

**MEDISAFE – IDENTIFY MAJOR DISEASES SPREAD  
RATE IN SRI LANKA**

2022 - 143

Final Report

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

Department of Information Technology  
Sri Lanka Institute of Information Technology  
Sri Lanka

September 2022

**MEDISAFE – IDENTIFY MAJOR DISEASES SPREAD  
RATE IN SRI LANKA**

2022 - 143

Senanayaka S.A.M.A.B.M  
IT19011608

Mr. Supunya Swarnakantha

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

Department of Information Technology  
Sri Lanka Institute of Information Technology  
Sri Lanka

September 2022

## Declaration

I declare that this is my own work and this proposal does not incorporate without acknowledgement 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 acknowledgement is made in the text.

Name	Student Id	Signature
<i>Senanayaka S.A.M.A.B.M</i>	<i>IT19011608</i>	

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

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

.....

Signature of the Supervisor

.....

Date

## **Abstract**

As the world have become more civilized, a range of diseases started to spread over the world. The evolution of medicine has resulted in the cure of several diseases. Considering Sri Lanka's history, several diseases have spread throughout the country. Malaria, Hepatitis, Covid-19 and Measles are just a few examples. Vaccines can easily prevent certain diseases, whereas some diseases, such as Leptospirosis and Dengue Fever, are not vaccine-preventable Sri Lanka's health spending accounted for about 3.76 percent of the country's gross domestic product in 2018. (GDP) Despite spending a significant amount of money, Sri Lanka has yet to identify some diseases that are high risk and low risk, as well as how to avoid some diseases utilizing current technologies. Traditional ways of recording diseases are still used in the health-care sector. This research focuses on the recent spreading diseases covid -19 and dengue fever in Sri Lanka. We can discover out the number of patients, deaths, and other danger areas based on statistics from the previous year. We can predict the data for the following year by analysing past years' data. Use the machine learning module to identify various variables. Create a healthcare intelligence dashboard with analytical tools to make all of the data easier to see in graphs and tables. Even if you have not much technical understanding, this will make it easy to view information and this intelligent dashboard for healthcare also shows which months have the most infections on the diseases. To deal with forecasting challenges, a variety of prediction approaches are widely utilized. In this research, common forecasting models were applied to forecast the COVID-19 and dengue fever: linear regression (LR) and support vector machine (SVM) Finally, all of this data is easily accessible via a mobile app. As a result, the health-care sector receives a lot of predictions about disease in the coming years, and they can effectively prepare for it.

**key words:** forecast, disease, healthcare intelligence dashboard

## **Acknowledgement**

We received a lot of advice and help from several people when we first started this research project. We would like to thank the Sri Lanka Institute of Information Technology for encouraging us to finish this project and the academic staff, our supervisors for guiding us with helpful insights. Enthusiastic encourage and suggestions for improvement of this research work So, these endeavours would not be possible without my supervisor, Mrs. Ravi Supunya, and our co-supervisor. Samantha, for their immensely valuable continuous support, guidance, and encouragement given to me throughout Next, I'd like to thank my team members for their heroic efforts in ensuring that this research project a success, as well as all of my SLIIT colleagues in my batch and my SLIIT lecturers who guided and helped me in numerous ways during this time. In addition, I'd like to thank everyone who helped me by taking online surveys and contributing valuable ideas for requirement gathering and assessment. Finally, I'd like to thank my family for their financial and time assistance in carrying out this research.

## Table of content

Declaration.....	i
Abstract.....	ii
Acknowledgement .....	iii
Table of content.....	iv
List of Figures .....	vi
List of Tables .....	vi
List of abbreviations.....	vii
1. Introduction .....	1
1.1 Background & Literature survey .....	1
1.2 Research Gap .....	6
2. Research Problem .....	8
3. Research Objective .....	9
4. Methodology.....	10
4.1 System Overview.....	10
4.2 System Architecture.....	10
4.3 Flow Chart .....	11
4.4 Commercialization .....	14
5. Testing And Implementation .....	16
5.1 Testing.....	16
5.1.1 Unit Testing .....	16
5.1.2 Integration Testing.....	16
5.1.3 System Testing .....	17
5.1.4 User Acceptance Testing.....	17
5.1.5 Testing Environment .....	17
5.2 Implementation .....	18
5.2.1 Technology selection .....	<b>Error! Bookmark not defined.</b>
5.2.2 Implementation Process .....	<b>Error! Bookmark not defined.</b>
6. Result And Discussion .....	21
6.1 Result .....	<b>Error! Bookmark not defined.</b>

6.2	Discussion.....	<b>Error! Bookmark not defined.</b>
6.3	Research Finding .....	<b>Error! Bookmark not defined.</b>
7.	References .....	24
8.	Appendices.....	26

## List of Figures

Figure 1:Current Diseases .....	1
Figure 2:Spread Where Disease .....	2
Figure 3:Healthcare Dashboard.....	2
Figure 4:System Architecture .....	10
Figure 5:Flow Chart .....	11
Figure 6:Level of Testing.....	17
Figure 7:Minmaxscaler .....	18
Figure 8:lstm .....	18
Figure 9:Deploy .....	19
Figure 10:Frontend.....	20
Figure 11:spread rate.....	26
Figure 12:Graph Analysis .....	27

## List of Tables

Table 1:Comparison Table.....	7
Table 2:Accuracy Table .....	21



## List of abbreviations

Abbreviation	Description
ML	Machine Learning
LSTM	Long short-term memory
RNN	Recurrent Neural Network

# 1. Introduction

## 1.1 Background & Literature survey

As the population of the country grew, so increased the number of diseases. Doctors and health officials tried to discover a treatment based on the disorders. It took approximately a year to discover the cure for the covid-19. Furthermore, when technology advances, it aids the medical field. Nowadays, IoT and machine learning provide a significant amount of assistance to the health sector and patients. This study demonstrates how to employ machine learning in the healthcare sector for their progress.

