Widening Scope of Principal Component Analysis: A Fuzzy Approach

D.S. Kalana Mendis, Asoka S. Karunananda, U Samarathunga

Department of Information Technology, Advanced Technological Institute,
Labuduwa, Sri Lanka
Faculty of Information Technology, University of Moratuwa,
Sri Lanka
Gampaha Wickramarachi Ayurveda Institute,
University of Kelaniya, Sri Lanka

E-Mail: kalanaatil@mail.com

Key words: Fuzzy statistics and data analysis; principal component analysis; knowledge modeling; Ayurveda

Abstract

Knowledge modeling is concerned with abstract model mapping using real world domains. Further all knowledge is tacit or rooted in tacit domains. Abstracting is mainly concerned with classification of such knowledge. In this issue statistical techniques can be issued and mainly concerned with multivariate statistical techniques. However using principal component analysis (PCA) as a multivariate technique makes problematic situation due to inability of classifying knowledge. Analysis using PCA is limited up to principal components extracted by PCA. Therefore existing algorithm of PCA to be addressed for modeling knowledge should be modified. This paper presents a novel mechanism for modifying PCA algorithm to address the problem in concerned using Fuzzy Principal component Analysis (FPCA). Here principal components have been used to define intervals for membership function. By doing so, knowledge classification is done effectively by constructing fuzzy memberships functions integrated with PCA. The experimental results using Ayurvedic medicine show that our approach is very promising.

1. Introduction

The existing system of tacit knowledge modeling in Ayurvedic medicine is not consistent [6, 5]. Experts use questionnaires that lead several problems like dependencies among the questions in the questionnaire and analysis of the constituent type. Although Principal Component Analysis (PCA) provides an algorithm to extract principal components from tacit knowledge for removing dependencies among questions in the questionnaire it is meant to address problems such as classification about the underline knowledge.

The ability to model tacit knowledge and represent modeled knowledge is critical to the success of experts, knowledge engineers and general users. It is a problem to experts to take decisions using the modeled tact knowledge as guidelines [9, 13, 15, 22, 25, 26]. It allows decision makers to take the changes into consideration and even take advantage of the changes when they make decisions. Knowing an effective knowledge-modeling algorithm in advance not only allows an expert to provide new guidelines and knowledge engineers to satisfy the requirement of its customers but also allows it to design corrective actions to stop inconsistencies or reduce uncertain decisions.

Existing principal component analysis as a multivariate technique, aimed at producing accurate models of the real world in an effective manner. But all real world domains are considered as tacit or rooted in tacit domains [14, 31, 32]. These are normally represented as questionnaires for human users to better understand the tacit problem domains and for modeling. However, regardless of how accurate a model can predict, it can only model based on the selected features and cannot lead to any knowledge classification because it will no longer be valid otherwise.

In this paper, we study the limitation of modeling tacit knowledge in the context of using principal component analysis (PCA) as a multivariate

technique for classification of knowledge. The problem of PCA above is defined as follows [7]. Let *S* be the set of all questions in the questionnaire and *P* be the set of all extracted principle components. Further, $S = \{S_1, S_2, ..S_{m-1}, S_m\}$, Then I^{th} principle Component is defined as

$$PC_i = a_{1i}S_1 + a_{2i}S_2 + .. + a_{m-1,i}S_{m-1} + a_{m,i}S_m$$

.For n number of extracted principal components, following computation is concluded.

$$X = \sum_{j=1}^{n} PC_j$$

$$\therefore X = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}S_i \ .$$

Many researchers have used it extensively and many algorithms have been proposed in PCA based multivariate statistics literature (e.g., [8]).

This problem definition is not developed for classifying knowledge in domains with tacit knowledge. Knowledge modeling without taking the classifications into consideration can result in severe degradation of performance of tacit knowledge. Since it is common that we need to model tacit knowledge based on a questionnaire, the modeling of scope in PCA is an important problem.

