Sri Lanka Institute of Information Technology

KidniFy - A Mobile based Chronic Kidney Disease Patient Care System Using ML and IoT

2023-032

STATUS DOCUMENT - I

Student Name: Samarawila D.R.N

Student ID: IT20235260

Group Details

Supervisor: Ms. Wishalya Vanshanee Tissera

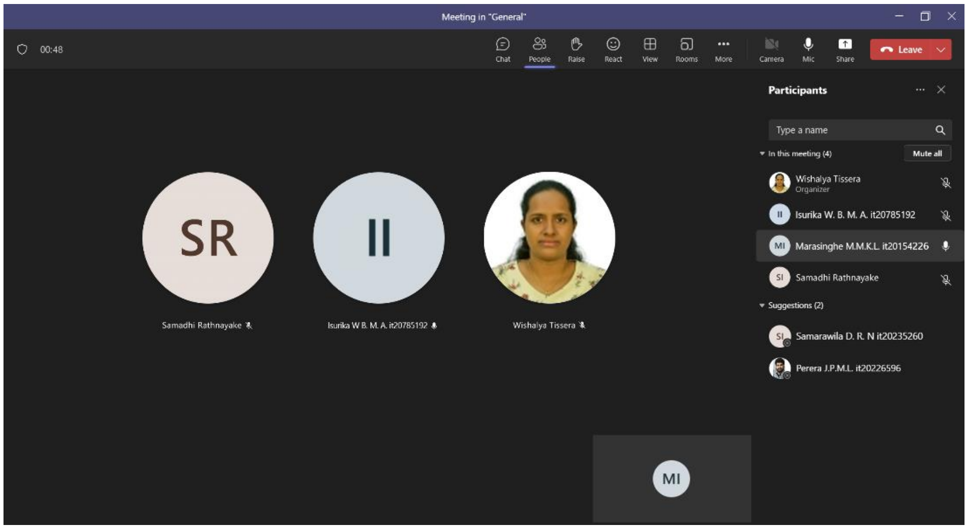
Co-Supervisor: Mr. Samadhi Chathuranga Rathnayake

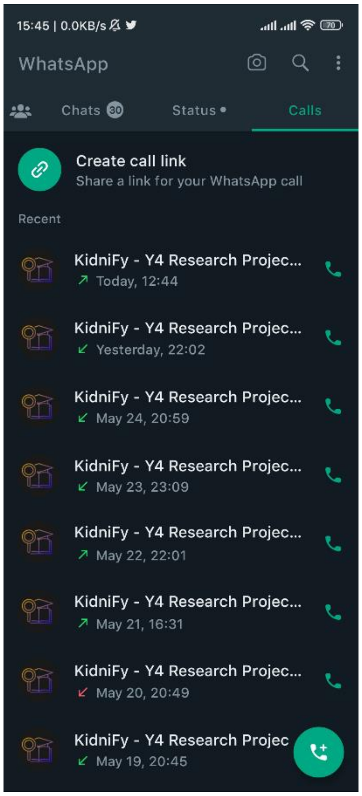
Student Name	Student ID	Contact No	Email Address
Marasinghe M.M.K.L.	IT20154226	0713037712	it20154226@my.sliit.lk
Isurika W.B.M.A.	IT20785192	0701484570	it20785192@my.sliit.lk
Perera J.P.M.L.	IT20226596	0776035479	it20226596@my.sliit.lk
Samarawila D.R.N.	IT20235260	0712421580	it20235260@my.sliit.lk

TABLE OF CONTENTS

1 Teams Meeting

1.1 Screenshots of Meetings & Calls





15:45 | 0.0KB/s

WhatsApp

Chats 30 Status • Calls

Create call link
Share a link for your WhatsApp call

Recent

KidniFy - Y4 Research Projec...
↗ Today, 12:44

KidniFy - Y4 Research Projec...
↙ Yesterday, 22:02

KidniFy - Y4 Research Projec...
↙ May 24, 20:59

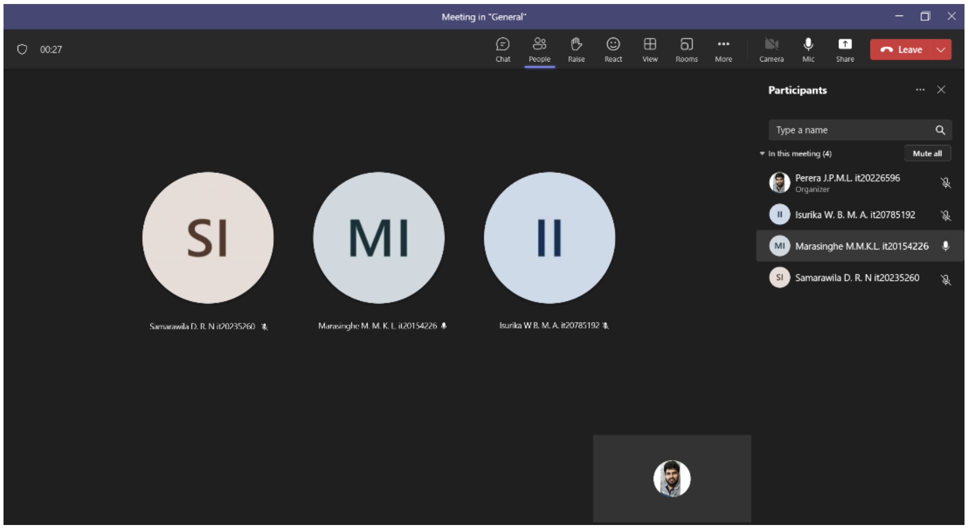
KidniFy - Y4 Research Projec...
↙ May 23, 23:09

KidniFy - Y4 Research Projec...
↗ May 22, 22:01

KidniFy - Y4 Research Projec...
↗ May 21, 16:31

KidniFy - Y4 Research Projec...
↙ May 20, 20:49

KidniFy - Y4 Research Projec
↙ May 19, 20:45


Meeting in "General"
00:27
Chat
People
Raise
React
View
Rooms
More
Camera
Mic
Share
Leave
Participants
Type a name
In this meeting (4)
Mute all
Perera J.P.M.L. it20226596
Organizer
Isurika W. B. M. A. it20785192
Marasinghe M.M.K.L. it20154226
Samarawila D. R. N it20235260
SI
MI
II
Samarawila D. R. N it20235260
Marasinghe M. M. K. L. it20154226
Isurika W B. M. A. it20785192

1.2 Meeting with the domain expert

Meeting with Consultant Nephrologist Dr. Pramil Rajakrishnan

At

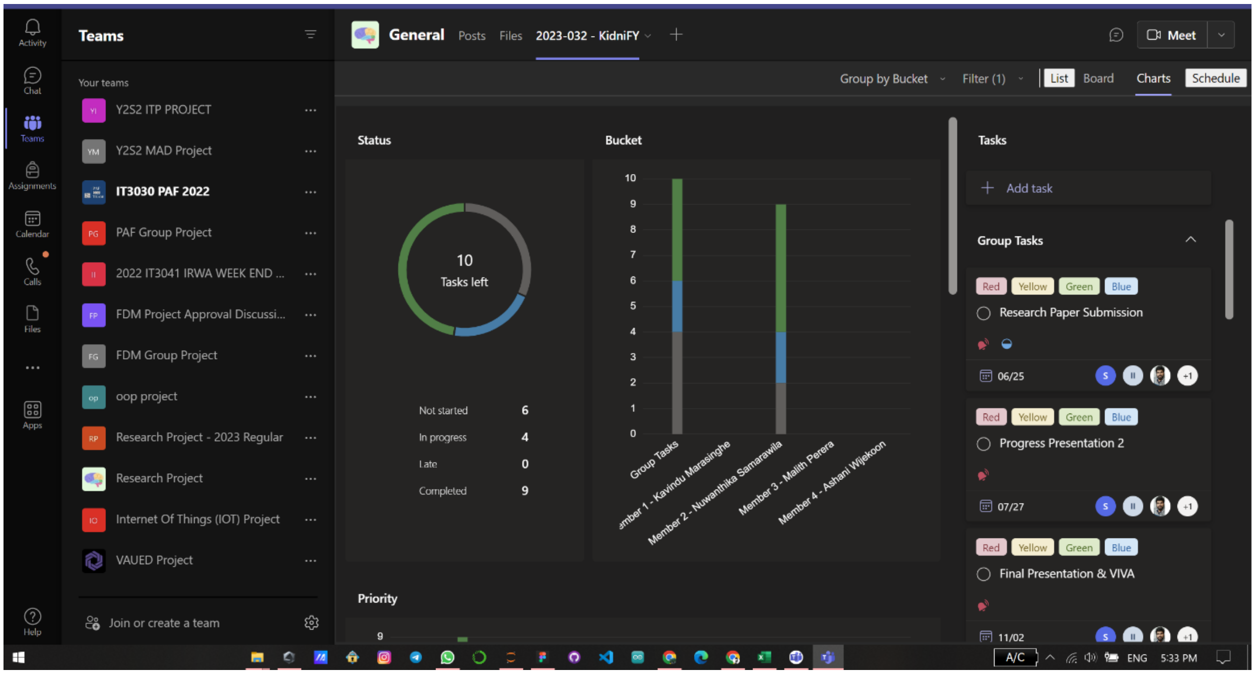
Kurunegala Teaching Hospital - Kidney dialysis Unit | 2023.04.29

