Expert System for Staging Breast Cancer

Adenike Adegoke-Elijah
Oduduwa University, Ipetumodu
Nigeria

Adewale Adisa
Obafemi University University
Teaching Hospital,
Ile-Ife, Nigeria

Chidiebere Raphael
Oduduwa University, Ipetumodu
Nigeria

ABSTRACT
The task of staging breast cancer is very essential in the management of the disease. This is because the stage of the disease determines the most appropriate treatment to be given to patients, and this consequently determines the rate of mortality due to the menace. Despite its importance, staging of breast cancer is cumbersome and requires experiential knowledge for it to be adequately done. New and inexperienced doctors therefore depend on consultant oncologists to carry out this task. This study therefore adopted a traditional rule-based approach to develop an expert system than could serve as a decision support system for inexperienced doctors in staging breast cancer. To achieve this aim, knowledge of the symptoms and the criteria for staging were elicited from the Association of American Cancer Society. The elicited knowledge was represented procedurally using IF-THEN rule. The Breast Cancer Staging System was later implemented using JAVA programming language, The developed system was evaluated for accuracy, reliability and usability using Mean Opinion Score; and the scores gotten were 4.12, 4.12 and 4.08 respectively. This study has therefore developed a computational tool that could reduce the bottlenecks in staging breast cancer and subsequently reduce mortality rate due to the disease.

General Terms
Expert Systems, Artificial Intelligence

Keywords
Breast Cancer; Breast Cancer Staging, Expert System; Rule-based system

1. INTRODUCTION
Cancer is caused by an abnormal growth of cells, which tends to proliferate in an uncontrolled way and in some cases, tend to metastasize [1]. There are many types of cancers divided into several groups and classifications, one of these is breast cancer which is now the most commonly diagnosed cancer [2], and so it is very important to know the operation of cancer and its level. Breast cancer, is the form of cancer, which originates from the breast tissue. It is the most common cancer in women in high, middle and low-income countries. According to [3], breast cancer resulted in 1.68million new cases and 522,000 deaths in 2012. As at in 2015, 2.1million cases of affected people were recorded which were mainly female and 533,600 death cases was recorded [4]. Although the disease affects both males and females, the estimated new cases of breast cancer in male very low compared to the estimated new-cases of the females with breast cancer. Apart from new cases reported, [5] also showed that there were approximately 40,480 deaths in United States.

Nigeria, which has approximately 20% of the population of Africa, and more than half of the population of West Africa contributed 15% of the estimated 681,000 new cases of cancer-related deaths in 2008 [6]. Also in Nigeria, about 100,000 new cases of cancer occur every year with high fatality ratio [7]. According to [8], 40% or more of breast cancer patients die due to late detection.

Breast cancer occurs in stages. According to [9], staging of breast cancer is dependent on three criteria with the acronym TNM, which means size of the tumor(T), whether or not the tumor has spread to the lymph nodes(N) in the armpits, and whether the tumor has metastasized, M(i.e. spread to a more distant part of the body); with larger size, nodal spread and metastasis having a larger stage number and worse prognosis. Staging of breast cancer is an important aspect in the management of the disease. This is because staging helps to determine the treatment suitable for a patient. Staging breast cancer is very cumbersome, as such inexperienced medical personnel depend on experts to understand the process. The shortage in the number of experts in the field makes the process often difficult. An expert system is a computer program that emulates the decision-making ability of a human expert [10]. Expert systems are designed to solve complex problems by reasoning through bodies of knowledge, represented mainly as IF–THEN rules rather than through conventional procedural codes. Expert systems have proven to be useful in several health-related areas. Despite the potential usefulness, the use of an expert system in breast cancer staging has not been explored. This study therefore developed an expert system that could aid the staging process. Section 2 gives the review of some work related to the study. Section 3 gives detail report on the method and tools used in the study. Section four discusses the results of the study, and the paper is concluded in Section 5.

2. LITERATURE REVIEW
In this section, the methods of staging breast cancer are examined. The section is concluded with review of some related work in the use of expert system for diagnosis.

