A PROJECT REPORT

ON

"EARLY DETECTION OF CHRONIC KIDNEY FAILURE "

Submitted to

UNIVERSITY OF MUMBAI

In Partial Fulfilment of the Requirement for the Award of

BACHELOR'S DEGREE IN
COMPUTER ENGINEERING

BY

Shaikh Almas Javed Nikhat	**16DCO45**
Shaikh Fatima Moin Khalida	**16DCO74**
Baig Mehnaaz Mushtaque Kulsum	**16DCO50**

UNDER THE GUIDANCE OF
Prof. Syed Aamer

DEPARTMENT OF COMPUTER ENGINEERING
Anjuman-I-Islam's Kalsekar Technical Campus
SCHOOL OF ENGINEERING & TECHNOLOGY
Plot No. 2 3, Sector - 16, Near Thana Naka,
Khandagaon, New Panvel - 410206

2018-2019

AFFILIATED TO

UNIVERSITY OF MUMBAI

A PROJECT II REPORT
ON

"EARLY DETECTION OF CHRONIC KIDNEY FAILURE"

Submitted to
UNIVERSITY OF MUMBAI

In Partial Fulfilment of the Requirement for the Award of

BACHELOR'S DEGREE IN
COMPUTER ENGINEERING

BY

Shaikh Almas Javed Nikhat	**16DCO45**
Shaikh Fatima Moin Khalida	**16DCO74**
Baig Mehnaaz Mushtaque Kulsum	**16DCO50**

UNDER THE GUIDANCE OF
Prof. Syed Aamer

DEPARTMENT OF COMPUTER ENGINEERING
Anjuman-I-Islam's Kalsekar Technical Campus
SCHOOL OF ENGINEERING & TECHNOLOGY
Plot No. 2 3, Sector - 16, Near Thana Naka,
Khandagaon, New Panvel - 410206

2018-2019
AFFILIATED TO

UNIVERSITY OF MUMBAI

Anjuman-i-Islam's Kalsekar Technical Campus

Department of Computer Engineering
SCHOOL OF ENGINEERING & TECHNOLOGY
Plot No. 2 3, Sector - 16, Near Thana Naka,
Khandagaon, New Panvel - 410206

CERTIFICATE

This is certify that the project entitled

"Early Detection Of Chronic Kidney Failure"

submitted by

Shaikh Almas Javed Nikhat	**16DCO45**
Shaikh Fatima Moin Khalida	**16DCO74**
Baig Mehnaaz Mushtaque Kulsum	**16DCO50**

is a record of bonafide work carried out by them, in the partial fulfilment of the requirement for the award of Degree of Bachelor of Engineering (Computer Engineering) at ***Anjuman-I-Islam's Kalsekar Technical Campus, Navi Mumbai*** under the University of MUMBAI. This work is done during year 2018-2019, under our guidance.

Date: / /

(Prof. Syed Aamer)
Project Supervisor

(Prof. Borke Kalpana)
Project Coordinator

(Prof. Khan Tabrez)
HOD, Computer Department

DR. ABDUL RAZAK HONNUTAGI
Director

External Examiner

Acknowledgements

We would like to take the opportunity to express my sincere thanks to my guide **Syed Aamer**, Assistant Professor, Department of Computer Engineering, AIKTC, School of Engineering, Panvel for his invaluable support and guidance throughout my project research work. Without his kind guidance & support this was not possible.

I am grateful to him for his timely feedback which helped me track and schedule the process effectively. His time, ideas and encouragement that he gave is help me to complete my project efficiently.

We would like to express deepest appreciation towards **DR. ABDUL RAZAK HONNUTAGI**, Director, AIKTC, Navi Mumbai, **Prof. Khan Tabrez**, Head of Department of Computer Engineering and **Prof. Bodke Kalpana**, Project Coordinator whose invaluable guidance supported us in completing this project.

At last we must express our sincere heartfelt gratitude to all the staff members of Computer Engineering Department who helped me directly or indirectly during this course of work.

Shaikh Almas Javed Nikhat

Shaikh Fatima Moin Khalida

Baig Mehnaaz Mushtaque Kulsum

Project I Approval for Bachelor of Engineering

This project entitled ***Ëarly Detection On Chronic Kidney Failure¨*** by ***Shaikh Almas Javed Nikhat,Shaikh Fatima Moin Khalida,Baig Mehnaaz Mushtaque Kulsum*** is approved for the degree of ***Bachelor of Engineering in Department of Computer Engineering.***

Examiners

1. .....
2. .....

Supervisors

1. .....
2. .....

Chairman

.....

Declaration

We declare that this written submission represents my ideas in my own words and where others ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been taken when needed.

Shaikh Almas Javed Nikhat
16DCO45

Shaikh Fatima Moin Khalid
16DCO74

Baig Mehnaaz Mushtaque Kulsum
16DCO50

ABSTRACT

The growing global burden of Non-communicable disease(NCD's) worldwide, is increasing day by day. The chronic renal failure is a type of NCD. These chronic disease are one of the leading cause of death. It becomes important for our society, to detect and cure it as early as possible. Our system aims to predict the possibility of chronic renal failure, i.e the chances of kidney failure of a patient. Huge amount of patient's data and their case histories are stored from years and years, and yet not being used.

This data can be used to predict, using massive data sets, by categorizing valid and unique patterns in data. We aim to make a system through which people can regularly check the risk to have chronic renal failure.

Keywords: CRF(Chronic Renal Failure), NCD(Non Communicable Disease), Massive Data sets, Prediction Algorithms, Machine Learning.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Purpose

Currently, kidney disease is a major problem and is an area of concern. It is a condition where kidney don't work as well as normal. CRF does not usually cause any symptoms until it has reached an advance stage. Kidney disease is very dangerous if not immediately treated on time, and may be fatal. If the doctors have a good tool that can identify patients who are likely to have kidney disease in advance, they can heal the patients in time. So, we are developing a system that can predict kidney damage at early stage. Our focus is on predicting CRF disease using algorithm. Large amount of data is being generated by medical industry that medical history is being unused. So it is required to process those unused data. It comes up with the set of techniques which when applied to this processed data, generates report for making appropriate decision and to make people aware about that the kidney failure can be predicted at early stages.

1.2 Project Scope

This Project will help to provide the immediate report to the patients that they are whether suffering from any kidney disease .It is to make people aware about that the kidney Failure can be detected at early stages.So that they can take precautions and start healing it at early stages rather ending up with either transplant of kidney or Dialysis.

1.3 Project Goals and Objectives

1.3.1 Goals

The goals of our system are:

- Helps to predict the CRF at early stages.
- Providing the immediate report to the patients.

- Maintains the patients report.

1.3.2 Objectives

Objective of this project is to predict the CRF at the early stage as well provides the immediate reports to the patients. Instead of waiting for the report to get ready for more than days the system will help to provide the immediate report on that day itself.Hence this will save the time of patients so that at early stage they get to know about any kidney disease from which they are suffering.

1.4 Organization of Report

- Chapter 1 : Gives a brief introduction about our project.
- Chapter 2 : Describes the literature review of the existing papers and the description about the system.
- Chapter 3 : Discuss about the project planning different roles and capability of the team members.Also discuss the budget of the project.
- Chapter 4 : Describe the brief description of the srs and the other requirement of the project.
- Chapter 5 : Shows the system architecture,functional requirements and different diagram of the project.
- Chapter 6 : Shows the implementation of the different coding.
- Chapter 7 : Shows the different testing performed and the problem faced.
- Chapter 8 : Discuss the step by step implementation.
- Chapter 9 : Discuss about the black book chapters.

Chapter 2

Literature Survey

2.1 Predictive Analytics for Chronic Kidney Disease Using Machine Learning Techniques

In this paper, Ho, Pai, Pheng, Lee and Chen present machine learning techniques for predicting the chronic kidney disease using clinical data. Four machine learning methods are explored including K-nearest neighbors (KNN), support vector machine (SVM), logistic regression (LR), and decision tree classifiers. These predictive models are constructed from chronic kidney disease data set and the performance of these models are compared together in order to select the best classifier for predicting the chronic kidney disease.They also presented a computer-aided diagnosis implement based on analyzing images.

2.1.1 Advantages of Paper

a. It will help in detecting and classifying different stages of CKD.

b. Use of different machine learning technique.

c. It also present a computer-aided diagnosis implement based on analyzing images.

2.1.2 Disadvantages of Paper

a. As they also present a computer-aided diagnosis implement based on analyzing images.So they required more than 1000 images to predict.

b. SVM has the highest sensitivity.

c. For prediction required more data set.

2.1.3 How to overcome the problems mentioned in Paper

a. All machine learning techniques have to be trained and tested by the proposed method.

2.2 CHRONIC KIDNEY DISEASE ANALYSIS USING DATA MINING CLASSIFICATION TECHNIQUES

In this paper, Veenita,Khushboo,A. Sai,Abhay present the different Data mining techniques which has been a current trend for attaining diagnostic results.For this Huge amount of unmined data is collected by the healthcare industry in order to discover hidden information for effective diagnosis and decision making. Data mining is the process of extracting hidden information from massive data set,categorizing valid and unique patterns in data . There are many data mining techniques like clustering, classification, association analysis, regression etc.The experimental results implemented in Rapidminer tool show that Naive Bayes produce more accurate results than Artificial Neural Network.The experimental results implemented in Rapidminer tool show that Naive Bayes produce more accurate results than Artificial Neural Network.

2.2.1 Advantages of Paper

a. Different use of data mining techniques.

b. Use of Naive Bayes Algorithm which indicate 100 percent of accuracy.

c. Provide effective diagnosis and decision making.

2.2.2 Disadvantages of Paper

a. It required huge amount of data set.

b. They have just compared the algorithms and not made a system which could predict these diseases.

2.2.3 How to overcome the problems mentioned in Paper

a. We will use Naive Bayes Algorithm and make a system which will predict CRF.

b. We will using this Rapidminer tool for checking the accuracy between algorithms.

2.3 Chronic Kidney Disease: A Research And Public Health Priority

In this paper, presents the importance of detection of CRF. Chronic renal failure is defined as either kidney damage or glomerular filtration rate less than 60ml/min for three months or more.This is invariably a progressive process that results in end stage renal disease.the importance of detection of CRF and detailed information of CRF is provided with causes and treatment.It is increasingly recognised that the burden of CKD not limited to its implications on demand.

2.3.1 Advantages of Paper

a. It defined the importance of CRF detection.

b. Statistics is provided .

c. It also provide the treatment and the prevention.

2.3.2 Disadvantages of Paper

a. Lots of clinical data of patients.

b. Patients with high Serum creatinine level can cause CRF.

c. Diabetes ,Hypertension are the risk factor of CRF.

2.3.3 How to overcome the problems mentioned in Paper

a. To prevent from the CRF it should be detected as early as possible

2.4 Technical Review

Our system is basically based on Data Science which involves :

a. Machine learning for making predictions.

b. Predictive causal analytics.

c. Prescriptive analytics.

For the Front End : Django , HTML .

For the Back End : Machine Learning , Mysql.

2.4.1 Advantages of Technology

a. Django helps in making it both cost effective and efficient.it is ensured that developers don't commit any mistakes related to security. To meet the heaviest traffic demand, the benefits of Django framework can be seen.When you are doing it in Django, it is ensured that developers don't commit any mistakes related to security.

b. HTML is widely used.Every browser supports HTML language.Easy to learn and use.It is by default in every window so you don't need to purchase extra software.You can integrate HTML with CSS, JavaScript, php etc.

c. Easily identifies trends and patterns.No human intervention needed. Continuous Improvement can be done.Handling multi-dimensional and multi-variety data.An also it is a Wide Applications.

2.4.2 Reasons to use this Technology

a. Django : When building a site that requires a strong and secure foundation that protects transactions and sensitive data, such as an ecommerce site, Django is a great framework to use. It hides your source code by default, and it's often one of the first frameworks to respond to a new vulnerability.

b. Machine Learning : Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

Chapter 3

Project Planning

3.1 Members and Capabilities

Table 3.1: Table of Capabilities

SR. No	Name of Member	Capabilities
1	Shaikh Almas Javed	UI Design
2	Shaikh Fatima Moin	Database
3	Baig Mehnaaz Mushtaque	UI Design

Work Breakdown Structure

3.2 Roles and Responsibilities

Table 3.2: Table of Responsibilities

SR. No	Name of Member	Role	Responsibilities
1	Shaikh Almas Javed	Team Leader	UI Design
2	Shaikh Fatima Moin	Member	Database
3	Baig Mehnaaz Mushtaque	Member	Coding

3.3 Assumptions and Constraints

3.3.1 Assumptions

a. Data provided by the user should be true.

b. The prediction would be 100 percent.

3.3.2 Constraints

a. Details of the patients should be maintained.

b. Pre-processing of report.

3.4 Project Management Approach

We have use Agile Methodology for the development of the project.AGILE methodology is a practice that promotes continuous iteration of development and testing throughout the software development life cycle of the project. Both development and testing activities are concurrent unlike the Waterfall model.Agile methodology is often compared with the waterfall model in the software development industry. However, agile approach is considered to be better. It uses an incremental approach where a sample prototype is discussed with the customer.The idea is to maintain product's quality in the entire phase of development.

Fig. Agile Model

3.5 Ground Rules for the Project

a. Show up on time and come prepared. Be prompt in arriving to the meeting and in returning from breaks.

b. Stay mentally and physically present.

c. Contribute to meeting goals.

d. Let everyone participate.

e. Listen with an open mind.

f. Think before speaking.

g. Stay on point and on time.

h. Attack the problem, not the person.

