ISSN 0975-3303
Mapana J Sci, 9, 2 (2010), 62-71
https://doi.org/10.12725/mjs.17.8

RULE BASED AND SUPERVISORY TRAINING APPROACH TO DEVELOP EXPERT SYSTEM TOOL FOR DETECTING LUNG CANCER DISEASE

K. Balachandran and R. Anitha***

ABSTRACT

Knowledge-based expert systems, or expert systems, use human knowledge to solve problems that normally would require human intelligence. These expert systems represent the expertise knowledge as data or rules within the computer. These rules and data can be called upon when needed to solve problems. Lung cancer is one of the dreaded disease in the modern era. It is responsible for the most cancer deaths in both men and women throughout the world. Early diagnosis and timely treatment are imperative for the cure. Longevity and cure depends on early detection. This paper gives an insight to identify the target group of people who are suffering or susceptible to suffer lung cancer disease. Seeking proper medical attention can be initiated based on the findings. Expert system tool developed, to find this target group based on the non-clinical parameters. Symptoms and

* *Christ University, Bangalore-29 Email : balachander63@gmail.com*
* *K. S. R. College of Technology, Tiruchengode-637 215 Email : aniraniraj@rediffmail.com*

risk factors associated with Lung cancer are taken as the basis of this study. This expert system basically works on the rule based approach to collect the data. Then Supervisory learning approach is used to infer the basic data. Once sufficient knowledge base is generated the system can be made to adopt in unsupervised learning mode.

Index Terms— Rule based expert system, supervisory learning, Neural Network, Radial basis Function Network

Introduction

Cancer Disease

Cancer is a group of cells usually derived from the single cell that has lost its normal control mechanism and thus has unregulated growth. Cancerous[malignant] cells can develop from any tissue within any organ. As cancerous cells grow and multiply, they form a mass of cancerous tissue – called a tumor – that invades and destroys normal adjacent tissues. The 'tumor' refers to an abnormal growth or mass; tumors can be cancerous or non-cancerous. Cancerous cells from the primary (initial) site can spread (metastasize) throughout the body.

Both external and internal factors cause cancer. Factors such as chemicals, radiation, viruses, hormones and inherited mutations may act together to start or further cancer. Ten or more years may pass between exposure and detectable cancer.

Anyone, from children to senior citizens can get this disease. What causes cancer in the first place, and how quickly the cells grow and spread, varies from person to person. While a large number of people with cancer overcome the disease and live fulfilled lives for many years, some succumb within a few months. The journey with cancer is a long, stressful and strenuous one.

Lung Cancer

Cancer of the lung, like all cancers, results from an abnormality in the body's basic unit of life, the cell. Normally, the body maintains a system of checks and balances on cell growth so that cells divide to produce new cells only when needed. Disruption of this system of checks and balances on cell growth results in an uncontrolled division and proliferation of cells that eventually forms a mass known as a tumor.

Lung cancer is responsible for the most cancer deaths in both men and women throughout the world. The American Cancer Society estimates that 213,380 new cases of lung cancer in the U.S. will be diagnosed and 160,390 deaths due to lung cancer will occur in 2007. Lung cancer is predominantly a disease of the elderly; almost 70% of people diagnosed with the condition are over 65 years of age, while less than 3% of cases occur in people under age 45.

Lung cancer can be classified into two major types viz. Small cell Lung cancer (SCLC) and Non-small cell Lung cancer (NSCLC).

Expert system

In an attempt to overcome limitations inherent in conventional computer-aided diagnosis, investigators have created programs that simulate expert human reasoning. Hopes that such a strategy would lead to clinically useful programs have not been fulfilled, but many of the problems impeding creation of effective artificial intelligence programs have been solved.

Knowledge-based expert systems, or expert systems, use human knowledge to solve problems that normally would require human intelligence. These expert systems represent the expertise knowledge as data or rules within the computer. These rules and data can be called upon when needed to solve problems.

Programs using artificial intelligence techniques have several major advantages over programs using more traditional methods. These programs have a greater capacity to quickly narrow the number of diagnostic possibilities, they can effectively use patho-physiologic reasoning, and they can create models of a specific patient's illness. Such models can even capture the complexities created by several disease states that interact and overlap. These programs can also explain in a straightforward manner how particular conclusions have been reached.

Lung Cancer diagnosis

Screening examinations, tests, or procedures are usually not diagnostic of cancer but instead indicate that a cancer may be present. The Diagnosis is then made following a workup that includes a biopsy and pathological confirmation.

Non-clinical symptoms and risk factors are some of the generic indicators of the cancer diseases. Early detection of the cancer disease is crucial in diagnosing and treating the patient. The cure, metastasis (spreading of cancer disease to completely new location), recurrence (relapse), Remission(absence of all evidence of a cancer after treatment) and survival rate. Percentage of people who survive for a given period of time after treatment all directly attributable to the phase of detection of the cancer disease.

