PROJECT REPORT DOCUMENTATION

EARLY DETECTION OF CHRONIC KIDNEY DISEASE

1. INTRODUCTION

- 1.1 Project Overview
- 1.2 Purpose

2. LITERATURE SURVEY

- 2.1 Existing problem
- 2.2 References
- 2.3 Problem Statement Definition

3. IDEATION & PROPOSED SOLUTION

- 3.1 Empathy Map Canvas
- 3.2 Ideation & Brainstorming
- 3.3 Proposed Solution
- 3.4 Problem Solution fit

4. REQUIREMENT ANALYSIS

- 4.1 Functional requirement
- 4.2 Non-Functional requirements

5. PROJECT DESIGN

- 5.1 Data Flow Diagrams
- 5.2 Solution and Technical Architecture
- 5.3 User Stories

6. PROJECT PLANNING & SCHEDULING

- 6.1 Sprint Planning
- 6.2 Sprint Estimation and Delivery Schedule

7. CODING & SOLUTIONING

- 8. TESTING
- 9. RESULTS
- 10. ADVANTAGES & DISADVANTAGES
- 11. CONCLUSION
- 12. FUTURE SCOPE
- 13. APPENDIX GitHub Link

1. INTRODUCTION

1.1 Project Overview: -

- The objective of the present project is to employ machine learning algorithms in an attempt to develop a prediction model for progression to detect the Chronic Kidney disease in earlier stage.
- By using the wrapper method,a feature reductionanalysis has been performed to find the attributes that detect this disease with high accuracy.
- By considering the parameters like albumin, specific gravity, diabetes mellitus, hemoglobin, and hypertension as features, we can predict the CKD at earlier stage.
- As the result of our project, following program objectives can be outlined as the fundamentals for research and practical work in the field of Nephrology,
 - o Deep research about the chronic kidney disease
 - o Conduct an enquiry about the causes of chronic kidney disease
 - Study the research observations of reported ways of treating the CKD
 - Select the most accurate machine learning method and carry out new research to trace the disease earlier
 - Create awareness among the ordinary people and to produce a user friendly model to assesstheir conditions at ease of their comfort.
 - State the limitations of the current program of CKD and to produce new ideas to make it better.
 - o To implement solutions for the limitations stated.
 - Introduce research findings to the nephrology and medical researchers to update their treatment techniques and the overallprocess o f finding CKD.
 - o Outline the directions for future enhancement.

1.2 Purpose: -

The purpose of the project is to alert doctors for an early detection of kidney disease and hence ensure speedy recovery or prevention of kidney disease. This Project aims at creating a model for early detection of Chronic Kidney Disease using Machine Learning technology. The output is integrated with Flask. The front end developed in html is used to receive user input on various parameters needed to decide on the early detection of kidney disease. The same model is deployed into IBM cloud using API keys and scoring endpoints.

2. LITERATURE SURVEY

2.1 Existing Problem:-

Presently kidney disease is detected at late stages in many countries leading to loss of precious lives. There are very few means to identify them at an early stage. Most of the user details remain unverified and it's difficult to track the fake users. The user interface of the application is not user friendly and the user must have a device with an android operating system with an active internet connection to interact with this application.

2.2 References:-

S.NO	PAPER	AUTHOR	YEA R	PROPOSED METHODS AND ALGORITH MS	ACCURAC Y
1.	Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study	Njoud Abdullah Almansour,Hajr a Fahim Syed, Nuha Radwan Khayat, Kanaan Altheeb, Renad Emad Juri, Jamal Alhiyafi, Saleh Alrashed, Sunday O.Olatunji	2019	Artificial Neural Network (ANN) and Support Vector Machine (SVM) techniques	97.75%
2.	Chronic Kidney Disease Prediction using Machine Learning Models	S.Revathy, B.Bharathi, P.Jeyanthi, M.Ramesh	2019	Decision Tree, Support Vector Machine (SVM) and Random Forest Classifier	98.33%
3.	A Machine Learning Methodology for Diagnosing Chronic Kidney Disease	Jiongming Qin, Lin Chen, Yuhua Liu,Chuanj un Liu,Changh ao Feng, and Bin Chen	2020	Logistic regression, random forest, support vector machine, k- nearest neighbour, naive Bayes classifier and feed forward neural network.	99.83%

