Cancer Care Insight

EMPOWERING WOMENS' HEALTH

	Student Name	Email
1	Aya Ahmed	aa6064@fayoum.edu.eg
2	Esraa Khamis	ek1217@fayoum.edu.eg
3	Sohaila Elbadry	sw1152@fayoum.edu.eg
4	Mariam Samuel	ms2946@fayoum.edu.eg
5	Yasmin Gamal	Yg1152@fayoum.edu.eg
6	Youstina William	yw1124@fayoum.edu.eg

Department/ Bioinformatics

Supervised By

Dr/ Mary Monir

2024

Abstract

Breast cancer is a type of cancer that originates in the cells of the breast. It occurs when abnormal cells in the breast tissue begin to grow and divide uncontrollably, forming a malignant tumor. Breast cancer can affect both women and men, although it is much more common in women. Traditional classification use morphology to divide tumors into separate categories with different behavior and prognosis. However, there are limitations of traditional classification systems, and new molecular methods are expected to improve classification systems. our website dedicated to supporting doctors in the field of breast cancer. We provide valuable resources and tools to aid medical professionals in their understanding and management of this complex disease. In addition to conventional approaches, we leverage the power of deep learning models to enhance diagnostic accuracy and treatment planning. These models analyze vast amounts of data, enabling us to extract meaningful insights and make informed decisions based on the individualized characteristics of each patient's tumor. Our website offers a comprehensive platform for doctors to access the latest research, clinical guidelines, and innovative tools incorporating deep learning algorithms. By harnessing the potential of these cutting-edge technologies, we aim to empower healthcare professionals in their mission to provide the best possible care for breast cancer patients. We already built a two successful models first model using CNN with accuracy of 96% and second model using Random Forest with accuracy of 96%.

Acknowledgement

The success of any project depends largely on the encouragement and guidelines of many others. We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project. The special thank goes to our helpful supervisor DR: Mary Mounir. The supervision and support he gave truly help in the progression of our graduation project. We also take this opportunity to express our gratitude to the Board of the Faculty of Computers and Information for their Support and efforts in providing us with all useful resources. In addition to the doctor who helped us in the field of breast cancer despite being busy with work. We express a Special thanks to the FCI stuff for the valuable information provided by them in their respective fields Lastly, we thank Almighty, Our Parents and all the people who helped, supported and encouraged us to successfully finish the graduation project whether they were in the university or in the industry.

Table of Contents

List of figures

List of tables

Abbreviations

ER	Estrogen Receptor
PR	Progesterone Receptor
HER2	Human epidermal growth factor receptor
TNBC	Triple Negative Breast Cancer
SDLC	Software Development Life Cycle
MTBF	Mean Time Between Failures
DFD	Data Flow Diagram
ERD	Entity Relationship Diagram
ERM	Entity-Relationship Model
Ki67	Ki67 cell proliferation marker
HRT	Hormone Replacement Therapy
CAR-T	Chimeric Antigen Receptor T-cell
GS	GreenSock
MVC	Model-view-controller
ORM	Object-Relational Mapping

CHAPTER 1

Introduction

Breast cancer is a serious and complicated challenge within the world of women's health. The existence of numerous subtypes underscores the complexity of the subject, with each subtype having an impact on diagnostic and therapeutic decision-making. The identification and classification of breast cancer subtypes, including estrogen receptor positive (ER+), progesterone receptor positive (PR+), human epidermal growth factor receptor positive (HER2+), and triple-negative breast cancer (TNBC), play a pivotal role in enhancing comprehension and treatment strategies for this challenging illness. The estrogen receptor (ER) holds a prominent position in the field of breast cancer diagnosis, as it plays a crucial role in the development of most invasive breast carcinomas. The expression patterns of progesterone receptor (PR), when coupled with it, not only enhance diagnostic precision but also provide vital insights into the functional pathways that underlie the disease. The diagnostic and prognostic value of ER and PR is emphasized by their influence on disease aggressiveness and patient outcomes. Human epidermal growth factor receptor 2 (HER2), which represents a significant proportion of breast malignancies, plays a crucial role in guiding treatment interventions. The overexpression of HER2 at an early stage plays a crucial role in the progression of breast cancer, leading to unfavorable clinical outcomes and shorter periods of being free from the disease. The inclusion of the proliferation marker Ki67 provides an additional level of understanding, as it indicates the level of cancer aggressiveness, responsiveness to treatment, and likelihood of recurrence. The necessity to customize treatment approaches, namely in the domains of hormone therapy and anti-HER2 therapy, gives rise to the demand for molecular classification. In recent years, significant advancements in molecular research have led to the discovery of distinct genes and markers that are linked to different subtypes of breast cancer. This development holds great potential for the identification of prognostic biomarkers and therapeutic targets in the field of breast cancer treatment. The focus of this documentation is to examine the primary subtypes of breast cancer, namely luminal A, luminal B, HER2-positive, and TNBC. The aim is to develop a comprehensive understanding of their distinct characteristics and the potential impact they have on tailored treatment approaches.

1.1 Overview

The website is designed for doctors also can be used hospitals providing an accessible medical history for any patient anywhere anytime, it also enables medical status and enables doctors to access patient`s history and pharmacies which improves medical service and help doctors by suggesting a treatment plan suitable for the patient's condition. The patient history enables pharmacist to add medicines of his pharmacy and laboratory. Our website also used to replace manual, paper-based system. This new system to control the following information, patient information, doctor`s information and pharmacy`s information in a secured maintained database to reduce the time and resources currently required for such tasks. Integrates all the systems within the hospital allowing the doctors have instance to his patient’s medical data from other physician’s offices.

1.2 Problem Definition

The precise classification of breast cancer is crucial for the successful identification and treatment of this prevalent issue in women's health. Nevertheless, it is important to acknowledge that classical classification systems, despite their historical significance, possess notable constraints. These methodologies face difficulties in effectively delineating the various aspects of breast cancer, hence hindering our capacity to customize treatments according to the unique requirements of individual patients. The tendency to oversimplify breast cancer subtypes in traditional methods frequently results in overgeneralization and a lack of accuracy, hence impeding accurate diagnosis and the development of effective treatment plans. The failure to integrate molecular heterogeneity is a common oversight in the study of breast cancer, as conventional classifications often neglect to account for significant genetic and molecular differences. The predictive capacity of traditional classifications for treatment response is limited, which hinders the customization of medicines based on individual cancer features and the optimization of treatment outcomes.

1.3 Proposed Solutions

To address the limitations of traditional breast cancer classification systems and improve the customization of treatments, here are some proposed solutions:

1. Incorporation of Molecular Profiling: Implementing molecular profiling techniques, such as genomic analysis, gene expression profiling, and proteomics, can provide a more comprehensive understanding of the genetic and molecular heterogeneity of breast cancer. By integrating these molecular characteristics into classification systems, we can identify more specific subtypes and tailor treatment approaches accordingly.

2. Development of Subtyping Criteria: Refine the existing classification systems by establishing more precise subtyping criteria that consider both histological and molecular features of breast cancer. This would involve identifying specific biomarkers and genetic alterations associated with different subtypes and incorporating them into the classification framework. This approach would help to capture the diverse characteristics of breast cancer and guide treatment decisions.

3. Utilization of Machine Learning and Artificial Intelligence: Leverage the power of machine learning and artificial intelligence algorithms to analyze large-scale genomic and clinical data. By training models on comprehensive datasets, these algorithms can identify patterns, correlations, and predictive markers that may not be apparent through traditional methods. Such models can assist in classifying breast cancer subtypes and predicting treatment response, enabling more personalized treatment plans.

4. Integration of Longitudinal Data: Incorporate longitudinal data, including information from multiple time points throughout a patient's diagnosis and treatment journey. By considering how the tumor evolves and responds to

therapy over time, classification systems can better capture the dynamic nature of breast cancer and guide treatment adjustments as needed.

5. Collaboration and Data Sharing: Foster collaboration among researchers, clinicians, and institutions to share data and knowledge. By pooling resources and data from large patient cohorts, we can enhance the statistical power and generalizability of findings, leading to more robust classification systems. This collaborative approach would also enable the validation of new subtypes and facilitate the development of standardized treatment guidelines.

6. Clinical Trials and Validation Studies: Conduct prospective clinical trials and validation studies to assess the effectiveness of new classification systems and treatment strategies. By involving diverse patient populations, these studies can provide evidence-based insights into the clinical utility of refined classification approaches and their impact on patient outcomes.

1.4 Website Objectives

The cancer care insight website has a lot of objectives for doctors, laboratories and pharmacies that deal with the website.

For Doctors:

1. Doctors can access medical and analyzes history of their patients.
2. Doctors can communicate with pharmacies.
3. Through medical and analyzes history of each patient, doctors can check patients correctly.
4. Doctors can check the patient quickly in situations of emergency which affects the patient`s life.
5. Through the medical history of the patient, the doctor can decide the suitable medicines which don`t conflict with other diseases the patient may have.

For pharmacies:

1. Pharmacy can communicate with patient introducing drugs the patient needs.
2. Pharmacy can discuss the way to deliver drugs to the patient.
3. Through our system, pharmacies communicate with a lot of patients which increases profit.

1.5 Website Features

This section provides the main features that the proposed website presents to any patient and doctor.

The website helps: doctor can access his patient`s medical history which helps him to check patient correctly and can communicate with pharmacies to decide if he should recommend an alternative medicine for his patient. website also enables doctor to check the patient quickly in situations of emergency giving him drugs which don`t conflict with other diseases the patient may have. Also Suggests a treatment plan for the patient based on the data. Pharmacies can reach a lot of patients introducing the available drugs. our website maintains all of this information in an efficient and cost wise fashion.

1.6 Stakeholders

This System will help in health management represented in:

- Patients.
- Doctors.
- Pharmacies.
- Laboratories.

1.7 Preview the next chapters

In the upcoming chapters, several critical topics will be discussed.

Chapter 2: focuses on Project Management, covering the Development Methodology, specifically the SDLC Waterfall Model and its design. The sequential phases and application of the Waterfall model are explored, along with its advantages and disadvantages. The chapter also delves into the Software Development Plan, emphasizing project scheduling, and highlights the importance of Project Management Activity. Risk Identification, including Schedule Risk and Requirement Risk, is addressed, followed by a comprehensive Feasibility Study encompassing technical, operational, schedule, and economic feasibility.

Chapter 3: shifts the focus to Background and Literature Review, providing background information on breast cancer diagnosis and treatment. A thorough literature review is conducted, and surveys conducted in the field are explored. The chapter also delves into the technologies employed in the project, including front-end, back-end, and deep learning selection.

Chapter 4: delves into System Analysis, starting with Business Modeling and stakeholder requests. System Requirements, both functional and non-functional, are discussed, followed by an exploration of use case analysis and scenario descriptions. Activity Diagrams, Sequence Diagrams, and Class Diagrams are utilized to illustrate the system's workflow, interactions, and structure. Data modeling is also covered, including the Data Flow Diagram (DFD), Context Diagram, Entity Relationship Diagram (ERD), and Relational Database Diagram, providing insights into the system's data management and flow.

Chapter 5: we will explore the development and analysis of two AI models. The chapter covers data exploration, preprocessing, and visualization, followed by a detailed methodology and model architecture for each model. It includes steps like data acquisition, augmentation, balancing, and handling missing data, concluding

with the results and performance evaluation of the models. The chapter provides a thorough understanding of the AI model lifecycle, from data preparation to inference.

Chapter 6: we will delve into website design, focusing on challenges and solutions. We will identify common challenges in website design and present our solutions to these issues. Additionally, we will explore user interface design, emphasizing creating a user-friendly and visually appealing interface. This chapter provides insights into the practical aspects of effective website design.

Chapter 7: we will summarize the key findings and insights from the entire work in the conclusion section. Following this, we will discuss potential future work, outlining areas for further research and development. This chapter wraps up the current study and sets the stage for ongoing exploration and improvements.

1.8 Summary

The website is a digital platform for doctors and hospitals to access and maintain patients' medical history, enabling suggested treatment plans and integrating with pharmacies. Benefits include improved medical service, reduced time and resources, and increased transparency and accessibility. Features encompass a secured maintained database, patient-pharmacy-doctor information integration, and replacement of manual paper-based systems.

CHAPTER 2

Project Management

Project planning details all tasks which need to be done, by whom and when .it also details the high-level milestones which are key checkpoints on the project. The purpose of project planning is to begin to define parameters of the project and to establish the appropriate project management and quality environment required to complete the project. The major deliverable for this process is the project Initiation plan. Development of the project Initiation plan is a pivotal starting point for the project that will serve as the foundation for all future efforts. The completion of this process is marked by the sign off and approval of the project Initiation plan. Successful projects begin with a detailed project definition that is understood and accepted by stakeholders. Putting everything down in writing helps ensure a commitment among project Team members and between the team and the stockholders. As part of project planning, Project planning begins with requirements that define the software to be developed. The plan includes information related to staffing, budgets, deadlines, goals and measurements. Much like a business plan serves as a road map for how a small business operates. A project plan also should progress as well, monitoring the progress begin made, ensuring activates are taking place When they should be. The plan also serves as a tool for keeping everyone associated with the project on track and focusing on the same details and information.

2.1 Development Methodology

2.1.1 SDLC Waterfall Model

The Waterfall Model was first Process Model to be introduced. It is also referred to as a linear sequential life cycle model. It is very simple to understand and use. In a waterfall model, each Phase must be completed before the next phase can begin and there is no overlapping in the phases. Waterfall model is the earliest SDLC approach that was used for software development. The waterfall Model illustrates the software development process in a linear sequential flow; hence it is also referred to as a linear-sequential life cycle model. This means that any phase in the

Development process begins only if the previous phase is complete. In waterfall model phases do not overlap.

2.1.2 Waterfall Model Design

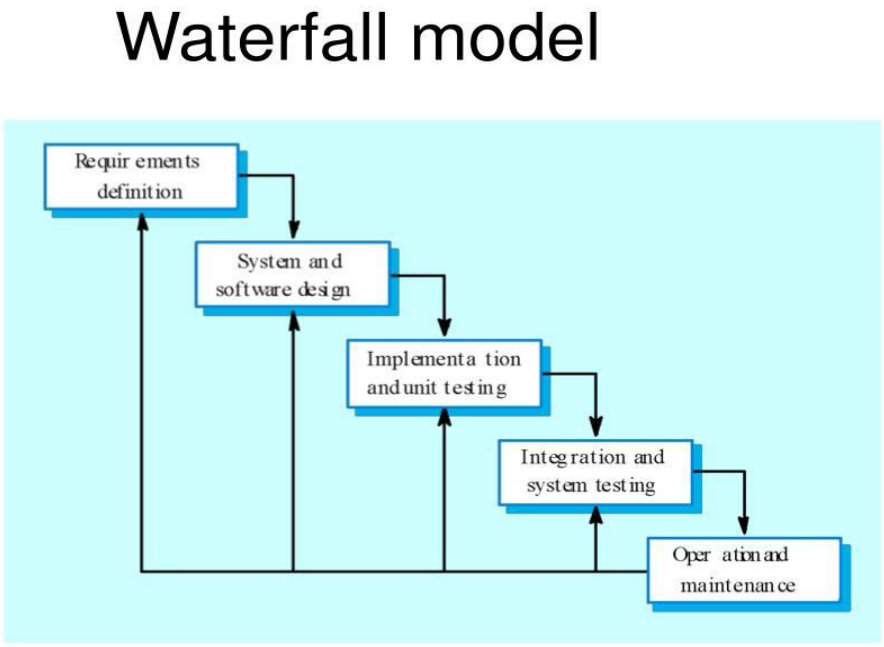

Figure 2-1 Waterfall model

Waterfall approach was first SDLC Model to be used widely in Software Engineering to ensure success of the project. In "The Waterfall" approach, the whole process of software development is divided into separate phases. In Waterfall model, typically, the outcome of one phase acts as the input for the next phase sequentially.

Following is a diagrammatic representation of different phases of waterfall model.

2.1.3 The sequential phases in Waterfall model are

Requirement Gathering and analysis: All possible requirements of the system to be Developed are captured in this phase and documented in a requirement specification doc.

System Design: The requirement specifications from first phase are studied in this phase and system design is prepared. System Design helps in specifying hardware and system Requirements and also helps in defining overall system architecture.

Implementation: With inputs from system design, the system is first developed in small programs called units, which are integrated in the next phase. Each unit is developed and tested for its functionality which is referred to as Unit Testing

Integration and Testing: All the units developed in the implementation phase are Integrated into a system after testing of each unit. Post integration the entire system is tested for any faults and failures.

Deployment of system: Once the functional and non-functional testing is done, the product is deployed in the customer environment or released into the market.

Maintenance: There are some issues which come up in the client environment. To fix those issues patches are released. also, to enhance the product some better versions are released. Maintenance is done to deliver these changes in the customer environment

All these phases are cascaded to each other in which progress is seen as flowing steadily downwards like waterfall through the phases. The next phase is started only after the defined set of Goals are achieved for previous phase and it is signed off, so the name "Waterfall Model". In this Model phases do not overlap.

2.1.4 Waterfall Model Application

Every software developed is different and requires a suitable SDLC approach to be followed based on the internal and external factors. Some situations where the use of Waterfall model is most appropriate are:

- Requirements are very well documented, clear and fixed.
- Product definition is stable.
- Technology is understood and is not dynamic.
- There are no ambiguous requirements.
- Ample resources with required expertise are available to support the product.
- The project is short.

2.1.5 Waterfall Advantages

The advantage of waterfall development is that it allows for departmentalization and control. A Schedule can be set with deadlines for each stage of development and a product can proceed through the development process model phases one by one. Development moves from concept, through design, implementation, testing, installation, Troubleshooting, and ends up at operation and maintenance.

Each phase of development Proceeds in strict order.

2.1.6 Waterfall Disadvantages

The disadvantage of waterfall development is that it does not allow for much reflection or revision. Once an application is in the testing stage, it is very difficult to go back and change something that was not well-documented or thought upon in the concept stage.

2.2 Software Development Plan

project scheduling

The following graphs contain a high-level schedule of some significant milestones for this project

Task Mode	Task Name	Duration	Start	Finish	Resource Names
	software development plan	**227 days**	**Tue 8/1/23 8:00 AM**	**Tue 6/11/24 5:00 PM**	
	planing	**13 days**	**Sat 9/16/23 8:00 AM**	**Tue 10/3/23 5:00 PM**	
	Define Project Goals and Scope	10 days	Sat 9/16/23 8:00 AM	Thu 9/28/23 5:00 PM	Mariam Samuel
	Risk Assessment	10 days	Tue 9/26/23 8:00 AM	Mon 10/9/23 5:00 PM	Youstina William
	divide project into tasks	10 days	Thu 10/5/23 8:00 AM	Wed 10/18/23 5:00 PM	Aya Ahmed
	anlysis,software requirements	**32 days**	**Wed 10/4/23 8:00 AM**	**Thu 11/16/23 5:00 PM**	
	Identify Data Sources	5 days	Tue 10/3/23 5:00 PM	Tue 10/10/23 5:00 PM	Sohaila Elbadry
	requirement gathering	**20 days**	**Fri 10/13/23 8:00 AM**	**Thu 11/9/23 5:00 PM**	
	functional reqirement	3 days	Fri 10/13/23 8:00 AM	Tue 10/17/23 5:00 PM	Esraa Khamis
	non functional requirement	3 days	Tue 10/17/23 8:00 AM	Thu 10/19/23 5:00 PM	Sohaila Elbadry
	data cleaning and intergration	5 days	Thu 11/2/23 8:00 AM	Wed 11/8/23 5:00 PM	Yasmin Gamal
	design digrams	10 days	Sat 11/11/23 8:00 AM	Thu 11/23/23 5:00 PM	Esraa Khamis, Mariam Samuel, Yasmin Gam
	design	**46 days**	**Thu 11/9/23 8:00 AM**	**Thu 1/11/24 5:00 PM**	
	system architecture	5 days	Thu 11/9/23 8:00 AM	Wed 11/15/23 5:00 PM	Sohaila Elbadry
	database design	15 days	Fri 12/1/23 8:00 AM	Thu 12/21/23 5:00 PM	Esraa Khamis
	user interface design	10 days	Tue 12/26/23 8:00 AM	Mon 1/8/24 5:00 PM	Mariam Samuel, Yasmin Gamal
	Design Diagrams	**4 days**	**Tue 12/26/23 8:00 AM**	**Fri 12/29/23 5:00 PM**	
	Use case analysis	3 days	Mon 1/1/24 8:00 AM	Wed 1/3/24 5:00 PM	Youstina William
	activity diagram	3 days	Thu 1/4/24 8:00 AM	Mon 1/8/24 5:00 PM	Mariam Samuel
	sequence diagram	3 days	Tue 1/9/24 8:00 AM	Thu 1/11/24 5:00 PM	Yasmin Gamal
	system sequence diagram	3 days	Fri 1/12/24 8:00 AM	Tue 1/16/24 5:00 PM	Yasmin Gamal
	class diagram	3 days	Wed 1/17/24 8:00 AM	Fri 1/19/24 5:00 PM	Yasmin Gamal
	data flow diagram	3 days	Mon 1/22/24 8:00 AM	Wed 1/24/24 5:00 PM	Sohaila Elbadry
	context diagram	3 days	Thu 1/25/24 8:00 AM	Mon 1/29/24 5:00 PM	Sohaila Elbadry
	relational database diagram	3 days	Thu 1/25/24 8:00 AM	Mon 1/29/24 5:00 PM	Esraa Khamis
	development	**82 days**	**Thu 1/11/24 8:00 AM**	**Fri 5/3/24 5:00 PM**	
		0 days	Thu 1/11/24 8:00 AM	Thu 1/11/24 8:00 AM	
	algorithm development	20 days	Thu 1/25/24 8:00 AM	Wed 2/21/24 5:00 PM	Aya Ahmed
	user interface development	20 days	Thu 2/22/24 8:00 AM	Wed 3/20/24 5:00 PM	Yasmin Gamal, Mariam Samuel
	develop code and unit testing	10 days	Fri 3/22/24 8:00 AM	Thu 4/4/24 5:00 PM	Sohaila Elbadry, Aya Ahmed
	testing	**20 days**	**Fri 5/3/24 8:00 AM**	**Thu 5/30/24 5:00 PM**	
	unit testing	10 days	Fri 5/3/24 8:00 AM	Thu 5/16/24 5:00 PM	Yasmin Gamal
	integration testing	10 days	Fri 5/17/24 8:00 AM	Thu 5/30/24 5:00 PM	Esraa Khamis
	validation testing	10 days	Wed 5/29/24 8:00 AM	Tue 6/11/24 5:00 PM	Mariam Samuel
	docunmentation	**6 days**	**Thu 5/30/24 8:00 AM**	**Thu 6/6/24 5:00 PM**	**Aya Ahmed, Youstina William**
	Meeting	**86 days**	**Tue 12/12/23 8:00 AM**	**Tue 4/9/24 5:00 PM**	**Dr/Mary Monir**

Figure 2-2 Scheduling

2.3 Project Management Activity

It is the formulation and completion of a defined set of goals. This is a highly complex process, and one that calls for good organizational skills and well-thought-out plan of action to follow.

