



Doctor Machine

(A Disease Prediction System Using Machine Learning)

Software Development Project-II

CSE 3200

PROJECT SUPERVISOR

Md.Abdus Salim Mollah

DEPT. OF CSE, KUET

DEVELOPED BY

Khandaker Tasnim Huq

Roll: 1307002

Md. Shakhawat Hossain Sajal

Roll: 1307057

LAYOUTS

- Introduction
- Motivation
- Objectives
- Working procedure
- Flow of Control
- Windows
- Advantages
- Limitation & Future plan
- Conclusion

Introduction

- An interactive Disease prediction system
- Based on basic Text Classification system
- A supervised learning task using four Learning Algorithms
- “Python 2.7” scripting language in windows 10 platform.

Motivation

- Interested in “Machine Learning”
- Inspired by a paper titled “Doctor Bayes” published on Machine Learning Summer project, 2015 from Stanford University [1][2]

Objectives

- To build an interactive disease classification application.
- To ease users finding out or predicting disease
- To recommend tests and diagnosis
- To prescribe medicines and treatments

Why We Used Python?

- Simple, elegant, consistent, mathlike
- Popular for applied data science and Machine learning
- “scikit-learn” most enriched, easiest and most advanced library
- NLTK (Natural Language Toolkit) & Matplotlib

Working Procedure

The whole project can be divided into two main period :