This paper presents the problem definition effectively. Furthermore, we generalize the work so that fuzzy PCA integrated system (FPCA) can be used for tackling this problem. Given principle component extracted from PCA, we propose to use fuzzy sets and fuzzy variable to model the knowledge mapped on a questionnaire. Furthermore, we propose to build a fuzzy membership functions to classify knowledge generated as output of PCA. The fuzzy membership functions can then be converted to a set of fuzzy rules.

These fuzzy rules are called *fuzzy multivariate -rules* because they are rules about multivariate technique. The *fuzzy multivariate -rules* can be used for human users to classify and for reasoning of tacit knowledge modeling.

The rest of this paper is organized as follows. In Section 2, we present the problem definition of principal component analysis in classification of knowledge. In Section 3, we describe an approach for classifying principal components.
To evaluate the performance of our approach, we applied it to Ayurvedic domain. The experimental results are provided in Section 4. Testing is described in section 5. In Section 6, we discuss the related work in the literature. Finally, we conclude this paper with a summary in Section 7.

2. The Problem Definition

Let S be the set of all questions in the questionnaire and P be the set of all extracted principal components.

Further, $P = \{PC_1, PC_2, ..PC_{n-1}, PC_n\}$

$$S = \{S_1, S_2, ..S_{m-1}, S_m\}$$

$$PC_i = a_{1i}S_1 + a_{2i}S_2 + .. + a_{m-1,i}S_{m-1} + a_{m,i}S_m \quad (1)$$

Let *M* be the principal components Matrix for filtered tacit knowledge.

$$M = \begin{bmatrix} a_{11}a_{12}..a_{1n} \\ a_{21}a_{22}..a_{2n} \\ . \\ . \\ a_{m1}a_{m2}..a_{mn} \end{bmatrix} \quad (2)$$

$$\therefore PC_1 = a_{11}S_1 + a_{21}S_2 + .. + a_{m1}S_m \quad (3)$$

$$PC_2 = a_{12}S_1 + a_{22}S_2 + .. + a_{m2}S_m \quad (4)$$

$$PC_{n-1} = a_{1n-1}S_1 + a_{2n-1}S_2 + .. + a_{mn-1}S_m \quad (5)$$

$$PC_n = a_{1n}S_1 + a_{2n}S_2 + .. + a_{mn}S_m \quad (6)$$

For n number of extracted principal components, following computation is concluded.

$$X = \sum_{j=1}^{n} PC_j \quad (7)$$

$$\therefore X = \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij} S_i \quad (8)$$

Further classification of knowledge in the questionnaire cannot be done using above described n number of extracted principle components.

Definition 1. The first principal component of the complex of sample values of the responses $X_1,\ldots X_p$ is the linear compound

$$Y_1 = a_{11}X_1 + \ldots + a_{p1}X_p$$

whose coefficients a_{i1} are the elements of the characteristic vector associated with the greatest characteristic root l_1 of the sample covariance matrix of the responses. The a_{i1} are unique up to multiplication by a scale factor, and if they are scaled so that $a'_1 a_1 = 1$, the characteristic root l_1 is interpretable as the sample variance of Y_1.

Example 1. Following table shows component matrix evaluated using SPSS

Component Matrix

	Component	
	1	2
VAR1	.988	.150
VAR2	.287	.958
VAR3	-.875	.483

Extraction Method: Principal Component Analysis.

2 components extracted.