Due to the multifaceted nature of projects, there are many activities that fall under the umbrella term “project management.” To oversee a project from its initial stages through to its completion requires a number of skills, including the ability to manage resources, set dates, and facilitate communication.

The following is a list of common project-management related activities: Setting goals within the overall framework of the project and ensuring that they are complete on time and in a satisfactory way.

1. Establishing timetables for the project and its various subtasks.

2. Monitoring the use of time for maximum efficiency.

3. Estimating the resources, both material and human, required by the project and ensuring that they are distributed and used properly.

4. Identifying potential risks to the project. Developing a risk management plan to deal with unfavorable contingencies.

5. Organizing relevant documents and records so that they may be conveniently consulted by those working on the project.

6. Analyzing the current conditions of the project and predicting future trends so as not to be caught off-guard by changes.

7. Assigning short-term tasks to specific groups or individuals and recording the progress made toward their completion.

8. Managing any issues that affect the project; keeping an issue log that organizes issues by type and priority.

9. Managing the quality of the products of the project. Providing adequate quality control and quality assurance. Finding ways to improve quality.

10. Facilitating communication among project members and between the project members and outside stakeholders.

2.4 Risk Identification

Project Risk Management includes the processes of conducting risk management planning, identification, analysis, response planning, and controlling risk on a project. The objectives of project risk management are to increase the likelihood and impact of positive events, and decrease the likelihood and impact of negative events in the project. Project Risk identification is the most important process in the Risk Management Planning. Risk Identification determines which risks might affect the project and documents their characteristics. However, as recommended by [Donna Ritter], we should not spend too much time in identifying risks. After the list is made, qualitative and quantitative analysis is done to figure out which risks you spend time and/or money on.

The main risks types facing development of this system are the follows:

2.4.1 Schedule Risk

- Schedule not realistic, only "best case".
- Important task missing from the schedule.
- A delay in one task causes cascading delays in dependent tasks.
- Unfamiliar areas of the product take more time than expected to design and implement.

2.4.2 Requirement Risk

- Requirements have been base lined but continue to change.
- Requirements are poorly defined, and further definition expands the scope of the project.
- Specified areas of the product are more time-consuming than expected.
- Requirements are only partly known at project start.
- The total features requested may be beyond what the development team can deliver in the time available.

2.4.3 Comprehensive Project Risk Management

Risk Identification in the project is critical in order to manage and complete the project successfully. The earlier the risk can be identified, the earlier the plan can be made to mitigate the effects of the potential risks. There are a lot of tools and techniques or method available to identify the project risks.

The method suggested in this article will complement the existing risk identification method to get a more comprehensive risk list for Risk Management Planning. Identifying the risk is an iterative process, and the entire project team should be involved from the beginning of the project. Comprehensive and good risk identification will produce a good project result.

2.5 Feasibility Study

Feasibility Study is a test of system proposal according to its workability, impact of the organization, ability to meet needs and effective use of the resources. It focuses on these major questions:

1. What are the user's demonstrable needs and how does a candidate system meet them?

2. What resources are available for given candidate system?

3. What are the likely impacts of the candidate system on the organization?

4. Whether it is worth to solve the problem? During feasibility analysis for this project, following primary areas of interest are to be considered. Investigation and generating ideas about a new system does this.

2.5.1 Technical Feasibility

A study of resource availability that may affect the ability to achieve an acceptable system. This evaluation determines whether the technology needed for the proposed system is available or not.

- Can the work for the project be done with current equipment existing software? Technology & available personal?

- Can the system be upgraded if developed?

- If new technology is needed then what can be developed? This is concerned with specifying equipment and software that will successfully satisfy the user requirement.

An important issue for the development of a project is the selection of suitable front-end and back-end. When we decided to develop the project, we went through an extensive study to determine the most suitable platform that suits the needs of the organization as well as helps in development of the project.

2.5.2 Operational Feasibility

It is mainly related to human organizations and political aspects. The points to be considered are:

- What changes will be brought with the system?
- What organization structures are disturbed?
- What new skills will be required? Do the existing staff members have these skills?

If not, can they be trained in due course of time? The system is operationally feasible as it very easy for the End users to operate it.

2.5.3 Schedule Feasibility

Time evaluation is the most important consideration in the development of project. The time schedule required for the developed of this project is very important since more development time effect machine time, cost and cause delay in the development of other systems. A reliable Cancer Care Insight Project can be developed in the considerable amount of time.

2.5.4 Economic Feasibility

Economic justification is generally the “Bottom Line” consideration for most systems. Economic justification includes a broad range of concerns that includes cost benefit analysis. In this we weight the cost and the benefits associated with the candidate system and if it suits the basic purpose of the organization i.e. profit making, the project is making to the analysis and design phase. The financial and the economic questions during the preliminary investigation are verified to estimate the following:

-The cost to conduct a full system investigation.

-The cost of hardware and software for the class of application being considered.

-The benefits in the form of reduced cost.

-The proposed system will give the minute information as a result,

-The performance is improved which in turn may be expected to provide increased profits.

-This feasibility checks whether the system can be developed with the available funds.

-The Cancer Care Insight does not require enormous amount of money to be developed.

-This can be done economically if planned judicially, so it is economically feasible.

-The cost of project depends upon the number of man-hours required.

2.6 Summary

Project management is the application of knowledge, skills, tools, and techniques applied to project activities in order to meet the project requirements. Project management is a process that includes planning, putting the project plan into action, and measuring progress and performance.

Managing a project includes identifying your project's requirements and writing down what everyone needs from the project. What are the objectives for your project? When everyone understands the goal, it's much easier to keep them all on the right path.

Make sure you set goals that everyone agrees on to avoid team conflicts later on. Understanding and addressing the needs of everyone affected by the project means the end result of your project is far more likely to satisfy your stakeholders.

Last but not least, as project manager, you will also be balancing the many competing project constraints. On any project, you will have a number of project constraints that are competing for your attention. They are cost, scope, quality, risk, resources, and time.

CHAPTER 3

Background and Literature review

3.1 Background information on breast cancer diagnosis and treatment

Breast cancer is a significant health concern affecting millions of women worldwide. Early detection remains paramount in effectively managing the disease and improving patient outcomes. Screening methods such as mammography, clinical breast exams, and self-exams play crucial roles in identifying breast abnormalities at their earliest stages, enabling timely intervention and treatment. However, the interpretation of screening results requires expertise, and additional diagnostic techniques, including imaging studies like mammograms, ultrasounds, and MRIs, are often employed to confirm or rule out a breast cancer diagnosis.

Once diagnosed, breast cancer treatment typically involves a multidisciplinary approach, with healthcare professionals from various specialties collaborating to develop personalized treatment plans tailored to each patient's specific circumstances. Surgical options, such as lumpectomy or mastectomy, may be recommended based on the size and location of the tumor, along with the patient's preferences and overall health. Advances in surgical techniques, including nipple-sparing mastectomy and oncoplastic surgery, aim to optimize cosmetic outcomes while effectively removing cancerous tissue.

In addition to surgery, adjuvant therapies such as chemotherapy, radiation therapy, hormone therapy, targeted therapy, and immunotherapy may be employed to reduce the risk of cancer recurrence and improve long-term survival rates. These treatment modalities are often used in combination, depending on the tumor's characteristics and the patient's individual risk factors. Targeted therapies, which specifically target cancer cells while sparing healthy tissue, have revolutionized breast cancer treatment, particularly for HER2-positive and hormone receptor-positive breast cancers.

The journey through breast cancer diagnosis and treatment can be physically and emotionally challenging, requiring ongoing support and care. Survivorship extends beyond treatment completion and encompasses ongoing surveillance, management of treatment-related side effects, and psychosocial support to address the holistic needs of breast cancer survivors. Through advances in early detection,

personalized treatment approaches, and supportive care, significant progress has been made in improving outcomes and quality of life for individuals affected by breast cancer.

3.2 Literature Review

The literature review demonstrates the valuable role of the website as a comprehensive resource for doctors specializing in breast cancer. By providing access to the latest scientific articles, clinical guidelines, and case studies, the website enables physicians to stay updated on cutting-edge advancements in the field. This wealth of knowledge enhances their understanding and practice, ultimately leading to improved outcomes for breast cancer patients.

The website also fosters collaboration among medical professionals, serving as a platform for sharing insights and experiences. This collaborative approach promotes a collective effort to improve patient care and treatment strategies. The user-friendly interface and personalized features of the website make it a valuable tool in supporting doctors in their work.

In the context of breast cancer diagnosis, the literature review highlights significant findings. Studies have shown that convolutional neural networks (CNNs) outperform multilayer perceptron (MLP) neural networks in breast cancer classification. Transfer learning techniques using pretrained models such as GoogLeNet, VGGNet, and ResNet have been proposed to classify malignant and benign cells with high accuracy.

To address overfitting with small datasets, deep CNN models incorporating transfer learning have been developed, achieving impressive results in terms of accuracy and area under the curve (AUC) on multiple datasets. Various CNN architectures and hybrid models have also been explored, demonstrating their efficacy in breast cancer classification.

Moreover, researchers have focused on optimizing the performance of CNN models through hyperparameter optimization and fine-tuning techniques. These approaches have shown promise in improving classification accuracy.

In this context, the proposed optimized CNN model based on the improved MPA algorithm aims to further enhance breast cancer classification. By leveraging the benefits of hyperparameter optimization, this model has the potential to support health professionals in accurate diagnosis and contribute to the ongoing efforts in breast cancer research and treatment.

We read a lot of papers to understand how to build a Deep learning model in breast cancer detection so we could know what type of data we should use in our model and we used mammogram images in our first model.

A new computer aided detection (CAD) system for classifying benign and malignant breast tumors is proposed, utilizing deep learning and segmentation techniques. The system employs manual and threshold-based segmentation approaches, with a fine-tuned Alex Net deep convolutional neural network (DCNN) for feature extraction. Achieving an accuracy of 73.6% on the CBIS-DDSM dataset, the system combines the DCNN with a support vector machine (SVM) classifier, yielding an 87.2% accuracy and the highest area under the curve (AUC) of 0.94 compared to previous studies [6].

Deep learning (DL) shows promise in medical applications, including breast abnormality diagnosis. Using the CBIS-DDSM dataset, CNN architectures were trained to classify mammograms into mass and calcification (88% accuracy) and further into benign or malignant (66% accuracy due to limited data). The VGG16 model achieved higher accuracy (90%) in the first task, highlighting its potential in breast image analysis. Improvement in the second task's accuracy requires a larger mammogram dataset [1].

Gene expression-based signatures and subtyping systems are important tools for stratifying breast cancer and guiding treatment decisions. This study highlights the existing signatures and subtyping systems for estrogen receptor-positive (ER+)

and triple-negative breast cancers (TNBC). However, the lack of consensus among these systems hinders their clinical translation. To address this issue, the study draws inspiration from the colorectal cancer community's approach and proposes the Breast Cancer Consensus Subtypes (BCCS) by integrating information from multiple gene expression platforms. The BCCS aims to provide a comprehensive and unbiased subtype classification for breast cancer. The study analyzes biological and clinical features of the subtypes across different cohorts, emphasizing the need for improved subtyping methods to guide treatment decisions in non-TNBC and HER2+ breast cancers [3].

Breast cancer is a significant global health concern and one of the leading causes of death. The use of DNA microarray technology, which examines gene expression levels, has shown great potential in aiding the early and effective detection of breast cancer. Deep learning algorithms have proven effective in various domains and are also being applied to diagnose chronic diseases and assist medical decision-making. In a recent study, different optimizers like SGD, RMSProp, AdaGrad, and AdaM were evaluated on an RNA sequence dataset to classify various forms of cancer, including breast invasive carcinoma (BRCA). The experimental results indicated that AdaGrad and AdaM performed well. The study highlights the importance of optimized feature selection methods in deep learning-based analysis of gene expression data for breast cancer detection [4].

In our model we used mammogram images as our dataset and our accuracy is (96%).Then we figured that’s not enough to distinguish our model so we searched what else we can do to help doctors in their work and facilitate also the process of choosing what kind of treatment plan should every patient go throw depending on many factors, such as the patient’s gender, whether he has a history of illness, or whether he has chronic diseases, this determines whether it will be hormonal treatment, chemotherapy, or anything else.

3.3 Dataset

We are fortunate to have access to two distinct datasets that hold immense value in the field of medical research and diagnosis. The first dataset, known as the Breast Histopathology Images dataset, is hosted on the National Center for Biotechnology Information. This exceptional collection of mammography images provides an opportunity to analyze and determine whether a tumor is benign or malignant. By scrutinizing these histopathology images, medical professionals can gain crucial insights into the nature and characteristics of breast tumors, aiding in accurate diagnoses and subsequent treatment decisions.

The second dataset comprises numerical data, conveniently stored in an Excel sheet, that represents the Histopathology test results of various patients. This dataset holds vital information such as ER (Estrogen Receptor), PR (Progesterone Receptor), HER2 (Human Epidermal Growth Factor Receptor 2), and Ki67 (%). These values play a pivotal role in determining the appropriate type of treatment and formulating an effective treatment plan for patients. The ER and PR values help gauge hormone receptor status, which is crucial for determining the suitability of hormone therapy. The HER2 value indicates the presence or absence of a specific protein that influences treatment decisions, and the Ki67 percentage serves as a measure of cell proliferation, aiding in the assessment of tumor aggressiveness.

Both datasets represent invaluable resources in the fight against breast cancer. The integration of mammography images and numerical test results fosters a comprehensive approach to diagnosis and treatment planning. By leveraging these datasets, medical professionals can enhance their understanding of breast cancer, refine diagnostic accuracy, and tailor personalized treatment strategies for patients. The synergy between these datasets presents an exciting avenue for further research and advancements in the field of breast cancer management.

3.4 Used Technologies

3.4.1 Front-end

1. It must have a graphical user interface that assists employees that are not from IT background.

2. Scalability and extensibility.

3. Flexibility.

4. Robustness.

5. According to the organization requirement and the culture.

6. Must provide excellent reporting features with good printing support.

7. Platform independent.

8. Easy to debug and maintain.

9. Event driven programming facility.

10. Front end must support some popular back.

We used React GS which is a library that integrates the GreenSock (GS) animation library with React, allowing you to create animations in your React applications. You can install GS with npm or yarn, and then import it and React in your component file. By using the useRef hook and the gsap.to function, you can animate elements in your React components.

According to the above stated features we selected Web design as the front-end for developing our project.

3.4.2 Back-end

1. Multiple user support.
2. Efficient data handling.
3. Provide inherent features for security.
4. Efficient data retrieval and maintenance.
5. Stored procedures.
6. Popularity.
7. Operating System compatible.
8. Easy to install.
9. Various drivers must be available.
10. Easy to implant with the Front-end.

We used PHP Laravel which is a popular open-source PHP web application framework that follows the model-view-controller (MVC) architectural pattern. It is known for its elegant syntax, expressive language, and comprehensive set of tools for building web applications. Laravel provides a simple and elegant way to build web applications, with features such as routing, middleware, and dependency injection. It also includes an ORM (Object-Relational Mapping) system for working with databases, and a powerful templating engine for generating HTML.

We also used Flask which is a lightweight web framework for building web applications in Python. It is based on the Werkzeug WSGI toolkit and Jinja2 template engine, and is designed to make it easy to build web applications quickly and easily. Flask is a microframework, which means that it does not include an ORM or other features that are not strictly necessary for building a web application.

The technical feasibility is frequently the most difficult area encountered at this stage. It is essential that the process of analysis and definition be conducted in parallel with an assessment to technical feasibility. It centers on the existing

computer system (Hardware, software and etc.) and to what extent it can support the proposed system.

3.4.3 Deep Learning selection

1. Define the problem and task.
2. Collect and prepare high-quality data.
3. Choose an appropriate deep learning architecture.
4. Design the model with suitable hyperparameters.
5. Train the model using a training set.
6. Implement regularization techniques to prevent overfitting.
7. Evaluate the model using appropriate metrics on a testing set.
8. Consider fine-tuning or transfer learning for better performance.
9. Deploy the model and monitor its performance.
10. Ensure scalability, flexibility, and user-friendliness.

Our data was a mamo images and there first there was image processing to handle our data before building our model we used in Deep learning, CNN (Conventional Neural Network), RNNs (Recurrent Neural Networks), DNA microarray (Gene Expression).

These techniques work together to make our model a successful one with accuracy of 96% .

3.5 Summary

Breast cancer is a major health concern affecting millions of women worldwide, and early detection is crucial for effective management and improving patient outcomes. Screening methods like mammography, clinical breast exams, and self-exams play crucial roles in identifying breast abnormalities at their earliest stages. However, interpretation of screening results requires expertise, and additional diagnostic techniques are often employed to confirm or rule out a breast cancer diagnosis.

Once diagnosed, breast cancer treatment typically involves a multidisciplinary approach, with healthcare professionals from various specialties collaborating to develop personalized treatment plans tailored to each patient's specific circumstances. Surgical options, such as lumpectomy or mastectomy, may be recommended based on the size and location of the tumor, along with the patient's preferences and overall health. Advances in surgical techniques, including nipple-sparing mastectomy and oncoplastic surgery, aim to optimize cosmetic outcomes while effectively removing cancerous tissue.

Adjuvant therapies, such as chemotherapy, radiation therapy, hormone therapy, targeted therapy, and immunotherapy, may be employed to reduce the risk of cancer recurrence and improve long-term survival rates. These treatment modalities are often used in combination, depending on the tumor's characteristics and the patient's individual risk factors.

The journey through breast cancer diagnosis and treatment can be physically and emotionally challenging, requiring ongoing support and care. Through advances in early detection, personalized treatment approaches, and supportive care, significant progress has been made in improving outcomes and quality of life for individuals affected by breast cancer.

CHAPTER 4

System Analysis

Systems analysis is the study of sets of interacting entities, including computer systems analysis. This field is closely related to requirements analysis or operations research. It's also " an explicit formal inquiring carried out to help someone [referred to as the decision maker identify a better course of action and make a better decision than he might otherwise have made ".

In this chapter we will present "patient recording system" system analysis including [functional requirements, nonfunctional requirements, use case diagrams and sequence diagram].

4.1 Business Modeling

4.1.1 Stakeholder Request

Breast cancer remains a significant health issue across the globe and has warranted an advanced and efficient means of early detection of the condition and personalized treatment planning. The traditional methods such as that in cancer institutes and general hospitals are often plagued with inefficiencies that lead to overcrowding, delays in conducting examination and trouble in providing proper alternatives for treatments. Our web-based system identifies such challenges and comes as a revolutionary solution that utilizes artificial intelligence in order to facilitate the process of detecting breast cancer and advancing treatment methodology.

Manual processes, paper-based records, as well as the limited real-time data analyses are challenges positioned against the conventional approaches to the detection and treatment management of breast cancer. Accumulative congestion and delay experienced in the process of conducting examinations may compromise the timeliness of diagnosis to the potential detriment of patient outcomes. Additionally, the absence of such a critical data-driven system curtails maximum personalized intervention treatments according to individual patient profiles for averting minimum success rates as mentioned.

4.1.2 Purpose

1. Efficiency increasing and speed in the detection of breast cancer.
2. Decongest health facilities and fast-tracking examination process.
3. Accuracy increasing and personalized treatment plan for patients detected with breast cancer.
4. Improve security for patient information and following them up through their treatment process.
5. Provide health care professionals with ideas of which treatment options are more suitable.
6. Gather post-treatment information helpful for further research and assessment.
7. How such data can be analyzed by tailoring the service towards a change in improving the service of breast cancer treatment.
8. contribute to the advancement of medical research in the field of breast cancer.

4.1.3 Scope

The project scope is doctor, clinics, data entry and admin. This project is about patient recording system web application; this application will consist of three important actors. Which will be like this: -

1. **Doctor**
   1) Doctors can access medical history of their patients.
   2) Through medical history of each patient, doctors can check patients correctly.
   3) Doctors can check the patient quickly in situations of emergency which affects the patient`s life.
   4) Through the medical history of the patient, the doctor can decide the suitable medicines which don`t conflict with other diseases the patient may have.
   5) Doctor monitors the patient's progress.
   6) Records observations in the group.

2. **Data entry personnel**
   1) Input and access comprehensive patient information.
   2) Review, edit, or delete entries as needed.
   3) Send messages to the patients.

3. **Clinical staff**

Statistical ratios related to patients' number and the types of treatment that will be utilized.

4. **Administrator**

Responsible for managing user access on the website.

4.1.4 System Overview

1. **Assessing the Problem**

Identification of Challenges: Traditional breast cancer detection methods are inefficient and prone to delays, leading to congestion in healthcare facilities. Treatment planning lacks personalization, impacting patient outcomes. Paper-based records result in data inefficiency and potential loss.