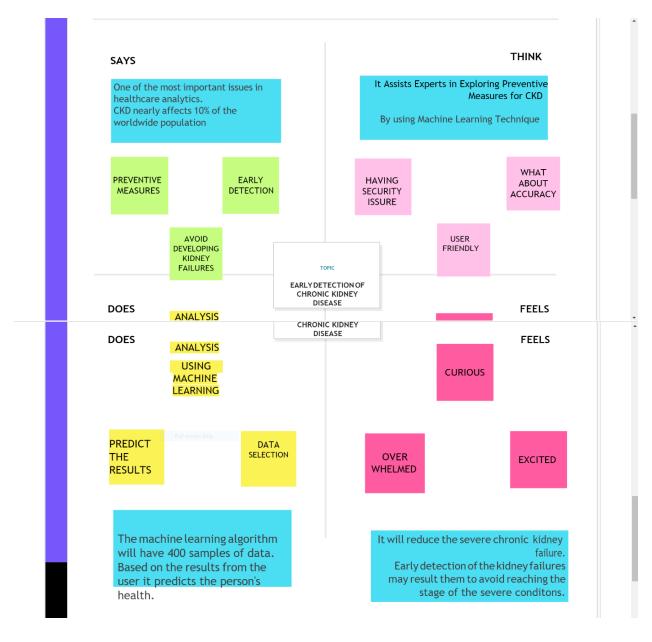
4.	An Empirical Evaluation of Machine Learning Techniques for Chronic Kidney Disease Proph ecy	Bilal Khan, Rashid Naseem, Fazal Muhammad, Ghulam Abbas, and Sunghwan Kim	2020	Support Vector Machine (SVM), Logistic Regression, Naïve Bayes, Artificial Neural Network (ANN) and Support Vector Machine (SVM) techniques	98.25%
5.	Prediction of Chronic Kidney Disease - A Machine Learning Perspective	Pankaj chittora, Sandeep chaurasia ,Prasun chakrabarti, Gaurav kumawat, Tulika chakrabarti, Zbigniew leonowicz ,Michał jasinski, Lukasz jasinski, Radomir gono, Elżbieta jasińska, and Vadim bolshev	2021	Artificial Neural Network (ANN), C5.0, Chi-square Automatic interaction detector, logistic regression, linear support vector machine with penalty L1 & with penalty L2 and random tree.	98.86%

2.3 Problem Statement Definition

Kidney diseases avert the normal function of the kidney. Mainly due to the large amount of alcohol consumption kidney disease arises. Early prediction of kidney disease using classification and regression algorithms are an efficacious task that can help the doctors to diagnose the disease within a short duration of time. Discovering the existence of kidney disease at an early stage is a complex task for the doctors. The main objective of this project is to analyze the parameters of various classification algorithms and compare their predictive accuracies so as to find out the best classifier for determining the kidney disease. This Project examines data from kidney patients concentrating on relationships between a key list of kidney enzymes, proteins, age and gender using them to try and predict the likeliness of kidney disease. Here we are building a model by applying various machine learning algorithms find the best accurate model. And integrate to flask-based web application.

3. IDEATION & PROPOSED SOLUTION

3.1. Empathy Map Canvas:-

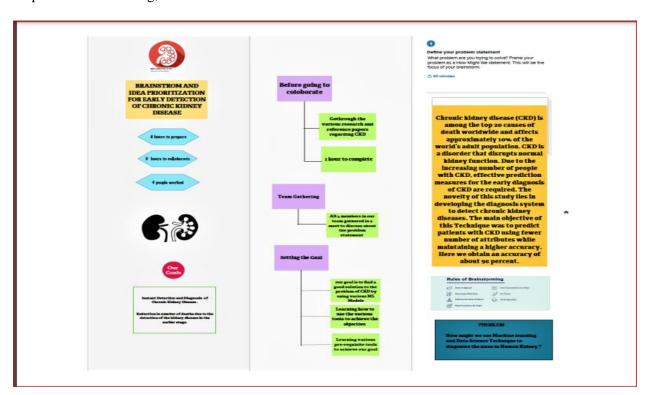


3.2 Ideation & Brainstorming:-

Brainstorming is an activity that will help you generate more innovative ideas. It's one of many methods of ideation—the process of coming up with new ideas—and it's core to the design thinking process. Brainstorming refers to a problem-solving technique used by teams or individuals. In this process, participants generate various ideas or solutions, then begin discussing and narrowing them down to the best options.

