	10.5	44	1.96	88	122.8	166	92.85	134.7	67.54	18.57	48.6	3.5	0.76	0	
34	R.S PANDEY	61	7.5	25.2	1.02	75.63	135.9	82.3	95.2	157.2	93.56	19.04	44.6	6.7	1.25	0	
35	KANTA PRASAD	50	10.2	54.1	1.01	67.91	124.3	124	93.5	141.6	74.2	18.7	48.7	8.9	1.49	2	
36	MAHENDRA NATH	54	13.5	38.3	1.41	85.34	119.4	66.4	91.6	129.5	62.18	18.32	49.1	13.4	2.82	1	
37	R.S PANDEY	59	14.5	28.4	0.98	82.61	131.9	121	84.5	128.4	68.34	19.05	45.8	12.1	2.46	3	
38	SANTOSH	55	9	33.7	0.77	95.41	129.6	276	118.63	154.75	93.69	21.73	38.33	9.8	2.3	1	
39	R.S TIWARI	62	9.2	32.6	0.88	83.98	130.8	130	135	135	82.98	27.65	42.92	10.4	2.15	1	
40	GYANENDAR	36	11.4	45.8	1.23	93.23	136.4	157	75.34	164	100	15.81	48	2.2	0.5	0	
41	DESH RAJ	32	6.8	27.3	0.94	85.34	121.5	166	170	155	74	34	45	6.7	1.41	0	
42	KRISHNA	50	8.6	34.8	1.11	94.56	123.9	149	130	158	80	26	52	11.8	2.75	0	
43	KANCHAN SHUKLE	46	11	25.6	0.77	95.34	131.7	163	120	156	81	24	51	13.3	3.13	0	
44	TARA DEVI	60	6.2	34.2	1.07	78.45	140.1	177	71.95	130.25	66.06	14.39	49.8	15.3	2.96	0	
45	GEETA KESARWAN	40	7.25	35.9	0.99	83.98	139.5	157	164.7	203	130.5	32.94	39.8	14.8	3.06	0	
46	GUDDU YADAV	40	11	43.8	1.08	76.92	123.8	135	188	143	64	38	40	3.5	0.66	1	0.9865
47	SAUGAR LAL	46	14	22.8	1.2	78.87	112.4	126	197	126	92	39	41	7.5	1.46	2	2.2376
48	NISHATH FATIMA	50	8.1	23.5	1.16	67.82	134.7	135	105	132	64	21	46	6.9	1.15	0	0.0083
49	ABDUL SALAAM	65	7.2	31.6	0.76	69.46	120.6	143	191	183	112	38	45	10.2	1.74	0	0.0063
50	RAM SUMAN	32	8.19	28.4	1.77	72.84	132.4	155	98	135	70	19	44	12.3	2.2	1	1.1765

51	RAM RAJ	54	8.4	34.9	1.41	84.65	143.9	128	64	157	97	13	47	14.7	3.07	1	1.0925
52	KAZIM	36	12	28.4	1.6	75.98	153.6	133	148	190	112	29	48	8.4	1.57	2	1.9742

IV. Result & Conclusion

The proposed artificial neural network technique has been implemented to have diagnosis of heart disease. We have considered the indirect relation of various parameters for time series data and presume that their relations are time invariant. The motivation of the study is to diagnose the heart disease, data are collected through various sampling techniques and obtained the diagnostic result through ANN using back propagation algorithm. A network requires information only on the input variable for diagnostic system. As values on test data are comparatively less the diagnostic model is reliable. The availability of data have tremendous amount of imprecision and uncertainty due to test reports of the patients are based on involvement of different electronic and mechanical equipments. Network performance could have been further improved by providing more training data. Moreover the considered connectionist models are robust, capable of handling the approximate data and therefore should be more reliable in worse situation. Optimal result will depend on the selection of parameters which is based on test results and symptoms of the patients. It is evident through the study that neural network model is even suitable over human diagnostic system

Heart disease prediction is of much use to the heart patient having pathological test data tending towards expected position of heart to avoid the heart attack. The motivation of the study is that the pathological test data are collected through various sampling and based on the reading of electronic machines, involving vagueness. Comparison of forecasted coronary angiogram test value of the patient obtained through ANN using back propagation algorithm with actual coronary angiogram value of the patients are listed in table-2.

