IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 06, Issue 02 (February. 2016), ||V2|| PP 01-06

Diagnosis and Detection of cancer cells in lungs & myocardial infarction using neural networks

Shikha Gangwal[1], Vasundhara Singh[1], Sara Gupta[1] and Udoshi Basavaraj[2*]
[1]Department of Electronics and Communication, P E S Institute of Technology- Bangalore South Campus, Bangalore -560100, Karnataka, India
[2]Department of Electronics and Communication, P E S Institute of Technology- Bangalore South Campus, Bangalore -560100, Karnataka, India

Abstract: The present paper reports the application of neural networks for the diagnosis of diseases like myocardial infarction and detection of cancer cells to lung cancer survival data. A survival model for lung cancer patients by Cox Regression Model was developed and clinical diagnosis of MCI was carried in larger patient sample using radiography and NMR scans. The testing model's data was tried to fit by non-linear square fit with baseline hazard rate function and tested with the help of Artificial Neural Network (ANN), as well as the analysis for MCI detection was done on basis of waveforms generated by electrocardiography (ECG) and phonocardiology. A feed forward back propagation network was constructed to stimulate the survival model whose data tend to have similar base line hazard by LSE. The error rate was 9% which defined the rate of false "alive-dead" judgments for 20 training cases.

On the other hand the clinical pattern sets to which the ANN was trained produced results which when compared to physicians caring confirmed the possibility to identify MCI. ANN had a sensitivity ranging from 80%-97.2%and specificity ranging from 83-98%. More layers of feed forward neural network resulted in higher accuracy computation of cancer cells. Non-linear methods were complex, and are computationally intensive whereas ANN configured for this specific applications of analysing pattern highlights the prospects of multifunctional application which helped the detection and diagnosis of cancer cells and MCI at early stage.

***Keywords:** -Myocardial Infarction (MCI), Lung Cancer, NMR, ECG, Radiographs, Non-linear square fit*

I. INTRODUCTION

In recent years, machine learning methods have been widely used in prediction, especially in medical diagnosis. It is one of the major problem in medical application. Neural network is being developed worldwide in medical diagnosis. They are used to increase accuracy and objectivity of medical diagnosis. This paper reports a systematic review that was constructed to access the benefit of artificial neural network (ANN) as decision making tools in the field of cancer and myocardial infarction. It is an information processing paradigm that is inspired by the way biological nervous systems such as brain, process information, and with its remarkable ability to derive meaning from complicated or imprecise data, it can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques.

Based on the power of artificial neural network ensemble, a survival model for detection of lung cancer named Cox Recognition Model was developed. The typical structure of a feed forward neural network consists of a layer of d (dimension) input units, a layer of output units and a variable number of hidden layers of units. More layers
result in higher accuracy, but simultaneously are time consuming on computation.

The study also shows that ANN can be used to develop a clinical score for risk assessment to determine the profile of patient with acute myocardial infarction. In hospitals, the prognosis of patients has progressively improved. However, despite advances in treatment of this

[redacted] -
blockers, angiotensin-converting enzyme ACE inhibitors into clinical practice, etc there is high in-hospital mortality.

Hence, the further section deals with the construction for ANN for the above purpose and provide test results that aids clinicians in accessing prognosis for each patients therefore being a useful guide.

II. METHODS

- **ANN FOR DETECTION OF LUNG CANCER:**

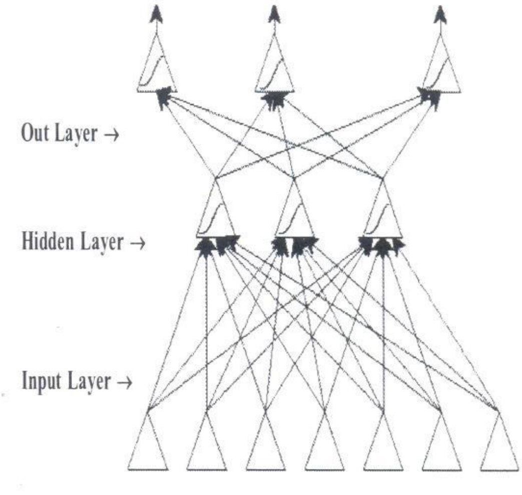


Figure 1: (Structure - Layers of Neural Network)

Fig 1: Structure-layers of neural network

Artificial Neural Networks have already been widely exploited in this area. Generally speaking, they are very useful in pattern recognition. Hornik showed that feed-forward ANN with one hidden layer can approximate functions in any accuracy. However until now there is no rigorous theory indicating how to do such error free approximations. Therefore, whether an artificial neural network based application will be useful or not depends on the user.