2. **Recap for Understanding**

Current Scenario: Existing systems face challenges in timely detection, personalized treatment planning, and efficient data management, hindering the overall quality of breast cancer care.

3. **Assessing Your Solution (if applicable)**

Developing software will reduce time and effort.

Software will detect breast cancer, patient treatment planning and maintain patient information in a proper way.

4.2 System Requirements

The medical system utilizes advanced machine learning and image processing for efficient healthcare management. Its dual-model architecture ensures swift and precise detection of medical conditions, optimizing time and resources. The first model analyzes X-ray images for rapid identification of conditions, enhancing diagnostic speed and resource efficiency. The system excels in secure data management, storing patient information and X-ray results. The second model interprets laboratory test results, offering insights into personalized treatment plans, with recording and tracking features for enhanced understanding of treatment distribution and valuable statistical data on medical cases.

4.2.1 Functional Requirements

In software engineering, a functional requirement defines a function of a software system or its component. A function is defined as a set of inputs, the behavior and output. Functional requirements may be calculations, technical details, data manipulation and processing and other specific functionality that define what the system is supposed to accomplish. Behavioral requirements describing all the cases where the system uses the functional requirements are captured in use cases.

Our system`s main functions represented as follow: -

Data entry personnel

1. Input and access comprehensive patient information.
2. Review, edit, or delete entries as needed.
3. Send messages to the patients.

Doctor

1. Read information of patient
2. Add notes

Administrator

Responsible for managing user access on the website.

Clinical staff

Read statistical ratios related to patients' number and the types of treatment that will be utilized.

4.2.2 Non-Functional Requirements

There are other requirements that don't actually DO anything, but are important characteristics nevertheless. These are called "non-functional requirements" or sometimes "Quality Attributes." For example, attributes such as performance, security, usability, compatibility. Aren't a "feature" of the system, but are a required characteristic. You can't write a specific line of code to implement them, rather they are "emergent" properties that arise from the entire solution. The specification needs to describe any such attributes the customer requires. You must

decide the kind of requirements that apply to your project and include those that are appropriate.

Non-functional requirements cover all the remaining requirements which are not covered by the functional requirements. They specify criteria that judge the operation of a system, rather than specific behaviors

Non-functional requirements specify the system's 'quality characteristics' or 'quality attributes.

-Each requirement is simply stated in English. Each requirement must be objective and quantifiable;

-There must be some measurable way to assess whether the requirement has been met.

-Often deciding on quality attributes requires making tradeoffs, e.g., between performance and maintainability. In the APPENDIX you must include an engineering analysis of any significant decisions regarding tradeoffs between competing attributes.

Here are some examples of non-functional requirements:

•Performance requirements

Requirements about resources required, response time, transaction rates, throughput, benchmark specifications or anything else having to do with performance.

Response Time: The system shall give responses in 1 second after the checking patient's information

Capacity: based on organization server.

• Operating constraints

List any run-time constraints. This could include system resources, people, needed software.

• Platform constraints

Discuss the target platform. Be as specific or general as the user requires. If the user doesn't care, there are still platform constraints.

• Accuracy and Precision

Requirements about the accuracy and precision of the data.

(Do you know the difference?) Beware of 100% requirements; they often cost too much.

• Modifiability

Requirements about the effort required to make changes in the software.

Often, the measurement is personnel effort (person- months).

Any modification (insert, delete, and update) for the Database shall be synchronized and done only by the administrator in the ward.

• Portability

The effort required to move the software to a different target platform.

The measurement is most commonly person-months or % of modules that need changing.

• Reliability

Requirements about how often the software fails. The measurement is often expressed in MTBF (mean time between failures). The definition of a failure must be clear. Also, don't confuse reliability with availability which is quite a different kind of requirement.

Be sure to specify the consequences of software failure, how to protect from failure, a strategy for error detection, and a strategy for correction.

Availability: The system shall be available all the time.

• Security

One or more requirements about protection of your system and its data. The measurement can be expressed in a variety of ways (effort, skill level, and time,) to break into the system. Do not discuss solutions (e.g., passwords) in a requirements document.

• Usability

Requirements about how difficult it will be to learn and operate the system.

The requirements are often expressed in learning time or similar metrics.

User-interface: The user-interface screen shall respond within 5 seconds.

• Maintainability

Back Up: The system shall provide the capability to back-up the Data.

Errors: The system shall keep a log of all the errors.

• Legal

There may be legal issues involving privacy of information, intellectual property rights, export of restricted technologies, etc.

4.3 Use case Analysis

A use case analysis is the most common technique used to identify the requirements of a system [normally associated with software | process design] and the information used to both define processes used and classes [which are a collection of actors and processes] which will be used both in the use case diagram and the overall use case in the development or redesign of a software system or program. The use case analysis is the foundation upon which the system will be built, as it provides a comprehensive understanding of the system's requirements, processes, and actors, guiding the development or redesign process and ensuring that the resulting system aligns with the intended functionality and goals.

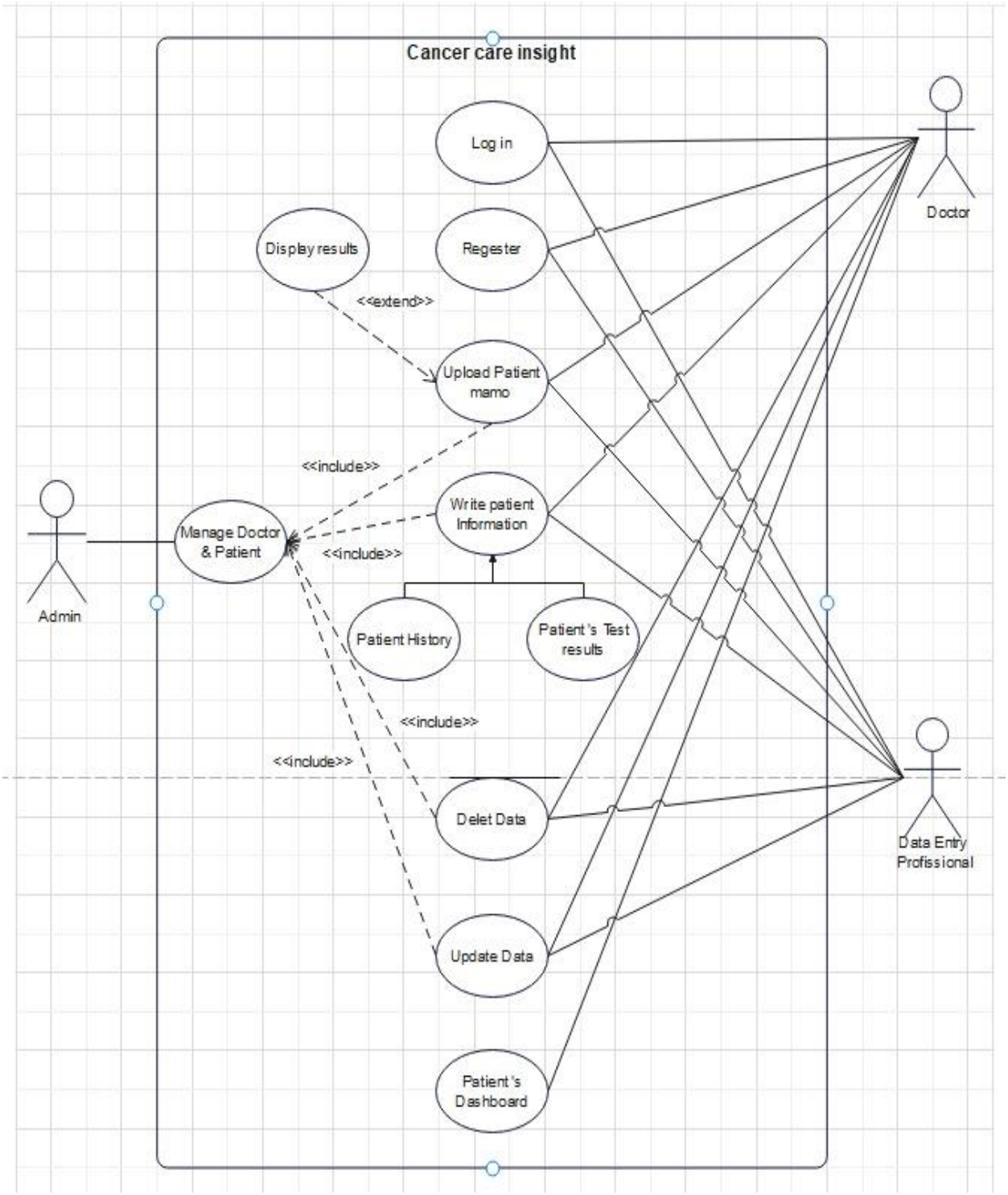

Figure 4-1 Use case

4.3.1 Use Case Scenario

Table 4-1 "Access patient medical history" use case scenario

Use case name	Access patient medical history.
Actor	Data entry professional, Doctor, Admin.
Stakeholders and Interests	Doctor.
Pre-condition	Login to system.
Post-condition	Doctor, Admin can update patient medical history.
Main success scenario	Doctor entered patient SSN/ID.

Table 4-2 "Register a personal account" use case scenario

Use case name	Register a personal account.
Actor	Data entry professional, Doctor.
Stakeholders and Interests	Data entry professional, Doctor.
Pre-condition	Confirm the registration from the activation message.
Main success scenario	The actor entered his, email address, name, SSN/ID, password…etc.
Extensions	The actor entered invalid email address or SSN/ID the system asks the user to re-enter the email address or SSN/ID and so on.

Table 4-3 "Login" uses case scenario

Use case name	Login.
Actor	Data entry professional, Doctor.
Stakeholders and Interests	Data entry professional, Doctor.
Pre-condition	Every actor registered in the system.
Post-condition	Actors entered system URL.
Main success scenario	The actor entered the right username and password and entered the system.
Extensions	The actor entered invalid username or password the system asks the user to re-enter the username or password.

Table 4-4 "Update patient medical history" use case scenario

Use case name	Update patient medical history.
Actor	Data entry professional, Doctor, Admin.
Stakeholders and Interests	Doctor, Patient.
Pre-condition	Login the system.
Main success scenario	Doctor or Data entry professional or Admin access patient medical history and enter his updates.

Table 4-5 "Results of patient ‘s Dashboard" use case scenario

Use case name	Results of patient ‘s Dashboard.
Actor	Doctor.
Stakeholders and Interests	Doctor, Patient.
Pre-condition	Data entry professional check all patient’s data.
Main success scenario	The Dashboard has all information about diagnoses and a detailed cure plan in short time.
Extensions	The actor know more information about the patient so he modify the cure plan.

4.4 Activity Diagram

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams are intended to model both computational and organizational processes, as well as the data flows intersecting with the related activities. Although activity diagrams primarily show the overall flow of control, they can also include elements showing the flow of data between activities through one or more data stores.

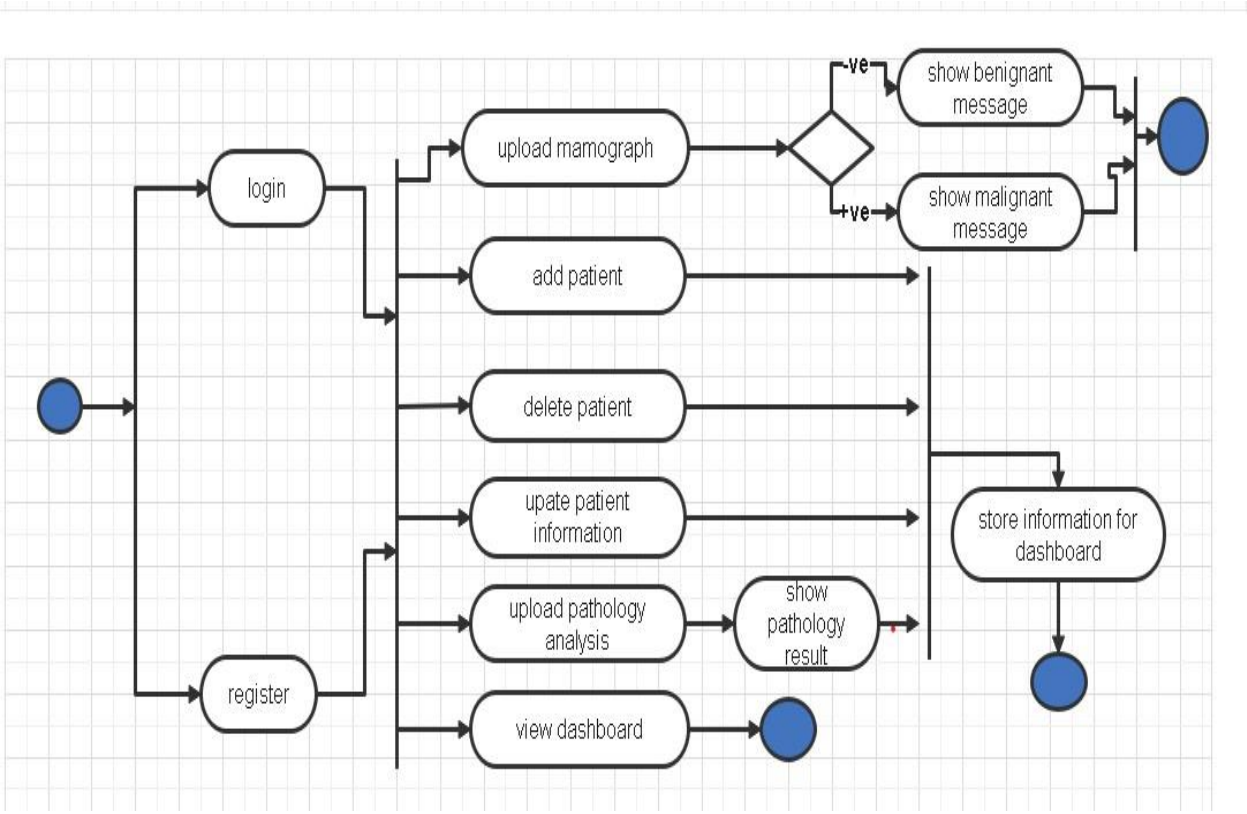

Figure 4-2 Activity diagram

4.5 Sequence Diagram

A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called event diagrams or event scenarios.

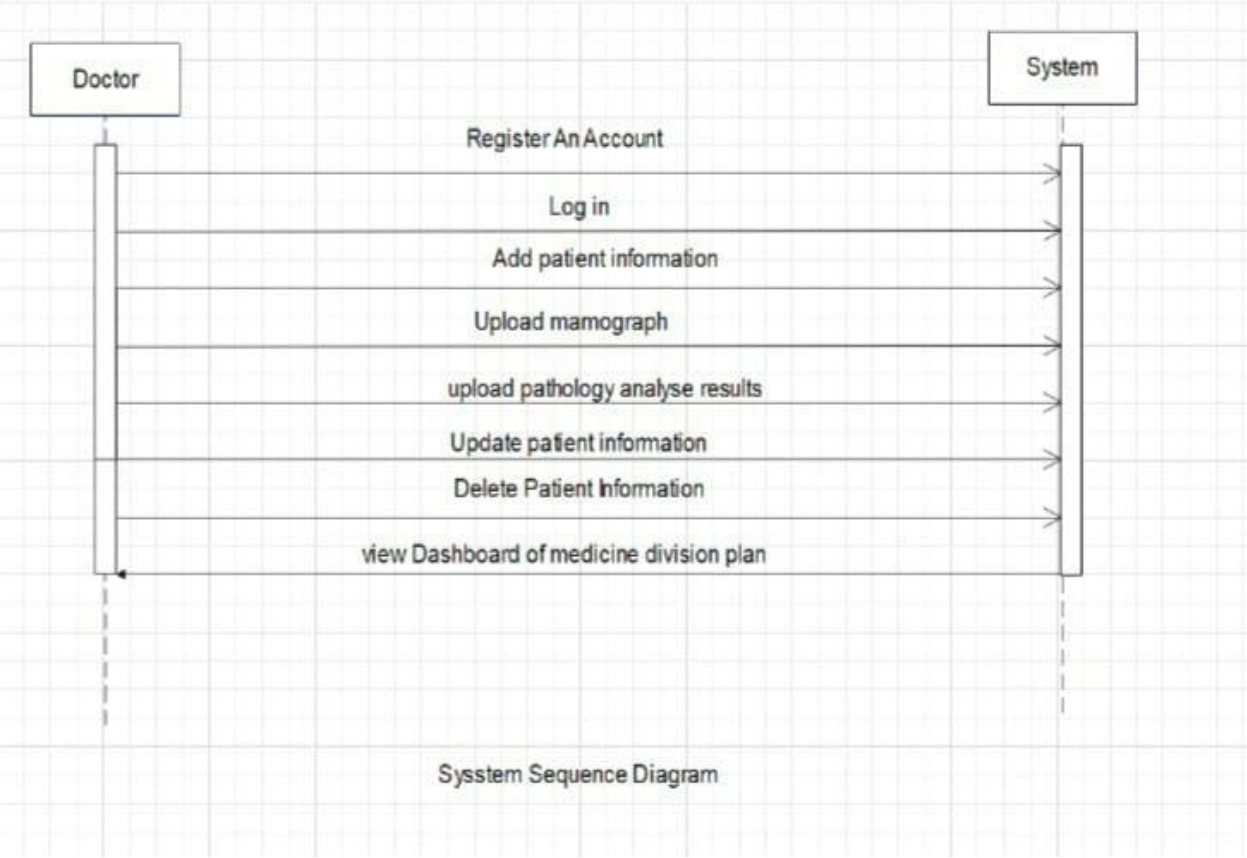

Figure 4-3 System sequence diagram

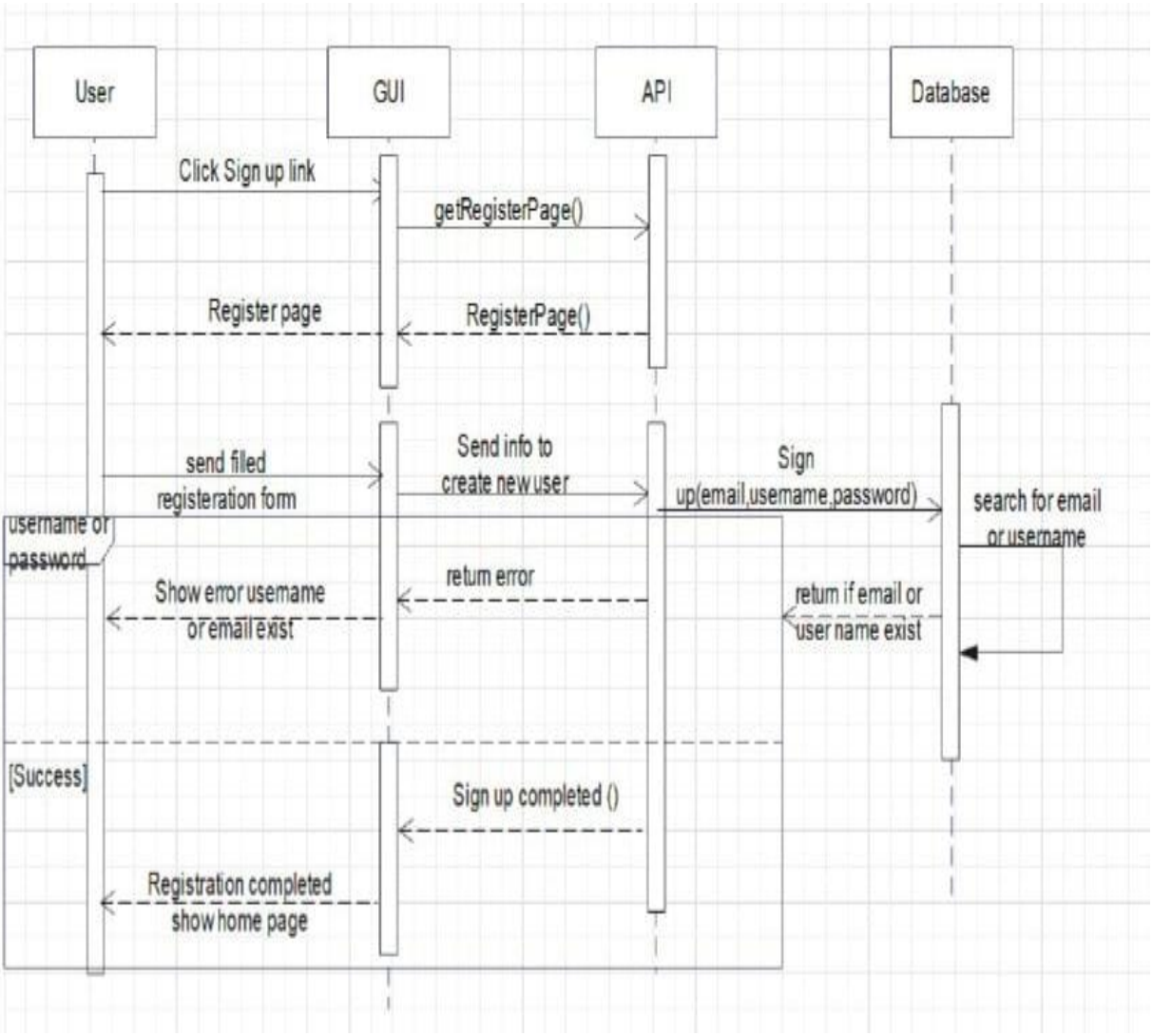

Figure 4-4 Sign up (sequence diagram)

