SFU
BIG DATA

Anomaly Detection and Surveillance System for Diagnosis in ICU Patients

Project Milestone Presentation

Team XYZ:

Bilal Hussain

Kyoun Huh

Hon Wing Eric Chan

Sakina Patanwala

CONTENTS

SFU BIG DATA

- Motivation
- Action Plan
- Progress
- Risk Mitigation

MOTIVATION: Importance

SFU BIG DATA

1. Predict the mortality rate and possible duration of a patient's stay for each disease.

2. Predict the efficiency of various antibiotics against distinctive pathogens.

3. Use anomaly detection, to know the patient's recent vital signs at are anomalous and send a warning to hospital staff if there is an emergency.

MOTIVATION: Challenges

SFU BIG DATA

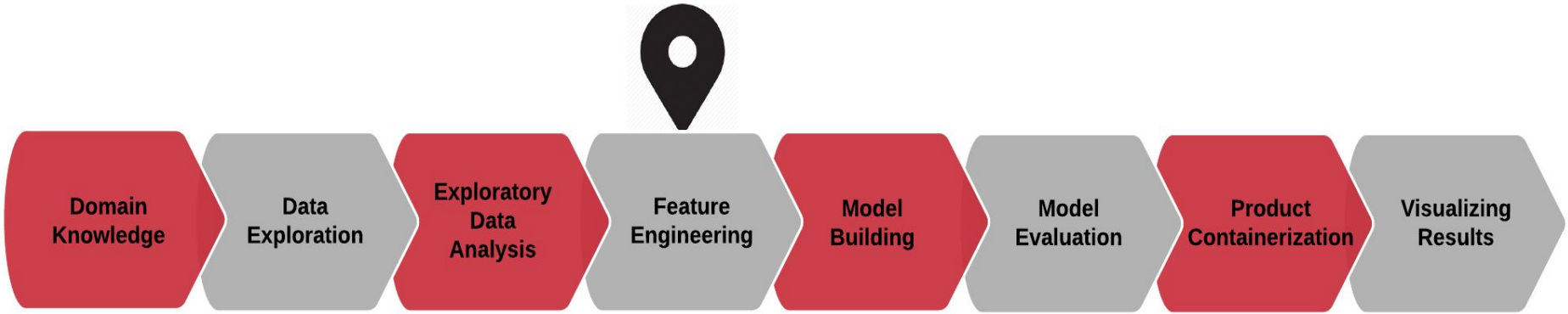
Incorporating warning tool functionalities because dataset provides dummy dates for patient details

Over 40 GBs of data with 534 features across 40 tables

Restrictions in determining predictive models for medical care paths as they depend on numerous patient features which are missing.

