PROJECT REPORT DOCUMENTATION

EARLY DETECTION OF CHRONIC KIDNEY DISEASE
TEAM ID - PNT2022TMID06255

TEAM MEMBERS:
Bharath Bala Murugan R
Karthick P
Kumaresan S
Madhan kumar S

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. Chronic Kidney Disease can be cured, if treated in the early stages. The main aim of the project is to predict whether the patient have chronic kidney disease or not in a painless, accurate and faster way based on certain diagnostic measurement like Blood Pressure(BP), Albumin(Al) etc., and then appropriate treatment can be given based on the details provided by the model.

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 :-

SNO	LITERATURE PAPER	AUTHOR	PROPOSED METHOD	ACCURACY	YEAR
1	Computer-Aided 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. Alturayeif			
4	Prediction of Chronic Kidney Disease Using Machine Learning Algorithm	Siddheshwar Tekale,Pranjal Shingavi,Sukanya 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,Hajra FahimSyed, Nuha Radwan Khayat,Rawan KanaanAltheeb,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,Sunday O.Olatunji			

2.3 Problem Statement Definition

Chronic kidney disease (CKD) is one of the most critical health problems due to its increasing prevalence. Chronic kidney disease, also known as chronic renal disease or CKD, is a condition characterized by a gradual loss of kidney function over time. It includes conditions that damage your kidneys and decrease their ability to keep you healthy by filtering wastes from your blood. Diabetes and high blood pressure, or hypertension, are responsible for two-thirds of chronic kidney disease cases. Anyone can get chronic kidney disease at any age. However, some people are more likely than others to develop kidney disease. Most people may not have any severe symptoms until their kidney disease is advanced.

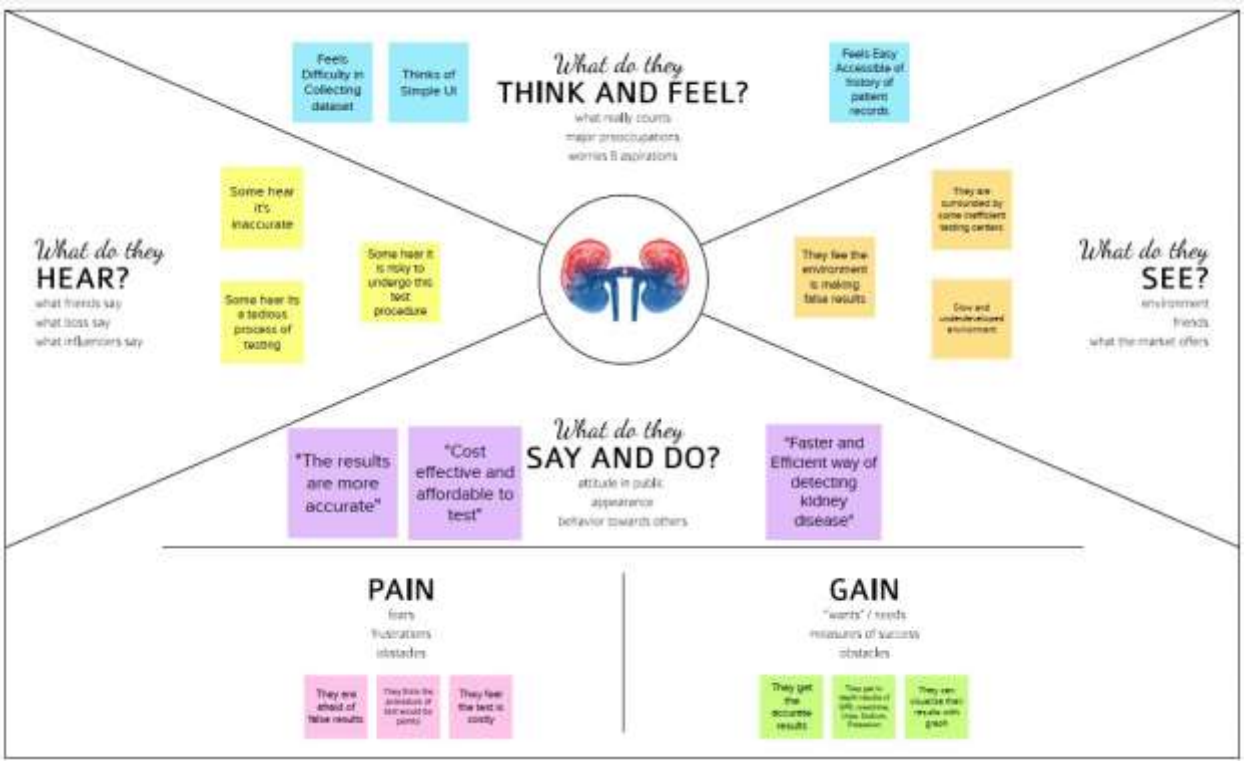
A better testing method which could possibly detect CKD in the early stages would be much more useful. Medical test results taken for other purposes are used to detect CKD at early stages. Various efforts have been undertaken to advance early therapy to prevent the condition from progressing to chronic disease. Recent research suggests that some of the negative outcomes can be avoided with early identification and treatment. Peculiar and contributing attributes from

the above-mentioned test results are combined to develop a Machine Learning Model.

This Machine Learning Model will be used to predict CKDs rather early than the presently existing methods.

3. IDEATION & PROPOSED SOLUTION

3.1. Empathy Map Canvas:-



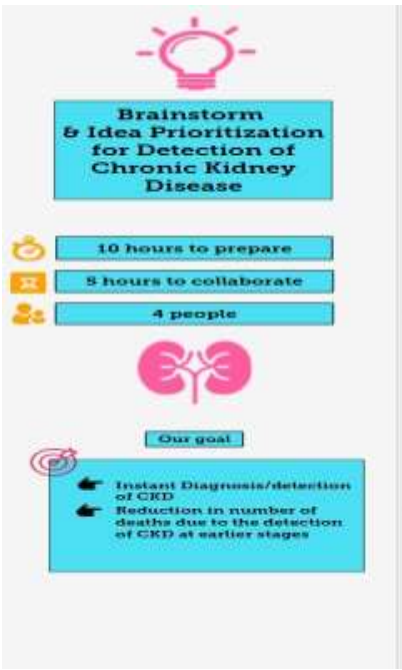
3.2 Ideation & Brainstorming:-

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.

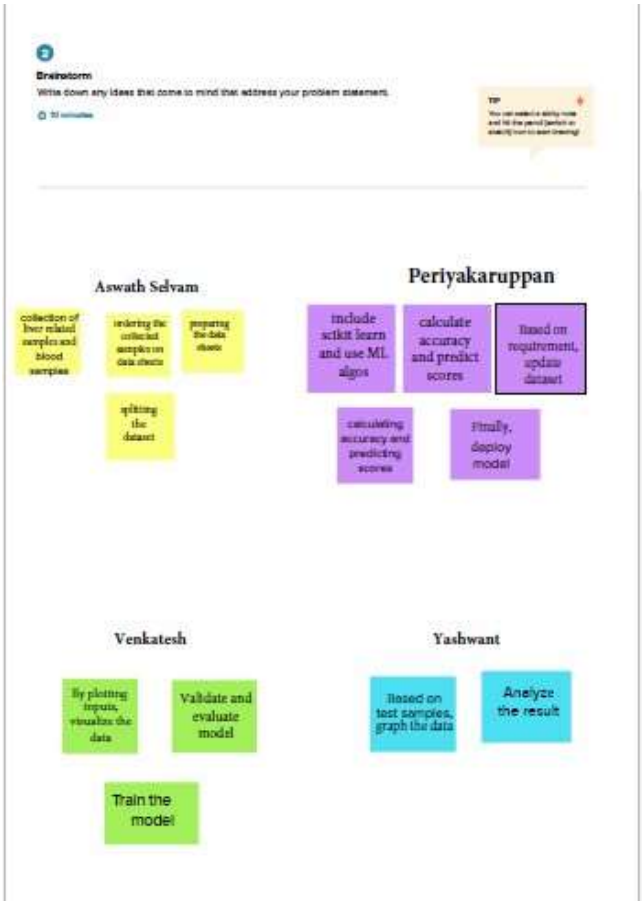
Ideation is often closely related to the practice of brainstorming, a specific technique that is utilized to generate new ideas. A principal difference between ideation and brainstorming is that **ideation is commonly more thought of as being an individual pursuit, while brainstorming is almost always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


Brainstorm & Idea Prioritization for Detection of Chronic Kidney Disease
10 hours to prepare
5 hours to collaborate
4 people
Our goal
Instant Diagnosis/detection of CKD
Reduction in number of deaths due to the detection of CKD at earlier stages

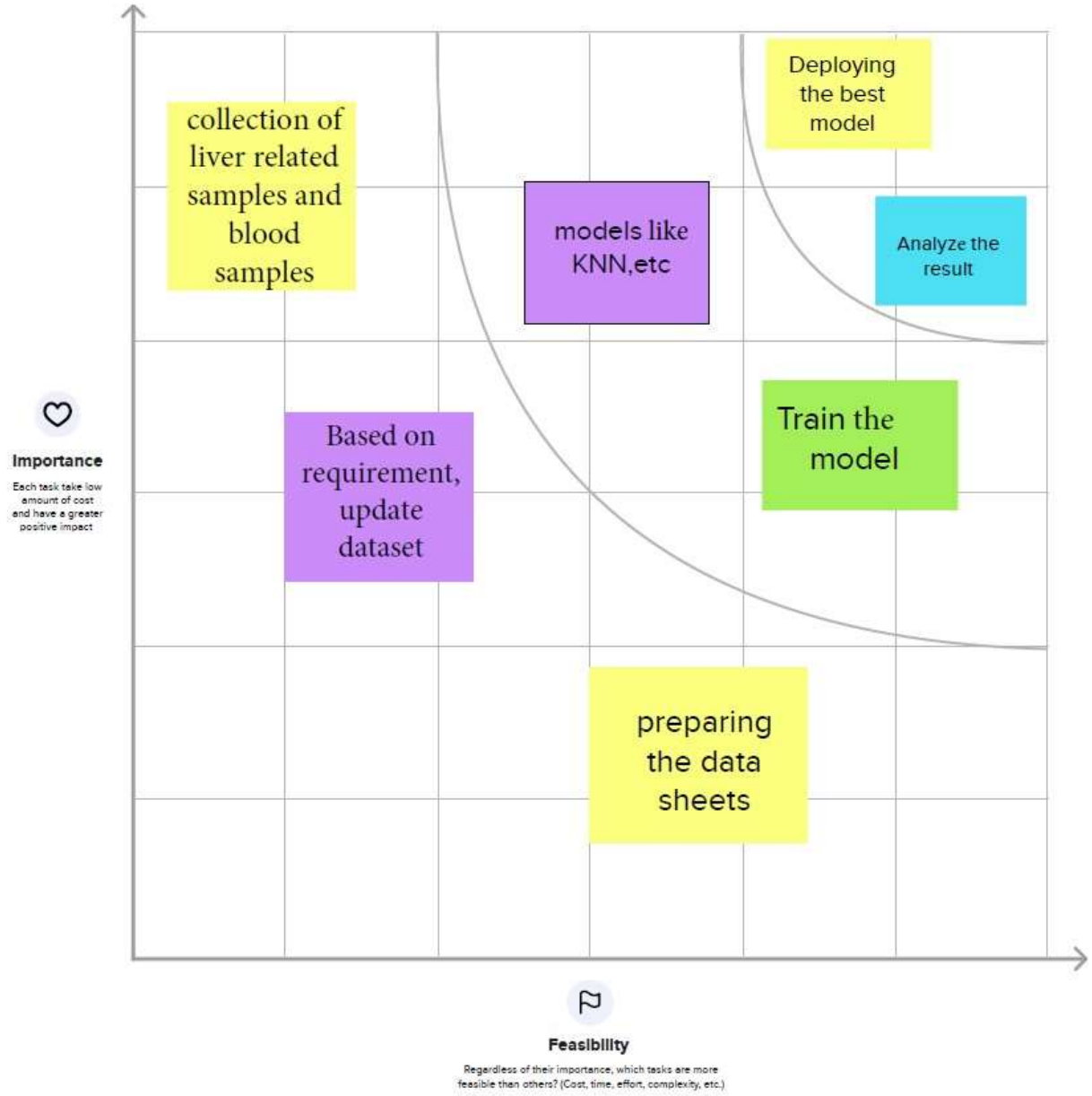

Before you collaborate
Reading various IEEE reference papers regarding CKD to gather detail information.
30 minutes
Team gathering
All 4 members in our team gathered in a meeting to breifly discuss about the problem statement.
Setting the goal
Our goal is to find a good solution to the problem of CKD by using various ML Models
Learning how to use the various tools to achieve the goal
Learning various pre-requisite tools to achieve our goal


Define your problem statement
Chronic kidney disease prediction is one of the most important issues in healthcare analytics.The most interesting and challenging taskadaylic-day lives as one third of American Adult population is affected by chronic kidney disease (CKD), and millions die each year because they do not have access to affordable treatment. Chronic Kidney disease can be cured, if treated in the early stages. The main aim of this project is to predict whether the patient have chronic kidney disease or not in a painless, accurate and faster way based on certain diagnostic measurement like Blood Pressure (Bp), Albumin(A), etc., and then appropriate treatment can be given based on the details provided by the model.
Key rules of brainstorming
To run an smooth and productive session
Stay in topic.
Encourage wild ideas.
Defer judgment.
Listen to others.
Go for volume.
If possible, be visual.
Problem
How might we diagnose the issues in a human kidney using the applications of Machine Learning and Data Science ?


2
Brainstorm
Write down any ideas that come to mind that address your problem statement.
Aswath Selvam
collection of liver related samples and blood samples
ordering the collected samples on data sheets
preparing the data sheets
splitting the dataset
Periyakaruppan
include scikit learn and use ML algos
calculate accuracy and predict scores
Based on requirement, update dataset
calculating accuracy and predicting scores
Finally, deploy model
Venkatesh
By plotting inputs, visualize the data
Validate and evaluate model
Train the model
Yashwant
Based on test samples, graph the data
Analyze the result


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 and break it up into smaller sub-groups.
Data collection and processing
collection of liver related samples and blood samples
preparing the data sheets
splitting the dataset
Visualizing the data by plotting inputs
Datasets and model
Based on requirement, update dataset
include scikit learn and use ML algos
models like KNN,etc
Training the model
Analyzing and Deployment
calculating accuracy and predicting scores
Analyze the result
Based on test samples, graph the data
Deploying the best model

Stage 4 - prioritize



3.3 Proposed Solution:-

The purpose of this tool is to provide a structured process for identifying a problem, understanding the root causes ascertaining solutionsteps, andprogress monitoring.

With a solution template, you can organize development content that you want to reuse for customer-specific solutions. Solution templates enable you 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)	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
2.	Idea / Solution description	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
3.	Novelty / Uniqueness	• Easy to use User interface (UI) • accurate accuracy by comparing the performance of different ml model technique
4.	Social Impact / Customer Satisfaction	• Greater cost reduction in hospitals for testing • Helps in early diagnosis of the disease • Chances of recovery is higher
5.	Business Model (Revenue Model)	• subscription based model with initial trial basis • charges/commission for the actual prediction and recovery of a person
6.	Scalability of the Solution	• 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

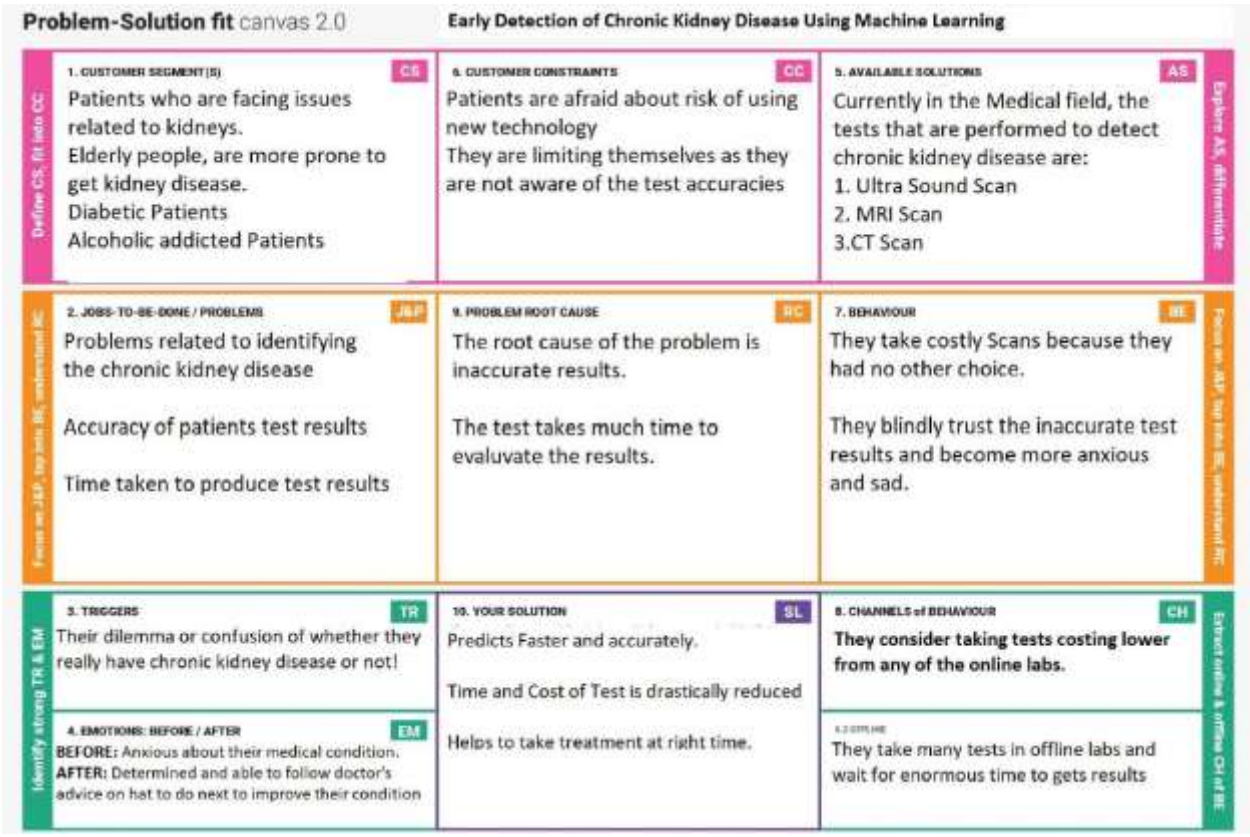
Key Partners	Key Activities	Value Propositions	Customer Relationships	Customer Segments
Our Key partners have a latest testing laboratory to collect samples of blood and urine to analyse various parameters with a testing capacity of almost 1000 persons a day. Our ML Models can test a large number of 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. **Key Resources** Our key resources are testing laboratories, ML Models and their data. To create an excel sheet from the data.	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. **Channels** 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.	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.