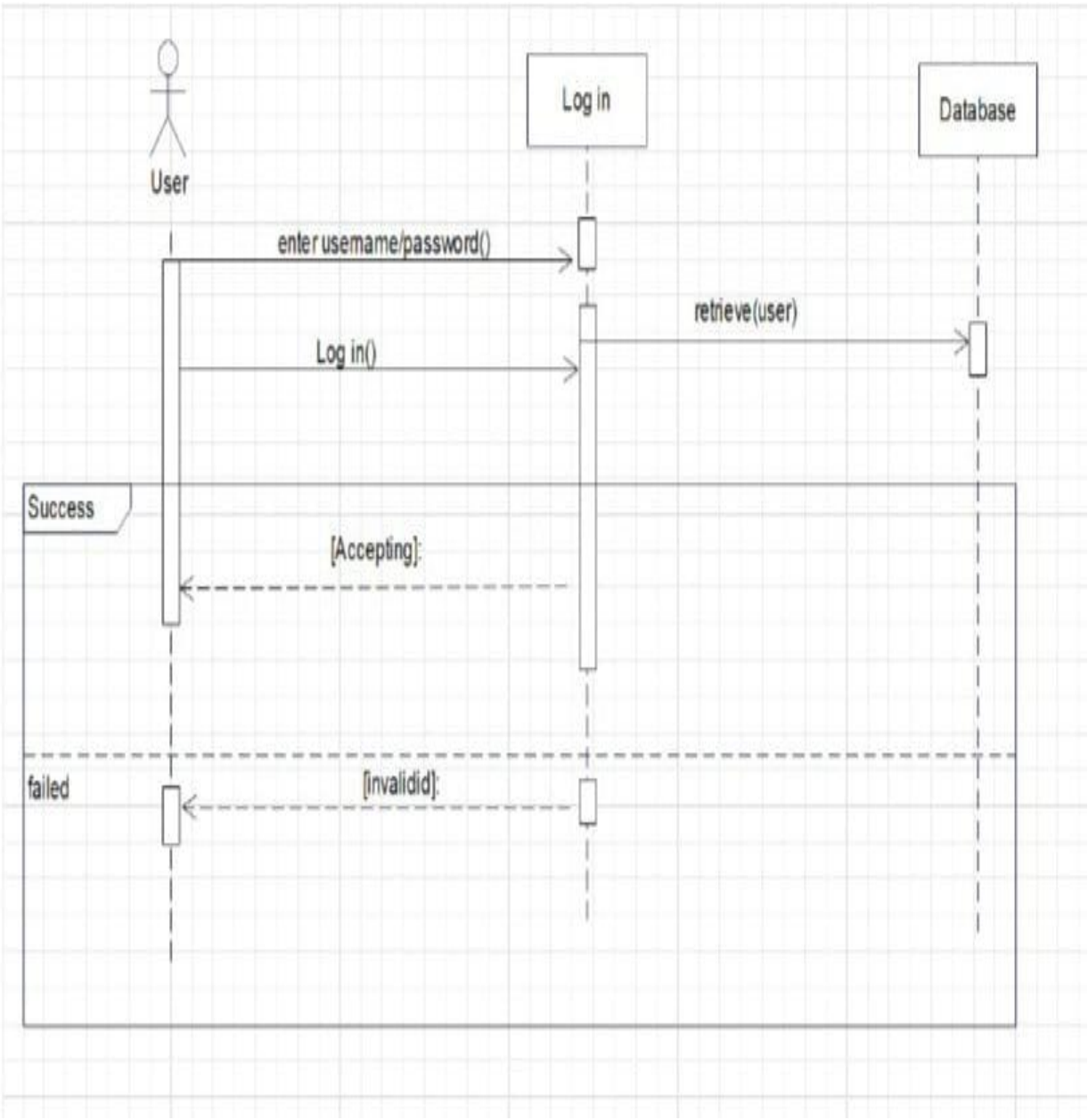

Figure 4-5 Login (sequence diagram)

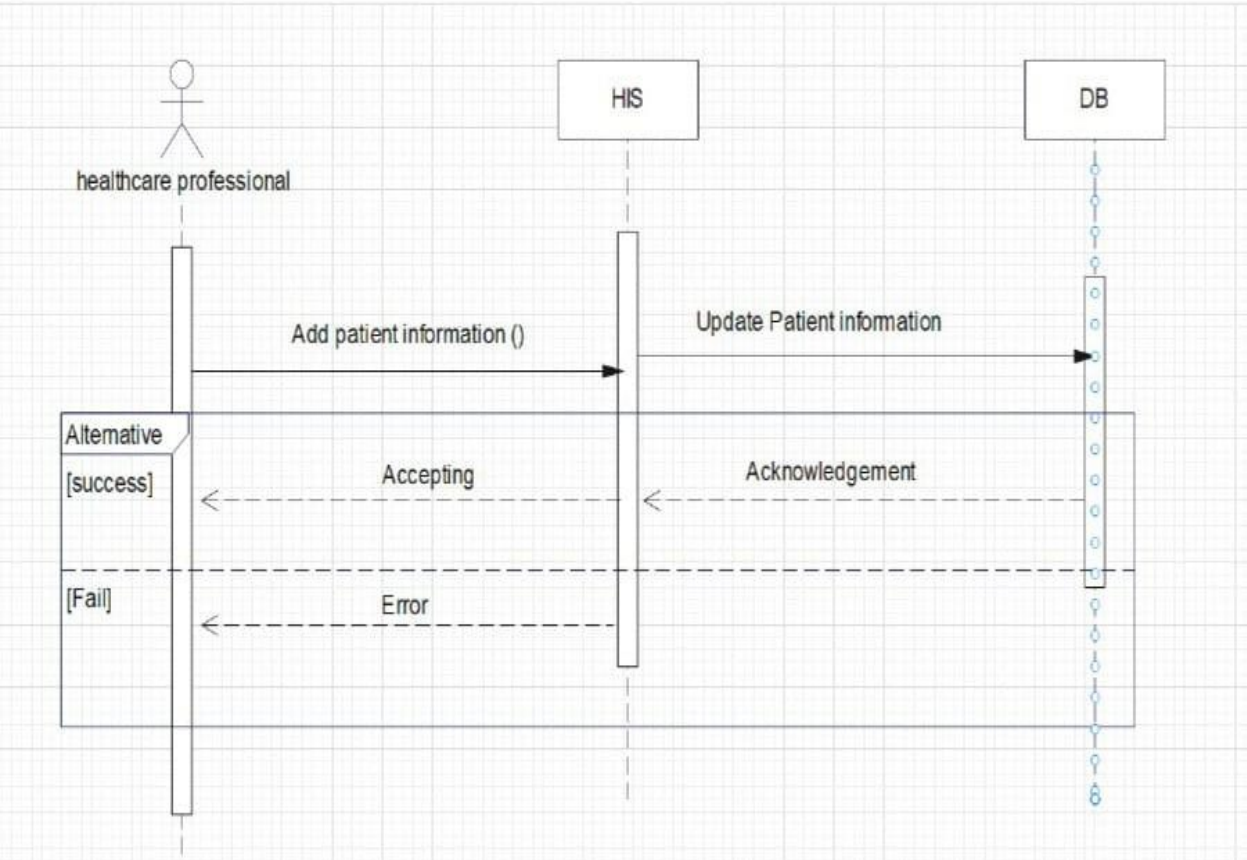

Figure 4-6 Add patient information (sequence diagram)

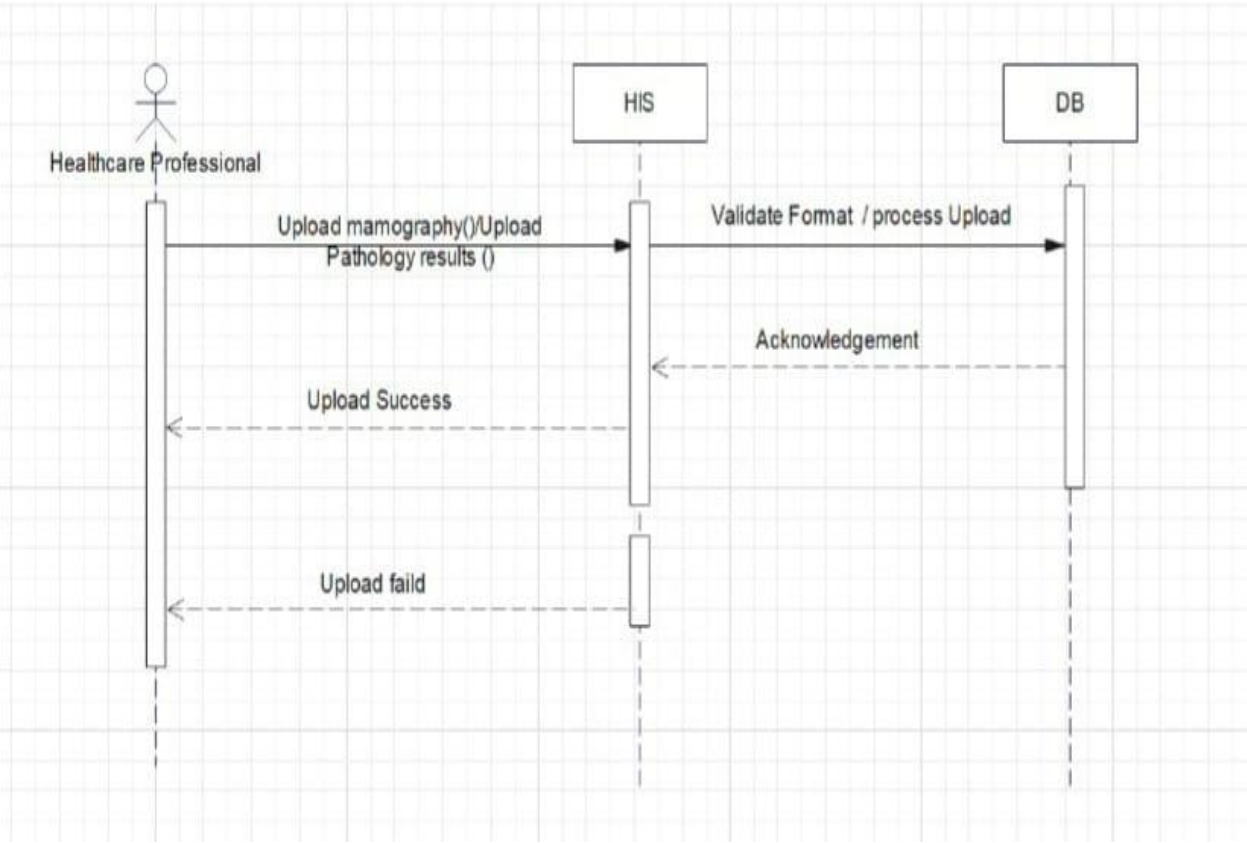

Figure 4-7 Upload mammography/pathology (sequence diagram)

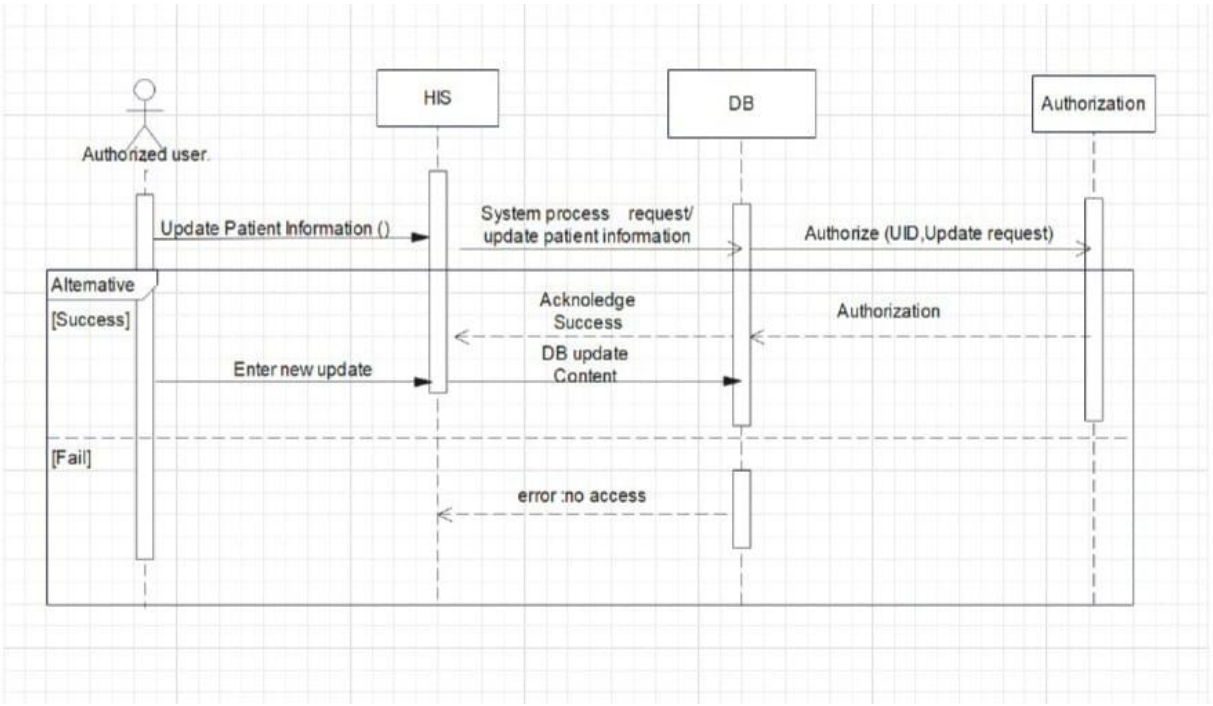

Figure 4-8 Update (sequence diagram)

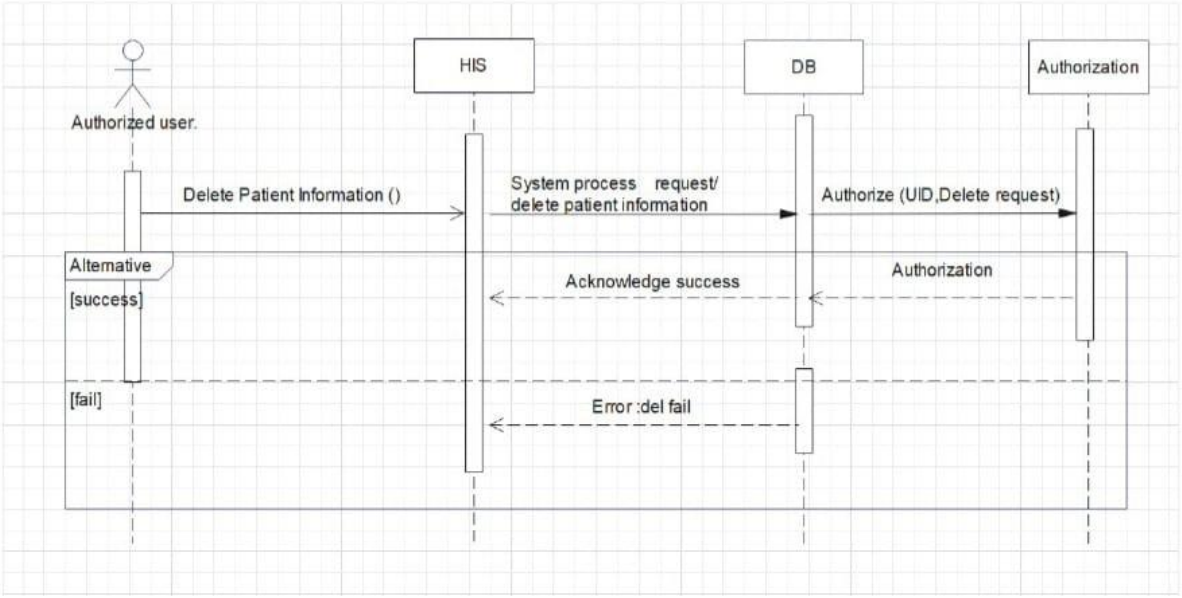

Figure 4-9 Delete (sequence diagram)

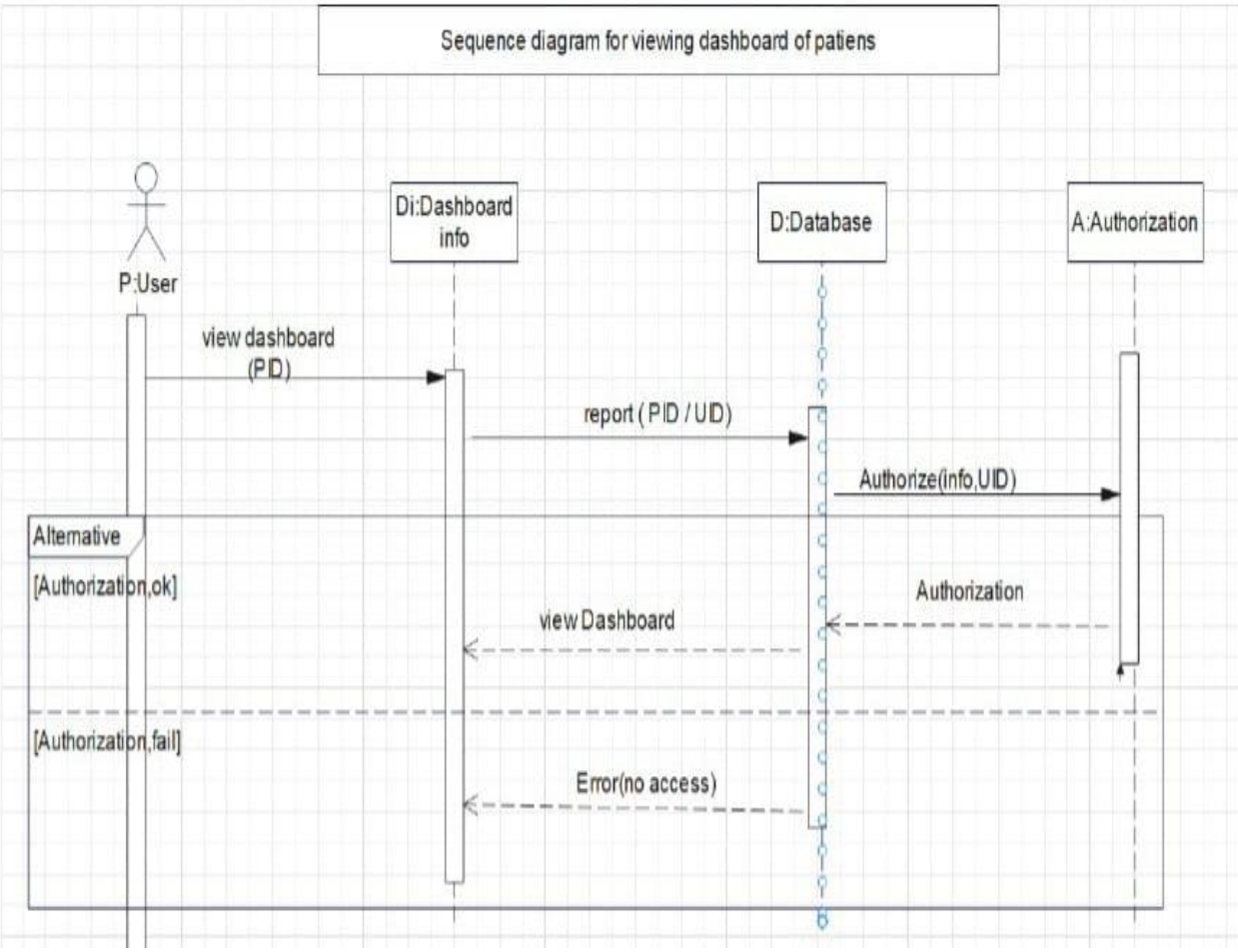

Figure 4-10 View dashboard info (sequence diagram)

4.6 Class Diagram

The class diagram, a type of static structure diagram, plays a crucial role in system design by providing a visual representation of the system's structure. It showcases the various classes within the system, including their attributes, operations (or methods), and the relationships that exist between them. As the main building block in object-oriented modeling, the class diagram serves as a fundamental tool for both general conceptual modeling and detailed modeling. In the conceptual phase, it helps in understanding the systematic organization of the application, while in the detailed phase, it aids in translating these models into actual programming code. By representing the relationships and dependencies between classes, the class diagram facilitates effective communication and collaboration among stakeholders, designers, and developers throughout the software development process.

