Project Report

1.INTRODUCTION

1.1 PROJECT OVERVIEW

Recent Covid-19 pandemic has raised alarms over one of the most overlooked areas to focus:Healthcare Management,while healthcare management has various usecases for using data science,patient length of stay is one of the critical parameter to observe and predict if one wants to improve the efficiency of the healthcare management in a hospital

This parameter helps hospitals to identify patients of high LOS-risk (patients who will stay longer) at the time of admission. Once identified, patients with high LOS risk can have their treatment plan optimized to minimize LOS and lower the chance of staff/visitor infection. Also, prior knowledge of LOS can aid in logistics such as room and bed allocation planning.

Suppose you have been hired as Data Scientist of Health Man – a not for profit organization dedicated to manage the functioning of Hospitals in a professional and optimal manner

1.2 PURPOSE

The purpose 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.

2.LITERATURE SURVEY

2.1 EXISTING PROBLEM

Due to lack of effective data governance procedures,capturing data is one of the biggest obstacles for healthcare organizations .To use data more efficient,it must be clean,presize,correctly formatted do that it can be used across various healthcare systems.The challenges of data analytics are how the amount of data being collected,collecting meaningfull and realtime data,vishual representation of data,data from multiple sources.

2.2 REFERENCES

AUTHOR NAME: Viceconti
YEAR OF PUBLISHING: 2015:-

AUTHOR NAME: Ritu,Rajesh et al.
YEAR OF PUBLISHING:2017

AUTHOR NAME:V.S.Tseng
YEAR OF PUBLISHING:2017

AUTHOR NAME:Prop.Nagarathna Kulennavar,Priyanka.K.
YEAR OF PUBLISHING:2014.

AUTHOR NAME:Dr.S.Smys
YEAR OF PUBLISHING:2019

2.3 PROBLEM STATEMENT DEFINITION

1.DESCRIPTION:-
Big data in healthcare and medicine refers to these various large and complex data, which they are difficult to analyse and manage with traditional software or hardware . Big data analytics covers integration of heterogeneous data, data quality control, analysis, modeling, interpretation and validation . Application of big data analytics provides comprehensive knowledge discovering from the available huge amount of data. Big data analytics in medicine and healthcare is very promising process of integrating, exploring and analysing of large amount complex heterogeneous data with different nature: biomedical data, experimental data, electronic health records data and social media data. Integration of such diverse data makes big data analytics to intertwine several fields, such as bioinformatics, medical imaging, sensor informatics, medical informatics, health informatics and computational biomedicine. As a further work, the big data characteristics provide very appropriate basis to use promising software platforms for development of applications.

2. DESCRIPTION:-
A Robust model proposed by Ritu, Rajesh et al., should be enhanced as the model has encompassed big data. Moreover, it may compromise Data Privacy and Security and decreases the consistency and the processing of Big Data. The key advantage in a predictive data analytics includes the principal phase which is the disease recognition, and also includes evaluating and treating the diseases in efficient ways. However, to attain more effective outcomes from medical domain is still an open demand for the future work. The scattered system should be

organized to share the information between the laboratories, hospital systems, clinical centres and also with the other participants. For instance, biomedical devices which are either HL7 or DICOM compatible can be interfaced with the Laboratory Investigation System (LIS) data and the Hospital Information system. Furthermore, the data analytics shall be enhanced through machine learning techniques to make the data analytics effective. Security solutions should guarantee protection for analytics and Big Data Frameworks.

3. DESCRIPTION:-

A Robust model proposed by Ritu, Rajesh et al., should be enhanced as the model has encompassed big data. Moreover, it may compromise Data Privacy and Security and decreases the consistency and the processing of Big Data. The key advantage in a predictive data analytics includes the principal phase which is the disease recognition, and also includes evaluating and treating the diseases in efficient ways. However, to attain more effective outcomes from medical domain is still an open demand for the future work. The scattered system should be organized to share the information between the laboratories, hospital systems, clinical centres and also with the other participants. For instance, biomedical devices which are either HL7 or DICOM compatible can be interfaced with the Laboratory Investigation System (LIS) data and the Hospital Information system. Furthermore, the data analytics shall be enhanced through machine learning techniques to make the data analytics effective. Security solutions should guarantee protection for analytics and Big Data Frameworks.