The first principal component accounts for the three dimensions as follows:

$$Y_1 = 0.988(VAR1) + 0.287(VAR2) - 0.875(VAR3)$$

Definition 2. The j^{th} principal component of the sample of p-variate observations is the linear compound

$$Y_j = a_{1j}X_1 + \ldots + a_{pi}X_p$$

whose coefficients are the elements of the characteristic vector of the sample covariance matrix S corresponding to the j[th] largest characteristic rot l_j. If $l_i \neq l_j$, the coefficients of the i[th] and J[th] components are necessary orthogonal; if $l_i = l_j$, the elements can be chosen to be orthogonal, although an infinity number of such orthogonal vectors exists. The sample variance of the j^{th} component is l_j, and the total system variance is thus

$$l_1 + \ldots + l_p = trS$$

The importance of the j^{th} component in a more parsimonious description of the system is measured by

$$\frac{l_j}{trS}$$

3. Classifying Principal components

In the exciting system, tacit knowledge modeling is not consistent [12]. In general, questionnaires are used to capture tacit knowledge, but the questions in such questionnaires can have various dependencies without notice. Although Principal Component Analysis (PCA) provides algorithm to extract principal components from tacit knowledge by removing dependencies among questions in the questionnaire, it is considered to address problems such as classification about the underline knowledge

Fuzzy Logic

Fuzzy logic handles situations, where conclusions do not fall into one extreme. As compared with classical logic, fuzzy logic can handle real world problems, which deal with more than two truth-values. In fuzzy logic, everything is a matter of degree. Therefore fuzzy logic can be used to make decisions in domains with tacit knowledge [13]. Individual classification in Ayurveda is a classic example, where the decision has more than one possible truth-value.

In our research we have used fuzzy logic for addressing the vagueness involved in tacit knowledge. For example, vagueness (dependencies & inconsistencies) involved in Ayurvedic classification of individuals has been manipulated using fuzzy logic.

3.1.1 What is a Fuzzy Set?

A fuzzy set can be simply defined as a set with fuzzy boundaries. Let X be the universe of discourse and its elements be denoted as *x*. Fuzzy set *A* of universe *X* is defined by function $\mu_A(x)$ called membership function of set *A*.

$\mu_A(x): X \longrightarrow [0.1]$

This degree, a value between 0 and 1, represents the degree of membership, also called membership value, of element *x* in set *A*.

3.1.2 Fuzzy Rules

A fuzzy rule can be defined as a conditional statement in the form:

IF *X*= A THEN *Y*= B

Where *X* and *Y* are linguistic variables; and *A* and *B* are linguistic values determined by fuzzy sets on the universe of discourses *X* and *Y*, respectively.

Building algorithm for Fuzzy Principle Component Analysis (FPCA)

Algorithm for FPCA has been designed by extending the existing mechanism of principle components analysis (PCA). There are 2 no. of stages used for FPCA in addition to PCA, which are described as follows.

3.2.1 Generatingmembership functions(fine–tuning analysis)

Let LS be the Likert scale [14], then

$$LS = [L,..,U] \quad (9)$$

X_L and X_U values are derived from results of the filtered tacit knowledge. It is computed as given below.

$$\therefore X_L = L \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}$$

$$\therefore X_U = \mathbf{U} \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij} \quad (11)$$

Let A be fuzzy set defined on a fuzzy concept using the interval of $[X_L,.., X_U]$. Then membership function is as follows.

$$A(X) = \begin{cases} 0 & X <= X_L \\ (X-X_L)/(X_U-X_L) & X_L < X < X_U \\ 1 & X >= X_U \end{cases} \quad (12)$$

3.4. Fuzzy rule base (reasoning)

Fuzzy rules can be constructed as follows,

Rule 1: If $X <= X_L$ then A (X)=0 %

Rule 2: If $X_L < X < X_U$ then $(X-X_L)/(X_U-X_L)$ %

Rule 3: If $X >= X_U$ then A (X)=100 %

Adding dynamically, in order to operate the reasoning process for answers given by the fuzzy rules, fuzzy rule base can be extended further.

4. Experimental results

Although Ayurvedic practitioners use questionnaire but those are faced with several problems like dependencies among the questions in the questionnaire and classification of constituent type [5,6]. We addressed these problems to solve them in following stages.

