Design Methodology of a Fuzzy Knowledgebase System to predict the risk of Diabetic Nephropathy

Rama Devi.E[1], Dr.Nagaveni.N[2]

**[1]Assistant Professor, Dept of Comp Science,
NGM College, Pollachi – 642 002, Coimbatore, India.**

**[2]Assistant Professor, Dept of Mathematics,
Coimbatore Institute of Technology, Coimbatore, India.**

Abstract

The main objective of the design methodology of a Fuzzy knowledgebase System is to predict the risk of Diabetic Nephropathy in terms of Glomeruler Filtration Rate (GFR). In this paper, the controllable risk factors "Hyperglycemia, Insulin, Ketones, Lipids, Obesity, Blood Pressure and Protein/Creatinine ratio" are considered as input parameters and the "stages of renal disorder" is the output parameter. The input triangular membership functions are Low, Normal, High and Very High and the output triangular membership functions are s1, s2, s3, s4 and s5. As the renal complications are now the leading causes of diabetes-related morbidity and mortality, a FKBS is designed to perform the optimum control on high risk controllable risk factors by acquiring and interpreting the medical experts' knowledge. Fuzzy logic is used to incorporate the available knowledge into intelligent control system based on the medical experts' knowledge and clinical observations. The proposed FKBS is validated with MatLab, and is used as a tracking system with accuracy and robustness. The FKBS captures the existence of uncertainty in the risk factors of Diabetic Nephropathy, resolves the renal failure with optimum result and protects the patients from End Stage Renal Disorder (ESRD).

Keywords: *Diabetes Mellitus, GFR, ESRD, Uncertainty, Fuzzy knowledgebase System, Inference Engine*

1. Introduction

Diabetes Mellitus is one of the most dangerous diseases in the modern society and represents not only a medical but a social problem. Diabetic is pandemic in both developed and developing countries. In 2000, there were an estimated 175 million people with diabetes worldwide and by 2030, the projected estimate of diabetes is 354 million [1]. In India alone, the prevalence of diabetes is expected to increase from 31.7 million in 2000 to 79.4 million in 2030 [1].

The proportion of patients with end stage renal disorder caused by diabetes has progressively increases day by day worldwide and Diabetic Nephropathy is the single most common cause of ESRD [7].

The major problem in diagnosing and predicting a disease is the uncertain variations in risk factors which occur due to the sedentary life style, food habit, stress, age, environment etc. Generally the risk factors in medical domain are classified as controllable (Blood Pressure, Lipids, Obesity, Hyperglycemia), non-controllable (age, heredity, sex) and contributing factors (smoking, alcohol, stress). Due to uncertainty, information is incomplete, fragmentary, not fully reliable, vague, contradictory, or deficient in some way. The existence of uncertain factors with errors is numerous in medical domain and it is to be solved in an efficient way.

The medical errors can be solved with the help of experts' knowledge. As the experts' knowledge are so valuable and tacit in nature which cannot be used as it is, therefore interpreting them in terms of rules is the best approach to resolve the uncertainty. Hence, in this paper, it is proposed to design a FKBS to overcome the uncertainties and to predict the risk of Diabetic Nephropathy. This will help to control the controllable risk factors for proper diagnosis and treatment, and to protect the Diabetic patients from ESRD [11].

2. Diabetic Nephropathy And Fuzzy Logic

Diabetes Mellitus is a group of metabolic diseases characterized by high blood sugar levels that result from defects in insulin secretion, or action or both. Diabetes leads to many major complications and renal disorders are now the leading causes of illness and death in diabetic patients. Diabetes Mellitus has been found to increase the risk of renal disorder which leads to End Stage Renal Disorder (ESRD) both in women and men.

The major risk factors considered in this paper for the impact of Diabetes Mellitus on renal are Hyperglycemia, Insulin, Ketones, Lipids, Obesity, Blood Pressure and Protien/Creatinine. The risk factors are uncertain due to the food habits, sedentary life style; stress and strain etc. depends on the individual life style, environment and occupation. The uncertain variations in these risk factors vary in each and every individual. The best logic to capture and resolve the existence of uncertainty to certainty is FUZZY LOGIC. Fuzzy logic yields results superior than conventional mathematical logic and also it is the suitable logic for handling words as linguistic variables rather than number. To estimate and predict the prevalence of Diabetes Mellitus based on the risk factors which inferences the renal disorder is validated using MATLAB to overcome the unexpected, sudden stop of the functioning of kidneys which in turn reduces the rate of mortality.