4. DESCRIPTION:-

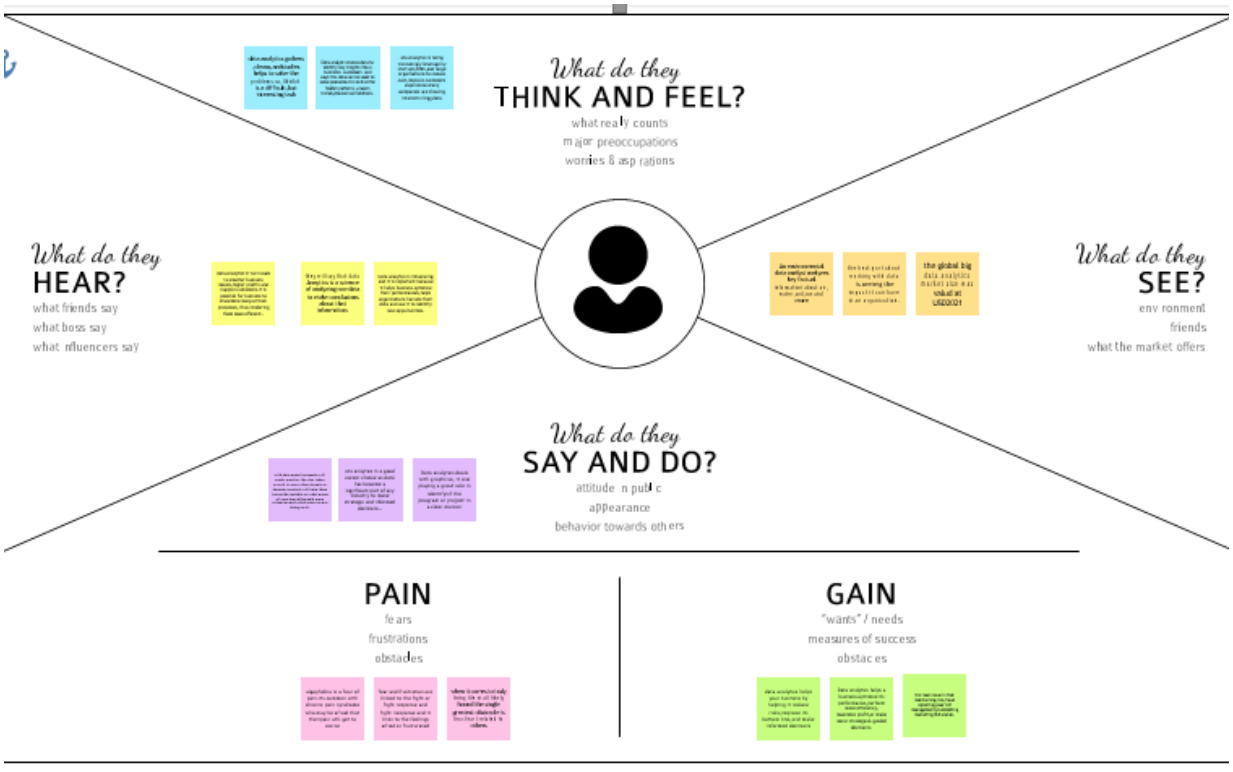
The rapid deployment of new emergency devices (i.e., wireless communications, mobile computing, and mobile devices) and patient monitoring systems has allowed for the focus to be on the design and delivery of digital health services that, leveraging real-time data, foster integrated and effective governance. It is essential to ensure a personalized health service, early disease diagnosis, and support for patient undergoing online care treatments . The gradual implementation of advanced digital solutions will support the clinical team’s decisions and release time for the most value- added clinical activities and treatment of the most complex cases. BD and AI not only have great potential in the fight against infectious diseases but can also be used for rapid drug
and vaccine development . Despite the important strides made in healthcare digitalization, there are numerous challenges to making the healthcare sector more resilient in the face of health crises. In this regard, it is necessary not only to strengthen the system but also to change its architecture toward a connected care model in which the organization, care, and assistance processes are redefined from a digital perspective and allow for making informed decisions using cutting-edge technology and BDA

5. DESCRIPTION:-

This paper gives a brief introduction about how we can uncover additional value from health information used in health care centers using a new information management approach called as big data analytics .Including big data analytics in health sector provides stakeholders with new insights that have the potential to advance personalized care, improve patient outcomes and avoid unnecessary costs. Analytics when applied in the context of big data is the process of examining large amounts of data, from a variety of data sources and in different formats, to deliver insights that can enable decisions in real or near real time. Various analytical concepts such as data mining, natural language processing, artificial intelligence and predictive analytics can be employed to analyze, contextualize and visualize the data.