4.1 PCA for removing dependencies

It consisted of 72 questions to analyze *vata*, *pita* and *kapha*. We have done a pilot survey for 100 students for statistical modeling using the questionnaire. Principal component analysis has been used to remove dependencies in the questionnaire. Twenty five (25) principal components have being identified using SPSS [17] as shown in matrix given below. Here V1, V2..V24,

K1, k2..K24, P1, P2..P24 denotes question-numbering system in the questionnaire. (10)

$$M = \begin{array}{c} \\ V1 \\ V2 \\ . \\ . \\ V23 \\ V24 \\ K1 \\ K2 \\ . \end{array} \begin{array}{c} V= \\ \\ \\ \\ \\ \\ K= \\ \\ \\ \end{array} \begin{array}{c} \begin{array}{cccc} 1 & 2 & .. & 24 \quad 25 \end{array} \\ \begin{pmatrix} -0.228622 & 0.249362 & . & -0.073945 & 0.058179 \\ 0.08431 & 0.20654 & . & -0.097192 & -0.112795 \\ & & & & \\ & & & & \\ -0.645803 & 0.232312 & . & 0.0067 & -0.083959 \\ -0.222147 & -0.06453 & . & -0.073514 & 0.084404 \\ 0.012511 & -0.096332 & . & 0.141314 & 0.25113 \\ -0.005642 & 0.268145 & . & -0.179992 & 0.111715 \\ & & & & \end{pmatrix} \end{array}$$

$$V = \begin{pmatrix} -0.228622 & 0.249362 & . & -0.073945 & 0.058179 \\ 0.08431 & 0.20654 & . & -0.097192 & -0.112795 \\ & & & & \\ & & & & \\ -0.645803 & 0.232312 & . & 0.0067 & -0.083959 \\ -0.222147 & -0.06453 & . & -0.073514 & 0.084404 \end{pmatrix}$$

24*25

4.2 Fuzzy logic for classifying knowledge

Human constituents can be computed in to *vata, pita* and *kapha* in percentages as shown below. Membership functions for *vata, pita* and *kapha* have been constructed using fuzzy logic based on out puts of principal component analysis.

4.21 Membership function for classifying Vata constitution

Boundary values of membership function have been constructed using the output of the principal component analysis.

$$\because X_L = 1 \sum_{i=1}^{25}\sum_{j=1}^{24} a_{ji} = 8.510004 \quad (13)$$

$$\because X_U = 6 \sum_{i=1}^{25}\sum_{j=1}^{24} a_{ji} = 51.06002 \quad (14)$$

Here X_L denotes lower bound value at the minimum level of evaluation scale (Does not apply) in the questionnaire. X_U denotes upper bound value at the maximum level of evaluation scale (Applies most) in the questionnaire

$$V(X) = \begin{cases} 0 & X =< X_L \\ (X-X_L)/(X_U-X_L) & X_L < X < X_U \\ 1 & X => X_U \end{cases}$$

V(x) denotes membership function for classifying *vata* constitution. This has been constructed using Visual Basic

4.2.2 Constructing fuzzy rules

Fuzzy rules can be constructed using the design model for fuzzy rules generator described in section 3.

So following fuzzy rules can be illustrated for classing humeral constitutions in to *vata, kapha* and *pita* in term of percentage values.

