

Project Report

Title: Early Detection of Chronic Kidney Disease using Machine Learning

Team ID: PNT2022TMID48683

Introduction

Project Overview

Chronic kidney disease, also called chronic kidney failure, describes the gradual loss of kidney function. Your kidneys filter wastes and excess fluids from your blood, which are then excreted in your urine. When chronic kidney disease reaches an advanced stage, dangerous levels of fluid, electrolytes and wastes can build up in your body.

In the early stages of chronic kidney disease, you may have few signs or symptoms. Chronic kidney disease may not become apparent until your kidney function is significantly impaired.

Treatment for chronic kidney disease focuses on slowing the progression of the kidney damage, usually by controlling the underlying cause. Chronic kidney disease can progress to end-stage kidney failure, which is fatal without artificial filtering (dialysis) or a kidney transplant.

Purpose

- To predict the disease early and to take the precautions in advance
- To have a better glance about the patient health condition
- To know the severity of chronic kidney disease in early stage
- Reducing the cost that taken to measure these disease
- Make a more reliable way to measure chronic kidney disease

Literature Survey

Existing Problem

1. To measure the glomerular filtration rate (GFR) in adults and children with cystic fibrosis (CF) using a radio-isotope technique as the gold standard and to compare this to serum creatinine based equations, serum cystatin C levels and tobramycin clearance, and to determine which method correlates most closely with measured GFR in this population.
2. The main limitation of the eGFR formulas is significant lack of accuracy in individuals as compared to groups. This causes problems in establishing the presence or absence of early CKD
3. Unfortunately, the initial creatinine tests can lead to misdiagnosis and inappropriate treatment." Misdiagnoses appear to be common, according to the researchers' review of the electronic health records of 3.8 million emergency and intensive care patients.

References

- [1] "9 Symptoms of Type 1 & Type 2 Diabetes: Complications, Causes & Diet", MedicineNet, 2019. [Online]. Available: https://www.medicinenet.com/diabetes_mellitus/article.htm [Accessed: 05- Jul- 2019].
- [2] "International Diabetes Federation - What is diabetes", Idf.org, 2019. [Online]. Available: <https://www.idf.org/aboutdiabetes/what-isdiabetes.html>. [Accessed: 08- Jul- 2019].
- [3] "Statistics About Diabetes", Diabetes.org, 2019. [Online]. Available: <https://www.diabetes.org/resources/statistics/statistics-about-diabetes>. [Accessed: 10- Jul- 2019].
- [4] R. Hira, M. Miah and D. Akash, "Prevalence of Type 2 Diabetes Mellitus in Rural Adults (> 31years) in Bangladesh", Faridpur Medical College Journal, vol. 13, no. 1, pp. 20-23, 2018. [Accessed 16 July 2019].
- [5] "Machine Learning - Definition and application examples", Spotlightmetal.com, 2019. [Online]. Available: <https://www.spotlightmetal.com/machine-learning-definition-andapplication-examples-a-746226/> [Accessed: 17- Jul- 2019].
- [6] A. Mir and S. Dhage, "Diabetes Disease Prediction Using Machine Learning on Big Data of Healthcare", in 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBE), Pune, India, 2018.
- [7] D. Dutta, D. Paul and P. Ghosh, "Analysing Feature Importances for Diabetes Prediction using Machine Learning", in 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), Vancouver, BC, Canada, 2018.
- [8] M. Raihan, Muhammad Muinul Islam, Promila Ghosh, Shakil Ahmed Shaj, Muhtasim Rafid Chowdhury, Saikat Mondal, Arun More, "A Comprehensive Analysis on Risk Prediction of Acute Coronary Syndrome Using Machine Learning

Approaches", in 2018 21st International Conference of Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 2018, pp. 1 – 6.

[9] D. Kinge and S. Gaikwad, "Survey on data mining techniques for disease prediction", International Research Journal of Engineering and Technology (IRJET), vol. 05, no. 01, pp. 630-636, 2018. [Accessed 21 July 2018].

[10] S. Manna, S. Maity, S. Munshi and M. Adhikari, "Diabetes Prediction Model Using Cloud Analytics", in 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Bangalore, India, 2018.

[11] D. Verma and N. Mishra, "Analysis and prediction of breast cancer and diabetes disease datasets using data mining classification techniques ", in 2017 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, India, 2017.

[12] W. Xu, J. Zhang, Q. Zhang and X. Wei, "Risk prediction of type II diabetes based on random forest model", in 2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), Chennai, India, 2017.

[13] H. Emblem, "When to use a Trimmed Mean", Medium, 2018. [Online]. Available: <https://medium.com/@HollyEmblem/when-to-use-a-trimmed-mean-fd6aab347e46> [Accessed: 24-Jul-2019].

[14] "TrimMean function R Documentation", Rdocumentation.org. [Online]. Available: www.rdocumentation.org/packages/sscore/versions/1.44.0/topics/trimMean [Accessed: 25-Jul-2019].

[15] "Median", RDocumentation, 2019. [Online]. Available: <https://www.rdocumentation.org/packages/stats/versions/3.5.2/topics/median> [Accessed: 01- Aug- 2019].

[16] H. Kandan, "Bagging the skill of Bagging(Bootstrap aggregating).", Medium, 2018. [Online]. Available: <https://medium.com/@harishkandan95/bagging-the-skill-of-baggingbootstrap-aggregating-83c18dcabdf1> [Accessed: 04-Aug- 2019].