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 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 ANALYSIS4.1 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:

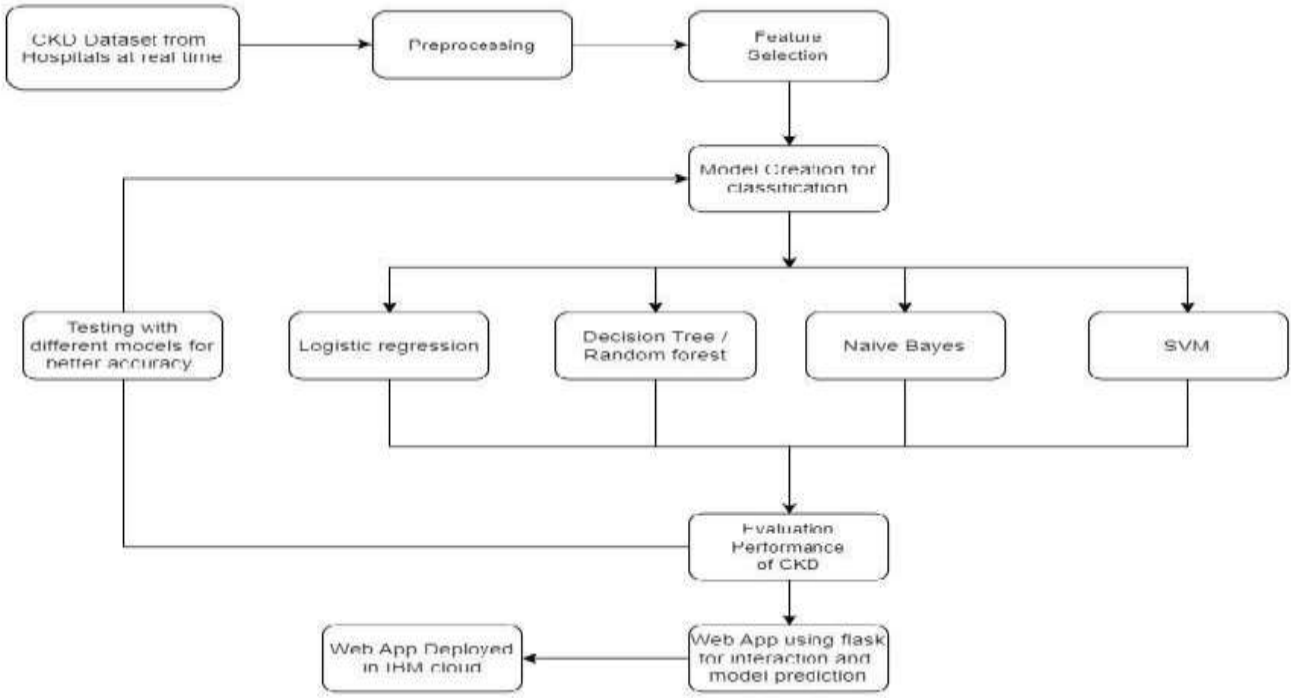
Following are the non-functional requirements of the proposed solution.

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

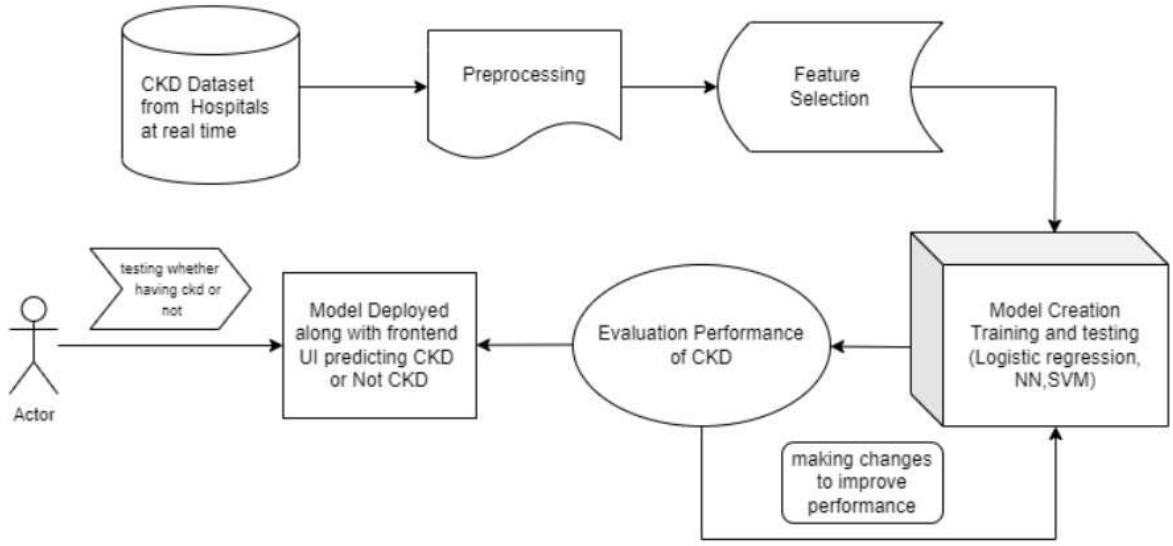
5.1 Data Flow Diagrams:-

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.



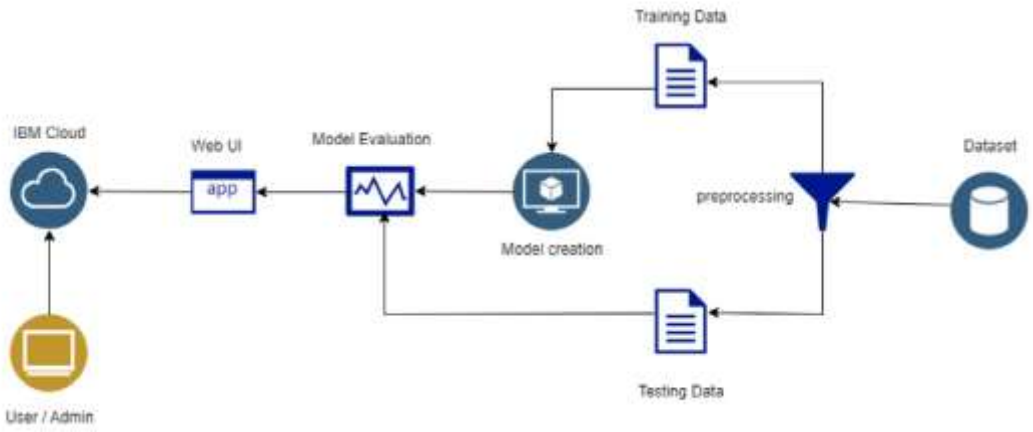
4.2 Solution & Technical Architecture: Solution

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 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.

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 , *React (Subsidiary)*
2.	Application Logic-1 (Registration)	User is redirected to register page for registering themselves by providing valid details	HTML, CSS , Flask , *React (Subsidiary)*
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 , *React (Subsidiary)*
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 , *React (Subsidiary)*
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	User Registration	USN-1	As a user, I can register for the application by entering my name, mobile number, email, password, and confirming my password.	10	High	Bharath Bala Murugan R Karthick P Kumaresan S Madhan Kumar S
Sprint-2		USN-2	As a user, I can register for the application through Gmail	5	Medium	Bharath Bala Murugan R Karthick P Kumaresan S Madhan Kumar S
Sprint-1	User Confirmation	USN-3	As a user, I will receive confirmation email once I have registered for the application	10	High	Bharath Bala Murugan R Karthick P Kumaresan S Madhan Kumar S
Sprint-2		USN-4	As a user, I will receive confirmation otp to verify the identity.	5	High	Bharath Bala Murugan R Karthick P Kumaresan S Madhan Kumar S

Sprint-2	Data Collection	USN-5	As a user, I will enter the input data for disease prediction in the form	10	High	Bharath Bala Murugan R Karthick P Kumaresan S Madhan Kumar S

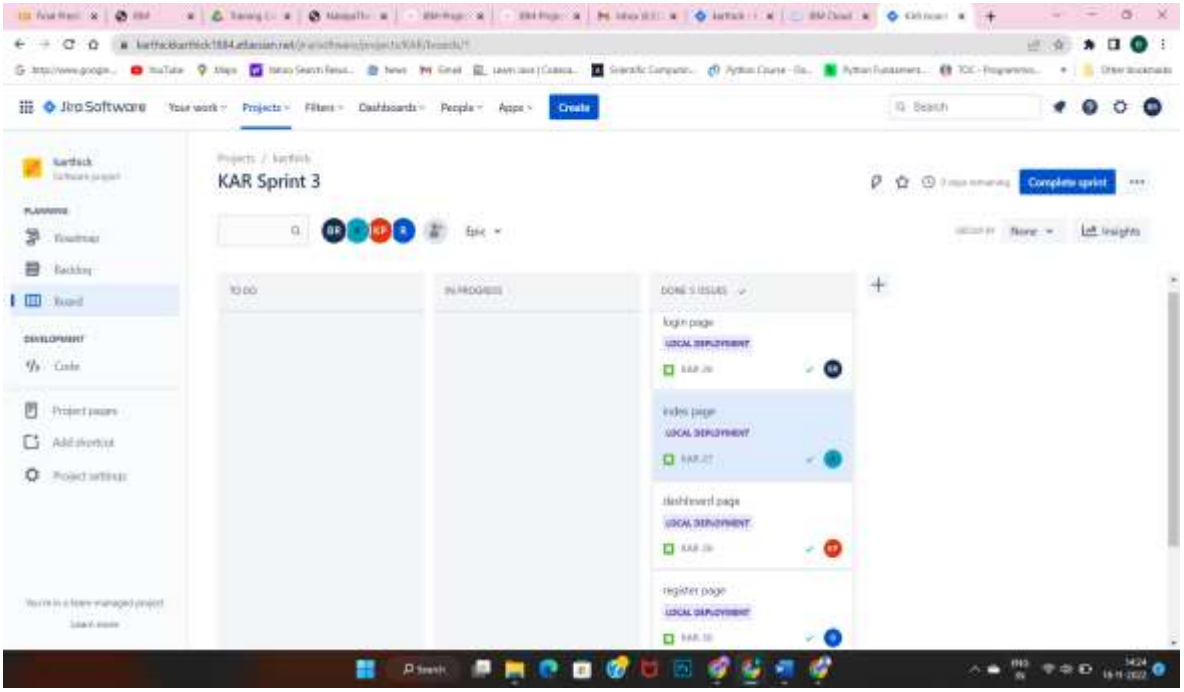
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-3	Provide output to the user	USN-6	As a user, I will get the result of disease prediction in the dashboard.	10	High	Bharath Bala Murugan R Karthick P Kumaresan S Madhan Kumar S
Sprint-3	Data Analysis	USN-7	As the admin, I will develop modules to preprocess and store the data.	10	High	Bharath Bala Murugan R Karthick P Kumaresan S Madhan Kumar S
Sprint-4	Prediction of disease	USN-8	As the admin, I will build a Machine Learning model to predict the disease	10	High	Bharath Bala Murugan R Karthick P Kumaresan S Madhan Kumar S
Sprint-4	Final Delivery	USN-9	Deploy the application in IBM cloud and make it available for use.	10	High	Bharath Bala Murugan R Karthick P Kumaresan S Madhan Kumar S

Velocity:

We have a 6-day sprint duration, and the velocity of the team is 20 (points per sprint). The team's average velocity (AV) per iteration unit (story points per day)