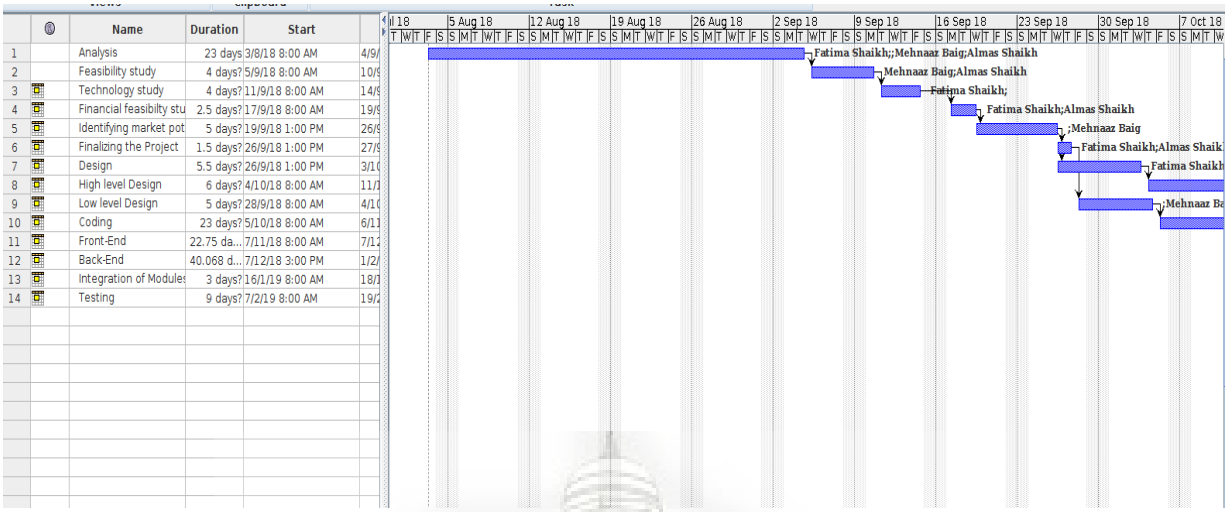
3.6 Project Budget

Django - Open Source

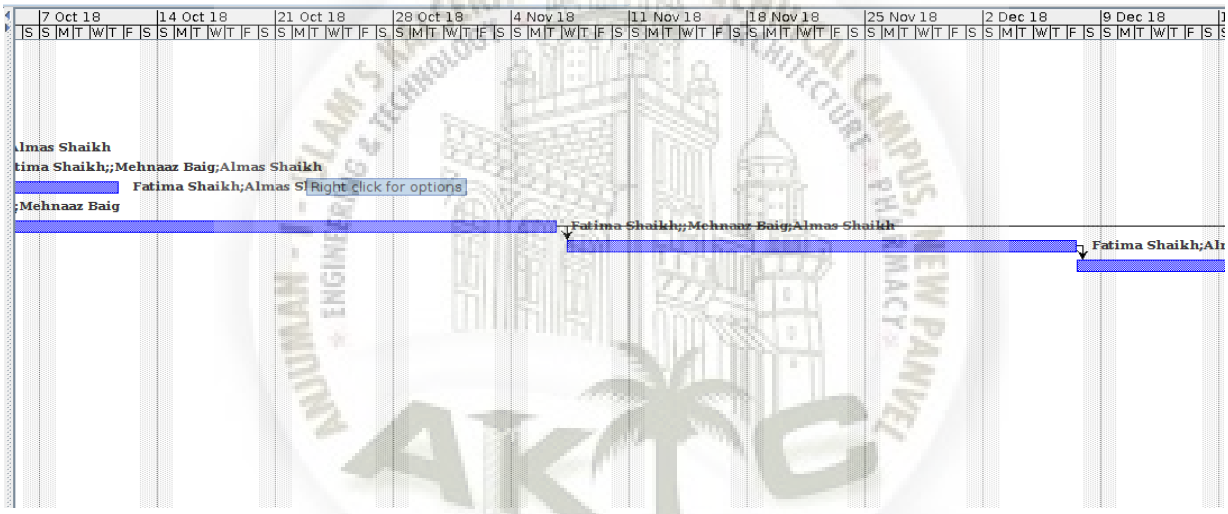
Python - Open Source

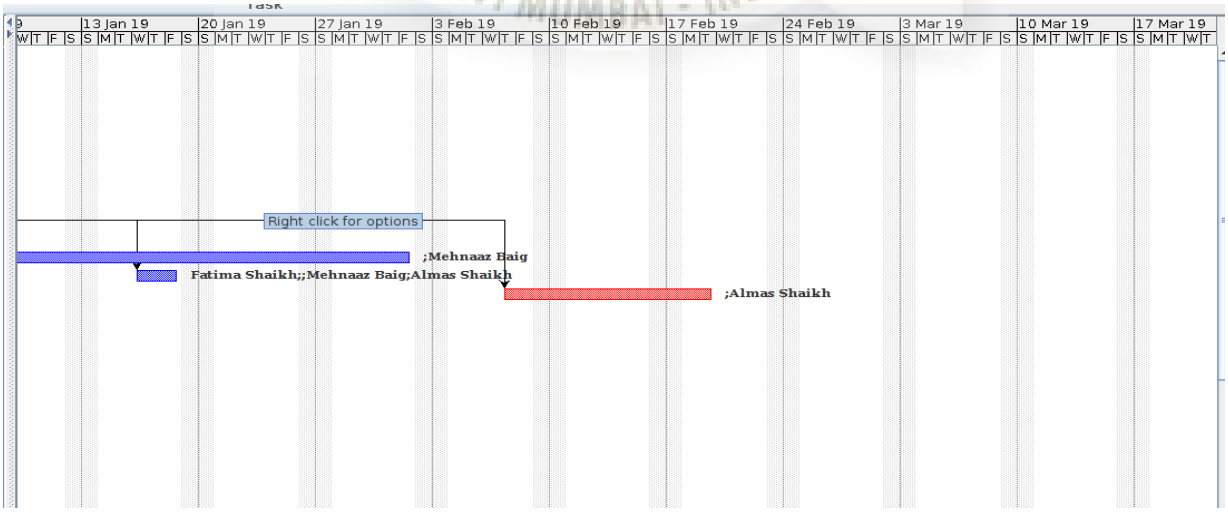
No Budget Cost.

3.7 Project Timeline



	Name	Duration	Start
1	Analysis	23 days	3/8/18 8:00 AM
2	Feasibility study	4 days?	5/9/18 8:00 AM
3	Technology study	4 days?	11/9/18 8:00 AM
4	Financial feasibilty stu	2.5 days?	17/9/18 8:00 AM
5	Identifying market pot	5 days?	19/9/18 1:00 PM
6	Finalizing the Project	1.5 days?	26/9/18 1:00 PM
7	Design	5.5 days?	26/9/18 1:00 PM
8	High level Design	6 days?	4/10/18 8:00 AM
9	Low level Design	5 days?	28/9/18 8:00 AM
10	Coding	23 days?	5/10/18 8:00 AM
11	Front-End	22.75 da...	7/11/18 8:00 AM
12	Back-End	40.068 d...	7/12/18 3:00 PM
13	Integration of Modules	3 days?	16/1/19 8:00 AM
14	Testing	9 days?	7/2/19 8:00 AM





Chapter 4

Software Requirements Specification

4.1 Overall Description

4.1.1 Product Perspective

The project is basically based on the prediction and prepossessing of the data set and providing the generated report to he patients.As we have collected the data set of the patients suffering from CRF.The details of the patients will be feeded in the system.We will provide this system to the doctor which will be having the data of the patients. We have design two test the Test 1 and Test 2. Test 1 is the general test for checking if the patients is suffering from diabetes or whether he/she smokes.After that the Test 2 will predict whether the patient is in early stage of CRF or the disease has been increase.Than as per that the immediate report will be provide to the patients.

4.1.2 Product Features

The system will make easier for the patients to predict the disease at the early stages.So to take precaution at the early stage.It will also provide the immediate data to the patients.The system will be provided to the doctors while taking the general test of the patients he/she can aware the patients whether they are suffering from CRF or not.

4.1.3 User Classes and Characteristics

The users of the system will be the Lab expert and Doctor.Lab Expert is an important actor in our system. He /She is going to upload the reports of patients suffering from chronic renal failure on a regular basis. This is important, as the increase in amount of data in the data set, more accurate prediction will be done.Doctor is going to use our system for prediction of chronic renal failure. The doctor will file a case,take all the necessary inputs from patient like patient details, symptoms and previous history. This data will get stored in our system which will be used for prediction.

4.1.4 Operating Environment

The environment in which the system will operate should be Django and the platform on which the software will run can be any browser. Further the required specification the system will need are :

a. Using of proper browser.

b. Proper internet connection.

4.1.5 Design and Implementation Constraints

The major challenge that will hurdle the development of the system is required the clean and huge amount of data set.With the help of accurate data set the system will pre-process the data which will than predict the CRF and provide the immediate report. The another constraints is that the doctor should know how to use the system.So that they can provide the accurate report to the patients.

4.2 System Features

The features of the system is to predict the CRF at the early stage, so that the patient can start taking the precautions.

4.2.1 System Feature

a. Provide accuracy.

b. Immediate Prediction.

Description and Priority

a. Provide Accuracy : To maintain the accuracy the system should provide with the accurate data set .So that the data can get pre-process and can predict the CRF disease.

b. Immediate Prediction : With the help of immediate prediction the patient will get to know at early stage whether they are suffering from the CRF or not.

Stimulus/Response Sequences

a. Doctor will login with his id and password.

b. As the patients comes to the doctor , doctor will make entry of the patient with the help of name and unique Id.

c. Now the patient is logged in , patients has to go under two test . The Test1 and Test 2.

d. Then the immediate report will be provided to the patient.

Functional Requirements

REQ-1. Doctor must login first.

REQ-2. Doctor has to make the entry of patients.

REQ-3. Patients has to take tests.

REQ-4. Immediate report will be provided to patient.

REQ-5. The server should response quickly.

4.3 External Interface Requirements

4.3.1 User Interfaces

As the user of our system is the Doctor and the Lab expert. We also has the patients as the user but as the system will be provided to the doctor because every one believe in doctors report.So the Doctor will login the system with details and will make patients enter with its unique id .Then the test will appear, with the help of the tests the data entered in the system will predict CRF.

4.3.2 Hardware Interfaces

a. 4 GB Ram.

b. 500 GB Hard Disk Minimum.

4.3.3 Software Interfaces

a. Operating System : Window,Linux.

b. Front end : Django , HTML.

c. Back end : Machine Learning , Mysql.

4.3.4 Communications Interfaces

a. The website will support all types of browser.

b. The interface between the database and the system will be done by using http protocol.

c. Proper internet connection.

4.4 Nonfunctional Requirements

4.4.1 Performance Requirements

As the system has two user the Doctor and the Lab Expert.The can logged in with it details and password and that is also applicable to the Lab Expert too.As the Doctor get logged in and the patients arrives the doctor will make the entry of the patient with help of name and unique id.As now the patients has been entered the test begins with test 1 which is the general test and followed by the test 2 which is a test related to the CRF.As the test get over the patients will get to whether they can suffer from the CRF disease or not. This will be done by the pre-processing and the prediction of the data set.

4.4.2 Safety Requirements

Backups is the major safety requirement if due the catastrophic failure, such as the disk crash,the recovery method restores a past copy of the database.Database Security applying Statistical Method.

4.4.3 Security Requirements

a. Access control : Access control is a security technique that regulates who or what can view or use resources in a computing environment.

b. Auditing : A computer security audit is a manual or systematic measurable technical assessment of a system or application.

c. Authentication : Authentication is distinct from authorization , which is the process of giving individuals access to system objects based on their identity

d. Encryption : The translation of data into a secret code. Encryption is the most effective way to achieve data security.

e. Integrity controls : Integrity refers to methods of ensuring that data is real, accurate and safeguarded from unauthorized user modification.

f. Application security : Application security is the use of software, hardware, and procedural methods to protect applications from external threats.

Chapter 5

System Design

5.1 System Requirements Definition

Our system is based on Data Science on a website ,the system function is to provide easy and immediate report to the patients.So the patients can take proper treatment as early as possible.

5.1.1 Functional requirements

a. The Doctor has to login in the system.

b. Make the entry of patients.

c. Perform two test.

d. Predict the CRF at early stage.

e. Generate immediate report.

Use-case Diagram

In the Use case of this system,we have two actor the Doctor and the Lab Expert, first both actor has to register themselves by providing the unique id and password.Now the Doctor and Lab Expert is logged in , the doctor will case a file of patient by entering there details.Then the data or test taken by the patient will be moved to the Training Engine which will train the data set,now the data will move to the Data Prepossessing where it will verify the data and predict the CRF and lastly will Generate Report.

