AUTOMATED BREAST CANCER DIAGNOSIS SYSTEM

BY

AJAYI JOSEPH FRIDAY

MATRIC NO: ICT/2251940330

BEING A PROJECT WORK SUMMITED TO THE DEPARTMENT OF COMPUTER SCIENCE, SCHOOL OF INFORMATION AND COMMUNICATION TECHNOLOGY, AUCHI POLYTECHNIC, AUCHI, EDO STATE.

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF HIGHER NATIONAL DIPLOMA (HND) IN COMPUTER SCIENCE

FEBRUARY, 2022

CERTIFICATION

We, the undersigned certify that this project work was carried out by Ajayi Joseph Friday with a Matriculation Number: **ICT/2251940330**of the department of computer science.

We also certify that the work is adequate in scope and quality in partial fulfillment of the requirements for the award of **Higher National Diploma (HND**) in computer science.

___________________________ ______________

Mr. Akhetuamen, S. O. **Date**
(Project Supervisor)

__________________________________ _______________

Mr. Akhetuamen S. O. **Date**
(Head, Department of Computer Science)

DEDICATION

This project work is dedicated to God Almighty for His grace, strength and good health through the period of my studies

ACKNOWLEDGEMENT

I am over helmed in all humbleness and gratefulness to acknowledge my depth to God and all those who have helped me to put these ideas, well above the level of simplicity and into something concrete.

I would like to express my special thanks of gratitude to my supervisor who is also the Head of Department of Computer Science, Mr. Akhetuamen S. O. who gave me the golden opportunity to do this wonderful project on the topic (Automated breast cancer diagnostic system), which also helped me in doing a lot of Research and also to the lecturers of the department of Computer Science to came to know about so many new things. I am really thankful to them.

Any attempt at any level can't be satisfactorily completed without the support and guidance of My parents (Mr. and Mrs. Ajayi Omoju) My beloved siblings and friends Paul Ojo, Ajala Rilwan, Mr. Gospel, Mr. Ford. Church of God mission in his presence.

I would like to thank my special gift(Rosemary) who helped me a lot in gathering different information, collecting data and guiding me from time to time in making this project, despite of her busy schedules.

Also I will like to thank Dr. and Mrs. Ihoyoka who gave me different ideas in making this project unique.

TABLE OF CONTENT

ABSTRACT

Breast cancer has affected women for a long time now due to lack of information, barriers of accessibility or affordability, it still poses as a significant threat to population. The aim of this work is to create a web based system that will assist its users in the cancer diagnosis process, which includes an interactive session between the system and the user through the system interface for a quick feedback and earlier detection as well as give more information and knowledge about breast cancer. The system has demonstrated more than 90% success rate. In addition, the performance issue has addressed time wastage in accessing medical services. The study outlines the main concepts of the analysis and design methodology of the proposed system, compares it to the existing and goes further to explain the design and implementation of the system. System was achieved using HTML, php, CSS and MySQL and database and run on a windows xp or higher versions. The fact finding techniques employed is interview, observation, online and library research.

Keywords: Breast, Patient, Cancer, Diagnosis, Automated

CHAPTER ONE

INTRODUCTION

1.0 Background to the Study

From the very earliest moments in the modern history of the computer, scientists have dreamed of creating an ‘electronic brain’. Of all the modern technological quests, this search to create artificially intelligent (AI) computer systems has been one of the most ambitious and, not surprisingly, controversial.

It also seems that very early on, scientists and doctors alike were captivated by the potential such a technology might have in medicine (Ledley and Lusted, 2015). With intelligent computers able to store and process vast stores of knowledge, the hope was that they would become perfect ‘doctors in a box’, assisting or surpassing clinicians with tasks like diagnosis. With such motivations, a small but talented community of computer scientists and healthcare professionals set about shaping a research program for a new discipline called Artificial Intelligence in Medicine (AIM). These researchers had a bold vision of the way AIM would revolutionize medicine, and push forward the frontiers of technology. Medical artificial intelligence is primarily concerned with the construction of all the programs that perform diagnosis and make therapy recommendations. Unlike medical applications based on other programming methods, such as purely statistical and probabilistic methods, medical AI programs are based on symbolic models of disease entities and their relationship to patient factors and clinical manifestations.

1.1 Statement of the Problem

It is unwise saying that the hospital is an exemption of the numerous life applications where manual operations pose some limitations with respect to effective service and productivity. Most women today face a lot of breast diseases due to lack of information or ignorance about several breast diseases, because there are no easy access to diagnose this disease on time. Therefore there is need for an implementation of an expert system on breast diseases diagnosis to detect these diseases on time and for early treatment.

1.2 Aim and Objectives of the Study

The aim of the project is to design an automated that can be used to diagnose breast diseases online.

The objective could be summarized as follows:

1. Aid health care providers responsible for large patient, the expert system can answer basic or general questions, leaving more time for individuals to patience with peculiar situations
2. Offering prompt feedback and self-evaluation.

3. Providing potentially infinite array of information of the steps to take in the eventuality of a particular occurrence.
4. Aiding the nurses and other staff to know what to in the case of emergency if the human expert is not present at that point in time.

1.3 Significance of the Study

Expert system proved greater diagnostic accuracy to physicians across a wide range of medical specialist. The success of such a system depends upon the strength of its knowledge base, and its easy use.

1.4 Scope of the Study

This research focus on breast cancer diagnosis using University of Benin as a case study.

1.5 Definition of Basic Terms

Web Portal: A website which uses the World Wide Web links to specify subject, or a website that offers a broad array of resources and services such as forums, emails, online shopping malls etc.

Database: A collection of information on a subject that is organizedin such a way that a computer program can quickly select desired piece of datafrom it. It can also be seen as an electronic filing system.

Cancer: A disease in which the cells of a tissue undergo uncontrolled proliferation.

Breast: Either of the two organs on the front of a woman's chest, which contain the mammary glands.

System: a collection of organized things.

Diagnosis: the identification of the nature and cause of an illness.

Medicine: a substance which specifically promotes healing when ingested or consumed in some way.

Injection: The act of injecting

CHAPTER TWO

LITERATURE REVIEW

2.0 The Evolution of Computers, Radiology and Decision-Making

Ever since man first learnt to communicate, knowledge that are to be shared and used by humans is most likely to be confined to what is stored in a person's head or what the person can learn from another (Swett, *et.al*. 2015).Consider the following words by William, (2012): "Medicine is a science of uncertainty and an art of probability".

Important components of the art of medicine are skills in repeatedly making decisions, formulating appropriate judgments and being comfortable with risk and uncertainty. Medical training, with its heavy emphasis on factual learning, often assigns a lesser priority to the study of decision making. Our own history of medicine contributes to dismissive attitudes about decision making. Before the later part of the 19th century, medical treatment was largely a matter of tradition, spurred on by a physician's need to do something for the patient (Umi, 2016).

2.1 Relevance of Health Informatics and Training

The future of Health Informatics (HI) as a profession is thus very promising. In other words, MI means managing medical and health care through information science and engineering technology. Like medicine, HI is also multidisciplinary.

HI deals with the entire domain of medicine and health care, from computer-based patient records to applications of image processing and from primary care practices to hospitals and regions of health care (Umi, 2016).

A few years ago, only a handful doctors had even heard of the term “health informatics.” Health informatics is a relatively new sub-speciality of medicine which uses information technology to manage clinical information. At a three-day eHealth Asia 2004 conference held at Kuala Lumpur in early April 2004, a local expert, Dr. H.M.Goh, council secretary of the Malaysian Health Informatics Association (MHIA) stated that there is space for growth in the local health informatics scene since few public and private hospitals have significant health management systems in place. The extraordinary thing about eHealth Asia 2004 was that it was attended by 350participants (compared to 250 participants in 2001) which featured 54 speakers from over 20 countries around the world. This indicates that the field of health informatics has made itself felt throughout the world. Globally, health informatics includes change management, artificial intelligence, messaging, mobile technology and the like. Only 10 of the 120 government hospitals are computerized, and only the Putrajaya and delayang hospitals have been fully-enabled with health informatics (Star, 2004). With this upcoming awareness, the field of health informatics is very relevant to the Malaysian market. It is very strongly felt and believed that the research involvement in this study addresses a portion of and fits into this niche of health informatics.

2.2 Radiology and the use of Technology

The rapid evolution of technology and clinical research makes it difficult even for the specialist to keep up. In the light of this 'information explosion', it has been demonstrated that physicians do not always make optimal decisions. It has been mentioned earlier in the introduction that immense knowledge needs to be dissipated amongst health providers through in-depth training. Specifically in radiology, this strategy has been fairly effective in large academic centers but realistically, much has to be done by radiologists to practice state-of-the-art radiology at the forefront of radiological practice, especially in Malaysia. Although computers have proven to be very efficient and helpful in carrying out mundane tasks and the processing of data into useful information, its potential as a powerful technology can be further exploited to assist radiologists in knowledge processing.

The utilization of computers in decision-making can be employed in many different forms. However, the basic understanding to be realized and engraved in each and everyone's mind is that these tools in decision-making have never been and are never intended in the first place to camouflage or belittle the decision makers in health care. Computers can be made as slaves to record huge amounts of detailed information. Simultaneously, these vast and abundant accumulated wealth of knowledge and information can be made available to radiologists at their disposal, put to use wherever or whenever abnormalities are encountered and ultimately arrive at a more consistent decision-making. Some diagnoses can be made in a more quantitative, algebraic fashion although it cannot be denied that most radiological decision-making is very subjective. An expert is usually consulted for solving a difficult diagnostic problem in this situation. Some diagnoses can be made in a more quantitative, algebraic fashion although it cannot be denied that most radiological decision-making is very subjective (Henry, 2015).

