CADS

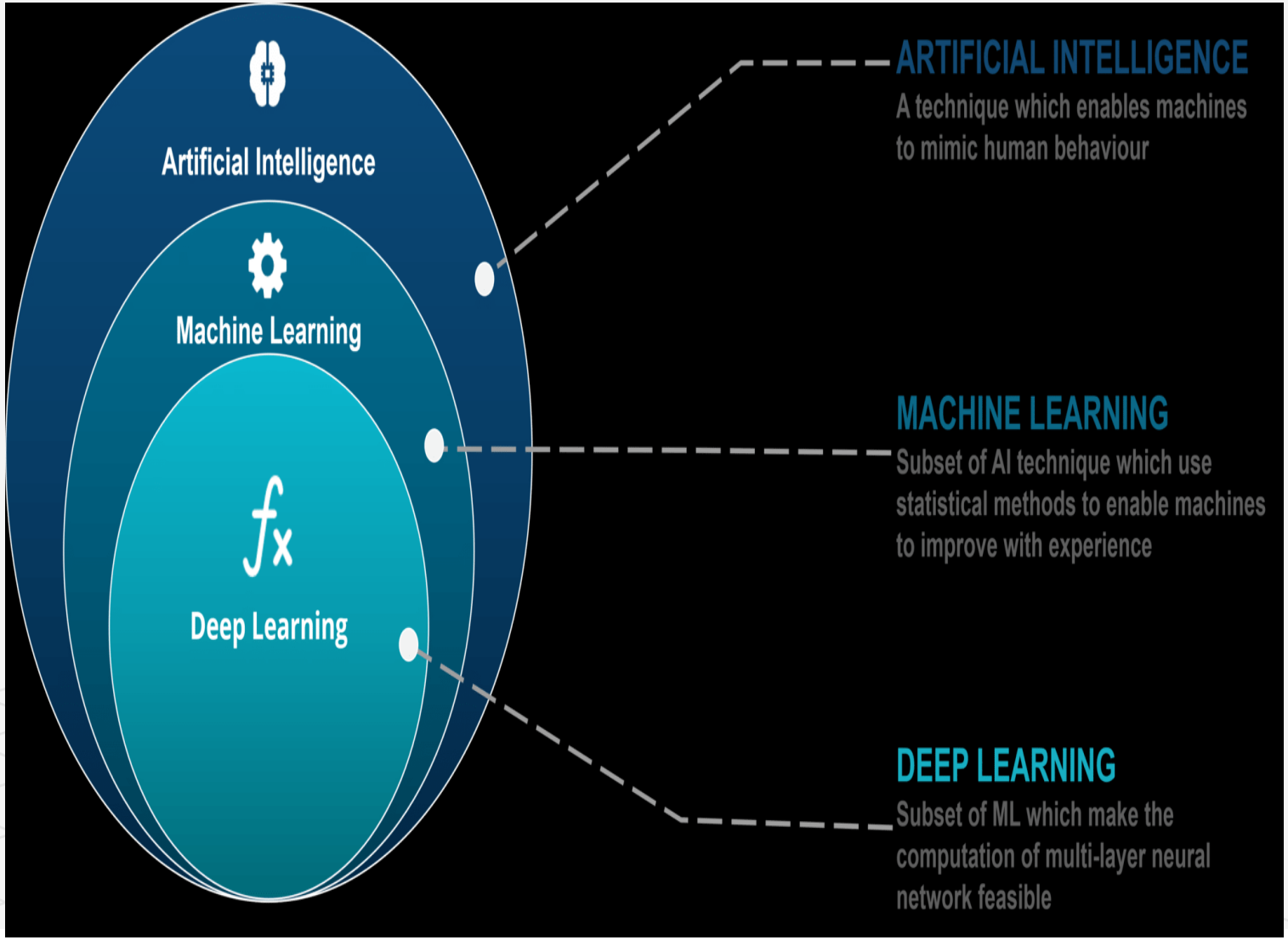
(COMPUTER AIDED DIAGNOSE SYSTEM)

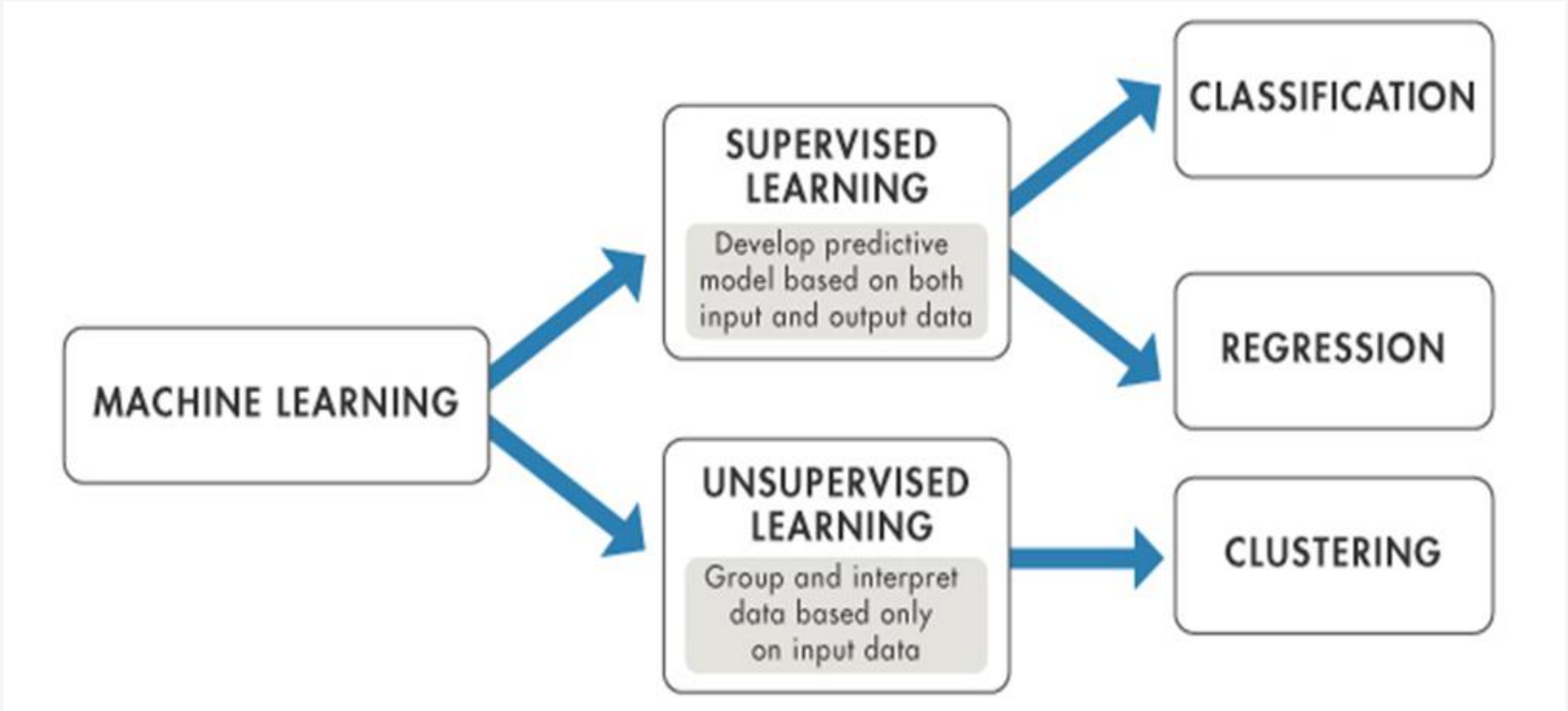
SUBMITTED BY
LALITA PRASAD PANT

Table of content

Introduction

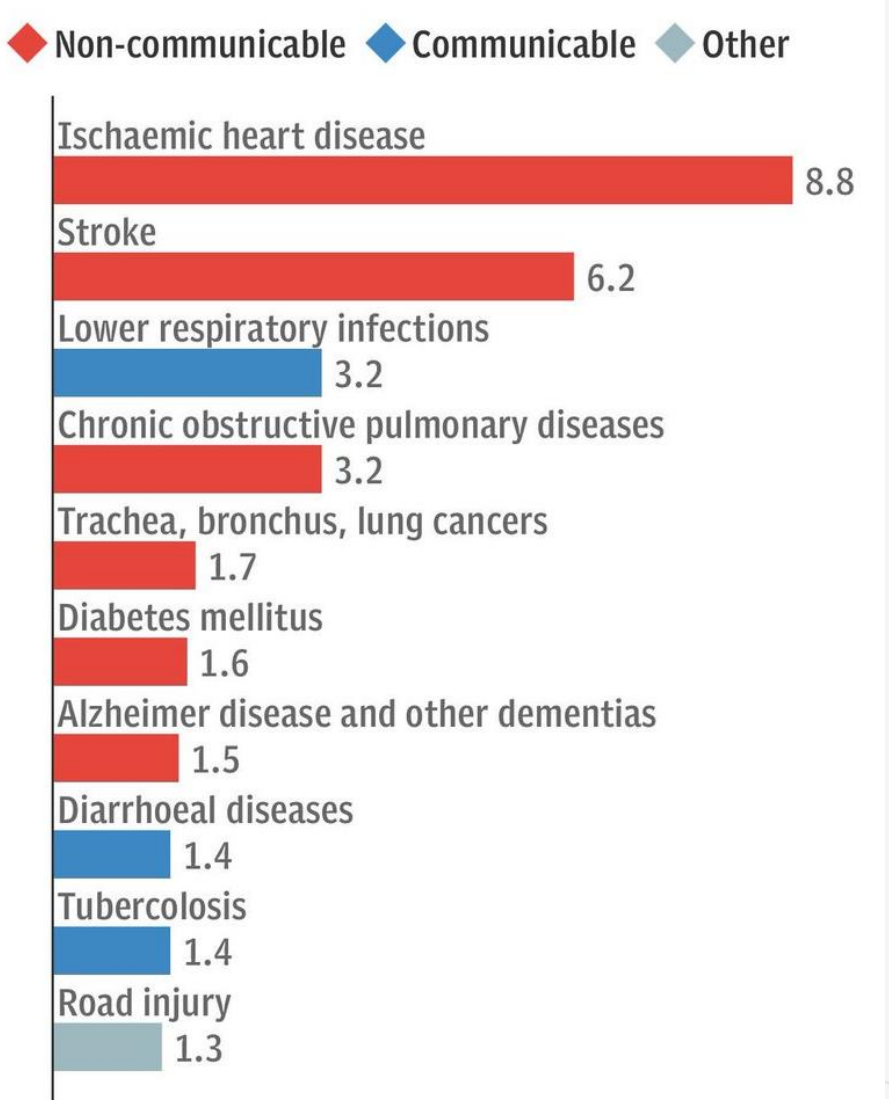
Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. **Machine learning** focuses on the development of computer programs that can access data and use it learn for themselves.