- Training period
- Test or Prediction period

Flow of Control

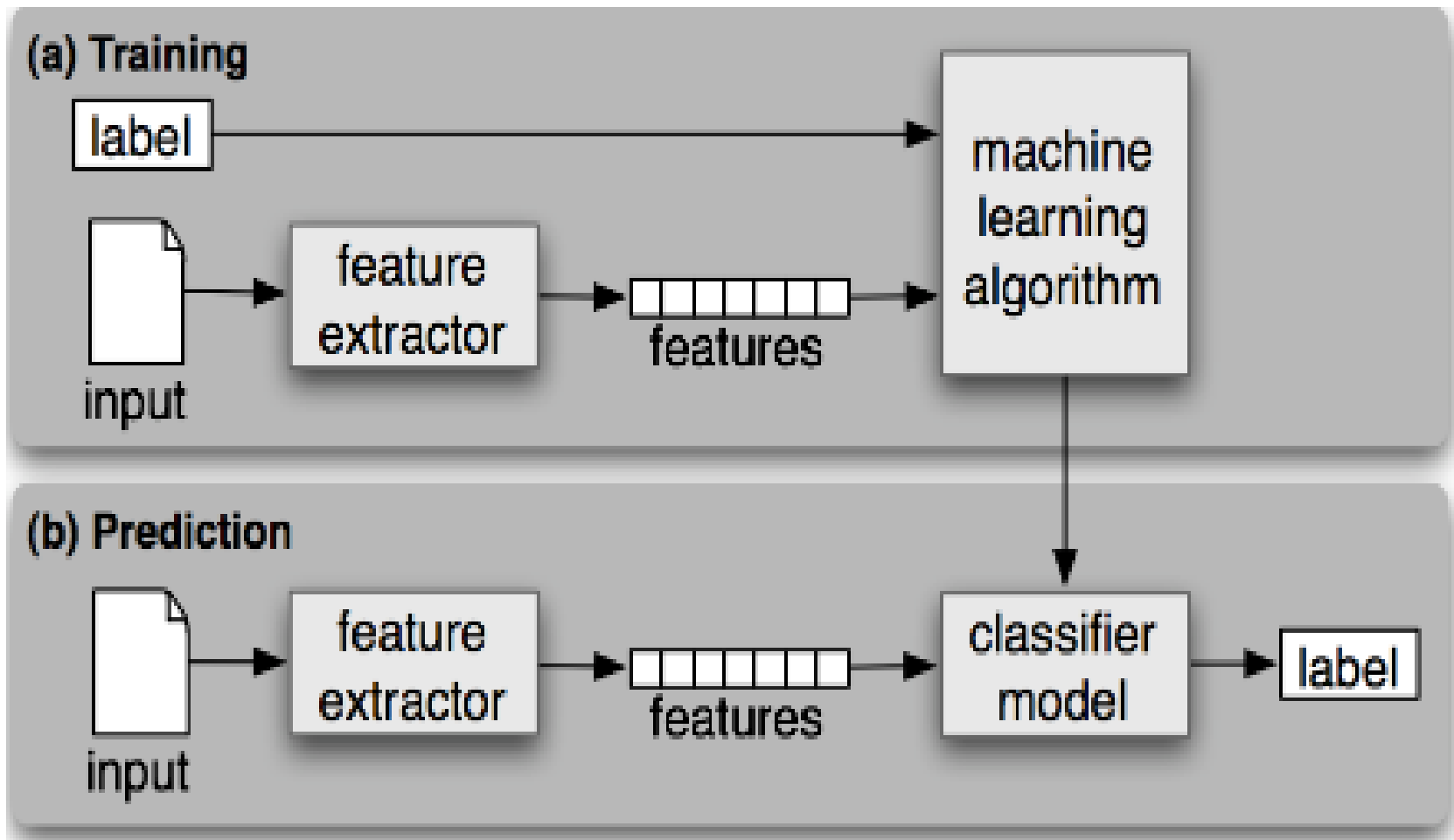


Figure 1: Supervised Classification

Text Data

Stemming

Punctuation Removal

Newline Removal

Stop Words Removal

Tf-Idf Weighting

Features

