ITCS 6112 Fall 2021
Software Systems Design and Implementation
Final Term Project
Cancer Prediction Web Application
VISION

Group Members:
Rahul Govindkumar
Parth Makwana
Vishwath Kamalanathan
Milindsinh Desai
Isha Shafqat

VISION

Table of Contents

Abstract

The aim for our project is to create a software solution proposed to aid healthcare specialists for Lung Cancer prediction in the department. Our application, Vision, is a web application that will help the specialists predict the stages of a cancer patient. It'll use data from the patients, then view the current status of the patient, the potential next stages, the probability of it being correct and the medications required for this patient. This application can further allow the specialist to create appointment for the patient, if necessary. The application overall aims to aid the Doctors and Oncologist professionals to make decisions on medications and understand the patients' cancer stages for further use.

Vision is created through the Back-end development, which utilized the Python language and using Flask. The front-end development is created using Angular. The project further utilizes Tableau for data visualization.

Introduction

The Vision application provides the necessary services for Doctors and Oncologist specialists to maintain their patient information. The healthcare system has an immense amount of patient information, so the system aims to help the specialists view the cancer stages and what predictions may be for the patient. Since data visualization is beneficial in providing concise analyses, it is used within our system.

The cancer specialization is Lung cancer, thus, there is information regarding the patient which displays their basic history, such as age, height, weight, what potentially caused the cancer, their tumor size, and the cancer stage the patient is in. This application allows the Doctors to view all their patients, patient statistics, create appointments for patients, view the patients' cancer stages, and provide medication and further recommendations in their notes.

The application allows the end -user to navigate through the website and be able to execute the services they deem necessary. The goal for the project is to make an easier system for Doctors to view the patients' lung cancer stages and be able to predict the next stage based on patient symptoms. The Doctors can also provide certain medication that may aid the patient in that certain stage.

Materials and Methods

The project has used methodologies and models which aim to describe the system. The software has used phases of Requirements analysis, Design and Implementation, Validation and Verification, and future ideas for Evolution and Maintenance.

Requirements Analysis

During the Requirements Analysis phase, it was necessary that we propose the problem that this software would solve. The problem is to create an effective way for Oncologists and Doctors to view their patient's data, and through that, be able to predict the stages of cancer that a patient is in. The product would use data from patients who are affected by lung cancer and are in a specific stage, the size of tumor and the spread. Further, there are statistics of drugs that would be effective for patients at a certain phase within cancer.

The core requirements for this software are,

- The HCPs (Health Care Professionals) should be able to view patient history, be able to get the patient's stage prediction (where the patient is more likely to be given the current medication)
- The HCPs must be able to view the efficacy of each drugs at different time points to make an informed decision on the which drug to be prescribed
- Predict the efficacy of a particular drug that is tailored for each patient

