REMOTE HEALTH MANAGEMENT TO HANDLE BASIC HEALTH DATA USING FUZZY APPROACH

SAYAN DAS

Abstract: In India, where nearly three quarters of 1.3 billion population lives in rural areas, providing efficient health care service is a significant challenge. Due to tremendous shortage of trained manpower and huge cost for setting up state-of-the-art facilities, it is often not possible to deliver proper healthcare services in the rural and remote areas. Lack of accurate and timely information further adds to the problem. In the paper a fuzzy based system has been proposed which analyze the basic data of the patients (blood pressure, pulse, height, weight, temperature) acquired and investigate whether a person is normal or needs any health care. Heterogeneity in the data set is handled using fuzzy variables in the rule set and severe condition of the patient is evaluated based on the presence of abnormal symptoms indicating the strength of the outcome i.e. normal or disease. Relational matrix between the combination of symptoms and each patient's outcome is built and used as training samples for classification. The discovered rule is more understandable to a human and more accurate to find normal or disease class. Using 10 fold cross validation 92.5% accuracy is achieved.

Keywords: Weightage of outcome, Lukasiewicz method, Frobenius Norm, Frobenius distance, Severity of disease.

Introduction: Rural health care sector is one of the biggest challenges in India due to lack of infrastructure, more specifically expert manpower. In order to provide primary health care services to the patients of remote villages health kiosks are set up in rural area for obtaining basic health related data like blood pressure, pulse, height, weight etc. There is dearth of trained manpower who can investigate the health of the patients using acquired data. Therefore for providing primary health service, large amount of data need to be analyzed using computational intelligent methods which can minimize involvement of the doctors, not attendant frequently in the kiosk.

Researchers have suggested different approaches and methodologies to deal with the diversified issues pertaining to medical diagnosis. In the reported studies, the need, importance, potential, necessity of fuzzification and approaches for designing of medical diagnosis expert systems are discussed [1-5]. Some studies have been conducted to verify the suitability of fuzzy set theory and its derived theories for developing knowledge based systems and fuzzy sets to model the medical concepts [6, 7]. Decision support systems for diagnostic purposes of human diseases helped patients and practitioners to a great extent [8-11]. The areas like diversified applications are developed using fuzzy logic includes: fuzzy models for illness, heart and cardiovascular disease diagnosis, asthma, abdominal pain, tropical diseases, neurological diseases ,medical analogy of consumption of drugs, malaria diagnosis, specific organs [12,13]. However, diagnosing a person as normal or need primary health care has not been yet addressed, by analyzing the basic health data, which is very much relevant in rural India due to dearth of infrastructure.

In the paper a fuzzy based system has been proposed which analyze the basic data of the person(blood pressure, pulse, height, weight, temperature) acquired and investigate whether the person is normal or need any health care. In rural area, patients coming in the health kiosk cannot express their discomforts or symptoms using proper vocabulary for various reasons. In most of the cases the query input is imprecise, incomplete and ambiguous using which diagnosis of the diseases are not possible. Fuzzy logic based system is used to incorporate human reasoning power and tolerant to imprecise data. The patients' data are non-homogeneous.For example, a sample patient data (patient id 10339) consist of Blood Pressure (156/96 in mmHg), Pulse (87 in bpm). The measuring unit is different in both the symptoms. Therefore it is necessary data to process in a homogeneous form. To deal with heterogeneity of the patients' dataset, fuzzy sets are defined with proper semantics. The patients’ symptoms are represented using categorical value with the help of standard value and expert knowledge. The fuzzy value of the fuzzy sets is evaluated. The important factor for finding a patient is disease or normal is weightage of outcome. Weightage of outcome depends on the severity of the disease and number of disease related to particular symptoms. Severity of the disease is calculated using severity function. Relation between each member of fuzzy sets and weightage of outcome are generated using Lukasiewcz method. Aggregation of different members of fuzzy set is done through fuzzy t-norm operator. Relational matrix between different combinations of symptoms and weightage of symptoms are built. Patients' query is applied to this matrix to find if patient's data belonging in disease or normal class level. Similarities between the matrixes are evaluated using the Frobenius method. The discovered rule is more understandable to a human and more accurate to find normal or disease class. Using 10 fold cross validation 92.5% accuracy is achived.