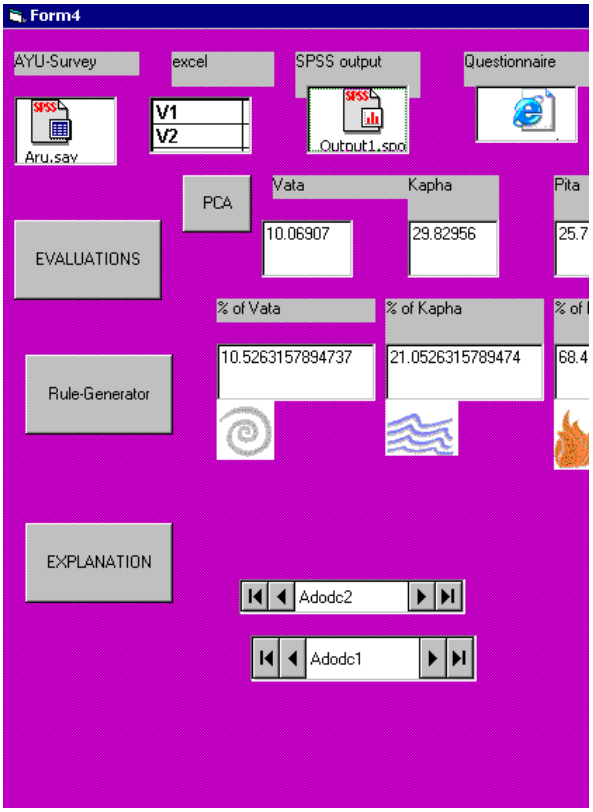

Fig.1. Analysis window

4.2.2.1 For *Vata* constitution

Rule 1: If $X <= X_L$ then V (X)=0 %

Rule 2: If $X_L < X < X_U$ then V (X)= $(X-X_L)/(X_U-X_L)$ %

Rule 3: If $X >= X_U$ then V (X)=100 %

5. Testing

The expert system developed using this approach was tested with a group of 30 persons of Ayurvedic experts and students (see Table 1).

id	vata	pitta	kapha	Expert_decision
14	25.71	20.71	53.57	KV
15	32.95	23.86	43.18	VP
16	39.88	23.81	36.31	VP
17	27.65	46.1	26.24	KP
22	25.69	29.36	44.95	KV
23	33.58	24.09	42.34	KV
24	25.71	34.28	40	KP
26	32.21	31.54	36.24	KV
27	22.51	29.8	47.68	KP
28	20.37	30.56	49.07	PK
35	30.6	35.52	33.88	PK
46	29.71	17.39	52.9	KV
47	41.07	10.71	48.21	KV
48	34.5	32.16	33.33	KV
49	23.46	28.57	47.96	PK
50	35.27	30.77	33.97	KV
52	42.36	36.11	21.53	VP
56	23.01	35.71	41.27	PK
65	47.94	19.86	32.19	KV
66	14.03	35.96	50	PK
67	19.15	36.88	43.97	PK
68	22.46	25.36	52.17	PK
69	40.47	26.78	32.74	PK
70	30.28	29.58	40.14	KV
71	12.71	44.92	42.37	PK
72	11.18	40	48.82	PK
73	11.24	40.24	48.52	PK
74	23.44	26.9	49.66	PK
93	17.09	36.75	46.15	KV
94	33.09	30.15	36.76	KV

Table 1: System Testing: Expert Vs. System

The evaluation was conducted to see that the answers generated by the system matches with the identification by Ayurvedic experts and the students. Further, the system's ability to fine-tune the answers was also tested. It has been investigated that 23 (77%) of conclusions matches with the system and expert (see Table 2), which leads to determine the accuracy of the system.

id	vata	pitta	kapha	Expert_ decision	conclusion
14	25.71	20.71	53.57	KV	matched
23	33.58	24.09	42.34	KV	matched
24	25.71	34.28	40	KP	matched
26	32.21	31.54	36.24	KV	matched
27	22.51	29.8	47.68	KP	matched
28	20.37	30.56	49.07	PK	matched
35	30.6	35.52	33.88	PK	matched
46	29.71	17.39	52.9	KV	matched
47	41.07	10.71	48.21	KV	matched
48	34.5	32.16	33.33	KV	matched
49	23.46	28.57	47.96	PK	matched
50	35.27	30.77	33.97	KV	matched
56	23.01	35.71	41.27	PK	matched
65	47.94	19.86	32.19	KV	matched
66	14.03	35.96	50	PK	matched
67	19.15	36.88	43.97	PK	matched
68	22.46	25.36	52.17	PK	matched
70	30.28	29.58	40.14	KV	matched
71	12.71	44.92	42.37	PK	matched
72	11.18	40	48.82	PK	matched
73	11.24	40.24	48.52	PK	matched
74	23.44	26.9	49.66	PK	matched
94	33.09	30.15	36.76	KV	matched