ACTION PLAN

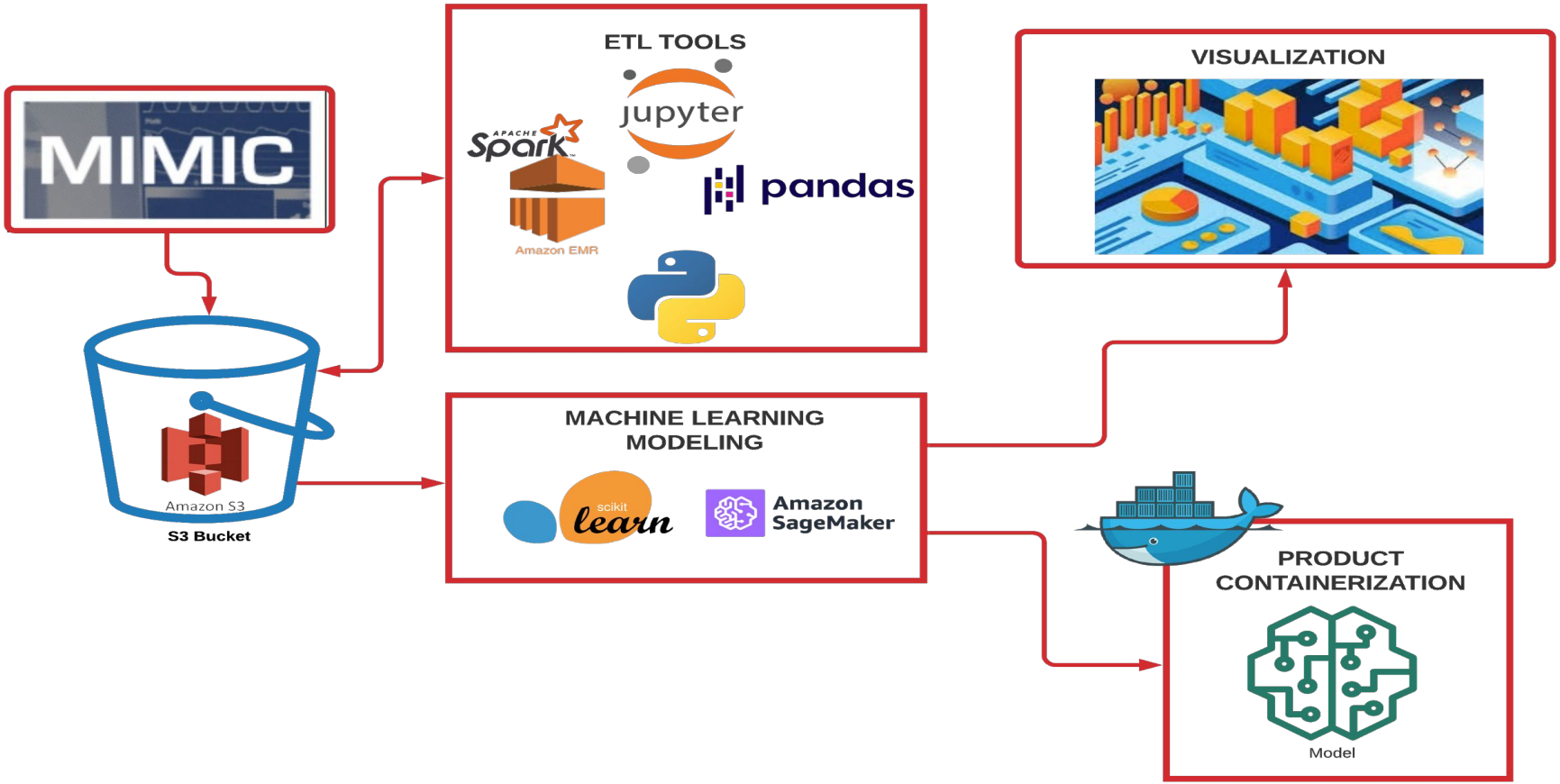
SFU BIG DATA



Note: Currently awaiting access to the full MIMIC-III dataset from PhysioNet. However, we are currently working using the demo dataset and it is enough to continue with our project work.

