Design of Pattern Recognition System for the Diagnosis of Gonorrhea Disease

Umoh, U. A., Umoh, A. A., James, G. G., Oton, U. U., Udoudo, J. J., B.Eng.

ABSTRACT: Sexually transmitted diseases (STDs) share common symptoms and can be classified as confusable disease, as such become difficult for physicians to correctly diagnose them. This work develops a pattern recognition system for the diagnosis of gonorrhea disease using genetic algorithm. Data on gonorrhea symptoms are collected and used in the development of the knowledge base. We classify the membership grade of the symptoms based on the mean of maxima and the derived membership function. The system accepts symptoms as input and provides the degree of membership of each symptom in any gonorrhea symptoms sets. We develop our system using PHP programming tool as back end and Java as front end platform. We explore Ms Access for the design of our database. The model helps the physicians to identify gonorrhea disease by its symptoms and provides solid basis for possibly determination of the ailment exactly if there are all symptoms.

Keyword: Diagnosis, Pattern Recognition, Sexually Transmitted Diseases, Symptoms,

1. INTRODUCTION

Expert system is a current and major segment of artificial intelligent system that is involved in solving complex professional problems that requires human experts. It is the software systems developed using different techniques of artificial intelligent that can act parallel to the “human” experts. The main role is consultative. These are intelligent information systems that use more than 2000 different rules and are capable to explain the decision. Considering the exiting experts systems, many medical expert systems assist the physicians in making diagnosing which may shorten the time spent in making correct diagnosing errors. At the same time, physician may obtain the information on the symptom of each of the diseases and pathologic syndromes contained therein (Rezers et al, 1984). In the wise, pattern recognition is considered as the knowledge representation that coordinates the diagnosis. In machine learning, pattern recognition is the assignment of some sort of output value (or Label) to a given input value (or instance), according to some specific algorithm (Girrantano, 2005). Medical diagnosis is the identification of a disease by its symptoms, which provides a solid basis for the treatment and prognosis of the patient (Critchely, 1986). The process of diagnosis begins when the patient consults the doctor and presents a set of complaints (the symptoms). The doctor obtains information from the patient about his symptoms, his previous state of health, living conditions, and so forth.

Umoh, A. A., Ph.D., *academic staff, department of Physics, University of Uyo, Nigeria, +23408023206450, draustineumoh@yahoo.com.*
Umoh, U. A., Ph.D, *academic staff, department of Computer Science, University of Uyo, Nigeria, +2348120064911, uduakaustynumoh@yahoo.com.*
James, G. G., M.Sc., *academic staff, department of Computer Science, College of Arts and Science, Akwa Ibom State.*
Oton, U. U., B.Sc. and Udoudo, J. J., B.Eng., *academic staff, department of Computer Engineering, Uyo City Polytechnic, Akwa Ibom State.*

Rather than considering the myriad disease that can affect the patient, the physician narrows down the possibilities likely to account for the apparent symptoms, making a loss of those conditions that could account for what is wrong with the patient (condo, 1987). According to Kent M. and De Graeff, (date), sexual transmitted disease (STDs) are contagious disease that affect the reproductive systems of both male and the female. These classes of disease are transmitted during sexual activity, and their frequency of occurrence in the society is in the increase. Gonorrhea is one example of STD. Gonorrhea, commonly called “clap”, is caused by the bacterium gonococcus, or Neisseria gonorrhea. In male, the disease causes inflammation of the urethra in some cases, accompanied by painful urination and frequently discharge of pus. In female, the condition is usually asymptomatic, and therefore many women may be unsuspecting carriers of the disease (katchadou Rian, 1985). Advanced stages of gonorrhea in female may infect the uterus and the urine tubes. A pregnancy woman with untreated gonorrhea may transmit the disease to the eyes of her new born during its passage through the birth canal, possibly causing blindness (Lager-krantz, 1986). Sometimes the disease may not be properly treated and can linger for long to cause more harms in the reproductive system in the future such as weak erection, quick ejaculations as well as erectile disfunctioning. Medical diagnosis is defined as the identification of a disease by its symptoms, which provides a solid basis for the treatment and prognosis of the patient (Critchely, 1986). However, it has two distinct definitions. The first definition is “the recognition of a disease or condition by its symptoms” while the second definition is “the analysis of the underlying physiological/ biochemical cause(s) of a disease or condition”. The doctor obtains information from the patient about his symptoms, his previous state of health, living conditions, and so forth. However, rather than considering the myriad disease that could affect the patient, the physician narrows down the possibilities likely to account for the apparent symptoms, making a lost of those conditions that could account for what is wrong with the patient (condo, 1987). Using gonorrhea disease as ac case study, the manual method of diagnosis