Meet the Ward Master of Kidney Dialysis Unit

At Anuradhapura Teaching Hospital | 2023.05.01

2 Screenshots of the Tasks by Planner

2.1 Chart Overview



Not started	6
In progress	4
Late	0
Completed	9

Red Yellow Green Blue
○ Research Paper Submission
06/25 S +1

Red Yellow Green Blue
○ Progress Presentation 2
07/27 S +1

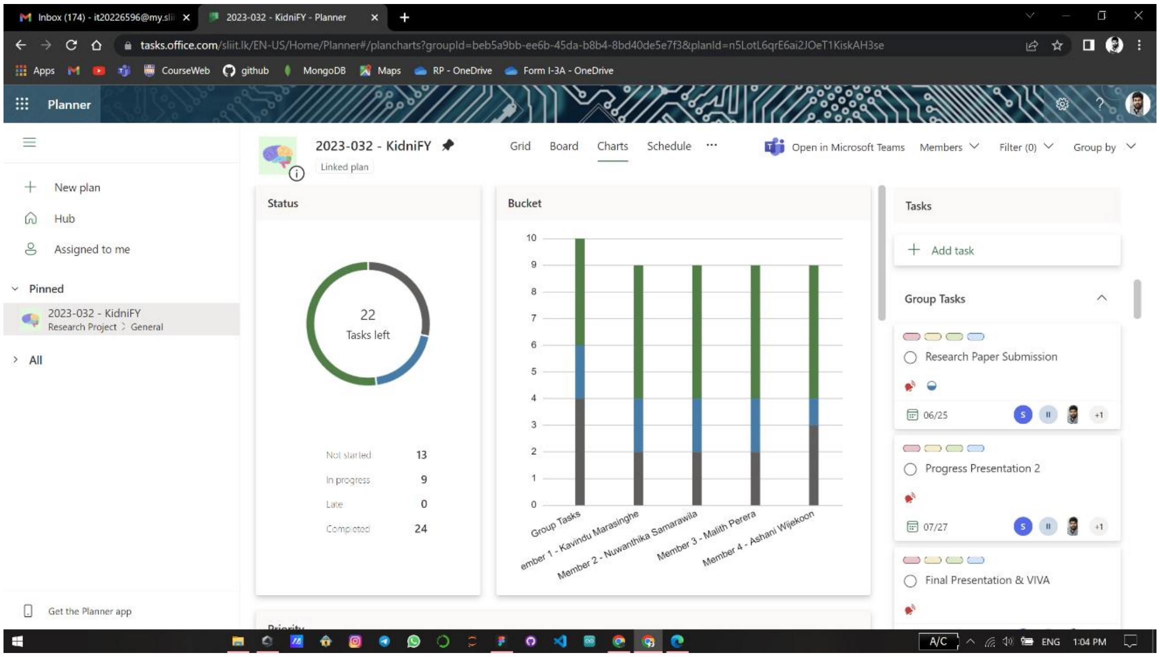
Red Yellow Green Blue
○ Final Presentation & VIVA
11/02 S +1