The rest of the paper is arranged as follows.Section II describes the methodology applied to detect patient diagnosis. Section III gives the experimental results. Finally, Section IV concludes the paper and points out some further research work.

Methodology: The fuzzy based expert system is designed using the health records of rural patients collected from the health kiosks situated in remote villages. The patient database contains different attributes among which six basic health related attributes (patient_id, age, blood pressure, pulse, height, and weight) are used for primary diagnosis of the patients. However, the raw data set is imprecise due to measurement error and heterogeneous, therefore transformed into variables, considering medical science data as reference point. Thevalue of the fuzzy variables is represented using different fuzzy sets with proper membership functions. The proposed methodology after investigating the basic attributes classify the patient either disease or normal.

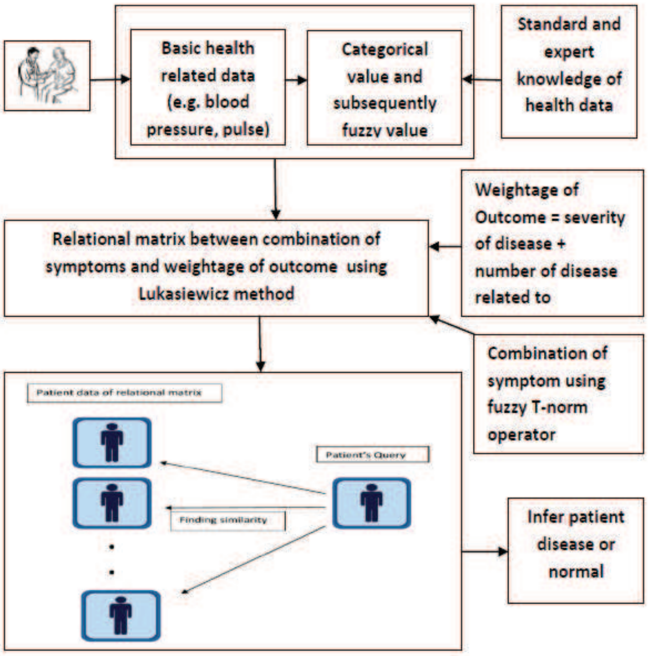

Fig. 1: Block Diagram of the Proposed Primary level Disease Detection System

Data Preparation: Few terminologies are defined below for data preparationin fuzzy rule based system

Membership function: A fuzzy set A on the universe of discourse X is defined as μ_A:X→ [0,1], where each element of X is mapped to a value between 0 and 1. The membership value of different fuzzy set is evaluated using membership function. In the paper Gaussian membership function are used for calculating membership value of the fuzzy sets. The membership value is

$$\mu_A^{\ i} : (x) = \exp(-\frac{(c_i - x)^2}{2\sigma_i^{\ 2}})$$

where,

c_i and σ_i are the centre and width of the i^{th} fuzzy set A_i, respectively and x is the linguistic variable.

Rule base: Fuzzy rule base has been designed consisting of symptoms as antecedents and class label as consequent. Fuzzy variables are used to represent the symptoms while class labels are assigned based on the experts' opinion.

The patients' data which are acquired by the sensors are non-homogeneous, inconsistent. Therefore it is necessary data to process in a homogeneous form. Here the patients' symptoms are represented using categorical value. To deal with heterogeneity of the patients' dataset, fuzzy sets are defined with proper semantics. Membership value of the fuzzy sets is evaluated. Data preparation phase consist of different steps:

1) Basic health related data like BP, Pulse, height, and weight are collected through sensors by the practitioner.
2) Using standard and expert knowledge data are converted into categorical value.
3) Next, using Gaussian function membership value of the data is evaluated.