2.1 Staging of Breast Cancer
Methods of Staging Breast cancer can be classified into two. These are Clinical staging and Pathological staging. Clinical staging is based on all available information obtained before a surgery done to remove the tumor. This may include information about the tumor obtained through physical examination, blood tests, radiologic examination, biopsy and endoscopy. Pathologic staging adds additional information gained by examination of the tumor microscopically by a pathologist after it has been surgically removed. Some of the screening procedures include Mammogram, Clinical Breast Examination, Magnetic Resonance Imaging (MRI), Genetic Test, Positron-Emission Tomography (PET)

Mammography makes use of low-dose X-ray system to

examine the breast tissue. The X-rays are connected into electrical signals. These electrical signals are then converted to digital images using different computer software. Two separate views are taken of the right breast, and two of the left. Cranio-caudal view and medio-lateral oblique view are the two separate views used to derive four images [11]. Mammography can detect changes in the breast up to nearly two years before a physician or a patient can be aware of them. Although mammography has been considered the gold standard for screening and early detection of breast cancer, it is not always feasible in many low- and middle income settings due to high cost of purchasing and maintaining the equipment, as well as difficulty in training and retraining skilled technologists.

Clinical Breast Examination is used by a certified health practitioner to physically observe and feel the breast. This in conjunction with mammograms is used to detect presence of lumps, and check for other breast abnormalities such as Mastitis of Fibro Adenoma.

Magnetic Resonance Imaging (MRI) of the breast is primarily used as a supplemental tool to breast screening with mammography or ultrasound. A breast MRI is mainly used for women who have been diagnosed with breast cancer to help measure the size of the cancer and look for tumor in the other breast [12].

Breast Ultrasound has been proved to be an exceptionally effective tool for imaging palpable abnormalities in the breast. It distinguishes cystic from solid masses and demonstrates those features of solid masses that would denote the mass of suspicious and warranting biopsy [13]

According to [9], there are five defined stages of breast cancer, each with further subcategories. However, because stage 0 has no evidence of cancer cells, and it is not cancerous, only four stages are usually considered.

2.2   Expert System

Expert Systems are designed to solve complex problems by reasoning through body of knowledge, represented using IF-THEN rules, rather than through conventional program. [14] revealed that expert systems emerged as a division of Artificial Intelligence (AI), which is an amalgamation of disciplines such as Computer Science, Engineering, Mathematics, Philosophy and Psychology. Expert system is a key area in Artificial Intelligence. The term Artificial Intelligence is applied when a machine mimics cognitive functions that are associated with other human minds, such as learning and problem solving. An expert system's information is obtained from expert sources and coded in a form suitable for reasoning process. An expert is someone who has prolonged or intense experience through practice and education in a particular field. Most times, experts are the ones to issue solution to some problems. However, when they are unavailable, it is desired or required for an expert system which could assist medical students in order to guide him/her on the kind of treatment to be given after being diagnosed [15]. In situations where oncologists are scarce or unavailable, an expert system can aid in appropriate staging of breast cancer.

An expert system is divided into two sub-systems: the inference engine and the knowledge base [16]. The knowledge base represents facts and rules. The inference engine applies the rule to the known to deduce new facts. Inference Engine can also include explanation and debugging activities. [17] divides expert system applications into ten (10) categories which are Interpretation, Prediction, Diagnosis, Design & Planning, Monitoring, Debugging, Repair, Instruction and Control. Expert based systems are the mainly common category of Artificial Intelligence in Medicine (AIM) system in routine clinical use. In fact, it was in the medical field that Expert Systems have made their existence felt first [18]. Basically, there are three types of expert Systems. These are Fuzzy, Rule-Based and Hybrid expert system.

The traditional rule-base expert systems are such in which the knowledge is represented using production rules, which defines how to solve a problem. The rules consist of two-parts: the IF part called the antecedent (Premise/Condition) and the THEN part, called the Consequent (Conclusion/Action). Rule based Expert systems are based on production model. The production model is based on the idea that humans solve problems by applying their knowledge (expressed as production rules) to a given problem represented by problem specific information. The production rules are stored in the long-term memory, and the problem-specific information or facts are stored in the short-term memory. Traditional rule-based expert systems have been used over the years for the diagnosis of medical conditions. MYCIN was the first well known medical expert system developed at Standard University to help doctors, not experts in anti-microbial drugs to prescribe drugs for blood infections. Clinical knowledge in MYCIN is represented as a set of IF-THEN rules, with certainty factors attached to diagnosis. It used a basic backward chaining reasoning strategy, resulting in exhaustive depth-first search of the knowledge base for relevant rules. It was the first convincing demonstration of the rule-based approach to the development of robust clinical decision support system. [19] developed a rule-based expert system for the diagnosis of lung disease using the diseases symptoms as system inputs. [20] developed a rule-based expert system for Neurological Disorders. The system diagnosed and suggested treatment for more than ten types of neurological disorders. The expert rules were built on the symptoms of each type of neurological disease and they were offered using decision tree and deduced using backward-chaining techniques. The knowledge base contains information gathered from practitioners about neurology and disorders. [21] developed an expert system for the diagnosis of Lassa fever using rule-based approach. [22] developed an expert system which is applied to detect major kidney diseases. The diagnosis is made using the readings obtained from clinical and para-clinical exams. The developed system eliminates the difficulties associated with diagnosing kidney disease. It has a very well-built knowledge base built around twenty-seven diseases from nine categories. Lung diseases have many regular symptoms and some of them are very much alike. This creates complications for the physicians to reach the right decision. [19] developed an expert system to identify the most important lung disease among patients. The medical expert system aids the doctor and it is acquainted with thirty-two lung diseases.

