MEDISAFE - STAY AWAY AND DEFEAT DISEASE

Project Id: 2022 - 143

Final Project Thesis

B.Sc. (Hons) Degree in Information Technology Specializing in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

2022 September

MEDISAFE - STAY AWAY AND DEFEAT DISEASE

Project Id: 2022 - 143

Perera B.A.A.W.S

IT19015422

Dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science in Information Technology specializing in Information Technology

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

2022 September

Declaration

I declare that this is my own work and this thesis report does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgment is made in the text.

Name	Student Id	Signature
Perera B.A.A.W.S	IT19015422	

The supervisor/s should certify the thesis report with the following declaration.

The above candidates are carrying out research for the undergraduate Dissertation under my supervision.

....

Signature of the Supervisor Date

....

Signature of the Co-Supervisor Date

Abstract

A nation's wealth lies in its healthy citizens. It is obvious that, in comparison to other resources, it is one of the most precious resources in the nation. A major determinant in the strength of the labour market is the condition of the nation's labour force. Young, vibrant vitality is dependent on being in good health. People must therefore maintain their health in order to keep their capacity. With the increase in the global population, heart and lung ailments are becoming more common. A system for early human healthcare diagnostics will be necessary for the future.

Pulse diagnosis is important in both eastern medicine and Ayurvedic medicine. People from Asia usually think that using pulse testing can detect and treat human ailments. Humans sought treatment from Ayurveda or handmade medicines before the development of Western medicine. This is due to the fact that Ayurveda is based exclusively on herbs, which people are aware are not chemically formulated and are easily accessible. "Pulse diagnosis" is also known as "Nadi Parikshawa" among Sri Lankans. Since ancient times, people have used the pulse to determine whether or not they are ill. The patient's wrist is frequently used to assess the patient's pulse.

Three fingers are used by ayurvedic physicians to monitor the pulse and determine its activity. Many people have switched to using Western treatment instead of Ayurvedic medicine as a result of current technological improvements. This study aims to clarify pulse analysis, then find hidden disorders in the human body and offer essential suggestions. Giving Ayurveda medicine a fresh start and carrying out the project by using ancient medical practices to offer alternatives to the current Western medical approaches are what is intended to be done here.

The goal of this study is to use web development to give an active packaging function and to enable a person to receive a sense of their health state by obtaining information from a sensor device without having to visit the hospital. Additionally, this research process offers a device that combines several sensors to understand the patient's condition, after which the data is analyzed in accordance with the risk level to identify the disease, by using an X-ray test to understand the disease's current state, and finally by predicting the disease. The patient is able to follow the suggestions once the disease

has been identified using these approaches, at which point recommendations is written for the identified condition.

Keywords: Machine learning, Pulse analysis, Recommendations, Suggestions

Acknowledgement

I would like to offer my sincere gratitude to a number of people and organizations for their support during my undergraduate studies. I want to start by sincerely thanking my supervisor, Mr. Ravi Supunya, and my co-supervisor, Mr. Samantha Rajapaksha, for their excitement, perseverance, valuable suggestions, valuable tips, practical guidance, and never-ending ideas, all of which have been of great assistance to me throughout the course of my studies and writing for this thesis. My ability to successfully accomplish this research was made possible by their vast knowledge, extensive experience, and professional research project management skills. This initiative would not have been accomplished without their assistance and leadership. I found better study supervisors than I could have hoped for.

I would also like to express my heartfelt thanks to Dr. Shashika Liyanage for his support and guidance in the development of this research. I am also thankful to the following lecturers: Mr. Dhammika Silva, Mr. S.M.B. Harshanath, and Miss. Pipuni Wijesiri as a panel of judges for all their ideas and suggestions and I am very thankful to Research Project (RP) team for accepting my research.

Finally, also all at Sri Lanka Institute of Information Technology (SLIIT), it has been great to share the premises with all my friends, colleagues, and research team for all their kind help and support over the last four years and I am very thankful to them. Additionally, I want to express my gratitude to my entire family for their unwavering support and patience as I conducted my research and wrote my project.

Table of contents

List of tables

List of Figures

List of abbreviations

Table 0.1: Abbreviations

Abbreviation	Description
NCD	Non-Communicable Diseases
CVD	Cardiovascular Disease
WHO	World Health Organization
ML	Machine Learning

Chapter 1: Introduction

1.1 Background literature

Nearly 70% of all fatalities worldwide occur each year as a result of non-communicable diseases, which claim 41 million lives annually. Chronic illnesses, commonly referred to as non-communicable diseases, disproportionately impact persons in low- and middle-income nations. Studies show that as a result, one out of every five people passes away too soon [1].

Those disorders collectively referred to as NCDs have long-term health effects, are frequently treated and cared for on an ongoing basis, and are not primarily brought on by an acute infection. The majority of non-communicable diseases have their roots in behaviors started in childhood and adulthood, despite the fact that they typically emerge in maturity. The chance of dying from a non-communicable disease is increased by using tobacco, not exercising enough, not eating poorly, and consuming too much alcohol.

A fatality from tobacco occurs every six seconds or about 6 million individuals each year. Over 600 000 of the fatalities are caused by non-smokers inhaling second-hand smoke, compared to more than 5 million deaths directly related to tobacco use. The primary NCDs—heart disease, cancer, chronic obstructive pulmonary disease—as well as other illnesses like tuberculosis and neurological conditions—are all commonly influenced by tobacco smoking. Around the world, tobacco use is to blame for 14% of all NCD-related deaths in persons 30 years and older [2]. According to that, there must be a way to prevent people with those disabilities by providing proper guidelines.

The bulk of NCD is caused by four main risk factors. Poor diet is the one that contributes the most, more so than alcohol, smoke, and inactivity put together. Generally, poor intakes of nutritious foods including fruit and vegetables, whole cereals, nuts, legumes, and seafood are exacerbated by increased consumption of processed foods that include excessive levels of sugar, salt, saturated, and trans fats [3].

Although COVID-19 can infect anyone, the World Health Organization (WHO) has found certain populations are more likely to develop serious illnesses. Non-communicable diseases (NCDs), such as lung and heart disease, diabetes, and cancer, are one such category. Poor diet top the list of the numerous risks for NCDs because, according to research, they cause more NCDs than inactivity, alcohol use, and smoking put together. The availability of different foods in retail food locations has a big impact

on population diets. Although there are several chances to affect the food ecosystems, these options are yet mostly unexplored. Who is responsible to take necessary actions on are contested topics [4]. Therefore, there must be a better way to manage people's dietary patterns and provide how to manage them and prevent NCDs for a better life.

People should have to be well mentally as well as physically. In this running world, everyone is in a busy schedule. Therefore, they don't have enough time to think about their physical activities. Without having better exercise can be harmful to our life in the future. Even in those who don't have any other risk factors, a lack of physical activity can result in heart disease. Additionally, it can make other risk factors for heart disease, such as obesity, high blood pressure, and high cholesterol, more likely to appear. So, people have to follow necessary physical workouts according to their health.

By lowering common risk factors like tobacco use, hazardous alcohol use, physical inactivity, and eating unhealthily, many NCDs can be avoided. In Sri Lanka, non-communicable diseases (NCDs) account for more fatalities than all other causes put together. Nevertheless, despite their rapid expansion, NCD-related morbidity and mortality account for a significant portion of the annual burden on people and society and are preventable through well-planned, financially viable, and practical interventions. The National NCD Policy Framework lists several initiatives to lessen the burden of disease in the nation, including early identification of key risk factors and health advice.

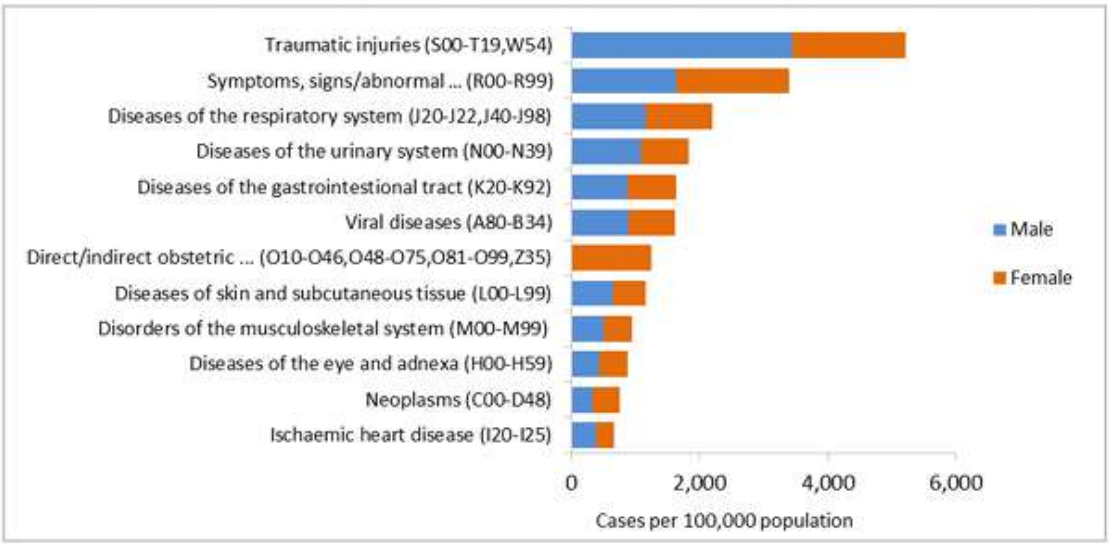


Figure 2.3 : Leading Causes of Hospitalization, 2019

Source: Medical Statistics Unit

Figure 1.1: Leading Causes of Hospitalization, 2019

The above Figure 1.1 clearly shows that respiratory system diseases are leading in Sri Lanka. There is no difference between gender because it depends on both genders. Accordingly, Pneumonia and Wheezing belong to diseases related to the respiratory system.

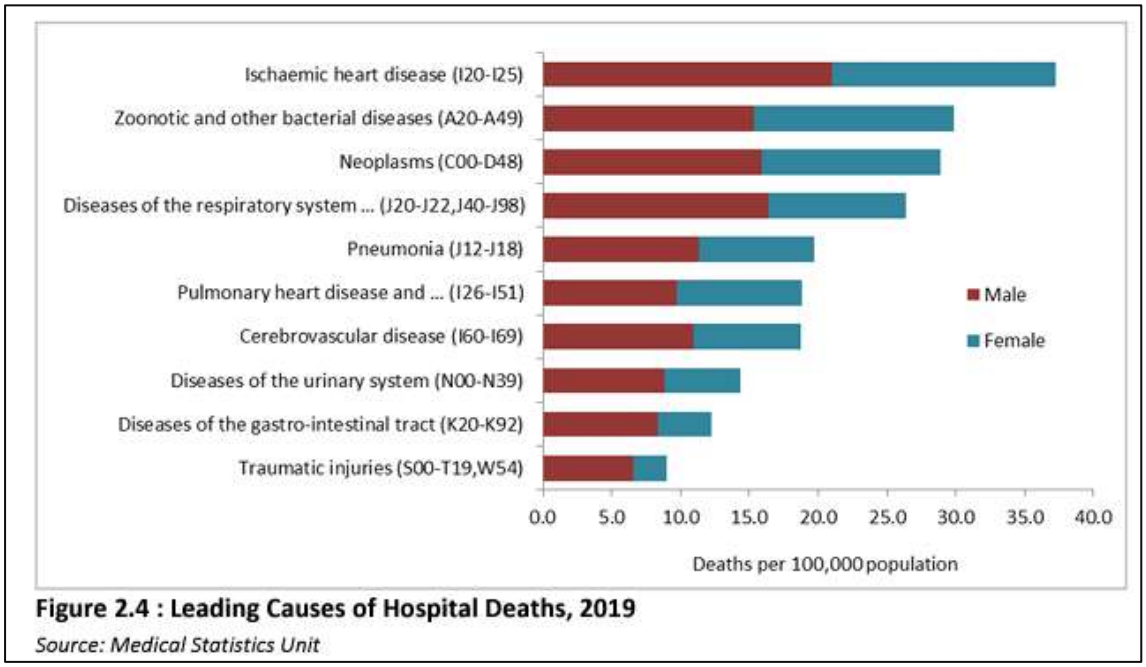
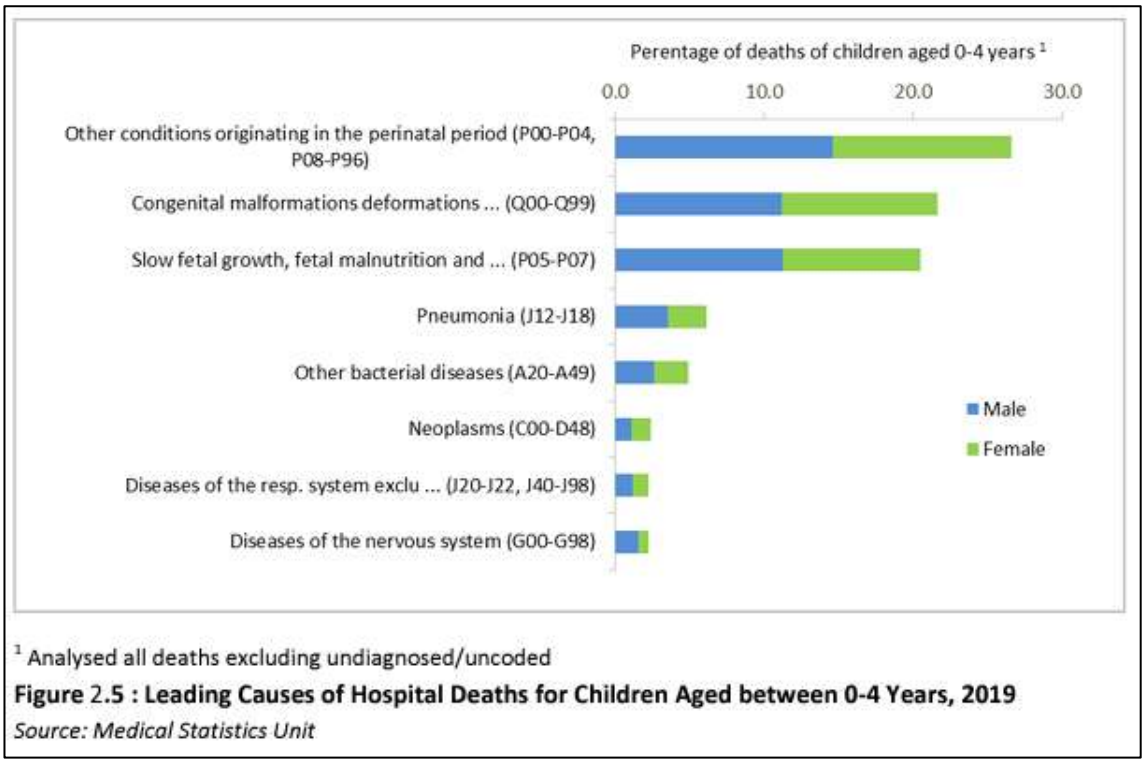


Figure 1.2: Leading causes of Hospital Deaths, 2019

According to above Figure 1.2, depicts the number of deaths per 100,000 population for the 10 leading causes. We can see more clearly that the number of people suffering from heart-related diseases are high. Figure 1.2 shows a significant gender difference in the number of deaths per 100,000 population. For most of the primary leading causes of mortality, male deaths are significantly larger than female deaths. The reasons for that have already been given above. Pneumonia, a disease related to the respiratory system, is also in fifth place. Based on those facts, we can decide that the number of people who die from diseases related to the heart and respiratory system is higher than the number of people who die from other diseases.



