

PROJECT REPORT

DOCUMENTATION

EARLY DETECTION OF CHRONIC KIDNEY DISEASE TEAM ID - PNT2022TMID12981

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 & Technical Architecture
- 5.3 User Stories

6. PROJECT PLANNING & SCHEDULING

- 6.1 Sprint Planning
- 6.2 Sprint Estimation and Delivery Schedule

7. CODING & SOLUTIONING

- 7.1 User Registration and Login
- 7.2 Dashboard and Result

8. TESTING

8.1 Test Cases

8.2 Sample Testing

9. RESULTS

10. ADVANTAGES & DISADVANTAGES

11. CONCLUSION

12. FUTURE SCOPE

13. APPENDIX

GitHub Link

1. INTRODUCTION

1.1 Project Overview: -

Chronic kidney disease prediction is one of the most important issues in healthcare-analytics. The most interesting and challenging tasks in day-to-day lives as one third of the adult population is affected by chronic kidney disease (CKD), and millions die each year because they do not have access to affordable treatment. This project is about predicting whether a person has signs of chronic kidney disease at the early stages using machine learning. It takes levels of red blood cells (RBCs), level of pus cells, blood pressure, whether the person has anemia or not, whether it is a coronary artery disease or not, blood glucose level, whether the person has diabetes mellitus or not has the input to predict is the person has early signs of chronic kidney disease. The dataset provided is run on three different machine learning models, namely, KNN, Random Forest and Linear Regression. Accuracy of the three models are used as a measure to compare the performances.

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. Early identification of chronic kidney disease is necessary since the number of individuals with end-stage renal disease is continuously rising (CKD). The largest risk factors for developing chronic kidney disease are diabetes, hypertension, or cardiovascular illness. Renal disease often goes unnoticed until it has progressed to this level in affected persons (CKD). To reduce the prevalence of problems and expenses associated with dialysis, patients with chronic illnesses like hypertension and diabetes need to adopt healthy attitudes and behaviors as well as learn about prevention and early identification of kidney disease. Therefore, CKD is associated with extremely widespread diseases like diabetes mellitus, which causes diabetic nephropathy and has high morbidity and mortality. On the basis of preventive measures, challenges include optimizing diagnosis and prognosis as well as further enhancing therapy

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 :-

SNO	LITERATURE PAPER	AUTHOR	PROPOSED METHOD	ACCURACY	YEAR
1	Computer-Ai ded Diagnosis of Chronic Kidney Disease in Developing Countries: A Comparative Analysis of Machine Learning Techniques	Andressa C. M. Da S. Queiroz, Alvaro Sobrinho, Leandro Dias Da Silva, Evandro De Barros Costa, Maria Eliete Pinheiro, Angelo Perkusich	J48 decision tree is a suitable machine learning technique for such screening in developing countries, due to the easy interpretation of its classification results	95.00%	2020

2	Chronic Kidney Disease Prediction using Machine Learning Models	S.Revathy, B.Bharathi, P.Jeyanthi, M.Ramesh	Decision tree, Random Forest and Support Vector Machine learning models are constructed to carry out the diagnosis of CKD	99.16%	2019
3	Preemptive Diagnosis of Chronic Kidney Disease Using Machine Learning Techniques	Reem A. Alassaf, Khawla A. Alsulaim, Noura Y. Alroomi, Nouf S. Alsharif, Mishael F. Aljubeir, Sunday O. Olatunji, Alaa Y. Alahmadi, Mohammed Imran,	ANN, SVM, Naïve Bayes along with k-NN comparison approach	ANN,SVM ,Naïve Bayes - 98% k-NN - 93.9%	2018

		Rahma A. Alzahrani, Nora S. Alturayef			
4	Prediction of Chronic Kidney Disease Using Machine Learning Algorithm	Siddheshwar Tekale,Pranj al Shingavi,Su kanya Wandhekar, Ankit Chatorikar	Decision tree algorithms along comparison with SVM	Decision tree – 91.75% SVM-96.75%	2018
5	Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study	Njoud Abdullah Almansour,H ajra FahimSyed, Nuha Radwan Khayat,Raw an KanaanAlthe eb,Renad Emad Juri,Jamal Alhiyafi,Sale	Comparative analysis was carried out on the two models-ANN and SVM	ANN - 99.75% SVM - 97.75%	2019

		h Alrashed,Su nday O.Olatunji			
--	--	-------------------------------------	--	--	--