An expert is usually consulted for solving a difficult diagnostic problem. This situation and paradigm has served as a model for the birth of a class of computer systems that are known as expert systems. A KBS is designed to meet the knowledge gaps of the individual physician with specific patient problems. KBS and such other ES can be a boon to the rural health centres because even general medical practitioners can operate the systems. These are ideal examples of AI. has grown steadily since their introduction. Zhang, Pham & Chen (2016) used applications of fuzzy logic in rule-based expert systems involving the problem of auto-focusing camera lens system and also another on a financial decision system. Tocatlidou *et al.* (2016) built an ES that was capable of diagnosing plant diseases and disorders while Park & Storch (2016) shared a representation of expert system in the shipbuilding industry which was able to downsize sizable development costs. Craker & Coenen (2006) proposed Knowledge Bazaar, the concept of which a paradigm for the development of expert system and knowledge bases are created dynamically using knowledge supplied by self-appointed internet communities. The philosophy underpinning the Knowledge Bazaar is the observation that knowledge can be accumulated, not from a limited number of experts or expert sources, but dynamically from internet users as they solve problems and offer advice. Mahmod *et al.* (2014) had shown the usage of neural networks combined with an expert system environment. Perhaps, all the relevant studies are best encapsulated in the paper (Liao, 2013) where expert system methodologies in almost all applications have been reviewed by the author for a span of a decade beginning from the year 1995.

2.3 Expert Systems in Medicine and Medical Application Areas

Expert or knowledge-based systems are the most common type of artificial intelligence in medicine (AIM) system in routine clinical use. Indeed, it was in the medical area that expert systems have made their presence felt in the first place.

AIMs contain medical knowledge, usually about a very specifically defined task and are able to reason with data from individual patients to eventually emerge with reasoned conclusions. Although there are many variations, the knowledge within an expert system is typically represented in the form of a set of rules (Keles et. al., 2013).

2.4 Expert Systems Evolution

Expert systems emerged as a branch of artificial intelligence – an amalgam of disciplines such as computer science, mathematics, engineering, philosophy and psychology. From the efforts of AI researchers, computer programs are developed that can reason as humans. expert system are one of the most commercially viable branches of AI and although there have been reports of expert system failures, surveys show that many parties have remained enthusiastic proponents of the technology and continue to develop important and successful applications in various fields (Duan *et al.*, 2013).

2.5 The Application Areas of Expert Systems

From its early days of infancy when MYCIN (Negnevitsky, 2013) was first pioneered, expert system have been developed in broad walks of life, in various areas and disciplines ranging from geology, statistics, electronics to medicine. In fact, the sky has no limit! To emphasize on this matter, a kaleidoscope of the expert systems developed in their respective fields is mentioned here. Williams (2015) suggested a prototype expert system for the design of complex statistical experiments. Geoplay is a knowledge based expert system developed by the U.S. Geological Survey that is available for explorations in the oil and gas industry. Yang *et al.* (2013) developed an expert system for vibration fault diagnosis of rotating machinery using decision tree and decision table and Duan *et al.* (2013) addressed the issues associated with the design, development and use of web-based expert system from a standpoint of the benefits and challenges of developing and using them. Wagner *et al.* (2001) applied ES to various problem domains and for the entry decisions of new products in business applications. The use of expert system in business Other areas of specific medical applications of expert systems are in Obstetrics and Gynaecology (Medical Decisions, 2003), for leukemia management (Chae *et al.*, 2012), for estimating the prognosis of head injured patients in intensive care unit (Sakellaropoulos & Nikiforidis, 2014), for heart valve diseases (Turkoglu *et al.*, 2016), applied to brain MRI (Zhang & Maeda, 2014), even as early as the 1980's to determine the irreversible cessation of all functions of the entire brain before any other organ transplantation (Pfurtscheller *et al.*, 2012). Alonso *et al.* (2016) developed a medical diagnosis system, obtained by combining the expertise of a physician specialized in isokinetic and data mining techniques where patients may exercise one of their knee joints using basically a physical support machine according to different ranges of movement and at a constant speed. Lee *et al.* (2012), introduced a holistic system, which amalgamates case-based reasoning, rule-based reasoning, causal-based reasoning and an ontological knowledge base for managing clinical incidents in general practice enabling health professionals to share medical incident information, which has caused harm and may or can cause potential harm. The re-use of such information may prevent or mitigate human or medical errors. Morelli *et al.* (2012) and Solano *et al.* (2006) both described computational systems for automated diagnosis of depression and as an aid to clinical decision making in the mental health field. Verdaguer *et al.* (2012) investigated the application of ES in patients suffering from pneumonia, while Leong (2012) developed a system to detect irregularities by analyzing heart sounds through the interpretation and analysis of auscultatory findings. Costly and sometimes deadly clinical incidents may occur during the provision of health care, such as errors in dispensing inappropriate drugs due to the similarity of medication names to a patient for example. Lee *et al.* (2012) developed a prototype for this situation. In the same year, Li (2012) proposed a system to diagnose AIDS risky patients. While Lee & Lee (2015) suggested that future medical E.S. be specifically developed having at least one, if not all of these three characteristics i.e. simulate the performance of group of human experts, deal with chronic diseases and deal with several diseases simultaneously, Lhotska *et al.* (2001) focused on efficiency enhancements on rule based systems.

2.6 Breast Cancer Scenario

Breast cancer is among the leading causes of deaths in women worldwide. Its incidences have been rising at an alarming rate. More and more women have been subjected to the misery, suffering and pain caused by the disease. In Malaysia alone, approximately one in 20 women will be afflicted with breast cancer by the age of seventy, and by the age of 85, women have a one in eight chance of developing breast tumour. In the year 2014, almost 4,000 newly diagnosed cases emerge in the country. Of these, nearly 45% result in deaths, making it the number one cause of cancer-related deaths among Malaysian women (Sunday, 2003). In a technical report drafted by the Ministry of Health Malaysia in the year 2001, 20% of patients afflicted by all kinds of the 1392 cancer cases have died from breast cancer alone. In its first report, the National Cancer Registry stated that 26,089 people were diagnosed with cancer in Peninsular Malaysia in the year 2016, of which 14,274 (55%) cases were cancers among women and 30.4% of it, were cancer of the breast (Mat, 2004). In Europe, 2004 estimates indicated

371,000 new cases with 129,900 breast cancer deaths. Mortality rates rose from 1951 until 1990 but fell noticeably in Western Europe, especially in the United Kingdom. However, this is not the case in Eastern and central Europe. Although rates in Hong Kong and Japan have been lower than those in Europe, they have also been increasing. Rates in North and South America are similar to Western Europe and so is Australia. The reasons for this decline in mortality rates in Western Europe, Australia and the Americas include the widespread practice of mammographic screening (Boyle *et al.*, 2013).

2.7 The Necessity for this Work

Looking at the previous facts, we are immersed in a war, where the latent 'enemies' that we are confronted with in the battle against this killer disease is lying dormant out there, lurking and striking from unexpected corners. Indeed, we find ourselves in a difficult situation. In formulating the strategy best taken in this 'war', certain points as in the following, are noteworthy. Diagnosticians with the training and experience to interpret mammographic images are scarce. Therefore, there is an emphasis in training new radiologists to be able to interpret the mammographic images. The situation would be more crucial if mass screening were to be adopted as a national policy in this country as has been practiced in certain countries in the west. In the early period of a doctor's professional activity, an expert system would prove valuable in minimizing the troubles that he or she might face due to inexperience. The existence of such facilities could be helpful especially for young radiologists or non-specialists. The existence of a diagnostic tool to aid in the interpretation process has been proven to be more useful for the junior than for the senior radiologists (Baileyguier *et al.*, 2013). It is also a valuable teaching tool for the junior radiologists. Sensitivity also improved slightly for the senior radiologists. However, specificity remained unchanged in the study. An expert system for this application would make diagnostic expertise more widely and readily available in the clinical community. Therefore, the success of medical imaging depends on subjective factors that influence the ability of the observer to ' interpret the information'. These factors can be summarized into two broad classifications:

1. Those factors that are image dependent and relate to the visual conspicuity of features relevant to the clinical problem; and

2. Those that are image independent; are primarily cognitive in nature and relate to what the observer knows about the visual information in front of him.

Variation between readers was greater than the differences between imaging techniques (Manning *et al.*, 2013). There are many image acquisition, display and processing parameters, and their effects on optimizing images for human interpretation are largely unknown. But we know less still, allowing the observer to structure the task of interpreting image features; perhaps a better understanding of these factors now deserves our research attention so that we can achieve a better match of image displays to cognitive/perceptual skills. The availability of Computer Aided Detection (CAD) and Computer Aided Diagnosis (CADx) should be employed with a word of caution. That is, it is important that the development and availability of such systems do not detract from quality and the need for radiological skills across the imaging workforce. In other words, the skill of radiologists using such CAD and CADx remains paramount. Maintaining high radiological skill levels whilst using technology efficiently and effectively to formulate correct diagnostic decisions quickly is a key issue for the future (Manning *et al.*, 2013).