For example, a sample patient data (patient id 10339) consist of BP (156/96 in mmHg), Pulse (87 in bpm).The standard value of blood pressure and pulse are as follow:

TABLE I. STANDARD VALUE OF BP AND PULSE

Input Variable	Range	Fuzzy Sets
Blood Pressure (Systolic/diastolic) In mmHG	(90-120)/(60-80)	Normal
	(121-139)/(81-89)	High(Pre Hypertension)
	(140-159)/(90-99)	High(Stage 1)
	160/100	High(Stage 2)
	Less than 90/60	Low
Pulse In bpm	(60-100)	Normal
	More than 100	High
	Less than 60	Low

Membership value of the particular patient is as follow:

TABLE II. MEMBERSHIP VALUE OF SYMPTOMS

Patient-id	BPN	BPVH	BPH	BPVL	BPL	PN	PVH	PH	PVL
10339	0.179	0.821	0.612	0.017	0.018	0.566	0.09	0.1	0.217

Weightage of outcome: The main challenge of weightage of outcome is severity of disease and the number of symptoms related to disease. Thus the weightage of outcome is a function related to severity of disease andthe number of symptoms related to the disease.

Weightage of outcome = severity of disease + number of disease related to symptoms.

Severity occurs when extreme condition occurs between symptoms.For example, a patient possesshigh blood pressure and low pulse have more severe than a patient have high blood pressure andhigh pulse. The following are severe condition:

TABLE III. DIFFERENT SEVERITY CONDITION

Blood Pressure	Pulse
High	Low
Low	High
High	Very Very High
Low	Very Very Low
Very Very High	High
Very Very Low	Low

Severity of disease can find using

$$\frac{k}{\max(p,q)} \oplus \frac{1}{P^{d+1}}$$

where,

kis random number between [1 – 2].

max() is a function to find the maximum membership value.

p,q are the membership value of symptoms.

d indicate the difference between type. The value assign with the type such as ‘low’:0, ‘normal’:1, ‘high’:2, ‘veryhigh’: 3 and ‘very very high’:4. For example, a patient who has very high Blood Pressure and high pulse the value of d is (3-2) =1.

The severity function works as follow:

Let x denotes $\frac{k}{\max(p,q)}$ and y denotes $\frac{1}{P^{d+1}}$.

i) If the value of x is greater than y then the severity value is value of x.

ii) If the value of y is greater than x then the severity value is value of y.

iii) If the value of x is nearly equal to y then the severity value is zero.

The second portion for calculating of weightage of outcome is number of diseases related to symptoms.

For example, a patient has high blood pressure and normal pulse and corresponding membership value is 0.821 and 0.566 respectively. Severity of the disease can find

$$\frac{1.42}{\max(0.821, 0.566)} \oplus \frac{1}{0.821^2}$$

where k=1.42 (randomly chosen) and direction = (2-1)=1. Thus the severity is 1.73.

Symptoms like high blood pressure related to 9 types of diseases viz. heart disease, kidney damage, brain damage, aneurysm, eye damage, sexual dysfunction, bone lose, sleeping trouble, nausea or vomiting. Patient has high blood pressure and has symptoms of headache and vomiting. The weightage of outcome is (2/9) =0.22 with respect to disease related to symptoms.

The total weightage of outcome is the summation of severity of disease and number of disease related to symptoms. For example, for that particular patient the weightage of outcome is (1.73+0.22) = 1.95.

Relational matrix: Few terminologies are related to preparation of relational matrix in fuzzy rule based system.

Lukasiewicz method: The relationship between particular symptoms and class label in a rule is evaluated using Lukasiewicz method as given bellow,

x ϴ y $= \max\{0, x+y-1\}$,for all elements $x, y \in [0,1]$.

The operation ϴ means logical conjunction applied on elements x and y representing the membership value of the symptom and weightage of outcome.

