Analytics for Hospital's Health-Care Data

Team ID: PNT2022TMID23753

1.INTRODUCTION

1.1 PROJECT OVERVIEW

Data analytics in clinical settings attempts to reduce patient wait times via improved scheduling and staffing, give patients more options when scheduling appointments and receiving treatment, and reduce readmission rates by using population health data to predict which patients are at greatest risk.

.The healthcare industry generates a tremendous amount of data but struggles to convert that data into insights that improve patient outcomes and operational efficiencies. Data analytics in healthcare is intended to help providers overcome obstacles to the widespread application of dataderived intelligence:

1. Making healthcare data easier to share among colleagues and external partners, and easier to visualize for public consumption
2. Providing accurate data-driven forecasts in real time to allow healthcare providers to respond more quickly to changing healthcare markets and environments
3. Enhancing data collaboration and innovation among healthcare organizations to convert analytics-ready data into business-ready information by automating low-impact data management tasks.

1.2 PURPOSE

Data analytics in health care is vital. It helps health care organizations to evaluate and develop practitioners, detect anomalies in scans and predict outbreaks in illness, per the Harvard Business School. Data analytics can also lower costs for health care organizations and boost business intelligence. Most importantly, it helps health care companies to make better care decisions for patients.

2. LITERATURE SURVEY

2.1 EXISTING PROBLEM

When it comes to big data analytics, the healthcare industry faces numerous challenges. These challenges may include security, visualization, and a wide array of data integrity concerns. Over the years, big data analytics in healthcare has emerged as one of the most challenging undertakings for the healthcare industry. For instance, healthcare professionals, who may not be well-versed with managing electronic health records, now need to gather actionable insights as well. Also, they are expected to apply

those learnings to complex initiatives that enhance their overall reimbursement rates.

2.2 REFERENCE

1)Mohammad Alkhatib , Amir Talaei-Khoei (University of Nevada,Reno)Amir Talaei-Khoei University of Nevada, Reno | UNR · Department of Accounting and Information Systems PhD of Information Systems-Amir Ghapanchi

2) From:"Book of Data Analytics" Chandank Reddy(Wayne State University) Charu C.Aggarwal(Watson Research Center)

3) From: Hoyt,RE,Yoshihashi,A,Eds.(2014).Health Informatics:Practical Guide for Healthcare and formation Technology Professionals,Sixth Edition.Pensacola,FL,Lulu.com

4) Panagiota Galetsia , Korina Katsaliakia , Sameer Kumarb,⁎ a School of Economics, Business Administration & Legal Studies, International Hellenic University, 14th km Thessaloniki-N. Moudania, Thessaloniki, 57001, Greece b Opus College of Business, University of St. Thomas Minneapolis Campus, 1000 LaSalle Avenue, Schulze Hall 435, Minneapolis, MN 55403, USA

5) from"n book: Innovative Data Communication Technologies and Application (pp.83-96)"P. Nagaraj-Professor (Assistant) at Kalasalingam University.

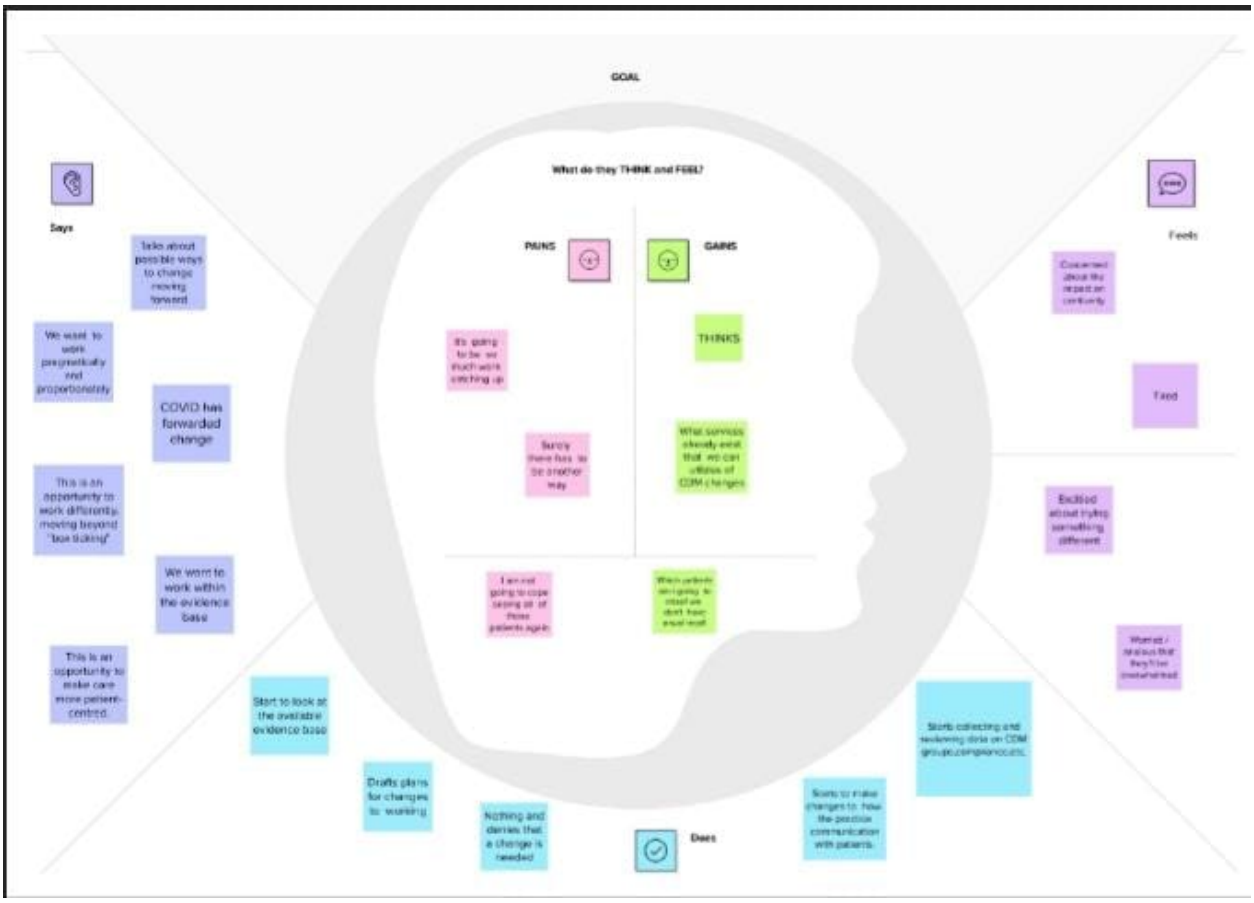
2.3 PROBLEM STATEMENT DEFINITION

Humanity today is suffering from one of the most dangerous pandemics in history, the Coronavirus Disease of 2019 (COVID-19). Although today there is immense advancement in the medical field with the latest technology, the COVID-19 pandemic has affected us severely. The virus is spreading rapidly, resulting in an escalation in the number of patients admitted. We propose a contextual patient classification system for better analysis of the data from the discharge summary available from the research hospital. The classification was done using the Knuth–Morris–Pratt algorithm. We have also analyzed the data of COVID-19 and nonCOVID-19 patients. During the analysis, studies on the medicines, medical services and tests, pulse count, body temperature, and the overall effect of age and gender was done. The death versus survival ratio for the COVID-19 positive patients has also been studied. The classification accuracy of the contextual patient classification system achieved was 97.4%. The combination of data analysis and contextual patient classification will be helpful to all the sectors to be better prepared for any future waves of the COVID-19 pandemic. The goal is to accurately predict the Length of Stay for each patient on case by case basis so that the Hospitals can use this information for optimal resource allocation and better functioning. The length of stay is divided into 11 different classes ranging from 0-10

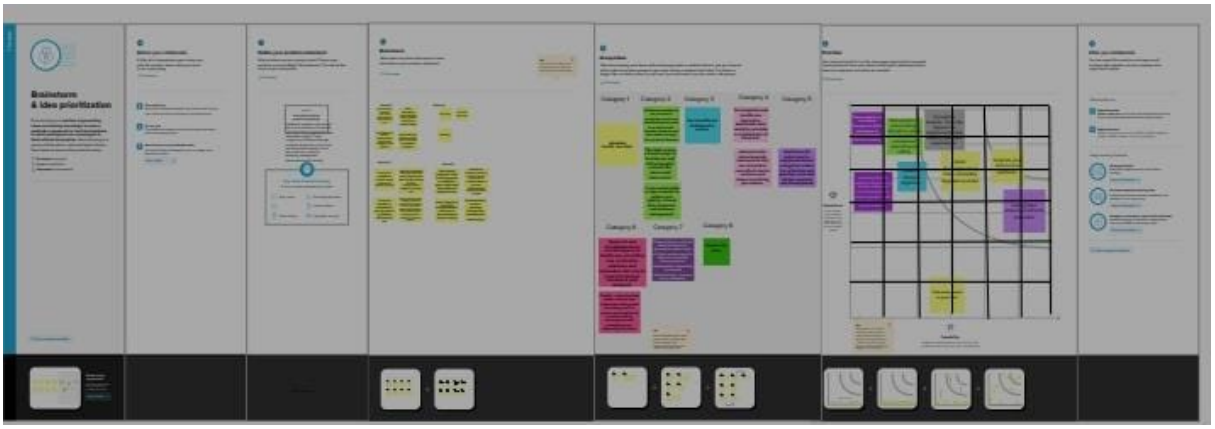
days to more than 100 days. The tools that we are using for data analytics is Cognos Analytics from IBM.