Even though the Breast Imaging Reporting and Data System (BI-RADS) was introduced to help standardize feature analysis and final management of breast modality findings, there still exists variations in their interpretations. Continued efforts to educate radiologists to promote maximum consistency still need to be carried out (Lehman *et al.*, 2016). The risk of breast cancer increases with age. Considerable evidence indicates that older women frequently do not undergo mammography. Offering on-site mammography at community-based sites where older women gather is an effective method for increasing breast cancer screening rates among older women (Reuben *et al.*, 2016). It is hoped that this work may be useful in filtering only the abnormal cases to be further scrutinized by specialists. Routine and repetitive use of computer-based systems developed for experiments would bring several benefits. Radiologists could be trained to evaluate the perceptual features appropriately (D'Orsi *et al.*, 2012). In clinical practice, only 15-30% of patient referred for biopsy are found to have a malignancy (Petrick et.al, 2016). Unnecessary biopsies increase health care costs and may cause patient anxiety and morbidity. It is therefore important to improve the accuracy of interpreting mammographic lesions (Petrick et.al., 2016), thereby improving the positive predictive values of detection modalities.

As the expert system contains specific rule base for the differentiation of breast diseases, it may be utilized both to help train physicians in breast cancer modalities and to promote a more consistent mammographic and ultrasound interpretation. The criterion for interpreting imagery is subjective and variable. With the help of an expert system, the diagnostic criteria can be made more explicit. This would serve as a basis for consistent and reproducible diagnoses. At the same time, it would also form the basis for discussion and further research to improve the validity of the diagnostic criteria. Expert systems would serve as models with intelligent behaviour in cognitive and perceptual realms and skills to solve problems thought to require human intelligence. People are better at clarifying a problem, suggesting kinds of procedures to follow, judging the reliability of facts and deciding if a solution is reasonable. The problem solver must know how to use knowledge and see patterns in the signals presented. To sum up, the following points are relevant:

1. The human heuristic approach of combining evidence to reach a prognosis can deal successfully with a limited amount of evidence. The proliferation of large databases of patient findings, due to the increased use of computers in clinical settings, offers an abundance of available data, challenging the limited human capacity for indirect inference. Decision support systems that are able to model uncertainty and analyze diverse sources of information can therefore become a useful tool for medical experts (Sakellaropoulos & Nikiforidis, 2014).
2. Some of the most successful applications have been for instruction e.g. use of a medical expert system to develop diagnostic skills thus encouraging students to structure knowledge and process it systematically in response to a problem or abnormality. Also, as precise analytical models of knowledge

and through the ways in which they are used, expert systems can enhance our understanding of human decision-making processes.

3. As clinical decision making inherently requires reasoning under uncertainty, expert systems will be suitable techniques for dealing with partial evidence and with uncertainty regarding the effects of proposed interventions (Shortliffe, 2012).
4. Radiology is gradually developing a more systematic approach to training, replacing the traditional mixture of ad hoc apprenticeship and formal lectures with a combination of structured tuition and case-based experiential learning. This is intended to meet a long-recognized need for clinicians to encapsulate general medical knowledge within the development of skills through diagnostic practice. A structured approach to training can have the additional benefit of equipping learners with a coherent 'conceptual framework': an appropriately defined and organized notation that enables them to externalize, reflect on and share diagnostic knowledge (Structured Computer-based Training, 2013).
5. Radiological expertise is based on two kinds of skills: the swift and accurate processing of normal appearance, and the ability to distinguish disease from normal variation in appearance. Thus, skill development in radiology

requires exposure to, and reporting of a large range of images, so that recognition of varied normal anatomy are firmly etched in the minds of the skilled interpreters and cognitive resources can be devoted to the process of describing abnormal appearances (Structured Computer-based Training, 2013).

6. Despite the wide applications of AI techniques to a range of clinical activities, few expert systems have been implemented in the field of medical imaging; its scarcity possibly due to the inherent difficulty in high-level vision. The data acquired from medical scanners can be noisy and ambiguous. Nevertheless, the potential benefits make it tempting to aim at designing expert systems using the digital images provided by the various modalities, especially with the advent of networking of medical images through PACS and the DICOM format. DICOM is an international standard, recognized by most hardware and software manufacturers for the storing and transmission of medical images acquired with all modalities (Chabat *et al.*, 2014). This work is an attempt to fulfill or partially fulfill the dearth of imaging expert systems and with the inadvertent and inevitable emergence of digital

mammography, radiologists would need to undergo pertinent retraining (Digital Imaging, 2004). Hence, this work will all the more be relevant.

2.8 Artificial Intelligence and Medicine

Man strives to augment his abilities by building tools, from the invention of the club to lengthen his reach and strengthen his blow to the refinement of the electron microscope to sharpen his vision, tools have extended his ability to sense and to manipulate the world about him. (Szolovits, 2016) Today we stand on the threshold of new technical developments which will augment man's reasoning, the computer and the programming methods being devised for it are the new tools to effect this change.

According to Schwartz, 2016 in his book AI and Medicine he opined that medicine is a field in which such help is critically needed, our increasing expectations of the highest quality health care and the rapid growth of ever more detailed medical knowledge leave the physician without adequate time to devote to each case and struggling to keep up with the newest developments in his field. For lack of time, most medical decisions must be based on rapid judgments of the case relying on the physician's unaided memory. Only in rare situations can a literature search or other extended investigation be undertaken to assure the doctor (and the patient) that the latest knowledge is brought to bear on any particular case.

Continued training and recertification procedures encourage the physician to keep more of the relevant information constantly in mind, but fundamental limitations of human memory and recall coupled with the growth of knowledge assure that most of what is known cannot be known by most individuals, it is the opportunity for new computer tools: to help organize, store, and retrieve appropriate medical knowledge needed by the practitioner in dealing with each difficult case, and to suggest appropriate diagnostic, prognostic and therapeutic decisions and decision making techniques (Schwartz, 2016).

The key technical developments leading to this reshaping will almost certainly involve exploitation of the computer as an ‘intellectual,’ ‘deductive’ instrument–a consultant that is built into the very structure of the medical-care system and that augments or replaces many traditional activities of the physician. Indeed, it seems probable that in the not too distant future the physician and the computer will engage in frequent dialogue, the computer continuously taking note of history, physical findings, laboratory data, and the like, alerting the physician to the most probable diagnoses and suggesting the appropriate, safest course of action (Doyle, 2013).

This vision is only slowly coming to reality. The techniques needed to implement computer programs to achieve these goals are still elusive, and many other factors influence the acceptability of the programs.

CHAPTER THREE

SYSTEM ANALYSIS AND DESIGN

3.0 Introduction

This chapter gives an overview of the general design process undergone in order to develop the application. It entails the architecture of the proposed system, and the design of the various components of the system. These components are categorized into three layers which include front layer, middle layer and the back layer.

3.1 Description of the Existing System

The existing system is one that has been manually operated over the years. It is a system in which the registration and breast diagnosis of patient is of a manual approach. Critical analysis of this system reveals that it is prone to errors or slow in operation. Careful analysis also shows that due to the complexities of the manual system, records of patient kept are mostly inaccurate and manually operated in such a way that requires nurse to register each patient in a book registry before seeing the patience for diagnosis, some patient usually complain that this process is stressful. Sometimes patient wait tirelessly to be diagnosed by the doctor and because of this so many person will refuse to go to the hospital.

3.2 Description of the Proposed System

The major fact taken into consideration in this design is to automate the breast diagnosis process. In this system, patient don't have to be on queue to see the doctor. One just need to sign up and use the user name and password created to get access to the series of question.

3.3 System Design

This is the process of defining the architecture, components, modules, interfaces and data for a system to satisfy specified requirements. Systems design could be seen as the application of systems theory to product development. There is some overlap with the disciplines of systems analysis, systems architecture and systems engineering.

Unified Modeling Language Diagrams:

UML stands for Unified Modeling Language. UML is a language for specifying, visualizing and documenting the system. This is the step while developing any product after analysis. The goal from this is to produce a model of the entities involved in the project which later need to be built. The representation of the entities that are to be used in the product being developed need to be designed.

Used Case Diagram

Use case Diagram for Patient

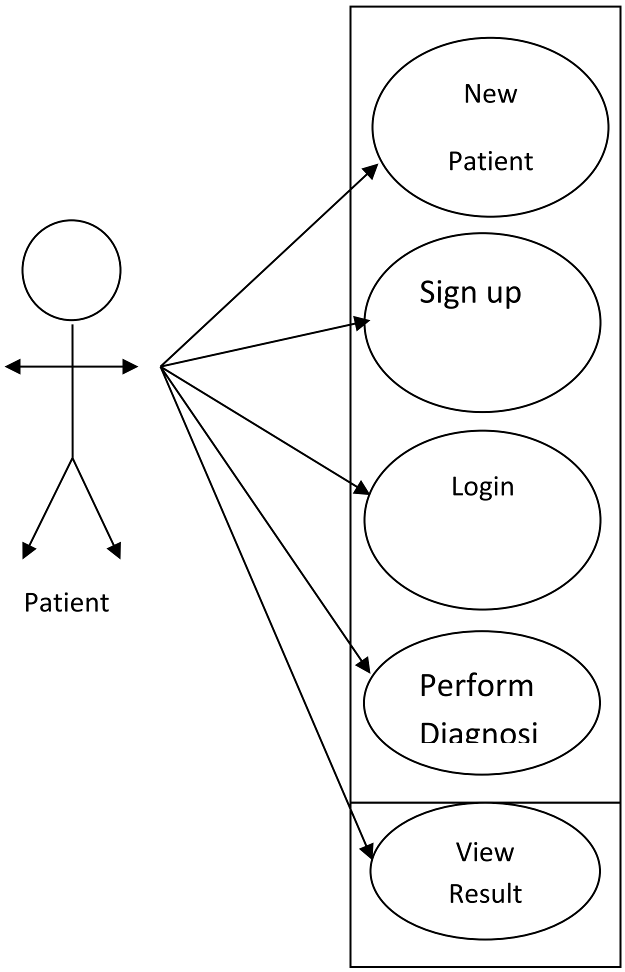


Fig 1: Patient use case diagram

DATA FLOW DIAGRAMS:

The DFD takes an input-process-output view of a system i.e. data object flow into the software, which are transformed by processing elements, and resultant data objects flow out of the software.