[1] Analysed all deaths excluding undiagnosed/uncoded

Figure 2.5 : Leading Causes of Hospital Deaths for Children Aged between 0-4 Years, 2019

Source: Medical Statistics Unit

Figure 1.3: Hospital death rate for Children aged between 0 – 4

According to Figure 1.3 above, we can conclude that age does not affect these diseases, which means people of any age are likely to have these diseases. Therefore, we should ensure to follow proper health habits and disease prevention.

Table 22. Leading Causes of Hospital Deaths, 2010 - 2019

Disease and ICD (10th Revision) Code		2019		2018		2017		2016		2015		2014		2013		2012		2011 [2]		2010 [2]	
		Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%
Ischaemic heart disease	I20 - I25	1	15.1	1	15.0	1	14.2	1	14.1	1	14.2	1	14.8	1	14.7	1	14.4	1	13.4	1	12.8
Zoonotic and other bacterial diseases	A20 - A49	2	12.1	3	10.9	2	11.5	3	11.6	3	9.7	3	9.1	6	7.9	6	7.1	6	6.7	6	6.6
Neoplasms[1]	C00 - D48	3	11.7	2	11.7	3	10.5	2	12.0	2	11.0	2	11.7	2	11.2	2	11.6	2	11.8	2	11.1
Diseases of the respiratory system excluding diseases of upper respiratory tract, pneumonia and influenza	J20 - J22, J40 - J98	4	10.7	4	9.9	4	9.8	5	8.3	4	9.2	6	8.0	5	7.9	5	7.2	5	6.9	5	7.0
Pneumonia	J12 - J18	5	8.0	7	7.8	6	8.2	7	6.4	7	7.5	7	6.6	8	6.1	8	5.7	9	5.2	9	5.2
Pulmonary heart disease and diseases of the pulmonary circulation	I26 - I51	6	7.6	6	7.9	5	8.5	4	8.7	5	8.3	4	8.6	4	8.4	3	9.0	4	8.7	3	8.7

Figure 1.4: Leading causes of Hospital Deaths, 2010 – 2019

When we take the years from 2010 to 2019 as shown in Figure 1.4 above, it becomes clear to us that the death rate of patients in 2019 has increased. Among them, heart-related diseases have taken the first rank and actual pneumonia is in the fifth position. Based on these facts, there is a division of responsibility that we must carry out in order to control

the growth of the death of patients. I hope to provide solutions to that through this research.

As a result, it is successful for medical facilities around the nation to build centers for healthy living. The main objective of MediSafe is to lower people's risk of NCDs by providing access to specialized care for patients who are more likely to develop cardiovascular disease and by detecting diseases and risk factors early on (CVD). Referrals are made for selected customers to have clinical exams and tests.

Thus, there are treatments for chronic non-communicable diseases in Ayurvedic methods as well as in western medicine. Ayurveda is a major medical practice that is founded on the conventional medical system of India. It uses a holistic approach that prioritizes nutrition, herbal remedies, exercise, rest, and exercise to treat and unite the body, mind, and spirit [5]. Western doctors reach decisions about the best medication for their patients based on systematic, scientific investigations. This method is known as evidence-based medicine. Evidence-based treatment approaches include, for example, prescribed medications, procedures, infusions, and other conventional therapy.

The entire country was under lockdown as COVID-19 spread, and hospitals were crowded with patients. The death rate increased as a result of many people being unable to access the necessary medical care [6]. Then, in an effort to prevent both their current and earlier ailments, people tried to develop their own medical patterns. Without understanding the proper Ayurvedic and Western medical dosage, they used improper therapies. Concerns about the coronavirus were also raised. Because of the emergence of the COVID-19 condition and the rising death rate among the elderly globally, people's lifestyles have altered, which has led to an increase in chronic diseases. It should be possible for people to get their health checked at home.

This study demonstrates the development of a portable prototype for pulse activation detection. The quality of people's lives will improve if they use contemporary medical technologies to prevent ailments. The best performance for this device comes via wireless operation. By examining and analyzing the pulse, Ayurvedic practitioners can use it to diagnose the illness and offer treatments for prevalent non-communicable diseases.

According to the poll, the following are responses from those who have suffered illnesses.

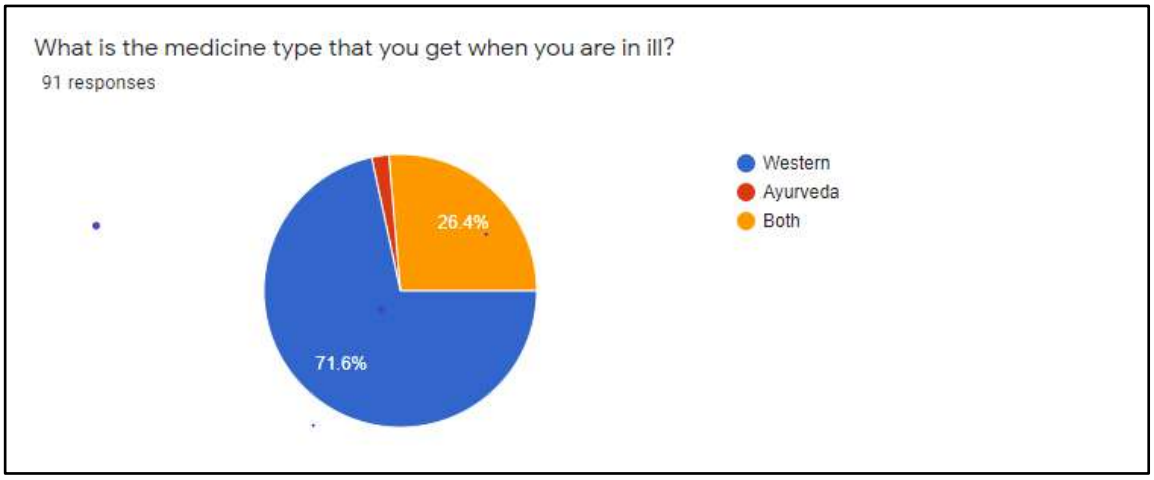


Figure 1.5: Drug type that people often get

Many people use western medicine, according to the data presented above (Figure1.5). We might deduce from this that individuals are progressively turning away from Ayurvedic medicine and toward Western medicine. The reason for this is that Ayurvedic treatments take a long time to produce effects, but Western medical approaches produce results quickly. However, the drugs employed in Ayurvedic medicine are not damaging to the body, although Western medicine can have negative effects.

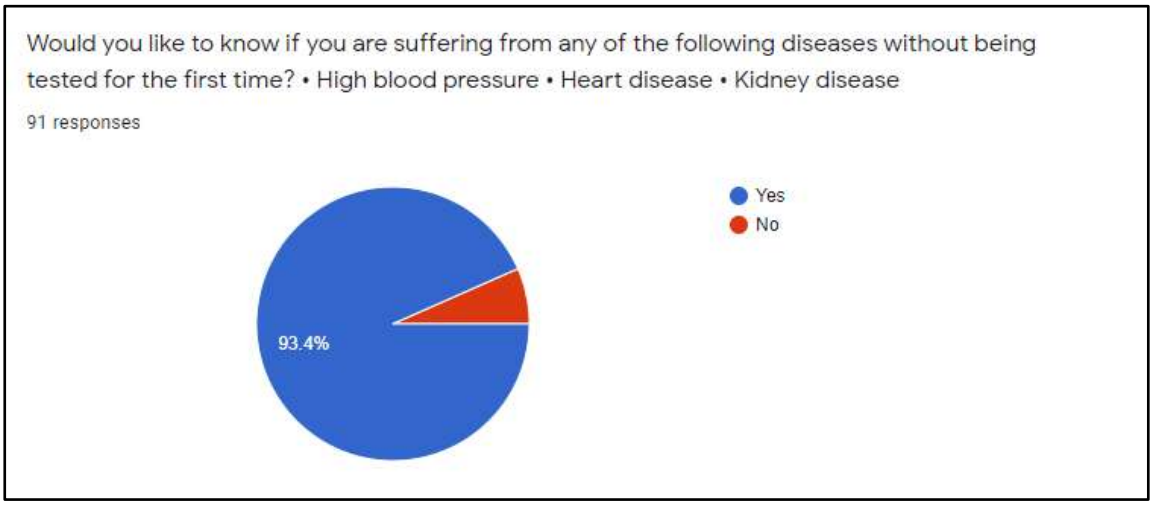


Figure 1.6: Willingness to obtain information about the diseases

Considering the above figure (Figure 1.6) shows that people like to know about their hidden diseases. As a result, we hope to use this method to diagnose what diseases are affecting the body.

Here the diseases in my component looking to diagnose are Wheezing, Pneumonia, and Heart Attack. That's helpful to people who don't have enough idea about their health.

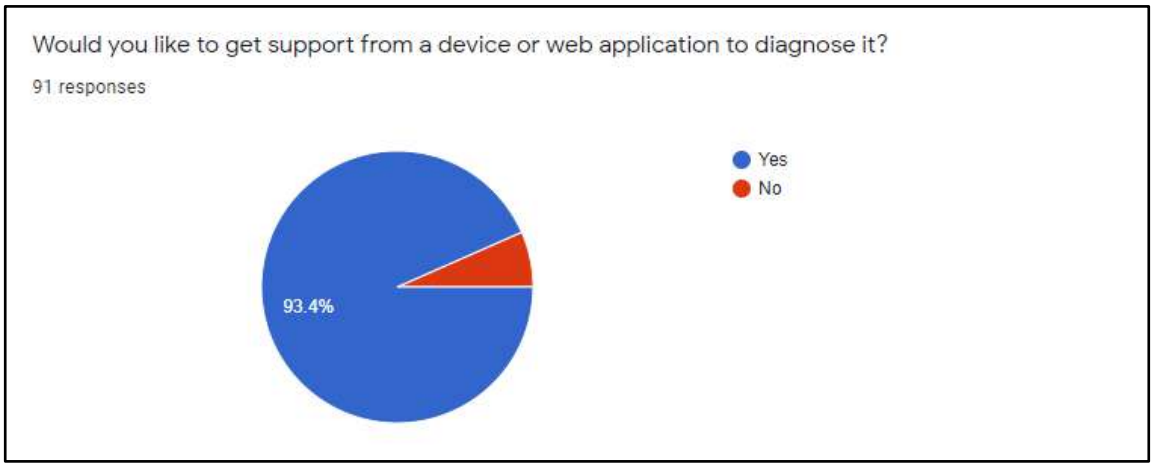


Figure 1.7: A preference for getting help from a gadget or web application

Based on the information supplied, we decided that a system is necessary (Figure 1.7). The illness detection system also includes a hardware component for reading the pulse. This enables us to collect accurate patient data in order to treat them. Once the pulse data has been read, it will employ machine learning algorithms to transfer the measured values to a web application and recommend treatment.

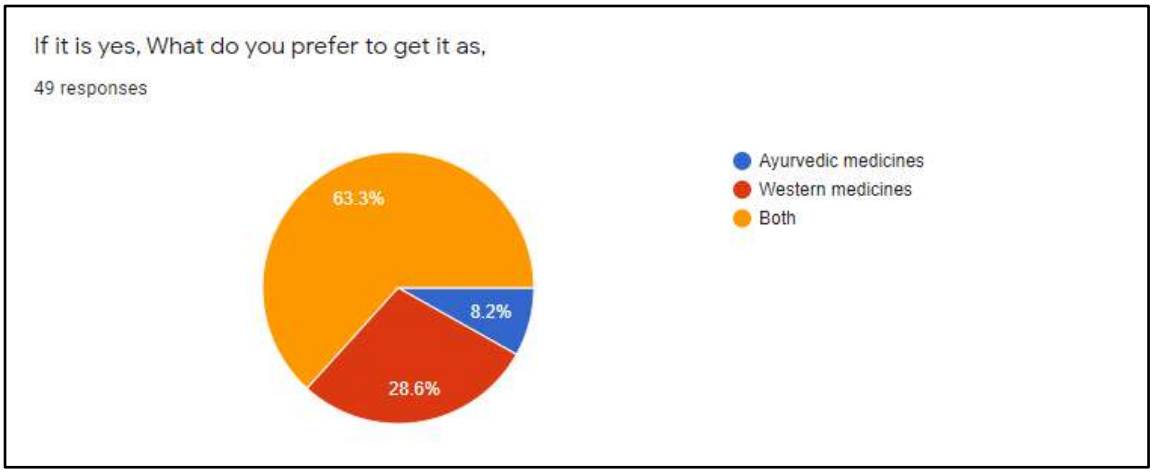


Figure 1.8: A preference to use specified drug type

According to data collected (Figure 1.5), roughly 71% of people use Western technique medications. However, looking at the statistics above (Figure 1.8), we can see that the majority are interested in both Western and Ayurvedic treatments. So, as with Ayurvedic approaches, our answer is to read the pulse and educate the community and

people about the importance of Ayurvedic medicines. However, by employing this method, people can obtain the Ayurvedic or Western medical prescription they choose.