Here are the results of a survey performed by persons who have had experience dealing with the spread of disease:

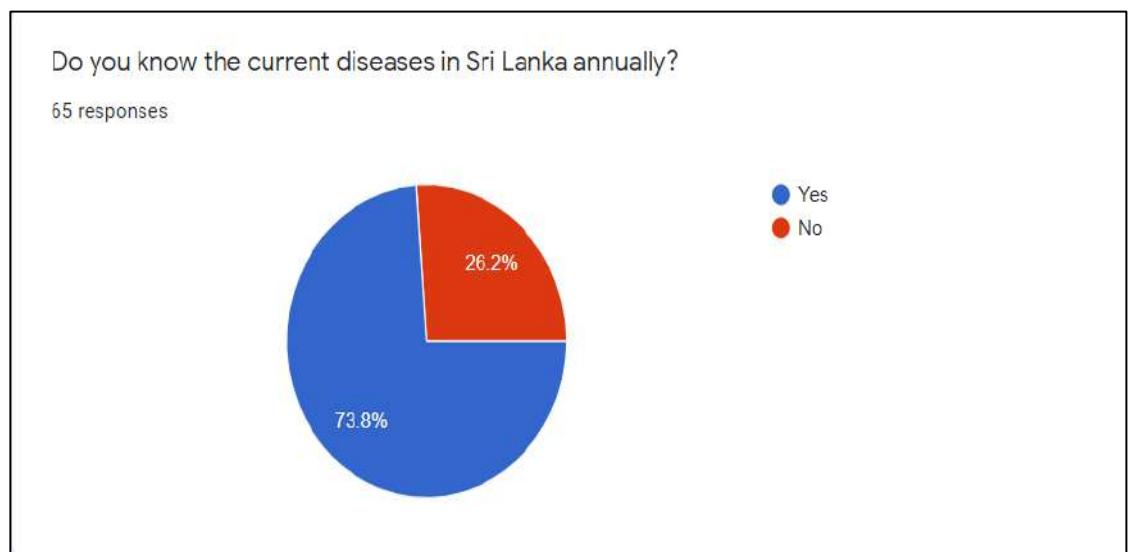


Figure 1: Current Diseases

According to the following data (Figure 1), most people are aware of current and past Sri Lankan diseases. Therefore, it is significant to be aware of disease because it enables easy identification of symptoms.

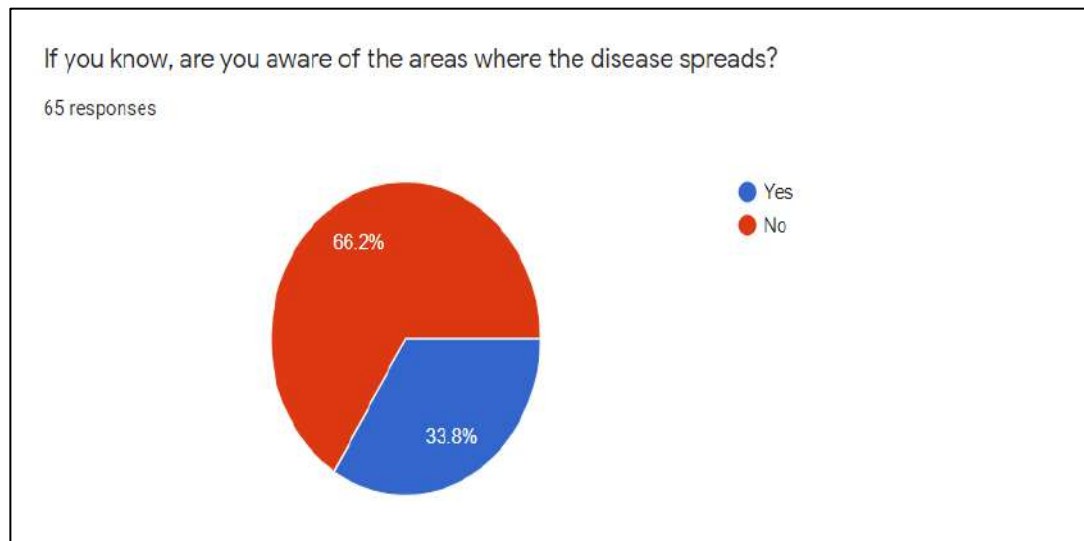


Figure 2: Spread Where Disease

Considering the data above (Figure 2), most people are aware of the disease's specifics but are unaware of its geographic distribution, which is why our research focuses on this topic.