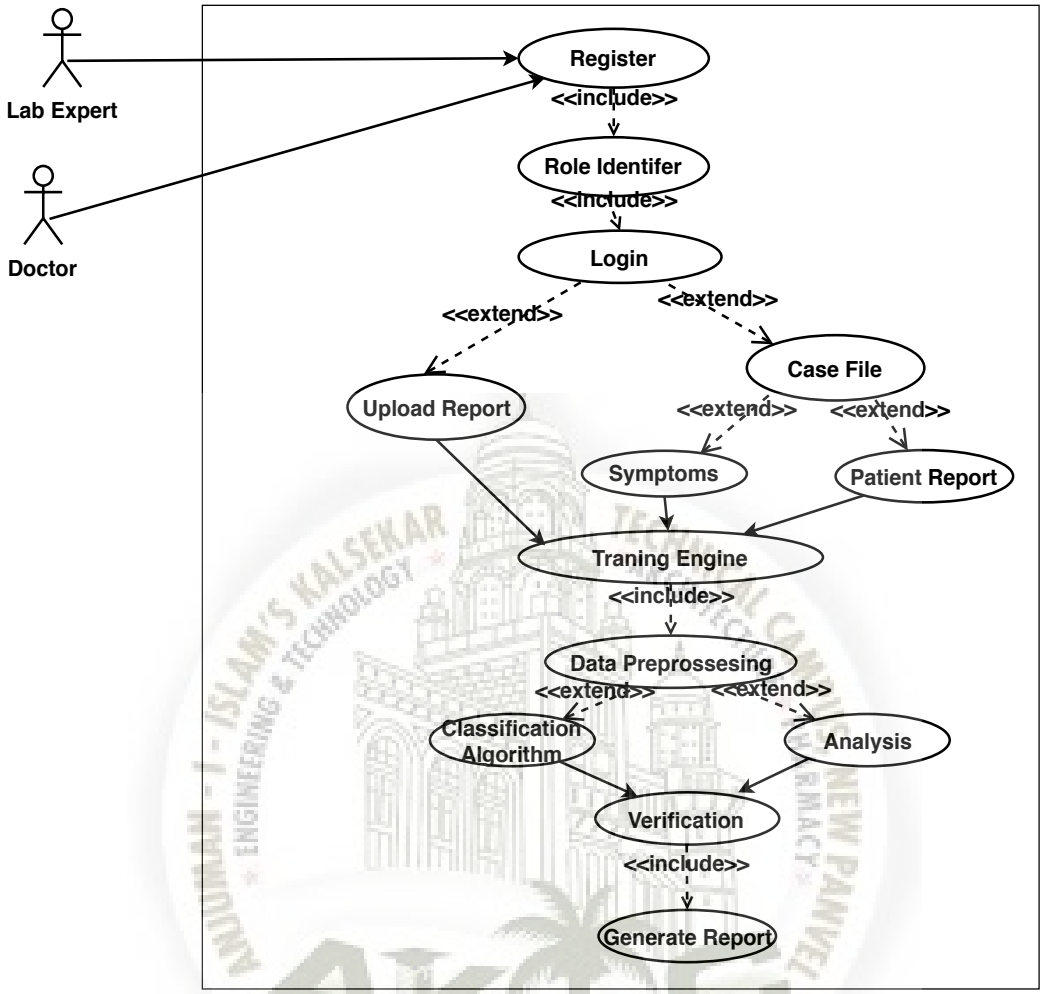


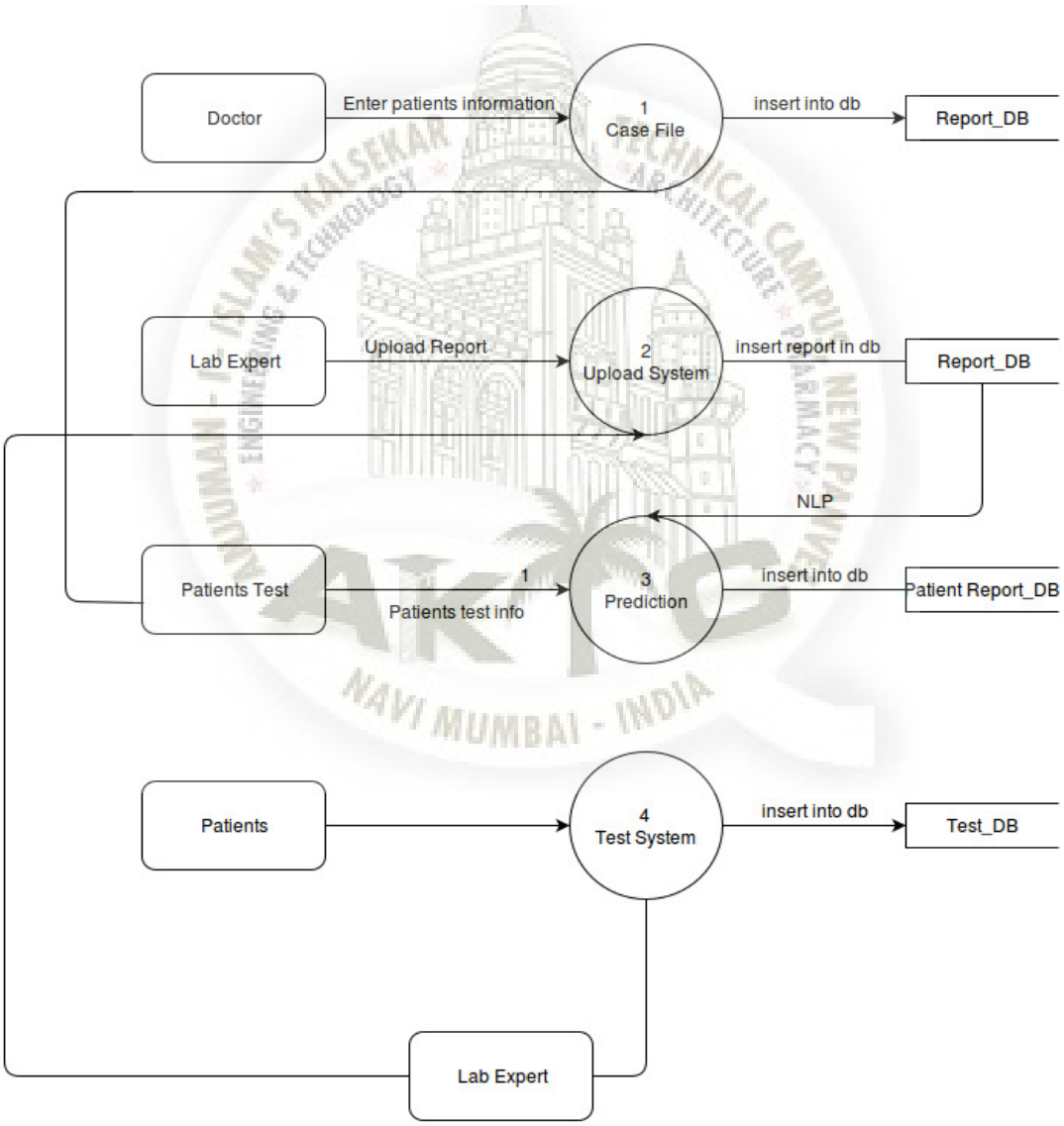
Figure 5.1: Use case

Data-flow Diagram

LEVEL 0

Figure 5.2: DFD Level 0

DFD 1



LEVEL 1

Figure 5.3: DFD Level 1

DFD 2

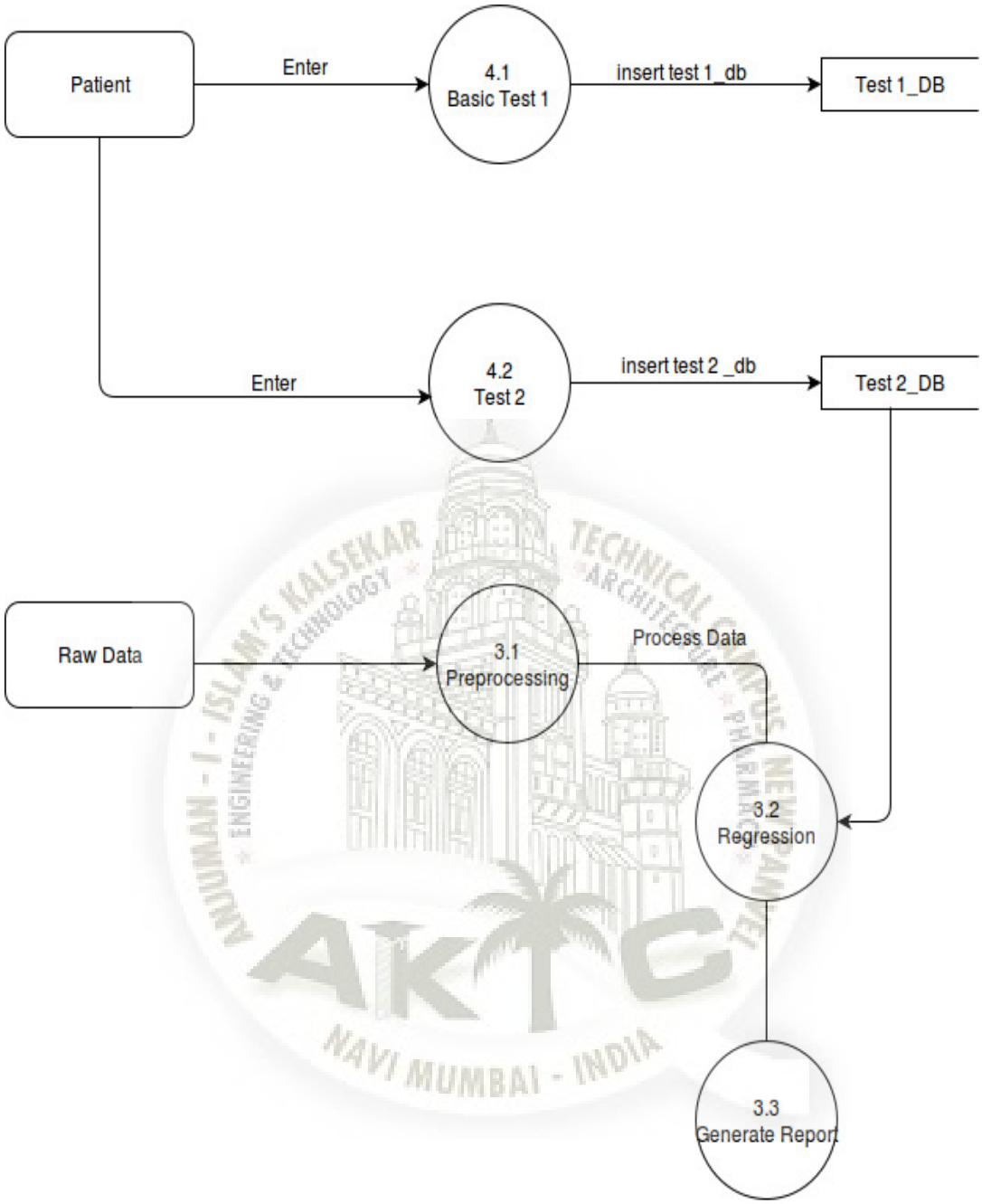


Figure 5.4: DFD Level 2

In the DFD of this system , we have two level like Level 0,Level 1 and Level 2. In Level 0 we are predicting the CRF with the help of symptoms and predicting the report.As in Level 2 we have Doctor, Lab Expert and Patients Test where will case a file upload in the system predict the CRF and maintain the data in data base.Lastly we have Level 2 in which we are performing tests , we have two test the general test and the CRF test which will help to predict the CRF at early stage.

5.1.2 System requirements (non-functional requirements)

5.1.3 Performance Requirements

The system should have high performance and low failure rates. The hardware and software should be able to transmit/receive data from databases.The immediate report generation is depend on the two test.

Algorithm to Naive Bayes

a. Step 1: Convert the data set into a frequency table

b. Step 2: Create Likelihood table by finding the probabilities.

c. Step 3: Now, use Naive Bayesian equation to calculate the posterior probability for each class. The class with the highest posterior probability is the outcome of CRF prediction.

5.2 System Architecture Design

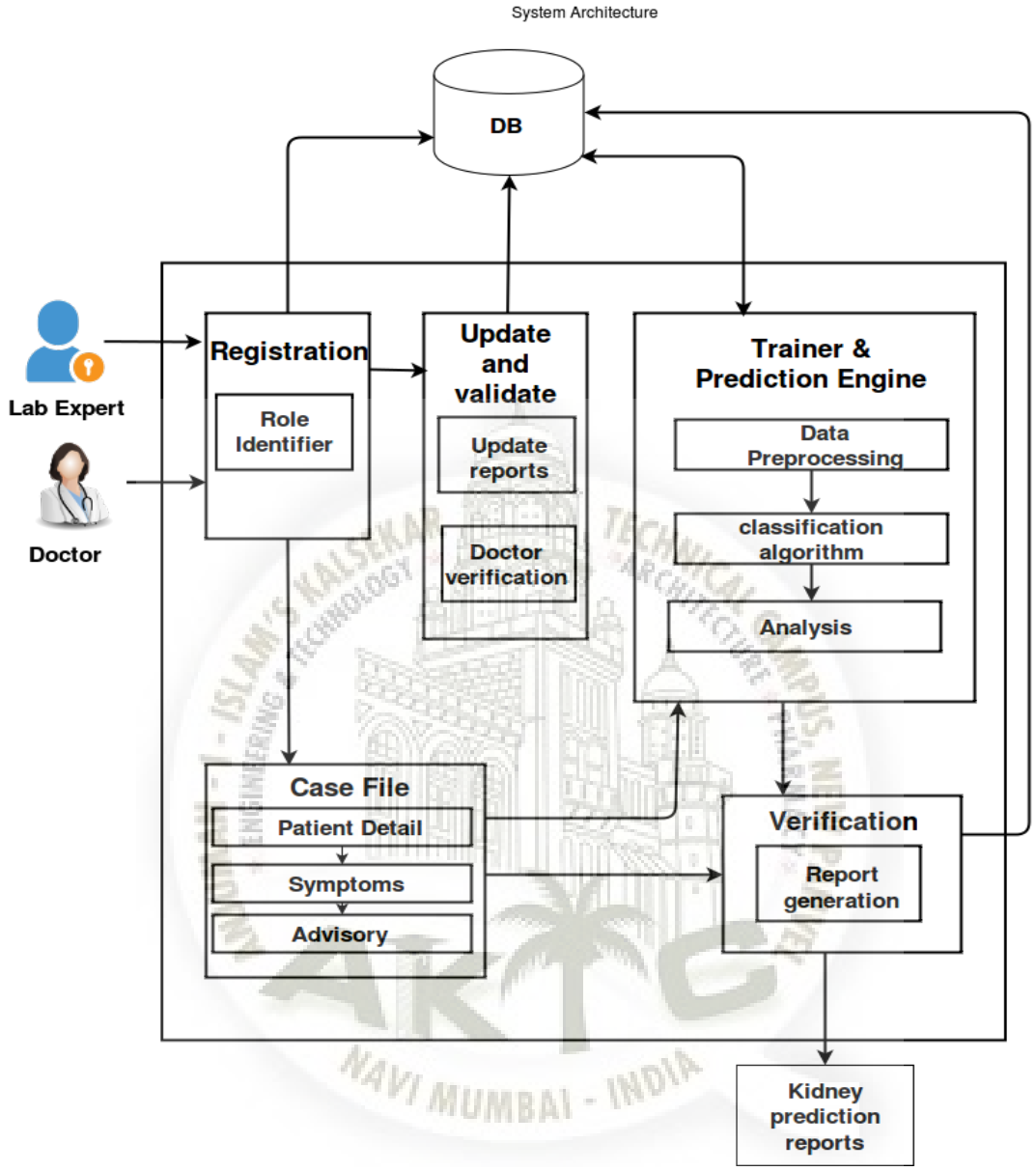


Figure 5.5: System Architecture

A system architecture or systems architecture is the conceptual model that defines the structure, behavior, and more views of a system. An architecture description is a formal description and representation of a system, organized in a way that supports reasoning about the structures and behaviors of the system.In our system we have two actor the Doctor and the Lab Expert . The First module is the Registration module were the Doctor can get register. Second is the update and validate which is for updating of the report which will be done by the Lab Expert.The most important module is the Trainer and Prediction Engine which will help to predict the CRF at early stage.

5.3 Sub-system Development

The module description provides the detailed information about the system and its supported components which is accessible in different manner.Following are the modules.

a . Login.

b . Test.

c . File Upload.

5.3.1 Login Module

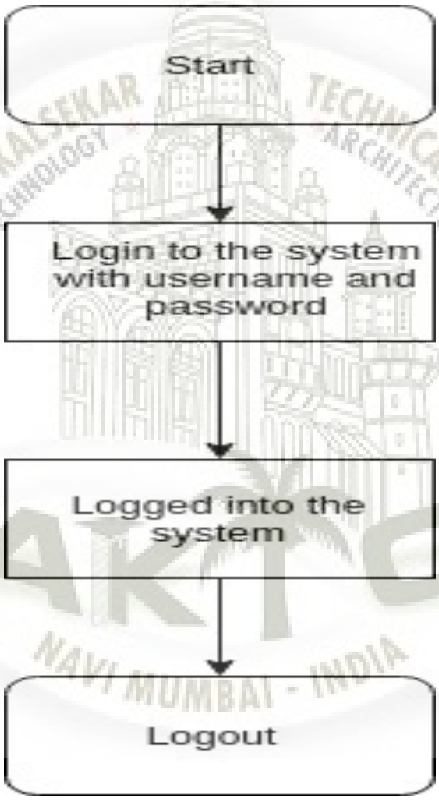


Figure 5.6: Login

In this Module we have two login , doctor login and lab expert login. The Doctor can login to the system with help of user name and password.After login the doctor will introduce the system to the patients by entering the details and unique id of patients.The second login is of Lab expert, the lab expert can login to the system with its unique id and password and will help in uploading the file of patients.