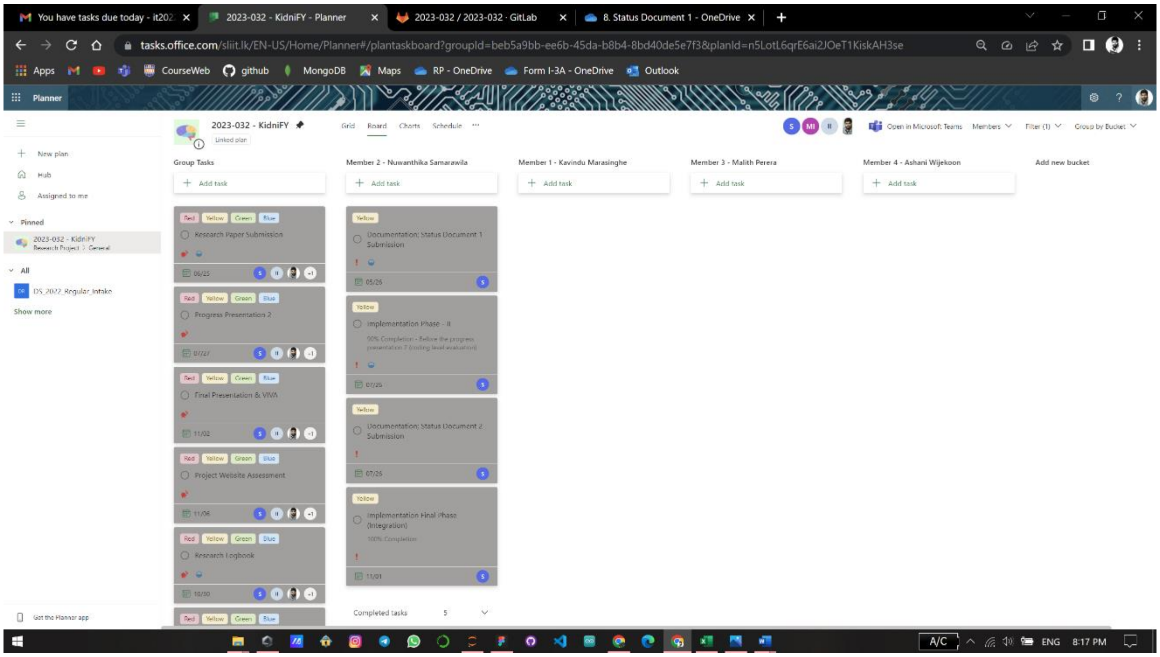
🟩 - To Do

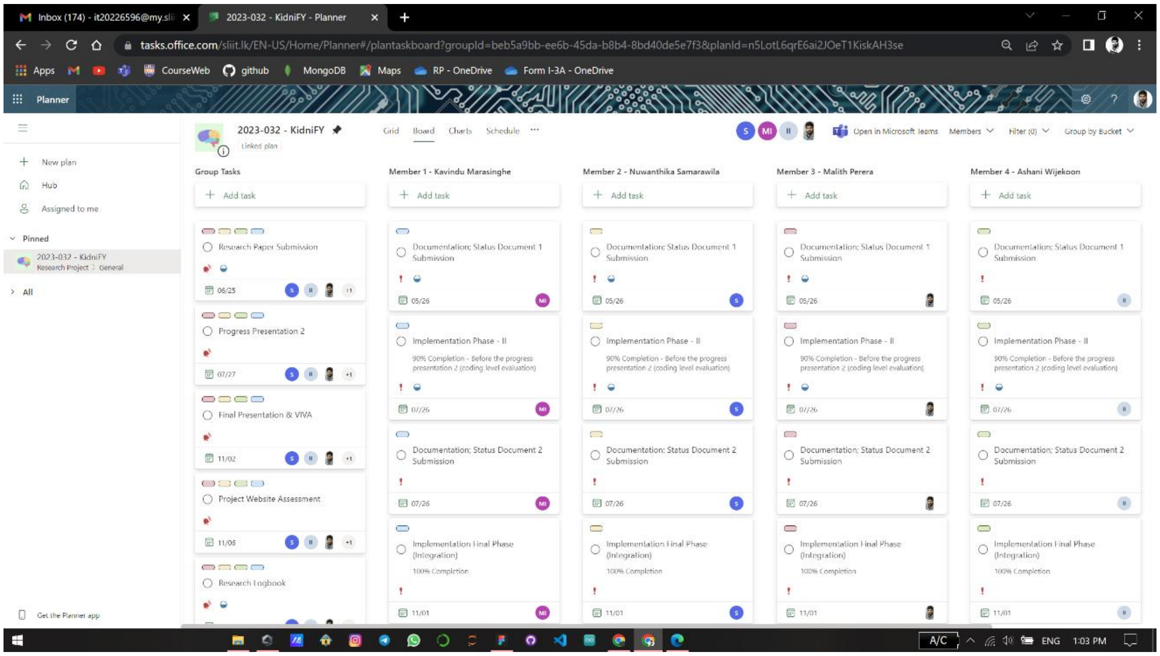
🟦 - In Progress

⬛ - Completed

2.2 Bucket List



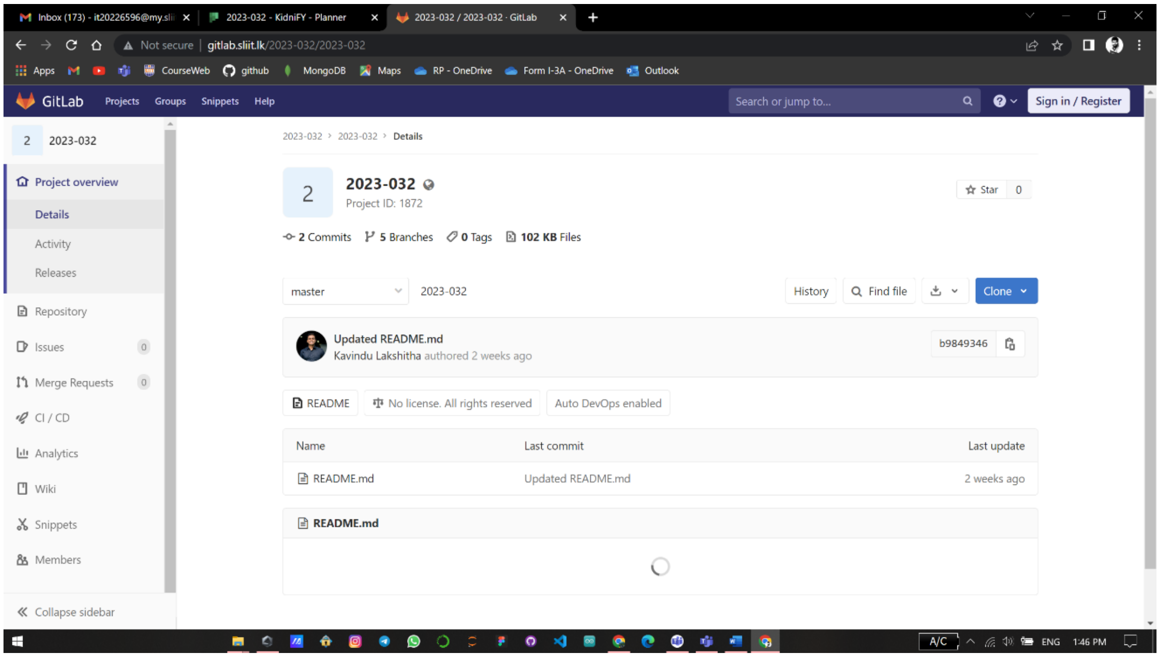

You have tasks due today - it202
2023-032 - KidniFY - Planner
2023-032 / 2023-032 · GitLab
8. Status Document 1 - OneDrive
tasks.office.com/sliit.lk/EN-US/Home/Planner#/plantaskboard?groupId=beb5a9bb-ee6b-45da-b8b4-8bd40de5e7f3&planId=n5LotL6qrE6ai2JOeT1KiskAH3se
Apps
CourseWeb
github
MongoDB
Maps
RP - OneDrive
Form I-3A - OneDrive
Outlook
Planner
New plan
Hub
Assigned to me
Pinned
2023-032 - KidniFY
Research Project > General
All
DS_2022_Regular_Intake
Show more
Get the Planner app
2023-032 - KidniFY
Linked plan
Grid
Board
Charts
Schedule
Open in Microsoft Teams
Members
Filter (1)
Group by Bucket
Group Tasks
Member 2 - Nuwanthika Samarawila
Member 1 - Kavindu Marasinghe
Member 3 - Malith Perera
Member 4 - Ashani Wijekoon
Add new bucket
Add task
Red
Yellow
Green
Blue
Research Paper Submission
06/25
Progress Presentation 2
07/27
Final Presentation & VIVA
11/02
Project Website Assessment
11/06
Research Logbook
10/30
Documentation: Status Document 1 Submission
05/26
Implementation Phase - II
90% Completion - Before the progress presentation 2 (coding level evaluation)
07/26
Documentation: Status Document 2 Submission
07/26
Implementation Final Phase (Integration)
100% Completion
11/01
Completed tasks 5
A/C
ENG
8:17 PM



Group Tasks	Member 1 - Kavindu Marasinghe	Member 2 - Nuwanthika Samarawila	Member 3 - Malith Perera	Member 4 - Ashani Wijekoon
Research Paper Submission 06/25	Documentation; Status Document 1 Submission 05/26	Documentation; Status Document 1 Submission 05/26	Documentation; Status Document 1 Submission 05/26	Documentation; Status Document 1 Submission 05/26
Progress Presentation 2 07/27	Implementation Phase - II 90% Completion - Before the progress presentation 2 (coding level evaluation) 07/26	Implementation Phase - II 90% Completion - Before the progress presentation 2 (coding level evaluation) 07/26	Implementation Phase - II 90% Completion - Before the progress presentation 2 (coding level evaluation) 07/26	Implementation Phase - II 90% Completion - Before the progress presentation 2 (coding level evaluation) 07/26
Final Presentation & VIVA 11/02	Documentation; Status Document 2 Submission 07/26	Documentation; Status Document 2 Submission 07/26	Documentation; Status Document 2 Submission 07/26	Documentation; Status Document 2 Submission 07/26
Project Website Assessment 11/05	Implementation Final Phase (Integration) 100% Completion 11/01	Implementation Final Phase (Integration) 100% Completion 11/01	Implementation Final Phase (Integration) 100% Completion 11/01	Implementation Final Phase (Integration) 100% Completion 11/01
Research Logbook				

