Low Level Design (LLD)

Your Project Name

Your Name

Partner Name

Document Version Control

Date Issued	Version	Description	Author
Date	1.2	Added Workflow chart	Name
Date	1.3	Added Exception Scenarios Overall, Constraints	Name
Date	1.1	First Draft	Name
Date	1.4	Added KPIs	Name
Date	1.5	Added user I/O flowchart	Name
Date	1.6	Added EHR, LSTM model diagrams	Name
Date	1.7	Added dataset overview and updated user I/O flowchart.	Name
Date	1.8	Restructure and reformat LLD	Name

Contents

Document Version Control 2

Abstract 4

1 4

1.1 4

1.2 5

1.3 5

1.4 6

1.5 6

2 6

2.1 Predicting Disease 7

2.2 Logging 7

2.3 Database 7

3 9

4 9

5 10

6 11

7 11

Abstract

With high increase in vehicles on the road, problem of traffic congestion and accidents has increased substantially. To overcome these problems, the in-depth analysis of causes such as number of traffic rules followed, is required. Therefore, continuous monitoring of traffic on highways and huge roads is mandatory. An Automatic Traffic Control System can prove to be a solution to above mentioned problems.

Automatic Traffic Counter Control can also help in drawing inferences from the recorded data

1 Introduction

1.1 Why this Low-Level Design Document?

The purpose of this document is to present a detailed description of the Deep EHR System. It will explain the purpose and features of the system, the interfaces of the system, what the system will do, the constraints under which it must operate and how the system will react to external stimuli. This document is intended for both the stakeholders and the developers of the system and will be proposed to the higher management for its approval.

The main objective of the project is to predict if a person can get a chronic disease in his/her future based on the EHR. EHR stands for Electronic Health Record, EHR is nothing but a dataset of medical history of the patients.

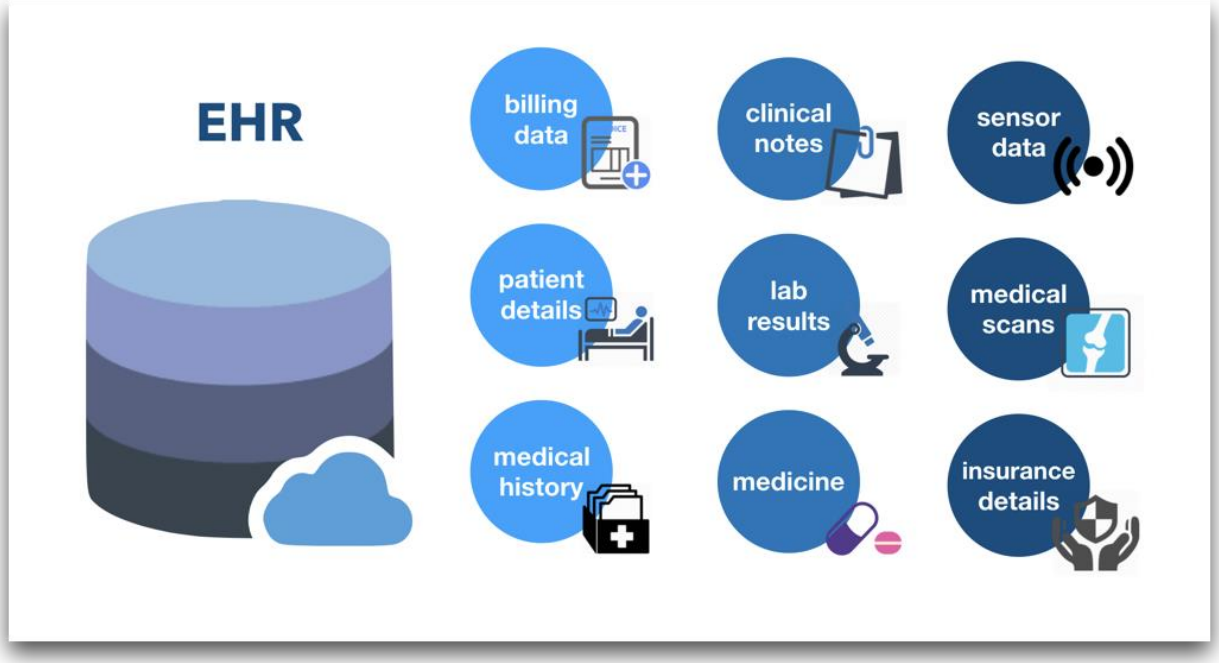
EHRs are a vital part of health IT and can:

- Contain a patient's medical history, diagnoses, medications, treatment plans, immunization dates, allergies, radiology images, and laboratory and test results
- Allow access to evidence-based tools that providers can use to make decisions about a patient's care
- Automate and streamline provider workflow

An electronic health record (EHR) contains patient health information, such as:

- Patient demographics
- Progress notes
- Vital signs
- Medical histories
- Diagnoses
- Medications

- Immunization dates
- Allergies
- Radiology images
- Lab and test results



This project shall be delivered in two phases:

Phase 1: All the functionalities with PyPi packages.

