CHRONIC DISEASE PREDICTION

An ML Based Application to predict Risk of Chronic Disease

- 19Z202 -Anita Priyadarshini A
- 19z211-Divya Darshini R
- 19z217- Hemavarshini B
- 19Z227 - Monirhithikka S.P
- 19Z240 - Samyuktha Sreekanth
- 19Z243 - Sarayu Miththira V.C

BATCH - 4

TABLE OF CONTENTS

INTRODUCTION

In the field of biomedical and healthcare communities the accurate prediction plays the major role to find out the risk of the disease in the patient. The prediction gives the benefits of early disease detection. Here we use a certain machine learning algorithm to state the risk of disease. Prediction process is done using the dataset provided online from certain hospitals, the entire dataset will be preprocessed and the missing values will be reconstructed. Compared to several types of prediction algorithms, the Logistic Regression model gives the highest accuracy of prediction.

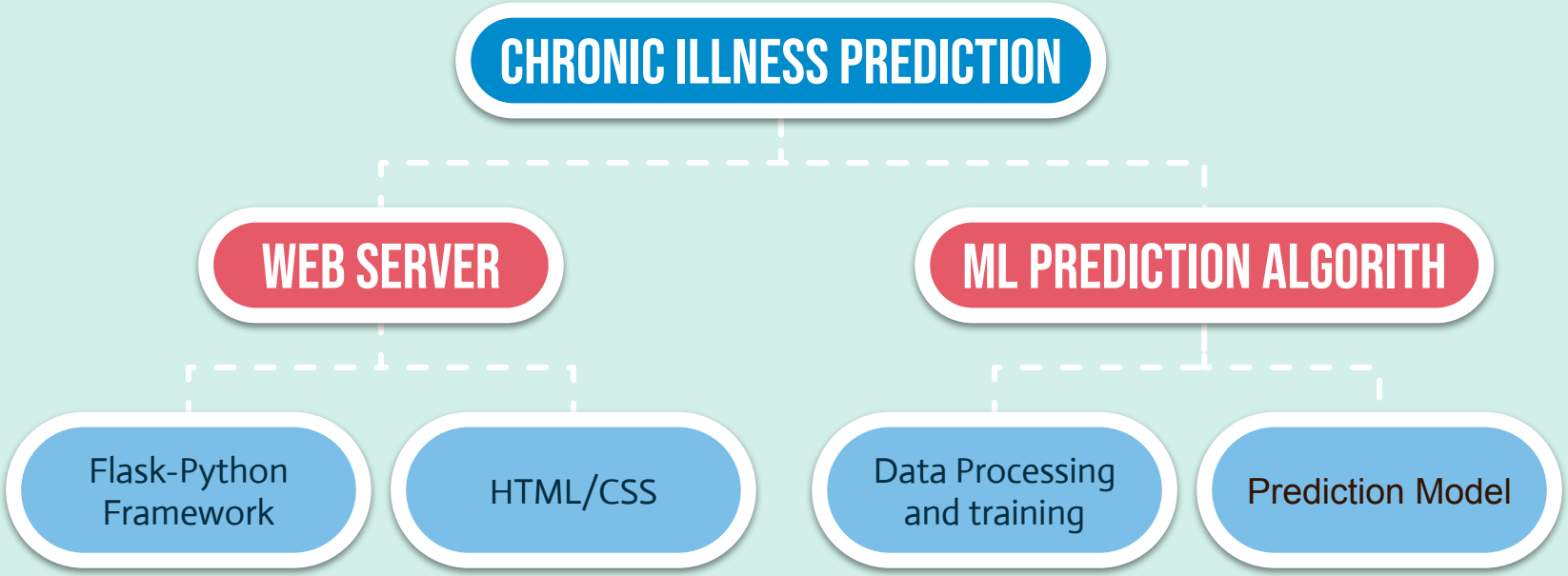
```
The model 'Logistic Regression' has an Average Accuracy: 76.94 %
Maximum Accuracy that can be obtained: 85.71
```

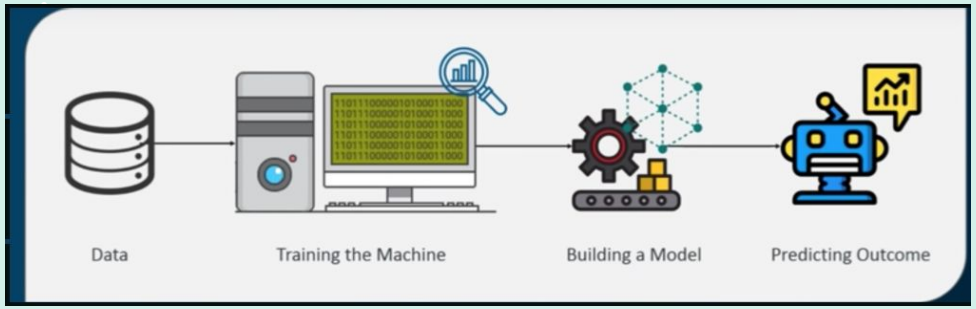
Finally, This was integrated to our Web Application, Built on Visual Studio Code and MySQL Work bench

PROBLEM STATEMENT

To Predict The risk of Chronic Disease in User, By analysing Collected parameters