**Feature
Extraction**

Figure 2:
Function inside
Feature Extractor

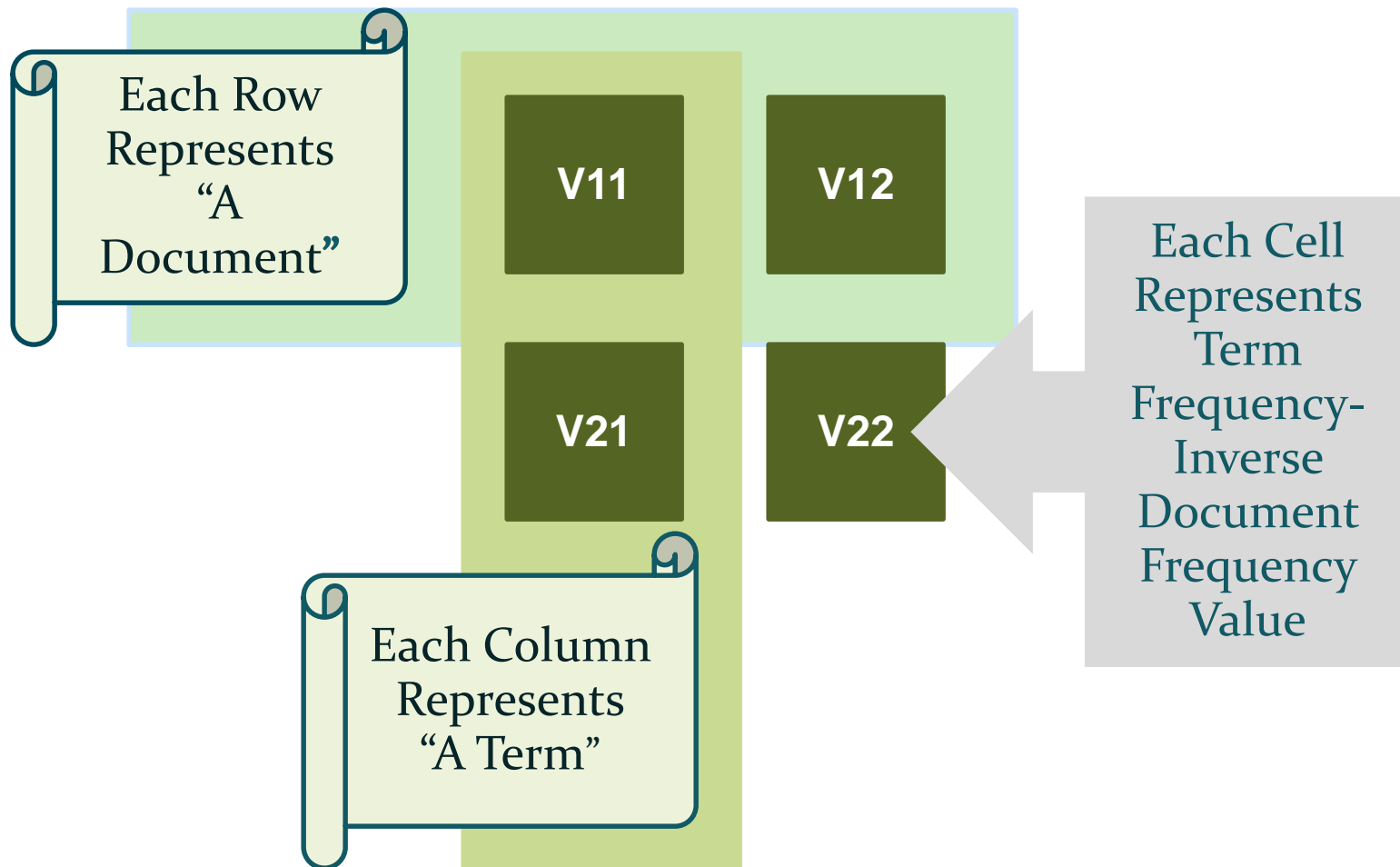


Figure 3: Document-Term Matrix

Algorithms

- Naive Bayes
- Multinomial Naive Bayes
- Logistic Regression
- One Vs Rest Classifier on Linear SVC