In the field survival analysis, there are two popular ways to test a model. One is to use half or two third of the time scale in survival data to determine parameters and then use the whole data set to examine the model. In our study, however, because of the short length of data it is given by,

$l=\log L_{i=1}\sum(f(n_1:\beta_1,\beta_2,\beta_3......\beta_v)$. Another way is to use whole data set to set up the model then use a resample method to check the model. But, if we randomly resample the original data, the selected data for testing may be far from the "pattern" of whole data set, hence we use maximum likelihood estimator (MLE), especially when data is not too long.

B.>COX-RECOGNITION MODEL:

In the COX model, which is given by $h(t/x_i)=h_0(t)\exp(f(x_i))$, where $h_0(t)$ is the hazard rate function and $\phi(\xi_\iota)=\xi_\iota\beta=\xi\iota\beta_{+}\ .\xi_{\iota\pi}\ \beta\ \iota_\pi$, the main interest is usually about the parameter (the covariates) vector β.However, when one assumes a parametric family for baseline hazard function, it is important to test that h_0 is equal to some specified hazard rate function.

In the following, we propose an artificial neural network 's (ANN) testing model by first learning for the patient's survival pattern from hospital data and then using ANN generate a list of "virtual data" and "stimulate" the survival pattern to test our covariate estimation and baseline hazard estimation.

C.> APPLICATION TO LUNG CANCER DATA:
(i) Data Structure:

A data set records the survival times (ST) of lung cancer patients at a hospital. It also records patients hospital conditions.

Table 1: Data of patients suffering through lung cancer

S.NO	PT	PN	STAGE	STAT	ST	DT	GRADE
1	T1	N2	IIIA	D	11	5	Mod
2	T1	N1	IV	D	17	0	Poor
3	T2	N0	IB	A	24	24	Well-mod
4	T3	N0	IIB	D	2	0	Mod-poor
5	T2	N2	IIIA	A	21	7	Well-mod
6	T1	N0	IA	D	3	3	Mod

Where:-

PN : Occurrence of lymph notes, a symptom of lung invasion, ranges from N0 toN2. PT: patients term, ranges from T1 to T4. STAGE: pathological diagnosis of cancer and it is ordinal, ranges from 1A to IV. ST and DT: survival time and disease free times. GRADE: fitness condition when patient is in hospital.

STAT: status of patient indicating dead (D) or alive

(A).

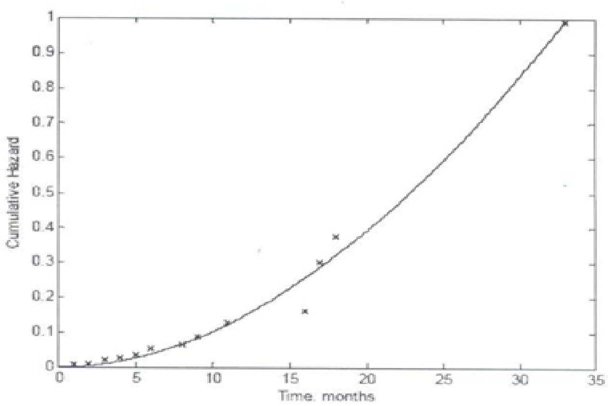
(ii) Estimation For Covariates:

The Cox regression model gives the mean and standard deviation for covariate in data. The β is estimated at a significant level. For "patient term" and "grade", β is positive which means higher hazard or risk of death whereas for "disease free time", β assumes a negative value less likely for person to die shortly.