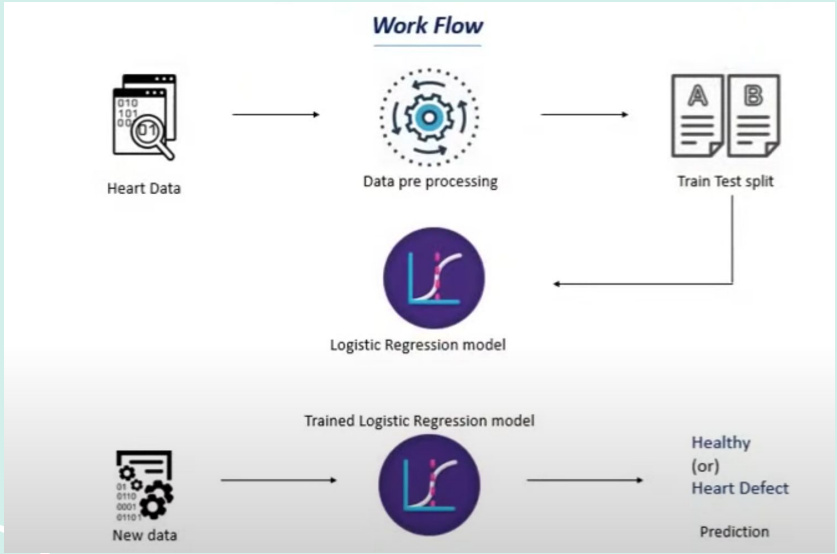
OUR OBJECTIVE

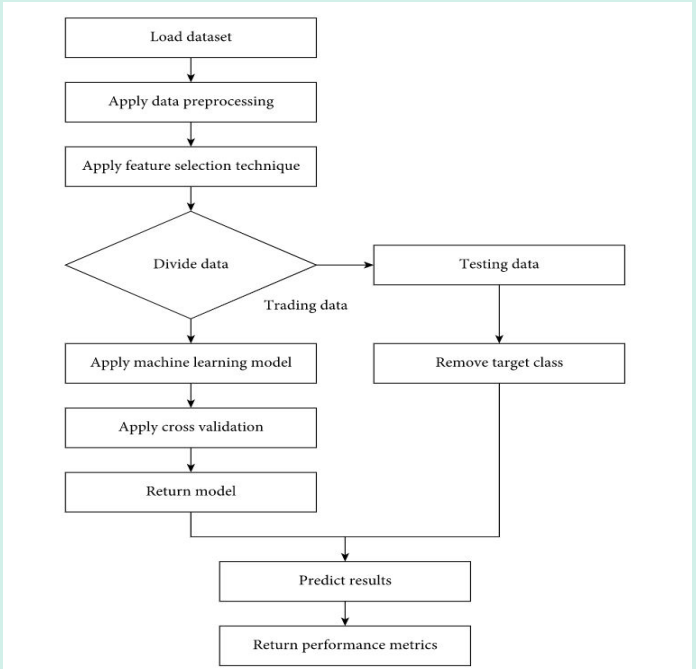


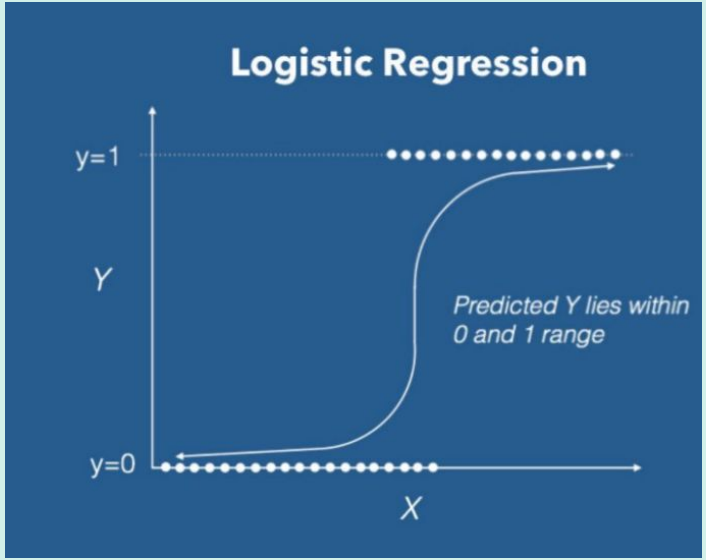


Classification

- Supervised Learning
- Output is a categorical quantity
- Main aim is to compute the category of the data
- Eg: Classify emails as spam or non-spam
- Algorithm: Logistic Regression




Load dataset
Apply data preprocessing
Apply feature selection technique
Divide data
Testing data
Trading data
Apply machine learning model
Remove target class
Apply cross validation
Return model
Predict results
Return performance metrics


Logistic Regression
y=1
Y
Predicted Y lies within
0 and 1 range
y=0
X

Different models Diabetes.ipynb

File Edit View Insert Runtime Tools Help Last edited on October 30

+ Code + Text

```
The model 'Logistic Regression' has an Average Accuracy: 76.94 %

Maximum Accuracy that can be obtained: 85.71 %

Standard Deviation: 0.053713908554226225


The model 'K Nearest Neighbour' has an Average Accuracy: 75.26 %

Maximum Accuracy that can be obtained: 82.89 %

Standard Deviation: 0.05441598084625328


The model 'Decision Tree' has an Average Accuracy: 69.52 %

Maximum Accuracy that can be obtained: 76.62 %

Standard Deviation: 0.0477305251276215


The model 'Random Forest' has an Average Accuracy: 76.82 %

Maximum Accuracy that can be obtained: 84.42 %

Standard Deviation: 0.051902513486997225


The model 'SVM' has an Average Accuracy: 75.9 %

Maximum Accuracy that can be obtained: 84.42 %

Standard Deviation: 0.05894736252068271


The model 'XGB' has an Average Accuracy: 76.43 %

Maximum Accuracy that can be obtained: 81.82 %

Standard Deviation: 0.040319660624529706


The model 'LightGBM' has an Average Accuracy: 73.96 %

Maximum Accuracy that can be obtained: 81.82 %

Standard Deviation: 0.062171867376286175
```

Stroke.ipynb

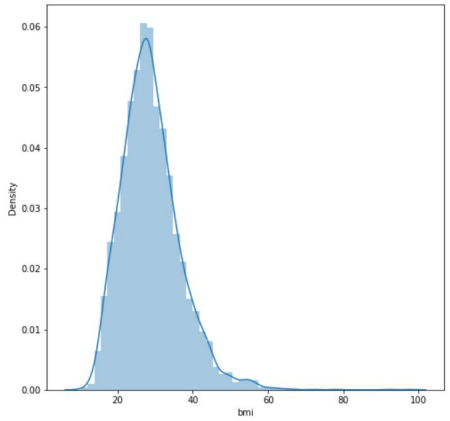
File Edit View Insert Runtime Tools Help Last edited on November 11

+ Code + Text

We have null values in BMI, so we will replace all null values with the median value

```
fig, ax = plt.subplots(figsize=(8,8))
sns.distplot(heart_data.bmi)
```

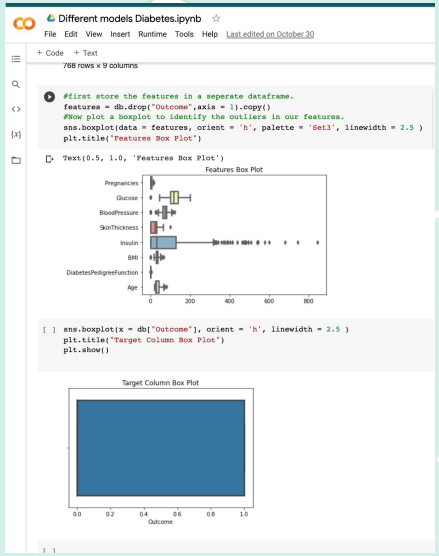
```
/usr/local/lib/python3.7/dist-packages/seaborn/distributions.py:2619: FutureWarning
  warnings.warn(msg, FutureWarning)
<matplotlib.axes._subplots.AxesSubplot at 0x7f9a756d2e50>
```

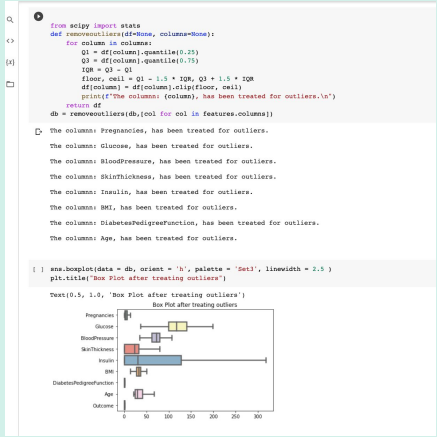



```
heart_data['bmi'].fillna(heart_data['bmi'].median(),inplace=True)
```

```
heart_data.isnull().sum()
```

```
id                   0
gender               0
age                  0
hypertension         0
heart_disease        0
ever_married         0
work_type            0
Residence_type       0
avg_glucose_level    0
bmi                  0
smoking_status       0
stroke               0
dtype: int64
```



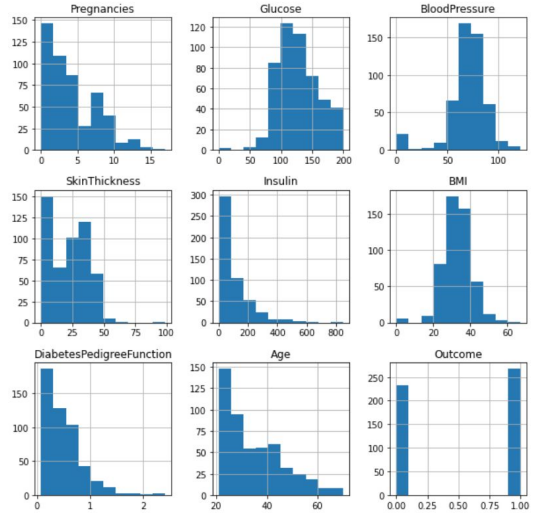


Diabetes.ipynb

File Edit View Insert Runtime Tools Help All changes saved

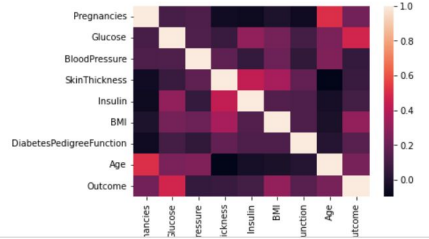
+ Code + Text

```
nds.hist(bins = 10, figsize = (10,10))
plt.show()
```




```
sns.heatmap(nds.corr())
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f48668022d0>
```