AV = Sprint duration / velocity = 20 / 6 = 3.33

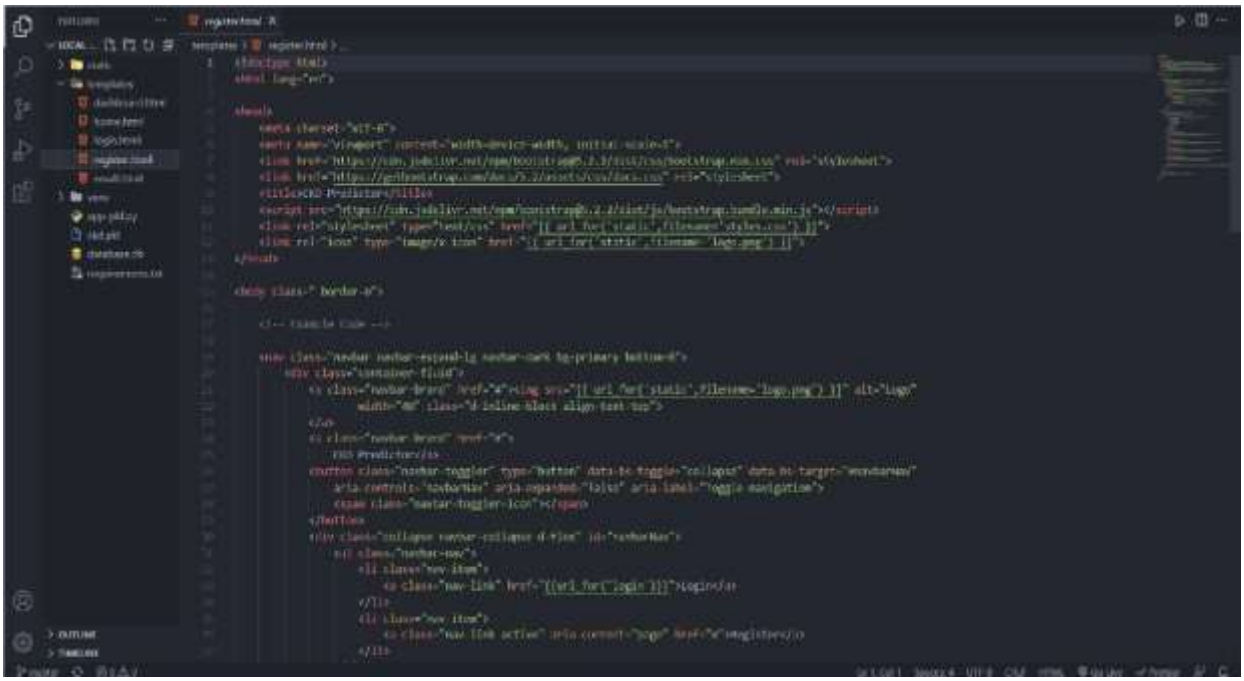
6.3 Reports from JIRA

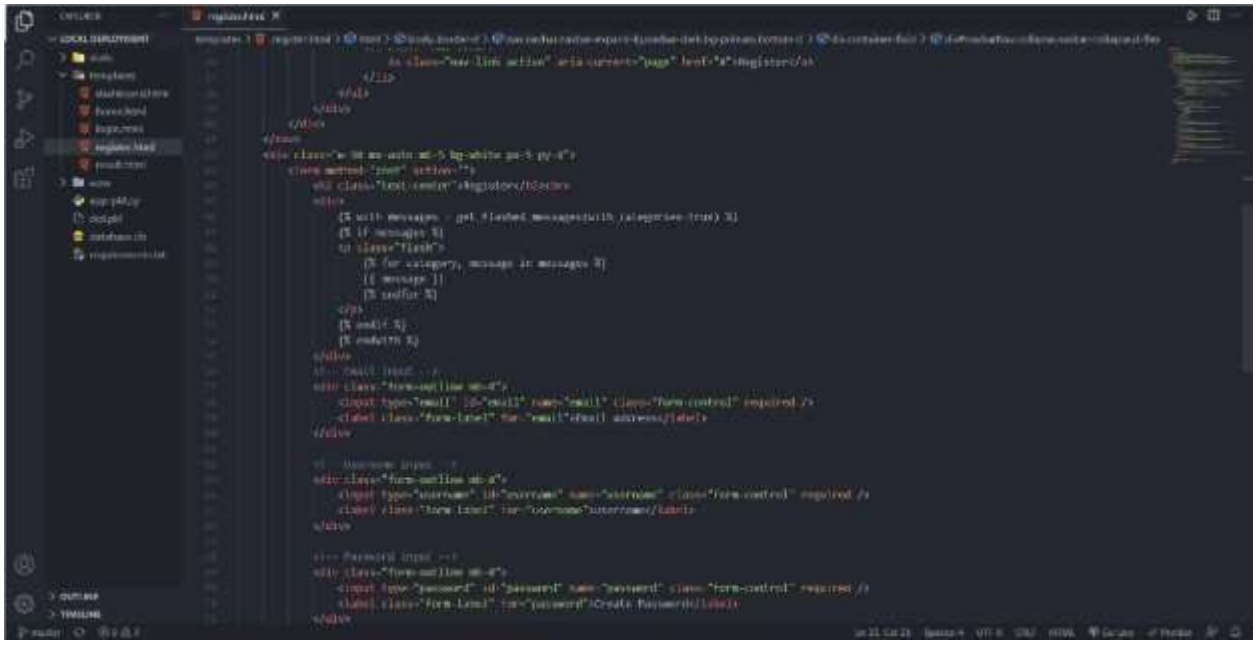


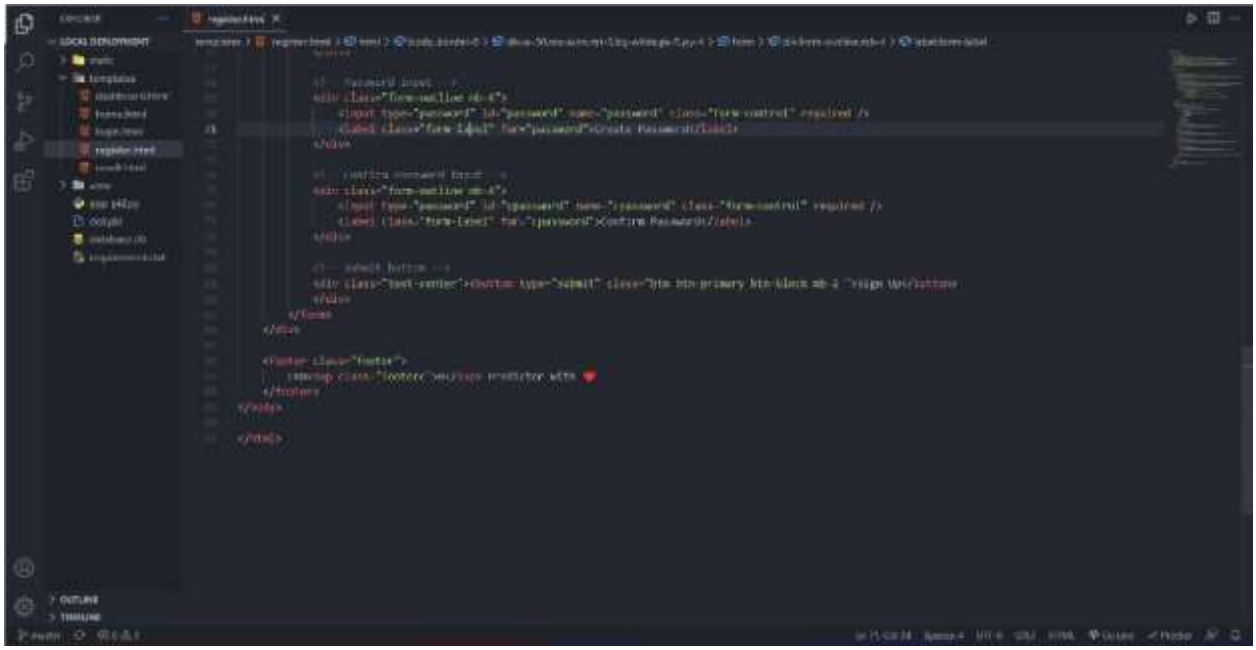
7. CODING & SOLUTIONING

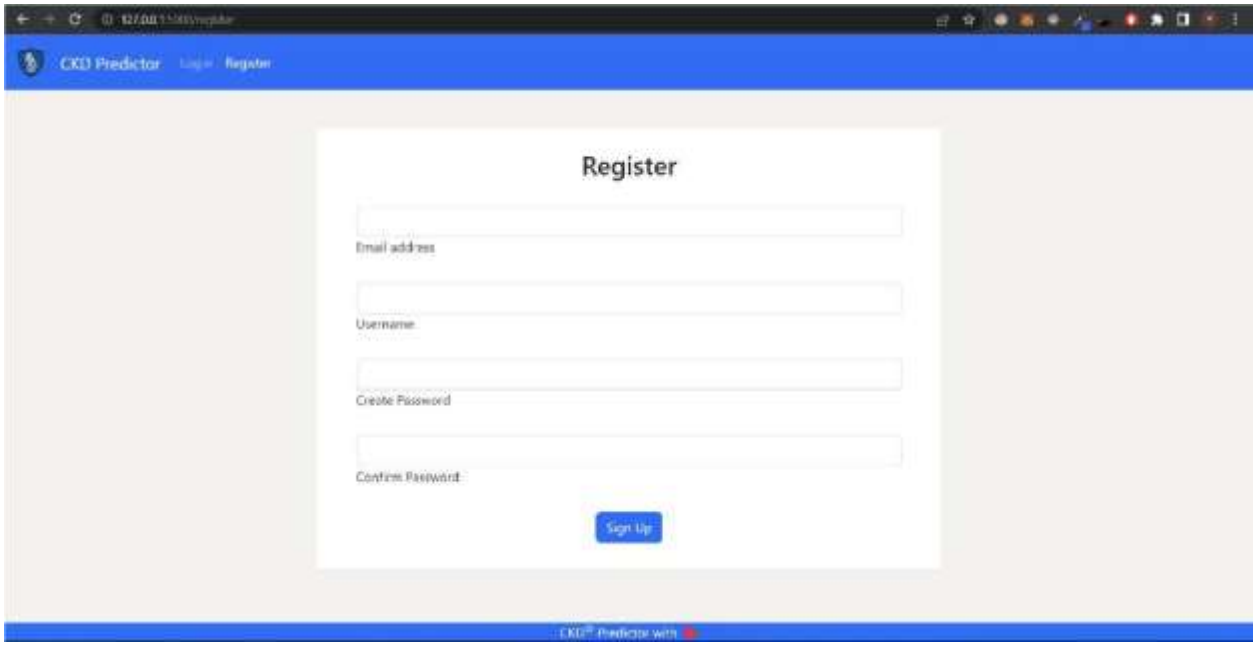
7.1 User Registration and login :

Register.html:

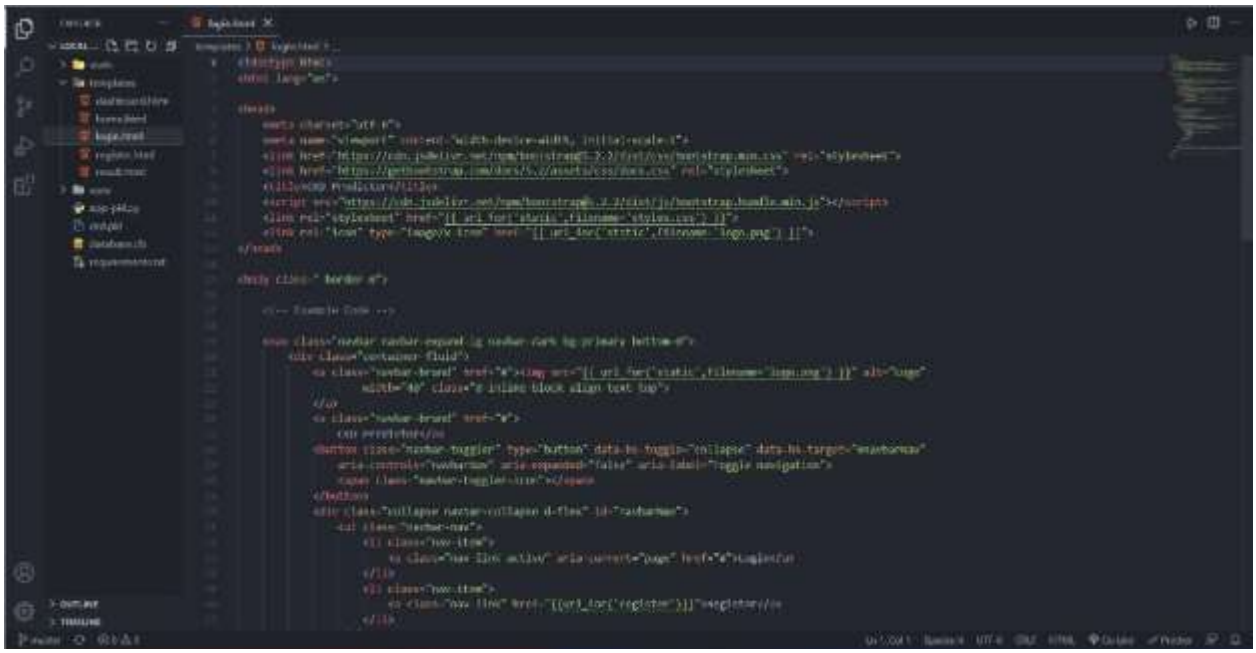


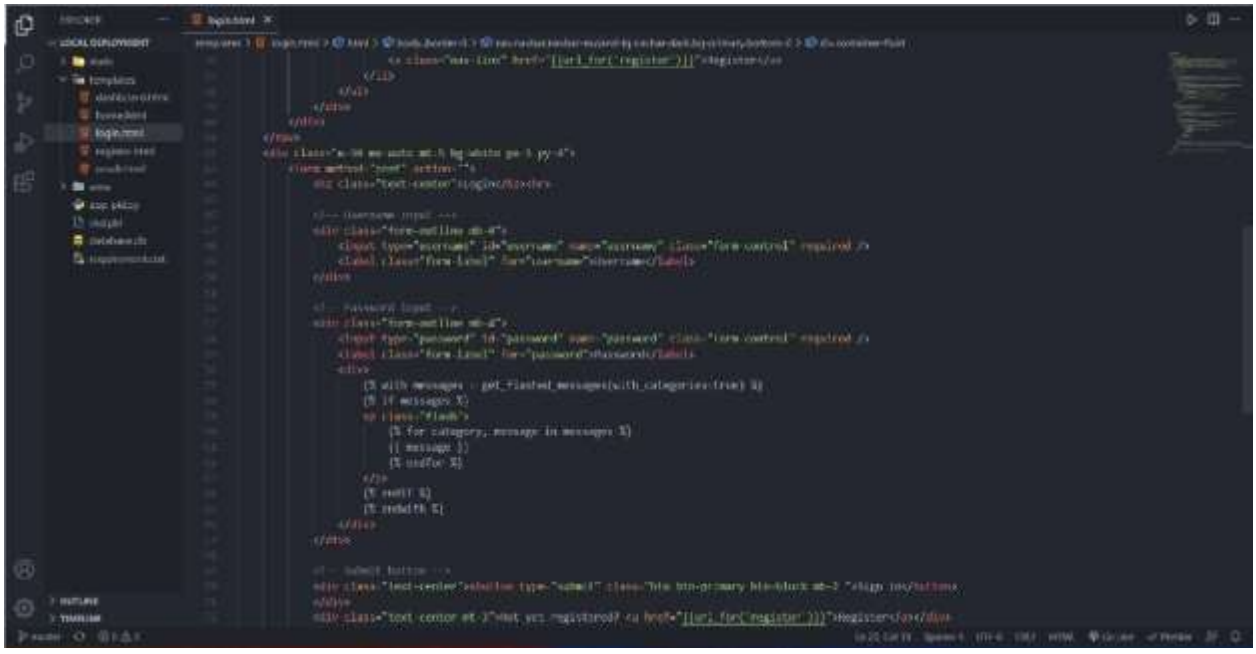


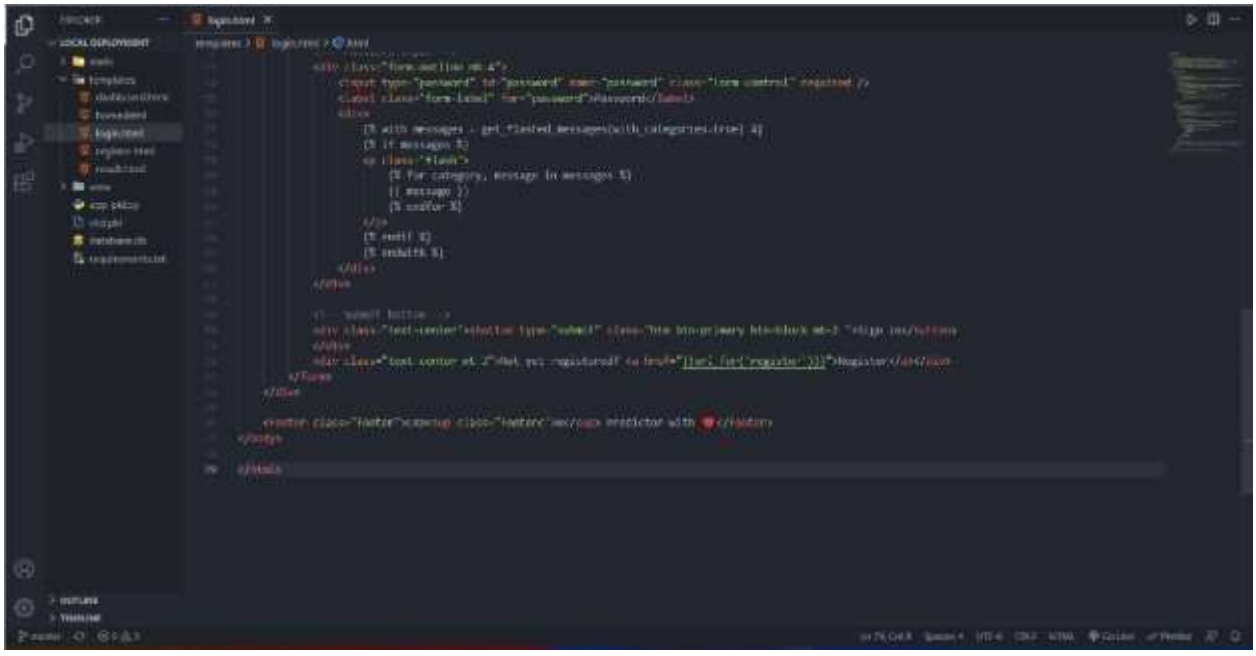


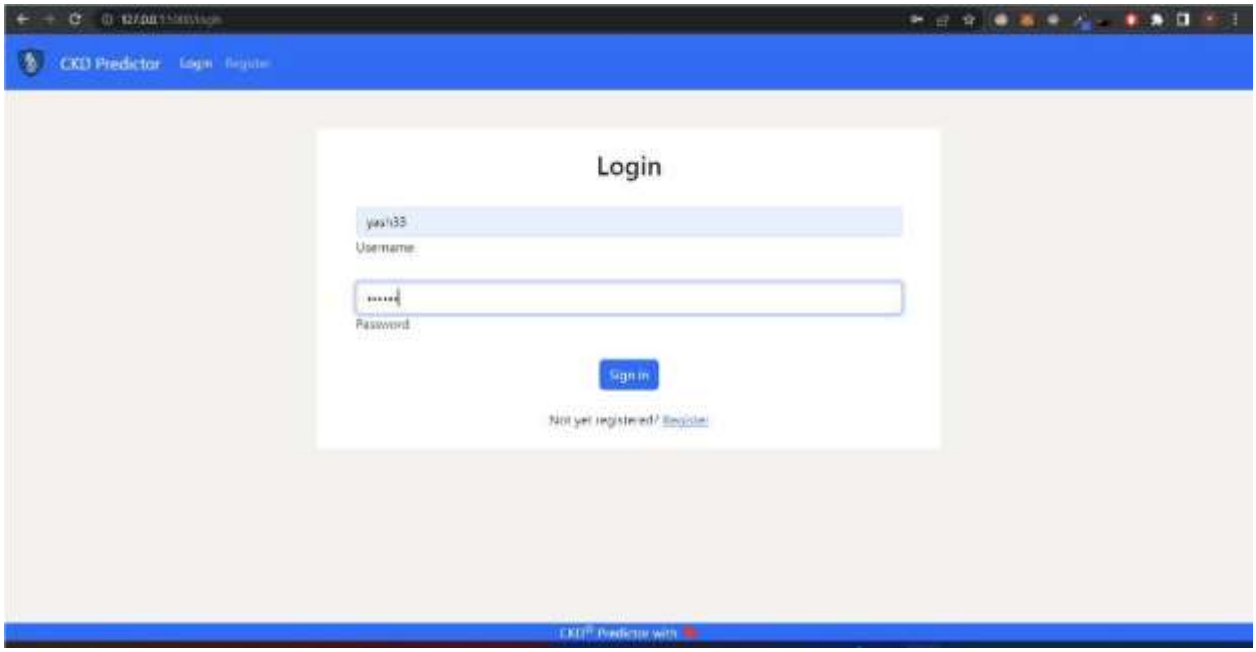

CKD Predictor
Register
Email address
Username
Create Password
Confirm Password
Sign Up