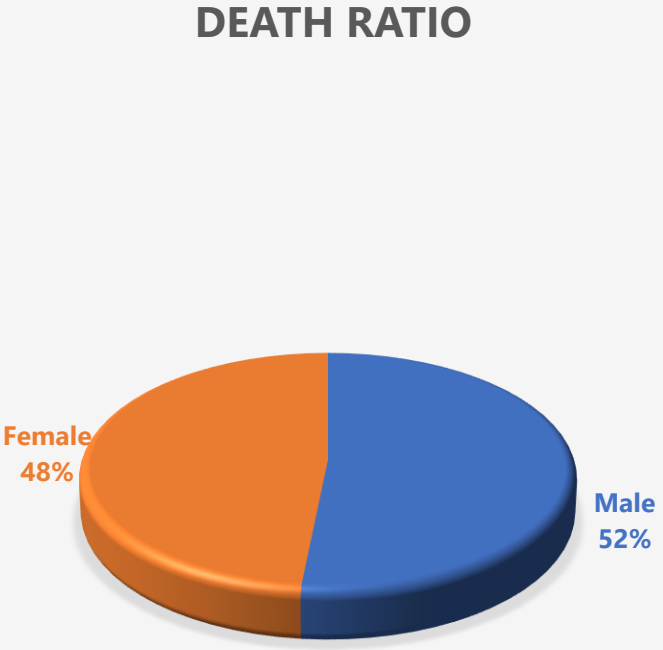

MACHINE LEARNING
SUPERVISED LEARNING
Develop predictive model based on both input and output data
UNSUPERVISED LEARNING
Group and interpret data based only on input data
CLASSIFICATION
REGRESSION
CLUSTERING

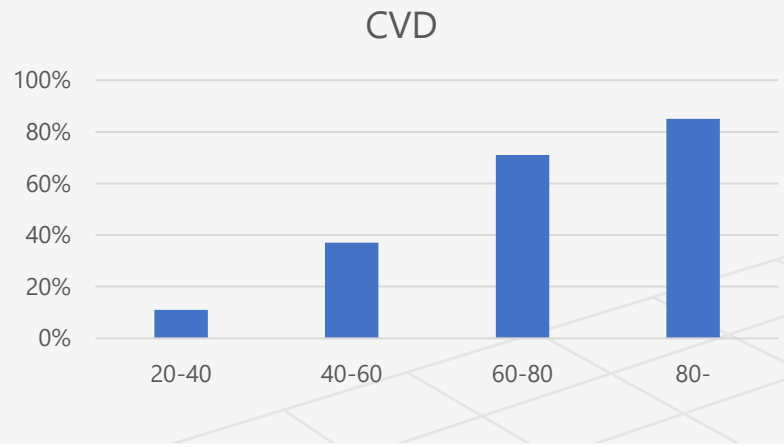
Heart disease and strokes are the biggest global killers

- According to WHO (World Health Organization) estimated that **17.9 million (31%)** deaths occur worldwide.
- In United States every **34 seconds** the heart disease kills one person.
- The important heart diseases are Cardiovascular Disease, Cardiomyopathy and Coronary heart disease.
- Cardiovascular disease generally refers to conditions that involve narrowed or blocked blood vessels that can lead to a heart attack, chest pain (angina) or stroke. Other heart conditions, such as those that affect your heart's muscle, valves or rhythm, also are considered forms of heart disease.



- **Coronary artery disease** and stroke account for 80% of **CVD** deaths in males and 75% of **CVD** deaths in females. Most **cardiovascular disease** affects older adults. In the United States 11% of people between 20 and 40 have **CVD**, while 37% between 40 and 60, 71% of people between 60 and 80, and 85% of people over 80 have **CVD.**
- Based on Heart Disease the WHO (World Health Organization) estimated that 17.9 million (31%) deaths occur worldwide.





Breast Cancer

- According to WHO (World Health Organization) estimated that impacting 2.1 million women each year, deaths occur worldwide.

- In 2018, it is estimated that 627,000 women died from breast cancer – that is approximately 15% of all cancer deaths among women.

- In order to improve breast cancer outcomes and survival, early detection is critical.

- There are two early detection strategies for breast cancer:-
- early diagnosis
- screening.
    - *Mammography*
    - *Clinical Breast Exam (CBE)*

Developed System and Its Advantages

- Accurate information about the disease when input given by user according to the parameters.
- With this help, accurate information of multiple diseases will be available through single window system.
- Under this, all the main algorithms (ML & DL) will be used and accurate information will be available.
- We have also kept an option to find out the current COVID19 disease through the demo data for accurate information.

Note: We have created the demo data itself, keeping in mind the main parameters of this disease, which are being released by the scientist in research.

- We know that if the information is in the form of graphics, along with the number, it is better understood, that is why the information will be fully visualized.
- Along with this, this system also gives you the facility that will be able to tell that you have which kind of disease when the symptoms are provided.
- This system will give you all the information that is required such as
- Accuracy, precision, recall, aucroc curve, train loss, valid loss etc
- With this help, a final result can be given from the accuracy of multiple algorithms, which will give more accurate results.
- This system also has the feature that you can compare multiple algorithms for one data set and get all the necessary information.

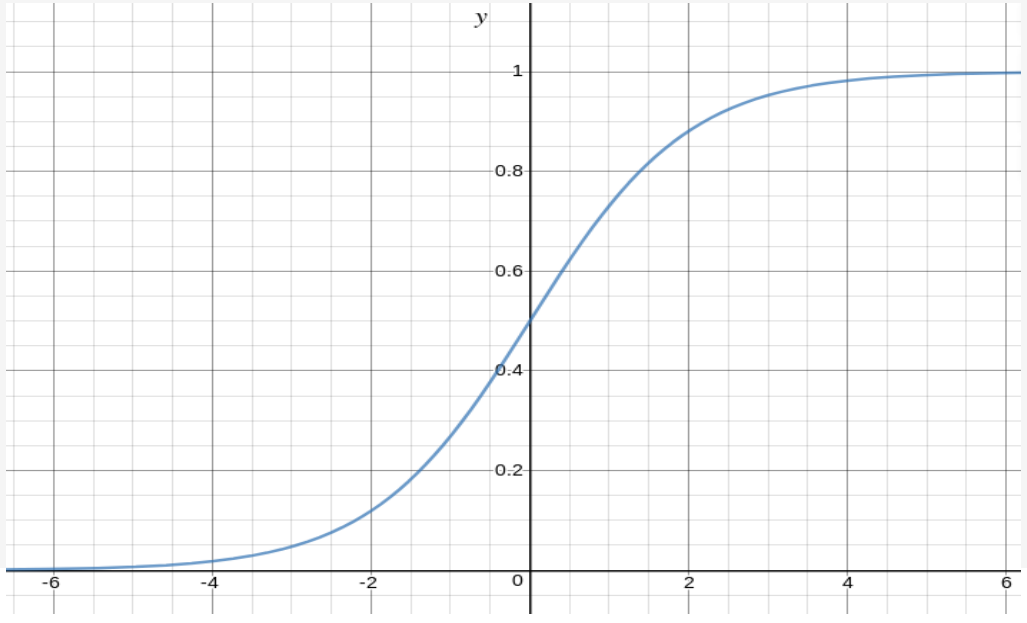
Regression

Regression is a statistical measurement used to determine the strength of the relationship between one dependent variable (usually denoted by Y) and a series of other changing variables (known as independent variables (X)).

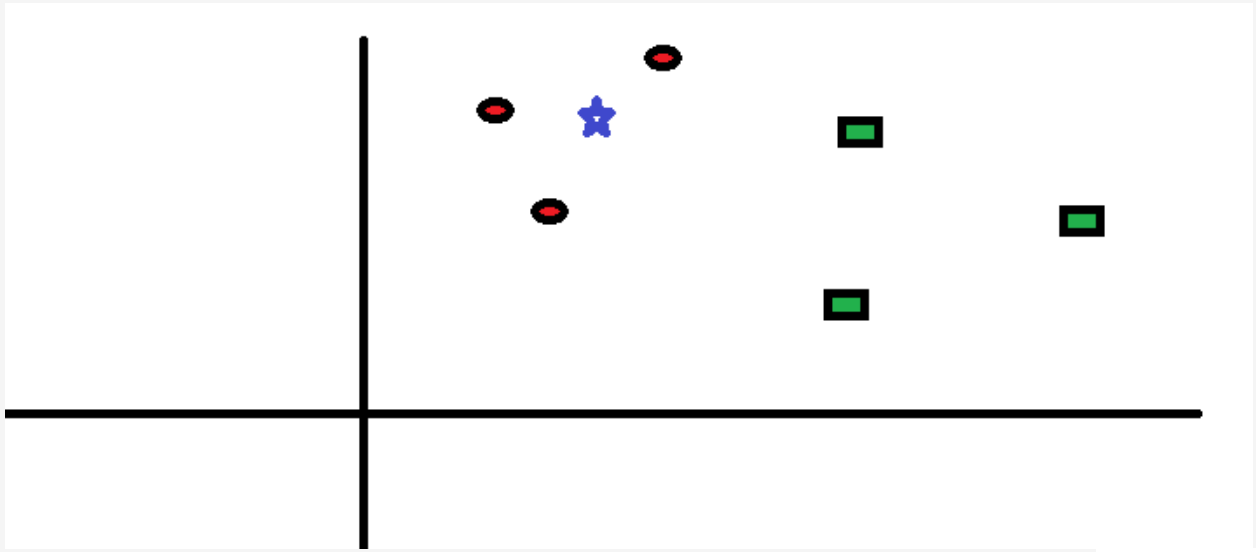
- **Logistic Regression** is basically a supervised classification algorithm. In a classification problem, the target variable (or output), y, can take only discrete values for given set of features (or inputs), X.
- Instead of predicting exactly 0 or 1, **logistic regression** generates a probability—a value between 0 and 1, exclusive.
- You might be wondering how a logistic regression model can ensure output that always falls between 0 and 1.

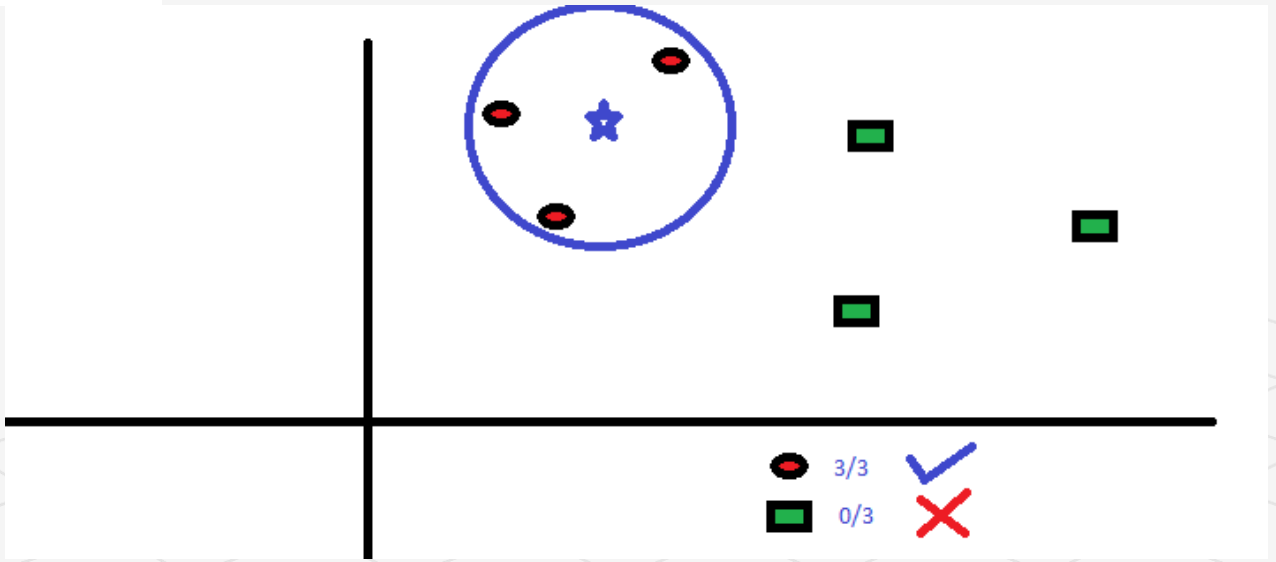
$$y = \frac{1}{1 + e^{-z}}$$

- Is called **logistic function** or the **sigmoid function**. Here is a plot showing g(z):