Data objects represented by labeled arrows and transformation are represented by circles also called as bubbles. DFD is presented in a hierarchical fashion i.e. the first data flow model represents the system as a whole. Subsequent DFD refine the context diagram (level 0 DFD), providing increasing details with each subsequent level.

Dataflow diagrams:

Database:

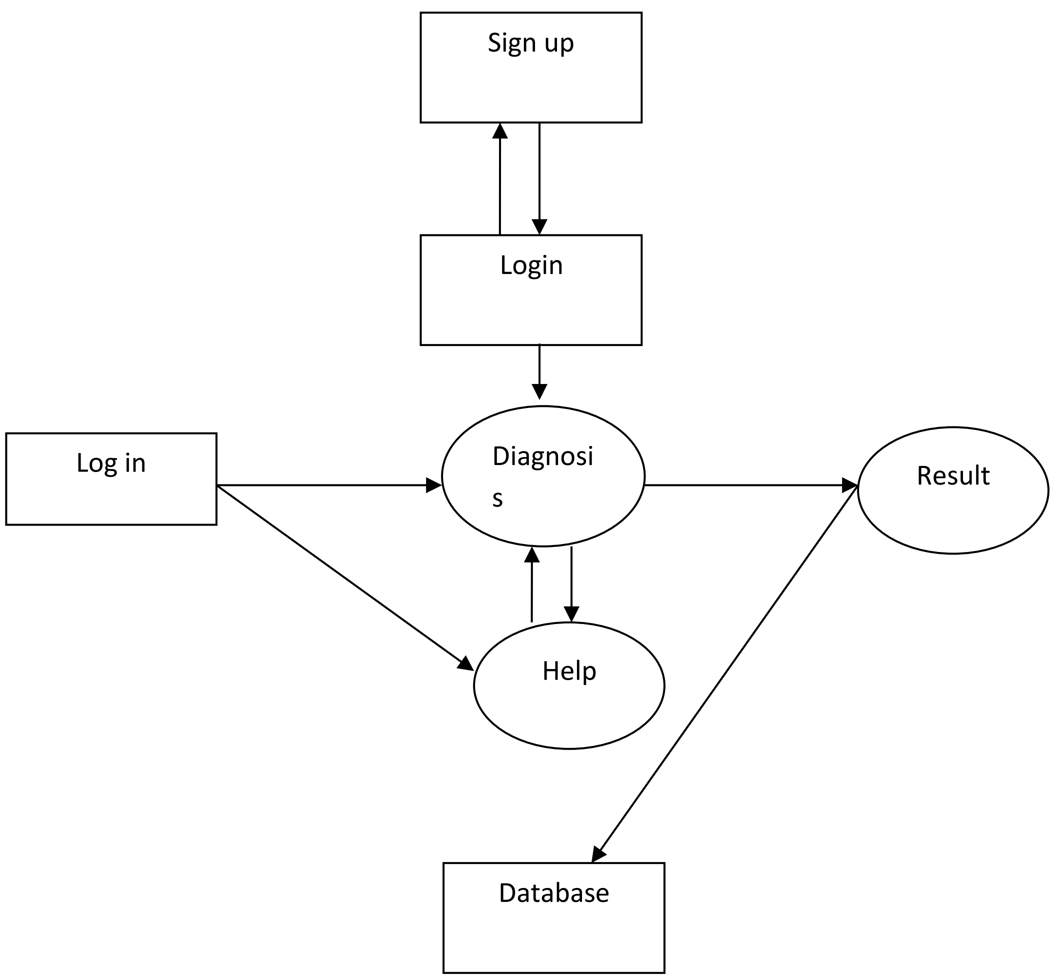


Fig 2: Dataflow diagram

User information database structure:

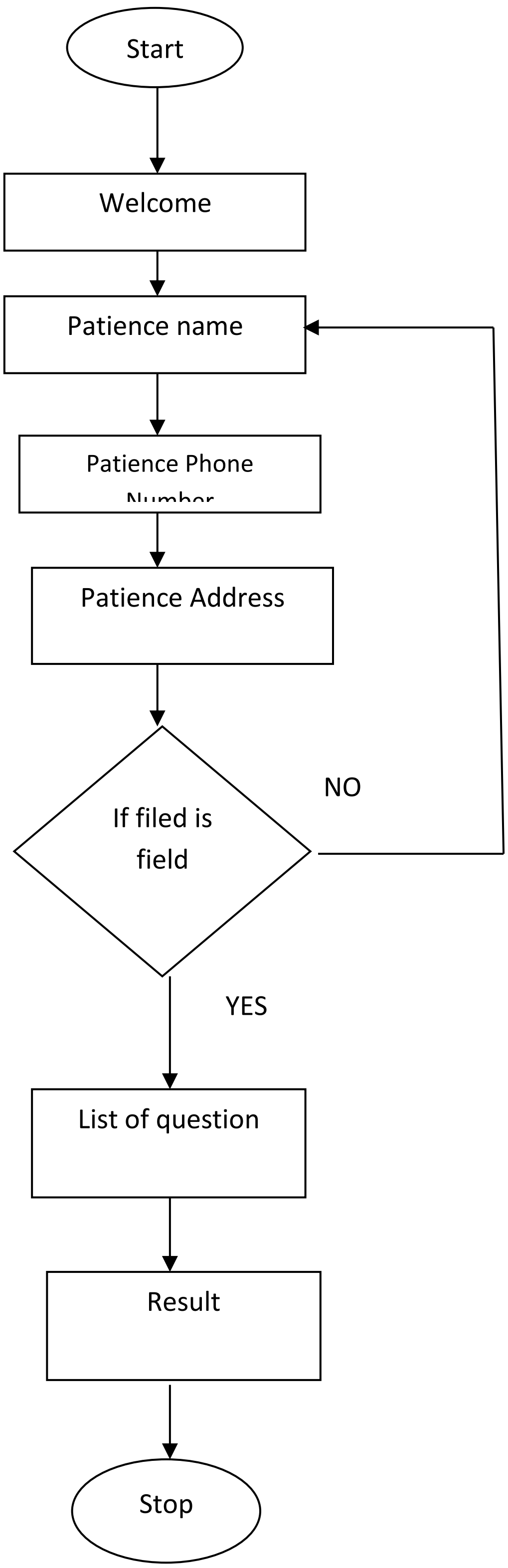
Table 3.1

FIELDNAME	DATATYPE	DESCRIPTION
Sn	Integer	This is an auto incrementing field to hold the count of all records in the database
Patience name	Text	This field holds the complete name of the user.
Patience Phone Number	Text	This field will hold the phone number of patience
Address	Text	This field will hold the full address of patience

3.5 Program Flowchart

Flowchart is a grammatical representation of an algorithm. Program flowchart spells out dramatically the detailed step by step procedure and control logic necessary for an algorithm solution of a data processing problem. Below is the program flowchart of this work.

Fig 3 Program Flowchart



3.5 System Architecture

It is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures of the system, which comprise system components, the externally visible properties of those components, the relationships (e.g. the behavior) between them, and provides a plan from which products can be procured, and systems developed, that will work together to implement the overall system.

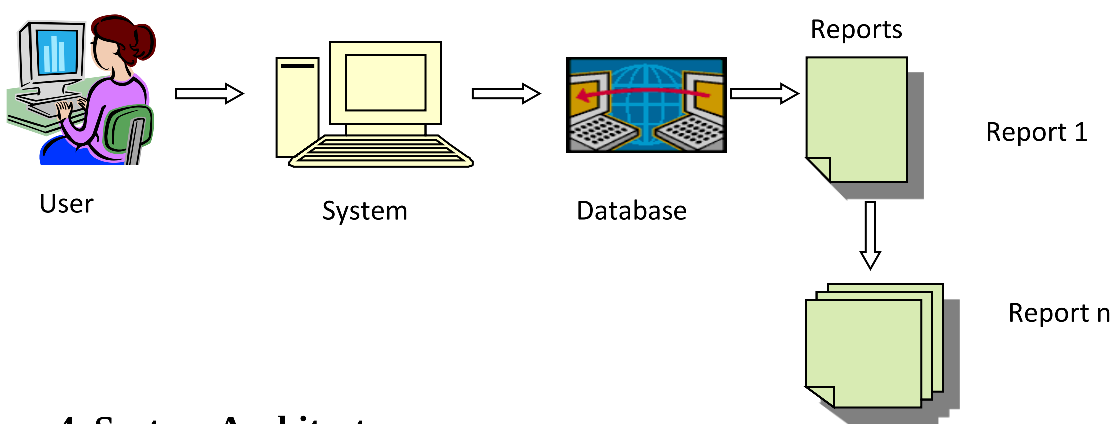


Figure 4. System Architecture

3.6 Site Structure

The proposed site has a very simple structure with easy navigation. Figure 3.2 below depict the structure using block diagram.