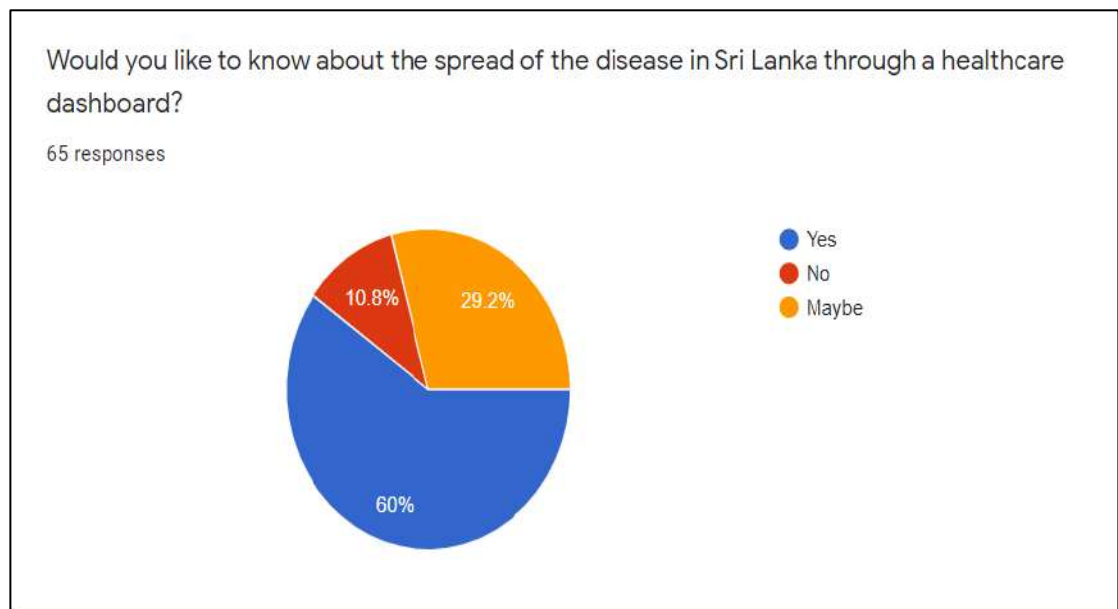


Figure 3: Healthcare Dashboard

Considering the facts above (Figure 3), most people prefer to see diseases on a single health care dashboard. This way, they can learn how diseases spread and the many details they can obtain.

One of the branches of artificial intelligence is machine learning. Machine learning can be used to make predictions. Machine learning also aids in the creation of various patterns. Calculations can be performed quickly using this method. Machine learning calculations are more accurate, faster to train, validate, and evaluate, and perform better than traditional physical calculations. ML has been used to make predictions in a variety of disciplines. And, for the most part, those forecasts have come to reality. Machine learning classifiers fall into three primary categories.

- 1)Supervised machine learning
- 2)Unsupervised machine learning
- 3)Semi-supervised learning

When creating predictions, their accuracy is important. When utilizing machine learning to make predictions, the accuracy of those forecasts is extremely high. Nowadays, almost every firm and industry use machine learning in their research and development. Medical services, financial services, government, and transportation are just a few examples.

In the year 2021, research was conducted at Daffodil International University in Dhaka, Bangladesh. Measuring the Risk of Heart Attack Using Different Types of Machine Learning Algorithms was the goal of this project [3]. The major goal of this study is to investigate the possibility of a heart attack. They proposed a model with three primary components: input data, correlation metrics, and split data, for which they applied a variety of algorithms. Bagging, logistic regression, random forest, decision tree, random forest

and logistic regression, the best result is taken into consideration. It has an accuracy of 80%. This side did more machine learning research, but they focused on how to measure the risk of heart attack using new machine learning methodologies.

A research report published in 2021 by the School of Computing Science and Engineering at Galgotias University in India reveals a study on disease prediction using big data for healthcare institutions [4]. They aimed to diagnose or anticipate diseases as soon as possible in order to save human lives. To do so, they used emerging technologies such as machine learning, the internet of things (IoT), and artificial intelligence (AI). They designed a web application that takes some symptoms as input and can successfully forecast diseases based on a fixed dataset in this study report. They use comprehensive databases to increase their disease accuracy. To obtain those symptoms, they applied the python flask framework and Java, as well as decision tree algorithms for machine learning.

The Faculty of Computer Science and Engineering at Patuakhali Science and Technology University in Patuakhali, Bangladesh researched Predicting the Probability of Covid-19 Recovering in South Asian Countries in February 2021[5]. The purpose of this study is to determine whether a healthy eating style can aid in the fight against the Corona Virus. They did so by examining ten cases (patients) and retrieving percentage forecasts from Covid-19 using three distinct types of machine learning algorithms. Three machine learning techniques are used to forecast the recovery rate: Random Forest, KNN, and SVM. They used kaggle.com to obtain the datasets. They selected several criteria from the available data, including Alcoholic Drinks, Animal Foods, Cereals - Excluding Beer, Meat, Vegetable Goods, Others, and Recovered. They observed that when a patient consumes a greater proportion of plant-based foods (e.g., cereals, vegetables) and less animal products, the restored percentage is higher; otherwise, the restored percentage is lower.

five researchers from the Regentrify Institute, Inc. Indiana University – Fairbanks School of Public Health built the Daily Representation of State COVID-19 Health Care dashboard [6]. They launched a community healthcare dashboard for corona viruses such as covid-19, influenza, and others. They employ visual design techniques such as headline graphics and a color palette to draw attention to various sections of the dashboard, as well as the filter type to select geographical locations.