Problem Statement Definition

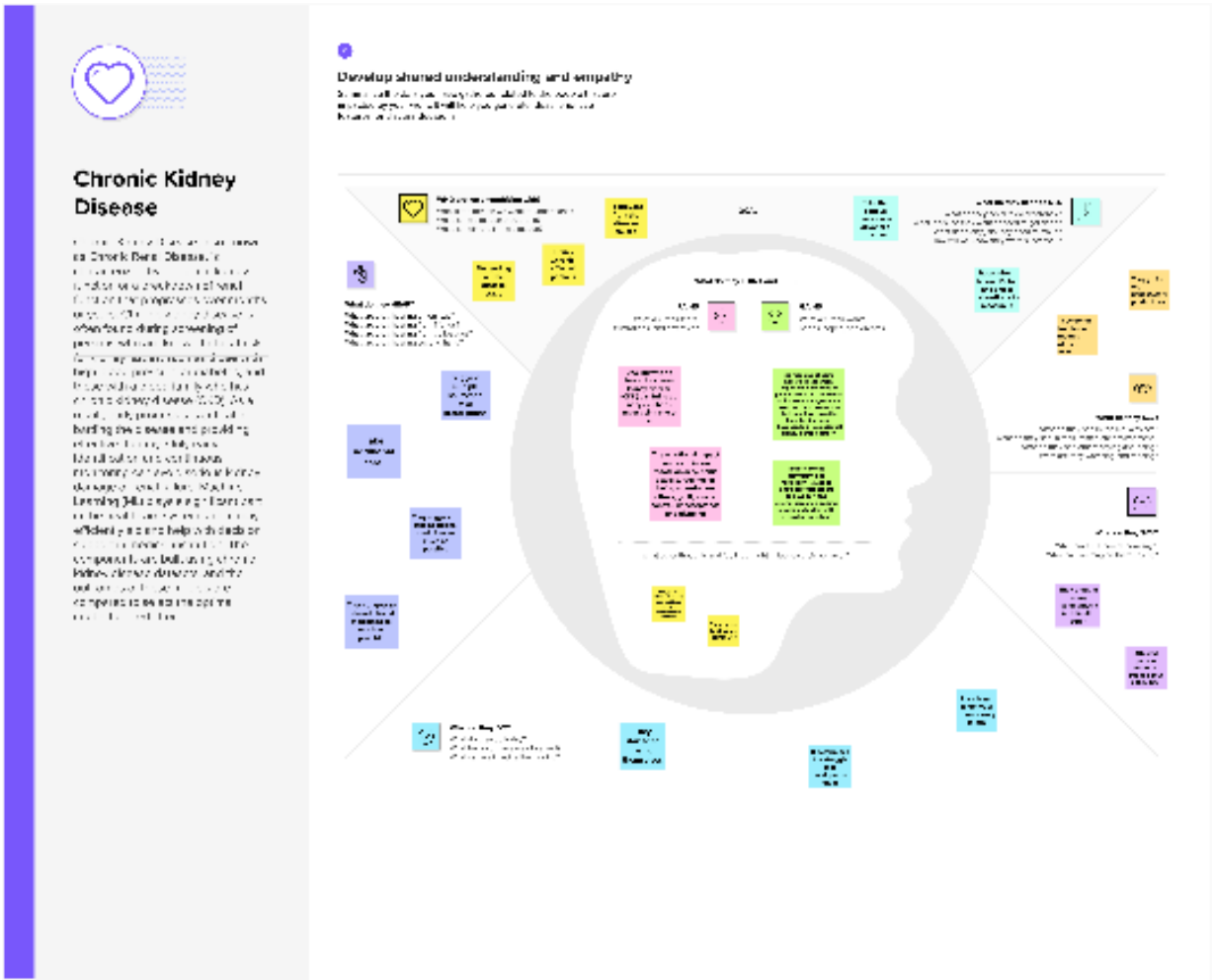
Chronic kidney disease includes conditions that damage your kidneys and decrease their ability to keep you healthy by filtering wastes from your blood. If kidney disease worsens, wastes can build to high levels in your blood and make you feel sick. You may develop complications like:

- high blood pressure
- anemia (low blood count)
- weak bones
- poor nutritional health
- nerve damage

Kidney disease also increases your risk of having heart and blood vessel disease. These problems may happen slowly over a long time. Early detection and treatment can often keep chronic kidney disease from getting worse. When kidney disease progresses, it may eventually lead to kidney failure, which requires dialysis or a kidney transplant to maintain life.

Ideation & Proposed Solution

Empathy Map Canvas



Proposed Solution

Chronic Kidney Disease, also known as Chronic Renal Disease, is characterized by abnormal kidney function or a breakdown of renal function that progresses over months or years. Chronic kidney disease is often found during screening of persons who are known to be at risk for kidney issues, such as those with high blood pressure or diabetes, and those with a blood family who has chronic kidney disease (CKD). As a result, early prognosis is critical in battling the disease and providing effective therapy. Only early identification and continuous monitoring can avoid serious kidney damage or renal failure. Machine Learning (ML) plays a significant part in the

<div>3. TRIGGERS</div> <div>TR</div> <div>What triggers customers to act?</div> <div>Diabetes is the most common cause of kidney disease. Both type 1 and type 2 diabetes. But also heart disease and obesity can contribute to the damage that causes kidneys to fail. Urinary tract issues and inflammation in different parts of the kidney can also lead to long-term functional decline</div>	<div>10. YOUR SOLUTION</div> <div>SL</div> <div>If you are working on an existing business, write down your current solution first, fill in the canvas, and check how much it fits reality.</div> <div>End stage renal disease (ESRD) describes the most severe stage of chronic kidney disease (CKD), when patients need dialysis or renal transplant. There is often a delay in recognizing, diagnosing, and treating the various etiologies of CKD. The objective of the present study was to employ machine learning algorithms to develop a prediction model for progression to ESRD based on a large-scale multidimensional database</div> <div>we will be going through the Chronic kidney disease dataset and doing the complete analysis on the same our main goal will be to predict whether an individual will have chronic kidney disease or not based on the data provided.</div>	<div>8. CHANNELS of BEHAVIOUR</div> <div>CH</div> <div>8.1 ONLINE What kind of actions do customers take online? Extract online channels from #7</div> <div>Via a web application the prediction goes on and it takes some parameters of persons to calculate</div> <div>8.2 OFFLINE Taking offline requirements such as health data of persons makes the prediction much more accurate to predict</div>
<div>4. EMOTIONS: BEFORE / AFTER</div> <div>EM</div> <div>How do customers feel when they face a problem or a job and afterwards?</div> <div>They started to lose their control when they face the problem before and after that condition they lose their ability to take control over situations</div>		

healthcare system, and it may efficiently aid and help with decision support in medical institution. The components are built using chronic kidney disease datasets, and the outcomes of these models are compared to select the optimal model for prediction.

Problem Solution Fit

<div>1. CUSTOMER SEGMENT(S)</div> <div>Who is your customer?</div> <div>Globally kidney chronic disease is a threat and also it is hard to know whether the person is affected or not hence the customer in this topic is global persons who want to know he/she is affected by chronic disease</div>	<div>6. CUSTOMER CONSTRAINTS</div> <div>What constraints prevent your customers from taking action or limit their choices</div> <div>of solutions?</div> <div>The began to reduce their chance of finding best solutions and also take the words of neighbors, others could limit their possibilities of taking best decisions under those circumstances</div>	<div>5. AVAILABLE SOLUTIONS</div> <div>or need to get the job done? What have they tried in the past? What pros & cons do these solutions have?</div> <div>Analyzing a sample of your urine (urinalysis) may reveal abnormalities that suggest kidney failure and taking blood test but it not much reliable because it doesn't show exactly whether the person is under worse condition.</div>
---	--	---

Requirement Analysis

Functional Requirement

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Google account Registration through Google forms
FR-2	User Confirmation	Confirmation via Google
FR-3	User Medical Data	Amount of blood sugar, Blood urea, Coronary Disease, Pus cells, Petal edema, Diabetes mellitus, Red blood