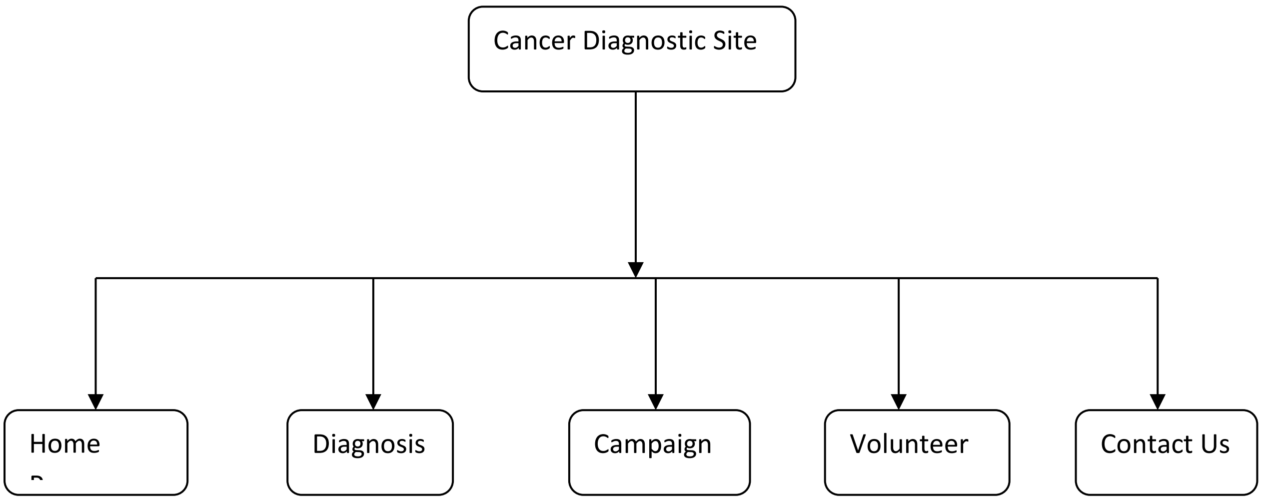


Figure 5: Structure Chart for Proposed System

The proposed architecture has 5 main modules. The Home page module provides general information to the end user. It also contains links to other modules of the system. The Campaign module gives specific information about the subject matter. These include factors that contribute to the growth of Cancerous cells and ways to prevent such. It also showcases the various activities of the group. The volunteer module manages the collection of end user intents to be active part of the group in order to achieve the objective of the system. The last module allows end users to give feedback on the performance of the system, contribute to knowledge and also make enquiries.

3.7 Front Layer Design

The front layer design combines the input design, the interface design and the output design. A number of special design considerations should be taken into account during the design. These include sign-on procedures, interactive processing, and error detection.

i. Sign-On Procedures: These include identification numbers, passwords, and other safeguards needed for an individual to gain access to computer resources

ii. Interactive Processing: With systems, people directly interact with the processing components of the system through terminals or networked PCs. The system and the user respond to each other in a real-time mode, which means within a matter of seconds. The design approach can include menu-driven systems, help commands, table lookup facilities, and restart procedures.

iii. Error Detection: The best and least expensive time to deal with potential errors is early in the design phase. Good system design attempts to prevent errors before they occur. This process includes developing a good backup system that can recover from an error.

The interfaces are simplified to allow the ease of location of vital information by targeted audience. The golden rules for the design of good interface were applied in the design process.

3.8 Database Design

Relational database is used. It made use of logically related tables as units for storing and retrieving data. Tables used in this system are: symptoms, member, admin, and results. The database tables in design view are made of five columns:

i. Field: This holds the value obtained with the input form in the system.

ii. Type: This consists of the data type in the field name, that is, it determines the kind of values that users can store in the given field such as varchar, tinyint, text, date, smallint, mediumint, int, etc; and the field size holds the size of the data in the system.

iii. Null: This field determines whether an empty value is permitted for a field or not. Usually some fields are not required while others are required. The required fields are specified as null.

iv. Default: This is the default value that appears in any field when a new table row is to be created.

v. Primary key: To avoid data duplication and referential integrity, primary key is used in the design of the tables. It also links one table to another in the database.

Symptoms

The symptoms database file contains the list of questions that can help the system in decision making in order to correctly diagnose a Breast Cancer patient. The table contains the symptom id (a unique identifier for each symptom), the question, the id of the next question to be asked if the user's response is Yes, the id of the next question to be asked if the user's response is No, Associated image (if applicable) and the weight. Due to the fact that some diagnostic question requires physical examination or observation, images are attached to some of the question to help the user give correct answer. The structure for the symptoms table is as shown in Table 3.1 below.

Table 3.2: Structure of Symptom Database File

Field	Type	Null	Default	Primary
symptomid	int(11)	No		Yes
question	varchar(255)	No		No
positivenext	int(3)	No		No
negativenext	int(3)	No		No
assoc_image	varchar(255)	No		No
weight	int(3)	No		No

Figure 1: Symptoms Table

Result

The Result Database file stores the result of diagnosis carried out by the system. This can be used to evaluate the performance of the system and also to monitor patient. The structure is as shown in table 3.2 below

Table 3.3: Structure of Result database file

Field	Type	Null	Default	Primary
resultid	int(11)	No		Yes
userid	int(11)	No		No
testdate	Date	No	0000-00-00	No
score	int(5)	No		No
systemadvice	varchar(255)	No		No

Fig 2. Database Structure

User

The user database file keeps record of the system users. The structure is as shown in table 3.3 below

Table 3.4: Structure of User database file

Field	Type	Null	Default	Primary
userid	int(11)	No		Yes
username	varchar(50)	No		No
password	varchar(50)	No		No
age	int(3)	No		No
stateo	varchar(50)	No		No
country	varchar(50)	No		No
phoneno	varchar(20)	No		No

Fig 3: User Table

Administrator

The member database contains the list of system administrators. The structure is as shown in the table below.

Table 3.5: Structure of Member Database File

Field	Type	Null	Default	Primary
member_id	int(11)	No		Yes
fullname	varchar(100)	No		No
imagepath	varchar(100)	No		No
gender	varchar(10)	No		No
phoneno	varchar(20)	No		No
email	varchar(50)	No		No
stateo	varchar(50)	No		No

