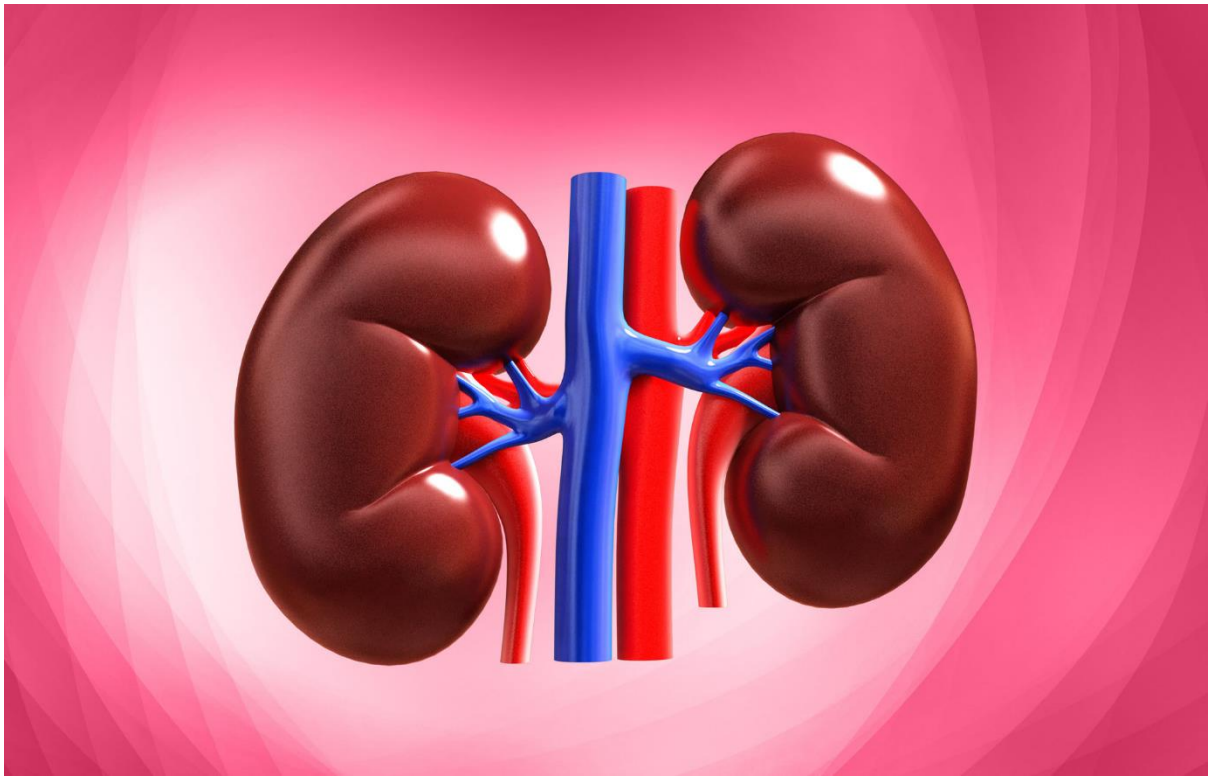
IBM PROJECT
EARLY DETECTION OF CHRONIC KIDNEY DISEASE USING MACHINE LEARNING



TEAM ID PNT2022TMID04324

TEAM MEMBERS DIVYA S
DHARANI R
DHANVARSINI P
DEEPIKA P

INDEX

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 & Estimation Delivery Schedule
6.2 Sprint
6.3 Reports from JIRA

7) CODING & SOLUTIONING (Explain the features added in the project along with code)

7.1 Feature 1
7.2 Feature 2
7.3 Database Schema (if Applicable)

8) TESTING

8.1 Test Cases
8.2 User Acceptance Testing

9) RESULTS

9.1 Performance Metrics

10) ADVANTAGES & DISADVANTAGES

11) CONCLUSION

12) FUTURE SCOPE

13) APPENDIX

Source Code

GitHub & Project Demo Link

1) INTRODUCTION TO PROJECT

1.1)Project Overview:

Early kidney disease usually doesn't have any symptoms. Testing is the only way to know how well your kidneys are working. Get checked for kidney disease If you have diabetes, get checked every year. If you have high blood pressure, heart disease, or a family history of kidney failure, talk with your health care provider about how often you should get tested. The sooner you know you have kidney disease, the sooner you can get treatment to help protect your kidneys.A blood test that checks how well your kidneys are filtering your blood, called GFR. GFR stands for glomerular filtration rate.A urine test to check for albumin. Albumin is a protein that can pass into the urine when the kidneys are damaged. If you have kidney disease, your health care provider will use the same two tests to help monitor your kidney disease and make sure your treatment plan is working.Your health care provider will use a blood test to check your kidney function. The results of the test mean the following:a GFR of 60 or more is in the normal range. Ask your health care provider when your GFR should be checked again.a GFR of less than 60 may mean you have kidney disease. Talk with your health care provider about how to keep your kidney health at this level.a GFR of 15 or less is called kidney failure. Most people below this level need dialysis or a kidney transplant. Talk with your health care provider about your treatment options.

1.2)Purpose:

Early detection might help prevent kidney disease from progressing to kidney failure. If you have a medical condition that increases your risk of kidney disease, your doctor may monitor your blood pressure and kidney function with urine and blood tests during office visits.Chronic kidney disease, also called chronic kidney failure, involves a gradual loss of kidney function. Your kidneys filter wastes and excess fluids from your blood, which are then removed in your urine. Advanced chronic kidney disease can cause dangerous levels of fluid, electrolytes and wastes to build up in your body.In the early stages of chronic kidney disease, you might have few signs or symptoms. You might not realize that you have kidney disease until the condition is advanced.Treatment for chronic kidney disease focuses on slowing the progression of kidney damage, usually by controlling the cause. But, even controlling the cause might not keep kidney damage from progressing. Chronic kidney disease can progress to end-stage kidney failure, which is fatal without artificial filtering (dialysis) or a kidney transplant.Chronic kidney disease can affect almost every part of your body. Potential complications include:Fluid retention, which could lead to swelling in your arms and legs, high blood pressure, or fluid in your lungs (pulmonary edema) Pregnancy complications that carry risks for the mother and the developing fetus.Irreversible damage to your kidneys (end-stage kidney disease), eventually requiring either dialysis or a kidney transplant for survival.To reduce your risk of developing kidney disease:Follow instructions on over-the-counter medications**.** When using nonprescription pain relievers, such as aspirin, ibuprofen (Advil, Motrin IB, others) and acetaminophen (Tylenol, others), follow the instructions on the package. Taking too many pain relievers for a long time could lead to kidney damage.

2) LITERATURE SURVEY

2.1)EXISTING PROBLEMS:

Chronic kidney disease has been neglected as a public health issue despite its significant burden. Worldwide, there are in excess of 850 million people living with kidney disease, of which 700 million have CKD. 4 5 As a comparison, 463 million people globally live with diabetes, the prevalence of cancer worldwide is 85.8 million and 36.8 million people are currently living with AIDS/HIV. CKD disproportionally affects disadvantaged populations and reduces the number of productive years of life. The disease can have major adverse consequences for the living conditions of affected individuals and their families, including severe financial difficulties. Many who progress to dialysis face a potential loss of employment. 4 Furthermore, the prospect of financial burden discourages many patients from treatment, thereby leading to preventable morbidity and death.

Related Work

The objective variable of the study in is the resource consumption such as medical and long-term care expenses and a predictive model for medical care using a random forest machine learning algorithm . This method uses data of more than 100 pieces that includes preventive activities, clinical tests, and medical practices. This model uses mean decrease Gini for classification and for regression mean square error (MSE) is used . The training model uses a grid search for hyperparameter tuning and is validated using *K*-fold cross-validation. Along with the objective variable, exploratory variables such as age, gender, and analysis period are also included, since the aim of this paper is proper management of the budget for medical care . A review that highlights the applications of machine learning techniques in various medical practices such as predicting, diagnosing, and prognosis of diseases such as multiple sclerosis, autoimmune chronic kidney disease, autoimmune rheumatic disease, and inflammatory bowel disease and for the selection of treatments and stratification of patients; drug development; drug repurposing; target interpretation; and validation has been given in . This paper also provides a detailed description of the challenges faced by the machine learning approaches such as the need for quality data in preparation of robust models, external model validation using the independent data set, difficulties faced during implementation of a model, and ethical concerns. A predictive model for chronic kidney disease is explained in .

2.2.REFERENCES

1)Neural network and support vector machine for the prediction of chronic kidney disease:

A comparative study this paper aims to assist in the prevention of Chronic Kidney Disease (CKD) by utilizing machine learning techniques to diagnose CKD at an early

stage. Kidney diseases are disorders that disrupt the normal function of the kidney. As the percentage of patients affected by CKD is significantly increasing, effective prediction pro-cedures should be considered. In this paper, we focus on applying different machine learning classificationalgorithms to a dataset of 400 patients and 24 attributes related to diagnosis of chronic kidney disease.He empirical results from the experiments indicated that ANN performed better than SVM, with accuracies of 99.75% and 97.75%, respectively, indicating that theoutcome of this study is very promising.