3.IDEATION AND PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS

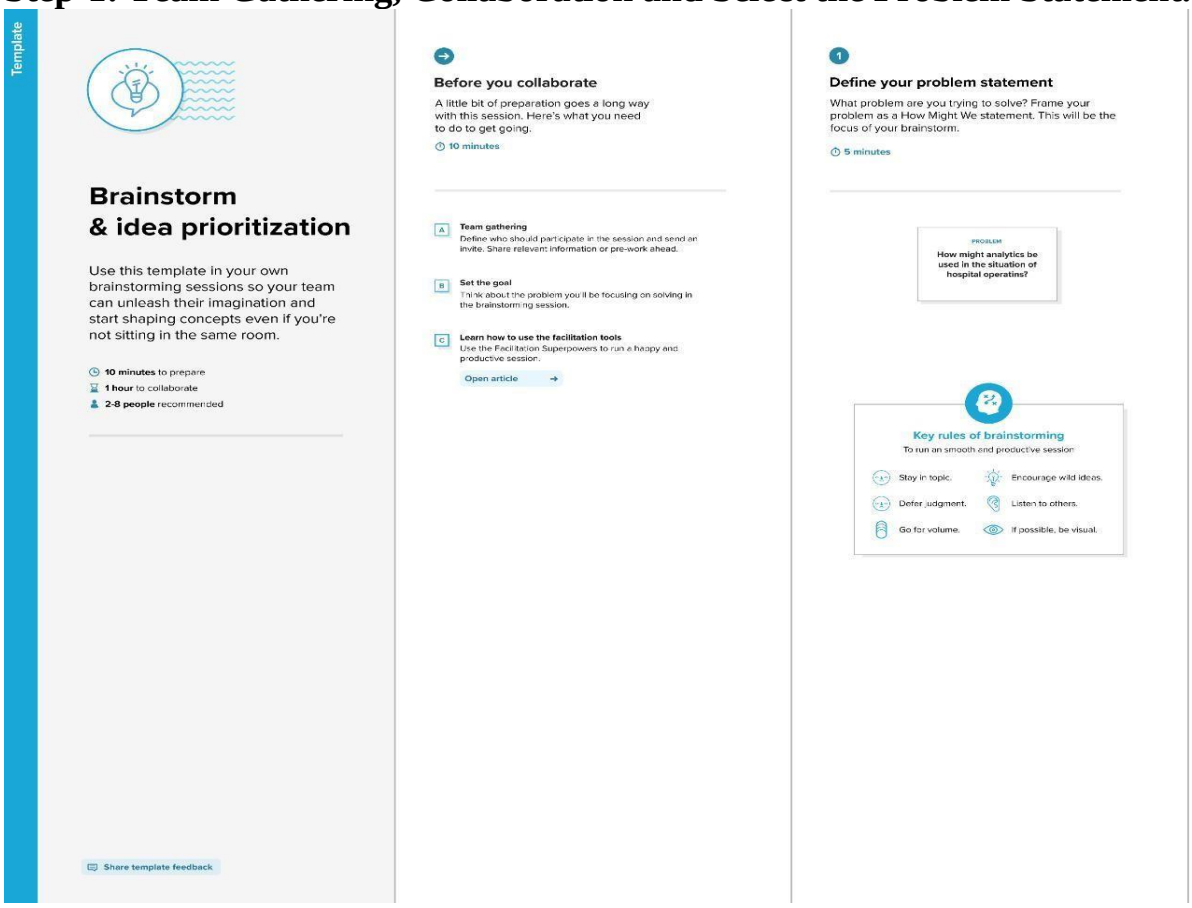


3.2 IDEATION AND BRAINSTROMING

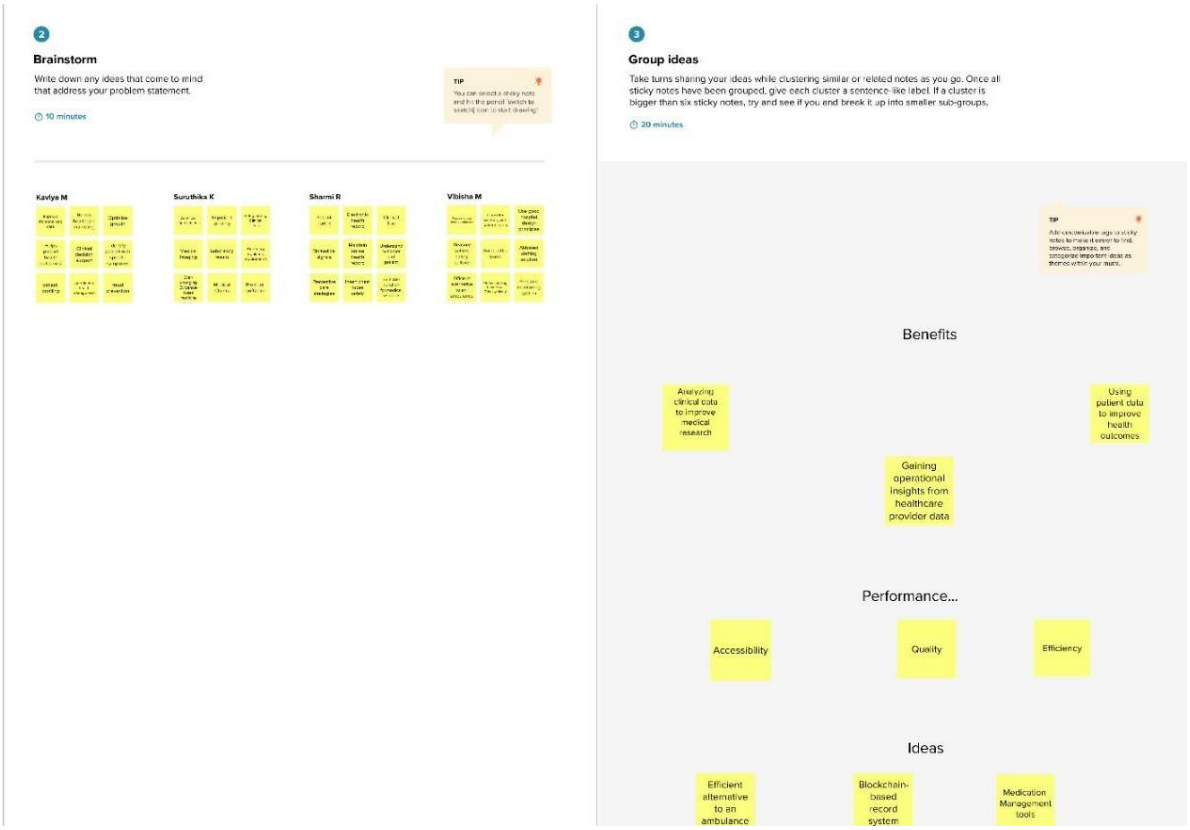
Brainstorm & Idea Prioritization :