Fig 4: Member Table

3.9 SYSTEM FLOW CHART

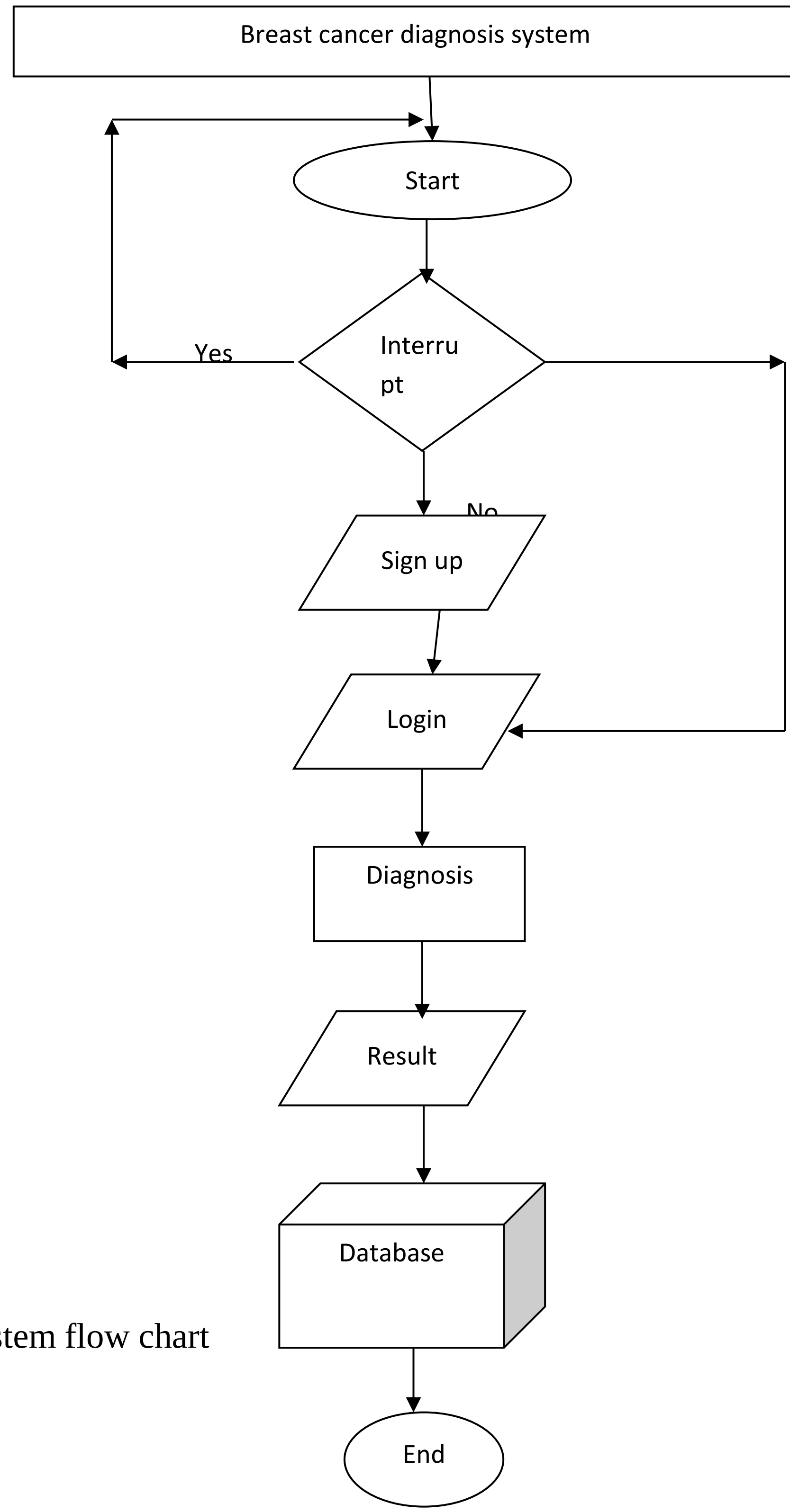


Fig 5 System flow chart

CHAPTER FOUR

SYSTEM IMPLEMENTATION

4.1 Introduction

In the previous chapter, we have been able to develop the architecture for the Web-based breast cancer diagnosis which has been delineated using structure charts and flow charts in order to achieve the aims and objectives. This chapter presents the implementation of the proposed design.

System Requirement

4.2 Hardware Requirement

The hardware requirements are the set of component of the system which can be seen or felt. For effective and effective functioning of the system, the following hardware specifications are recommended for the web server:

1. A VGA/EGA Colour Monitor
2. A LaserJet or DeskJet printer
3. An Uninterruptible Power Supply (UPS) unit
4. A stabilizer of about 1KVA
5. External storage device (Flash, Diskette, CD-R etc.)
6. A Central Processing Unit with the following configuration
7. A Pentium IV Processor with a speed not less than 1.6GHz

8. 1TB Hard disk size
9. 1GHz RAM (Random Access Memory) size
10. Floppy disk drive/ USB port
11. A standby Power Generating Set, and
12. A modem

4.3 Software Requirement

The Software Requirement comprises of programs and data that make the hardware to carry out the specified task. The following software packages are required for the application to function.

1. SQL (Structured Query Language)
2. PHP 4.1 Software
3. Web Server (Apache, IIS, App. Server, etc.)
4. Operating System (UNIX, Window, etc.)

Other system requirement includes

1. A reliable and up-to-date Antivirus Software
2. An internet Connection

4.4 Choice of Programming Language

There are various language used in instructing the computer to carryout specified tasks. Most standalone applications are implemented using high level languages such as Visual BASIC 6.0, FORTRAN, QBASIC, C++, etc. but in case of application that runs on the web; scripting languages are used as tools for the implementation.

Some commonly used scripting languages are HTML, JavaScript, VB Script, JSP (Java Server Pages) Java, Applets, ColdFusion, .NET(s), PHP, etc.

In order to implement the proposed system, PHP (Hypertext Preprocessor) was used as the implementation tool.

The following are the justification for the selection of PHP among others for the implementation:

Hypertext Processor is a powerful server-side scripting language for creating dynamic and interactive websites.

Hypertext Processor is the widely-used, free, and efficient alternative to competitors such as Microsoft's ASP. Hypertext Processor is perfectly suited for Web development and can be embedded directly into the HTML code.

The Hypertext Processor syntax is very similar to Perl and C. PHP is often used together with Apache (web server) on various operating systems. It also supports ISAPI and can be used with Microsoft's IIS on Windows.

In addition, the Online Course Registration Application is a database driven application. Among the available database application which includes cybase, dbase, Microsoft Access, Oracle, MySQL, etc,MySQL was used for the database and the choice was based on the following reasons

SQL (Structured Query Language) is

1. MySQL is a database server
2. MySQL is ideal for both small and large applications
3. MySQL supports standard SQL
4. MySQL compiles on a number of platforms
5. MySQL is free to download and use

4.5 Installation Process

Web applications are being made available to the end user by hosting it on a system called the web server. A web server is a computer system that is connected to the internet, and can be viewed by any other computer connected to the internet with a browser such as Mozilla Firefox, Microsoft Internet Explorer, etc.

4.6 Program Analysis

The Web-based Breast Cancer diagnosis System can be accessed by user located at any part of the world after the site has been hosted and published but for the unit and system testing, a local server will be used.

The application has modules which includes; the Main page, Diagnosis, Campaign, Volunteer and About Us. Each section is explored below.

4.6.1 The Main Page

The application homepage is the first contact to the system hence the need for user friendliness and accuracy of information. Also the ease of navigation is an important quality that must be met. All these features were put into consideration in the design of the system. The result colourful interface is presented in figure 4.1. The homepage has the links to other pages located at the upper part of the application.

Login

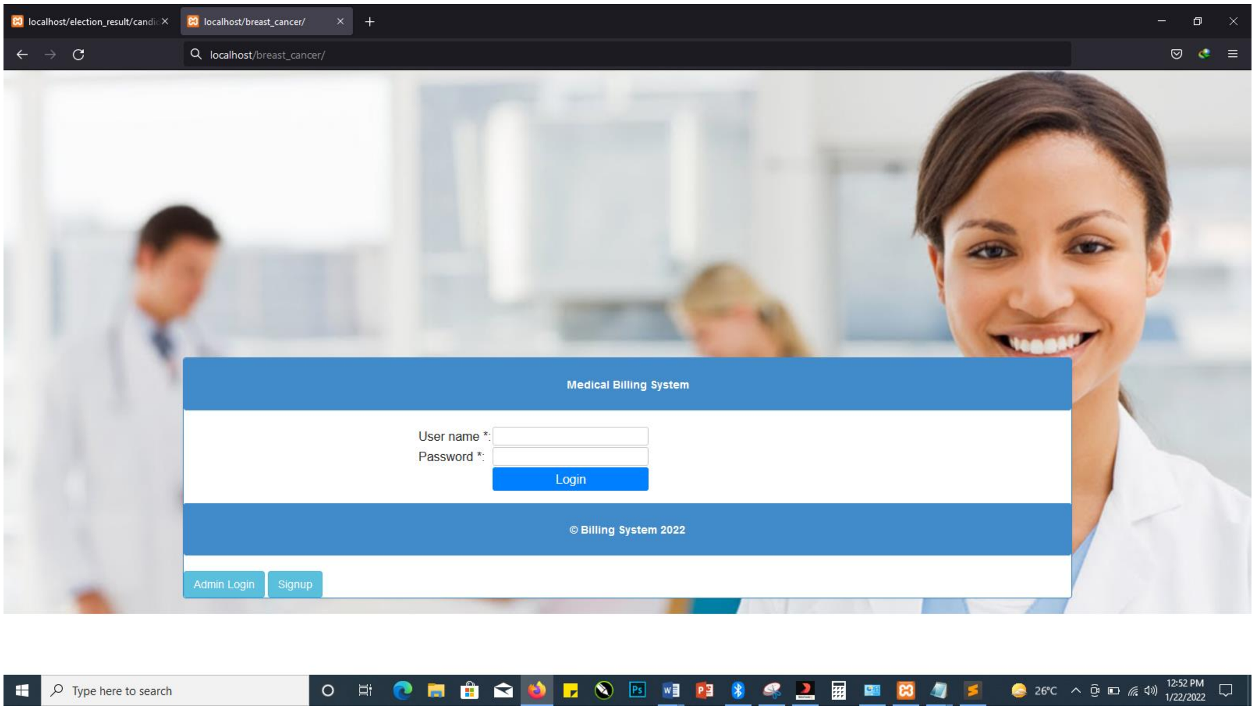


Figure 4.1 Application Homepage

4.6.2 Signup Page

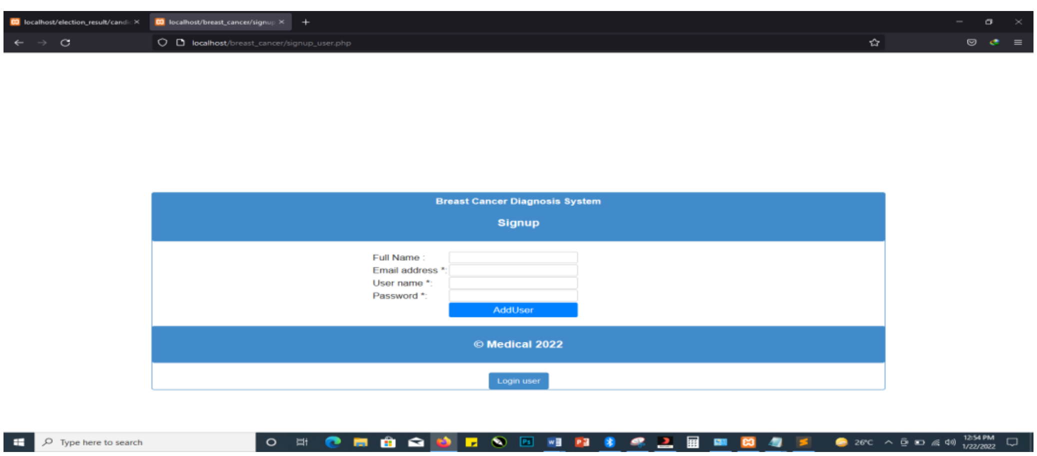


Fig 4.2 Signup page

4.6.3 Diagnosis Page

The diagnosis page is the core component of the application. The user will have to click on the start button in order to initiate the diagnosis process. Figure 4.5 shows the diagnosis page. Questions are asked from the user as shown in figure 4.6. Some questions require illustration in order to aid the user to know what exactly to be done. Such questions have images attached to the question for illustration purpose as shown in figure 4.7. Once the user completes the answering of the question, the result is shown to the user without further delay and suggestions will be displayed to the user.

