PNT2022TMID36002

Early Detection Of Chronic Kidney Disease Using Machine Learning

IBM-Project-2573-1658475170

NALAIYA THIRAN PROJECT BASED LEARNING ON PROFESSIONAL READLINESS FOR INNOVATION, EMPLOYMENT AND ENTERPRENEURSHIP

PROJECT REPORT

Submitted By

POOJA BALASUBRAMANIAN	2019506061
SANJAY KANNAN M	2019506081
GNANAVARSHINI M	2019506031
NITHYAKAMAL ILAMURUGU	2019506059
SANDHYA S	2019506080

TEAM ID **:** PNT2022TMID36002

INDUSTRY MENTOR : Ms. Lalitha Gayathri

FACULTY MENTOR **:** Ms. Eliza Femi Sherley S

EVALUATOR **:** Dr. P. Kola Sujata

MADRAS INSTITUTE OF TECHNOLOGY, ANNA UNIVERSITY

CHENNAI - 600044

PNT2022TMID36002

PROJECT REPORT

CHAPTER - 1
INTRODUCTION

1.1 PROJECT OVERVIEW

- CKD is one of the growing concerns of the current generation with many people following an unhealthy lifestyle. This project focuses on identifying Chronic Kidney Disease at an earlier stage, by observing various parameters like BP, sugar, rbc. Identifying CKD proves crucial because early identification can reduce the costs significantly, and is trivial which otherwise can be fatal.
- Firstly, we analyzed the data using various methods like univariate,multivariate analysis .Then the data was pre-processed like handling missing values, label encoding was performed before proceeding forward.
- The processed data was then further fed into machine-learning models like KNN, random forest, logistic regression, where the data was split with the ratio of 70 for training and 30 for testing.
- Then a front-end was developed using HTML, CSS and the trained model was converted into a pickle file and using flask, the model was deployed, where the values can be taken in and the prediction will be done based on the values and the result will be displayed as a web page.

1.2 PURPOSE

- Chronic Kidney Disease prediction is one of the most important issues in healthcare analytics. The most interesting and challenging tasks in day-to-day life is prediction in medical field. 10% of the population worldwide is affected by chronic kidney disease (CKD), and millions die each year because they do not have access to affordable treatment. Chronic kidney Disease can be cured, if treated in the early stages.

- The main aim of this project is to predict whether the patient have chronic kidney disease or not, in more accurate and faster way based on certain diagnostic measurements like Blood Pressure (Bp), Albumin(Al).

- Early detection of kidney disease can help in treatment which could save lives. Analyzing various medical tests, would give us an idea about which attributes help us distinguish the disease.

CHAPTER - 2
LITERATURE SURVEY

2.1 EXISTING PROBLEM

Chronic Kidney Disease is a global health problem with high morbidity and mortality rate and which can increases the probability of being affected by a cardiovascular disease .Chronic Kidney Disease can grow worse over time, if left untreated and unidentified so it becomes important to identify the disease at the early stages, the main complication with identifying CKD is that is difficult to identify it in the early stages, where it can be important to treat the person, so machine learning models are used to identify CKD economically and accurately.

2.2 REFERENCES

1. Qin, Jiongming, et al. "A machine learning methodology for diagnosing chronic kidney disease." *IEEE Access* 8 (2019): 20991-21002.

2. Almansour, Njoud Abdullah, et al. "Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study." *Computers in biology and medicine* 109 (2019): 101-111.

3. Vásquez-Morales, Gabriel R., et al. "Explainable prediction of chronic renal disease in the colombian population using neural networks and case-based reasoning." *Ieee Access* 7 (2019): 152900-152910.

4. Tekale, Siddheshwar, et al. "Prediction of chronic kidney disease using machine learning algorithm." International Journal of Advanced Research in Computer and Communication Engineering 7.10 (2018): 92-96.

5. Chimwayi, Kerina Blessmore, et al. "Risk level prediction of chronic kidney disease using neuro-fuzzy and hierarchical clustering algorithm (s)." (2017).

6. Dulhare, Uma N., and Mohammad Ayesha. "Extraction of action rules for chronic kidney disease using Naïve bayes classifier." *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*. IEEE, 2016

7. Rubini, L. Jerlin, and P. Eswaran. "Generating comparative analysis of early Stage prediction of Chronic Kidney Disease." *International Journal of Modern Engineering Research (IJMER)* 5.7 (2015): 49-55.

8. Rubini, L. Jerlin, and P. Eswaran. "Generating comparative analysis of early stage prediction of Chronic Kidney Disease." *International Journal of Modern Engineering Research (IJMER)* 5.7 (2015): 49-55.

9. Dubey, Abhinandan. "A classification of ckd cases using multivariate kmeans clustering." *International Journal of Scientific and Research Publications* 5.8 (2015):1-5.

10. Neves, José, et al. "A soft computing approach to kidney diseases evaluation." *Journal of medical systems* 39.10 (2015): 1-9.

2.3 PROBLEM STATEMENT DEFINITION

Many people of the current age follow a very unhealthy lifestyle, which is one of the main causes of many diseases, like cancer, diabetes, heart-attacks etc. One of the most prominent diseases of modern times is Chronic kidney disease (CKD), which is caused due to a sedentary lifestyle, unhealthy eating habits, and excessive smoking. CKD is not just restricted to the elderly people but many of the younger generations also develop CKD which is very concerning. CKD brings with it a number of other problems like increased chances of heart-stroke, weakened immune system, shortness of breath etc. If CKD is left untreated for a long time or is not discovered early can lead to death of the person. Hence machine learning models KNN, Logistic Regression are used to help in early detection of CKD which could prove trivial in the treatment of CKD.

CHAPTER - 3
IDEATION AND PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS

- An empathy map is a collaborative tool teams can use to gain a deeper insight into their customers. Much like a user persona, an empathy map can represent a group of users, such as a customer segment.

- The empathy map was originally created by Dave Gray and has gained much popularity within the agile community.

- An empathy map is an effective visualization template that helps analyse the behaviour and emotions of customers and users.

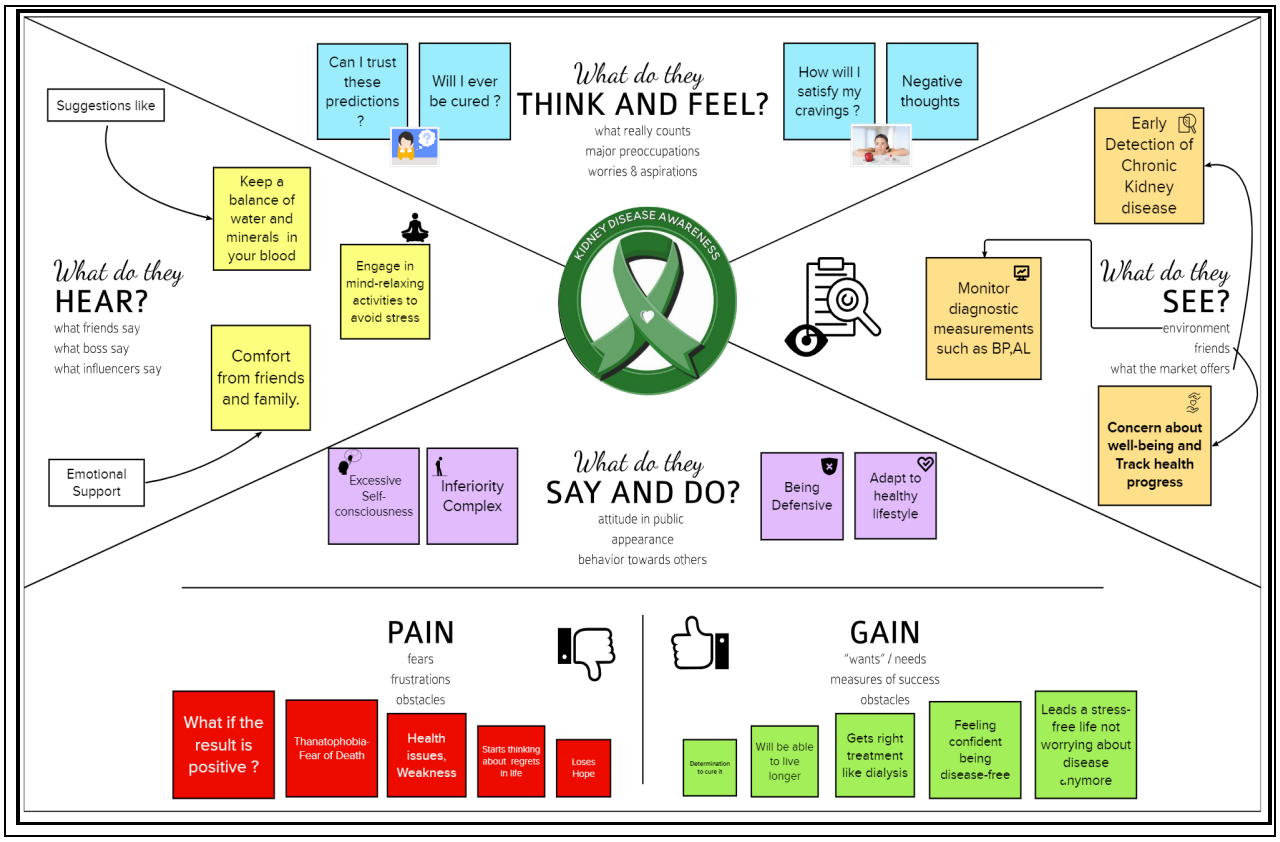
- Empathy maps not only detect the behaviours but highlight possible mediums for brands to communicate with their customers in a better way.