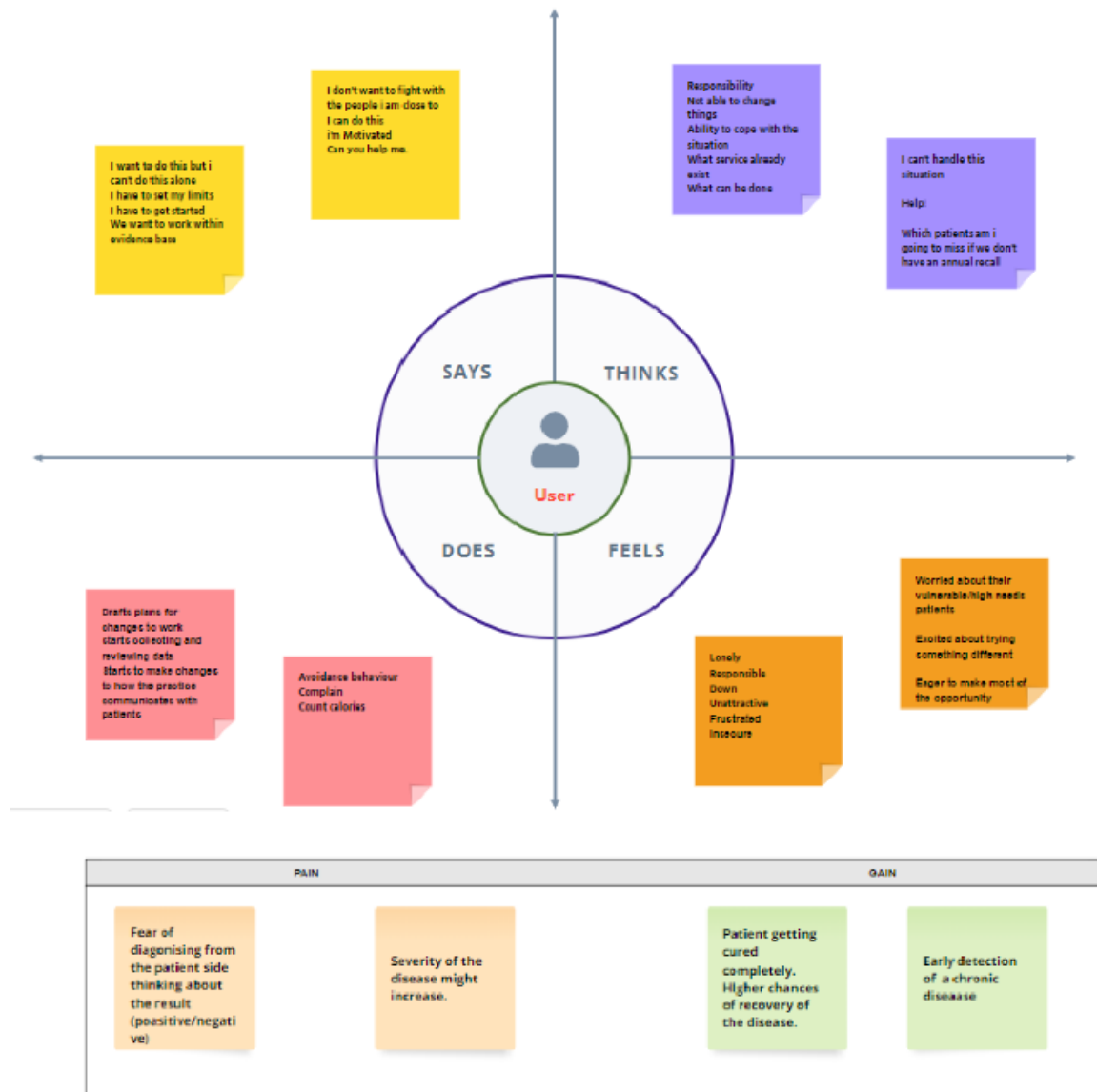
2.3 Problem Statement Definition

Chronic Kidney Disease (CKD) is a serious medical condition that, if diagnosed early enough, is curable. Most individuals are unaware that the various medical tests we undergo for various reasons may provide important information about kidney disorders. As a result, characteristics of numerous medical tests are examined to see which characteristics might contain useful information about the disease. According to the information, doing so enables us to assess the problem's severity, and we utilize this data to create a machine learning model that forecasts chronic kidney disease. If chronic kidney disease is addressed early on, it may be cured. This project's primary goal is to more accurately and quickly identify whether a patient has chronic kidney disease using diagnostic data including Blood Pressure (Bp), Albumin, and other parameters (Al).

3. IDEATION & PROPOSED SOLUTION

3.1. Empathy Map Canvas:-

Empathy Map for Early Detection of Chronic Kidney Disease



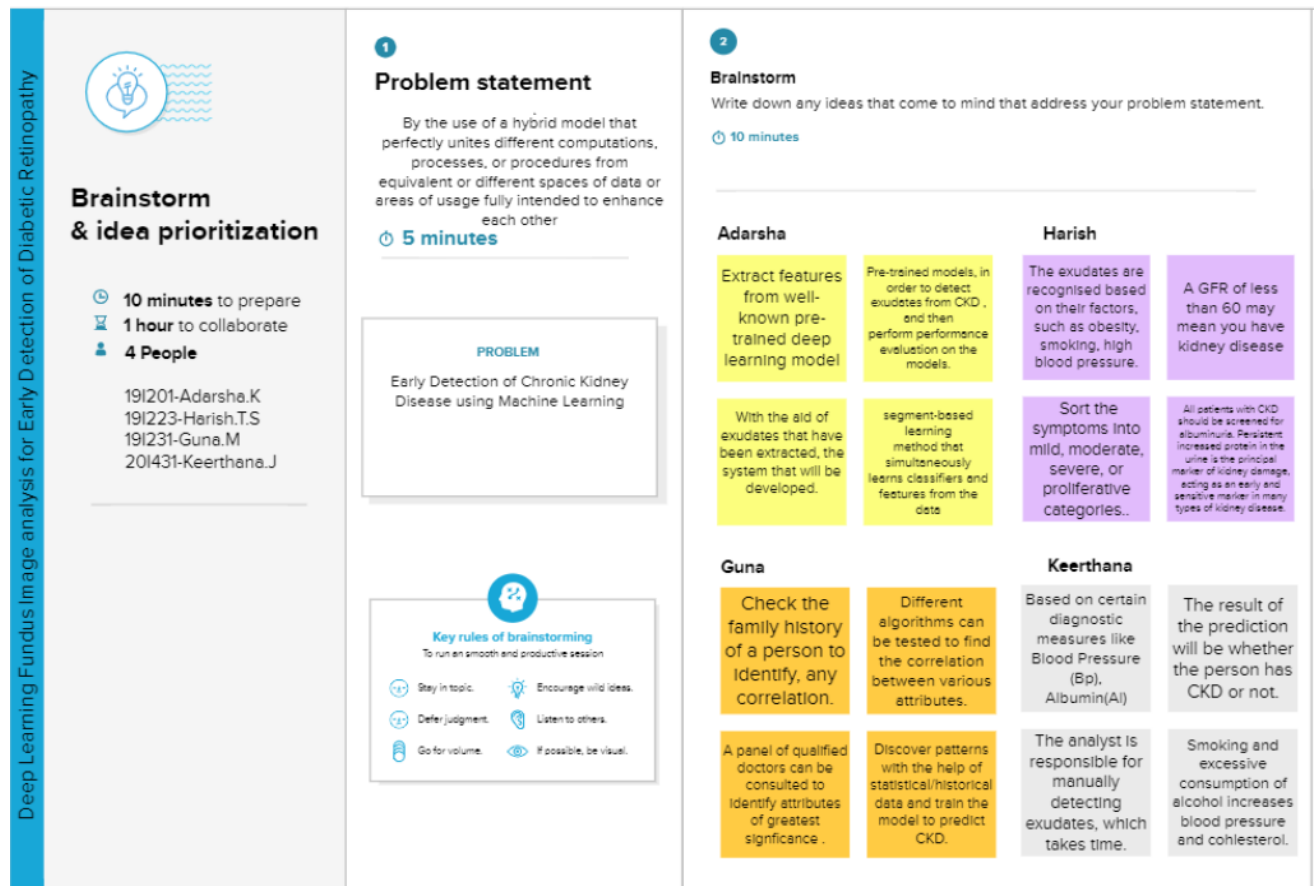
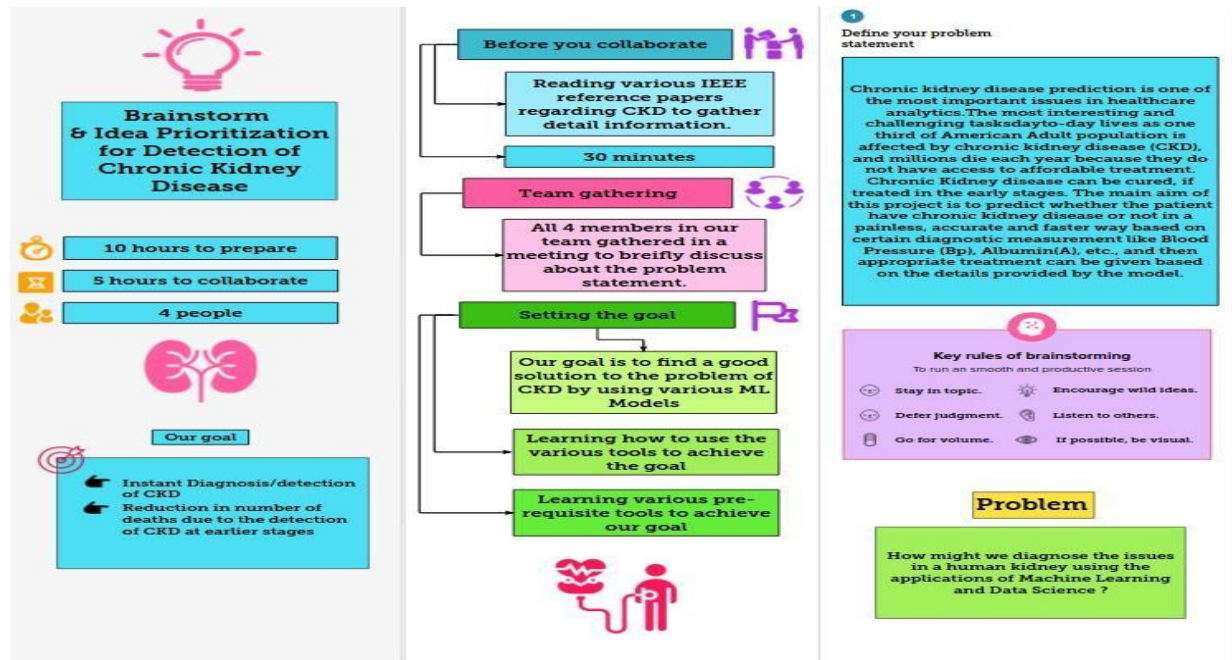
3.2 Ideation & Brainstorming:-

Brainstorming is an activity that helps to 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 ideas and solutions are generated through sessions such as Sketching, Prototyping, Brainstorming, Brainwriting, Worst Possible Idea, and a wealth of other ideation techniques.