Table 1: Possible Symptoms and the Conditional Codes

BODY PART	FEATURE	Possible states	Conditional code
Breast	Cancerous tumor	(a) Positive	A
		(b) Negative	B
	If positive:		
	(i) size	(a) < 2cm	C
		(b) >2cm and < 5cm	D
		(c) >5cm	E
	(ii) Attachment	(a) To Underline muscle	F
		(b) To overlying skin	G
(Axilla)Armpit	Lymph nodes	(a) Swollen	H
		(b) Not swollen	I
	If swollen	(a) Mobile	J
		(c) Fixed/Matted	L
Distant organs (Lung) (Brain) (Bones) (Liver)	Cancer cells	(a) Positive (b) Negative	M N

Fuzzy Expert System is based on Fuzzy Logic which is used to calibrate vagueness. It is a logic that is based on the idea that all things admit to degrees. This is in contrast with Boolean Logic which makes use of sharp distinctions. Fuzzy Logic makes use of mathematical principle for knowledge representation, based on degree of membership. Fuzzy Expert System is based on membership functions and fuzzy rules. These functions and rules are used to reason concerning the data. It accepts members as inputs, and thereafter transform the input number into linguistic variable describing the degree of variation. These variations are called fuzzification [23]. Fuzzy rules and linguistic variables are stored in Fuzzy Knowledge base. It is the duty of the rules to map the input linguistic terms that concerns the outputs. This is done by the inference engine. The last step involves defuzzification which involves the translation of the output linguistic expressions into an output number. [24] developed a fuzzy Expert System for the diagnosis of back pain disease using body mass index, age, gender and clinical observations as inputs to the syatem. The system was tested using clinical data that corresponds to 20 patients with difference back pain disease, with accuracy of 90%. [25] developed a fuzzy expert system for the diagnosis of Hepatitis. The system utilizes fuzzy cluster system using a set of 10 symptoms which are Malaise, Biliousness, Lethargy, Orange urine, pale faeces, kidney failure, yellowish eye, Diarrhea and nausea. These symptoms serve as basic set of parameters needed for the diagnosis.

Hybrid Expert Systems integrate the similarities of Artificial Neural network and Fuzzy inference system into a single framework that increases the performance of the expert system. Neuro-fuzzy logic explores approximation techniques from neural networks to find the parameters of a fuzzy system. [26] developed a neuro fizzy expert system for the diagnosis of Leukemia. The study uses the predominant symptoms of leukemia which are paleness, fatigue, bone pain, asthemia, palpitation, frequent infection, nose bleeding and thrombocytopenia to classify the disease into Acute Lymphoblastic Leukemia (ALL), acute myeloid leukemia (AML), Chronic Lymphoblastic Leukemia(CLL) or Chronic Myeloid Leukemia(CML)

3.   METHODOLODY
The section discusses the method and tool used to address the set objectives. It consists of methods and tools used for the elicitation of knowledge, the design of the model, system implementation and the evaluation of the developed system. Each of these is discussed in the subsections below:

3.1   Elicitation of Knowledge
As mentioned earlier, all expert systems depend on knowledge elicited from a human expert. The knowledge used for the development of breast cancer staging system was elicited from Dr Adisa, a consultant Oncologist in Obafemi Awolowo University, and it aligns with [32]. The guidelines for staging breast cancer are shown below:

Stage I cancer is characterized by the following

i.   The tumour is non-palpable or is palpable and measures less than or equal to 2cm
ii.   No lymph nodes are involved

Stage II cancer is characterized by the following

i.   No tumour is palpable in breast but lymph nodes are swollen
ii.   Tumour measures 2cm or less, and has spread to the lymph nodes
iii.   Tumour larger than 2cm but smaller than 5cm, but has not spread to the lymph nodes
iv.   Tumour is larger than 5cm and lymph node(s) are swollen but mobile

Stage III cancer is characterized by the following

(i)   No tumour is palpable in the breast, but found in the lymph nodes that have clumped together or sticky to other structures.