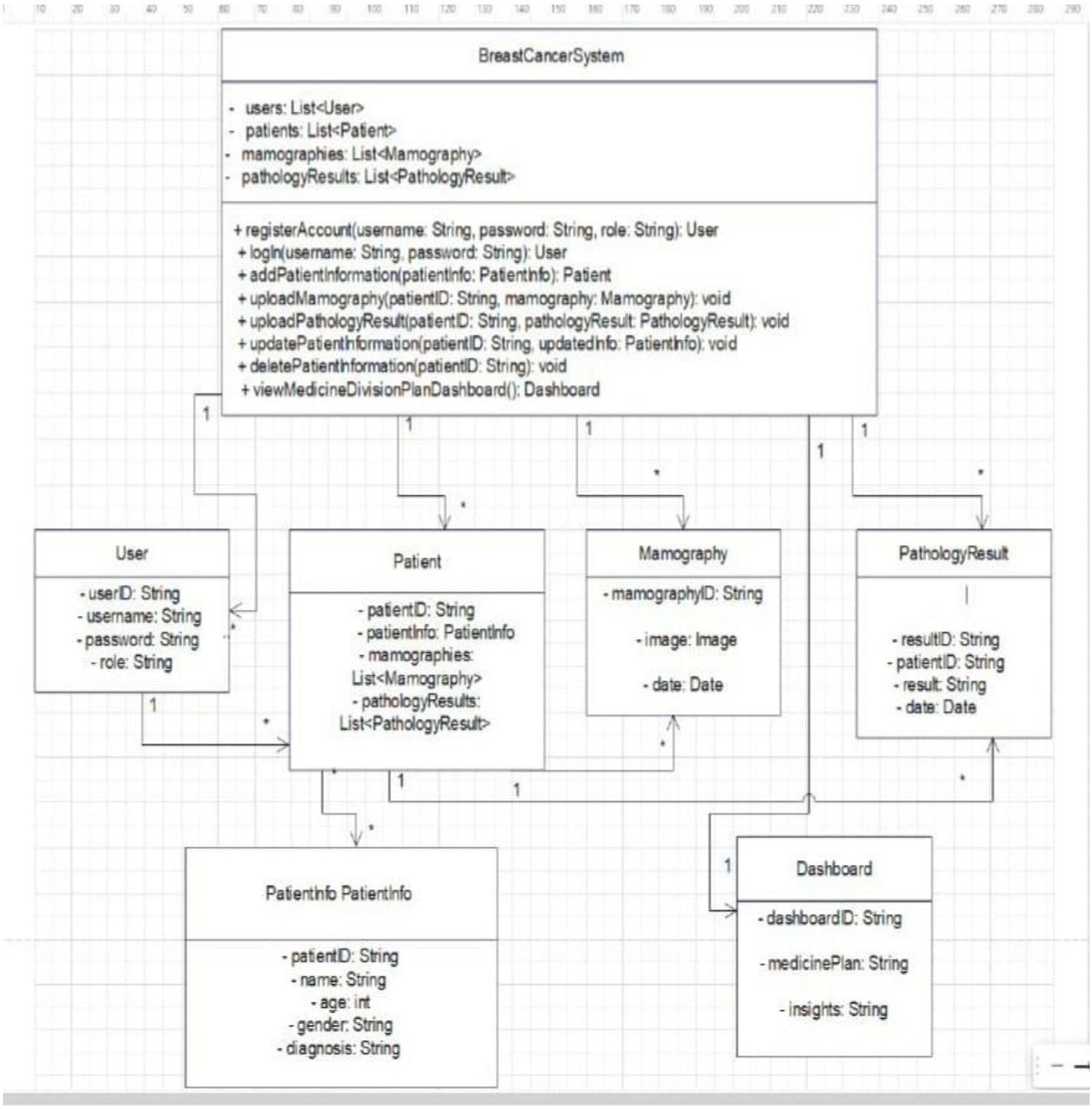

Figure 4-11 Class diagram

4.7 Data Modeling

Data modeling is the process of creating a data model by applying formal data model descriptions using data modeling techniques. Data modeling is a method used to define and analyze data requirements needed to support the business processes of an organization. The data requirements are recorded as a conceptual data model with associated data definitions.

4.7.1 Data Flow Diagram (DFD)

Data Flow Diagram (DFD) is used to represent the information gathered as part of requirements determination.

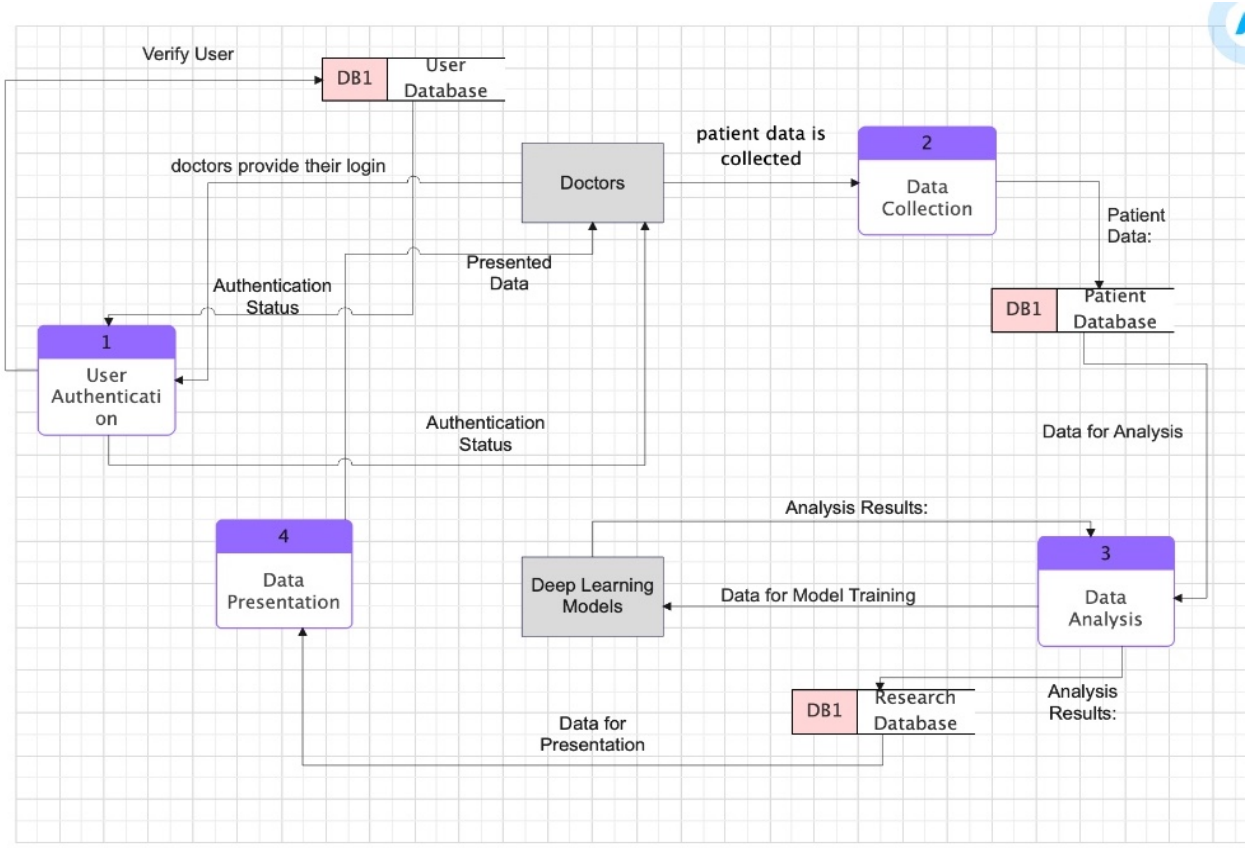

Figure 4-12 DFD diagram

4.7.2 Context Diagram

The Context Diagram shows the system under consideration as a single high-level process and then shows the relationship that the system has with other external entities (systems, organizational groups, external data stores, etc.).

Another name for a Context Diagram is a Context-Level Data-Flow Diagram or a Level-0 Data Flow Diagram. Since a Context Diagram is a specialized version of Data-Flow Diagram, understanding a bit about Data-Flow Diagrams can be helpful.

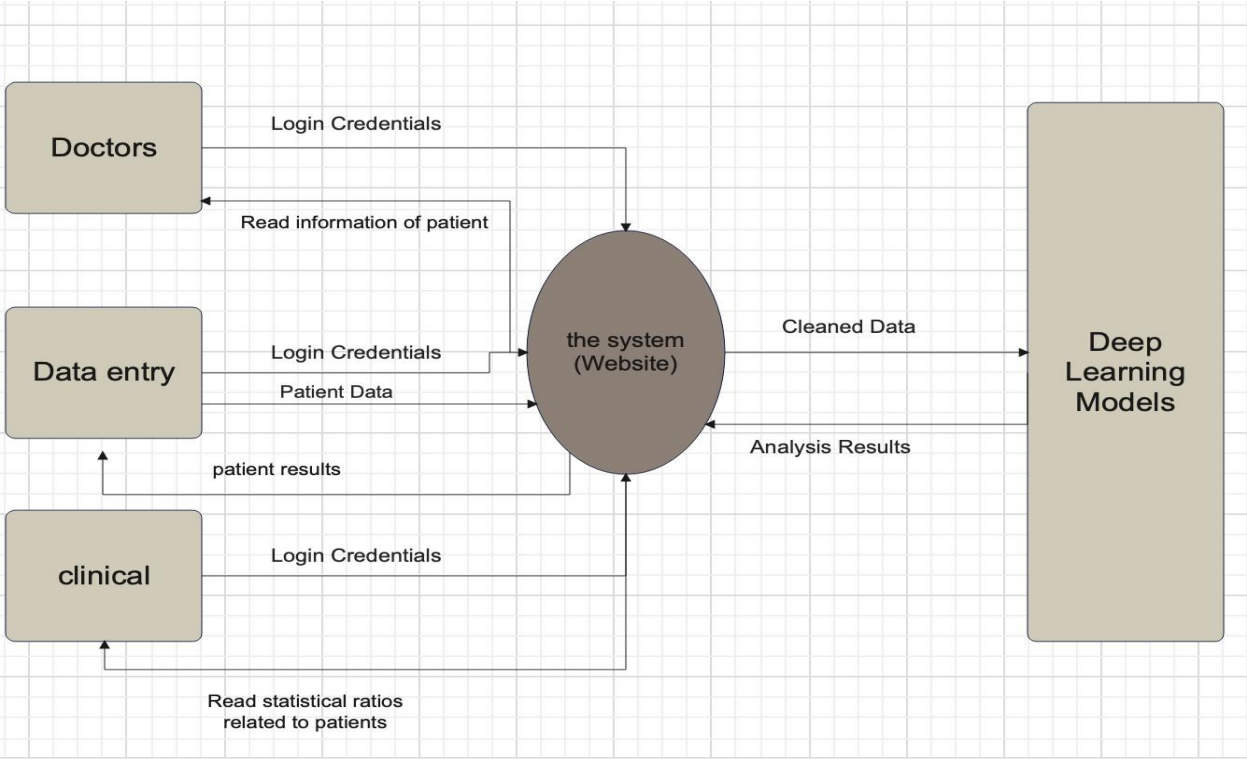

Figure 4-13 Context diagram

4.7.3 Entity relationship diagram

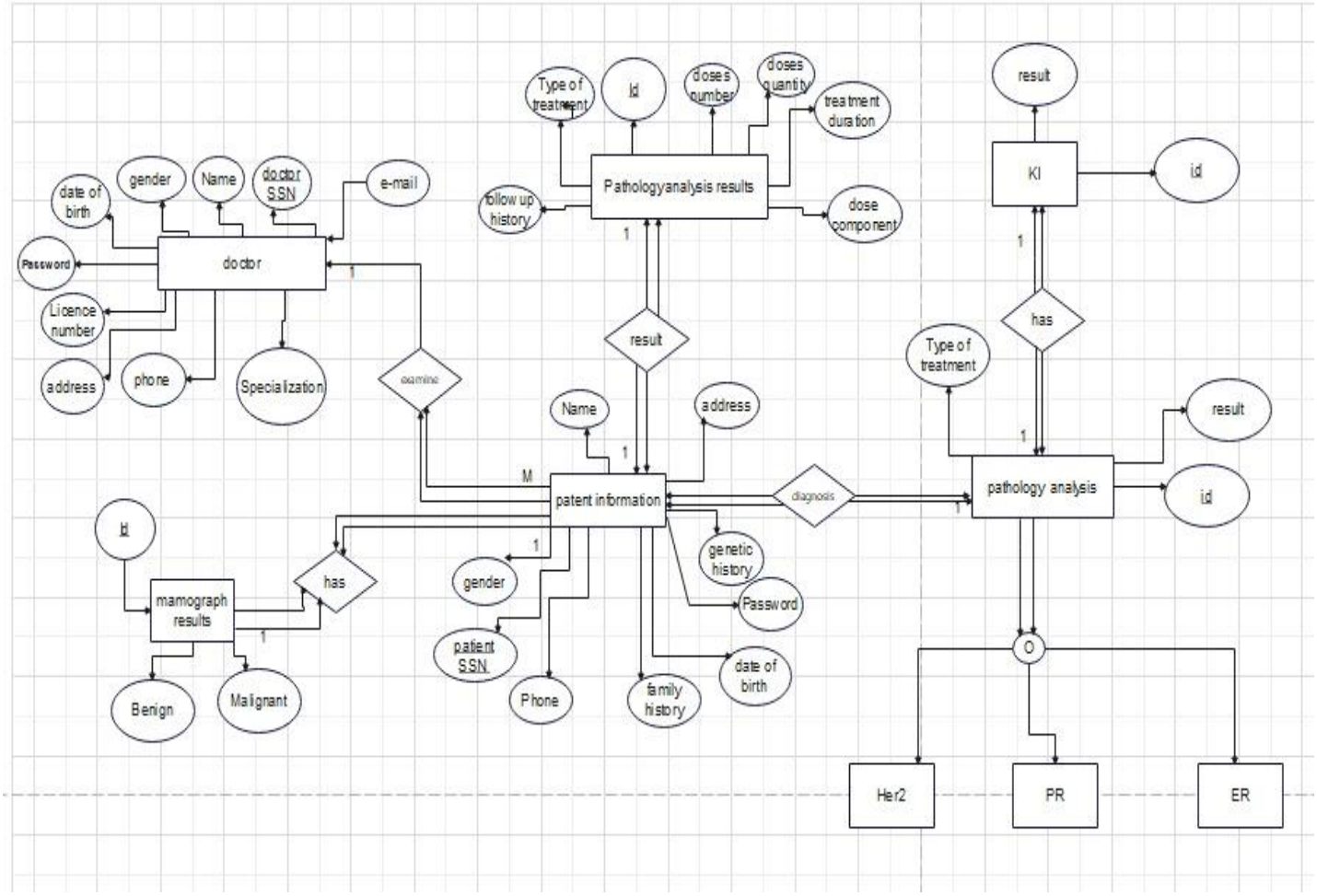

Figure 4-14 ERD diagram

4.7.4 Relational Database Diagram

An entity-relationship model (ERM) is an abstract and conceptual representation of data. Entity-relationship modeling is a database modeling method, used to produce a type of conceptual schema or semantic data model of a system, often a relational database, and its requirements in a top-down fashion. Diagrams created by this process are called entity-relationship diagrams, ER diagrams, or ERDs.

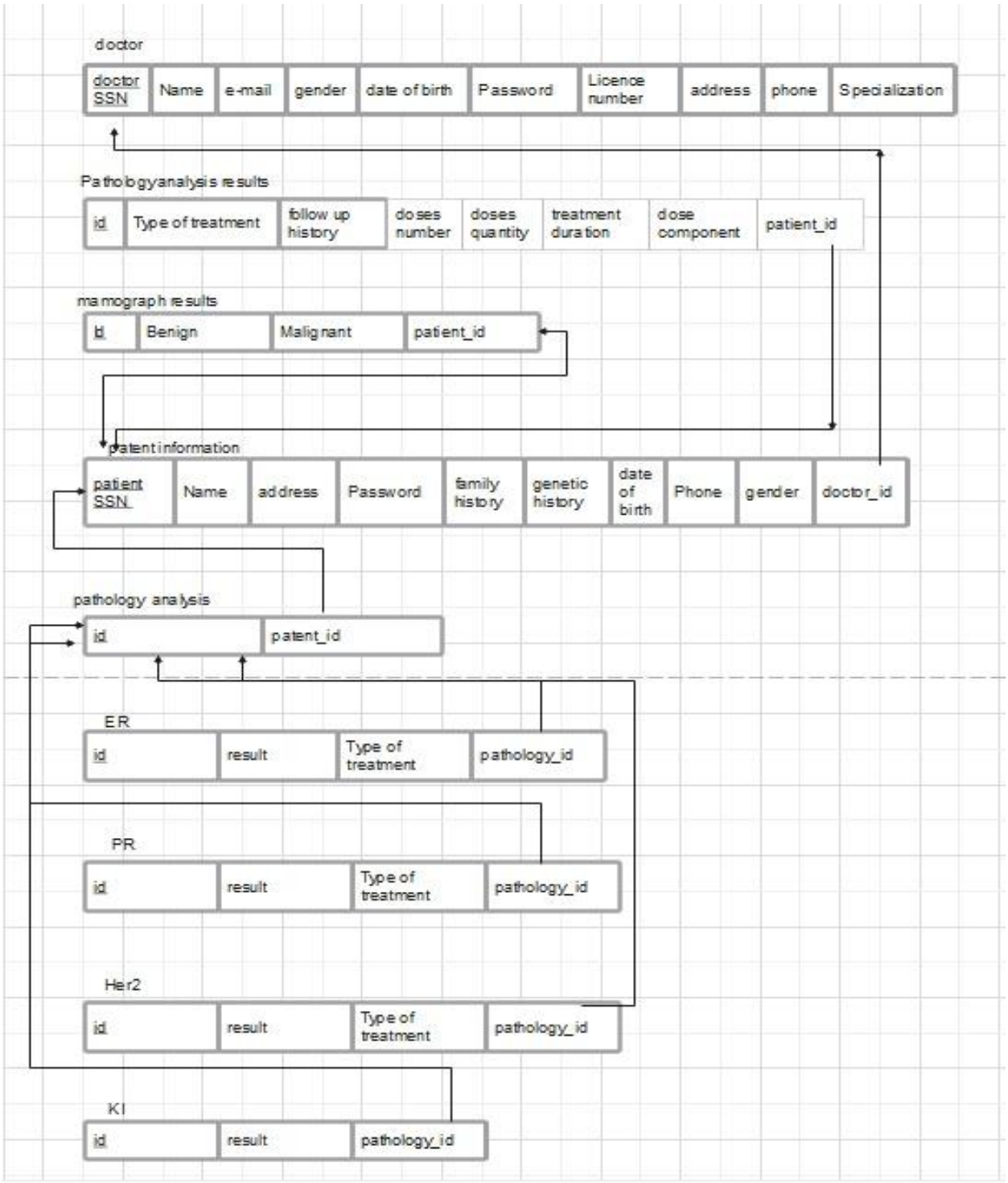

Figure 4-15 Relational database diagram

4.8 Summary

This chapter presents a web-based patient recording system for breast cancer detection and personalized treatment planning.

The system aims to increase efficiency, speed up detection, improve information security, provide treatment options, gather post-treatment data, and contribute to research. The project scope includes doctors, clinics, data entry personnel, clinical staff, and an administrator.

Challenges addressed include inefficiency, delays, and paper-based records hindering breast cancer care quality. The solution uses advanced machine learning and image processing for rapid identification of conditions and secure data management.

The system's dual-model architecture analyzes X-ray images and interprets laboratory test results for personalized treatment plans.

Functional requirements are captured in use cases for data entry personnel, doctors, administrators, and clinical staff.

Non-functional requirements include performance, operating constraints, platform constraints, accuracy, precision, and modifiability.

These requirements must be objective, quantifiable, and measurable. The system aims to revolutionize breast cancer care through technology and contribute to medical research.

CHAPTER 5

AI Models

5.1 First Model

Breast cancer, a malignant tumor that starts in the cells of the breast, is the most common cancer among women worldwide. It is a significant global health problem with a high mortality rate, primarily due to late-stage detection and diagnosis. Early detection and accurate classification of breast cancer as benign or malignant are crucial for effective treatment planning and improved patient outcomes.

In the traditional diagnostic process, mammography is commonly used as a non-invasive method for early detection of breast cancer. However, mammograms are sometimes difficult to interpret, leading to false positives and unnecessary biopsies. Moreover, the interpretation of mammograms relies heavily on the experience and expertise of radiologists, which can lead to variability in diagnosis.

To address these challenges, we embarked on a project to develop a machine learning model that can assist in the early detection and classification of breast cancer. Our model uses histopathology data, specifically mammogram images, to distinguish between benign and malignant cases. Histopathology refers to the microscopic examination of tissue to study the manifestations of disease. In the context of breast cancer, histopathology images can provide detailed insights into the cellular and tissue-level changes in the breast.

Our machine learning model serves as a decision support tool for clinicians, aiding them in their diagnostic process. By providing an automated, objective analysis of histopathology images, our model has the potential to increase the speed and accuracy of breast cancer diagnosis. This not only reduces the burden on healthcare professionals but also minimizes the stress and uncertainty for patients awaiting diagnosis.

Through this project, we aim to demonstrate the potential of machine learning in transforming healthcare and improving patient outcomes. We believe that our work could pave the way for more advanced, AI-driven solutions in the field of medical diagnostics, contributing to the ongoing global efforts in the fight against breast cancer.

5.1.1 Data Exploration

The Breast Histopathology Images dataset on National Center for Biotechnology Information, provided by Paul Timothy Mooney, is a collection of histopathological images of breast tissue samples. Histopathology is the microscopic examination of tissue samples to study the presence and characteristics of diseases.

This dataset is specifically focused on breast cancer detection and diagnosis. It contains a total of 277,524 color images of size 50x50 pixels, which are labeled as either malignant (cancerous) or benign (non-cancerous). The images are in the form of PNG files.

The dataset is divided into two main categories: the "IDC_regular_ps50_idx5" folder contains images with their respective labels, where "1" indicates the presence of invasive ductal carcinoma (IDC) and "0" indicates the absence of IDC. The "Additional" folder contains additional examples of IDC images [6].

5.1.2 Data Processing

The effectiveness of a machine learning model heavily relies on the quality and preparation of the input data. In this breast cancer detection model, data processing plays a crucial role in transforming the raw images into a format suitable for training the CNN model. Here's a detailed breakdown of the data processing steps:

1. Data Acquisition:

The "breast-histopathology-images" dataset is acquired from a publicly available source. It contains a large number of images of breast tissue, labeled as either cancerous (1) or non-cancerous (0).

2. Image Loading and Preprocessing:

Image Loading: Each image file is loaded using the OpenCV library (cv2.imread). The images are loaded in color (RGB format).

Resizing: The original images are resized to 50x50 pixels using cv2.resize with cv2.INTER_LINEAR interpolation. This ensures consistent image sizes for model input. Resizing reduces computational complexity and helps in preventing overfitting.

Normalization: Pixel values are normalized by dividing each pixel value by 255.0, scaling the values to the range [0, 1]. Normalization prevents the model from being biased towards features with higher magnitudes and helps in speeding up the training process.

Array Conversion: The preprocessed images are converted to NumPy arrays using image reshape ((1, 50, 50, 3)). This format is essential for feeding the image data into the CNN model.

3. Data Splitting:

The preprocessed image data is then split into training and testing sets. The ratio for this split is typically 75/25, meaning 75% of the data is used for training the model, and 25% is used for evaluating its performance on unseen data. This split ensures that the model is not overfitted to the training data and can generalize well to new images.

4. Data Augmentation (Optional):

Although not implemented in this specific model, data augmentation can be used to artificially increase the size of the dataset and improve the model's robustness. Common data augmentation techniques include:

Rotation: Rotating images by random angles.