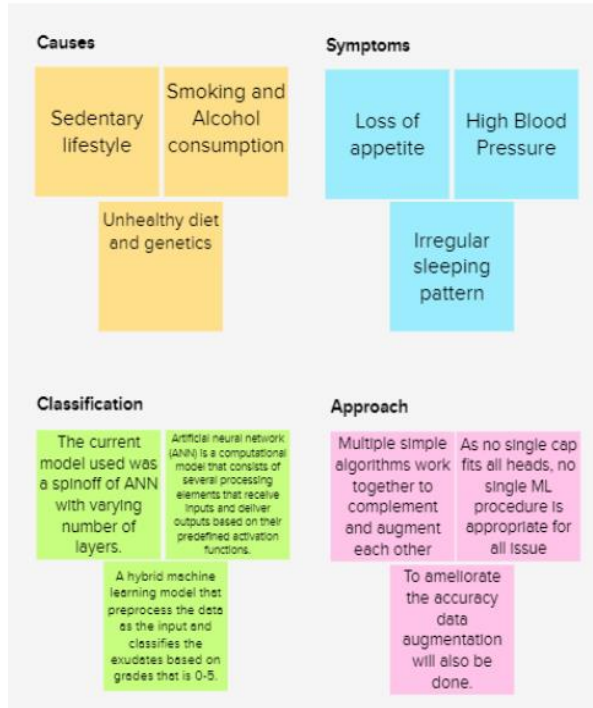


3

Group ideas

Take turns sharing your ideas while clustering similar or related notes as you go. In the last 10 minutes, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you can break it up into smaller sub-groups.

⌚ 20 minutes

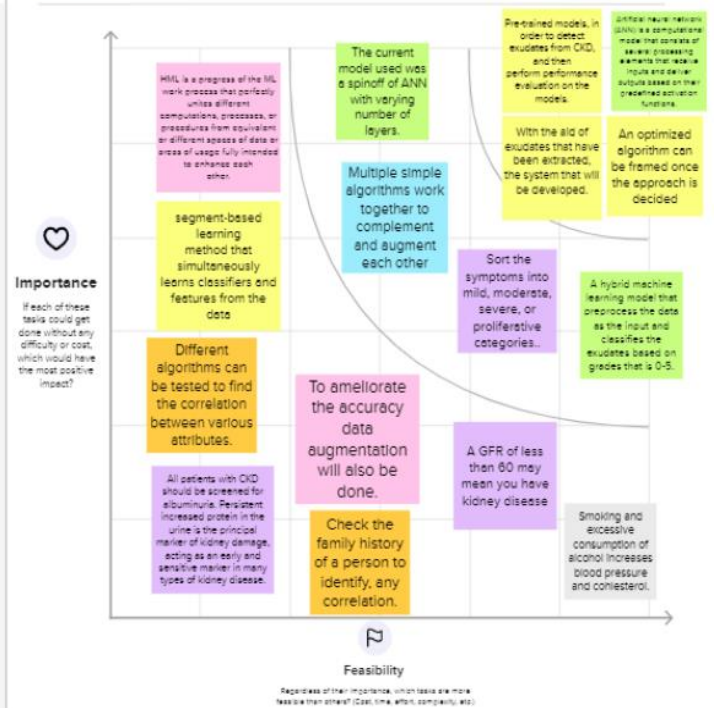


4

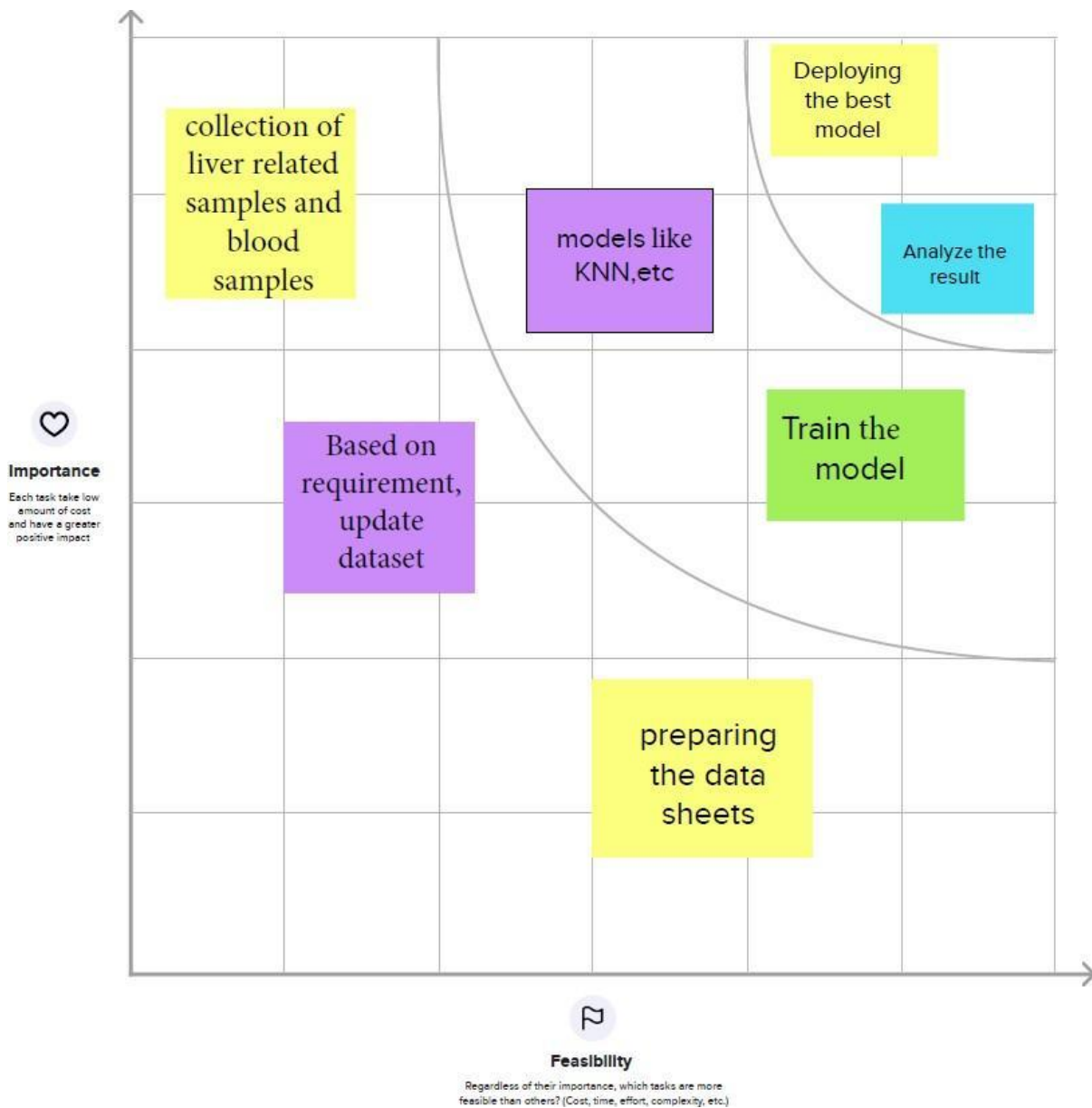
Prioritize

Your team should all be on the same page about what's important moving forward. Place your ideas on this grid to determine which ideas are important and which are feasible.