System Design

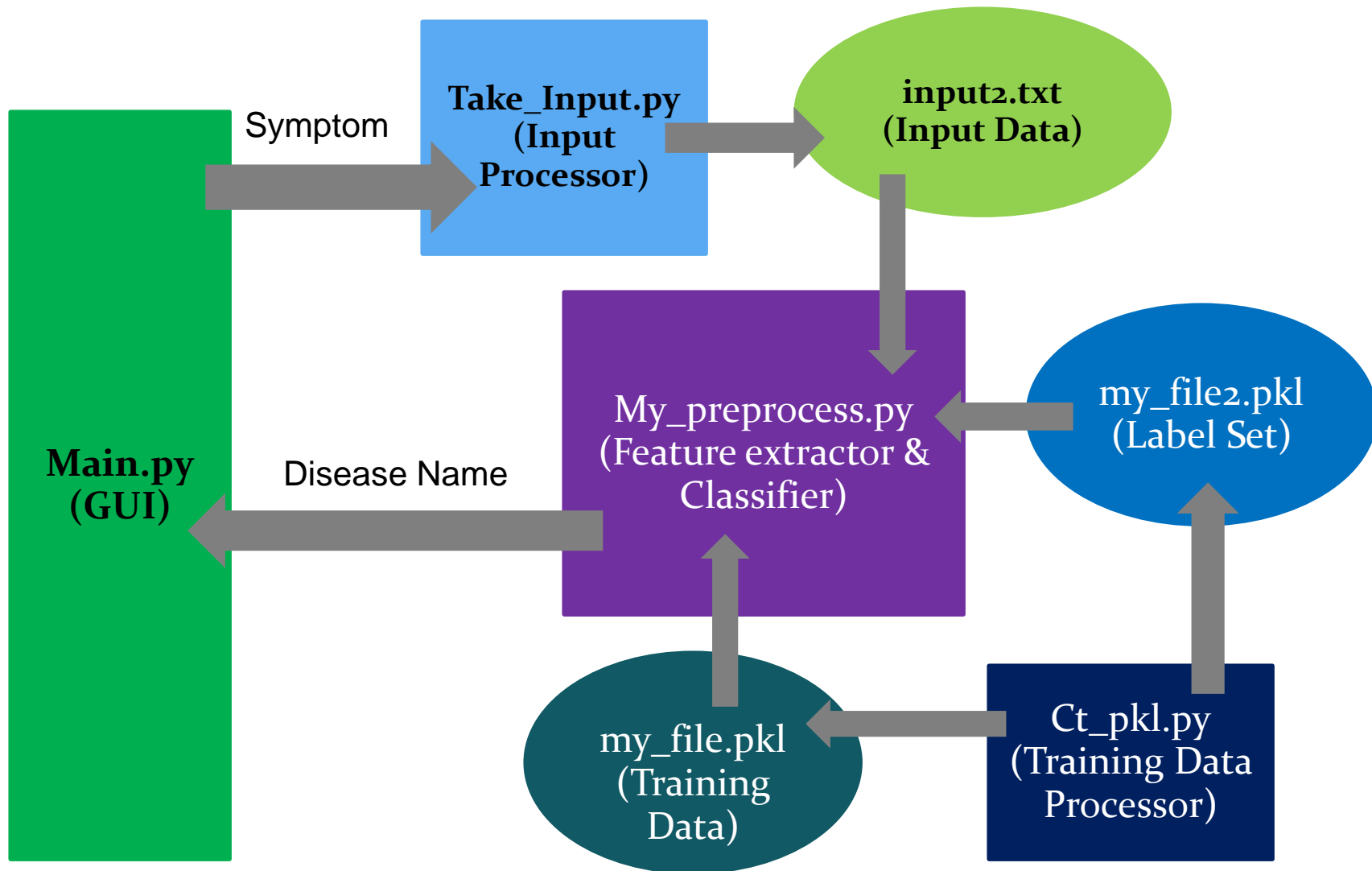