Brainstorming provides a free and open environment that encourages everyone within a team to participate in the creative thinking process that leads to problem solving. Prioritizing volume over value, out-of-the-box ideas are welcome and built upon, and all participants are encouraged to collaborate, helping each other develop a rich amount of creative solutions.Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

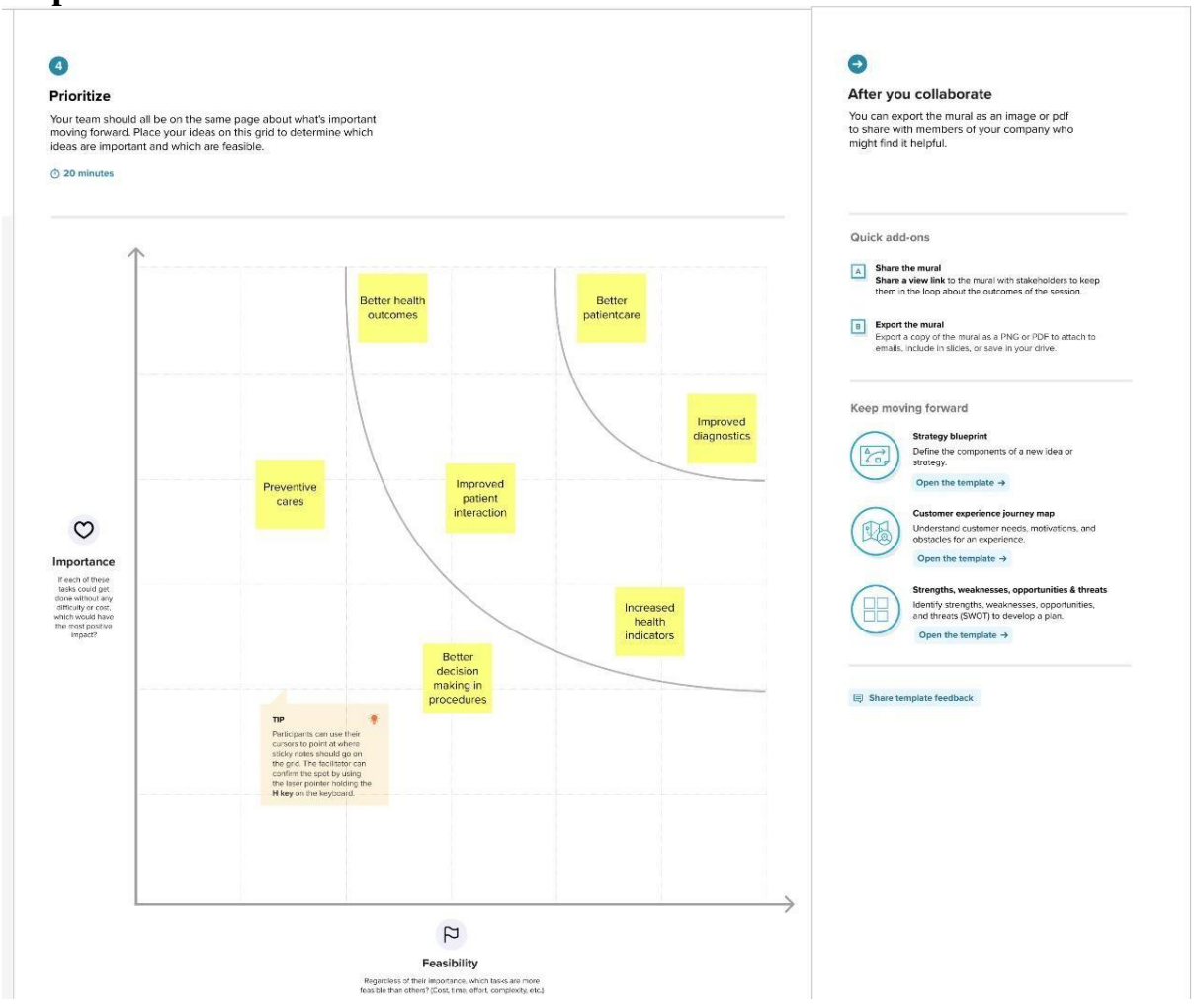
Step-1: Team Gathering, Collaboration and Select the Problem Statement.