Ideation is often closely related to the practice of brainstorming, a specific technique that is utilized to generate new ideas. A principal difference between ideation and brainstorming is that ideation is commonly more thought of as being an individual pursuit, while brainstorming is almost always a group activity.

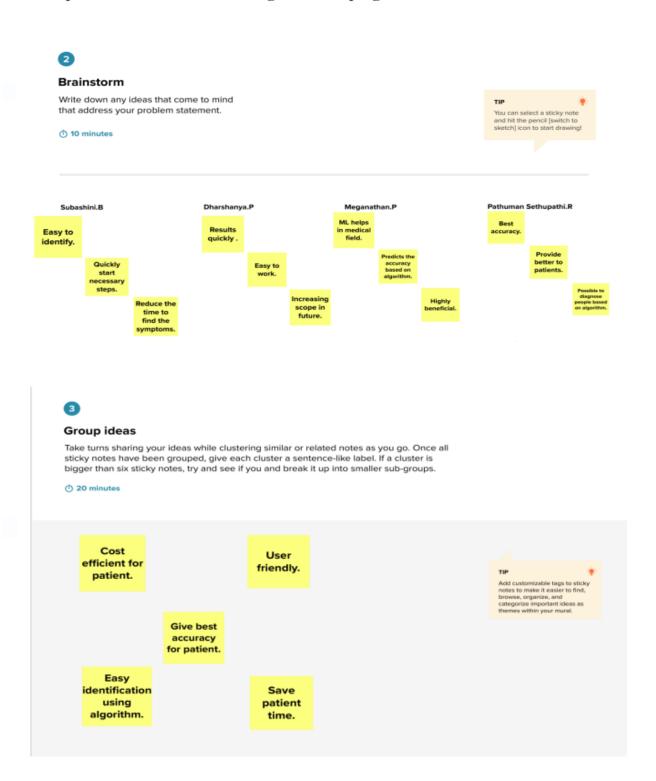
Ideation is the process where you generate ideas and solutions through sessions such as Sketching, Prototyping, Brainstorming, Brainwriting, Worst Possible Idea, and a wealth of other ideation techniques.



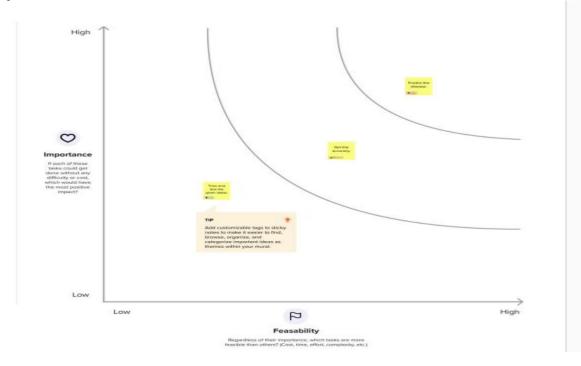
Step-1: Team Gathering, Collaboration and Select the Problem Statement