⌚ 20 minutes



Stage 4 - prioritize



3.3 Proposed Solution:-

NOVELTY: Compared to the conventional method that requires the patient to visit the doctor's office physically to test for chronic kidney disease, this solution using machine learning helps notice any patterns present once information regarding the patient is given. It is not a replacement to a doctor, however, it gives the opportunity for more patients to check if they are likely to have CKD since they can access this tool online. There many symptoms and features that effect the likeliness of having CKD, so using a ml tool that

can track these said symptoms for a large crowd can help establish which symptoms are stronger indicators.

FEASIBILITY OF IDEA: Considering that there is already a large dataset on the internet for chronic kidney disease, it is very feasible to implement this solution. Having a reliable dataset to work with makes it easier to implement a model and train it. There are multiple online tools and virtual GPUs that make it even more easier to compute large volumes of data during the training of the model. Since the data is labelled as well, the testing and validation process will not be very long. Along with the online support tools to use ml models, it is feasible to make the product given the allocated time and resources.

The purpose of this tool is to provide a structured process for identifying a problem, understanding the root causes, ascertaining solution steps, and progress monitoring.

With a solution template, development content can be organized that can be reused for customer-specific solutions. Solution templates enable to easily start the development of customer-specific solutions, for example, for a specific industry.

The term business model refers to a company's plan for making a profit. It identifies the products or services the business plans to sell, its identified target market, and any anticipated expenses. Business models are important for both new and established businesses.

Proposed Solution Template:

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	<p>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.</p> <p>A better testing method which could possibly detect CKD in the early stages would be much more useful using machine learning algorithm</p>
2.	Idea / Solution description	<p>The idea of approaching the problem is by creating a suitable machine learning model which involves deep understanding of the data which needs to be collected from real time , handle the missing data and standardizing the data by preprocessing technique which makes it suitable for ml model training and prediction using different approach of model creation depending on the dataset and output</p>
3.	Novelty / Uniqueness	<ul style="list-style-type: none">• Easy to use User interface (UI)• accurate accuracy by comparing the performance of different ml model technique

4.	Social Impact / Customer Satisfaction	<ul style="list-style-type: none">• Greater cost reduction in hospitals for testing• Helps in early diagnosis of the disease• Chances of recovery is higher
5.	Business Model (Revenue Model)	<ul style="list-style-type: none">• subscription based model with initial trial basis• charges/commission for the actual prediction and recovery of a person
6.	Scalability of the Solution	<ul style="list-style-type: none">• The server in which the app is deployed containing the ml model must be capable of handling concurrent request and handle multiple request

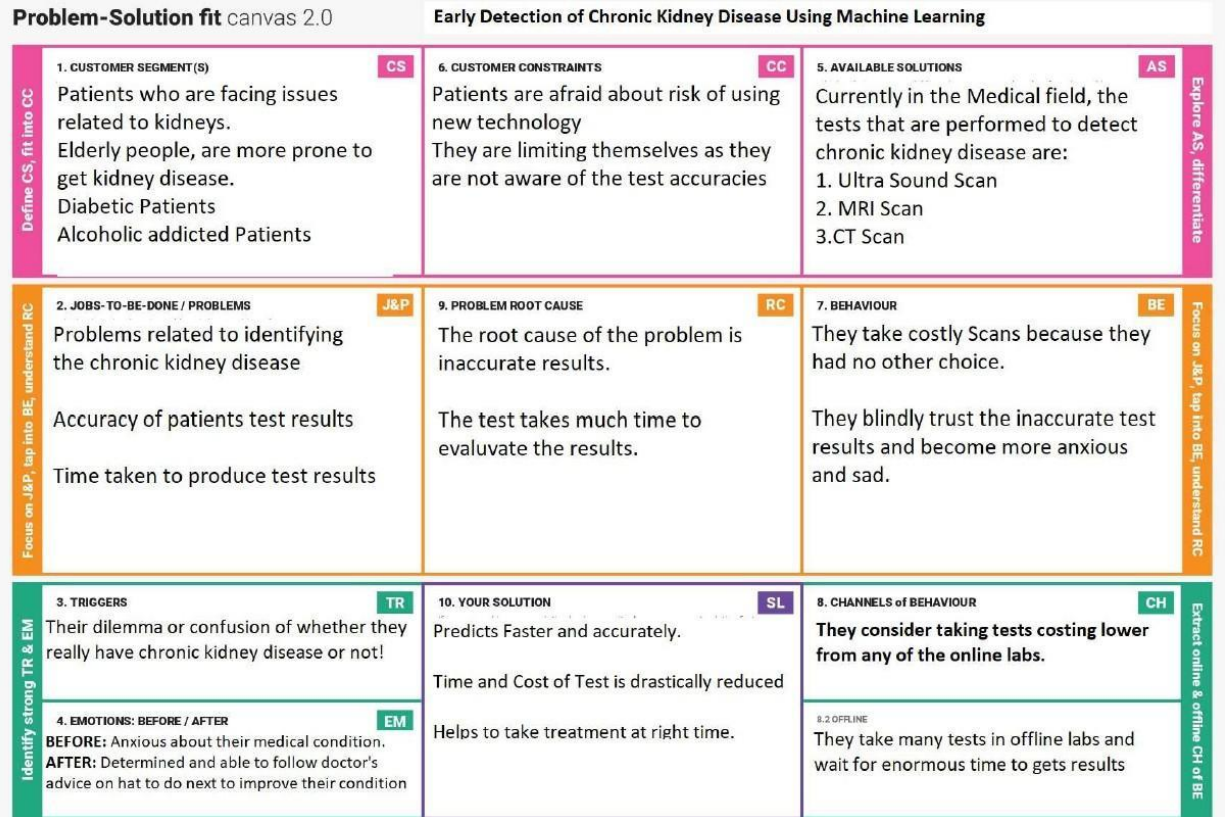
BUSINESS MODEL

Business Model							
Key Partners	Key Activities	Value Propositions	Customer Relationships		Customer Segments		
Our Key partners have a latest testing laboratory to collect samples of blood and urine to analyse various parameters with a testing capacity of almost 1000 persons a day. Our ML Models can test a large number of samples and give accurate output within a short period of time.	Our key activities are to find out whether persons have chronic kidney disease or not.	We are trying to solve chronic kidney disease at an early stage thereby helping them to recover at a faster rate. We are targeting persons liable to get chronic kidney disease using statistical analysis of the collected data.	We need to have a cordial relationship with the persons coming forward for giving blood test. They expect accurate result from our ML model. The cost of building the model may outfit the early detection of chronic kidney disease.		We are creating value for humans. Our customers are common people who have work culture or wrong lifestyle habits which may lead to chronic kidney disease. Customer base will be a mass market as everyone will give importance to health. We will also be providing a diet regulatory counselling to them.		
	Key Resources		Channels				
	Our key resources are testing laboratories, ML Models and their data. To create an excel sheet from the data.		Need to reach our patients through doctors. Using ML model to suggest best practices which can help doctors and patients in avoiding and recovering chronic kidney disease.				
Cost Structure			Revenue Streams				
Most important cost is mainly for data collection through laboratories. This model is cost driven as it has maximum automation with only a few questions needed to be asked to patients. As compared to older methods, ML Detection technique is cost effective.			Patient's visiting doctors are willing to pay for knowing if they have chronic kidney disease or not. Overall revenue is good and it depends on the accuracy of ML model which will help the doctors to do a correct prediction.				