The Team brainstormed and came up with additional requirements,

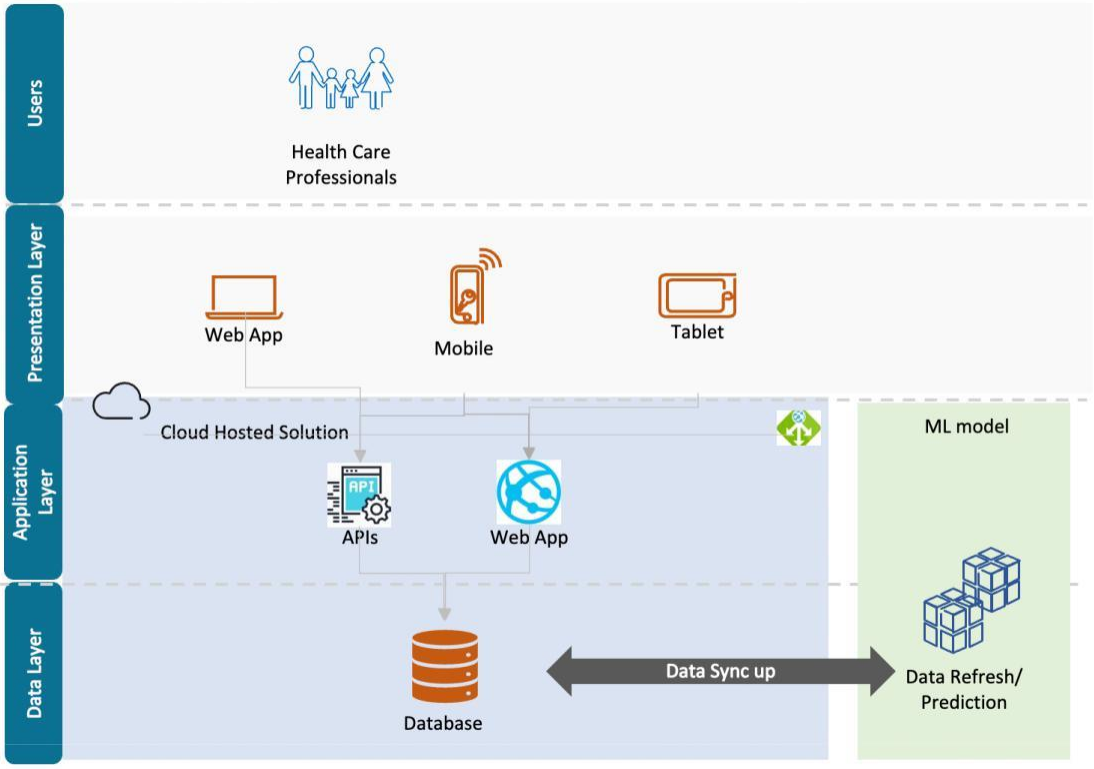
- Creating authentication so only authorized HCPs would be allowed access to the system
- Statistics page where patient and drug statistics are used to give an overall picture for the HCP
- Creating appointments directly form the application, which could be tied back to the HER system for ease
- Ability for the doctor to save notes during each appointment

Design and Implementation

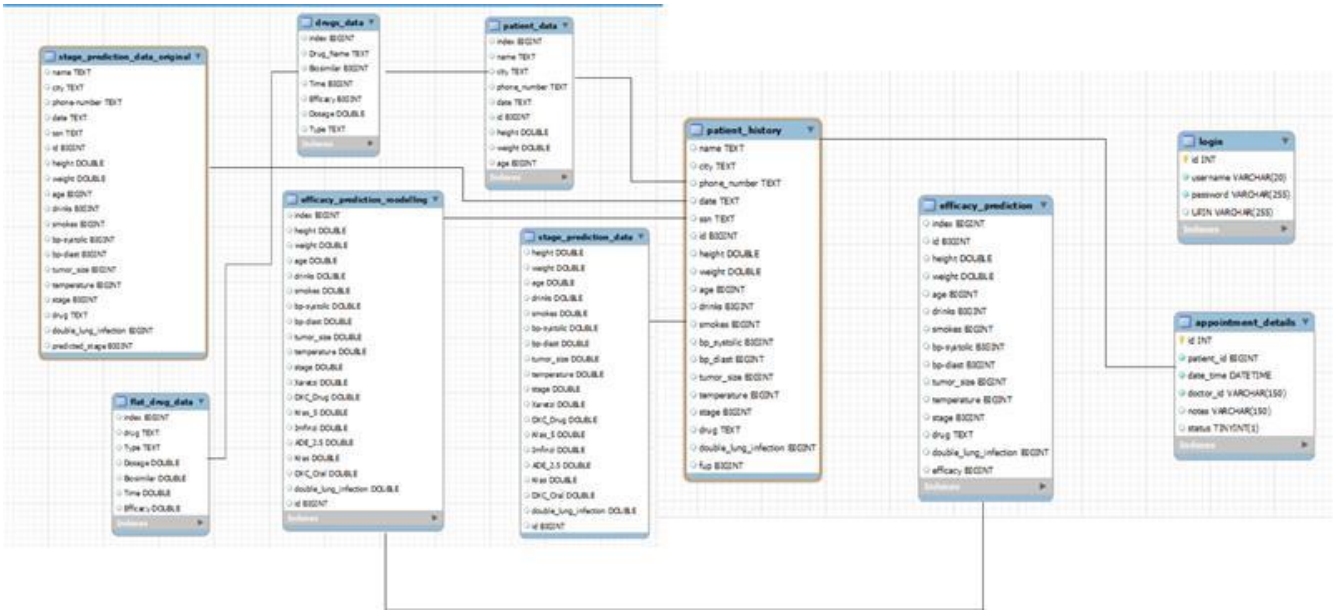
There were a variety of designs used, which are; System Architecture, ER Diagram, Use Case Diagram, Domain Model Class Diagram, Sequence Diagram, Deployment Diagram, and Component Diagram. Each diagram has helped display the models to develop the software product.