2.3 Screenshots of GitLab

Repository in GitLab



2023-032 › 2023-032 › Details

2023-032

Project ID: 1872

2 Commits 5 Branches 0 Tags 102 KB Files

master 2023-032

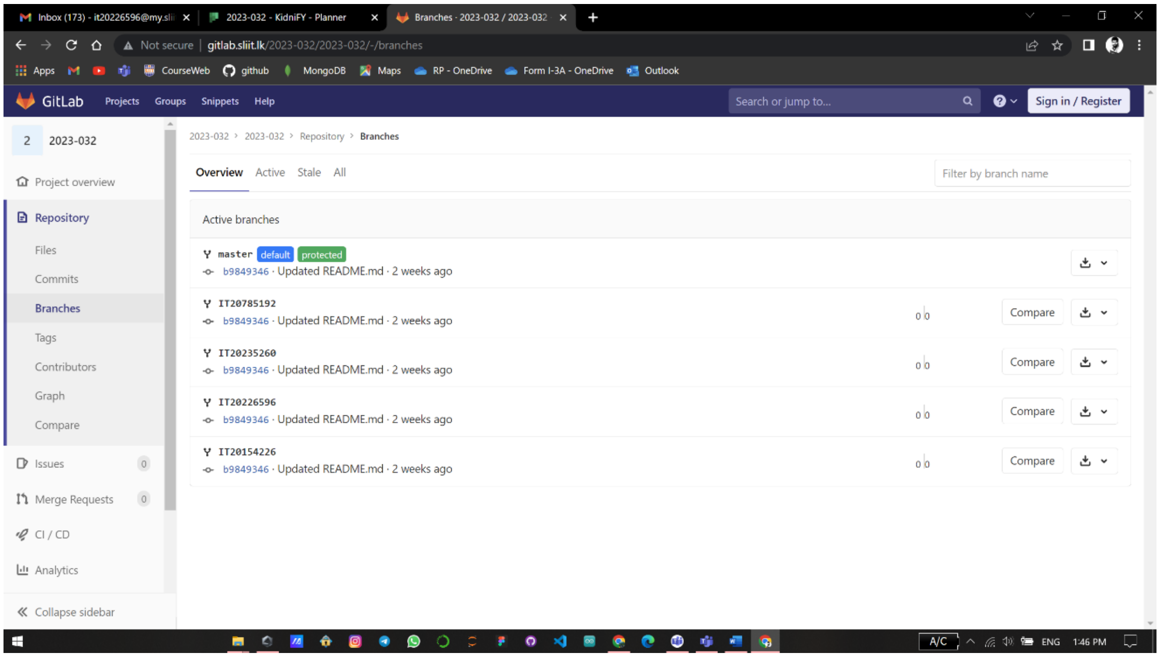
Updated README.md
Kavindu Lakshitha authored 2 weeks ago
b9849346

README No license. All rights reserved Auto DevOps enabled

Name	Last commit	Last update
README.md	Updated README.md	2 weeks ago

README.md

Create GitLab Branches & Starting Implementations



Active branches

master default protected
b9849346 · Updated README.md · 2 weeks ago

IT20785192
b9849346 · Updated README.md · 2 weeks ago 0 0

IT20235260
b9849346 · Updated README.md · 2 weeks ago 0 0

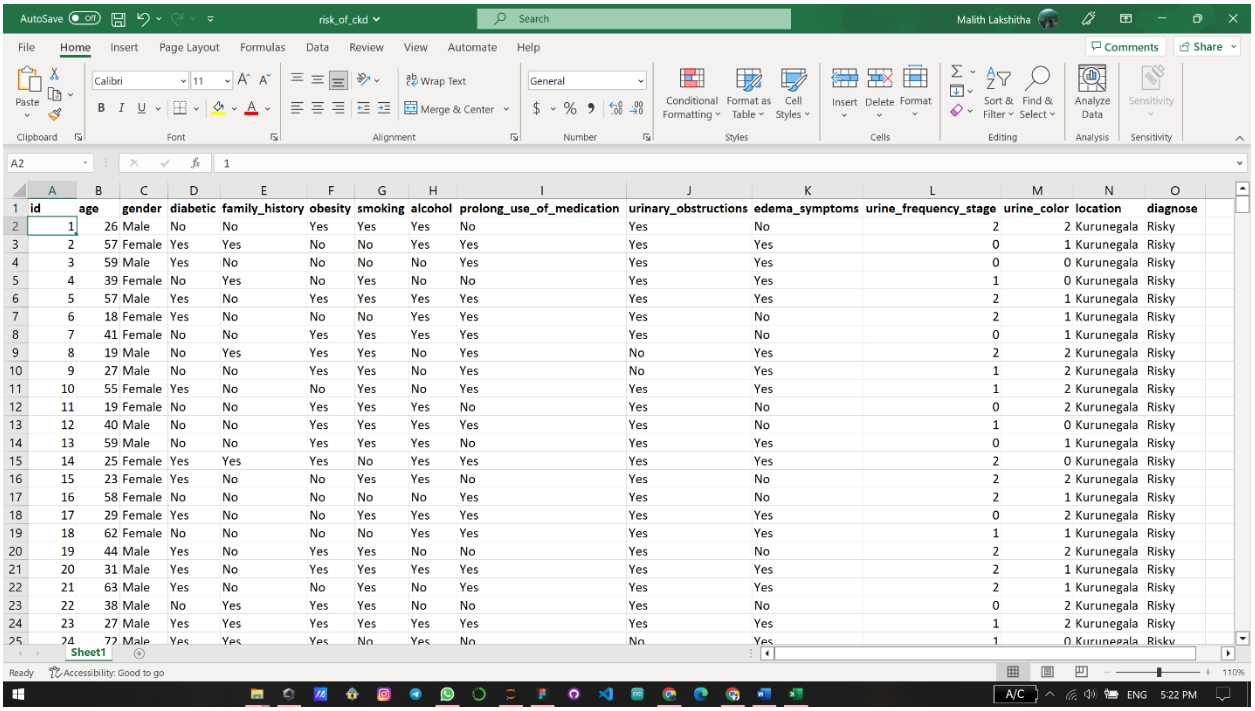
IT20226596
b9849346 · Updated README.md · 2 weeks ago 0 0

IT20154226
b9849346 · Updated README.md · 2 weeks ago 0 0

3 Project Implementation

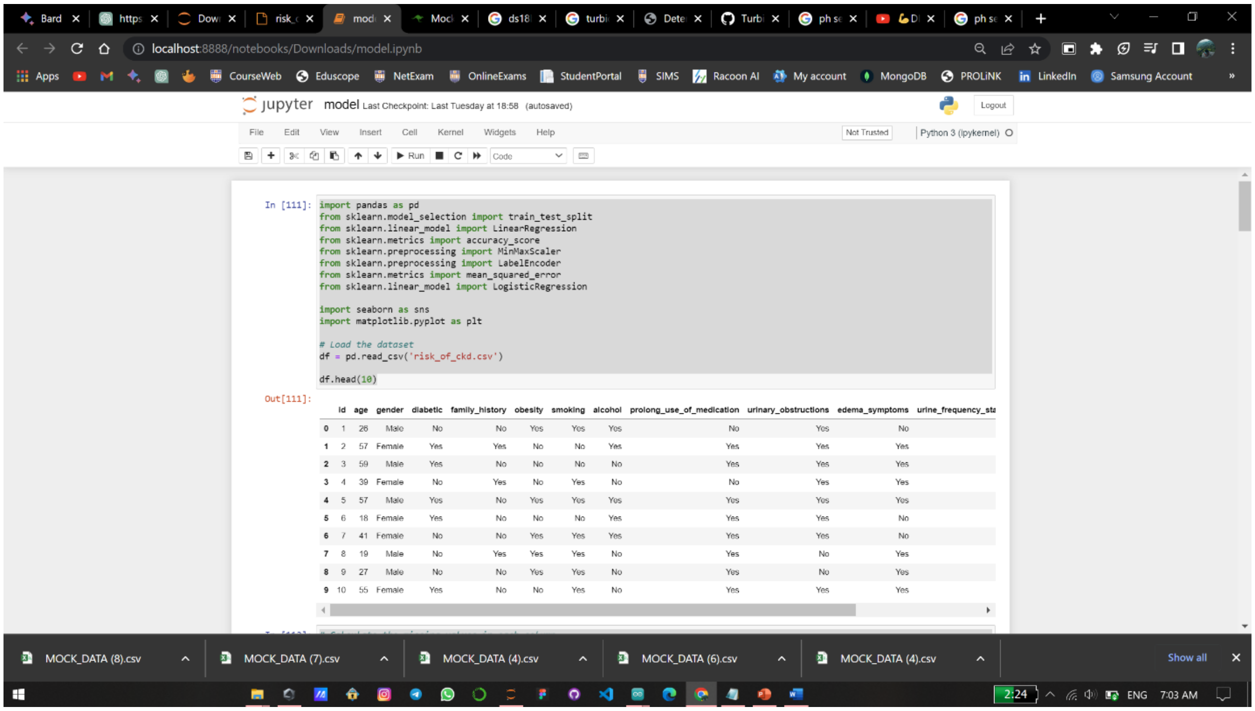
Predictive analysis for risks of having a kidney disease.