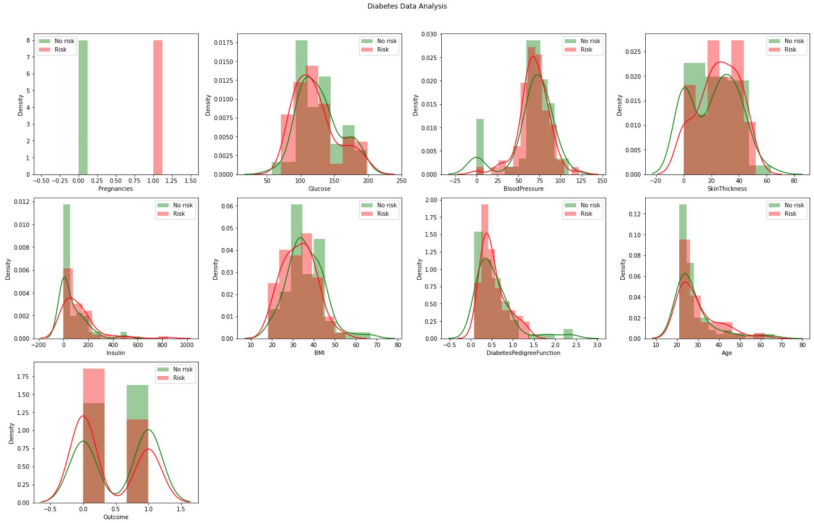


Diabetes.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
fig.subplots_adjust(top=0.95)
plt.show()
```



MAJOR MODULES

DISEASE PORTAL

Select The Chronic Disease to be tested

INPUT PARAMETERS

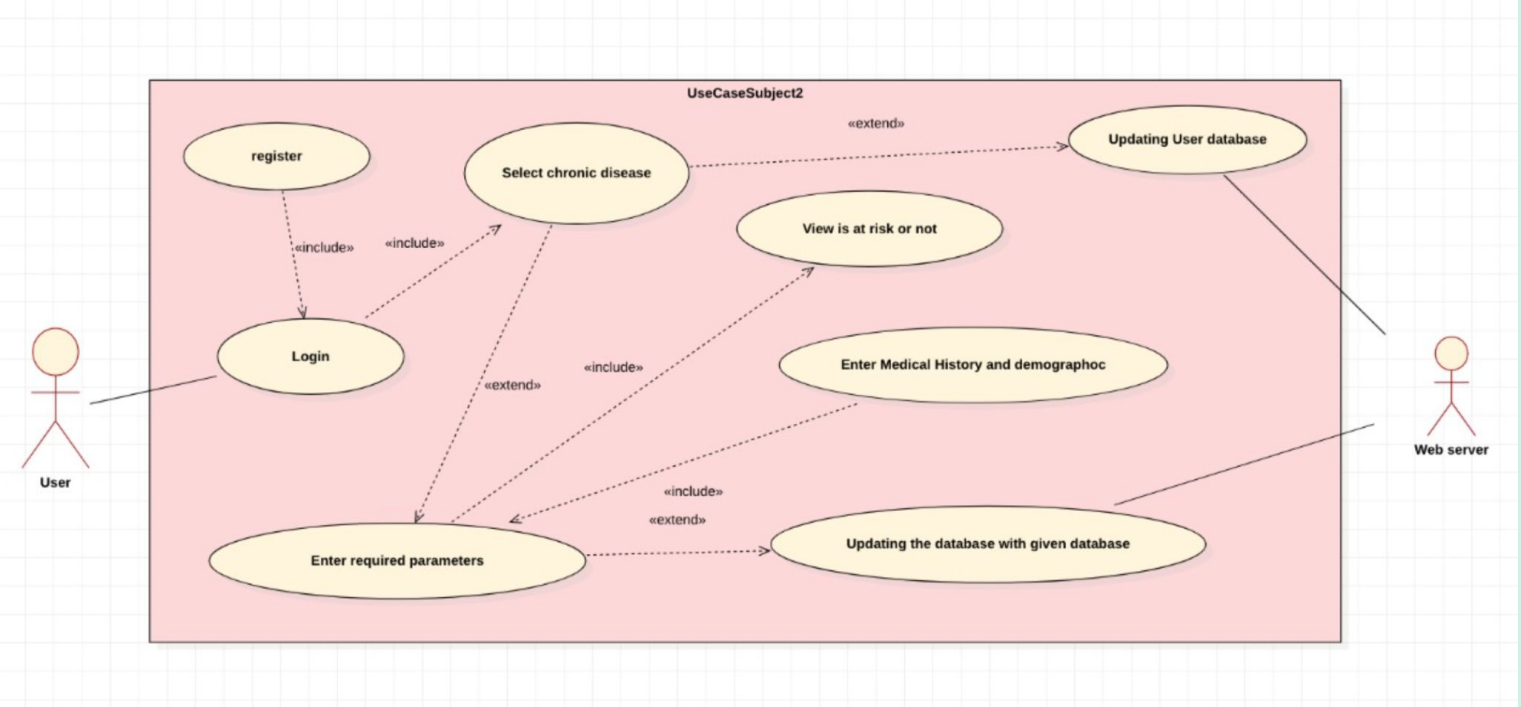
Collect Data from User as required by trained model's dataset

ANALYSIS AND OUTPUT

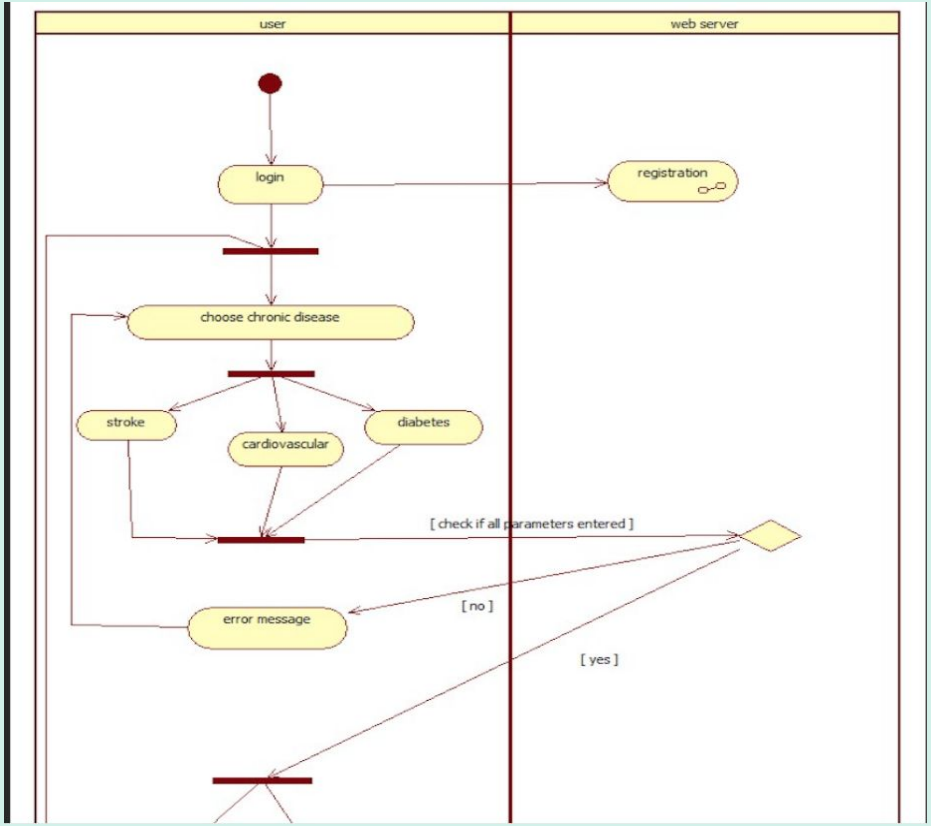
Display the Outcome of the analysis made to the user


DESIGN

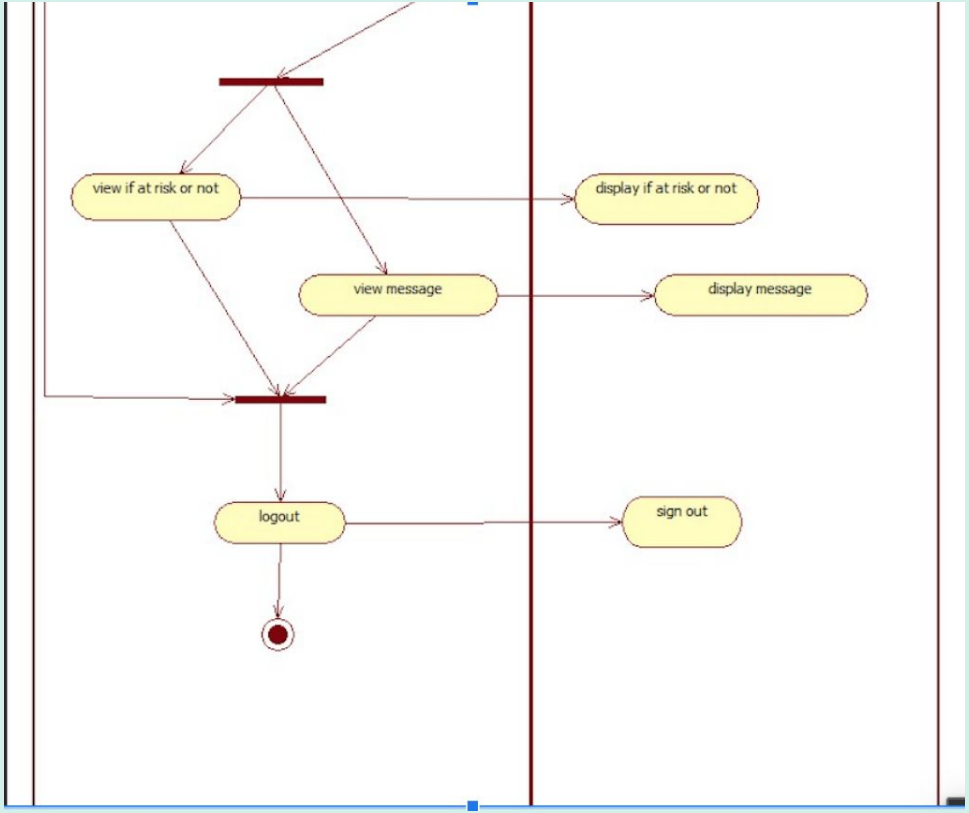
USECASE DIAGRAM



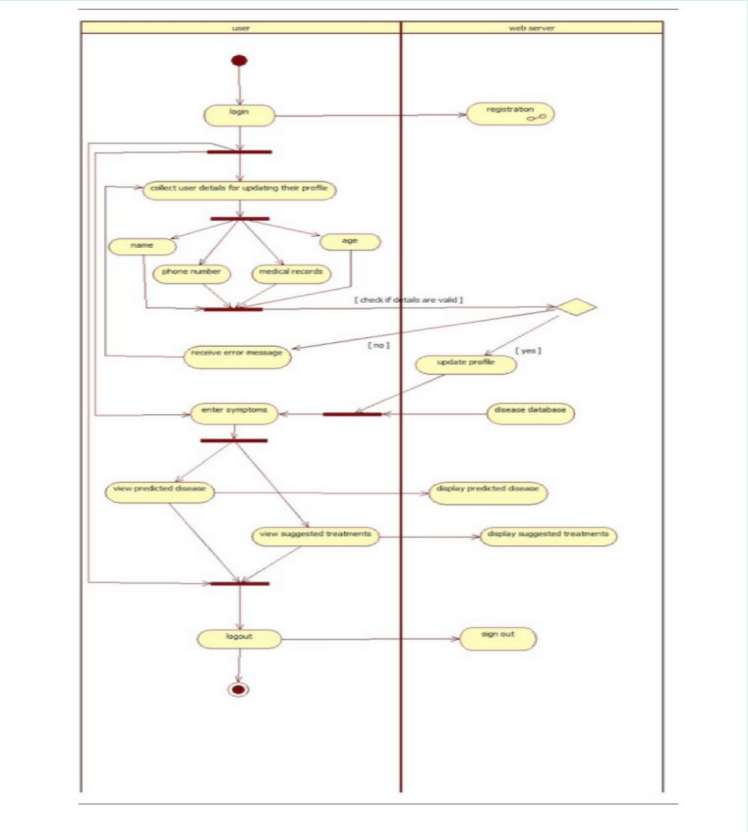
ACTIVITY DIAGRAM



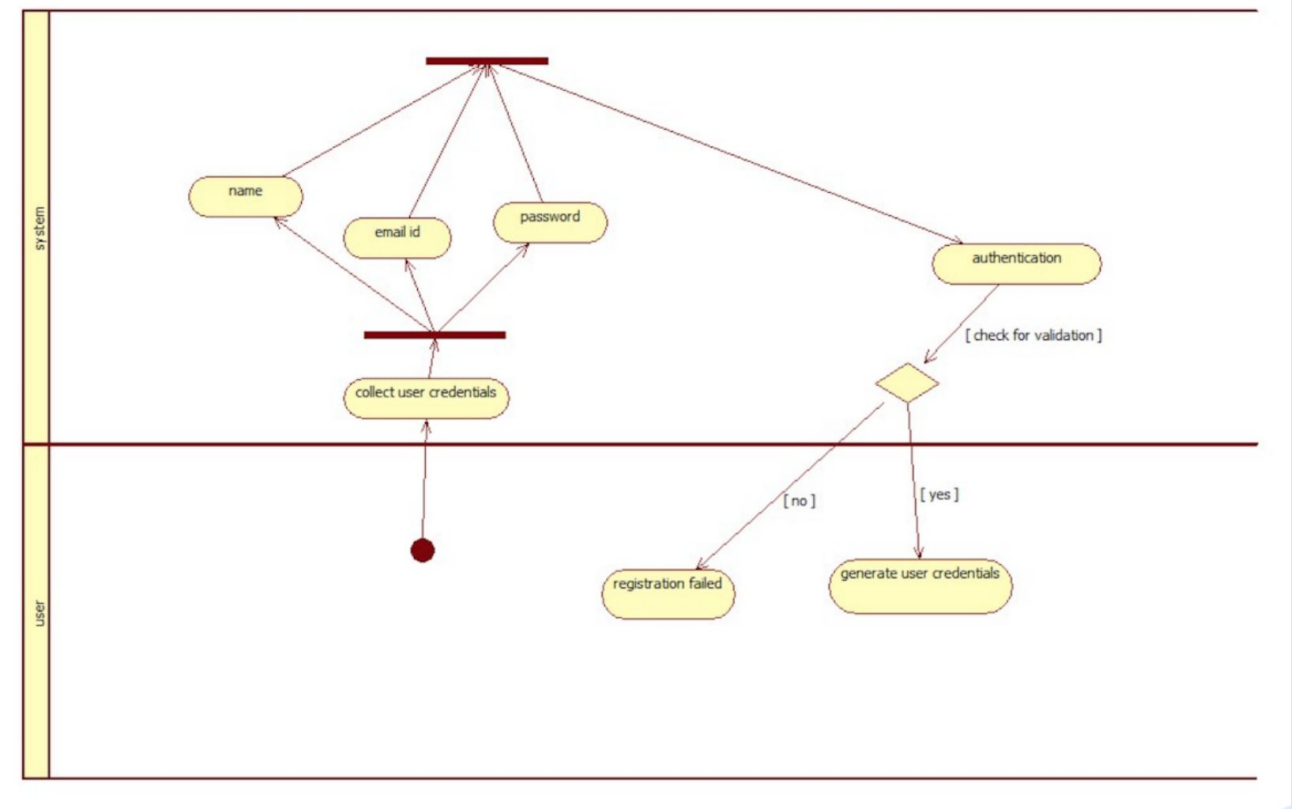
ACTIVITY DIAGRAM



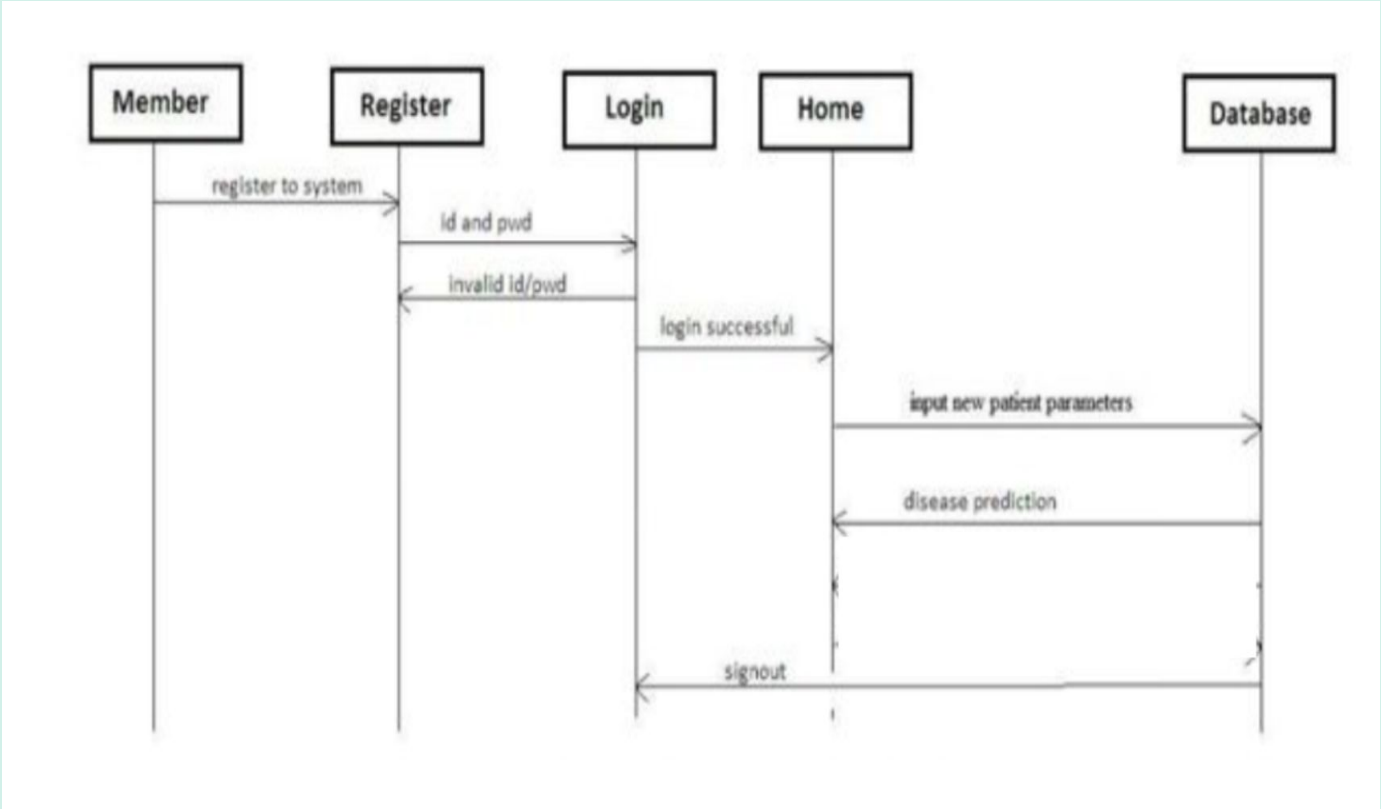
ACTIVITY DIAGRAM



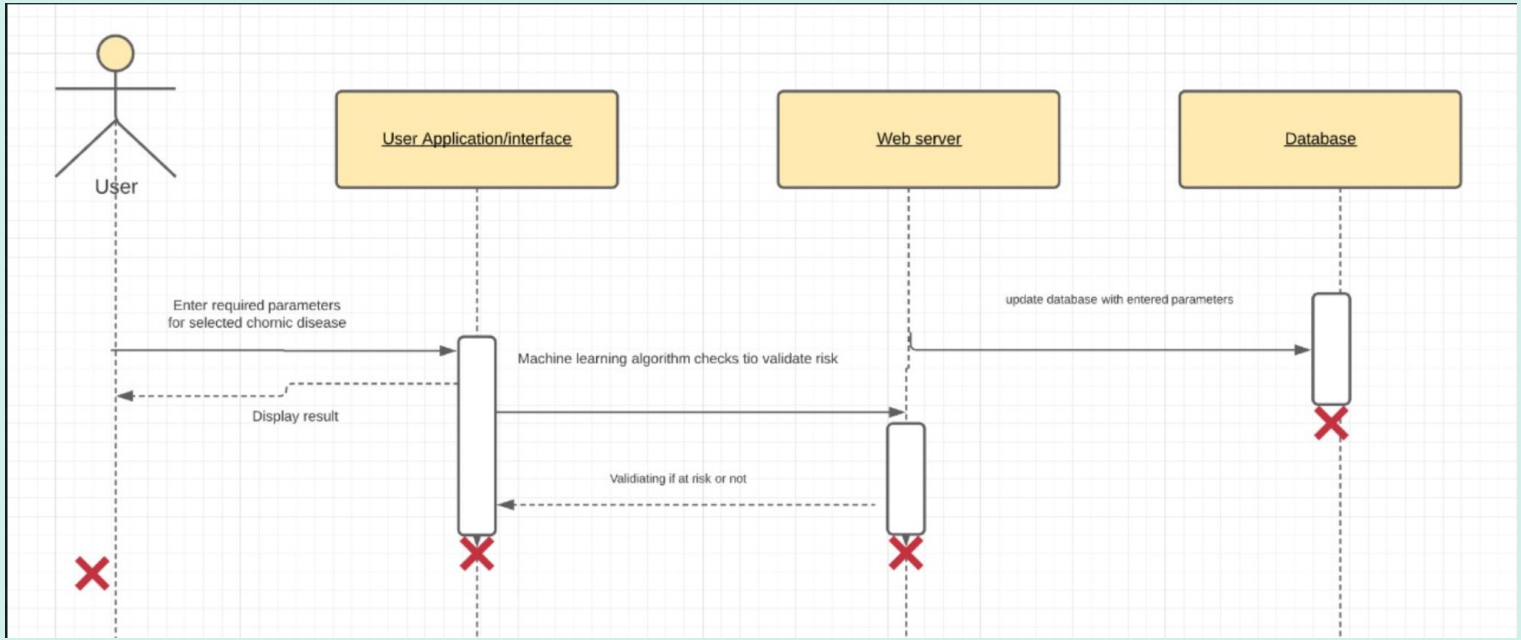
ACTIVITY DIAGRAM



SEQUENCE DIAGRAM

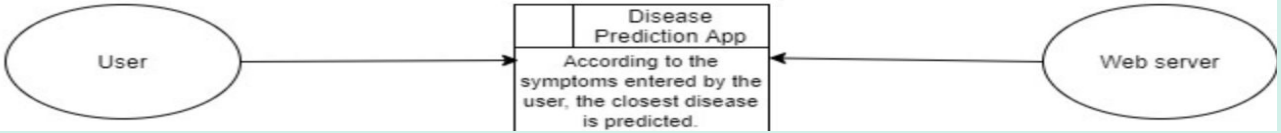


SEQUENCE DIAGRAM