Cancer diagnosis is a specialized field. The pathological facilities and the experts availability are very restrictive. Most of the time due these problems cancer patients are not able to get their disease diagnosed in time.

Hence in our approach we want to identify the target group of peoples who are either suffering or susceptible for Lung cancer disease, so that they can be directed for specialized treatment at the early onset stage of cancer.

II. Methodology

The Methodology we followed is to identify, whether the patient has a possibility of suffering with lung cancer based on the symptoms and risk factors as primary source of data. We create a knowledge base based on the known Lung cancer data and try to project on the sample data. The approach we adopted here is to list the various factors of risk and symptoms assign each factor with some static weights (refer Figure 1 & 2). These static weights are assigned based on the current available statistics. We followed a hybrid model of rule based system and supervised learning system in order to populate the knowledge base.

Rule based approach:

Presence or absence of the factors are initially obtained bases on either

If < weight loss > <continuous fever>... <fatigue present> then

Confirmation-of-status logic1

Else

Confirmation-of-status logic2

Or

If < condition 1 > then

Confirmation-of-status logic1

Else

Confirmation-of-status logic2

approach.

For some of the factors cascaded logic sequence is followed. For Example in the case of passive smoking factor, the sequence of exploring the logic is:

If < passive smoking > then

If < extent of exposure greater than equivalent of 6 cigarettes2> then

If < exposure duration more than 3 years> then

Confirmation-of-status logic1

else

Confirmation-of-status logic2

else

Confirmation-of-status logic3

else

Confirmation-of-status logic4

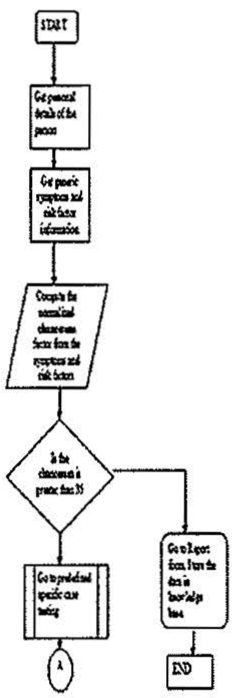

Figure 1- Generic Cancer Flowchart

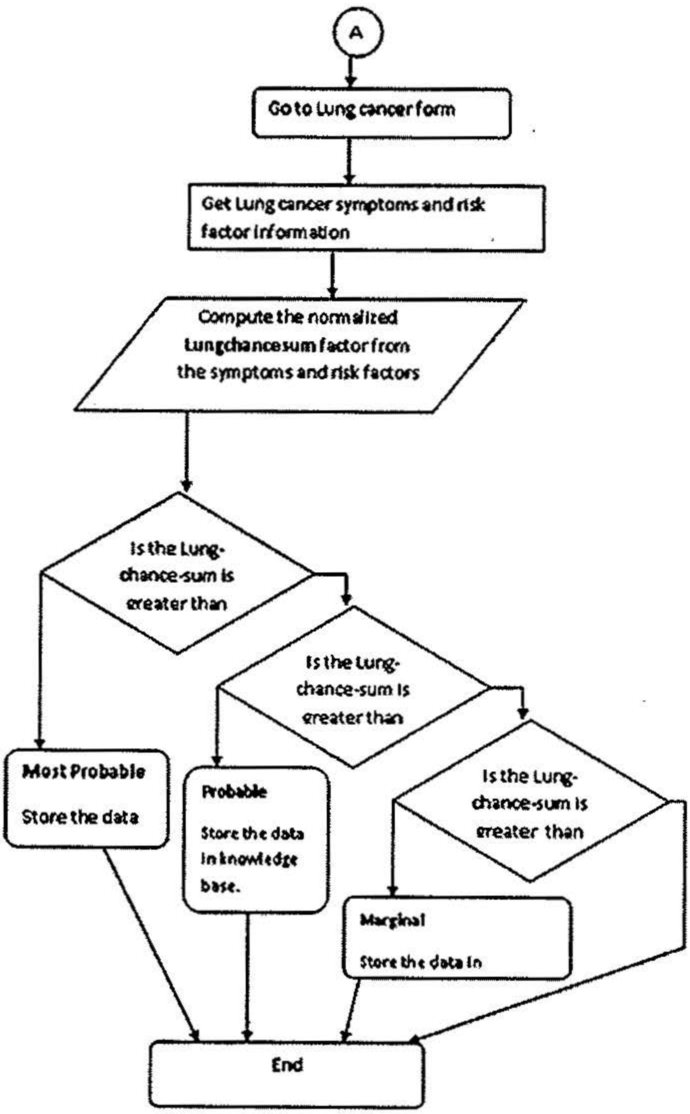

Figure 2- Lung cancer flow chart

Table 1- Tabulation of Symptoms & Risk factors with static weight factors

	Logic	Weights	ICD code
Symptoms			
Do you have Cough with expectoration	1	40	C34SYM01
Do you Experience Chest pain	1	20	C34SYM02
Do you Experience Loss of weight	2	0	C34SYM03
Do you Experience fatigue?	2	0	C24SYM04
Do you Experience Fever	2	0	C34SYM05
Do you Experience Hoarseness of voice	2	0	C34SYM06
Do you Experience Puffiness of face	2	0	C34SYM07
Do you Experience Loss of appetite	2	0	C34OYM00
Do you Experience Nausea and vomiting	2	0	C34SYM08
Risk factors			
Smoking	1	12	C34RIS01
High dose of ionizing radiation	1	3	C34RIS02
Occupational exposure to musterd gas, chloromethyl ether, inorganic arsenic, chromium, nickel, vinyl chloriede, radon asbestos	1	3	C34RIS03
Radon exposure	2	0	C34RIS04
Air pollution	2	0	C34RIS05
Insufficient consumption of fruits & vegetables	2	0	C34RIS06
Sum:		70	
Result		Normalised score	
Probable		67.24138	

Supervised Learning approach