This study recommends implementing a dashboard known as a healthcare analytical dashboard. Because what's going on here is about making future decisions. Present state of dashboard should also be used since this examination is based on previous data. Drill downs and pivot tables are utilized here. Using a healthcare dashboard has several advantages.

- 1)Easy to use, good navigation
- 2)Better decision making
- 3)Interactivity

As a result, in any country, having a healthcare dashboard like this is important. The health-care industry will thereafter be able to compare the most important healthcare metrics and forecast some of them. It is simple to use and simulate situations using the forecast feature to test the accuracy of predicted values. There's no need to waste time processing data with various ways.

## 1.2 Research Gap

The prediction has been made in various medical fields using machine learning. Most previous research papers in the country were done on only one disease. This research paper focuses on various diseases in Sri Lanka (malaria, covid-19, dengue)

In addition, we concentrate on high-risk and low-risk areas throughout the country, previous year's death count, infection rates, and vaccination center locations based on the COVID-19. Previous research papers did not result in the creation of healthcare dashboards. For this research paper, we created a medical dashboard for Sri Lanka in order to identify various details. We can also apply the most up-to-date machine learning techniques, such as K-Nearest Neighbours, Learning Vector Quantization, and Support Vector Machines. Also, previous research papers did not use mobile applications to view disease predictions, so in this research, we focused on implementing a mobile app for that purpose as well. Using the mobile app, you can easily view Sri Lankan disease predictions as well as medical center details. Below table display the briefly research gap.

Table 1: Comparison Table

Research products	Features			
	Healthcare dashboard	Identify the several diseases	High risk areas and death counts	Identify the geo location
Research A	✗	✗	✓	✗
Research B	✗	✗	✗	✗
Research C	✓	✗	✓	✗
Proposed system (MediSafe)	✓	✓	✓	✓



## **2. Research Problem**

Despite the reality that Sri Lanka spends a lot of money on health care, traditional data collection and analysis methods are still used. However, because the information technology industry is evolving at a breakneck pace these days, we must adapt as well. and other countries, new technologies are being used to numerous industries, including the health-care sector. Since Sri Lanka has yet to adopt cutting-edge technologies such as machine learning, its disease information is still shown in a table format on their health.gov.lk website [9]. According to the data gathered through a survey (figure 1.2), most people are unaware with the disease's spread areas and major diseases in Sri Lanka.

The approach is to develop a health-care dashboard for health-care providers, then use tools and machine learning to make predictions. It collects data and forecasts based on prior years' results. Sri Lankan doctors and the general public will then have a good understanding of how diseases are spread in Sri Lanka. Considering COVID-19's current situation, Because the government does not define where vaccination centers are located, we recommend using the Google Maps API and our mobile application to locate them.

### **3. Research Objective**

Many people are locking themselves in their homes to prevent the spread of the covid-19, which is now under control, but new covid-19 variations are emerging all the time. It's important to note that the main objective of this research is to provide patients with information and health advice based on a patient's IOTt device results, so that they can better understand their own health and avoid unnecessary hospitalizations. This is especially important for the elderly, who suffer from a wide range of diseases. In addition, the latest technologies are being used to obtain patient details and analyse the data, to detect diseases in their earliest stages. Based on information gleaned from the questionnaire and the patients' descriptions of their symptoms and conditions, suggest medical advice and instructional videos for patients Sri Lankan disease spreading information can also be accessed by anyone who wants to learn more about using our implementation. Web and mobile applications so that everyone can access any application at any time based on their preference.

I'm working on the spread rate considering the previously mentioned key points. What happens here is that several machine learning algorithms are used to forecast the details of future diseases based on historical data. Predictive data is displayed on a healthcare dashboard.

## 4. Methodology

### 4.1 System Overview

This chapter elaborates data collection methodologies, methodologies, and the implementation of the proposed system, including the variety of technologies, frameworks, libraries, programming languages, and integrated development environments (IDEs) which are used to implement the system. The previous chapter provided a brief introduction to the problem and background and discussed the literature review and the system. Additionally, this chapter will include explanatory screenshots of the architecture and code blocks.