I. ***System Architecture*** - The overall system architecture portrays the users, the presentation layer of the system, the application layers, and the data layer. The users are Health care professionals. The system is presented through a web application but can further be evolved into a mobile application and tablet view. The system has used

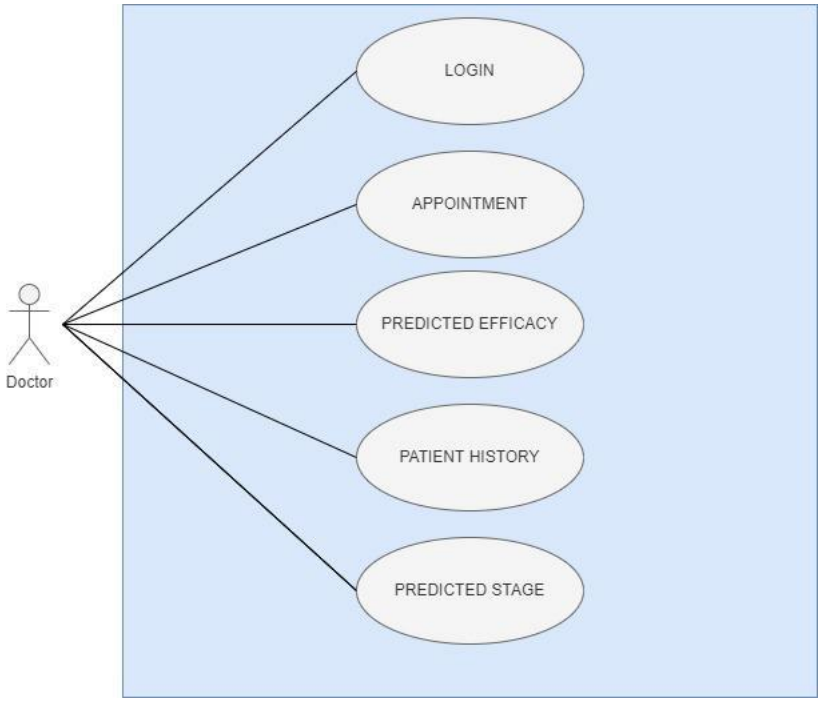
APIs and web application, along with a Machine learning model to apply the system features. And the system has used data from its database, which was synced up from the Machine learning model.