does not guarantee precision in the diagnosis of gonorrhea. It is by trial and error, for instance, apart from chronic gonorrhea; there are other diseases like diabetes that cause weak erection or erectile disfunction. The potential of Artificial Intelligent (A1) technique in medicine has been expressed by a number of researchers to include: provision of a laboratory for the examination, organization, representation, and cataloging of medical knowledge, produces new tools to support medical decision making, training and research, integrates activities in medical, computer, cognitive and other sciences, offers a content-rich discipline for future scientific medical specialist, etc. (Hong, 1981), (Wan et al, 2009),. Many expert systems have been developed for the purpose of enhancing health-care and providing better health facilities. As expressed by many studies (Mahabala et al, 1992; maniekam etal, 1999; Ruseckaite et al, 1999; Bourles et al, 1999), expert systems are developed to assist users (particularly doctors and patients) and provide early diagnosis and prediction to prevent serious illness. However, the field of A1 has provided several approaches to defining systems that may be useful for medical diagnosis. (Mrein, et al, 2008) (Duda, et al, 2007). This highlights the need to design a good and effective pattern recognition system for diagnosis of gonorrhea diseases using genetic algorithm. In order to achieve our objective, a study of a knowledge based system for effective pattern recognition system for diagnosis of gonorrhea diseases is carried out. We classify the membership grade of the symptoms based on the mean of maxima and the derived membership function. The system accepts symptoms as input and provides the degree of membership of each symptom in any gonorrhea symptoms sets. Details on gonorrhea symptoms are collected and used in the development of the knowledgebase. We develop our system using PHP programming tool as back end and Java as front end platform. We explore Ms Access for the design of database. The model helps to identify gonorrhea disease by its symptoms and provides solid basis for possibly determination of the ailment exactly if there are all symptoms. The obtained simulation and implementation pattern recognition system results are investigated and discussed. In section 2 presents the research methodology is presented. Section 3 presents the model experiment while in Section 4 results of findings are discussed. Finally in Section 5, some recommendations are made and conclusion is drawn.

Research Methodology

The architecture of the pattern recognition system (PRS) for diagnosis of gonorrhea disease is shown in Figure 1. This system consists of three major components; Knowledge Base, Inference Engine and Pattern classifier.