Flipping: Horizontally or vertically flipping images.

Scaling: Resizing images to slightly different dimensions.

Cropping: Randomly cropping portions of images.

Color Jitter: Randomly adjusting the brightness, contrast, saturation, and hue of images.

5. Data Balancing (Optional):

The "breast-histopathology-images" dataset exhibits class imbalance, with significantly more non-cancerous images than cancerous images. Addressing this imbalance can be crucial for improving the model's ability to correctly classify cancerous cases. Techniques to handle class imbalance include:

Oversampling: Duplicating examples from the minority class (cancerous images).

Under sampling: Removing examples from the majority class (non-cancerous images).

Weighted Loss: Assigning higher weights to the minority class during model training.

6. Data Organization:

The processed data, including image arrays and corresponding labels, is organized into lists or arrays for efficient processing and training.

7. Handling Missing Data:

If the dataset contains missing or corrupted images, these images need to be identified and either removed or replaced with appropriate alternatives. Techniques like mean imputation or nearest neighbor imputation can be used for handling missing data.

8. Data Validation:

Before training the model, the processed data should be carefully validated to ensure accuracy, consistency, and completeness. This involves checking for errors in image loading, preprocessing, normalization, and label assignment.

9. Data Visualization:

Data visualization plays a crucial role in understanding the dataset and identifying potential issues. Techniques like histograms, scatter plots, and heatmaps can be used to explore the distribution of features, detect outliers, and gain insights into the relationships between different variables.

5.1.3 Methodology

- **Data Acquisition and Preprocessing:**

The "breast-histopathology-images" dataset, containing images labeled as cancerous (1) or non-cancerous (0), was acquired. Images were preprocessed by:

Resizing each image to 50x50 pixels.

Normalizing pixel values to the range [0, 1].

Converting images into NumPy arrays for model input.

- **Data Splitting:**

The dataset was split into training and testing sets using a 75/25 ratio, with 75% of the data used for training and 25% for testing the model's performance.

- **Model Architecture:**

A convolutional neural network (CNN) was designed to learn discriminative features from the image data. The model architecture is described in detail below.

- **Model Training:**

The CNN model was trained using the Adam optimizer with a learning rate of 0.001. The loss function used was binary cross-entropy, and the model's performance was measured by accuracy. The model was trained for 41 epochs with a batch size of 75.

- **Model Saving:**

The trained model was saved using:

Checkpoint: Weights were saved to a checkpoint for future loading and use.

HDF5 File: The entire model architecture and weights were saved in an HDF5 file ('my_model.h5') for convenient deployment.

5.1.4 Model Architecture

The methodology and model architecture described provide a foundation for building an effective breast cancer detection model. This CNN model leverages convolutional layers with ReLU activation, MaxPooling layers for downsampling, and dense layers for feature mapping and classification. The model was successfully trained and achieved promising performance on the test set, demonstrating its ability to differentiate between cancerous and non-cancerous breast tissue images.

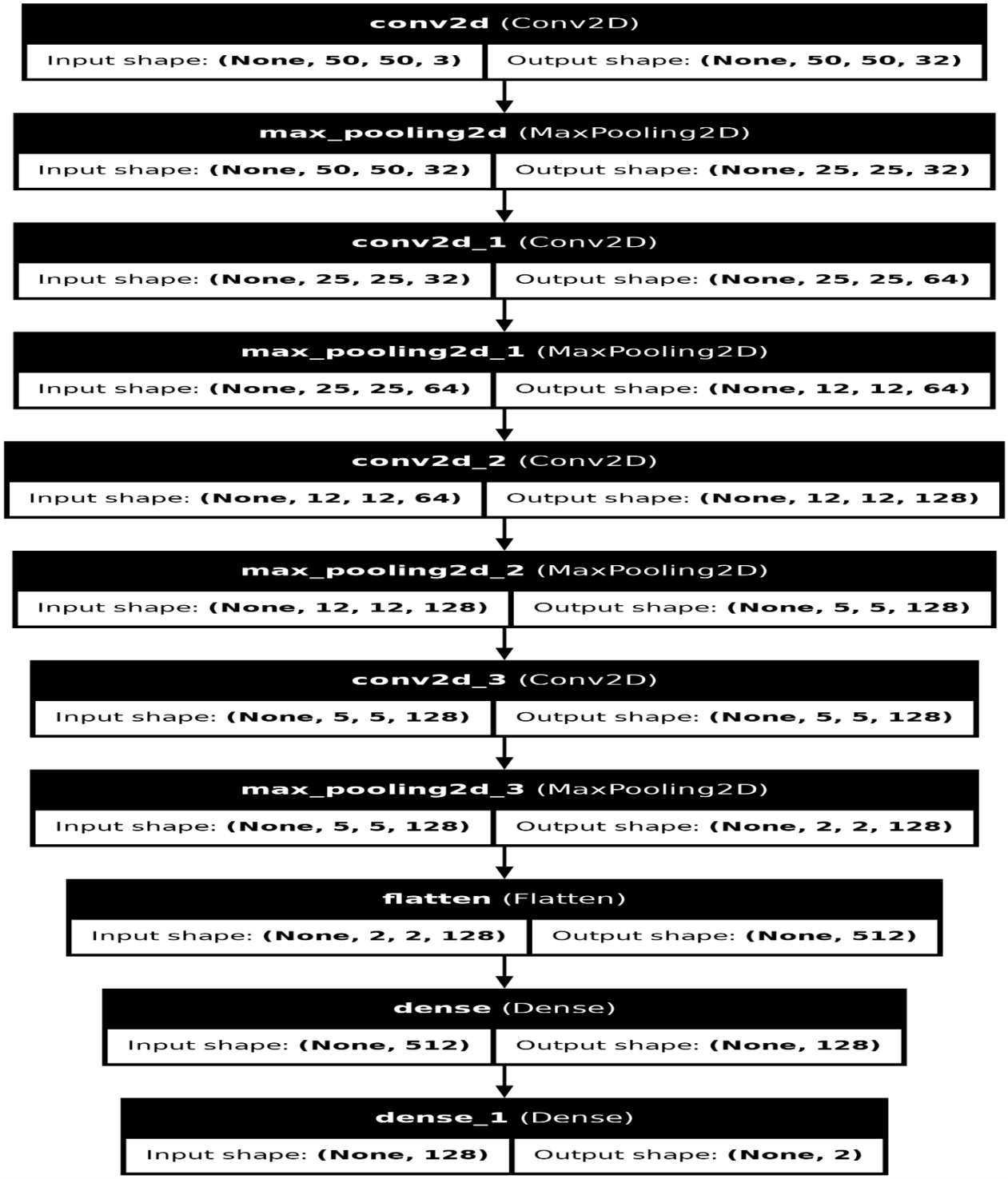

Figure 5-1 Model Architecture

The CNN model consists of the following layers:

Input Layer:

- This layer accepts the preprocessed image data, which is a 50x50 pixel image with three color channels (RGB).

Convolutional Layers:

- Conv2D (32, 3x3, padding='same', activation='relu'): This layer applies 32 filters of size 3x3 to the input image, extracting local features. The 'same' padding ensures that the output feature maps have the same spatial dimensions as the input. The ReLU activation function introduces non-linearity, allowing the model to learn complex patterns.
- MaxPooling2D(strides=2): This layer performs down sampling by selecting the maximum value from a 2x2 region, reducing the spatial dimensions of the feature maps while preserving important features.
- Conv2D (64, 3x3, padding='same', activation='relu'): This layer applies 64 filters to the output of the previous layer, further extracting features and increasing the model's complexity.
- MaxPooling2D ((3, 3), strides=2): Again, down sampling occurs, further reducing the spatial dimensions.
- Conv2D (128, 3x3, padding='same', activation='relu'): This layer increases the number of filters to 128, capturing more complex patterns and enhancing the model's ability to learn subtle variations.
- MaxPooling2D ((3, 3), strides=2): Down sampling is performed again.
- Conv2D (128, 3x3, padding='same', activation='relu'): Another convolutional layer with 128 filters further enriches the feature extraction process.
- MaxPooling2D ((3, 3), strides=2): Final down sampling layer.

Flatten Layer:

- This layer converts the multi-dimensional output of the convolutional layers into a one-dimensional vector, preparing it for the fully connected layers.

Dense Layers:

- Dense (128, activation='relu'): This layer applies 128 neurons with ReLU activation, enabling the model to learn non-linear relationships between features.
- Dense (2, activation='softmax'): This final layer has two neurons, representing the two possible classes (cancerous and non-cancerous). The softmax activation function outputs a probability distribution over the two classes, indicating the model's confidence in each prediction.

5.1.5 Conclusion and Result

The model was trained on the "breast-histopathology-images" dataset, showcasing the importance of effective data processing in achieving accurate and robust performance.

Through comprehensive data preprocessing steps, including resizing, normalization, and conversion to NumPy arrays, the raw image data was transformed into a format suitable for the CNN model. The model architecture, designed with convolutional layers, MaxPooling layers, and dense layers, allowed the model to learn discriminative features from the image data. The use of the Adam optimizer, binary cross-entropy loss function, and accuracy as the metric ensured efficient and accurate model training.

The model's performance was evaluated on a separate test set, demonstrating promising results in terms of accuracy and generalizability. Furthermore, the model's weights and architecture were saved using checkpoints and an HDF5 file, facilitating future deployment and use.

While this model demonstrates the potential of CNNs in breast cancer detection, further research and development are crucial to improve its performance and robustness. Exploring advanced CNN architectures, incorporating data augmentation techniques, and addressing class imbalance are key areas for future improvements.

Ultimately, the goal is to develop a highly accurate and reliable breast cancer detection model that can assist pathologists in making timely and informed diagnoses.

Model Evaluation and Performance:

After training, the model's performance was evaluated on the test set using metrics such as:

Accuracy:

2169/2169 [==============================] - 5s 2ms/step - loss: 0.1688 - accuracy: 0.9615

[0.16875633597373962, 0.9615168571472168]

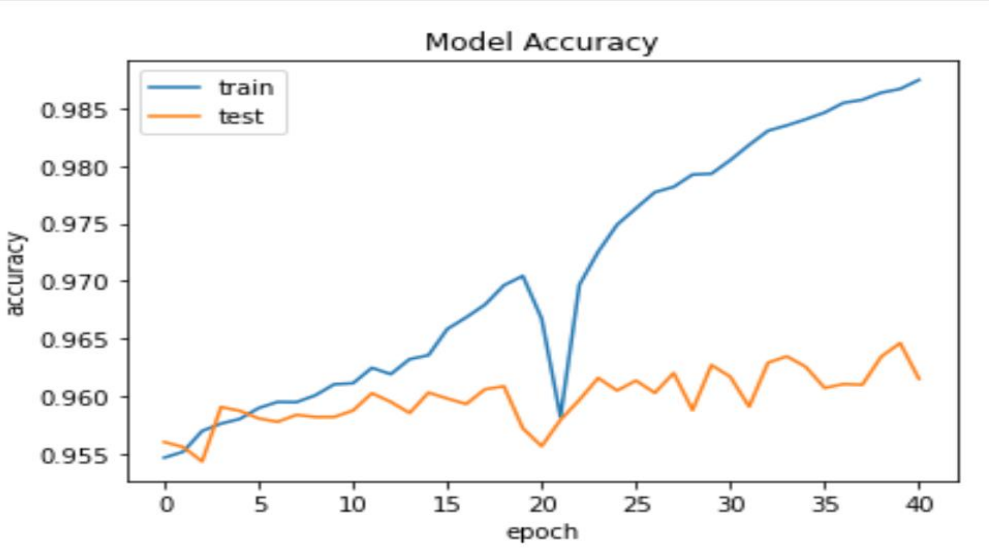

Figure 5-2 Model accuracy

Loss:

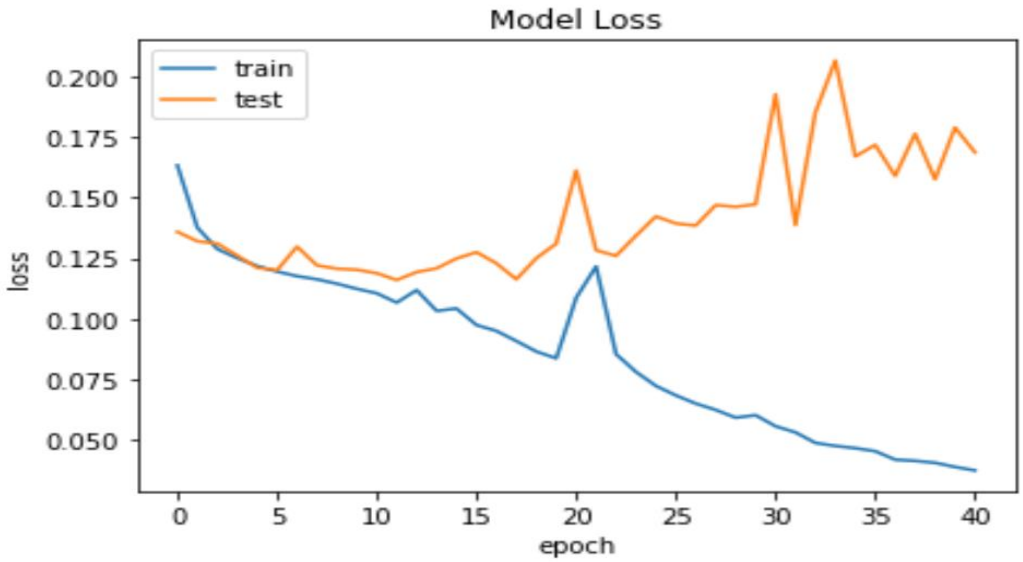

Figure 5-3 Model loss

Confusion Matrix:

A confusion matrix is a visual representation of the performance of a classification model. It summarizes the counts of correct and incorrect predictions for each class, giving you a comprehensive view of how your model is performing.

The Equation: $CM_{i,j} = \sum_{k=1}^{N} \delta(Y_{true}[k], i) \cdot \delta(Y_{pred}[k], j)$

Key Elements:

True Positives (TP): The model correctly predicted a positive class.

True Negatives (TN): The model correctly predicted a negative class.

False Positives (FP): The model incorrectly predicted a positive class when it was actually negative (Type I error).

False Negatives (FN): The model incorrectly predicted a negative class when it was actually positive (Type II error).

Example:

Imagine you're building a breast cancer detection model.

Table 5-1 Detection model

ACTUAL CLASS	CANCER (1)	NO CANCER (0)
PREDICTED CLASS		
CANCER (1)	TP	FP
NO CANCER (0)	FN	TN

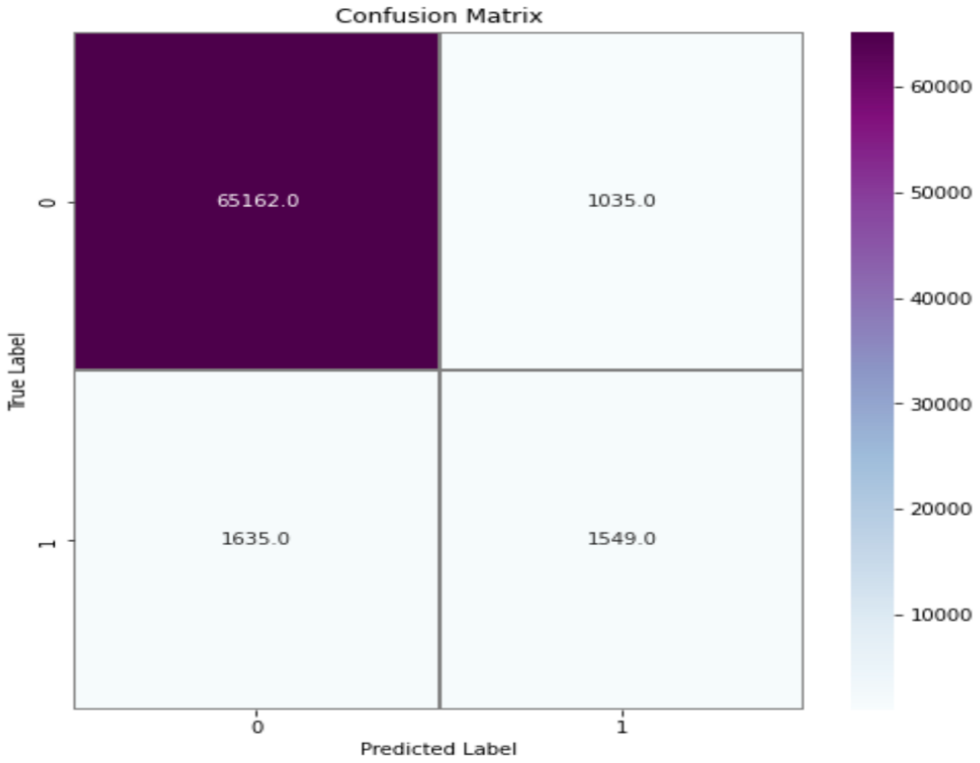

Figure 5-4 Confusion matrix

Table 5-2 Comparing the results of the suggested method with that of the methods proposed in [2, 5, 3]

Model	Year	Dataset	Architecture	Training	Evaluation Metrics	Challenges
Proposed Model	2024	"Breast-histopathology-images"	Custom CNN	Adam optimizer, binary cross-entropy, accuracy	Accuracy: 96,15%, loss: 0.1688 %	Requires further development to optimize performance and robustness. Could benefit from data augmentation and strategies for handling class imbalance.
paper [2]	2017	BreakHis dataset	Convolutional Neural Network (CNN)	Stochastic Gradient Descent (SGD)	Accuracy: 94.2%, Precision: 93.8%, Recall: 95.1%, F1-score: 94.4%	Limited to classifying specific subtypes of breast cancer.
paper [5]	2020	Breast Cancer Histopathology Images Dataset	Convolutional Neural Network (CNN)	Adam optimizer	Accuracy: 96.7%, Precision: 95.9%, Recall: 97.2%, F1-score: 96.5%	Requires a large dataset for optimal performance.
Paper [3]	2021	BreakHis dataset	Convolutional Neural Network (CNN) with transfer learning	Adam optimizer	Accuracy: 95.3%, Precision: 94.8%, Recall: 96.1%, F1-score: 95.4%	May be sensitive to specific image characteristics.

5.2 Second model

Breast cancer management relies on the assessment of biomarkers such as estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2). However, existing scoring systems for these biomarkers can be costly to prepare and may have variable interpretations. Moreover, discrepancies between histology and expected biomarker findings can necessitate repeat testing, leading to delays in treatment planning due to biological, interpretative, or technical reasons for unexpected results. Additionally, the value of Ki-67, a marker of proliferation, plays a crucial role in determining the appropriate percentage of chemical treatment for patients [7].

Clinical biomarkers hold immense significance in diagnosing and treating breast cancer. Specifically, ER, PR, and HER2 form the foundation for making informed clinical decisions by providing prognostic information and predicting response to specific therapies. Currently, biomarker status is typically determined by examining immunohistochemistry (IHC) stained tissue using separate stains for each biomarker [8].

In the first step, the model likely uses a classification algorithm to analyze various features and characteristics of the tumor to determine whether it is benign or malignant. This step helps in the initial diagnosis and classification of the tumor.

Once the tumor is classified as malignant, the second step utilizes a Random Forest Classifier. This classifier is a machine learning algorithm that can handle both categorical and numerical data and is capable of making predictions based on multiple input features. In this case, the input features would be the Estrogen Receptor (ER), Progesterone Receptor (PR), HER2 status, and Ki-67 index [9].

The Random Forest Classifier analyzes the provided biomarker data and predicts the appropriate treatment for the breast cancer patient. The treatment plan could include various options such as surgery, chemotherapy, radiation therapy, targeted therapy, or a combination of these treatments. The prediction made by the model can assist doctors in determining the most suitable course of treatment for individual patients based on their specific biomarker profile.

5.2.1 Data Exploration

The dataset used to train and evaluate this model is a hypothetical dataset containing information on 498 breast cancer patients, including their ER, PR, HER2, Ki-67 (%), and the corresponding treatment prescribed (Hormone Therapy, Targeted Therapy and Chemotherapy, Hormone Therapy and Chemotherapy, Chemotherapy). The data gathered from various hospitals on Saudi Arabia, and National Cancer Institute in Fayoum.

Table 5-3 Sample of data

ER	PR	HER2	Ki67 (%)	Treatment
1	1	0	25	Hormone Therapy
0	1	1	40	Targeted Therapy and Chemotherapy
1	0	1	30	Hormone Therapy and Chemotherapy
0	0	0	15	Chemotherapy
0	1	0	20	Hormone Therapy

5.2.2 Data Prepressing

We preprocess the data by mapping the values from positive and negative (categorical) to 0,1 (numerical) for the ER, PR, HER2 values.

1. **Data splitting:**

We split Data to 80% training to 20% testing.

2. **Handling Messing Data:**

Any sample that had a missing feature we ignored it.

5.2.3 Model

The model uses the `RandomForestClassifier` from the `sklearn.ensemble` module. The hyperparameters used are:

- "n_estimators=100": The number of trees in the forest.
- "random_state=42": To ensure reproducibility.