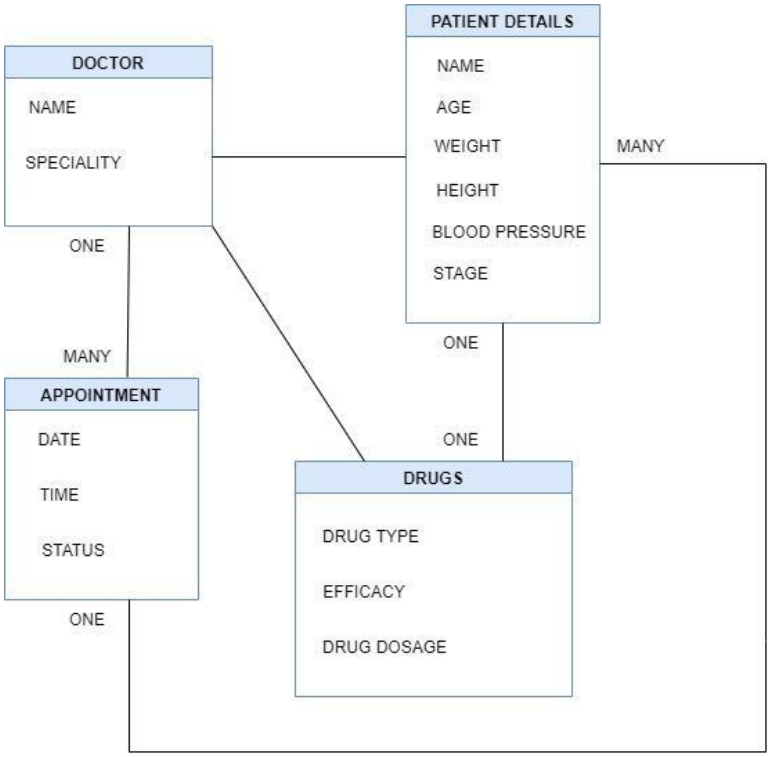
II. ***ER Diagram***–This diagram displays the entities within the system correlate to one another. These entities contain fields about the software, which further contain data from the database.



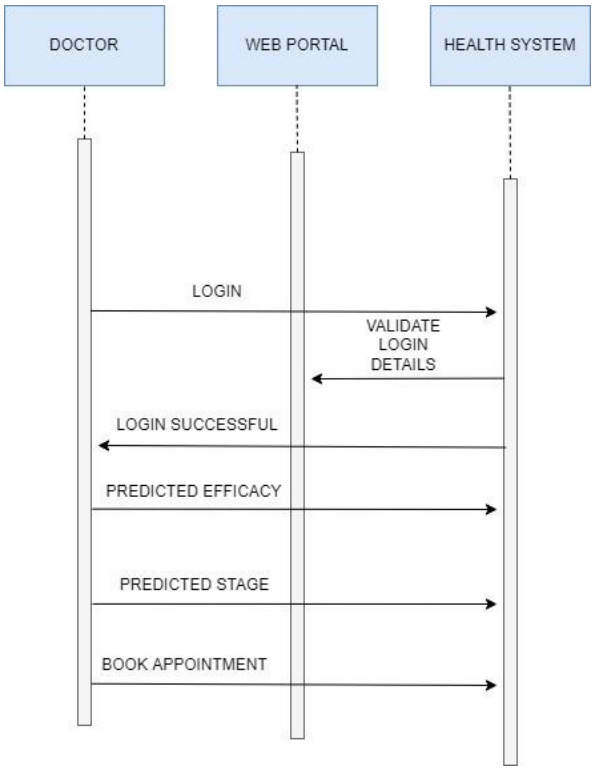
III. ***Use case Diagram*** – This diagram displays the interaction between the Doctors and Oncologists. Each are portrayed as an "actor" and there are the actions that the actor may take in this software. The "actors" are able to login, create a patient's appointment, predict the drug's efficacy, view the patient's history and view the predicted stage of lung cancer.



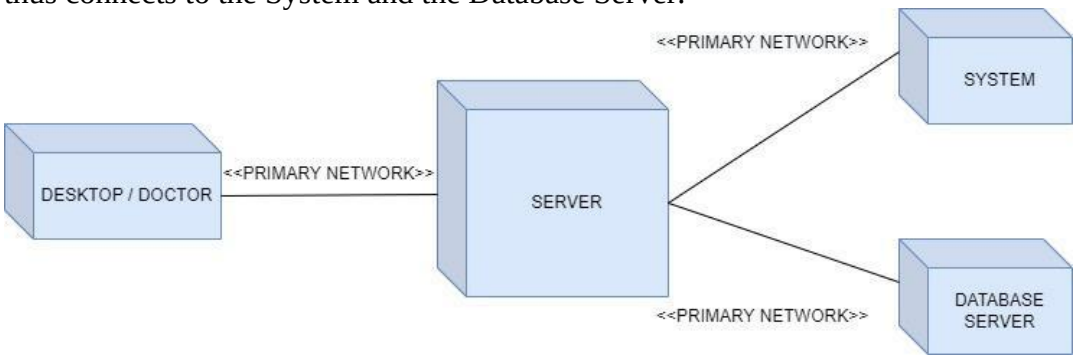
IV. ***Domain Model Class Diagram*** – This diagram shows the users' work domain and defines the requirements for the object orientation of this software. There are classes for Doctor, Appointments, Drugs, and Patient Details. Each class has fields that relate to it and define the data information. Also, this diagram shows the relations between classes.



