#### ACKNOWLEDGEMENT

First and foremost, we would express our deepest gratitude and appreciation to our advisor **Dr. Ahmed Hesham** for his support, outstanding guidance, and encouragement throughout our graduation project.

We would also like to thank our families, especially our parents, for their encouragement, patience, and assistance over the years. We are forever indebted to our parents, who have always kept us in their prayers and pushed us to Success. Finally, for our faculty for providing the suitable environment that leaded us to represent the best image that computer science graduates of Helwan University are supposed to represent.

#### Table Of Content

Contents

[Table Of Content **3**](#_bookmark0)

[Abstract **5**](#_bookmark1)

1. [Introduction 6](#_bookmark2)
   1. [Overview 6](#_bookmark3)
   2. [Problem statement 7](#_bookmark4)
   3. [Objectives 7](#_bookmark5)
   4. [Scope of work 8](#_bookmark6)
   5. [General constraints: 8](#_bookmark7)
   6. [Purpose 9](#_bookmark8)
2. [project “planning and analysis” 10](#_bookmark9)
   1. [System Development Life Cycle 10](#_bookmark10)
   2. [Methodology 11](#_bookmark11)
   3. [Software development phases: 11](#_bookmark12)
      1. [Planning 11](#_bookmark13)
      2. [Analysis and Limitation of existing system 12](#_bookmark14)
      3. [Need for the new system 13](#_bookmark15)
      4. [Analysis 15](#_bookmark16)
3. [Software design 22](#_bookmark17)
   1. [Class diagram 22](#_bookmark18)
      1. [Option 1 22](#_bookmark19)
      2. [Option 2 23](#_bookmark20)
      3. [Class diagram description 24](#_bookmark21)
   2. [Use Case diagram 25](#_bookmark22)
      1. [Login $ Registration 25](#_bookmark23)
      2. [System Use Cases 27](#_bookmark24)
   3. [Sequence diagram 30](#_bookmark25)
      1. [Login 30](#_bookmark26)
      2. [Register 31](#_bookmark27)
      3. [Symptoms checker 32](#_bookmark28)
      4. [Prescription 34](#_bookmark29)
      5. [Appointment 35](#_bookmark30)
      6. [Medical Image 36](#_bookmark31)
   4. [Activity Diagrams 37](#_bookmark32)
      1. [System Activity Diagram 37](#_bookmark33)
      2. [Login/register Activity Diagram 38](#_bookmark34)
4. [Implementation 40](#_bookmark35)
   1. [Machine and deep learning models: 40](#_bookmark36)
   2. [Flow Chart 83](#_bookmark37)
   3. [Pseudocode 84](#_bookmark38)
5. [Testing 86](#_bookmark39)
   1. [Functional Testing: 86](#_bookmark40)
   2. [Unit testing: 86](#_bookmark41)
   3. [Integration testing: 87](#_bookmark42)
   4. [Additional testing: 88](#_bookmark43)
6. [Graduation Project 2023 90](#_bookmark44)
   1. [Results and Discussion 90](#_bookmark45)
      1. [Expected Result 90](#_bookmark46)
      2. [Actual Result: 90](#_bookmark47)
   2. [Discussion 91](#_bookmark48)
7. [Conclusion 92](#_bookmark49)
8. [Future work 94](#_bookmark50)

#### Abstract

The proposed medical system is designed to provide a wide range of features to users, including a symptom checker that can diagnose diseases based on a series of questions, the ability to book appointments with doctors, and the option to upload prescriptions and receive reminders for treatment dates. Additionally, the system allows users to upload medical images to aid in the diagnosis of diseases. The system provides users with a convenient and efficient way to manage their healthcare needs and make informed decisions about their health.

# Chapter one:

Introduction

In this chapter we’re going to discuss and go deeper in the overview of the project and know more about its scope and limitations and explain some terminologies we will find throughout the document.

## Introduction

### Overview

The proposed medical system is an all-in-one platform that offers a range of features to help users manage their healthcare needs. The system includes a symptom checker that can diagnose diseases by asking users a series of questions related to their symptoms. This feature allows users to obtain an initial diagnosis and determine whether they need to seek medical attention from a doctor.

In addition to the symptom checker, the system offers clinical management features, allowing users to book appointments with doctors easily. Users can also upload prescriptions to the system, and the system will

remind them of their treatment dates, ensuring they do not miss their appointments or medication. This feature helps users to manage their treatment plan more efficiently and reduces the risk of missing

important appointments.

Moreover, the system allows users to upload medical images to aid in the diagnosis of diseases. This feature enables doctors to review the images and provide an accurate diagnosis, helping to save users time and money by eliminating the need for additional medical tests.

Overall, the proposed medical system provides users with a comprehensive and convenient way to manage their healthcare needs. The system's features enable users to make informed decisions about their health and receive timely medical care, ultimately improving their overall health outcomes.

### Problem statement

The healthcare industry faces several challenges, including the need for efficient and accurate diagnosis of diseases, effective management of clinical appointments and medication, and timely access to medical care. These challenges can result in delayed diagnosis, missed appointments, and reduced health outcomes for patients.

To address these challenges, the proposed medical system aims to provide a comprehensive solution that can diagnose diseases efficiently and accurately using a symptom checker. The system also aims to provide clinical management features, allowing users to book appointments with doctors easily and manage their treatment plans effectively. Additionally, the system allows users to upload medical images to aid in the diagnosis of diseases and receive timely medical care.

The problem statement for the medical system is the lack of a comprehensive and convenient platform that provides users with efficient and accurate diagnosis of diseases, effective clinical management, and timely access to medical care. The proposed medical system aims to address these challenges and provide users with a reliable and efficient way to manage their healthcare needs, ensuring better health outcomes for patients.

### Objectives

* + - To provide an efficient and accurate diagnosis of diseases using a symptom checker, enabling users to identify potential health issues and seek medical attention promptly.
    - To facilitate clinical management by allowing users to book appointments with doctors easily and manage their treatment plans effectively, ensuring timely access to medical care.
    - To provide a user-friendly platform for users to upload prescription information, allowing the system to remind them of their treatment dates and reduce the risk of missed appointments and medication.
    - To enable users to upload medical images to the system to aid in the diagnosis of diseases, enabling doctors to provide accurate diagnoses and reduce the need for additional medical tests.
    - To improve the overall health outcomes of users by providing a comprehensive and convenient platform to manage their healthcare needs, reducing the risk of delayed diagnosis, missed appointments, and reduced health outcomes.
    - To ensure the privacy and security of user data by implementing robust security measures, such as data encryption and access controls, and complying with relevant data protection laws and regulations.

#### Scope of work

* Planning: This involves identifying the project's objectives, defining the system's requirements, and outlining the project's timeline, budget, and resources.
* Designing: This involves creating a user-friendly interface for the system, developing the symptom checker, designing the clinical management features, and creating a secure platform for users to upload their medical information.
* Coding: This involves developing the system's functionalities, including the symptom checker, clinical management features, and medical image upload feature. The coding process will adhere to industry best practices and standards and will prioritize the system's security and performance.
* Testing: This involves testing the system's functionalities and identifying and fixing any bugs or issues that may arise. This will include testing the symptom checker's accuracy, clinical management features, prescription upload feature, and medical image upload feature.
* Documentation: This involves creating user manuals, system specifications, and technical documentation for the system. This documentation will provide users with clear instructions on how to use the system and will help to ensure the system's maintainability and scalability.

#### General constraints:

* Regulatory compliance: The system must adhere to relevant regulatory requirements, such as data protection laws and regulations, to ensure the privacy and security of user data.
* Technical constraints: The system must be developed using appropriate technologies and adhere to best practices for system architecture, security, and performance.
* Time constraints: The project must be completed within a specified timeline to ensure timely delivery of the system.
* Resource constraints: The project must be completed within a specified budget and resource allocation.
* User accessibility: The system must be designed to accommodate users with varying levels of technological expertise and accessibility needs, such as users with disabilities.
* Integration constraints: The system must integrate with existing healthcare systems, such as electronic health record systems, to ensure compatibility and interoperability.
* Accuracy and reliability: The symptom checker and medical image upload feature must be accurate and reliable to ensure that users receive appropriate diagnoses and treatment.
* User data security: The system must ensure the privacy and security of user data, including medical records, prescriptions, and personal information.

these constraints must be considered when developing the medical system to ensure that it meets the needs of users and regulatory requirements while maintaining technical standards and adhering to project constraints.

#### Purpose

The main Purpose of our project is to help people to maintenance their health by frequently check up.

# Chapter Two:

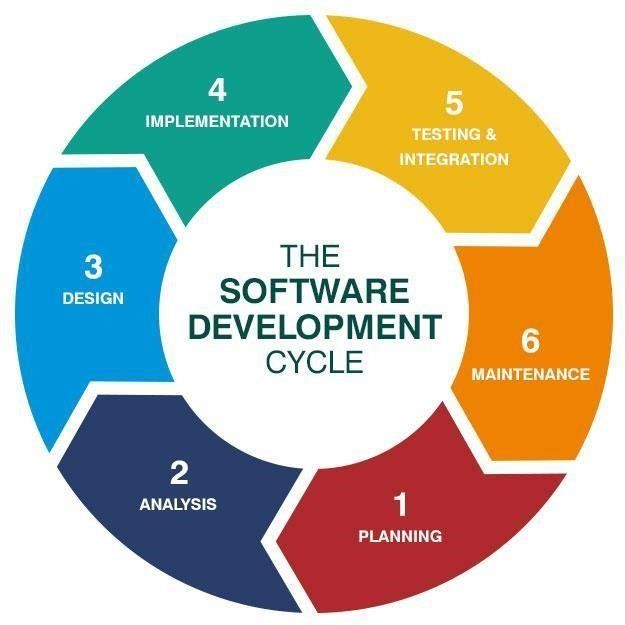
Planning and analysis

In this chapter we’re going to discuss and go deeper in the overview of the project and know more about its scope and limitations, and explain some terminologies we will find throughout the document

## project “planning and analysis”

### System Development Life Cycle

We followed system development life cycle approach in developing this project with its 7 phases, planning, analysis, design, development, testing, deployment and maintenance.



### Methodology

We followed Agile analysis and design methodology in developing, this project and developed all related UML diagrams, represented in this document later.

### Software development phases:

* + 1. **Planning**

Very important phase through the phases to build our software. In this phase we will define the problem and show the following:

* + - * Define problem.
      * Define Project Scope and purpose.
      * Estimate the cost, schedule.
      * Identify the project team, team leader, and project manager.
      * How will the software be used?
      * What data will serve as the input of the software?
      * What will be the output of the software?
      * Who is going to use the software (who the stakeholders)?

### Analysis and Limitation of existing system

#### User Experience Analysis:

Evaluate the user experience of the existing system by gathering feedback from users. Assess factors such as ease of use, intuitiveness of the interface, clarity of instructions, and overall satisfaction with the system's functionalities.

#### Functionality Evaluation:

Analyze the effectiveness and accuracy of the symptom checker in diagnosing diseases based on user input. Assess the system's ability to provide relevant recommendations or next steps. Evaluate the appointment booking process for its efficiency and ease of use.

#### Performance Assessment:

Evaluate the system's performance in terms of response time, availability, and scalability. Identify potential bottlenecks or performance issues that may impact user experience or system reliability.

#### Data Management and Security:

Assess the existing system's data management practices, including how it handles and stores user data, prescriptions, and medical images. Evaluate the system's security measures to protect user data from unauthorized access or breaches.

#### Diagnostic Accuracy and Reliability:

Evaluate the accuracy and reliability of the system's medical image diagnosis feature. Assess the success rate of correctly identifying diseases or abnormalities from uploaded medical images. Analyze any limitations or challenges faced in accurately interpreting the images.

#### Integration and Interoperability:

Analyze the system's ability to integrate with other healthcare systems or electronic health record (EHR) systems. Assess the level of interoperability and data exchange capabilities to ensure seamless communication and continuity of care.

#### System Limitations:

Identify any limitations or shortcomings of the existing system, such as lack of support for specific medical conditions, insufficient medical knowledge database, or limited availability of healthcare professionals for appointments.

#### User Feedback and Satisfaction:

Analyze user feedback and satisfaction levels through surveys, interviews, or user feedback mechanisms. Identify areas where users express dissatisfaction or suggest improvements to enhance the system's functionality and user experience

Based on the analysis, limitations of the existing system can be identified. These limitations can include technical issues, user experience challenges, inaccuracies in diagnosis, limitations in data management, security vulnerabilities, and lack of comprehensive medical knowledge. Addressing these limitations may involve system enhancements, such as improving the accuracy of the symptom checker, enhancing the user interface for a more intuitive experience, implementing stronger security measures, expanding the medical knowledge database, and ensuring compatibility with other healthcare systems. It is crucial to consider user feedback, engage with healthcare professionals, and conduct thorough testing and validation to mitigate the limitations and improve the overall performance and effectiveness of the medical system.

### Need for the new system

#### Improved Accessibility:

The new system provides users with increased accessibility to healthcare services. Users can access the system from anywhere, at any time, eliminating the need for physical visits to healthcare facilities for preliminary assessments or appointment bookings.

#### Efficiency and Time Savings:

The new system streamlines the process of diagnosing diseases and booking appointments. Users can conveniently use the symptom checker to assess their symptoms and receive prompt recommendations. They can also schedule appointments with doctors directly through the system, saving time and effort.

#### Enhanced Diagnostic Capabilities:

The new system includes the ability to upload medical images for disease diagnosis. This feature expands the diagnostic capabilities beyond the symptom checker, allowing healthcare professionals to review and analyze uploaded images to make accurate diagnoses.

#### Better Treatment Management:

The new system enables users to upload prescriptions and stores them for future reference. It also includes a treatment date reminder feature, ensuring users remember their scheduled treatments. This improves treatment adherence and overall management of healthcare appointments and prescriptions.

#### Centralized Health Information:

The new system centralizes user health-related information, including symptoms, medical history, prescriptions, and medical images. This allows healthcare professionals to access comprehensive user profiles, leading to better-informed decision-making and improved continuity of care.

#### Empowerment and Patient Engagement:

The new system empowers users to actively participate in their healthcare journey. It provides them with valuable information, recommendations, and tools to make informed decisions about their health. This promotes patient engagement and self-care. By introducing the new system, healthcare organizations can improve the efficiency of their operations, enhance patient experiences, increase accessibility to healthcare services, and promote better health outcomes through early detection and timely treatment.

#### Technological Advancements:

The new system takes advantage of advancements in technology, such as artificial intelligence and data analytics. These technologies enhance the accuracy of symptom checking, medical image analysis, and appointment scheduling, leading to more precise diagnoses and improved healthcare outcomes.

#### Improved communication:

Patients can communicate with doctors and healthcare providers through the system, which can improve communication and reduce the risk of miscommunication.