3. Methodology

3.1 Fuzzy Logic and Impact of Diabetes Mellitus on Renal

Diagnosis of a disease is a significant and tedious task in medical domain. The detection of renal disorders due to Diabetes Mellitus from various factors or symptoms is a multilayered issue which is not free from false presumptions often accompanied by unpredictable efforts.

The main objective of this paper is to identify the risk factors of Diabetes Mellitus which causes the severe impact on renal. Based on the medical expert's knowledge, the risk factors may be grouped as

- Controllable risk factors which can be controlled
- Non-Controllable risk factors which cannot be controlled
- Contributing risk factors which can be controlled or cannot be controlled depends on individual's contribution.

In this proposed work, the risk factors identified are Hyperglycemia, Insulin, Ketones, Lipids, Obesity, Blood Pressure and Protien/Creatinine which are controllable. The values of risk factors based on Universe of Discourse varies patient to patient depends on the symptoms that leads to so many complications.

To overcome the uncertainty, fuzzy logic is used which yields results superior than the conventional crisp value. By applying fuzzy logic the evaluation of risk factors can be interpreted properly for effective diagnosis to minimize the rate of mortality with maximum efficiency.

3.2 Design and Validation of Fuzzy Rule Base System

The essential part of the proposed Fuzzy knowledgebase System is to convert the given set of linguistic control strategy based on expert's knowledge into an automatic control strategy. It consists of the four major components namely Fuzzification Interface, Rule Base, Inference Engine and Defuzzification. Each component explains how the risk factor values are validated to achieve the aim and objective of the proposed Fuzzy knowledgebase System.

3.2.1 Fuzzification Interface

Fuzzy Interface converts crisp values into fuzzy value. There is a degree of membership for each linguistic term that applies to that input variable.

Fuzzy Representation of Input Parameters

The membership function for Insulin and P/C Ratio are shown below. In the same way the membership function are constructed for other risk factors also.

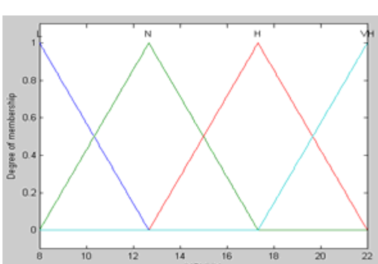

Fig: 1 Membership function for Insulin

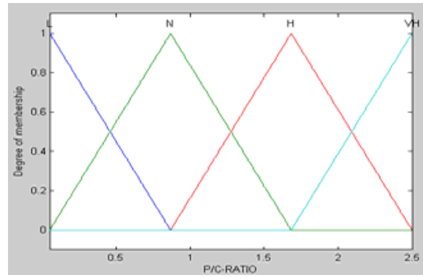

Fig 2 Membership function for P/C ratio

Fuzzy Representation of Output Parameter

The membership function for output parameter stages of renal is shown below. The output represents the different stages of renal disorder.

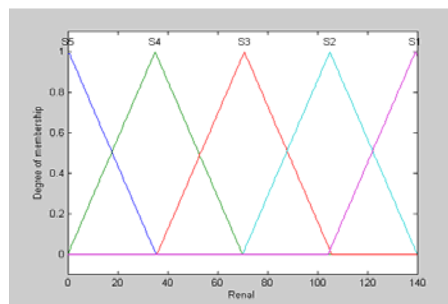

Fig 3 Membership function for stages of Renal

3.2.2 Knowledgebase

The Fuzzy Knowledgebase System consists of a set of fuzzy propositions and is derived from the knowledge of medical expertise [17]. A fuzzy proposition establishes a relationship between different input fuzzy sets and output fuzzy sets. The Fuzzy Knowledgebase System is constructed based on the formula $\mathbf{n^m}$ where n is the risk factor and m is linguistic variable. The proposed FKBS consists of 7 risk factors and 4 linguistic variables, which leads to 16384 rules. This will effectively evaluate all the possible variations in the Universe of Discourse for the patients with renal disorder due to Diabetes Mellitus [16]. It also predicts and forecast the risk at the earliest of any stage to protect the patient from fatal end. Here only 16 rules have shown for reference.