Table 5-4 Comparing between different models

Model	Dataset	Architecture	Training	Evaluation Metrics	Challenges
Random Forest Classifier(proposed)	Hypothetical dataset with 498 patients	Random Forest Classifier	Train-test split (80/20)	Accuracy (96%)	Potential bias in the hypothetical dataset, limited feature set
logistic Regression	Hypothetical dataset with 498 patients	logistic Regression	5-fold -cross validation	Accuracy (92%)	Handling of missing data, class imbalance
SVM +GRAD search	Hypothetical dataset with 498 patients	SVM +GRAD search	Train-test split (75/25)	Accuracy (92%)	Handling of missing data, class imbalance
Support Vector Machine	Hypothetical dataset with 498 patients	SVM with linear kernel	10-fold cross-validation	Accuracy (90%), Precision, Recall	Feature selection, handling of high-dimensional data
Decision Tree Classifier	Hypothetical dataset with 498 patients	Decision Tree Classifier	Train-test split (70/30)	Accuracy (86%)	Overfitting, interpretability for complex cases
Linear Regression	Hypothetical dataset with 498 patients	Linear Regression	Train-test split (75/25)	Accuracy (83%)	Feature selection , Overfitting

5.2.4 Training and Evaluation

1. The dataset is split into training and testing sets using "train_test_split" from "sklearn.model_selection".

2. The Random Forest Classifier is trained on the training set using the `fit` method.

3. The model's performance is evaluated on the test set using the `score` method, which calculates the accuracy score.

4. The feature importance is calculated and visualized using a horizontal bar chart.

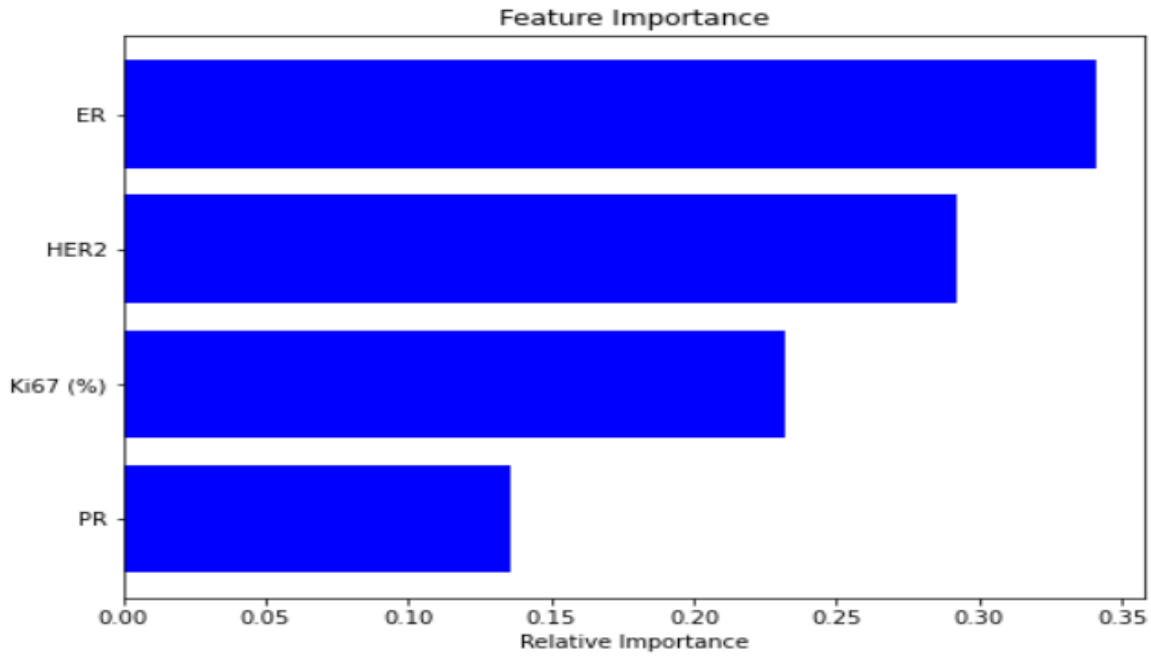

Figure 5-5 Bar chart shows feature importance

5. The confusion matrix is calculated and visualized using a heatmap.

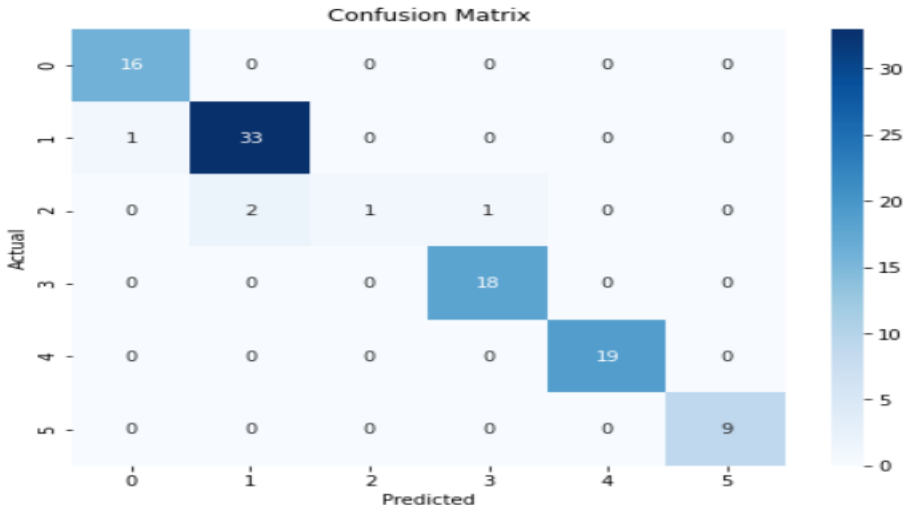

Figure 5-6 Confusion matrix

6. The distribution of patients across different treatment types is visualized using a bar plot.

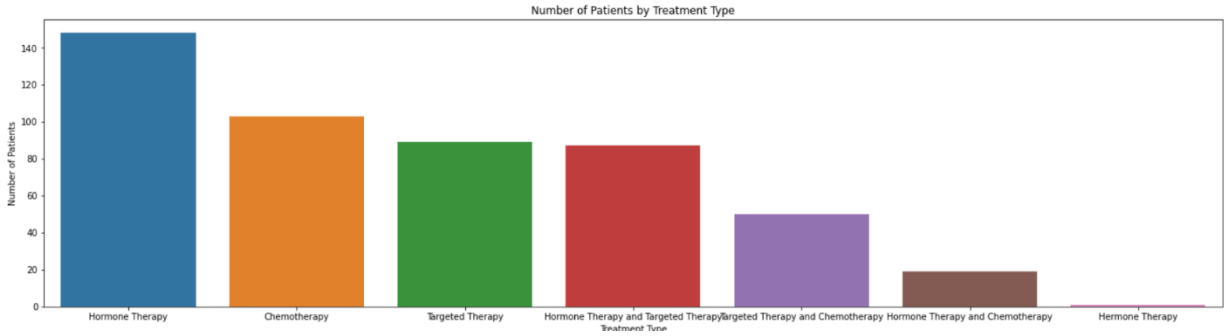

Figure 5-7 Bar plot show distribution of data

5.2.5 Inference

The trained model can be used to predict the treatment for a new patient with the following steps:

1. Create a NumPy array with the new patient's ER, PR, HER2, and Ki-67 (%) values.

2. Use the `predict` method to get the predicted treatment type.

Example:

```
input_data = np.array ([[1, 1, 0, 50]])

prediction = rf.predict (input_data)

print("Predicted treatment:", prediction)
```

5.2.6 Conclusion

The Random Forest Classifier model achieved an accuracy of 96% on the test set, demonstrating its effectiveness in predicting the appropriate treatment for breast cancer patients based on their pathological test results. The feature importance plot and confusion matrix provide insights into the model's decision-making process and potential areas for improvement.

CHAPTER 6

Website Design

In the rapidly evolving healthcare industry, technology plays a pivotal role in enhancing the efficiency and effectiveness of medical professionals. This chapter, "Cancer Care Insight" focuses on creating a website that caters specifically to the needs of doctors, helping them streamline their daily work, improve patient care, and stay connected with their peers.

We begin by exploring the unique requirements and challenges faced by doctors when interacting with digital platforms. Understanding these needs is crucial in designing a website that offers practical solutions, whether it's managing patient records, scheduling appointments, accessing medical research, or facilitating communication within the medical community.

Next, we will delve into the core principles of web design, with a special emphasis on usability, accessibility, and responsive design tailored for healthcare professionals. A doctor’s time is valuable, and a well-designed website must be intuitive, easy to navigate, and accessible on various devices, ensuring that critical information is always at their fingertips.

The chapter will then examine the essential components of a healthcare-focused website. We will discuss effective layout designs, color schemes that promote a professional yet calming atmosphere, and the strategic use of typography and imagery to convey trust and reliability. Real-world examples and case studies will illustrate how these elements come together to create a cohesive and effective web presence.

User experience (UX) and user interface (UI) design take center stage as we discuss how to create a seamless and engaging experience for doctors. We will explore strategies for designing interfaces that facilitate quick access to important features, support decision-making processes, and enhance overall productivity. Understanding the specific workflow of medical professionals will guide you in developing interfaces that are both practical and user-friendly.

6.1 Challenges and Solutions

Doctors face numerous challenges when adapting to new websites, especially those intended to assist in their daily tasks. Understanding these challenges is essential for designing an effective interface that minimizes friction and enhances usability. This section will explore common challenges and demonstrate how our website addresses these issues with a user-friendly interface.

6.1.1 Common Challenges

1. **Complex Navigation**: Doctors often deal with time-sensitive tasks and require quick access to information. Complex or unintuitive navigation can hinder their efficiency and increase frustration.
2. **Information Overload**: A cluttered interface with too much information can overwhelm users, making it difficult to find relevant data quickly.
3. **Learning Curve**: New systems often come with a steep learning curve, which can be time-consuming and counterproductive for busy medical professionals.
4. **Inconsistent Interfaces**: Doctors frequently use multiple systems. Inconsistent interfaces between these systems can cause confusion and errors.
5. **Accessibility Issues**: Healthcare professionals need websites that are accessible on various devices, including desktops, tablets, and smartphones, to facilitate on-the-go access.
6. **Security Concerns**: Ensuring the security and privacy of sensitive medical data is paramount. Complicated security protocols can deter usability.
7. **Integration with Existing Systems**: Difficulty in integrating new websites with existing EHR (Electronic Health Records) systems can disrupt workflow and data continuity.

6.1.2 Our Solutions

1. **Intuitive Navigation**: Our website features a clean, well-organized layout with a simple, intuitive navigation menu. Key functions are easily accessible from the homepage, reducing the number of clicks needed to reach important

information. Clear labels and logical groupings help users find what they need quickly.

2. **Minimalist Design**: We employ a minimalist design approach, focusing on essential elements and avoiding clutter. This reduces cognitive load and allows doctors to concentrate on critical tasks. Important information is highlighted, while secondary details are accessible but not overwhelming.
3. **User-Friendly Tutorials and Support**: To address the learning curve, we provide comprehensive tutorials, quick-start guides, and user support. Interactive walkthroughs and tooltips help users familiarize themselves with the interface without needing extensive training.
4. **Consistent User Interface**: Our design adheres to consistent UI conventions, making it easier for doctors to transition from other systems. Consistency in design elements, such as buttons, icons, and terminology, helps users navigate seamlessly.
5. **Responsive Design**: Our website is fully responsive, ensuring optimal functionality across all devices. Whether on a desktop, tablet, or smartphone, the interface adapts to provide a smooth experience, allowing doctors to access information on the go.
6. **Enhanced Security with Simplicity**: We implement robust security measures while maintaining ease of use. Secure login procedures, data encryption, and compliance with healthcare regulations are integrated without complicating the user experience. Features like single sign-on (SSO) simplify access while ensuring security.
7. **Seamless Integration**: Our website is designed to integrate smoothly with existing EHR systems and other healthcare applications. This ensures continuity of data and streamlines workflow, allowing doctors to access all necessary information in one place.

By addressing these common challenges through thoughtful design and user-centered solutions, our website aims to provide doctors with a tool that enhances their efficiency and effectiveness in their daily work. The easy interface not only simplifies the adoption of new technology but also improves overall user satisfaction, making it an invaluable asset in the healthcare environment.

6.2 User interface

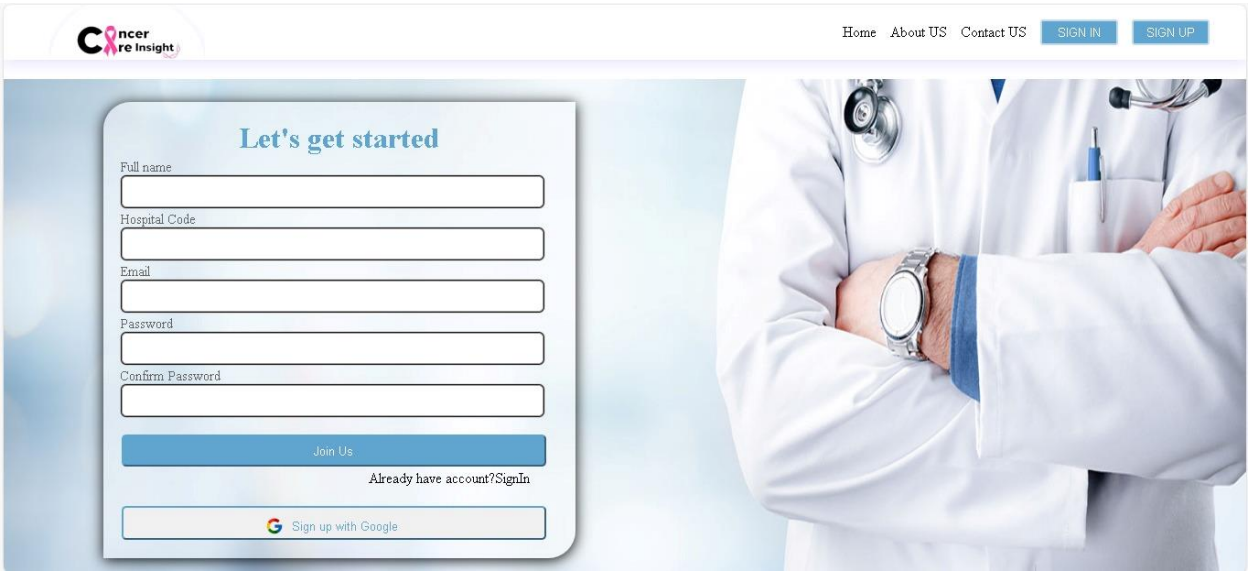

Figure 6-1 Sign-up page

The Sign-Up page on the Cancer Care Insight website enables new users to create an account by entering their name, email, hospital code, and a secure password. It includes fields for confirming the password to ensure accuracy. A "Sign Up" button is provided to complete the registration process. Additionally, there is an option for users who already have an account to sign in, facilitating easy access and user-friendly navigation.

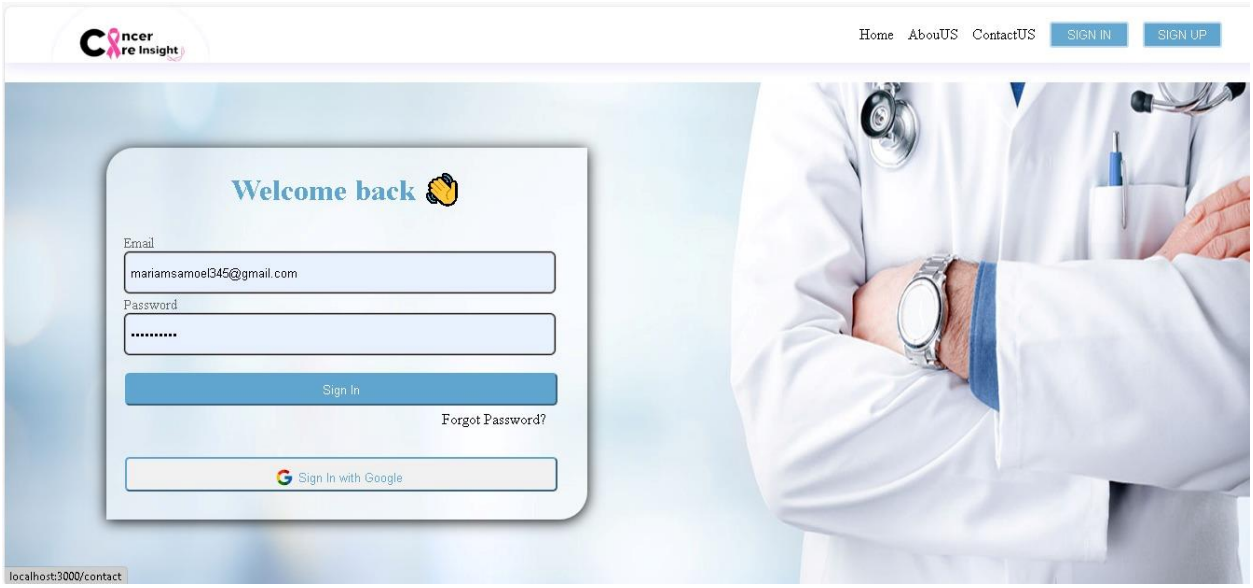

Figure 6-2 Sign-in page

The Sign In page on the Cancer Care Insight website allows users to access their accounts securely using their email and password. It features input fields for email and password, a "Sign In" button, and a link to reset the password if forgotten. Additionally, it provides an option to sign up for new users who don't have an account yet, ensuring easy access and navigation for all users.

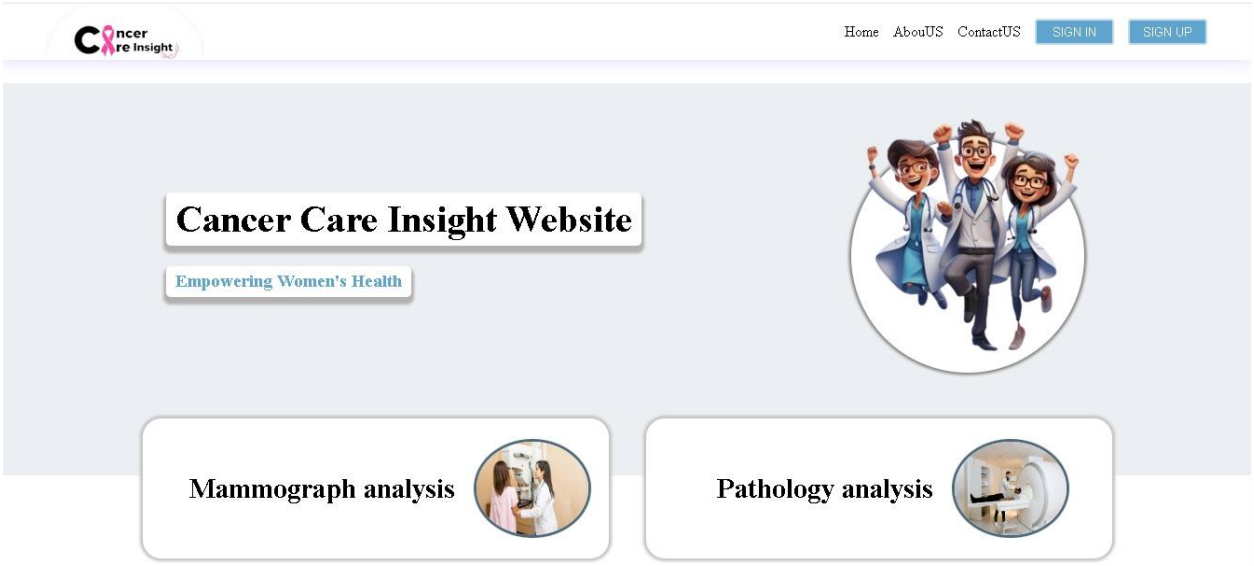

Figure 6-3 Home page

Welcome to the Cancer Care Insight Website, your comprehensive resource for breast cancer care. Our mission is to empower women's health through advanced diagnostic tools and supportive information. Explore our services, including mammograph analysis and pathology analysis, designed to provide accurate and timely insights into breast cancer detection and treatment. Join our community of dedicated healthcare professionals committed to your wellness journey.

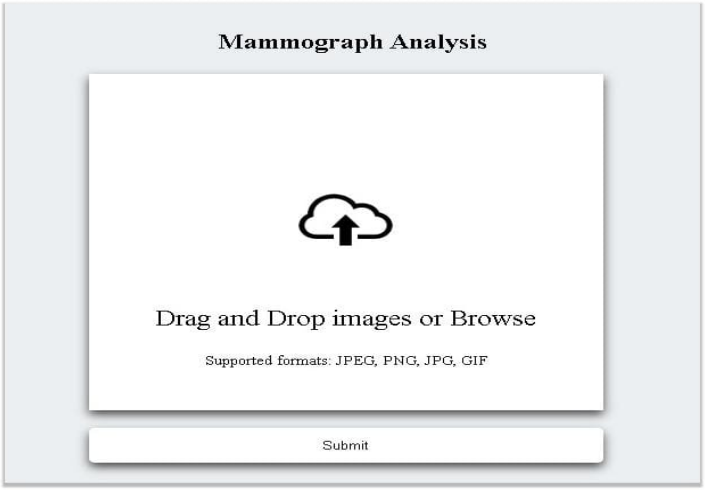

Figure 6-4 Upload mammograph image

Uploading only images in theses format (JPG, PNG, GIF, JPEG).

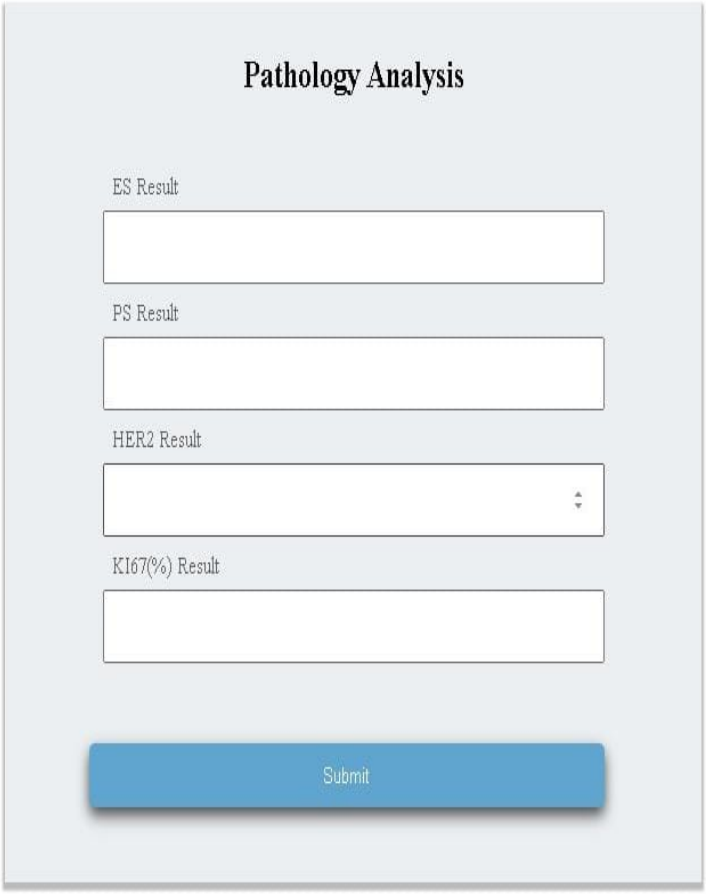

Figure 6-5 Upload pathology test results

Pathology Results Overview: -

Welcome to the Pathology Results section on Cancer Care Insight. Here, you'll find detailed information on key biomarkers essential for understanding breast cancer:

ER (Estrogen Receptor): Indicates sensitivity to hormone therapies targeting estrogen.