Gathered Date Set



id	age	gender	diabetic	family_history	obesity	smoking	alcohol	prolong_use_of_medication	urinary_obstructions	edema_symptoms	urine_frequency_stage	urine_color	location	diagnose
1	26	Male	No	No	Yes	Yes	Yes	No	Yes	No	2	2	Kurunegala	Risky
2	57	Female	Yes	Yes	No	No	Yes	Yes	Yes	Yes	0	1	Kurunegala	Risky
3	59	Male	Yes	No	No	No	No	Yes	Yes	Yes	0	0	Kurunegala	Risky
4	39	Female	No	Yes	No	Yes	No	No	Yes	Yes	1	0	Kurunegala	Risky
5	57	Male	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	2	1	Kurunegala	Risky
6	18	Female	Yes	No	No	No	Yes	Yes	Yes	No	2	1	Kurunegala	Risky
7	41	Female	No	No	Yes	Yes	Yes	Yes	Yes	No	0	1	Kurunegala	Risky
8	19	Male	No	Yes	Yes	Yes	No	Yes	No	Yes	2	2	Kurunegala	Risky
9	27	Male	No	No	Yes	Yes	No	Yes	No	Yes	1	2	Kurunegala	Risky
10	55	Female	Yes	No	No	Yes	No	Yes	Yes	Yes	1	2	Kurunegala	Risky
11	19	Female	No	No	Yes	Yes	Yes	No	Yes	No	0	2	Kurunegala	Risky
12	40	Male	No	No	Yes	Yes	Yes	Yes	Yes	No	1	0	Kurunegala	Risky
13	59	Male	No	No	Yes	Yes	Yes	No	Yes	Yes	0	1	Kurunegala	Risky
14	25	Female	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	2	0	Kurunegala	Risky
15	23	Female	Yes	No	No	Yes	Yes	No	Yes	No	2	2	Kurunegala	Risky
16	58	Female	No	No	No	No	No	Yes	Yes	No	2	1	Kurunegala	Risky
17	29	Female	Yes	No	No	Yes	Yes	Yes	Yes	Yes	0	2	Kurunegala	Risky
18	62	Female	No	No	No	No	Yes	Yes	Yes	Yes	1	1	Kurunegala	Risky
19	44	Male	Yes	No	Yes	Yes	No	No	Yes	No	2	2	Kurunegala	Risky
20	31	Male	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	2	1	Kurunegala	Risky
21	63	Male	Yes	No	No	Yes	No	Yes	Yes	Yes	2	1	Kurunegala	Risky
22	38	Male	No	Yes	Yes	Yes	No	No	Yes	No	0	2	Kurunegala	Risky
23	27	Male	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	1	2	Kurunegala	Risky
24	72	Male	Yes	Yes	Yes	No	Yes	No	No	Yes	1	0	Kurunegala	Risky

Import Libraries & load the dataset.



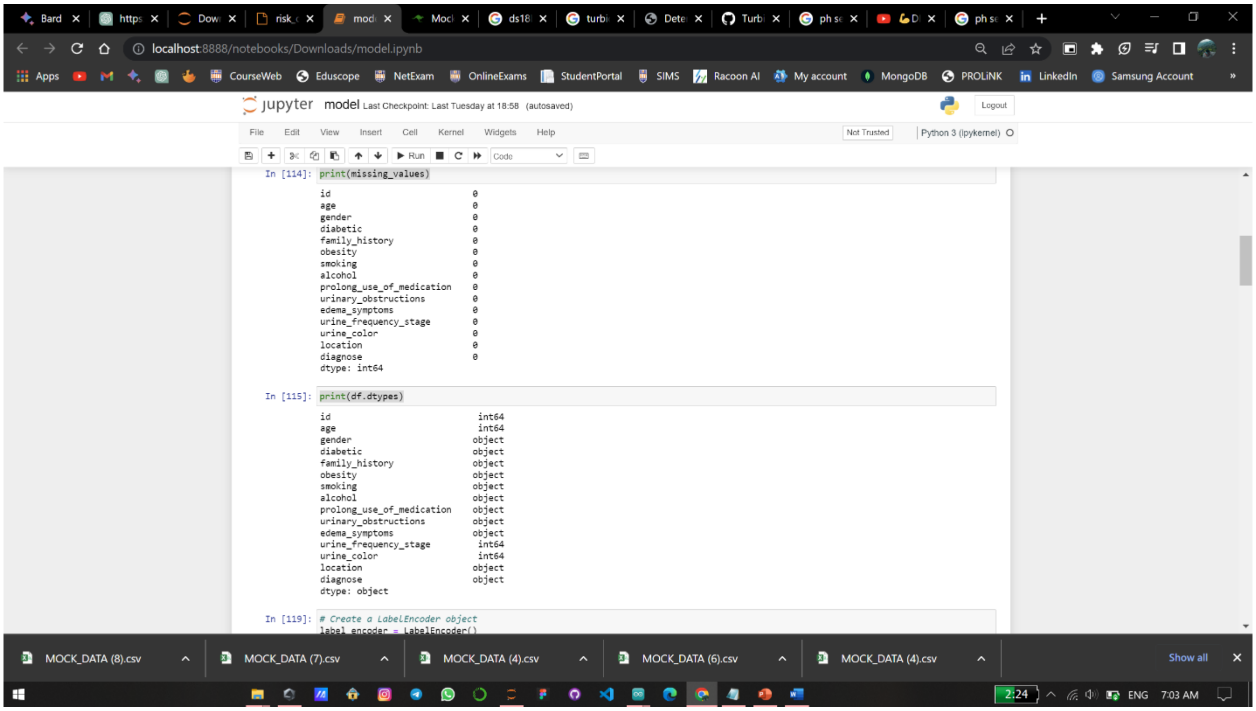
localhost:8888/notebooks/Downloads/model.ipynb

jupyter model Last Checkpoint: Last Tuesday at 18:58 (autosaved)

In [111]:
```python
import pandas as pd
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import accuracy_score
from sklearn.preprocessing import MinMaxScaler
from sklearn.preprocessing import LabelEncoder
from sklearn.metrics import mean_squared_error
from sklearn.linear_model import LogisticRegression

import seaborn as sns
import matplotlib.pyplot as plt

# Load the dataset
df = pd.read_csv('risk_of_ckd.csv')

df.head(10)
```

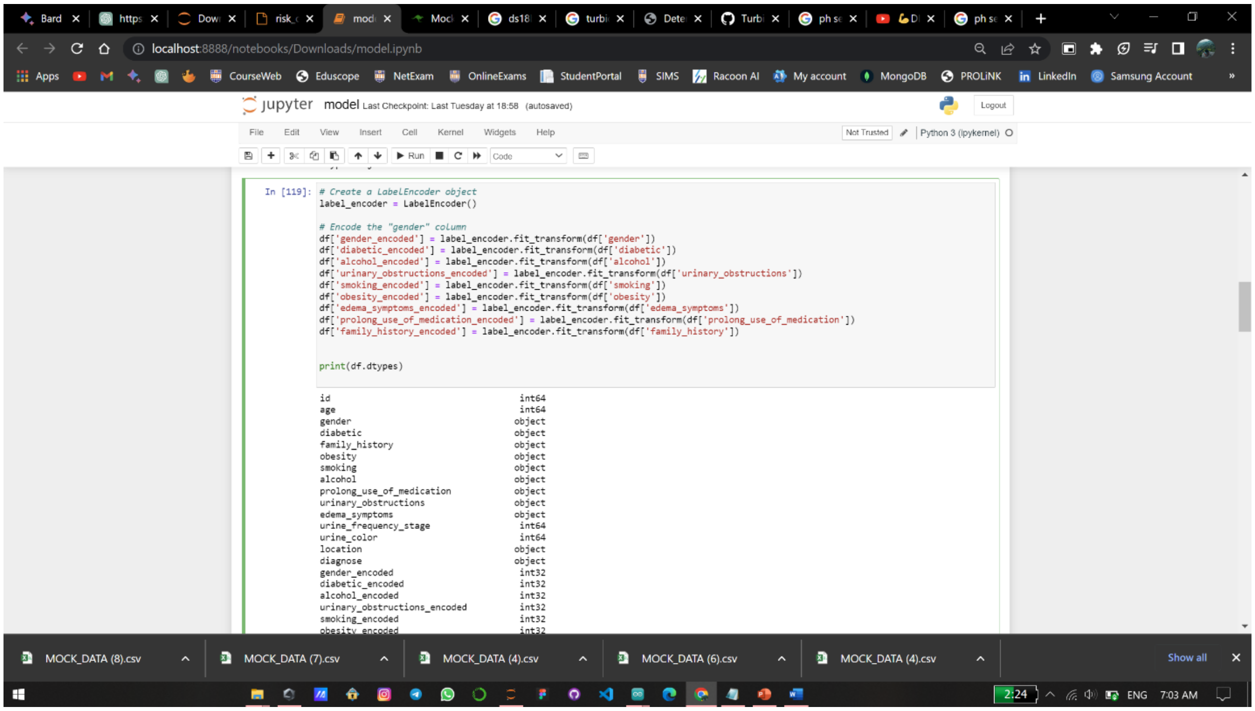
Out[111]:
```
   id  age  gender  diabetic  family_history  obesity  smoking  alcohol  prolong_use_of_medication  urinary_obstructions  edema_symptoms  urine_frequency_sta
0   1   26    Male        No              No      Yes      Yes      Yes                         No                   Yes              No
1   2   57  Female       Yes             Yes       No       No      Yes                        Yes                   Yes             Yes
2   3   59    Male       Yes              No       No       No       No                        Yes                   Yes             Yes
3   4   39  Female        No             Yes       No      Yes       No                         No                   Yes             Yes
4   5   57    Male       Yes              No      Yes      Yes      Yes                        Yes                   Yes             Yes
5   6   18  Female       Yes              No       No       No      Yes                        Yes                   Yes              No
6   7   41  Female        No              No      Yes      Yes      Yes                        Yes                   Yes              No
7   8   19    Male        No             Yes      Yes      Yes       No                        Yes                    No             Yes
8   9   27    Male        No              No      Yes      Yes       No                        Yes                    No             Yes
9  10   55  Female       Yes              No       No      Yes       No                        Yes                   Yes             Yes
```