Step-2: Brainstorm, Idea Listing and Grouping

Step-2: Brainstorm, Idea Listing and Grouping



Step-3: Idea Prioritization



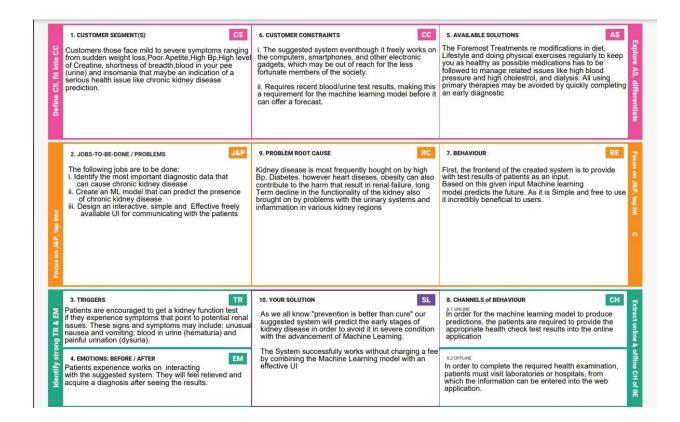
3.3 Proposed Solution:-

S.No.	Parameter	Description			
1.	Problem Statement (Problem to be solved)	Early detection of chronic diseases using machine learning.			
2.	Idea / Solution description	Strategies that are known to reduce mortality or prevent progressive loss of kidney function in CKD. 1. Control of BP use of an angiotensin -converting enzyme inhibitor or angiotensin Receptor blocker and strain treatment. 2. Control of blood glucose and use of sodium-glucose cotransporter inhibitors in people with diabetes. Early detection of CKD should be Beneficial because it enables clinicians to initiate effective treatment of mild prevention loss of kidney function and delaying or avoiding progression to kidney failure.			
3.	Novelty / Uniqueness	Novelty of this project is to predict whether the patient has chronic kidney disease or not, in a more accurate and faster way. Not only providing the nutrition diet chart for chronic kidney disease. That nutrition diet chart helps to improve any other diagnosed severe diseases also.			

4.	Social Impact / Customer Satisfaction	Social impacts such as education, unemployment status, income, familiarity and social stress have also a major role in determining chronic diseases both directly and indirectly by increasing susceptibility to behavioral risk factors.
5.	Business Model (Revenue Model)	Can generate revenue through direct customers. Can collaborate with the health care sector and generate revenue from their customers
6.	Scalability of the Solution	In Existing system the severe of kidney disease measured by common symptoms, such as blood in your pee (urine), an increased need to pee particularly at night, difficulty sleeping (insomnia), itchy skin so its takes time to find out the disease. But in our Proposed system to check 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) .This helps kidney patients to cure in early stages to take prescribed activities and foods . This method is very helpful for poor people.

3.4 Problem Solution Fit:-

The Problem-Solution Fit simply means that you have found a problem with your customer and that the solution you have realized for it actually solves the customer's problem. Problem-Solution Fit - this occurs when you have evidence that customers care about certain jobs, pains, and gains. At this stage you've proved the existence of a problem and have designed a value proposition that addresses your customers' jobs, pains and gains.



4. REQUIREMENT ANALYSIS

4.1 Functional Requirements:-

Functional Requirements:

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)	
FR-1	User Registration	Registration through Gmail	
FR-2	User Verification	Confirmation via Email	
FR-3	User Login	Login through Email	
FR-4	User Dashboard	Log in to the application	
FR-5	User Help	Report issues through Email	

4.2 Non- Functional Requirements:-

Non-functional Requirements:

Following are the non-functional requirements of the proposed solution.

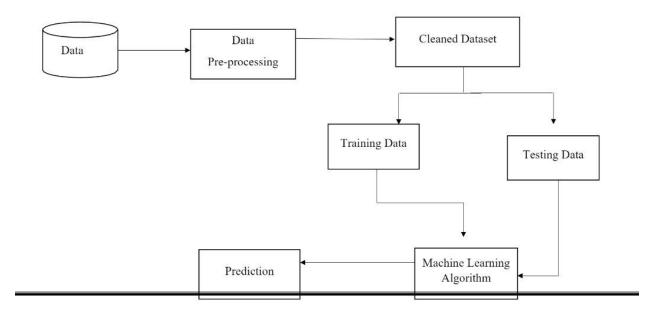
FR No.	Non-Functional Requirement	Description
NFR-1	Usability	This project provides the best accuracy for predicting the kidney disease.
NFR-2	Security	It will give security for the users in the way of Unique Id and password to login.
NFR-3	Reliability	It helps to identify the kidney disease in effective way, reduce the cost and user friendly.
NFR-4	Performance	Using both classification and regression algorithms it helps the doctors to diagnose the disease within a short duration of time.
NFR-5	Availability	It is accessible for web users.
NFR-6	Scalability	This project can be improved by giving medical suggestion for patients.