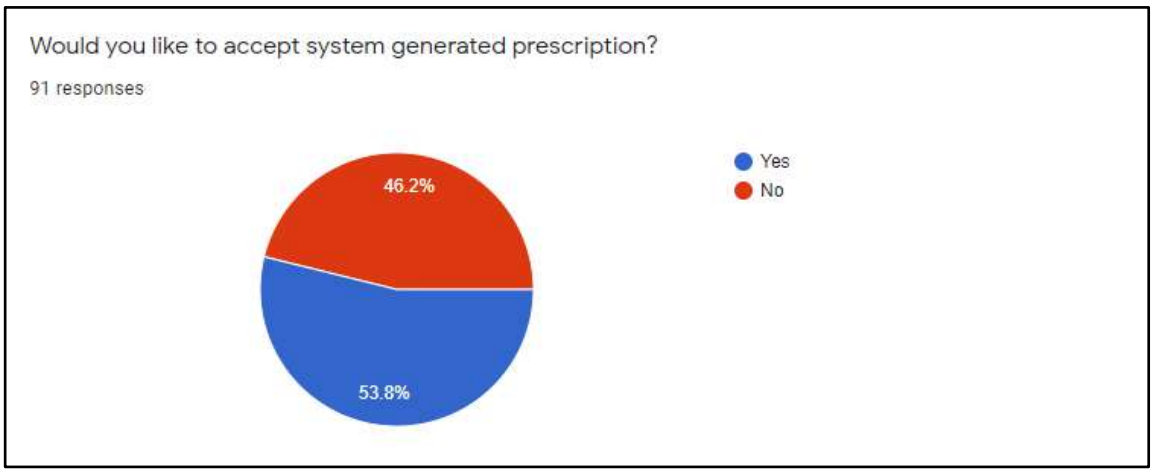


Figure 1.9: Acceptance of system generated prescription

However, the majority of people are still hesitant to order prescription medications online. There have been stories of persons being misdiagnosed by surgeons and struggling with physical and mental disabilities. A diagnostic inaccuracy can be the difference between life and death for a patient. While estimates vary, it is likely that over 100,000 people die or are permanently crippled each year as a result of medical diagnoses that first miss conditions, are incorrect, or are delayed [7]. Given the above, patients are reluctant to accept a system-generated prescription for their ailment. Also, they don't want to take it in proper dosage. For those reasons, without presenting a prescription as a solution to this, it is aimed to give a series of instructions on how the patient should deal with the level of illness related to his disease in such a case. Also, it does not harm the patient and has the opportunity to reduce the risk to a certain extent.

1.2 Research gap

There have been accounts of persons suffering from numerous maladies as a result of their usage of tobacco, alcohol, and an improper diet. According to the knowledge we can glean from people suffering from such disorders, the majority of research has focused on the following metabolic parameters.

- Pulse rate
- Gender
- Age
- Alcohol usage
- Cholesterol level
- Anxiety

Among the studies that have been conducted are the,

- Recognizing aberrant pulse rhythms
- Gender and age monitoring
- Pulse rate study indicates various diseases

According to current research papers [8] - [15] and resources, various research papers on the concept of MediSafe (Proposed system) have been established to monitor human pulse rate. The majority of study is focused on pulse measuring using pulse analysis those three Ayurvedic traditional methods (Vata, Pitta, Kapha) [8], [11], [14]. The goal of Research A [11] is to uncover the possibility of employing the same traits that are used to determine age and gender to diagnose diabetics. Advances in instrumentation technology have allowed Research B [14] to contemplate building a high-quality pulse acquisition system. Implementation of Research C [12] is estimating the gathered pulse rate and identifying gender and age. Taking that into account reveals whether or not there is a danger of heart disease. The mentioned research papers [8] – [15] couldn't have a way to identify heart-related diseases from their risk level (High, Medium, and Low). Therefore, that is also a drawback of those research.

The below table will describe the comparison of previous research and our suggested system.

Table 1.1: Comparision of previous research and proposed system

Research products	Features				
	Identify gender & age	Development of advanced equipment	Identify chronic diseases	Disease risk level-wise identify	Provide recommendations and suggestions in western & Ayurveda methods
Research A	✓	×	✓	×	×
Research B	×	✓	×	×	×
Research C	✓	×	✓	×	×
Proposed system (MediSafe)	✓	✓	✓	✓	✓

My research component is in the diagnosis of heart disease. These include heart attacks, pneumonia, and shortness of breath. This system helps in diagnosing the above diseases from home or any other place using data models to ensure the continuous connection between the device and the pulse monitor. It monitors the patient's pulse patterns based on the patient's pulse data and based on their age, gender and other inputs given by the patient tells the risk level of the disease and in which parts of the body the disease is present and gives a set of instructions for the patient to follow. The aim is to help reduce the risk levels of diseases by developing an algorithm for diagnosis.

Chapter 2: Research problem

An increasing number of people are suffering from stress and other chronic ailments as a result of their stressful work schedules, tobacco use, alcohol consumption, bad dietary habits, and strenuous work schedules, or mismanaged lifestyles. The current lifestyle of today's youth is a red flag for significant health problems. Because muscular movements are restricted, the body weakens, increasing the risk of obesity, cardiovascular disease, metabolic syndrome, and early death, hypertension, kidney disease, and diabetes, all of which are serious concerns in today's culture [16]. This is rapidly growing among those aged 20 to 40.

The corona outbreak, which lasted over two years, produced a slew of issues for humans. The inability to see a doctor and obtain medicine for an illness, the development of concealed ailments as a result of the virus's spread and the necessity to avoid physical exercise due to having to stay at home were all issues. As a result, this has grown into a global issue, and it is our obligation as undergrads to provide new solutions. When a person is stressed, the body creates adrenaline, which causes the pulse rate to rapidly increase [17]. There are now ways for assessing pulse rate that uses mobile applications and other digital technology [17], but these electronic gadgets are pricey. To retrieve the pulse rate, we must hold the apparatus in such a way that it can detect the pulse rate.

As mentioned above, nowadays there is a chance of developing many diseases related to the heart. The main reasons for suffering from such ailments are mentioned above. Among them, heart attack, pneumonia, and wheezing are the most common diseases. Therefore, there is a need to get rid of such sick conditions and to know if a person is showing such a sick condition [18]. With busy lifestyles, people lose opportunities to look after their health conditions, and thus we have heard many cases of minor medical conditions becoming severe [19]. Therefore, many advantages can be gained in the early detection of such a disease and the opportunity to control the severe conditions that may arise from it. One of the problems that can be presented here is the premature death of patients due to the lack of proper attention to the conditions of these patients.

Another problem is that nowadays, with the advancement of technology, people suffering from these diseases can find many procedures to follow. But people who do not have good knowledge of technology and do not have smart software lose the ability to know about such situations [20]. This software works to be placed in a place where there are

many people. That is, any person has a chance to know about their current situation. And it does not require any special knowledge to use it. That's one of the advantages here. Therefore, as a remedy for that, I would like to present the remedies to know the disease through my component through this system and to get rid of it.

Another research question is whether people improperly control their eating patterns [21]. That is, they follow many methods such as skipping meals, eating foods that are harmful to the body, and eating foods that are not related to their body's immunity. Some people resort to unhealthy diets to lose weight. But it will affect some kind of anomalies in the future. It will affect the internal organs of the body [22], [23]. Therefore, they may also affect heart disease. Therefore, among the procedures to be followed, comments are also made about the way to get food.

As adults age, walking becomes difficult. In such a case, you have to spend some time and wait to go to the hospital and get medicines for your body's discomfort. This makes them even more tired. In such a situation, there is an opportunity to save time and get solutions without going to a hospital without getting tired.

As a remedy to the aforementioned issue, I advocated implementing solutions based on proper technology to prevent such diseases. My contribution is to propose a solution through recommendations based on pulse analysis. Finally, this study proposes "MediSafe" as a tool and web application that can assist people in improving their quality of life by identifying concealed disorders.

Chapter 3: Research Objectives

3.1 Main Objective

The primary goal of this research is to provide patients with the essential remedies and requirements for their existing ailments, as well as to discover chronic disorders in their bodies. The covid-19 virus is one of today's most serious social challenges. So far, the issue has been kept under control to some extent, but it is once again on the rise. It is tough for patients in this situation to understand their existing physical condition and to find answers to it. The project's goal is to present the proposed device design as well as the outcome of a mobile application and online application needed to operate the device.

The device is planned to be available in areas where the majority of people congregate. Appropriate for a nursing facility, for example. That is, the majority of these adults suffer from multiple diseases, and others have trouble walking. It is tough to go to the clinic and have a medical assessment in such a condition. There are benefits to acquiring their diagnosis quickly in this case. It is also appropriate for an office. That is, due to the employees' hectic schedules, they are left to maintain their correct physical condition as well as receive the essential therapy. In this instance, they can examine their health status at any time by configuring this gadget.

Accurate analysis of the information gathered from the sensors will provide them with an understanding of their present situation, the procedures to be followed to prevent them, the opportunity to respond appropriately to them through a dialogue (via chatbot), and their perspectives on the current situation in Sri Lanka. It is hoped that this would allow accurate information to be provided.

3.2 Specific Objectives

The specific aims are as follows, according to the preceding major objective explanation:

1) Expert gadget creation to gather and analyze sensor data.
2) Disease level-wise analysis and provide recommendations
3) Heart and lung diseases identify using image processing techniques
4) Provides the current situation diseases information in the country.

My implementation for the aforesaid second step in this study project. The device retrieve data that can be used to determine a patient's pulse. And also, this will gather some additional details about the patient to provide a better solution for the disease situation. Following that, this technique comprises taking into account the person's pulse rate data before records are obtained from various sources and used to identify ailments using an algorithm. Finally, it will provide the appropriate guidelines to be followed. This will also help individuals to learn about chronic conditions (Such as heart disease - Pneumonia, Heart attack, and Wheezing).

Chapter 4: Methodology

4.1 Overall System Architecture

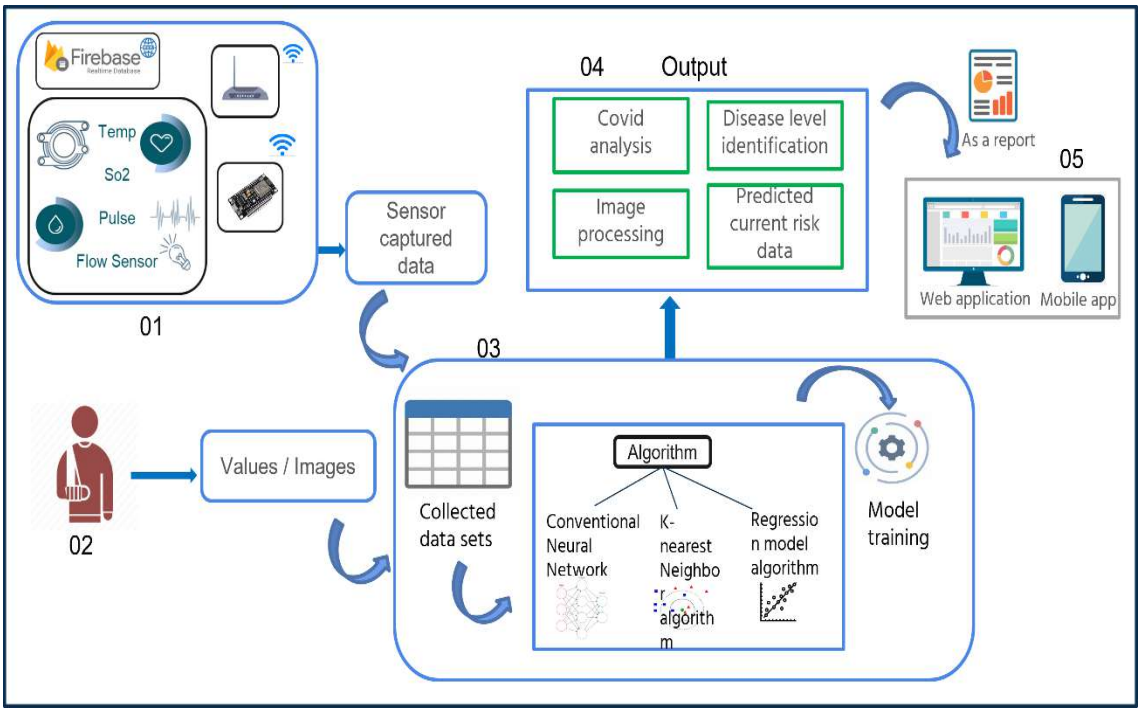


Figure 4.1: Overall system diagram

The proposed "MediSafe" system is a device implementation with a web application that can perform the following tasks: In the brief introduction of each component.

1) Measuring the current situation parameters of the patient via the proposed device and analyzing data and covid 19 identification.
   - The proposed device should be able to detect data. The proposed system includes temperature detectors, breath detectors, and blood pressure detectors. The sensors should provide correct values for the parameters. The captured data should be transferred to the cloud database via Wi-Fi. Finally, it will retrieve the sent data from the cloud database and present the relevant analyzed data to the user via the web application.

2) Frequency distribution of pulse rate variability data use to identify the heart and lung-based diseases analyzed by risk level and providing recommendations.

- Patients may check their pulse rate via the implemented device that is connected to the web application. Additionally, it is getting some user inputs from the patient and processes it according to the trained model and it provides the disease risk level and suggestions to prevent those ailments.

3) Image processing techniques to identify lung diseases
   - Patients have to upload their X-ray images to the web application. Then according to the trained model of image processing-based techniques, it will provide the affected disease and its probability as a percentage to the user via the web application.

4) Identify the Sri Lankan current diseases spread rate diseases information.
   - Users have to insert their current location and date via the web application. Then the use of machine learning prediction and trained data sets system will provide the future prediction that number of patients can be affected in the selected location.