Table 2: Comparison of Conclusions: Expert Vs. System

The system facilitated to derive constituents types in percentages while Ayurvedic experts obtain only the constituent type. As recommendation given by the Ayurvedic experts, determining constituent's types in percentages is an important criterion for prescribing drugs for a disease. Further, our system provides as an option to find out possible diseases. In general, the system can be used as a self-assessment for finding constituents. According to Ayurvedic medicine, regiments can be done easily by knowing the constituent type. The human constituents can be computed as a combination. So it would help to find the effectiveness of minimum type in a diagnosis.

5. Discussion

Present mechanism in PCA [7] is not valid in knowledge modeling because it can reduce dependencies but not further classification of knowledge. However by integrating fuzzy logic module, present algorithm of PCA is extended for classification of knowledge.

However FPCA gives novel approach in classifications of Principal Components in principle components analysis.

Construction of Fuzzy membership functions using PCA gives an extensive advantage for tacit knowledge modeling. Here we use a case study of classification of human constitutions in Ayurvedic medicine and gained significant results to reduce existing inconsistencies of diagnosis of human constitutions. In the first instance tacit knowledge involved for classification of human constitutions has been mapped in to a questionnaire. PCA has been used to reduce dependencies among questions in the questionnaire. But existing PCA algorithm could be used up to this stage. However we needed to classify human constitutions in to 3 categories. Existing PCA algorithms has been integrated with fuzzy logic module to classify constitutions in to vata, pita and kapha by giving percentages.

6. Conclusion

In this paper, we present the definition of the problem of classification of principal components in principle component analysis. The proposed approach allows a fuzzy principle component analysis to be used for tackling this problem. In tacit knowledge modeling, it is required to (i) map the tacit knowledge into a questionnaire and reduce dependencies used; (ii) classify the knowledge in the questionnaire using extracted principal components; and (iii) reasoning for knowledge classification using fuzzy rules.
To model tacit knowledge, to do classification and reasoning in a comprehensible fashion, we propose to use fuzzy membership functions. The fuzzy membership functions can then be converted to fuzzy rules. These fuzzy rules are called fuzzy PCA-rules because they are rules about principal components. Furthermore, the discovered fuzzy PCA-rules can be used to suggest decisions using modeled tacit knowledge. To evaluate the performance of our approach, we make use of a domain of tacit knowledge of which is a set of rules collected in analysis of human constitutions in Ayurvedic domain. Tacit knowledge in human constitutions analysis of Ayurvedic domain is mapped in to a questionnaire. Fuzzy membership functions are then constructed in the classification of tacit knowledge modeling and fuzzy rules for reasoning the output of classification of human constitutions analysis in Ayurvedic domain. The experimental results show that our approach is very promising.

7. References

[1] Karunananda A.S., "How To Do Research " Author Publication, 2000.

[2] Jonson L., "Expert system Architectures", Kopan Page Limited, 1988.

[3] George J., "Fuzzy sets and Fuzzy Logic", Prentice Hall of India, 1997.

[4] Noak V., "Discovering the world with Fuzzy logic", A Springer – Verlag Company, 2000, PP. 3 – 50.

[5] Dubey G.P.,"The Physiological concepts in Indian medicine", Science and Philosophy of Indian medicine, 1978.

[6] Tripathi S.N., "Clinical Diagnosis", Science and Philosophy of Indian medicine, 1978.

[7] Chatfied C.," Introduction to Multivariate Analysis", Chapman and Hall, 1996.

[8] Morrison D.F., "Multivariate Statistical Methods", Mc Graw-Hill, Inc, 1990, PP. 312 - 357.