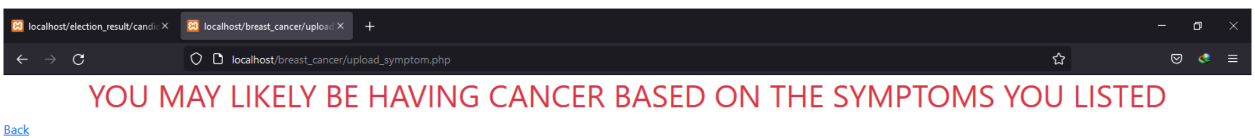


Fig 4.3 Diagnostic Result

CHAPTER FIVE

SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.0 Summary

Breast Cancer is a significant public health issue and is largely curable when detected in its early stages. The largest barriers to timely diagnosis appear to be constraints of access and affordability. It is imperative that a more accessible and affordable solution be proffered in order to reduce the mortality rate due to breast cancer. This study examined some alternatives to mammogram that can be used to detect breast cancer through simple self-examination. The proposed solution modeled the problem domain as a set of symptoms S with corresponding weight W. The average of the weight is evaluated along the path of diagnosis examination and the result is compared with a threshold value. The inclusion of a flexible threshold value was in line with findings of Krishnamurthy (2007) who explored the differences in breast cancer prognostic factors: age, at diagnosis, histology type of cancer, and summary stage diagnosis, and how they contribute to the diminished survival rate of African-American women diagnosed with breast cancer, an noted that the race/ ethnicity is a factor to be considered during diagnosis as indicated in their theoretical model and backed up by their findings. The proposed solution was implemented as a web-based application that is easily accessible to the general public.

5.1 Conclusion

While public health represents a myriad of discipline, the core principles strive to improve health and wellbeing for the populace. This study has focus on the application of computer science tools and techniques in promoting the well-being of the general population by presenting an alternative to the inaccessibility and affordability of breast cancer screening equipment. It should be noted that this system is not a replacement for mammogram but a better alternative to absolute lack of screening facilities.

5.2 Recommendations

Due to the high rate of mortality due to late detection of breast cancer, a solution has been proposed which do not only carry out diagnosis but also educate people on how to reduce the risk of being victim of breast cancer. Hence, it is recommended that the system should be used not only by those who felt they might be having cancer but the general populace. Also, further research should be carried out to model the distribution of population of people suffering from breast cancer in order to dynamically generate the threshold value for the system. Finally, research of this nature should be funded by the government and other authorities to improve on the quality and reliability of the system result.

REFERENCES

Alonso, L. Ledley, P. and Lusted, M. (2016). Predictiva Value of Tc sestamibi scintigraphy in the evaluation of doxorubicin based chemotherapy response in patients with advanced breast cancer. *Nuclear Medicine Communication* 23: 765-771

Baileyguier, J. Boyle, P. Veronesi, U. Goldhirsch, A. Orecchia, R. Viale, G. (2013). Breast cancer**.** *Lancet* 2005, 365**:**1727-1741.

Blaning, W. (2003). Readings in Medical Artificial Intelligence: *The first decade.*

Boyle, Y. Joe, F. Henry, S. (2013): The power of knowledge in radiologic education. *Journal of Digital Imaging.* 4(4) 99-101

Chae, M. Chabat, K. Hinton, P. Roebuck, J, Williams, M. Blamey, R. (2014). Mammographic parenchymal patterns: value as a predictor of hormone dependency and survival in breast cancer. ***American Journal of*** *Roentgenol* 1985**,** 144**:**1103-1107.

Craker. C. (2006) Expert system based on neuro-fuzzy rules for diagnosis breast cancer," *Expert Systems with Applications*, 38: 5719-5726.

D'Orsi, J. Ledley, R. S. & Lusted, L.B., (2012). Reasoning Foundations of Medical Diagnosis. *American Association for the Advancement of Science* 130 (3366): 9-21.

Diuan, J. Mak, B. and Blanning, R. (2013) A logic-based approach to rule induction in expert systems. Expert Systems, 20: 141–149. doi: 10.1111/1468-0394.00237

Duan, et al. (2013).Thermal conductivity of superconductive Y-Ba-Cu-O. *Phys. Rev. B* Vol. 36, No. 7, 3917-3919

Geo, U. (2003) "Breast cancer detection: Evaluation of a massdetection algorithm for computer-aided diagnosis—Experience in 263 patients," *Radiology*, vol. 224, pp. 217–224.

Henry, P. (2015). Overview of ship-design expert systems. Expert Systems, 19: 136–141. doi: 10.1111/1468-0394.00199

Keles, *et al* (2013). Coastal Lithofacies and Biofacies Associated with Syndepositional Dolomitization and Silicification (Draken Formation, Upper Riphean, Svalbard). Precambrian Research 53(3-4): 165-197

Ledley, K. and Lusted, O. (2015) Reasoning under uncertainty for plant diseases diagnosis. Expert Systems, 19: 46–52. doi: 10.1111/1468-0394.00188

Lee, K. *et al* (2012). A mammogram and breast ultrasound-based expert system with image processing features for breast diseases. Ph.D research work, University of Malaysia, (Unpublished)

Leong, L (2012) Decision making in Radiology. *Perspective, Division of Medical Education, University of Southern Clifornia.*

Lhotska . Broglio, K. Gilcrease, M. Dawood, S. Dempsey, J. Valero, V. Hortobagyi, G. Atchley, D. Arun B. (2001): Mammographic features of triple receptor-negative primary breast cancers in young premenopausal women. *Breast Cancer Res Treat*, **111:**405-410.

Li, U. (2012). "A Fuzzy Genetic Algorithm for Creative Shape Design", Proceedings of the Fourth International Symposium on Tools and Methods of Competitive Engineering.315-329, Wuhan, Hubei, China

Liao, K. (2013) Epidemiology and diagnosis. UpToDate. 2021. Accessed at https://www.uptodate.com/contents/gestational-breast-cancerepidemiology-and-diagnosis.

Mahmod Litton JK, Theriault RL. (2014).Chapter 65: Breast Cancer During Pregnancy and Subsequent Pregnancy in Breast Cancer Survivors. In: Harris JR, Lippman ME, Morrow M, Osborne CK, eds. Diseases of the Breast. 5th ed

Manning et al 2013 Breast Cancer Treatment During Pregnancy (PDQ). 2019. Accessed at https://www.cancer.gov/types/breast/hp/pregnancy-breast-treatment-pdq

Mat, L. (2004). NCCN Clinical Practice Guidelines in Oncology. Breast Cancer Screening and Diagnosis. Version 1

Moreli, O. Sechopoulos I, Suryanarayanan S, Vedantham S, D'Orsi CJ, Karellas A. (2012). Radiation dose to organs and tissues from mammography: Monte Carlo and phantom study. Radiology. 2008;246(2):434-443.

Nikiforidis, K. (2014). Reducing delay in the diagnosis of pregnancy-associated breast cancer: How imaging can help us. J Med Imaging Radiat Oncol. 55(1):33-42

Park, L. & Storch, L. (2016). Breast imaging for cancer screening: Mammography and ultrasonography. UpToDate. 2021. Accessed at https://www.uptodate.com/contents/breast-imaging-for-cancer-screeningmammography-and-ultrasonography

Petrick, B. (2016) Viswanathan S, Ramaswamy B. Pregnancy-associated breast cancer. Clin Obstet Gynecol. 2011;54(4):546-555

Pfurscheller, Yang WT, Dryden MJ, Gwyn K, (2012) Imaging of breast cancer diagnosed and treated with chemotherapy during pregnancy. Radiology. (1):52-60.Schwartx

Shortliffe, K. (2012). New insights into the metastatic behavior after breast cancer surgery, according to well-established clinicopathological variables and molecular subtypes. PLOS ONE 12, e0184680

Solano *et al* (2006) Time trends of overall survival among metastatic breast cancer patients in the real-life ESME cohort. Eur. J. Cancer 96, 17–24
Star 2004 Clinical value of circulating ESR1 mutations for patients with metastatic breast cancer: a meta-analysis. Cancer Manag. Res. 10, 2573–2580

Sunday, D. (2003). Genomic evolution of breast cancer metastasis and relapse. Cancer Cell 32, 169–184.e7

Sweet, H. *et al* (2012) Liquid biopsy: will it be the ‘magic tool’ for monitoring response of solid tumors to anticancer therapies? Curr. Opin. Oncol. 27, 560–567