#### Faster diagnosis:

By uploading medical images to the system, doctors can quickly and accurately diagnose diseases, which can lead to faster treatment and better outcomes.

### Analysis

#### User requirements

* + - * 1. User-friendly interface: The system should have an easy-to-use interface that is accessible to people of all ages and backgrounds. It should be intuitive and straightforward to navigate.
        2. Symptom checker: The system should have a symptom checker that allows users to answer some questions about the symptoms to predict diseases.
        3. Appointment booking: The system should allow users to book appointments with doctors or healthcare providers. Users should be able to see the availability of doctors and book appointments at their convenience.
        4. Prescription management: The system should allow users to upload their prescription and keep track of their medications. The system should also send reminders to users when it's time to take their medication.
        5. Medical image diagnosis: The system should allow users to upload medical images and receive a diagnosis from a doctor or healthcare provider. The system should be able to handle different types of medical images, such as X-rays, CT scans, and MRI scans.
        6. Security and privacy: The system should be secure and protect users' personal and medical information. The system should comply with all relevant laws and regulations, such as HIPAA.
        7. Availability and support: The system should be available 24/7 and provide support to users when needed. The system should have a help desk or customer support team to assist users with any issues or questions they may have.

#### Software requirement

* + - * 1. Database: The system should have a database that stores user information, medical history, appointments, prescriptions, and medical images. The database should be secure, reliable, and scalable.
        2. Symptom checker algorithm: The system should have an accurate and reliable symptom checker algorithm that can provide potential diagnoses based on user inputs. The algorithm should be regularly updated and maintained by healthcare professionals.
        3. Appointment scheduling module: The system should have an appointment scheduling module that allows users to book appointments with doctors or healthcare providers. The module should integrate with the system's database and provide real-time availability of doctors' schedules.
        4. Prescription management module: The system should have a prescription management module that allows users to upload their prescription and keep track of their medications. The module should integrate with the system's database and send reminders to users when it's time to take their medication.
        5. Medical image diagnosis module: The system should have a medical image diagnosis module that allows users to upload medical images and receive a diagnosis from a healthcare provider. The module should be able to handle different types of medical images and integrate with the system's database.
        6. Security and privacy measures: The system should have security and privacy measures in place to protect users' personal and medical information. These measures should comply with all relevant laws and regulations, such as HIPAA.
        7. User interface: The system should have a user-friendly interface that is easy to use and navigate. The interface should be designed to provide a seamless user experience and integrate all the system's modules.

##### Domain requirements

* + - * 1. Healthcare domain knowledge: The system should have a deep understanding of the healthcare domain, including medical terminology, diseases, symptoms, treatments, and medications. The system should be developed and maintained by healthcare professionals, such as doctors, nurses, and pharmacists.
        2. Regulatory compliance: The system should comply with all relevant regulations and standards in the healthcare industry, such as HIPAA, GDPR, and FDA guidelines. The system should also comply with local laws and regulations in different jurisdictions.
        3. Medical ethics: The system should adhere to medical ethics principles, such as patient autonomy, confidentiality, informed consent, and non-maleficence. The system should prioritize the wellbeing and privacy of patients.
        4. Interdisciplinary collaboration: The system should facilitate interdisciplinary collaboration among healthcare professionals, such as doctors, nurses, pharmacists, and radiologists. The system should support communication and information sharing among different specialists.
        5. User-centered design: The system should be designed with a user centered approach that takes into account the needs, preferences, and limitations of different user groups, such as patients, caregivers, and healthcare providers. The system should be accessible to users with different abilities and backgrounds.
        6. Continuous improvement: The system should be continuously improved based on user feedback, clinical evidence, and technological advancements. The system should be updated regularly to reflect the latest medical knowledge and best practices.

##### Functional Requirements

|  |  |
| --- | --- |
| **1. Symptom checker module:** | * Allows users to input their symptoms and receive potential diagnoses. * Uses an accurate and reliable algorithm based on medical knowledge and clinical evidence. * Provides users with information on potential diagnoses, including causes, symptoms, and treatments. * Offers users the option to schedule an appointment with a doctor or healthcare provider based on their potential diagnosis |
| **2. Appointment scheduling module:** | * Allows users to book appointments with doctors or healthcare providers * Integrates with the system's database to provide real-time availability of doctors' schedules * Allows users to select their preferred doctor or healthcare provider based on their availability and expertise * Sends reminders to users before their appointment to reduce no shows and improve attendance rates |
| **3. Prescription management module:** | * Allows users to upload their prescription and keep track of their medications * Integrates with the system's database to provide a comprehensive view of users' medical history and medications * Sends reminders to users when it's time to take their medication * Allows users to request refills of their medication and sends notifications to healthcare providers |

|  |  |
| --- | --- |
| **4. Medical image diagnosis module:** | * Allows users to upload medical images and receive a diagnosis from a doctor or healthcare provider * Supports different types of medical images, such as X-rays, CT scans, and MRI scans * Integrates with the system's database to provide a comprehensive view of users' medical history and diagnostic results * Offers users the option to schedule an appointment with a doctor or healthcare provider based on their diagnostic results |
| **5. User account management module:** | * Allows users to create and manage their accounts * Provides users with a dashboard to view their medical history, appointments, medications, and diagnostic results * Enables users to update their personal and medical information, such as their contact details, insurance information, and allergies * Offers users the option to delete their account and remove their personal and medical information from the system |

* + - 1. **Non- Functional Requirements:**

|  |  |
| --- | --- |
| **1.**  **Performance:** | * The system should be fast and responsive to provide a seamless user experience * The symptom checker algorithm should provide potential diagnoses within seconds * The appointment scheduling module should load quickly and enable users to book appointments efficiently * The system should be able to handle large volumes of data, such as medical images and diagnostic reports |
| **2. Reliability:** | * The system should be reliable and available 24/7 to provide continuous service * The system should have backup and recovery mechanisms to prevent data loss and minimize downtime * The system should be able to handle errors and exceptions gracefully to minimize disruptions to users |

|  |  |
| --- | --- |
| **3. Security:** | * The system should be secure and protect users' personal and medical information from unauthorized access and breaches * The system should comply with relevant security standards, such as HIPAA, GDPR, and ISO 27001 * The system should use encryption and authentication mechanisms to ensure confidentiality, integrity, and availability of data |
| **4. Usability:** | * The system should be easy to use and navigate for users of different ages, backgrounds, and abilities * The system should have a user-friendly interface that provides clear and concise information to users * The system should provide users with guidance and feedback to help them complete tasks efficiently and effectively |
| **5. Scalability:** | * The system should be scalable and able to handle growing demands from users and healthcare organizations. * The system should be able to accommodate new features and functionalities without compromising performance, reliability, or security * The system should use cloud-based infrastructure and technologies to enable flexible and costeffective scaling |
| **6.**  **Interoperability:** | * The system should be interoperable and able to exchange data with other healthcare systems, such as electronic health records (EHRs), medical devices, and telemedicine platforms * The system should support different data formats and standards to facilitate data sharing and collaboration among healthcare organizations |

##### Advantages of the new system:

* + - * 1. Improved access to healthcare: The new system allows users to access healthcare services from the comfort of their homes or workplaces, without the need to travel to clinics or hospitals. This is especially beneficial for users with mobility issues, chronic conditions, or busy schedules.
        2. Faster diagnosis and treatment: The symptom checker algorithm can provide potential diagnoses within seconds, enabling users to receive timely and accurate diagnosis and treatment. The appointment scheduling module and prescription management module can also streamline the process of booking appointments and managing medications, reducing waiting times and improving healthcare outcomes.
        3. Enhanced efficiency and productivity: The new system can automate many administrative tasks, such as appointment scheduling, prescription management, and medical image diagnosis, freeing up healthcare professionals' time to focus on patient care and clinical decision-making. This can improve efficiency and productivity in healthcare organizations and reduce healthcare costs.
        4. Better patient engagement and satisfaction: The new system can improve patient engagement and satisfaction by providing users with easy and convenient access to healthcare services, transparent information on potential diagnoses and treatment options, and personalized care plans. The system can also send reminders to users about their appointments and medications, improving adherence and continuity of care.
        5. Enhanced collaboration and communication among healthcare professionals: The new system can facilitate interdisciplinary collaboration and communication among healthcare professionals, such as doctors, nurses, pharmacists, and radiologists, by providing a secure and private channel for communication and information sharing. This can improve the quality and effectiveness of healthcare services and reduce medical errors.
        6. Improved data analytics and insights: The new system can provide healthcare professionals with data analytics and insights on patient trends, disease patterns, and treatment outcomes, enabling them to make informed decisions and improve healthcare quality. This can also support research and development in the healthcare industry

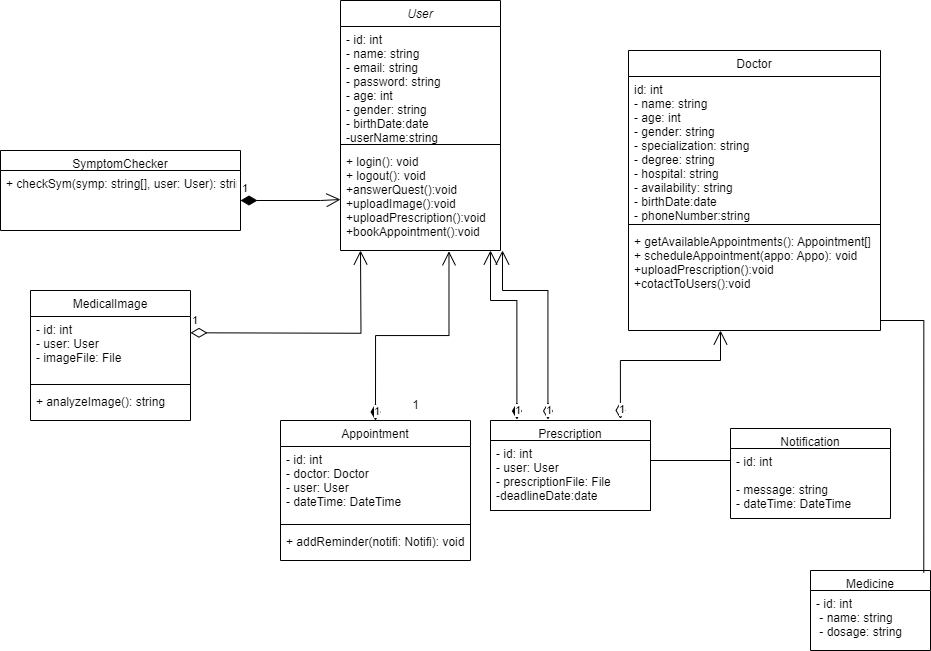
# Chapter Three: Software Design

In this chapter we’re going to discuss and go deeper in how we plan the project and show the steps and the instructions that we’ve followed to plan the application.

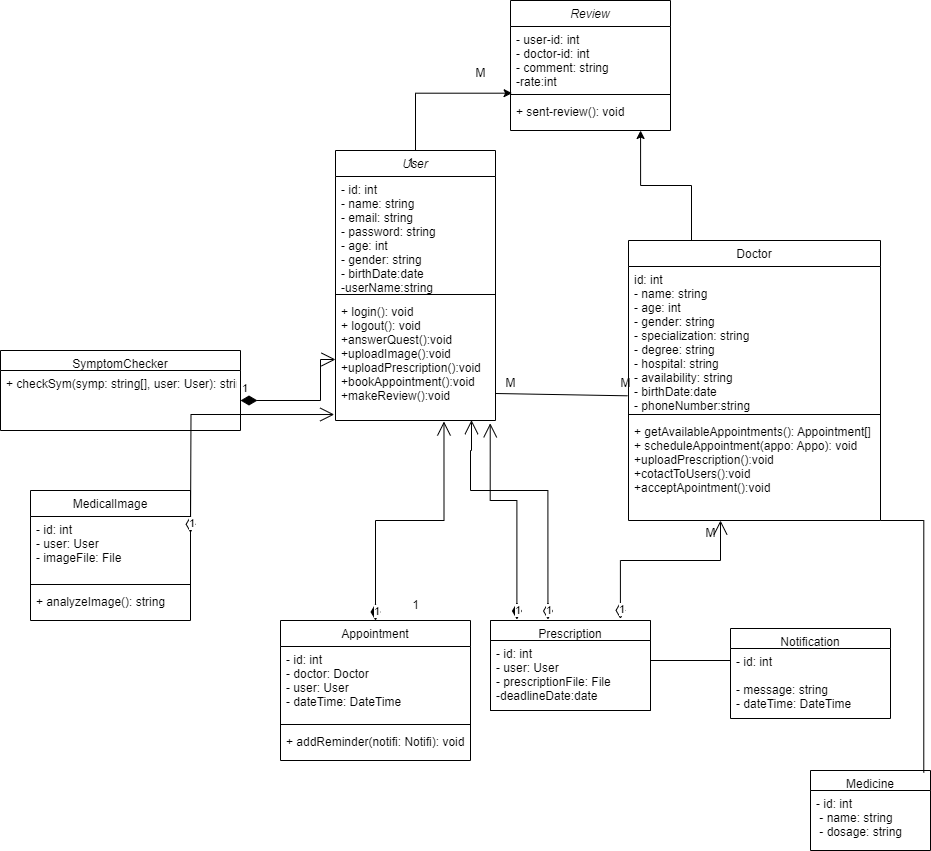
## Software design

### Class diagram

#### Option 1



* + 1. **Option 2:**



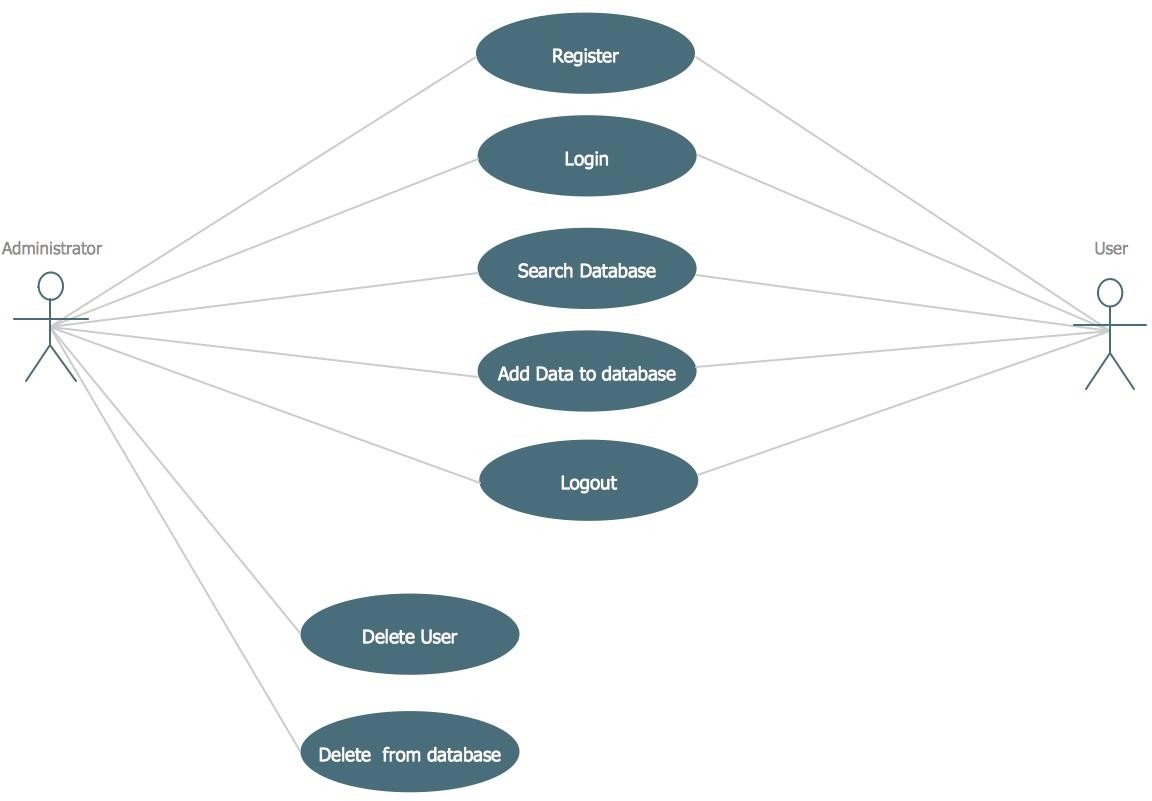
#### Class diagram description:

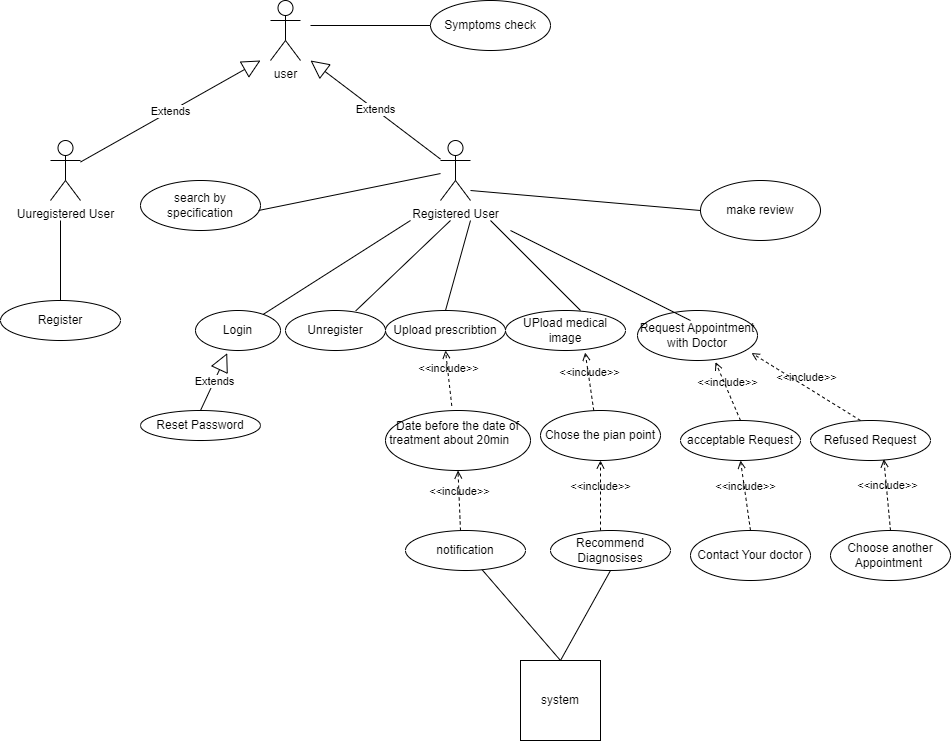
Here is a description of the updated classes and their attributes and methods:

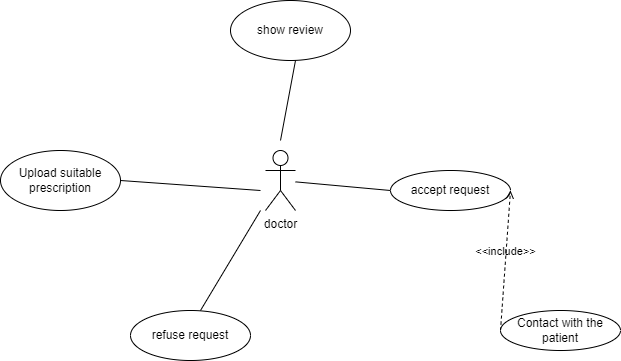
|  |  |
| --- | --- |
| **1. User.** | : This class represents the user of the system and has attributes such  as id, name, email, password, age, and gender. It has methods such as login, logout, and register |
| **2. SymptomChecker:.** | This class represents the symptom checker component of the system. It has methods such as getDiagnosis and analyzeSymptoms |
| **3. Appointment:.** | This class represents the appointment booked by the user. It has attributes such as patient, doctor, date, and location. It has methods such  as createAppointment, cancelAppointment, and getAppointmentDetails |
| **4. Doctor:.** | This class represents the doctor involved in the appointment. It has attributes such as id, name, speciality, and schedule. It has methods such as getSchedule, checkAvailability, and interpretImage |
| **5. Prescription:** | This class represents the prescription uploaded by the user. It has attributes such as patient, doctor, medication, dosage, and treatmentDate. It has methods such as createPrescription, updatePrescription, and getPrescriptionDetails. |
| **6. MedicalImage:** | This class represents the medical image uploaded by the user. It has attributes such as patient, doctor, image, and diagnosis. It has methods such as uploadImage, interpretImage, and getDiagnosis. |

### Use Case diagram

#### Login & Registration:







In this use case scenario diagram, the user actor interacts with the system to perform various tasks. They can use the "Symptom Checker" to input their symptoms and then receive a diagnosis through the "Diagnose Diseases" use case. They can also book an appointment with a doctor through the "Book Appointment" use case, upload a prescription using the "Upload Prescription" use case, and upload a medical image for diagnosis using the

"Upload Medical Image" use case. Additionally, the user can directly choose to diagnose diseases without using the symptom checker.

The "Book Appointment" use case involves the user requesting an appointment with a doctor, who is represented by the doctor actor.

* + 1. **System Use Cases**

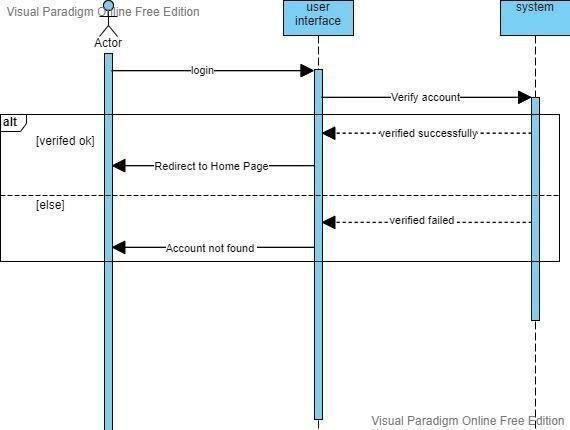
|  |  |
| --- | --- |
| **Identifier and Name:** | |
| **Initiator:** | User |
| **Goals:** | help people to diagnosis their diseases by questions |
| **Precondition:** | register in the application |
| **Postcondition:** | Get accurate diagnosis |

|  |  |  |  |
| --- | --- | --- | --- |
| **use case name Precondition: Postcondition: Actor:** | | | |
| **1. login** | user has registered, enter e-mail and password | verify password | user |
| **2. register** | enter his data with unique phone number and e-mail | send data to the database | user |
| **3. edit profile** | user has an account | change users’ data | user |
| **4. verify**  **password** | user enter password and e-mail | display an error message or login to home page | system |
| **5. Error** | user entered wrong password or e- mail | user receive notification to login again with valid data | system |
| **6. symptoms checker** | System asks questions to clarify symptoms | System provides a list of possible diagnoses | system |
| **7. Book**  **appointment** | user select a doctor | confirmed appointment | user |
| **8. Show**  **appointment** | System shows available appointment | User selects an appointment | system |
| **9. Prescription Upload** | User or Doctor uploads a prescription | System stores the prescription | user |
| **10. system**  **notifications** | prescription is Uploaded | System reminds the User of the treatment date | system |

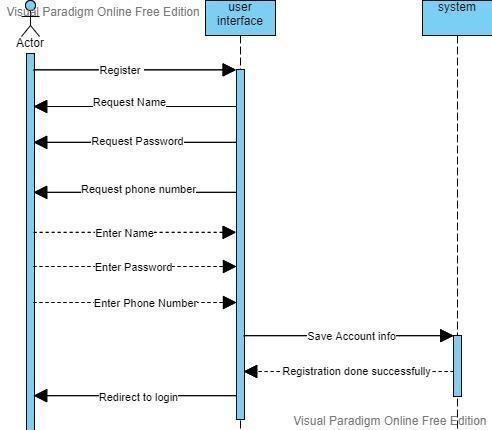
|  |  |  |  |
| --- | --- | --- | --- |
| **11. Medical**  **Image Upload** | User uploads a Medical Image | System analyzes the image | user |
| **12. analysis the medical image** | Medical image is uploaded | System shows the User the possible diagnosis | system |