<div>2. JOBS-TO-BE-DONE / PROBLEMS</div> <div>J&P</div> <div>Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one; explore different sides.</div> <div>The major threat of kidney chronic disease is that it is unknown before till it becomes worse condition so our solution implies the best method to predict whether the host is affected by chronic disease in the real time and choose to take further cause for that</div>	<div>9. PROBLEM ROOT CAUSE</div> <div>PRC</div> <div>What is the real reason that this problem exists? What is the back story behind the need to do this job?</div> <div>Diabetes and high blood pressure, or hypertension, are responsible for two-thirds of chronic kidney disease cases. Doesn't taking much precautions for those disease will lead them to renal affection</div>	<div>7. BEHAVIOUR</div> <div>BE</div> <div>What does your customer do to address the problem and get the job done?</div> <div>The best method to reduce the risk of kidney chronic is to predict it earlier which will make great impact to take much cause to prevent or safe guard from the hazardous stage. The internet is wide open now a days so make the prediction via online is much more reliable for persons</div>
---	---	---

NFR-1	Usability	The usability of the project is easy and user friendly environment to the user to interact with it
NFR-2	Security	The data of users are well preserved and there is no any other third parties to gather the information of user
NFR-3	Reliability	User can use the product across any platform in the market

Project Planning & Scheduling

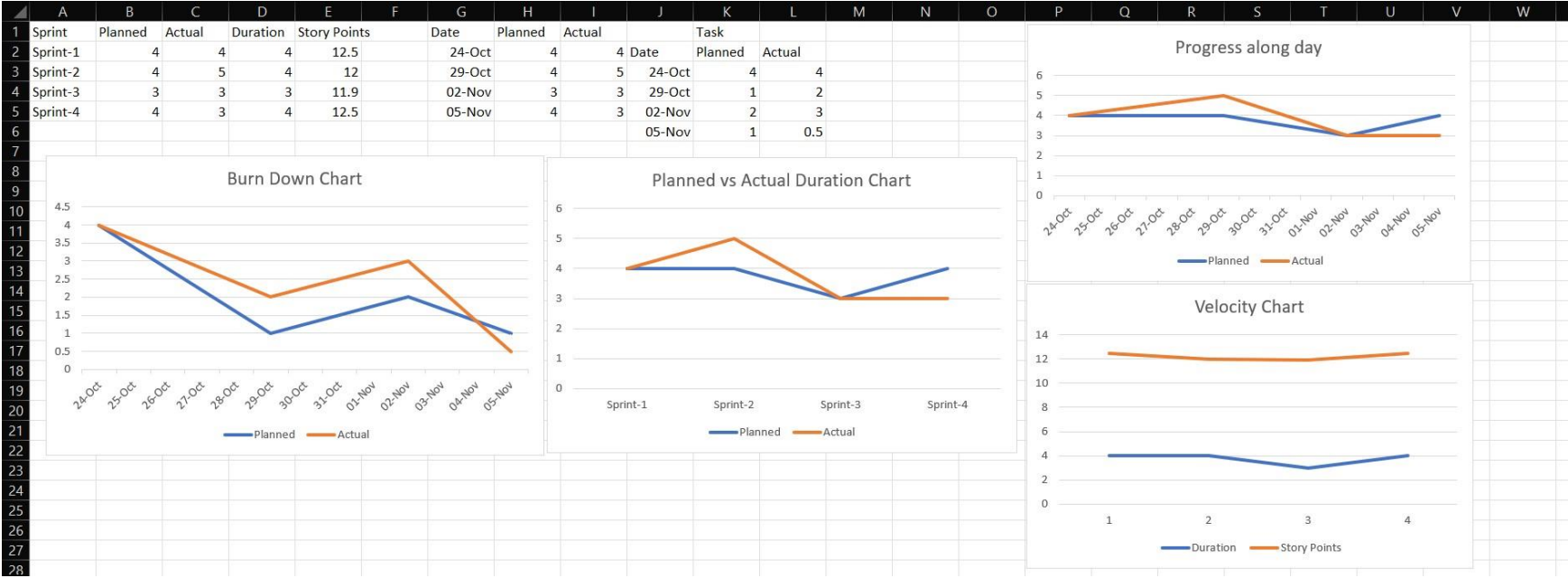
Sprint Planning and Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	User can sign in to our web application by using google account	12.5	High	Mohamed Ibrahim, Mohamed Asif, Prabhu, Kamalesh
Sprint-1		USN-2	Users are considered to fill the google form for further authentication	12.5	Medium	Mohamed Ibrahim, Mohamed Asif, Prabhu, Kamalesh
Sprint-1		USN-3	User log in data is stored and surveillance by google firebase authentication service	12.5	High	Mohamed Ibrahim, Mohamed Asif, Prabhu, Kamalesh
Sprint-1	Login	USN-4	User now successfully registered using google authentication service	12.5	High	Mohamed Ibrahim, Mohamed Asif, Prabhu, Kamalesh
Sprint-2	Dashboard	USN-5	Now user is preferred to enter the medical details to analyse their disease severity or whether they have that disease	12	High	Mohamed Ibrahim, Mohamed Asif, Prabhu, Kamalesh
Sprint-3	Dashboard	USN-6	There are three values which should be given in float or numerically	11.8	High	Mohamed Ibrahim, Mohamed Asif, Prabhu, Kamalesh
Sprint-3	Dashboard	USN-7	There are some other valuers which are fulfilled by alpha values whether they have or not (yes or no questions)	12	High	Mohamed Ibrahim, Mohamed Asif, Prabhu, Kamalesh
Sprint-4	Dashboard	USN-8	After fulfilment of user medical values the prediction of disease is analysed by machine learning model that present in the IBM Watson machine learning platform	12.5	High	Mohamed Ibrahim, Mohamed Asif, Prabhu, Kamalesh

Sprint Delivery Schedule

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	12.5	4 Days	20 Oct 2022	24 Oct 2022	12.5	24 Oct 2022
Sprint-2	12.5	4 Days	24 Oct 2022	28 Oct 2022	12	29 Oct 2022
Sprint-3	12.5	3 Days	28 Oct 2022	2 Nov 2022	11.9	2 Nov 2022
Sprint-4	12.5	4 Days	2 Nov 2022	6 Nov 2022	12.5	5 Nov 2022