Checking missing values



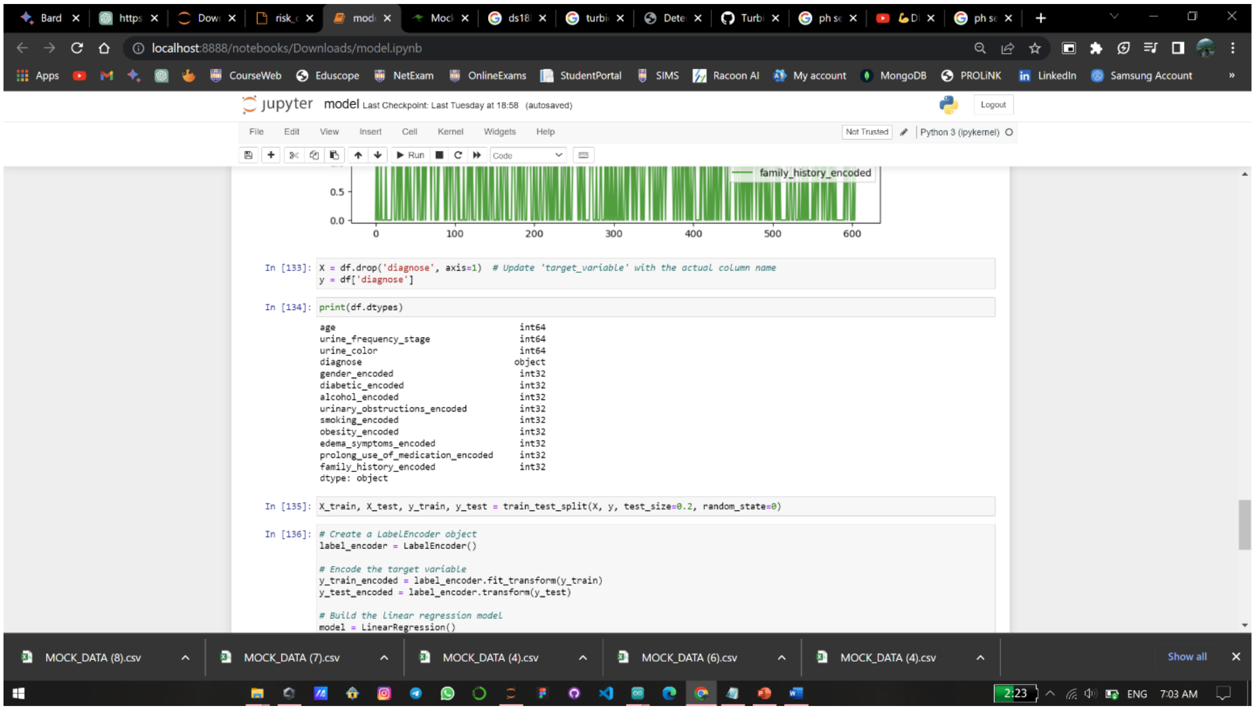
```
In [114]: print(missing_values)

          id                          0
          age                         0
          gender                      0
          diabetic                    0
          family_history              0
          obesity                     0
          smoking                     0
          alcohol                     0
          prolong_use_of_medication   0
          urinary_obstructions        0
          edema_symptoms              0
          urine_frequency_stage       0
          urine_color                 0
          location                    0
          diagnose                    0
          dtype: int64

In [115]: print(df.dtypes)

          id                           int64
          age                          int64
          gender                      object
          diabetic                    object
          family_history              object
          obesity                     object
          smoking                     object
          alcohol                     object
          prolong_use_of_medication   object
          urinary_obstructions        object
          edema_symptoms              object
          urine_frequency_stage        int64
          urine_color                  int64
          location                    object
          diagnose                    object
          dtype: object

In [119]: # Create a LabelEncoder object
          label_encoder = LabelEncoder()
```

Using Label Encoding



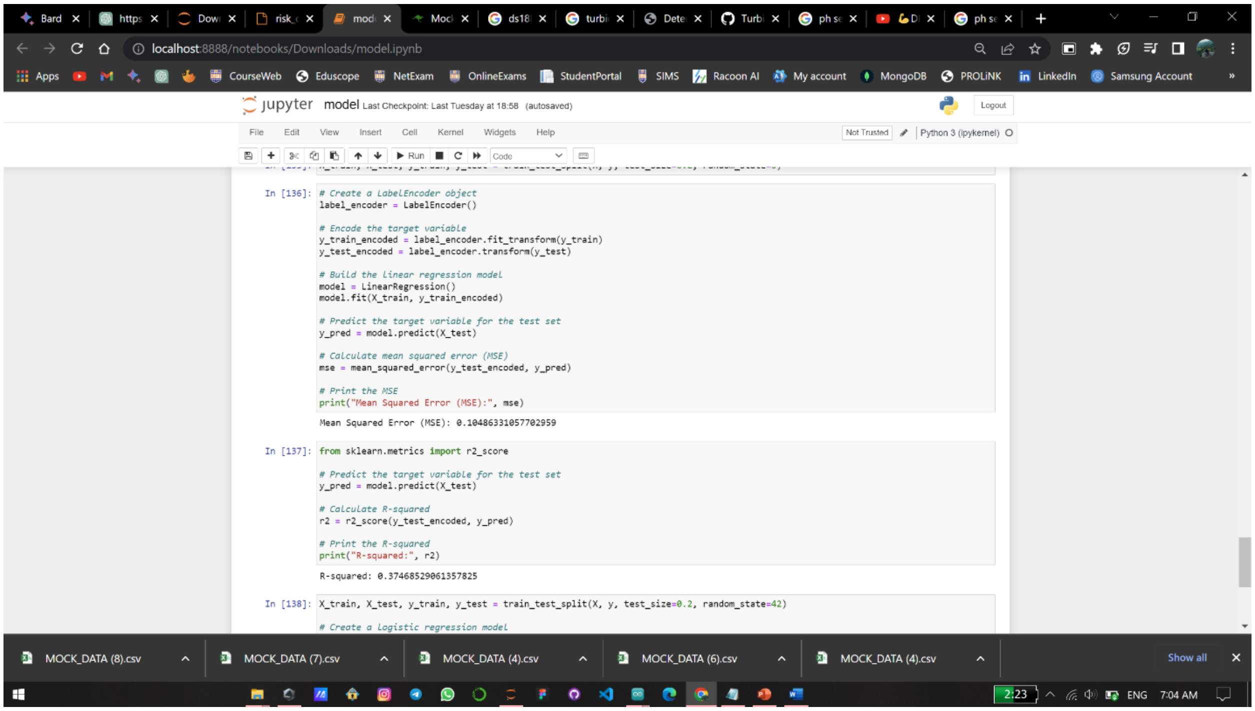
```
In [119]: # Create a LabelEncoder object
          label_encoder = LabelEncoder()

          # Encode the "gender" column
          df['gender_encoded'] = label_encoder.fit_transform(df['gender'])
          df['diabetic_encoded'] = label_encoder.fit_transform(df['diabetic'])
          df['alcohol_encoded'] = label_encoder.fit_transform(df['alcohol'])
          df['urinary_obstructions_encoded'] = label_encoder.fit_transform(df['urinary_obstructions'])
          df['smoking_encoded'] = label_encoder.fit_transform(df['smoking'])
          df['obesity_encoded'] = label_encoder.fit_transform(df['obesity'])
          df['edema_symptoms_encoded'] = label_encoder.fit_transform(df['edema_symptoms'])
          df['prolong_use_of_medication_encoded'] = label_encoder.fit_transform(df['prolong_use_of_medication'])
          df['family_history_encoded'] = label_encoder.fit_transform(df['family_history'])


          print(df.dtypes)

          id                                int64
          age                               int64
          gender                           object
          diabetic                         object
          family_history                   object
          obesity                          object
          smoking                          object
          alcohol                          object
          prolong_use_of_medication        object
          urinary_obstructions             object
          edema_symptoms                   object
          urine_frequency_stage             int64
          urine_color                       int64
          location                         object
          diagnose                         object
          gender_encoded                    int32
          diabetic_encoded                  int32
          alcohol_encoded                   int32
          urinary_obstructions_encoded      int32
          smoking_encoded                   int32
          obesity_encoded                   int32
```