### 4.2 System Architecture

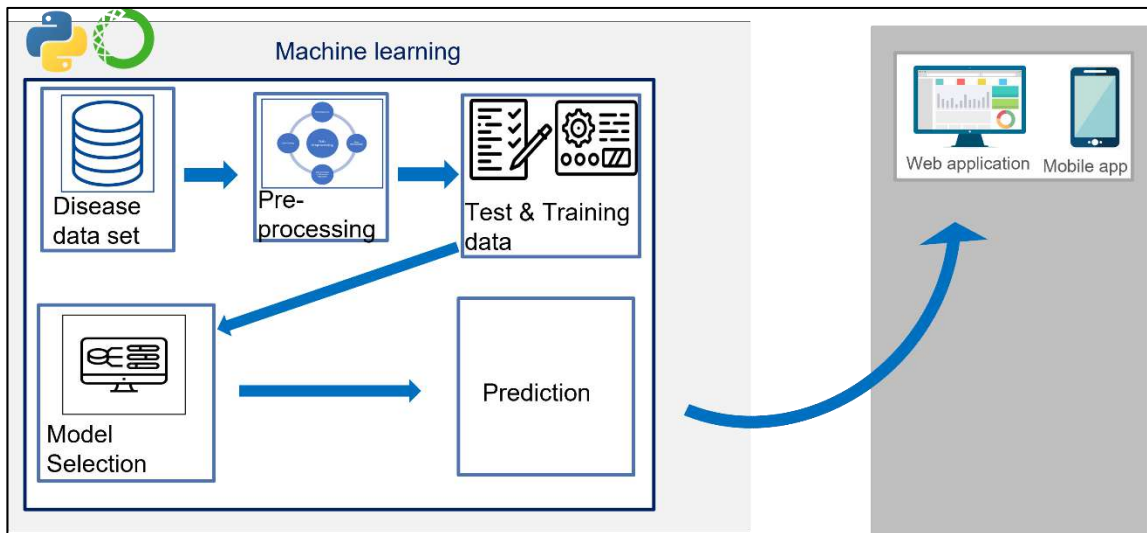


Figure 4: System Architecture

### 4.3 Flow Chart

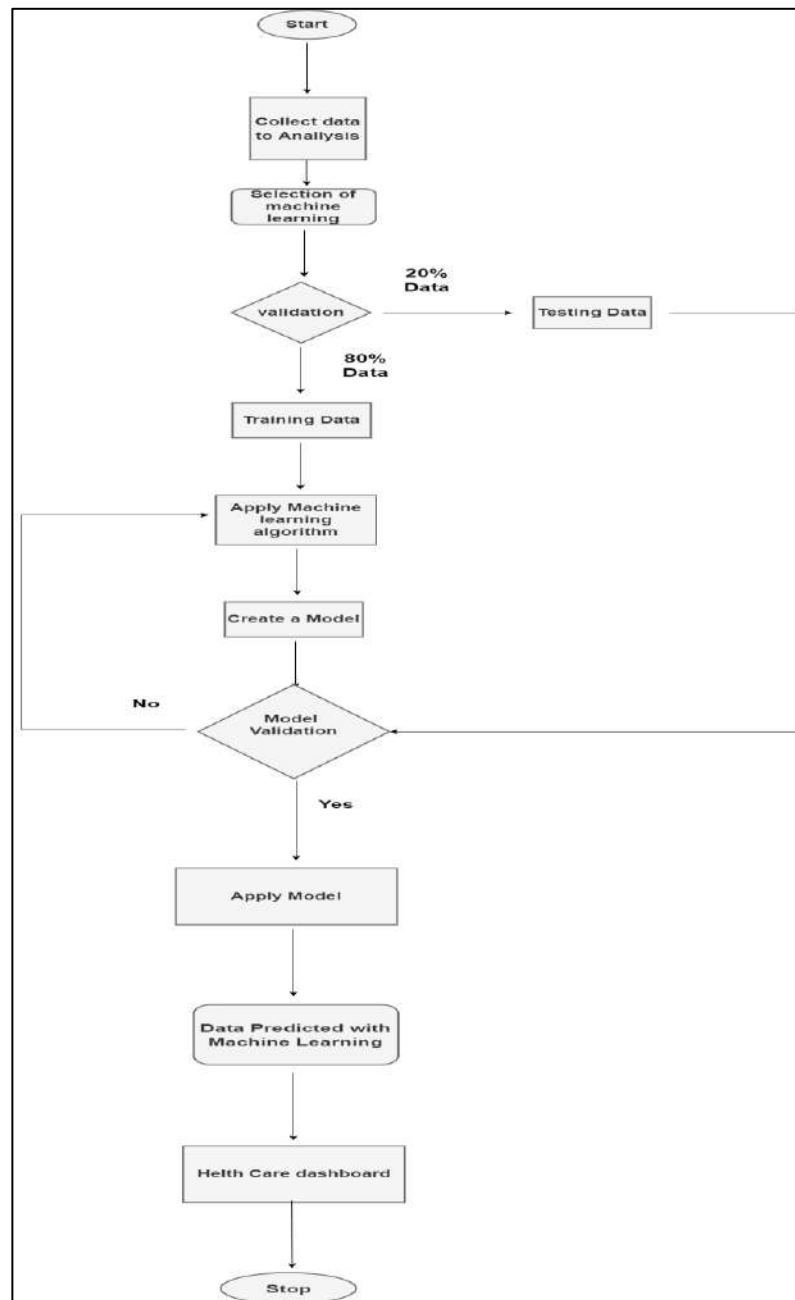


Figure 5:Flow Chart

## 4.4 Selection Technologies

### **Recurrent Neural Network**

Today, various machine learning algorithms have been introduced to handle various data types. However, forecasting sequential data is more difficult than other data types. Traditional neural networks have input and outcomes that are independent each other. In the run, what happens is based on the previous step's forecast of the current step. For instance, if we need to predict the next word in a sentence and remember the previous ones, rnn has a layer it calls the hidden Layer. This layer preserves the data as a series. then anticipate the next word

### **LSTM**

It is a special kind of Recurrent neural network, and it has the ability to learn long-term dependencies. RNN has a short-term memory problem, so it can't recall long and complex sequences. Long-term memory (LSTM) utilizes an internal mechanism produced of gates that only permits relevant data to flow along the architecture.