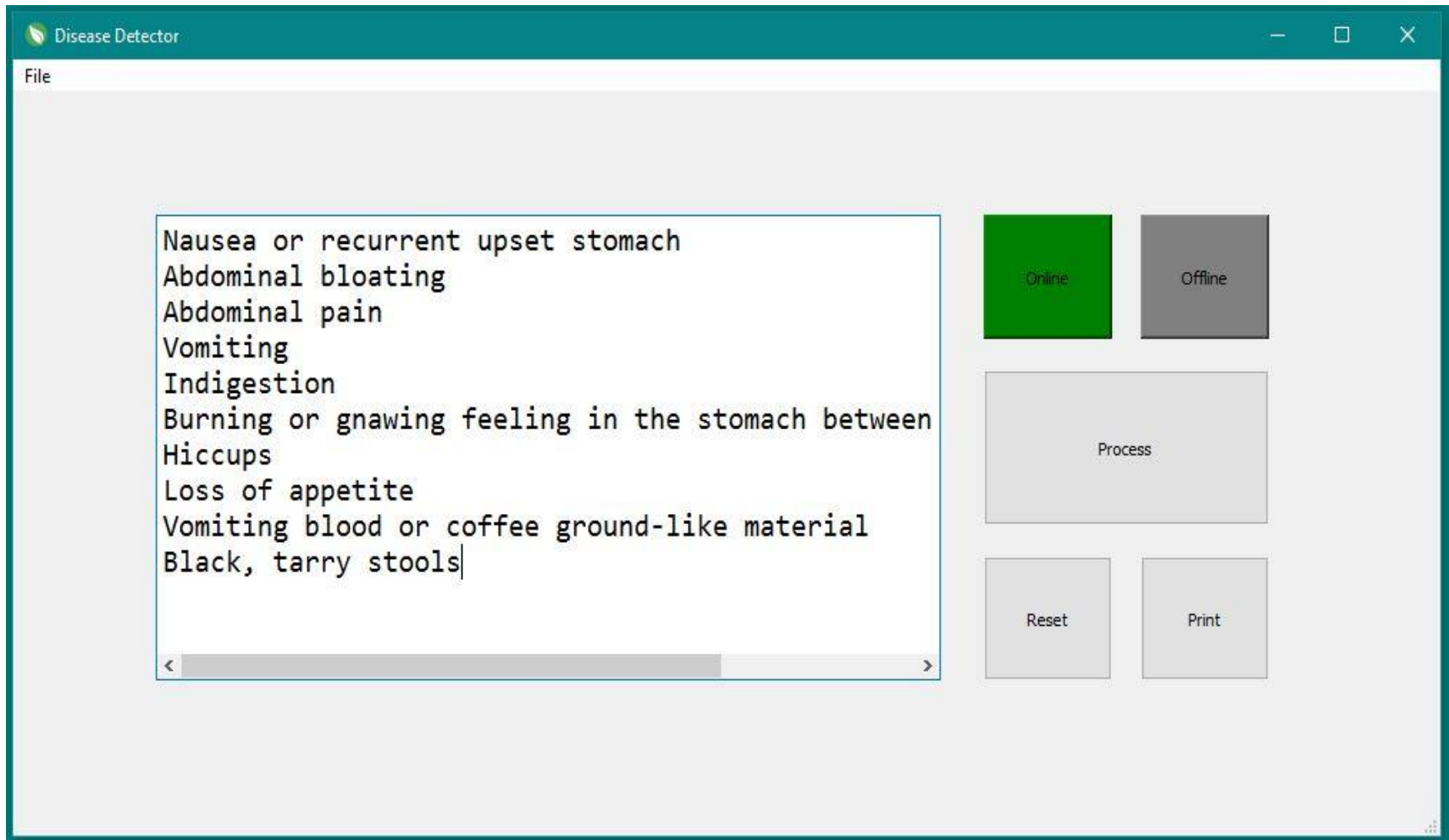
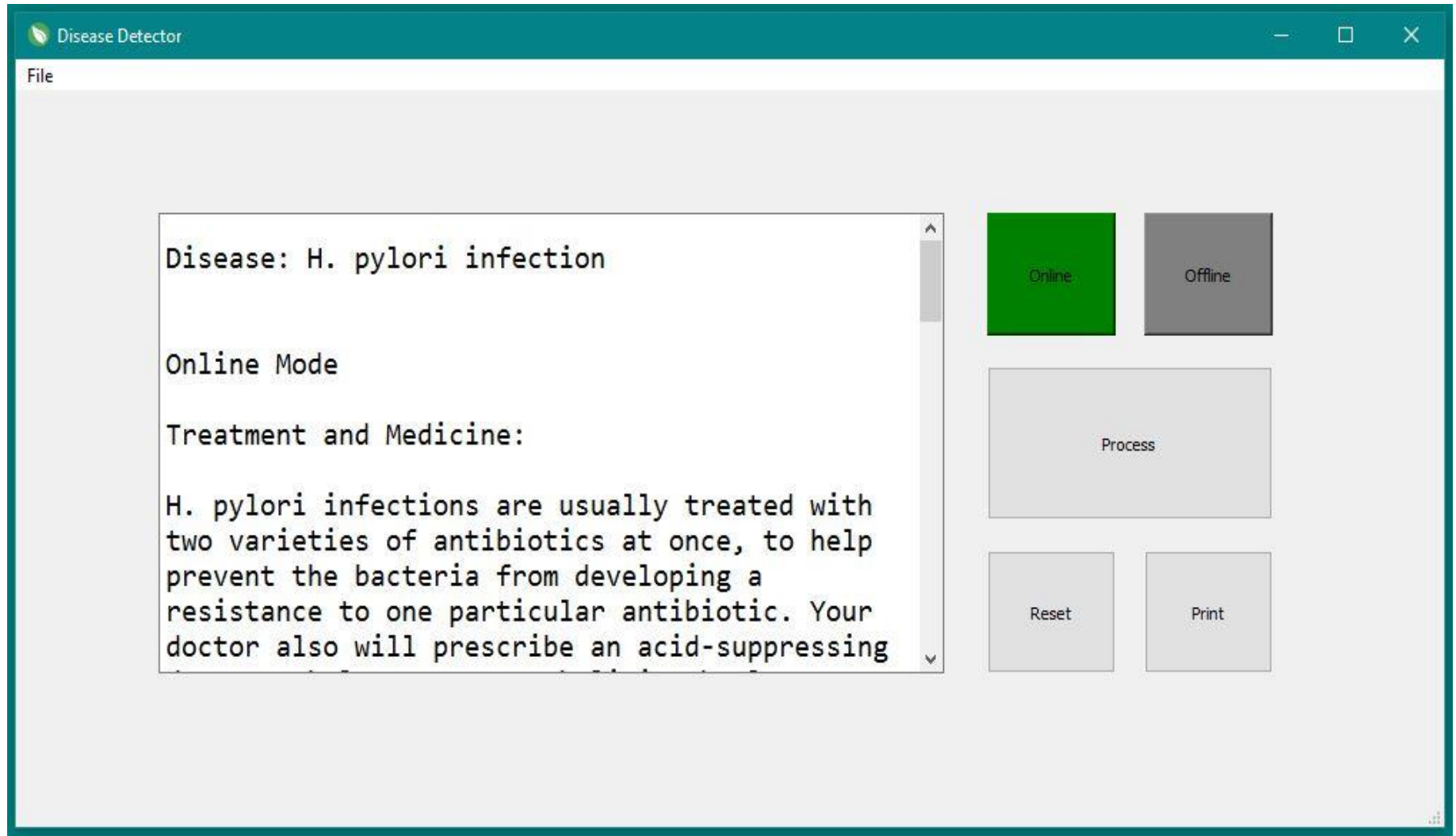