2) Chronic Kidney Disease Prediction Using Machine Learning Models

The field of biosciences have advanced to a largerextent and have generated large amounts of information from Electronic Health Records.Data mining methods and machine learning play a major role in this aspect of biosciences. Chronic Kidney Disease(CKD) is a condition in which the kidneys are damaged and cannot filter blood as they always do. A family history of kidney diseases orfailure, high blood pressure, type 2 diabetes may lead to CKD.This is a lasting damage to the kidney and chances of getting worser by time is high. The very common complications that results due to a kidney failure are heart diseases, anemia, bone diseases, high potasium and calcium. The worst case situation leads to complete kidney failure and necessitates kidney transplant to live. An early detection of CKD can improve thequality of life to a greater extent.

3) A Machine Learning Methodology for Diagnosing Chronic Kidney Disease

Chronic kidney disease (CKD) is a global health problem with high morbidity and mortality rate, and it induces other diseases. Since there are no obvious symptoms during the early stages of CKD, patients often fail to notice the disease. Early detection of CKD enables patients to receive timely treatment to ameliorate the progression of this disease. Among these machine learning models, random forest achieved the best performance with 99.75% diagnosis accuracy. By analyzing the misjudgments generated by the established models, we proposed an integrated model that combines logistic regression and random forest by using perceptron, which could achieve an average accuracy of 99.83% after ten times of simulation. Hence, we speculated that this methodology could be applicable to more complicated clinical data for disease diagnosis.

4)Prediction of Chronic Kidney Disease - A Machine Learning Perspective

Chronic Kidney Disease is one of the most critical illness nowadays and proper diagnosis is required as soon as possible. Machine learning technique has become reliable for medical treatment. With the help of a machine learning classifier algorithms, the doctor can detect the disease on time. From the results, it is marked that LSVM with penalty L2 is giving the highest accuracy of 98.86% in synthetic minority over-sampling technique with full features. Along with accuracy, precision, recall, F-measure, area under the curve and GINI coefficient have been computed and compared results of various algorithms have been shown in the graph. Least absolute shrinkage and selection operator regression selected features with synthetic minority over-sampling technique gave the best after synthetic minority over-sampling technique with full features. In the synthetic minority over-sampling technique with least absolute

shrinkage and selection operator selected features, again linear support vector machine gave the highest accuracy of 98.46%. Along with machine learning models one deep neural network has been applied on the same dataset and it has been noted that deep neural network achieved the highest
accuracy of 99.6%.

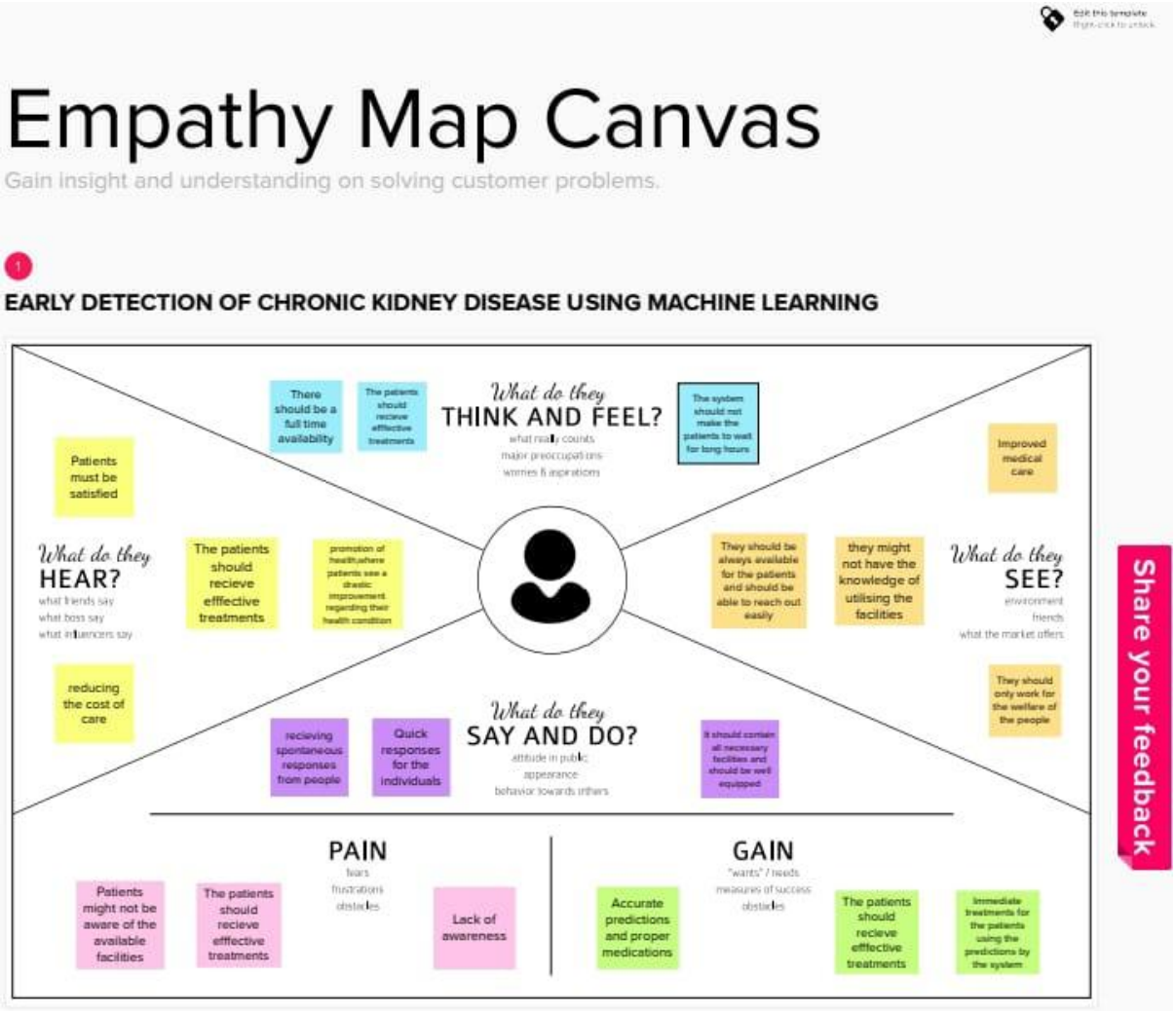
PROBLEM STATEMENT DEFINITION:

I am	The customer group here includes the people who are affected with chronic kidney disease (CKD). People with high blood pressure, diabetes are most likely to be affected with CKD.
I'm trying to	The people affected with CKD are concerned more about getting cured at early stages itself without making things complicated. They aim at getting best treatment with cost effective methods and preventing further damages to their kidneys
But	Some people are not aware of CKD and it does not show any early symptoms, even if people notice some abnormalities, they neglect them. Hence only after getting their medical tests they come to know that they have CKD. Some medical tests are also expensive because of which people belonging to different economies are not able to afford.
Because	The main reason is that people are unaware of the symptoms and the disease. Hence, it becomes difficult for people to get cured at earlier stages itself and this becomes a root cause for not being able to detect at early stages.
Which makes me feel	As a result, people feel anxious and insecure about their health conditions and after knowing that they are affected with CKD the must take appropriate treatment

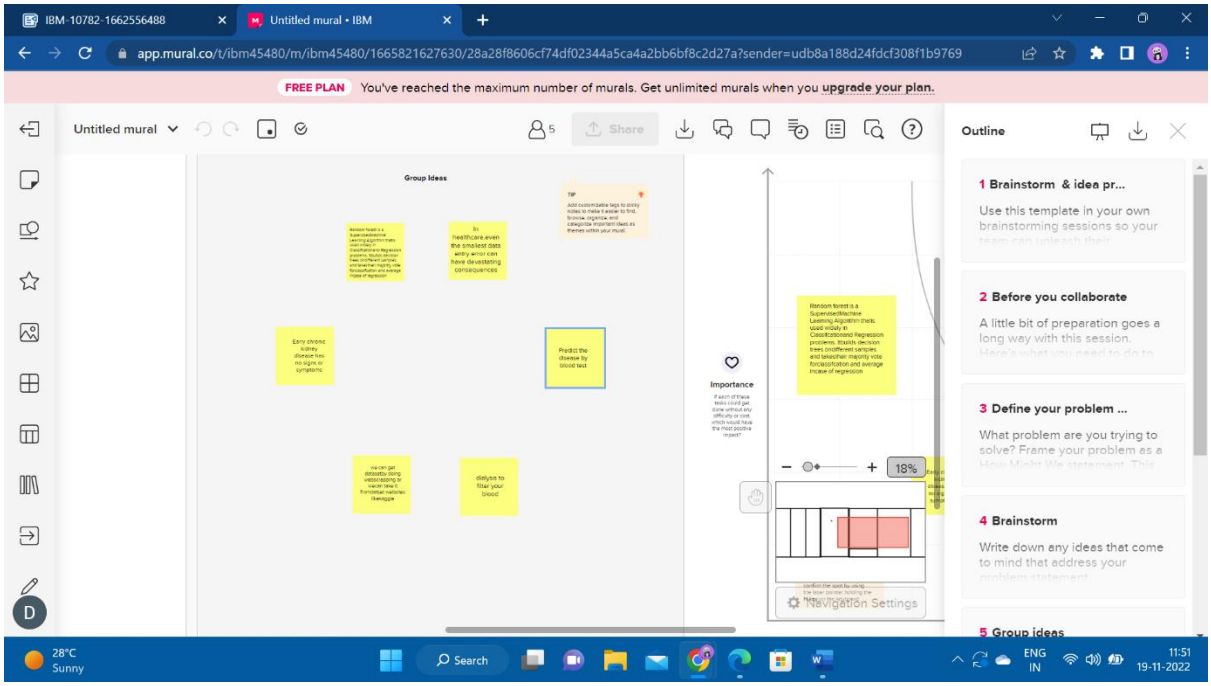
	to suppress the progression of this disease. So, it is better if people get to know about their health conditions at early stages itself.