Step-2: Brainstorm, Idea Listing and Grouping



Step-3: Idea Prioritization



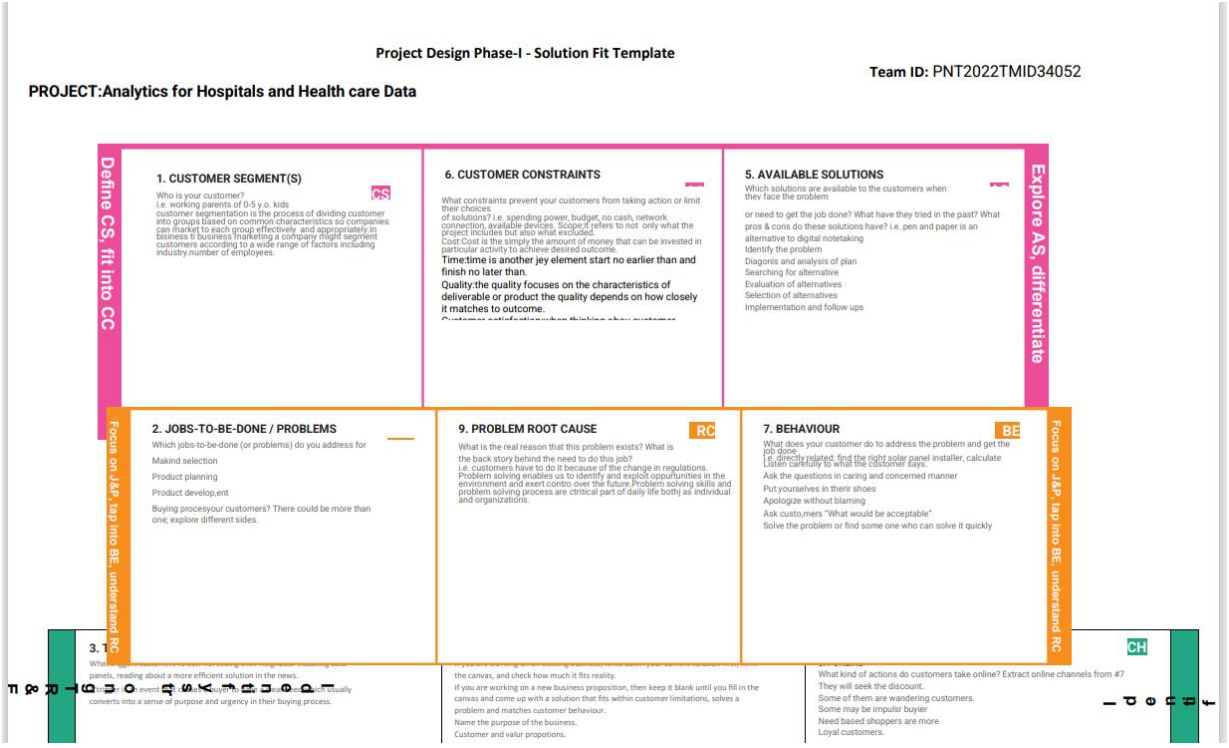
3.3 PROPOSED SOLUTION

Proposed Solution Template:

Project team shall fill the following information in proposed solution template.

Sl.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Recent covid-19 pandamic has raised alarms over one of the most overlooked areas to focus :Healthcare Management.While healthcare management has various use cases for using data science ,patient length of stay is one critical parameter to observe and predict if one wants to improve the efficiency of the health care management in the hospital. This parameter helps hospitals to identify patients of high -LOS risk ,once identified proper treatment can be given to them and logistics of bed allocation planning.
2.	Idea / Solution description	The short project forecasting surgery volume at a medical the portfolio project analysed twitter trends on covid-19 vassinations.some of them are the predicative analyst poroject,the out-of-the-box project.this is the purpose of healthcare data analysis using data-driven findings to predict and solve the problem before it is too late.
3.	Novelty / Uniqueness	Potential sources of information about health care numerous and diverse,but in practice four main sources are used medical records,certificates of the other health-related events responses in surveys and facts obtained in the course of conducting resources/
4.	Social Impact / Customer Satisfaction	Potential satisfaction is measured with the help of an HCAHPS survey(also known asCAHPS(customer assessment of healthcare providers and system) Hospital survey,which refers to a set of survey that collect patients data to measure patients experience about hospital care and services
5.	Business Model (Revenue Model)	A business model describes the resources ,processes and cost assumption that an organization makes that will lead to the delivery of a unique value propotion to the customer.
6.	Scalability of the Solution	Scalability is the ability of the healthcare intervision shown to be efficient on a small scale and/or under controlled condition to be expanded under real world conditions to reach a greater propotion of the eligible population while retaining effectiveness

3.4 PROBLEM SOLUTION FIT



4.REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENTS

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIN
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	Undertaking various Registration	HMS is able to facilitate various Registration to enter the details of patients.
FR-4	Visualizing Data	User can visualize Departments,ward types,bed availability,city id ,remaining rooms availablethrough Dashboard created using IBM cognos Analytics.
FR-5	Check Out	The HMS helps facilities in ensuring all formalities and commitments using unique ID.
FR-6	Generating report	User can view his/her health report and can make Decisions accordingly.