Burndown Charts



Coding & Solution

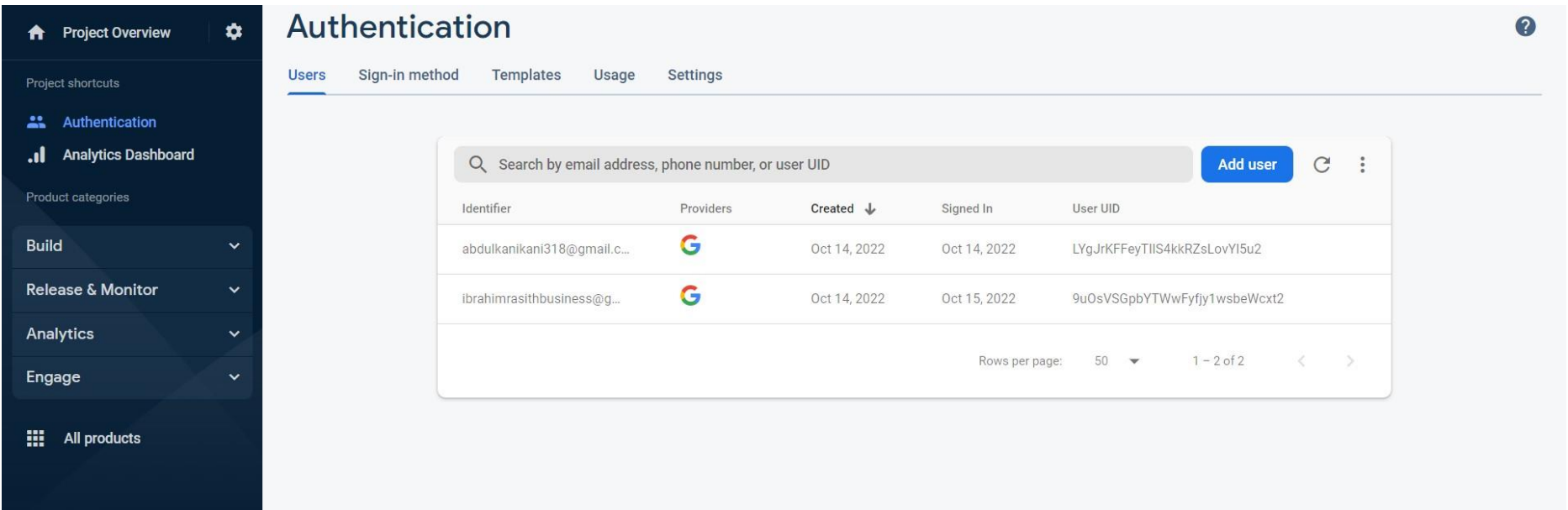
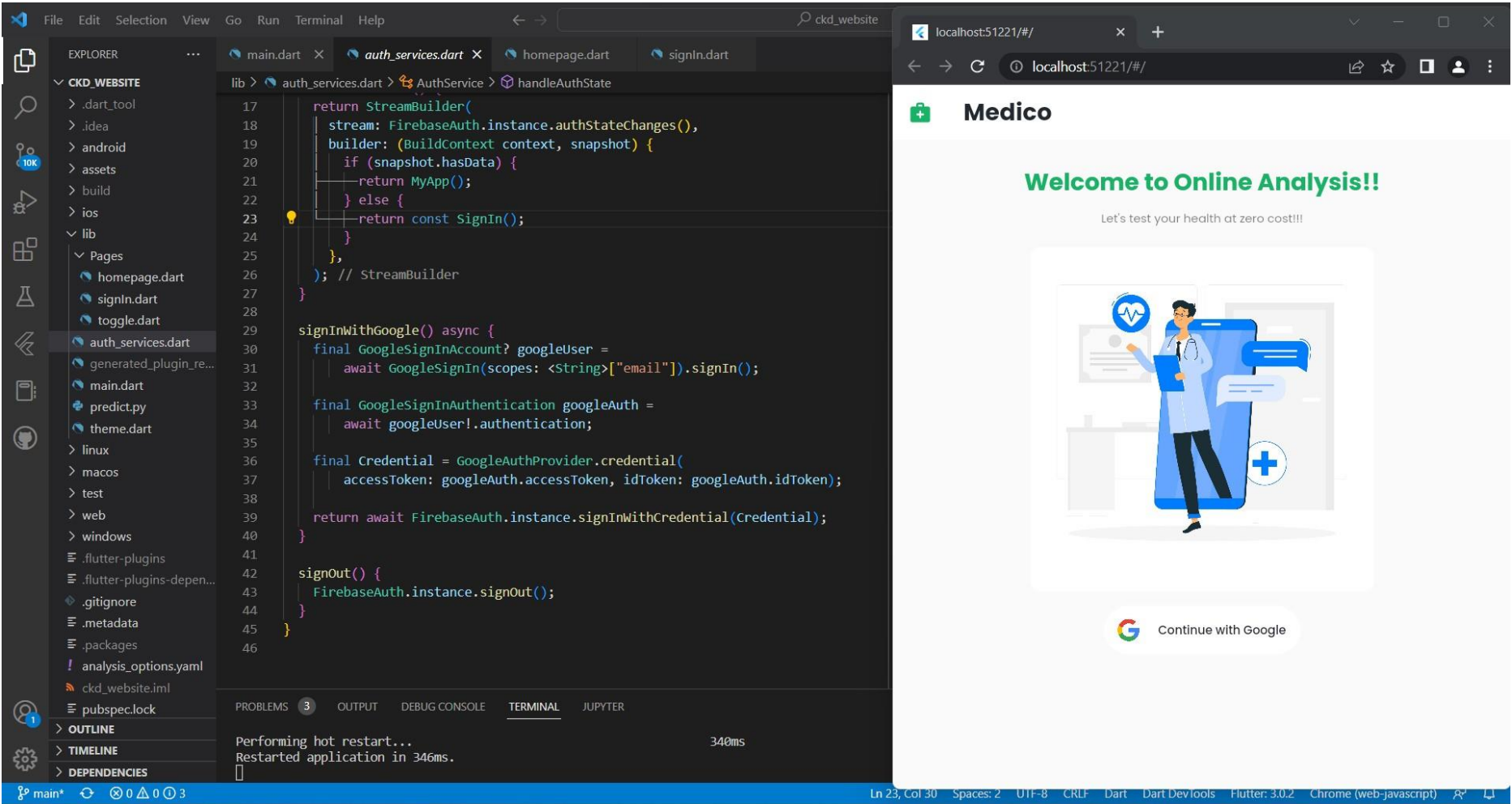
Feature 1