It also gives a baseline cumulative hazard and overall cumulative hazard v/s survival time. To estimate hazard function, we find covariates at mean value, and then we use least squares regression to estimate 'a' and 'b' from,
$H0(t)=\int_0^1 h0(x)dx= a/b[x-1/b(1-_{e}-^{bt})]$
From the estimation of the baseline hazard function it is found to be

$h_0(t)=0.1266(1-e^{(-0.1727t)})$

The following Figures show the fit for cumulative baseline hazard and baseline hazard as a function of time.



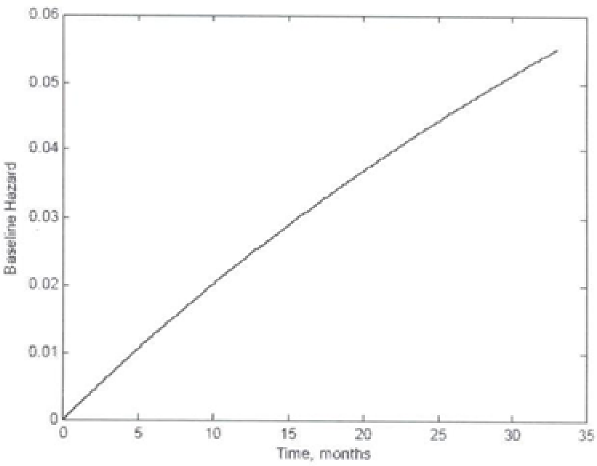


Fig 2: Graph for cumulative baseline hazard function as a function of time

D.> Detection of Myocardial

Infarction :

Since the number of patients attending emergency department show symptoms related to chest pain or acute myocardial infarction, early diagnosis helps to prevent permanent damage. But prognosis of MCI is difficult as elevated cardiac muscle enzyme helps in detection which is not helpful for short duration of symptom.

Most cases (approx. 25%) have shown that
patients were sent home due to misjudgement of the disease. Computer based ECG interpretation program are helpful but their performance can be improved using ANN. A recent study has also shown high accuracy in network result than in physicians diagnosis of acute myocardial infarction.

a) Study design and method:

The study consisted of a registry of AMI and UA patients admitted to hospitals and were designed for a limited time follow-up. The following variables were prospectively recorded: demographic data, history of hypertension, diabetes, peripheral vascular disease, Killip class etc.

The diagnosis included characteristic chest pain lasting >20 minutes, elevated creative kinase level and characteristic serial ECG changes which included new Q waves in at least 2 adjacent least. Two important criteria were fulfilled i.e., ANN with a multilayer perceptron which consisted one input layer, one hidden layer and one output layer, and secondly it was trained with the back propagation as learning algorithm.

The multilayer perceptron uses several neurons in multilayer structure. In this structure, output of neurons in one layer is in turn the input of the neurons in following layer with no direct connection or feedback taken in consideration. The connection weights were adjusted by use of langevin extension of back propagation updating

[redacted] of 0.5 and was decreased geometrically with every

[redacted] process. The momentum was set to 0.7 after 10th pattern.

To decide when to terminate the training process different parameters were set and for this the data was divided into 3 parts. One used as test and the other two used for training. This was repeated three times and the error in training was decreased, and finally mean was calculated.

The final procedure included 8 fold cross-validation procedure and the calculations were done using JETNET 3.0 package. Several problems were found described as below:

(i) Local minimum: The back propagation algorithm searches for closest minimum value in the error surface adjusting the weights using some gradient descending method which is done using inverse direction of gradient error surface curve. The problem arises because this curve may have more than one minimum and the adjustment starts over a local minimum instead of global minimum.

(ii) Saturation: The activation function works on its extremes causing learning algorithm to stop. This problem can be solved using a small target output or smaller weights at beginning of training.

(iii) Error Function: The back propagation algorithm depends upon error function to minimize and thus to get best model functions used were sum of squares, cross entropy, absolute value and Minkowski.