- Empathy maps can also be used to collect data directly from the users. Used alongside user interviews, survey answers, etc., you can also have a user fill in an empathy map themselves. This often reveals aspects of the user that may have remained unsaid or not thought of.

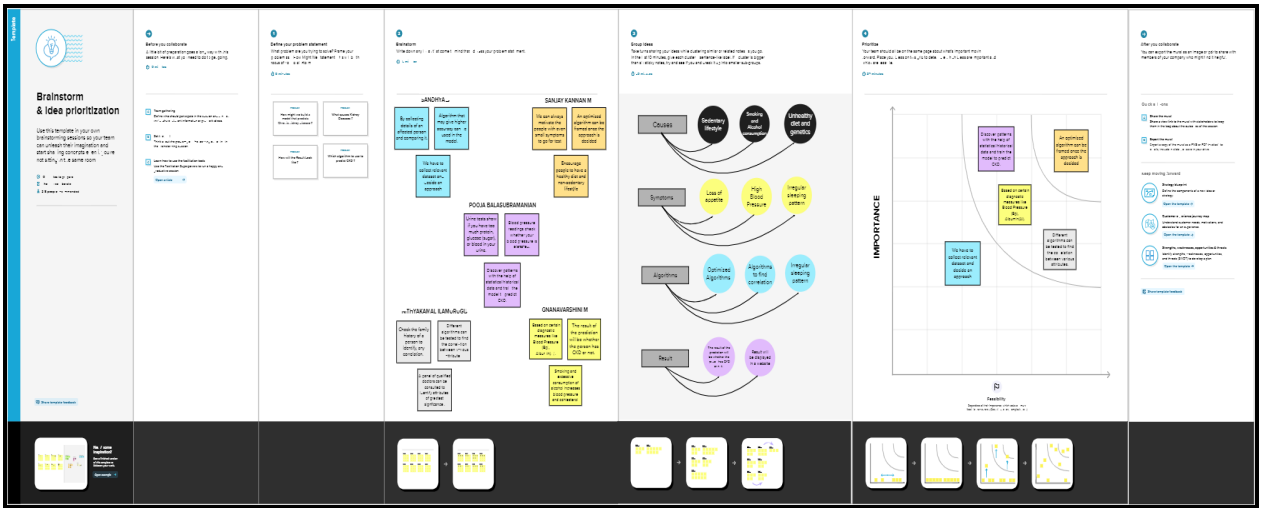
- Each of the four quadrants and additionally two pain and gain sections comprise a category that helps us delve into the mind of the user.

- The four empathy map quadrants look at what the user says, thinks, feels, and does. The pain and gain section are used to know the different positive and negative aspects of the project.

PNT2022TMID36002



3.2 IDEATION AND BRAINSTORMING



3.3 PROPOSED SOLUTION

S.NO.	PARAMETER	DESCRIPTION
1.	Problem Statement (Problem to be solved)	Chronic kidney disease (CKD) is a condition in which the kidneys are damaged and cannot filter blood as well as they should.It has beenof growing concern, as the kidney is one of the most important organs in the body required for filtering blood. 10% of the population worldwide is affected by CKD, and millions die each year because they do not have access to affordable treatment. Thus, it is important to be able to predict CKD using various machine learning techniques.
2.	Idea / Solution description	Various diagnostic measurements like Blood Pressure (Bp), Albumin (Al) etc., of the patients are collected and the data is processed and given to a ML model that will predict if the patienthas CKD or not. Among the various ML models that are present, the one that will give higher accuracy will be chosen to get better results.
3.	Novelty / Uniqueness	We aim to find the best machine learning model for the early prediction of CKD by analyzing the essential parameters and comparing their predictive accuracies. Then collaborate the best ML model to an interactive UI which helps in the early detection of CKD and provide cure.

4.	Social Impact / Customer Satisfaction	The main aim of this application is early prediction and proper treatments can possibly stop or slow the progression of this disease to the end stage.
5.	Business Model (Revenue Model)	We can generate revenue through direct customers or can also collaborate with the health care sector and generate revenue from their customers.
6.	Scalability of the Solution	We can build various models using machine learning algorithms and compare them to find the best accurate model. We can also use image data and apply Deep Learning techniques, Probabilistic Neural Networks (PNN), and Multilayer Perceptron (MLP) etc., which will provide an improved accuracy than the machine learning techniques.

PNT2022TMID36002

3.4 PROBLEM-SOLUTION FIT

Project Title: Early Detection of Chronic Kidney Disease using Machine Learning **Project Design Phase-I** - Solution Fit Template **Team ID:** PNT2022TMID36002

Define CS, fit into CC

1. CUSTOMER SEGMENT(S) CS

Margret is a 55-year-old diabetic patient, and has been observing many unusual symptoms recently like metallic-taste, high blood-pressure. She also has a family history of chronic kidney disease and wants to know if she is hereditarily affected by CKD.

6. CUSTOMER CONSTRAINTS CC

People are always skeptical about being diagnosed fearing it might cost them a fortune to get treated, as it would involve a lot of tests and medical procedures to cure the disease, so people consider better not to be diagnosed. People with less awareness get diagnosed with CKD only when the symptoms get worse and becomes more difficult to treat and is more deadly

5. AVAILABLE SOLUTIONS AS

Margret can consult a doctor, stating that she has been facing some unusual symptoms recently, following the prescription she could take tests to discover if she has CKD or not.

Explore AS, differentiate

Focus on J&P, tap into BE, understand RC

2. JOBS-TO-BE-DONE / PROBLEMS J&P

CKD can be fatal to health and a person has to change their lifestyle and can face unprecedented issues like

- Anemia
- Increased chances of stroke
- Decreased immune response
- Loss of appetite
- Depression
- Retention of fluids

9. PROBLEM ROOT CAUSE RC

Common causes of CKD include diabetes, high blood pressure, obesity etc. Though Marget is already diabetic, she followed a routine of unhealthy lifestyle like smoking, sedentary life, and having a sweet tooth. All these habits, could have played a major role, given that she has a family history of CKD.

7. BEHAVIOUR BE

Margret has been trying to change her habits and include healthier foods, and develop a healthier lifestyle like going for workouts etc. She has also been going to a rehabilitation center to get rid of her smoking habits.

Focus on J&P, tap into BE, understand RC

Identify strong TR & EM

3. TRIGGERS TR

Margret has noticed her change in her appearance which has caused her to develop insecurities. She has also lost her appetite, and developed insomnia.

4. EMOTIONS: BEFORE / AFTER EM

BEFORE:
Marget was incompetent to decide if she has been affected by the disease and this caused her to feel distressed and uneasy.

AFTER:
Marget, after she has been identified has a better understanding of what the disease is and feels more secure about the decisions, she should take to face CKD.

10. YOUR SOLUTION SL

The proposed solution is to identify the chronic kidney disease using machine learning techniques, in its earlier stages to facilitate timely treatment which would reduce the cost of the treatment significantly and also the fatality. In reality, it would be feasible as only a prescribed number of tests would be taken and based on it the disease can be predicted.

8. CHANNELS of BEHAVIOUR CH

8.1 ONLINE
Marget browsed online about the unusual symptoms and developed her suspicions for CKD and tried to alleviate her ailments using home treatments.

8.2 OFFLINE
Margret attends awareness programs to find out more about CKD and tries to conclude if she is affected, she later approaches a doctor to confirm her suspicions of CKD.

Identify strong TR & EM

CHAPTER - 4
REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENT

FR NO.	FUNCTIONAL REQUIREMENT (EPIC)	SUB REQUIREMENT (STORY / SUB-TASK)
FR-1	Home Page	• Symptoms and cure for chronic kidney disease can be displayed. • Test vitals required for prediction of chronic kidney disease. • Sign Up and Login options for the user accordingly. If new user the user needs to SIGN UP and gets directed to the registration page else, redirected to the LOGIN page.
FR-2	User Registration	The user needs to enter few account credentials and other personal details required for registering.
FR-3	User Confirmation	The user must confirm the details given in the registration page.
FR-4	User Login	The user needs to enter the login credentials such as name, username, and password.