Login.html



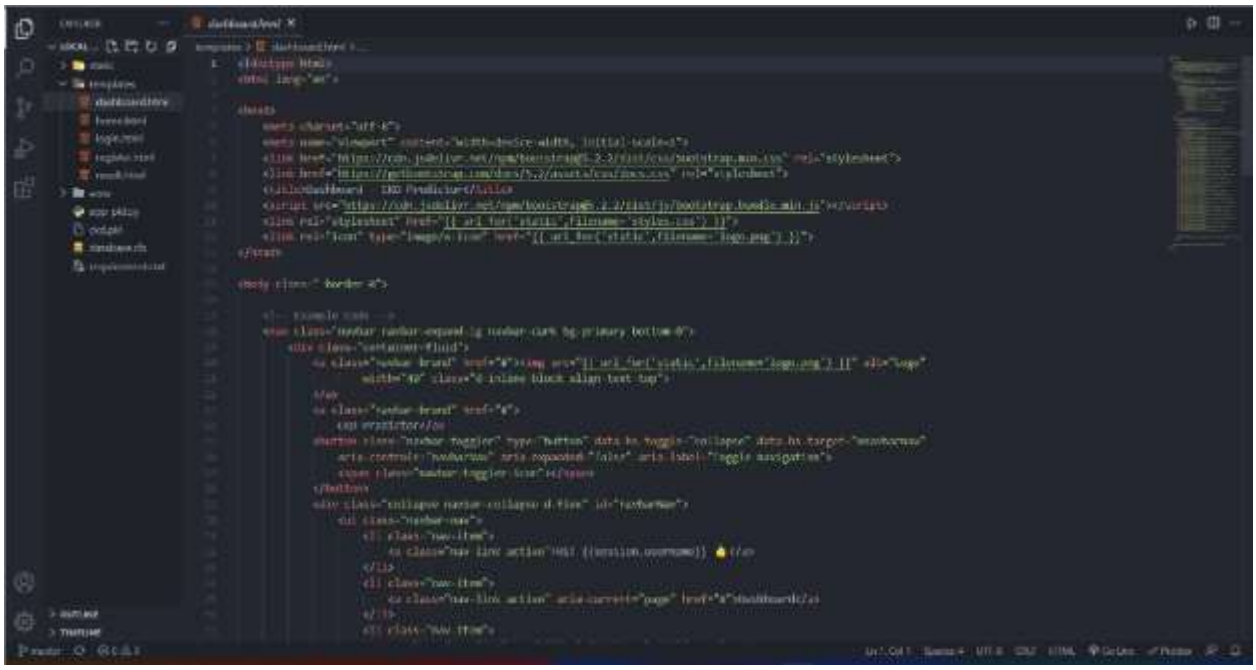


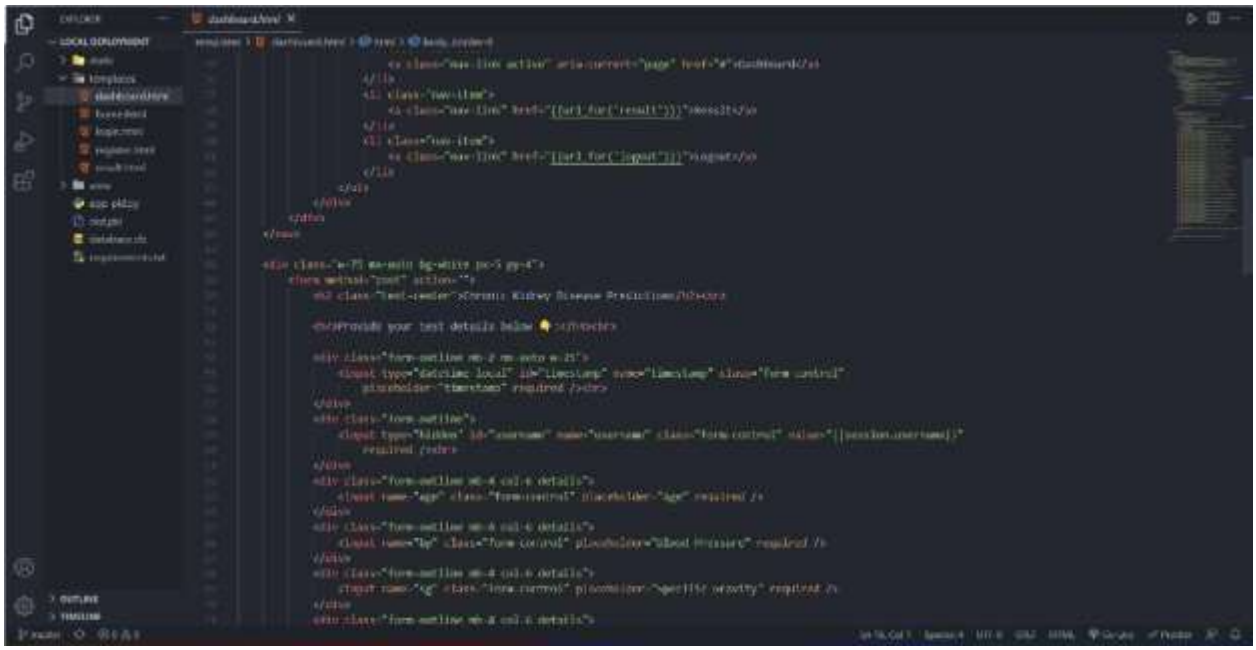


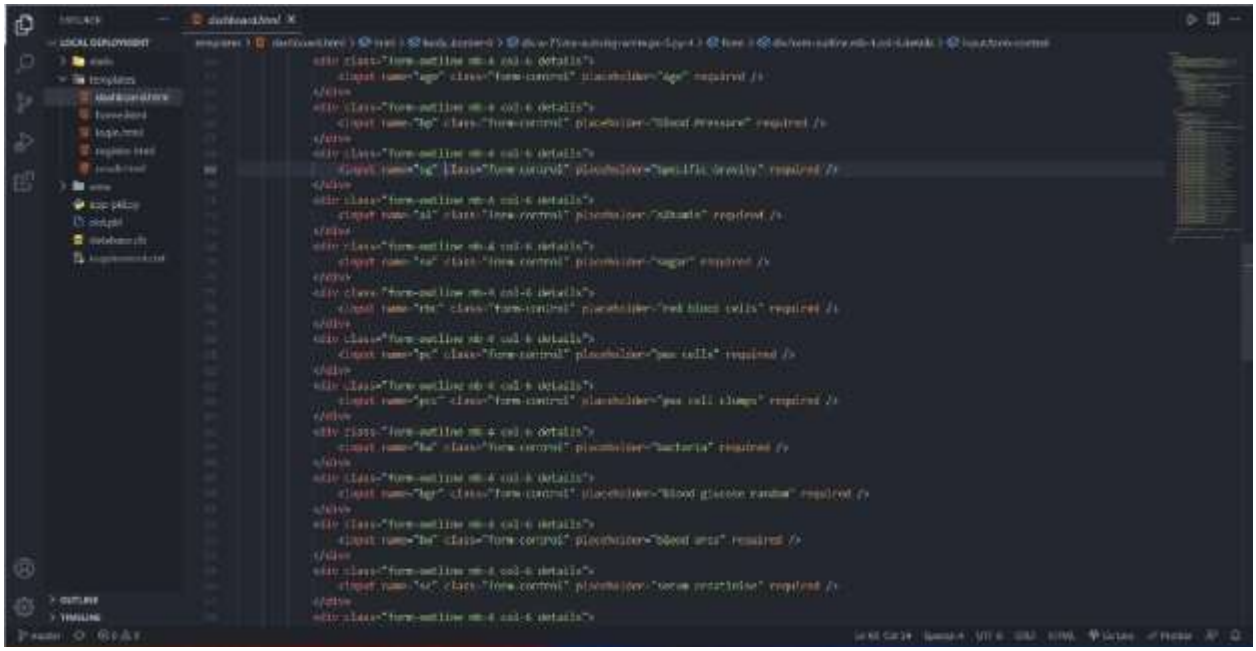


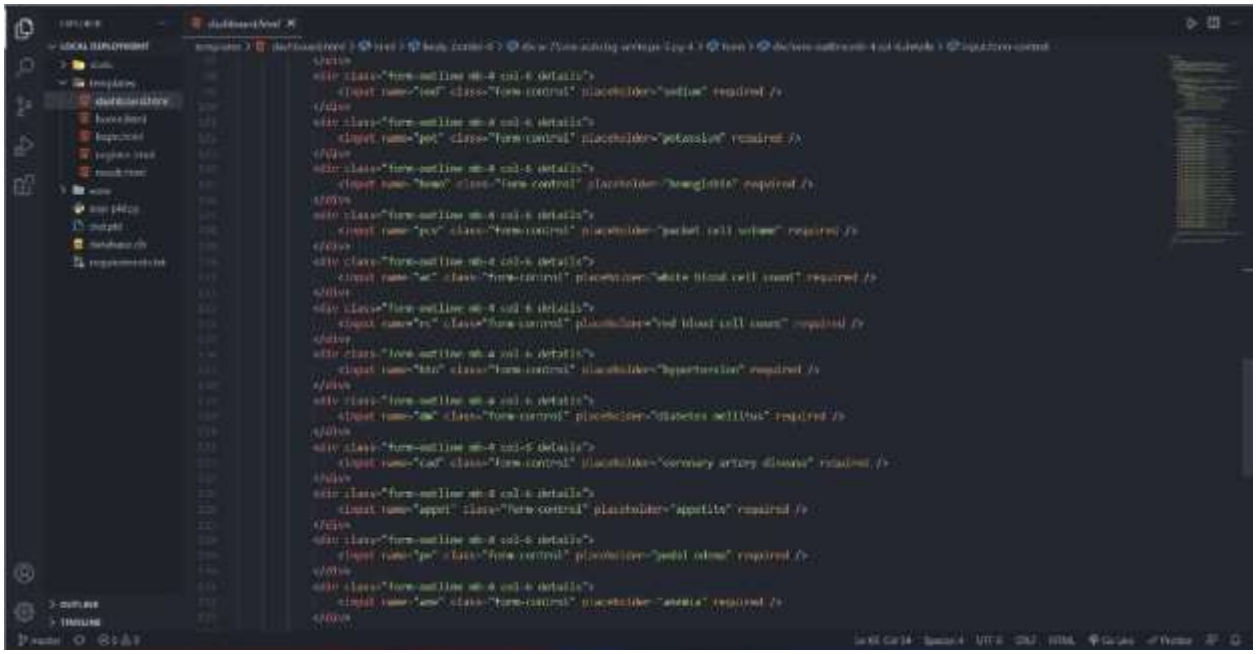
7.2 Dashboard and Result

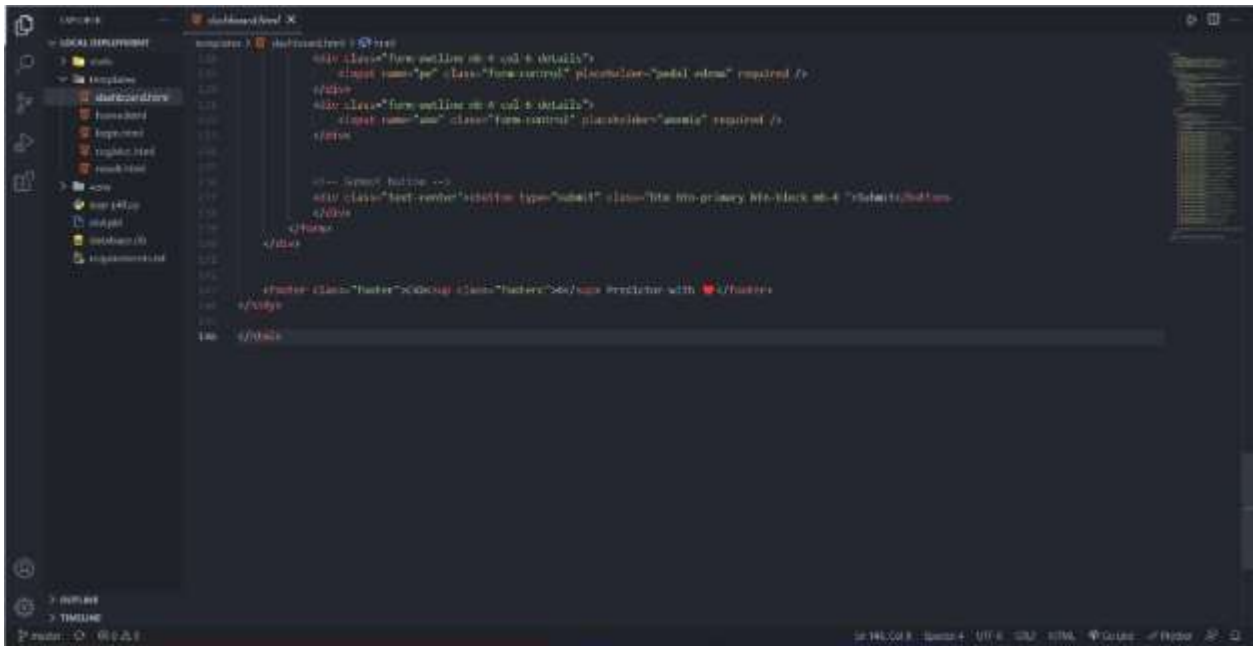
Dashboard.html

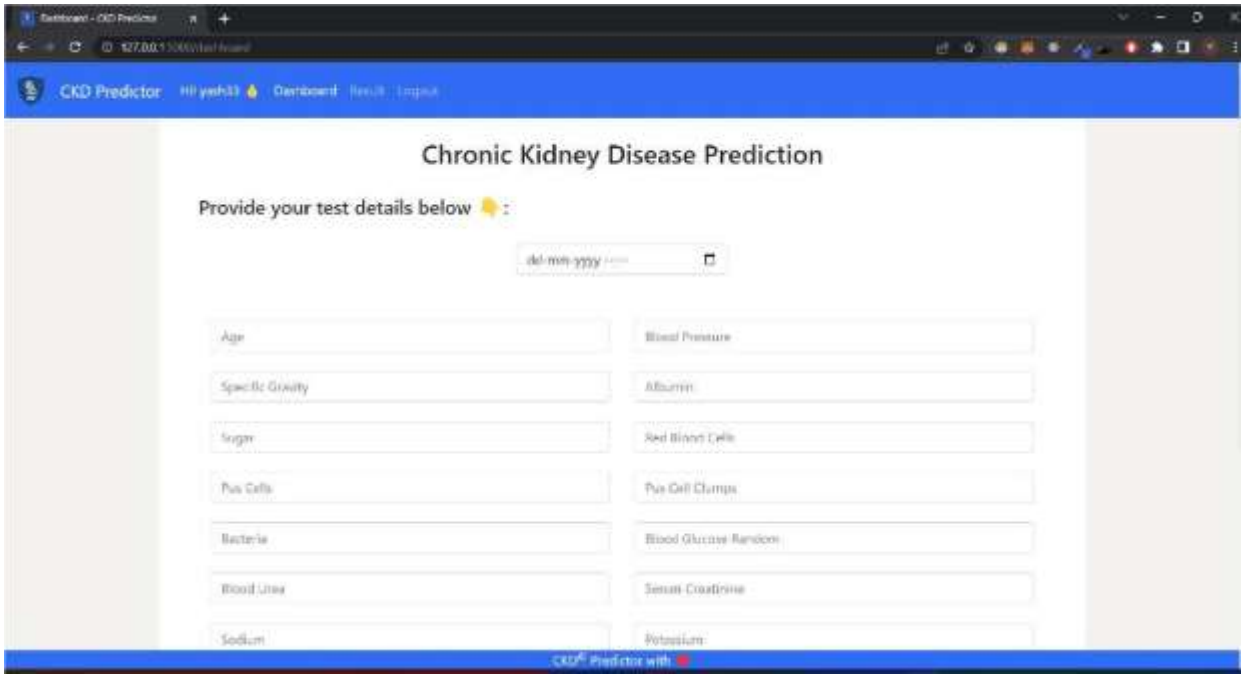




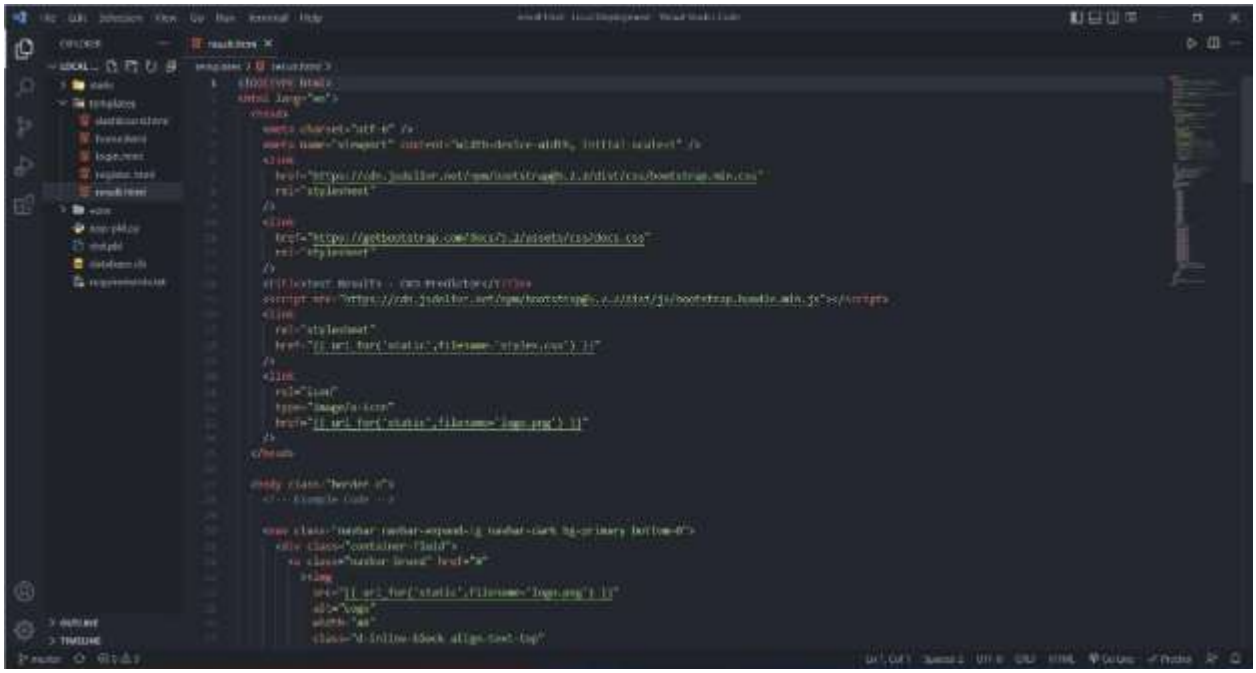


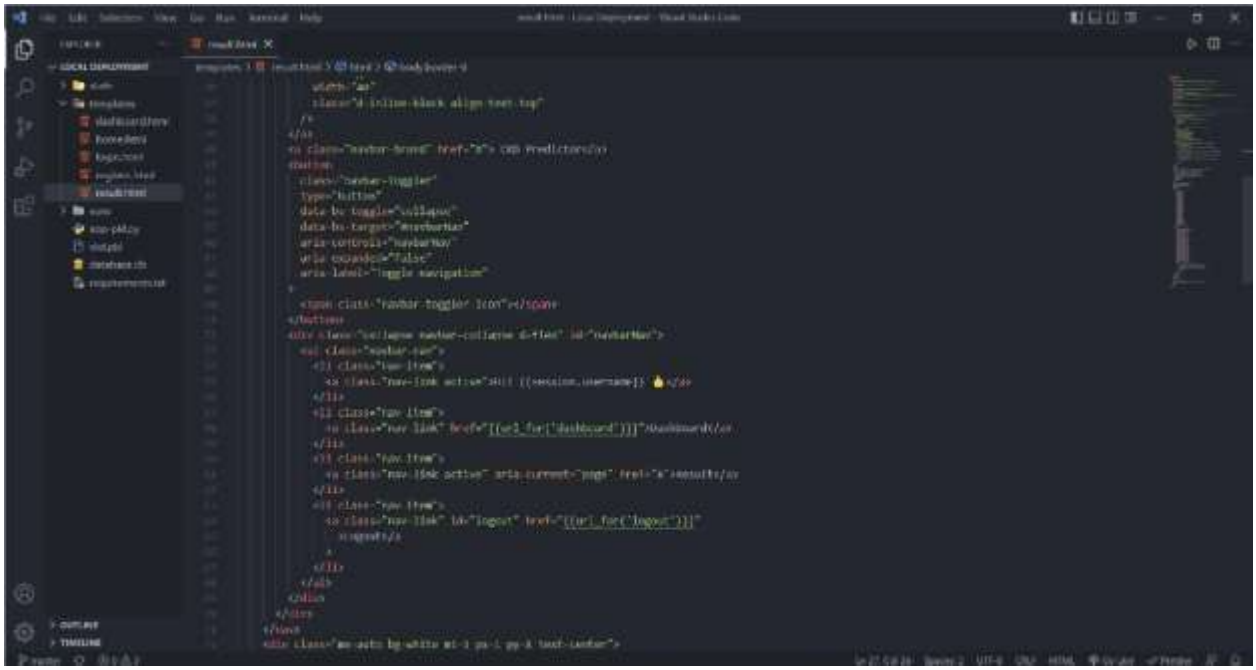


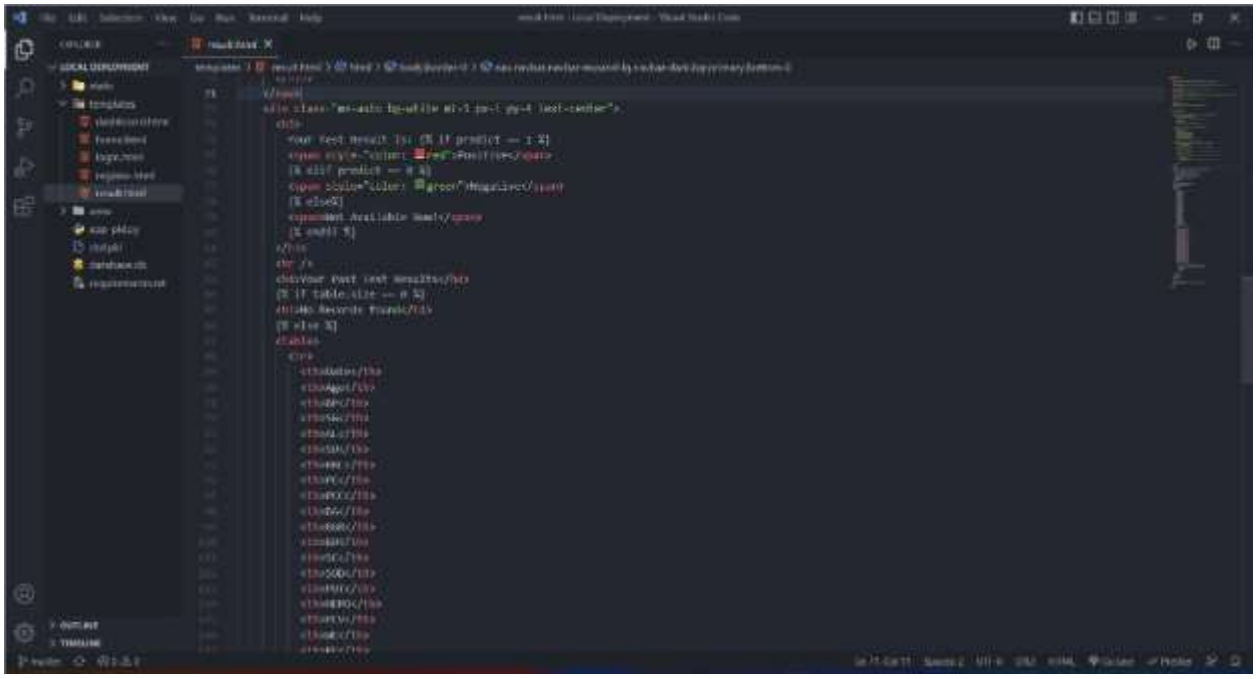


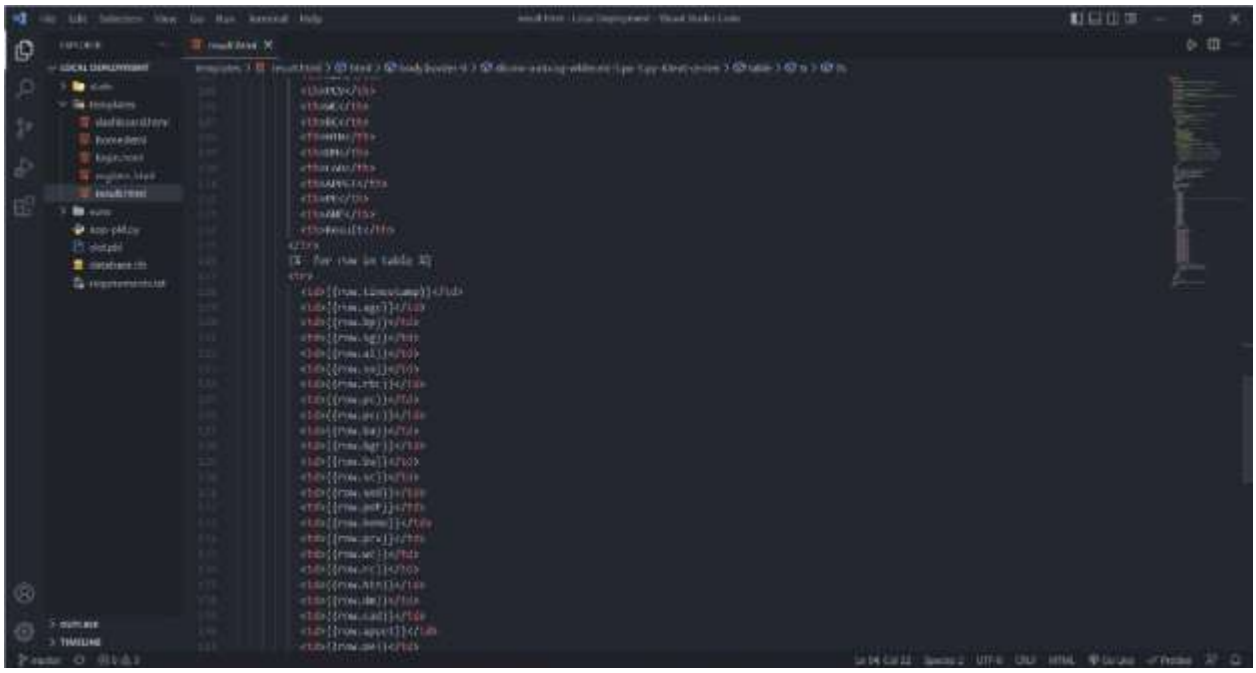

CKD Predictor
Chronic Kidney Disease Prediction
Provide your test details below 👇 :
Age
Blood Pressure
Specific Gravity
Albumin
Sugar
Red Blood Cells
Pus Cells
Pus Cell Clumps
Bacteria
Blood Glucose Random
Blood Urea
Serum Creatinine
Sodium
Potassium