As the effect of the particular condition may not yield the same type of outcomes in others rule based approach cannot be used here. After obtaining the data from the rule based approach the data is fed to the supervisory control wherein the relative static weights were specified. For the Symptoms the logic factor (L_i) and the weights (W_i) are multiplied. Similarly the logic factor (L_j) and the weights (W_j) are multiplied. and summation of these product terms are computed and normalized. If the number of symptom factors are 'n' and number of symptom factors are 'm' then

$$\text{Lung-Chance-sum-Factor} = \left[\left(\sum_{1}^{n} Wi.Li + \sum_{1}^{m} WjLj \right) \div \left(n+m \right) \right] x100$$

The weight assigning factors are based on current knowledge about the particular factor and also the available statistics pertaining to that factor. Weights also could be assigned by doing correlation study of the particular factor with that of the incidence of the cancer. The parameters are maintained in a database in a ICD (The International Statistical Classification of Diseases and Related Health Problems is a coding of diseases and signs, symptoms, abnormal findings, complaints, social circumstances and external causes of injury or diseases, as classified by the World Health Organization) code format(Ref. Table 1). Based on the parameters a neural network with multi layer Perceptron approach is framed. Next step is to choose a suitable neural network approximators. Radial basis Function Networks approach is chosen as it allows the application of successive differentiation for the local approximation of differential equations.

Un-supervised Learning approach

Once sufficient data is populated it is possible to derive the weights factor directly from the knowledge base, wherein the system can work in the Unsupervised learning methodology or Machine learning methodology.

III. Discussion and conclusion

In the first step, data is collected about the status of various generic factors of common cancer diseases. If the chance-sum computed warrants about the probability, then additional data pertaining to the specific symptoms and risk factors of Lung cancer are collected and chance sum is computed. The flow chart of Fig 1 and 2 depicts the processes involved. Collected data along with the assigned weights are shown in Table 1.

The data Results are inferred as follows.

1) Most probable (if Lung-Chance-sum-Factor > 80), 2) probable (if Lung-Chance-sum-Factor > 60 and <=80), 3) likely but further probing is required (if Lung-Chance-sum-Factor > 35 and <=60) and 4) unlikely(if Lung-Chance-sum-Factor <=35).

Based on the results, referral of the patient to the particular oncology specialist is done. The pathological tests suggested and the inference by the specialist can be augmented in the knowledge base. Once sufficient data is populated it is possible to derive the weights factor directly from the knowledge based wherein the system can work in the Unsupervised learning methodology.

This system is developed based on the proven Lung cancer patients data of Amala Cancer Research Hospital, Thrissur, Kerela. When this system has to be used for normal patients, some adaptability features have to be incorporated. As some of the symptoms and risk factors are also associated with other type of diseases, before entering the data ruling out the possibility of the generic diseases have to be done.