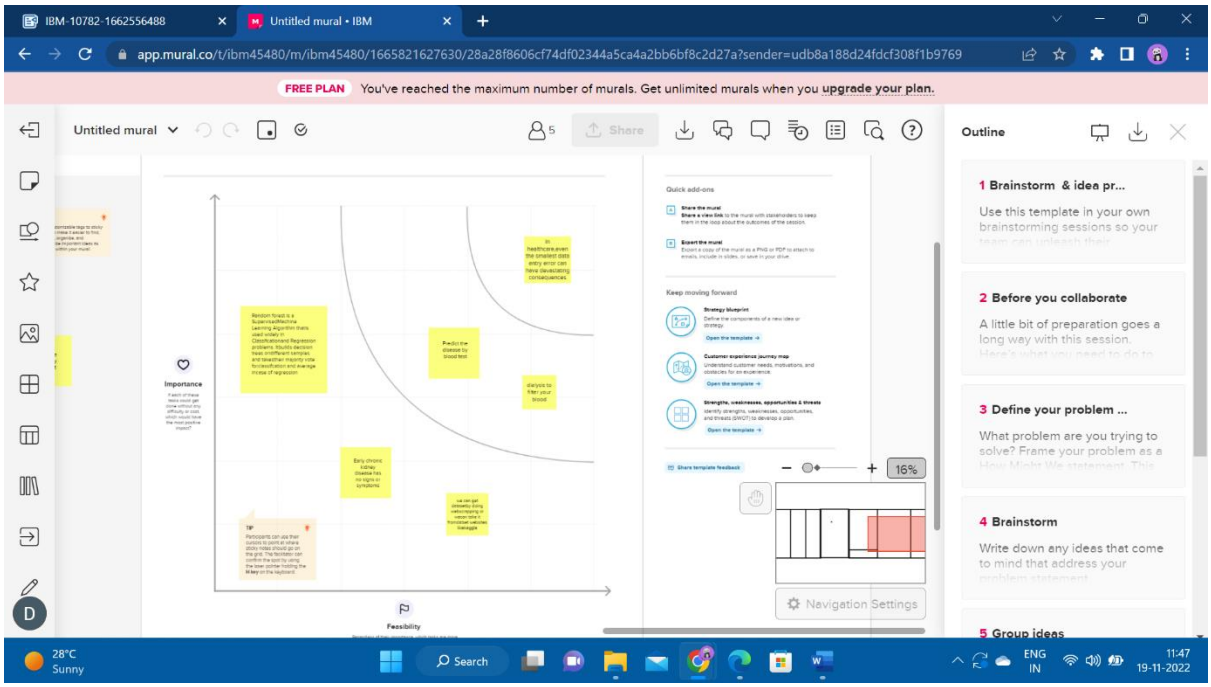
3) IDEATION AND PROPOSED SOLUTION:

3.1) EMAPATHY MAP CANVAS:



3.2) IDEATION AND BRAINSTROMING:


IBM-10782-1662556488
Untitled mural • IBM
app.mural.co/t/ibm45480/m/ibm45480/1665821627630/28a28f8606cf74df02344a5ca4a2bb6bf8c2d27a?sender=udb8a188d24fdcf308f1b9769
FREE PLAN
You've reached the maximum number of murals. Get unlimited murals when you upgrade your plan.
Untitled mural
Share
Outline
1 Brainstorm & idea pr...
Use this template in your own brainstorming sessions so your
2 Before you collaborate
A little bit of preparation goes a long way with this session.
3 Define your problem ...
What problem are you trying to solve? Frame your problem as a
4 Brainstorm
Write down any ideas that come to mind that address your
5 Group ideas
Group Ideas
Importance
18%
Navigation Settings
28°C
Sunny
Search
ENG
IN
11:51
19-11-2022


IBM-10782-1662556488
Untitled mural • IBM
app.mural.co/t/ibm45480/m/ibm45480/1665821627630/28a28f8606cf74df02344a5ca4a2bb6bf8c2d27a?sender=udb8a188d24fdcf308f1b9769
FREE PLAN
You've reached the maximum number of murals. Get unlimited murals when you upgrade your plan.
Untitled mural
Share
Outline
1 Brainstorm & idea pr...
Use this template in your own brainstorming sessions so your
2 Before you collaborate
A little bit of preparation goes a long way with this session.
3 Define your problem ...
What problem are you trying to solve? Frame your problem as a
4 Brainstorm
Write down any ideas that come to mind that address your
5 Group ideas
Importance
Feasibility
Quick add-ons
Keep moving forward
16%
Navigation Settings
28°C
Sunny
Search
ENG
IN
11:47
19-11-2022

3.3) PROPOSED SOLUTION

S.NO	Parameter	Description
1	Problem Statement (Problem to be solved)	Detection of CKD (chronic kidney disease) at an early stage, so that the issue does not get progressed over a period of time. Detecting the disease at an early stage so that the patients can be provided with appropriate treatment without developing any complications in their health conditions.
2	Idea / Solution description	So, the basic idea is that the real time data of the patients regarding their sugar levels and other prominent measures like GFR (Glomerular Filtration Rate), protein and creatinine levels etc., should be collected from the hospital records. Using these data, we will build an efficient ML (machine learning) model. First the collected data should be devoid of missing values hence pre-processing is done using various methods. Next, to identify only the key features of the dataset and to remove the inefficient and irrelevant features dimensionality reduction is performed. Finally based on the above-mentioned steps a ML model can be built to detect the presence of CKD and the results of the predicted model can be utilized for further treatments for the patients according to the level of severity in their health conditions.
3	Novelty / Uniqueness	The ML model provides high level of accuracy on comparing the performance with other ML models. When it comes to healthcare and management the proposed methodologies should provide an easy interface to the customer which is often described as user-friendly products.
4	Social Impact / Customer Satisfaction	Since people are not aware of the medical test attributes which concerns the kidney disease, using this business model solution the useful medical test attributes are gathered separately and henceforth it reduces the cost of by not repeating the same medical tests for detecting the kidney disease. As the kidney disease is also predicted at a very early stage it does not make things complicated for doctors to treat the patients and patients also get cured just by simple medications and regular doctor visits.
5	Business Model (Revenue Model)	It can be used for collaboration with healthcare sectors wherein they collect fees from the patients for various amenities provided by the hospitals. Generating charges for the early predictions of kidney disease and easy recovery of the patients. The proposed models can be used for getting

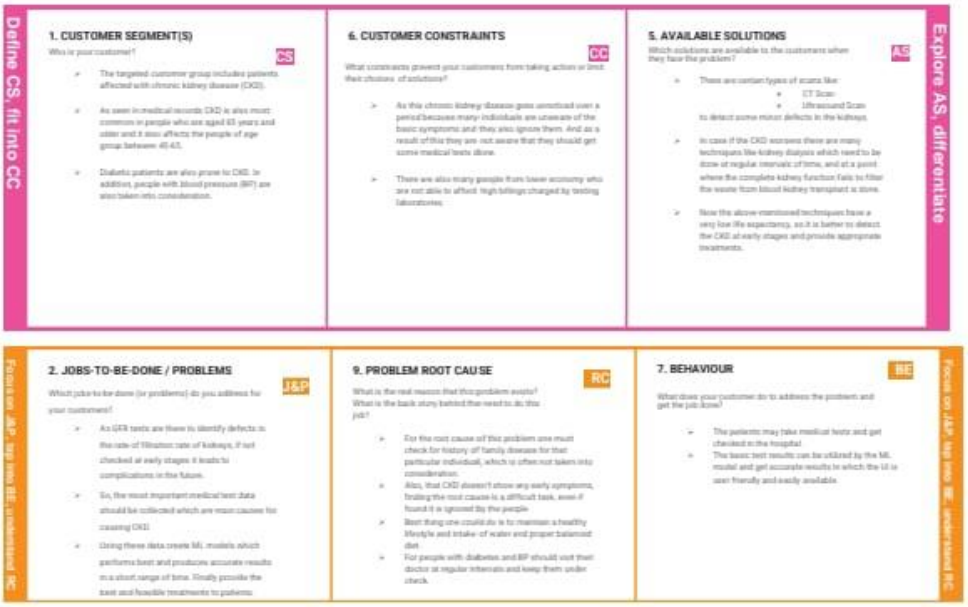
		direct income from the patients as their medical expenses as it produces intermediate outcomes.
6	Scalability of the Solution	Changes can be made at any point of time in the process of developing any solution hence addition of features in the future should be taken into consideration and it should not affect the existing one. On deploying the model on any server, it should be able to handle multiple requests from the clients. In consideration of the future aspects of patients it can be used to diagnose a wider variety of chronic disorders that the patients develop.