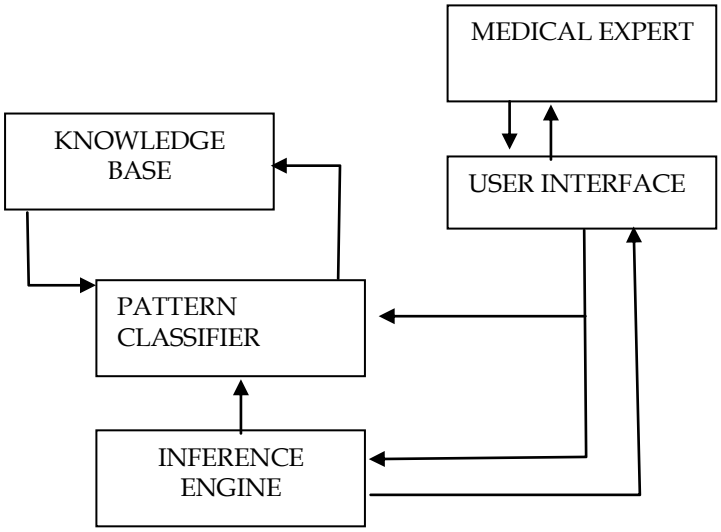

Fig. 2: Architecture of Pattern Classification System for Diagnosis of Gonorrhea

(i) Knowledge Base - keeps tracks of relevant knowledge required for the diagnosis of gonorrhea. The user through the user interface supplies fact and information to the expert system or receives expert advice from the system. The knowledge base contains knowledge about the problem domain and database as its component. The set of frame that holds knowledge about gonorrhea is presented in Table 1.

Table 1: Frame for holding patient symptoms

SLOT	DATA TYPE	VALUE
di	Float	*
di + 1	Float	**
. . .	. . .	. . .
Dn	Float	***

Table 2 contains disease as slot ranges from di, di+l, …, dn, where dn is the slots at infinity and '*' indicates any float value between o and 1. The slots di – dn, where n = 6 for acute gonorrhea and n = 5 for chronic gonorrhea. In the case of acute gonorrhea, the knowledge base holds symptoms about gonorrhea in the following categories; d_1 holds value for testis pains, d_2 holds value for watery sperm, d_3 holds value for anorexia, d_4 holds value for severe headache, d_5 holds value for painful urination, d_6 holds value for purse. Similarity for chronic gonorrhea involve; d_1 holds value for virginal discharge, d_2 holds value for watery sperm, d_3 holds

value for weak erection, d_4 holds value for painful menstruation, d_5 holds value for erectile disfunction

(ii) Inference Engine - the process of drawing conclusions from existing data is called inference. The PRS inference uses the knowledge in the knowledge base to draw conclusions and decide whether the patient is infected with gonococcus bacterium or not. The system applies a probabilities output, such that the probabilistic pattern – recognition algorithm is effectively incorporated into a larger machine learning tasks, in a way that partially or completely eliminate the problems or error propagation in the diagnostic process of sexual transmitted disease.The pattern recognition system algorithm classifies symptoms into; Number of classes, c = 2, Features vector dimension, d = 7, Classification coefficient, m = 2, Termination criteria, c = 0.01

The PRS algorithm for the diagnosis of gonorrhea diseases is designed using the formula:
Randomly initialized matrix

$$u = uij \quad (1)$$

where i = 1, 2, ……, N and j = 1, 2, ……., M
Computing the classification criterion cj,

$$Cj = \frac{\sum_{i=1}^{N} U^2 ij.xi}{\sum_{i=1}^{N} U^2 ij} \quad (2)$$

and

$$uij = \frac{1}{\sum_{i=1}^{N}((//xi - ci//)^2 /(//xi - cx//))} \quad (3)$$

(iii) Pattern Classifier - pattern recognition has to do with the assignment of some sort of output value (or Label) to the series of input value (for instance), according to some specific algorithm. We use the classification algorithm to provide some reasonable answer for all possible inputs and to do “fuzzy” matching of inputs. The pattern matching algorithms look for exact matches in the input with pre existing patterns. We categorize the features (also known as norminal, i.e, consisting of one of a set of unordered items, such as a gender of “male” or “female”, or a blood type of “A”, “B” “AB”, “O”), ordinal (consisting of one of a set of ordered items, e.g “large”, “medium”, or “small”), integer- valued (e.g a count of the number of occurrences of a particular word), or real-valued (e.g, a measurement of blood pressure). The recognizer algorithm is probabilistic in nature, in that they use statistical inference to find the best label for a given instance. Because of its probabilistic nature, it produces a probabilistic output of the instance as described by the labels in (4) and (5). Table 2 shows the endpoint of each membership function.