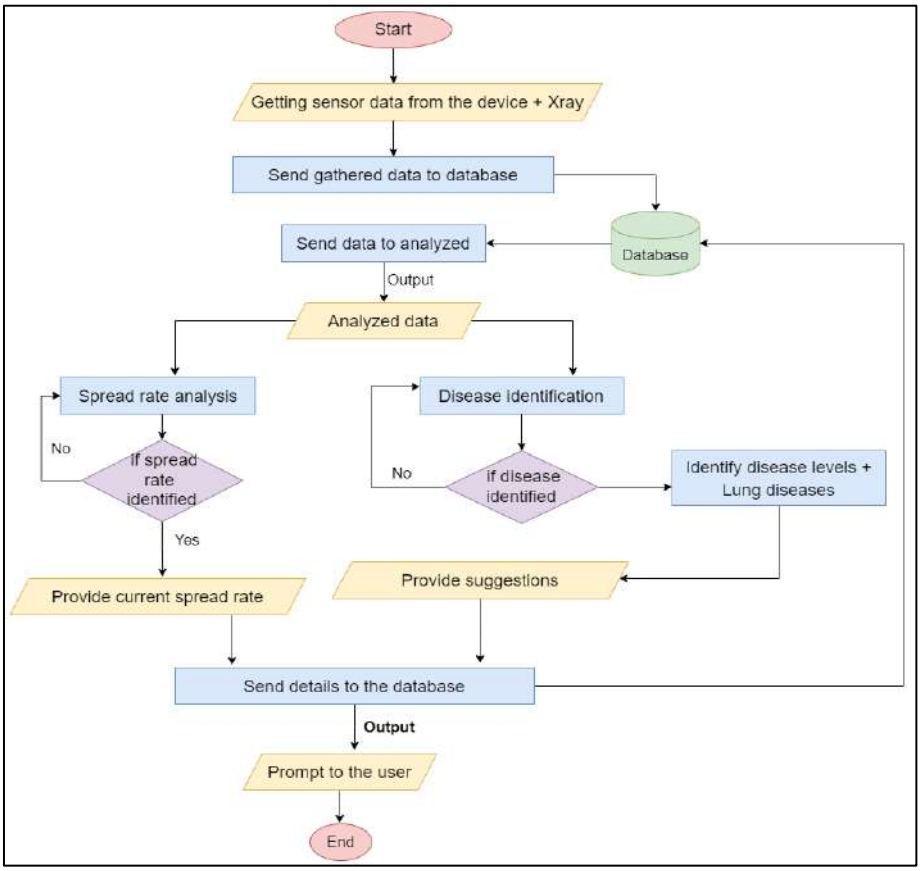


Figure 4.2: Entire system architecture working process

4.2 Individual component Architecture

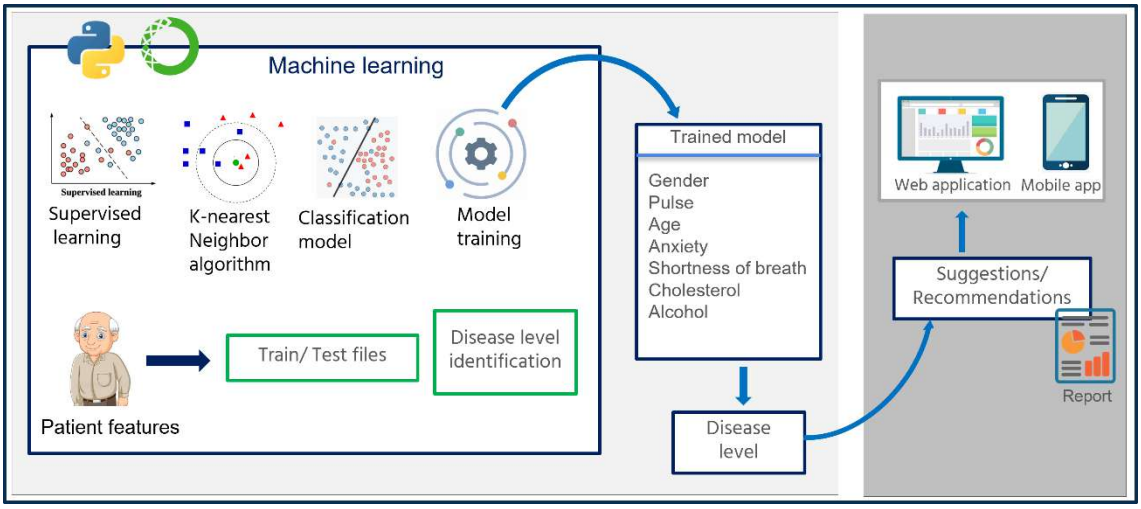


Figure 4.3: Individual component system architecture

4.2.1 Data collection and model selection

The starting point of the implementation, first of all, is to find a suitable dataset for the model training. In this research my component some functionalities are based on machine learning using Python language. To do that, I had to find a dataset related to non-communicable diseases (Heart Attack, Pneumonia, Wheezing). The datasets were collected from Kaggle.com and data.world. After finding,

1. Clean the dataset – The practice of correcting or deleting incorrect, corrupted, improperly structured, redundant, or incomplete data from a dataset using interprograte.
2. Data arrange – Merge the relevant features related to the relevant implementations.
3. Divide dataset – In here the whole dataset was divided into two categories. 80% for training, 20% for testing and validations.
4. The collected dataset was trained using the KneighborsClassifier model. Its accuracy was 96.25%.

```
    regr.fit(X_trainH, y_trainH)
    encoder2 = LabelEncoder()
    y_testH = encoder2.fit_transform(y_testH)
    r2_score = regr.score(X_testH,y_testH)
    print('confidence :',r2_score*100,'%')


confidence : 96.30732215636247 %
```

Figure 4.4: Accuracy of the selected model

4.2.2 Model training and data analysis using machine learning techniques

This implementation was completed by using the KneighborsClassifier model. Here Pandas, Joblib, and Sklearn libraries were used for this model training. First, need to import the .csv files and had to call for the features (Attributes of the dataset) of the separate dataset files. Then need to do normalization, where the training model is saved, and then need to call the normalization function to make a scaler for the separate diseases (Heart Attack, Pneumonia, Wheezing). Using the transform method will apply the scaler file. After applying the train values and test values those were normalized. Then had to import the KneighborsClassifier. Then train the classifier according to the previously introduced X and Y axes. After that, need to define the scaler file that should be saved and provide a proper naming. Finally using the jb. dump save the scaler file and model files. After model training, 20% of testing data was tested and it provides the expected output as shown in the dataset results.

4.2.3 Output generation of the trained model

In the trained model, the output will provide 3 categories. In the analyzed diseases output the disease risk level like High, Medium, and Low. The first user has to provide their current health status details then according to the trained model, the entered details will be analyzed and provide the disease risk level. Then according to the identified risk level, the system will provide suggestions to the user via the web application.

```
print('-------------------------------------------------')
print('                   USER DETAILS                  ')
print('-------------------------------------------------')
print('                                                 ')
age = input('Enter your age : ')
Gender = input('Enter your gender (1 - Male, 0 - Female) : ')
Cholesterol = input('Enter your Cholesterol value : ')
Pulse = input('Enter your pulse : ')
Smoke = input('Are you smoking (1 - Smoking , 0 - Not smoking ) : ')
Alcohol = input('Alcohol usage (1 - Yes, 0 - No)  : ')
Shortness_of_breath = input('Have any Shortness_of_breath (1 - Yes, 0 - No) : ')
Anxiety = input('Have any Anxiety (1 - Yes, 0 - No) : ')
```

Figure 4.5: Details that the user should have to provide

Note - The above Figure 4.5 is previously implemented python implementation for the 50% completion.

4.2.4 Web application development

The web application frontend development was done using React, the backend development Python language and the server is Flask.

4.3 Software solution

A software solution is a technology that aids in the execution of time-consuming tasks and easily automates the routine task of catering. Because the finished product must fulfill both non-functional and functional needs, a software solution is necessary. They work with all parties engaged in the process to get the best possible conclusion. The agile methodology is employed as a software development approach to carry out the procedure.

Agile software development refers to a set of methodologies that emphasize iterative development and the implementation of changes through collaboration among self-organizing cross-functional teams.

Scrum is a software development agile approach that is focused on incremental and iterative procedures. The ultimate benefit of scrum development is that it allows groups to offer product faster, with higher quality and predictability, as well as greater flexibility

to adapt to change. The diagram below depicts the essential stages of the agile software development life cycle.

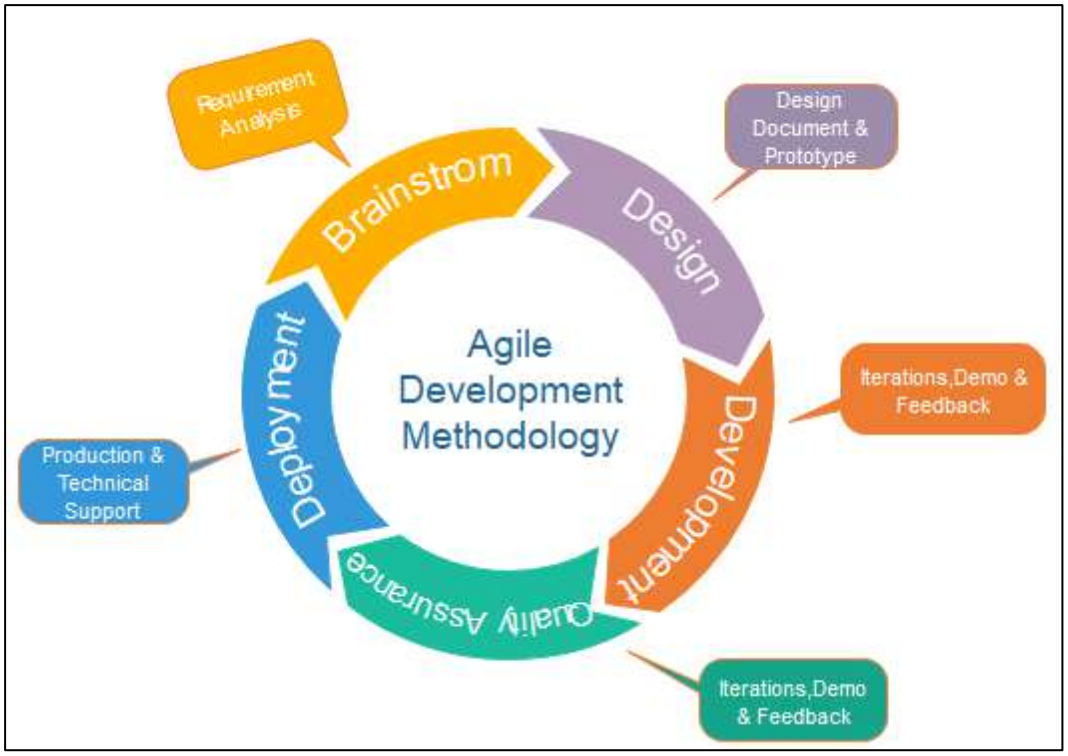


Figure 4.6: Agile methodology

- **Requirements analysis**

Based on the information and research required to complete this project, we did many surveys to ensure the success of our project. We gathered the information we required by collecting the necessary information and sharing it with others via a Google form. Furthermore, we looked at prior study papers written by researchers and gleaned a lot of ideas from them. We also had the opportunity to learn a lot from discussions with doctors who are experienced in these disorders and the industry. We are hoping to continue this project with the help of our doctors.

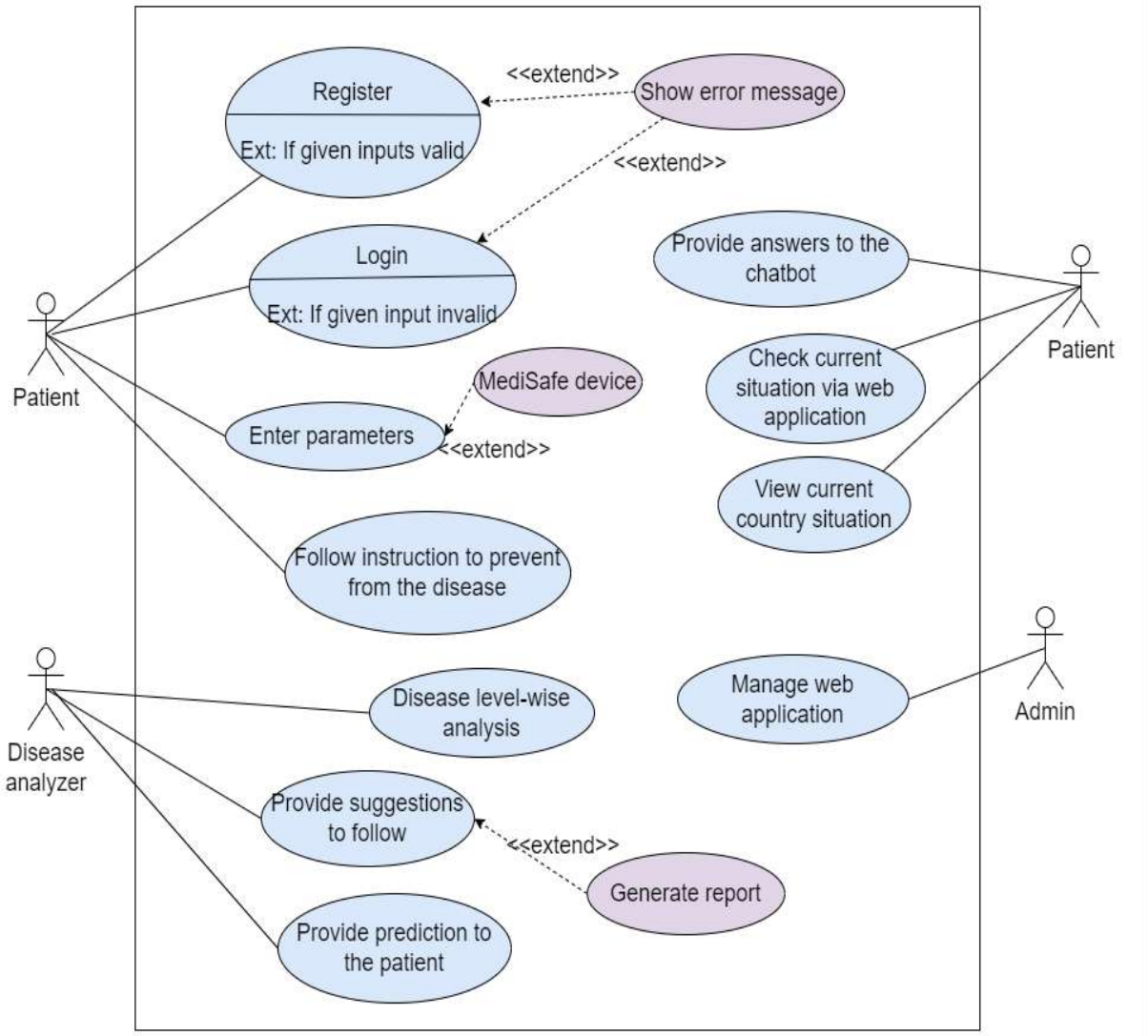


Figure 4.7: Use Case diagram for MediSafe

- **Design and development**

This design stage focused on how the final application should be looked like and how to implement that system as user-friendly by considering functional and non-functional requirements. Wireframe designing started and discuss those interface issues with the group members and finalized perfectly. MockFlow is the tool used for this design process. In the below appendices section, the wireframe designs are attached. And also, In this stage, all the members discussed what are the technologies we will need for these implementations. After the discussions, we choose to React for the web application development and Flask for the server-side development.

- **Feasibility study**

As a result of this project, our developed system will work with the web application to deliver the right solutions to problems that are working perfectly, producing correct outcomes with no omissions or mistakes.

Furthermore, the expense of designing these items keeps to a minimum. Furthermore, the cost of the components employed should be cheap, and the active outputs of those components should be consistent. As a result, action must be made to provide fair cost-of-use alternatives.

The creation of a device is a critical component of this undertaking. As a result, it is critical to understand the electronics utilized in the design to acquire the correct output by connecting the devices, as well as a full understanding of the equations needed to obtain that output. We must rely on specific sources to gather this information.