FR-5	User Authentication	Validate user credentials through password.
FR-6	Symptoms Form	The user must enter their symptoms and answer the questions properly for prediction.
FR-7	Test Result	• The page displays the test report. • If positive - displays the test report along the necessary measures to be taken to cure the disease. • If negative – displays the test report along with the preventive measures for the disease.

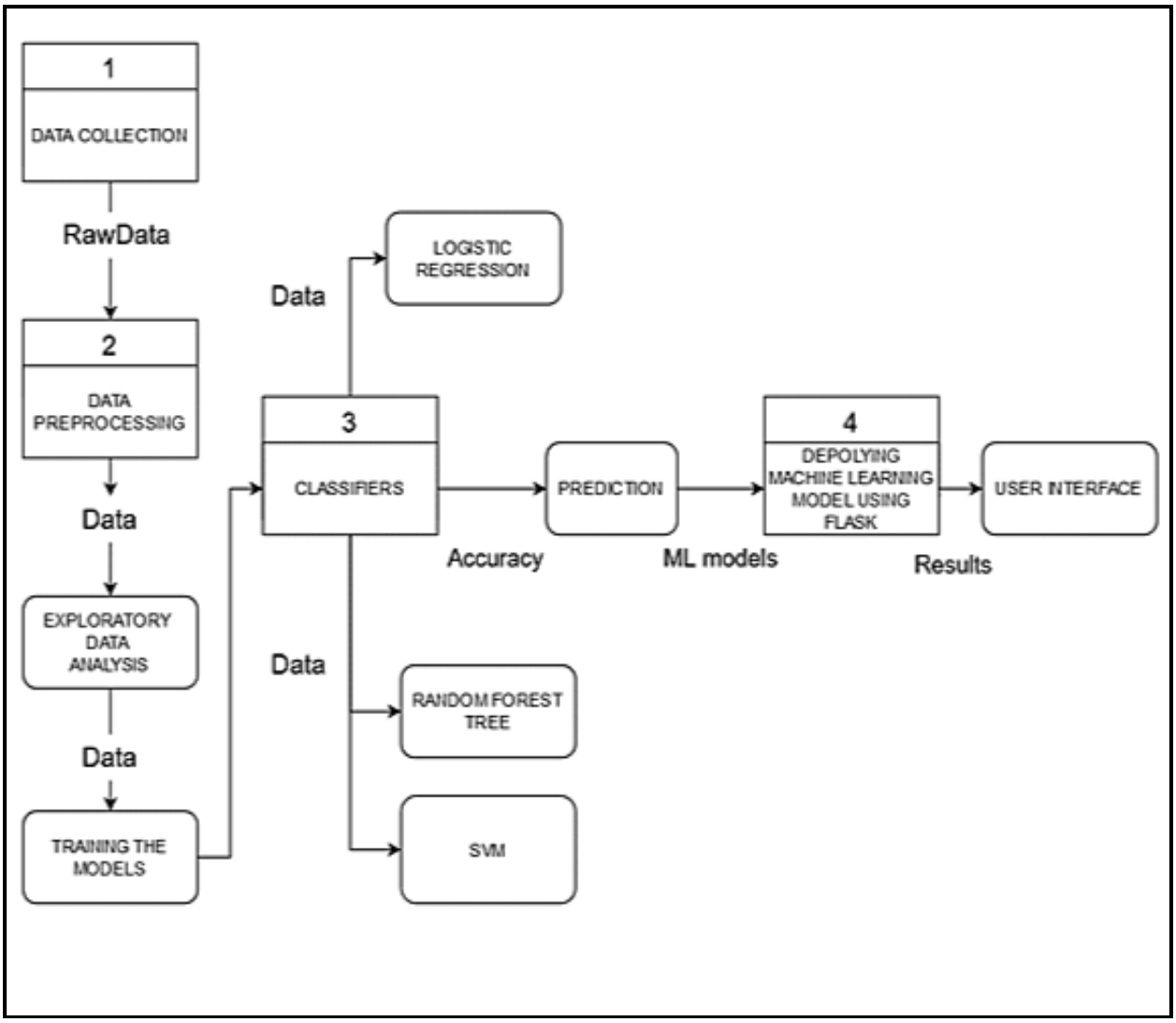
4.2 NON-FUNCTIONAL REQUIREMENT

FR NO.	**NON-FUNCTIONAL REQUIREMENT**	**DESCRIPTION**
NFR-1	Usability	Simple, interactive and user friendly interface of the website for easy navigation.
NFR-2	Security	User preferences and the predictions need to be confidentially maintained.
NFR-3	Reliability	Should be portable and cross-platform independent.

NFR-4	Performance	Traffic should be efficiently managed and the accuracy of the results should be good.
NFR-5	Availability	The application should be compatible on the device that the user chooses to use.
NFR-6	Scalability	The application should be developer friendly and there must be scope for more advancement in the application if the need be.