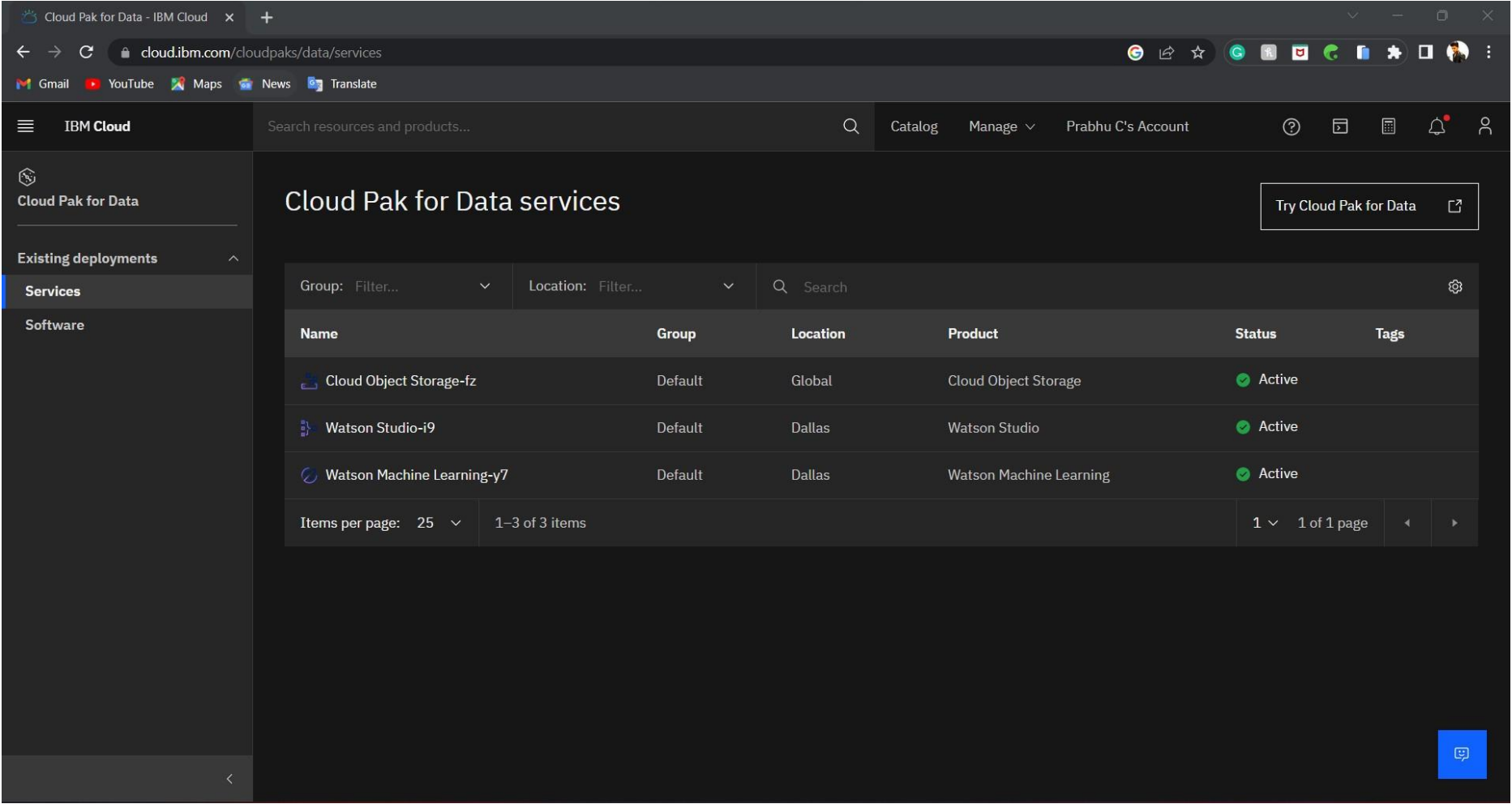
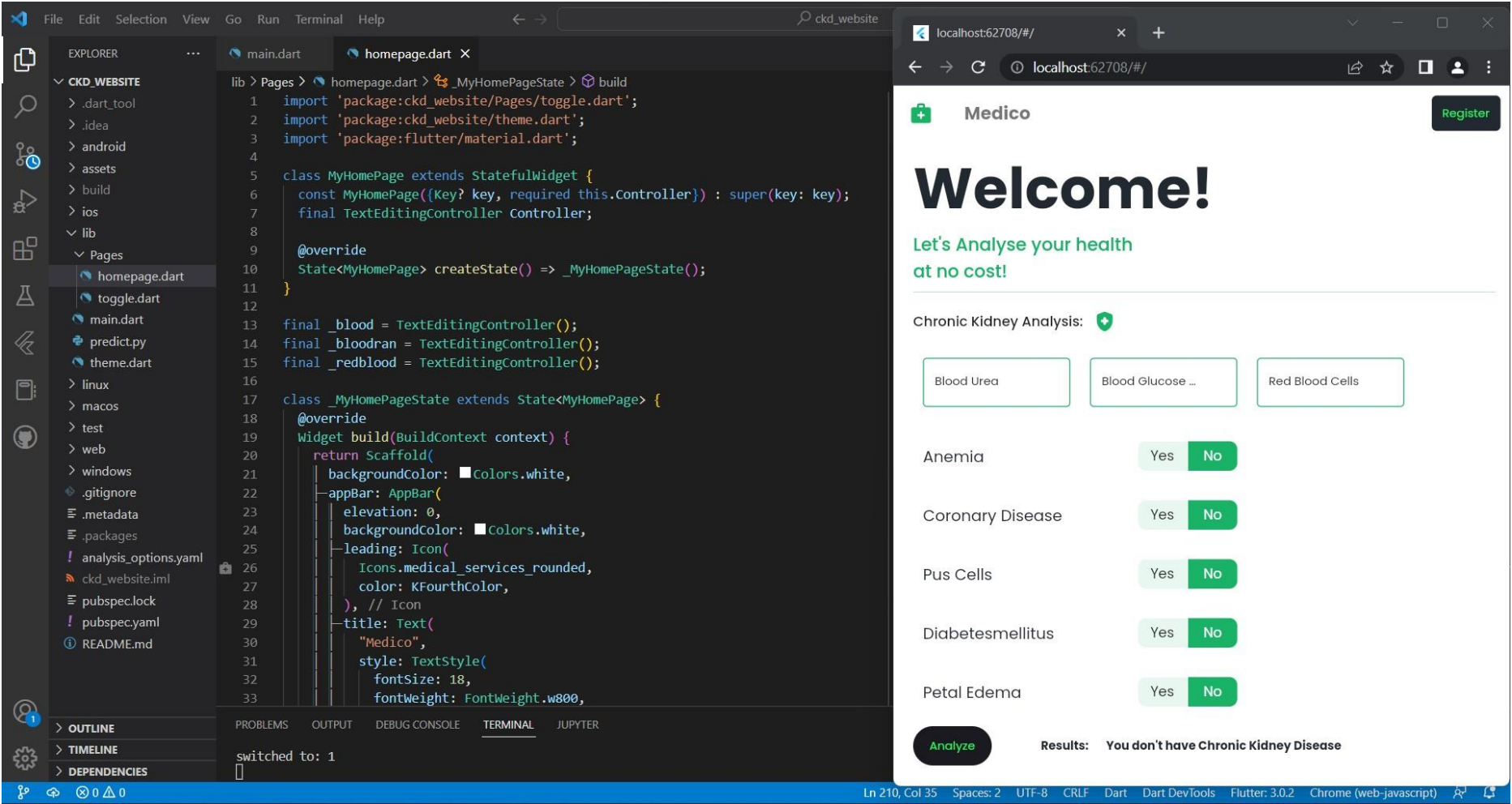
Full code link is attached below