- **Implementation**

The device is built with Arduino technology, while the web application is built with React. For development purposes, a cloud-based server is also used. The server is Flask.

This project took the following steps.

- ✓ Data will be gathered by the device's sensors and processed.
- ✓ Disease level identification and recommendations
- ✓ Disease detection using image processing
- ✓ Patients will be allowed to contribute additional information in addition to the information provided.

This is accomplished by using a web application to provide the user with diagnoses for the final diagnosed ailment.

- **Testing**

Each system functions process should be tested. Unit testing is utilized to detect early-stage defects in the system and guarantee that each unit is operationally sound. Each unit should be checked before being fully incorporated to prevent faults and system failure. As a result, any coding flaws will be obvious. Integration testing should be undertaken after the system has already been integrated to ensure that it is operationally sound. Integration testing is performed to detect any issues that may have happened after the system was connected.

- **Maintenance**

In the maintenance, the system should have some kind of version with some new features. And have to manage it without bugs and allow the user to use it independently without getting others' help to diagnose the disease.

4.4 Requirements

The implementing system is a collection of the below requirements.

4.4.1 Functional requirements

- Detection of pulse sensor data
- Analysis of collected data using the trained model
- Effectiveness
- Provision of proper recommendations/suggestions

4.4.2 Non- functional Requirements

- Serviceability
- Accuracy
- Security
- Performance

- Quick response
- Maintainability

4.5 Commercialization

The Medisafe system that we present here is most helpful for the elderly who suffer from frequent illnesses. Therefore, we are working to present this system to an elderly nursing house. In this way, referring the elderly to hospitals is a very difficult task and they do not need to exert more effort, and in any case, if the risk of their diseases is high, they have the opportunity to immediately go to a hospital. This saves their money and time.

Moreover, with the busy lifestyle, many people have limited time to look after their health. In such a case, by installing this system in one place in an office, office workers will have an opportunity to take care of their health in their free time. In such a case also, this can be used as a very advantageous system. We are expecting to place this system in a small office.

And we would like to take this opportunity to install this device and system in our external supervisor's medical center under his guidance. We have the opportunity to verify the accuracy of this device and system by examining incoming patients.

And also hope to take an opportunity to present this to an outsider. That means a good marketplace. Then we would be able to introduce this to a proper rate.

Social media platforms are also a good opportunity to promote our product in advance. By posting different kinds of articles about our product will be able to promote among people. And it's a better way to a promote product without spending more money or time. We no need to use promoters to promote our product. Our team members would be able to promote by using different platforms available on social media. (Facebook, Instagram, Web sites – daraz.lk, ikman.lk)

Chapter 5: Testing & Implementation

5.1 Testing process

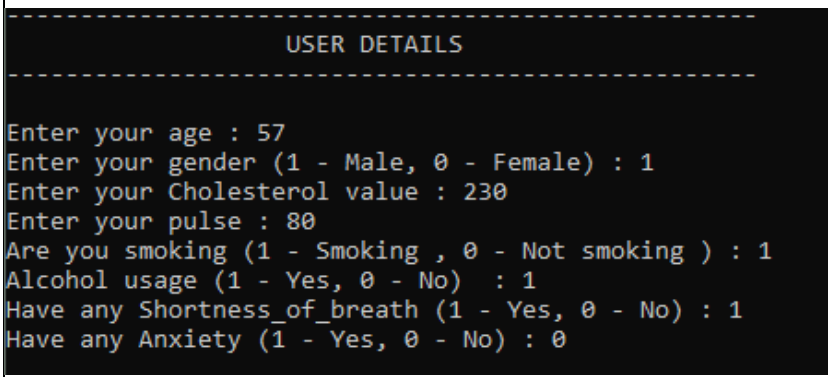
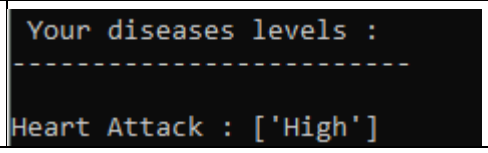
The system's functionality is tested individually. Unit testing is used to check each functionality of the system to identify bugs in the functions and to confirm those are working smoothly. This should be tested before integrating each function together to avoid system failure. As a consequence, any coding flaws will be obvious.

After completing the system integration, the whole integrated system can be tested by using integration testing. Integration testing is performed to uncover any errors that might have happened after the system has been integrated.

After the system has already been finished, it must be tested. System testing guarantees that the system produces the expected output and meets the needs of the user. Here we can verify that the expected outcome is the same as the actual outcome by discussing the external supervisor also. Then it allows using for anyone.

Task - Heart Attack, Pneumonia, and Wheezing diseases risk level-wise identification

Table 5.1: A test case for Heart attack risk level

Test Case ID	TestCase_01
Test Scenario	Check Heart attack risk level
Test Input	
Test output	

Expected output	Heart attack in High Level
Actual output	Heart attack in High Level

Table 5.2: A test case for Pneumonia risk level

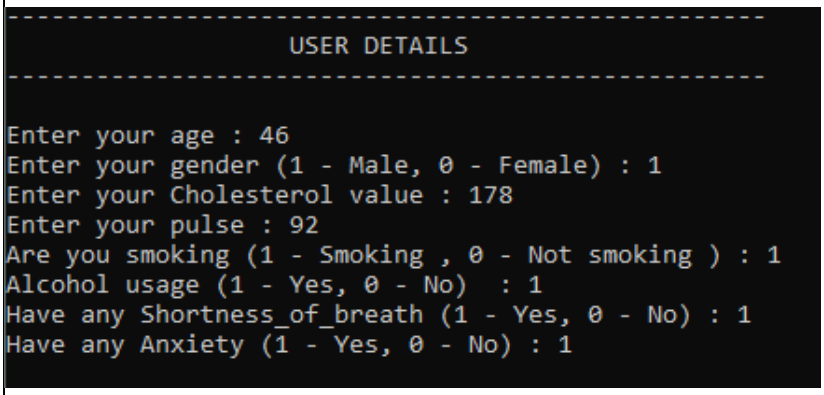
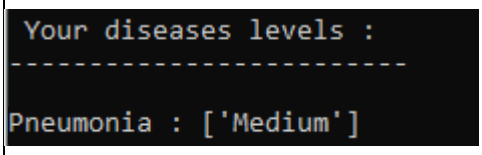
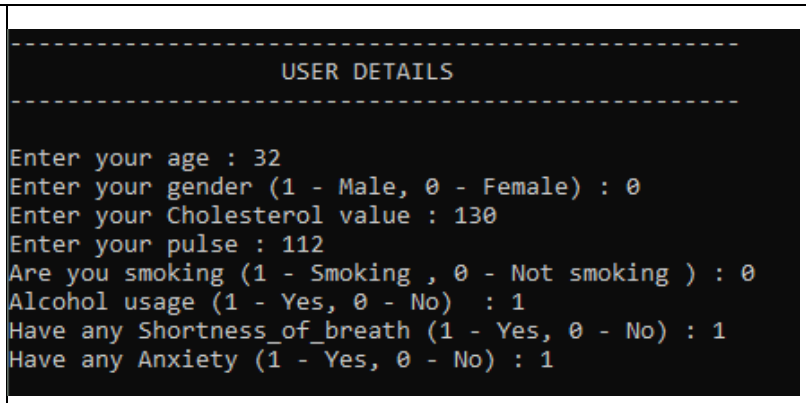
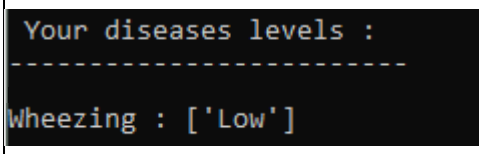
Test Case ID	TestCase_02
Test Scenario	Check Pneumonia risk level
Test Input	
Test output	
Expected output	Pneumonia in Medium level
Actual output	Pneumonia in Medium level

Table 5.3: A test case for Wheezing risk level

Test Case ID	TestCase_03
Test Scenario	Check the Wheezing risk level

Test Input	
Test output	
Expected output	Wheezing in Low level
Actual output	Wheezing in Low level

5.2 Implementation

The disease level-wise identification component act according to the user entered values. After inserting the values that should analyze the entered data according to the trained model. And then provide the risk level and according to the risk level that will give an output to the user as a recommendation. Accuracy and performance are non-functional requirements of the component. The technologies utilized to construct my component for illness risk level identification are listed in the table below. Python was used to implement the risk level detection and analysis. The implementation of the risk level identification and analysis function is the primary emphasis of this section.

Table 5.4: Project technology stack

Algorithm	K – nearest neighbor algorithm
Libraries	pandas, sklearn, joblib, numpy
Distribution	Anaconda Navigator
Language	Python, React
IDE	Jupyter Notebook, VS code

Data collection	Kaggle, Data.world

A. Disease level identification model training

The risk level identification doing according to the user entering details. So first, it should have to train the model by referencing the collected data sets.

```
disease > predict.py > ...
1  from sklearn.model_selection import train_test_split
2  from sklearn.preprocessing import StandardScaler
3  from sklearn.neighbors import KNeighborsClassifier
4  from sklearn.metrics import classification_report, confusion_matrix
5  import joblib as jb
6  import pandas as pd
7
8  dataframe_H = pd.read_csv('Heart_attack.csv')
9  dataframe_P = pd.read_csv('Pneumonia.csv')
10 dataframe_W = pd.read_csv('Wheezing.csv')
11
12 namesH = [
13     'age', 'Gender', 'Cholesterol', 'Pulse', 'Smoke', 'Alcohol', 'Risk'
14 ]
15
16 namesP = [
17     'age', 'Gender', 'Shortness_of_breath', 'Pulse', 'Smoke', 'Alcohol', 'Risk'
18 ]
19
20 namesW = [
21     'age', 'Gender', 'Anxiety', 'Shortness_of_breath', 'Smoke', 'Alcohol', 'Risk'
22 ]
23
24 XH = dataframe_H.iloc[:, :-1].values
25 yH = dataframe_H.iloc[:, 6].values
26
27 XP = dataframe_P.iloc[:, :-1].values
28 yP = dataframe_P.iloc[:, 6].values
29
30 XW = dataframe_W.iloc[:, :-1].values
31 yW = dataframe_W.iloc[:, 6].values
32
```

Figure 5.1: Importing libraries and datasets for model training

Above figure 5.1 describes the libraries importing and dataset importing to the model training and introduces the features of the datasets. (namesH means Heartattack features, namesP means Pneumonia features and namesW means wheezing features.) Those are the three main diseases that I analyzed.

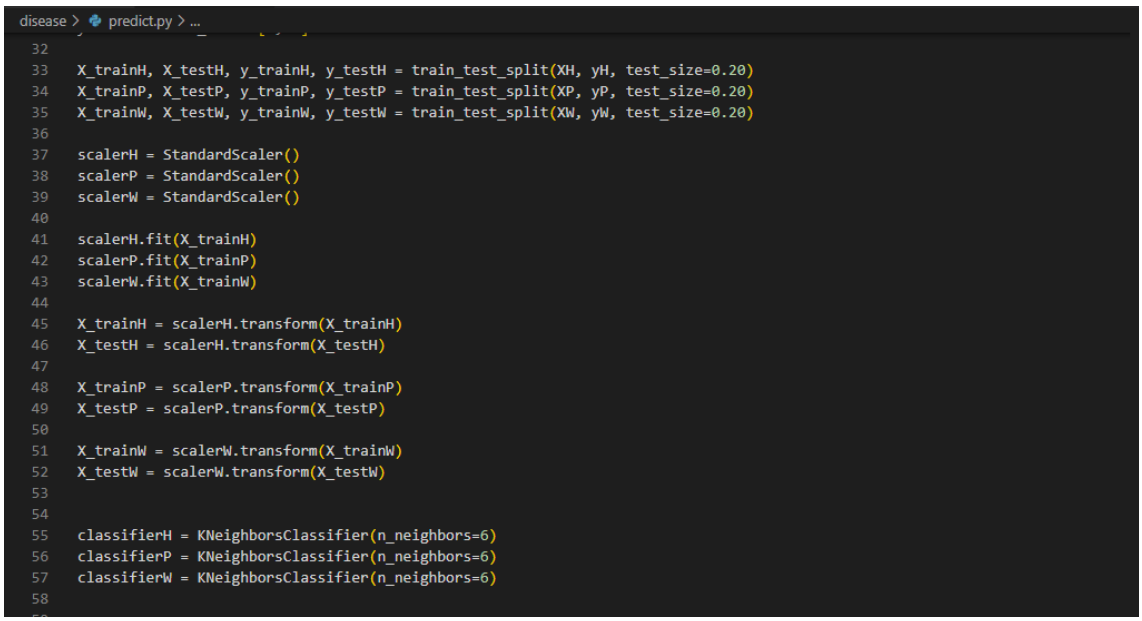


Figure 5.2: Split dataset for training and testing/scaler file creation

After importing the relevant libraries and datasets then should have to split those datasets for the training and testing. Here 80% were for training and 20% were used for testing. And that will describe in lines 33- 35. Then normalization was started, where the training model will have to be in a saving location. Then using the .fit command, Call the normalization function and make a scaler for the three diseases. After that using the .transorm method will apply the scaler file. After applying the train values and test values will be normalized according to the relevant algorithm. Here I used the KneighborsClassifier model to train the model. From this normalization, all the values are set to the 0-1 range. In 55 – 57 will import the KneighborsClassifier for the relevant 3 diseases.

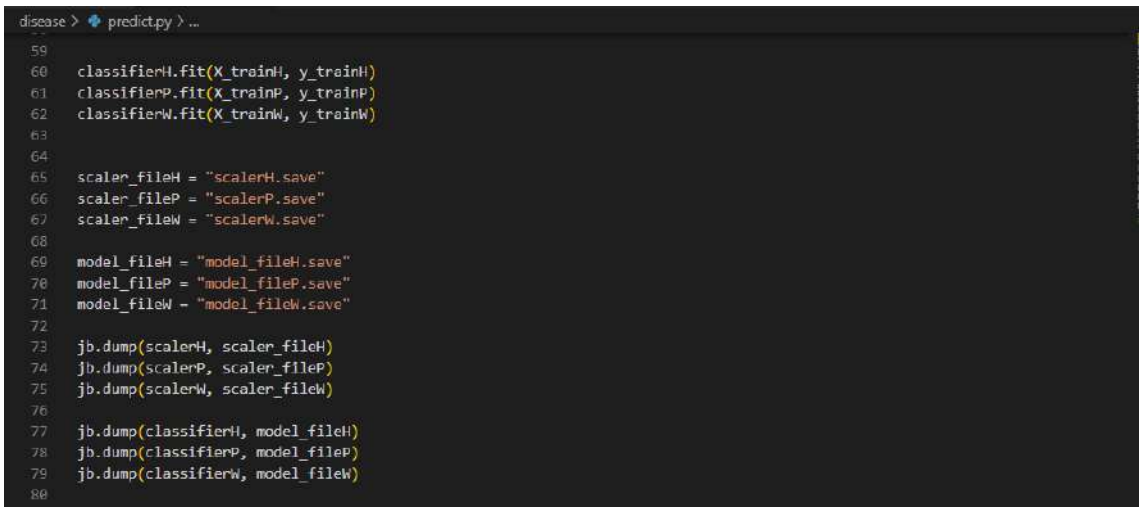


Figure 5.3: Train the classification

Then train the classification using X_train and Y_train. After that should have to define the scaler file that should have to be saved with a proper name. As shown in "schalerH.save ".