Phase2: Integration of UI to all the functionalities.

1.2 Scope

This software system will be a Web application This system will be designed to detect the diseases at earliest for better disease management, improved interventions, and more efficient health-care resource allocation using previous EHR records available. More specifically, Early detection of any preventable diseases is important for better disease management. This system is designed to predict the diseases from patient information such as demographics, disease history, lab results, procedures and medications.

1.3 Constraints

We will only be selecting a few of the chronic diseases.

1.4 Risks

Document specific risks that have been identified or that should be considered.

1.5 Out of Scope

Delineate specific activities, capabilities, and items that are out of scope for the project.

2 Technical specifications

2.1 Dataset

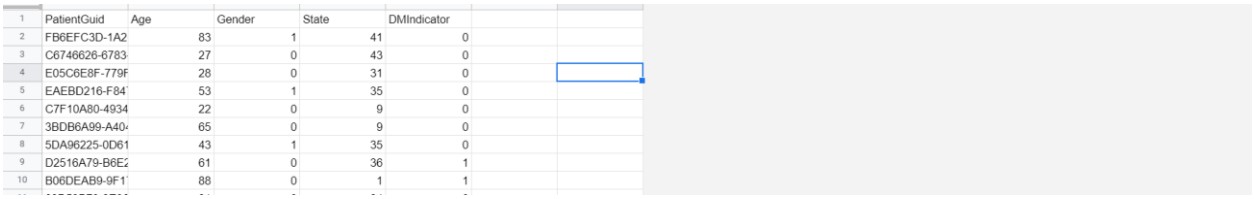
Disease	Finalized	Source
Diabetes	yes	https://github.com/kthouz/Diabetes-PracticeFusion/tree/master/agg_data
Stroke	Yes	
heart disease	Yes	
Cancer	Yes	

2.1.1 Diabetes dataset overview

Consists of 4 different tables, Patient table consists of the patient's personal information and most importantly we have the historic data of a patient in the table diagnosis. Whereas the transcript table consists of patient demographic data. Physician_speciality table consists of patient behaviour.

There are a total of 9,948 patients in the training set and 4,979 patients in the test set.

- Patient table



	PatientGuid	Age	Gender	State	DMIndicator
2	FB6EFC3D-1A2	83	1	41	0
3	C6746626-6783	27	0	43	0
4	E05C6E8F-779F	28	0	31	0
5	EAEBD216-F84	53	1	35	0
6	C7F10A80-4934	22	0	9	0
7	3BDB6A99-A40	65	0	9	0
8	5DA96225-0D61	43	1	35	0
9	D2516A79-B6E2	61	0	36	1
10	B06DEAB9-9F1	88	0	1	1

- Diagnosis table

PatientGuid	Icd9_001-139	Icd9_140-239	Icd9_240-279	Icd9_280-289	Icd9_290-319	Icd9_320-359	Icd9_360-389	Icd9_390-459	Icd9_460-519	Icd9_520-579	Icd9_580-629	Icd9_630-679	Icd9_680-709
00023761-9D8D	0	0	1	0	0	0	0	0	0	0	1	0	
0005D9BD-0247	0	0	1	0	0	0	0	0	0	0	0	0	
000B4862-7CE7	0	0	0	0	0	0	1	3	1	0	0	0	
00110ABC-DAB	0	0	1	0	0	0	0	2	0	0	0	0	
002667F4-B9A3	0	0	4	0	0	0	0	2	1	0	0	0	
0029BBC8-7C22	2	0	2	0	2	2	0	0	0	1	1	0	
003CEE1F-5BF	0	0	2	0	0	0	0	2	1	0	0	0	
004382BD-E31F	0	0	0	0	0	1	0	1	0	1	0	0	
004AACFF-5D6	0	0	0	0	1	0	0	0	2	1	0	0	
004CC712-BAD	0	0	0	0	0	0	0	0	3	0	0	0	
004EA6A7-6159	0	0	0	0	0	0	1	0	1	0	0	0	
005F61C9-E537	0	0	1	0	1	0	1	0	5	3	2	0	
00607D4B-5F8F	1	0	1	0	0	0	0	0	2	0	0	4	
0063B34F-C2C4	0	0	2	0	1	0	1	1	4	3	0	0	
0066553F-732C	0	0	1	0	0	1	0	2	0	0	0	0	
006948F2-1118-	2	0	0	0	0	0	0	0	0	0	0	0	
006A9198-1BB4	1	0	0	0	0	0	0	0	1	0	0	0	
006BF25A-315A	0	0	1	0	0	0	0	0	0	0	0	0	
006E3A23-F786	2	0	8	1	0	0	0	1	1	0	0	0	
006F8D62-651E	0	0	0	0	0	1	0	0	0	0	0	0	
0070A69E-281C	0	0	1	0	1	0	0	0	0	0	0	0	
0073C602-D9AE	0	0	2	0	0	0	0	0	0	0	0	0	
007CB9B1-F4DI	1	0	0	0	1	0	0	1	1	1	0	0	
008383B8-BF24	1	0	2	0	0	0	0	4	0	0	0	0	