3. IDEATION AND PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS



3.2 IDEATION AND BRAINSTORMING



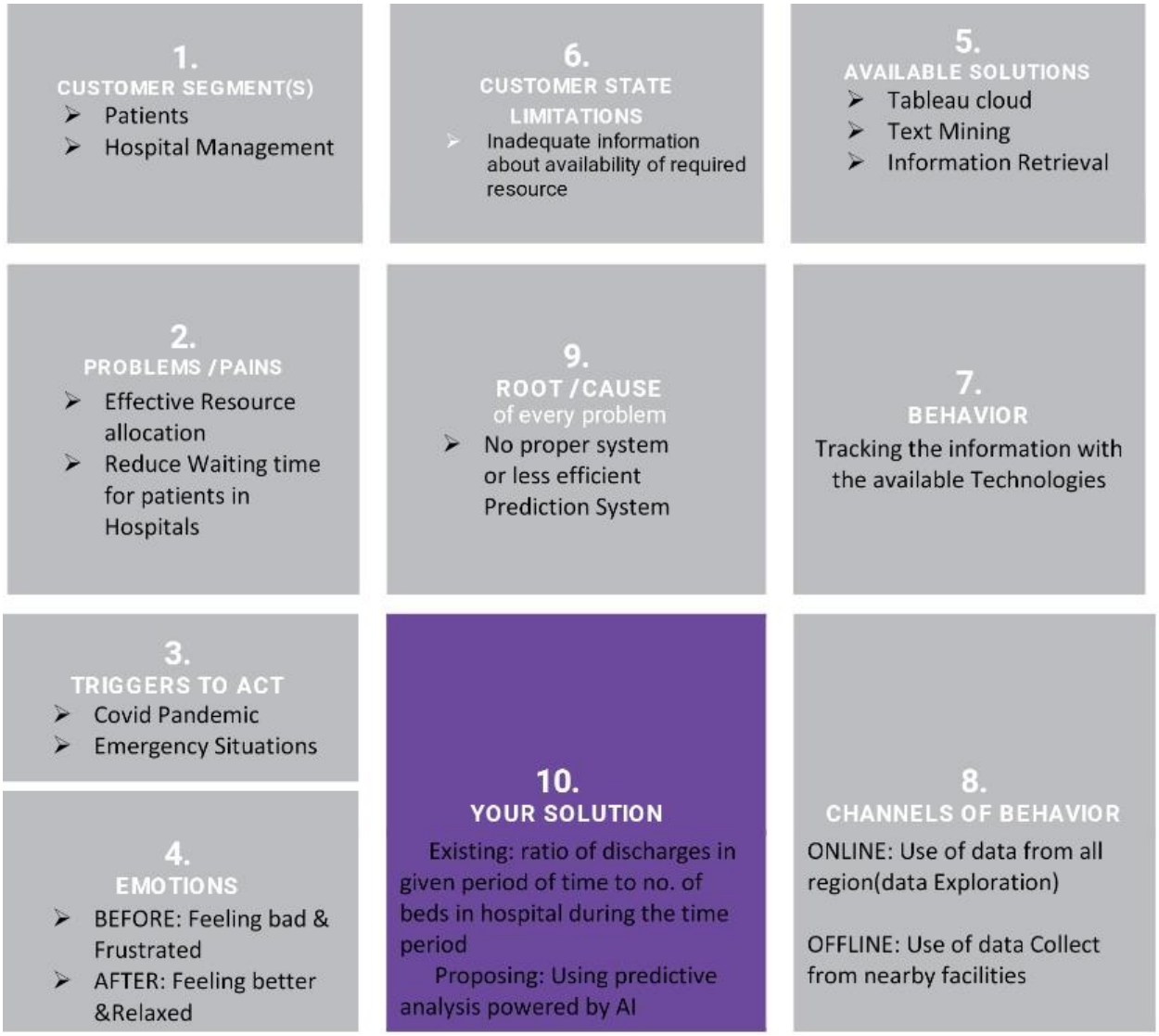
3.3 PROPOSED SOLUTION

S.NO	Parameters	Description
1	Problem Statement (Problem to be solved)	During the covid-19 pandemic, we have faced one of the difficult times of our life. Everyone seeks to survive from the great disaster. At the time of pandemic, noone get to know about which hospital has vacant beds(free beds) to admit themselves or others infected by covid. This situation made the death rate higher.
2	Idea / Solution description	Predictive analytics can create patient journey dashboards and disease trajectories that helps us to know about the patient's period of stay. It improves effective allocation of beds and other resources, treatment delivery, improves efficiencies, and so on.
3	Novelty / Uniqueness	Healthcare data frequently resides in several locations. The Collected data should be stored in central system(like centralized storage). This data becomes accessible and usable when it is combined into a single, central system, such as an enterprise datawarehouse

		(EDW). Uniqueness of our project is that we can able to use data for different things such as which medicine is more effective and for understanding behavioural pattern of particular disease.
4	Social Impact / Customer Satisfaction	effective use of resource Enhanced diagnosis Improved Treatment enhancing the overall quality of treatment and life of patients.
5	Business Model (Revenue Model)	With the gathered data, redirecting the patients to particular hospital based on the vacancy, leading retailers used methods like market-basket analysis to discover insights about consumer purchase behaviour and used these insights to optimize the physical store experience, target relevant ads and streamline the supply chain, among other strategic initiatives.
6	Scalability of the Solution	A variety of institutions must store, evaluate, and take action on the massive amounts of data being produced by the health care sector as it expands quickly. India is a vast, culturally varied nation with

		a sizable population that is increasingly able to access centralised healthcare services.

3.4 PROBLEM SOLUTION FIT



4. REQUIREMENTS ANALYSIS

4.1 FUNCTIONAL REQUIREMENTS

FR NO	Functional Requirement	Sub Requirement (Story /

	(Epic)	Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail
FR-2	User Confirmation	Confirmation via Email Confirmation via Message
FR-3	Interoperability	Dashboard helps to share the patient's information interoperable to the hospitals in timely manner.
FR-4	Accuracy	Dashboard helps predict the patient's Health risks accurately based on LOS (Length of Stay).
FR-5	Compliance	The compliance of a dashboard is like to use very interactively in real time by the hospitals.
FR-6	Concise	These dashboards are clear, intuitive, and customizable and interactive in manner

4.2 NON- FUNCTIONAL REQUIREMENTS

NFR NO	Non-Functional Requirement	Description
NFR-1	Usability	This Dashboards are designed to offer a comprehensive overview of patient's LOS, and do so through the use of data visualization tools like charts and graphs.
NFR-2	Security	The Dashboard helps to indicate the current threat level to the Hospitals; an