(ii)   The tumour is of any size, and has spread to chest wall, or skin of the breast
(iii)   The tumour is of any size, and has spread to the lymph nodes and are matted (clumped together)

Stage IV is described as invasive cancer that has spread to other parts like the second breast, liver, bone, brain or lungs, etc.

According to the guidelines shown above, the body parts that are physically examined for features needed for staging of breast cancer are breast and armpit; and other distant organs such as lungs, brain, liver, the second breast etc. The breast is examined for lump. The lump can either be palpable or non-palpable. If the lump is palpable, the size of the lump and whether it is attached to underlying muscle or overlying skin is checked. The sizes of the lump are segmented into three viz: less than 2cm (< 2cm), between 2cm and 5cm, and greater than 5cm. Next the armpit is examined for swollen lymph nodes. Swollen Lymph nodes are examined to check if they are fixed/matted. Lastly, the distant organs are checked to see if the tumour has spread to them. The possible. The possible states of the examines body parts are represented using conditional codes as shown in Table 1.

3.2 Design of the Model

To ease the design of a rule-based model for the developed system, conditional codes are attached to each of the possible symptoms seen in the breast. Conditional code A is used to indicate situation in which the breast in the lump is palpable. Conditional code B is used to represent the situation in which the lump is not palpable. Conditional code A_1 is used to denote a situation in which the palpable lump measures less than or equals to 2cm. A_2 denotes the instance in which the lump size is between 2cm and 5cm, while A_3 refers to cases in which the lump size is greater than 5m. Conditional code C is

IF (A₁ OR B) AND F THEN STAGE 1

IF B AND E THEN STAGE 2

IF A₁ AND E THEN STAGE 2

IF A2 AND F THEN STAGE 2

IF A3 AND E1 THEN STAGE 2

IF B AND E2 THEN STAGE 3

(IF A1 OR A2 OR A3) AND (C OR D) THEN STAGE III

(IF A1 OR A2 OR A3) AND E2 THEN STAGE III

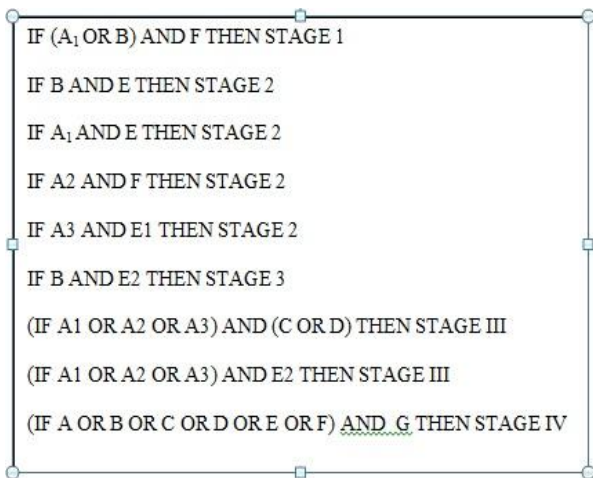(IF A OR B OR C OR D OR E OR F) AND G THEN STAGE IV

Fig. 1: Model Design Using IF-THEN rules

ttached to a situation in which the lump is attached to underlying muscle. Conditional code D denotes when the lump is attached to the overlying tissue. Code E refers to condition in which the lymph node found on the armpit is swollen. Code F refers to the condition in which the lymph node id not swollen. The code E_1 refers to the situation in which the swollen lymph node if mobile, while E_2 refers to the condition in which the other swollen lymph node is fixed/matted. Conditional code G depicts the situation in which the tumour has spread to distant organs, while the code H is used to represent where the tumour has not spread to distant organs. The symptoms and the conditional codes attached to them are shown in Table 1. Using the conditional codes shown in Table 1 and the criteria for staging described by American Joint Committee on Cancer (2009), the rules for staging breast cancer are reduced to logic statements as shown

in Figure 1.