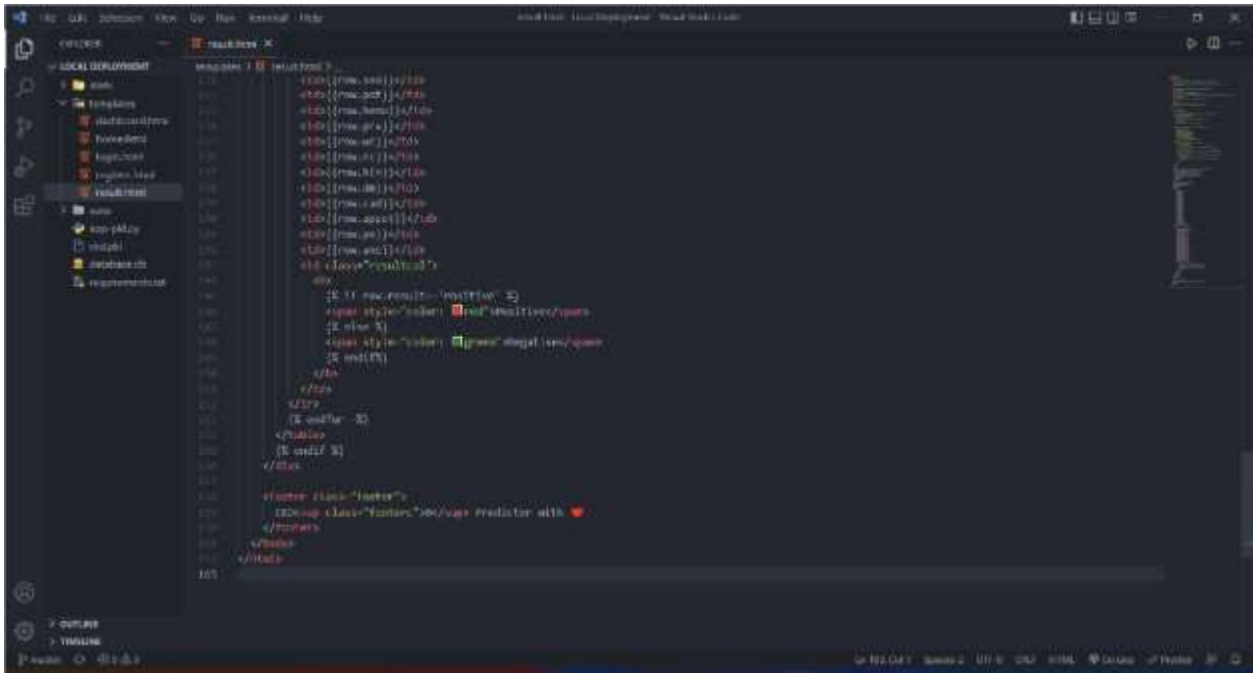
Result.html

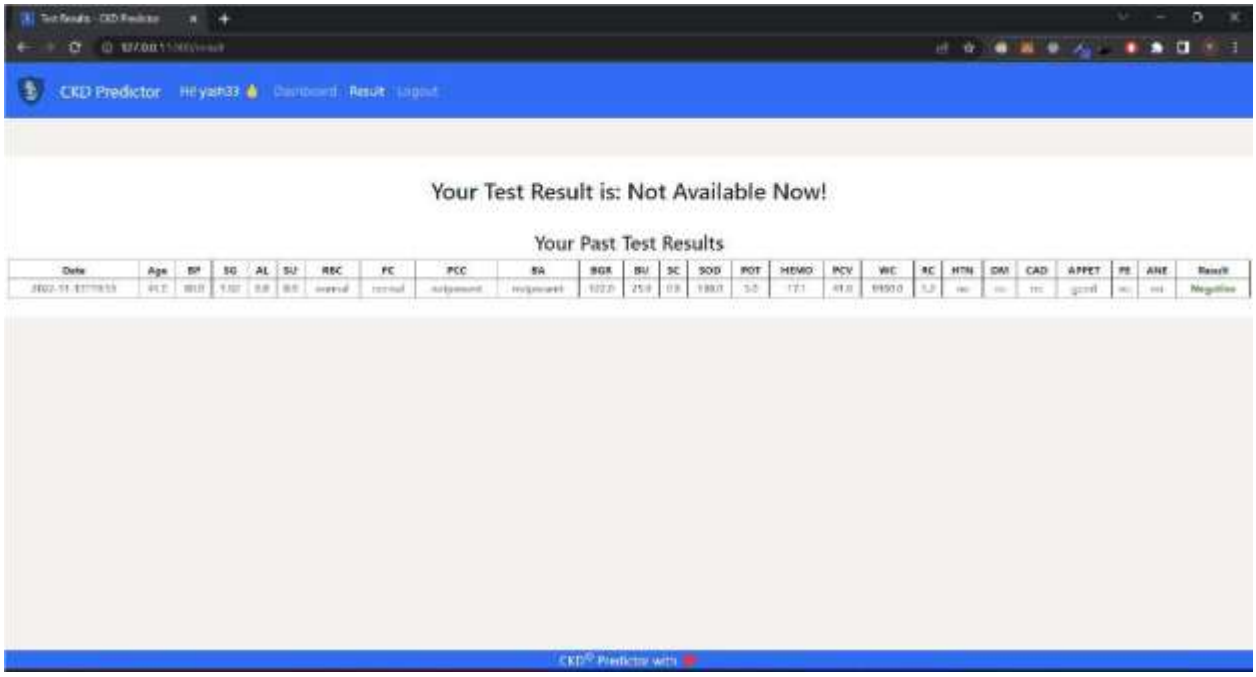




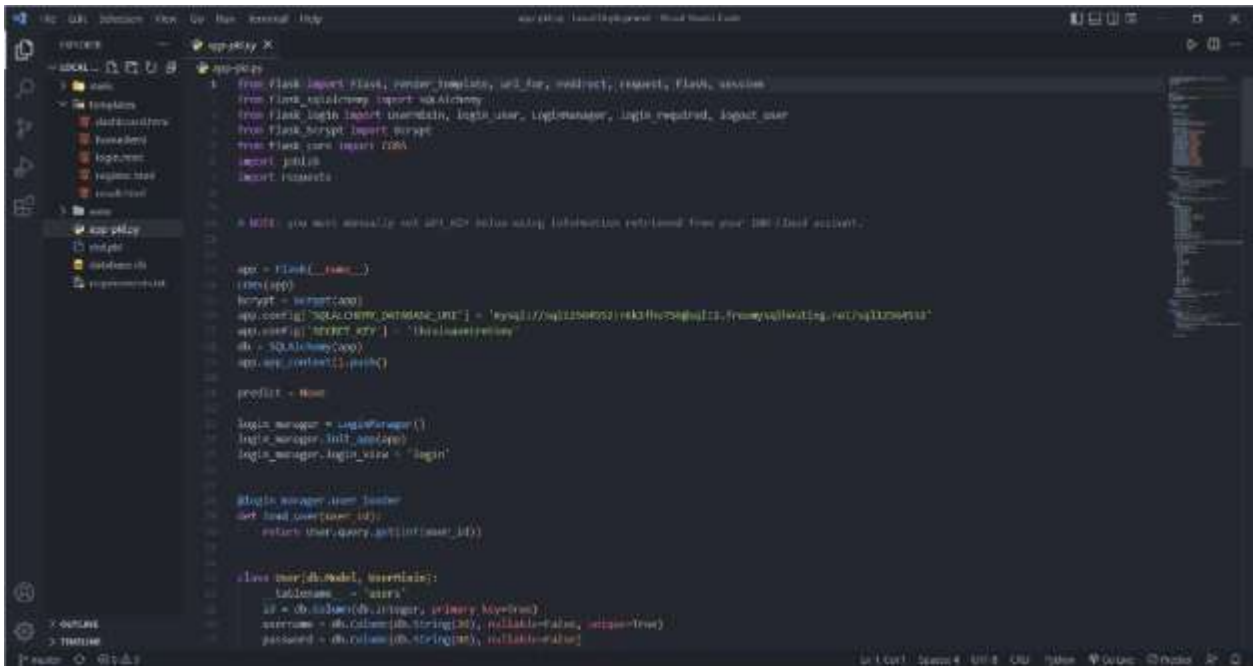


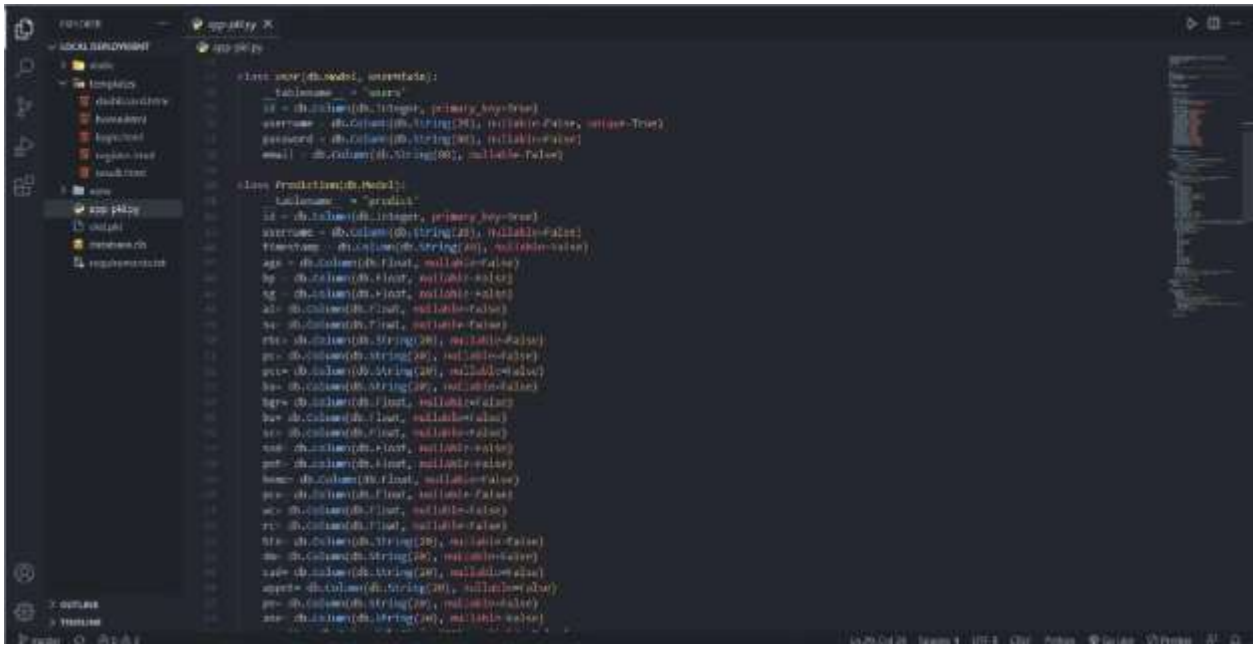


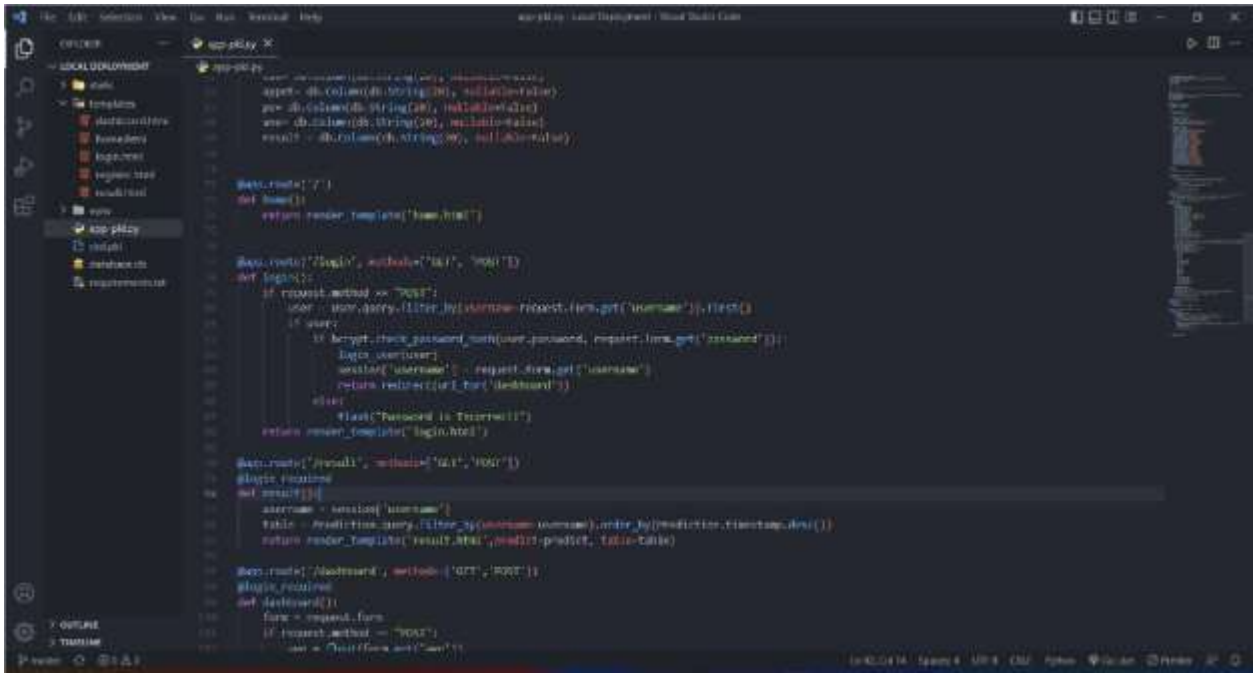


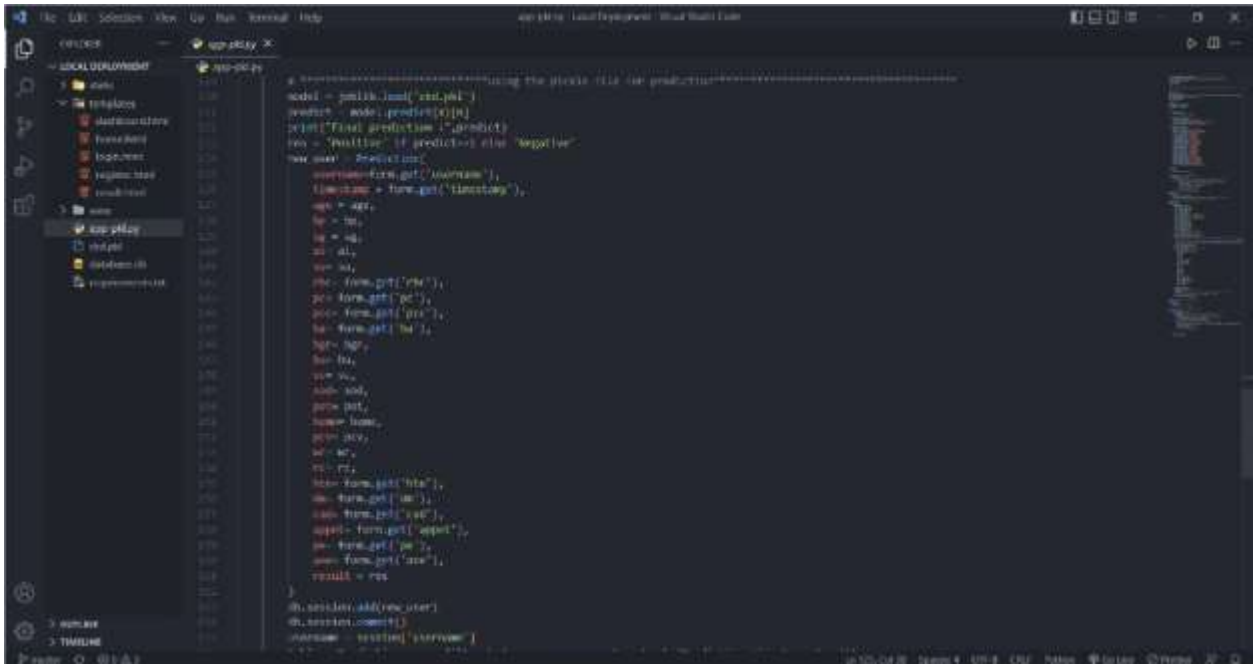


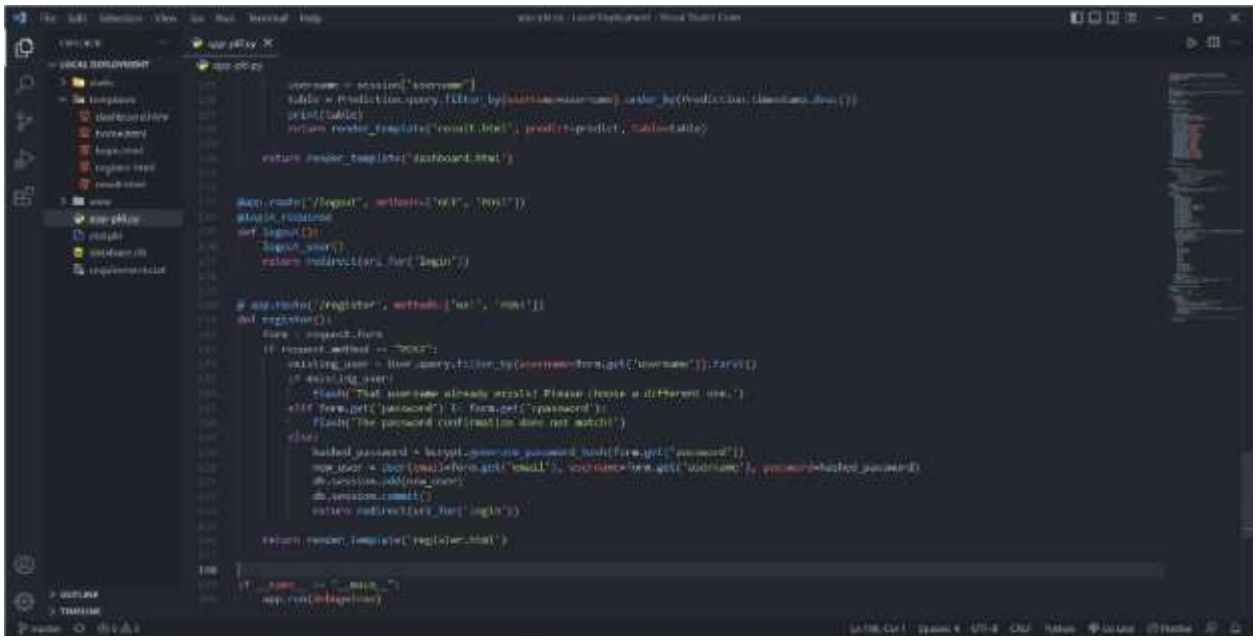
7.3 Flask Integration and Deployment





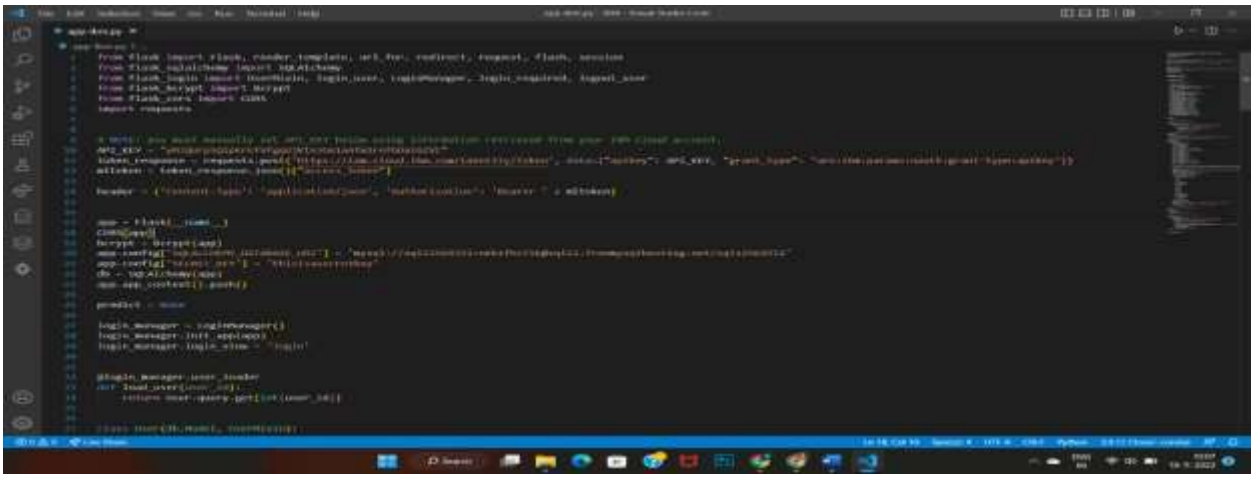


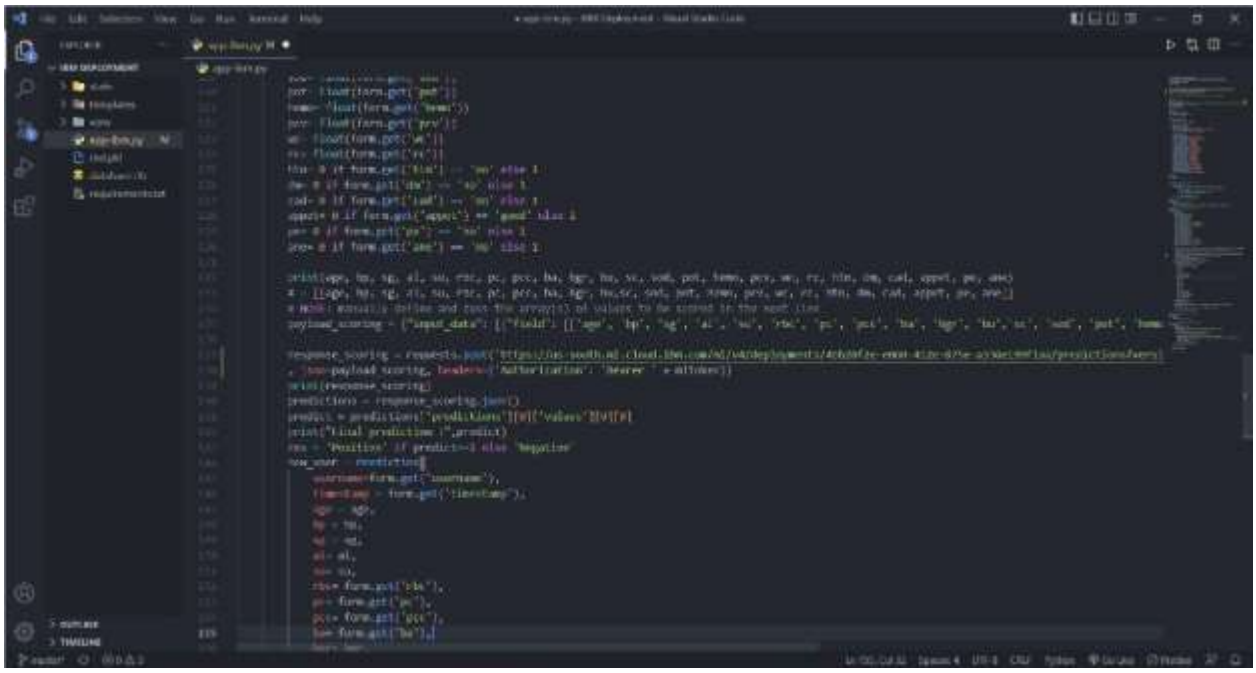


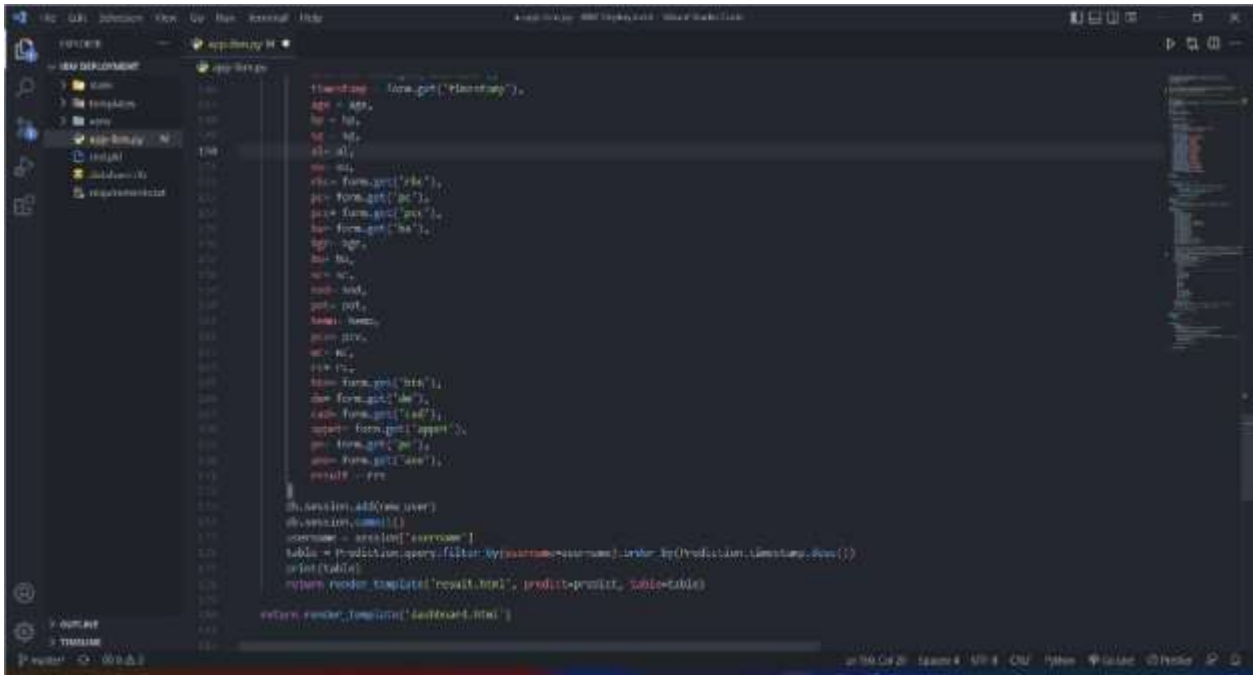


Using pickle to integrate with flask

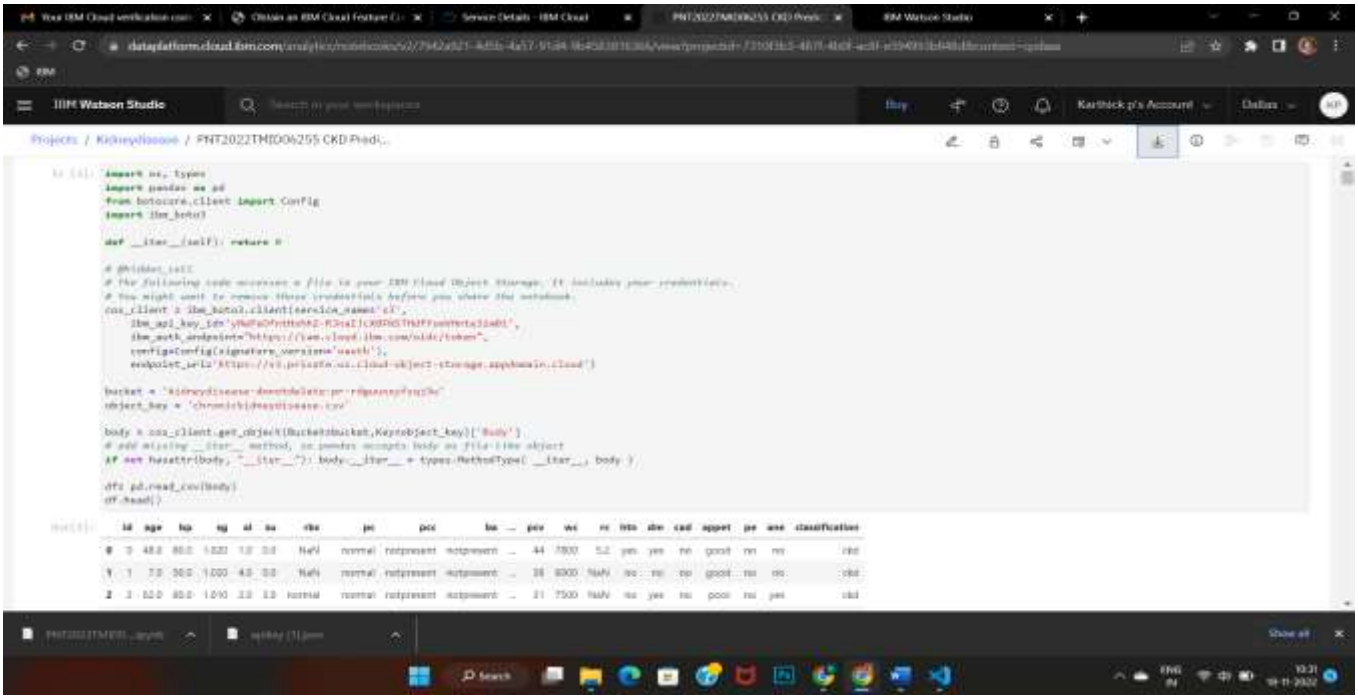
Flask changes for ibm deployment

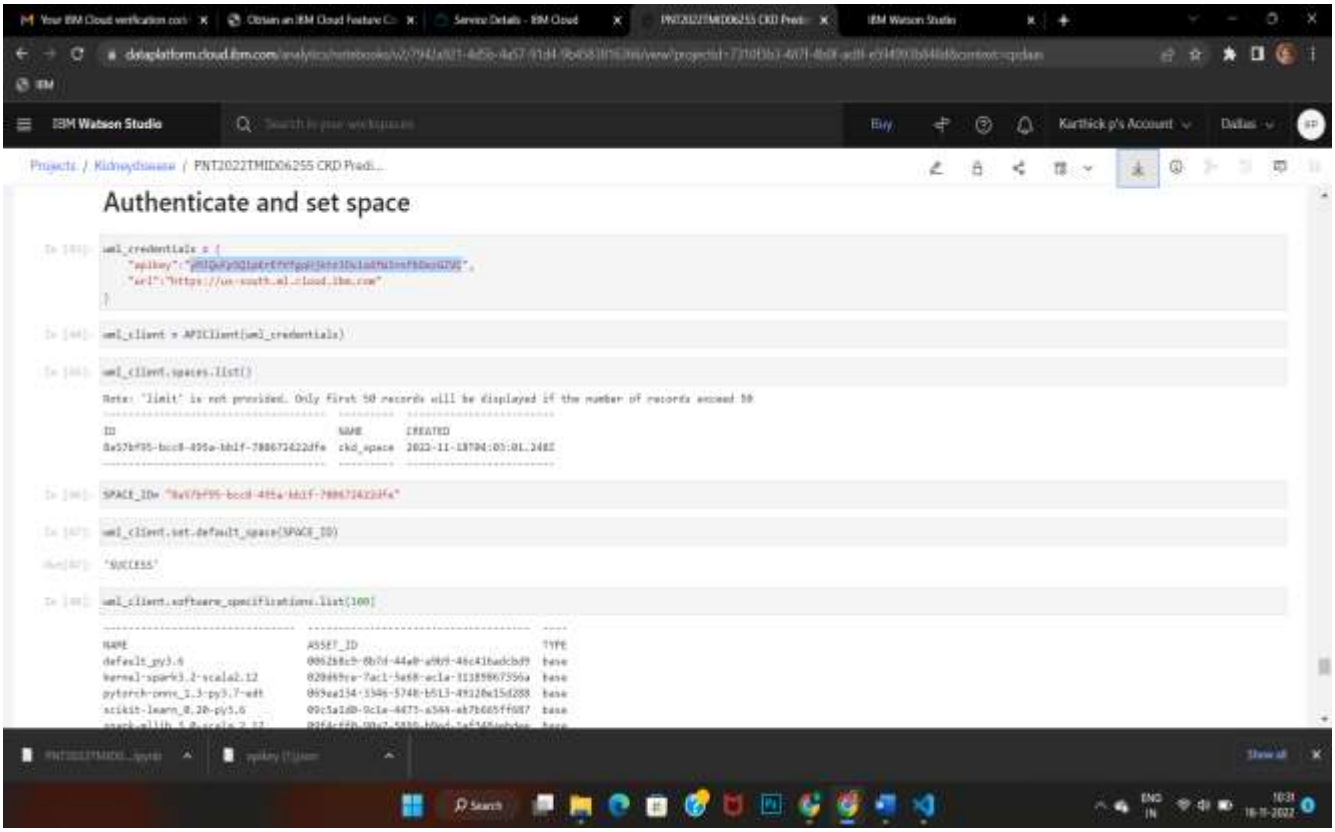


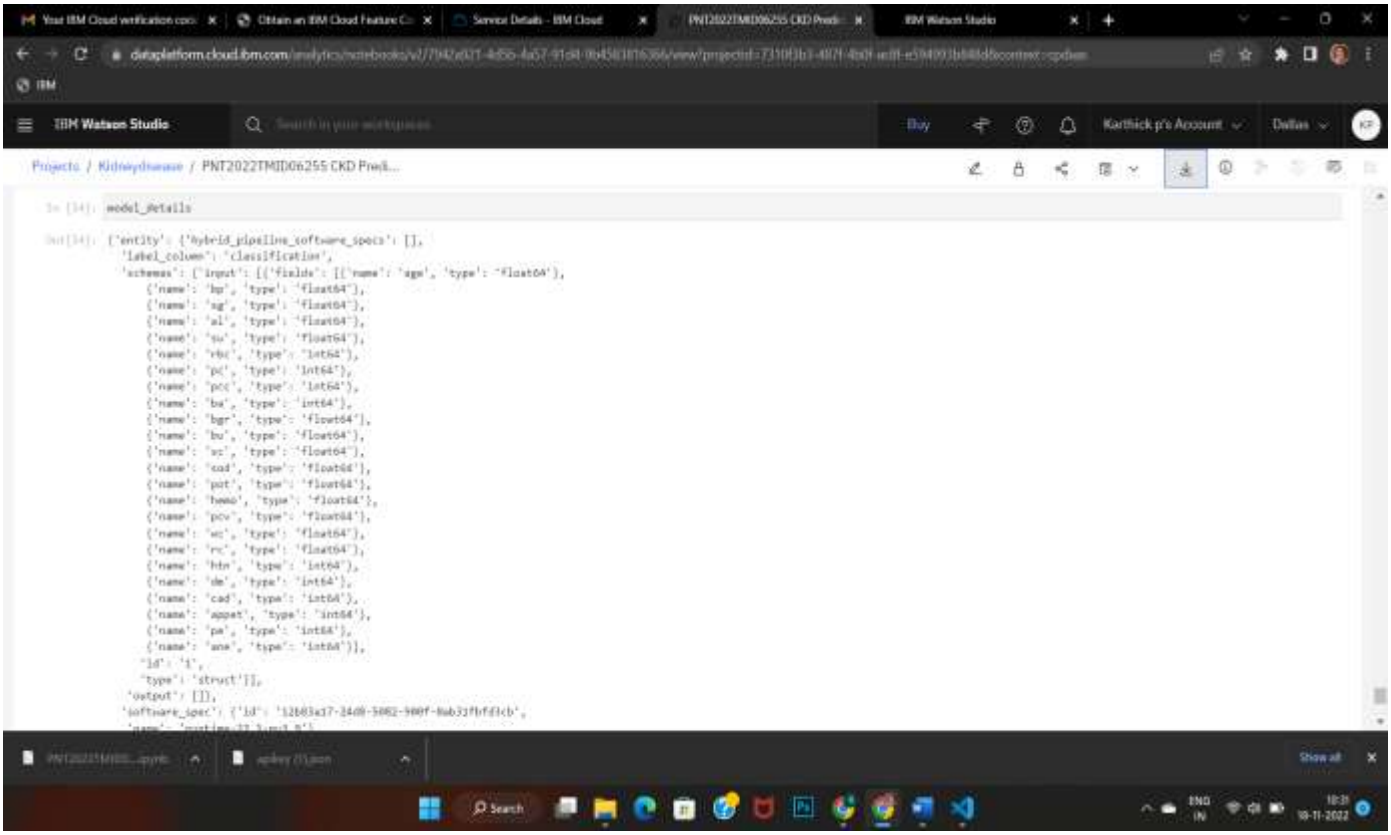


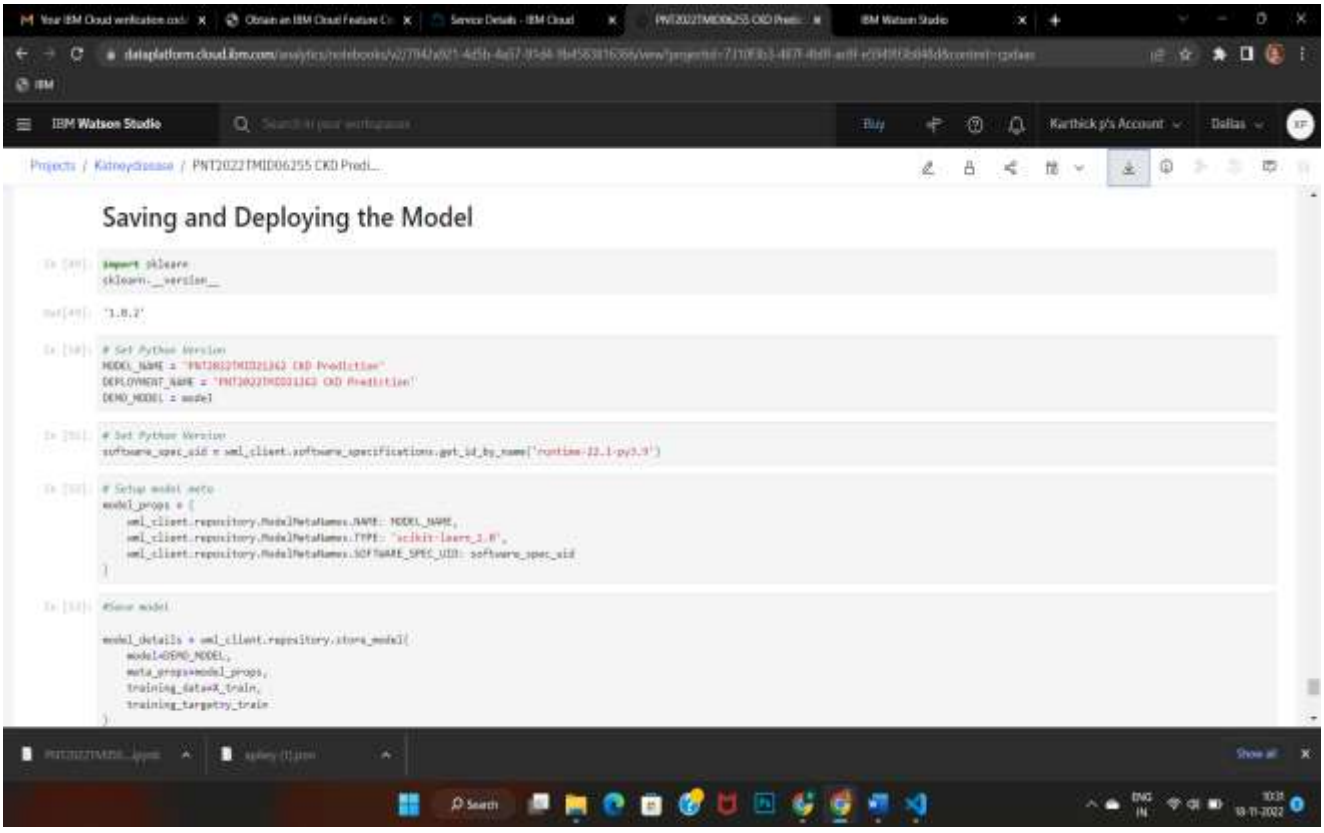


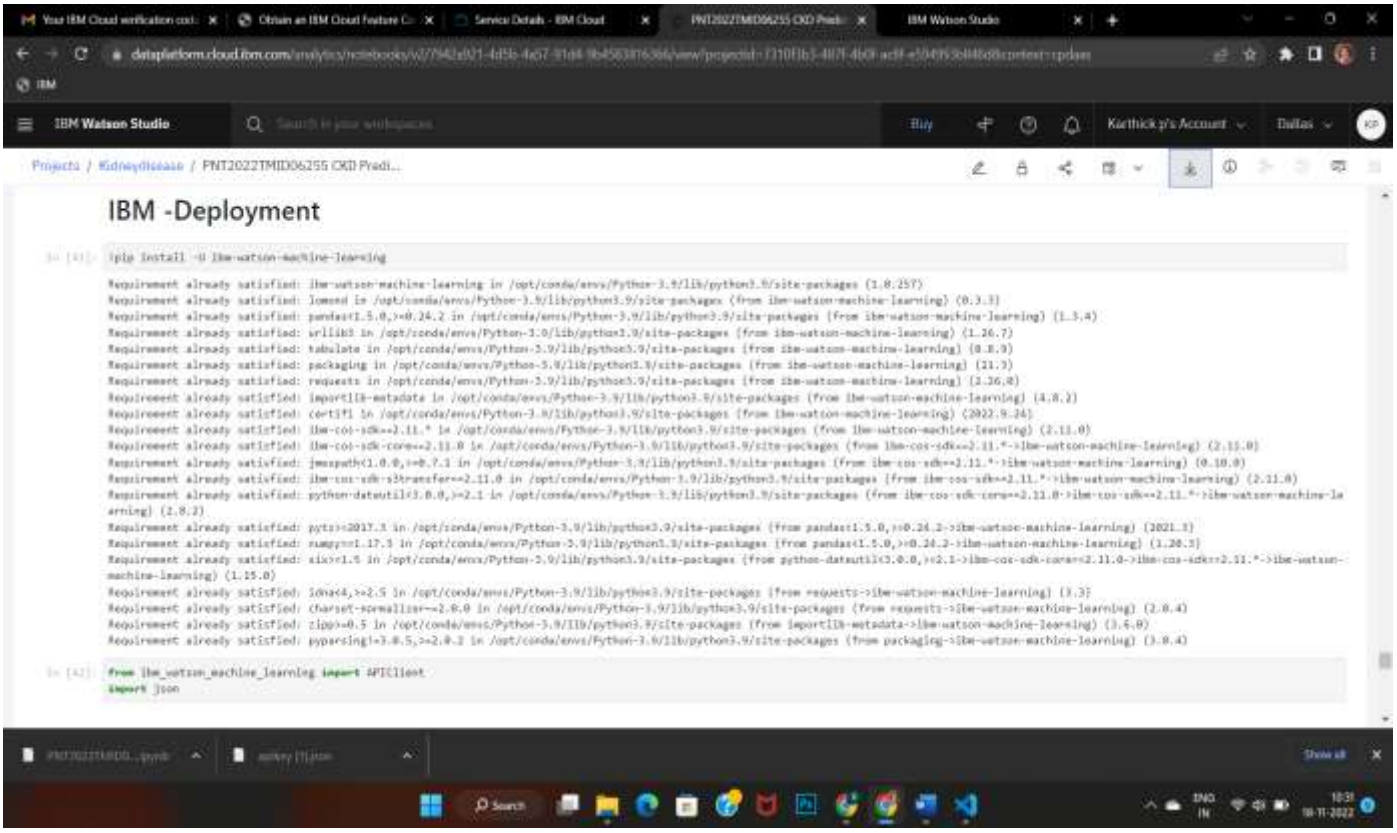
Ibm deployed ipynb file

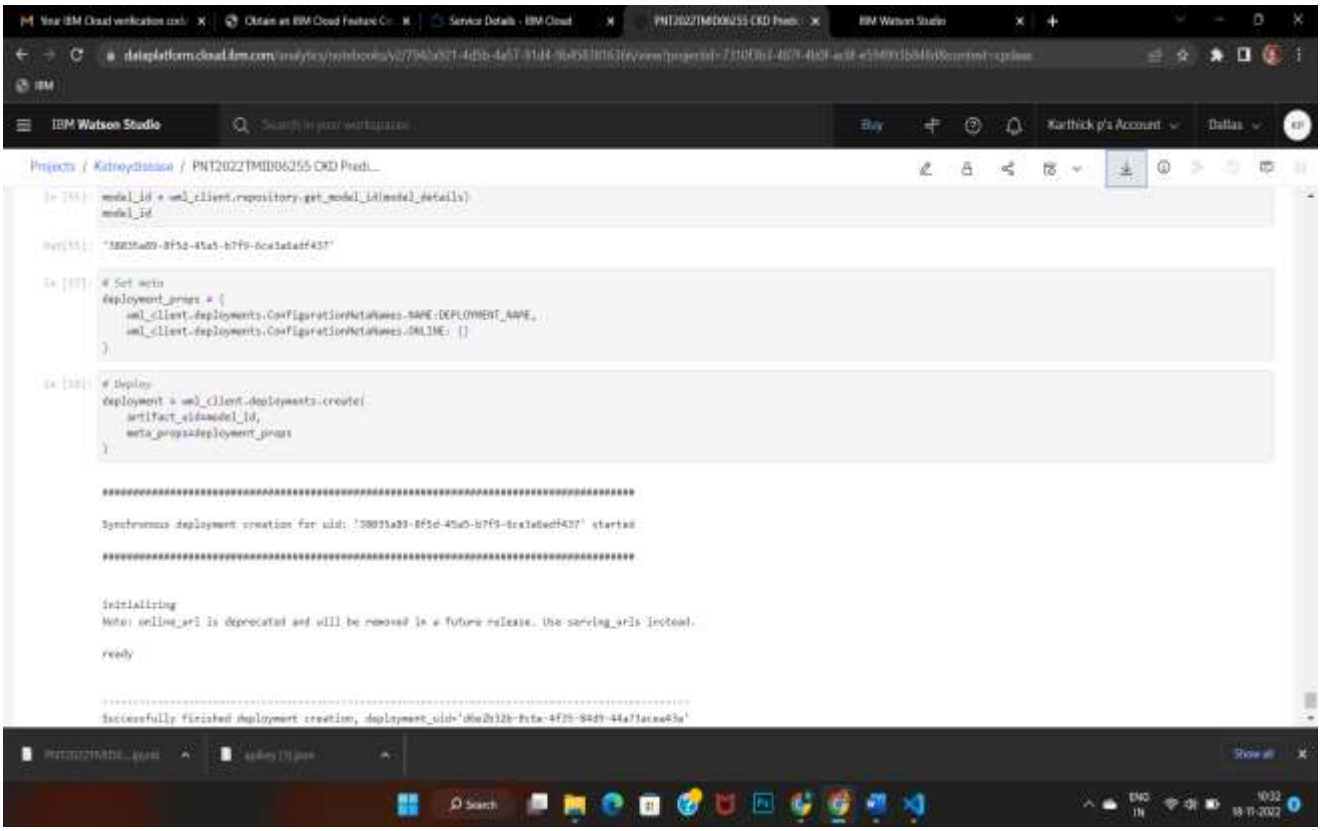




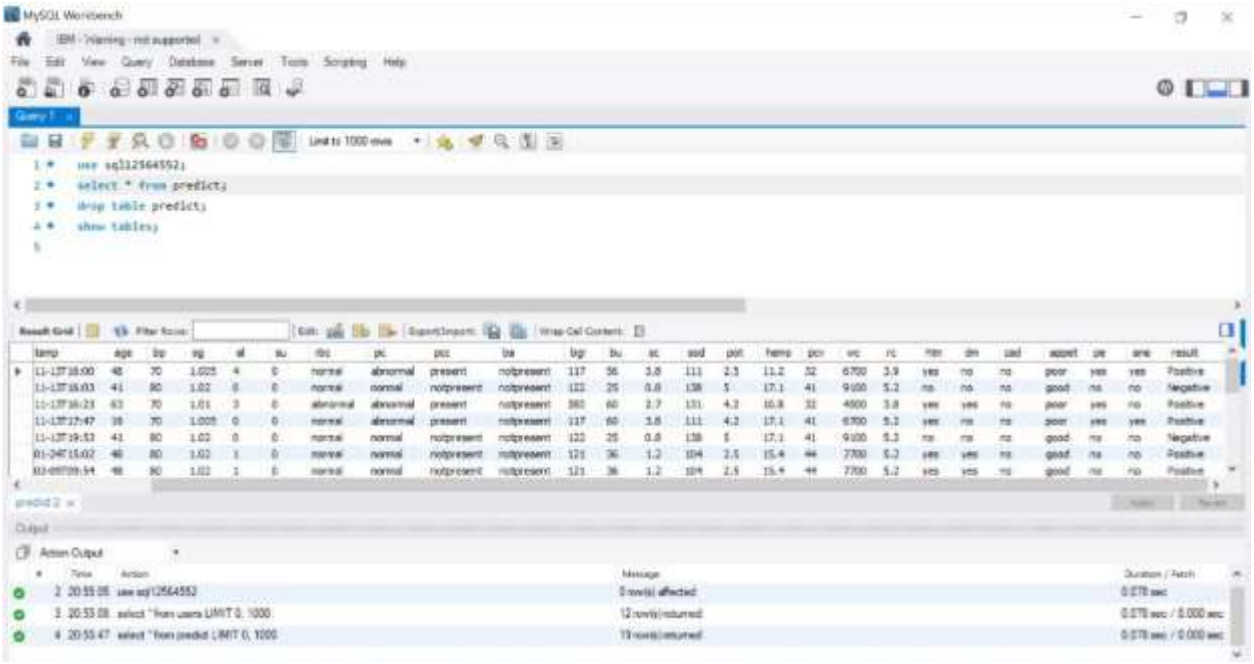

IBM Watson Studio
Projects / Kidneydisease / PNT2022TMID06255 CKD Predi...
In [34]: model_details
Out[34]: {'entity': {'hybrid_pipeline_software_specs': [],
'label_column': 'classification',
'schemas': {'input': [{'fields': [{'name': 'age', 'type': 'float64'},
{'name': 'bp', 'type': 'float64'},
{'name': 'sg', 'type': 'float64'},
{'name': 'al', 'type': 'float64'},
{'name': 'su', 'type': 'float64'},
{'name': 'rbc', 'type': 'int64'},
{'name': 'pc', 'type': 'int64'},
{'name': 'pcc', 'type': 'int64'},
{'name': 'ba', 'type': 'int64'},
{'name': 'bgr', 'type': 'float64'},
{'name': 'bu', 'type': 'float64'},
{'name': 'sc', 'type': 'float64'},
{'name': 'sod', 'type': 'float64'},
{'name': 'pot', 'type': 'float64'},
{'name': 'hemo', 'type': 'float64'},
{'name': 'pcv', 'type': 'float64'},
{'name': 'wc', 'type': 'float64'},
{'name': 'rc', 'type': 'float64'},
{'name': 'htn', 'type': 'int64'},
{'name': 'dm', 'type': 'int64'},
{'name': 'cad', 'type': 'int64'},
{'name': 'appet', 'type': 'int64'},
{'name': 'pe', 'type': 'int64'},
{'name': 'ane', 'type': 'int64'}],
'id': '1',
'type': 'struct'}],
'output': []},