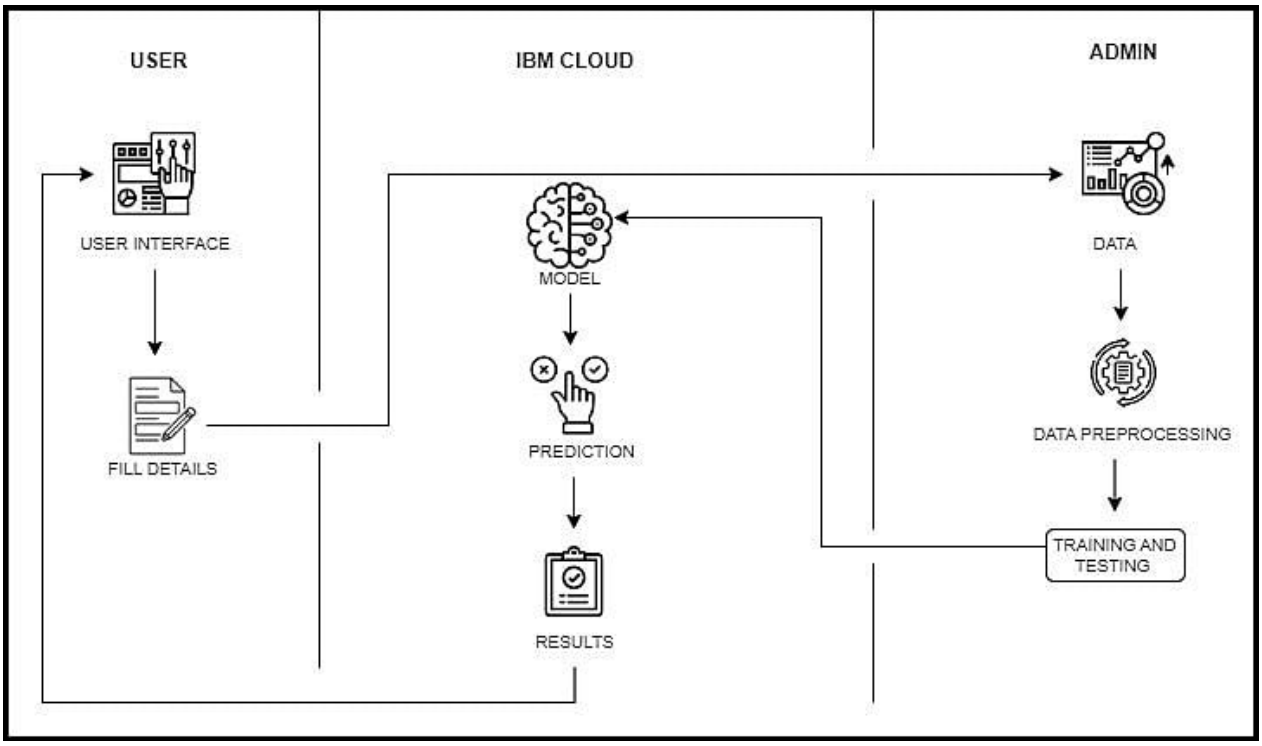
CHAPTER - 5
PROJECT DESIGN

5.1 DATA FLOW DIAGRAMS



5.2 SOLUTION AND TECHNICAL ARCHITECTURE

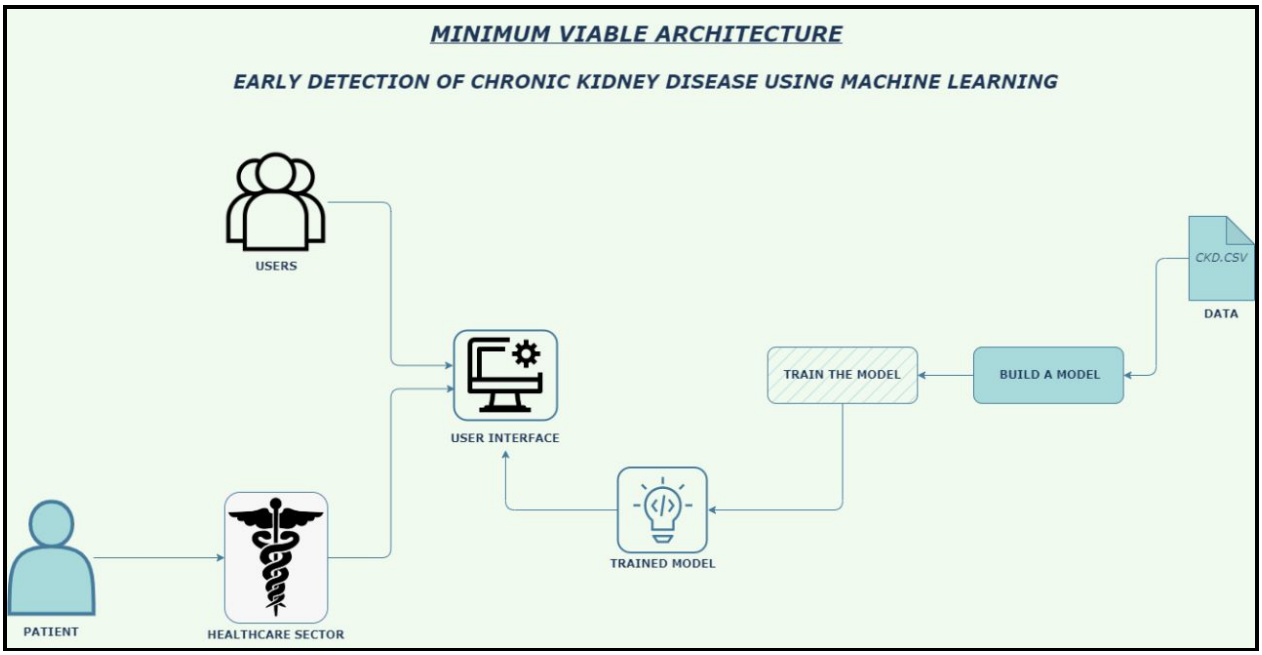
TECHNICAL ARCHITECTURE



GUIDELINES:

- All the processes are listed under the Application logic block.
- The trained model will be uploaded in the IBM cloud.
- API will get the data from the cloud when requested.
- Input from the user will be taken through the web page and will be processed in the cloud.
- Accuracy will be shown according to the inputs.

SOLUTION ARCHITECTURE



5.3 USER STORIES

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Data Collection Team	Collecting data	USN-1	As a user, the data collection team has to collect enough data to train the models.	X-Ray images can be collected.	High	Sprint-1
		USN-2	The data collected is	The data is cleaned and	High	Sprint-1

			cleaned and preprocessed.	converted into csv format.		
Model Training Team	Training the models	USN-3	Different classification models are trained	Models which have high accuracy are accepted.	High	Sprint-2
		USN-4	Using the trained models, the prediction is done.	Using the accepted model, the prediction is performed.	High	Sprint-2
Web Development Team	Deploying trained models	USN-5	The trained models are deployed using flask framework.	The trained models are deployed without any malfunctions.	Medium	Sprint-3
Customer (Web user)	Web-Pages	USN-6	User can visit the web page and detect if they have kidney disease.		Low	Sprint-4

CHAPTER - 6
PROJECT PLANNING AND SCHEDULING

6.1 SPRINT PLANNING AND ESTIMATION

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Data Collection	Task-1	To build the machine learning model, we begin with the process of downloading the dataset and then perform data analysis	4	Low	Sandhya S
Sprint-1	Data Analysis	Task-2	We import the required libraries and then perform data analysis on the given dataset.	3	Medium	Nithyakamal Ilamurugu
Sprint-1	Data Pre-processing	Task-3	Data cleaning, handling missing values and performing label encoding.	8	Medium	Gnanavarshini M
Sprint-1	Building Login Page	USN-1	As a user, I can log into the application through a mail and password	5	High	Pooja Balasubramanian

Sprint-2	Register Page	USN-2	As a new user, I can register for the application through email.	5	High	Sanjay Kannan M
Sprint-2	Splitting the dataset	Task-4	Splitting dataset into train and test split.	3	Medium	Gnanavarshini M
Sprint-2	Building the Model	Task-5	Build three different ML models for classification and prediction.	12	High	Sandhya S Nithyakamal Ilamurugu
Sprint-3	Home Page	USN-3	As a user, I can view the symptoms of CKD and test vitals required for its prediction.	5	Medium	Gnanavarshini M
Sprint-3	Comparing different ML Models	Task-6	Evaluating each model and choosing the one with better accuracy.	3	Low	Pooja Balasubramanian
Sprint-3	Creating User Database	Task-7	Storing the user login details in the database.	12	High	Sandhya S Sanjay Kannan M
Sprint-4	Prediction Page	USN-4	As a user, I can view the test results.	5	Low	Sanjay Kannan M
Sprint-	Train model on IBM	Task-8	Train the ML model on IBM	7	Medium	Nithyakamal

4	Cloud		Watson.			Ilamurugu
Sprint-4	Flask Integration	Task-9	Integrating the HTML files with the ML model.	8	High	Pooja Balasubramanian

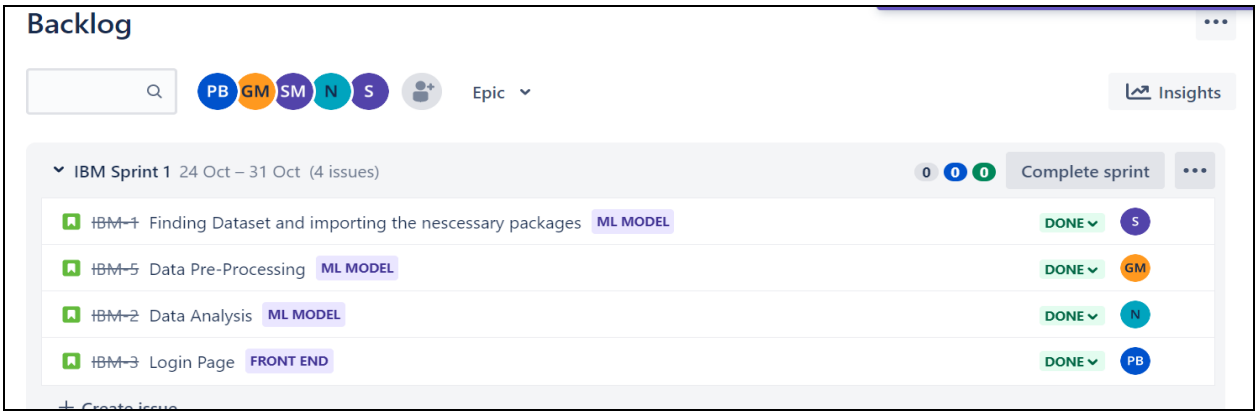
6.2 SPRINT DELIVERY SCHEDULE

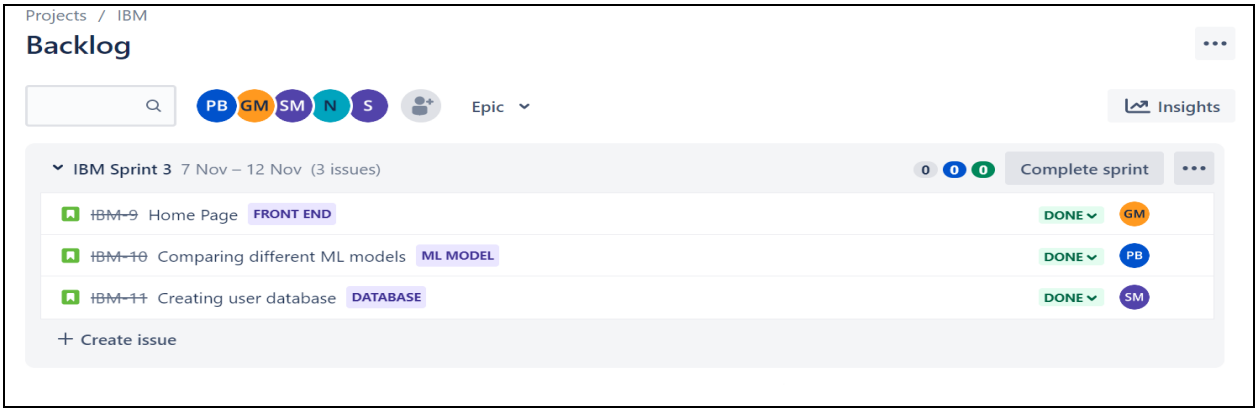
Sprint	Task
Sprint 1 (24th to 29th October)	● Dataset is downloaded and data analysisis performed ● Data is cleaned, missing values are handled, and labelencoding is performed ● Login in usinglogin credentials
Sprint 2 (31st to 5th October)	● Register into diagnosis tool ● Dataset is split into train and test ● Models are built for prediction and classification
Sprint 3 (7th to 12th October)	● Symptoms andtest vitals viewed ● Evaluating the models ● Storing login details
Sprint 4 (14th to 19th October)	● Test results can be viewed. ● ML model to be trained on IBM Watson ● Integrating the website and HTML model