Table-2

S No.	Name	COR. ANGIO.	ANN Forecasted COR. ANGIO.
1	GUDDU YADAV	1	0.9865
2	SAUGAR LAL	2	2.2376
3	NISHATH FATIMA	0	0.0083
4	ABDUL SALAAM	0	0.0063
5	RAM SUMAN	1	1.1765
6	RAM RAJ	1	1.0925
7	KAZIM	2	1.9742

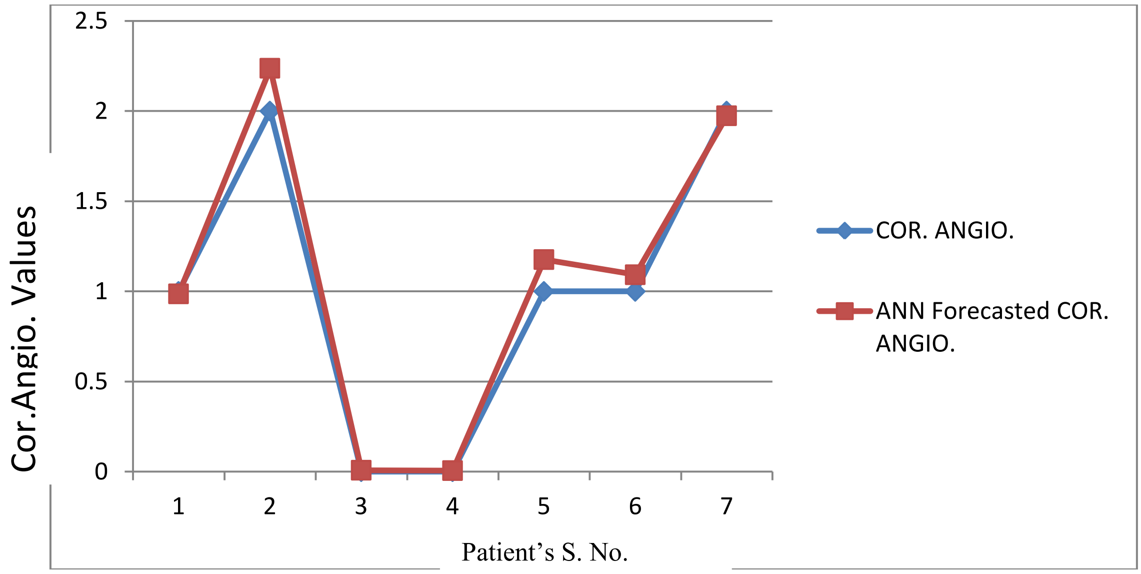

Fig.- Actual Cor.Angio. Vs Forecasted Cor.Angio. Values

In the study the target patient S.NO. 46 to 52 for the prediction of coronary angiogram value computed through the ANN method are quite impressive by comparison with actual value. Further, the computations shows that the predicted data through ANN method provides much better, suitable and reliable forecast for the heart disease patients.

References

[1] Chou-An Chen[2007], Neuro –Fuzzy Technology as a Predictor of Parathyroid Harmonse Level in Hemodialysis Patients,Tohku J,Exp.Med.,211,81-87

[2] Ferdous Ahemad Sohel, A new Neural Network with Fuzzy Technique, Disease diagnosis, A Case study,Dept. of Computer Science & Engg.,International Islamic University-Dhaka

[3] Holmstr,M. L. Koinstinen P, [1992],using Additive Noise in Backpropagation Training,IEEE Transaction on Neural Networks,pp 24-38,vol 3No. 1

[4] K.K.Shukla,Raghunath,[1999],An Efficient Global Algorithm for Supervised Training of Neural Networks,International Journal of Computers and EE,PERGAMON Press,25,pp 195-218

[5] R. Brause, F Friedrich[2000],A Neuro Fuzzy Approach as Medical Diagnostic Interface,pp 201-206

[6] S Moein,a Monadjemi[2006],A Novel Fuzzy-Neural Based Medical Diagnosis System,volume 1, Number 3

[7] V.S. Kodogiannis,MBoulougoura[2007], An Adaptive NeuroFuzzy Approach for the Diagnosis in Wireless Capsule endoscopy Imaging,International Journal of Information Technology,Vol. 13 No. 1

[8] Werbos Paul J. [1990], Back propagation through time: what it does and how to do it,*Proceeding of the IEEE*,78,1550-1559.

[9] Yang CC[1998],The use of back propagation neural network for the sim. & analysis of time series data ,*Transcations of the ASAE*,41,1181-1187