- Transcrip table

PatientGuid	Height_Max	Weight_Max	BMI_Max	SystolicBP_Max	DiastolicBP_Ma	RespiratoryRate	Temperature_M	Height_Min	Weight_Min	BMI_Min	SystolicBP_Min	DiastolicBP_Min	RespiratoryRate T
00023761-9D8D	68	185	28.972	134	86	16.91502579	97.9	67	182	27.67	120	62	16
0005D9BD-0247	67	137	22.796	152	82	16.91502579	98.6	65	0	0	125	59	15
000B4862-7CE7	65.89672602	100.09	16.654	126.728523	76.45129703	20	97.76533086	65	0	0	80	56	14
00110ABC-DAB	71.5	215	29.983	130	80	16.91502579	97.76533086	65.89672602	0	0	125	75	12
002667F4-B9A3	65.89672602	109	22.013	160	90	20	97.76533086	59	0	0	112	70	16.91502579
0029BBC8-7C22	68	186	28.278	163	98	16.91502579	99.8	65.89672602	0	0	101	70	16.91502579
003CEE1F-5BF	66	251.327	22.75	160	82	14	97.76533086	60	105	0	116	54	10
004382BD-E31F	63	242	42.864	152	92	16.91502579	97.76533086	63	230	40.738	134	80	16.91502579
004AACFF-5D6	67	180	26.623	128	96	16.91502579	98.6	65.89672602	0	0	102	80	16
004CC712-BAD	68.5	161.7	24.226	126.728523	78	18	97.76533086	65.89672602	0	0	115	72	16.91502579
004EA6A7-6159	67	142	22.238	126.728523	76.45129703	16.91502579	98.3	65.89672602	0	0	96	64	16.91502579
005F61C9-E537	65.89672602	180	32.341	152	98	20	98.9	61.5	0	0	110	60	15
00607D4B-5F8F	65.89672602	155	22.629	126.728523	76.45129703	16.91502579	97.76533086	65	0	0	84	52	16.91502579
0063B34F-C2C4	68	290	45.415	133	85	20	98.7	65.89672602	0	0	106	72	18
0066553F-732C	65.89672602	265	46.938	148	88	16.91502579	97.76533086	63	0	0	110	72	16.91502579
006948F2-1118-	67.5	170	27.436	130	82	18	102.6	66	164	26.23	110	66	14
006A9198-1BB4	65.89672602	96	18.747	126.728523	82	16.91502579	99.1	60	0	0	98	62	16.91502579
006BF25A-315A	65.89672602	155	28.347	137	81	16.91502579	97.76533086	62	0	0	116	66	16.91502579
006E3A23-F786	72.5	241	32.233	157	79	18	98.4	65.89672602	221	0	118	52	10
006F8D62-651E	67	155	24.481	139	76.45129703	16.91502579	97.76533086	65.89672602	0	0	126.728523	68	16.91502579
0070A69E-281C	65.89672602	218.2	38.648	138	88	16.91502579	97.76533086	62	0	0	118	76.45129703	14
0073C602-D9AE	66	154	24.854	126.728523	85	16.91502579	97.76533086	65.89672602	0	0	103	76.45129703	16.91502579
007CB9B1-F4DI	72	218	29.563	202	87	22	98	72	190	25.766	130	78	20
008383B8-BF24	68	300	46.982	150	90	22	98.6	65.89672602	0	0	110	76.45129703	16