3.4 Problem Solution Fit:-

The Problem-Solution Fit simply means that a problem has been found with the customer and that the solution that they have realized for it actually solves the customer's problem.

Problem-Solution Fit - this occurs when there is evidence that customers care about certain jobs, pains, and gains. At this stage it has been proved the existence of a problem and have designed a value proposition that addresses 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 Help	Report issues through Email
FR-5	Disease prediction	User can predict the disease based on the valid inputs given by him/her

4.2 Non- Functional Requirements:-

Non-functional Requirements:

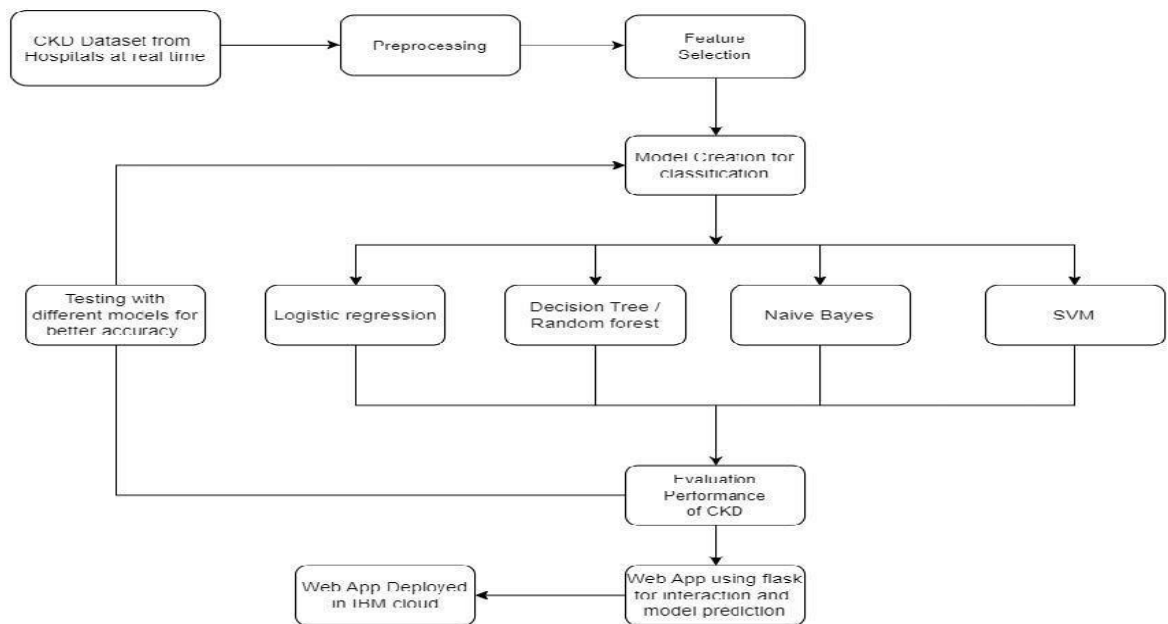
Following are the non-functional requirements of the proposed solution.

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	In terms of accuracy and efficiency, this project is the best for predicting kidney disease
NFR-2	Security	It is highly secured as there is separate user id and password to login
NFR-3	Reliability	It is reliable as it is user friendly and cost-efficient
NFR-4	Performance	In terms of time, this project is highly efficient as we have used very efficient algorithms like regression
NFR-5	Availability	Since it is web based one, it is available to those who have access to web
NFR-6	Scalability	If user demand increases, it can be scaled up easily as we made everything dynamic

5. PROJECT DESIGN

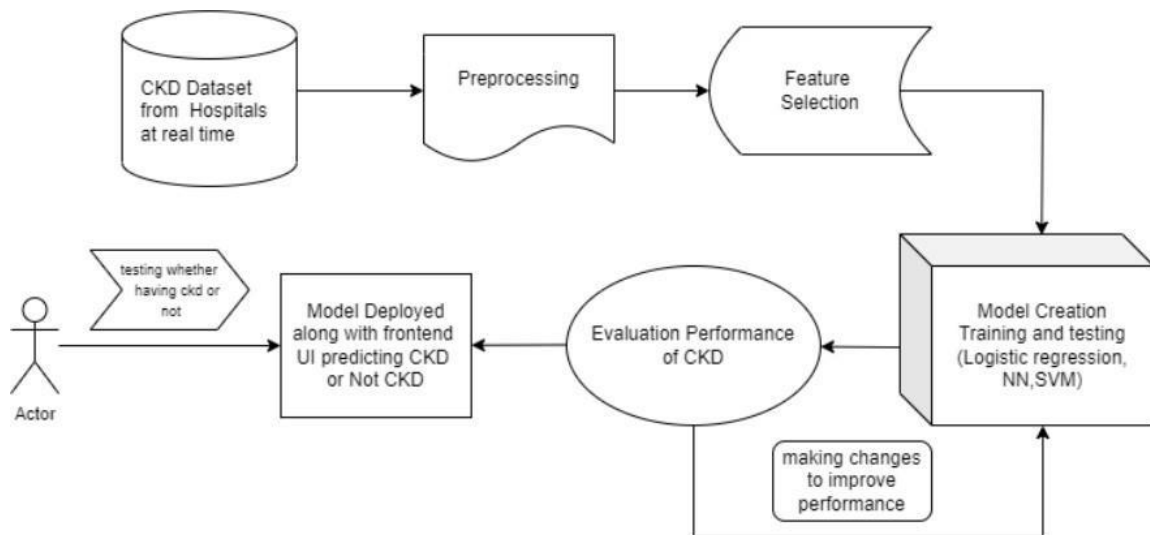
5.1 Data Flow Diagrams:-

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.



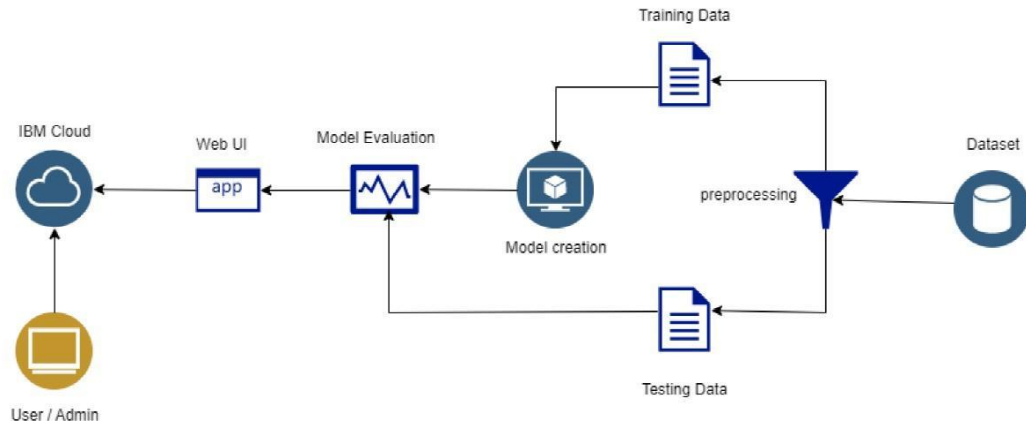
4.2 Solution & Technical Architecture: -

Solution Architecture:



Technical Architecture:

Technical Architecture:



5.3 User Stories :-

User Stories

Use the below template to list all the user stories for the product.