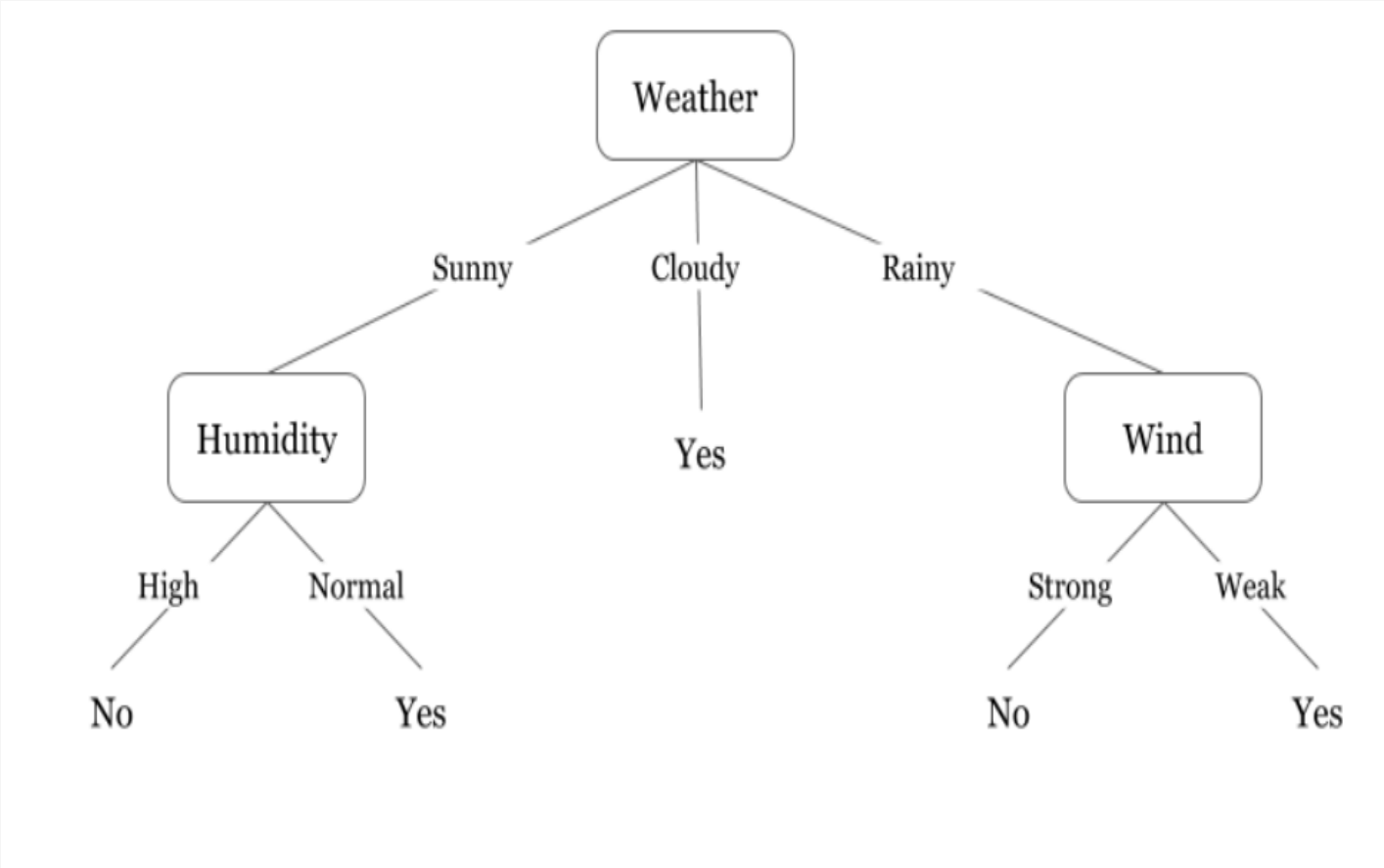


K-Nearest Neighbor(KNN) Algorithm :



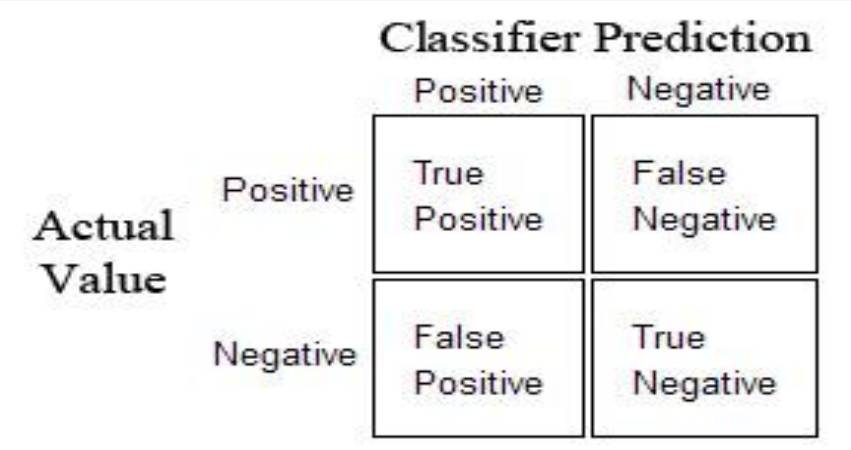


Decision Tree :



Confusion Matrix

A confusion matrix is a summary of prediction results on a classification problem.



Recall: Recall refers to the percentage of your results correctly classified.

$$Recall = \frac{TP}{TP + FN}$$

Precision: Precision refers to the percentage of your results which are relevant.

$$Precision = \frac{TP}{TP + FP}$$

Related Work

Research Papers	DOI
Study on the new design of computer-aided diagnosis system	10.1109/ITIME.2009.5236464
A development of computer-aided diagnosis system using fundus images	10.1109/VSMM.2001.969697
A Computer-aided diagnosis system for classifying prominent skin lesions using machine learning	10.1109/CEEC.2018.8674183
A proposed computer-aided diagnosis system for Parkinson's disease classification using ^{123}I-FP-CIT imaging	10.1109/ATSIP.2017.8075510
New Approach for Cancer Computer Aided Diagnosis and Treatment	10.1109/IIH-MSP.2009.67

Objective

- The main objective of this project is to create a software system that will give information about the disease using the latest technology and more accurate data, in this we will be better with the help of highly advanced algorithm of machine learning and the deep neural network ,There is an effort to create a better system with the help of updated data.
- With this help, we can treat the disease before it causes more harm.

TOOLS / TECHNOLOGY USED

Python:

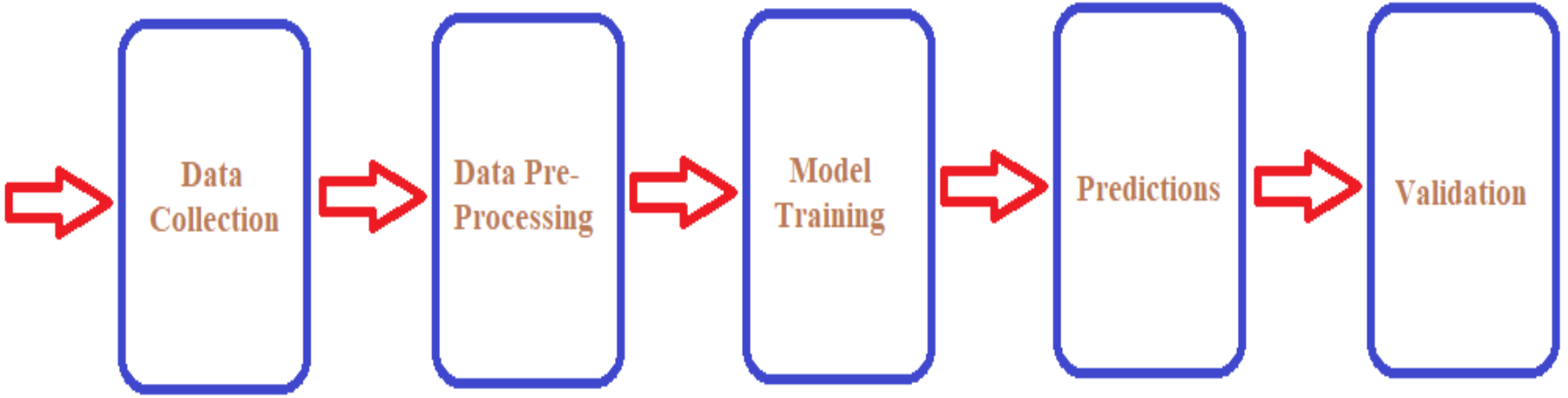
Python is a widely used general-purpose, high level programming language. It was initially designed by Guido van Rossum in 1991 and developed by Python Software Foundation.

Anaconda Spyder/VS Code:

Anaconda is a free and open-source distribution of the Python and R programming languages for scientific computing, that aims to simplify package management and deployment.

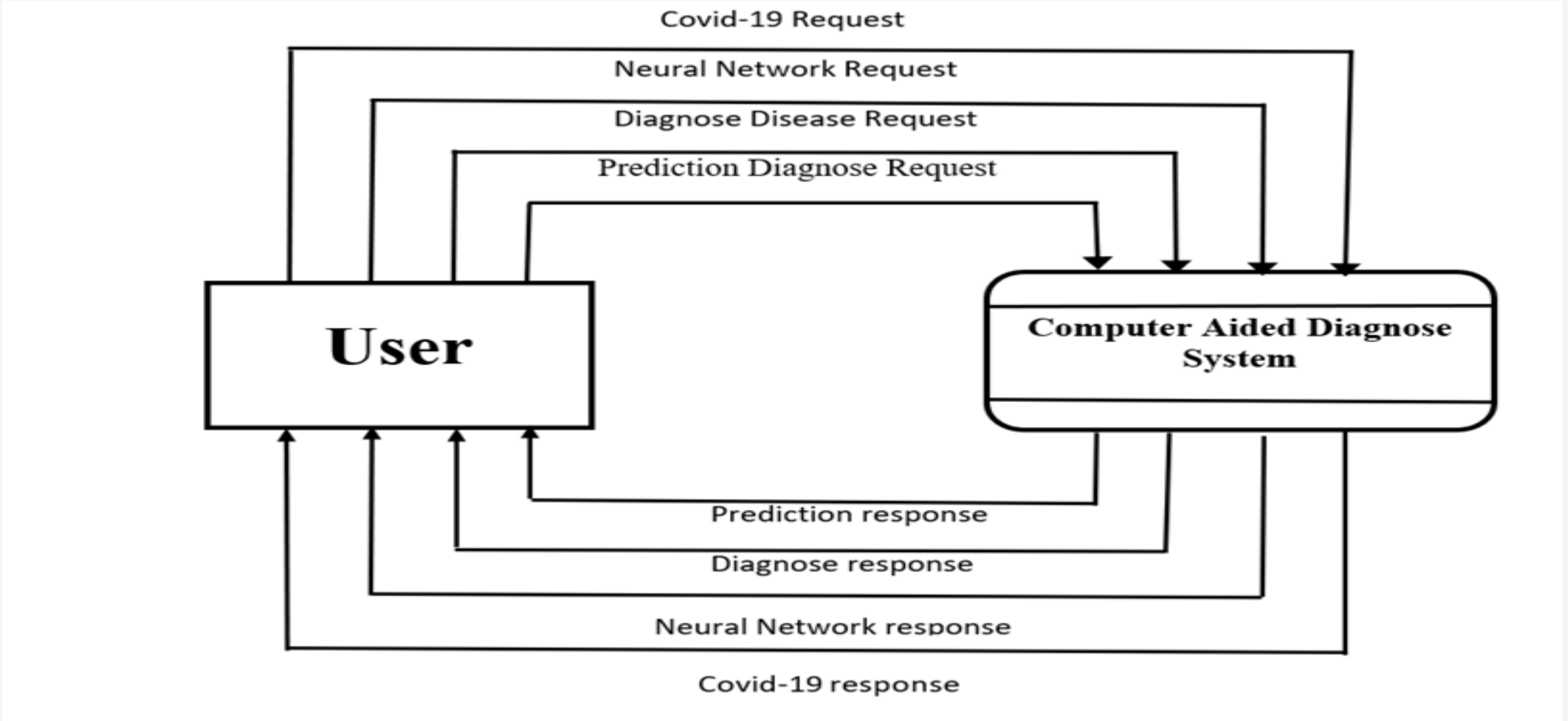
PROPOSED FRAMEWORK

- The proposed framework involved five phases- Data Collection, Data Pre-processing, Model Training, Predictions, Validation.