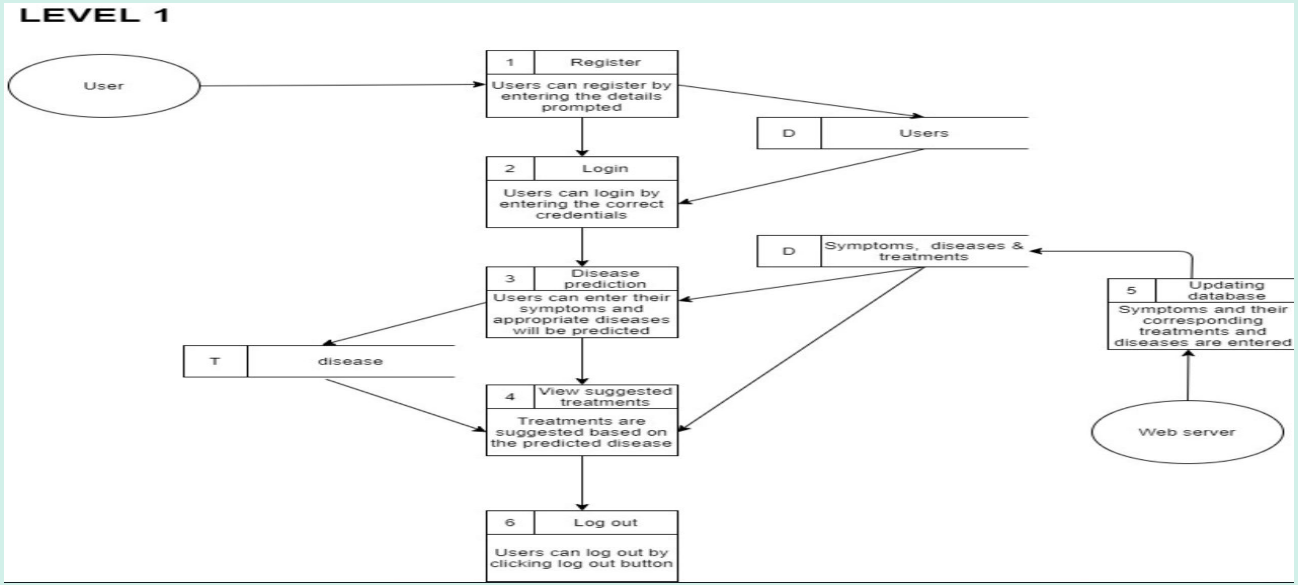


DFD DIAGRAM

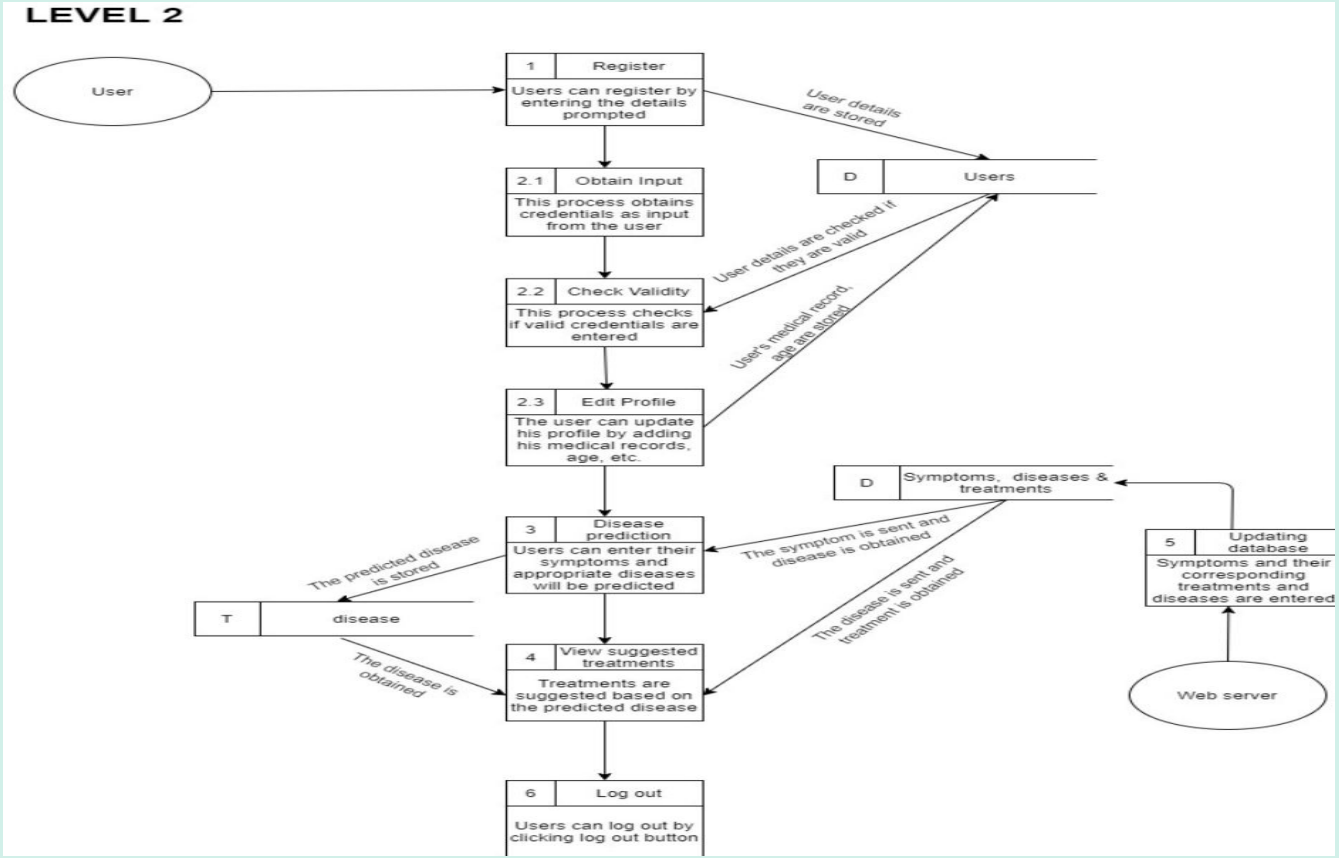
LEVEL 0



LEVEL 1



DFD DIAGRAM


LEVEL 2
User
1 Register
Users can register by entering the details prompted
User details are stored
D Users
2.1 Obtain Input
This process obtains credentials as input from the user
User details are checked if they are valid
2.2 Check Validity
This process checks if valid credentials are entered
User's medical record, age are stored
2.3 Edit Profile
The user can update his profile by adding his medical records, age, etc.
D Symptoms, diseases & treatments
3 Disease prediction
Users can enter their symptoms and appropriate diseases will be predicted
The symptom is sent and disease is obtained
The predicted disease is stored
T disease
The disease is obtained
4 View suggested treatments
Treatments are suggested based on the predicted disease