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 log into the application by entering email & password	I am authorized user to avail the web service	High	Sprint-1
	Dashboard	USN-4	As a user, I can navigate and interact with the web app to provide inputs for prediction and testing	I am entitled to enter only valid input for prediction	High	Sprint-1
Customer Care Executive	Assist	USN-5	Collecting the issues and reports from the user through various method of communication	The report or issue must be valid and fully verified	High	Sprint-2
Administrator	Manage	USN-6	Management head controlling all the web services as well as assigning task to improve the service	Complete proper working of web service including security aspect	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 before a sprint, the more likely to achieve the 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 that ensures the efforts are successful.

Technical Architecture:

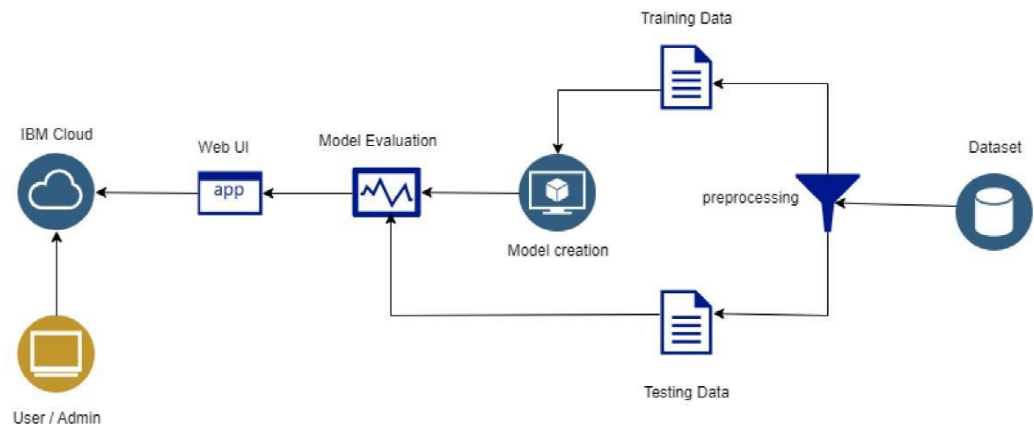


Table-1 : Components & Technologies:

S.No	Component	Description	Technology
1.	User Interface	User Interact with our application through web user interface	HTML, CSS , Flask , <i>React (Subsidiary)</i>
2.	Application Logic-1 (Registration)	User is redirected to register page for registering themselves by providing valid details	HTML, CSS , Flask , <i>React (Subsidiary)</i>
3.	Application Logic-2 (Login)	Once the user is registered he is now able to login to access the web service . There is an external login button to redirect to login page	HTML, CSS , Flask , <i>React (Subsidiary)</i>
4.	Application Logic-3 (Test / Prediction)	The test or prediction page is present for the user who has logged in and can predict the disease by providing input in the form.	HTML, CSS , Flask , <i>React (Subsidiary)</i>
5.	Database	Data Type - String , Numbers	MySQL
6.	Cloud Database	Database Service on Cloud	IBM DB2, IBM Cloudant etc.
7.	File Storage	File storage requirements	IBM Block Storage or Other Storage Service or Local Filesystem
8.	External API-1	Role Based Access is provided for using the API	Backend API
9.	External API-2	Purpose of External API used in the application	NIL
10.	Machine Learning Model	To predict the output based on the training and testing of the data from dataset	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	Frameworks are used in both in making web app and for model creation	Flask (micro web framework) , python (web framework) , Scikit-learn (ML framework)
2.	Security Implementations	passwords are hashed for the user , as well roles are provided for access based control system	SHA 256
3.	Scalable Architecture	The Scalability can be improvised by using three-tier architecture	Three tier architecture
4.	Availability	Scalability includes availability,the service must be available even if there are more user request ,load balancer is needed to do the above task	Load Balancer
5.	Performance	Performance is key for increased revenue , handling multiple requests and expanding it can be done using Load Balancer.	Load Balancer

6.2 Sprint Estimation and Delivery Schedule:

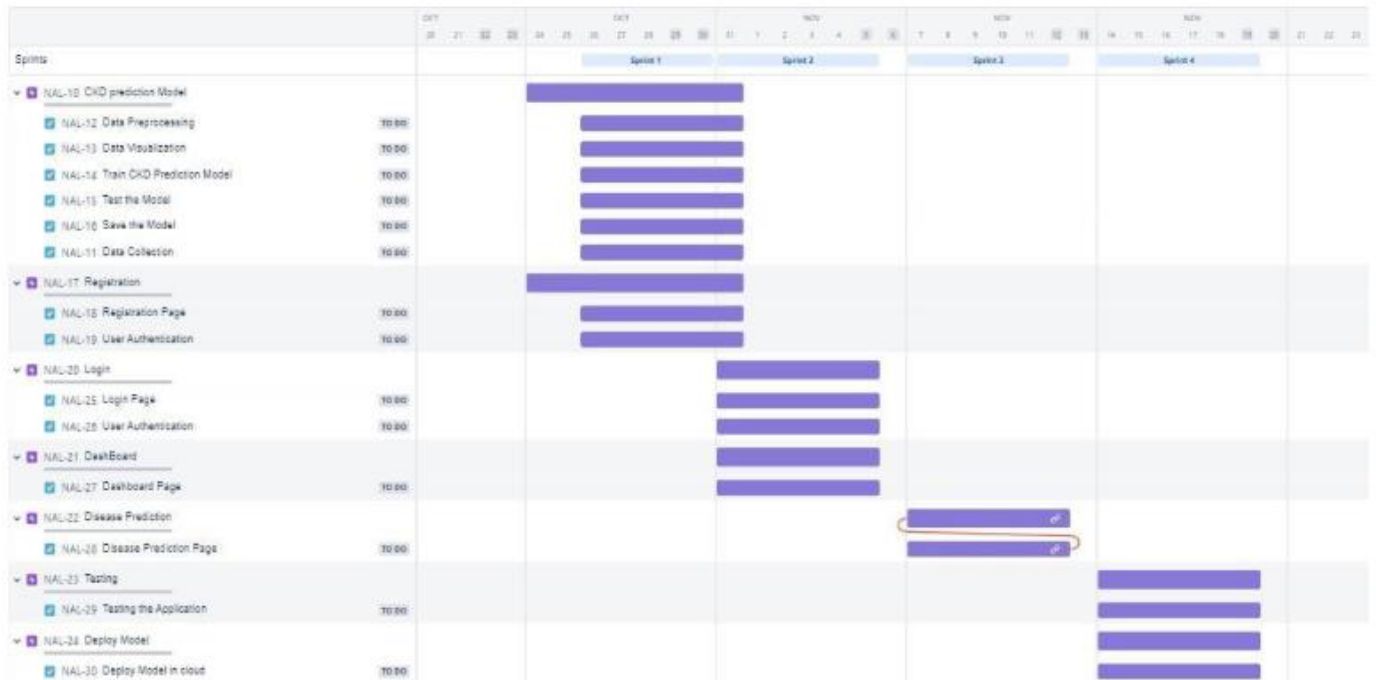
A sprint estimation shows how much effort a series of tasks require. It's based on assumptions, requirements, and dependencies of a project