5.3.2 Test Module

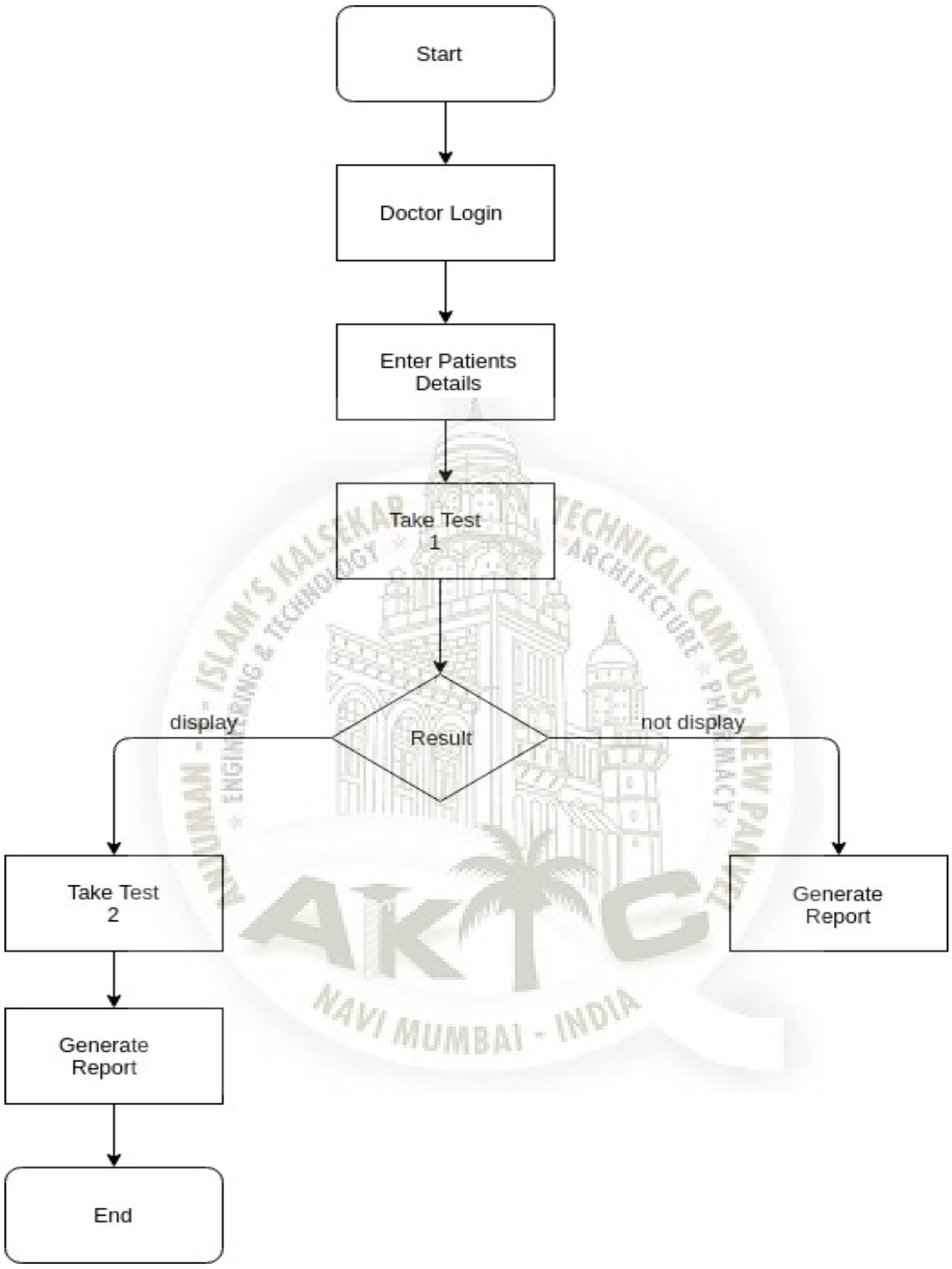


Figure 5.7: Test Module

In the Test Module , the Doctor will login first with the correct username and password.Then the doctor will make patient entry who has come for the check-up . Each patients entry will be unique.After that the Test 1 will appear which is a General Test after result an immediate report will get generate of the general test .If in Test 1 if any risk get predicted in the general report of the patients the doctor will

proceed to the Test 2 which is the important test for the patients, which will help to know the patients weather the patient is suffering from CRF.

5.3.3 File Upload Module

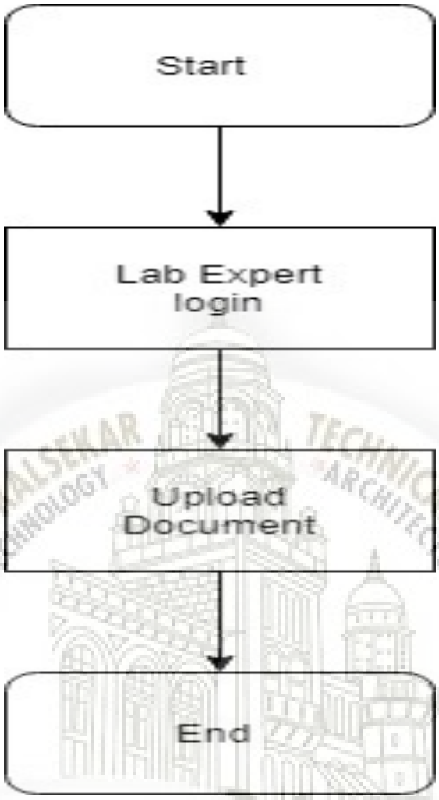


Figure 5.8: File Upload Module

In the File Upload Module, only the Lab Expert can login because the immediate report which will get generated after the test 1 and test 2 needs to get uploaded as per the patients details so to keep track of patients health.The upload fill of per patients will be done by the lab expert.

5.4 Systems Integration

5.4.1 Class Diagram

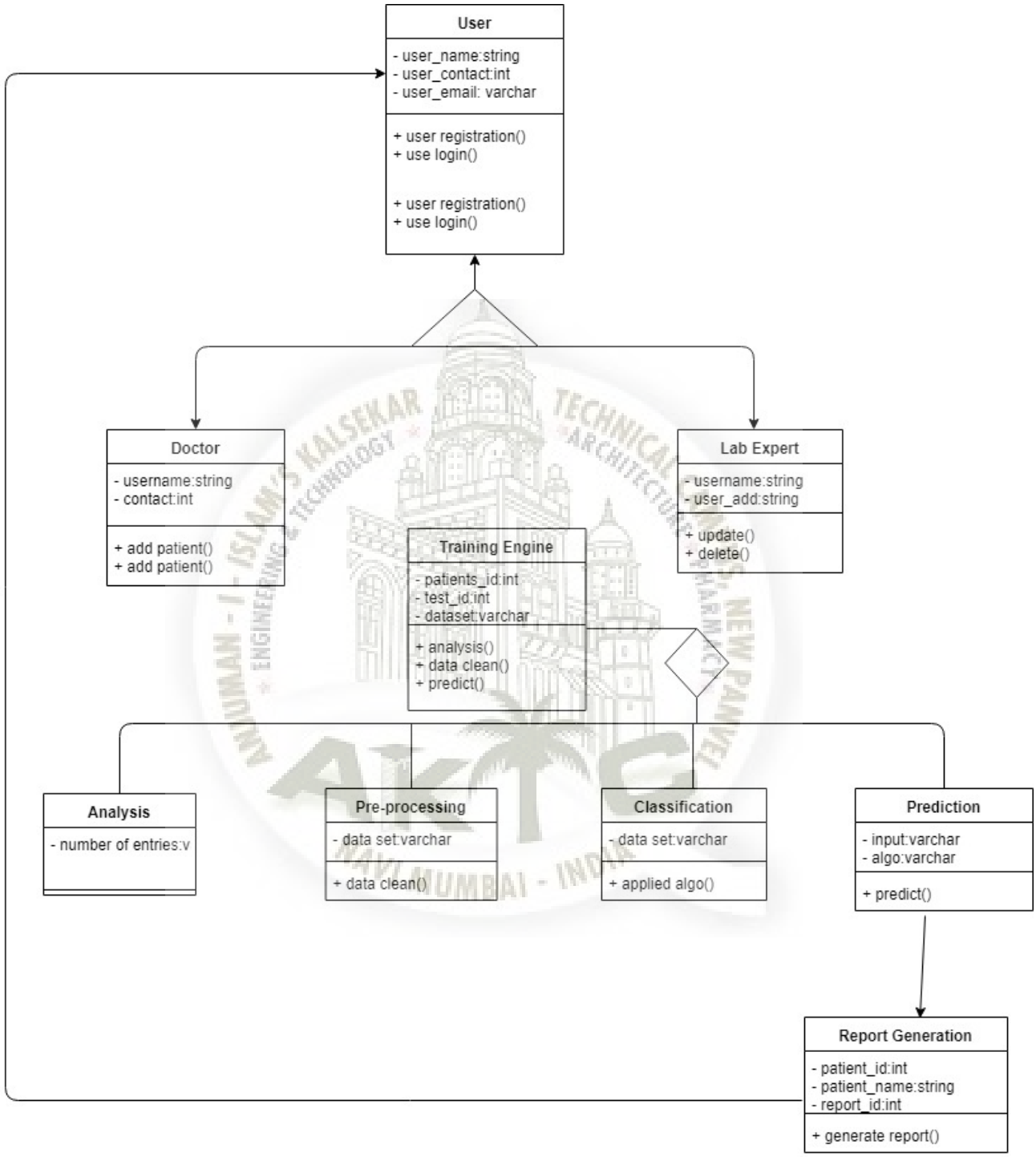


Figure 5.9: Class Diagram

In the Class Diagram , we have User,Doctor,Lab Expert and the most important part is the Training Engine.The Training Engine will help us to Analysis,Preprocess,Classification and Prediction.After the prediction the accurate report will be generated

5.4.2 Sequence Diagram

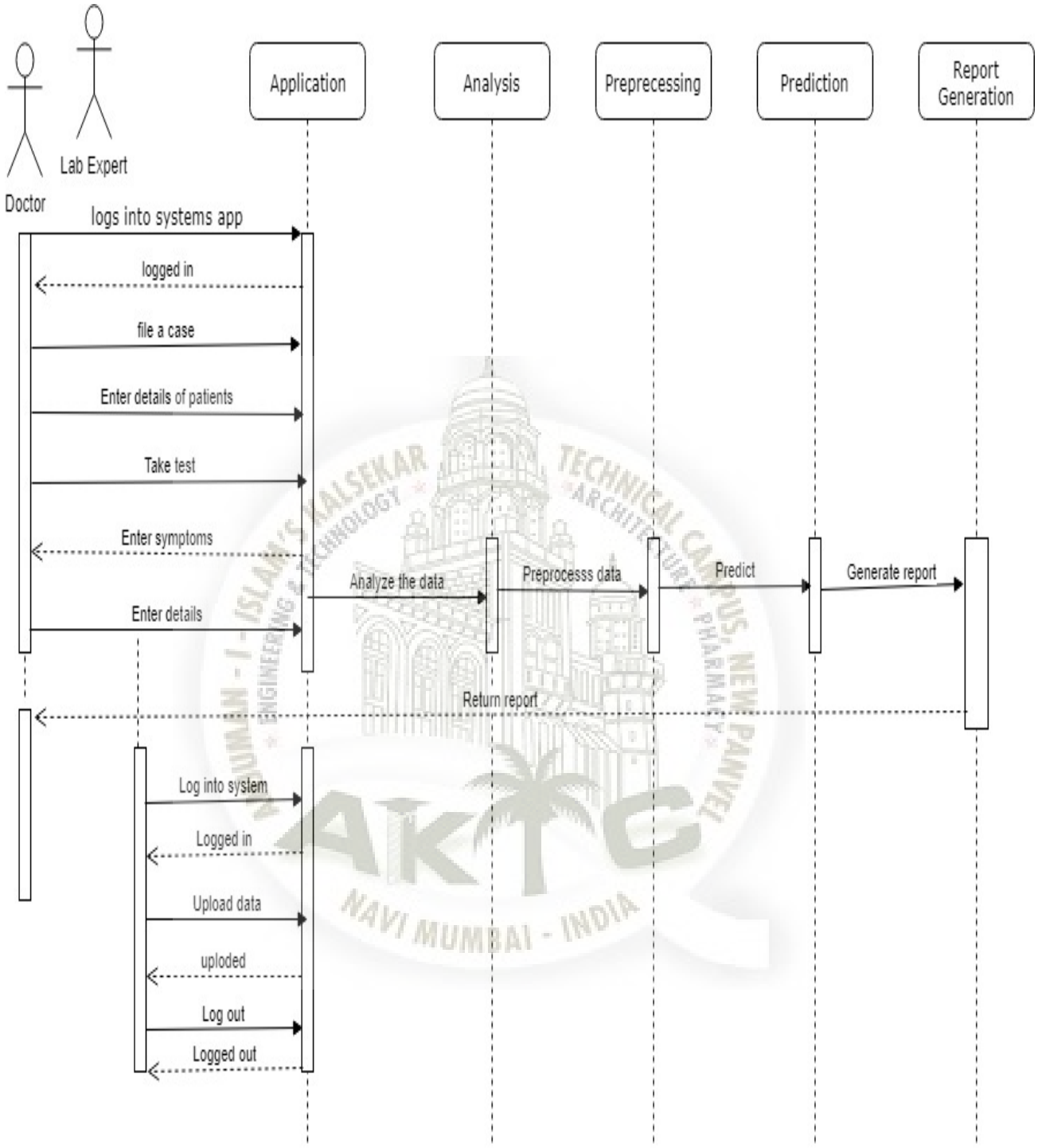


Figure 5.10: Sequence Diagram

A sequence diagram simply depicts interaction between objects in a sequential order i.e. the order in which these interactions take place. We can also use the terms event diagrams or event scenarios to refer to a sequence diagram. Sequence diagrams describe how and in what order the objects in a system function.

5.4.3 Component Diagram

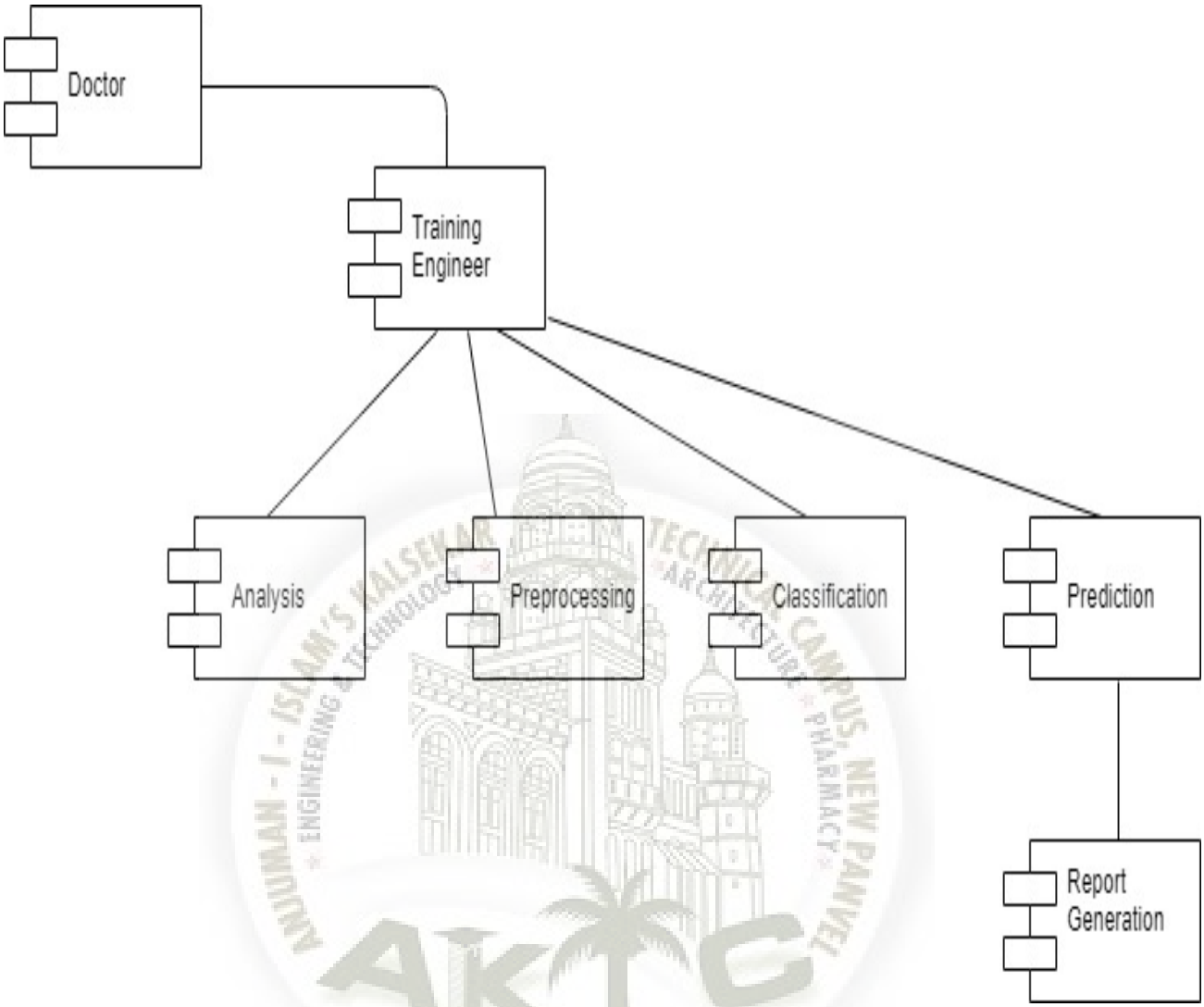


Figure 5.11: Component Diagram

Component diagram is a special kind of diagram in UML. The purpose is also different from all other diagrams discussed so far. It does not describe the functionality of the system but it describes the components used to make those functionalities.