ACTION PLAN: Architecture

SFU BIG DATA

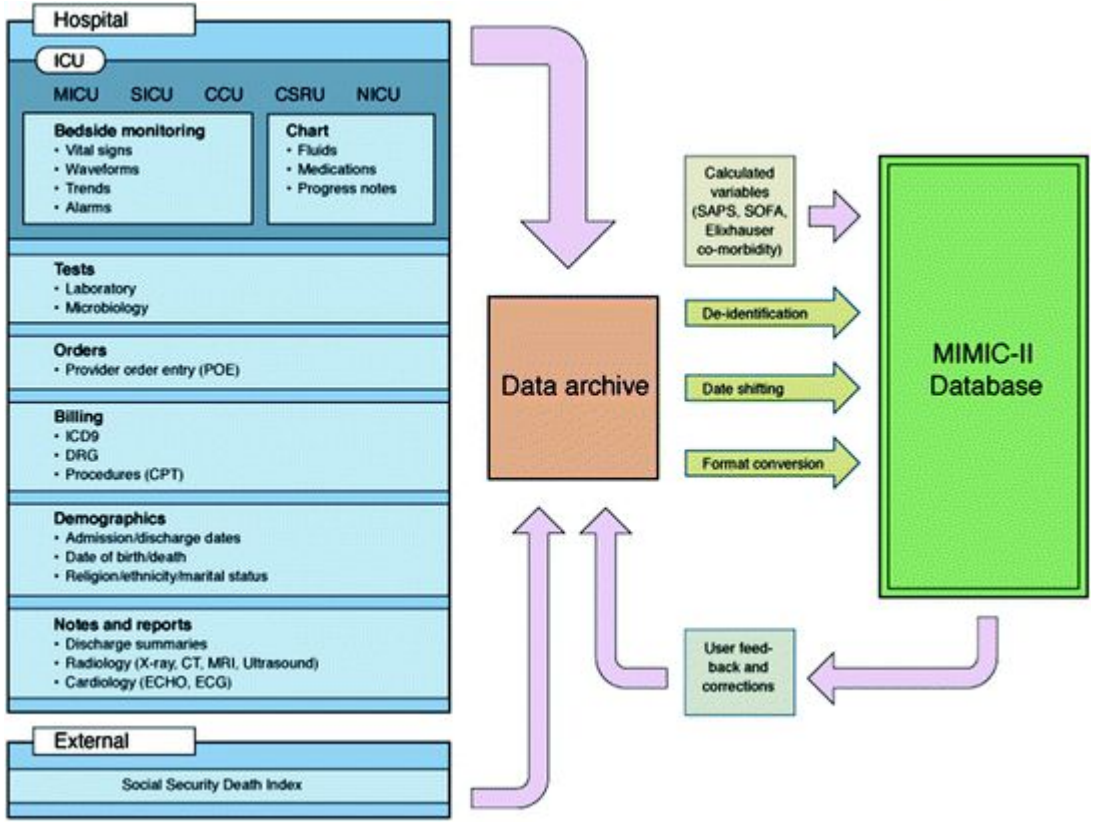


ACTION PLAN: Dataset

SFU BIG DATA

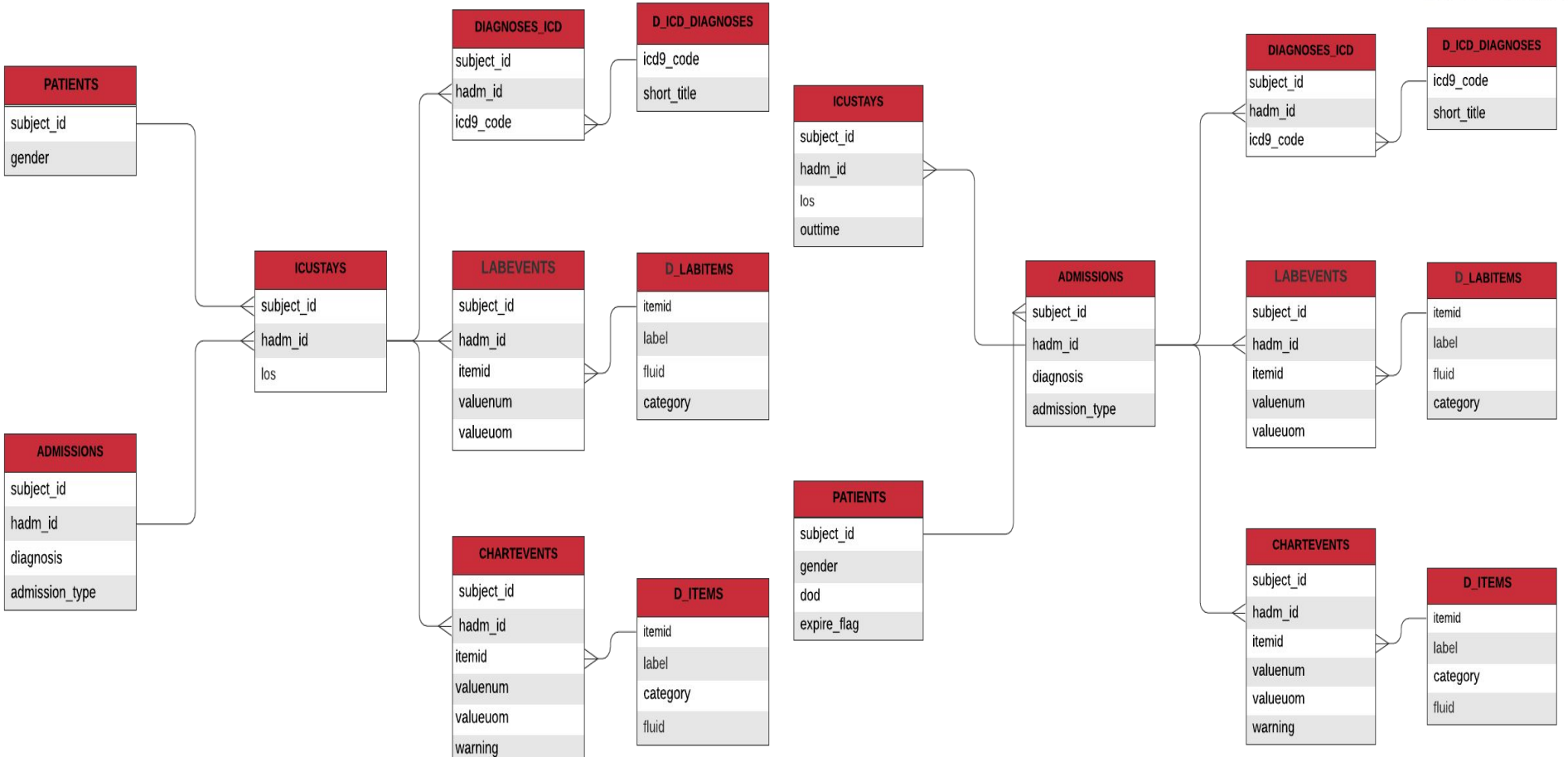
- MIMIC-III is a publicly available database consists of de-identified ICU patient data from Beth Israel Deaconess Medical Center in Boston, Massachusetts from 2001 to 2012.

- Data includes information of patients' demographics, laboratory tests, bedside monitoring, medications, microbiology procedures, notes, reports by caregivers, etc.