PNT2022TMID36002

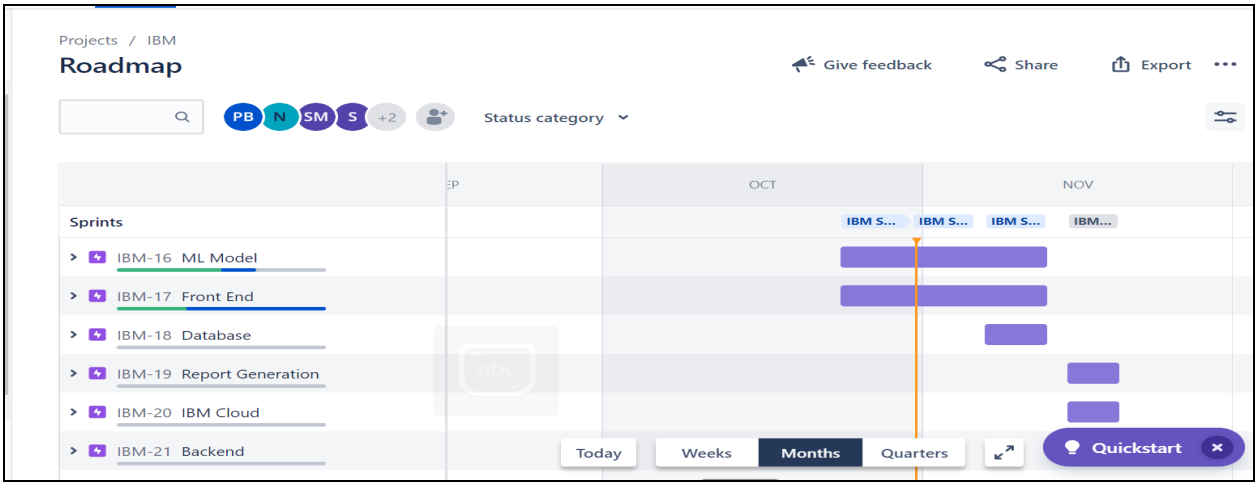
6.3 REPORTS FROM JIRA

BACKLOG:

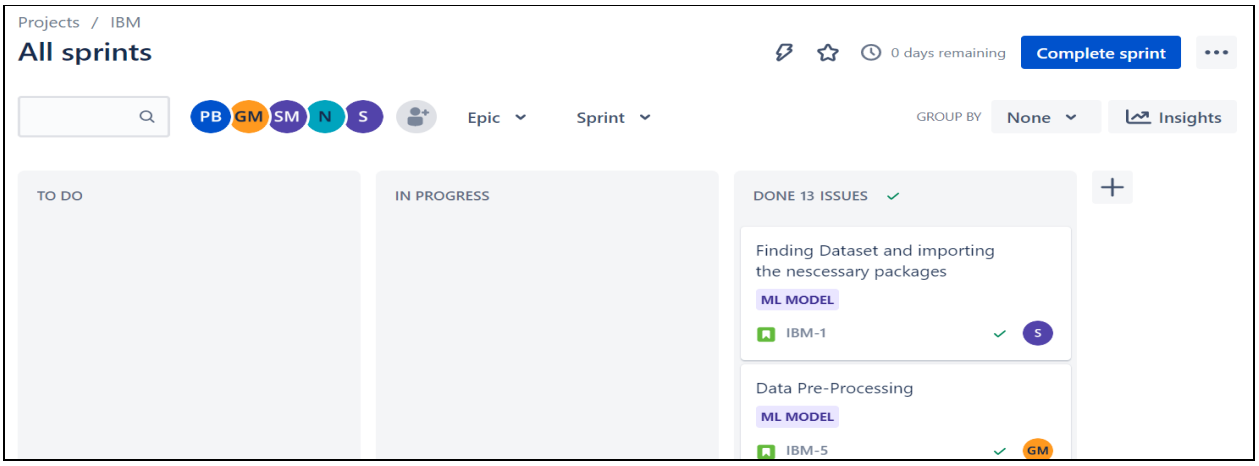




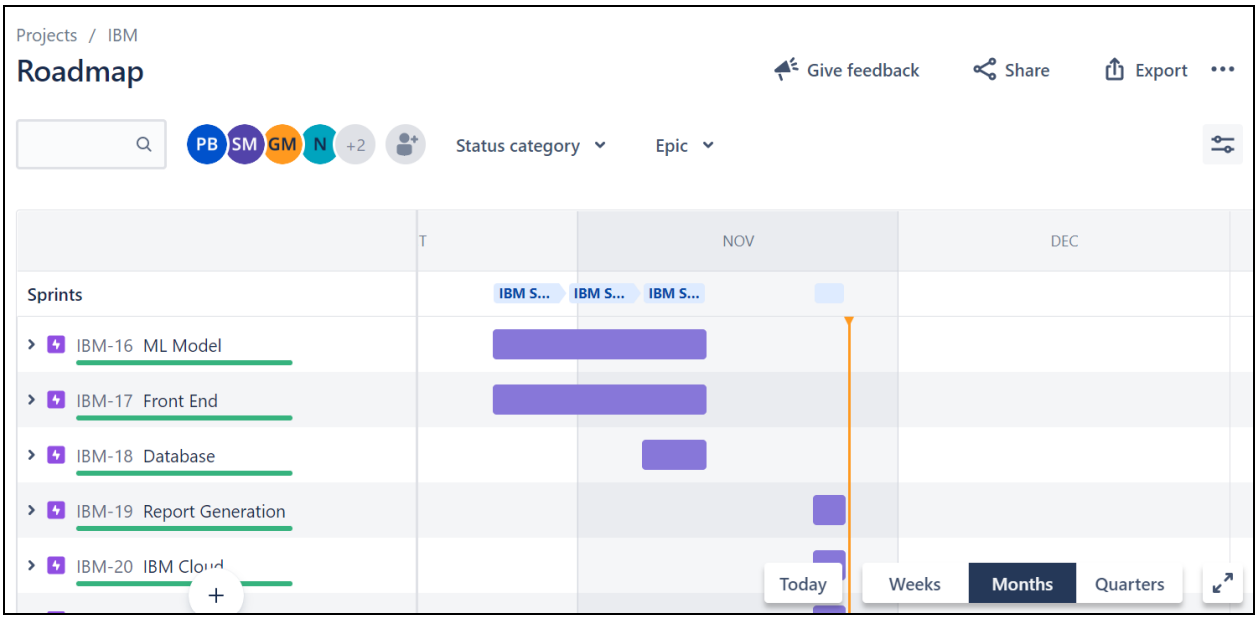
BOARDS:



PNT2022TMID36002



ROADMAP:



CHAPTER - 7
CODING AND SOLUTIONING

7.1 FEATURE 1: HOME PAGE

The Home page consists of description and statistics of Chronic Kidney Disease and acts as the navigation bars for the prediction form page.

TEMPLATE:

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <title>CKD Prediction</title>
  <link
href="https://fonts.googleapis.com/css2?family=Poppins:wght@400;600;700;900&dis
play=swap" rel="stylesheet">
  <link href="/static/styles/logincss.css" rel="stylesheet">
<style>
body:before {
content: '';
position: fixed;
  width: 100vw;
  height: 100vh;
  background-image: url({{url_for('static',filename='images/ckd.jpg')}});
  background-position: center center;
  background-repeat: no-repeat;
  background-attachment: fixed;
  -webkit-background-size: cover;
  background-size: cover;
```

```
  -webkit-filter: blur(10px);
  -moz-filter: blur(10px);
  filter: blur(10px);
}
</style>
</head>
<body>
  <div class="idx-page contact-form">
      <h3><center>CHRONIC KIDNEY DISEASE PREDICTOR</center></h3>
      <p align="justify">Chronic Kidney Disease prediction is one of the most important issues in healthcare analytics. 10% of the population worldwide is affected by chronic kidney disease (CKD), and millions die each year because they do not have access to affordable treatment. It is essential to check if you have CKD or not. Find out NOW!</p><br>
      <center><div class="txt"><a href="form.html">Take our CKD test now!!</a></div><br>
  </div>
</body>
</html>
```

7.2 FEATURE 2: PREDICTION

HTML PAGE - form.html

The form page consists of questionarrie to procure test vitals crucial for the prediction of CKD.