5. PROJECT DESIGN

5.1 Data Flow Diagrams:-

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

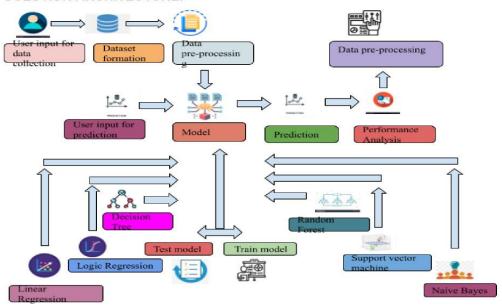
Data Flow Diagrams:



5.2 Solution & Technical Architecture:

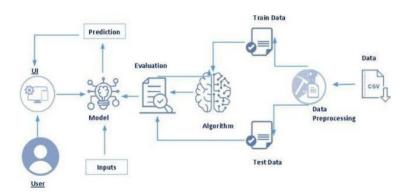
Solution Architecture:

SOLUTION ARCHITECTURE:



Technical Architecture:

Technical Architecture:



5.3 User Stories :-

User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Web user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High	Sprint-1
	Verification	USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High	Sprint-1
	Login	USN-3	As a user, I can login to the application by entering email and password.	Check whether password and email is correct	High	Sprint-1
	Dashboard	USN-4	If the email id and password is correct, the user can log in to the application otherwise it shows 'incorrect password or Id'.	View the dashboard of user who is log in	High	Sprint-1
Customer Care Executive	Help	USN-5	If the user faces any issues, he/she can report it to our mail id.	Report option will be available in web app	High	Sprint-2
Administrator	Verification	USN-6	Administrator also has unique Id and password to login. He has additional users to organize the users of this web app	Check whether password and email is correct	High	Sprint-3

6. PROJECT PLANNING AND SCHEDULING

6.1 Sprint Planning

Sprints are the backbone of any good Agile development team. And the better prepared you are before a sprint, the more likely you are to hit your goals. Spring planning helps to refocus attention, minimize surprises, and (hopefully) guarantee better code gets shipped. The main event during agile methodology is the sprint, the stage where ideas turn into innovation and valuable products come to life. On one hand, agile sprints can be highly effective and collaborative. At the same time, they can be chaotic and inefficient if they lack proper planning and guidance. And for this reason, making a sprint schedule is one of the most important things you can do to ensure that your efforts are successful.

Table-1: Components & Technologies:

S. No	Component	Description	Technology	
1.	User Interface	User interact with our application through web user interface	HTML, CSS and Python flask	
2.	Application Logic-1 Login.	When the user clicks on the login button, he/she is directed to login page, if they are registered already.	HTML, CSS and Python flask	
3.	Application Logic-2 Registration	When the user clicks on the Register button, he/she is directed to Register Page for further process.	HTML, CSS and Python flask	
4.	Application Logic-3 Test Vitals Form	After Logged in, when the user clicks on the test vital form button, he/she directed to the form page to enter the vitals for prediction.	Front end-HTML, CSS, MySQL, Python flask Backend-Python	
5.	Database	Data Type-String, Numeric	MySQL	
6.	Cloud Database	Database Service on Cloud	IBM	
7.	File Storage	File storage requirements	NIL	
8.	External API-1	Purpose of External API used in the application	NIL	
9.	External API-2	Purpose of External API used in the application	NIL	
10.	Machine Learning Model	Get the data from the user and predict the data with tested and trained dataset models	Data Recognition Model, etc.	
11.	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration:	NIL	

Table-2: Application Characteristics:

S. No	Characteristics	Description	Technology
1.	Open-Source Frameworks	International Business Machines.	Cloud
2.	Security Implementations	Access permission for login page using CAPTCHA	Encryptions
3.	Scalable Architecture	The key of Three tier architecture is improving scalability.	Three Tier architecture
4.	Availability	Load balancer or ADC is the key component that ensures high availability by sending request.	Load balancer
5.	Performance	The system should be able to handle large number of users at the time.	Load balancer