For example, a particular patient with id 10339 has high blood pressure with membership value 0.821. The relationship between symptom high blood pressure and weightage of outcome are max (0, 0.821+1.95 -1) = 1.771. Similarly the relationship between all other symptoms and weightage of outcome are calculated.

Fuzzy Aggregations: Fuzzy aggregations are logic operators like t-norm, t-conorm and averaging operators applied to fuzzy membership values.The min and max operators are defined as

Min {a,b}= $a \wedge b$, for all elements $a,b \in [0,1]$.

Max {a,b}= $a \vee b$, for all elements $a,b \in [0,1]$.

Combination of different symptoms is formed using fuzzy t-norm operator. Relational matrix for all the patients is built between combination of symptoms and strength of outcome.

Similarity Calculation

Few terminologies are related to similarity calculation in fuzzy rule based system.

Frobenius Norm: The Frobenius norm, sometimes called Euclidean norm is matrix norm,of an $m \times n$ matrix A defined as the square root of the sum of absolute squares of its elements,

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}$$

The Frobenius norm can also be considered as a vector norm.It is also equal to the square root of the matrix trace of AA^H, where A^H is the conjugate transpose, i.e.,

$$\|A\|_F = \sqrt{Tr(AA^H)}.$$

Frobenius distance:The Frobenius distance of $m \times n$ matrix A and B is define as,

Frobenius distance= $|(fnorm(A) - fnorm(B))|$,

where fnorm = square root of sum of squares of all singular values.

Frobenius distance finds the similarity between the two matrixes. Minimum the value of distance indicate the matrices are similar.

For example, frobenius norm of the particular patient 10339 is 5.9004. Similarly, frobenius norm of the patients id 10342, 10346, 10361 are 5.91, 6.52, 7.6 respectively. The frobenius distance between the particular patient 10339 and the others are 0.0096, 0.6196, and 1.6996. The minimum of this distance are 0.0096. Thus the patients 10339 and 10342 are similar.If the minimum distance is less than 1.00 classifies the patient as disease. Here the distance is less than 1.00 and the patient classify as disease.

Result and Discussion

We will demonstrate the process of 200 patients' records acquired by the practitioner of kiosk at rural area. In the following Fig. 2.a. and 2.b. show the membership curve of blood pressure and pulse. Table IV shows membership value of different symptoms of some patients.

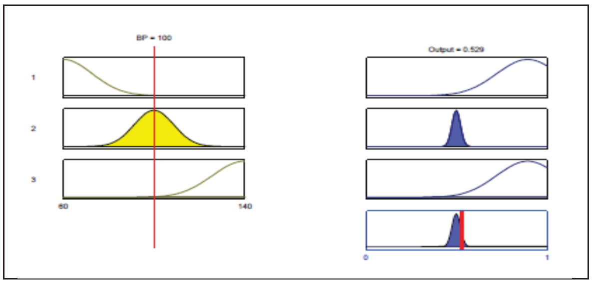

Fig 2.a.: Membership curve of blood pressure

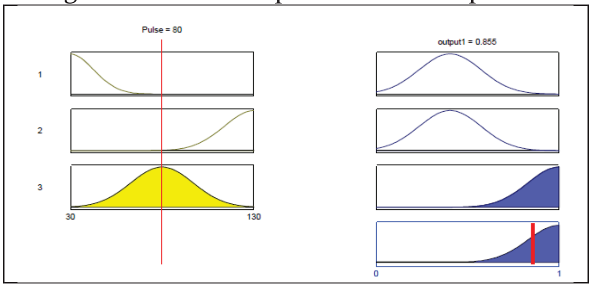

Fig 2.b.: Membership curve of pulse

Table Iv. Membership Value of Symptoms

Patient-id	BPN	BPVH	BPH	BPL	PN	PH
10338	0.506	0.194	0.201	0.017	0.564	0.09
10341	0.79	0.21	0.37	0.022	0.382	0.718
10340	0.508	0.192	0.209	0.023	0.594	0.732
10342	0.179	0.821	0.612	0.014	0.556	0.09

Here the symbols BPN, BPVH, BPH, BPL, PN, PH and W denotes Blood pressure normal, Blood pressure very high, blood pressure high, blood pressure low, pulse normal, pulse high and weightage of outcome respectively.