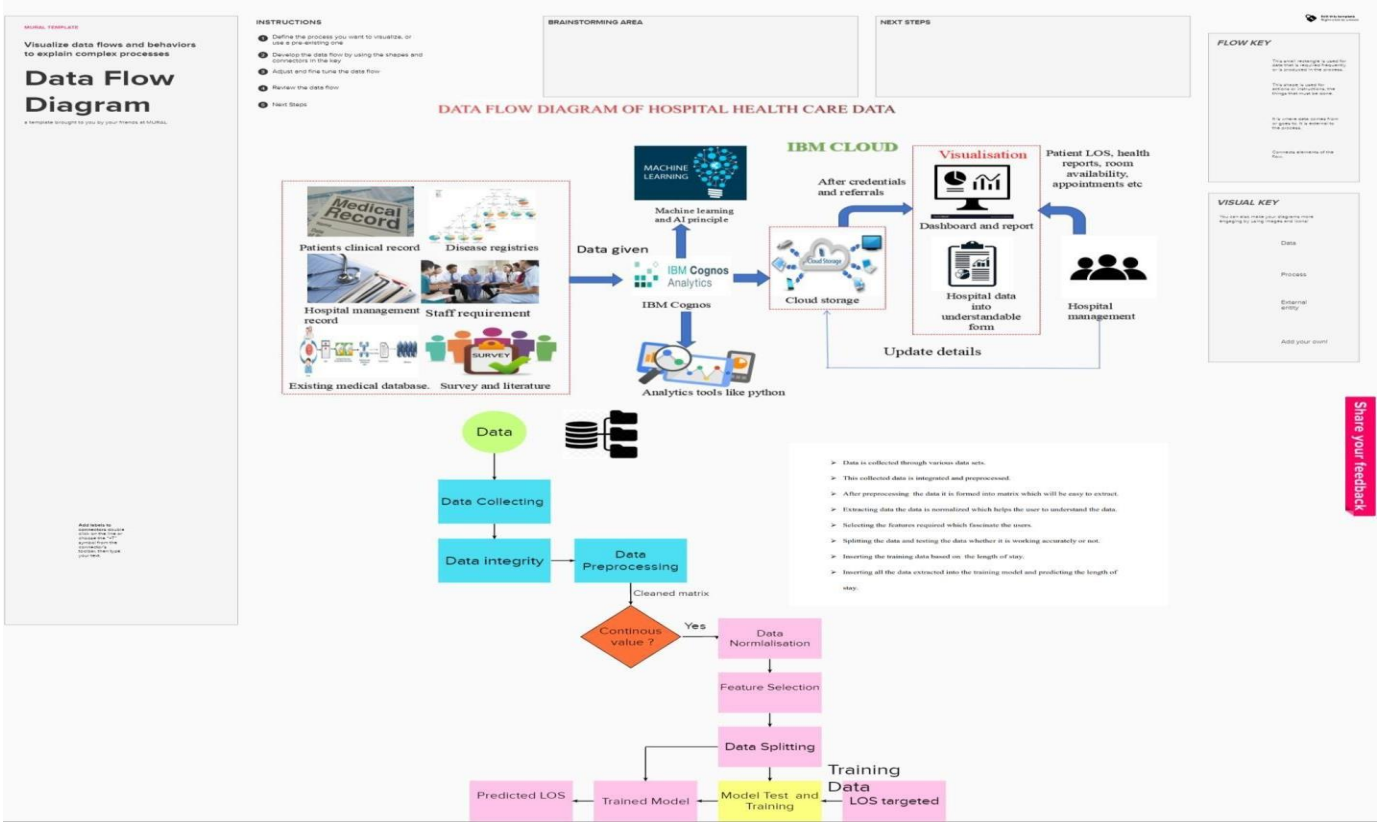
4.2 NON FUNCTIONAL REQUIREMENTS

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

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The application will have a simple and user friendly graphical interface.User will able to understand and use all the features of the application easily.Any action has to be performed with just a few clicks.
NFR-2	**Security**	For security of the application the technique known as database replication should be used so that all the important data should be Kept safe.Incase of crash the system should be able to backup and recover the data.
NFR-3	**Reliability**	The application has to be consistent at every scenario and has to work without failure in any environment.
NFR-4	**Performance**	Performance of the application depends on the response time and the speed of the data submission .The response time of the application is direct and faster which depends on the efficiency of implemented algorithm.
NFR-5	**Availability**	The application has to be available 24x7 for users without any interruption.
NFR-6	**Scalability**	The application can withstand the increase in the number of users and has to be able to develop Higher versions.