6.2 Sprint Estimation and Delivery Schedule:

PRODUCT BACKLOG, SPRINT DELIVERY, ESTIMATION:

Sprint	Functional requirement (Epic)	User story number	User story andtasks	Story point	priority	Team member
Sprint 1	Data collection	USN 1	Use dataset from Google and clean the dataset	110	High	Meganathan P
Sprint 1	Model	USN 2	Create, test and save the model	10	High	Meganathan P
Sprint2	Display	USN 3	Display user entry form to user	6.7	High	Subhashini B, Dharsanya P
Sprint2	Enter data	USN 4	Receive data from user as numeric values	6.7	High	Subhashini B, Dharsanya P
Sprint2	Enter data	USN 5	Receive data from user as selection from pull down menu	6.7	High	Subhashini B, Pathuman sethupathi R
Sprint 3	Select	USN 6	As a user can select prediction	10	Medium	Pathuman sethupathi R, Meganathan P
Sprint 3	View data	USN 7	As a user can view finalresult	10	Medium	Dharsanya P
Sprint 4	Application building for project	USN 8	Deploy into IBM cloud	20	High	Pathuman sethupathi R

Project tracker, velocity:

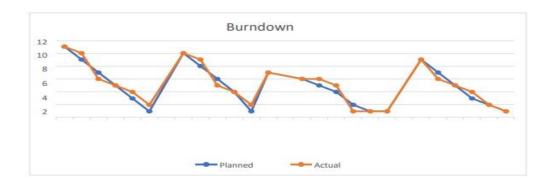
Sprint	Total story points	duration	Sprint start date	Sprint end date (planned)	Story point complete d (as on planned end date)	Sprint release date(actu al)
Sprint 1	20	6 days	24-oct - 2022	29-oct- 2022	20	29-oct- 2022
Sprint 2	20	6 days	31-oct- 2022	05-nov- 2022	20	05-nov- 2022
Sprint 3	20	6 days	07-nov- 2022	12-nov- 2022	20	12-nov- 2022
Sprint 4	20	6 days	14-nov- 2022	19-nov- 2022	20	19-nov- 2022

Velocity:

Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points).

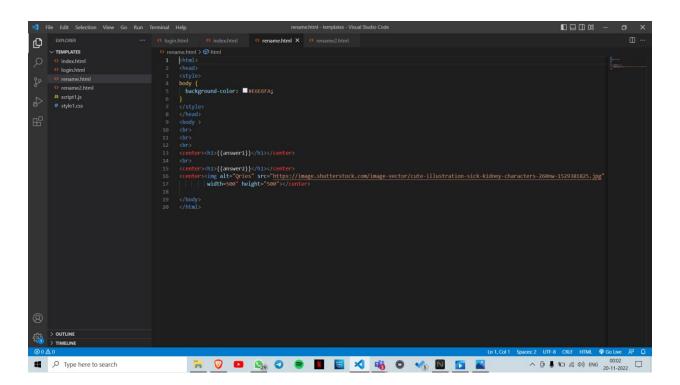
AV= SPRINT DURATION /VELOCITY = 20/10 = 2 AV of CKD Project = 20/6 = 3.33

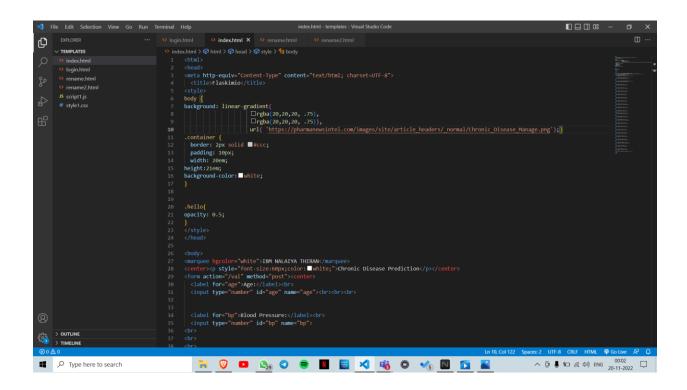
Burndown chart:

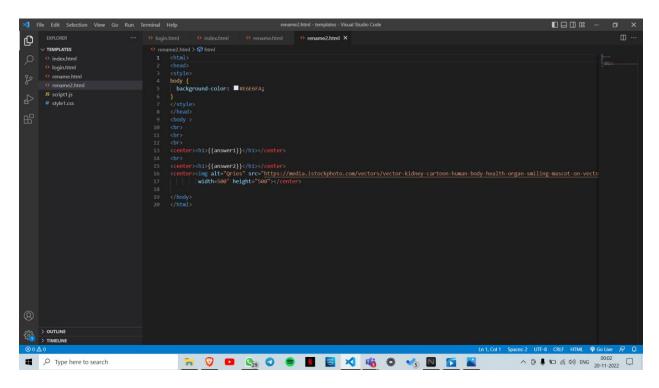


7. CODING & SOLUTIONING

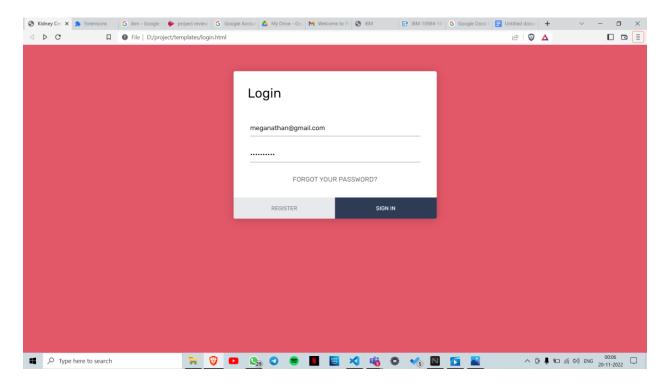
```
| The Last Selection | Vew | Go | Run | New | Ne
```