$$\rightarrow P(label / x, \theta) = f(x,j\ \theta) \quad (4)$$

Such that

$$P(label/x) = \int p(label/x, \theta)\ p(\theta/D) f\theta \quad (5)$$

Table 2: Endpoint of each Membership Function

	A	B	C
Class 1 (no testis pain)		36.45	36.62
Class 2 (slight testis pain)	27.21	38.20	30.12
Class 3 (high testis pain)	38.22	39.62	31.01
Class 4 (very high testis pain	40.26	42.22	

The function for each class is derived using the formula:

$$bj = \frac{yi*si+yi+1*+(si:+1)/2+A+Y_{k-1}*(si-2+si-1)/(2+yk+s_{k-1})}{si+(si+si+1)/2+(A+s_{k-2})/2+s_{k-1}+s_{k-1}} \quad (6)$$

$$aj = \frac{bj - bj - yi}{1-\mu j\,(yi)} \quad (7)$$

$$aj = \frac{bj + yk - bj}{1-\mu j\,(yk)} \quad (8)$$

The values for each bj, aj and cj are classified as; ntp = no testis pain, stp = slight testis pain, htp = high testis pain, and vhtp = very high testis pain. The classification of the membership grade is based on the mean of maxima as shown in (9). The membership functions evaluations for the endpoints are presented in equations 10-13 respectively.

$$U = \sum_{1i-1}^{R} Ui / R \quad (9)$$

$$M_{ntp}(x) = \begin{cases} 1 & \\ \frac{38.5 - x}{0.5} & \text{If } x < 37.0 \\ 0 & \text{If } x > 38.5 \\ & \text{If } 37.0 < x < 38.5 \end{cases} \quad (10)$$

$$M_{stp}(x) = \begin{cases} 0 & \\ \frac{x - 37.2}{1.5} & \text{If } x < 37.2 \\ & \text{If } 37.2 \le x < 38.2 \\ \frac{40.2 - x}{1.5} & \text{If } 38.7 < x < 40.2 \\ 0 & \text{If } x \ge 40.2 \\ & \text{If } x < 38.2 \end{cases} \quad (11)$$

$$M_{htp}(x) = \begin{cases} \frac{x - 38.2}{1.4} & \text{If } 38.2 \le x < 39.6 \\ \frac{41.0 - x}{1.4} & \text{If } 39.6 < x < 41.0 \\ 0 & \text{If } x \ge 41.0 \end{cases} \quad (12)$$

$$M_{vhtp}(x) = \begin{cases} 0 & \text{If } x < 40.2 \\ \frac{x - 40.2}{2.0} & \text{If } 40.2 \le x < 42.2 \\ 1 & \text{If } x > 42.2 \end{cases} \quad (13)$$

RESULT

In the pattern recognition system implementation, figure 3 shows the graph of endpoint of membership function. Pains in Testis-the class set for testis pain take the values 0, 1 and 2 respectively. This implies that: no pains = 0, slight pains = 1, very high pains = 2. Painful Urination- the class set for the painful urination of the membership grade a based on the values 0, 1 and 2 respectively: This implies that: no pains = 0, slight pains = 1, very high pains= 2. Weak Erection- the class set for the weak erection of the membership grade is based on the values 0 and 1; no erection = 0, and erection = 1. Virginal Discharge- the class set for virginal discharge of the membership grade is based on the value 0 and 1; no virginal discharge = 0, and virginal discharge = 1. Presence of Purse - the class set for present of purse of the membership grade is based on the value 0 and 1; no purse = 0, and present of purse = 1. Table 3 shows the database for storing knowledge in the knowledge representation. Table 4 shows the database for entering doctor ID and password for access permit. Table 5 shows the file structure for entering patient's information and Table 6 shows the file structure for patient's symptoms. Figure 4 presents the PRS Class Diagram.