5.PROJECT DESIGN

5.1 DATA FLOW DIAGRAMS



5.2 SOLUTION AND TECHNICAL ARCHITECTURE

S.No	Component	Description	Technology
	User Interface	How user interacts with application e.g. Web UI, Mobile App, Chatbot etc.	HTML, CSS, JavaScript / Angular Js
	Application Logic-1	Logic for a process in the application	Java / Python
	Application Logic-2	Logic for a process in the application	IBM Watson STT service
	Application Logic-3	Logic for a process in the application	IBM Watson Assistant
	Database	Data Type, Configurations etc.	MySQL, NoSQL, etc.
	Cloud Database	Database Service on Cloud	,IBM Cloudant etc.
	File Storage	File storage requirements	IBM Block Storage or Other Storage Service or Local Filesystem
	External API-1	Purpose of External API used in the application	IBM Weather API, etc.
	External API-2	Purpose of External API used in the application	Aadhar API, etc.
	Machine Learning Model	Purpose of Machine Learning Model	Object Recognition Model, etc.
	Infrastructure (Server / Cloud)	Application Deployment on Local System / Cloud Local Server Configuration: Cloud Server Configuration :	Kubernetes,

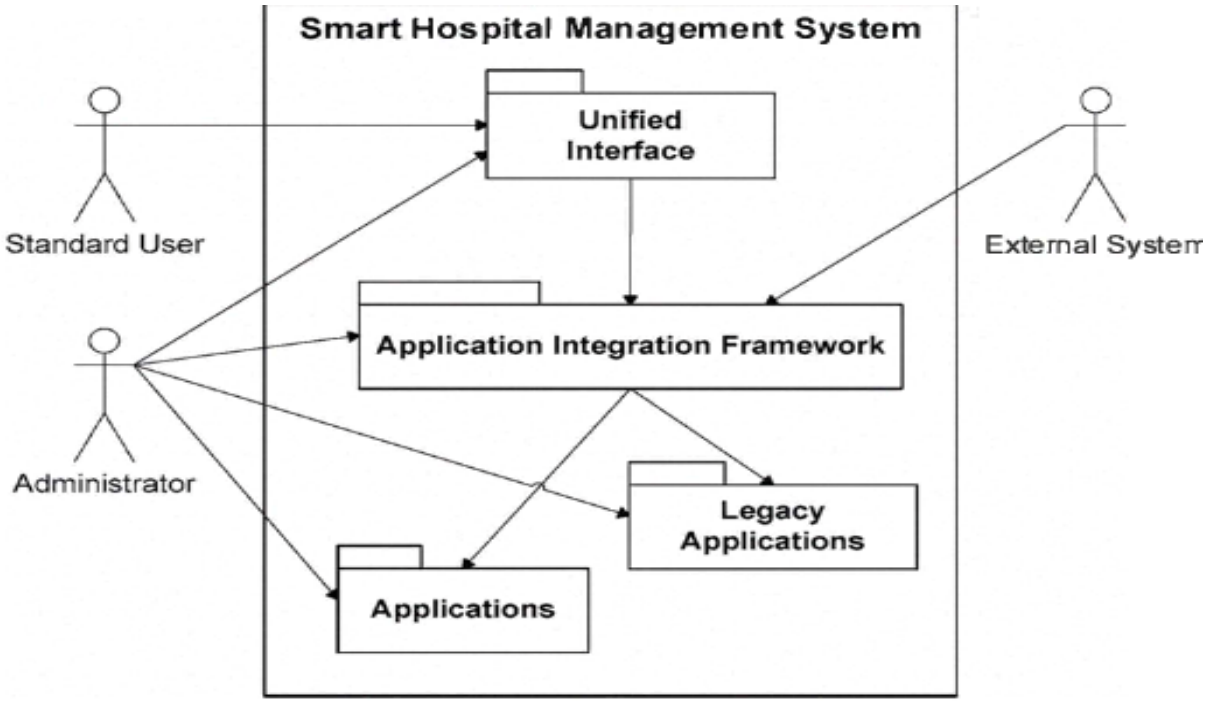


Table-2: Application Characteristics:

S.No	Characteristics	Description	Technology
	Open-Source Frameworks	List the open-source frameworks used	Technology of Opensource framework
	Security Implementations	List all the security / access controls implemented, use of firewalls etc.	e.g. SHA Encryptions
	Scalable Architecture	Justify the scalability of architecture (3 – tier, Micro-services)	Technology used
	Availability	Justify the availability of application (e.g. use of load balancers, distributed servers etc.)	Technology used
	Performance	Design consideration for the performance of the application (e.g use of load balancers,distributed servers etc)	Technology used