5.4.4 Deployment Diagram

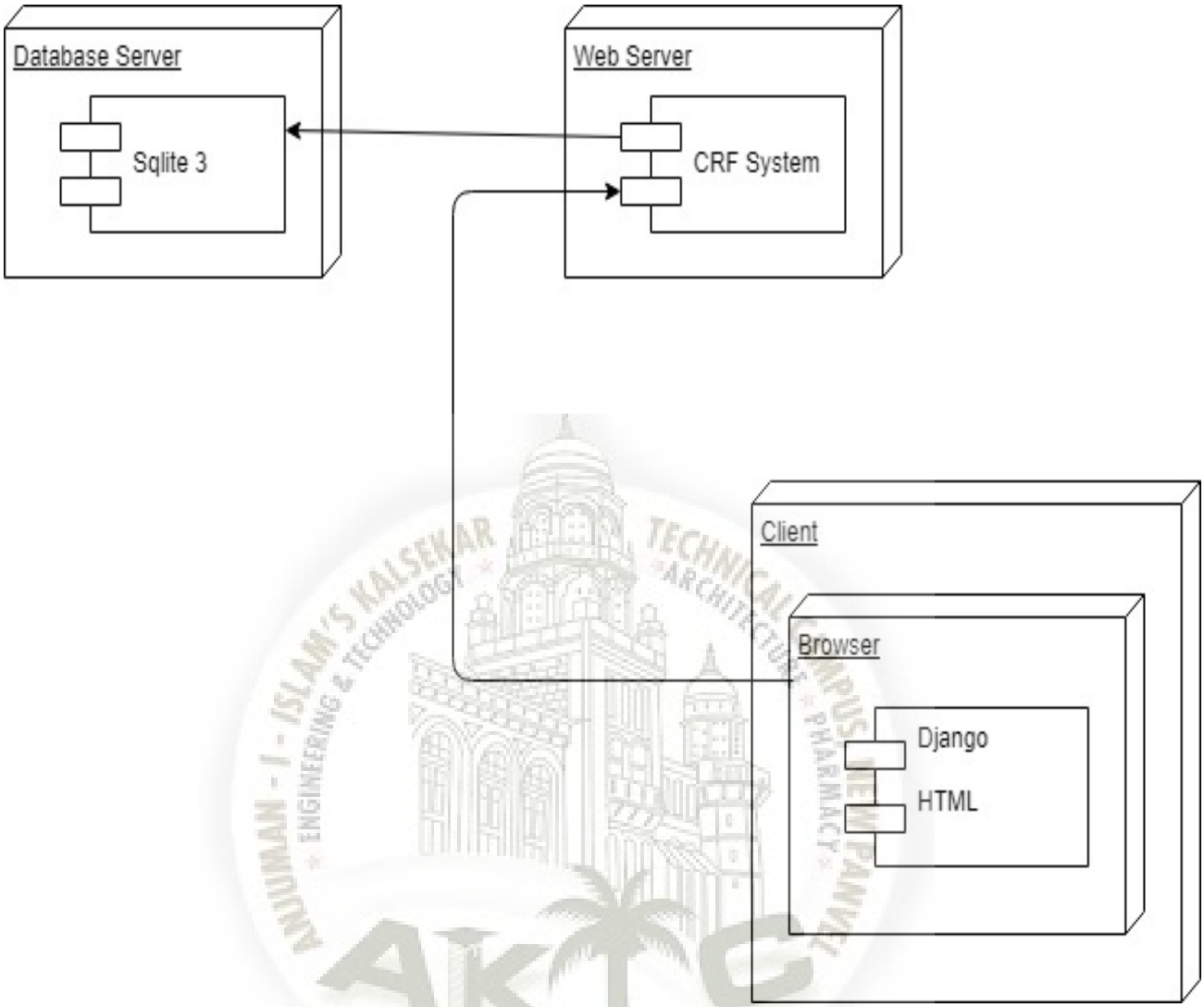


Figure 5.12: Deployment Diagram

A deployment diagram is a UML diagram type that shows the execution architecture of a system, including nodes such as hardware or software execution environments, and the middle ware connecting them. Deployment diagrams are typically used to visualize the physical hardware and software of a system.

Chapter 6

Implementation

6.1 Registration

Figure 6.1: Doctor Registration

```
1
2 {% extends "blog/base2.html" %}
3
4 {% block content %}
5 <div class="content-section">
6 <form method="POST">
7 {% csrf_token %}
8 <fieldset class="form-group">
```

```
{{ form.as_p }}
</fieldset>

<div class="form-group">
<button class="btn btn-outline-info" type="submit">Sign Up</button>
</div>
</form>
<div class="border-top pt-3">
<small class="text-muted">

Already have an account <a href="{% url 'login' %}">Sign In</a>
</small>
</div>
</div>
{% endblock content %}
```

6.2 Login

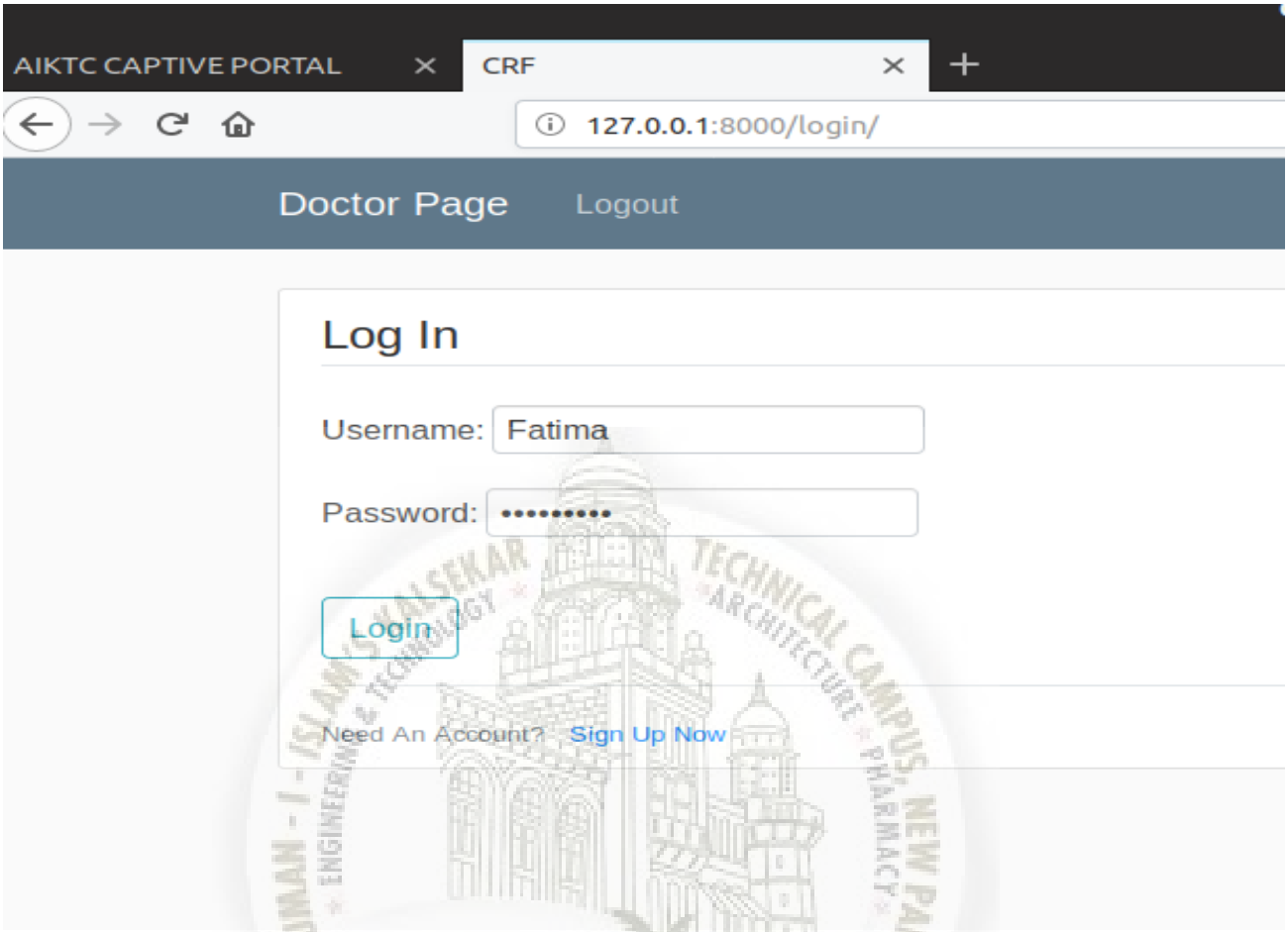


Figure 6.2: Login

```
{% extends "blog/base2.html" %}
{% block content %}
<div class="content-section">
<form method="POST">
{% csrf_token %}
<fieldset class="form-group">
<legend class="border-bottom mb-4">Log In</legend>
{{ form.as_p }}
</fieldset>
<div class="form-group">
<button class="btn btn-outline-info" type="submit">Login</button>
</div>
</form>
<div class="border-top pt-3">
<small class="text-muted">
Need An Account? <a class="ml-2" href="{% url 'register' %}">Sign Up Now</a>
</small>
</div>
</div>
{% endblock content %}
```

6.3 Patient Case File

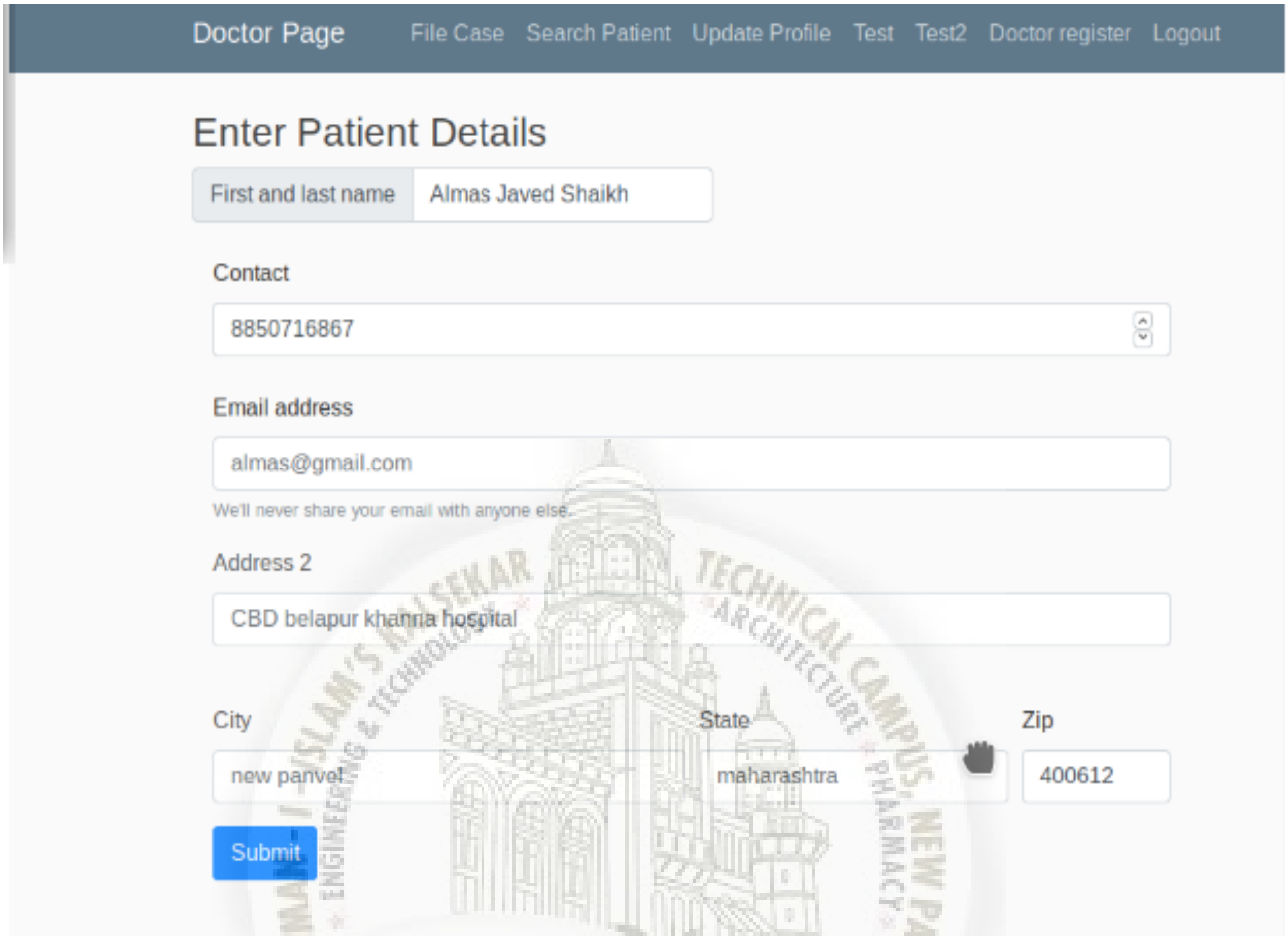


Figure 6.3: Patient Case File

```
{% extends "blog/base.html" %}
{% block content %}
<h3>Enter Patient Details</h3>

<main role="main" class="container">
<div class="row">
<div class="col-md-8">
</div>
<form method="POST" action="{% url 'user1' %}">
{% csrf_token %}
{{ form.as_p }}

<div class="input-group">
<div class="input-group-prepend">

<span class="input-group-text" id="">First and last name</span>
</div>
<input type="text" class="form-control" name="name">
</div>
<br>
</div>
<label for="inputContact">Contact</label>
<input type="number" class="form-control" id="inputContact" placeholder="
    1234567890" name="contact">
```

```
<br>
<div class="form-group">
<label for="exampleInputEmail1">Email address</label>
<input type="email" class="form-control" id="exampleInputEmail1" aria-
    describedby="emailHelp" placeholder="Enter email" name="email">
<small id="emailHelp" class="form-text text-muted">We'll never share your email
    with anyone else.</small>
</div>
<div class="form-group">
<label for="inputAddress2">Address 2</label>
<input type="text" class="form-control" id="inputAddress2" placeholder="
    Apartment, studio, or floor" name="address2">
</div>
<br>
<div class="form-row">
<div class="form-group col-md-6">
<label for="inputCity">City</label>
<input type="text" class="form-control" id="inputCity" placeholder="eg Mumbai"
    name="city">
</div>
<div class="form-group col-md-4">
<label for="inputState">State</label>
<input type="text" class="form-control" id="inputCity" placeholder="eg
    Maharashtra" name="state">
</div>
<div class="form-group col-md-2">
<label for="inputZip">Zip</label>
<input type="text" class="form-control" id="inputZip" placeholder="eg 400614"
    name="zip">
</div>
</div>
<button type="submit" class="btn btn-primary">Submit</button>
</form>

</div>
</div>
</main>

<!-- Optional JavaScript -->
<!-- jQuery first, then Popper.js, then Bootstrap JS -->

{% endblock content %}
```