* 1. **Sequence diagram**

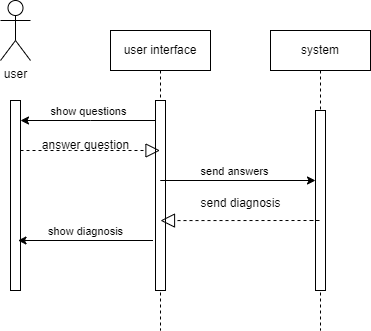
### Login

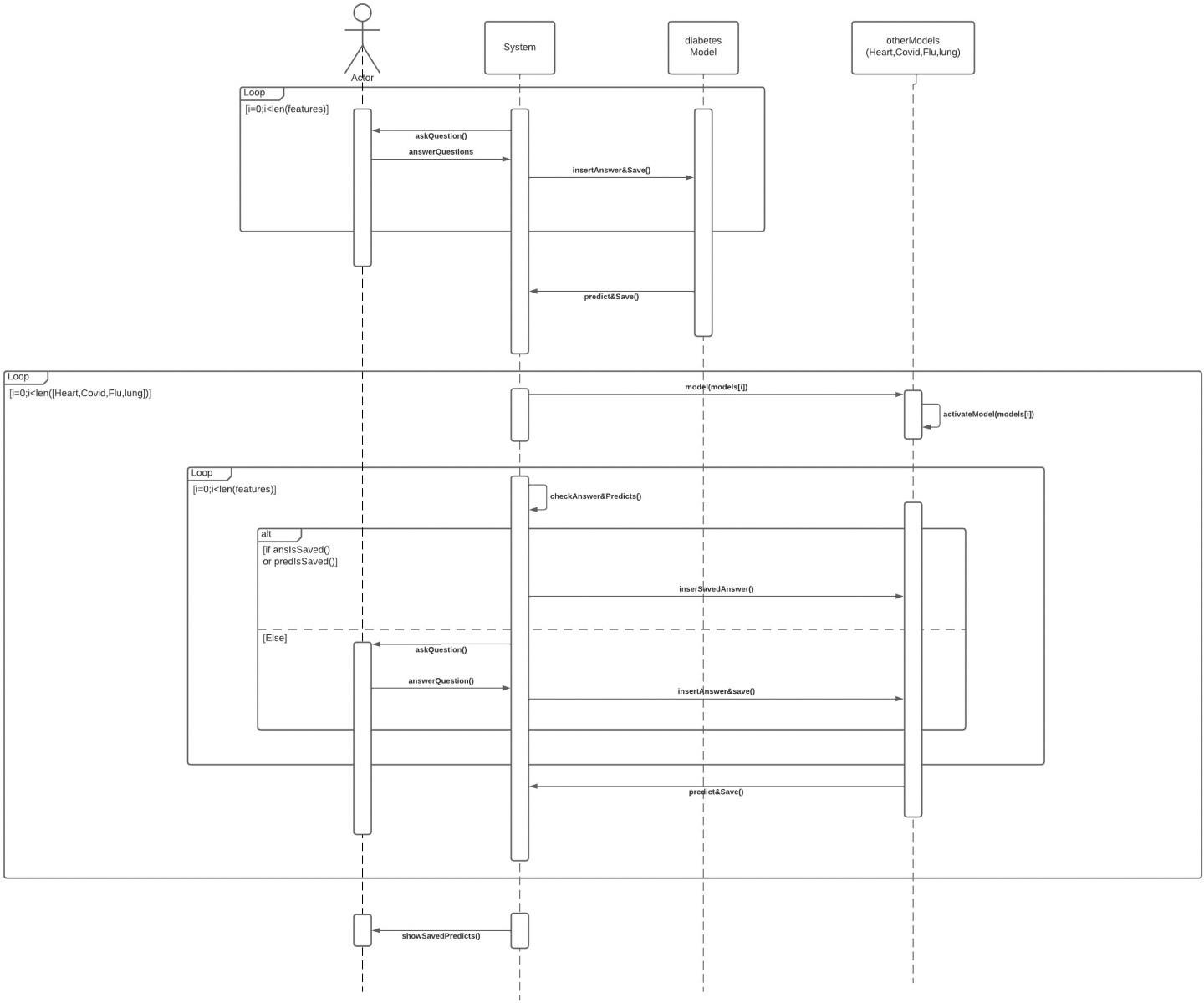


* + 1. **Register**

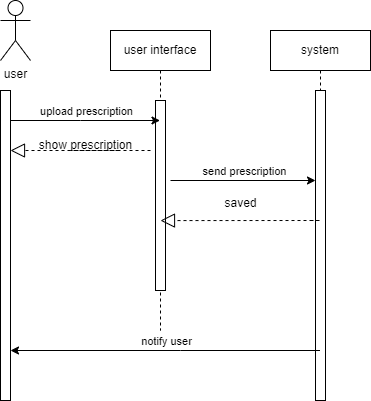


### Symptoms checker

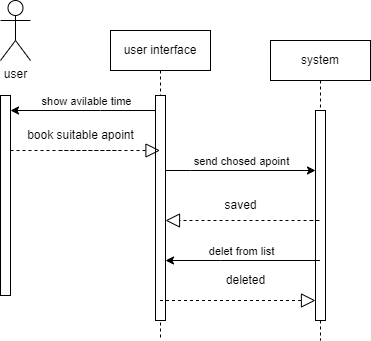




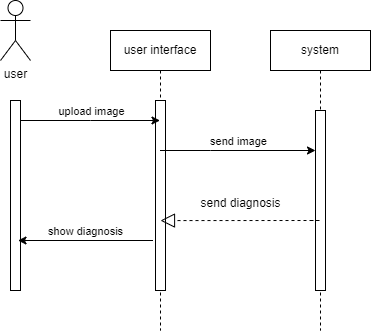
* + 1. **Prescription**



### Appointment

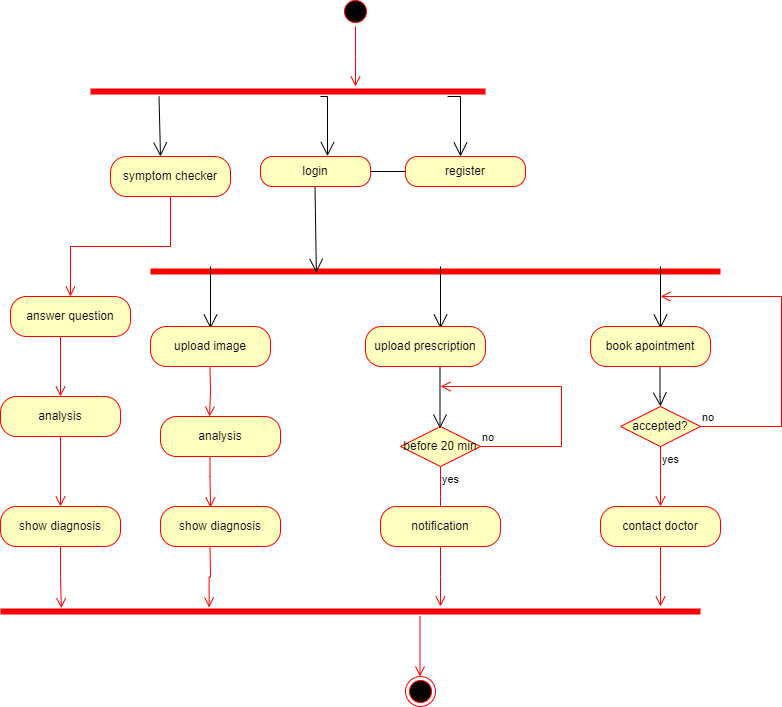


* + 1. **Medical Image**

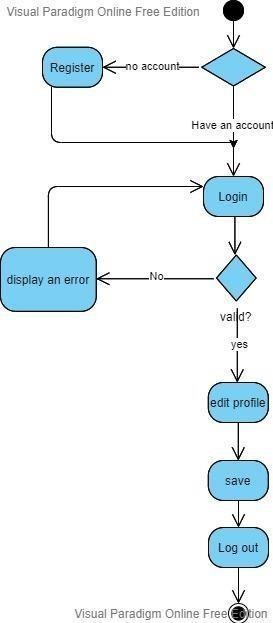


### Activity Diagrams

#### System Activity Diagram



* + 1. **Login/register Activity Diagram**



# Chapter Four: Implementation

In this chapter we’re going to discuss and go deeper in safe zone application’s implementation and present its code and the algorithms used to build it.

## Implementation

The first step in implementing such a system is to gather a team of experts in software development, healthcare, and machine/deep learning. The team should work together to define the system's requirements, design the architecture, and determine the necessary data sources and algorithms.

Next, the team can use Angular to create a user-friendly interface that allows users to access the system's features easily. The symptom checker can be a key component of the system, enabling users to enter their symptoms and receive a list of possible diagnoses. The system can use machine learning algorithms to analyze the entered symptoms and compare them to a database of known symptoms and diseases. The system can then provide the user with a list of possible diagnoses, along with information on each diagnosis, such as the symptoms, causes, and treatments.

Users can also use the system to book appointments with doctors. The system can provide a list of available doctors based on the user's location, specialty, and availability. Users can choose a doctor and schedule an appointment, and the system can send a reminder to both the user and the doctor before the appointment.

The system also allow users to upload prescriptions and remind them of the treatment date. The system can use deep learning algorithms to recognize the prescription's details, such as the medication name, dosage, and duration. The system can then store this information and send reminders to the user before the treatment date.

Finally, the system can allow users to upload medical images, such as X-rays or MRI scans, to diagnose diseases. The system use deep learning algorithms to analyze the images and provide a list of possible diagnoses. The system can also provide the user with information on each diagnosis, such as the symptoms, causes, and treatments.

### Machine and deep learning models:

#### Why we should use deep learning

**Automatic Feature Learning:** Deep learning algorithms automatically learn hierarchical representations, eliminating manual feature engineering .

**Performance on Large-Scale Data:** Deep learning excels in handling vast amounts of data, delivering exceptional performance on large-scale datasets.

**End-to-End Learning:** Deep learning enables direct learning from raw input to output, simplifying development by eliminating intermediate steps .

**Transfer Learning:** Deep learning models leverage pretraining on large datasets, enhancing training efficiency and performance on related tasks with limited labeled data .

**Continual Learning:** Deep learning models adapt and learn incrementally, accommodating changing data distributions and new information .

**Flexibility and Versatility:** Deep learning finds applications across diverse domains like computer vision, natural language processing, and autonomous driving, showcasing its broad applicability.

#### CNN

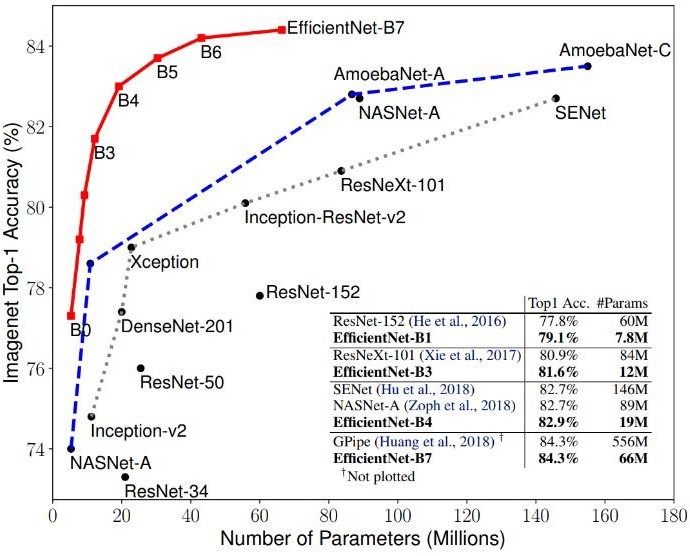
|  |  |
| --- | --- |
| **Component Description** | |
| **Convolutional Layers** | Apply filters to extract local features from input data |
| **Non-linear Activation Functions** | Introduce non-linearities for capturing complex relationships |
| **Pooling Layers** | Reduce spatial dimensions of feature maps for efficiency |
| **Fully Connected Layers** | Learn complex combinations of features in high-level representations |
| **Backpropagation** | Update network weights and biases using gradient-based optimization |

* + 1. **Optimization and Performance Enhancing**

|  |  |
| --- | --- |
| **Tool Description** | |
| **Learning Rate** | Controls the step size for weight updates during training |
| **Early Stopping** | Technique to prevent overfitting by stopping training at an optimal point |
| **Data Augmentation** | Technique to increase training data diversity for better generalization |
| **ImageDataGenerator** | Utility class for on-the-fly data augmentation in image datasets |

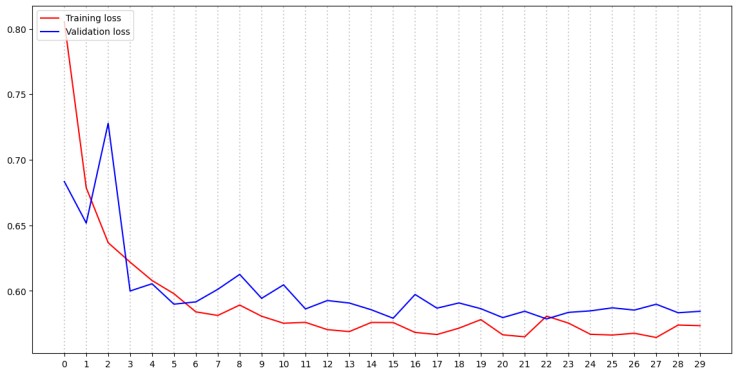
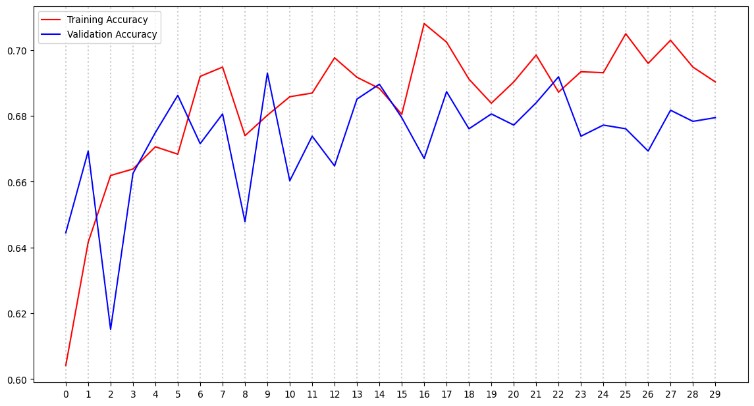
#### Different Pretrained Models

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Inception** | **VGG** | **EfficientNet** | **EfficientNet** | |
| **Year** | 2014 | 2014 | 2019 | 2019 | |
| **Architecture** | Deep CNN | Deep CNN | Compound scaling | Compound scaling | |
| **Strengths** | 1. Computational Efficiency 2. High Accuracy 3. Feature Extraction at Multiple Scales | 1. Simplicity 2. Strong   Performance | 1. Improved Efficiency 2. Scalability 3. High Accuracy | 1.  2.  3. | Improved Efficiency Scalability  High Accuracy |
| **Weaknesses** | 1. Increased Complexity 2. Risk of Overfitting | 1. High Computational   Requirements   1. Prone to Overfitting | 1. Increased Complexity 2. Training Data Requirements 3. Model Size | 1.  2.  3. | Increased Complexity Training Data Requirements Model Size |



* + 1. **Models**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Disease** | | | | | |
| Cardiomegaly | | | | | |
| **Data Source** | | **Author** | | **License** | |
| Source: The dataset was created by Rahimanshu, a data scientist based  in India.. | | Author: Rahimanshu | | License: The dataset is licensed under the Creative Commons Attribution- NonCommercial-ShareAlike 4.0  International License. | |
| **Model** | | | | | |
| VGG19 | | | | | |
| **Architecture** | | | | | |
|  | | | | | |
| **Accuracy** | **loss** | | **val\_accuracy** | | **val\_loss** |
| 0.6985 | 0.5649 | | 0.6840 | | 0.5844 |
| **Accuracy for each epoch** | | | **Loss for each epoch** | | |



|  |  |  |
| --- | --- | --- |
| **Disease** | | |
| Breast Tumor | | |
| **Data Source** | **Author** | **License** |
| The source of the Breast Ultrasound Images Dataset is the Data in Brief  journal | The author of the dataset is Wael Al-Dhabyani, Mohamed Gomaa, Hesham Khaled, and Ahmed  Fahmy | The dataset is licensed under the Creative Commons Attribution 4.0  International License. |
| **Model** | | |
| ResNet50 | | |
| **Architecture** | | |
|  | | |
| **Accuracy** | **loss val\_accuracy** | **val\_loss** |
| 0.8150 | 0.4818 0.8788 | 0.4551 |
| **Accuracy for each epoch** | **Loss for each epoch** |  |
|  | | |

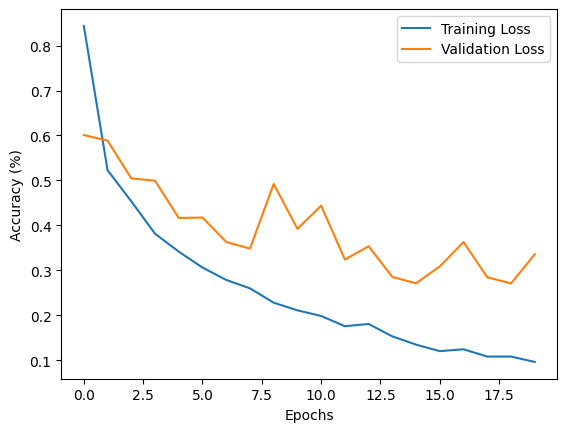
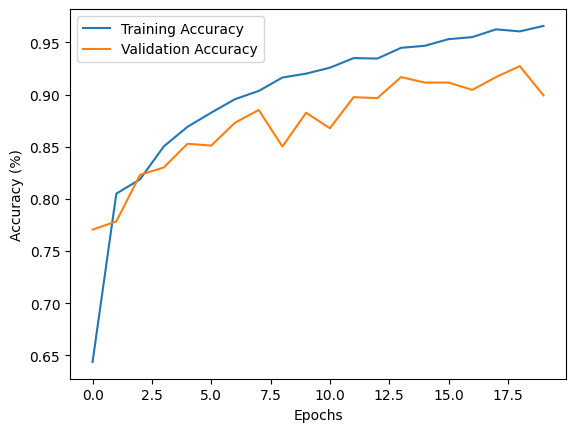
*AI Diagnoses Prediction & clinical system | Computer Science - Helwan University - Medical department*

|  |  |  |
| --- | --- | --- |
| **Disease** | | |
| Colon Disease | | |
|  | | |
| **Data Source** | **Author** | **License** |
| The dataset was curated by Francis Mon, a data scientist  based in Singapore. | Francis Mon | The dataset is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0  International License |
| **Model** | | |
| EfficientNetB2 | | |
| **Architecture** | | |
|  | | |
| **Accuracy** | **loss** | **val\_accuracy val\_loss** |
| 0.9563 | 0.1558 | 0.2396 0.9125 |
| **Accuracy for each epoch** |  | **Loss for each epoch** |
|  | | |