<https://www.mediafire.com/file/0pjvai3irreddfh/chronic-kidney-disease-clustering-and-prediction.ipynb/file>

Testing

Test cases





Results

File Edit Selection View Go Run Terminal Help

cdk_website

EXPLORER

CKD_WEBSITE

lib > Pages > homepage.dart

```
lib > Pages > homepage.dart > _MyHomePageState > build
1 import 'package:cdk_website/Pages/toggle.dart';
2 import 'package:cdk_website/theme.dart';
3 import 'package:flutter/material.dart';
4
5 class MyHomePage extends StatefulWidget {
6   const MyHomePage({Key? key, required this.Controller}) : super(key: key);
7   final TextEditingController Controller;
8
9   @override
10  State<MyHomePage> createState() => _MyHomePageState();
11 }
12
13 final _blood = TextEditingController();
14 final _bloodran = TextEditingController();
15 final _redblood = TextEditingController();
16
17 class _MyHomePageState extends State<MyHomePage> {
18   @override
19   Widget build(BuildContext context) {
20     return Scaffold(
21       backgroundColor: Colors.white,
22       appBar: AppBar(
23         elevation: 0,
24         backgroundColor: Colors.white,
25         leading: Icon(
26           Icons.medical_services_rounded,
27           color: KFourthColor,
28         ), // Icon
29         title: Text(
30           "Medico",
31           style: TextStyle(
32             fontSize: 18,
33             fontWeight: FontWeight.w800,
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL JUPYTER

switched to: 1

localhost62708/#/

localhost:62708/#/

Medico

Register

Welcome!

Let's Analyse your health at no cost!

Chronic Kidney Analysis:

Blood Urea Blood Glucose ... Red Blood Cells

Anemia Yes No

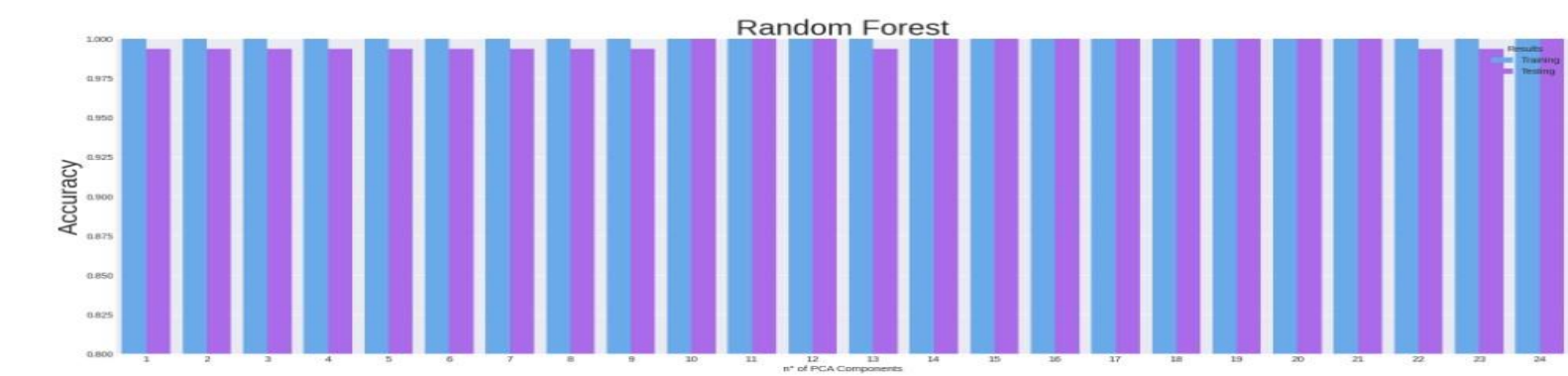
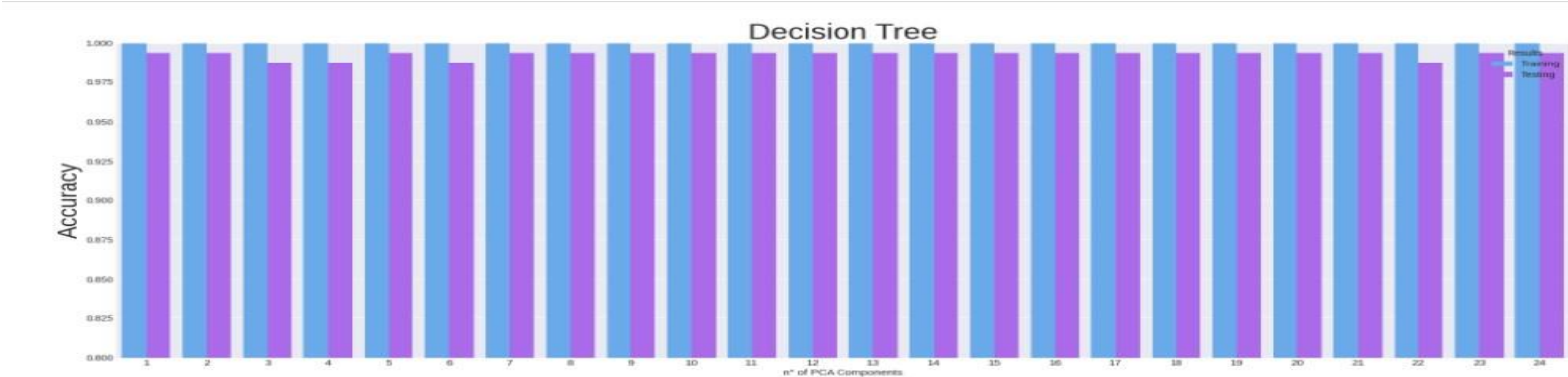
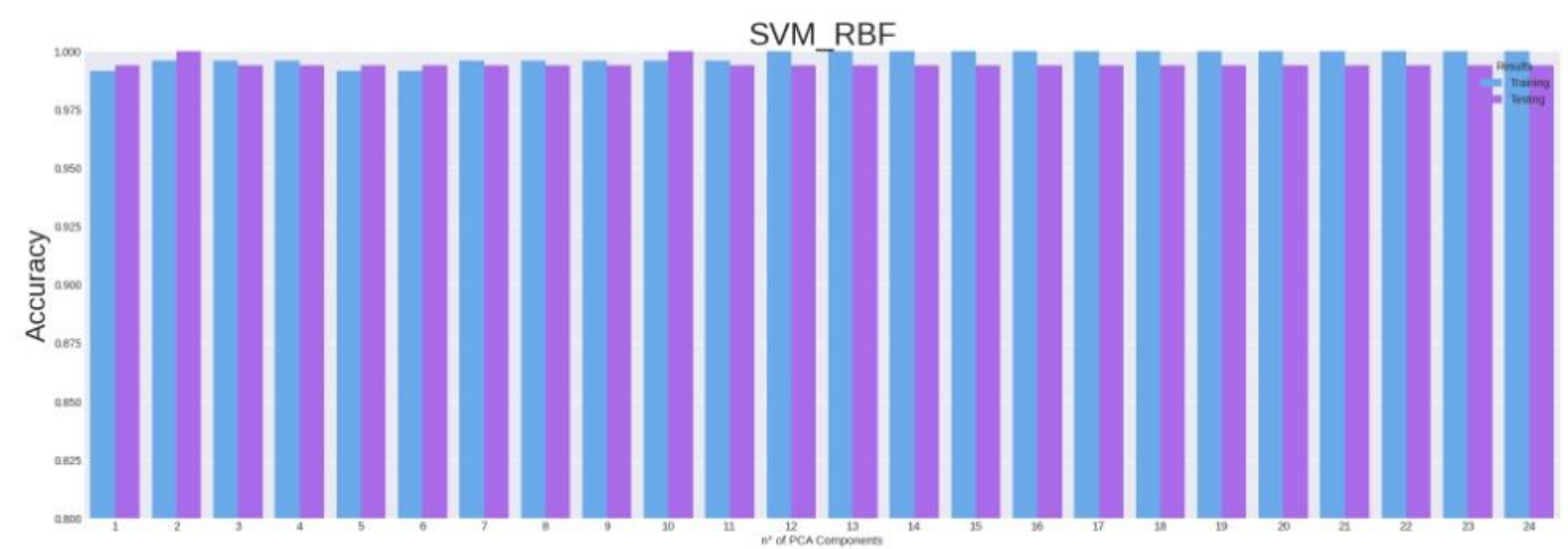
Coronary Disease Yes No

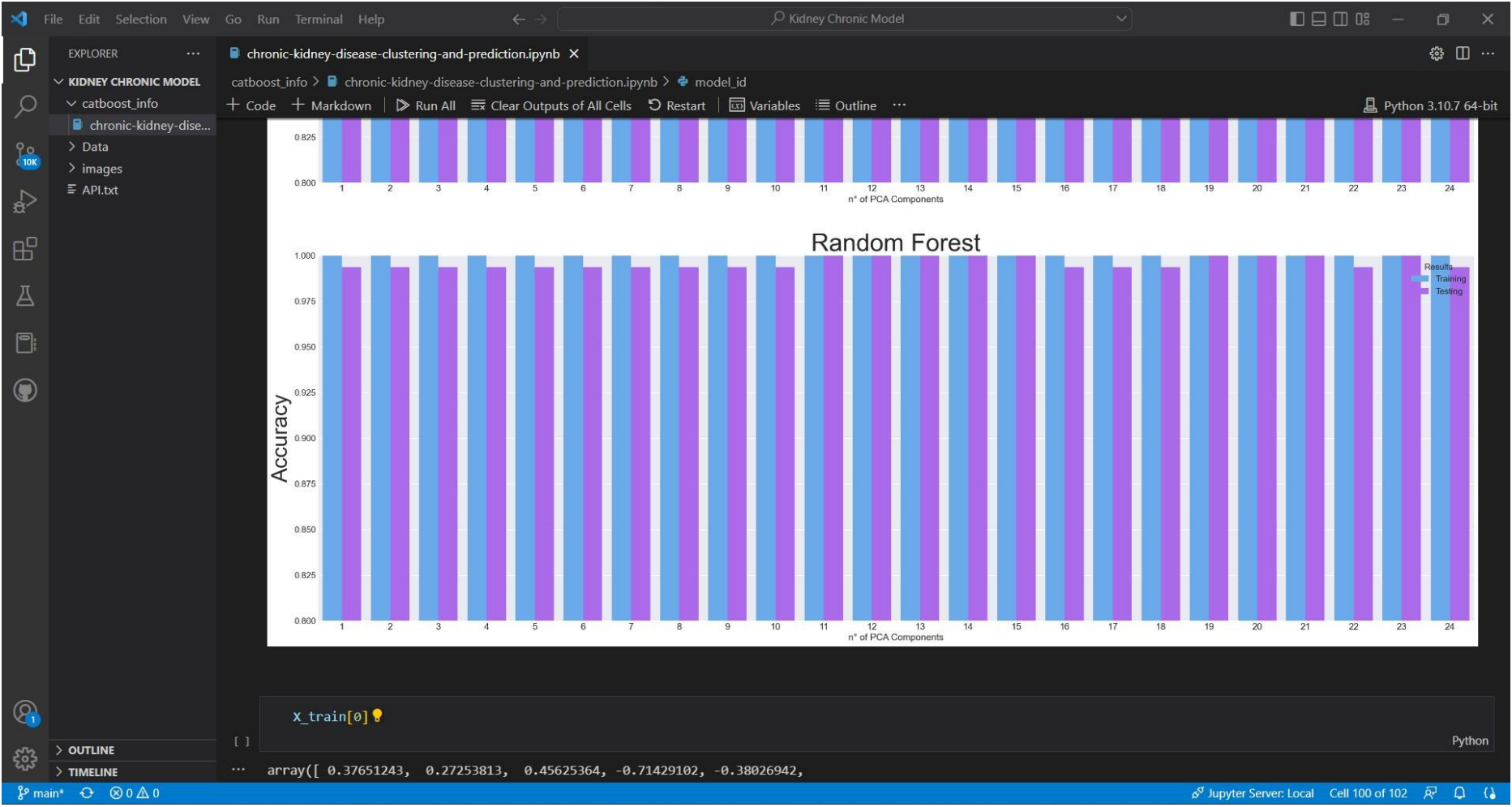
Pus Cells Yes No

Diabetesmellitus Yes No

Petal Edema Yes No

Analyze Results: You don't have Chronic Kidney Disease