Split the Dataset.



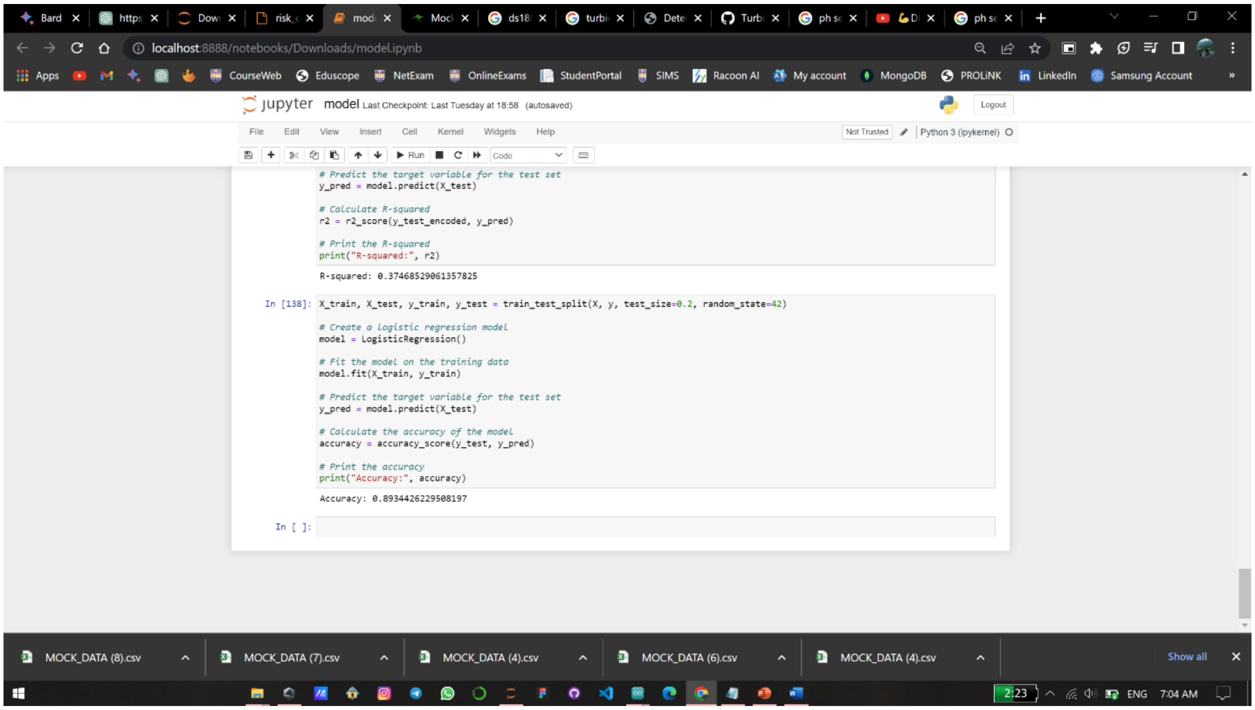
```
In [133]: X = df.drop('diagnose', axis=1)  # Update 'target_variable' with the actual column name
          y = df['diagnose']

In [134]: print(df.dtypes)

          age                                 int64
          urine_frequency_stage               int64
          urine_color                         int64
          diagnose                           object
          gender_encoded                      int32
          diabetic_encoded                    int32
          alcohol_encoded                     int32
          urinary_obstructions_encoded        int32
          smoking_encoded                     int32
          obesity_encoded                     int32
          edema_symptoms_encoded              int32
          prolong_use_of_medication_encoded   int32
          family_history_encoded              int32
          dtype: object

In [135]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=0)

In [136]: # Create a LabelEncoder object
          label_encoder = LabelEncoder()

          # Encode the target variable
          y_train_encoded = label_encoder.fit_transform(y_train)
          y_test_encoded = label_encoder.transform(y_test)

          # Build the linear regression model
          model = LinearRegression()
```

Build the Linear regression Model & get the accuracy.



```
In [136]: # Create a LabelEncoder object
          label_encoder = LabelEncoder()

          # Encode the target variable
          y_train_encoded = label_encoder.fit_transform(y_train)
          y_test_encoded = label_encoder.transform(y_test)

          # Build the linear regression model
          model = LinearRegression()
          model.fit(X_train, y_train_encoded)

          # Predict the target variable for the test set
          y_pred = model.predict(X_test)

          # Calculate mean squared error (MSE)
          mse = mean_squared_error(y_test_encoded, y_pred)

          # Print the MSE
          print("Mean Squared Error (MSE):", mse)

          Mean Squared Error (MSE): 0.10486331057702959

In [137]: from sklearn.metrics import r2_score

          # Predict the target variable for the test set
          y_pred = model.predict(X_test)

          # Calculate R-squared
          r2 = r2_score(y_test_encoded, y_pred)

          # Print the R-squared
          print("R-squared:", r2)

          R-squared: 0.37468529061357825

In [138]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

          # Create a Logistic regression model
```

Build the Logistic regression Model & get the accuracy.



localhost:8888/notebooks/Downloads/model.ipynb
jupyter model Last Checkpoint: Last Tuesday at 18:58 (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
Run
Code
MOCK_DATA (8).csv
MOCK_DATA (7).csv
MOCK_DATA (4).csv
MOCK_DATA (6).csv
MOCK_DATA (4).csv
Show all
2:23
ENG
7:04 AM

```
# Predict the target variable for the test set
y_pred = model.predict(X_test)

# Calculate R-squared
r2 = r2_score(y_test_encoded, y_pred)

# Print the R-squared
print("R-squared:", r2)
```