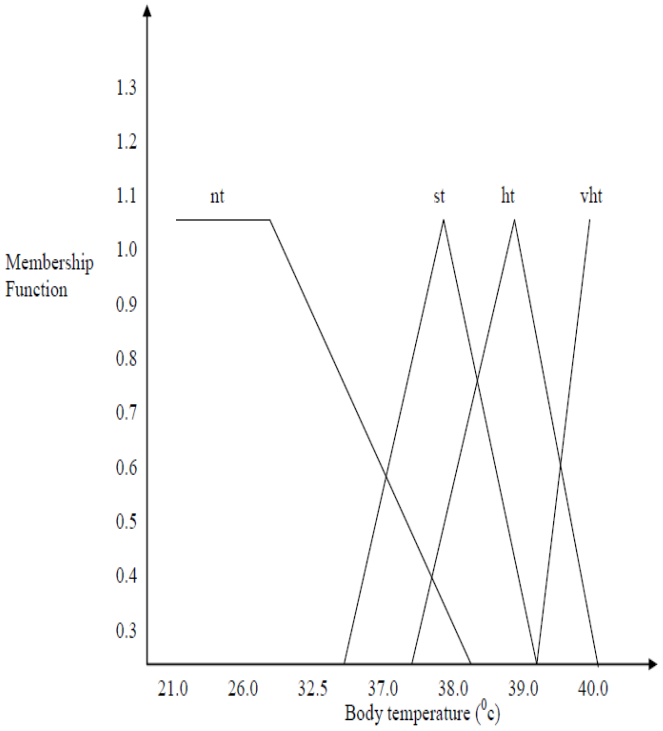

Fig. 3: Graph Showing Endpoint of Each Membership Function

Table 3 shows the database for storing knowledge in the knowledge representation.

Field Name	Data Type	Description	Length
Di	Float	Holds value for pains	8
di+1	Float	holds value for weak erection	8
Dn	Float	Holds value for virginal discharge	8

Table 4 shows the database for entering doctor ID and password for access permit

Field Name	Data Type	Description	Length
Doctor – ID	Alphanumeric	Doctor's identification No.	5
Password	Alphanumeric	Doctor's password	10

Table 6 shows the file structure for entering patient's information

Field Name	Data Type	Description	Length
Title	Text	Patient's Title	3
Name	Text	Patient's Name	30
Age	Text	Patient's Age	4
Sex	Text	Patient's Sex	2
Religion	Text	Patient's Religion	15
Patient ID	Text	Identification Number	

Table 6 shows the file structure for patient's symptoms

FIELD NAME	DATA TYPE	DESCRIPTION	LENGTH
Index	Integer	Store row number	5
Patient's Id	Text	Patient's ID number	5
Testis pains	Float	Holds value for testis pains	4
Abdominal pains	Float	Holds values for abdominal pains	4
Anorexia	Float	Holds value for anorexia	4
Headache	Float	Holds value for headache	4
Painful urine	Float	Holds value for painful urination	4
Purse	Float	Holds value for purse	4
Virg-discharge	Float	Hold value for virginal discharge	4
H20-sperm	Float	Holds value for watery sperm	4
Weak erection	Float	Holds value weak erection	4
Menstrual –pains	Float	Holds value for menstrual pains	4