3.3 System Implementation

Using the IF-THEN rules shown in Fig. 1, an android based breast cancer staging system was developed using JAVA programming language. The mobile version is preferred because of the popularity and portability of mobile devices. Figure 2 shows the staging area of the software, where the user is allowed to choose the symptoms presented by a patient. It shows the highlights of the symptoms presented by a 35 years old patient, Adesoga

Fig. 2: Cancer Staging Page

Abioye, who presented a case of palpable lump breast which has a size of between 2cm to 5cm. The lump is attached to the

underlying skin. Upon checking her armpit, the lymph was discovered to be swollen and matted. This breast cancer was diagnosed to be at stage 3. Figure 3 provides the user with an interface for keeping the record of the patient into a database, with detailed symptoms of the patient outlined. The stored record of the patient is shown in Figure 4.

Fig. 3: Entering Patient's Information

Fig. 4: Stored Patient's record

4. RESULTS AND DISCUSSION

The system was thereafter evaluated using Mean Opinion Score. Mean Opinion Score is a numerical measurement of the human judgement of an overall event or experience, and it gives a value ranging from 1 to 5(1-Bad, 2-poor, 3-fair, 4-good, 5-Excellent). In this study, the mean opinion score was used to judge the performance of the system to check it was able to mimic the intelligence of a medical expert. The system was tested using three performance metrics. These are accuracy, responsiveness and usability. Responsiveness measures how timely the software responds to user's request. Usability is concerned with testing how easy and user-friendly a software application is. Accuracy in the context of this study, examines the degree to which the developed system can accurately determine the stage of breast cancer, when

compared with what a human expert will do. Questionnaires containing 15 questions, with 5 questions targeted at each performance metrics were administered to ten medical doctors. The first five questions were related accuracy, the next five pertain to reliability while the last five questions were associated with usability. The evaluation data are shown in Table 2.

Table 2: System Evaluation Data

Question No	No of Respondent	Sum of values
1	10	42
2	10	45
3	10	40
4	10	39
5	10	40
6	10	41
7	10	43
8	10	42
9	10	40
10	10	41
11	10	38
12	10	42
13	10	45
14	10	42
15	10	40

The average result gotten for usability, accuracy and responsiveness were, 4.1, 4.1 and 4. This gives an average of 4.2 mean score for the developed system. Given an accuracy of 4.1, the developed system can therefore mimic the intelligence ability of humans.

5. CONCLUSION

The problem associated with staging of breast cancer has caused a lot of health hazard and affected families, psychologically. This is more pronounced in Nigeria where there is shortage of medical doctors in our hospitals. The introduction of an expert system in the diagnosis of breast cancer will therefore improve the treatment of breast cancer by providing an accessible method to effectively stage breast cancer, and thus improve the process of treating cancer among females. The study focused on the use of symptoms obtained from clinival examination alone, which are the ones done prior to the surgical removal of the tumor. This may be inadequate to accurately stage breast cancer in females. Future study could therefore consider the use of pathological information of the tumor, to enhance the accuracy of the staging system.

6. REFERENCES

[1] Zeichner, S; Herna, S; Mani, A; Ambros, T; Montero, A; Mahtani, R; Vogel, C. 2015. Survival of Patients with De-novo Metastatic Breast Cancer: Analysis of Data from a Large Breast Cancer-specific Private Practice, Breast Cancer Research and Treatment. 153(3), Pages 617-624.

[2] Sung H., Ferlay J., Siegel RL, Laversanne M. Soerjomataram I., Jemal A., Bray F. 2021 Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 cancers in 185 Countries, CA Cancer J Clin 71(3) pages 209-249

[3] McGuire, S., 2016. World cancer report 2014. Geneva, Switzerland: World Health Organization, international agency for research on cancer, WHO Press, 2015. Advances *in nutrition*, *7*(2), pp.418-419.

[4] Fitzmaurice, C., Dicker, D., Pain, A., Hamavid, H., Moradi-Lakeh, M., MacIntyre, M. F., ... & Global Burden of Disease Cancer Collaboration. 2015. JAMA oncology, 1(4), 505-527.

[5] Jemal A., Siegel R, Ward E., Hao Y., Xu J,. Murray T., Thun M.J, (2008) Cancer Statistics 2008, CA: A Cancer Journal for Clinicians Volume 58, Issue 2. Pgs 71-96

[6] Sylla, B. and Wild, C. 2011. A million Africans a year dying from cancer by 2030: What can cancer research and control offer to the continent? International Journal for Cancer. 4. 568.