PR (Progesterone Receptor): Provides insights into additional hormone receptor status for treatment planning.

Ki-67: Measures cell proliferation rate, helping gauge cancer aggressiveness.

HER2 (Human Epidermal Growth Factor Receptor 2): Identifies HER2 protein overexpression, guiding targeted therapy options.

Explore each biomarker's significance, implications for treatment decisions, and links to related resources to empower informed discussions with your healthcare team.

Figure 6-6 Mammogram result (positive)

The "Mammogram Result - Positive" page on the Cancer Care Insight website provides essential information for patients with positive mammogram results. It includes patient details, test specifics, and a clear summary of the positive result. The page outlines next steps for further diagnostics, offers doctor's notes, and provides resources for support and education. Contact information for the medical team and helpline is also available, ensuring patients receive the guidance and assistance they need promptly.

Figure 6-7 Mammogram result (negative)

The "Mammogram Result - Negative" page on the Cancer Care Insight website provides key information for patients with negative mammogram results. It includes patient details, test specifics, and a clear summary indicating no abnormal tissue was detected. The page offers reassurance, outlines any recommended follow-up actions, and provides educational resources on breast health. Contact information for the medical team and support services is also available, ensuring patients can easily reach out for any further questions or peace of mind.

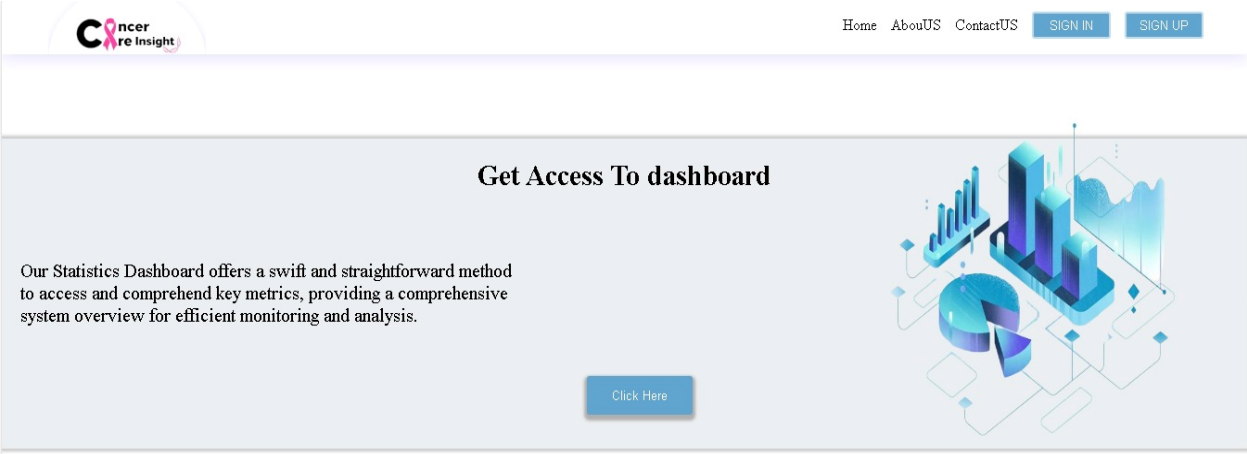

Figure 6-9 Get access to dashboard

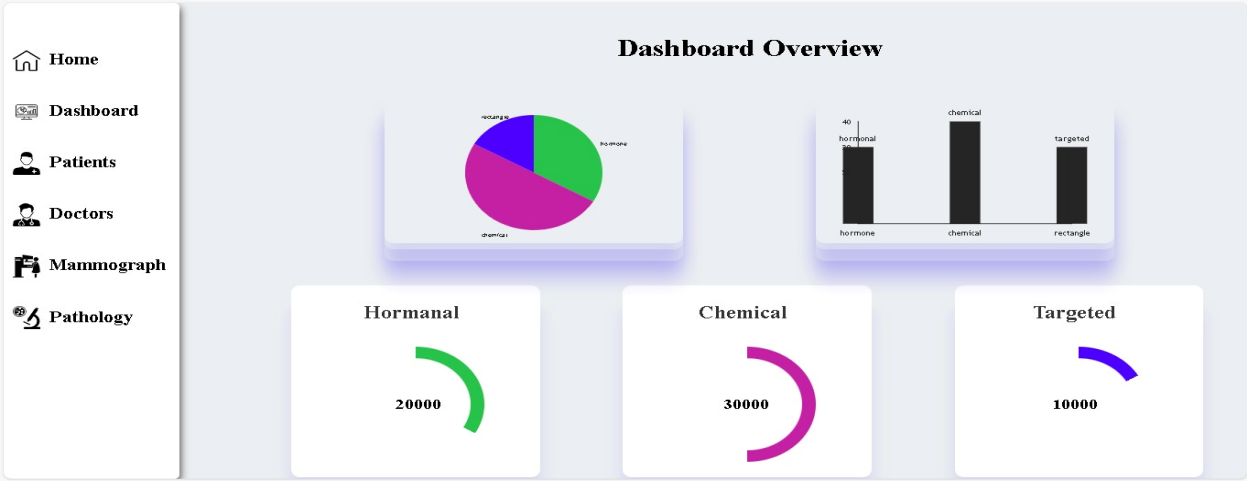

Figure 6-8 Dashboard

Our Statistics Dashboard provides a swift and straightforward method to access and understand key metrics, offering a comprehensive system overview for efficient monitoring and analysis. By clicking the access button, users can delve into detailed statistics and visualizations that aid in informed decision-making and enhanced patient care management.

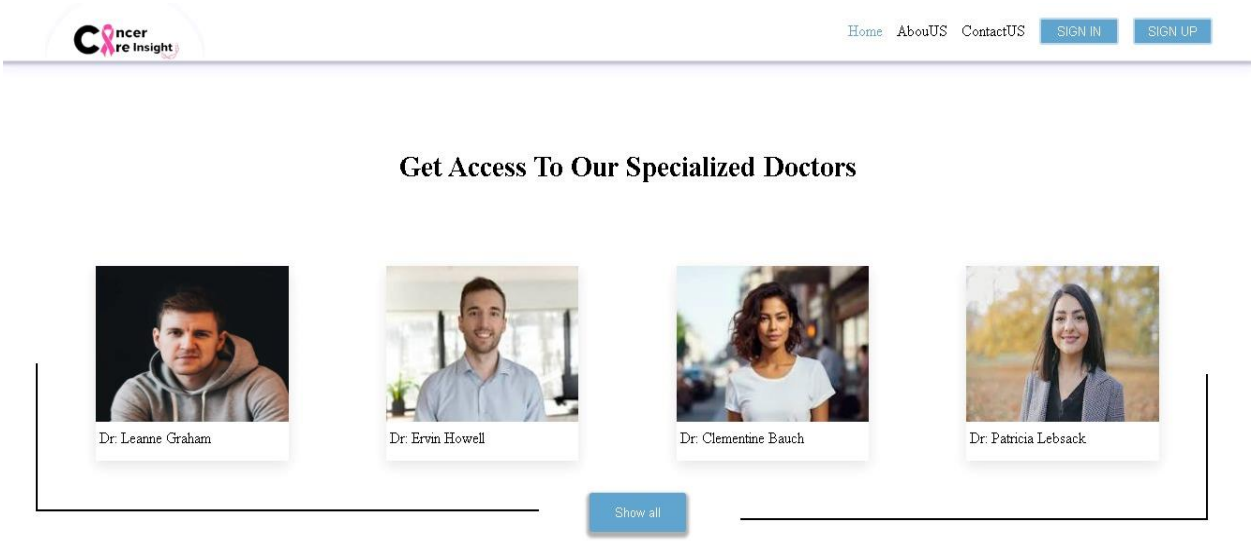

Figure 6-10 Team of specialized doctors

Meet our team of specialized doctors dedicated to providing exceptional breast cancer care. Our experts are here to offer personalized and comprehensive medical support. Click the "Show all" button to view the full list of our skilled healthcare professionals committed to your health and well-being.

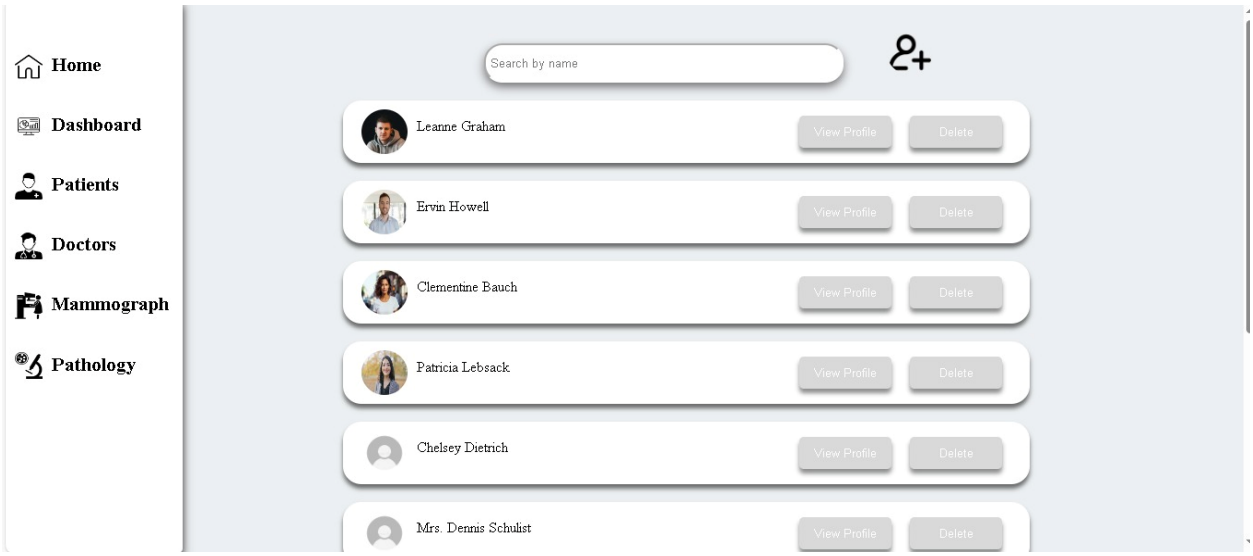

Figure 6-11 Search for a doctor

Figure 6-12 Doctor's information

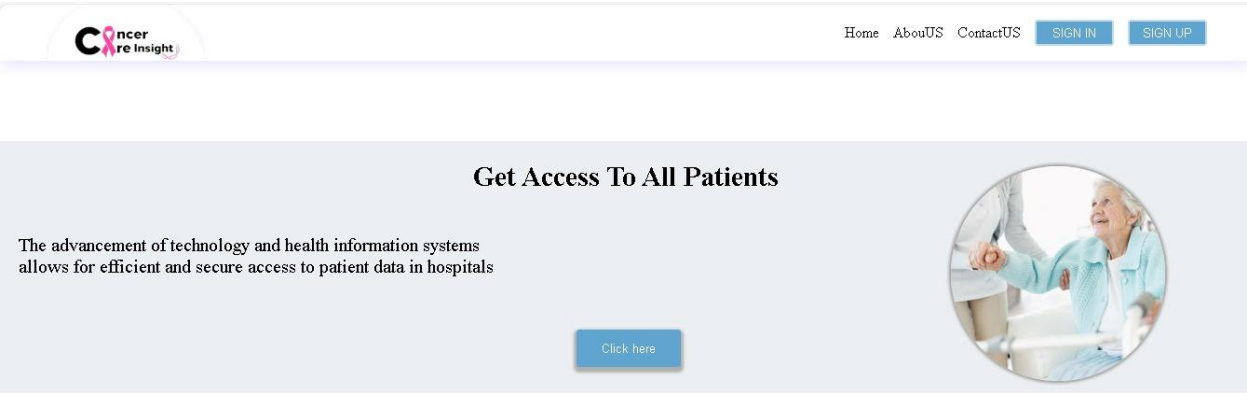

Figure 6-13 Get access to all patients

The "Get Access to All Patients" feature on the Cancer Care Insight website provides healthcare professionals with streamlined, secure access to patient data within hospitals. Utilizing advanced health information systems, this feature enhances efficiency and ensures that providers have the necessary information to deliver optimal care. This innovative approach supports improved patient outcomes by enabling timely and informed decision-making in clinical settings.

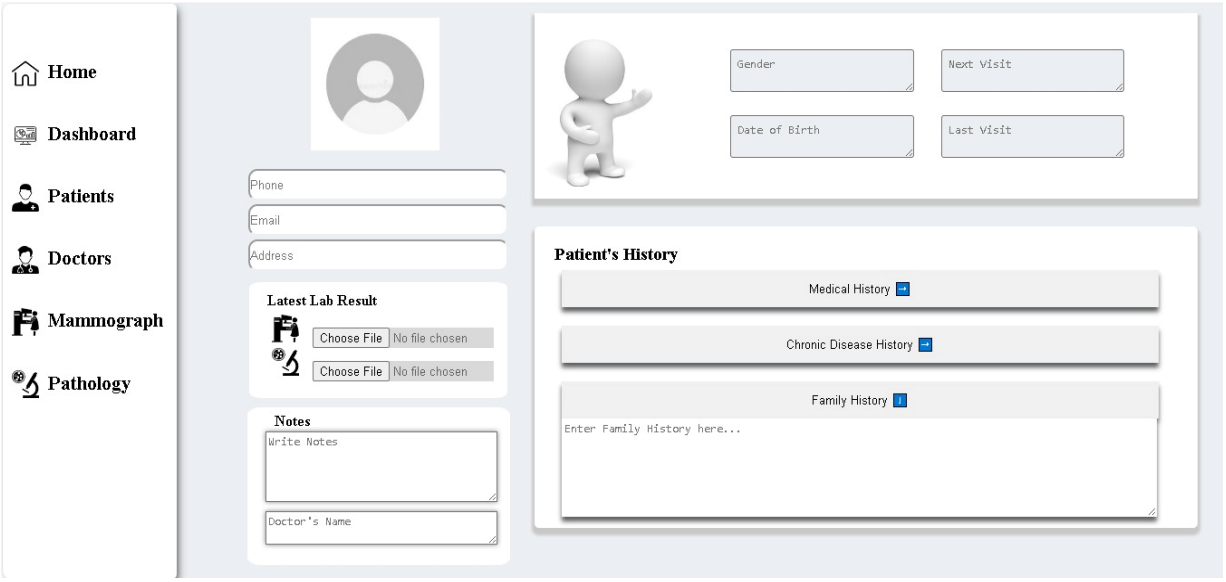

Figure 6-14 Patient dashboard

The "Patient Dashboard" feature on the Cancer Care Insight website offers a comprehensive interface for managing patient information. Key elements include:

Profile Information: Basic details such as phone, email, address, gender, date of birth, and visit history.

Latest Lab Results: Options to upload and access recent laboratory results.

Medical History: Sections to document medical history, chronic disease history, and family history.

Notes: Space for adding notes and doctor's name.

This feature enables efficient organization and access to vital patient information, enhancing the quality of care and facilitating effective communication among healthcare providers.

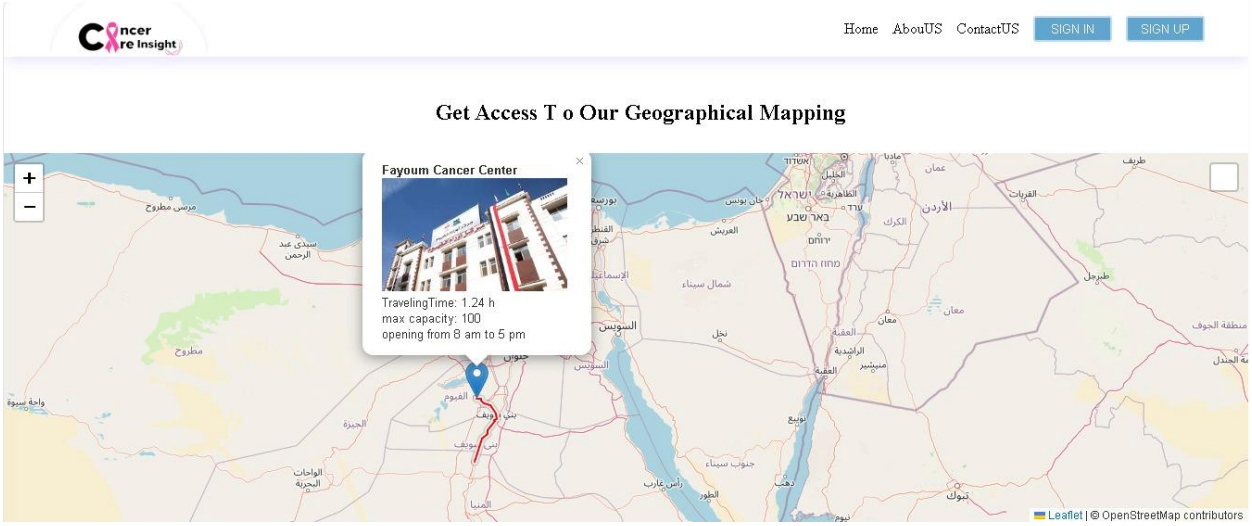

Figure 6-15 Geographical mapping

The "Get Access to Our Geographical Mapping" feature on the Cancer Care Insight website provides users with an interactive map showcasing various cancer care centers. This tool displays essential information for each center, such as traveling time, maximum capacity, and operating hours. This feature aims to help patients and healthcare providers locate and choose the most suitable cancer treatment facilities based on proximity and capacity, enhancing accessibility and planning for cancer care services.

CHAPTER 7

Conclusion And Future Work

In this chapter, we will provide a comprehensive overview of the project as a whole. We will summarize the key aspects, achievements, and contributions of the project. Additionally, we will discuss potential future works that aim to enhance the availability, performance, and security features of the project.

7.1 Conclusion

The main aim of our project is to provide an accessible medical history anywhere anytime for each citizen and keep the patient connecting with doctors, it also enables doctors to track the medical status of each patient. Pharmacies can also introduce available medicines and laboratory can add analyzes which doctor want from patient.

Our system has a lot of objectives for patient, doctors, pharmacies and laboratories that deal with the system

- For patient

1. Patient can get a message on his phone of the results of his tests.
2. Patient can schedule his appointment from any doctor according to his medical status he also can remove or delay this appointment.
3. Patient can keep touch with pharmacies to buy his drugs and decide the way to receive it.

- For Doctor

1. Doctor can register on our medical website system to have his own account using his license.
2. Through his account, he can receive appointments from patients.
3. Doctor can access patient`s medical profile and update it.
4. Doctor can access pharmacies to see whether drugs are available or he needs to order an alternative drug.
5. The doctor can determine the required tests and see them.

6. Through our system, doctor can give the patient suitable drugs which don`t conflict with other diseases the patient may have.

➢ For Pharmacist

1. Pharmacist can register on our medical website system to have his own account using his pharmacy`s license.
2. Pharmacist can set available drugs in his pharmacy.

➢ For laboratory

- Can register on our medical website system to have his own account.

7.2 Future work

Medical management and business processes in medical fields have changed considerably over the previous years, as did the use of Information Technology (IT) in their daily works. It was found in our analysis that the use of IS in the medical fields did not develop according to the needs and developments in the medical organizations over the past decade.

We put a plan to develop our website:

1. Development of Deep Learning models website development launch of the website milestone.
2. User base expansion collaboration with medical institutions establishment of partnerships with pharmaceutical companies.
3. Advanced personalization features implementation of user feedback mechanism.
4. Expansion of disease scope peer-reviewed publications.
5. Integration of Telemedicine services achievement of global recognition.

References

1- Hoteit, Hawraa, et al. "Breast Abnormalities' Classification Using Convolutional Neural Network." 2022 International Conference on Smart Systems and Power Management (IC2SPM). IEEE, 2022.

2- Wei, Benzheng, et al. "Deep learning model based breast cancer histopathological image classification." 2017 IEEE 2nd international conference on cloud computing and big data analysis (ICCCBDA). IEEE, 2017.

3- Horr, Christina, and Steven A. Buechler. "Breast Cancer Consensus Subtypes: A system for subtyping breast cancer tumors based on gene expression." NPJ breast cancer 7.1 (2021): 136.

4- Gupta, Surbhi, et al. "Deep learning techniques for cancer classification using microarray gene expression data." Frontiers in Physiology 13 (2022): 952709.

5- Gami, Bhavin, Khushi Chauhan, and Brijeshkumar Y. Panchal. "Breast Cancer Detection Using Deep Learning." Mobile Radio Communications and 5G Networks: Proceedings of Third MRCN 2022. Singapore: Springer Nature Singapore, 2023. 85-95.

6- Dataset (Available online): https://www.kaggle.com/datasets/paultimothymooney/breast-histopathology-images

7- GEO. Accessed: 2024. [online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8522027/

8- GEO. Accessed: 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8430879/

9- Gamble, Paul, et al. "Determining breast cancer biomarker status and associated morphological features using deep learning." Communications medicine 1.1 (2021): 14.