The disease is sent and treatment is obtained
5 Updating database
Symptoms and their corresponding treatments and diseases are entered
Web server
6 Log out
Users can log out by clicking log out button

SYSTEM REQUIREMNT

Software and Hardware Requirements

SOFTWARE REQUIREMENTS

FRONTEND

1. HTML
2. CSS

BACKEND

1. Python

FRAMEWORK

1. Flask

SOFTWARE REQUIREMENTS

IDE

1. Visual Studio Code

DATABASE

1. MySQL

ML MODEL

1. Google Collab
2. Jupyter

HARDWARE REQUIREMENTS

OS-WINDOWS

HARDDISK-40GB

RAM-256 MB

PROCESSOR-PENTIUM ®
DUAL CORE CPU

CONTRIBUTIONS

WEB APPLICATION

FRONTEND AND BACKEND

19Z227 - Monirhithikka S.P

19Z240 - Samyuktha Sreekanth

19Z243 - Sarayu Miththira V.C

ML MODEL AND TRAINING

CODE AND DATASETS

19Z202 -Anita Priyadarshini

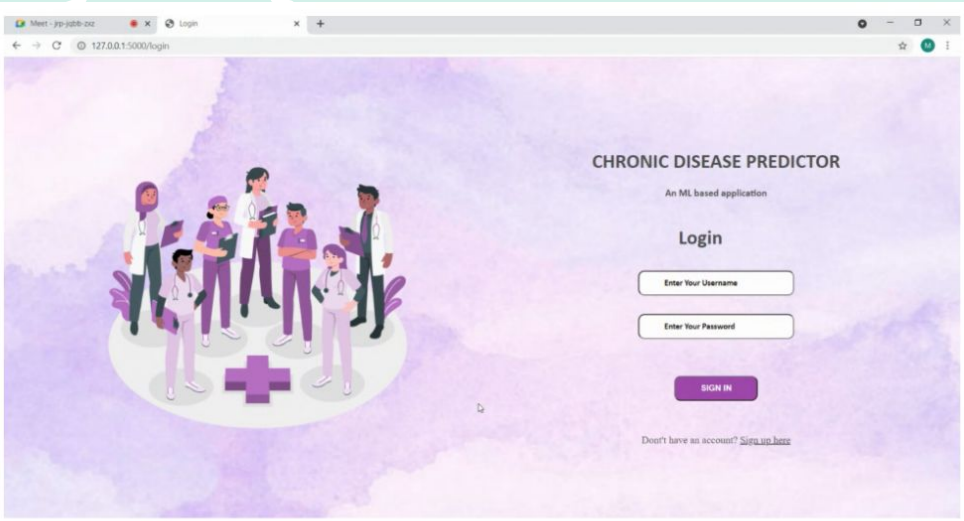
ML MODEL AND DOCUMENTATION