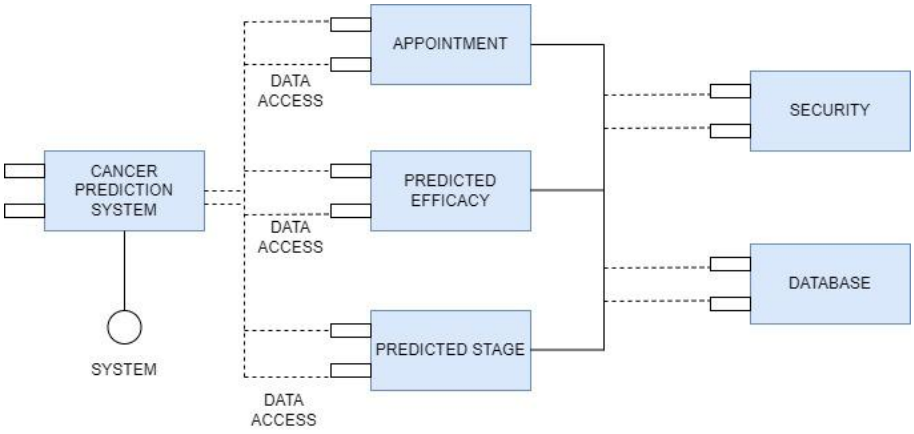
V. ***Sequence Diagram*** – This diagram shows the process of sequences that occur within the system. The doctor is able to login to the system, predict the efficacy for the drugs, view the predicted cancer stage, and book further appointments for the patient. The web portal allows the Doctors to utilize these features. Finally, the health system validates the login details for doctors.



VI. ***Deployment Diagram*** – The deployment diagram shows the architecture of the system and the components within the system. This diagram displays the component of the Desktop and Doctors, who's network is being displayed through the Server. The Server thus connects to the System and the Database Server.



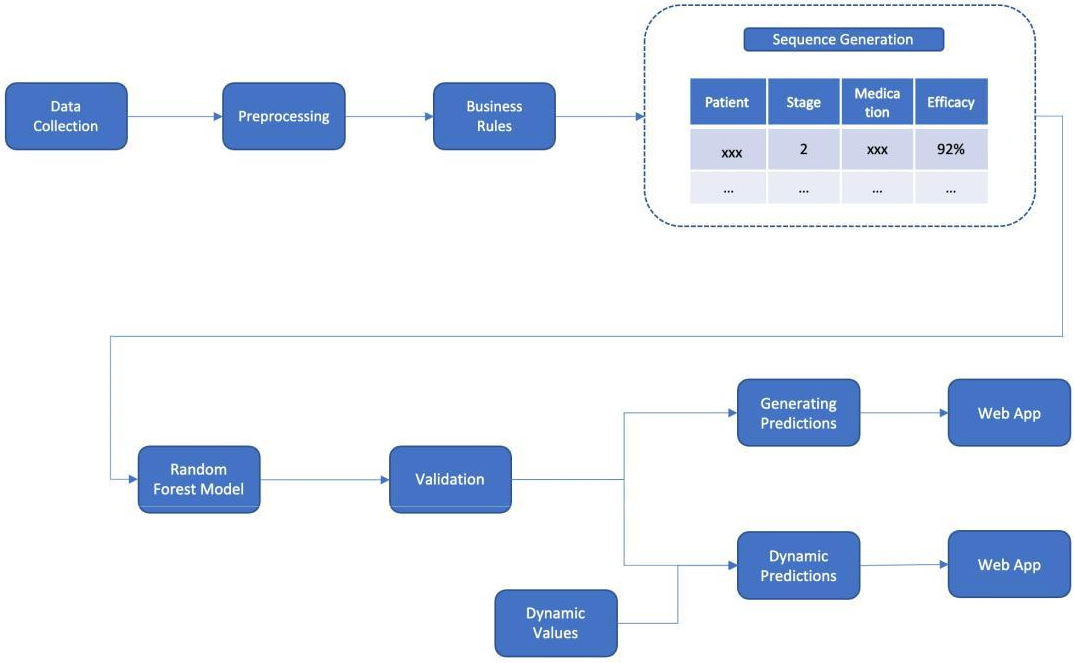
VII. ***Component Diagram*** – This diagram shows the interaction and wiring within each component and how they are connected to other components. The Cancer Prediction System uses data from the appointment, the predicted drug's efficacy, and the predicted cancer stage. Then, these are all connected to the systems security and the overall systems database.