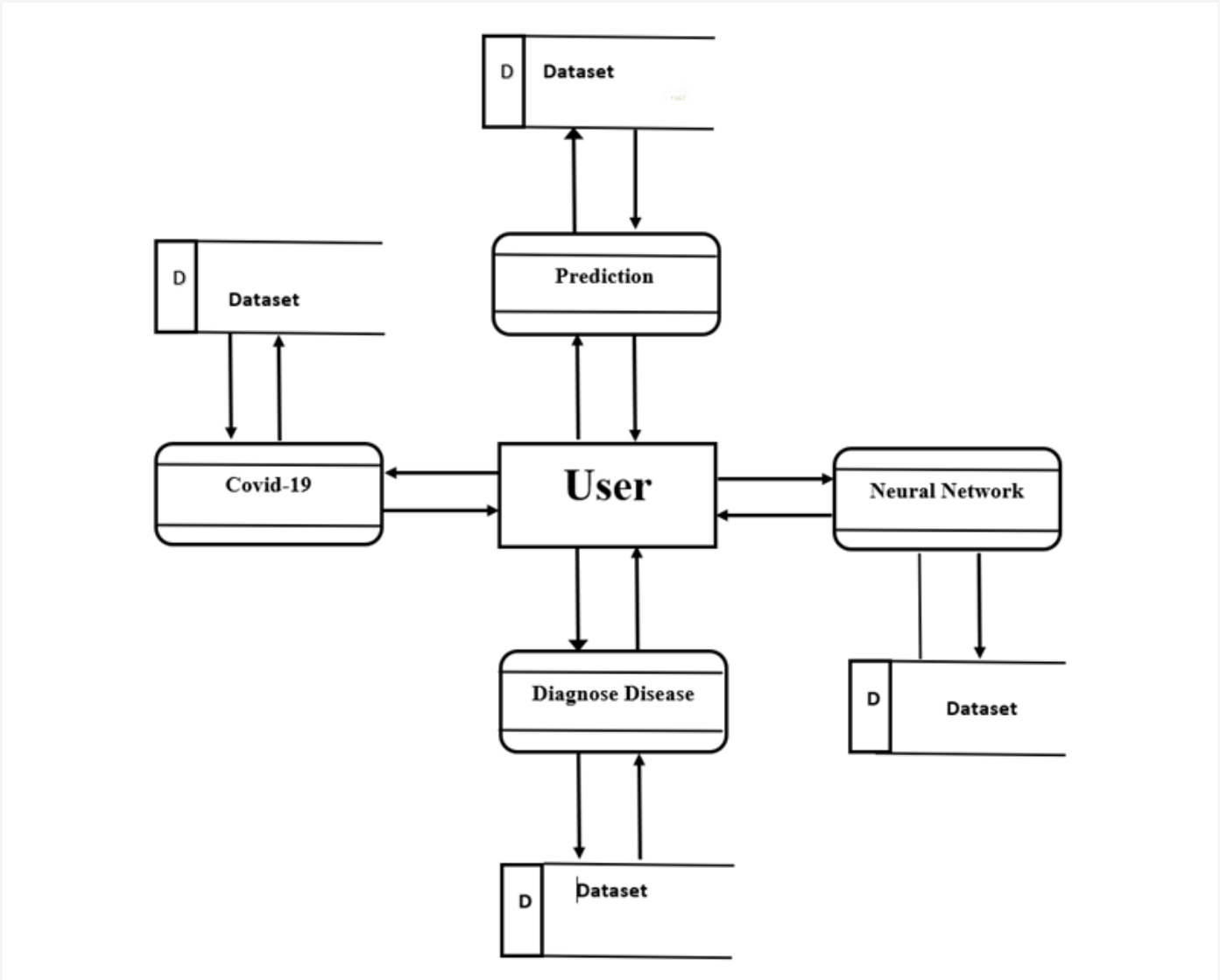


Prediction Workflow

DFD level 0


Covid-19 Request
Neural Network Request
Diagnose Disease Request
Prediction Diagnose Request
User
Computer Aided Diagnose System
Prediction response
Diagnose response
Neural Network response
Covid-19 response

DFD level 1


D
Dataset
Prediction
D
Dataset
Covid-19
User
Neural Network
Diagnose Disease
D
Dataset
D
Dataset

Software and Hardware Requirement

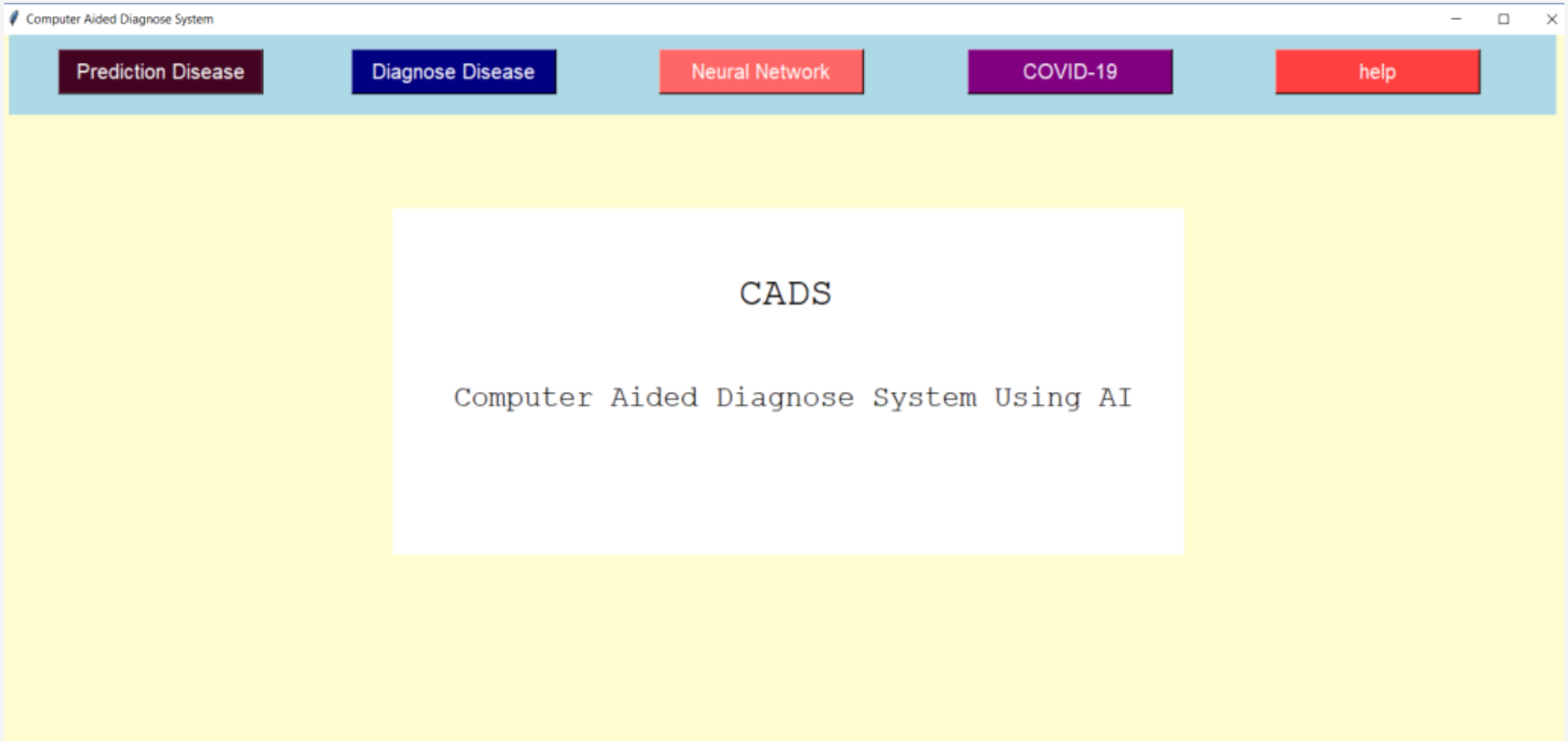
Software

Operating System	**:** Windows* 7 or later, macOS, and Linux
Developmental Tool	: Anaconda (Spyder)
Programming Language	: Python * versions (3.6.X)
Python packages	: NumPy, scikit-learn*, pandas, Matplotlib