- physician speciality

PatientGuid	PhySp_Academ	PhySp_Addictio	PhySp_Adolesc	PhySp_Adult_R	PhySp_Allergy_I	PhySp_Alternati	PhySp_Anesthe	PhySp_Cardiova	PhySp_Chiropra	PhySp_Clinical_	PhySp_Clinical_	PhySp_Critical_	PhySp_Dermato
00023761-9D8D	0	0	0	0	0	0	0	0	0	0	0	0	0
0005D9BD-0247	0	0	0	0	0	0	0	0	0	0	0	0	0
000B4862-7CE7	0	0	0	0	0	0	0	0	0	0	0	0	0
00110ABC-DAB	0	0	0	0	0	0	0	0	0	0	0	0	0
002667F4-B9A3	0	0	0	0	0	0	0	0	0	0	0	0	0
0029BBC8-7C22	0	0	0	0	0	0	0	0	0	0	0	0	0
003CEE1F-5BF	0	0	0	0	0	0	0	0	0	0	0	0	0
004382BD-E31F	0	0	0	0	0	0	0	0	0	0	0	0	0
004AACFF-5D6	0	0	0	0	0	0	0	0	0	0	0	0	0
004CC712-BAD	0	0	0	0	0	0	0	0	0	0	0	0	0
004EA6A7-6159	0	0	0	0	0	0	0	0	0	0	0	0	0
005F61C9-E537	0	0	0	0	0	0	0	0	0	0	0	0	0
00607D4B-5F8F	0	0	0	0	0	0	0	0	0	0	0	0	0
0063B34F-C2C4	0	0	0	0	0	0	0	0	0	0	0	0	0
0066553F-732C	0	0	0	0	0	0	0	0	0	0	0	0	0
006948F2-1118-	0	0	0	0	0	0	0	0	0	0	0	0	0
006A9198-1BB4	0	0	0	0	0	0	0	0	0	0	0	0	0
006BF25A-315A	0	0	0	0	0	0	0	0	0	0	0	0	0
006E3A23-F786	0	0	0	0	0	0	0	0	0	0	0	0	0
006F8D62-651E	0	0	0	0	0	0	0	0	0	0	0	0	0
0070A69E-281C	0	0	0	0	1	0	0	0	0	0	0	0	0
0073C602-D9AE	0	0	0	0	0	0	0	0	0	0	0	0	0
007CB9B1-F4DI	0	0	0	0	0	0	0	0	0	0	0	0	0
008383B8-BF24	0	0	0	0	0	0	0	0	0	0	0	0	0

2.1.2 Input schema

Feature name	Datatype	Size	Null/Required
Age	int	3	Required

2.2 Predicting Disease

- The system displays the choices of the disease.
- The User chooses the target disease by clicking one of the available diseases.
- The User selects the disease.
- The system presents the set of inputs required from the user.
- The user gives required information.
- The system should be able to predict whether infected for the chosen disease based on the user information.

2.3 Logging

We should be able to log every activity done by the user.

- The System identifies at what step logging required
- The System should be able to log each and every system flow.
- Developers can choose logging methods. You can choose database logging/ File logging as well.
- System should not be hung even after using so many loggings. Logging just because we can easily debug issues so logging is mandatory to do.

2.4 Database

System needs to store every request into the database and we need to store it in such a way that it is easy to retrain the model as well.

1. The User chooses the disease.
2. The User gives required information.
3. The system stores each and every data given by the user or received on request to the database. Database you can choose your own choice whether MongoDB/ MySQL.

2.5 Deployment

1. AWS

3 Technology stack

Front End	HTML/CSS/JS/React
Backend	Python Django
Database	MongoDB/MySql
Deployment	AWS