The weightage of outcome is calculated by adding severity of disease and number of disease related to symptoms. To calculate severity of disease the random value of k plays an important role. The value of k lies in between [1 – 2]. It is found experimentally severity is maximum if the value of k lies in between 1.40 - 1.50. Fig. 3 shows the maximum severity occurs at 1.43 and Table V shows different disease related to symptoms. Weightage of outcomevalue of 5 different patients is shown in Table VI.

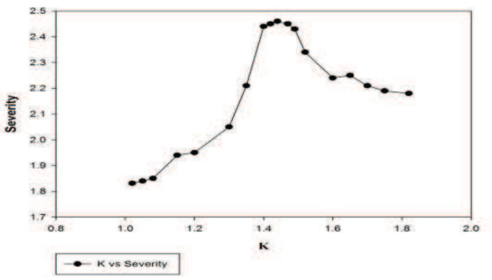

Fig 3: K v/s severity plot

K denotes random number between [1-2].

TABLE: V. SYMPTOMS RELATED TO DIFFERENT DISEASE

Symptoms	Related disease
High blood Pressure	Heart disease, kidney damage, brain damage, aneurysm, eye damage, sexual dysfunction ,bone lose, sleeping trouble, nausea or vomiting.
Low blood pressure	Dizziness,lack of concentration,blurred vision,nausea,paleskin,rapid breathing, fatigue,depression, thrist.
High pulse	Dizziness, shortness of breath, heart palpitation, chest pain,fainting.
Low pulse	Heart failure, chest pain,hypotension.

TABLE VI WEIGHTAGE VALUE OF PATIENTS

Patient-id	Weightage
10338	2.643
10341	2.077
10340	1.998
10342	1.807
10343	1.84

Relational ship between symptoms and weightage are found using Lukasiewicz method. Relational matrix is built for different combination of symptoms. Table VII show a portion of relational matrix of different patients.

Patient query is applied to find a patient is disease or normal. The similarity between the patient query matrix and different patients' relational matrix are found using frobenius method. The distance which is minimum is similar to the query patient and if the distance is less than 1 then the patient is in disease class otherwise normal. Table VIII shows frobenius norm and distance of different patients.

TABLE VII. RELATIONAL MATRIX OF SYMPTOMS

Patient -id	BPN & W	BPV H& W	BP H& W	BP L& W	PN & W	PH& W
10338	2.149	1.837	1.844	1.66	2.207	1.733
10341	1.867	1.287	1.447	1.098	1.459	1.795
10340	1.506	1.19	1.207	1.021	1.592	1.73
10342	0.987	1.628	0.824	0.821	1.363	0.897
10343	1.024	1.656	0.858	0.863	1.407	0.93

Here the symbols have usual meaning as describe earlier in the paper.

TABLE VIII. RELATIONAL MATRIX OF SYMPTOMS

Patient-id	Frobenius norm	Frobenius norm of Patient Query	Minimum Frobenius distance	Remark
10338	10.06	6.5297 of patient id 10346	0.33	The patient 10346 is similar to 10340 and in disease class.
10341	7.3582			
10340	6.8597			
10342	5.9108			
10343	6.0878			

Finally, classify the patient data set using 10-fold cross validation technique. Table IX shows different statistical measure of 10 fold classification.

TABLE IX STATISTICAL MEASURE

Accuracy	Misclassification Rate	True Positive rate	False Positive rate	Prevalence	Precision
92.5	0.0972	0.972	0.5	0.9	0.946

Fig 4. shows the graph of number of patients and accuracy.