Figure 4: How System Be Designed

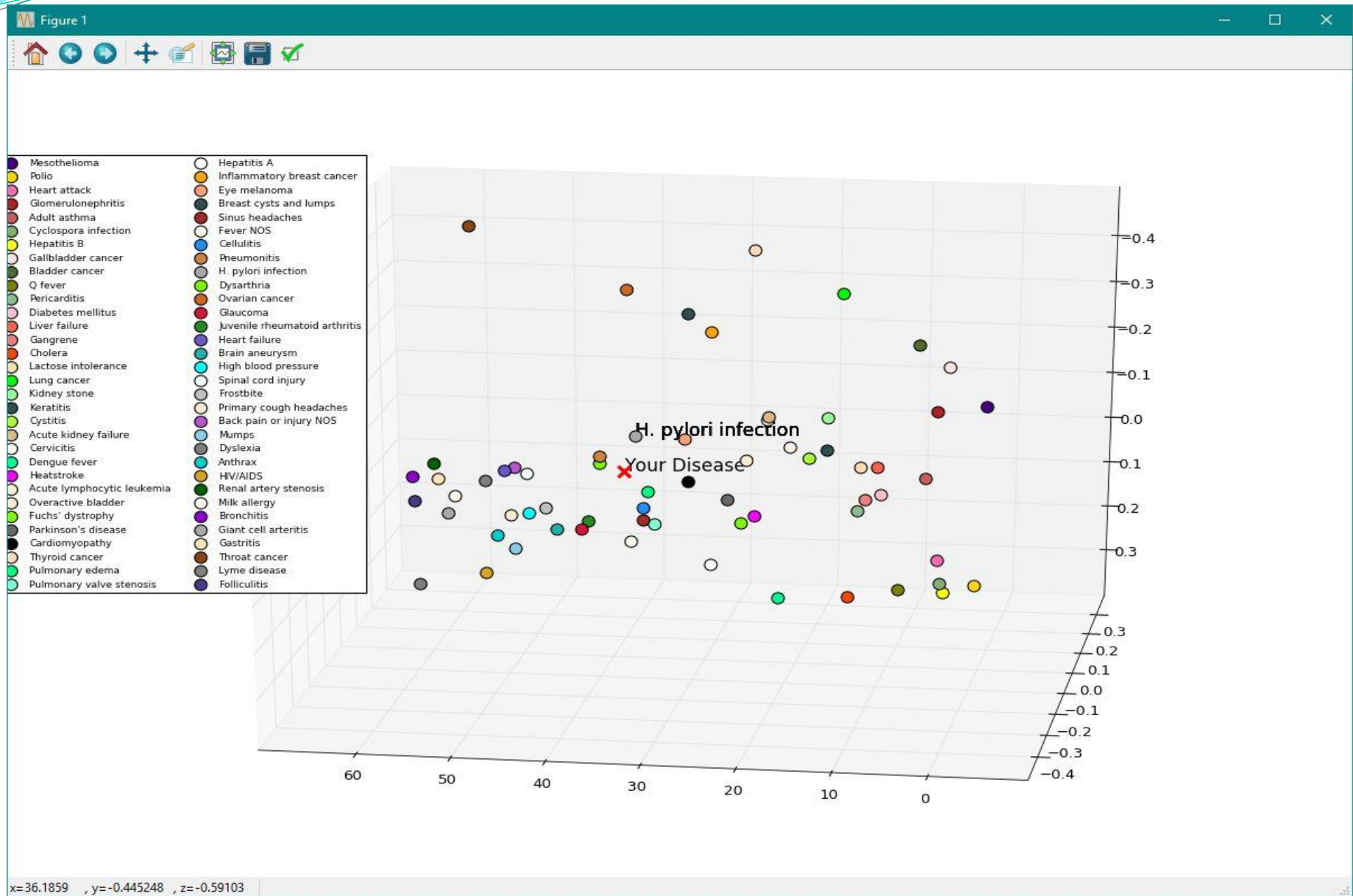
Project window



Project window



Project window



Advantages

- Get idea about disease type.
- Be acknowledged about level of risk of disease .
- Serious diseases can be detected earlier
- Works both online and offline

Limitations

- Many disease dataset are not trained yet.
- Can not detect invalid input.

Future Plan

- Increasing the accuracy of the prediction .
- Making it handy and more sophisticated for actual Doctors

Conclusion

- Although doctors are irreplaceable, it would be beneficial to the healthcare system to have a way to reliably get an assessment of health without human interaction.

References

1. “CS Project: Doctor Bayes” by Brandon Beckhardt, Leonid Keselman, Anthony Perez in Machine Learning Summer project, 2015 from Stanford University. (Link: [\[cs229.stanford.edu/proj2015/271_report.pdf\]](http://cs229.stanford.edu/proj2015/271_report.pdf))
2. CS 229 Machine Learning Final Projects, Autumn 2015 (Link: [\[http://cs229.stanford.edu/projects2015.html\]](http://cs229.stanford.edu/projects2015.html))
3. “Intro to Machine Learning Course” by Udacity. (Link: [\[https://www.udacity.com/course/intro-to-machine-learning--ud120\]](https://www.udacity.com/course/intro-to-machine-learning--ud120))
4. “Your First Machine Learning Project in Python Step-By-Step” by Jason Brownlee. (Link: [\[http://machinelearningmastery.com/machine-learning-in-python-step-by-step/\]](http://machinelearningmastery.com/machine-learning-in-python-step-by-step/))
5. Stackoverflow.com

Thanks To All