Then using the "jb. dump" will save the scaler file and model files. That's the end of the model training and scaler file creation. Without using the scaler file we are unable to run the model and couldn't see any implementation.

B. Output generate as a risk level

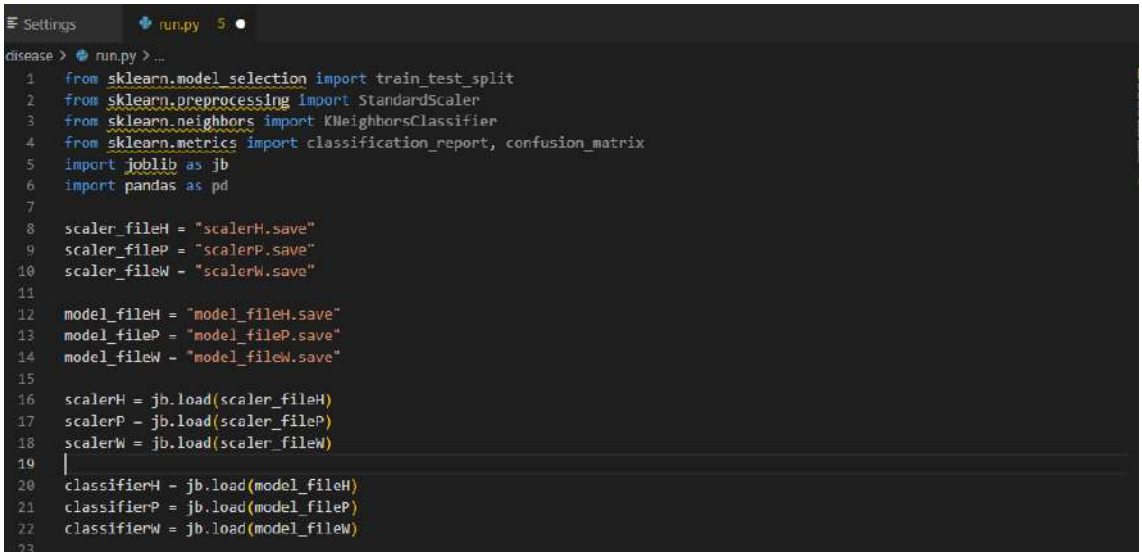


Figure 5.4: Staring getting risk level as an output

Here, by referring to the previously implemented model and scaler files the output will provide the user with 3 types of risk levels. These are High, Medium, and Low.

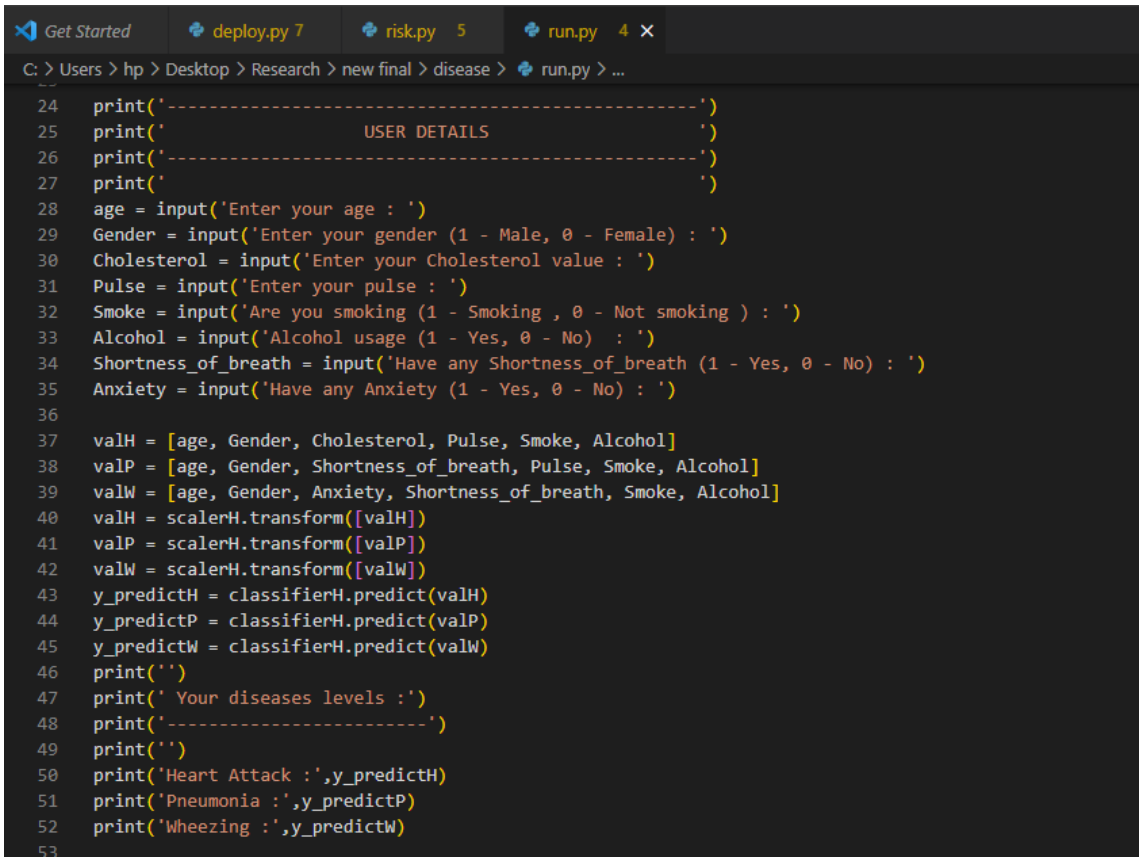


Figure 5.5: The way output should print

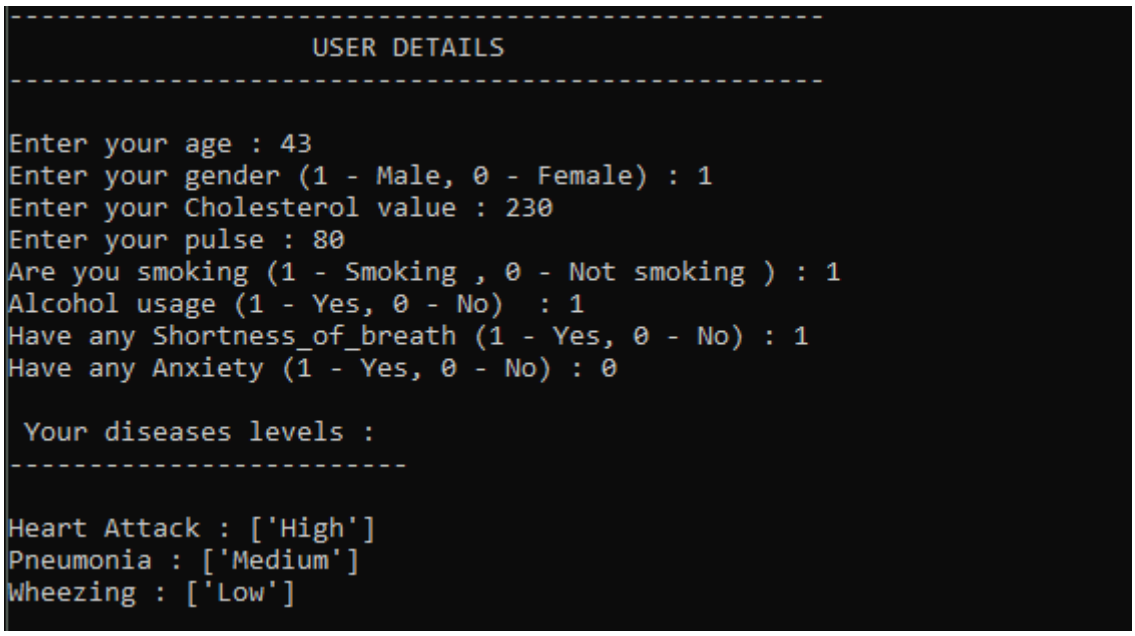


Figure 5.6: Cmd output of risk analysis

According to Figure 5.5 steps, the output will display to the user as shown in Figure 5.6. This implementation is only for 50% completion. Other progress will describe in the future.

C. Web application implementation

From this onwards will describe the individual implementation of this web application.

➢ Frontend development

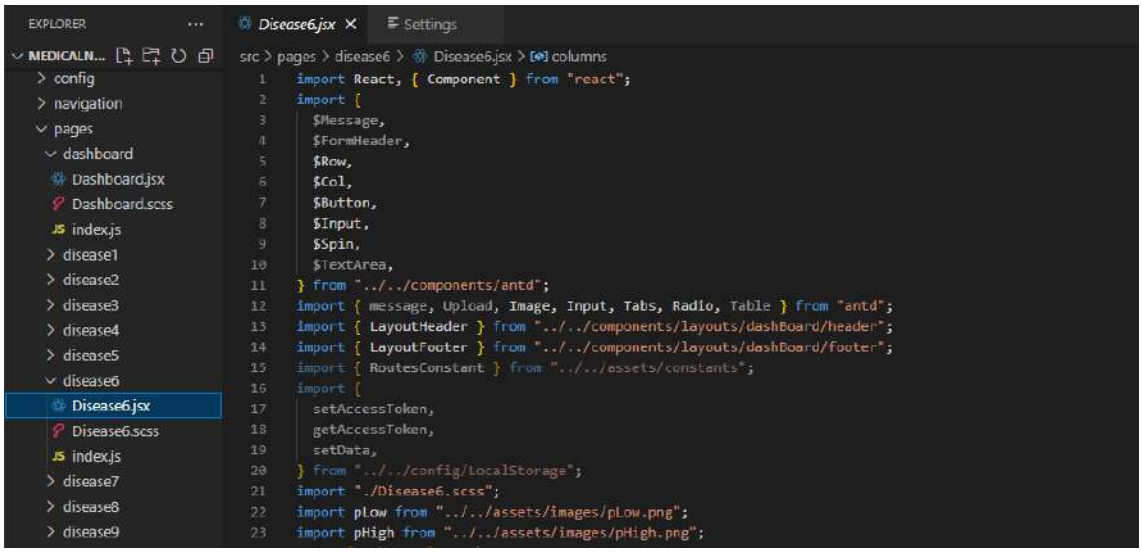


Figure 5.7: Pneumonia identification imports

For disease identification, first of all, need to import relevant libraries, files, and other implemented files in the existing folder. Here the identified disease risk suggestions will provide as an image to the user. And also style sheets imported from the outside. Not in inline implementations.

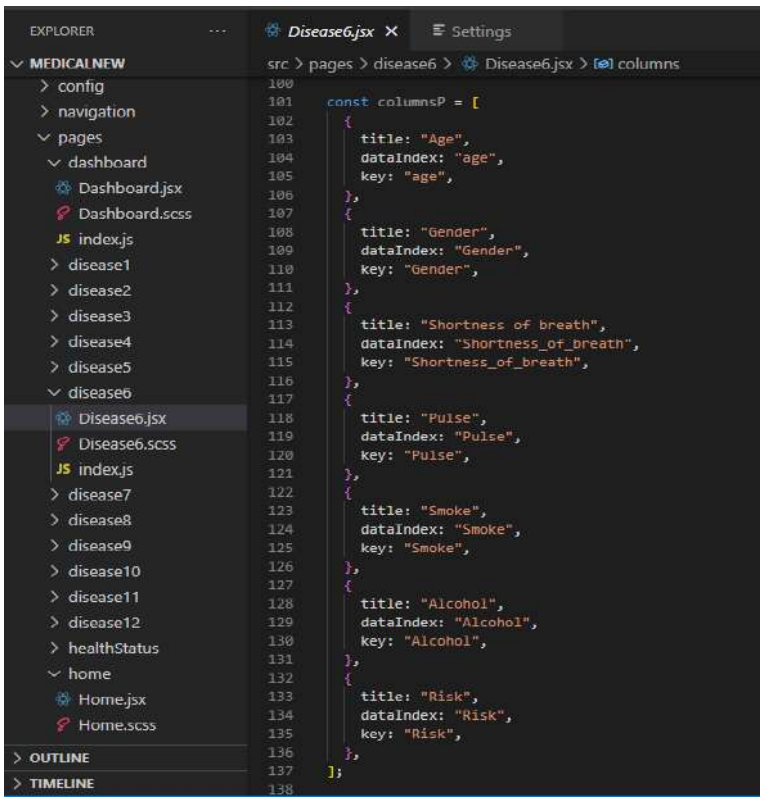


Figure 5.8: Constant Column creation

Whether there are any targeting or output features with no mappings to the source attributes, a Constant Column is generated. This is disease identification frontend development So in order to do that the relevant fields that we going to collect data from the user should be created like figure 5.8. So those are the details that I hope to collect from the user and analyze.