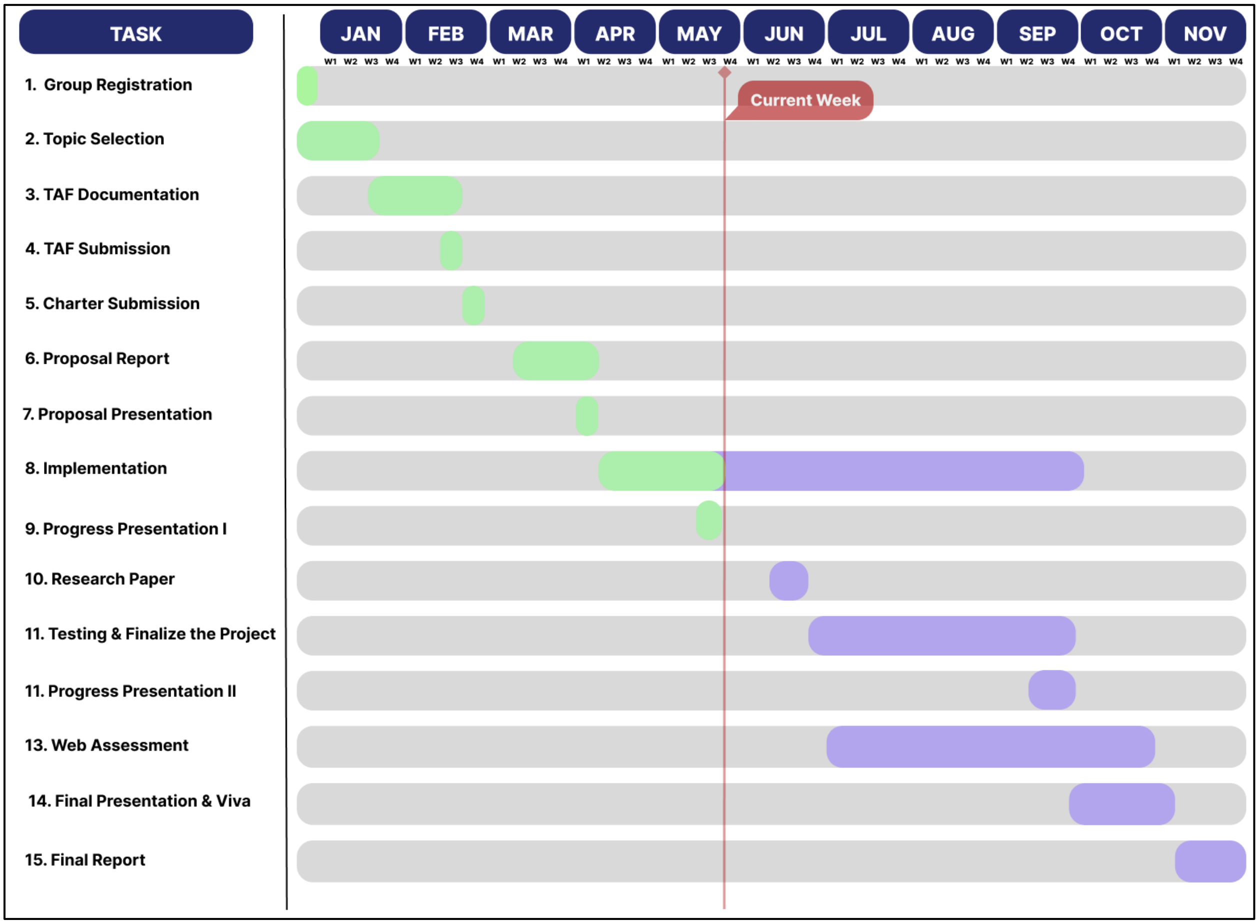
```
R-squared: 0.37468529061357825
```

```
In [138]: X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

          # Create a logistic regression model
          model = LogisticRegression()

          # Fit the model on the training data
          model.fit(X_train, y_train)

          # Predict the target variable for the test set
          y_pred = model.predict(X_test)

          # Calculate the accuracy of the model
          accuracy = accuracy_score(y_test, y_pred)

          # Print the accuracy
          print("Accuracy:", accuracy)
```

```
Accuracy: 0.8934426229508197
```

```
In [ ]:
```

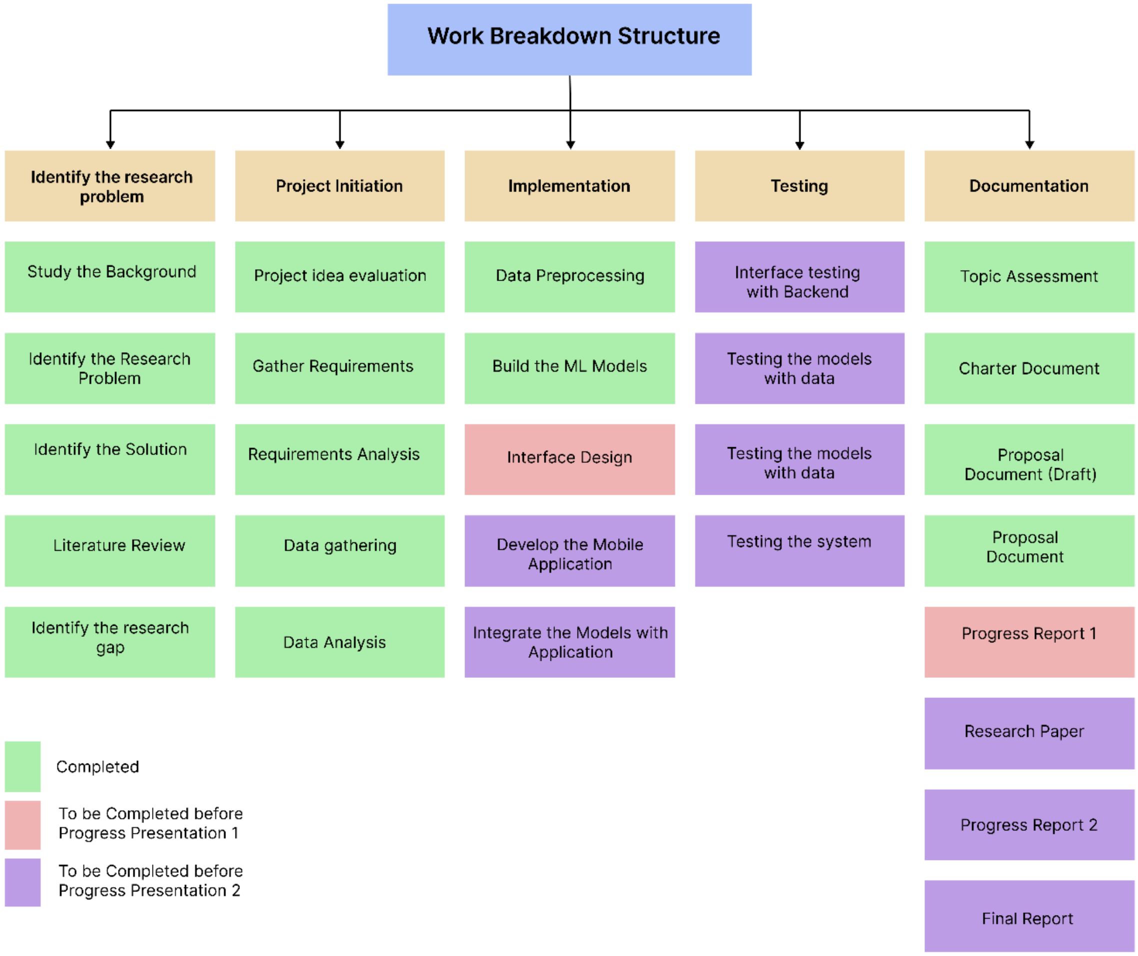
4 Gantt Chart



🟩 - Completed

🟪 - In Progress

5 Work Breakdown Structure


Work Breakdown Structure
Identify the research problem
Study the Background
Identify the Research Problem
Identify the Solution
Literature Review
Identify the research gap
Project Initiation
Project idea evaluation
Gather Requirements
Requirements Analysis
Data gathering
Data Analysis
Implementation
Data Preprocessing
Build the ML Models
Interface Design
Develop the Mobile Application
Integrate the Models with Application
Testing
Interface testing with Backend
Testing the models with data
Testing the models with data
Testing the system
Documentation
Topic Assessment
Charter Document
Proposal Document (Draft)
Proposal Document
Progress Report 1
Research Paper
Progress Report 2
Final Report
🟩 Completed
🟥 To be Completed before Progress Presentation 1
🟪 To be Completed before Progress Presentation 2