### **Data Preparation**

Before the values are passed as inputs, the neural network needs a shape. thereby, a three-dimensional array is required. this includes:

1. Samples
2. Timing Steps
3. Features

Before a model is fitted and a forecast is generated, the dataset is subjected to changes.so that the techniques below were used.

data transform the specific scale: Scaling the data to predictions is necessary for training a neural network, such as a Long Short-Term Memory recurrent neural network. Thus, there are two types of scaling.

1. normalization
2. standardization

for that can achieved using the scikit-learn library. When data are normalized, they are scaled down from its original range to ensure that all value of variance between 0 and 1. For normalization to work, the minimum and maximum observable values must be precisely estimated. Utilize the Scikit-Learn object MinMaxScaler to normalize the dataset. This research project uses these methods as best practices.

fit function (): this will be used to estimate the minimum and maximum observable values

transform () function: use the normalized data to train your model

### **Hyperparameter Description**

LSTM has number of hyper - parameters that can be optimized. Best known of them are

1. Epochs: The total number of training dataset iterations.
2. Time Steps: It refers to the length of the input pattern or the length of prior data contained in each input sample.
3. Loss function: the function that is supposed to be the goal of minimization. This action basically computes the difference between the actual and predicted value. Typically, MSE is chosen. Its formula is [image]
4. Optimizer: It is a function that determines the amount to reduce the lost function.
5. Batch Size and iteration:

## **Keras Functional API**

The Keras Python library allows making deep learning models convenient and easy. A variety of Keras models can be built sequentially or functionally. For the majority of issues, the sequential API enables you to build models layer by layer. Its limitation is that it prevents you from building models with many inputs or outputs or shared layer. the functional API enables you to create models that have a lot more flexibility.

### **4.5 Commercialization**

People's busy lives make it difficult to visit hospitals on a regular basis to determine whether they are infected with certain diseases. Various types of digital applications have been introduced to the world over time, but none of them have been recognized as being accurate and complete. As a result, we devised a solution that is "Medi safe". It is extremely beneficial in any pandemic situation, not just the covid. What happens in this application is that it identifies various types of lung disease and also checks whether a human is affected or not, so that we don't have to go to hospitals to get a clear picture. We also display Sri Lankan major diseases spread rate count in Sri Lanka for educational purposes because it is very helpful for medical sectors and students to get a clear idea about the diseases spread rate count in the every year and predict to the next few years so that we can get ready for that one using this future forecasting details. We intend to commercialize this application in two stages, one of which is non-paid and the other of which is paid. As a result, we have added more features to the paid version. The table below clearly shows that one.

free version limited offers includes

1. one lung disease identified
2. only identify human temperature in the free version

paid version included all features in our product. We intend to publish our product on social media once we have completed our project. so that we can create a "medisafe" page on Instagram and anyone interested in this product can use the free or paid version that they prefer.



## 5. Testing And Implementation

### 5.1 Testing

According to the Software Development Life Cycle, the testing phase is the most significant because the This phase must be completed to ensure that all requirements are met in the testing environment. Application must be tested before it is made available to users. This software testing phase is divided into four phases which are

- Unit Testing
- Integration Testing
- System Testing
- User Acceptance Testing

#### 5.1.1 Unit Testing

As a developer, you must ensure that your individual component adheres to the test cases. The biggest advantage of using unit testing is that we can find errors early on and merge the code without bugs to GitLab, allowing us to assume that there will be no errors in the production project.

#### 5.1.2 Integration Testing

Considering that our project has four major components, we must integrate those modules into a single project. And our project is made up of several software modules written by various programmers using python and react languages. The goal of this level of testing is to discover failings in the interaction of such software modules when they are combining.

### 5.1.3 System Testing

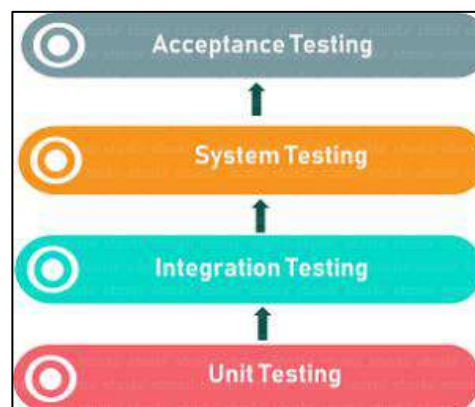
Following system testing, we must examine the entire flow of the software product, from beginning to end, as well as the desired output based on the inputs, as well as the user experience of our product.

### 5.1.4 User Acceptance Testing

Following system testing, we must examine the entire flow of the software product, from beginning to end, as well as the desired output based on the inputs, as well as the user experience of our product.