6.4 Health Checkup-Test 1

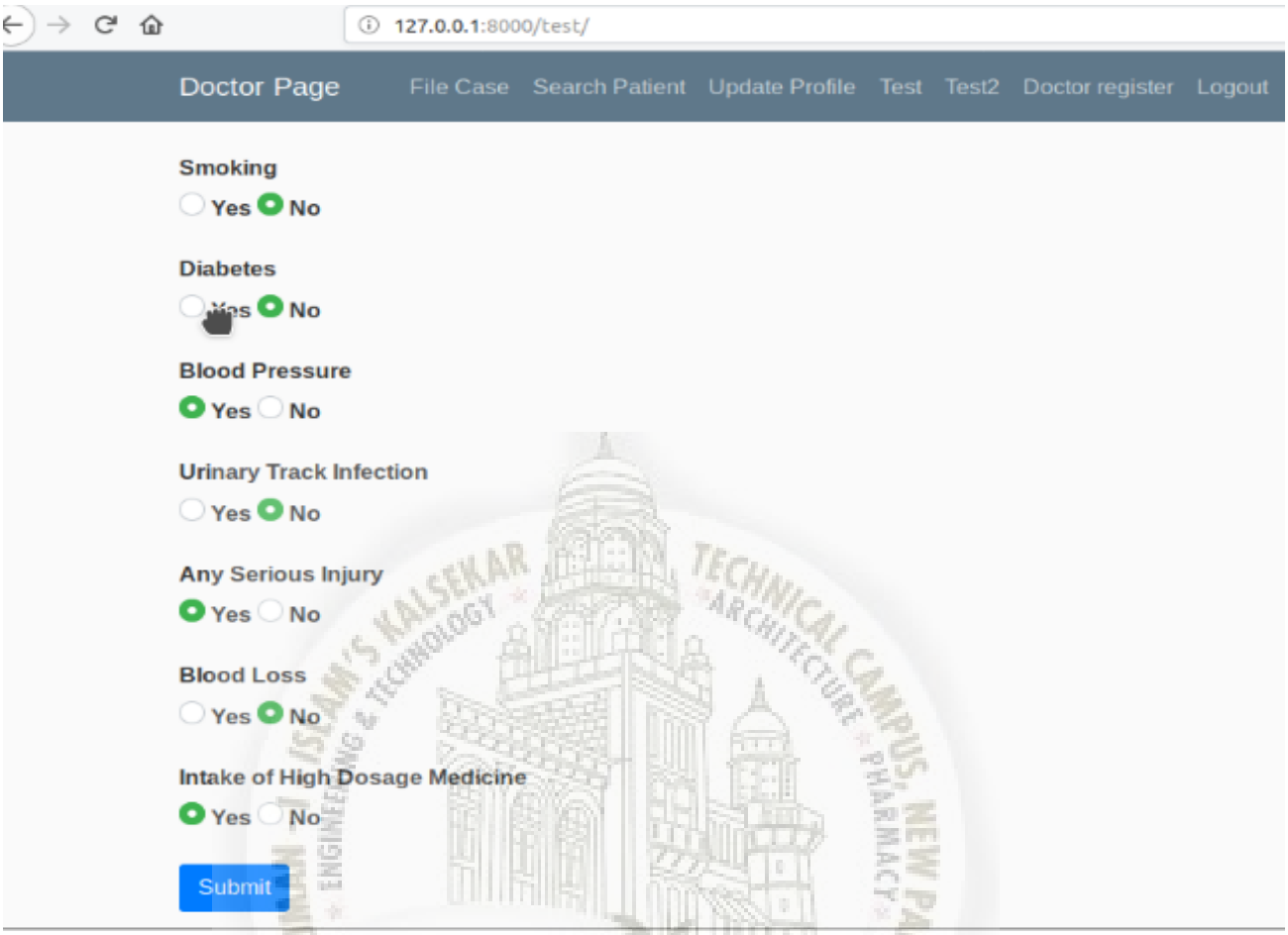


Figure 6.4: Test 1

```
  {% extends "blog/base.html" %}
{% block content %}

{% for x in xx:
print("Your Result",xx)
%}

{{x}}
{% endfor %}
                                <form method = "POST" action="{% url 'blog-test'
    %}">
{% csrf_token %}
<div class="form-group">
<b>

<label >Smoking</label>
<br>
<input type="radio" name="smoking" value="0,1"> Yes</t>
<input type="radio" name="smoking" value="1,0"> No<br>
<br>
<label >Diabetes</label>
<br>
<input type="radio" name="diabetes" value="0,1"> Yes
<input type="radio" name="diabetes" value="1,0"> No<br>
```

```
<br>
<label >Blood Pressure</label>
<br>
<input type="radio" name="bp" value="0,1"> Yes</t>
<input type="radio" name="bp" value="1,0"> No<br>
<br>
<label >Urinary Track Infection</label>
<br>
<input type="radio" name="uti" value="0,1"> Yes</t>
<input type="radio" name="uti" value="1,0"> No<br>
<br>
<label >Any Serious Injury</label>
<br>
<input type="radio" name="injury" value="0,1"> Yes</t>
<input type="radio" name="injury" value="1,0"> No<br>
<br>

<label >Blood Loss</label>
<br>
<input type="radio" name="bloodloss" value="0,1"> Yes
<input type="radio" name="bloodloss" value="1,0"> No<br>
<br>

<label >Intake of High Dosage Medicine</label>
<br>
<input type="radio" name="medicine" value="0,1"> Yes</t>
<input type="radio" name="medicine" value="1,0"> No<br>
<br>
</b>
<butt{% for x in xx:
print("Your Result",xx)
%}
{{x}}
{% endfor %}
on type="submit" class="btn btn-primary">Submit</button>
</div>

</form>

{% endblock content %}
```

6.5 Test 2

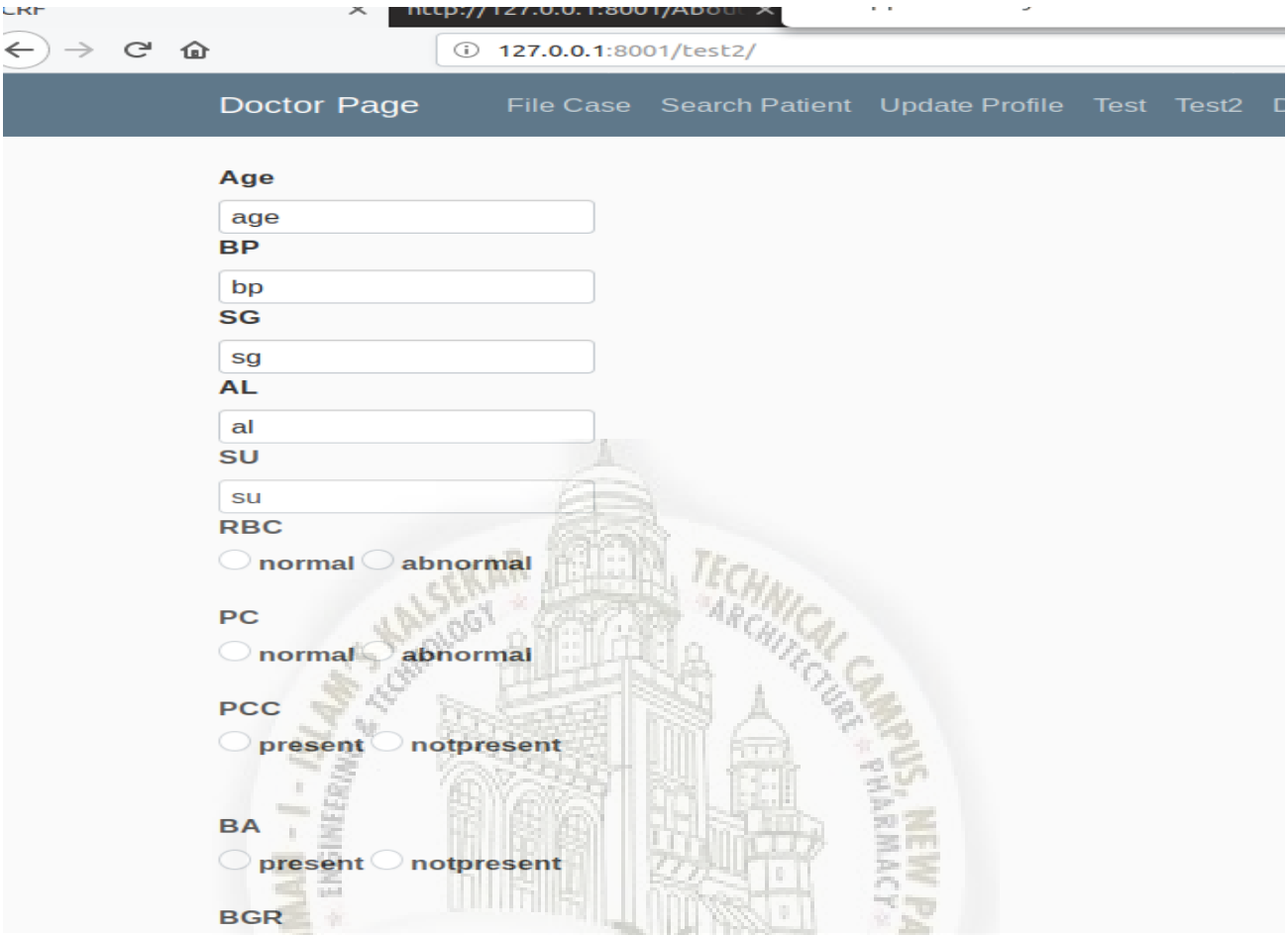


Figure 6.5: Test 2

```
  {% extends "blog/base.html" %}
{% block content %}

{% for x in xx %}
{{x}}
{% endfor %}
                                <form method = "POST" action="{% url 'blog-test2
    ' %}">
{% csrf_token %}
<div class="form-group">
<b>
<label >Age</label>
<br>
<input type="text" name="age" value="age">
<br>
<label >BP</label>
<br>
<input type="text" name="bp" value="bp">
<br>
<label >SG</label>
<br>
<input type="text" name="sg" value="sg">
<br>
<label >AL</label>
```

```
<br>
<input type="text" name="al" value="al">
<br>
<label >SU</label>
<br>
<input type="text" name="su" value="su">
<br>

<label >RBC</label>
<br>
<input type="radio" name="normal" value="normal"> normal
<input type="radio" name="abnormal" value="abnormal"> abnormal<br>
<br>
<label >PC</label>
<br>
<input type="radio" name="normal" value="normal"> normal</t>
<input type="radio" name="abnormal" value="abnormal"> abnormal<br>
<br>
<label >PCC</label>
<br>
<input type="radio" name="present" value="present"> present</t>
<input type="radio" name="notpresent" value="notpresent"> notpresent<br>
<br>

<br>
<label >BA</label>
<br>
<input type="radio" name="present" value="present"> present</t>
<input type="radio" name="notpresent" value="notpresent"> notpresent<br>
<br>
<b>
<label >BGR</label>
<br>
<input type="text" name="bgr" value="bgr">
<br>
<label >BU</label>
<br>
<input type="text" name="bu" value="bu">
<br>
<label >SC</label>
<br>
<input type="text" name="sc" value="sc">
<br>
<label >SOD</label>
<br>
<input type="text" name="sod" value="sod">
<br>
<label >POT</label>
<br>
<input type="text" name="pot" value="pot">
<br>
<label >HEMOGLOBIN</label>
<br>
<input type="text" name="hemo" value="hemo">
<br>
<label >PVC</label>
<br>
<input type="text" name="pvc" value="pvc">
<br>
<label >WBCC</label>
<br>
```

```
<input type="text" name="wbcc" value="wbcc">
<br>
<label >RBCC</label>
<br>
<input type="text" name="rbcc" value="rbcc">
<br>
<br>
<label >HTN</label>
<br>
<input type="radio" name="yes" value="yes"> yes</t>
<input type="radio" name="no" value="no"> no<br>
<br>
<br>
<label >DM</label>
<br>
<input type="radio" name="yes" value="yes"> yes</t>
<input type="radio" name="no" value="no"> no<br>
<br>
<br>
<label >CAD</label>
<br>
<input type="radio" name="yes" value="yes"> yes</t>
<input type="radio" name="no" value="no"> no<br>
<br>

<br>
<label >APPET</label>
<br>
<input type="radio" name="good" value="good"> good</t>
<input type="radio" name="poor" value="no"> no<br>
<br>

<br>
<label >PE</label>
<br>
<input type="radio" name="yes" value="yes"> yes</t>
<input type="radio" name="no" value="no"> no<br>
<br>

<br>
<label >ANA</label>
<br>
<input type="radio" name="yes" value="yes"> yes</t>
<input type="radio" name="no" value="no"> no<br>
<br>


</b>
<button type="submit" class="btn btn-primary">Submit</button>
</div>

</form>

{% endblock content %}
```

Chapter 7

System Testing

7.1 Test Cases and Test Results

Test ID	Test Case Title	Test Condition	System Behavior	Expected Result
T01	Login	should be register user	connect to mysql database	login to system
T02	Test	provide details of patient	prediction of CRF	provide report
T03	File Upload	data of patient	upload files in databases	File Uploade

7.2 Sample of a Test Case

Title: Login Page – Authenticate Successfully
Description: A registered user should be able to successfully login.

Precondition: the user must already be registered with an email address and password.

Assumption: a supported browser is being used.

Test Steps:

1. Navigate to web app.
2. In the 'email' field, enter the email of the registered user.
3. Click the 'Next' button.
4. Enter the password of the registered user
5. Click 'Sign In'

Expected Result: When the two test is performed the prediction should be done properly. So that the patients get to know about the CRF at the early stage and also provide the immediate report.

Actual Result: The Prediction should be done 100 percent correct . Provide the proper report to the patients.