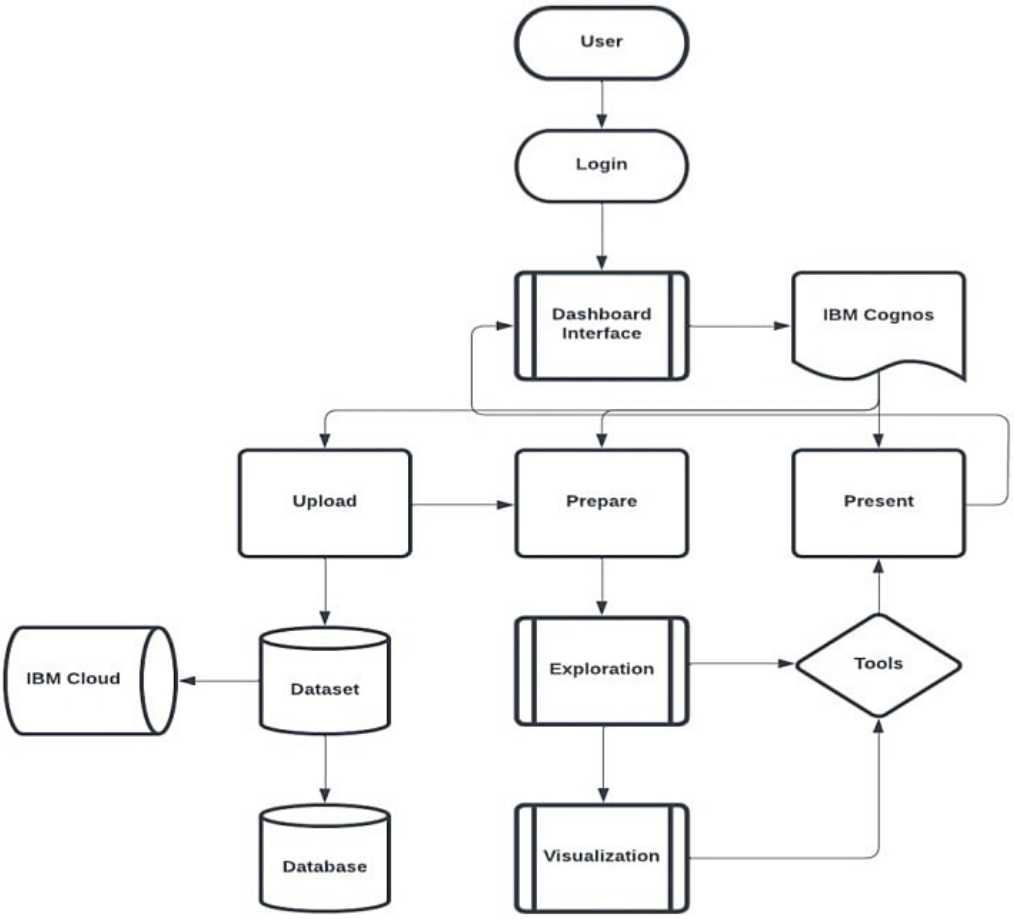
		indication of events and incidents that have occurred; a record of authentication errors; unauthorized access
NFR-3	Reliability	This dashboard will be consistent and reliable to the users and helps the user to use in effective, efficient and reliable manner
NFR-4	Performance	This dashboard can scan the backend users and analyzing the frequency in which they visit the dashboard helps understand how useful and helpful the data displayed is for tasks.
NFR-5	Availability	The dashboard can available to meet user's demand in timely manner and it is also helps to provide necessary information to the user's dataset
NFR-6	ScalabilitY	The layers used in the dashboard are a hosted feature layer, feature layer view, or hosted tile layer.

5. PROJECT DESIGN

5.1 DATA FLOW DIAGRAM

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

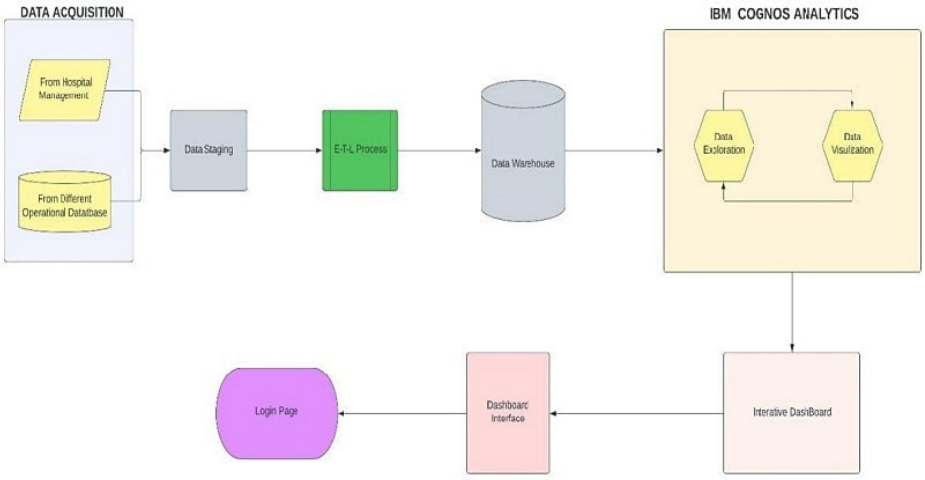
the information, and where data is stored.



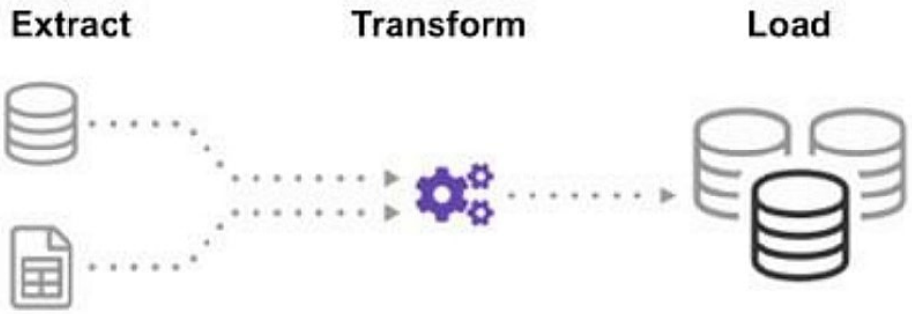
5.2 SOLUTION &TECHNICAL ARCHITECTURE

SOLUTION ARCHITECTURE

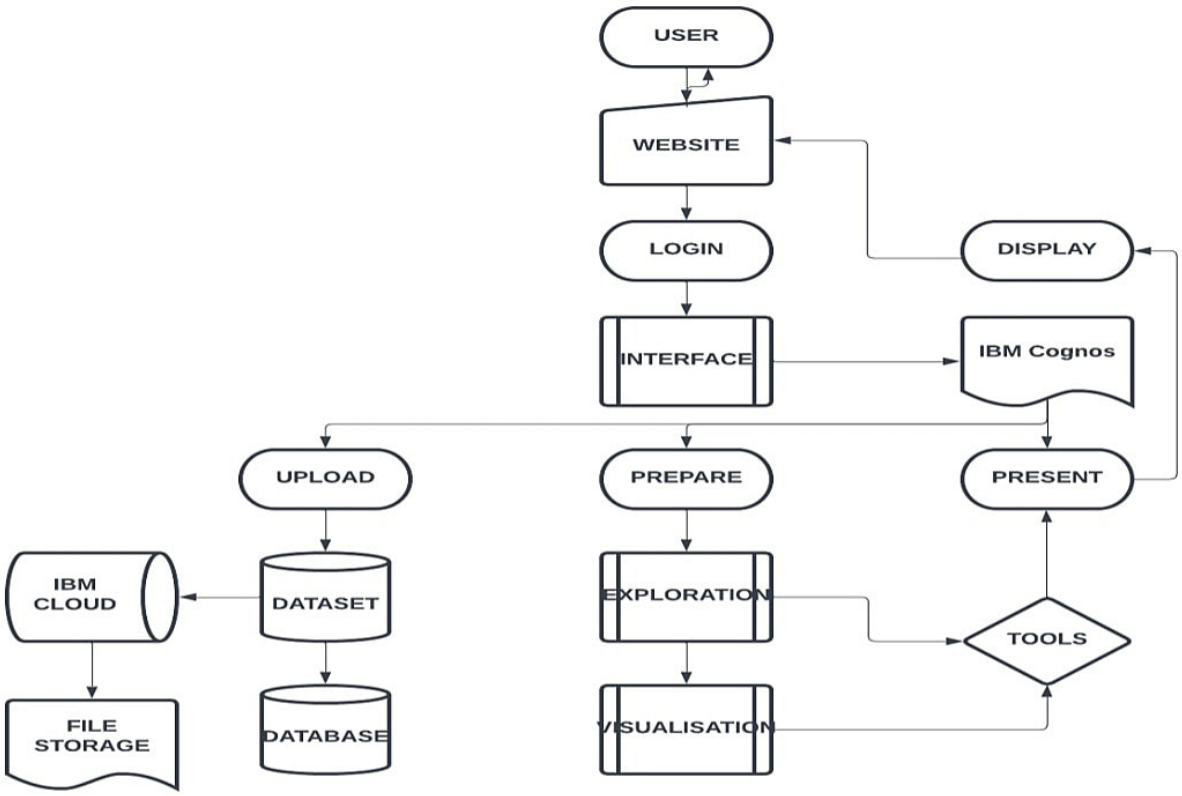
SYSTEM ARCHITECTURE:



ETL PROCESS (DATA INTEGRATION PROCESS):



TECHNICAL ARCHITECTURE



5.3 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 (Mobile user)	Registration	USN-1	As a user, I can register for the dashboard by entering my email, and password, and confirming my password.	I can access my account in the dashboard	High	Sprint-1
		USN-2	As a user, I will receive a confirmation email once I have registered for the dashboard	I can receive a confirmation email & click confirm	High	Sprint-1
		USN-3	As a user, I can register for the dashboard through Social Media	I can register & access the dashboard with Social Media Login	Low	Sprint-2
		USN-4	As a user, I can register for the dashboard through Gmail	I can register and access dashboard with Gmail	Medium	Sprint-2
	Login	USN-5	As a user, I can log into the application by entering email & password	I can login to the account in my email login.	High	Sprint-2
	Dashboard	USN-6	As a user ,I can use my account in my dashboard for uploading dataset.	I can login to the account for uploading dataset.	Medium	Sprint-3
Customer (Web user)	Website	USN-7	As a user ,I can use my dashboard in website	I can login into the dashboard by visiting website.	Medium	Sprint-3
Customer Care Executive		USN-8	As a user ,I can contact Customer care Executive for my login.	I can contact customer executive for my login.	High	Sprint-4
Administrator		USN-9	As a user ,I can contact administrator for my queries.	I can contact administrator for solving my queries.	High	Sprint-4
Exploration	Dashboard	USN-10	As a user, I can prepare data by using Exploration Techniques.	I can prepare data by using Exploration Techniques.	High	Sprint-3
Presentation	Dashboard	USN-11	As a user, I can Present data in my dashboard.	I can present data by using my account in dashboard.	High	Sprint-4
Visualization	Dashboard	USN-12	As a user, I can Prepare Data by using Visualization Techniques	I can prepare data by using Visualization Techniques.	High	Sprint-3

6.PROJECT PLANNING AND SCHEDULING

6.1 SPRINT PLANNING AND ESTIMATION

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	2	High	Priyadharshini D, Narmadha Varshini N
Sprint-1		USN-2	As a user,	1	High	Priyadhars

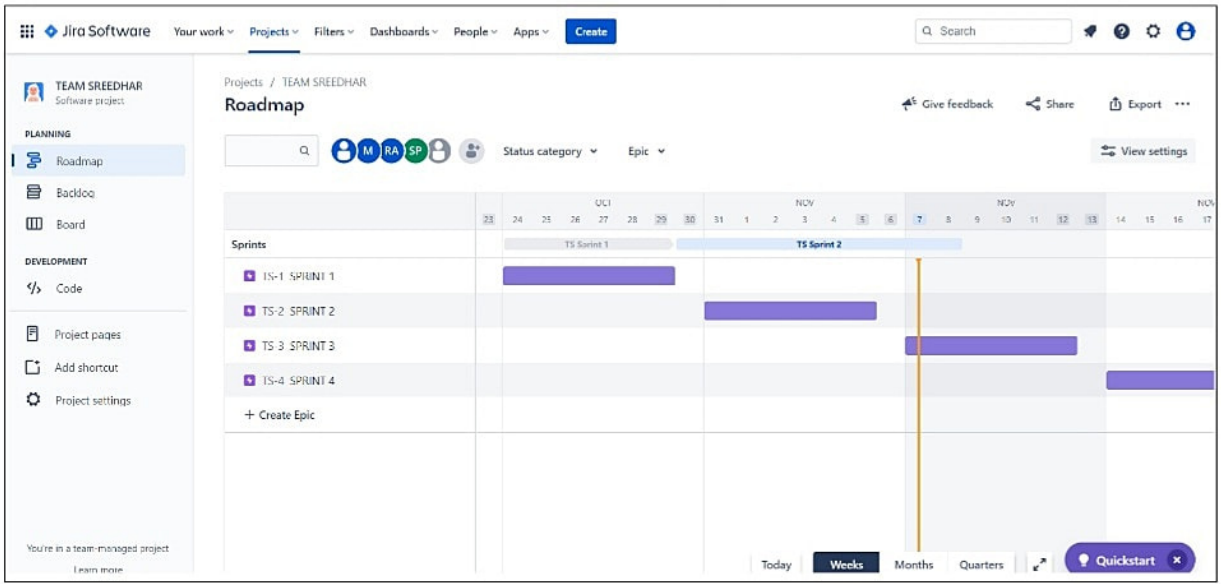
			I will receive confirmation email once I have registered for the application			hini D, Yasotha S
Sprint-2		USN-3	As a user, I can register for the application through Facebook	2	Low	Priyadharshini D, Yasotha S
Sprint-1		USN-4	As a user, I can register for the application through Gmail	2	Medium	Yamuna Devi V,Priyadharshini D
Sprint-1	Login	USN-5	As a user, I can log into the application by entering email & password	1	High	Yamuna Devi V,Priyadharshini D
	Dashboard					

6.2 SPRINT DELIVERY SCHEDULE

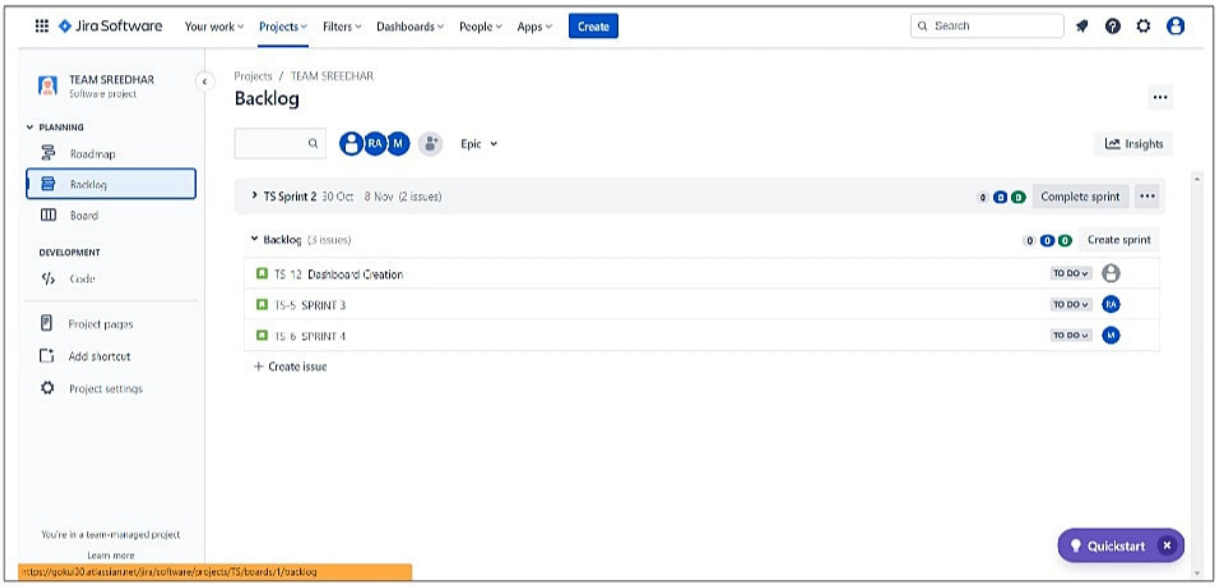
Sprints	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	11 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	15 Nov 2022