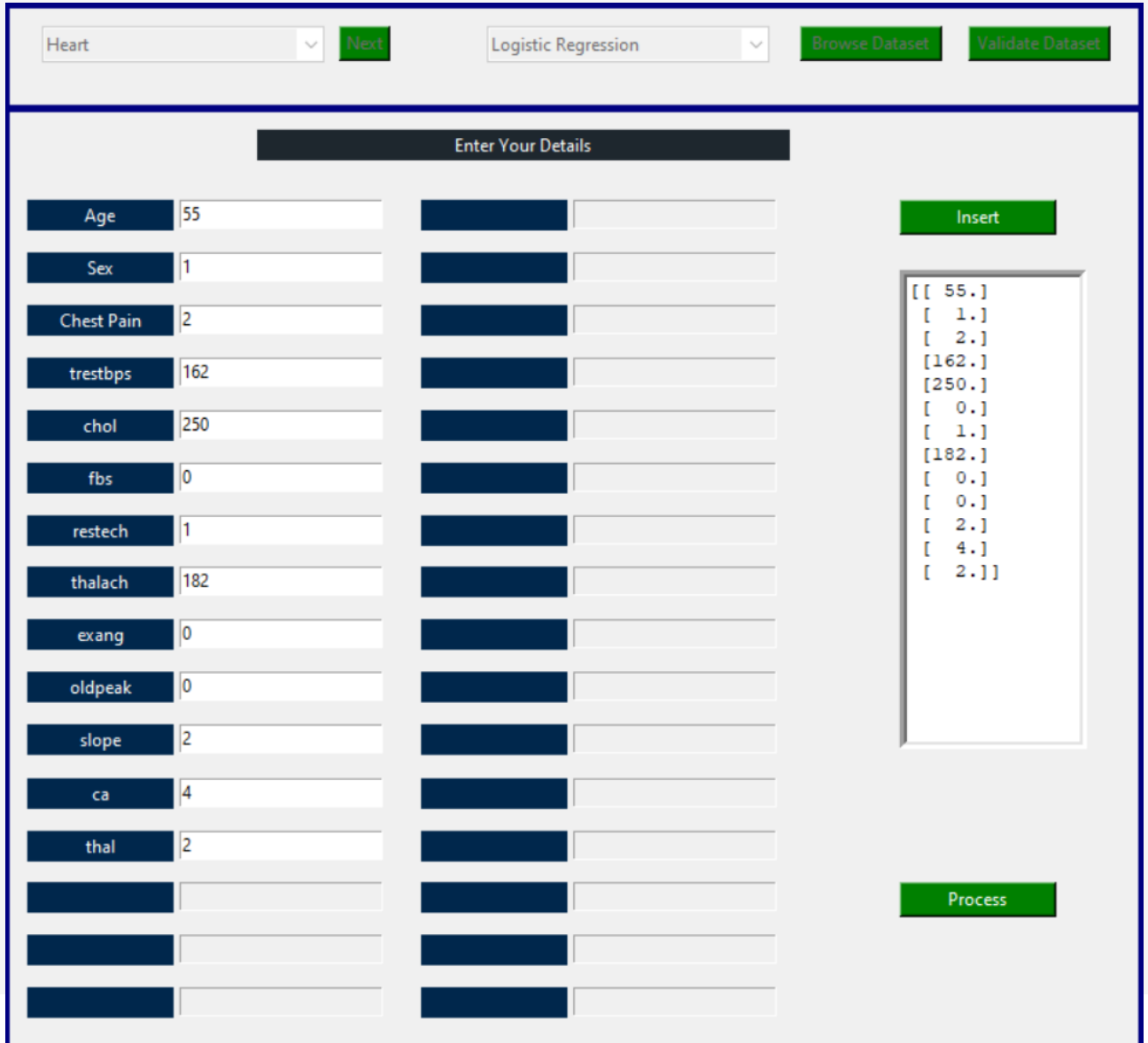
Hardware

Processor	: Intel® Core™ i5 processor
RAM	: 8 GB
HDD	: 100 GB

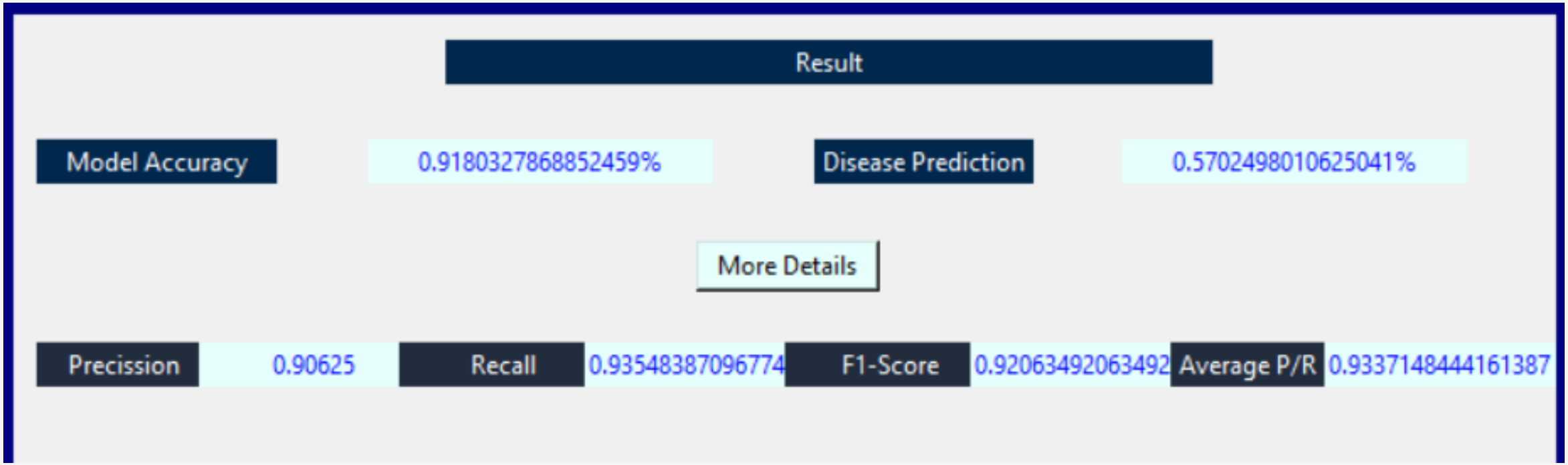
PROJECT DESCRIPTION



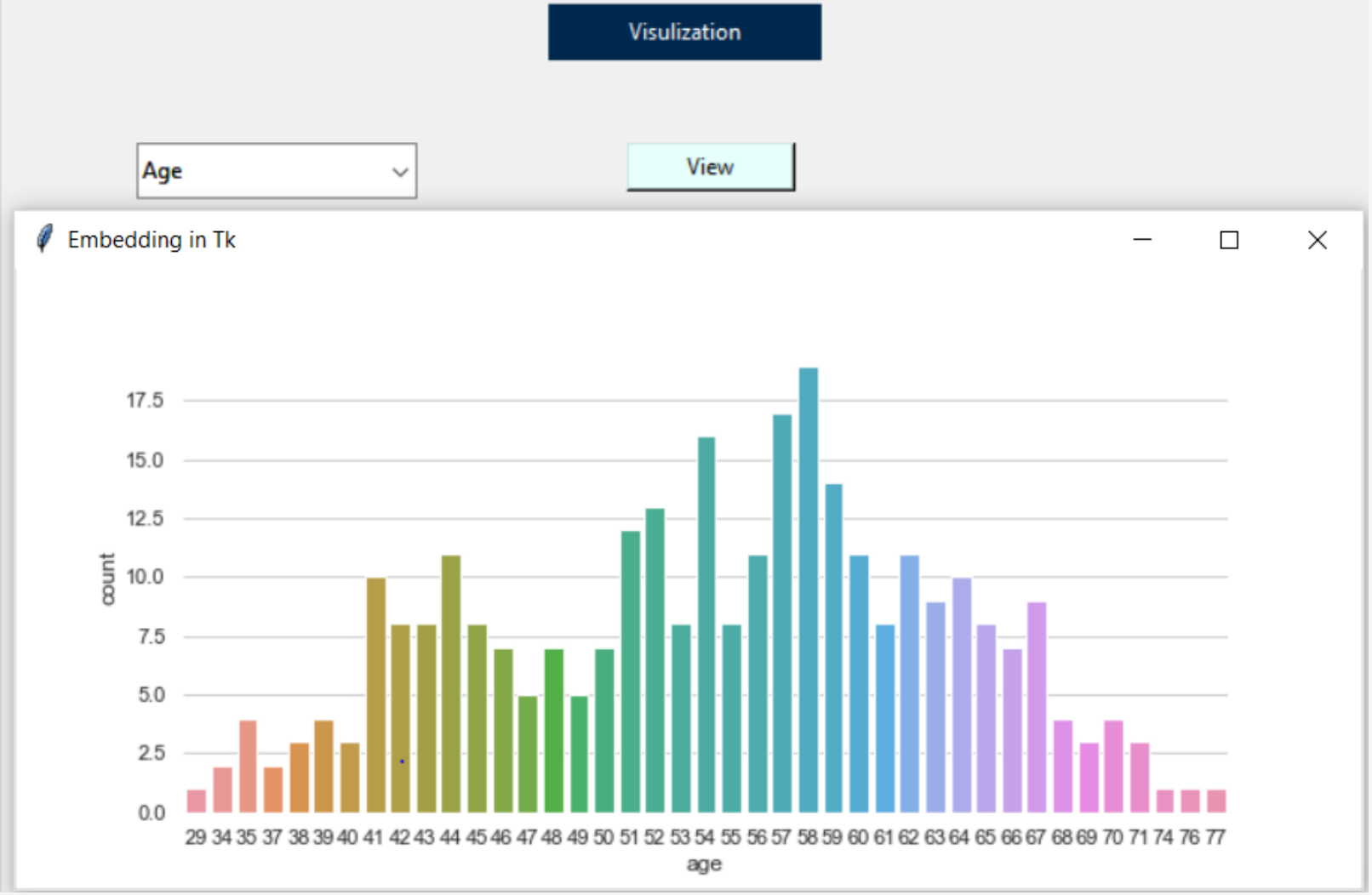
Prediction Disease: Input:

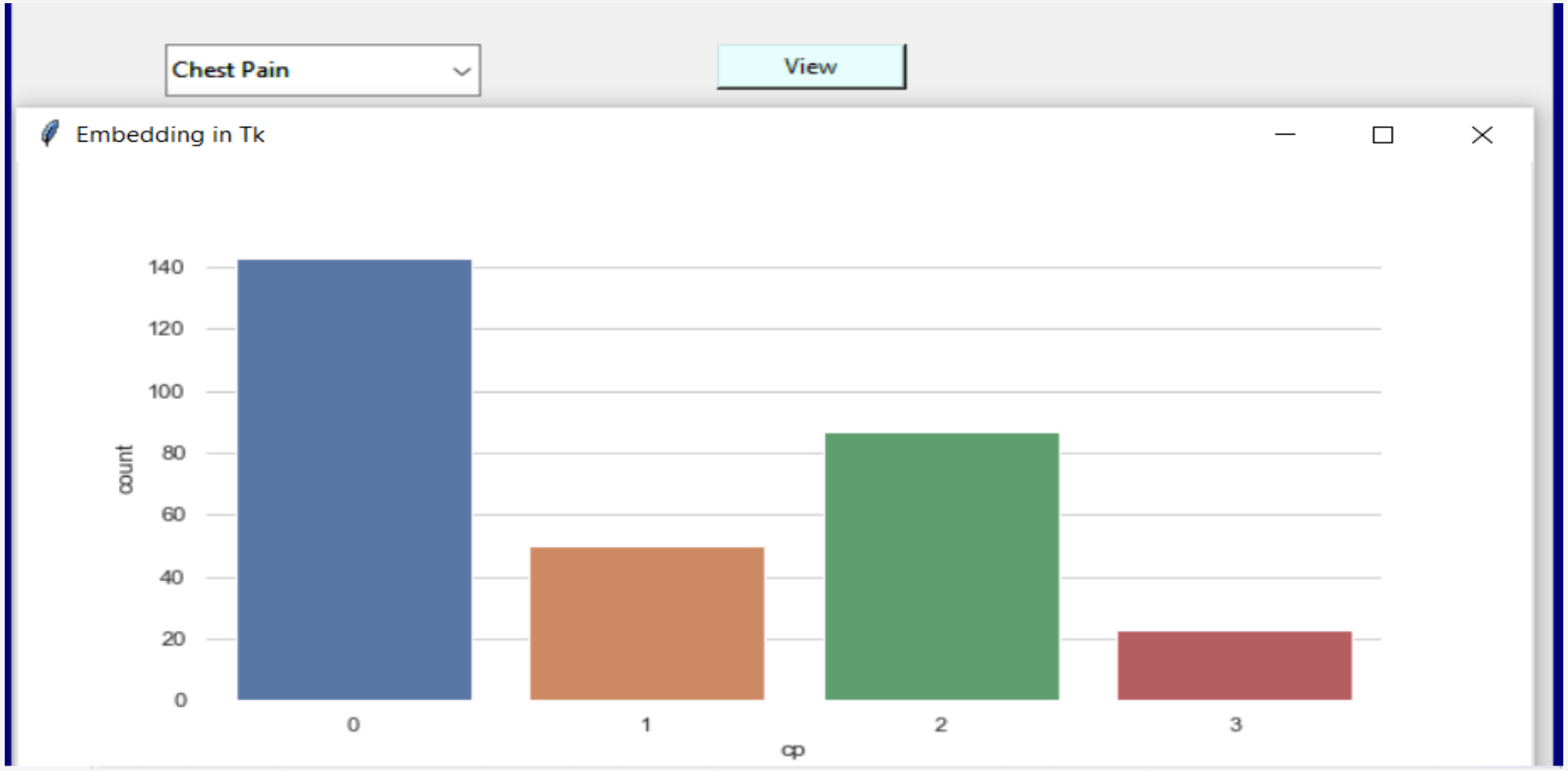


Output:

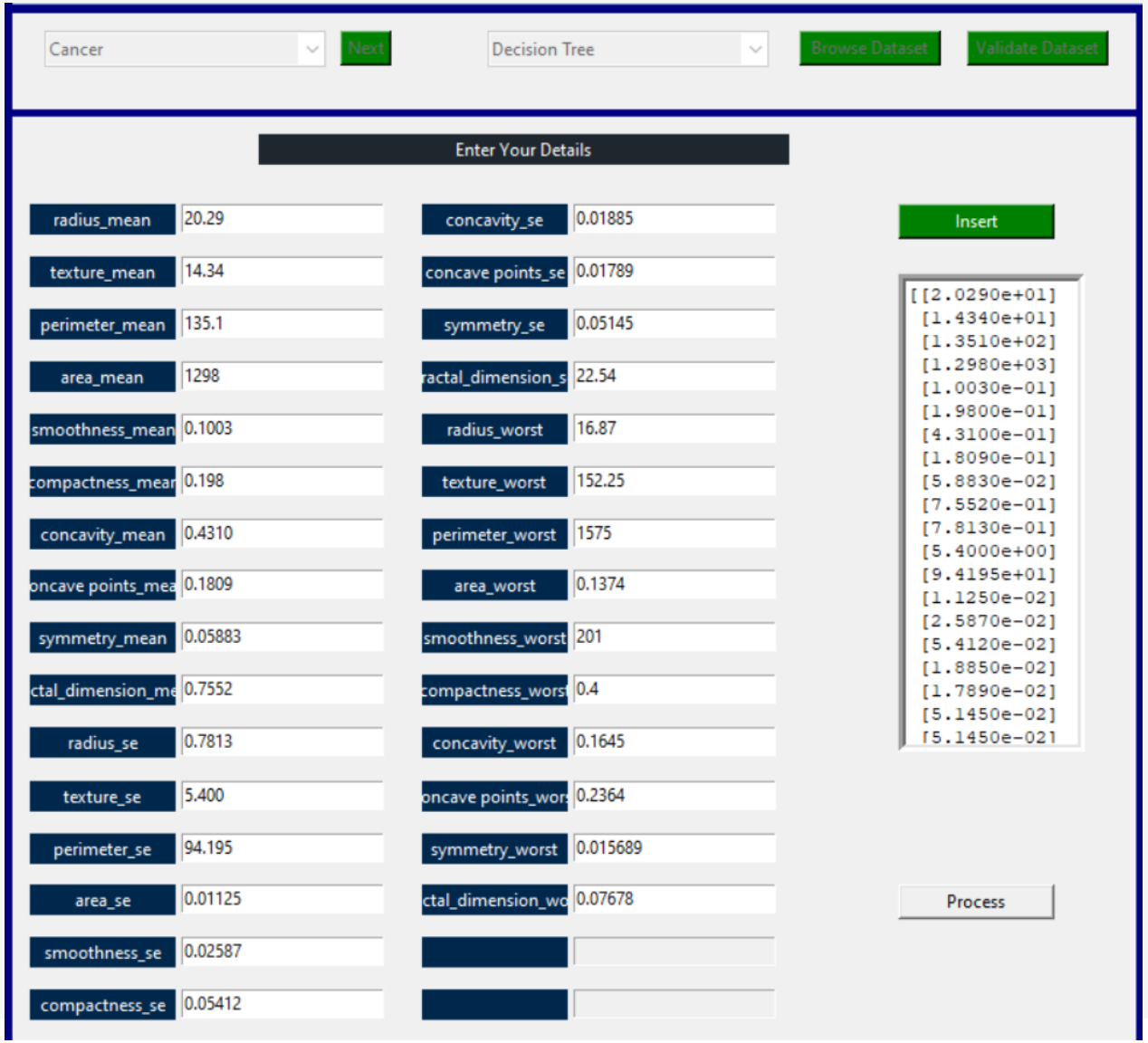

Result
Model Accuracy
0.91803278688524 59%
Disease Prediction
0.5702498010625041%
More Details
Precission
0.90625
Recall
0.935483870967 74
F1-Score
0.9206349206349 2
Average P/R
0.9337148444161387