```
mae = mean_absolute_error(actual, pred)
print("MAE without outliers:", mae)
mse = mean_squared_error(actual, pred)
print("MSE without outliers:", mse)
```

[8] ✓ 0.8s

MAE without outliers: 14.44

MSE without outliers: 334.4

```
mae = mean_absolute_error(actual, pred)
print("MAE with outliers:", mae)
mse = mean_squared_error(actual, pred)
print("MSE with outliers:", mse)
```

[11] ✓ 0.5s

MAE with outliers: 16.8

MSE with outliers: 560.4

```
print("RMSE", np.sqrt(mean_squared_error(actual, pred)))
```

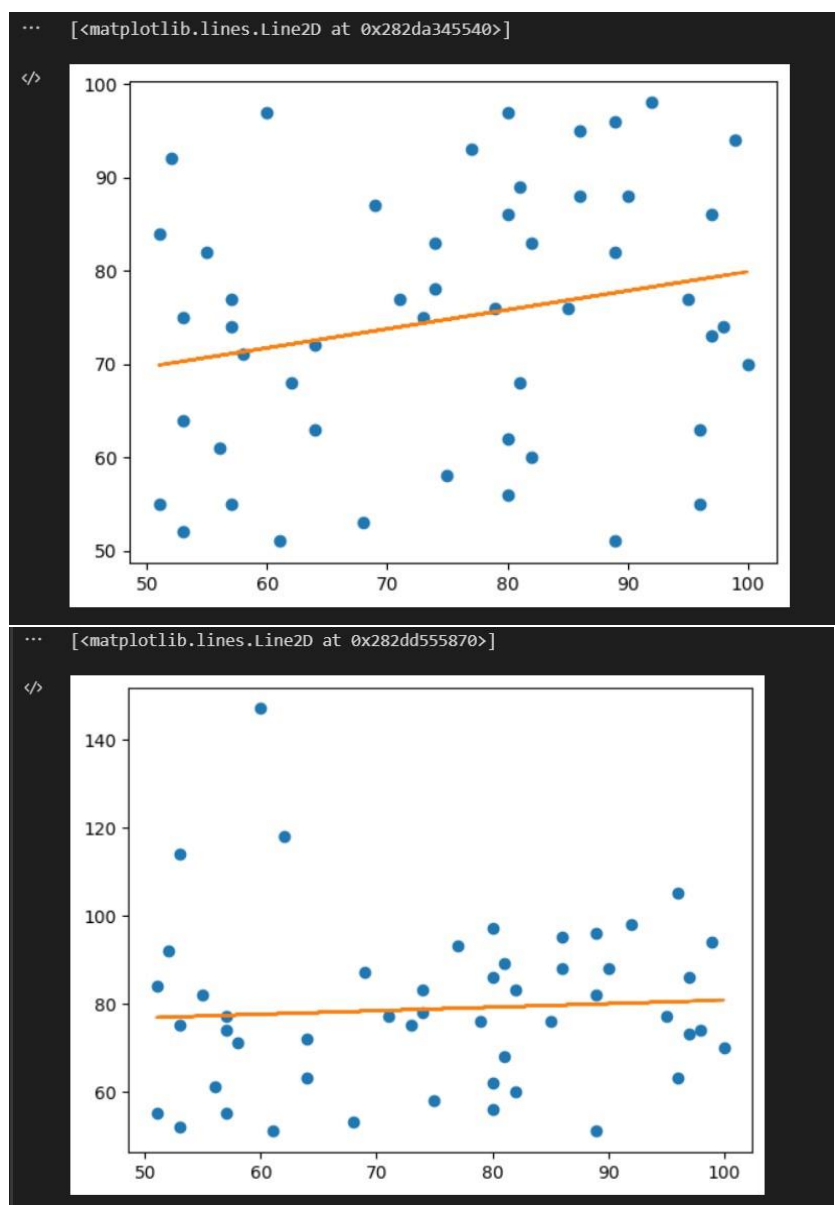
[16] ✓ 0.1s

RMSE 23.672769166280485

```
from sklearn.metrics import r2_score
r2 = r2_score(actual, pred)
print(r2)
```

[17] ✓ 0.6s

-1.3730317894793895



Advantages & Disadvantages

Advantages

- ML based algorithm has high accuracy in predicting kidney disease
- Using Big Data the chance of working progression is improved
- The model build has high performance in accuracy and gives better results
- Thus introducing into clinical field can prevent patients from severeness of chronic kidney disease
- The indication of kidney chronic disease shows the severeness according to that treatment can be done
- The time consuming is also reduced hence relevant action can be taken
- Constant remembrance of kidney health via online
- Machine Learning has enormous efficiency in data handling
- The chance of human error is less
- Machine learning in healthcare can also be used to improve the quality of patient care.

Disadvantages

- Gaps in healthcare data can lead to inaccurate forecasts
- ML requires high-quality structured data to produce accurate predictions.

Conclusion

ML algorithms are a tool for unearthing the rules of big data, and prediction models which incorporate them have exceptional accuracy in predicting kidney disease patients' poor prognosis during clinical practice. The use of ML algorithms can help clinicians detect patients at high risk of kidney function progression in the early stages. In this way, they can receive treatment and management in time. In sum, we suggest the gradual incorporation of ML algorithm-based prediction models into clinical practice.

Future Scope

- Machine Learning is the type of technology that keeps on evolving. There is a lot of scope in ML to become the top technology in the future. The reason is, it has a lot of research areas in it. This helps us to improve both **hardware** and **software**.
- In hardware, we have various laptops and GPUs. These have various ML and Deep Learning networks in them. These help in the faster processing power of the system. When it comes to software we have various UIs and libraries in use. These help in designing more efficient algorithms

- o By improving and using ML in clinical sector will have huge potential over medical issues
- o The medical issues can be prevented in time rather than comparing with traditional method of predicting
- o Introducing bid data in medical field can improve effective way of handling and using it in useful progress over human health
- o Further introducing blockchain in medial field can solve clinical data security and cause users a helpful way of growth by their own medical data

Appendix

Source code

Model Building - <https://github.com/mohamed8270/Machine-Learning-Kidney-Disease-Prediction-/tree/main/Kidney%20Chronic%20Model/Model%20Building>

Frontend Features - <https://github.com/mohamed8270/Machine-Learning-Kidney-Disease-Prediction-/tree/main/Kidney%20Chronic%20Model/Application%20Building>

GitHub

IBM Project repo - <https://github.com/IBM-EPBL/IBM-Project-48048-1660804120>

GitHub repo - <https://github.com/mohamed8270/Machine-Learning-Kidney-Disease-Prediction->

Demo Link - <https://www.mediafire.com/file/dlrc9826dgkj73i/Demo.mp4/file>

Video Link - <https://youtu.be/nTeslfRZvk8>