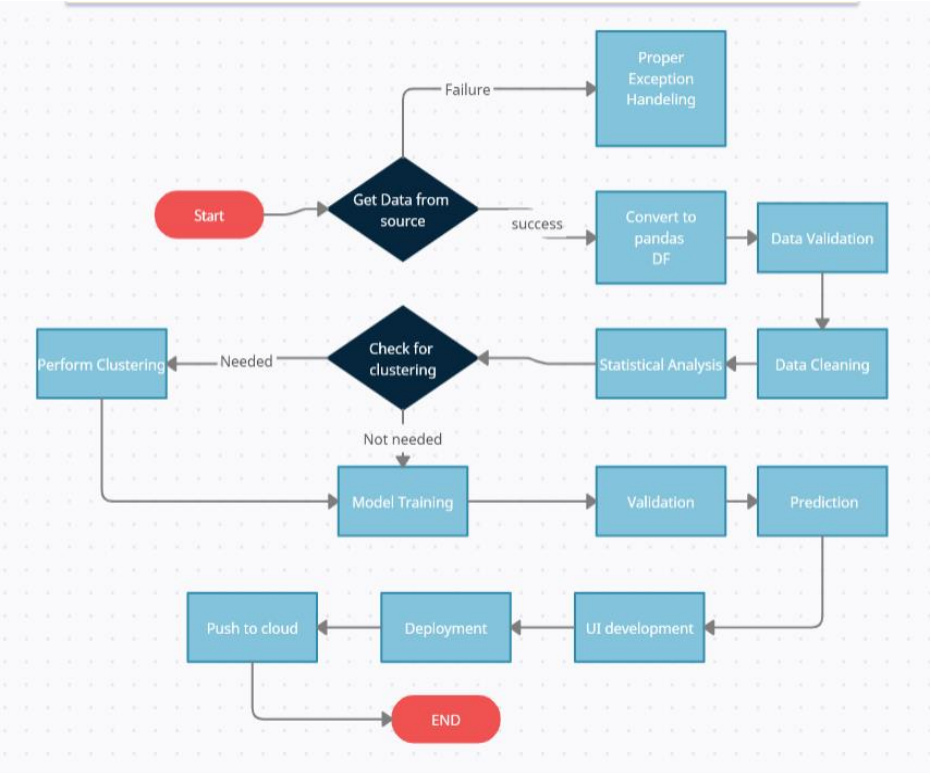
4 Proposed Solution

refer: https://arxiv.org/abs/1808.04928

Based on the actual research paper, if we are using history of the patient to predict the future then we might want to consider using LSTM. However, drawing a baseline in the form of some Machine Learning algorithm would be helpful. Why making a baseline model important? Well, to compare the performance of our actual model, let say LSTM in this case, is very important to ascertain that we are in the right direction as if performance of LSTM is not better than the baseline model then there is no point of using LSTM.