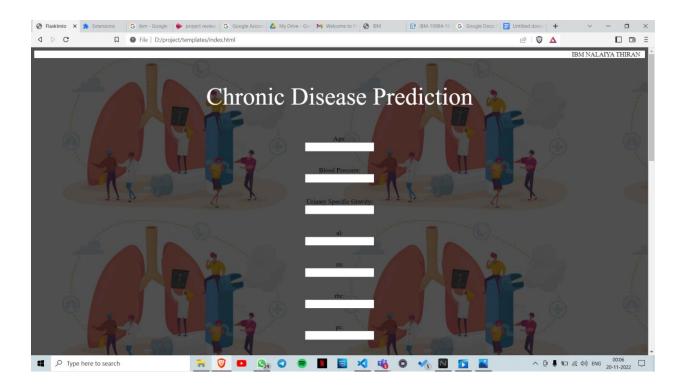


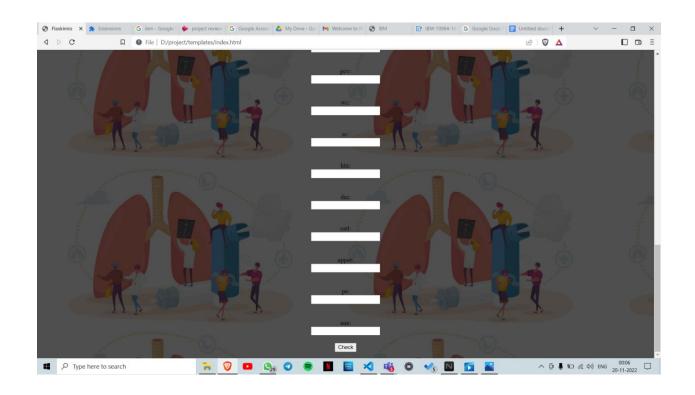


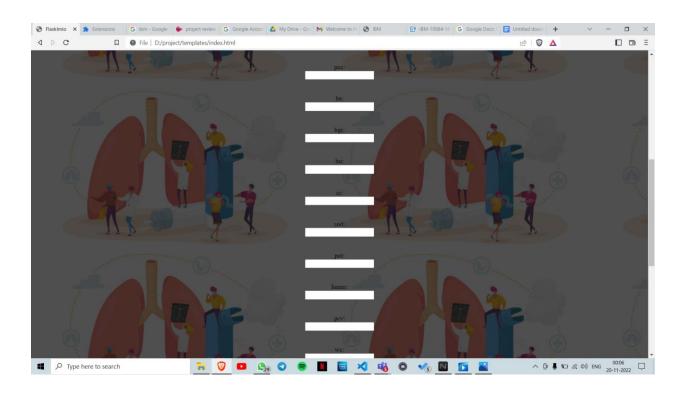


8. TESTING

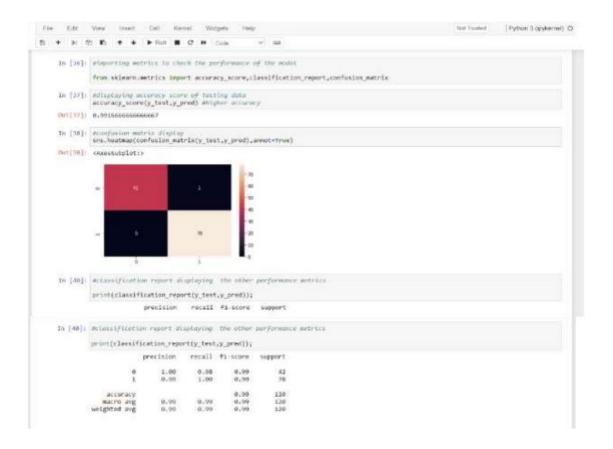








9. RESULTS



10. ADVANTAGES & DISADVANTAGES

Advantages:

Chronic kidney disease (CKD) is one of the most critical health problems due to its increasing prevalence. It is also known as chronic renal disease which is a condition characterized by a gradual loss of kidney function over time. A better testing method which could possibly detect CKD in the early stages would be much more useful using machine learning algorithm

- Greater cost reduction in hospitals for testing
- Helps in early diagnosis of the disease
- Chances of recovery is higher

Disadvantages:

Even Though the CKD prediction model web application consists of a lot of advantages but it comes with certain disadvantages here are some of them .

• Chances of prediction to be wrong for least number of time which can cause problems

- Vast feature in dataset on discovery of time for the disease making the model inefficient to keep up the metrics
- Since its a web application it requires scaling of web application to handle concurrent requests after certain threshold

11.CONCLUSION

Chronic Kidney Disease as the name suggests it's a chronic disease, any chronic disease would make the person miserable and last longer till their livelihood. If in such cases the disease gets unnoticed in early stages which can be cured by medical facilities it's a huge carelessness and risking a person's life. In such cases finding an optimal solution is important, thus there comes the use of a machine learning model for early detection and prediction of the chronic kidney disease which can greatly reduce the potential risk of getting the disease and get cured immediately if it is detected in early stages of the disease. Think of the traditional way of diagnosing kidney disease, it is through blood test, and blood test reports take longer than expected, but blood test is not the only step for diagnosing there are still many more tests taken, which can be time consuming. In those cases the model prediction plays an important role in predicting the disease sooner and faster for the medical team to treat the person if he/she is vulnerable.

Thus early detection of chronic kidney disease is very much necessary in current hospital functioning to diagnose the patient in no time and do necessary treatment to cure if found.

12. FUTURE WORK:

The current work remains the base for the prediction model primarily used by everyone extending from hospitals to normal users .

The future aspects can be as follows:

- subscription based model can be created with initial trial basis.
- Scaling the existing application for simultaneous user to request.

Modifying the model based on adding new feature in the existing dataset based on the hospitals input and standards.

13. Appendix:

https://ieeexplore.ieee.org/abstract/document/8029917

https://iopscience.iop.org/article/10.1088/1742-6596/1255/1/012024/meta

https://start.atlassian.com/

https://ieeexplore.ieee.org/abstract/document/9333572

GITHUB LINK:

https://github.com/IBM-EPBL/IBM-Project-10984-1659250312