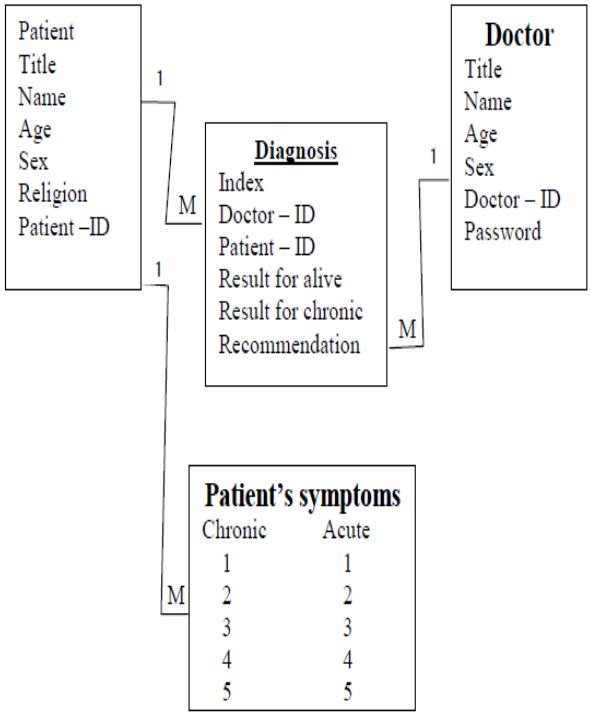

Fig. 4: PRS Class Diagram

Screen Shots Of The System

In this section, we present some of the screen shots of the system. Figure 5 shows the application's Login Page. This is the only entry point for the system by all users through the supply of identification password. Once the user authenticates the application successfully, the Splash screen is loaded within the browser as shown in Figure 6. Figure 7 shows the diagnosis screen, where patients' information is entered and submitted for diagnosis by the system.