### 5.1.5 Testing Environment

This system is tested using the laptops listed below, and the web application loading process is impacted by the laptop's speed. Table view



*Figure 6:Level of Testing*

## 5.2 Implementation

Explain how to implement the key functionality at the project level in the implementation section.

```
1 scaler = MinMaxScaler(feature_range=(0, 1))
2 saved_scaler = scaler.fit(values)
3 scaled = scaler.fit_transform(values)
```

*Figure 7:Minmaxscaler*

Previously explained in the methodology section, this is the implementation of the minmaxscaler function. [Figure 7]

Explain how to apply the lstm algorithms to the dataset in this code segment. additionally, an optimizer for the decrease a loss function is used here. [Figure 8]

```
1 model = Sequential()
2 model.add(LSTM(50, input_shape=(train_X.shape[1], train_X.shape[2])))
3 model.add(Dropout(0.1))
4 model.add(Dense(9))
5 model.compile(loss='mae', optimizer='adam')
6
7 history = model.fit(train_X, train_y, epochs=10000, batch_size=4, validation_data=(test_X, test_y), verbose=2, shuffle=False)
8
9 pyplot.plot(history.history['loss'], label='train')
10 pyplot.legend()
11 pyplot.show()
```

*Figure 8:lstm*

This section of code explains the backend API call used to get the spread rate count using GET AND POST API Calls. use parameters like "disease," "city," and "date," and output JSONFIY data [Figure 9]

```
1 @app.route('/spread', methods=['GET', 'POST'])
2 def predictS():
3     data = {}
4     post_data = request.json
5
6     sickness = str(post_data['sickness'])
7     city = str(post_data['city'])
8     Date = post_data['Date']
9
10    print(sickness, city, Date)
11
12    responseArray = []
13    for x in Date:
14        response = get_prediction(sickness, city, str(x).split('T')[0])
15        responseArray.append(response)
16
17    data['details'] = responseArray
18    return jsonify(data)
```

*Figure 9:Deploy*

This section demonstrates how to retrieve data from the backend that was submitted to the frontend. the frontend state was updated using the set State method after the POST API was called and the result.[frontend.png]

```

1  submit = async () => {
2    const form = { ...this.state.form };
3    const sickness = form.sickness;
4    const city = form.city;
5    const Date = form.Date;
6
7    const data = {
8      sickness: sickness,
9      city: city,
10     Date: Date,
11   };
12   this.setState({ loading: true });
13   try {
14     await fetch("/spread", {
15       method: "POST",
16       headers: {
17         "Content-Type": "application/json",
18       },
19       body: JSON.stringify(data),
20     })
21     .then((response) => response.json())
22     .then((response) => {
23       var spread = response["details"];
24
25       this.setState({
26         spread: spread,
27       });
28     });
29     this.setState({ loading: false });
30   } catch (error) {
31     this.setState({ loading: false });
32   }
33   this.resetFields();
34   this.setState({ loading: false });
35 };

```

Figure 10:Frontend

## 6. Result And Discussion

The accuracy level for this Sri Lankan spread rate count is 85%, and to get that level of accuracy, the lstm algorithms are trained 50 times.

*Table 2:Accuracy Table*

Developed Application	Accuracy
COVID-19 Fatality Rate Classification (SMOTE) [10]	78%
Medisafe	85%

The study mentioned above demonstrate that the suggested model's accuracy was higher than other appears to work. The correct prediction has a significant impact on the model's accuracy. Among the techniques being used enhance the model's accuracy were hyperparameter optimization and data augmentation. In accordance with the information augmentation, hyper-parameter optimization operations were utilized to optimize the model. To get the best fit to the model with a high level of accuracy, parameters such as kernel size, activation function, normalization method, padding method, and other aspects that directly affect the performance and accuracy of the generated neural networks were modified. Details regarding the final accuracy report are available in

## 7. Conclusion

The primary goal of this study was to use machine learning algorithms to determine the major illness spread rate in Sri Lanka. Still, the Sri Lankan health care sector is pursuing conventional methods to overcome the difficulties. As researchers, we developed the "medisafe" to address the problem. The section of my research that is most important to the health care sector is the one where people learn how the disease spreads across the county on an annual basis. then they can discover a solution for it as well, and medical students will also have a clear understanding disease spread about Sri Lanka for their studies. LSTM Algorithms were mostly used for the spread rate section. And data set obtained from Sri Lanka's Ministry of Health website 75% of the dataset was used for training purposes. 25% were used for testing. Overall, this spread rate accuracy 85%. a number of strategies were employed for the model fitted. Used fit and transform functions as well. Finally, these backend data are sent to the frontend in json format. React is a well-known library that was utilized in the front end. The spread rate count is displayed using the react bar chart library.

As the world have become more civilized, a range of diseases started to spread over the world. The evolution of medicine has resulted in the cure of several diseases. Considering Sri Lanka's history, several diseases have spread throughout the country. Malaria, Hepatitis, Covid-19 and Measles are just a few examples. Vaccines can easily prevent certain diseases, whereas some diseases, such as Leptospirosis and Dengue Fever, are not vaccine-preventable [1] Sri Lanka's health spending accounted for about 3.76 percent of the country's gross domestic product in 2018. (GDP) [2] Despite spending a significant amount of money, Sri Lanka has yet to identify some diseases that are high risk and low risk, as well as how to avoid some diseases utilizing current technologies. Traditional ways of recording diseases are still used in the health-care sector. This research focuses on the recent spreading diseases covid -19 and dengue fever in Sri Lanka. We can discover

out the number of patients, deaths, and other danger areas based on statistics from the previous year. We can predict the data for the following year by analysing past years' data. Use the machine learning module to identify various variables. Create a healthcare intelligence dashboard with analytical tools to make all of the data easier to see in graphs and tables. Even if you have not much technical understanding, this will make it easy to view information and This intelligent dashboard for healthcare also shows which months have the most infections on the diseases. To deal with forecasting challenges, a variety of prediction approaches are widely utilized. In this research, common forecasting models were applied to forecast the COVID-19 and dengue fever: linear regression (LR) and support vector machine (SVM) Finally, all of this data is easily accessible via a mobile app. As a result, the health-care sector receives a lot of predictions about disease in the coming years, and they can effectively prepare for it.