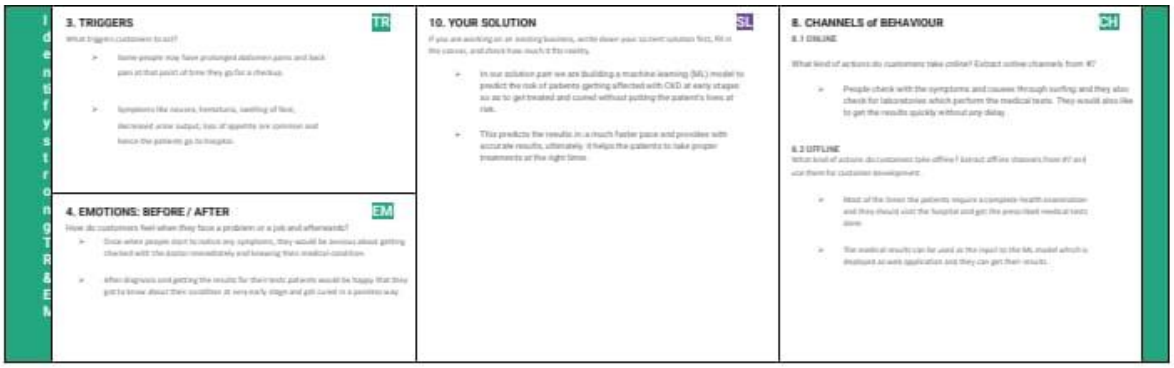
3.4) PROBLEM SOLUTION FIT:

Project Title: Early Detection of Chronic Kidney Disease using Machine Learning

Project Design Phase-I – Solution Fit Template

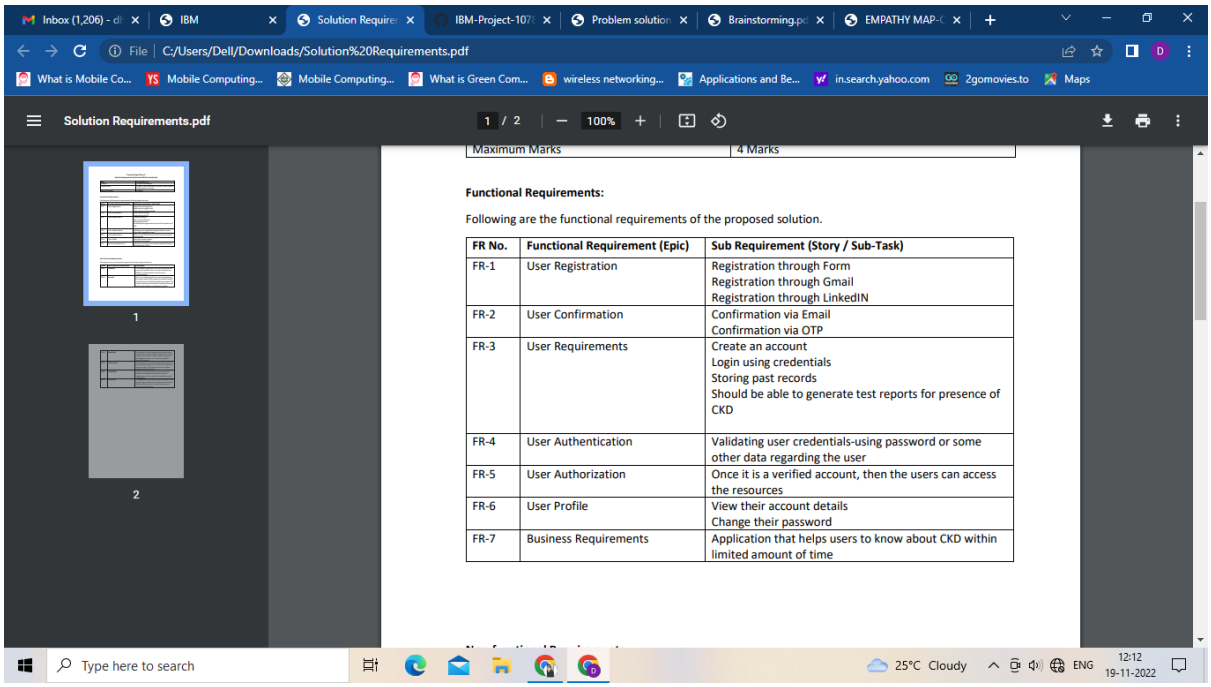
Team ID: PNT2022TMID04324





4) REQUIREMENT ANALYSIS

4.1)FUNCTIONAL REQUIREMENT



Solution Requirements.pdf

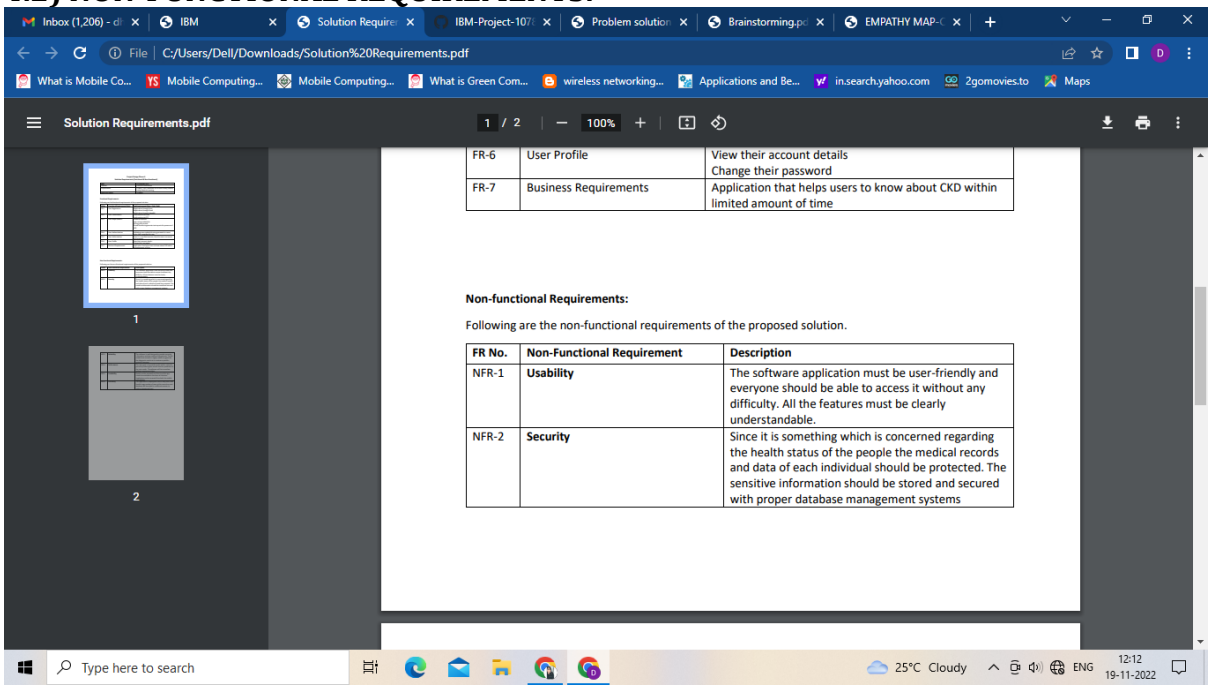
Maximum Marks	4 Marks

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 Form Registration through Gmail Registration through LinkedIN
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	User Requirements	Create an account Login using credentials Storing past records Should be able to generate test reports for presence of CKD
FR-4	User Authentication	Validating user credentials-using password or some other data regarding the user
FR-5	User Authorization	Once it is a verified account, then the users can access the resources
FR-6	User Profile	View their account details Change their password
FR-7	Business Requirements	Application that helps users to know about CKD within limited amount of time

4.2) NON-FUNCTIONAL REQUIREMENTS:



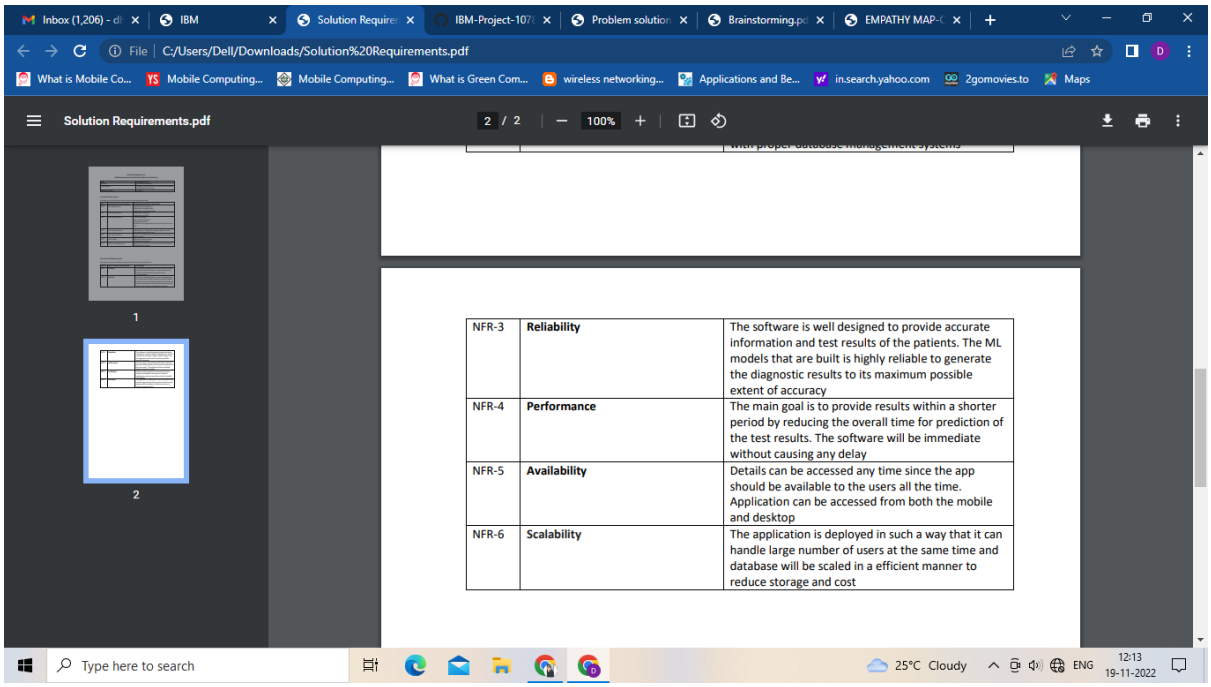
Solution Requirements.pdf

FR-6	User Profile	View their account details Change their password
FR-7	Business Requirements	Application that helps users to know about CKD within limited amount of time