'software_spec': {'id': '12b83a17-24d8-5082-900f-0ab31fbfd3cb',


IBM Watson Studio
Projects / Kidneydisease / PNT2022TMID06255 CKD Predi...
Saving and Deploying the Model
In [49]: import sklearn
sklearn.__version__
Out[49]: '1.0.2'
In [50]: # Set Python Version
MODEL_NAME = 'PNT2022TMID22362 CKD Prediction'
DEPLOYMENT_NAME = 'PNT2022TMID22362 CKD Prediction'
DEMO_MODEL = model
In [51]: # Set Python Version
software_spec_uid = wml_client.software_specifications.get_id_by_name('runtime-22.1-py3.9')
In [52]: # Setup model meta
model_props = {
wml_client.repository.ModelMetaNames.NAME: MODEL_NAME,
wml_client.repository.ModelMetaNames.TYPE: 'scikit-learn_1.0',
wml_client.repository.ModelMetaNames.SOFTWARE_SPEC_UID: software_spec_uid
}
In [53]: #Save model
model_details = wml_client.repository.store_model(
model=DEMO_MODEL,
meta_props=model_props,
training_data=X_train,
training_target=y_train
)


IBM Watson Studio
Projects / Kidneydisease / PNT2022TMID06255 CKD Predi...
IBM -Deployment
!pip install -U ibm-watson-machine-learning
Requirement already satisfied: ibm-watson-machine-learning in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (1.0.257)
Requirement already satisfied: lomond in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-watson-machine-learning) (0.3.3)
Requirement already satisfied: pandas<1.5.0,>=0.24.2 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-watson-machine-learning) (1.3.4)
Requirement already satisfied: urllib3 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-watson-machine-learning) (1.26.7)
Requirement already satisfied: tabulate in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-watson-machine-learning) (0.8.9)
Requirement already satisfied: packaging in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-watson-machine-learning) (21.3)
Requirement already satisfied: requests in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-watson-machine-learning) (2.26.0)
Requirement already satisfied: importlib-metadata in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-watson-machine-learning) (4.8.2)