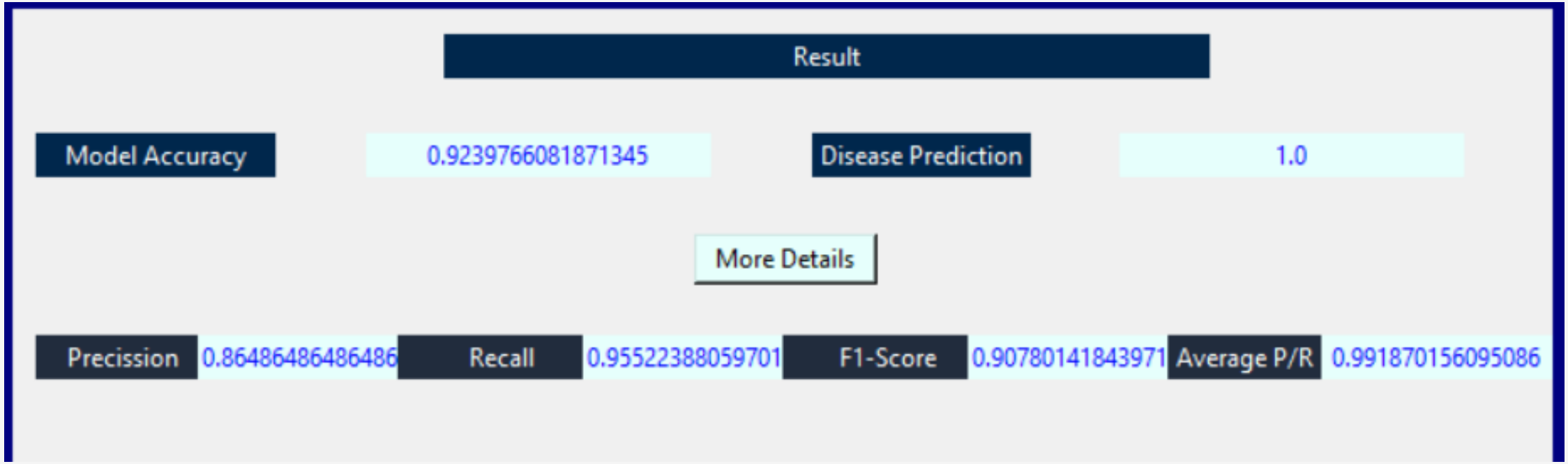
Graphs:


Visulization
Age
View
Embedding in Tk
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
count
29 34 35 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 74 76 77
age


Chest Pain
View
Embedding in Tk
140
120
100
80
60
40
20
0
count
0
1
2
3
cp

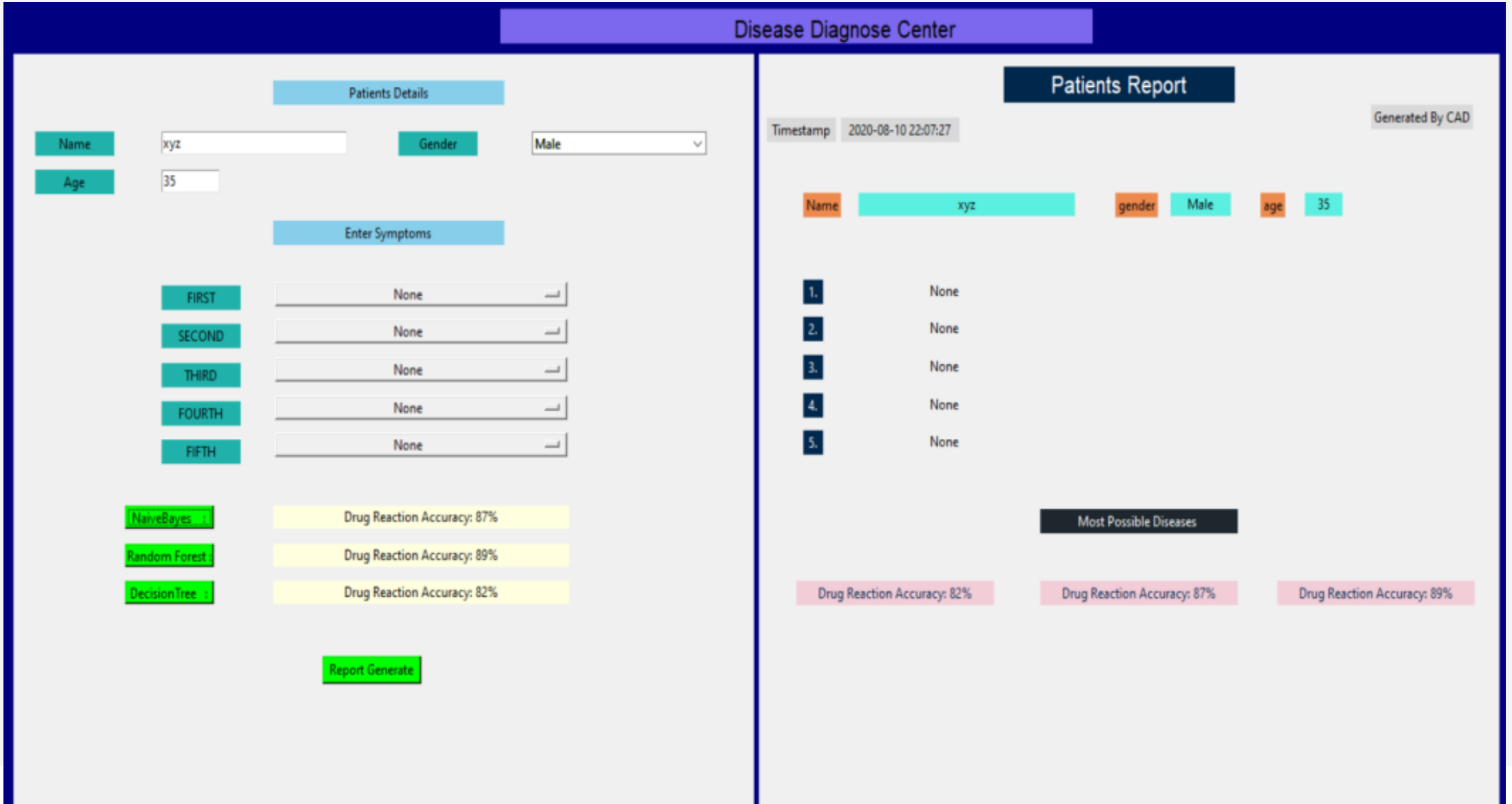
Cancer:Input:



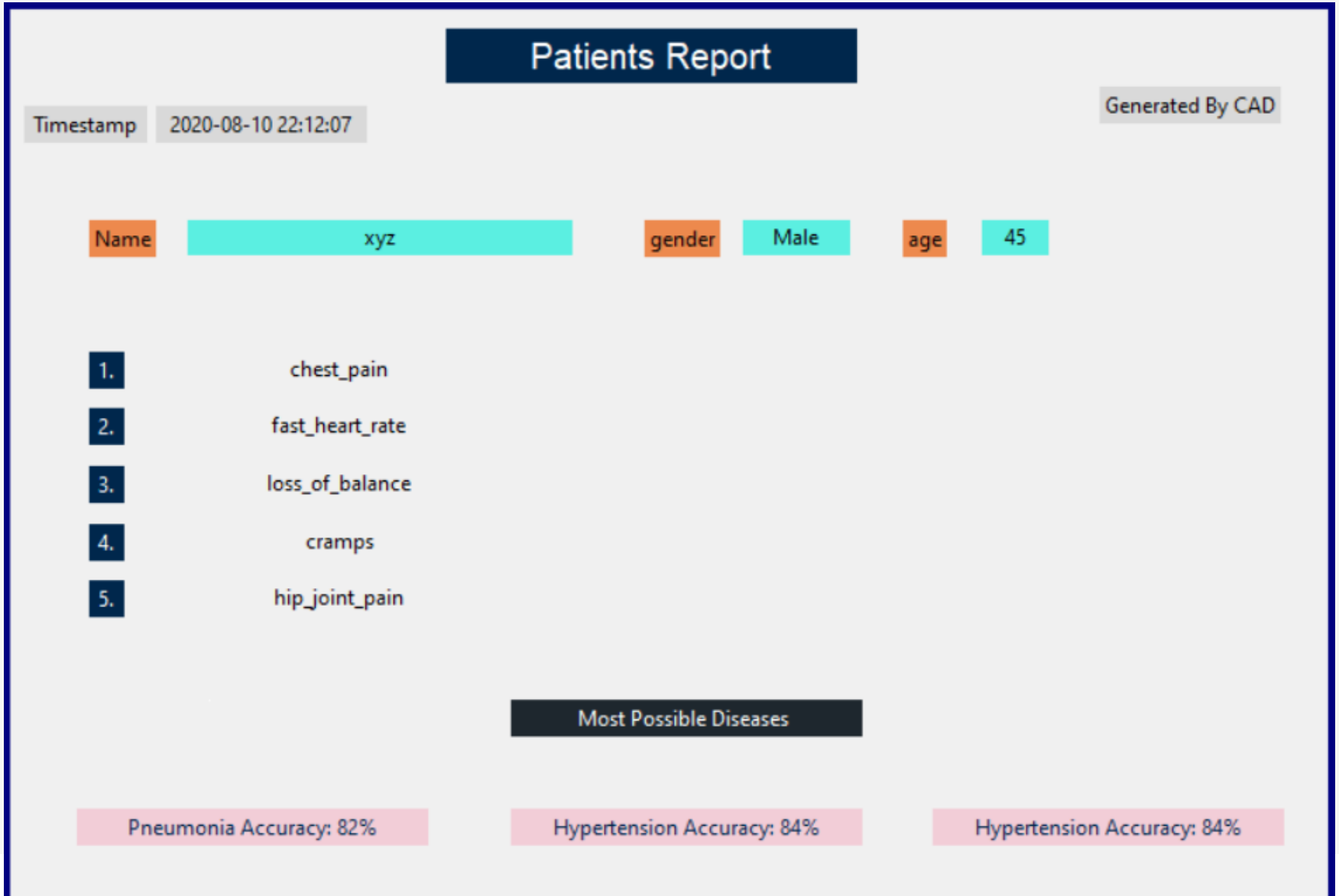

Result
Model Accuracy
0.9239766081871345
Disease Prediction
1.0
More Details
Precission
0.86486486486486
Recall
0.95522388059701
F1-Score
0.90780141843971
Average P/R
0.991870156095086

Diagnose Disease:

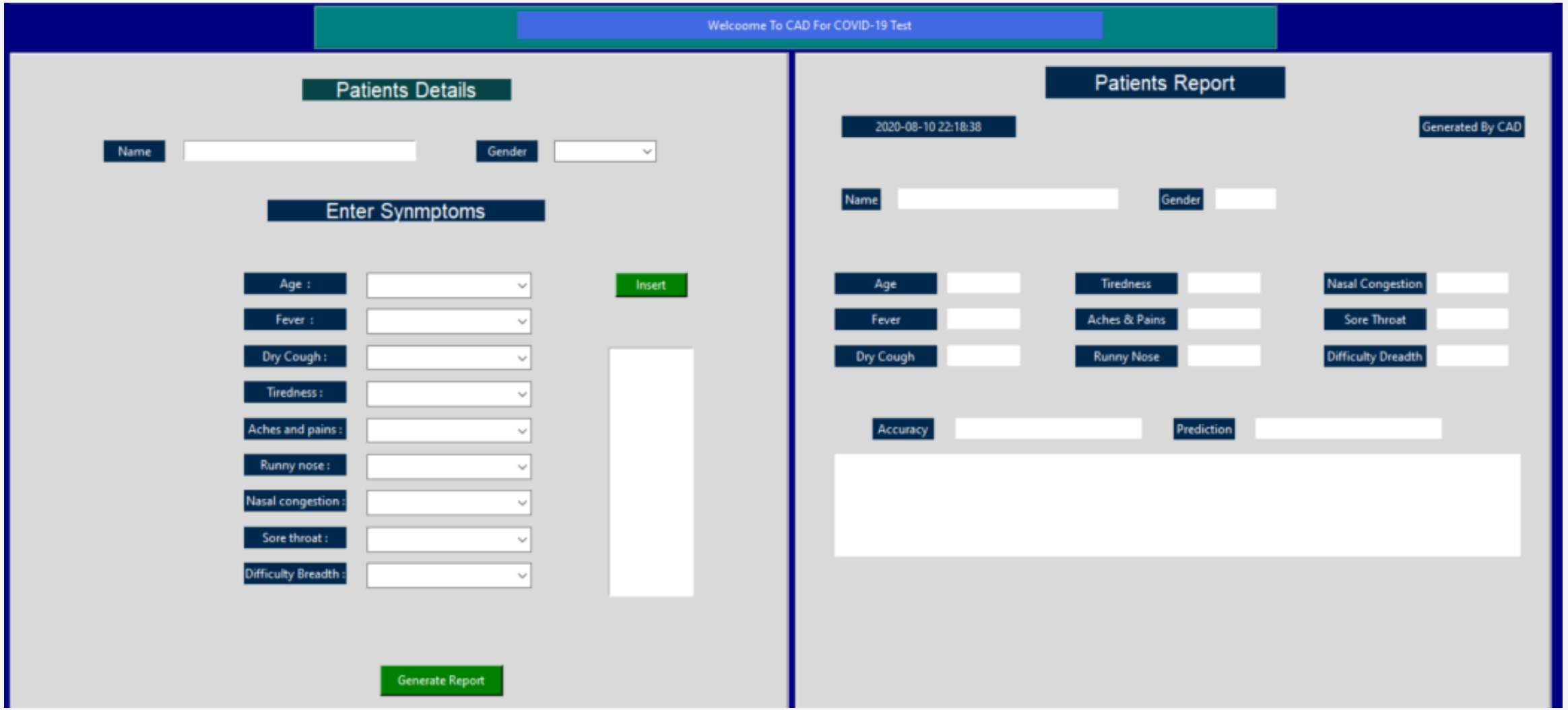
Input(without any symptoms):