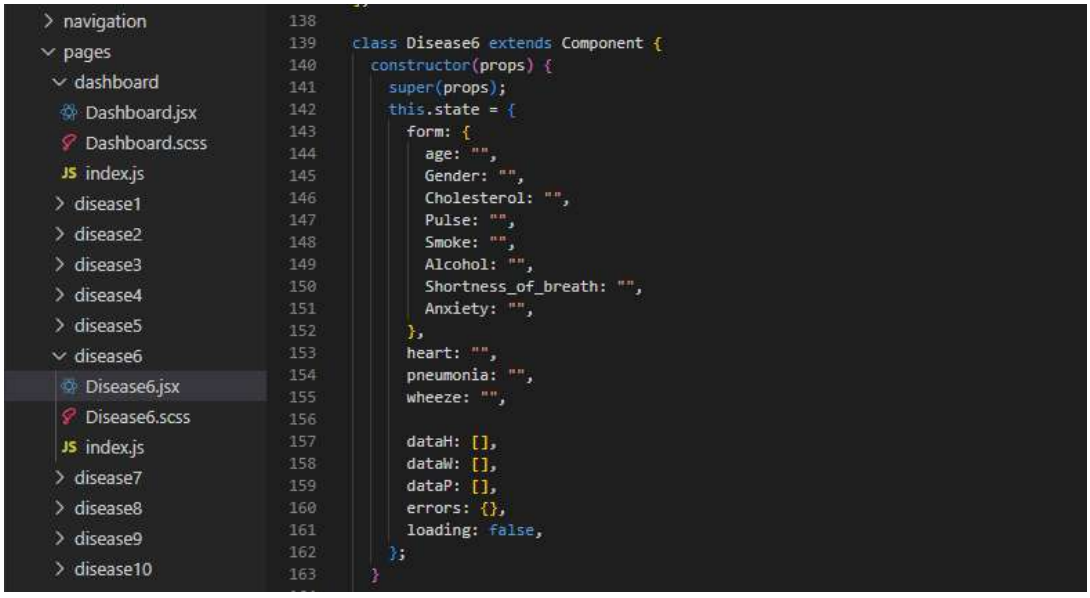


Figure 5.9: Constructor implementation

A React component's constructor is performed before the component is deployed. When writing a React component's constructor, I just need to call the super(props) function before every other statement. If the super(props) method is not called, this. Props will be unspecified in the constructor, which may result in errors.

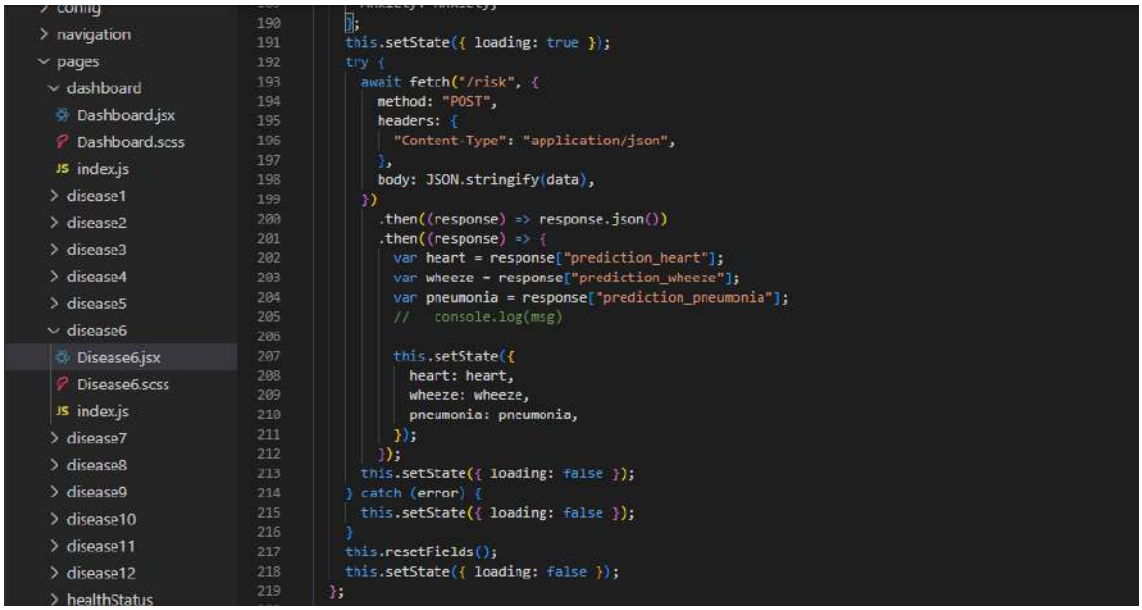


Figure 5.10: Sending data to the server

Using the POST method the data should be sent to the server. After sending, using the "body: JSON.stringify(data)," technique that transforms a JavaScript value into a JSON string. Then json(), The Response interface's json() process utilizes a Response stream and examines it to fulfillment. Then the user-entered data will be read and ready for the process to respond to those.

Figure 5.11: Display fields to the user

Above igure 5.11 will describe how the interface fields should be implemented. Like wise, all the fields user input fields were developed.

Figure 5.12: Final output implementation

This is the how final output will display to the user. The output will display as an image then the user can zoom in and zoom out to see the recommendations. Here if the user risk level is low and medium the same suggestions will provide to the user and if it is high it will display another suggestion to the user for their reference.

In this documentation, important snapshots are only included.

- **Backend development**

  From the front end entered data will be analyzed according to the trained model by usoing the below coding implementation.
  According to the GET and POST method, the data will gather from the server and return the analyzed identified risk level via frontend to the user. This is how backend development was implemented.