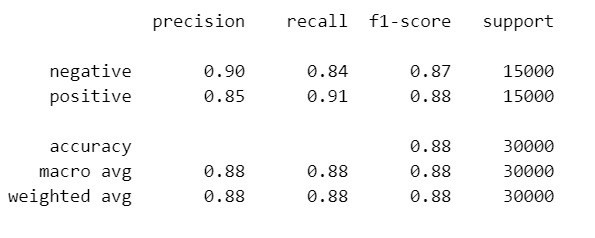
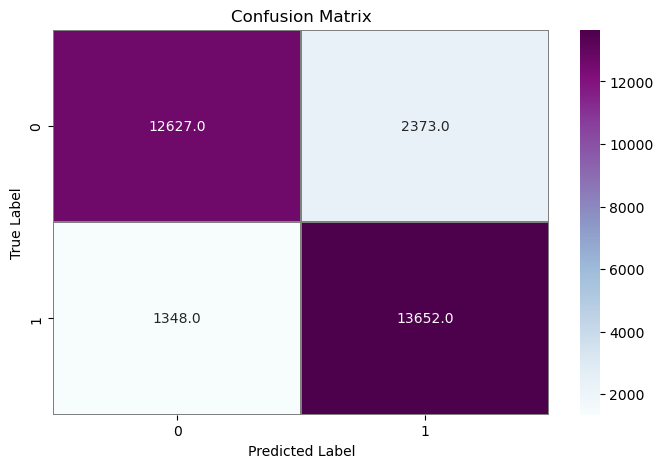
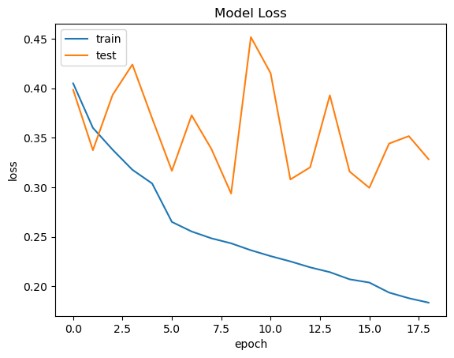
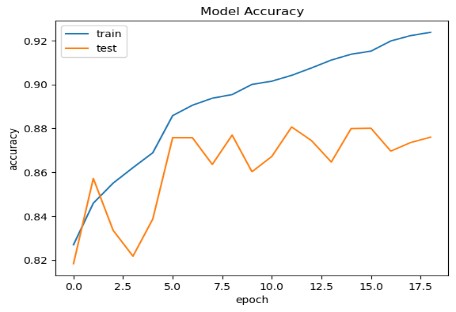
*AI Diagnoses Prediction & clinical system | Computer Science - Helwan University - Medical department*

|  |  |  |
| --- | --- | --- |
| **Disease** | | |
| Brain Tumor | | |
| **Data Source** | **Author** | **License** |
| The dataset was created by Masoud Nickparvar, a data scientist based in  Iran | Author: Masoud Nickparvar | License: The dataset is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0  International License. |
| **Model** | | |
| CNN | | |
| **Architecture** | | |
|  | | |
| **Accuracy** | **loss** | **val\_accuracy val\_loss** |
| 0.9591 | 0.1114 | 0.9238 0.2705 |
| **Accuracy for each epoch** |  | **Loss for each epoch** |

*AI Diagnoses Prediction & clinical system | Computer Science - Helwan University - Medical department*



|  |
| --- |
| **Disease** |
| Breast Histopathology |
| **Data Source Author License** |
| Source: The dataset was Paul Mooney. He is a research scientist The Breast Histopathology Images dataset is created by Rahimanshu, a at Google AI, where he works on the licensed under the CC0 Public Domain  data scientist based in development of machine learning Dedication.  India.. algorithms for medical image analysis. He has a PhD in computer science from the University of California, Berkeley. |
| **Collect method** |
| The Breast Histopathology Images dataset was collected from a variety of sources, including:   * Hospitals: The dataset includes images from breast cancer patients who were treated at hospitals. * Research laboratories: The dataset includes images from breast cancer patients who were participating in research studies. * Public databases: The dataset includes images from public databases of breast cancer images. The images in the dataset were collected using a variety of methods, including: * Whole slide imaging: Whole slide imaging is a technique for digitizing entire tissue sections. This allows researchers to view and analyze images of tissue at high magnification. * Microscopy: Microscopy is a technique for magnifying and viewing small objects. This allows researchers to view and analyze images of tissue at lower magnification. |
| **Model** |
| CNN |
| **Architecture** |
|  |



**Accuracy**

**loss**

**val\_accuracy**

**val\_loss**

0.9041

**Accuracy for each epoch**

0.2252

0.8806

**Loss for each epoch**

0.3079

Classification report

Confusion matrix

##### Disease

Lung and Colon Cancer

##### Author License

Biplob Dey. He is a data scientist at the University of Waterloo. He has a Master of Applied Science degree in Data Science from the University of Waterloo

The license for the dataset you linked to is the Creative Commons Attribution 4.0 International License. This

license allows you to do the following:

* Share the dataset with others.
* Remix the dataset, meaning you can create new datasets by combining this dataset with other datasets.
* Adapt the dataset, meaning you can modify the dataset to fit your needs.
* Build upon the dataset, meaning you can use the dataset to create new products or services.

##### Collect method

1. HIPAA compliant and validated sources: These sources include hospitals, clinics, and other medical institutions that have been certified by the Health Insurance Portability and Accountability Act (HIPAA). HIPAA is a federal law that protects the privacy of patient health information.
2. Augmentor package: The Augmentor package is a Python library that can be used to augment images. Augmentation is the process of artificially increasing the size of a dataset by creating new data points from existing data points. This can be done by applying a variety of transformations to the data points, such as cropping, flipping, and rotating.

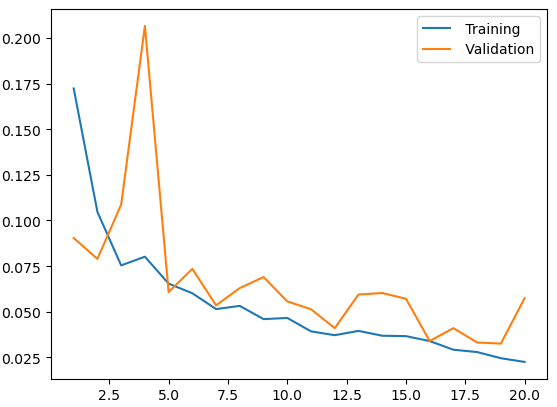
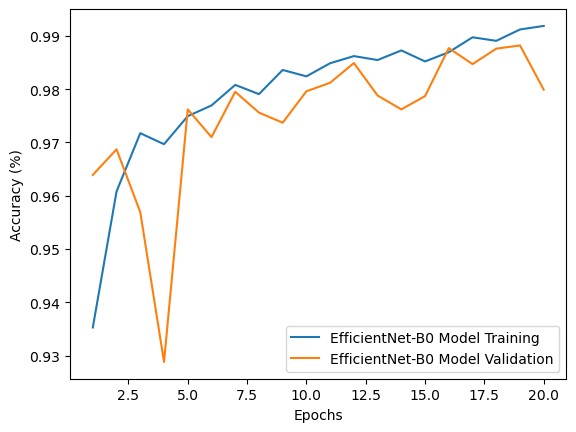
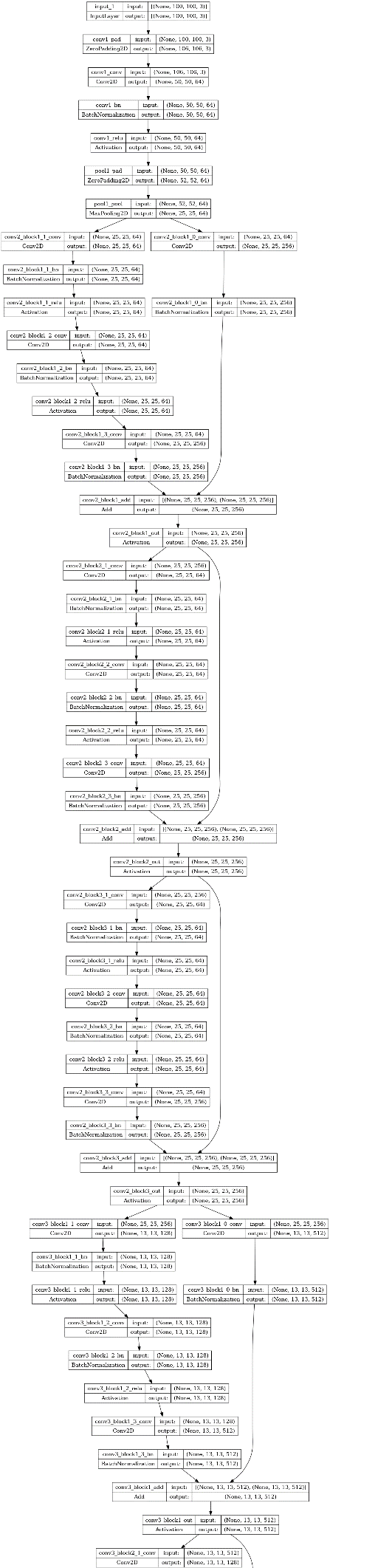
The data was collected in a variety of ways, including:

* + Surgical specimens: These are tissue samples that are removed during surgery.
  + Biopsies: These are small samples of tissue that are removed from the body using a needle.
  + Cytological samples: These are samples of cells that are collected from the body using a brush or swab.

##### Model

ResNet50

##### Architecture



**Accuracy**

**loss**

**val\_accuracy**

**val\_loss**

0.9912 0.0246

**Accuracy for each epoch**

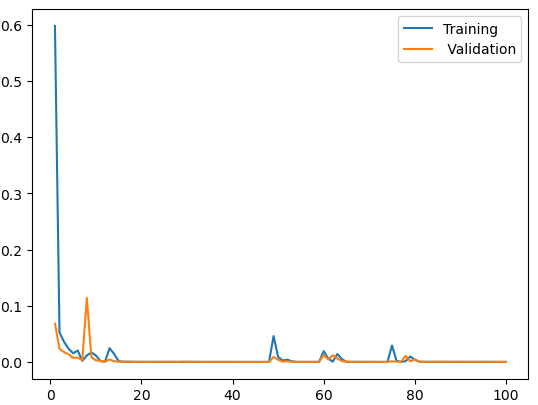
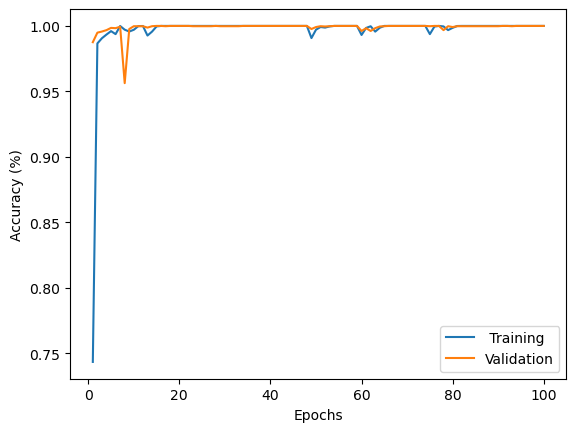
0.9882

**Loss for each epoch**

0.0326

|  |
| --- |
| **Disease** |
| Kidney |
| **Author License** |
| s Md. Nazmul Islam. He is a Research Assistant at the Department of The license for the CT Kidney Dataset: Normal-Cyst- Computer Science and Engineering, North South University, Dhaka, Tumor and Stone is © Original Authors. This means Bangladesh. He received his Master of Science degree in Computer that the authors of the dataset retain the copyright and Science from North South University in 2021. His research interests grant users the right to use the dataset for research include machine learning, deep learning, and medical image analysis purposes only. The dataset cannot be used for  commercial purposes without the express permission  of the authors |
| **Collect method** |
| The CT Kidney Dataset: Normal-Cyst-Tumor and Stone was collected from Picture archiving and communication system (PACS) from different hospitals in Dhaka, Bangladesh. PACS is a system that stores and retrieves medical images and associated patient data. The images in the dataset were collected from patients who had already been diagnosed with kidney  tumors, cysts, stones, or normal findings. |
| **Model** |
| CNN |
| **Architecture** |
|  |

*Diagnoses Prediction & clinical system | Computer Science - Helwan University - Medical department*



**Accuracy**

**loss**

**val\_accuracy**

**val\_loss**

0.9999 0.0019

**Accuracy for each epoch**

0.9994

**Loss for each epoch**

0.0027

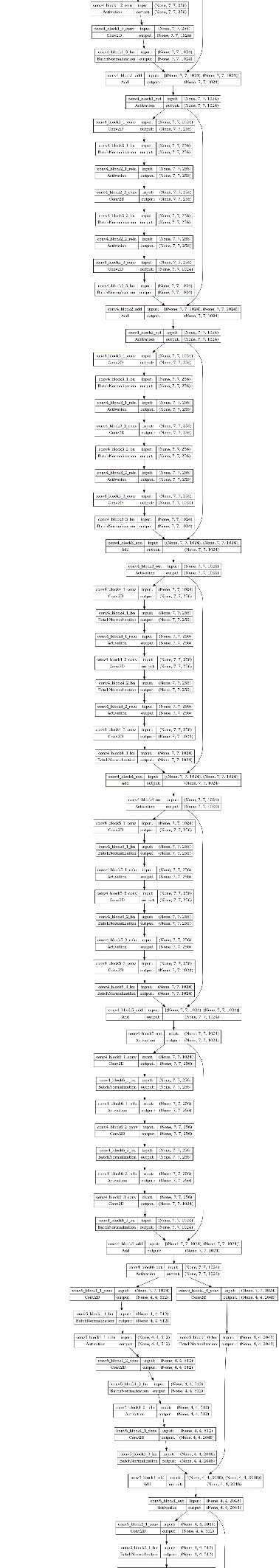
|  |  |
| --- | --- |
| **Disease** | |
| Malaria | |
| **Author** | **License** |
| Devakumar K.P. He is a Data Scientist at IBM Research AI, India. He received his Master of Science degree in Computer Science from the Indian Institute of Technology Madras in 2017. His research interests include machine learning, deep learning, and natural language processing. | The license for the CT Kidney Dataset: Normal-Cyst- Tumor and Stone is © Original Authors. This means that the authors of the dataset retain the copyright and grant users the right to use the dataset for research purposes only. The dataset cannot be used for commercial purposes without the express  permission of the authors |
| **Collect method** | |
| The dataset was created by collecting data from the World Health Organization (WHO) and the Malaria Atlas Project (MAP). The data includes information on the number of malaria cases and deaths, as well as the risk of malaria transmission, in  different countries and regions around the world. | |
| **Model** | |
| CNN | |
| **Architecture** | |

*Diagnoses Prediction & clinical system | Computer Science - Helwan University - Medical department*

|  |  |  |  |
| --- | --- | --- | --- |
|  | | | |
| **Accuracy** | **loss** | **val\_accuracy** | **val\_loss** |
| 0.9554 | 0.1343 | 0.9543 | 0.1429 |
| **Accuracy for each epoch** | | **Loss for each epoch** | |
|  | |  | |

|  |  |
| --- | --- |
| **Disease** | |
| Oral cancer | |
| **Author** | **License** |
| Ashenafi Fasil Kebede. He is a research scientist at Google AI, where he works on developing machine learning algorithms for medical image analysis | is licensed under the CC0 Public Domain Dedication. This means that the dataset is free to use for any purpose,  without any restrictions.  The CC0 license is a creative commons license that waives all copyright and related or neighboring rights to the dataset. This means that anyone can use the dataset for any purpose, without asking permission from the  author or publisher. |
| **Collect method** | |
| Cancer Detection using CNNs" dataset was collected from two sources:   * The University of Pennsylvania Oral Cancer Tissue Bank * The National Cancer Institute (NCI) Head and Neck Histology Data Repository   The University of Pennsylvania Oral Cancer Tissue Bank collected tissue samples from patients with oral cancer and normal tissue. The tissue samples were stained with hematoxylin and eosin (H&E) and then imaged using a microscope. The images were then digitized and stored in a database.  The NCI Head and Neck Histology Data Repository collected tissue samples from patients with head and neck cancer, including oral cancer. The tissue samples were stained with H&E and then imaged using a microscope. The images were then digitized and stored in a database. | |
| **Model** | |
| ResNet50 | |
| **Architecture** | |