Table 1: Sample rules framed for the impact of Diabetes Mellitus on Renal

Sno	HG (FPG/ 2hPG)	Insulin	Ketones	Lipids (LDL/HDL)	Obesity	BP (sys/dias)	P/C	Stages of Renal/ GFR
1.	L	L	L	L	L	L	L	S3
2.	L	L	L	L	L	L	N	S2
3.	L	L	L	L	L	L	H	S3
4.	L	L	L	L	L	L	VH	S4
5.	L	L	L	L	L	N	L	S2
6.	L	L	L	L	L	N	N	S1
7.	L	L	L	L	L	N	H	S2
8.	L	L	L	L	L	N	VH	S3
9.	L	L	L	L	L	H	L	S3
10.	L	L	L	L	L	H	N	S2
11.	L	L	L	L	L	H	H	S3
12.	L	L	L	L	L	H	VH	S4
13.	L	L	L	L	L	VH	L	S3
14.	L	L	L	L	L	VH	N	S2
15.	L	L	L	L	L	VH	H	S4
16.	L	L	L	L	L	VH	VH	S4

3.2.3 Inference Engine

The purpose of the inference engine is to infer information based on the knowledgebase to take major decisions in diagnosis. The knowledgebase is time variant, adaptive and robust in structure [18]. The number of fuzzy sets describing input variables increases, the number of proposition also increases. Since, writing all propositions is a tedious job, a matrix like representation may be followed and is known as Fuzzy Associative Memory. Based on the knowledgebase framed, inferences are acquired to take major decisions regarding therapy and treatment.

Table 2: FAM for the sample rules framed

LP,OB,PC,BP / HG/INS/KET	LLLL	LLLN	LLLH	LLLVH
LLL	S3	S2	S3	S4
LLN	S2	S1	S2	S3
LLH	S3	S2	S3	S4
LLVH	S3	S2	S4	S4

Based on the inferences, the knowledgebase may be restructured for proper, perfect and fast diagnosis of the disease.

Decision making logic component

The decision rules are constructed for input parameters and control output values to find the active cells so that what control actions should be taken as a result of firing several rules. Finally, the aggregation of minimum control outputs is taken into consideration to maximize the grade of output to resolve the uncertain linguistic input to produce crisp output. The proposed system consists of 7 parameters with 4 triangular membership functions for each parameter. According to the formula n^m , 4^7 = 16384 rules are framed in which 128 rules are fired for the sample input values of each parameter as, HG (L) = 65/72, Insulin (N) = 14.995, Ketones (H) = 30, Lipids (L) = 80/180, Obesity (L) = 12, Blood Pressure (H) = 130/82, Protien/Creatinine (L) = .08.
According to Mamdani Inference Method [19], the aggregation of control output is

$$\mu agg(b)=\max\{\min(1/4,\mu_{s3(b)}), \min(1/2.41,\mu_{s2(b)}), \min(1/2.41,\mu_{s1(b)})\}$$

3.2.4 Defuzzification Interface

Defuzzification is a process to get a non-fuzzy control action that represents the best possibility distribution of an inferred fuzzy control action. Unfortunately there is no systematic procedure for choosing a good defuzzification strategy. So, by considering the properties of application, any one of the five methods can be selected for defuzzification methods. In this paper, the "MEAN OF MAXIMUM' defuzzification method is applied to find the intersection point of μ = 1/2.41 with the triangular fuzzy number $\mu_{s2\ (b)}$ and $\mu_{s1(b).}$ By substituting (i) μ = 1/2.41 in s2 (b),

μ = 90-b/30, 30 μ = 90-b, 30*1/2.41 = 90-b, 12.45 = 90-b, 12.45-90 = -b, b = 77.5

μ = b-60/30, 30 μ = b-60, 30*1/2.41 = b-60, 12.45 = b-60, 12.45+60 = b, b= 72.45

after aggregation, the values are Ç1 = 77.5, Ç2 = 72.45, the output parameter i.e. the stage **s2** in terms of Glomeruler Filtration Rate (GFR) of a patient [6] is identified to produce the crisp output as

Zm* = (77.5+72.5)/2 = 149.95/2 = 74.98

(ii) μ = 1/2.41 in s1 (b),

μ = 140-b/50, 50μ =140-b, 50*1/2.41 = 140-b, 20.75 = 140-b, 20.75-140 = -b, b = 119.25

μ = b-90/50, 50μ = b-90, 50*1/2.41 = b-90, 20.75= b-90, 20.75+90 = b, b = 110.75

after aggregation, the values are Ç1 = 119.25, Ç2 = 110.75, the output parameter i.e. the stage **s1** in terms of Glomeruler Filtration Rate (GFR) of a patient is identified to produce the crisp output as

Zm* = (119.25+110.75)/2 = 230/2 = 115.