[7] Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. 2010. Estimates of worldwide burden of Cancer in 2008, International Journal on Cancer, Volumer 127, pages 2893-2917

[8] Cheng H., Shi X., Min R., Hu L., Cai X., Du H., 2006. Approaches for automated detection and classification if masses in mammograms, Pattern Recognition 39 (4) 646-668.

[9] Carlson, R.W., Allred, D.C., Anderson, B.O., Burstein, H.J., Carter, W.B., Edge, S.B., Erban, J.K., Farrar, W.B., Goldstein, L.J., Gradishar, W.J. and Hayes, D.F., 2009. Breast cancer. *Journal of the National Comprehensive Cancer Network*, *7*(2), pp.122-192.

[10] Jackson, Peter 1998 Introduction to Expert Systems(3rd Edition), Addison Wesley, ISBN 978-0-201-87686-4

[11] Sampat, M.P., Markey, M.K. and Bovik, A.C., 2005. Computer-aided detection and diagnosis in mammography. *Handbook of image and video processing*, *2*(1), pp.1195-1217.

[12] Radhakrishna, S., Agarwal, S., Parikh, P.M., Kaur, K., Panwar, S., Sharma, S., Dey, A., Saxena, K.K., Chandra, M. and Sud, S., 2018. Role of magnetic resonance imaging in breast cancer management. *South Asian journal of cancer*, *7*(02), pp.069-071.

[13] Sood, R., Rositch, A.F., Shakoor, D., Ambinder, E., Pool, K.L., Pollack, E., Mollura, D.J., Mullen, L.A. and Harvey, S.C., 2019. Ultrasound for breast cancer detection globally: a systematic review and meta-analysis. *Journal of global oncology*, *5*, pp.1-17.

[14] Levin, Mark 2012. Towards Combinatorial Evolution of Composite Systems. Expert Systems with Applications, Volume 40, Number 4 pages 1342-1351

[15] Abu-Naser, S, Kashkash K and Fayyad, M. 2010. Development of an Expert System for Plant Disease diagnosis. Journal of Artificial Intelligence; Volume 3, Issue 4, pages 269-276.

[16] Nwigbo Stella, N. and Chuks, A.O., 2011. Expert system: a catalyst in educational development in Nigeria. Proceedings of the 1st International Technology, Education and Environment Conference (c) African Society for Scientific Research (ASSR).

[17] Haye-Roth, F ; Waterman D, Lenat Douglas 1983. Building Expert Systems. Addison-Wesley. ISBN 0-201-106686-8

[18] Han, C.W., 2016. Breast Cancer Diagnosis using Logic-based Fuzzy Neural Networks. *Digital Contents & Applications*, pp.69-72.

[19] Jimmy Singla, 2013. The Diagnosis of Some Lung Diseases in PROLOG Expert System, International Journal of Computer Applications, Vol 78. Issue 15. Pgs 37-40

[20] Ahmad A. Hajji , 2012. Rule Based Expert System for Diagnosis and Symptom of Neurological disorders, proceedings of ICCIT.

[21] Hambali Moshood Abiola, Akinyemi Adesina Alaba, Luka Joy D, 2017 Expert System For Lassa Fever Diagnosis Using Rule Based Approach, Annals Computer Science Series, Vol. XV fasc. 2.

[22] Roventa, E., & Rosu, G. 2009. The diagnosis of some kidney diseases in a small prolog Expert System. In 2009 3rd International Workshop on Soft Computing Applications (pp. 219-224). IEEE

[23] Noura, A.(2015), Heart Disease Diagnosis using Artificial Neural Network, IISTE Network and Complex Syestems, 5(4).

[24] Kadhim, Mohammed Abbas, M. Afshar Alam, and Harleen Kaur.Design and implementation of fuzzy expert system for back pain diagnosis. International Journal of Innovative Technology & Creative Engineering 1, no. 9 (2011): 16-22.

[25] Imianvan, A.A, and Obi J.C, 2011. Diagnostic Evaluation of Hepatitis Utilizing Fuzzy Clustering Means. World Journal of Applied Science and Technology (WOJAST), The Official Publication of the Faculty of Science, University of Uyo, *3*(1), pp.23-30.

[26] Obi, J.C. and Imianvan, A.A., 2011. Interactive neuro-fuzzy expert system for diagnosis of leukemia. *global journal of computer science and technology*.