*Diagnoses Prediction & clinical system | Computer Science - Helwan University - Medical department*



P a g e 60 | 95

|  |  |  |  |
| --- | --- | --- | --- |
| **Accuracy** | **loss** | **val\_accuracy** | **val\_loss** |
| 0.9723 | 0.0725 | 0.9222 | 0.2183 |
| **Accuracy for each epoch** | | **Loss for each epoch** | |
|  | |  | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Disease** | | | | |
| Pneumonia | | | | |
| **Author** | | | | **License** |
| Paul Timothy Mooney. He is a Research Scientist at Google AI, where he works on developing machine learning algorithms for medical imaging. He has a PhD in Computer Science from the University of California, Berkeley. The dataset was created by Mooney and his colleagues at the Guangzhou Women and Children's Medical Center in China. | | | | The Chest X-Ray Images (Pneumonia) dataset on Kaggle is licensed under the Creative Commons Attribution 4.0  International License. |
| **Collect method** | | | | |
| The Chest X-Ray Images (Pneumonia) dataset on Kaggle was collected from the Guangzhou Women and Children's Medical Center in China. The images were collected as part of routine clinical care, and were graded by two expert physicians before being included in the dataset. Here are the steps involved in collecting the data:   1. Patients were admitted to the Guangzhou Women and Children's Medical Center with suspected pneumonia. 2. Chest X-rays were taken of all patients. 3. The images were graded by two expert physicians. 4. Images that were clearly labeled as pneumonia or normal were included in the dataset. 5. The images were anonymized before being released to the public. | | | | |
| **Model** | | | | |
| CNN | | | | |
| **Accuracy** | **loss** | **val\_accuracy** | **val\_loss** | |
| 0.9509 | 0.1405 | 0.6250 | 1.0401 | |
| **Accuracy and Loss** | | | | |
|  | | | | |

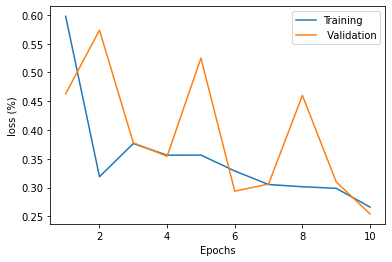
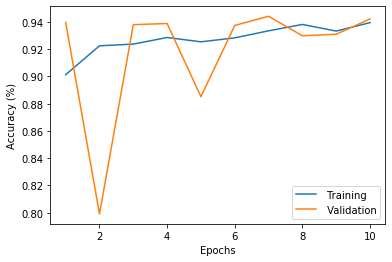
*AI Diagnoses Prediction & clinical system | Computer Science - Helwan University - Medical department*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Disease** | | | | |
| Face Images of Acute Stroke and Non-Acute Stroke | | | | |
| **Author** | | | **License** | |
| Danish J. Al-Khateeb. He is a research scientist at the University of California, San Diego. He is interested in using machine learning to improve the diagnosis and treatment of stroke.  The dataset was created by collecting face images from a variety of sources, including public image databases, medical research studies,  and social media. The images were then labeled by experts as either "acute stroke," "non-acute stroke," or "control." | | | he Chest X-Ray Images (Pneumonia) dataset on Kaggle is licensed under the Creative Commons Attribution 4.0 International License. | |
| **Collect method** | | | | |
| The dataset "Face Images of Acute Stroke and Non-Acute Stroke" is licensed under the Creative Commons Attribution 4.0  International License. This means that you can freely use, share, and adapt the dataset for any purpose, as long as you give credit to the original authors.  Here are some of the terms of the license:   * You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use. * You may not adapt the license for commercial purposes. * No additional restrictions: You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits. | | | | |
| **Model** | | | | |
| ResNet50 | | | | |
| **Accuracy** | **loss** | **val\_accuracy** | | **val\_loss** |
| 0.9890 | 0.0303 | 0.9768 | | 0.0637 |
| **Accuracy for each epoch** | | **Loss for each epoch** | | |
|  | |  | | |

*Diagnoses Prediction & clinical system | Computer Science - Helwan University - Medical department*

|  |  |  |  |
| --- | --- | --- | --- |
| **Disease** | | | |
| **Melanoma** | | | |
| **Author** |  | **License** |  |
| Dr. Andrei Scarlat. He is a postdoctoral researcher at the University The melanoma dataset on Kaggle is licensed under the  of Basel, Switzerland. His research interests include machine Creative Commons Attribution-NonCommercial-ShareAlike learning, image analysis, and medical imaging. He has published 4.0 International License.  several papers in top academic journals and conferences on these  topics. | | | |
| **Collect method** | | | |
| e melanoma dataset on Kaggle was collected from a variety of sources, including:   * The International Skin Imaging Collaboration (ISIC) * The University of Basel's Skin Imaging Database (SID) * The University of Pennsylvania's Skin Lesion Image Archive (SLIA) * The University of Texas MD Anderson Cancer Center's Melanoma Image Archive (MIA) | | | |
| **Model** | | | |
| EfficientNetB7 | | | |
| **Architecture** | | | |
|  | | | |
| **Accuracy** | **loss** | **val\_accuracy** | **val\_loss** |
| 0.9332 | 0.2987 | 0.9309 | 0.3099 |
| **Accuracy for each epoch** |  | **Loss for each epoch** |  |

*Diagnoses Prediction & clinical system | Computer Science - Helwan University - Medical department*



|  |  |
| --- | --- |
| **Disease** | |
| Alzheimer | |
| **Author** | **License** |
| Sarvesh Dubey. He is a Data Scientist at IBM. He has a Master of Science degree in Data Science from the Indian Institute of Technology, Bombay. He is interested in machine learning, natural language processing, | The dataset is licensed under the Creative Commons Attribution-Share Alike 4.0 International License. |
| **Collect method** | |
| The data in the Alzheimer's Dataset (4 class of Images) on Kaggle was collected from a variety of sources, including:   * The Alzheimer's Disease Neuroimaging Initiative (ADNI) * The Open Access Series of Imaging Studies (OASIS) * The Parkinson's Disease Imaging Consortium (PDiC) * The Alzheimer's Disease Research Centers (ADRCs)   The data was collected using a variety of methods, including:   * MRI scans * PET scans * Diffusion tensor imaging (DTI) * Functional MRI (fMRI) * Magnetic resonance spectroscopy (MRS) | |
| **Model** | |
| CNN | |
| **Architecture** | |

*Diagnoses Prediction & clinical system | Computer Science - Helwan University - Medical department*

|  |  |  |  |
| --- | --- | --- | --- |
|  | | | |
| **Accuracy** | **loss** | **val\_accuracy** | **val\_loss** |
| 0.9795 | 0.0602 | 0.9531 | 0.1548 |
| **Accuracy for each epoch** | | **Loss for each epoch** | |
|  | |  | |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Disease** | | | | |
| **Autism** | | | | |
| **Author** | | | **License** | |
| The author of the autism image dataset on Kaggle is Cihan Senol. He is a software engineer at Google AI, and he is also a member of the Autism Research  Group at Bogazici University in Turkey. | | | he Autism\_Image\_Data dataset on Kaggle is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. | |
| Method | | | | |
| The dataset was created by collecting images of children with autism and children without autism from a variety of sources, including:   * The Autism Speaks website * The National Autistic Society website * The Simons Foundation Autism Research Initiative website | | | | |
|  | **accuracy** | **loss** | **val\_accuracy** | **val\_loss** |
| **EfficientNet B7** | 0.9641 | 0.1060 | 0.8400 | 0.5154 |
| **VGG 19** | 0.8592 | 0.3124 | 0.8000 | 0.4074 |
| **Inception V3** | 0.7236 | 0.5491 | 0.7200 | 0.5158 |

*AI Diagnoses Prediction & clinical system | Computer Science - Helwan University - Medical department*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **EfficientNet B0** | 0.7946 | 0.4392 | 0.8000 | 0.4131 |
|  | Accuracy for each Epoch | | Loss for each Epoch | |
| **EfficientNetB7** |  | |  | |
| **VGG16** |  | |  | |
| **Inception V3** |  | |  | |

|  |  |  |
| --- | --- | --- |
| **EfficientNet B0** |  |  |
| **Result** | | |
|  | | |

**Machine learnning algorithms** :

* + **K-nearest neighbors (KNN) algorithm :**
    - It classifies a new data point based on the majority class of its K nearest neighbors in the training data.
    - The neighbors are determined by calculating distances (e.g., Euclidean distance) between the data points.
  + **Decision tree algorithm:**
    - It is a supervised learning algorithm that recursively splits the dataset based on feature conditions to create a tree-like structure for classification or regression.
    - At each node, the algorithm selects the best feature to split the data, optimizing a criterion such as Gini impurity or information gain.
    - The process continues until a stopping condition is met, such as reaching a maximum depth or a minimum number of samples, resulting in a tree that can be used for prediction.
  + **Support Vector Machine (SVM) algorithm:**
    - It is a supervised learning algorithm used for classification and regression tasks.
    - SVM finds the optimal hyperplane that maximally separates data points of different classes in a high- dimensional feature space.
    - It achieves this by defining support vectors, which are the data points closest to the decision boundary, and using them to classify new data points.
  + **Random Forest (RF) algorithm:**
    - It is an ensemble learning method that combines multiple decision trees to make predictions for classification and regression tasks.
    - RF randomly selects subsets of features and data samples to train each decision tree, reducing overfitting and increasing robustness.
    - The final prediction is determined by aggregating the predictions of all the individual trees, either through voting (for classification) or averaging (for regression).
  + **Voting Classifier algorithm**:
    - It is an ensemble learning method that combines multiple individual classifiers or models to make predictions.
    - Each classifier/model provides its prediction, and the voting classifier algorithm aggregates the predictions using voting (e.g., majority voting for classification).
    - It allows for combining different types of classifiers (e.g., decision trees, support vector machines) to improve overall prediction performance.
  + **Logistic Regression algorithm:**
    - It is a supervised learning algorithm used for binary classification tasks.
    - Logistic Regression models the relationship between the input variables and the binary output using the logistic function, which maps the input to a probability value between 0 and 1.
    - It estimates the model parameters through the maximum likelihood estimation and uses a decision boundary to classify new data based on the predicted probabilities.
  + **MLP (Multilayer Perceptron) algorithm:**
    - It is a type of artificial neural network widely used for supervised learning tasks such as classification and regression.
    - MLP consists of multiple layers of interconnected nodes (neurons) and uses backpropagation to adjust the weights to minimize the error between predicted and actual outputs.
    - It employs activation functions, typically sigmoid or ReLU, to introduce non-linearity and can handle complex relationships between input features and output predictions.

|  |  |  |
| --- | --- | --- |
| **Metric Definition Calculation** | | |
| **Recall** | Ability to identify positive instances | TP / (TP + FN) |
| **Precision** | Ability to identify true positives among predicted  positives | TP / (TP + FP) |
| **F1 Score** | Harmonic mean of precision and recall | 2 \* (Precision \* Recall) / (Precision + Recall) |

**Grid search**

1. Grid search is an optimization technique used to find the best hyperparameters for a model.
   * It systematically explores different combinations of hyperparameter values.
   * It evaluates the model's performance using a validation set or cross-validation.
2. Process of grid search:
   * Define the hyperparameter grid with feasible values or ranges.
   * Generate combinations of hyperparameters from the grid.
   * Train and evaluate the model for each combination.
   * Record the performance metric for each combination.
   * Select the combination with the best performance.
3. Grid search helps identify optimal hyperparameters for the model.
   * It improves the model's performance by finding the best hyperparameter values.
   * It is widely used in machine learning for hyperparameter tuning.
4. Considerations for large hyperparameter spaces:
   * Grid search can be computationally intensive.
   * Alternatives like random search or Bayesian optimization can be more efficient.

By systematically exploring hyperparameter combinations, grid search allows for the identification of the best hyperparameters, leading to improved model performance.

### CHALLENGINGS

#### Duplicated Dataset:

* + - * + Dataset: <https://www.kaggle.com/datasets/neelima98/disease-prediction-using-machine-learning> Author: Neelima98

License: CC BY-SA 4.0

* + - * + The dataset was created by Neelima98 and is licensed under the Creative Commons AttributionShareAlike

4.0 International License. This means that you can share, adapt, and build upon the dataset, as long as you give credit to the original author and license your derivative works under the same terms.