According to Mean of Maximum method, the maximum value is considered, for obtaining the optimum result. In this paper, two values obtained are 74.98 and 115 in which 115 is the maximum which represents the stage s1 of renal disorder in terms of GFR .Hence the diagnosis may be concluded as, the patient is in stage s1, which indicates that the patient's renal is in normal state .

Therefore the proposed Fuzzy knowledgebase System is proven to be a rapid rule base application to monitor, predict and forecast the nature of renal disorder due to the impact of Diabetes Mellitus. It also helps to maintain the P/C Ratio in renal system, to avoid the renal disorders.

4. Results

In order to validate the fuzzy logic approach used in construction of FKBS, the extensive simulation is carried out using MATLAB. The system responses with

- variation defined in the membership function as a rule viewer, surface view, cluster formation and preservation .
- data training, checking, testing with sample data to capture the error
- modulation of risk factors reading which varies from patient to patient with the level of risk due to the impact of DM on renal disorder and the results are favorable.

4.1 Micro View of the simulated knowledgebase for Diabetic Nephropathy

The constructed knowledgebase is simulated using **MATLAB** to identify the output parameter- the stages of Renal Disorder of a diabetic patient.

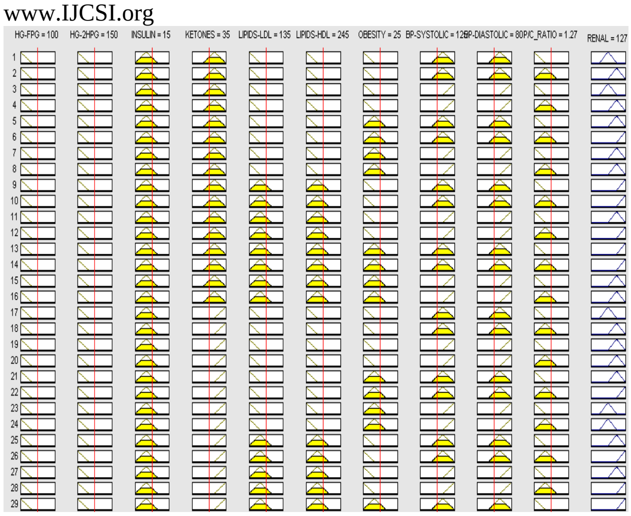

Fig 4: Simulated view of the rule base – Impact of |Diabetes Mellitus on Renal

4.2 Surface View of Mapping with Input Parameters vs. Output Parameter

Surface view shows the three dimensional view curves that represent the mapping with input parameters vs. the output parameter, the stages of Renal disorder. Sample surface view of the proposed Fuzzy knowledgebase System is shown as per the rules framed.

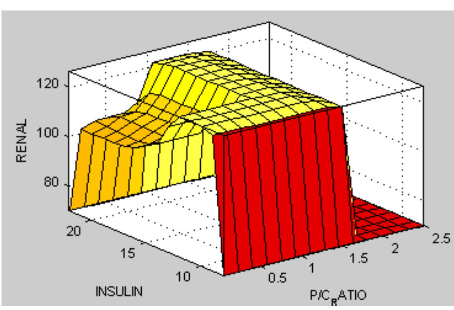

Fig 5: Surface view of Insulin Vs P/c Ratio

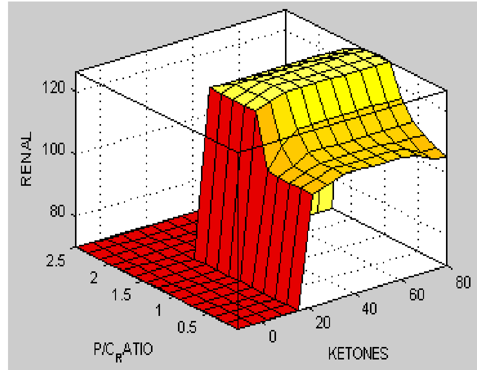

Fig 6: Surface view of P/C Ratio Vs Ketones

Likewise the surface view can be generated for all the seven input parameters with the output parameter which helps to monitor the stages of renal disorder based on the variations of the risk factors from time to time.

4.3 Grouping o f Patients Data according to the Risk Factors as Clusters

Clustering is a way to examine similarities and dissimilarities of observations or objects. Data often fall naturally into groups or clusters of observations, where the characteristics of objects in the same cluster are similar and the characteristics of objects in different clusters are dissimilar. In this proposed FKBS, the cluster analysis is used for the purpose of grouping the patients with high risk and low risk. Grouping of clusters also identifies the patients who need the emergency care. Using the Adaptive Neuro-Fuzzy Inference System (ANFIS) editor, membership functions are shaped by training them with input/ output data rather than specifying them manually.