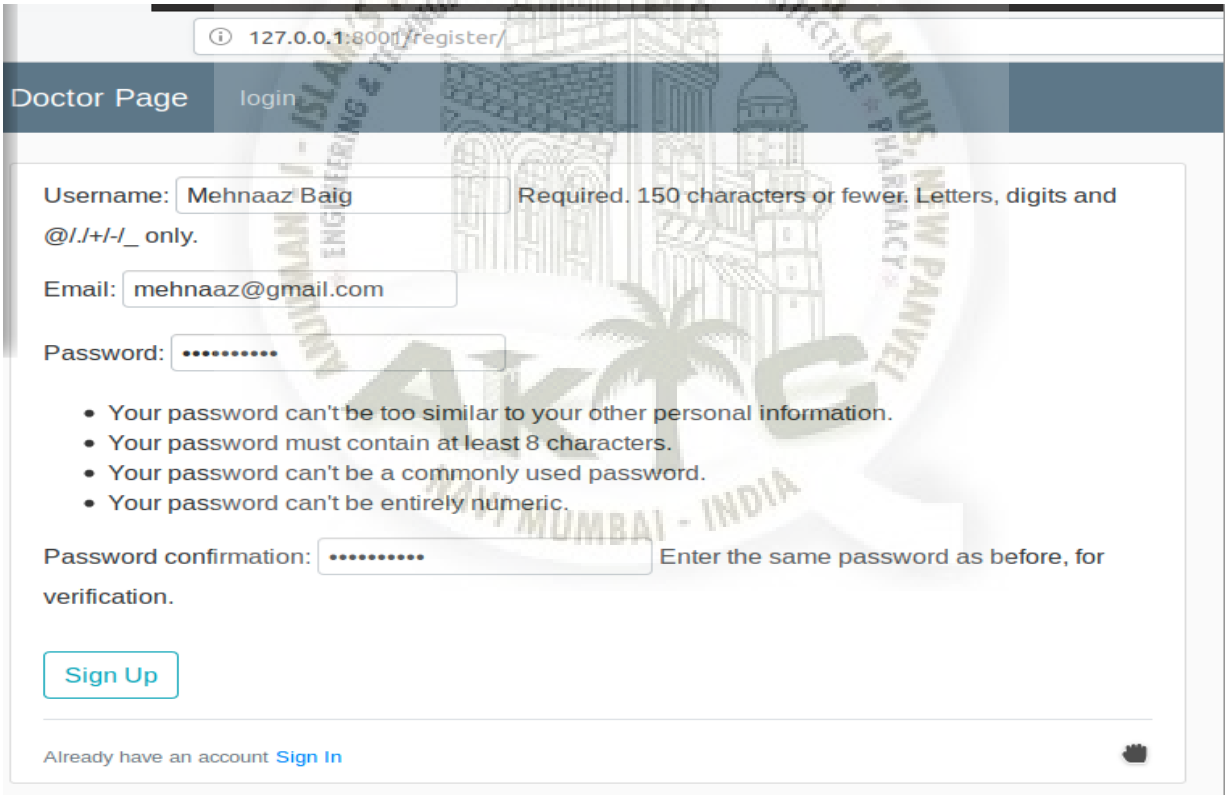
7.2.1 Software Quality Attributes

Specify any additional quality characteristics for the product that will be important to either the customers or the developers. Some to consider are: adaptability, availability, correctness, flexibility, interoperability, maintainability, portability, reliability, re usability, robustness, test-ability, and usability. Write these to be specific, quantitative, and verifiable when possible. At the least, clarify the relative preferences for various attributes, such as ease of use over ease of learning.

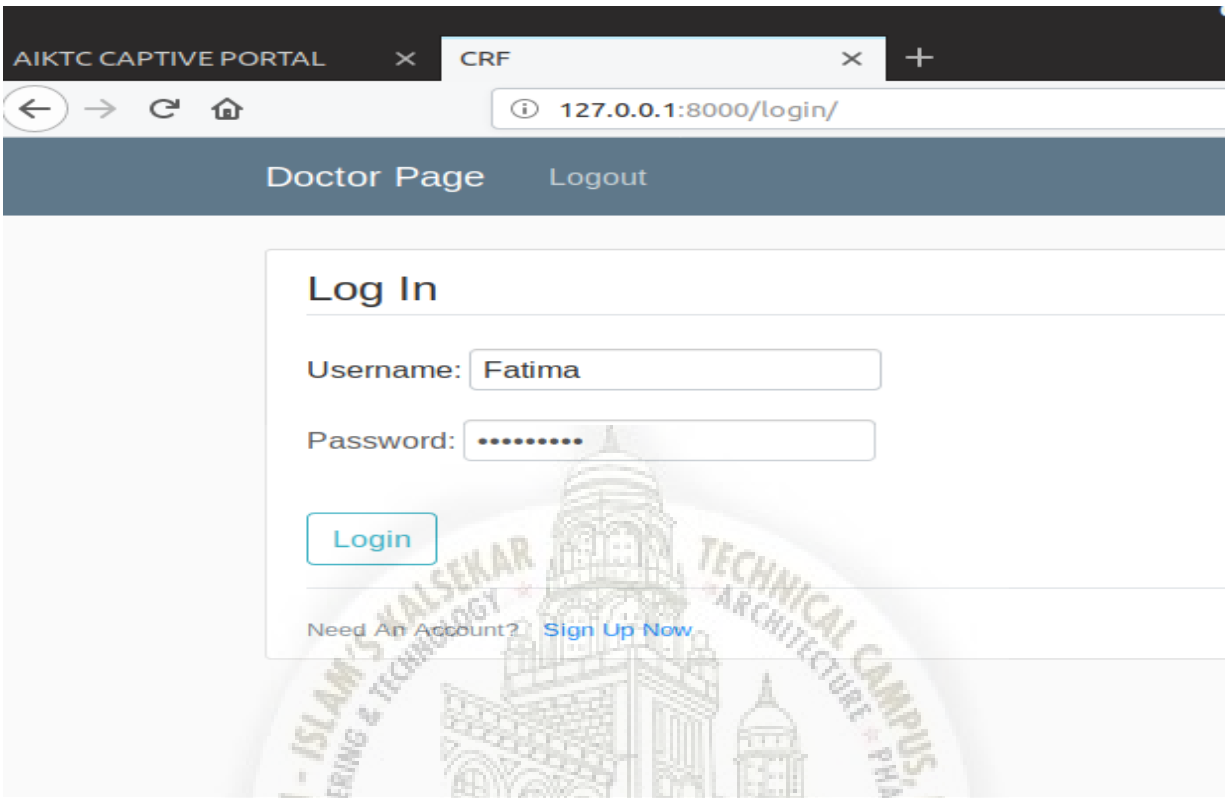
Chapter 8

Screenshots of Project

8.1 Registration

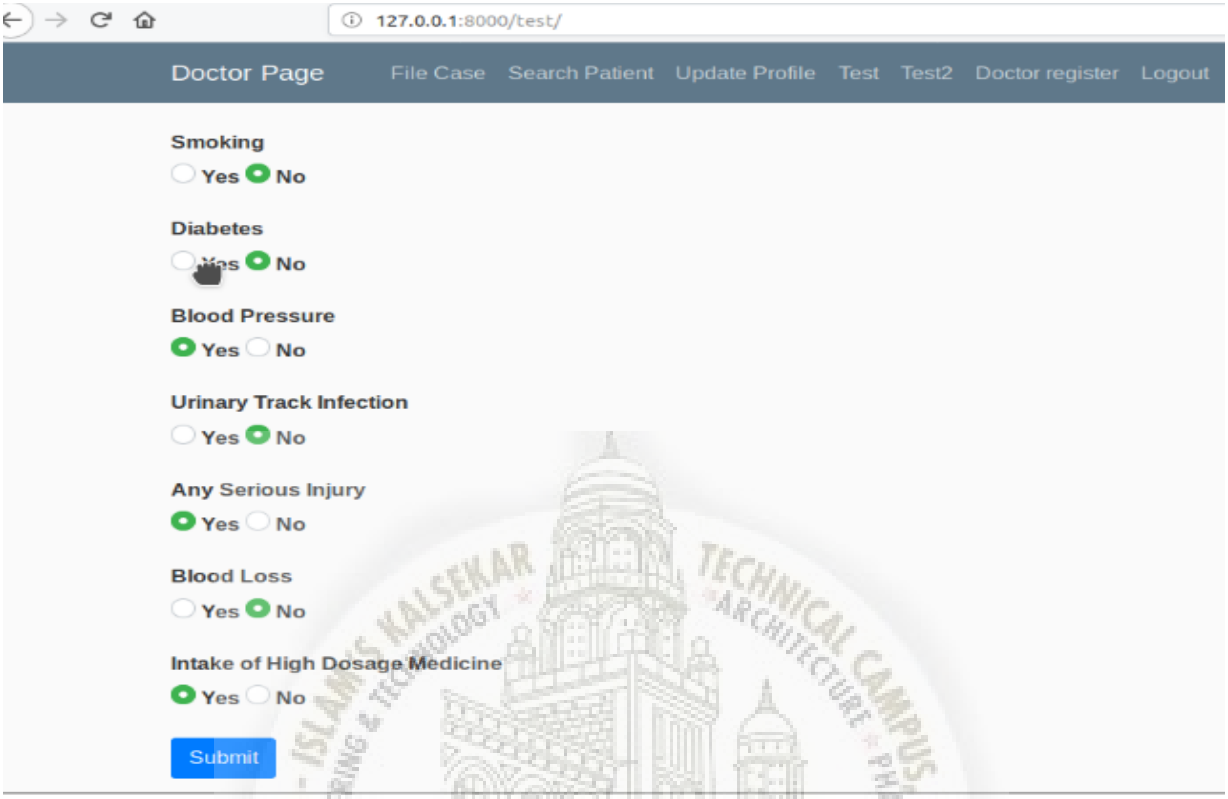


8.2 Login

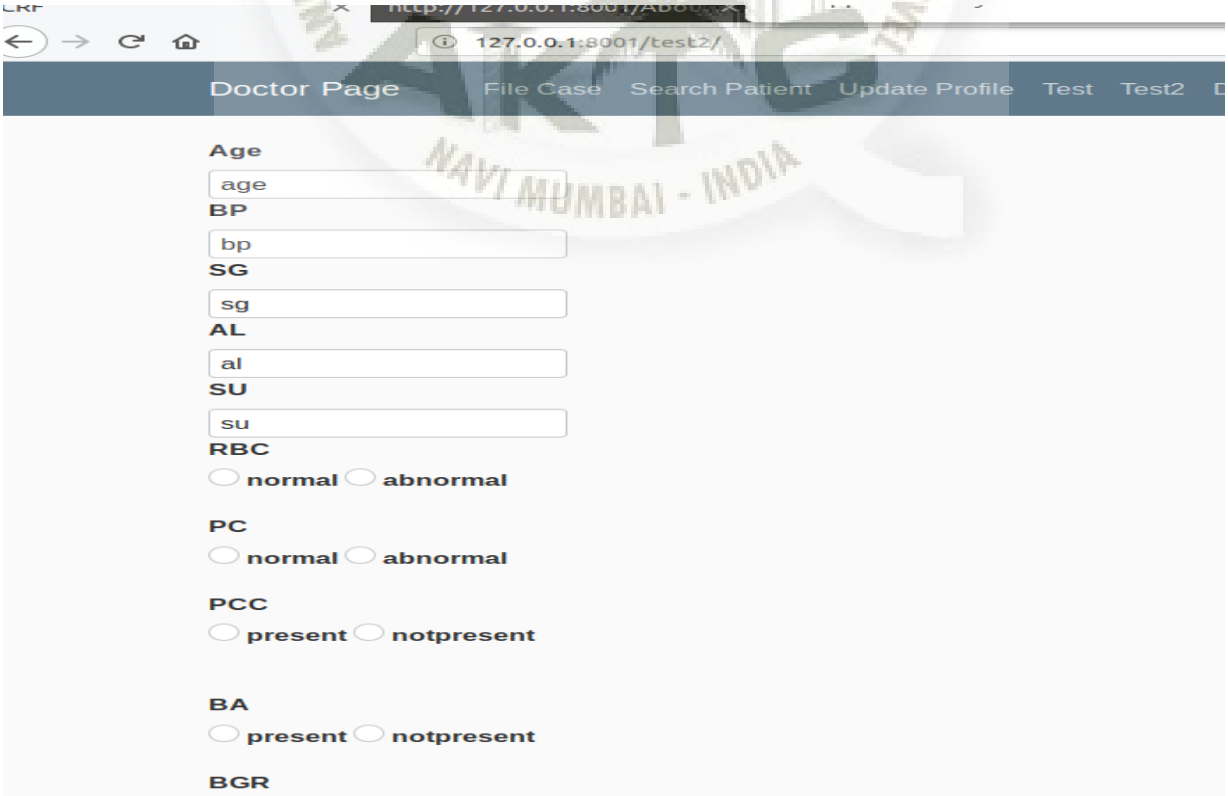


8.3 Case File

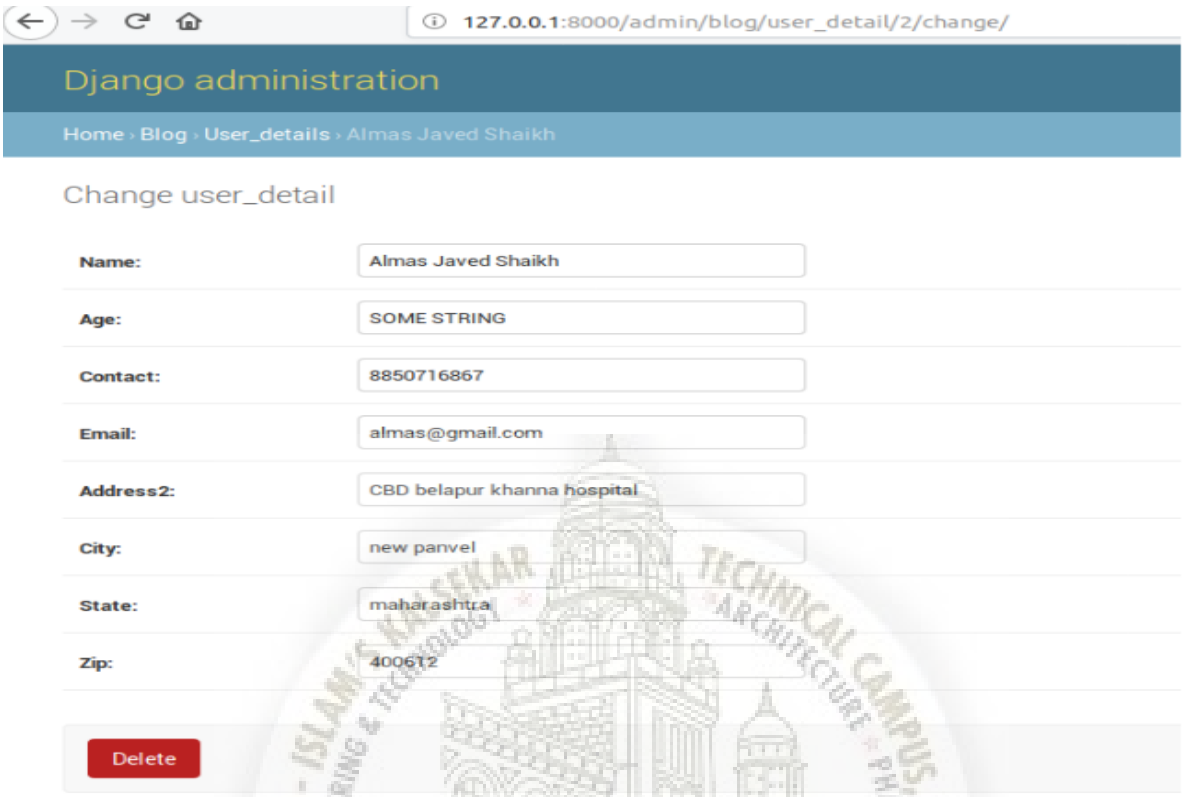
8.4 Test 1



8.5 Test 2



8.6 Database



Chapter 9

Conclusion and Future Scope

9.1 Conclusion

Chronic renal failure represents a critical period in the evolution of chronic renal disease and is associated with complications that begin early in the course of the disease. These conditions are initially sub clinical but progress relentlessly and may eventually become symptomatic and irreversible. Therefore, our system aims to predict this at an early stage. The system uses Naive Bayes algorithm to predict CRF, we use Naive Bayes algorithm as it gives highest accuracy. In this system we have some consider age,diabetes, blood pressure, RBC count etc. The efficiency of the system can be increase considering other parameters like food type,working environment, living conditions, availability of clean water, environmental factors etc for kidney disease prediction.