Non-functional Requirements:

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

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The software application must be user-friendly and everyone should be able to access it without any difficulty. All the features must be clearly understandable.
NFR-2	**Security**	Since it is something which is concerned regarding the health status of the people the medical records and data of each individual should be protected. The sensitive information should be stored and secured with proper database management systems

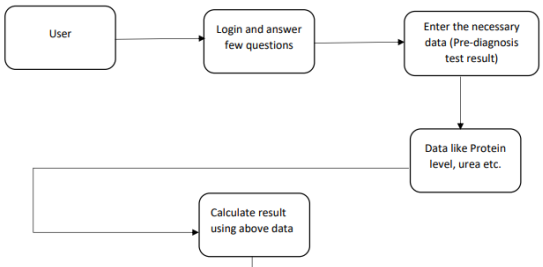


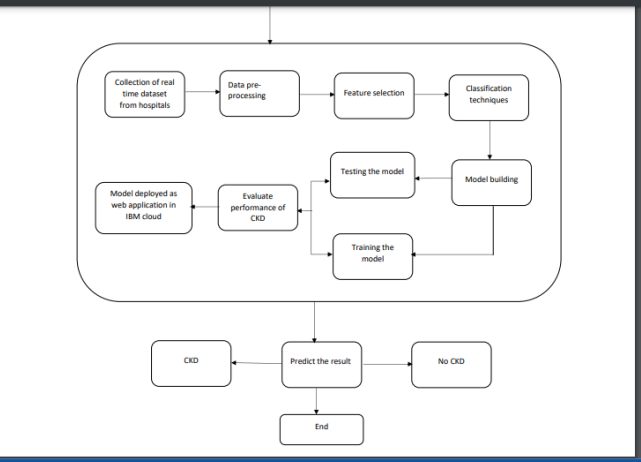
5)PROJECT DESIGN

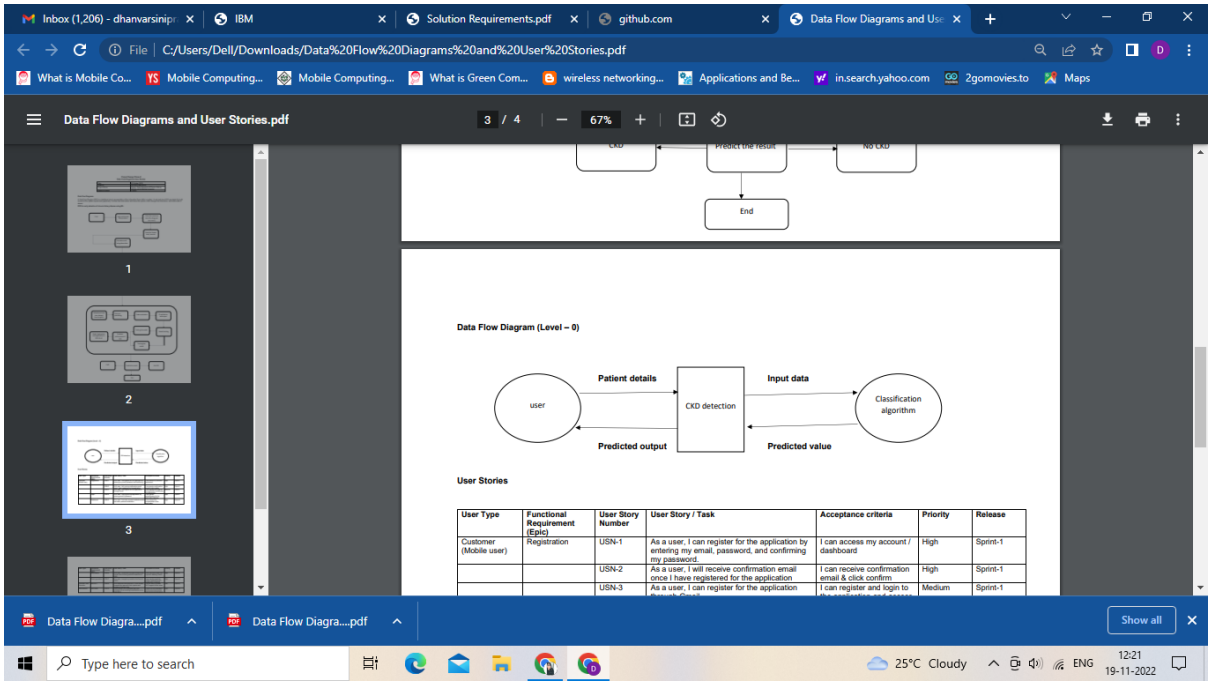
5.1) DATA FLOW DIAGRAM:

amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

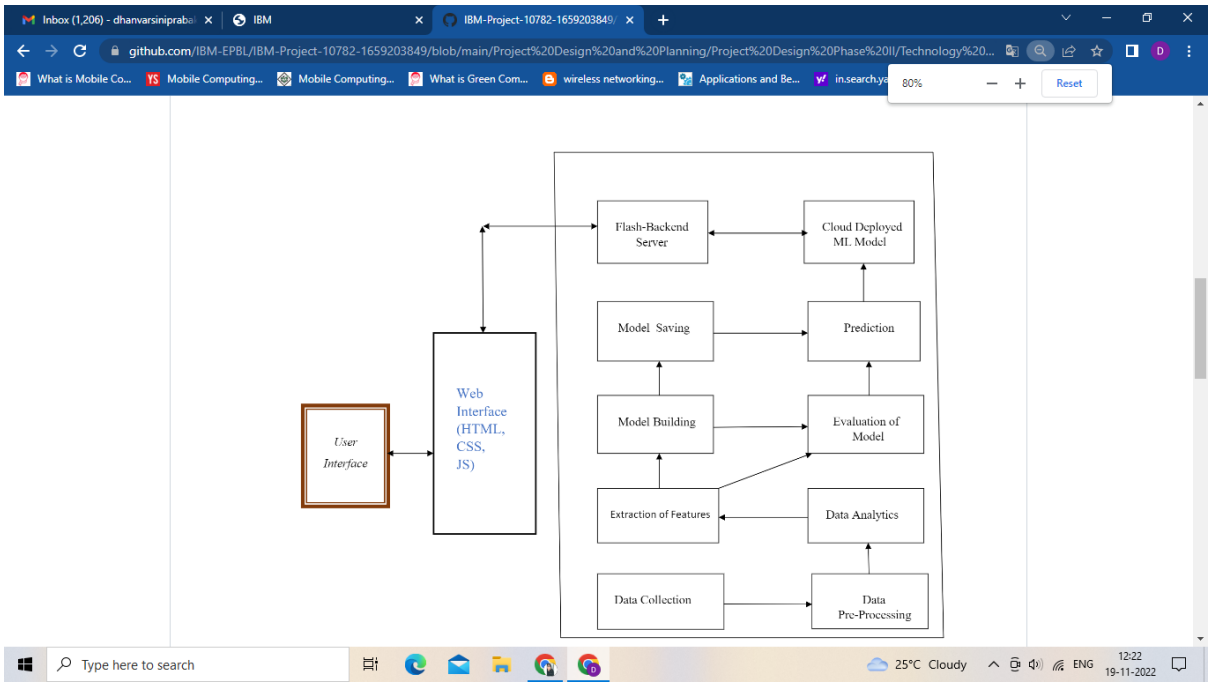
DFD for early detection of chronic kidney disease using ML:



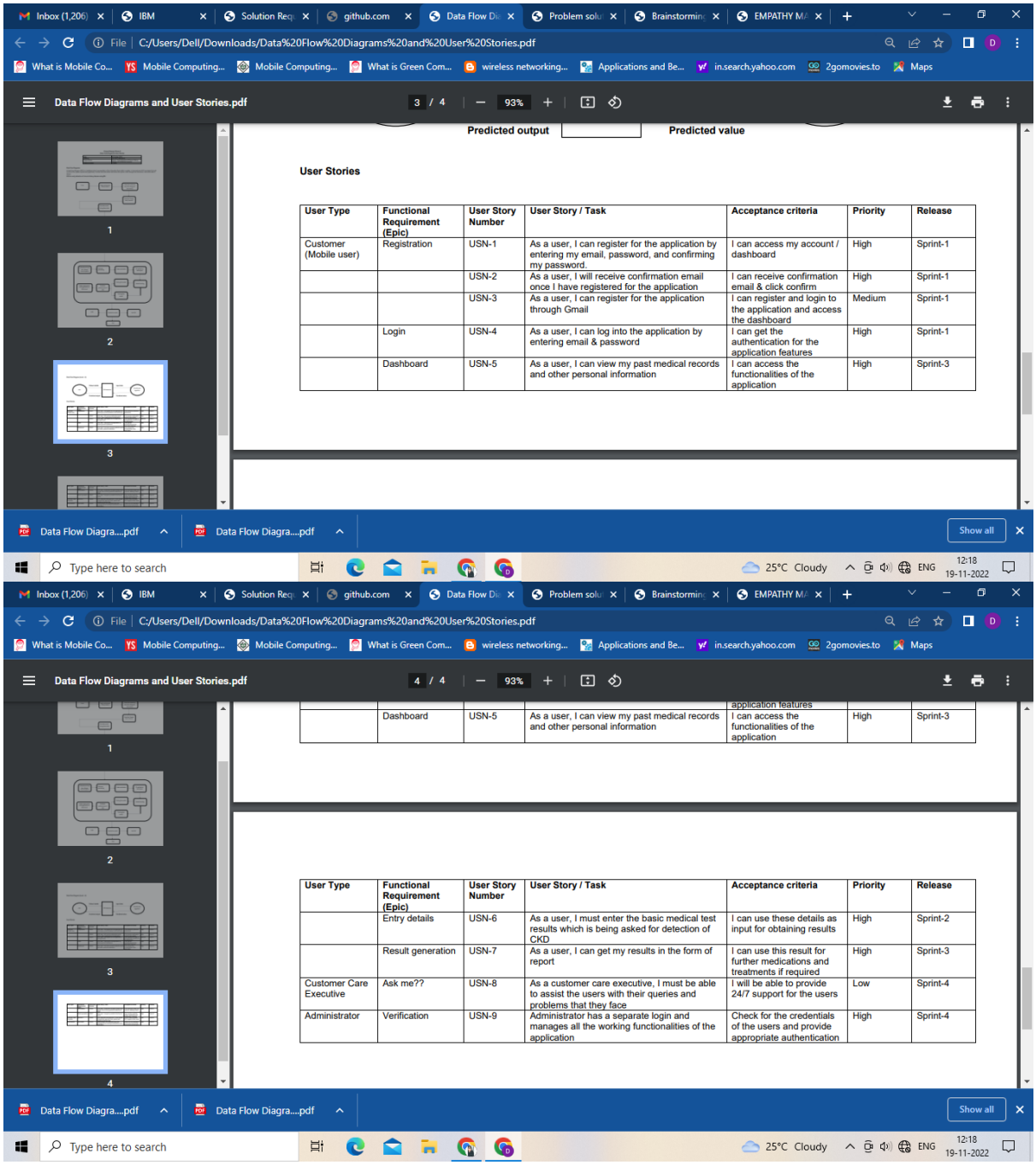




5.2)SOLUTION AND TECHNICAL ARCHITECTURE:

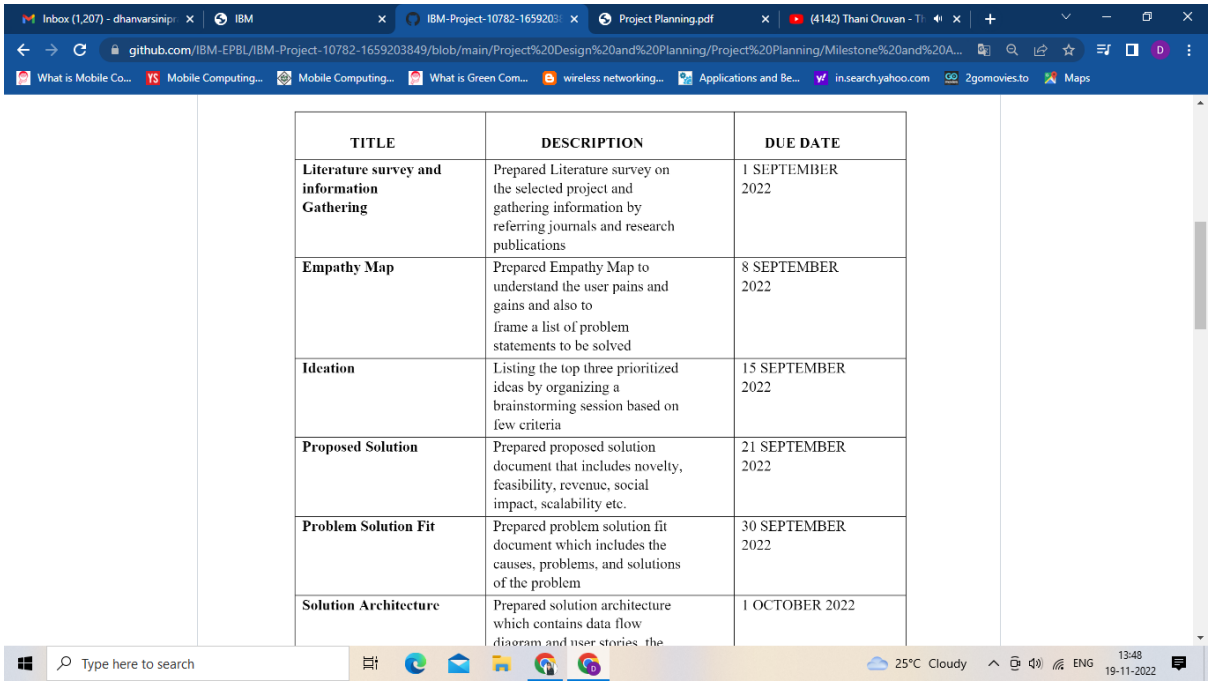


5.3)USER STORIES:



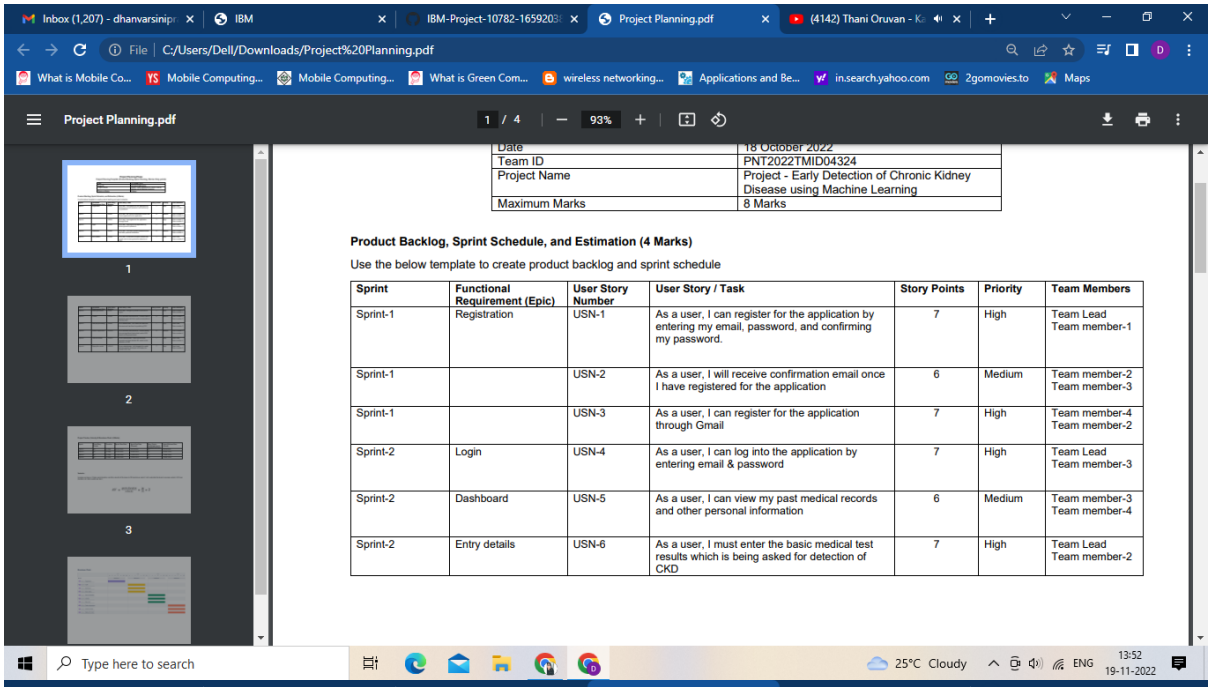
6)PROJECT PLANNING AND SCHEDULING

6.1)SPRINT PLANNING AND ESTIMATION:



TITLE	DESCRIPTION	DUE DATE
Literature survey and information Gathering	Prepared Literature survey on the selected project and gathering information by referring journals and research publications	1 SEPTEMBER 2022
Empathy Map	Prepared Empathy Map to understand the user pains and gains and also to frame a list of problem statements to be solved	8 SEPTEMBER 2022
Ideation	Listing the top three prioritized ideas by organizing a brainstorming session based on few criteria	15 SEPTEMBER 2022
Proposed Solution	Prepared proposed solution document that includes novelty, feasibility, revenue, social impact, scalability etc.	21 SEPTEMBER 2022
Problem Solution Fit	Prepared problem solution fit document which includes the causes, problems, and solutions of the problem	30 SEPTEMBER 2022
Solution Architecture	Prepared solution architecture which contains data flow diagram and user stories, the model, and the website flow	1 OCTOBER 2022
Customer Journey	Prepared customer journey map which helps to understand the user needs	5 OCTOBER 2022
Functional Requirement	Prepared the functional requirements that contains the necessary features for the application	12 OCTOBER 2022
Data Flow Diagrams	Prepared the data flow diagrams and user stories for the application which depicts the information flow within the system	15 OCTOBER 2022
Technology Architecture	Prepared the technology architecture that defines the technologies and IBM cloud	19 OCTOBER 2022

6.2)SPRINT DELIVERY SCHEDULE:

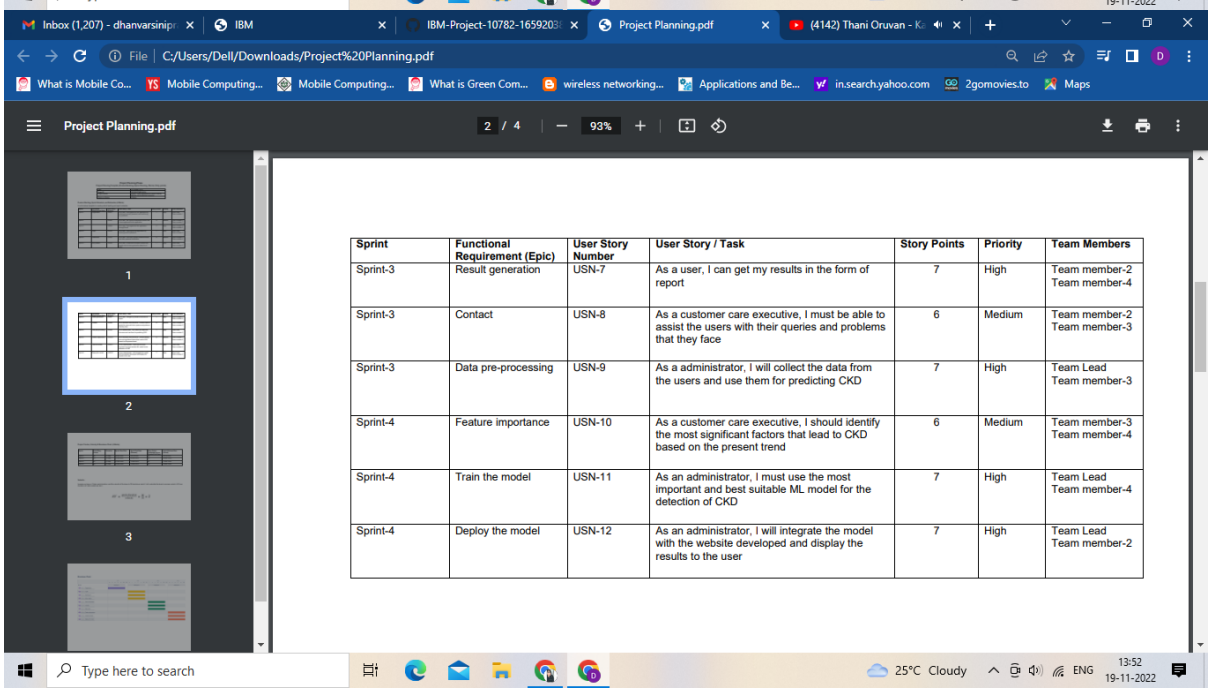


Date	18 October 2022
Team ID	PNT2022TMID04324
Project Name	Project - Early Detection of Chronic Kidney Disease using Machine Learning
Maximum Marks	8 Marks

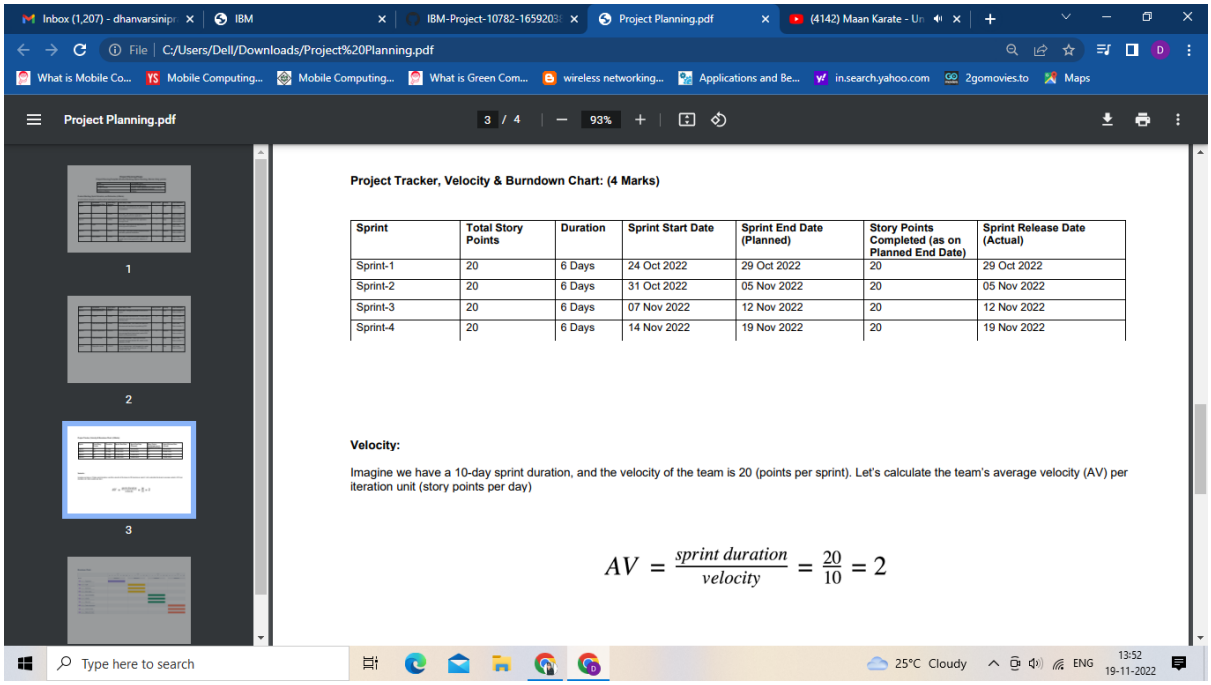
Product Backlog, Sprint Schedule, and Estimation (4 Marks)

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	As a user, I can register for the application by entering my email, password, and confirming my password.	7	High	Team Lead Team member-1
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	6	Medium	Team member-2 Team member-3
Sprint-1		USN-3	As a user, I can register for the application through Gmail	7	High	Team member-4 Team member-2
Sprint-2	Login	USN-4	As a user, I can log into the application by entering email & password	7	High	Team Lead Team member-3
Sprint-2	Dashboard	USN-5	As a user, I can view my past medical records and other personal information	6	Medium	Team member-3 Team member-4
Sprint-2	Entry details	USN-6	As a user, I must enter the basic medical test results which is being asked for detection of CKD	7	High	Team Lead Team member-2



Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-3	Result generation	USN-7	As a user, I can get my results in the form of report	7	High	Team member-2 Team member-4
Sprint-3	Contact	USN-8	As a customer care executive, I must be able to assist the users with their queries and problems that they face	6	Medium	Team member-2 Team member-3
Sprint-3	Data pre-processing	USN-9	As a administrator, I will collect the data from the users and use them for predicting CKD	7	High	Team Lead Team member-3
Sprint-4	Feature importance	USN-10	As a customer care executive, I should identify the most significant factors that lead to CKD based on the present trend	6	Medium	Team member-3 Team member-4
Sprint-4	Train the model	USN-11	As an administrator, I must use the most important and best suitable ML model for the detection of CKD	7	High	Team Lead Team member-4
Sprint-4	Deploy the model	USN-12	As an administrator, I will integrate the model with the website developed and display the results to the user	7	High	Team Lead Team member-2



Project Tracker, Velocity & Burndown Chart: (4 Marks)

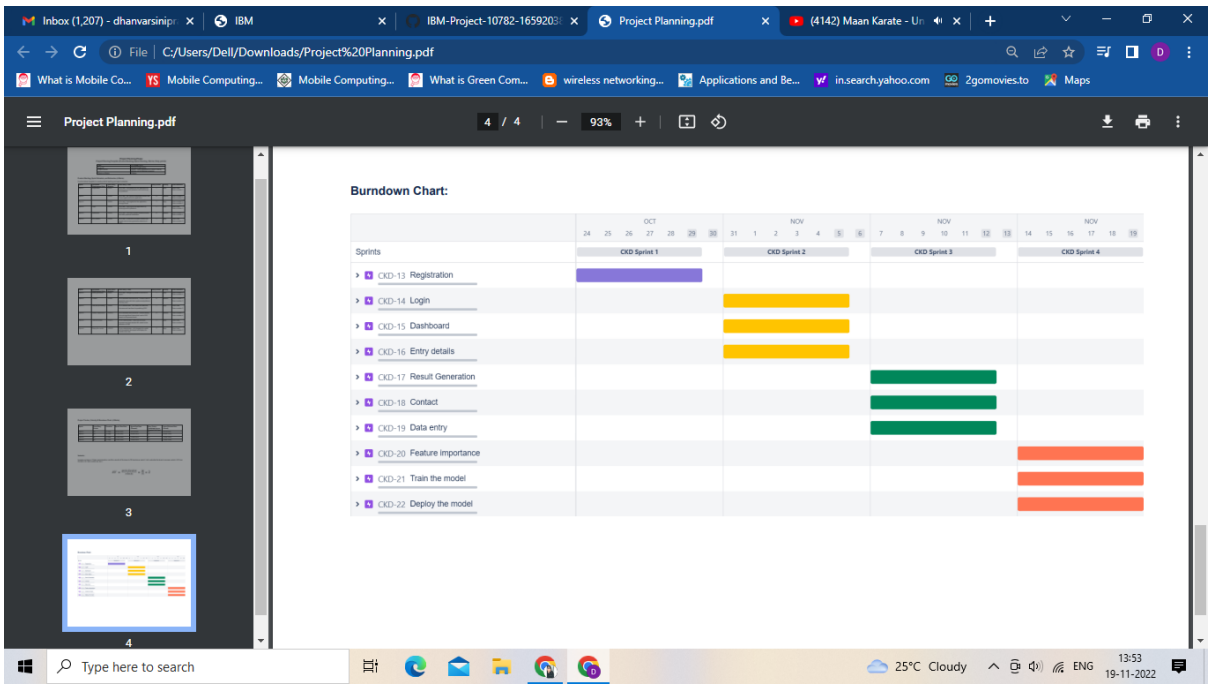
Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	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 per day)