* + - * + The dataset contains data on patients with various diseases. The data includes the following features: - ID: The patient's ID number

Age: The patient's age

Gender: The patient's gender

Blood Pressure: The patient's blood pressure

Cholesterol: The patient's cholesterol level

Glucose: The patient's glucose level

Anemia: Whether or not the patient has anemia

Heart Disease: Whether or not the patient has heart disease

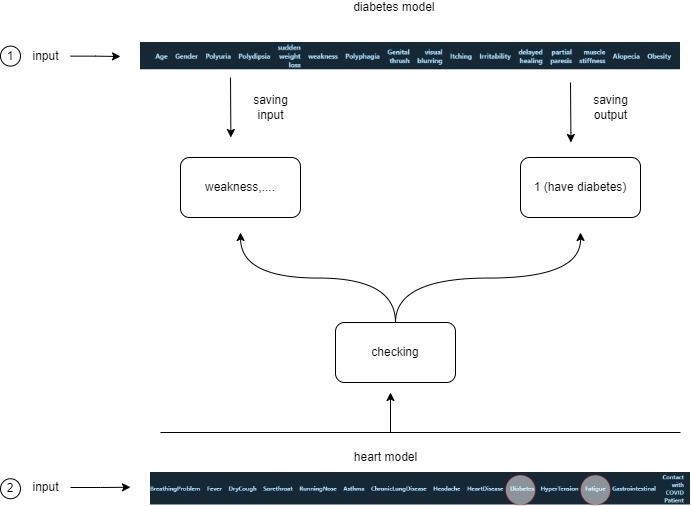
Stroke: Whether or not the patient has had a stroke

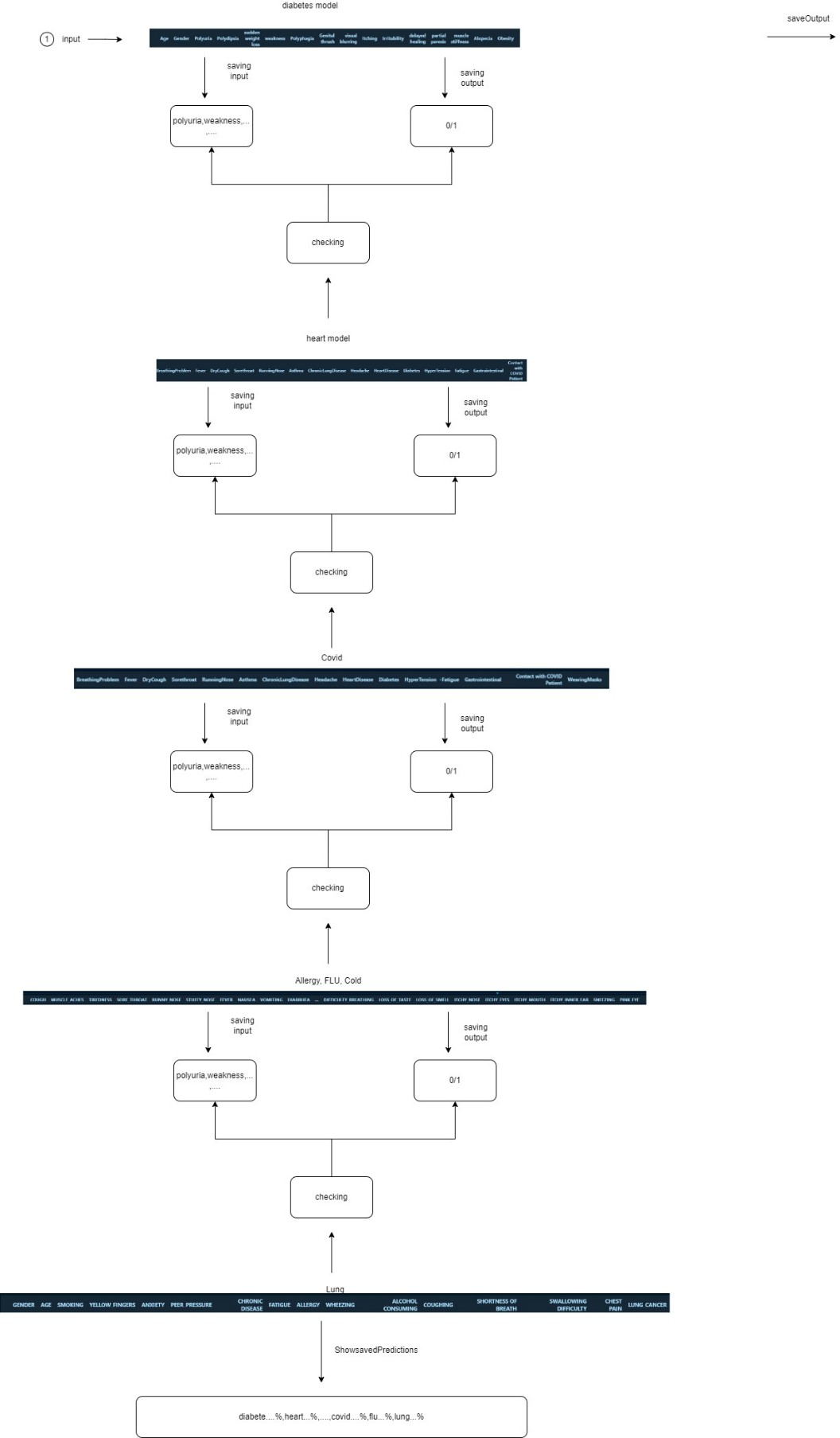
Diabetes: Whether or not the patient has diabetes

* + - * + This dataset has 4921 record but containing 304 records after removing duplicates. However, you have 41 different diseases represented in this dataset, which means each disease has only around 6 records. You're concerned that this small number of records per disease may not be sufficient for an algorithm to effectively learn from the data.



* + - * + Having a small number of records per disease can indeed pose challenges for training machine learning algorithms. Insufficient data can lead to overfitting, where the algorithm learns the specific examples in the dataset too well but fails to generalize to new
* **Solution**
* As a solution we gather more than 5 data, these are related with each other data so they may have the same inputs and outputs.

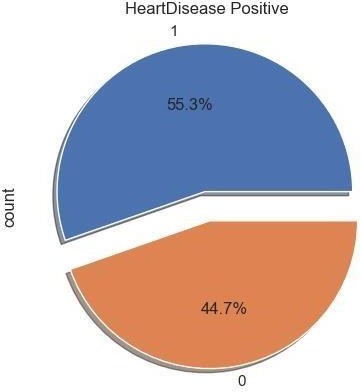




* **heart dataset:**

Dataset on Kaggle is a collection of 11 clinical features that can be used to predict the likelihood of a patient developing heart failure. The dataset was created by combining five different datasets that were previously available independently

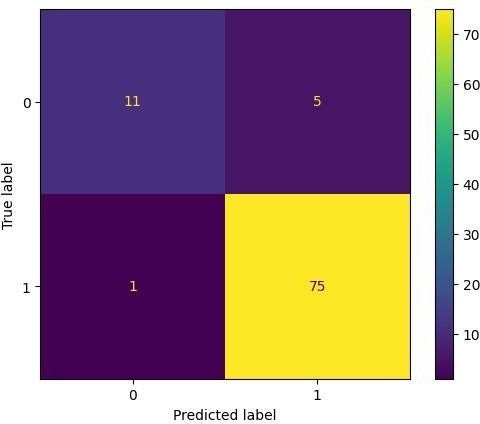
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algo** | **precision** | **recall** | **F1score** | **accuracy** |
| **KNN** | N: 0.91 /P:0.81 | N: 0.77 / P:0.92 | N:0.84 /P:0.86 | 0.85 |
| **RF** | N: 0.95 /P:0.85 | N: 0.82 / P:0.96 | N:0.88 /P:0.90 | 0.89 |
| **DT** | N: 0.92 /P:0.79 | N: 0.75 / P:0.93 | N:0.83 /P:0.86 | 0.84 |
| **MLP** | N: 0.84 /P:0.81 | N: 0.79 / P:0.84 | N:0.81 /P:0.82 | 0.82 |
| **Voting** | N: 0.90 /P:0.81 | N: 0.78 / P:0.91 | N:0.84 /P:0.86 | 0.85 |
| **logistic** | N: 0.85 /P:0.81 | N: 0.79 / P:0.86 | N:0.82 /P:0.84 | 0.83 |
| **SVM** | N: 0.96 /P:0.86 | N: 0.84 / P:0.96 | N:0.90 /P:0.91 | 0.90 |



* **Covid dataset:**

The COVID-19 dataset on Kaggle was created by Meir Nizri and contains anonymized patient data from Mexico and other origins.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algo** | **precision** | **recall** | **F1score** | **Accuracy** |
| **KNN** | N :1.00 /P: 0.88 | N :0.38 /P: 1.00 | N :0.55 /P: 0.94 | 0.89 |
| **RF** | N :0.86 /P: 0.88 | N :0.38 /P: 0.99 | N :0.52 /.93 | 0.88 |
| **DT** | N :0.92 /P: 0.97 | N: 0.75 / P: 0.80 | N: 0.83 / P: 0.79 | 0.891 |
| **MLP** | N: 0.90 /P:0.91 | N: 0.56 / P:0.99 | N:0.69 /P:0.95 | 0.88 |
| **Voting** | N: 0.90 /P:0.91 | N: 0.56 / P:0.99 | N:0.69 /P:0.95 | 0.88 |
| **logistic** | N: 0.90 /P:0.91 | N: 0.56 / P:0.99 | N:0.69 /P:0.95 | 0.92 |
| **SVM** | N: 0.90 /P:0.91 | N: 0.56 / P:0.99 | N:0.69 /P:0.95 | 0.86 |



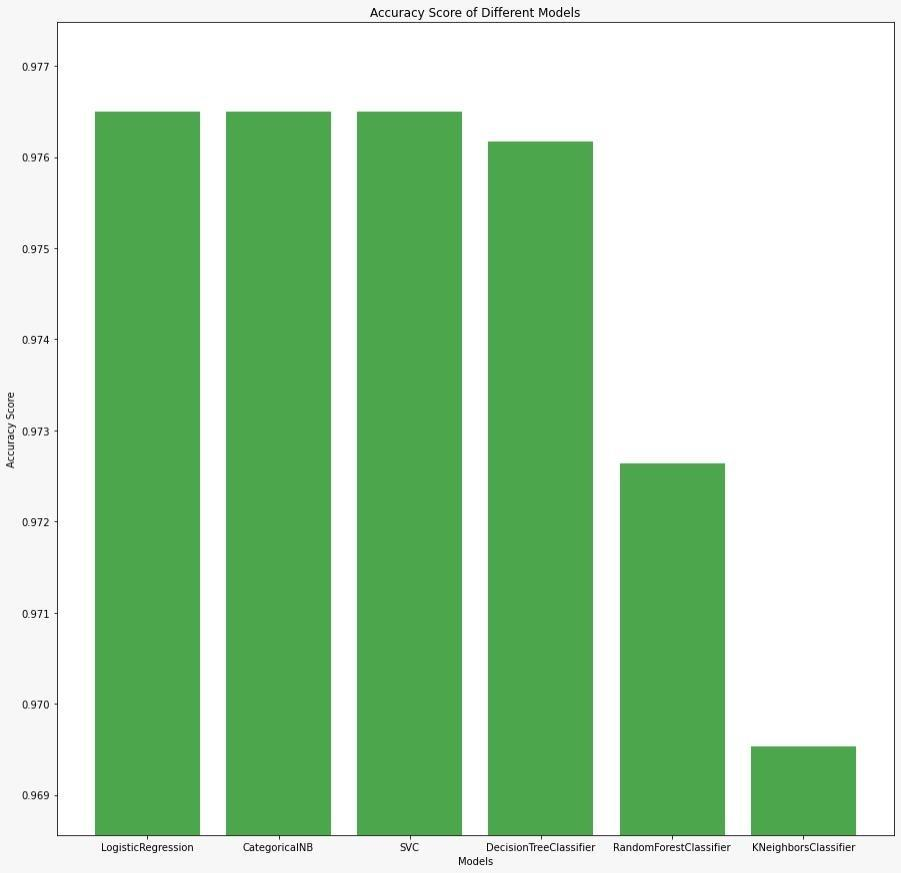
* **Flu, cold ,allergy dataset:**

The data was collected from people who visited the Mayo Clinic website. The data is not a representative sample of the population.

The data is not intended to be used for medical diagnosis.

If you are experiencing any of the symptoms in this dataset, it is important to see a doctor consist of : 11853 record .

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algo** | **Precision** | **recall** | **F1-score** | **accuracy** |
| **KNN** | 0.98 | 0.98 | 0.98 | 0.97 |
| **RF** | 0.94 | 0.99 | 0.97 | 0.96 |
| **DT** | 1.00 | 0.97 | 0.98 | 0.98 |
| **MLP** | 0.90 | 0.56 | 0.69 | 0.88 |
| **Voting** | 0.90 | 0.78 | 0.84 | 0.86 |
| **logistic** | 1.00 | 0.97 | 0.98 | 0.98 |
| **SVM** | 1.00 | 0.97 | 0.98 | 0.98 |



* **Lung Cancer Dataset:**

The Lung Cancer Dataset was created by Mysara Mahdbha in 2021. It contains information on 1000 patients with lung cancer, including their age, gender, smoking history, and other medical conditions. The data was

collected from a variety of sources, including online surveys and medical records. The dataset is freely available for download on Kaggle.com.

The dataset can be used to train machine learning models to predict whether a patient has lung cancer. It can also be used to study the risk factors for lung cancer and to develop new treatments for the disease consist of

: 274 record

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algo** | **precision** | **recall** | **F1-score** | **accuracy** |
| **KNN** | N :0.60 /P: 0.96 | N :0.75 /P: 0.92 | N :0.67 /P: 0.94 | 0.97 |
| **RF** | N :0.86 /P: 0.88 | N :0.38 /P: 0.99 | N :0.52 /.93 | 0.96 |
| **DT** | N :0.92 /P: 0.97 | N: 0.75 / P: 0.80 | N: 0.83 / P: 0.79 | 0.90 |
| **MLP** | N: 0.90 /P:0.91 | N: 0.56 / P:0.99 | N:0.69 /P:0.95 | 0.88 |
| **Voting** | N: 0.90 /P:0.91 | N: 0.56 / P:0.99 | N:0.69 /P:0.95 | 0.86 |
| **logistic** | N: 0.90 /P:0.91 | N: 0.56 / P:0.99 | N:0.69 /P:0.95 | 0.93 |
| **SVM** | N: 0.90 /P:0.91 | N: 0.56 / P:0.99 | N:0.69 /P:0.95 | 0.98 |

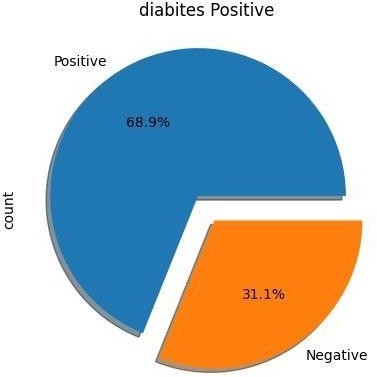
* **Diabetes:**

Data Set Information:

This has been collected using direct questionnaires from the patients of Sylhet Diabetes Hospital in Sylhet, Bangladesh and approved by a doctor.