## 8. References

- [1] Epidemiology unit, ministry of health, List of Notifiable Diseases. Available: [https://www.epid.gov.lk/web/index.php?option=com\\_content&view=article&id=145&Itemid=446&lang=en](https://www.epid.gov.lk/web/index.php?option=com_content&view=article&id=145&Itemid=446&lang=en)
- [2] "Health expenditure as a share of gross domestic product in Sri Lanka from 2009 to 2018", statista.com, Available: <https://www.statista.com/statistics/780525/health-expenditure-share-of-gdp-sri-lanka/>
- [3] M. S. Keya, M. Shamsojjaman, F. Hossain, F. Akter, F. Islam and M. U. Emon, "Measuring the Heart Attack Possibility using Different Types of Machine Learning Algorithms," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS), 2021, pp. 74-78, doi: 10.1109/ICAIS50930.2021.9395846.
- [4] S. Vijayalakshmi, A. Saini, A. Srinivasan and N. K. Singh, "Disease prediction over big data from healthcare institutions," 2021 International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), 2021, pp. 914-919, doi: 10.1109/ICACITE51222.2021.9404567.
- [5] M. S. Hossen and D. Karmoker, "Predicting the Probability of Covid-19 Recovered in South Asian Countries Based on Healthy Diet Pattern Using a Machine Learning Approach," 2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI), 2020, pp. 1-6, doi: 10.1109/STI50764.2020.9350439.
- [6] B. E. Dixon, S. J. Grannis, U. Tachinardi, J. L. Williams, C. McAndrews and P. J. Embí, "Daily Visualization of Statewide COVID-19 Healthcare Data," 2020 Workshop

on Visual Analytics in Healthcare (VAHC), 2020, pp. 1-3, doi: 10.1109/VAHC53729.2020.00007.

[8] P. Podder and M. R. H. Mondal, "Machine Learning to Predict COVID-19 and ICU Requirement," 2020 11th International Conference on Electrical and Computer Engineering (ICECE), 2020, pp. 483-486, doi: 10.1109/ICECE51571.2020.9393123.

[9] Ministry of Health , Available : [http://www.health.gov.lk/moh\\_final/english/](http://www.health.gov.lk/moh_final/english/)

[10] T. Oladunni, J. Stephan and L. A. Coulibaly, "COVID-19 Fatality Rate Classification Using Synthetic Minority Oversampling Technique (SMOTE) for Imbalanced Class," 2021 IEEE 2nd International Conference on Pattern Recognition and Machine Learning (PRML), 2021, pp. 21-26, doi: 10.1109/PRML52754.2021.9520700.

## 9. Appendices

### Spread Rate Analysis

Select Disease

Dengue

Select District

Colombo

Select the Date

2022-09-06 → 2022-10-20

Figure 11:spread rate

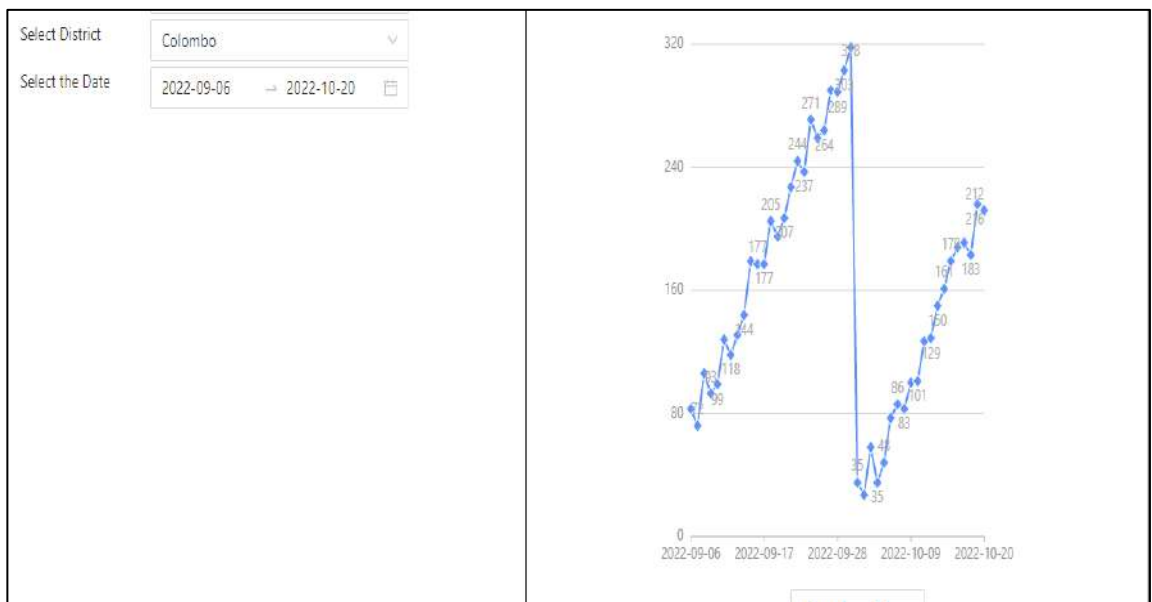


Figure 12: Graph Analysis