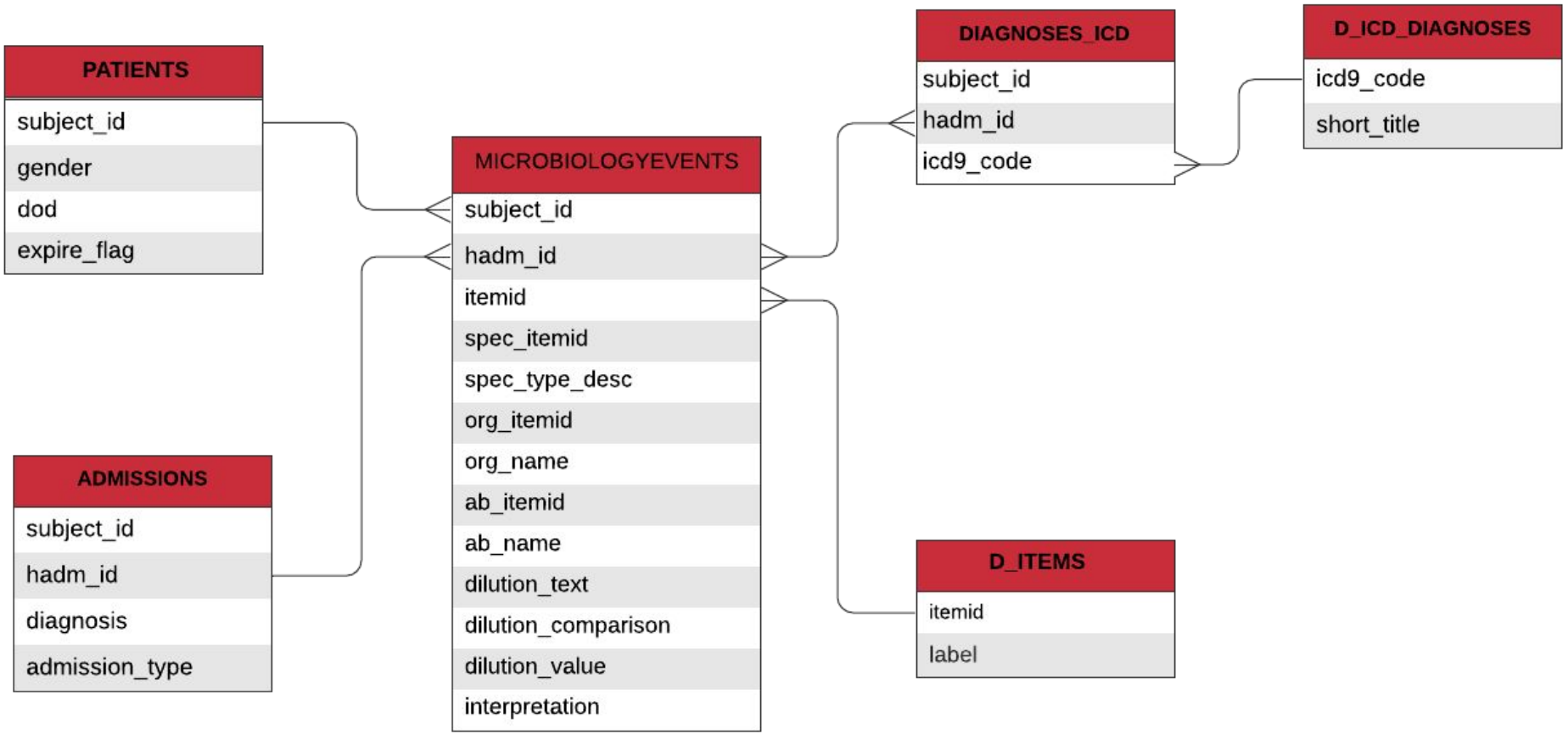
ACTION PLAN: Schemas for Objective 1&3

SFU BIG DATA



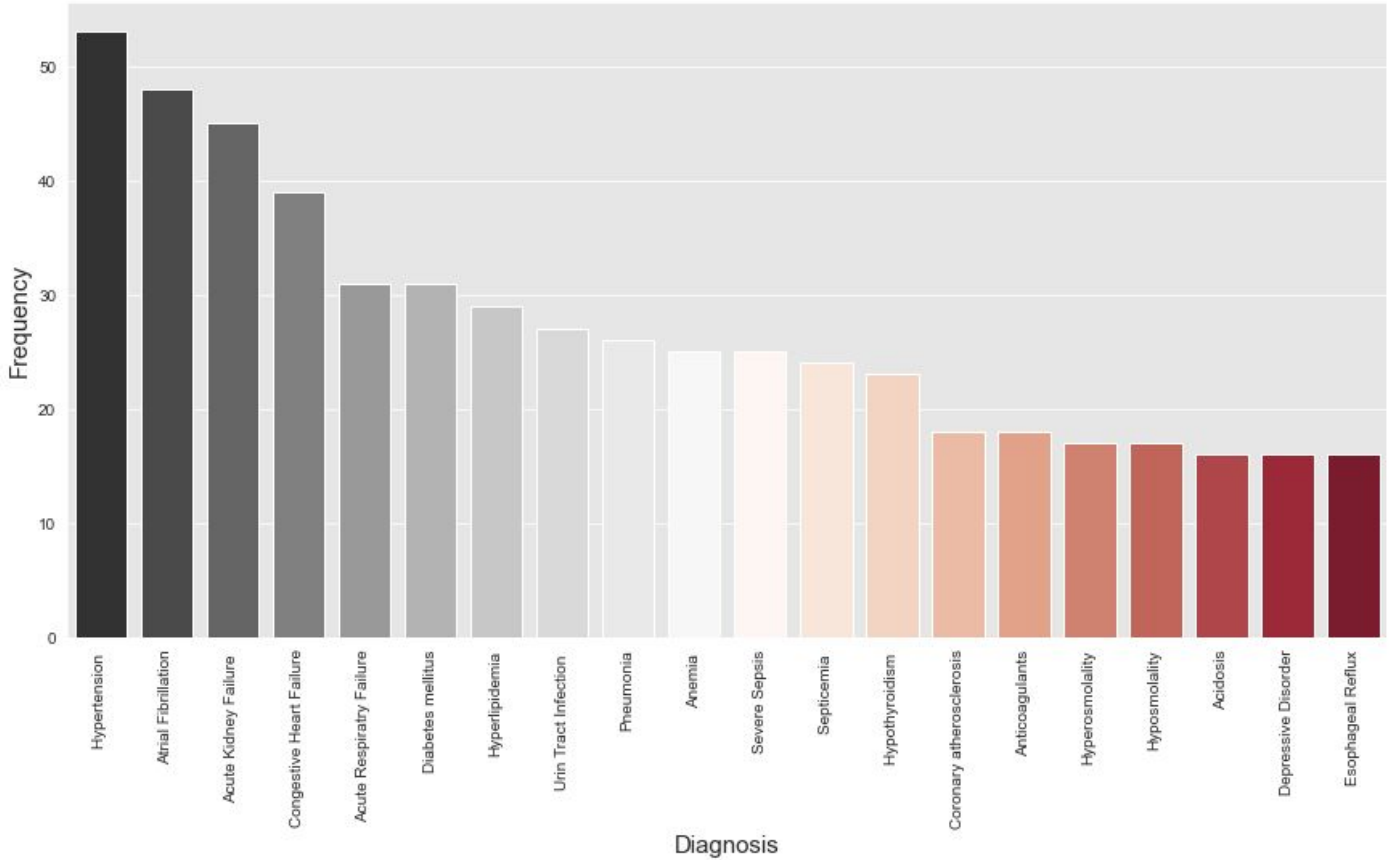
ACTION PLAN: Schema for Objective 2

SFU BIG DATA



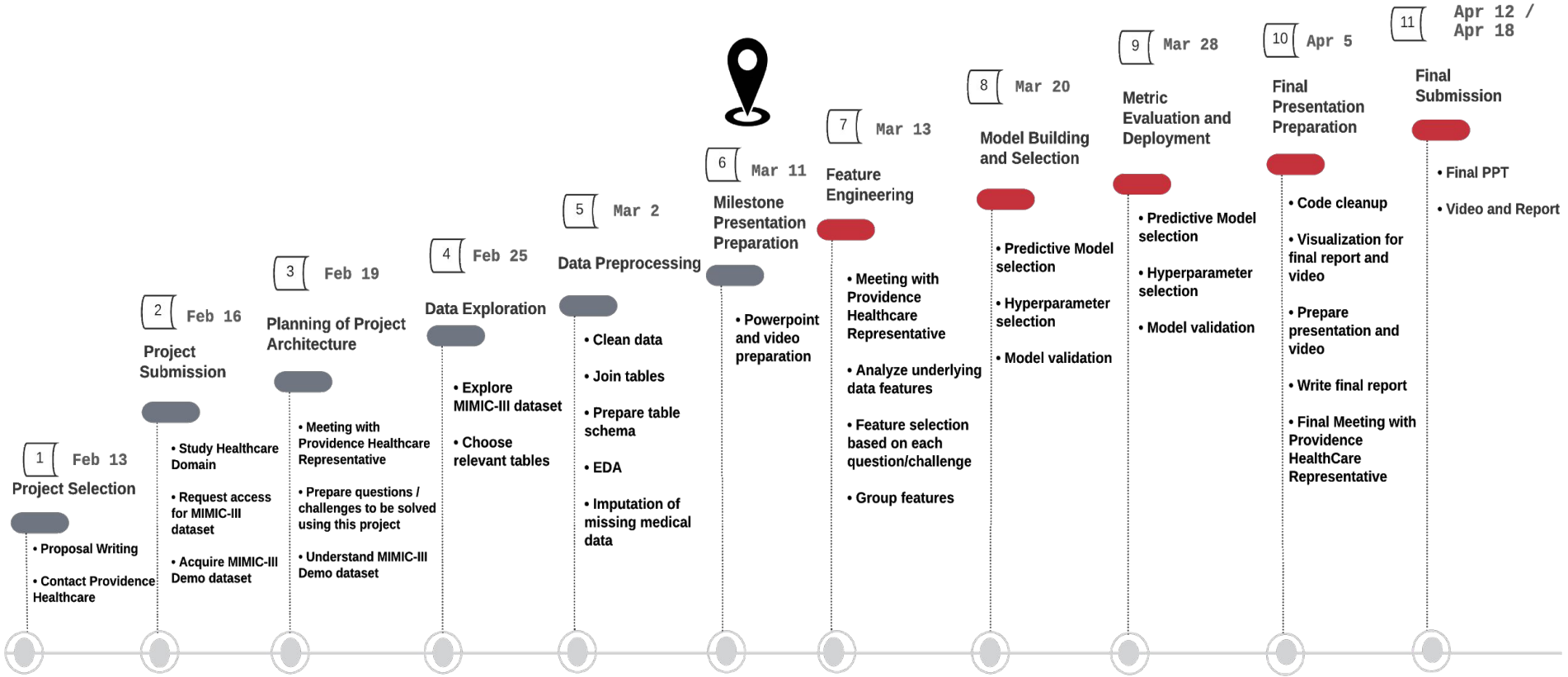
ACTION PLAN: Diagnosis Frequency Distribution

SFU BIG DATA



PROGRESS: w.r.t Project Architecture

SFU BIG DATA



PROGRESS: Tracked using Kanban Board

SFU BIG DATA

CMPT733-Kanban

Updated 11 hours ago

9 To do

- Selecting predive model and hyperparameter tuning
  #13 opened by qyune
- Model Validation and Metric Evaluation
  #21 opened by qyune
- Containerize the final outcome (on Docker).
  #18 opened by qyune
- Visualization of obtained results
  #19 opened by qyune
- Code Cleanup
  #20 opened by qyune