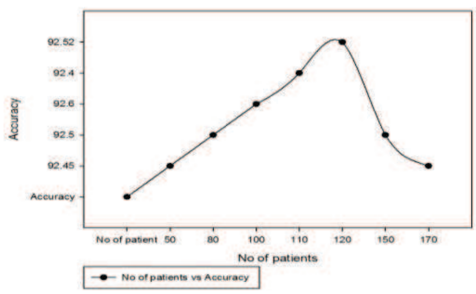

Fig 4: Number of patients v/s accuracy plot

Conclusion: In this paper we have discussed the design and implementation of a Fuzzy based System for provisional diagnosis. The system helps to investigate whether the person is normal or need any health care. In the implementation, we have taken the most basic symptoms of patient and evaluate the severity to each of the symptom using weightage of outcome. We get satisfiable result which help primary diagnosis of rural people. But, the availability of patient related data is not easy for different legal and ethical reasons which initially created bottleneck to assess the proposed method. However, presently performance of the proposed method is remarkable and replicated by establishing kiosks in rural villages under a project scheme sponsored by Govt. of India. Studying the nature of the patients coming in the kiosks including other symptoms and increase the accuracy of severity function is further designed in our future work. However, it is worth to mention that the system is for primary healthcare of the people in rural area and in case of emergency is referred to the experts.

References:

1. Bruce G. Buchanan and Edward H. Shortliffe, “Rule Based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project”, Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA, 1984, ISBN:0201101726.
2. Engle, R.L., Flehinger, B.J. 1987. Why Expert Systems for Medical Diagnosis are not being generally used: A valedictory opinion. Bulletin of the New York Academy of Medicine, 63, 2: 193-198.
3. Ali, Z., Singh, V. 2010. Potentials of Fuzzy Logic : An Approach To Handle Imprecise Data. American Medical Informatics Association, 2, 4: 358-361.
4. Ameri, A., Moshtaghi H., Design and Development of an Expert System in Differential Diagnosis of Maxillofacial Radio-lucent Lesions. Web URL,http://www.idt.mdh.se/ kurser/ct3340/archives/ht08/papersRM08/21.pdf.
5. Adlassnig, K.P., Kolarz, G., Scheithauer, W. 1985. Present State of the Medical Expert System CADIAG-2. Methods of Information in Medicine, 24: 13-20.
6. Ahmad, M.R., Mahdi, A.A., Salih, A.A. 2009. Designing a Disease Diagnosis System by Using Fuzzy Set Theory. Proceedings of 5th Asian Mathematical Conference, Malaysia: 256-260.
7. Schumann, A. 2010. Unconventional Probabilities and Fuzziness In Cadiag's Computer Assisted Medical Expert Systems. Studies in Logic, Grammer& Rhetoric, 22, 35: 113-124.
8. Jeharon, H., Seagar, A., Seagar, N. 2005. Feature Extraction From Phonocardiogram For Diagnosis Based on Expert System. Int. Conf. of Proc. of IEEE Engineering in Medicine & Biology Society, 5: 5479-5482.
9. Patra, P.S., Sahu, D.P., Mandal, I. 2010. An Expert System For Diagnosis of Human Diseases. Int. J. of Computer Applications, 1, 13, 71-74.
10. Bates, J.H.T., Young, M.P. 2003. Applying Fuzzy Logic to Medical Decision Making in the Intensive Care Unit. American Journal of Respiratory & Critical Care Medicine, 167: 948-952.
11. Saritas, I., Allahverdi, N., Sert, I.U. 2003. A Fuzzy Expert System Design for Diagnosis of Prostate Cancer. Int. Conf. on Computer Systems & Technologies: 1-7.
12. Holzmann, C.A., Perez, C.A., Rosselot, E. 1988. A Fuzzy Model For Medical Diagnosis. Medical Progress Through Technology, 13: 171-178.
13. Ali, A., Mehdi, N. 2010. Fuzzy Expert System for Heart Disease Diagnosis. Proc. of Int. Multi Conference of Engineers & Computer Scientists, I: 1-6.

Sayan Das
Indian Institute of Engineering Science and Technology,
Shibpur, India