9.2 Future Scope

a . The Future Scope of our project is to predict the CRF stages.

b . Help the patients to take necessary treatment so that they can medicate as early as possible.

c . Provide a proper diet plan.

References

[1] *Predictive Analytics for Chronic Kidney Disease Using Machine Learning Techniques* ; Anusorn Charleonnan, Thipwan Fufaung, Tippawan Niyomwong, Wandee Chokchueypattanakit, Sathit Suwannawach, Nitat Ninchawee ,The 2016 Management and Innovation Technology International Conference (MITiCON-2016) , May 2016

[2] *Chronic kidney disease: a research and public health priority*; Norberto Perico and Giuseppe Remuzzi1 ,Advance Access publication 3 July 2012

[3] *CHRONIC KIDNEY DISEASE ANALYSIS USING DATA MINING CLASSIFICATION TECHNIQUES*; Veenita Kunwar,Khushboo Chandel,A.Sai Sabitha,Abhay Bansal,IEEE Conference, July 2016

Achievements

1. Publications
   (a) *Chronic Renal Failure Prediction System*; Almas Shaikh,Fatima Shaikh,Mehnaaz Baig,At AVALON,February 2019(Venue:Terna Engineering College) published(`http://EXAMPLE.com`)
2. Project Competitions
   (a) *Chronic Renal Failure Prediction System*; Almas Shaikh,Fatima Shaikh,Mehnaaz Baig,At AVALON,February 2019(Venue:Terna Engineering College)

TERNA PUBLIC CHARITABLE TRUST'S
TERNA ENGINEERING COLLEGE
AN ISO 9001 : 2000 | NBA ACCREDITATION

terna
Knowledge to Empower

CERTIFICATE OF PARTICIPATION

This is to certify that

Shaikh Amras Javed

of

A.I. K.T.C

has participated in

Avalon 2019, A National Level
(Technical Paper Presentation / Project Competition)
conducted on 5th & 6th March, 2019
at **Terna Engineering College, Nerul**

Prof. D.M. Bavkar
Avalon co-ordinator

AVALON

Dr. L.K. Ragha
Principal

TERNA PUBLIC CHARITABLE TRUST'S
TERNA ENGINEERING COLLEGE
AN ISO 9001 : 2000 | NBA ACCREDITATION

terna
Knowledge to Empower

CERTIFICATE OF PARTICIPATION

This is to certify that

Shaikh Fatima Moin

of

A.I.K.T.C

has participated in

Avalon 2019, A National Level
(Technical Paper Presentation / Project Competition)
conducted on 5th & 6th March, 2019
at **Terna Engineering College, Nerul**

Prof. D.M. Bavkar
Avalon co-ordinator

TERNA ENGINEERING COLLEGE, NERUL, NAVI MUMBAI
AVALON

Dr. L.K. Ragha
Principal

TERNA PUBLIC CHARITABLE TRUST'S
TERNA ENGINEERING COLLEGE
AN ISO 9001 : 2000 | NBA ACCREDITATION

terna
Knowledge to Empower

CERTIFICATE OF PARTICIPATION

This is to certify that

Baig Mehnaz

of

A.I.K.T.C

has participated in

Avalon 2019, A National Level
(Technical Paper Presentation / Project Competition)
conducted on 5th & 6th March, 2019
at **Terna Engineering College, Nerul**

Prof. D.M. Bavkar
Avalon co-ordinator

ENGINEERING COLLEGE, NERUL, NAVI MUMBAI
AVALON

Dr. L.K. Ragha
Principal

TERNA PUBLIC CHARITABLE TRUST'S
TERNA ENGINEERING COLLEGE
AN ISO 9001 : 2000 | NBA ACCREDITATION

terna
Knowledge to Empower

CERTIFICATE OF PARTICIPATION

This is to certify that

Shaikh Amras Javed

of

A.I. K.T.C

has participated in

Avalon 2019, A National Level
(Technical Paper Presentation / Project Competition)
conducted on 5th & 6th March, 2019
at **Terna Engineering College, Nerul**

Prof. D.M. Bavkar
Avalon co-ordinator

AVALON

Dr. L.K. Ragha
Principal

TERNA PUBLIC CHARITABLE TRUST'S
TERNA ENGINEERING COLLEGE
AN ISO 9001 : 2000 | NBA ACCREDITATION

terna
Knowledge to Empower

CERTIFICATE OF PARTICIPATION

This is to certify that

Shaikh Fatima Moin

of

A.I.K.T.C

has participated in

Avalon 2019, A National Level
(Technical Paper Presentation / Project Competition)
conducted on 5th & 6th March, 2019
at **Terna Engineering College, Nerul**

Prof. D.M. Bavkar
Avalon co-ordinator

TERNA ENGINEERING COLLEGE, NERUL, NAVI MUMBAI
AVALON

Dr. L.K. Ragha
Principal

TERNA PUBLIC CHARITABLE TRUST'S
TERNA ENGINEERING COLLEGE
AN ISO 9001 : 2000 | NBA ACCREDITATION

terna
Knowledge to Empower

CERTIFICATE OF PARTICIPATION

This is to certify that

Baig Mehnaz

of

A.I.K.T.C

has participated in

Avalon 2019, A National Level
(Technical Paper Presentation / Project Competition)
conducted on 5th & 6th March, 2019
at **Terna Engineering College, Nerul**

AVALON

Prof. D.M. Bavkar
Avalon co-ordinator

Dr. L.K. Ragha
Principal

Chronic Renal Failure Prediction System

Almas Shaikh, Mehnaaz Baig, Mehvash Shaikh, Fatima Shaikh

Computer Department, Mumbai University
A I Kalsekar Technical Campus, New Panvel
Email : almasshaikh9813@gmail.com, baigmehnaaz@gmail.com, shaikhmehvash1998@gmail.com, manalshaikh116@gmail.com
Phone: 8425907018, 9930378372, 8850716867.

Abstract— The growing global burden of Non-communicable disease(NCD's) worldwide, is increasing day by day. The chronic renal failure is a type of NCD. These chronic disease are one of the leading cause of death. It becomes important for our society, to detect and cure it as early as possible. Our system aims to predict the possibility of chronic renal failure, i.e the chances of kidney failure of a patient. Huge amount of patient's data and their case histories are stored from years and years, and yet not being used. This data can be used to predit, using massive data sets, by categorizing valid and unique patterns in data. We aim to make a system through which people can regularly check the risk to have chronic renal failure.

Keywords— CRF(Chronic Renal Failure), NCD(Non Communicable Disease), Massive Datasets, Prediction Algorithms, Machine Learning.

I. INTRODUCTION

Currently, kidney disease is a major problem and is an area of concern. It is a condition where kidney don't work as well as normal. CRF does not usually cause any symptoms until it has reached an advance stage. Kidney disease is very dangerous if not immediately treated on time, and may be fatal. If the doctors have a good tool that can identify patients who are likely to have kidney disease in advance, they can heal the patients in time. So, we are developing a system that can predict kidney damage at early stage. Our focus is on predicting CRF disease using algorithm. Large amount of data is being generated by medical industry that medical history is being unused. So it is required to process those unused data. It comes up with the set of techniques which when applied to this processed data, generates report for making appropriate decision and to make people aware about that the kidney failure can be predicted at early stages.

In this paper, we present predictive models using machine learning techniques that can predict chronic kidney disease including Decision tree, Naive Bayes and Support Vector Machine Models.

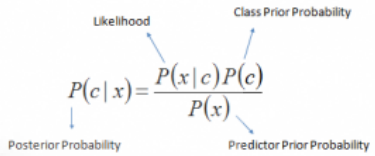
In part 2, we present the details of these models. Part 3 will present the proposed method. In part 4, the results of experiments are described.

II. Machine Learning Techniques

A.Naive Bayes

It is a classification technique based on Bayes Theorem with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. For example, a fruit may be considered to be an apple if it is red, round, and about 3 inches in diameter. Even if these features depend on each other or upon the existence of the other features, all of these properties independently contribute to the probability that this fruit is an apple and that is why it is known as 'Naive'.

Bayes theorem provides a way of calculating posterior probability P(c|x) from P(c), P(x) and P(x|c). Look at the equation below:



$$P(c \mid X) = P(x_1 \mid c) \times P(x_2 \mid c) \times \cdots \times P(x_n \mid c) \times P(c)$$

Above,

- $P(c|x)$ is the posterior probability of *class* (c, *target*) given *predictor* (x, *attributes*).
- $P(c)$ is the prior probability of *class*.
- $P(x|c)$ is the likelihood which is the probability of *predictor* given *class*.
- $P(x)$ is the prior probability of *predictor*.

The detailed of the Naive Bayes is described in algorithm.

B.Support Vector Machine

For the classification problem, the support vector machine (SVM) is the popular data mining method used to predict the category of data. The main idea of SVM is to find the optimal hyperplane between data of two classes in the training data. SVM finds the hyperplane by solving optimization problem:

$$\max Q(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j d_i d_j \mathbf{x}_i^T \mathbf{x}_j$$
$$\text{where } 0 \le \alpha_i \le C \text{ for } i = 1, 2, \ldots, n$$

SVM uses the decision function f(x) defined in form of the kernel function for calculating the output as

$$f(x) = \text{sign}\left[\sum_{i=1}^{l} \alpha_i d_i K(x, x_i) + b\right]$$

where $K(x, x_i)$ is the kernel function.

C. Decision Tree

Decision tree is the classification method frequently used in data mining task. A decision tree is a structure that includes a root node, branches, and leaf nodes. It divides the data into classes based on the attribute value found in sample.

Algorithm

Step 1: Convert the data set into a frequency table.

Step 2: Create Likelihood table by finding the probabilities.

Step 3: Now, use Naive Bayesian equation to calculate the posterior probability for each class. The class with the highest posterior probability is the outcome of prediction.

Naive Bayes uses a similar method to predict the probability of different class based on various attributes. This algorithm is mostly used in text classification and with problems having multiple classes.

III. The Proposed Method

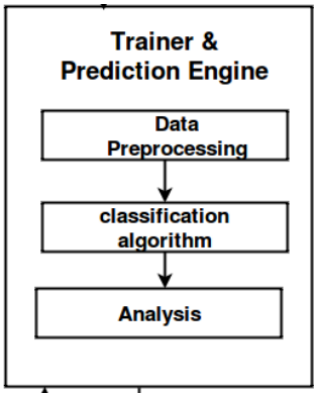
A. Working Of System

Our system consists of six working modules which includes login, update, filing case, training and prediction engine, verification and lastly the report generation. We have two actors in our system Doctor and the lab expert. Both will get logged in to the system through the login page. Doctor will file a case and enter all patient details which will get stored in the database. After the patient's details are entered, doctor will take test and all the data will go to the training and prediction engine, where prediction is done. After the predication, verification will be done and finally report will be generated.

B. Working Of Agents

- Lab Expert
  Lab Expert is an important actor in our system. He / She is going to upload the reports of patients suffering from chronic renal failure on a regular basis. This is important, as the increase in amount of data in the data set, more accurate prediction will be done.
- Doctor
  Doctor is going to use our system for prediction of chronic renal failure. The doctor will file a case,take all the necessary inputs from patient like patient details, symptoms and previous history. This data will get stored in our system which will be used for prediction.

C. Training And Prediction Engine



- Data

Prepossessing

Data Prepossessing involves transforming raw data into an understandable format. Real-world data is often incomplete, inconsistent, and/or lacking in certain behaviors or trends, and is likely to contain many errors. Data prepossessing is a proven method of resolving such issues. It is use for finding unknown patterns and information in health data.

- Classification

Naive Bayes is the probabilistic classifier based on Bayes theorem. It assume variables are independent of each other. This algorithm is easy to build and works well with huge amount of data. The numerical attributes includes age, blood pressure , blood urea, etc. On the basis of this attribute, its evaluation of classification is done.

- Analysis

On the basis of the symptoms and all the data provided by the patient, a test will be done which will analyze the data to predict CRF.

IV. EXPERIMETAL RESULTS

In this experiment, the machine learning methods described in section 3 are trained to predict the chronic kidney disease. Three classifier methods are used in this experiment consisting of Support Vector Machine (SVM) with Naive Bayes and Decision tree.

A. Dataset

The Indian's Chronic Kidney Disease (CKD) dataset consistsof linear and non linear data. The Attribute of this dataset consists of two types of attributes which are numeric and nominal attributes. This dataset is collected from the Apollo Hospital. The dataset is divided into two groups, one for training and another for testing. The ratio of training and testing data is 70% and 30% respectively.

B. Performance Measurement

In this paper, the performance of the proposed method is measured by accuracy described as follows.

• Accuracy (ACC) is the overall success rate of the classifier defined as

ACC = (TP+TN) / (P + N) (6)

where TP is the true value of positive rate, P is the positive class or yes class, N is the negative class or no class.

C. Results

The results of machine learning techniques are compared in the experiments. All machine learning techniques are trained and tested by the proposed method. Finally, we will get a report and the advisory from the system.

V. CONCLUSIONS

Chronic renal failure represents a critical period in the evolution of chronic renal disease. These conditions are initially subclinical but progress relentlessly.It may becomes irreversible. Therefore, our system aims to predict this at an early stage. Chronic Renal Failure has been predicted and diagnosed using Naive Bayes algorithm. Naive Bayes algorithm gives the highest accuracy as compared to decision tree.

ACKNOWLEDGMENT

The authors would like to express their gratitude to Apollo Hospital for providing the dataset and facilitating anonymised access to it. We would thank then for making this system possible.

REFERENCES

[1] Chronic renal disease - Early identification and active management of patients with renal impairment in primary care can improve outcomes by Malvinder S Parmar.

[2] Chronic Kidney Disease Analysis using Data Mining Classification Techniques by ASET

[3] Predictive analytic chronic kidney disease using machine learning technique

[4] Performance Analysis of Naive Bayes and J48 Classification Algorithm for Data Classification

[5] Scaling up the accuracy of Naive Bayes Classifier:A Decision-Tree Hybrid by R Kohavi.

[6] Naive Bayes: http://www.statsoft.com/textbook/naive-bayes-classifier

[7] https://ieeexplore.ieee.org/document/7919649

[8] https://www.analyticsvidhya.com/blog/2017/09/understaing-support-vector-machine-example-code/