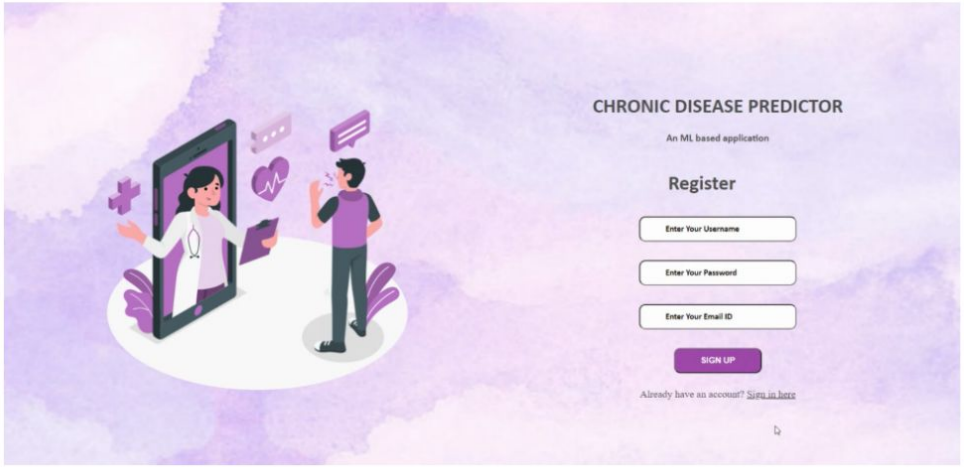
SYSTEM ARCHITECTURE AND ML ALGORITHM

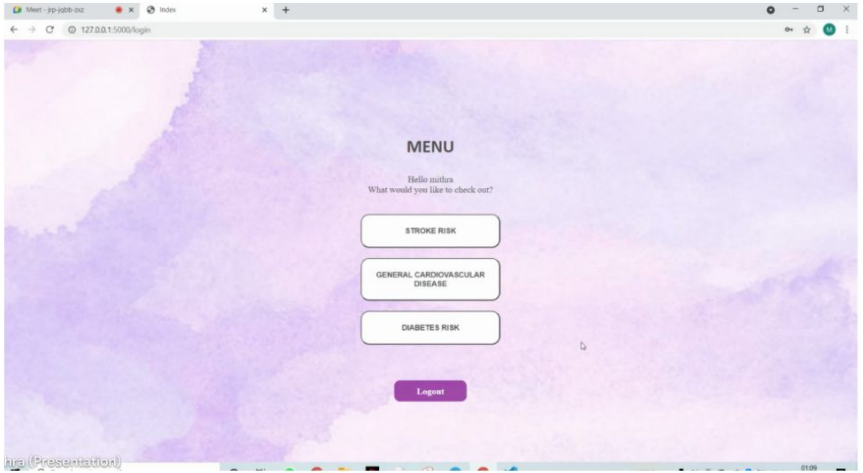
19z211-Divya Darshini R

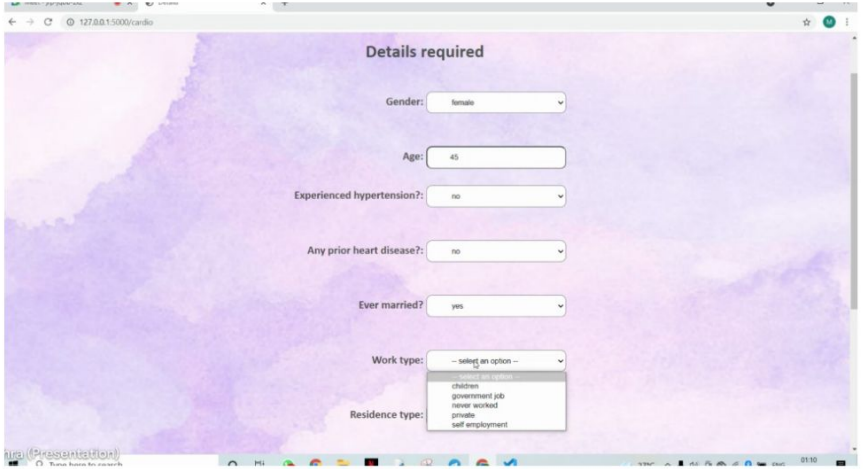
19z217- Hemavarshini B

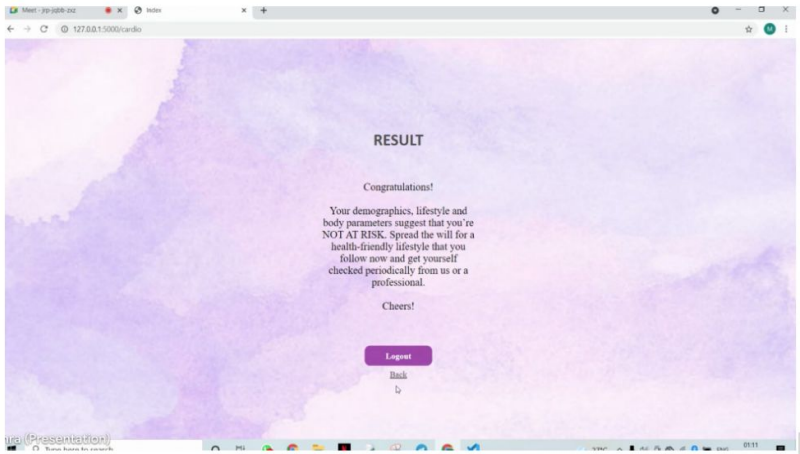
INTERFACE AND DEMO

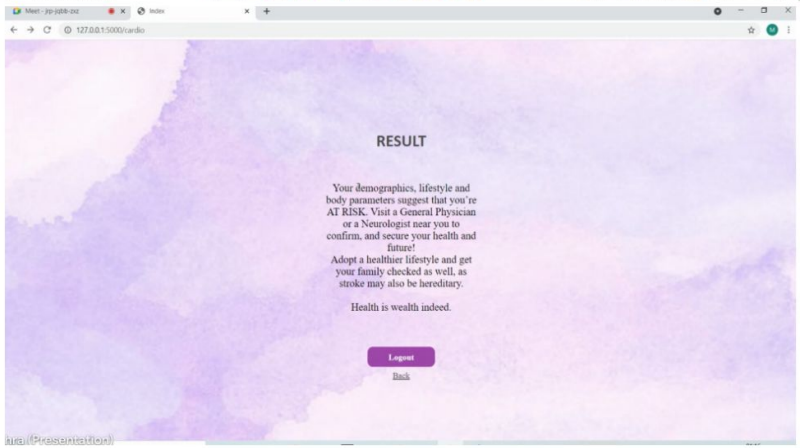

127.0.0.1:5000/login
CHRONIC DISEASE PREDICTOR
An ML based application
Login
Enter Your Username
Enter Your Password
SIGN IN
Don't have an account? Sign up here

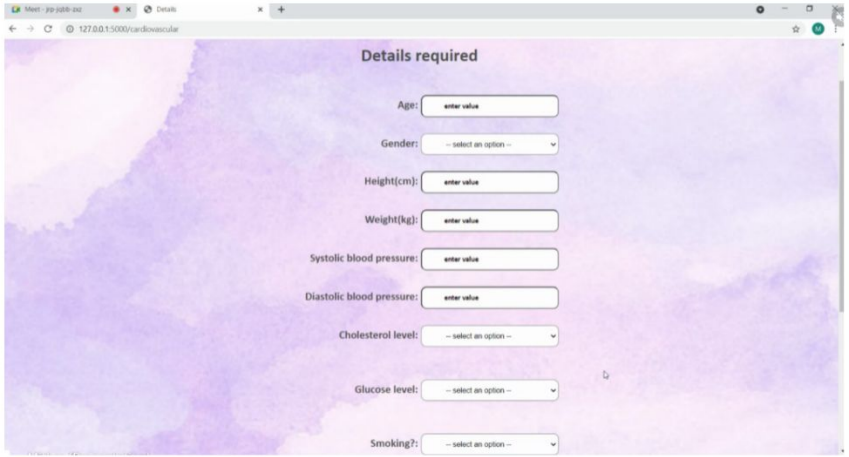

CHRONIC DISEASE PREDICTOR
An ML based application
Register
Enter Your Username
Enter Your Password
Enter Your Email ID
SIGN UP
Already have an account? Sign in here

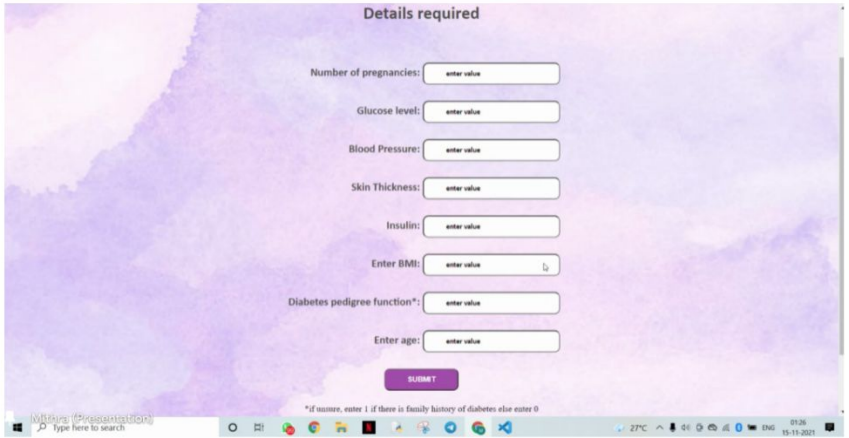

MENU
Hello mithra
What would you like to check out?
STROKE RISK
GENERAL CARDIOVASCULAR DISEASE
DIABETES RISK
Logout


Details required