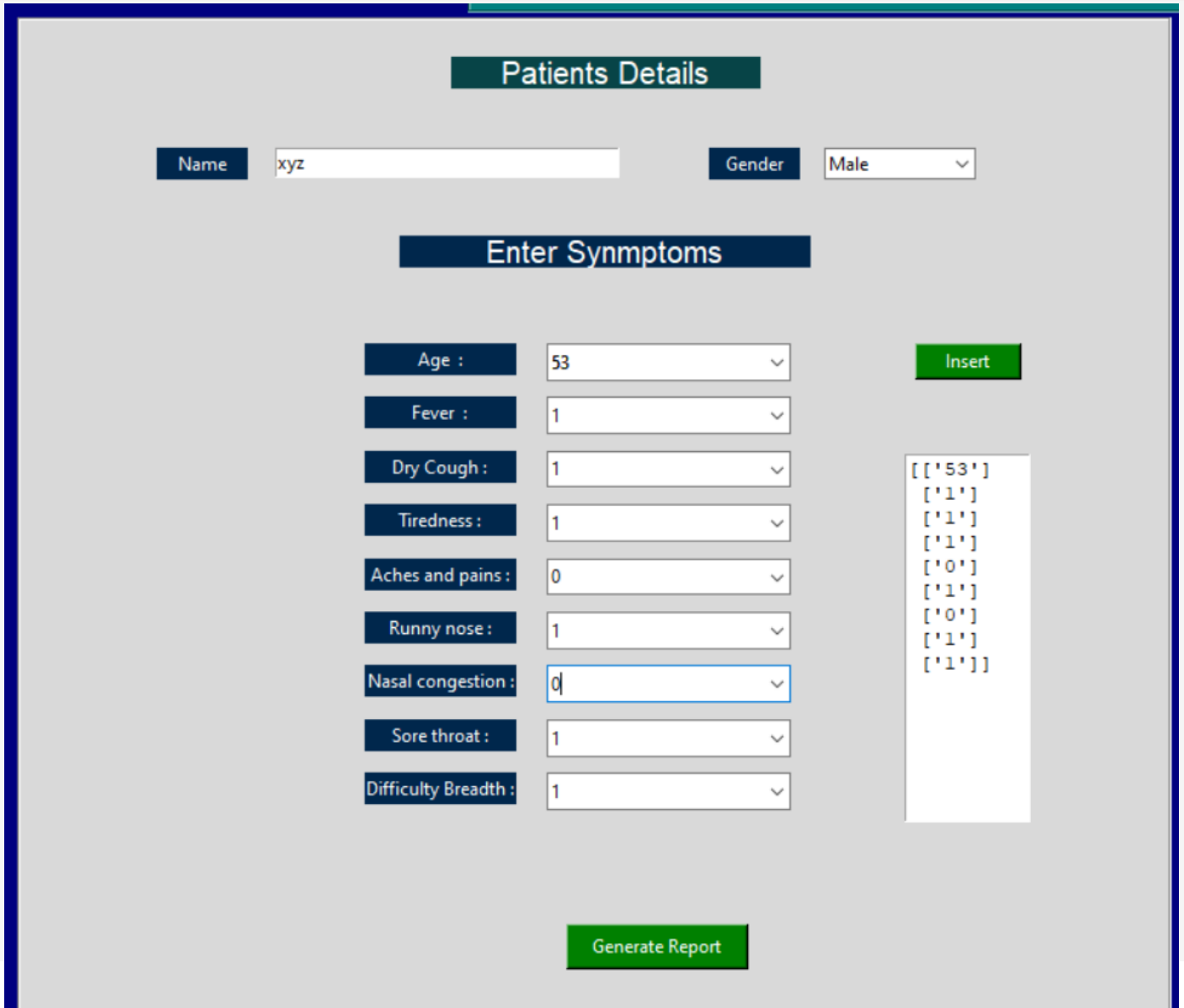
With User Synmptoms:

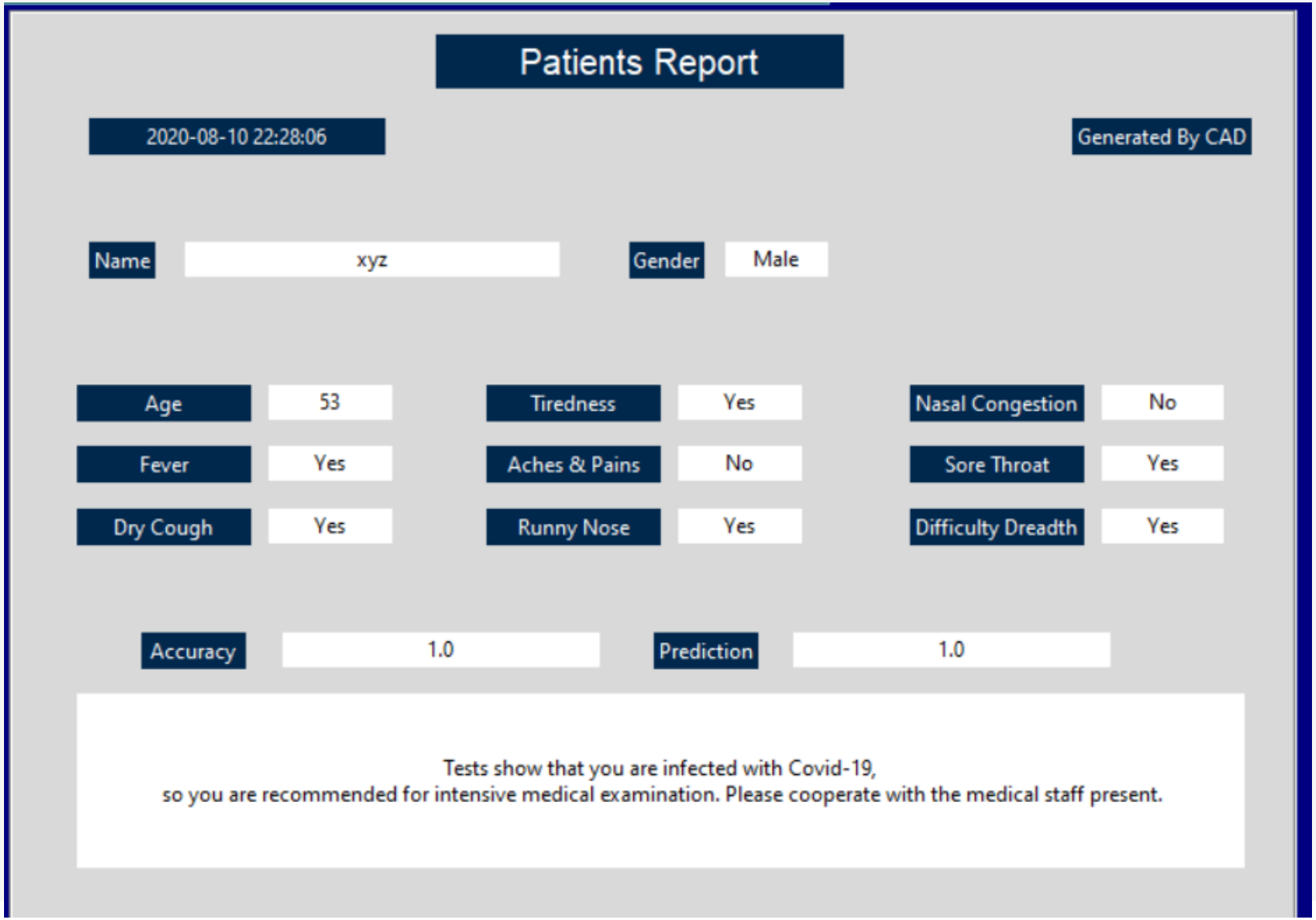

Patients Report
Timestamp
2020-08-10 22:12:07
Generated By CAD
Name
xyz
gender
Male
age
45
1.
chest_pain
2.
fast_heart_rate
3.
loss_of_balance
4.
cramps
5.
hip_joint_pain
Most Possible Diseases
Pneumonia Accuracy: 82%
Hypertension Accuracy: 84%
Hypertension Accuracy: 84%

Covid-19:

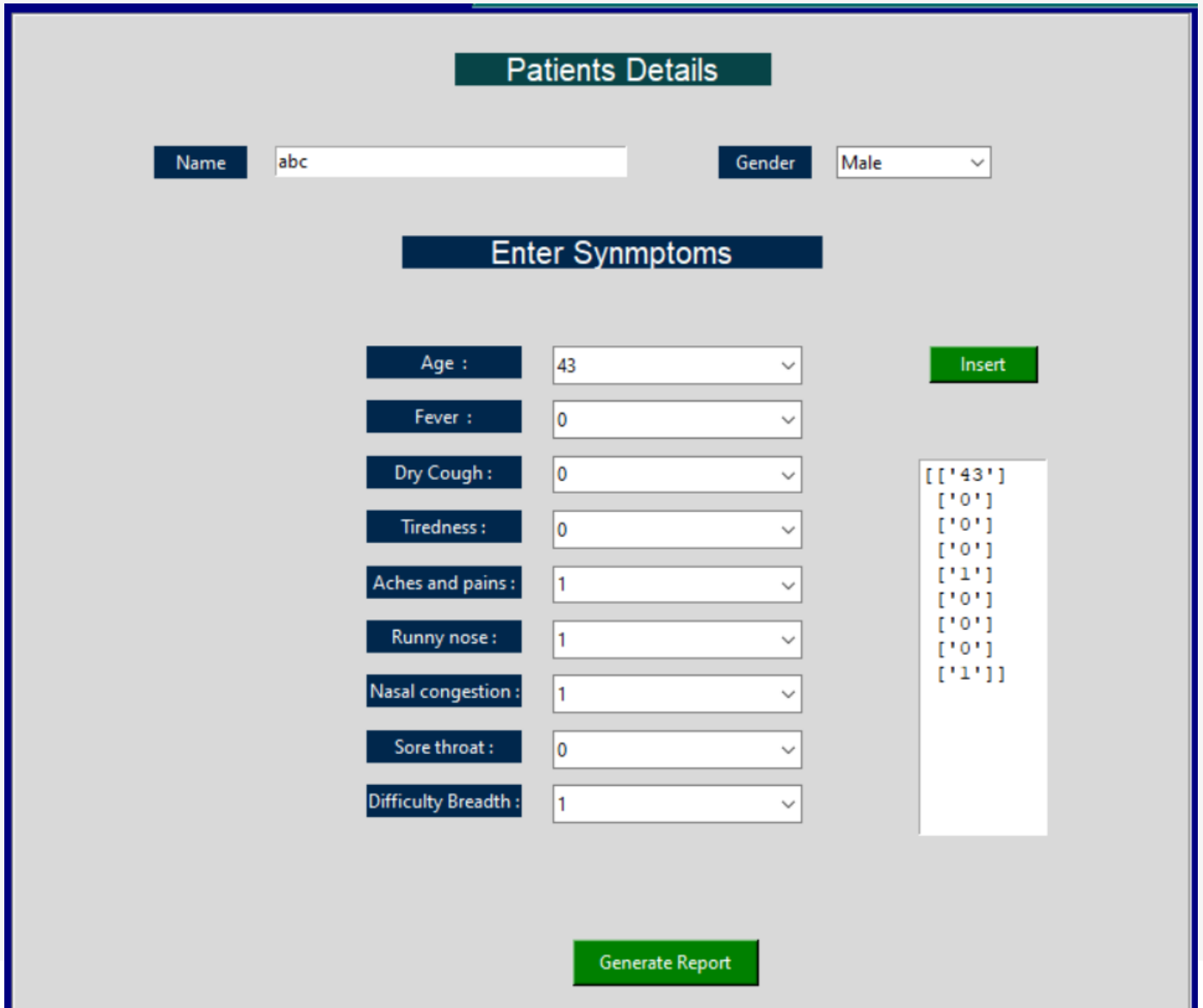

Welcome To CAD For COVID-19 Test
Patients Details
Name
Gender
Enter Synmptoms
Age :
Fever :
Dry Cough :
Tiredness :
Aches and pains :
Runny nose :
Nasal congestion :
Sore throat :
Difficulty Breadth :
Insert
Generate Report
Patients Report
2020-08-10 22:18:38
Generated By CAD
Name
Gender
Age
Fever
Dry Cough
Tiredness
Aches & Pains
Runny Nose
Nasal Congestion
Sore Throat
Difficulty Dreadth
Accuracy
Prediction