1. Baseline Model: Logistic Regression, since this is a classification problem.
2. Actual model: LSTMs.

5 Model training/validation workflow



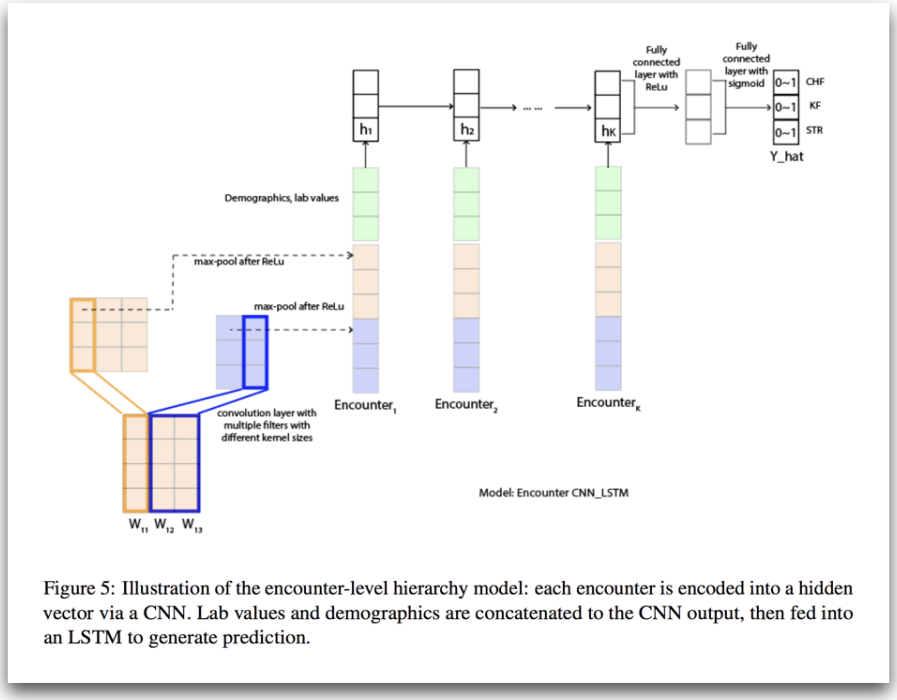
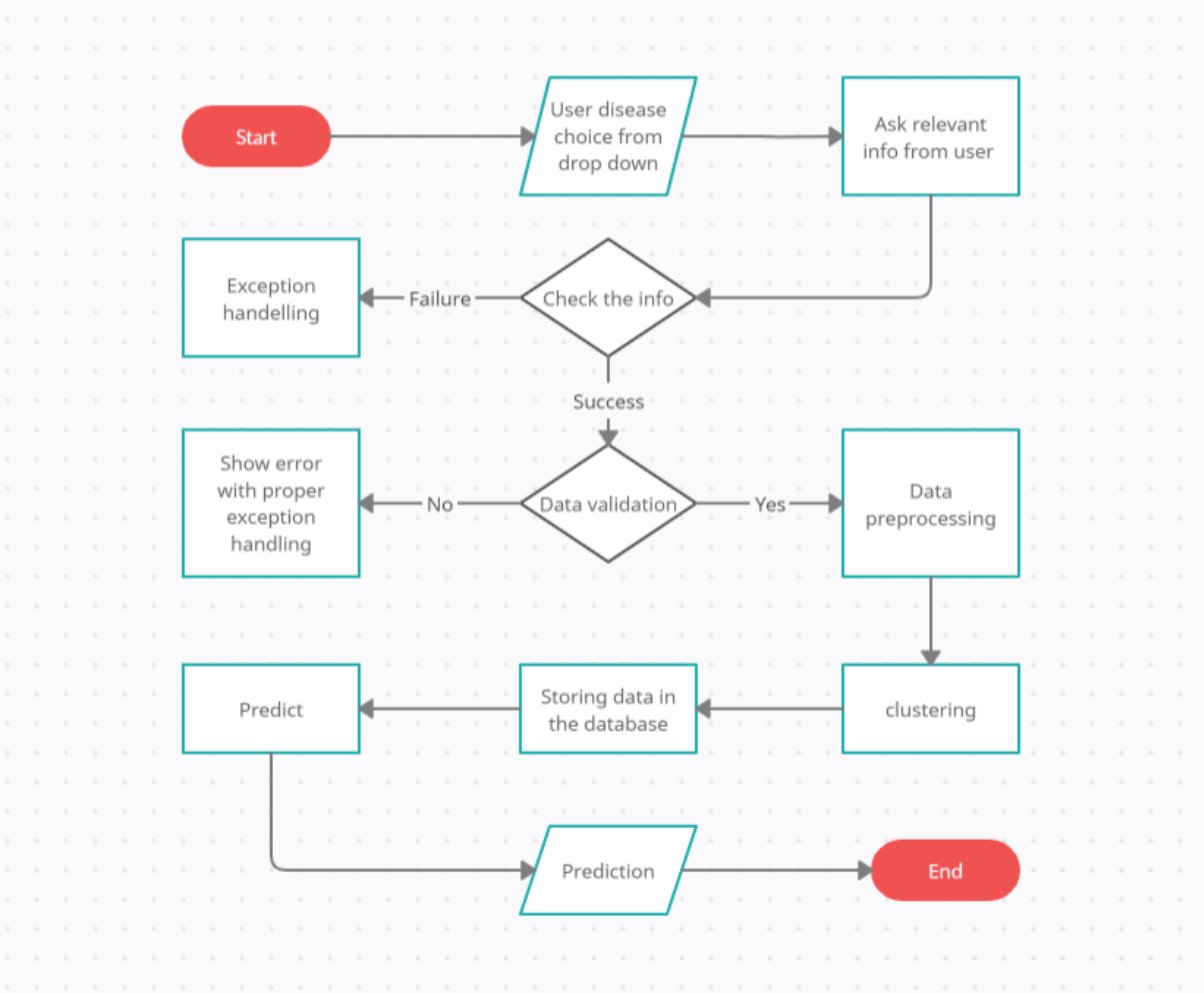


Figure 5: Illustration of the encounter-level hierarchy model: each encounter is encoded into a hidden vector via a CNN. Lab values and demographics are concatenated to the CNN output, then fed into an LSTM to generate prediction.

6 User I/O workflow



7 Exceptional scenarios

Step	Exception	Mitigation	Module
Date	1.1	First Draft	Name
Date	1.2	Added Workflow chart	Name

8 Test cases

Test case	Steps to perform test case	Module	Pass/Fail

9 Key performance indicators (KPI)

- Time and workload reduction using the EHR model.
- Comparison of accuracy of model prediction and doctor's prediction.
- Number of times a patient visits the hospital.
- Time between symptom onset and detection of illness/visit to hospital.
- Immunity of patient (based on previous illnesses).
- Vaccines the patient has taken.
- Length of stays in hospital.