Gender:
female
Age:
45
Experienced hypertension?:
no
Any prior heart disease?:
no
Ever married?
yes
Work type:
-- select an option --
children
government job
never worked
private
self employment
Residence type:

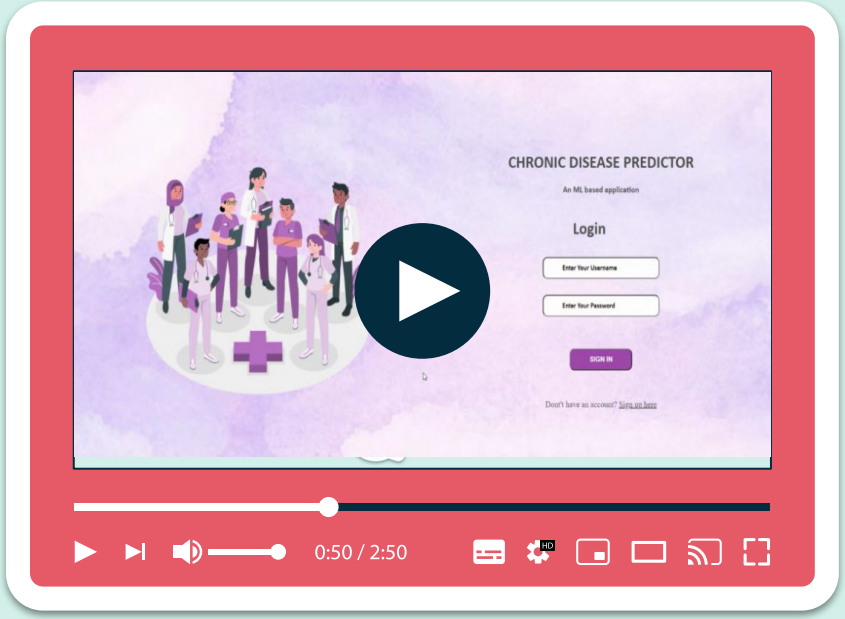

RESULT
Congratulations!
Your demographics, lifestyle and body parameters suggest that you're NOT AT RISK. Spread the will for a health-friendly lifestyle that you follow now and get yourself checked periodically from us or a professional.
Cheers!
Logout
Back


RESULT
Your demographics, lifestyle and body parameters suggest that you're AT RISK. Visit a General Physician or a Neurologist near you to confirm, and secure your health and future!
Adopt a healthier lifestyle and get your family checked as well, as stroke may also be hereditary.
Health is wealth indeed.
Logout
Back


Details required
Age:
Gender:
Height(cm):
Weight(kg):
Systolic blood pressure:
Diastolic blood pressure:
Cholesterol level:
Glucose level:
Smoking?:


Details required
Number of pregnancies:
Glucose level:
Blood Pressure:
Skin Thickness:
Insulin:
Enter BMI:
Diabetes pedigree function*:
Enter age:
SUBMIT
*if unsure, enter 1 if there is family history of diabetes else enter 0

DEMO


CHRONIC DISEASE PREDICTOR
An ML based application
Login
Enter Your Username
Enter Your Password
SIGN IN
Don't have an account? Sign up here
0:50 / 2:50

THANK YOU!

We Hope To develop This Project further by adding more Chronic Diseases and training More accurate Analysis Models. We would also like deploy it and have it serve as a stand-alone application so that it can be accessed by all and Detect Chronic diseases in a Timely manner with the power of Data and Efficient Computation