6.3 REPORTS FROM JIRA

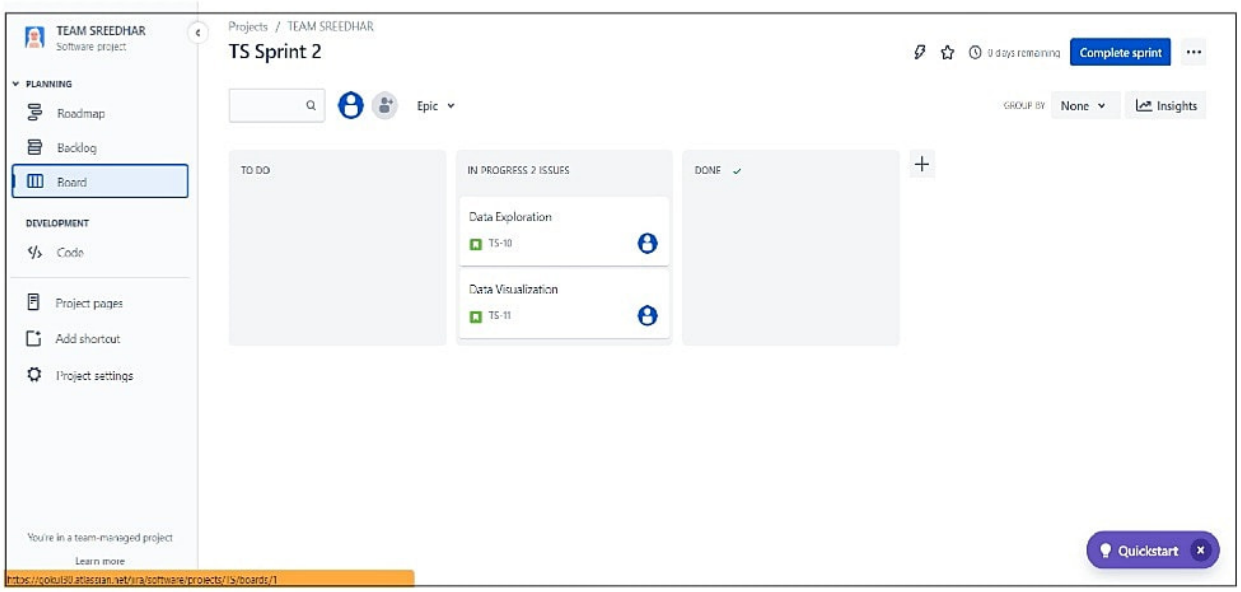
ROADMAP



BACKLOGS



BOARD



7. CODING AND SOLUTIONING

7.1 FEATURE 1

Data Collection

Download the dataset here

```
[ ]  from google.colab import drive
     drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

```
[ ]  cd/content/drive/MyDrive/Colab Notebooks
```

```
/content/drive/MyDrive/Colab Notebooks
```

```
[ ]  # Unzipping the dataset
     !unzip 'Dataset.zip'
```

Image Preprocessing

```
[ ]  from keras.preprocessing.image import ImageDataGenerator
```

Image Data Augmentation

```
[ ]  train_datagen = ImageDataGenerator(rescale=1./255,shear_range=0.2,zoom_range=0.2,horizontal_flip=True)
     test_datagen=ImageDataGenerator(rescale=1./255)
```

Applying Image DataGenerator Functionality To Trainset And Testset

```
x_train = train_datagen.flow_from_directory(
    r'/content/drive/MyDrive/Colab Notebooks/Dataset/TRAIN_SET',
    target_size=(64, 64),batch_size=5,color_mode='rgb',class_mode='sparse')
x_test = test_datagen.flow_from_directory(
    r'/content/drive/MyDrive/Colab Notebooks/Dataset/TEST_SET',
    target_size=(64, 64),batch_size=5,color_mode='rgb',class_mode='sparse')
```

Model Building

1. Importing The Model Building Libraries

```
[ ]  import numpy as np
     import tensorflow as tf
     from tensorflow.keras.models import Sequential
     from tensorflow.keras import layers
     from tensorflow.keras.layers import Dense,Flatten
     from tensorflow.keras.layers import Conv2D,MaxPooling2D,Dropout
```

2. Initializing The Model

```
[ ]  classifier = Sequential()
```

3. Adding CNN Layers

```
[ ]  classifier = Sequential()
     classifier.add(Conv2D(32, (3, 3), input_shape=(64, 64, 3), activation='relu'))
     classifier.add(MaxPooling2D(pool_size=(2, 2)))
     classifier.add(Conv2D(32, (3, 3), activation='relu'))
     classifier.add(MaxPooling2D(pool_size=(2, 2)))
     classifier.add(Flatten())
```

4. Adding Dense Layers

```
[ ]  classifier.add(Dense(units=128, activation='relu'))
     classifier.add(Dense(units=5, activation='softmax'))
```

```
     classifier.summary()
```

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 62, 62, 32)        896
```

5. Configure The Learning Process

```
[ ]  classifier.compile(optimizer='adam', loss='sparse_categorical_crossentropy', metrics=['accuracy'])
```

6. Train The Model

```
[ ]  classifier.fit_generator(generator=x_train,steps_per_epoch = len(x_train),epochs=20, validation_data=x_test,validation_steps = len(x_test))
```

```
/usr/local/lib/python3.7/dist-packages/ipykernel_launcher.py:2: UserWarning: `Model.fit_generator` is deprecated and will be removed in a future version. Pl

Epoch 1/20
494/824 [==========>.....] - ETA: 6:52 - loss: 0.7194 - accuracy: 0.7174
```

7. Saving The Model

```
[ ]  classifier.save('nutrition.h5')
```

8. Testing The Model

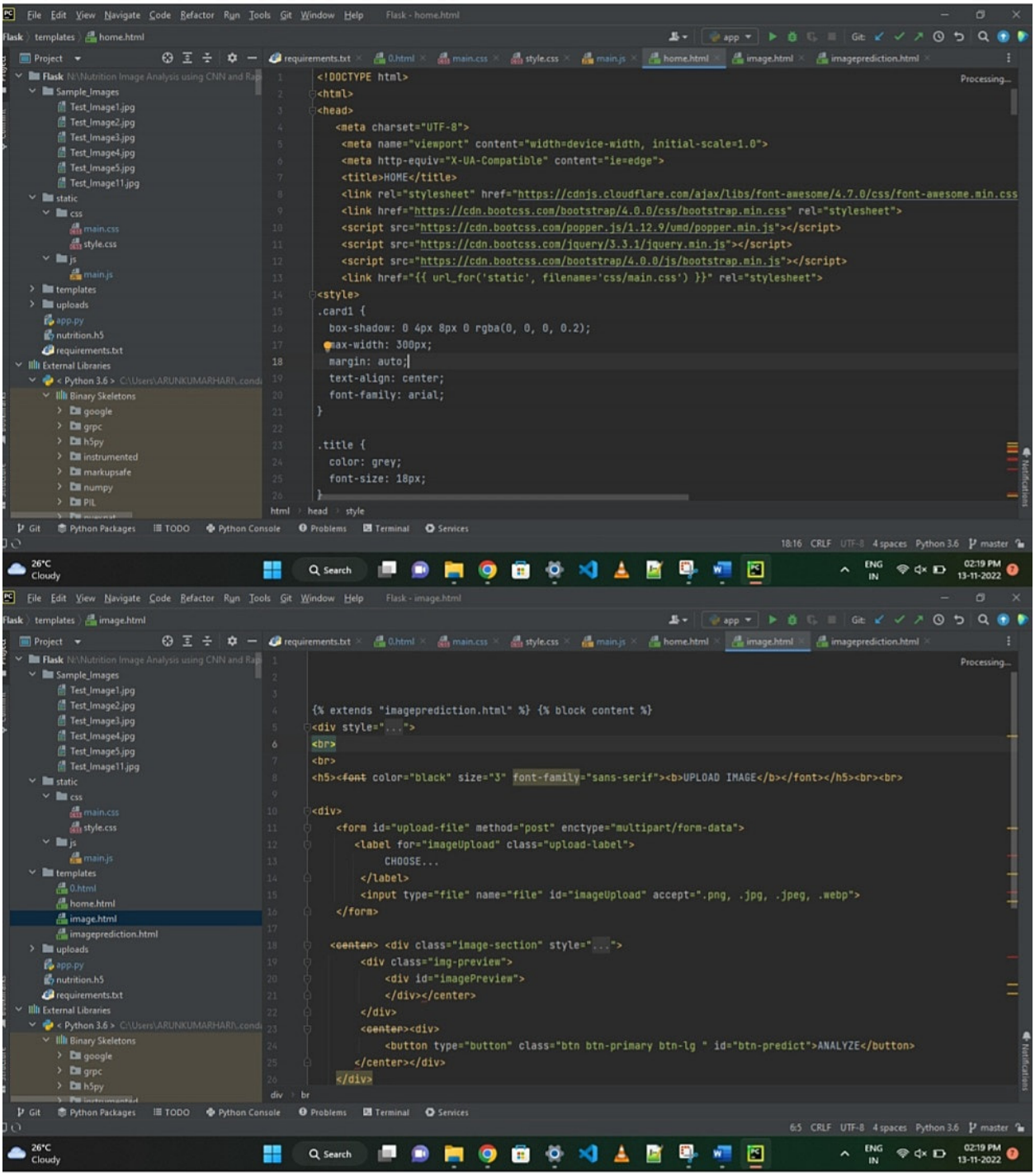
```
[ ]  from tensorflow.keras.models import load_model
     from keras.preprocessing import image
     model = load_model("nutrition.h5")
```

```
from tensorflow.keras.models import load_model
from tensorflow.keras.preprocessing import image
model = load_model("nutrition.h5")
img = image.load_img(r'/content/drive/MyDrive/Colab Notebooks/Sample_Images/Test_Image1.jpg',grayscale=False,target_size= (64,64))
x = img_to_array(img)
x = np.expand_dims(x,axis = 0)
predict_x=model.predict(x)
classes_x=np.argmax(predict_x,axis=-1)
classes_x
```

```
1/1 [==============================] - 0s 62ms/step
array([0])
```

```
[ ]  index=['APPLES', 'BANANA', 'ORANGE','PINEAPPLE','WATERMELON']
     result=str(index[classes_x[0]])
     result
```