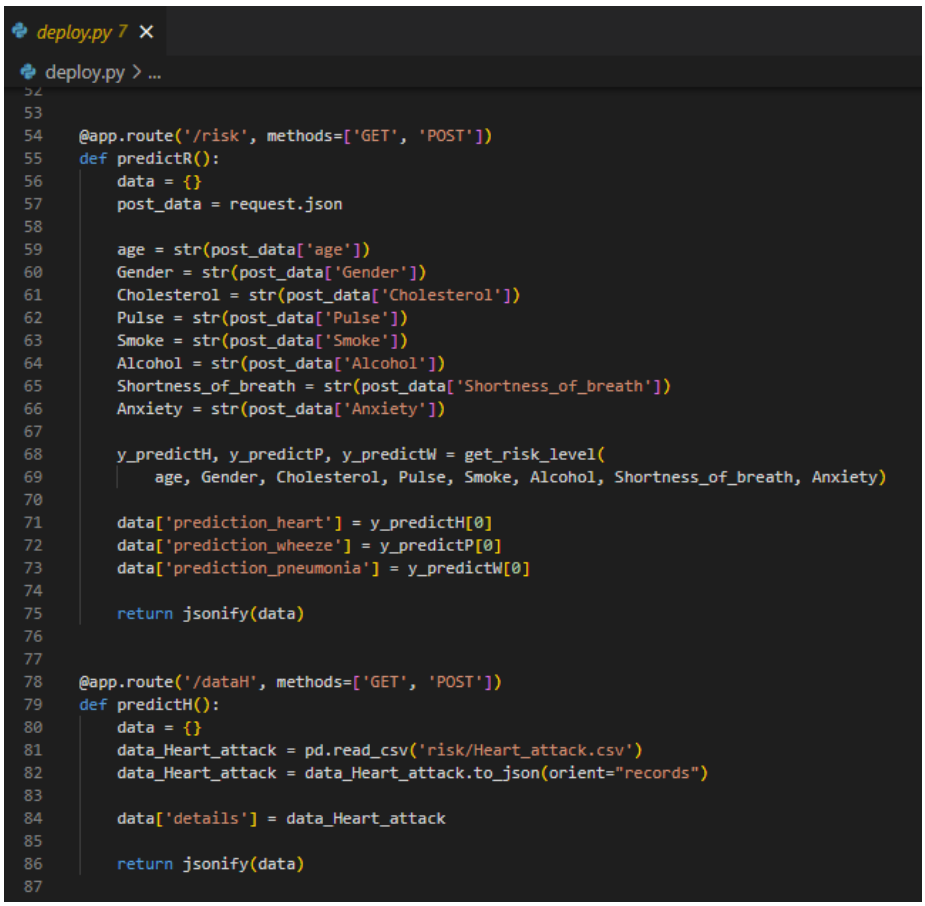


Figure 5.13: Backend development I

Figure 5.14: Backend development I

- **User Interfaces**

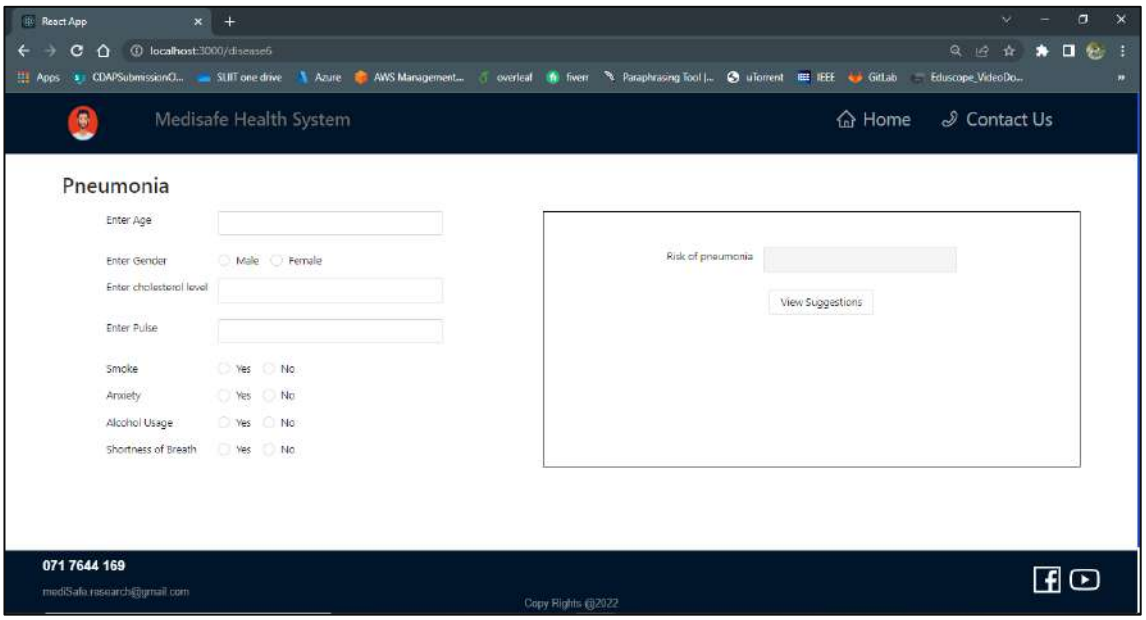


Figure 5.15: Pneumonia UI implementation

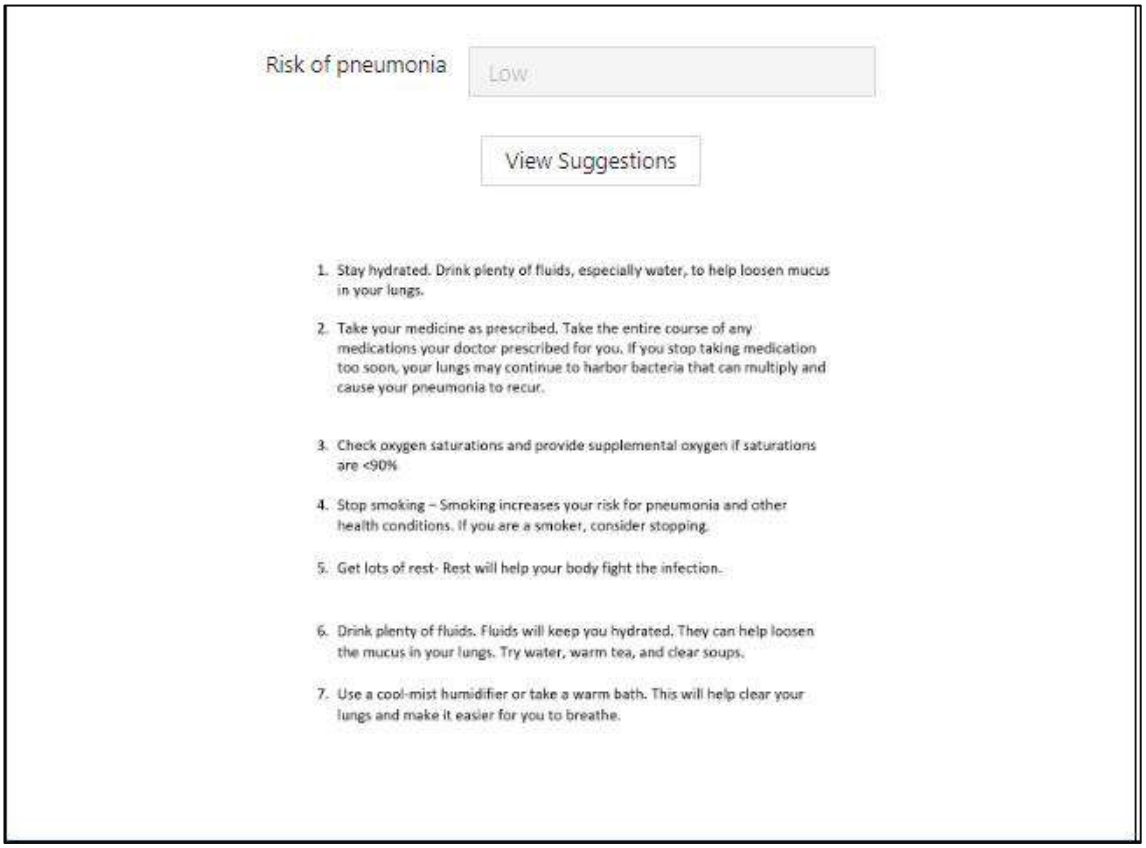


Figure 5.16: Pneumonia risk level suggestion according to the entered details

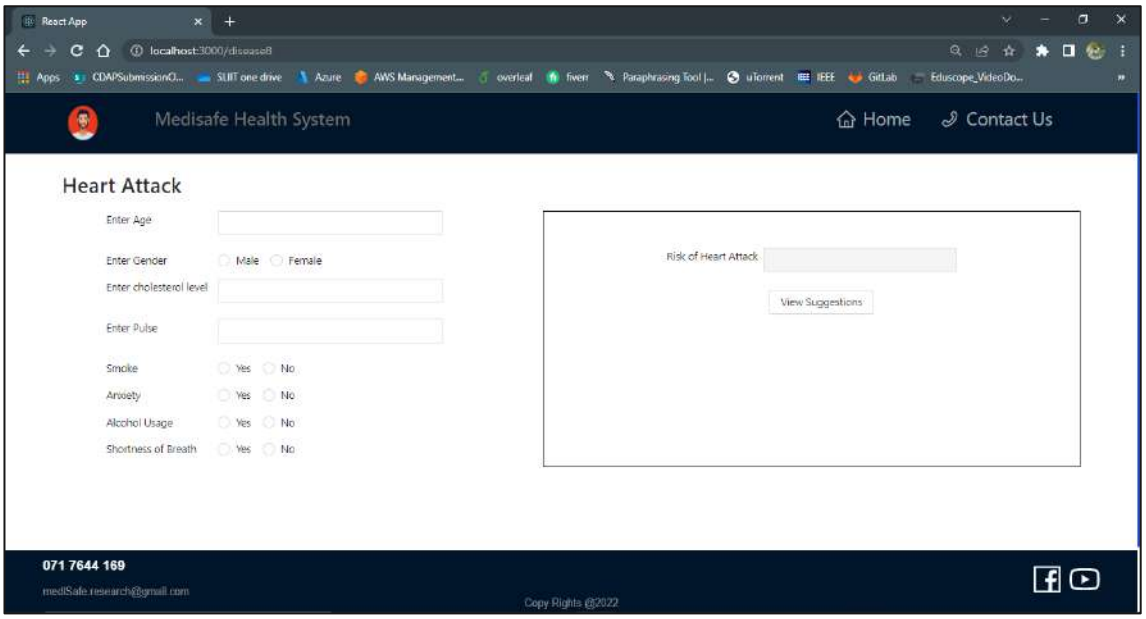


Figure 5.17: HeartAttack UI implementation

Figure 5.18: HeartAttack risk level suggestion according to the entered details

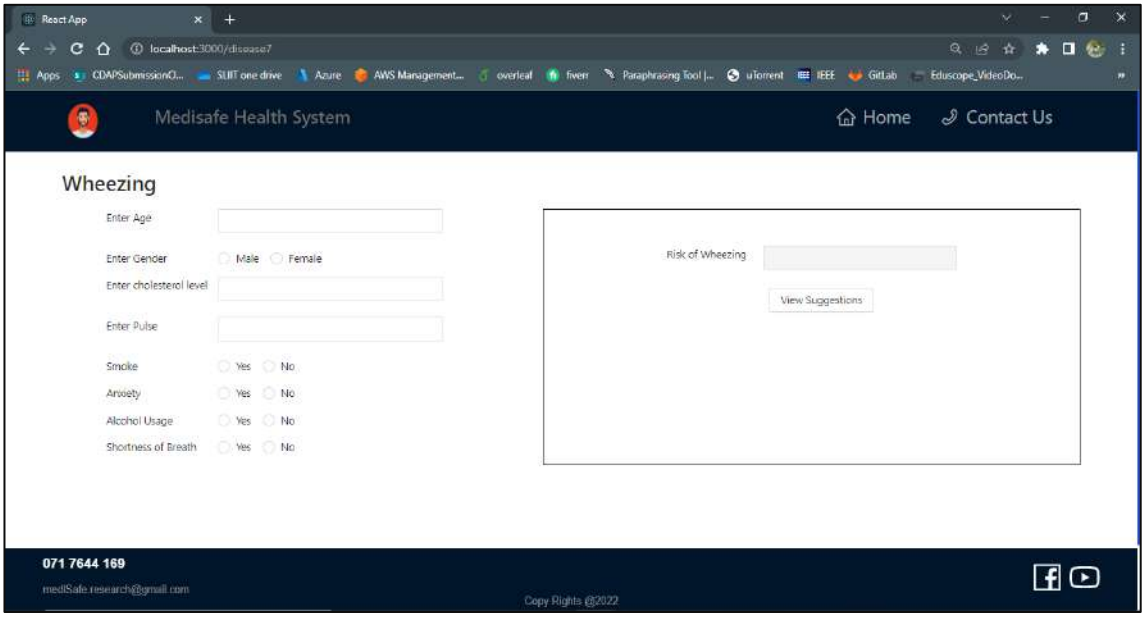


Figure 5.19: Wheezing UI implementation

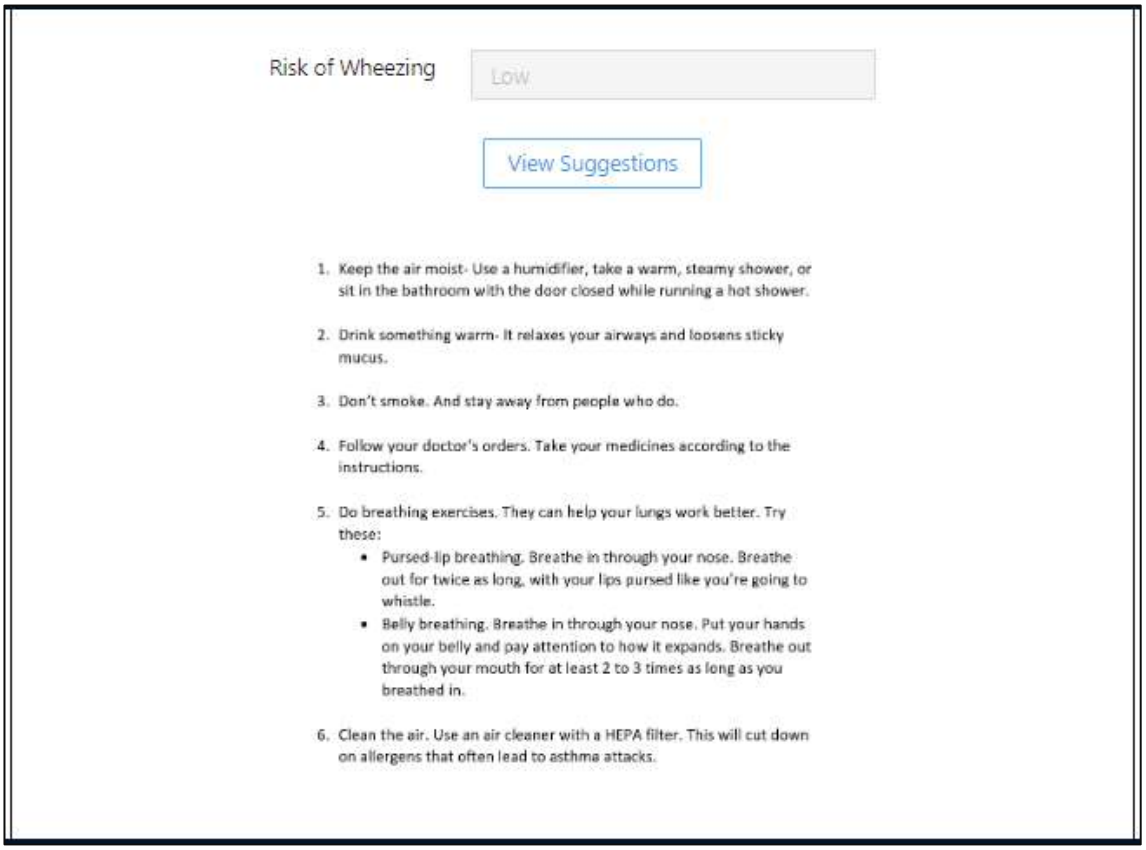


Figure 5.20: Wheezing risk level suggestion according to the entered details

Chapter 6: Result & Discussion

6.1 Result

Non-communicable diseases (NCDs) have had a rapidly increasing impact on human health in recent times and are disproportionately high in developing countries. NCDs are the leading cause of death worldwide and pose a serious public health hazard to developing countries.

This research component will provide a better solution for human beings matters who are suffering from these kinds of diseases (Heartattack, Pneumonia, Wheezing). The final outcomes are already attached above. From that, the success of this research shows how it works and how it provides solutions for ailments. The K – nearest neighbor algorithm was done as the main part of this research component and the testing was also included.

6.2 Research findings

The main intention of this system is to prevent the spread of non-communicable diseases throughout Sri Lanka. Absence of a process in this country to analyze the risk and provide certain procedures to be followed. Therefore, with the help of this system, people behave according to the instructions to slow down the risk of disease.

This component finished, in the above-identified research problem according to the research gap. This covers all the functionalities of the research problem.

6.3 Discussion

Companies create products based on the concepts of others, and those products must suit the requirements of the end users. The MediSafe technology will be extremely beneficial to users. One advantage of this strategy is that it informs people about their disease risk and introduces them to the steps to take in the case of sickness.

During the Covid pandemic, most of the patients had to stay at home and follow what are the things they know about to reduce those sicknesses. But it's not the better solution for that matter. Therefore some researchers introduced different kinds of things but all products only provide the disease is impacted or not like this only. But in this case, this will identify the diseases as risk levels and provide necessary guidelines that people can follow. So that is one of the great achievement of this research component.

Chapter 7: Conclusion

Noncommunicable diseases are becoming increasingly common in developing countries, where nearly double the infectious disease burden. If current trends continue, health systems in low- and middle-income nations will not be able to support the illness burden. The most common causes of cardiovascular disease and respiratory illness can be avoided, but immediate (preventive) action is required, and effective programs must address risk factors such as smoking, drinking, physical inactivity, and the Western diet.

We want to offer our project here to a medical center or an elderly home. We are currently carrying out the tasks assigned to us, with the permission of those institutions' officials and preparations for future work. Furthermore, by further improving this concept, we will be able to commercialize it in both domestic and international markets.

References

[1]	Who.int. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases.
[2]	P. Majumder, P. P. Ray, S. Ghosh, and S. K. Dey, "Potential effect of tobacco consumption through smoking and chewing tobacco on IL1beta protein expression in chronic periodontitis patients: In silico molecular docking study," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, pp. 1–1, 2019.
[3]	L. Hyseni *et al.*, "The effects of policy actions to improve population dietary patterns and prevent diet-related non-communicable diseases: scoping review," *Eur. J. Clin. Nutr.*, vol. 71, no. 6, pp. 694–711, 2017.
[4]	"Talking economics - an appetite for health: Regulating Sri Lanka's food environments to fight NCDs," *Ips.lk*. [Online]. Available: https://www.ips.lk/talkingeconomics/2021/02/01/an-appetite-for-health-regulating-sri-lankas-food-environments-to-fight-ncds/.
[5]	R. M. Abeyrathne, "Current trends in the popular sector traditional medicine in Sri Lanka," *Pdn.ac.lk*. [Online]. Available: https://arts.pdn.ac.lk/socio/staff/articles/abeyrathnayake%202019.pdf.
[6]	"COVID - 19 confirmed deaths - Weekly analysis," Gov.lk. [Online]. Available: https://www.epid.gov.lk/web/index.php?option=com_content&view=article&id=233&Itemid=489&lang=en
[7]	*Hopkinsmedicine.org*. [Online]. Available: https://www.hopkinsmedicine.org/news/newsroom/news-releases/johns-hopkins-medicine-researchers-identify-health-conditions-likely-to-be-misdiagnosed
[8]	M. S. Begum and R. Poonguzhali, "Noi Kanippaan: Nadi diagnosing system," 2011 International Conference on Recent Trends in Information Technology (ICRTIT), 2011, pp. 1049-1054, doi: 10.1109/ICRTIT.2011.5972318.
[9]	Z. Jiang, D. Zhang and G. Lu, "A Robust Wrist Pulse Acquisition System Based on Multisensor Collaboration and Signal Quality Assessment," in IEEE Transactions on Instrumentation and Measurement, vol. 68, no. 12, pp. 4807-4816, Dec. 2019, doi: 10.1109/TIM.2019.2899514.
[10]	N. Arunkumar and K. M. Mohamed Sirajudeen, "Approximate Entropy based ayurvedic pulse diagnosis for diabetics - a case study," 3rd International Conference on Trendz in Information Sciences & Computing (TISC2011), 2011, pp. 133-135, doi: 10.1109/TISC.2011.6169099.

[11]	H. E. J. Umasha, H. D. F. R. Pulle, K. K. R. Nisansala, R. D. B. Ranaweera and J. V. Wijayakulasooriya, "Ayurvedic Naadi Measurement and Diagnostic System," 2019 14th Conference on Industrial and Information Systems (ICIIS), 2019, pp. 52-57, doi: 10.1109/ICIIS47346.2019.9063271.
[12]	H. Pogadadanda, U. Shwetha Shankar and K. R. Jansi, "Disease Diagnosis Using Ayurvedic Pulse and Treatment Recommendation Engine," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 2021, pp. 1254-1258, doi: 10.1109/ICACCS51430.2021.9441843.
[13]	Sukesh Rao M and R. Rao, "Investigation on pulse reading using flexible pressure sensor," 2015 International Conference on Industrial Instrumentation and Control (ICIC), 2015, pp. 213-216, doi: 10.1109/IIC.2015.7150740.
[14]	A. Joshi, A. Kulkarni, S. Chandran, V. K. Jayaraman and B. D. Kulkarni, "Nadi Tarangini: A Pulse Based Diagnostic System," 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2007, pp. 2207-2210, doi: 10.1109/IEMBS.2007.4352762.
[15]	B. Thakker, A. L. Vyas and D. M. Tripathi, "Radial pulse analysis at deep pressure in abnormal health conditions," 2010 3rd International Conference on Biomedical Engineering and Informatics, 2010, pp. 1007-1010, doi: 10.1109/BMEI.2010.5639735.
[16]	Sukesh Rao M and R. Rao, "Investigation on pulse reading using flexible pressure sensor," 2015 International Conference on Industrial Instrumentation and Control (ICIC), 2015, pp. 213-216, doi: 10.1109/IIC.2015.7150740.
[17]	Yourhormones.info. (2018). Adrenaline | You and Your Hormones from the Society for Endocrinology. [online] Available at: http://www.yourhormones.info/Hormones/Adrenaline.aspx [Accessed 07 Feb. 2022].
[18]	S. S, S. S, V. S and A. R, "Health and Disease Prognosis System using Machine Learning," 2021 IEEE International Conference on Intelligent Systems, Smart and Green Technologies (ICISSGT), 2021, pp. 192-196, doi: 10.1109/ICISSGT52025.2021.00048.
[19]	D. K. Ravish, K. J. Shanthi, N. R. Shenoy and S. Nisargh, "Heart function monitoring, prediction and prevention of Heart Attacks: Using Artificial Neural Networks," 2014 International Conference on Contemporary Computing and Informatics (IC3I), 2014, pp. 1-6, doi: 10.1109/IC3I.2014.7019580.
[20]	D. Zhou, "Research on the Best Resource of Teacher's Tacit Knowledge Acquisition," 2009 Second International Symposium on Knowledge Acquisition and Modeling, 2009, pp. 319-323, doi: 10.1109/KAM.2009.225.

[21]	K. E. Demos, W. M. Kelley and T. F. Heatherton, "Dietary Restraint Violations Influence Reward Responses in Nucleus Accumbens and Amygdala," in Journal of Cognitive Neuroscience, vol. 23, no. 8, pp. 1952-1963, Aug. 2011, doi: 10.1162/jocn.2010.21568.
[22]	J. -F. Chen, L. -L. Wu and S. -C. Chou, "Effects of Journaling Dietary Intake App on the Health Outcomes of Chronic Kidney Disease Stage 3B-5," 2016 49th Hawaii International Conference on System Sciences (HICSS), 2016, pp. 3379-3388, doi: 10.1109/HICSS.2016.421.
[23]	M. Bhattacharyya, S. Maity and S. Bandyopadhyay, "Exploring the Missing Links Between Dietary Habits and Diseases," in IEEE Transactions on NanoBioscience, vol. 16, no. 3, pp. 226-238, April 2017, doi: 10.1109/TNB.2017.2654121.

Appendices

- Created poster

- Designed wireframes for web application