Data Set Information : This has been collected using direct questionnaires from the patients of Sylhet Diabetes Hospital in Sylhet, Bangladesh and approved by a doctor.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Algo** | **precision** | **recall** | **F1score** | **accuracy** |
| **KNN** | N: 0.52 /P:0.88 | N: 0.81 / P:0.66 | N:0.63 /P:0.75 | 0.71 |
| **RF** | N: 0.84 /P:1.00 | N: 1.00 / P:0.91 | N:0.91 /P:0.96 | 0.94 |
| **DT** | N: 0.68 /P:0.97 | N: 0.94 / P:0.80 | N:0.79 /P:0.88 | 0.84 |
| **MLP** | N: 0.71 /P:0.88 | N: 0.75 / P:0.86 | N:0.73 /P:0.87 | 0.82 |
| **Voting** | N: 0.87 /P:0.92 | N: 0.81 / P:0.94 | N:0.84 /P:0.93 | 0.90 |
| **logistic** | N: 0.67 /P:0.93 | N: 0.88 / P:0.80 | N:0.79 /P:0.88 | 0.82 |
| **SVM** | N: 0.68 /P:0.97 | N: 0.94 / P:0.80 | N:0.79 /P:0.88 | 0.84 |



##### Imbalance datasets:

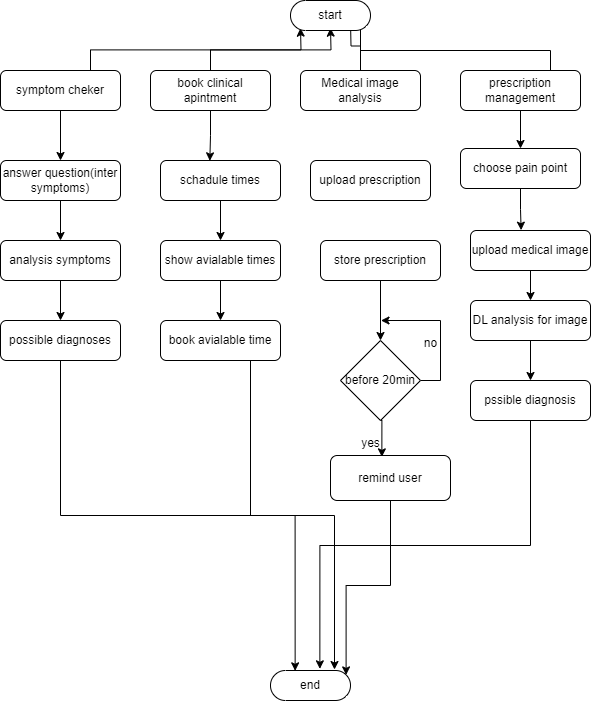
Classes that make up a large proportion of the data set are called majority classes. Those that make up a smaller proportion are minority classes.

* **Solution**

|  |  |  |
| --- | --- | --- |
| **Oversampling**   | **Technique Description** | |
| SMOTE (Synthetic Minority Technique) | Randomly duplicates instances from the minority class to increase its  representation in the dataset. Can lead to overfitting if the same instances are repeatedly used during training. |
| ADASYN (Adaptive Synthetic) | An extension of SMOTE that introduces a density-based distribution to generate synthetic examples. |
| **Undersampling** | Undersampling | Randomly removes instances from the majority class to balance the class  distribution. |
| Cluster Centroids | Identifies clusters within the majority class and replaces each cluster with its centroid. |
| Tomek Links | Pairs of very close instances but of opposite classes. Removing the instances of the majority class of each pair increases the space between the two  classes.  Tomek's link exists if the two samples are the nearest neighbors of each other. |

### Flow Chart

In this Figure we show the communication between client side and server side.



### Pseudocode

Start

User selects option:

* + Symptom checker
  + Book appointment
  + Prescription management
  + Medical image analysis

If user selected Symptom checker:

User enters symptoms

diagnoses = machine\_learning\_algorithm(symptoms) display(diagnoses)

If user selected Book appointment: doctors = get\_available\_doctors() display(doctors) user\_selects\_doctor = user\_selects(doctors) appointment\_times = get\_available\_appointment\_times(user\_selects\_doctor

) display(appointment\_times)

user\_selects\_appointment\_time = user\_selects(appointment\_times) send\_appointment\_confirmation(user\_selects\_doctor, user\_selects\_appointment\_time)

If user selected Prescription management: prescription = user\_upload\_prescription() prescription\_details = deep\_learning\_algorithm(prescription) store\_prescription\_details(prescription\_details) send\_treatment\_reminder(prescription\_details.treatment\_date) If user selected Medical image analysis:

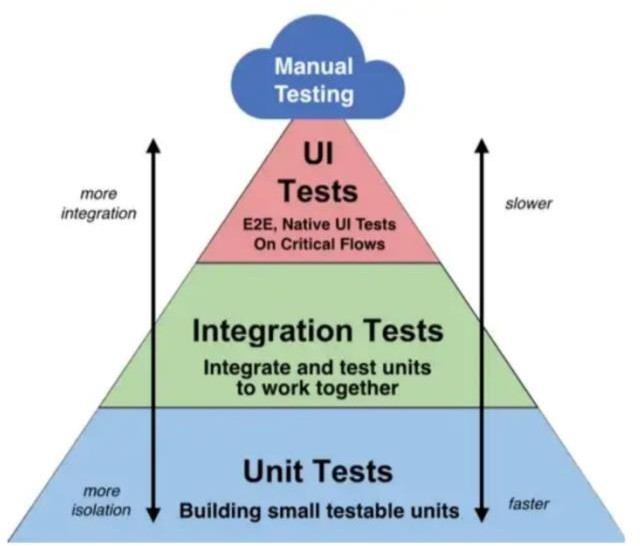
medical\_image = user\_upload\_medical\_image() diagnoses = deep\_learning\_algorithm(medical\_image) display(diagnoses) End

*department*

# Chapter Five: Testing

In this chapter we’re going to discuss and go deeper in Safe zone application’s testing and present the types of testing to be used and test cases we examined our application through.

## Testing



### Functional Testing:

#### Unit testing:

1. Test the symptom checker function to ensure that it correctly identifies known symptoms and provides accurate diagnoses:
   * Input: "headache, fever, cough"
   * Expected output: ["Migraine", "Flu", "Common cold"]
2. Test the appointment scheduling function to ensure that it correctly displays available doctors and appointment times:
   * Input: user selects "Book appointment"
   * Expected output: list of available doctors and appointment times
3. Test the prescription upload function to ensure that it correctly recognizes prescription details and stores them in the system:
   * Input: user uploads a prescription for "Ibuprofen 200mg, 1 tablet, 3 times a day for 5 days"
   * Expected output: prescription details are correctly recognized and stored in the system
4. Test the treatment reminder function to ensure that it sends reminders to users before their treatment date:
   * Input: user uploads a prescription with a treatment date of 2023-06-15
   * Expected output: reminder is sent to the user on 2023-06-14
5. Test the medical image analysis function to ensure that it correctly identifies known diseases based on uploaded images:
   * Input: user uploads an X-ray image of a bone Fracture
   * Expected output: ["Fracture "]

### Integration testing:

* + 1. Test the integration between the symptom checker and medical image analysis functions to ensure that the system provides accurate diagnoses based on both symptoms and medical images:
       - Input: user enters symptoms of "fever, headache" and uploads an X-ray image of a fractured bone
       - Expected output: ["Fracture", "Flu"]
    2. Test the integration between the appointment scheduling function and prescription management function to ensure that the system correctly schedules appointments based on user prescriptions:

Input: user uploads a prescription with a treatment date of 2023-06-15 and schedules an appointment for 2023- 06-20

Expected output: the system sends a reminder to the user on 2023-06-14 and confirms the appointment with the doctor on 2023-06-20

* + 1. Test the integration between the symptom checker and medical image analysis functions to ensure that the system provides accurate diagnoses based on both symptoms and medical images:
       - Input: user enters symptoms of "chest pain, shortness of breath" and uploads an X-ray image of the chest
       - Expected output: ["Pneumonia", "Heart attack"]
    2. Test the integration between the prescription management and treatment reminder functions to ensure that the system correctly sends reminders to users before their treatment date:
       - Input: user uploads a prescription with a treatment date of 2023-06-15
       - Expected output: the system sends a reminder to the user on 2023-06-14
    3. Test the integration between the symptom checker and appointment scheduling functions to ensure that the system correctly displays available doctors based on user symptoms:
       - Input: user enters symptoms of "stomach pain, nausea" and selects "Book appointment"
       - Expected output: the system displays a list of available doctors who specialize in gastroenterology

### Additional testing:

In addition to unit and integration testing, here are some additional testing methods that could be used to ensure the quality and reliability of a medical system that incorporates a symptom checker, clinical management, prescription management, and medical image analysis:

* + 1. Usability testing: This involves testing the system with actual users to evaluate how easy it is to use and navigate. Usability testing can help identify areas of the system that may be confusing or difficult to use, allowing developers to make improvements to the user interface and user experience.
    2. Security testing: This involves testing the system for vulnerabilities and potential security risks. Security testing can help ensure that the system is protected against unauthorized access, data breaches, and other security threats.
    3. Performance testing: This involves testing the system's performance under different conditions, such as high traffic or heavy usage. Performance testing can help identify potential bottlenecks or performance issues that may affect the system's responsiveness and reliability.
    4. Regression testing: This involves testing the system after changes or updates have been made to ensure that existing functionality has not been affected. Regression testing can help ensure that new updates or features do not introduce new bugs or issues to the system.
    5. Acceptance testing: This involves testing the system with stakeholders or end-users to ensure that it meets their requirements and expectations. Acceptance testing can help ensure that the system meets the needs of its intended users and that it is ready for deployment.

By performing a combination of these testing methods, developers can ensure that the medical system is reliable, secure, and meets the needs of its users.

# Chapter Six: Results and Discussion

In this chapter we’re going to find out the results of the project whether they’re achieved or not and the differences between the desired results and the actual ones.

## Graduation Project 2023

### Results and Discussion

#### Expected Result:

* + - * Faster diagnosis: By asking targeted questions about symptoms, the system can help users determine the likely cause of their health issues more quickly than traditional methods.
      * Improved access to care: Users can book appointments with doctors through the system, which can be especially helpful for those who are unable to visit a doctor in person due to location or mobility issues.
      * Greater convenience: Users can upload prescriptions and medical images to the system, allowing them to easily access their medical records and treatment plans from anywhere.
      * More accurate diagnosis: By allowing users to upload medical images for analysis by healthcare professionals, the system can help ensure that diagnoses are as accurate as possible.

#### Actual Result:

1. Improved speed and accuracy of diagnosis: Through the use of a symptom checker and medical image upload, the system could assist healthcare professionals in making more accurate diagnoses in a more timely manner.
2. Increased convenience for users: The ability to book appointments with doctors and upload prescriptions and medical images through the system could make accessing healthcare more convenient for users.
3. Better record-keeping and organization: The system's ability to remember treatment dates and store medical records could help healthcare professionals keep track of patients' health histories and treatment plans more efficiently.
4. Improved patient outcomes: By enabling faster and more accurate diagnosis and providing more convenient access to healthcare, the system could potentially lead to better health outcomes for patients.

### Discussion

The concept of a medical system that can diagnose diseases by asking questions, allow users to book appointments with doctors, and upload prescriptions and medical images is an exciting development in

healthcare technology. There are several benefits to this type of system, as well as potential challenges that should be considered.

One of the main benefits of a symptom checker is that it can help users identify potential health issues more quickly. This can lead to earlier diagnosis and treatment, which can improve health outcomes. Additionally, the ability to book appointments with doctors through the system and upload prescriptions and medical images can make accessing healthcare more convenient for users.

However, there are also potential challenges to this type of system. For example, if the symptom checker is not accurate or does not ask the right questions, it could lead to incorrect diagnoses. Additionally, there may be concerns around privacy and security when it comes to uploading medical images and records to the system.

Another potential challenge is ensuring that healthcare professionals are able to effectively use the system. They will need to be properly trained in how to interpret the results from the symptom checker and how to use the system's features to provide the best possible care for their patien

# Chapter Seven:

Conclusion

## Conclusion

In conclusion, the new system has the potential to transform the healthcare industry by providing accessible, efficient, and personalized healthcare services to patients. The system's achievement lies in its ability to improve patient outcomes by providing accurate and timely diagnoses, enabling patients to conveniently book appointments and manage their medications, and facilitating communication and collaboration among healthcare professionals. Additionally, the system's non-functional requirements prioritize performance, reliability, security, usability, scalability, and interoperability, ensuring that the system operates optimally and is accessible to a wide range of users.

However, there is still room for improvement, and if given the right resources, the system could be enhanced with additional functionalities and capabilities to further improve the quality and effectiveness of healthcare services. Firstly, the system could benefit from incorporating natural language processing (NLP) to improve the accuracy and reliability of the symptom checker. This would enable the system to diagnose more complex medical conditions and provide personalized treatment options to patients. Secondly, the system could be enhanced with video conferencing and telemedicine capabilities to enable virtual consultations and remote monitoring of patients. This would increase access to healthcare services for patients who are unable to physically visit a healthcare facility, and reduce healthcare costs associated with frequent in-person visits.

Thirdly, the system could integrate with wearable devices and mobile apps to collect and analyze real-time data on users' health and lifestyle. This would enable the system to provide proactive and preventive healthcare services, and identify potential health risks before they become serious medical conditions. Fourthly, the system could leverage blockchain technology to enhance data security and interoperability. This would ensure that patient data is secure and protected from unauthorized access, while at the same time making it easier for different healthcare providers to share information and collaborate on patient care. Finally, the system could be enhanced with AIpowered chatbots to provide 24/7 support to patients, answer their questions, and provide guidance on medical conditions and treatments.

in conclusion, the new system has the potential to significantly improve the quality and accessibility of healthcare services. However, to fully unlock its potential, the system needs to be enhanced with additional functionalities and capabilities, including NLP and ML algorithms, video conferencing and telemedicine capabilities, wearable devices and mobile apps integration, blockchain technology, and AIpowered chatbots. With these enhancements, the system can provide patients with personalized and proactive healthcare services, while at the same time ensuring data security, interoperability, and compliance with healthcare regulations and standards

# Chapter Eight: Future Work

## Future work

1. Enhancing the accuracy and reliability of the symptom checker: The system could incorporate more advanced algorithms and machine learning models to improve the accuracy and reliability of the symptom checker. This would require the system to collect and analyze more data on medical conditions and symptoms, and continuously update the algorithm based on new insights and feedback from healthcare providers and patients.
2. Incorporating natural language processing (NLP) and voice recognition: The system could incorporate NLP and voice recognition technologies to enable patients to describe their symptoms in natural language and receive personalized treatment options. This would improve the user experience and reduce the risk of misdiagnosis or incorrect treatment.
3. Expanding the coverage of medical specialties: The system could expand its coverage of medical specialties beyond primary care to include more specialized fields such as oncology, cardiology, and neurology. This would require the development of more advanced symptom checkers and algorithms that are tailored to specific medical conditions and treatments.
4. Introducing remote monitoring and telemedicine capabilities: The system could introduce remote monitoring and telemedicine capabilities to enable patients to receive healthcare services from the comfort of their own homes. This would reduce healthcare costs and improve access to healthcare services for patients living in remote or underserved areas.
5. Personalizing healthcare services based on patient data: The system could leverage big data and machine learning algorithms to provide personalized healthcare services to patients based on their medical history, lifestyle, and preferences. This would enable the system to provide more targeted and effective treatment options to patients and improve patient outcomes.
6. Enhancing data security and privacy: The system could introduce more advanced security and privacy features to protect patient data from unauthorized access or breaches. This would require the system to

comply with relevant regulations and standards and implement state-of the-art encryption and authentication mechanisms.

1. Collaborating with healthcare organizations and stakeholders: The system could collaborate with healthcare organizations and stakeholders such as hospitals, clinics, insurers, and pharmaceutical companies to share data and insights, facilitate research and development, and improve healthcare outcomes. This would enable the system to tap into a wider network of expertise and resources and provide more comprehensive and integrated healthcare services to patients.