7.2 FEATURE 2

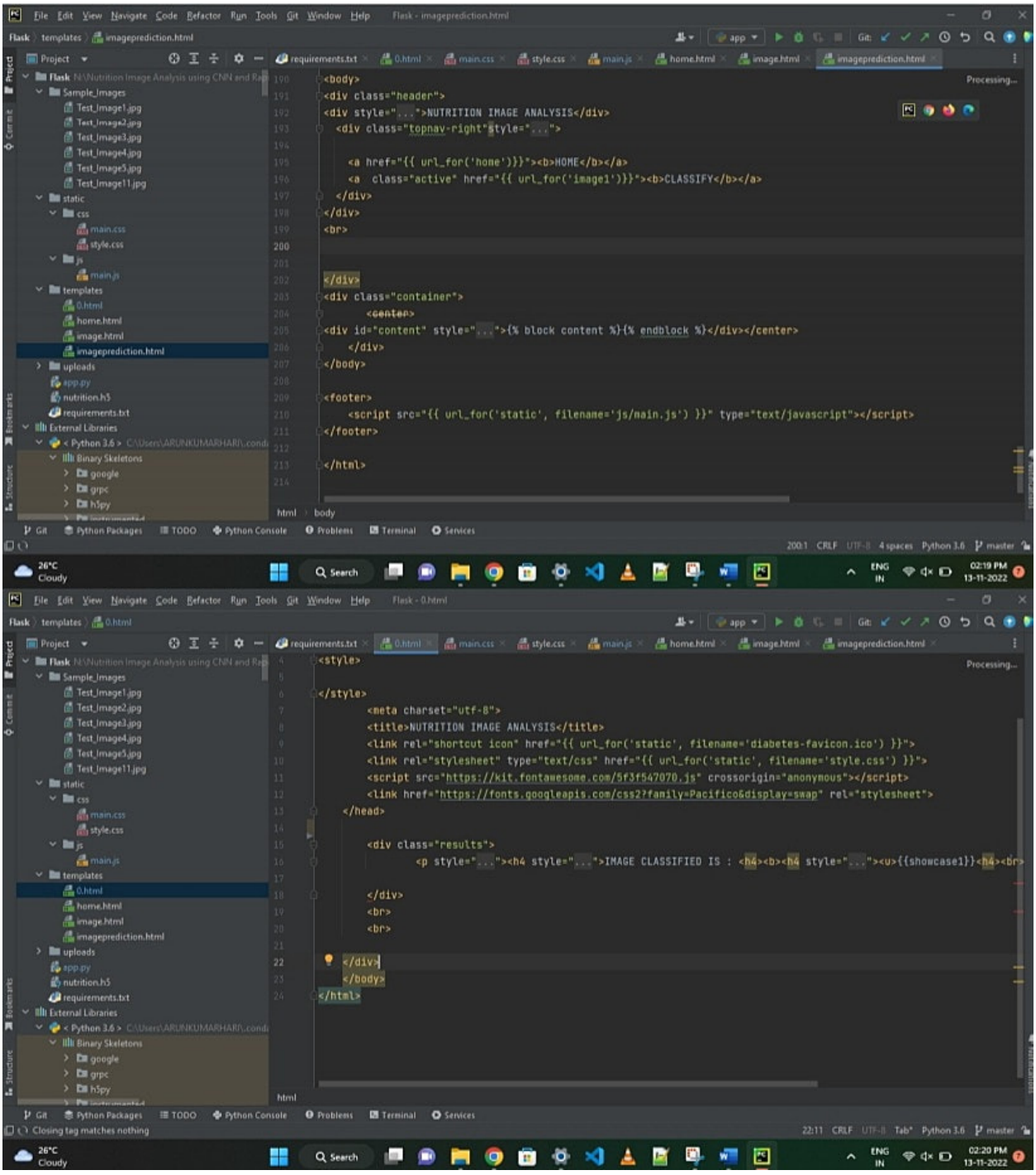

File Edit View Navigate Code Refactor Run Tools Git Window Help
Flask - home.html
Flask templates home.html
app
Git:
Project
requirements.txt
0.html
main.css
style.css
main.js
home.html
image.html
imageprediction.html
Processing...
Flask N:\Nutrition Image Analysis using CNN and Rap
Sample_Images
Test_Image1.jpg
Test_Image2.jpg
Test_Image3.jpg
Test_Image4.jpg
Test_Image5.jpg
Test_Image11.jpg
static
css
main.css
style.css
js
main.js
templates
uploads
app.py
nutrition.h5
requirements.txt
External Libraries
< Python 3.6 > C:\Users\ARUNKUMARHARI\.conda
Binary Skeletons
google
grpc
h5py
instrumented
markupsafe
numpy
PIL
<!DOCTYPE html>
<html>
<head>
<meta charset="UTF-8">
<meta name="viewport" content="width=device-width, initial-scale=1.0">
<meta http-equiv="X-UA-Compatible" content="ie=edge">
<title>HOME</title>
<link rel="stylesheet" href="https://cdnjs.cloudflare.com/ajax/libs/font-awesome/4.7.0/css/font-awesome.min.css
<link href="https://cdn.bootcss.com/bootstrap/4.0.0/css/bootstrap.min.css" rel="stylesheet">
<script src="https://cdn.bootcss.com/popper.js/1.12.9/umd/popper.min.js"></script>
<script src="https://cdn.bootcss.com/jquery/3.3.1/jquery.min.js"></script>
<script src="https://cdn.bootcss.com/bootstrap/4.0.0/js/bootstrap.min.js"></script>
<link href="{{ url_for('static', filename='css/main.css') }}" rel="stylesheet">
<style>
.card1 {
box-shadow: 0 4px 8px 0 rgba(0, 0, 0, 0.2);
max-width: 300px;
margin: auto;
text-align: center;
font-family: arial;
}
.title {
color: grey;
font-size: 18px;
}
html head style
Git Python Packages TODO Python Console Problems Terminal Services
18:16 CRLF UTF-8 4 spaces Python 3.6 master
26°C Cloudy
Search
ENG IN
02:19 PM 13-11-2022
File Edit View Navigate Code Refactor Run Tools Git Window Help
Flask - image.html
Flask templates image.html
{% extends "imageprediction.html" %} {% block content %}

<h5>UPLOAD IMAGE</h5>

<form id="upload-file" method="post" enctype="multipart/form-data">
<label for="imageUpload" class="upload-label">
CHOOSE...
</label>
<input type="file" name="file" id="imageUpload" accept=".png, .jpg, .jpeg, .webp">
</form>

<button type="button" class="btn btn-primary btn-lg " id="btn-predict">ANALYZE</button>

div br
6:5 CRLF UTF-8 4 spaces Python 3.6 master
26°C Cloudy
Search
ENG IN
02:19 PM 13-11-2022


Flask - imageprediction.html
requirements.txt
0.html
main.css
style.css
main.js
home.html
image.html
imageprediction.html
Processing...
<body>