- Final Submission: Visualizations for reports
  #14 opened by qyune
- Final Submission: Video
  #15 opened by qyune
- Final Submission: Report
  #16 opened by qyune
- Final Submission: PPT
  #17 opened by qyune

2 In progress

- Request for credentials on physionet
  #11 opened by qyune
- Feature selection and grouping of features
  #12 opened by qyune

10 Done

- Contact representative from Providence HealthCare
  #3 opened by qyune
- HealthCare Domain Knowledge
  #2 opened by qyune
- Explore MIMIC-III dataset
  #1 opened by qyune
- Prepare questions/ challenges to be solved using the MIMIC-III dataset
  #4 opened by qyune
- Choose relevant tables from MIMIC-III dataset
  #5 opened by qyune
- Removal of rows that are missing data
  #6 opened by qyune
- Joining relevant tables and preparing schema
  #7 opened by qyune
- Milestone Submission: PPT
  #8 opened by qyune
- Milestone Submission: Visuals for Submission
  #9 opened by qyune
- Milestone Submission: Video no longer than 5 mins but less than 4 mins
  #10 opened by qyune

RISK MITIGATION

SFU BIG DATA

Unsuccessful in building a predictive model for a particular objective	Have a backup project objective/challenge to conquer
Unable to process complete MIMIC-III dataset (size: 40GBs)	Use smaller samples of the MIMIC-III dataset
Predictive model not standing true in real world	Minimize the number of features to reduce error and consult healthcare professionals
Time conflict with other classes/coursework	Better time management and communication

SFU
BIG DATA

Thank You!