```
<!DOCTYPE html>
<html>
  <head>
```

```
    <title>CKD Prediction Form</title>
    <link
href='https://fonts.googleapis.com/css?family=Open+Sans:400,300,300italic,400italic,60
0' rel='stylesheet' type='text/css'>
    <link href="https://fonts.googleapis.com/css?family=Roboto:300,400,500,700"
rel="stylesheet">
    <style>
      html, body {
      min-height: 100%;
      }
      body, div, form, input, select, p {
      padding: 0;
      margin: 0;
      outline: none;
      font-family: 'Poppins', sans-serif;
      font-size: 14px;
      color: black
      }
      h1 {
      margin: 0;
      font-weight: 400;
      font-family: sans-serif;
      font-weight: bolder;

      }
      h3 {
      margin: 12px 0;
      color: black;
      }
```

```
.main-block {
display: flex;
justify-content: center;
align-items: center;
background: #fff;
background-image: url({{url_for('static',filename='images/ckd.jpg')}});
height: 100%;
background-position: center;
background-repeat: no-repeat;
background-size: cover;
}
form {
width: 100%;
padding: 20px;
}
fieldset {
border: none;
border-top: 1px solid black;
}
.account-details, .personal-details {
display: block;
flex-wrap: wrap;
/*justify-content: space-between;*/
}
.account-details >div, .personal-details >div >div {

display: flex;
align-items: center;
margin-bottom: 10px;
```

```
}
.account-details >div, .personal-details >div, input, label {
width: 100%;
}
label {
padding: 0 5px;
text-align: right;
vertical-align: middle;
}
input {
padding: 5px;
vertical-align: middle;
}
.checkbox {
margin-bottom: 10px;
}
select, .children, .gender, .bdate-block {
width: calc(100% + 26px);
padding: 5px 0;
}
select {
background: transparent;
}
.gender input {
width: auto;
}
.gender label {
padding: 0 5px 0 0;
}
```

```
.bdate-block {
display: flex;
justify-content: space-between;
}
.birthdate select.day {
width: 35px;
}
.birthdate select.mounth {
width: calc(100% - 94px);
}
.birthdate input {
width: 38px;
vertical-align: unset;
}
.checkbox input, .children input {
width: auto;
margin: -2px 10px 0 0;
}
button {
width: 100%;
padding: 10px 0;
margin: 10px auto;
border-radius: 5px;
border: none;
background: red;
font-size: 14px;
font-weight: 600;
color: #fff;
box-shadow: 0 8px 16px 0 rgba(0,0,0,0.2), 0 6px 20px 0 rgba(0,0,0,0.19);
```

```
    }
    button:hover {
    background: red;
    opacity: 0.4;
    }
    @media (min-width: 568px) {
    .account-details >div, .personal-details >div {
    width: 50%;
    }
    label {
    width: 40%;
    }
    input {
    width: 60%;
    }
    select, .children, .gender, .bdate-block {
    width: calc(60% + 16px);
    }
    }
  </style>
</head>
<body>
  <div class="main-block">
  <form method="POST", action="/predict">
    <h1 align="center">ENTER YOUR DETAILS</h1>
    <fieldset>
      <legend>
        <h3>Personal Details</h3>
      </legend>
```

```
      <div  class="account-details">
       <div><label><b>Name</b></label><input type="text" name="a" required></div>

       <div>
            <label><b>Gender</b></label>
        <div class="gender">
         <input type="radio" value="none" id="male" name="b" required/>
         <label for="male" class="radio">Male</label>
         <input type="radio" value="none" id="female" name="b" required/>
         <label for="female" class="radio">Female</label>
        </div>
       </div>

       <div><label><b>Age</b></label><input type="number" name="c" required
min="3" max="100"></div>

      </div>
    </fieldset>
    <fieldset>
      <legend>
       <h3>Medical Details</h3>
      </legend>
      <div  class="personal-details">
       <div>
        <div><label><b>Blood Pressure</b></label><input type="number" name="d"
required></div>
        <div><label><b>Specific Gravity</b></label><input type="number" name="e"
required step=".01"></div>
        <div><label><b>Albumin</b></label><input type="number" name="f"></div>
```

```
        <div><label><b>Sugar</b></label><input type="number" name="g"
required></div>
        <div>
              <label><b>How about RBC's?</b></label>
          <div class="gender">
            <input type="radio" value="1" id="normal" name="h" required/>
            <label for="normal" class="radio">Normal</label>
            <input type="radio" value="0" id="abnormal" name="h" required/>
            <label for="abnormal" class="radio">Abnormal</label>
          </div>
        </div>
        <div>
              <label><b>Pus Cell</b></label>
          <div class="gender">
            <input type="radio" value="1" id="normal" name="i" required/>
            <label for="normal" class="radio">Normal</label>
            <input type="radio" value="0" id="abnormal" name="i" required/>
            <label for="abnormal" class="radio">Abnormal</label>
          </div>
        </div>
        <div>
              <label><b>Any clumps in your Puss cells?</b></label>
          <div class="gender">
            <input type="radio" value="1" id="present" name="j" required/>
            <label for="present" class="radio">Present</label>
            <input type="radio" value="0" id="notpresent" name="j" required/>
            <label for="notpresent" class="radio">Not Present</label>
          </div> </div>
        <div>
```

```
            <label><b>Bacteria?</b></label>
        <div class="gender">
          <input type="radio" value="1" id="present" name="k" required/>
          <label for="present" class="radio">Present</label>
          <input type="radio" value="0" id="notpresent" name="k" required/>
          <label for="notpresent" class="radio">Not Present</label>
        </div>
      </div>
      <div><label><b>Blood Glucose Random</b></label><input type="number"
name="l" required></div>
      <div><label><b>Blood Urea</b></label><input type="number" name="m"
required></div>
      <div><label><b>Serum Creatinine</b></label><input type="number" name="n"
required step=".01"></div>
      <div><label><b>Sodium</b></label><input type="number" name="o"
required></div>
      <div><label><b>Potassium</b></label><input type="number" name="p"
required step=".01"></div>
      <div><label><b>Hemoglobin</b></label><input type="number" name="q"
required step=".01"></div>
      <div><label><b>Packed Cell Volume</b></label><input type="number"
name="r" required></div>
      <div><label><b>WBC Count</b></label><input type="number" name="s"
required></div>
      <div><label><b>RBC Count</b></label><input type="number" name="t"
required step=".01"></div>
      <div>
            <label><b>Do you have hypertension</b></label>
        <div class="gender">
```

```
    <input type="radio" value="1" id="yes" name="u" required/>
    <label for="yes" class="radio">Yes</label>
    <input type="radio" value="0" id="no" name="u" required/>
    <label for="no" class="radio">No</label>
  </div>
</div>
<div>
      <label><b>Do you have Diabetes Mellitus?</b></label>
  <div class="gender">
    <input type="radio" value="1" id="yes" name="v" required/>
    <label for="yes" class="radio">Yes</label>
    <input type="radio" value="0" id="no" name="v" required/>
    <label for="no" class="radio">No</label>
  </div>
</div>
<div>
      <label><b>Do you have CAD?</b></label>
  <div class="gender">
    <input type="radio" value="1" id="yes" name="w" required/>
    <label for="yes" class="radio">Yes</label>
    <input type="radio" value="0" id="no" name="w" required/>
    <label for="no" class="radio">No</label>
  </div>
</div>
<div>
      <label><b>How's your apetite?</b></label>
  <div class="gender">
    <input type="radio" value="0" id="good" name="x" required/>
    <label for="good" class="radio">Good</label>
```

```
            <input type="radio" value="1" id="poor" name="x" required/>
            <label for="poor" class="radio">Poor</label>
          </div>
        </div>
        <div>
              <label><b>Do you have Pedal Edema?</b></label>
          <div class="gender">
            <input type="radio" value="1" id="yes" name="y" required/>
            <label for="yes" class="radio">Yes</label>
            <input type="radio" value="0" id="no" name="y" required/>
            <label for="no" class="radio">No</label>
          </div>
        </div>
        <div>
              <label><b>Do you have Anaemia?</b></label>
          <div class="gender">
            <input type="radio" value="1" id="yes" name="z" required/>
            <label for="anaemia" class="radio">Yes</label>
            <input type="radio" value="0" id="no" name="z" required/>
            <label for="anaemia" class="radio">No</label>
          </div>
        </div>
      </div>
    </div>
  </fieldset>
  <input type="submit" , value='Predict'>
 </form>
 </div>
</body> </html>
```

HTML PAGE - result.html