Product Backlog, Sprint Schedule, and Estimation (4 Marks)

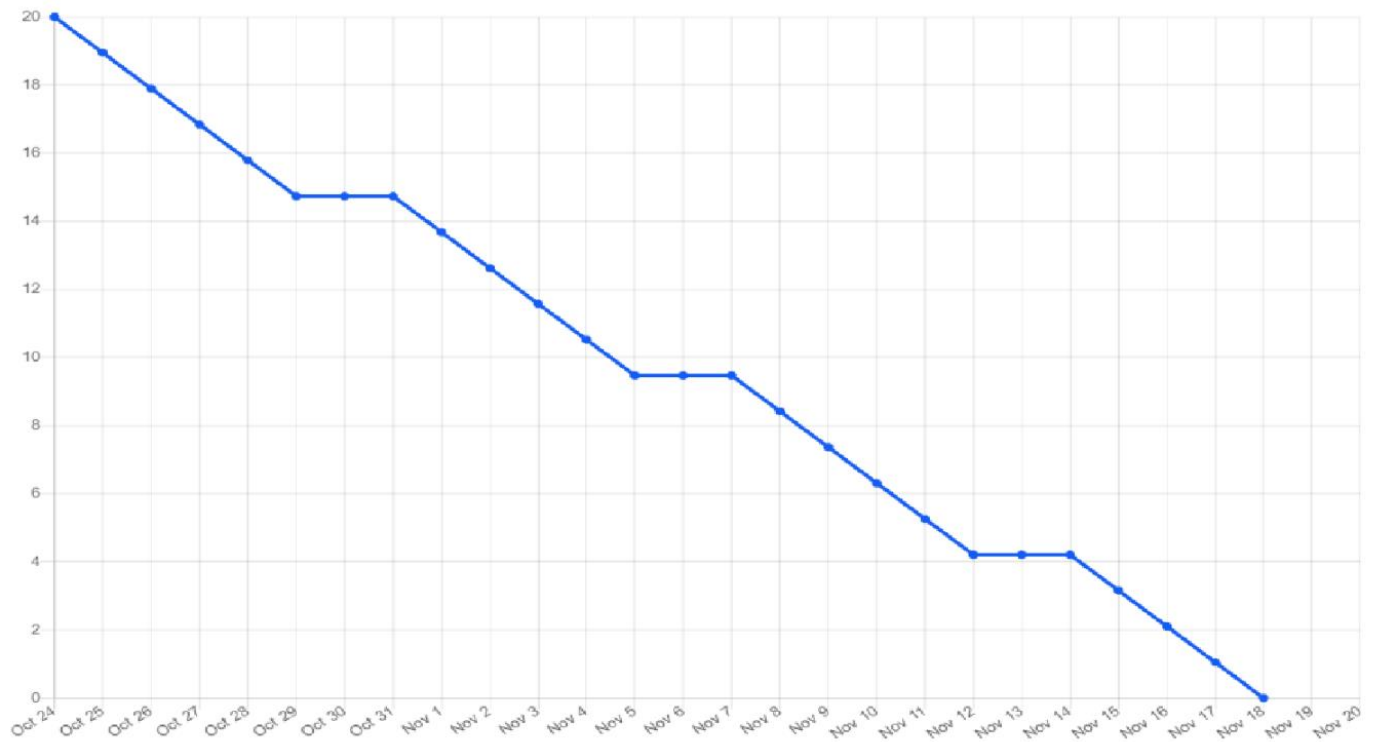
Use the below template to create product backlog and sprint schedule

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	New user enters into the System He/ She can register into the Application by entering mail Id and Password.	8	High	Team Lead
Sprint-1		USN-2	The user will receive conformation Email	5	High	Team Member -1
Sprint-1	Login	USN-3	After Successful registration the user can Log into the application by entering the registered Mail Id and Password	8	Low	Team Member- 2
Sprint-2	Dashboard	USN-4	User can get into the Dashboard only when the Verification Successful. After the user can access the displayed information in the Dashboard.	8	Medium	Team Member-3
Sprint-2	Diagnosis Form	USN-5	As a user, I must enter my pre-diagnostic test results to give as required.	10	High	Team Lead, Team Member-1
Sprint-3	Report	USN-6	As a user, I can view the report generated by the tool	7	High	Team Member- 3


Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-4	Quality Assurance	USN-8	As a user they have some credibility issues while using application.	8	High	Team Member- 2
Sprint-3	Train Model	USN-9	As an administrator, I must use the most suitable ML model for detection of CKD	9	High	Team lead, Team Memebr-1



Burndown Chart



6.3 Reports from JIRA:

 Jira Software

Your work ▾

Projects ▾


Filters ▾

Dashboards ▾


People ▾


Apps ▾

Create


 **CKD predictor**
Software project


PLANNING


 Roadmap

 Board

DEVELOPMENT

 Project pages



 Add shortcut

 Project settings

You're in a team-managed project
[Learn more](#)

Projects / CKD predictor

CP

TO DO 7 ISSUES

USN-3 Register using GMail account
☒ CP-3

USN-5 View Dashboard
☒ CP-5

USN-7 View diagnosis report
☒ CP-7

USN-8 Remedies for diagnosis
☒ CP-8

USN-9 Customer Q&A
☒ CP-9

USN-10 Customer feedback
☒ CP-10

USN-12 Select optimal model
☒ CP-12

IN PROGRESS 3 ISSUES

USN-11 Feature selection for model
☒ CP-11

USN-2 Confirmation Email on registration
☒ CP-2

USN-6 Entry form for pre-diagnostic test
☒ CP-6


DONE 2 ISSUES ✓

USN-1 Registration Page
☒ CP-1 ✓

USN-4 Login page for registered users
☒ CP-4 ✓

7. CODING & SOLUTIONING

7.1 User Registration and login :

 CKD Predictor [Login](#) [Register](#)

Login


Username

Password

[Sign in](#)

Not yet registered? [Register](#)

7.2 Dashboard and Result

 CKD Predictor [Hi! Guna 🙌](#) [Dashboard](#) [Result](#) [Logout](#)

Your Test Result is: **Positive (Treatment Required)**

Your Past Test Results

Date	Age	BP	SG	AL	SU	RBC	PC	PCC	BA	BGR	BU	SC	SOD	POT	HEMO	PCV	WC	RC	HTN	DM	CAD	APPET	PE	ANE	Result
2022-11-23T19:57	6767.0	67.0	7.0	6.0	76.0	7	6	7	6	766.0	76.0	77.0	6.0	76.0	76.0	7.0	6.0	76.0	7	6	7	6	6	76	Positive (Treatment Required)
2022-11-23T19:57	6767.0	67.0	7.0	6.0	76.0	7	6	7	6	766.0	76.0	77.0	6.0	76.0	76.0	7.0	6.0	76.0	7	6	7	6	6	76	Positive (Treatment Required)
2022-11-23T19:57	6767.0	67.0	7.0	6.0	76.0	7	6	7	6	766.0	76.0	77.0	6.0	76.0	76.0	7.0	6.0	76.0	7	6	7	6	6	7676	Positive (Treatment Required)
2022-11-23T19:57	6767.0	67.0	7.0	6.0	76.0	7	6	7	6	766.0	76.0	77.0	6.0	76.0	76.0	7.0	6.0	76.0	7	6	7	6	6	7676	Positive (Treatment Required)
2022-11-19T14:15	14.0	34.0	34.0	232.0	13.0	464	34	46	24	35.0	24.0	35.0	2.0	2.0	424.0	24224.0	24265.0	6.0	24	64	543	3443	23	2	Positive (Treatment Required)
2022-11-19T14:15	12.0	34.0	34.0	232.0	13.0	464	34	46	24	35.0	24.0	35.0	2.0	2.0	424.0	24224.0	24265.0	6.0	24	64	543	3443	23	2	Positive (Treatment Required)