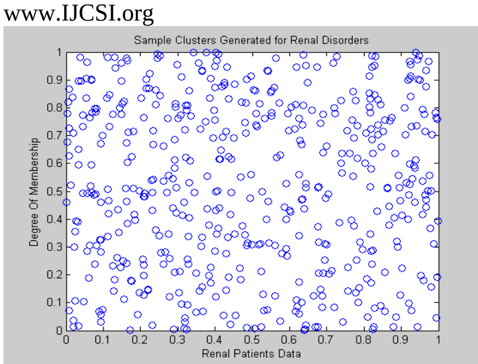

Fig 7: Sample clusters generated for renal disorder

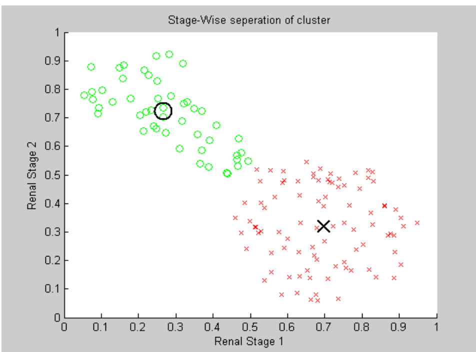

Fig 8 Disease-wise seperation of clusters

4.4 Generation of Fis for Cluster Shape Preservation Using Adaptive Neuro-Fuzzy Inference System (Anfis) .

The shape preservation is applied to organize the same size and shape clusters, to maintain the characteristics and properties of data in such a way to identify and differentiate the patient's level of risk for emergency treatment.

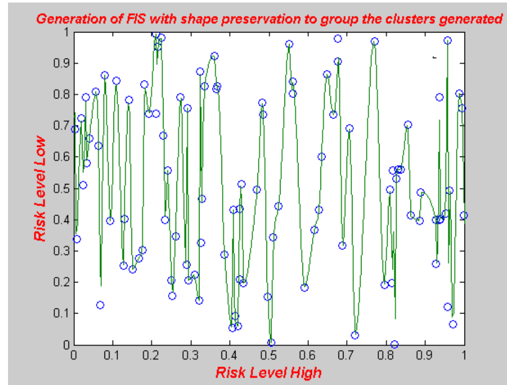

Fig 9 Generation of FIS with shape preservation to group the clusters generated

5. Conclusion

The diagnosis of renal disorders for diabetic patients represents a serious clinical problem. The medical knowledge in this field is characterized by uncertainty, imprecision and vagueness. As the symptoms of diabetic patients are very fuzzy in nature, the use of linguistic approximation enables knowledge to be represented in a more meaningful way. Fuzzy logic is very useful to deal with the absence of sharp boundaries of the sets of symptoms, diagnosis and phenomena of diseases. The use of this approach is contributed to medical decision-making and the development of computer-assisted diagnosis in medical domain.

In this paper, the major risk factors for diabetic patients are identified and the FKBS is designed with 16384 rules which may adopt as an aid for finding the patient's risk of renal disorder for proper treatment and care in time to avoid the ESRD. The designed knowledgebase is simulated with Mat Lab and the results are promising with the values considered

This fuzzy knowledgebase captures all the variations of the symptoms and so, it will be useful to infer the exact stage of renal patient as per the expert's knowledge. This is also useful to take proper decision at the right time for giving various types of treatments which ultimately reduces the rate of mortality. The designed FKBS may be extended to devise a MEMS device for monitoring the status of the risk factors with changing life style, food habits, occupation of human beings as further enhancement of this work.

6. References

[1] Sarah Wild, Gojka Roglic, Anders Green, Richard Sicree, and Hilary King, "Global Prevalence of Diabetes - Estimates for the Year 2000 and Projections for 2030", Diabetes Care 27, pp 1047-1053, 2004.

[2] Constantinos Koutsojannis and Ioannis Hatzilygeroudis, "FESMI: A Fuzzy Expert System for Diagnosis and Treatment of Male Impotence", Knowledge-Based Intelligent Systems for Health Care, pp 1106–1113, 2004.

[3] Rahim F, Deshpande A, Hosseini A, "Fuzzy Expert System For Fluid Management In General Anaesthesia", Journal of Clinical and Diagnostic Research, pp 256-267,2007.