```
<!DOCTYPE html>
<html lang="en">
<head>
  <meta charset="UTF-8">
  <title>Results</title>
  <link
href="https://fonts.googleapis.com/css2?family=Poppins:wght@400;600;700;900&dis
play=swap" rel="stylesheet">
  <link href="static/styles/logincss.css" rel="stylesheet">
<style>
body:before {
content: '';
position: fixed;
  width: 100vw;
  height: 100vh;
  background-image: url({{url_for('static',filename='images/ckd.jpg')}});
  background-position: center center;
  background-repeat: no-repeat;
  background-attachment: fixed;
  -webkit-background-size: cover;
  background-size: cover;
  -webkit-filter: blur(10px);
  -moz-filter: blur(10px);
  filter: blur(10px);
}
</style>
</head>
<body>
```

```
  <div class="idx-page contact-form">
      <h3><center>CHRONIC KIDNEY DISEASE PREDICTOR</center></h3>
      <p font-size="20px" align="center">{{ data }}</p><br></center>
      <br><br><br><br><br><br><br><br>
      <center><div class="txt"><a href="/">Take CKD test
again!!</a></div><br></center>
  </div>
</body>
</html>
```

7.3 FEATURE 3: APPLICATION

With the help of the application, the user will be able to check if he/she has CKD and take necessary measures accordingly.

FLASK CODE:

```
#importing libraries
import os
import numpy as np
import flask
import pickle
from flask import Flask, render_template, request
from flask_ngrok import run_with_ngrok
import requests

# NOTE: you must manually set API_KEY below using information retrieved from your
IBM Cloud account.
API_KEY = "xsc0DgO5B50sSt04GVW3Lnrcm1Bou0Bn9oFqlEK7g01N"
token_response = requests.post('https://iam.cloud.ibm.com/identity/token',
data={"apikey":
```

```
 API_KEY, "grant_type": 'urn:ibm:params:oauth:grant-type:apikey'})
mltoken = token_response.json()["access_token"]

header = {'Content-Type': 'application/json', 'Authorization': 'Bearer ' + mltoken}
#model = pickle.load(open('model.pkl', 'rb'))
#creating instance of the class
app=Flask(__name__, template_folder="/content/")
run_with_ngrok(app)

#to tell flask what url shoud trigger the function index()
@app.route('/')
def man():
    return flask.render_template('index.html')
@app.route('/form.html')
def forw():
    return flask.render_template('form.html')

@app.route('/predict',methods = ['POST'])
def home():
    x=[param for param in request.form.values()]
    x=[float(p) for p in x[2:]]
    print(x)
    lis = [['age', 'blood_pressure', 'specific_gravity', 'albumin', 'sugar',
      'red_blood_cells', 'pus_cell', 'pus_cell_clumps', 'bacteria',
      'blood_glucose_random', 'blood_urea', 'serum_creatinine', 'sodium',
      'potassium', 'hemoglobin', 'packed_cell_volume',
      'white_blood_cell_count', 'red_blood_cell_count', 'hypertension',
      'diabetesmellitus', 'coronary_artery_disease', 'appetite','pedal_edema','anemia']]
    payload_scoring = {"input_data": [{"field": lis, "values": [x]}]}
```

```
    response_scoring = requests.post('https://us-
south.ml.cloud.ibm.com/ml/v4/deployments/f4af460b-49e4-4b53-9c7a-
25929a2729b8/predictions?version=2022-11-24', json=payload_scoring,
    headers={'Authorization': 'Bearer ' + mltoken})
    pred = response_scoring.json()
    p = pred['predictions'][0]['values'][0][0]
    print(p)
    if int(p)== 0:
      prediction ='YOU HAVE CHRONIC KIDNEY DISEASE'
    else:
      prediction ='YOU DON\'T HAVE CHRONIC KIDNEY DISEASE'
    return render_template('result.html', data=prediction)

if __name__ == "__main__":
  app.run()
```

CHAPTER - 8
TESTING

8.1 TESTCASES

Model

● The models used for the project are: Logistic Regression, RandomForest, and K-Nearest Neighbors(KNN).

● Each of the above models were trained and accuracies were obtained for the same.

● The model with the best training and testing accuracy was used for further processing. In our case, logistic regression has the maximum accuracy of 90.8%.

Web pages (Home page, Prediction input page, Prediction output page):

● Home page with description - working well without issues.

● Prediction input page - is able to get all the inputs without any issues.

● Prediction output page - able to display the prediction accurately.

Flask App:

All the connections made in the python script have been tested. All the links are working properly without any issues. The app was successful when tested for functionality.

Model deployment:

The model was trained on IBM cloud and deployed with no issues. The model can be used publicly.

				Date	3-Nov-22								
				Team ID	PNT2022TMID36002								
				Project Name	Project - Early Detection of Chronic Kidney Disease using								
				Maximum Marks	4 marks								
Test case ID	Feature Type	Component	Test Scenario	Pre-Requisite	Steps To Execute	Test Data	Expected Result	Actual Result	Status	Comments	TC for Automation(Y/N)	BUG ID	Executed By
Home_page_T1	Functional	Home Page	Verify if user is able to see the	-	1.Click the flask app URL	-	Home page should be opened	Working as	Pass	The home page opens			Pooja Balasubramanian
Home_page_T2	UI	Home Page	Verify all the UI elements in home	-	1.Click the flask app URL	-	CKD description and statistics should	The page is	Pass	All the components of the			Sanjay Kannan M
Home_page_T3	Functional	Home page	Verify if user is able to navigate to the Predictor page from the home page where the test vitals must be entered	-	1. Click the flask app URL 2. Click on the predictor button to open the predictor input page	-	The predictor input page should get opened	Working as expected	Pass	The predictor input page was opened			Sandhya S
Predictor_page_T1	UI	Predictor	Verify all the UI elements in		1. Click the flask app URL	-	All the input boxes from the predictor	The page is	Pass	All the input boxes are			Nithyakamal Ilamurugu
Predictor_page_T2	Functional	Predictor	Navigate from this page to result		1. Click the flask app URL	-	The connections are all accurate and	Working as	Pass	The result page is displayed			Pooja Balasubramanian
Predictor_page_T3	Functional	Predictor	Give all inputs to test for accurate		1. Click the flask app URL	[65.0, 70.0, 1.01, 0.0, 0.0,	All inputs must be given to the result	Working as	Pass	The predictor is working			Gnanavarshini M
Predictor_page_T4	Predictor_page_T5	Predictor_pa	Navigate from this page to result		1. Click the flask app URL	-	Result box must be displayed properly	Displaying	Pass	Result page is opened.			Sanjay Kannan M
Result_page_T1	UI	Result Page	Verify all the UI elements in result		1. Click the flask app URL	-	The connections are all accurate and	Working as	Pass	Result page is displayed			Sandhya S
Result_page_T2	UI	Result Page	Verify if the prediction is getting		1. Click the flask app URL	-	The output obtained from the	The output	Pass	CKD Prediction is displayed			Nithyakamal Ilamurugu
Result_page_T3	Functional	Result Page	Verify if user is able to navigate to		1. Click the flask app URL	Prediction : You Have CKD	Home page should get displayed	Working as	Pass	Home page is displayed			Gnanavarshini M

8.2 USER ACCEPTANCE TESTING

DEFECT ANALYSIS

This report shows the number of resolved or closed bugs at each severity level, and how they were resolved

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Sub-total
By Design	0	2	1	1	4
Duplicate	0	1	0	1	2
External	1	1	1	0	3
Fixed	7	3	2	5	17
Not Reproduced	0	0	1	0	1
Skipped	0	0	1	1	2
Won't Fix	0	2	2	1	5
Totals	8	9	8	9	34

TESTCASE ANALYSIS

This report shows the number of test cases that have passed, failed, and untested.

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	8	0	1	7
Client Application	10	0	2	8
Security	2	0	1	1
Outsource Shipping	2	0	0	2
Exception Reporting	3	0	1	2
Final Report Output	6	0	0	6
Version Control	1	0	0	1

CHAPTER - 9
RESULTS

9.1 PERFORMANCE METRICS

- The models used for the project are Logistic Regression, RandomForest, and K-Nearest Neighbors(KNN). Comparing the accuracies of the above three models:

```
[76] print("ACCURACY")
     print("Logistic Regression: ",(log_acc)*100,"%")
     print('Random Forest: ',(rand_forest_acc)*100,"%")
     print("K-Neighest Neighbors: ",(knn_acc)*100,"%")

     ACCURACY
     Logistic Regression:  90.83333333333333 %
     Random Forest:  89.16666666666667 %
     K-Neighest Neighbors:  85.0 %
```

- Logistic Regression has the maximum accuracy of 90.8%. The training score of this model:

```
#training score
train_score=logreg.score(x_train,y_train)
print('Training Score: ',(train_score)*100,"%")

Training Score:  98.92857142857143 %
```

- **Model Summary:**

```
{'C': 1.0,
 'class_weight': None,
 'dual': False,
 'fit_intercept': True,
 'intercept_scaling': 1,
 'l1_ratio': None,
 'max_iter': 100,
 'multi_class': 'auto',
 'n_jobs': None,
 'penalty': 'l2',
 'random_state': None,
 'solver': 'lbfgs',
 'tol': 0.0001,
 'verbose': 0,
 'warm_start': False}
```

- **Classification Report:**

It contains the precision, recall, F1-score, and support of the Logistic Regression model.