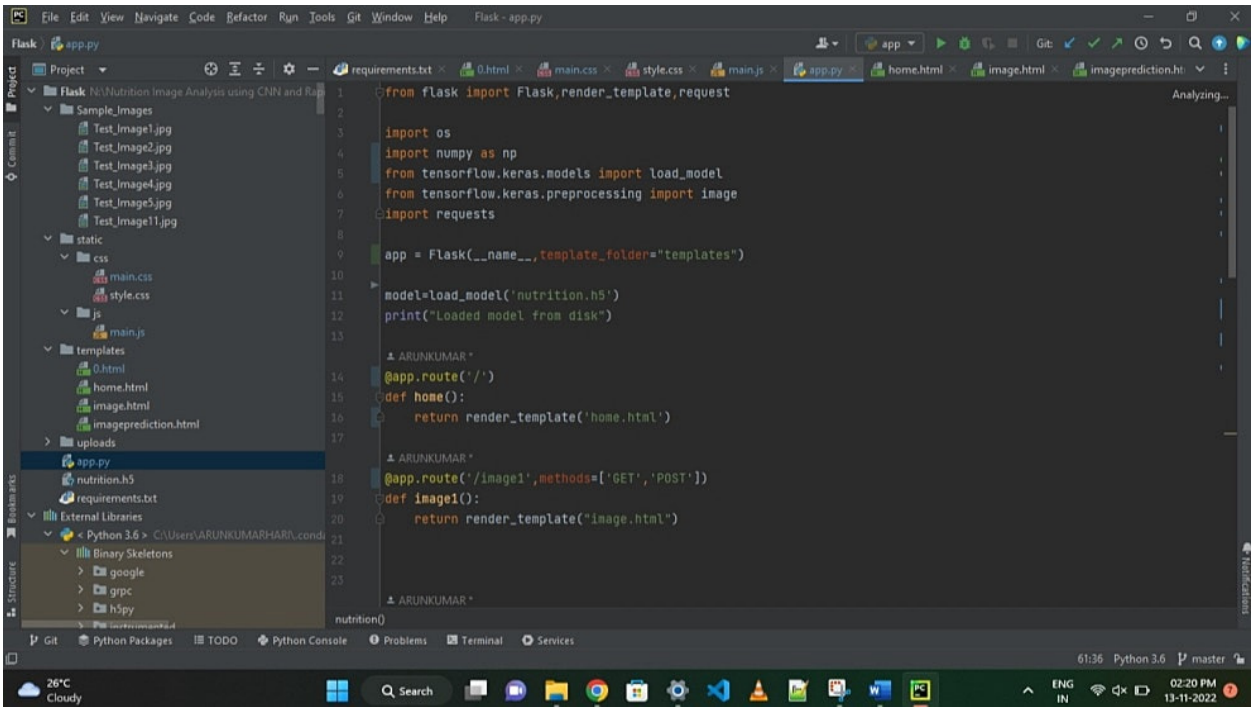
NUTRITION IMAGE ANALYSIS
<div class="topnav-right"style="...">
HOME
CLASSIFY

{% block content %}{% endblock %}

</body>
<footer>
<script src="{{ url_for('static', filename='js/main.js') }}" type="text/javascript"></script>
</footer>
</html>
Sample_Images
Test_Image1.jpg
Test_Image2.jpg
Test_Image3.jpg
Test_Image4.jpg
Test_Image5.jpg
Test_Image11.jpg
static
css
main.css
style.css
js
main.js
templates
0.html
home.html
image.html
imageprediction.html
uploads
app.py
nutrition.h5
requirements.txt
External Libraries
Binary Skeletons
google
grpc
h5py
Git
Python Packages
TODO
Python Console
Problems
Terminal
Services
26°C
Cloudy
Search
ENG
IN
02:19 PM
13-11-2022
Flask - 0.html
<style>
</style>
<meta charset="utf-8">
<title>NUTRITION IMAGE ANALYSIS</title>
<link rel="shortcut icon" href="{{ url_for('static', filename='diabetes-favicon.ico') }}">
<link rel="stylesheet" type="text/css" href="{{ url_for('static', filename='style.css') }}">
<script src="https://kit.fontawesome.com/5f3f547070.js" crossorigin="anonymous"></script>
<link href="https://fonts.googleapis.com/css2?family=Pacifico&display=swap" rel="stylesheet">
</head>

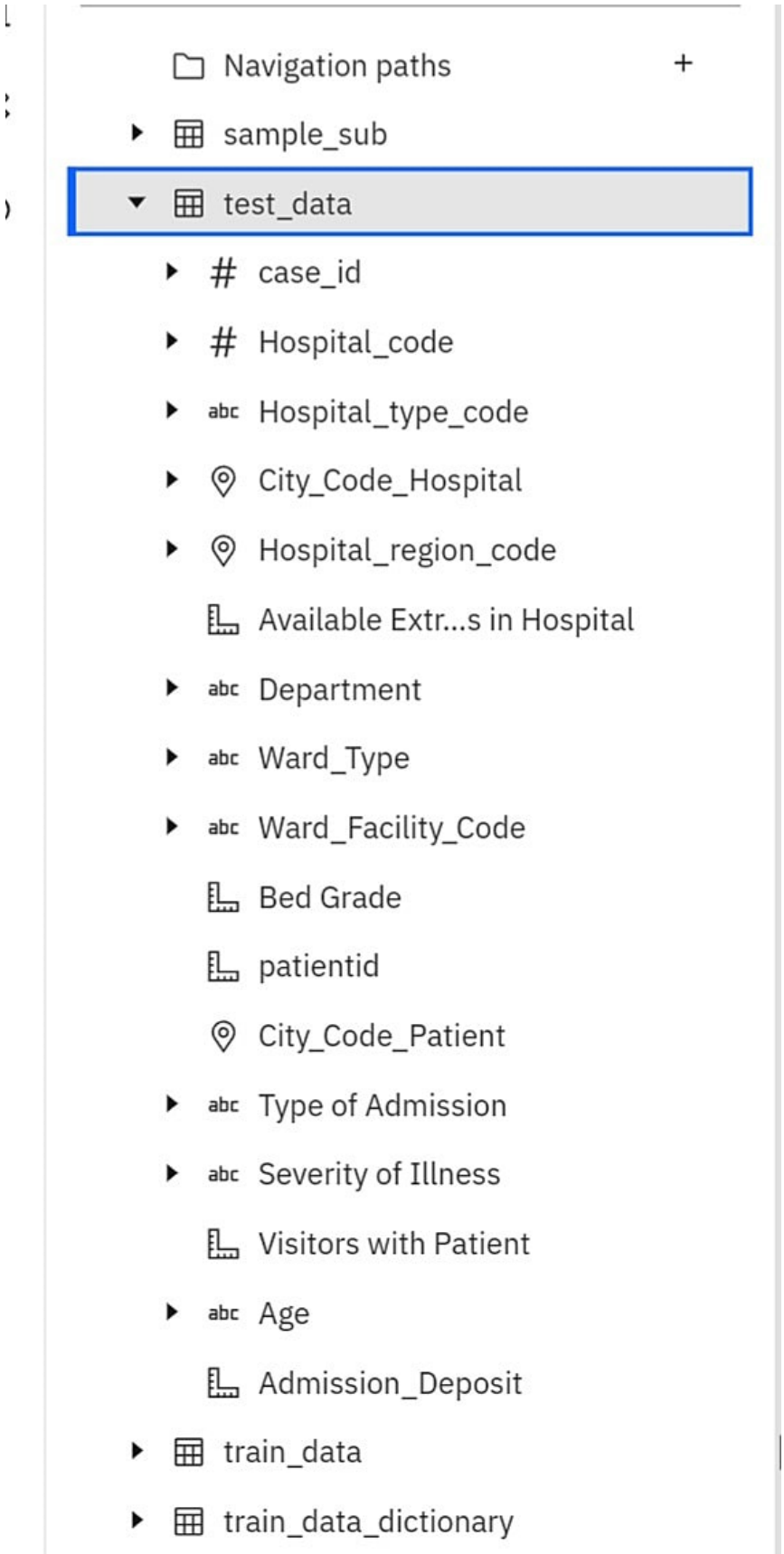
<p style="..."><h4 style="...">IMAGE CLASSIFIED IS : <h4><h4 style="..."><u>{{showcase1}}<h4>