[4] Muhammad zubair Asghar, Abdur Rashid Khan, Muhammad Junaid Asghar: "Computer Assisted Diagnosis for Red Eye", International Journal on Computer Science and Engineering, Volume 1(3), pp 163-170, 2009.

[5] Nazmi Etik, Novruz Allahverdi, Ibrahim Unal Sert and Ismail Saritas, "Fuzzy expert system design for operating room air-condition control systems", Expert Systems with Applications – An International Journal, Volume 36, Issue 6, pp 9753-9758, 2009

[6] Hamidreza Badeli, Mehrdad Sadeghi, Elias Khalili Pour, Abtin Heidarzadeh, "Glomerular Filtration Rate Estimates Based on Serum Creatinine Level in Healthy People", Iranian Journal of Kidney Diseases, Volume 3 ,Number 1, 2009.

[7] Niall G. McLaughlin, Udayan Y. Bhatt, Gil A. Cu, Ronald L. Mars, N. Stanley Nahman, "A Review of the Diagnosis and Therapy of Diabetic Nephropathy", Northeast Florida Medicine Journal, pp 27-31, 2005

[8] Parving HH. "Diabetic nephropathy: prevention and treatment." Kidney Int. 60(5), pp 2041-55,2001.

[9] Mauer S, Steffes M, Brown D., " The kidney in diabetes."Am J Med. 70 pp 603, 1981.

[10]Brenner BM. Regarding: "Management of glomerular proteinuria: a commentary." J Am Soc Nephrol.15(5), pp 1354-57, 2004.

[11] Hostetter T. "Prevention of end-stage renal disease due to type 2 diabetes", N Engl J Med 345, pp 910-911,2001.

[12] Maduka Ignatius C, Neboh Emeka .E, Kwubiri Uchenna N, "The Prevalence of Diabetic Nephropathy in Diabetic Patients", European Journal of Scientific Research,Volume 26, Issue 2,pp 225-259, 2009.

[13] Nader Antonios, MD and Scott Silliman, MD, "Diabetes Mellitus and Stroke", Northeast Florida Medicine Journal, pp 17-22, 2005.

[14] Mahfouf M, Abbod MF & Linkens DA, "A survey of fuzzy logic monitoring and control utilization in medicine", Artificial Intelligence in Medicine 21, pp 27-42, 2001.

[15]Uncu, U. Koklukaya, E. Gencsoy, A. Annadurdiyew, O, "A fuzzy rule-base model for classification of spirometric FVC graphs in chronical obstructive pulmonary diseases", Volume 4, pp 3866-3869, 2001.

[16] Novruz Allahverdi Ismail Saritas Ilker Ali Ozkan Mustafa Argindogan, "An Example of Determination of

Medicine Dose in the Treatment by Fuzzy Method**", International** Conference on Computer Systems and Technologies - CompSysTech' 2006.

[17] Adam E. Gawedal, Alfred A. Jacobs, Michael E. Brierl, "Fuzzy Rule-based Approach to Automatic Drug Dosing in Renal Failure", The IEEE International Conference on Fuzzy Systems, 2003.

[18] H.F Kwokab, D.A Linkensa, M Mahfoufa, G.H Millsb, "Rule-base derivation for intensive care ventilator control using ANFIS", Artificial Intelligence in medicine, Volume 29, Issue 3, pp 185-201,2003.

[19] Feng Liu, Chai Quek, Geok See Ng, " A Novel Generic Hebbian Ordering – Based Fuzzy Rule Base Reduction Approach to Mamdani Nuero-Fuzzy System", Neural Computation, Volume 19, Issue 6, pp 1656-1680, 2007.

E.RamaDevi: Mrs. E. RamaDevi received MPhil degree in Computer Science from Bharathiar University, Coimbatore. Currently She is an Assistant Professor in Computer Science at NGM College, Pollachi, India. She has got 5 years of research experience. Her research interest includes areas like Data Mining, Knowledgebase system, Intelligent and Control systems and Fuzzy Logic. Presented papers in National and International conferences. Published 2 research papers and has more than 15 years of teaching experience.

Dr. N. Nagaveni: Dr. N. Nagaveni received PhD from Bharathiar University, Coimbatore, India. She is an Assistant Professor in Mathematics at Coimbatore Institute of Technology since 1992. She has got 14 years of research experience and 21 years of teaching experience. Her research interest includes cryptography, cloud computing, data mining, privacy issues and web mining. Published more than 30 research papers in different journals. Presented papers both in National and International conferences. Member in ISPE, Indian Science Congress and Kerala Mathematical Association.