Requirement already satisfied: certifi in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-watson-machine-learning) (2022.9.24)
Requirement already satisfied: ibm-cos-sdk==2.11.* in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-watson-machine-learning) (2.11.0)
Requirement already satisfied: ibm-cos-sdk-core==2.11.0 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-cos-sdk==2.11.*->ibm-watson-machine-learning) (2.11.0)
Requirement already satisfied: jmespath<1.0.0,>=0.7.1 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-cos-sdk==2.11.*->ibm-watson-machine-learning) (0.10.0)
Requirement already satisfied: ibm-cos-sdk-s3transfer==2.11.0 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-cos-sdk==2.11.*->ibm-watson-machine-learning) (2.11.0)
Requirement already satisfied: python-dateutil<3.0.0,>=2.1 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from ibm-cos-sdk-core==2.11.0->ibm-cos-sdk==2.11.*->ibm-watson-machine-learning) (2.8.2)
Requirement already satisfied: pytz>=2017.3 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from pandas<1.5.0,>=0.24.2->ibm-watson-machine-learning) (2021.3)
Requirement already satisfied: numpy>=1.17.3 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from pandas<1.5.0,>=0.24.2->ibm-watson-machine-learning) (1.20.3)
Requirement already satisfied: six>=1.5 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from python-dateutil<3.0.0,>=2.1->ibm-cos-sdk-core==2.11.0->ibm-cos-sdk==2.11.*->ibm-watson-machine-learning) (1.15.0)
Requirement already satisfied: idna<4,>=2.5 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from requests->ibm-watson-machine-learning) (3.3)
Requirement already satisfied: charset-normalizer~=2.0.0 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from requests->ibm-watson-machine-learning) (2.0.4)
Requirement already satisfied: zipp>=0.5 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from importlib-metadata->ibm-watson-machine-learning) (3.6.0)