</body>
</html>
Closing tag matches nothing
02:20 PM
13-11-2022

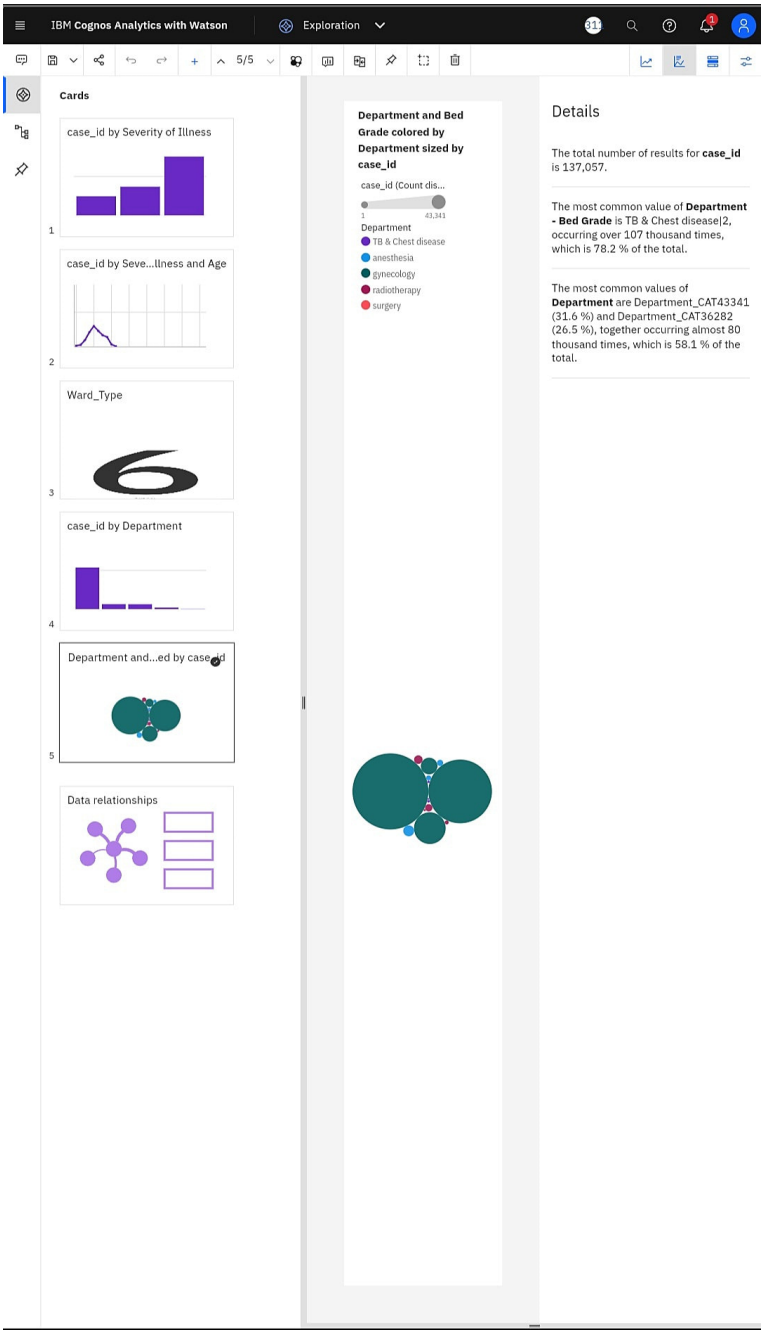


8. TESTING

8.1 TEST CASES


Navigation paths
sample_sub
test_data
case_id
Hospital_code
Hospital_type_code
City_Code_Hospital
Hospital_region_code
Available Extr...s in Hospital
Department
Ward_Type
Ward_Facility_Code
Bed Grade
patientid
City_Code_Patient
Type of Admission
Severity of Illness
Visitors with Patient
Age
Admission_Deposit
train_data
train_data_dictionary

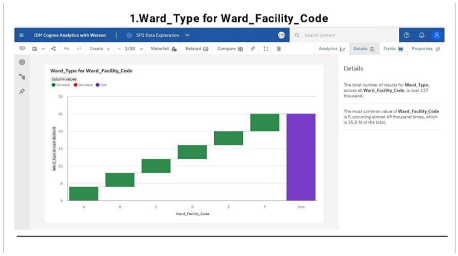
8.2 USER ACCEPTANCE TESTING



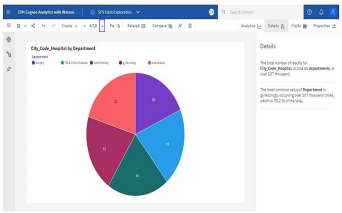
9.RESULTS

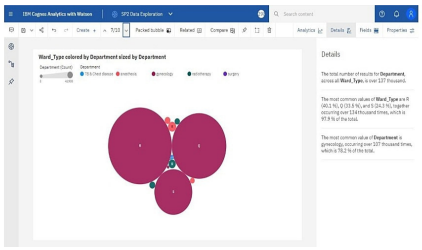
9.1 PERFORMANCE TESTING

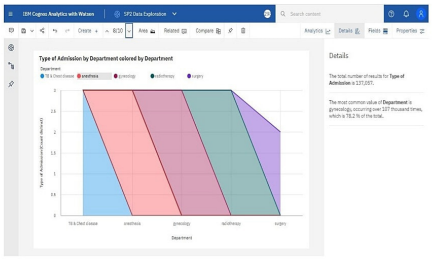
1.Ward_Type for Ward_Facility_Code



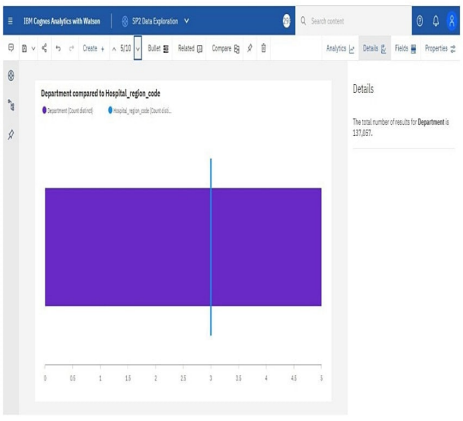
4. City_Code_Hospital by Department

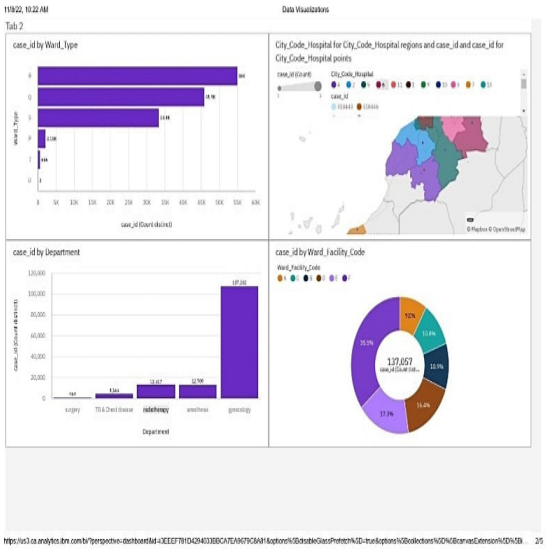




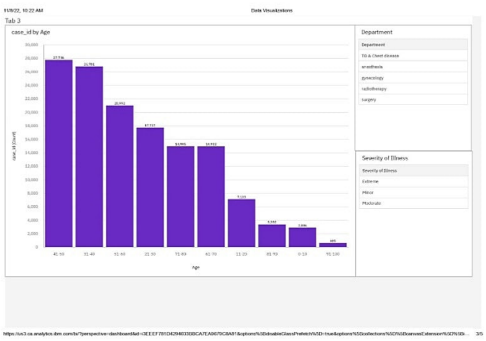


5. Department compared to Hospital_region_code

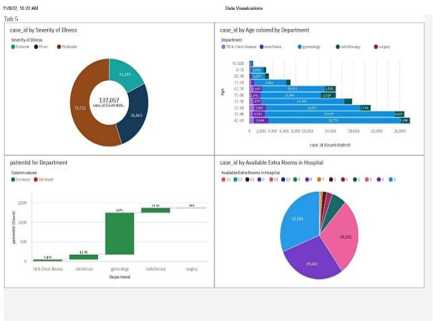


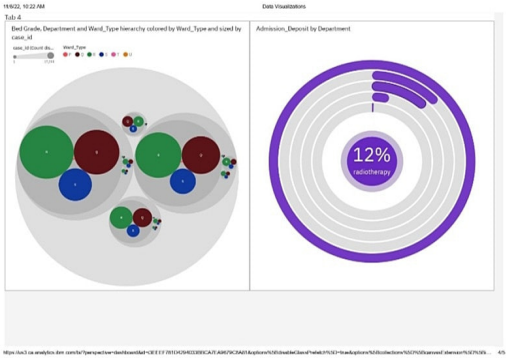


3.Age Wise Patients With Department And Severity Filters



5.Dashboard Showing Pie, Stacked Bar, Waterfall And Pie Charts





10. ADVANTAGES &DISADVANTAGES

10.1 ADVANTAGES

1. Improved research efforts
2. Improved health outcomes
3. Obtain operational insights
4. Improved staffing
5. Informed strategic planning

10.2 DISADVANTAGES

1. Lack of alignment within teams
2. Lack of commitment and patience
3. Low quality of data
4. Privacy concerns
5. Complexity & Bias

11.CONCLUSION

Data analytics has the potential to transform the way healthcare providers use sophisticated technologies to gain insight from their clinical and other data repositories and make informed decisions. In the future we'll see the rapid, widespread implementation and use of big data analytics across the healthcare organization and the healthcare industry.

12. FUTURE SCOPE

Data analysts can develop software to automatically inform patients about recommended lifestyle changes to prevent certain conditions. This helps improve performance by delivering data-based quality patient care which, in turn, improves patient satisfaction