8. TESTING

8.1 Test cases

Pre-Requisite	Steps To Execute	Test Data	Expected Result	Actual Result	Status	Comments	TC for Automation(Y/N)
Kaggle	1.Enter into kaggle website 2.Download the dataset	https://www.kaggle.com/	Download the Dataset	Working as expected	Pass		NO
Anaconda prompt , Jupyter Notebook	1.Enter Anaconda prompt 2.Enter Jupyter Notebook & do Data pre-processing		Pre-processing the dataset using machine learning Algorithm	Working as expected	pass		NO
Anaconda prompt , Jupyter Notebook	1.Enter Anaconda prompt 2.Enter Jupyter Notebook & do Model Building	Model building using logistic regression	Build a Machine Learning Model	Working as expected	pass		NO
Visual Studio Code	1.Click on VS code ,create html pages . Run html pages on app.py by using live server .	Run a website in localhost server http://127.0.0.1:5000/	Appears a Prediction page on local host server	Working as expected	pass		NO
Visual Studio Code	Click on the http link Enter the values as in the dataset Click on submit	Gives prediction result as patient have CKD or NOT http://127.0.0.1:5000/predict	Predict the Result	Working as expected	Pass		NO
	1.Enter IBM Cloud using login credentials 2.Use jupyter notebook in IBM	Deploy the project in IBM CLOUD	Application should show same resut as vs code flask integration				

Sample tests:

Dashboard - CKD Predictor

127.0.0.1:5000/dashboard

CKD Predictor | Hi! yash33 | Dashboard | Result | Logout

Chronic Kidney Disease Prediction

Provide your test details below 🗨 :

17-11-2022 21:04

41	80
1.02	0
0	normal
abnormal	present
notpresent	122
32	0.8
139	5

CKD® Predictor with

Your Test Result is: **Positive**

Your Past Test Results

Date	Age	BP	SG	AL	SU	RBC	PC	PCC	BA	BGR	BU	SC	SOD	POT	HEMO	PCV	WC	RC	HTN	DM	CAD	APPET	PE	ANE	Result
2022-11-17 12:04	41.0	80.0	1.02	0.0	0.0	normal	abnormal	present	notpresent	122.0	32.0	0.8	139.0	5.0	14.1	41.0	9100.0	5.2	yes	yes	no	good	no	no	Positive
2022-11-13 11:53	41.0	80.0	1.02	0.0	0.0	normal	normal	notpresent	notpresent	122.0	25.0	0.8	138.0	5.0	17.1	41.0	9100.0	5.2	no	no	no	good	no	no	Negative

CKD Predictor with ❤️

8.2 User Acceptance Testing (UAT)

1. Purpose of Document

The purpose of this document is to briefly explain the test coverage and open issues of the [Early Detection of Chronic Kidney Disease] project at the time of the release to User Acceptance Testing (UAT).

2. 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	Subtotal
By Design	3	1	1	1	6
Duplicate	4	0	2	0	6
External	2	2	0	1	5
Fixed	1	1	1	1	4
Not Reproduced	0	0	0	0	0
Skipped	0	0	0	0	0
Won't Fix	0	0	0	0	0
Totals	10	4	4	3	21

3. Test Case Analysis

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

Section	Total Cases	Not Tested	Fail	Pass
Home Screen	1	0	0	1
User Input	3	0	0	3
Chronic Kidney Disease testing	2	0	0	2
No Chronic Kidney Disease testing	2	0	0	2
Version Control	2	0	0	2

9. RESULTS

The application is evaluated based on metrics to verify the correct working of the application against the business requirements.

Accuracy is the simple ratio between the number of correctly classified points to the total number of points.

Confusion Matrix is a summary of predicted results in specific table layout that allows visualization of the performance measure of the machine learning model for a binary classification problem (2 classes) or multi-class classification problem (more than 2 classes).

9.1 Performance Metrics (Random Forest Classifier)

```
print(classification_report(y_test,y_pred));
```

	precision	recall	f1-score	support
0	1.00	0.98	0.99	42
1	0.99	1.00	0.99	78
accuracy			0.99	120
macro avg	0.99	0.99	0.99	120
weighted avg	0.99	0.99	0.99	120

9.2 Performance Metrics (Logistic Regression)

```
print(classification_report(y_test,y_pred1))
```

	precision	recall	f1-score	support
0	0.90	0.86	0.88	42
1	0.93	0.95	0.94	78
accuracy			0.92	120
macro avg	0.91	0.90	0.91	120
weighted avg	0.92	0.92	0.92	120

9.3 Performance Metrics (Decision Tree)

```
print(classification_report(y_test,y_pred5))
```

	precision	recall	f1-score	support
0	0.98	1.00	0.99	42
1	1.00	0.99	0.99	78
accuracy			0.99	120
macro avg	0.99	0.99	0.99	120
weighted avg	0.99	0.99	0.99	120

9.4 Performance Metrics (Gaussian NB)

```
print(classification_report(y_test,y_pred3))
```

	precision	recall	f1-score	support
0	0.93	1.00	0.97	42
1	1.00	0.96	0.98	78
accuracy			0.97	120
macro avg	0.97	0.98	0.97	120
weighted avg	0.98	0.97	0.98	120

9.5 Performance Metrics (Support Vector machine)

```
print(classification_report(y_test,y_pred4))
```

	precision	recall	f1-score	support
0	0.00	0.00	0.00	42
1	0.65	1.00	0.79	78
accuracy			0.65	120
macro avg	0.33	0.50	0.39	120
weighted avg	0.42	0.65	0.51	120

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:

- Sobrinho, A., Queiroz, A. C. D. S., Da Silva, L. D., Costa, E. D. B., Pinheiro, M. E., & Perkusich, A. (2020). Computer-aided diagnosis of chronic kidney disease in developing countries: A comparative analysis of machine learning techniques. *IEEE Access*, 8, 25407-25419.
- Revathy, S., Bharathi, B., Jeyanthi, P., & Ramesh, M. (2019). Chronic kidney disease prediction using machine learning models. *International Journal of Engineering and Advanced Technology*, 9(1), 6364-6367.
- Almansour, N. A., Syed, H. F., Khayat, N. R., Altheeb, R. K., Juri, R. E., Alhiyafi, J., ... & Olatunji, S. O. (2019). Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study. *Computers in biology and medicine*, 109, 101-111.
- Tekale, S., Shingavi, P., Wandhekar, S., & Chatorikar, A. (2018). Prediction of chronic kidney disease using machine learning algorithm. *International Journal of Advanced Research in Computer and Communication Engineering*, 7(10), 92-96.

GITHUB LINK :

<https://github.com/IBM-EPBL/IBM-Project-12906-1659499575>

DEMO LINK:

<https://youtu.be/IjcyFUTL8zs>