Szolovits, J. (2016). Panel on breast cancer screening. the benefits and harms of breast cancer screening: an independent review. Lancet 380, 1778–1786

Tocatlidou, D. (2016). Breast-cancer screening — viewpoint of the IARC Working Group. N. Engl. J. Med. 372, 2353–2358

Turkoglu et (2016). Overdiagnosis of breast cancer in population screening: does it make breast screening worthless? Cancer Biol. Med

Umi, E. (2016). Effect of mammography screening on surgical treatment for breast cancer in Norway: comparative analysis of cancer registry data. BMJ 343, d4692–d4692

Verdaguer, E. et al (2012). Effect of mammography screening on surgical treatment for breast cancer: a nationwide analysis of hospitalization rates in Germany 2005–2009. Eur. J. Epidemiol. 28, 689–696

Wagner, D. *et al* (2013). Evidence that an early pregnancy causes a persistent decrease in the number of functional mammary epithelial stem cells — implications for pregnancy-induced protection against breast cancer. Stem Cells 26, 3205–3209

While, L. and Le, K. (2015). Causes of cancer in the world: comparative risk assessment of nine behavioural and environmental risk factors. Lancet 366, 1784–1793

William, E. (2012). Tennekoon, K. H. & Samarakoon, S. R. Emerging role of histone deacetylase inhibitors as anti-breast-cancer agents. Drug Discov. Today 24, 685–702

Yang, D. *et al* (2013). A phase II study of the histone deacetylase inhibitor vorinostat combined with tamoxifen for the treatment of patients with hormone therapy-resistant breast cancer. Br. J. Cancer 104, 1828–1835

Zhang, F. and Maeda, D. (2014). Oestrogen receptors in breast cancer: basic mechanisms and clinical implications. Ecancermedicalscience 7, 370

Zhang, C. (2016). Estrogen receptors outside the nucleus in breast cancer. Breast Cancer Res. Treat. 108, 351–36

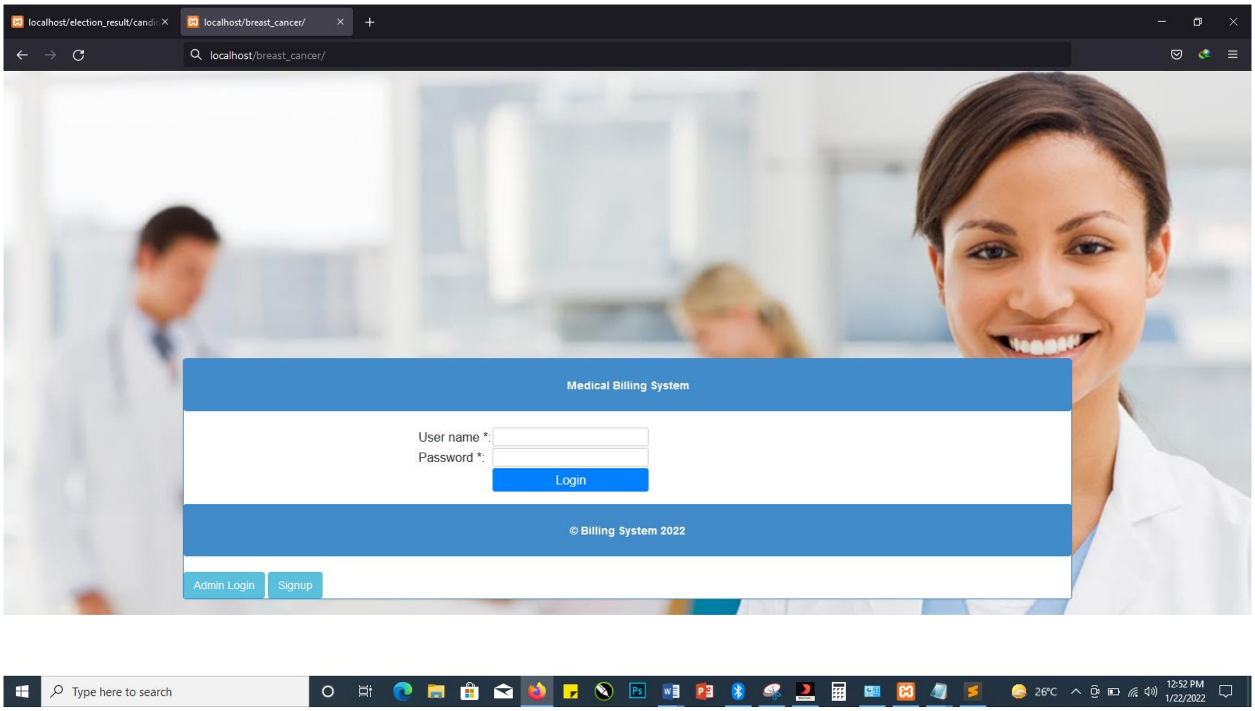
Appendix I

```
<!DOCTYPE html>

<html lang="en">

	<head>

		<title>Home</title>

		<meta charset="utf-8">

		<meta name = "format-detection" content = "telephone=no" />

		<link rel="icon" href="images/favicon.ico">

		<link rel="shortcut icon" href="images/favicon.ico" />

		<link rel="stylesheet" href="css/owl.carousel.css">

		<link rel="stylesheet" href="css/slider.css">

		<link rel="stylesheet" href="css/style.css">

		<script src="js/jquery.js"></script>

		<script src="js/jquery-migrate-1.1.1.js"></script>

		<script src="js/script.js"></script>

		<script src="js/jquery.ui.totop.js"></script>

		<script src="js/superfish.js"></script>

		<script src="js/jquery.equalheights.js"></script>

		<script src="js/jquery.mobilemenu.js"></script>
```

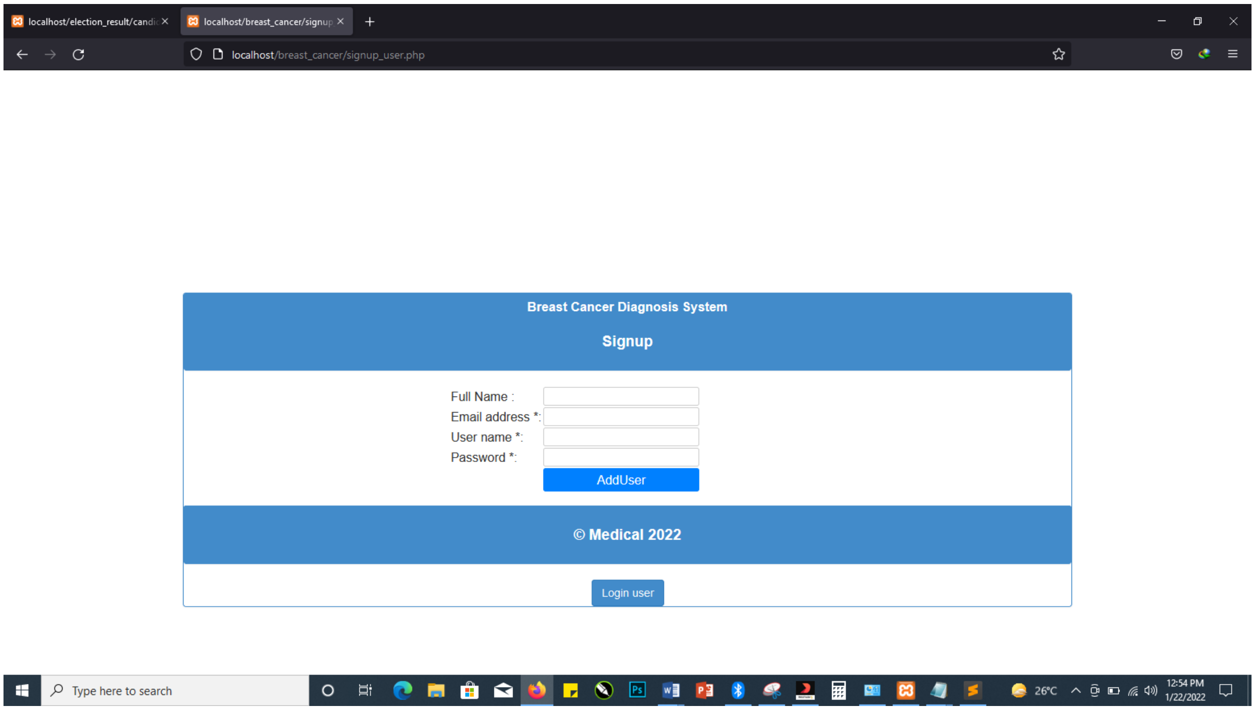
```
<script src="js/jquery.easing.1.3.js"></script>
<script src="js/owl.carousel.js"></script>
<script src="js/jquery.flexslider-min.js"></script>
<script src="js/kwiks.js"></script>
<script>
    $(document).ready(function(){
        $().UItoTop({ easingType: 'easeOutQuart' });
        var owl = $("#owl");
        owl.owlCarousel({
            items : 4, //10 items above 1000px browser width
            itemsDesktop : [995,3], //5 items between 1000px and 901px
            itemsDesktopSmall : [767, 2], // betweem 900px and 601px
            itemsTablet: [700, 2], //2 items between 600 and 0
            itemsMobile : [479, 1], // itemsMobile disabled - inherit from itemsTablet option
            navigation : true,
```

Appendix II

Login



Signup



Diagnostic Result