In come cases even in the advanced level Lung cancer patients does not show the symptoms associated with the Lung cancer. Leaving aside these extremities the system can be used

IV. Acknowledgment

Authors would like to extend their sincere thanks to Dr. Devuru, Dr. Ramachandra Reddy of Kidwai Institute of oncology, Bangalore for providing domain knowledge, Fr. Paul Achandy & Fr. Jossy for providing logistic help in collecting data from Amala Cancer Research Hospital, Thrissur, Kerela. Authors would like to express their gratitude to Dr.(Fr.) Thomas C. Mathew, Vice chancellor, Dr. (Fr.). V.M. Abraham, Pro-vice chancellor, Prof. K.A.Chandrasekharan, Peronal officer, Dr. Nanje Gowda, Dean of Sciences, Prof. Joy Paulose and other faculty members of Computer Science department of Christ University, Bangalore for providing technical support in carrying out the work

References

[1]. Eugene Braunwald, Anthony S.Fauci, Dennis L.Kasper, Stephen L.Hauser, Dal L. Longo & J.Larry Jameson, Harrison's Principles of Internal Medicine Volume I (15th Edition) (part six: Oncology & Hematology) 2001, Page 562-571.

[2. George F.Luger , Artificial Intelligence Structures and strategies for complex Problem solving,2004

[3]. Janakiraman.V.S, Sarukesi, Gopalakrishnan, Foundations of Artificial Intelligence and Expert systems, 2006

[]4. Henscheke CI etal.,Early Lung cancer Action project: overall design and findings from baseline screening .Lancet 354:99,1999

[5]. Mark Beers, Andrew J.Fletcher, Robert Porter, Michael Berkwits & Justin L.Kaplan, The Merck Manuel of Medical Information Second Home Edition, 2002

[6]. Whyatt and Spiegelhalter, Edward Feigenbaum, Pamela McCorduck, & H. Penny Nii , Medical Expert system,2001

[7]. Indian council of Medical Research's National Cancer Registry Programme: Two year report of the Population based Cancer Registries 1999-2000 Incidence and Distribution of cancer-2005

[8]. JS Temel, J McCannon, JA Greer, VA Jackson, P Ostler, WF Pirl, TJ Lynch, JA Billings, Aggressiveness of Care in a Prospective Cohort of Patients With Advanced NSCLC, Journal of Oncology Stat Cancer 2008 Aug 15;113(4):826-833

[9]. Chute et al., Twenty years of phase III trials for patients with extensive-stage-small-cell lung cancer: perceptible progress, Journal of Clinical oncology 17:1794,1999

[10]. Balachandran.K, Dr. R.Anitha, Expert system tool to study the feasibility of diagnosing Lung cancer disease by non-clinical method, Journal of Computer Applications, 9-13, Vol. I issue No.2 Apr-Jun 2008

[11]. Sang Min Park, Min Kyung Lim, Soon Ae Shin & Young Ho Yun (2006) Impact of prediagnosis smoking, Alcohol, Obesity and Insulin resistance on survival in Male cancer Patients: National Health Insurance corporation study Journal of clinical Oncology, Vol 24 Number 31 November 2006

[12]. Yongqian Qiang, Youmin Guo, Xue Li, Qiuping Wang, Hao Chen, & Duwu Cuic(2007) The Diagnostic Rules of Peripheral Lung cancer Preliminary study based on Data Mining Technique Journal of Nanjing Medical University, 21(3): 190-195

[13]. Murat Karabhatak, M.Cevdet Ince (2008) Expert system for detection of breast cancer based on association rules and neural network Journal: Expert systems with Applications

[14]. ICMR Report(2006): Cancer Research in ICMR Achievements in Nineties

[15]. Ta-Cheng Chen, Tung-Chow Hsu (2006)A GAs based approach for mining breast cancer pattern Journal: Expert systems with Applications 30(2006) 674-681

[16]. Petra Perner (1992) Mining Knowledge in X-ray images for Lung cancer diagnosis Journal: Computer vision and applied computer Sciences

[17]. W.Z.Liu, A.P.White, M.T.Hallissey, J.W.LFielding (1995) Machine learning techniques in early screening for gastric and oesophageal cancer Artificial Intelligence in Medicine 8(1996) 327-341

[18]. Edward H.Shortliffe, A.Carlisle Scott, Miriam B.Bischoff, A.Bruce Campbell, William vanMelle, Charlotte D Jacobs (~1982) Oncocin: An Expert system for Oncology protocol Management Proceedings of the 7th IJCAI 1981

[19]. Edward H.Shortliffe, A.Carlisle Scott, Miriam B.Bischoff, A.Bruce Campbell, William vanMelle, Charlotte D Jacobs (1981) An Expert system for Oncology protocol Management Proceedings of the 7th IJCAI 1981 chapter 35 653-665

[20]. James S.Gordon(2008) Mind-body, Medicine and Cancer Journal: Hematology Oncology clinic N.America 22(2008) 683-708

[21]. Kemal Polat, Salih Gunes (2008) Computer aided medial diagnosis system based on Principal Component analysis and artificial immune recognition system classifier algorithm Expert systems with Applications 34(2008) 773-779