Requirement already satisfied: pyparsing!=3.0.5,>=2.0.2 in /opt/conda/envs/Python-3.9/lib/python3.9/site-packages (from packaging->ibm-watson-machine-learning) (3.0.4)
from ibm_watson_machine_learning import APIClient
import json


IBM Watson Studio
Projects / Kidneydisease / PNT2022TMID06255 CKD Predi...
model_id = wml_client.repository.get_model_id(model_details)
model_id
'38835a89-8f5d-45a5-b7f9-6ca3a6adf437'
# Set meta
deployment_props = {
wml_client.deployments.ConfigurationMetaNames.NAME:DEPLOYMENT_NAME,
wml_client.deployments.ConfigurationMetaNames.ONLINE: {}
}
# Deploy
deployment = wml_client.deployments.create(
artifact_uid=model_id,
meta_props=deployment_props
)
##
Synchronous deployment creation for uid: '38835a89-8f5d-45a5-b7f9-6ca3a6adf437' started
##
Initializing
Note: online_url is deprecated and will be removed in a future release. Use serving_urls instead.
ready
Successfully finished deployment creation, deployment_uid='d0e2b12b-8cba-4f35-94d9-44a71acaa43a'

7.4 Database (mysql)

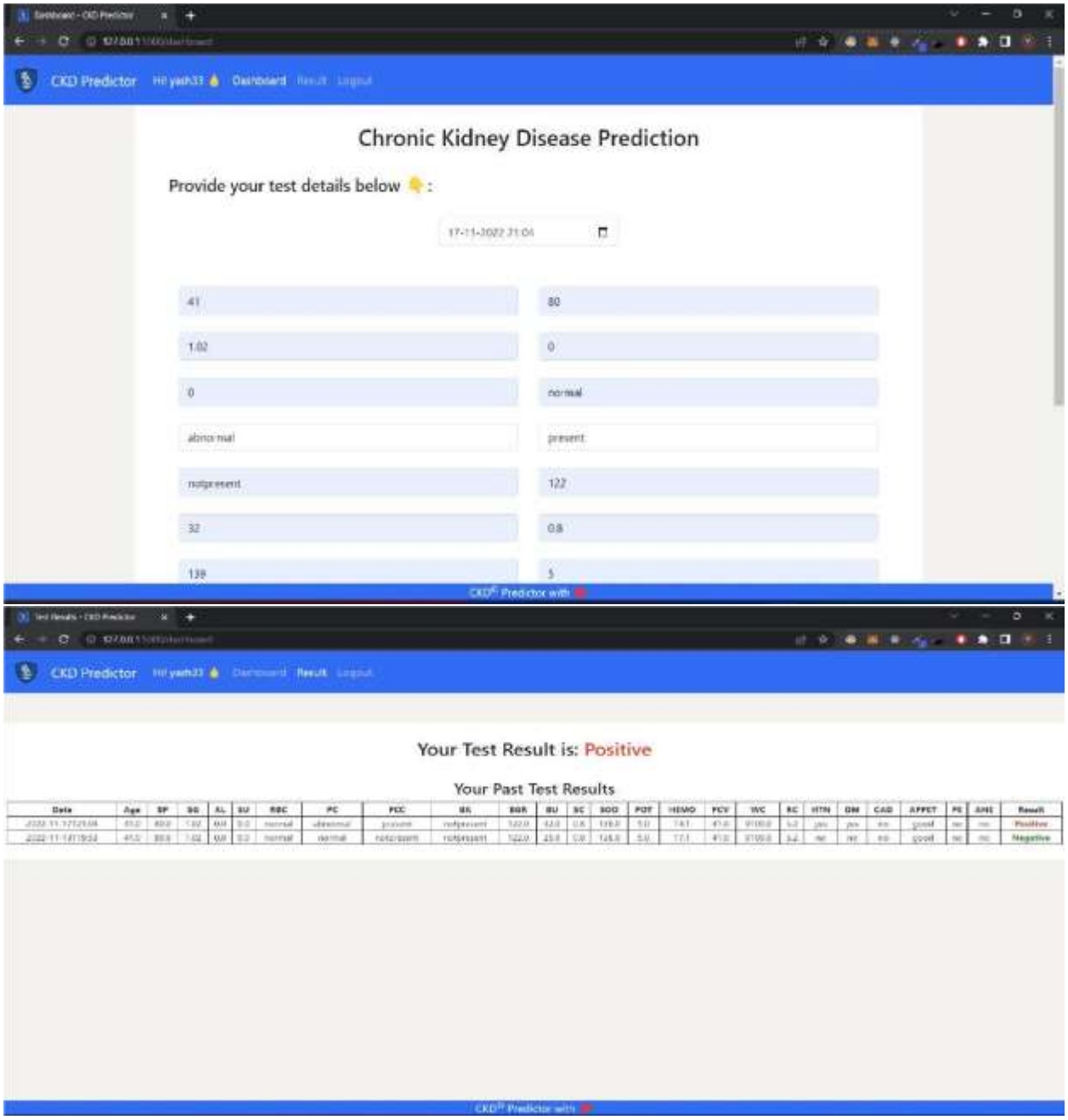


8. TESTING

8.1 Test cases

Pre-Requisite	Steps To Execute	Test Data	Expected Result	Actual Result	Status	Commnets	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 flue server .	Ran 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:


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
CKD Predictor
Hi yash33
Dashboard
Result
Logout
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 21:03:58
41.0
80.0
1.02
4.0
0.0
normal
abnormal
present
notpresent
122.0
32.0
0.8
139.0
5.0
7.4
41.0
9100.0
3.2
yes
yes
no
good
no
no
Positive
2022-11-17 21:19:24
41.0
80.0
1.02
4.0
0.0
normal
normal
notpresent
notpresent
122.0
25.0
0.8
145.0
5.0
7.7
41.0
9100.0
3.2
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

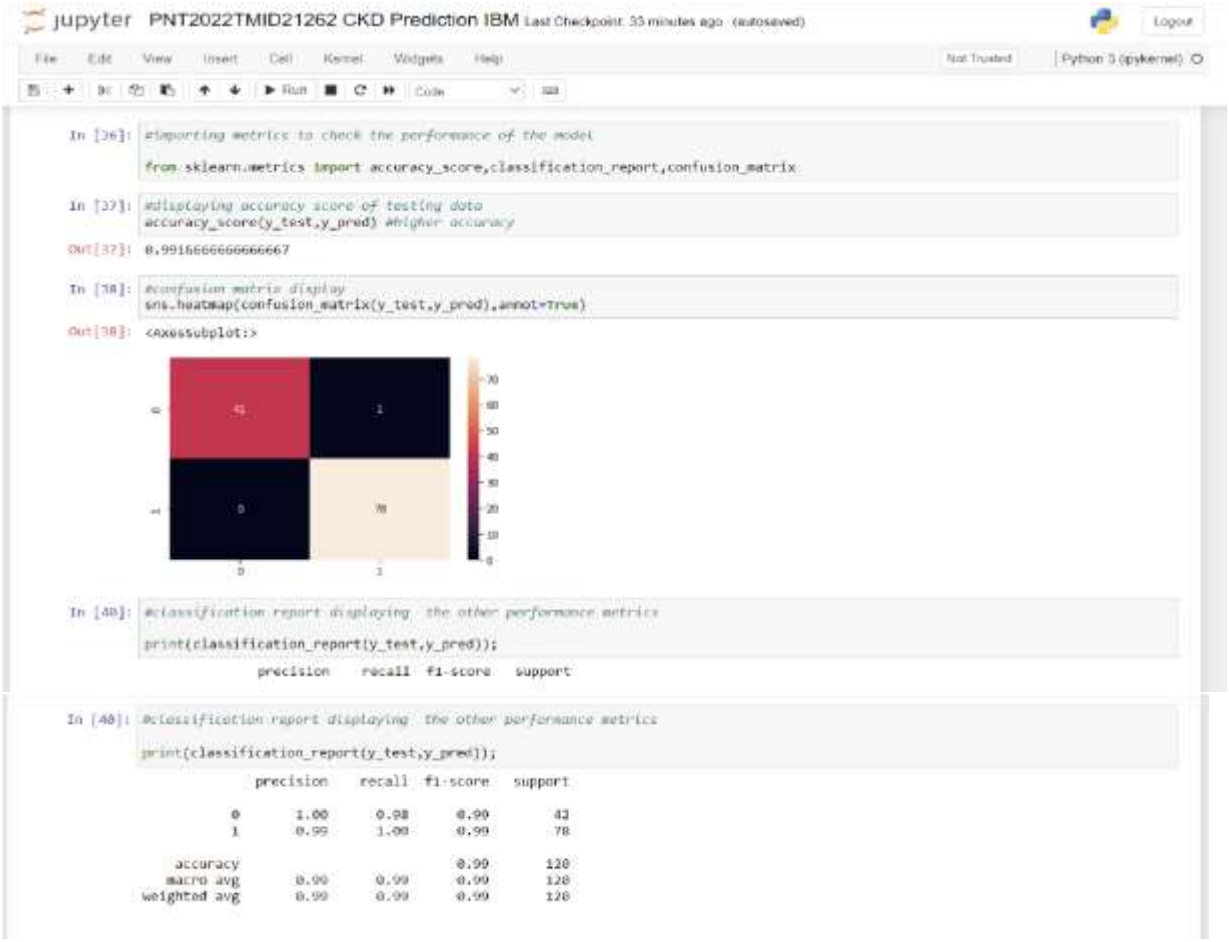
3. Test Case Analysis

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

9.1 Performance Metrics (Random Forest Classifier)



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-369-1658297986