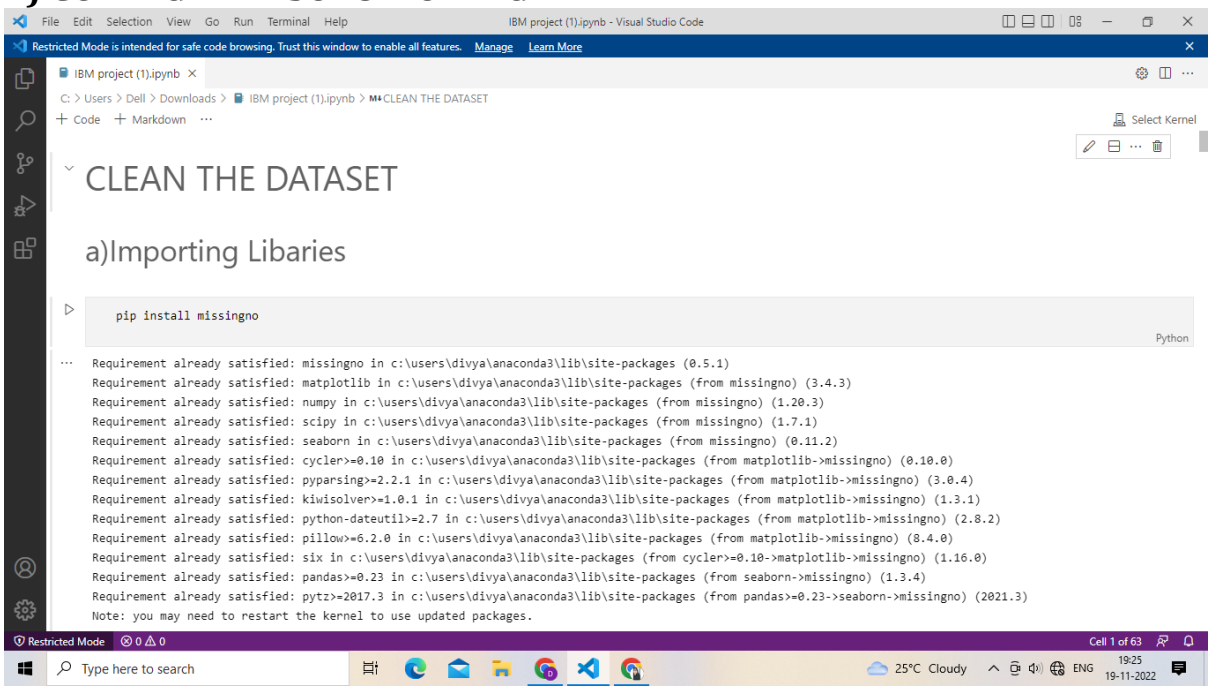
$$AV = \frac{sprint\ duration}{velocity} = \frac{20}{10} = 2$$

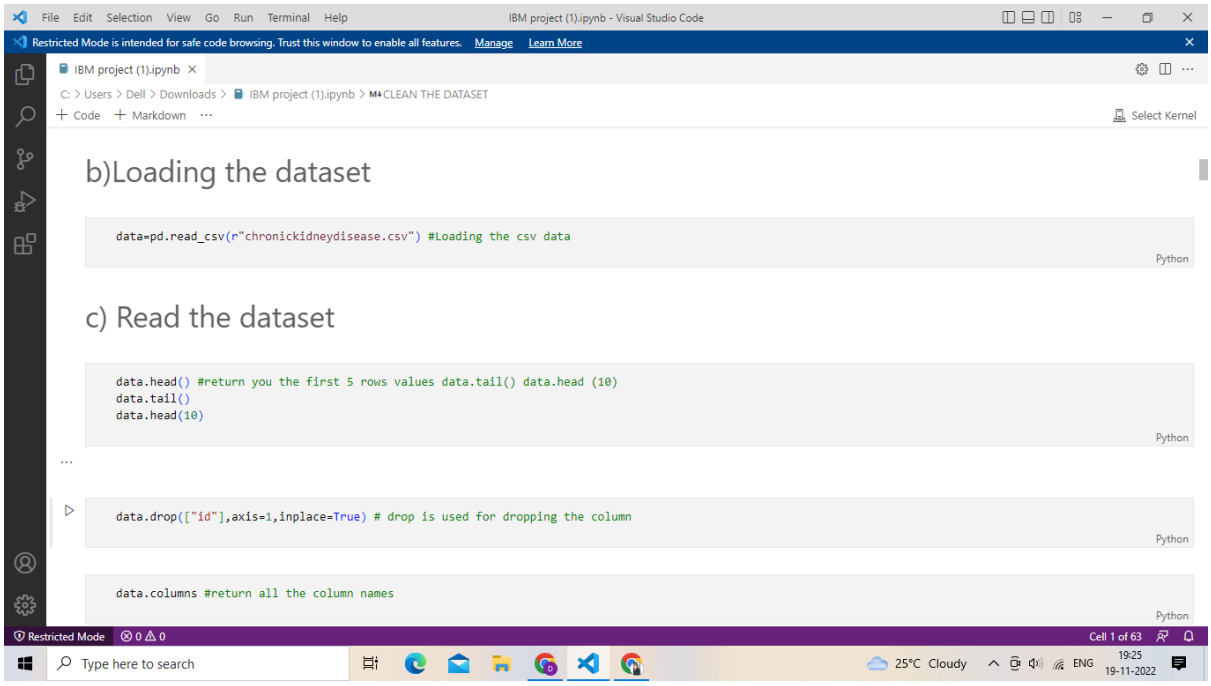
6.3)REPORTS FROM JIRA SOFYWARE:



Burndown Chart:

7) CODING AND SOLUTIONING





10)ADVANTAGES AND DISADVANTAGES

Kidneys are bean shaped organs which are located at the back of the abdominal cavity; they are necessary because they filter waste products such as nitrogen from the bloodstream, reabsorb necessary products (e.g. sodium and water) and remove the waste as urea via the ureter. The specific part of the Kidney that filters waste products is called the nephrons. There are millions of these filters within the kidney tissue, which take blood from the renal vein, transport it through a tubular filtration system (where necessary products are removed and reabsorbed into the body) and remove the waste via the ureter. When the renal system fails, homeostatic balance becomes disrupted, due to the fact that water and salt cannot be reabsorbed as required and the fact that waste products cannot be removed from the body. This homeostatic imbalance creates toxicity and eventually affects all bodily functions. There are two methods that are used to treat kidney failure; dialysis (using a machine to filter waste products from the blood stream) and kidney transplantation (taking a functioning kidney from another person and surgically implanting it into the body). This essay will explore the advantages and disadvantages of dialysis and kidney transplantation in order to make a judgement on the most effective treatment for renal failure. The kidneys are responsible for filtering blood and producing urine as waste from your body to the bladder. But there are instances of kidneys losing their functionality, this is called either acute disease were kidneys lose function for less than 3 months or chronic kidney disease(chronic kidney failure) for persistent loss of functions. So what happens when the kidneys lose function? this causes uremic syndrome liquids and waste build up in the body which can reach dangerous levels and cause death without a transplant or dialysis treatment.Patients will go through one of two forms of this treatment continuous renal replacement therapy (CRRT) or slow low efficiency dialysis (SLED). SLED being the faster method as it cuts the session time in half from the average time in CRRT of 6 to 12 hours a session down to 3-4 hours. But each method has its own advantages and disadvantages. SLED being cheaper and faster method that is gaining more favor in recent times. CRRT is used to treat critically ill patients with acute kidney injury and being more user friendly while having better fluid removal.

11)CONCLUSION

Chronic renal failure represents a critical period in the evolution of chronic renal disease and is associated with complications and comorbidities that begin early in the course of the disease. These conditions are initially subclinical but progress relentlessly and may eventually become symptomatic and irreversible. Early in the course of chronic renal failure, these conditions are amenable to interventions with relatively simple treatments that have the potential to prevent adverse outcomes. It summarises strategies for effective management of chronic renal disease. By acknowledging these facts, we have an excellent opportunity to change the paradigm of management of chronic renal failure and improve patient outcomes.

12)FUTURE SCOPE

The increasing prevalence of chronic kidney disease is well known, as it is a fact that recorded data in all countries show continuing growth in the number of patients that need substitutive treatment for their renal function.
Here are different reasons for such high mortality rates, amongst them significant increase in the age of patients undergoing treatment, restoration with haemodialysis and peritoneal dialysis of only 15 to 20ml/min of kidney function, and a significant associated co-morbidity.
The requirements for new technologies in dialysis are, therefore, based on the following objectives:

1. Continuous function.

2. Elimination of molecular weight solutes similar to kidney function.

3. Elimination of water and solutes according to patient's needs.

4. Biocompatibility.

5. Portable, or even better, implantable.

6. Low cost.

7. Safety.

GIT HUB LINK
https://github.com/IBM-EPBL/IBM-Project-10782-1659203849

PROJECT DEMO LINK
https://detection-chronic-kidney-disease.glitch.me/