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)	HIV/AIDS Risk Smart Form for Data Entry	USN-1	As a Clinician,I need to review ,consoldate and update HIV/AIDS Risk Form. So that I can Determine the patient's risk of HIV/AID(risk caregory),and ensure proper remedy accordingly.	I can access patient record or data accurately	High	Sprint-1
	BPA to prompt ordering HIV/AIDS on Admission	USN-2	As an Inpatient, I want to be prompted to order HIV/AIDS on admission.so that I remember to place my patient on AIDS	Maintain the record for correct preference.	High	Sprint-1
	HIV/AIDS dynamic order group in Admit order sets	USN-3	As an inpatient ,I want to view only risk-appropriate HIV/AID options in Admission order sets So that I can ensure my patient is getting optimal HIV/AID prophylaxis.		Low	Sprint-2
		USN-4	As a user, I can access the data in visualise mode.		Medium	Sprint-1
	Dashboard	USN-5	As a user, I can access the data from the queries,graph ,pie char		High	Sprint-1
Customer (Web user)			Get the older information from the hospitals			
Customer Care Executive			Have data in graph modes		Medium	
Administrator			Access in cloud easily		Medium	

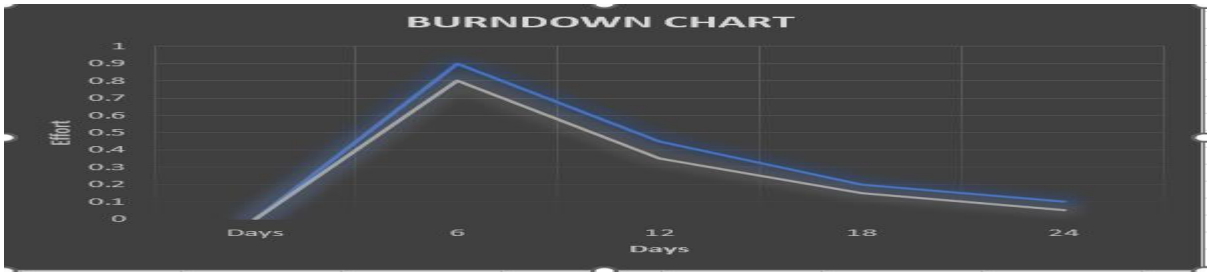
6.PROJECT PLANNING AND SHEDULING

6.1 PROJECT 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.	8	High	M.kaviya R.Sharmi
		USN-2	As a user ,I will receive conformation email once I have registered for the application.	8		M.Vibisha
	Login	USN-3	As a user,I can log into the application by entering email, password.	4	High	K.Suruthika R.Sharmi
Sprint-2	Working with Dataset	USN-4	To work with Dataset,Understood and load the dataset.	10	Low	R,Sharmi M.Vibisha
		USN-5	Exploration of current health condition including patients byward types,Departments,city,bed etc.	5		K.Suruthika M.Kaviya
Sprint-3	Data Visualization	USN-6	Visualization of average age for issue in health condition types ,and exercise for that.	5	Medium	R.Sharmi K.Suruthika
		USN-7	Exploration of the type of Disease.	7		
		USN-8	Exercise to get away from Disease.	6		
Sprint-4	Dashboard Creation	USN-9	Dashboard showing different types of Visualization.	20	High	R.Sharmi K.Suruthika M.Kaviya M.Vibisha

6.2 SPRINT DELIVERY SCHEDULE

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 N0v 2022



6.3 REPORTS FROM JIRA

7.CODING AND SOLUTIONING

7.1 FEATURE 1

Login using HTML

Features of HTML:

HTML is the most common used language to write web pages. It has recently gained popularity due to its advantages such as, It is the language which can be easily understood and can be modified.Effective presentations can be made with the HTML with the help of its all formatting tags.It provides the more flexible way to deign web pages along with the text.. Links can also be added to the web pages so it helps the readers to browse the information of their interest.You can display HTML documents on any platforms such as Macintosh, Windows and Linux etc. Graphics, videos and sounds can also be added to the web pages which give an extra attractive look to your web pages.

PYTHON

Flask comes with built-in development server as well as fast debugger

It also contains the integrated support required for unit testing

It has the feature of restful request dispatching

Comes with Jinja2 templating technique

Flask supports secure cookies i.e. client-side sessions

Also has the WSGI 1.0 compliant feature.

It is based on Unicode.

Python Flask is extensively documententation .

ANACONDA

It is free and open-source.

It has more than 1500 Python/R data science packages.

Anaconda simplifies package management and deployment.

It has tools to easily collect data from sources using machine learning and AI.

JUPYTER

Data visualizations:

Most people have their first exposure to Jupyter Notebook by way of a data visualization, a shared notebook that includes a rendering of some data set as a graphic. Jupyter Notebook lets you author visualizations, but also share them and allow interactive changes to the shared code and data set.

Code sharing. Cloud services like GitHub and Pastebin provide ways to share code, but they're largely non-interactive. With a Jupyter Notebook, you can view code, execute it, and display the results directly in your web