```
Classification Report:
                precision    recall  f1-score   support

           0       0.99      0.95      0.97        78
           1       0.91      0.98      0.94        42

    accuracy                           0.96       120
   macro avg       0.95      0.96      0.95       120
weighted avg       0.96      0.96      0.96       120
```

- **Regression Error Metrics:**

The below error metrics are calculated for the Logistic Regression model:

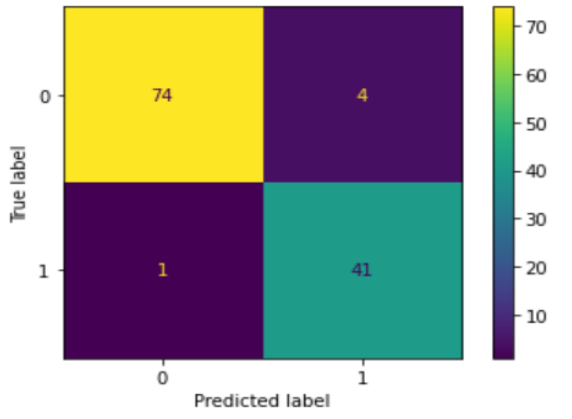
  - Mean Absolute Error(MAE)
  - Mean Squared Error(MSE)
  - Root Mean Squared Error(RMSE)

```
Regression Metrics:
MAE                    :  0.041666666666666664
MSE                    :  0.2041241452319315
RMSE                   :  0.45180100180492244
```

- **Confusion Matrix:**

It summarizes and visulaizes the performance of the classification model(Logistic Regression).

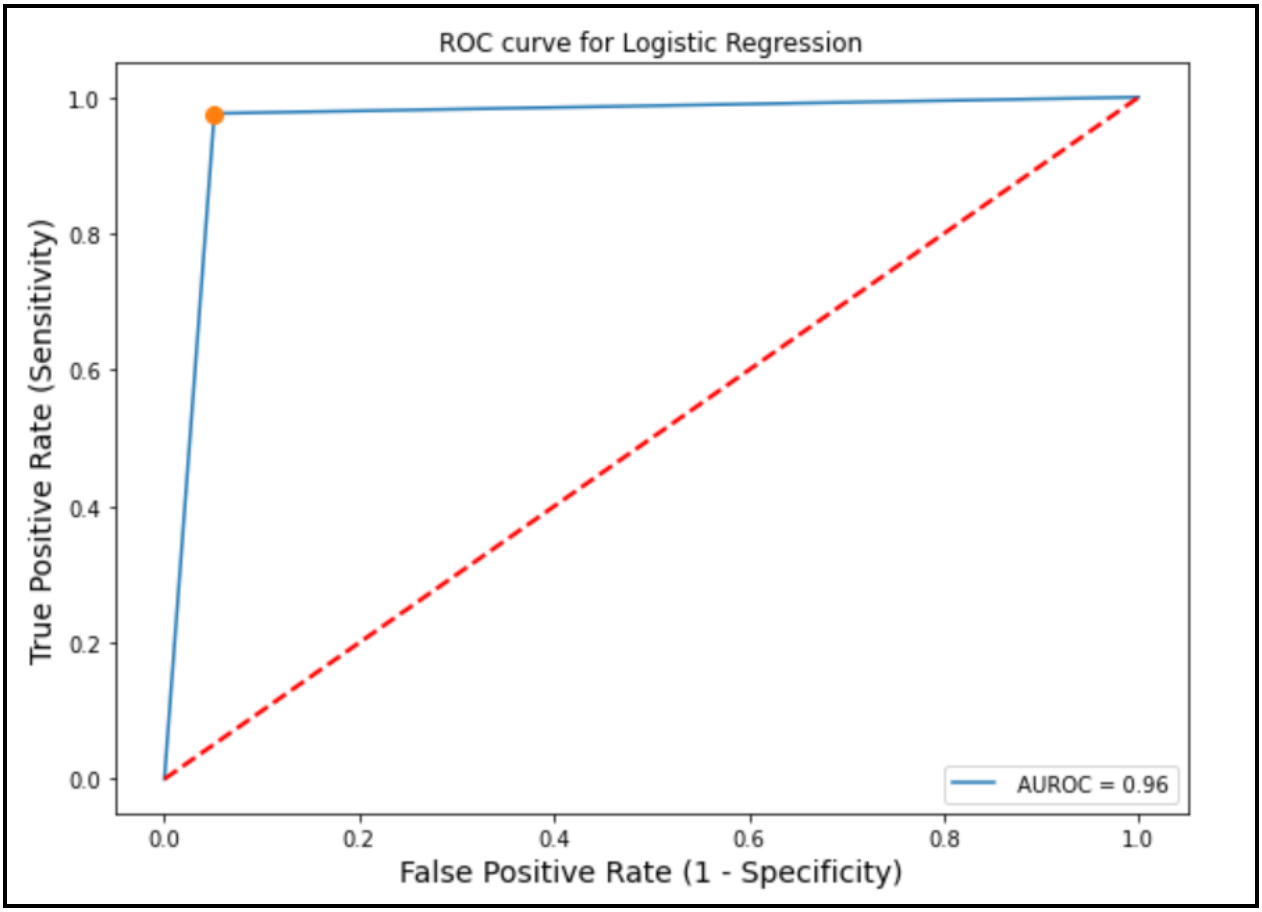
Confusion Matrix for Logistic Regression:



PNT2022TMID36002

ROC Curve:

A graph to show the performance of a classification model(Logistic Regression) at all classification thresholds.



CHAPTER - 10
ADVANTAGES AND DISADVANTAGES

10.1 ADVANTAGES

- It is crucial to identify Chronic Kidney Disease at an early stage, as it can tremendously reduce the cost of the treatment, and increase the chance of treating CKD.
- It helps people become more aware of their health conditions and to follow a healthier lifestyle.
- Training a model with different symptoms and parameters can help us identify different correlations and causes of CKD which can help with further research.
- People who are pessimistic about meeting a doctor and discussing their symptoms can give their symptoms as input and get an accurate result.
- The workload of people to book an appointment with a doctor for a consultation would be reduced and would simplify the process.
- The web-page is user friendly which makes it very simple for a normal user to navigate and identify they have CKD.

10.2 DISADVANTAGES

- All chronically ill patients face a common set of challenges, which imply difficult lifestyle adjustments. Some of these adjustments include complex medication regimens, obtaining helpful medical care,dealing with symptoms, disability and emotional impacts, all of which involve significant psychological processes.

- Chronic illness involves recognition of the worlds of pain and suffering, possibly even of death, which are normally only seen as distant possibilities or the plight of others.
- There is a chance that incorrectly entered values will result in an incorrect assessment of the patient's health profile.
- It has been suggested that chronic illness diagnoses negatively affect relevant functionings for following long-term treatments, such as affiliation. This could create a cycle of disadvantage for those who are chronically unwell.

CHAPTER - 11
CONCLUSION

Data science and machine learning are used in the application area of chronic kidney disease prediction to foretell the patients' kidney health. In order to safeguard our kidney and decrease kidney-related disease mortality, it is crucial to forecast kidney disease. In order to make predictions based on different patients and their diagnosis, this study analyzed several machine learning methods, including Random Forest, KNN and Logistic Regression. These were displayed and put to the test using information provided by IBM.

CHAPTER - 12
FUTURE SCOPE

The ability to predict chronic kidney disease can also be extended to predicting other related diseases.The project can be further developed by providing life advice to people on how to lead a healthy lifestyle which can help reduce the probability of getting affected by CKD. Based on the results people can be referred to a doctor for further analysis and treatment.

CHAPTER - 13
APPENDIX

13.1 SOURCE CODE

- **Source Code Link:** https://github.com/IBM-EPBL/IBM-Project-2573-1658475170/tree/main/Final%20Deliverables

13.2 GITHUB AND PROJECT DEMO LINK

- **GitHub Link:** https://github.com/IBM-EPBL/IBM-Project-2573-1658475170
- **Demo Video Link:** https://github.com/IBM-EPBL/IBM-Project-2573-1658475170/blob/main/Final%20Deliverables/Demo%20Video.mp4