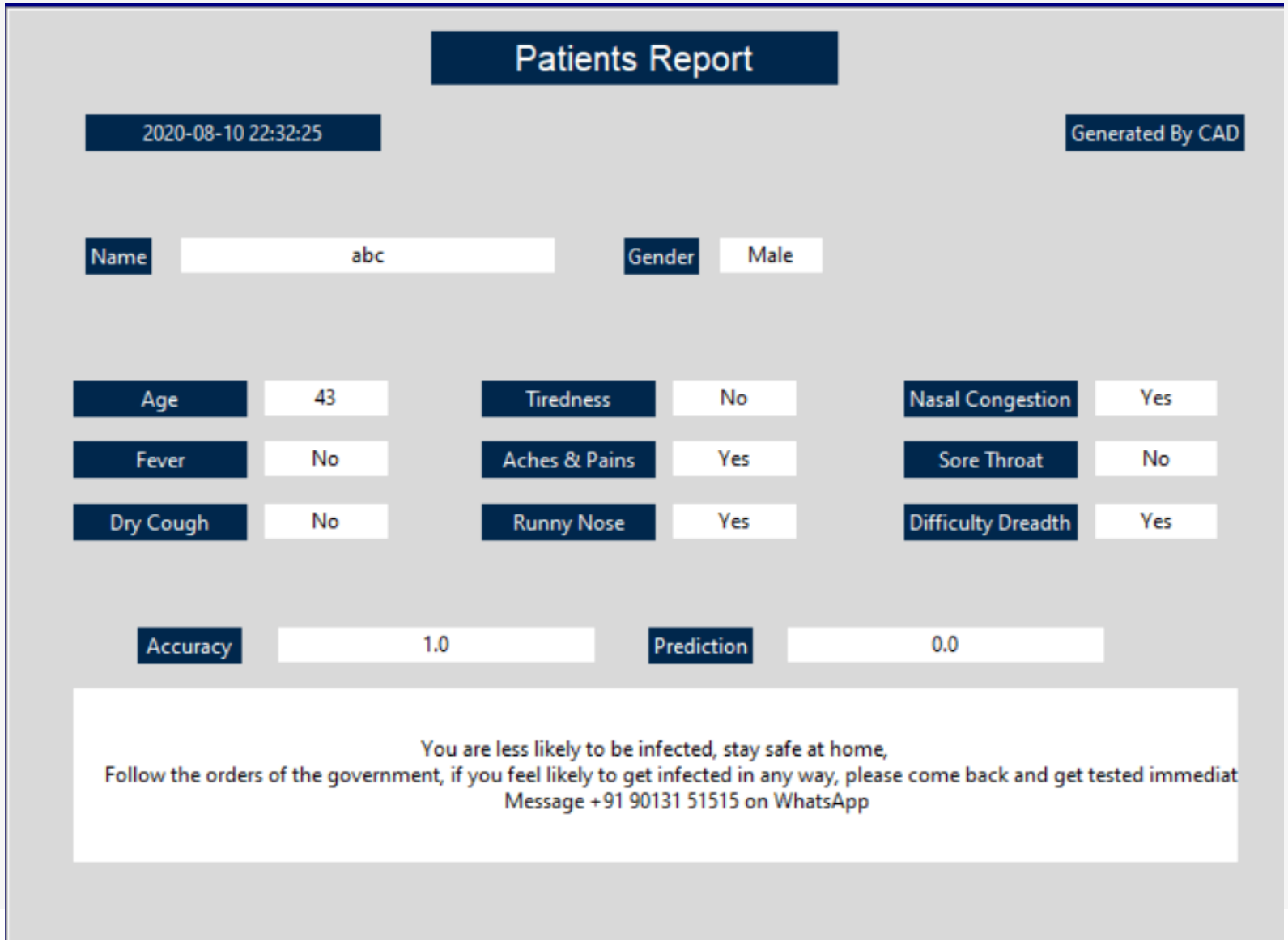
User Input(infected):



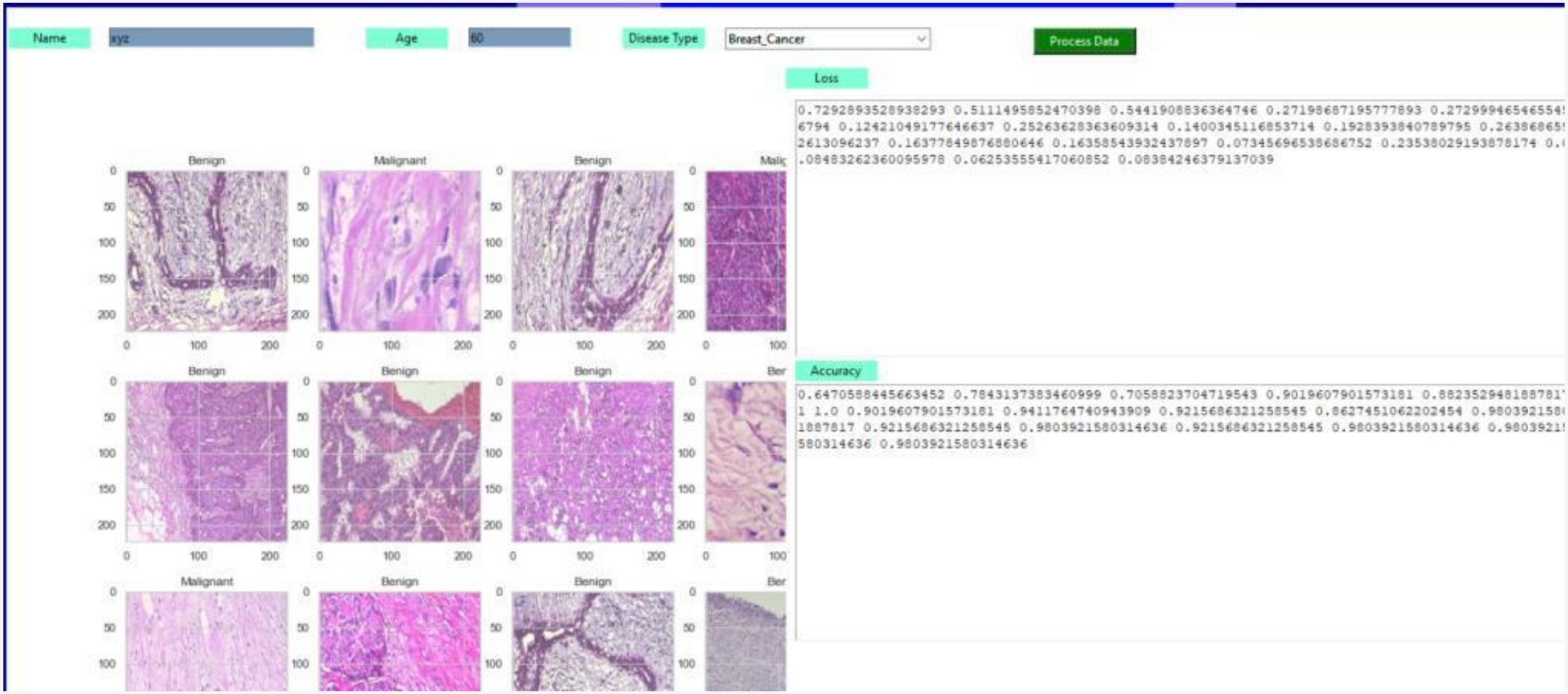

Patients Report
2020-08-10 22:28:06
Generated By CAD
Name
xyz
Gender
Male
Age
53
Tiredness
Yes
Nasal Congestion
No
Fever
Yes
Aches & Pains
No
Sore Throat
Yes
Dry Cough
Yes
Runny Nose
Yes
Difficulty Dreadth
Yes
Accuracy
1.0
Prediction
1.0
Tests show that you are infected with Covid-19,
so you are recommended for intensive medical examination. Please cooperate with the medical staff present.

Uninfected:


Patients Details
Name
abc
Gender
Male
Enter Synmptoms
Age :
43
Insert
Fever :
0
Dry Cough :
0
Tiredness :
0
Aches and pains :
1
Runny nose :
1
Nasal congestion :
1
Sore throat :
0
Difficulty Breadth :
1
[['43']
['0']
['0']
['0']
['1']
['0']
['0']
['0']
['1']]
Generate Report


Patients Report
2020-08-10 22:32:25
Generated By CAD
Name
abc
Gender
Male
Age
43
Tiredness
No
Nasal Congestion
Yes
Fever
No
Aches & Pains
Yes
Sore Throat
No
Dry Cough
No
Runny Nose
Yes
Difficulty Dreadth
Yes
Accuracy
1.0
Prediction
0.0
You are less likely to be infected, stay safe at home,
Follow the orders of the government, if you feel likely to get infected in any way, please come back and get tested immediat
Message +91 90131 51515 on WhatsApp

Disease Image Processing CNN:



Prediction Disease | Diagnos Disease | Neural Network | COVID-19 | help

system software /hardware requirment: minimum 4 GB RAM ,8 GB recommended,python 3.7+ and supporting liberaries

Follow the steps given one by one for good user experience

After starting process wait for result may be it will take some time for completion

After starting process wait for result may be it will take some time for completion

In time of Image classification it will take time according to data so you have to be patient about that.

ex if you using 32 images for training and validation each then it may take time according to your system specification.

CONCLUSION

Machine learning is an interdisciplinary field combining computer science and mathematics to develop model with the *intent of delivering maximal predictive accuracy results*. after completion of this project we can say that we find satisfactory results from this Software. Death due to Heart Disease & Breast Cancer can be reduced. People who are suffering from these disease and who treat this disease will benefit from it.

FUTURE SCOPE OF PROJECT

Nothing is perfect in this world. So, we are also no exception. Although, we have tried our best to present the information effectively, yet, there can be further enhancement in the Application.

We know that the way human is messing with nature, in the result of that, we have to keep on facing new diseases, as we can see in the present, it would not be wrong to say that in future we will have to fight against such diseases. If it is necessary then why should we not be prepared by using the right place of technology.

There is also an option to convert it into mobile application and web application with image processing in future.

BIBLIOGRAPHY

Books :

Python -Martin C. Brown , “Software Engineering ” by Pankaj Jalote

Online Sources:

DOI: 10.1109/ITCE.2018.8327948
DOI: 10.1109/ITIME.2009.5236464
DOI: 10.1109/CEEC.2018.8674183
DOI: 10.1109/ITCE.2018.8327948
DOI: 10.1109/CEEC.2018.8674183 https://www.forbes.com/sites/nicolemartin1/2019/09/30/artificial-intelligence-is- being-used-to-diagnose-disease-and-design-new-drugs/#3162134b44db
https://www.mygreatlearning.com/blog/what-is-tensorflow-machine-learning- library-explained/#1
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1955762/
https://ieeexplore.ieee.org/document/4318787
https://ieeexplore.ieee.org/document/8075510
https://ieeexplore.ieee.org/document/5337226/
https://www.udemy.com/course/complete-python-bootcamp/learn/lecture/20447423#overview
https://www.coursera.org/learn/machine-learning?

THANK YOU