Table 2: Comparison of NN, conventional criteria and cardiologist performance:-

Specificity	Criteria A	Criteria B	Discriminant	Power
95.2	47.1	28.8	1.59	1.15
95.4	46.2	30.7	1.60	1.22
99.6	16.8	3.8	2.16	1.26
86.3	65.9	55.4	1.67	1.13
99.8	13.4	22.7	2.40	2.73

III. RESULTS

With the help of feed-forward back propagation network, survival model was survival model was simulated and a total of three layers gives the output representing "dead" and "alive" with "0" and "1" respectively. When having "traingda/learngdm" as training learning function, the ANN reaches best performance and the error rate for the training set was 9%. After the Cox Regression and a least square fit for cumulative baseline hazard for ANN, generated data is plotted as a of time.

The results show that detection of acute myocardial infarction using trained ANN can be done in 12 load ECG at a sensitivity much higher than two sets of conventional rule based criteria and even better than an experienced cardiologists, but does not work every time with cardiologists findings in definite MCI.

It is compared to the original hospital data as below:-

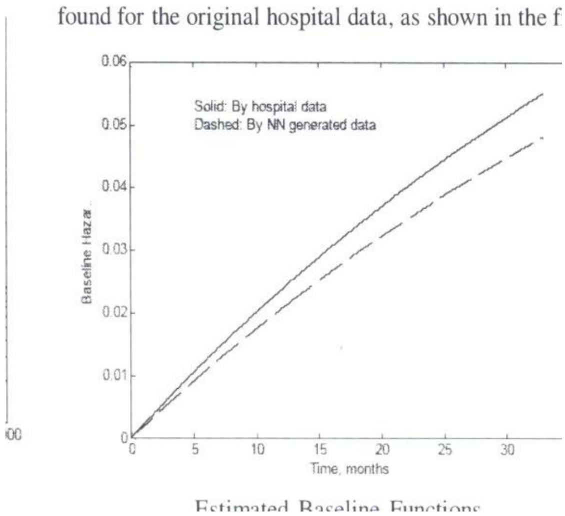


Fig 3: Estimated baseline function

IV. CONCLUSION

Cancer is one of the most common and deadly diseases in the world especially lung cancer. Detection in early stage is key of its cure and Automatic diagnosis is an important, real world problem.

When NN generated data assume the same mean and SD with original data, they tend to have similar baseline hazard function by LSE. Thus in our paper we have shown a model can be constructed for diagnostic system that performs at an accuracy level with painless detection for patients and help to oncologists with which they can plan for better medication.

Early detection of acute MCI is important at the same time. Since, myocardial infarction showed symptoms concomitant with chest pain, etc they were misjudged or overloaded by physicians in 80% of the cases. Trained ANN could detect acute MCI at a sensitivity higher than conventional rule based criteria with ECG and radiographs.

These results show that neural network application is survival analysis has promising which can even be extended with improved automated ECG and more number of data, from which we are able to diagnosis and detect lung cancer cells and MCI.

V. ACKNOWLEDGEMENT

This project would have been impossible if it wasn't for a few people who stuck with us and guided us through. First and foremost, we would like to thank our parents for being our backbone, as without their support and encouragement, nothing would have been possible. We would also like to thank our teachers who guided us without losing their patience. We would be failing in our duty if we forget to thank our friends for constant encouragement and for being our support system. Last but not the least, we would like to thank PESSE, BSC for giving us this opportunity to display our work and talent.

REFERENCE

[1]. Killip T, Kimball JT. Treatment of myocardial infarction in a contrary in a coronary care unit ;a two year experience with 250 patients Am J Cordial 1967; 20:457-60

[2]. Effectiveness of intravenous thrombolitic therapy in acute myocardial infarction. Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardio (GISSI). Lancet 1986; 1:397-401.

[3]. Ijcaoline.net/journal/...../pxc387783. pdf

[4]. Circ.ahajournals.org/content/.........../ 1798.long

[5]. Andy Sloggett, M. Hills Regression Models for Survival Statistics in Medicine ,23, 51-64

[6]. Kevin F.O' Brein a Life Table Regression Analysis for Cancer Survival Data , East Carolina University and by C. Suchindaran, University of North Carolina.

[7]. Table 1 Reference: Detection of cancer cells by neural network testing of P.K. Chenniappan and N. Anusheela

[8]. http://m.circle.ahajurnals.org/conten t/96/6/1798.long Time:6:30 17/04/2015