VIII. *Model Architecture*

The model architecture gives and through overview of data ingestion to prediction. The various stages of ML architecture are as follows,

- **Data Collection -** The data is collected from various source (Patient History, Drug Data)
- **Preprocessing** – The data is preprocessed to handle missing values and standardized to have values on the same scale
- **Business Rules** – Certain business rules are applied (Currently not implementing any business rules, the code is modular to accommodate future changes)
- **Sequence Generation** – The input data is then flattened out to be passed as input into the mode
- **Random Forest Model** – Random Forest classification algorithm is used to train the mode and for prediction generation. The idea of using Random Forest algorithm is it works best for the given distribution and relatively simple and less complex to implement.
- **Prediction Generation** – The predictions are generated on the fly through API calls. The saved model from training is deployed for prediction in the server side
- **Data Refresh** – The model will be retrained at the end of every month and will as new data points come into the database, thus creating a more accurate model that can generalize better



Validation and Verification

During the validation and verification process, the product was tested to assure the correctness when relating to the software requirements. The software product must be ensured that it is the "correct" product, and that the product is "right". The results were correct as we expected the program to run, and the results were also valid, in reference to having the patient's data and lung cancer stage prediction to be correct.

At each stage of the integration process, test cases were generated and tested before components were integrated into the system.

Technologies Used

Frontend – Angular, HTML

Backend – Flask, Python

Data Base – MySQL, SQL

Version Control – Git

Machine Learning – Python

Evolution

For further development and maintenance of this software, the product could evolve to help the clients better. An example for maintenance would be to alter the database for any excess information for the patients or the cancer. If the lung cancer has developed an immunity to a certain drug, that would need to be recorded into the database. There can also be evolution of the Presentation of the system. This system could be used in a mobile application for effectiveness for Doctors. Lastly, the system could be further developed for patient-use. The application would let the patient's families be able to view their cancer stage and how to have certain treatments at home.

Results

The software system resulted in the users, such as the Doctors and Oncologists, to view the patient's data and the prediction of their cancer stages, the probability of how accurate that prediction will be, and which drug would be most effective for a certain patient. The Doctor is able to search a patient and choose a drug, then view the efficacy of that drug for that patient. The Doctor can view the previous history of a patient in order to see their details and previous medication. Also, the Doctor is able to schedule an appointment for that patient and write any notes during their appointment.

Conclusion

The Cancer Prediction Web Application, Vision, helps Doctors and Oncologists to support patients' data. These individuals are cancer patients who are in certain stages of lung cancer. The web application has patients' history and information, such as their age, height, weight, stage of cancer, their tumor size, and how much the cancer has spread. The end user has the ability to view patients' statistics and use the prediction feature to view the patient's cancer stage and the probability of how accurate that prediction is. Further, the user has the ability to predict drug would be most effective for a patient.

References

1. Breiman, L. Random Forests. *Machine Learning* **45,** 5–32 (2001). https://doi.org/10.1023/A:101093340