Fig. 5: Login Page

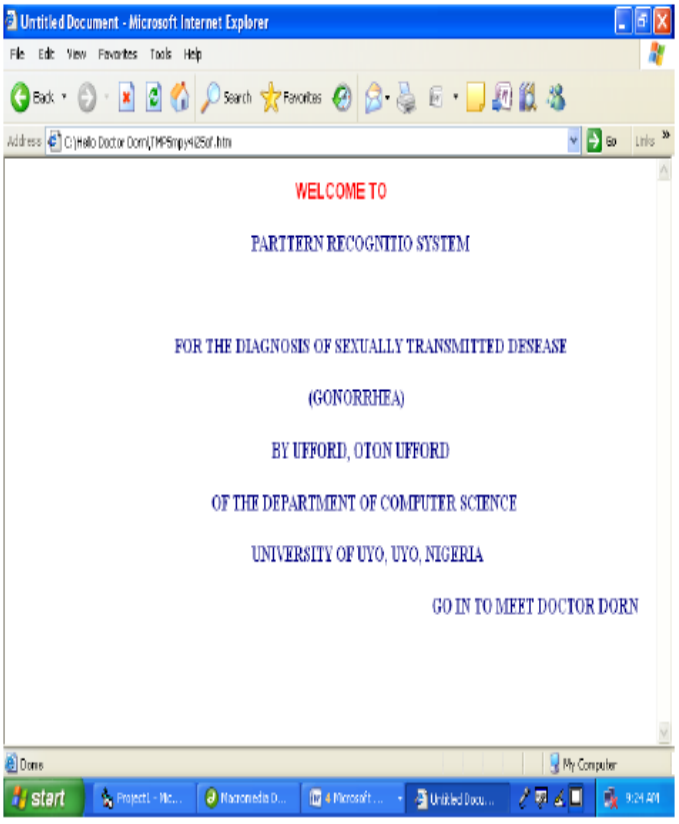

Fig. 6: Splash Screen

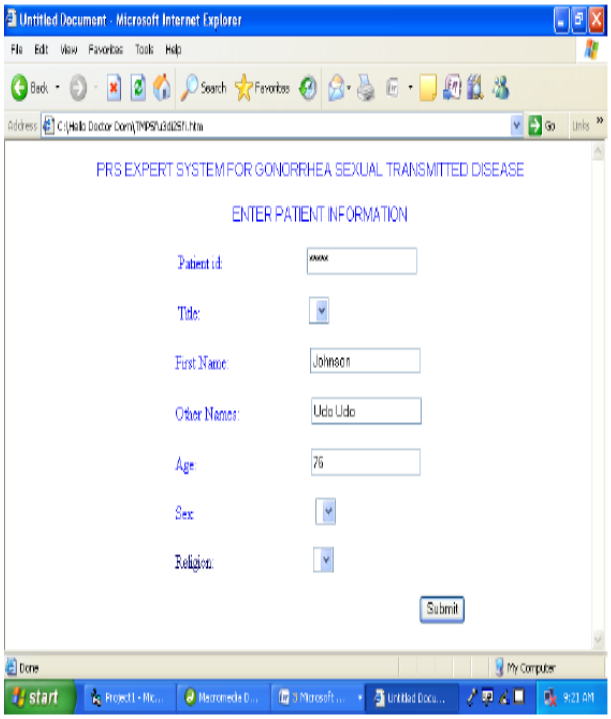

Fig. 7: Diagnosis Screen

Conclusion

With the rate of sexual transmitted diseases in our generation today, medical doctors find it difficult to handle the diagnosis of a given class of STD and as such, patients are being forwarded to the laboratory. Sometimes, the patient may be in a window states as such can prove such diagnosis wrong. This project has presented the application of a pattern recognition system for the diagnosis of gonorrhea with better performance, reliability and increase efficiency and availability. Our system, (PRS) serves as an important supportive tool for the medical doctors in diagnosing of sexual transmitted diseases (STDs) like gonorrhea. It will indeed assist the physicians to determine precisely whether or not a patient is infected with gonorrhea. The benefits of applying artificial intelligence techniques to medical diagnosis are numerous.

References

Adlessnig K. P. (1996): Fuzzy Set Theory in Medical Diagnosis. IEEE Transactions on systems, Man, Cybernetics, SMC – 16, No. 2,, pp.260 -265

AIJSCERT: Brochure, http://ftp.auscert.org.au/pub/auscertiauscert_brochure.ps.Z 1995 M. *Enterprise System Architectures,* CRC press, 2000.

A. Webb (2002): Statistical Pattern Recognition, 2nd edition, John Wiley, Hoboken, NJ.

Beckhard, *A.,* A model for Executive Management of 'Iranstormationat Change, Human Resource Strategies, London:Stage IOpen University, 1992

Castagnetto, J. M., Rawat, H. ,Schumann, S. Scollo, c., Veliath, D.: 1999 Professional PHP Programming, Wrox Press Ltd, UK

[Cern] No author or date information: Simple Digest Security Scheme, http://Www.w3.org/hypertextlWww/protocols/HTTP/digest_specitication.h tm1

Chapman, D. B. & E. D. Zwicky (1995). _Building Internet Firewalls_. O'Reilly & Associates (Sebastopol, *CA).* ISBN 1-56592-124-0. xxvi + 517. Index. See p. 352 ff.

George Becks, M. Dotoli, Diedrich Graf Keyserlingk, Jan Jantzen (2000): Fuzzy Clustering. A Versatile Means to Explore Medical Database, ESIT, Aachen Germany.

Girrantano Riley (2005): Expert Systems. The Development and Implementation of Expert System, McGraw Hill inc. USA.

James C. Bezdek, R. Krishnampuram, N. R. Pal (1999): Fuzzy Models and Algorithms for Pattern Recognition, Kluwer Academic Publishers, Dordercht.

James, C. Bezek (1981); Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, new York.

Jones, K. L. et al (1985) Dimensions of Human Sexuality. Dubuque: Win. C. Brown.

Katchadourian, A. A., and D. T. Lunde, (1985). Fundamentals of Human Sexuality, 4th ed. New York; Holt, Rinehort, and Winston.

Karkanis S. A. (1999): Integration of Neural Networks and Knowledge-Based System in Medicine.

L. A. Zadech (1994): Fuzzy Sets and their Application to Medical Diagnosis and Pattern Recognition.

Lagerkrantz, A. and T. A. Slotkin (1986). The “Stress” of being born scientific American.

McWhinney p. C., (2005): Health and Disease Problems of Definition.

Richard Robertson and Friendman J. H. (2008): A Fuzzy System for Helping Medical Diagnosis of Malformations of Cortical Development.