[9] Ross D.,"Developing Knowledge-Based Client Relationship", Butterworth Heinemann, 2001.

[10] Karunananda A.S.,"Pinki – A motivation for Artificial Intelligence projects",Author publication, 2001.

[11] Mendis D.S.K, Karunananda A.S, Samaratunga U., "Reasoning with Uncertainty", Health Informatics Conference, Colombo, 2003.

[12] Mendis D.S.K, Karunananda A.S, Samaratunga U., "Tacit knowledge modeling", Sri Lanka Association for Artificial Intelligence Annual Session and Annual General Meeting, Colombo, 2003.

[13] Bellman R.E, Zadeh L.A., "Decision making in a fuzzy environment", Management Science 17, 1970, pp. 141–164.

[14] Richards D, Bush P., "Measuring, Formalizing and Modeling Tacit Knowledge" IEEE/Web Intelligence Conference (WI-2003) Bejing, 2003.

[15] M. Dzbor., "Explication of Design Requirements Through Reflection on Solutions", In Proc. of 4th IEEE Conference on Knowledge-based Intelligent Engineering Systems. Brighton, UK, 2000,pp. 141-144.

[16] Dave W., Flex reference guide, LPA, U.K, 2000.

[17] Matei Ciobanu Morogan., SPSS for windows, release 8.0.0, DSV KTH/ SU, Sweden, 1997.

[18] Evangelos P., "Mastering Visual Basic 6", BPP Publications, New Delhi

[19] George J.K, Yuan B., Fuzzy sets and Fuzzy logic, prentice hall of India, 1995, pp. 280–300.

[20] XpertRule Knowledge Builder, www.attar.com.

[21] Peter J., Introduction to expert systems, Pearson education Asia, 1995, pp. 182–199.

[22] Karl B, Klaus P, A,Yoichi H, Thomas E. Rothenfluh, Harald L., "Knowledge acquisition in the fuzzy knowledge representation framework of a medical consultation system" International Journal of Artificial Intelligence in medicine, 2004,pp. 1–26.

[23] Rajeev K., "Integrating knowledge bases in expert system shells: an open system approach", International Journal of computer application in technology, 1995, pp. 78–89.

[24] Coppin G, Skrzyniarz A., "Human–centered processes : Individual and distributed decision support", IEEE Intelligent systems, 2003, pages 27–33.

[25] Hayashi Y, Tazaki E, Yoshida K, Dey P., "Medical diagnosis using simplified multi-dimensional fuzzy reasoning", In: Proceedings of IEEE 1988 International Conference on Systems, Man and Cybernetics, Beijing, China, 1988. pp. 58–62.

[26] Kolousek G., ″The system architecture of an integrated medical consultation system and its implementation based on fuzzy technology", Doctoral thesis, Technical University of Vienna, Austria, 1997.

[27] Binaghi E., De Giorgi O., Maggi G., Motta T., Rampini A., "A knowledge acquisition tool for computer assisted diagnosis of postmenopausal osteoporosis using a fuzzy expert system shell", Computers and Biomedical Research, 1993,pp.498-516.

[28] Zadeh L. A., Fuzzy Logic, Computer IEEE, 1998, pp. 83-93.

[29] Senkar J., ″Network and tacit knowledge in innovation″, Economics et societies, 1995, pp. 99-118.

[30] Zimmerman H.J., Zimmerman Fuzzy set theory and its applications, Kluwer–Nijhoff, Boston, 1995.

[31] Ravi V, Reddy P. J. and Zimmermann H.J., Pattern classification with principal component analysis and fuzzy rule bases European Journal of Operational Research, November 2000, pp. 526-533

[32] Paolo G, Henk A. L. Kiers., Principal Component Analysis of symmetric fuzzy data, Computational Statistics & Data Analysis, Volume 45, Issue 3, 10, April 2004, Pages 519-548