

**A COMPARATIVE STUDY ON PREDICTION OF DENGUE FEVER USING  
MACHINE LEARNING ALGORITHM**

**BY**

**SAIF MAHMUD KHAN DOURJOY**

**ID: 161-15-7098**

**AND**

**ABU MOHAMMED GOLAM RABBANI RAFI**

**ID: 161-15-6719**

This Report Presented in Partial Fulfillment of the Requirements for the  
Degree of Bachelor of Science in Computer Science and Engineering

Supervised By

**Zerin Nasrin Tumpa**

Lecturer

Department of CSE

Daffodil International University

Co-Supervised By

**Dewan Mamun Raza**

Lecturer

Department of CSE

Daffodil International University



**DAFFODIL INTERNATIONAL UNIVERSITY  
DHAKA, BANGLADESH  
DECEMBER 2019**

## APPROVAL

This Project titled "A Comparative Study on Prediction of Dengue Fever using Machine Learning Algorithm", submitted by Saif Mahmud Khan Dourjoy, ID No: 161-15-7098, Abu Mohammed Golam Rabbani Rafi, ID No: 161-15-6719 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 07 December, 2019.

## BOARD OF EXAMINERS



**Dr. Syed Akhter Hossain**  
**Professor and Head**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

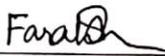
**Chairman**



**Abdus Sattar**  
**Assistant Professor**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

**Internal Examiner**



**Farah Sharmin**  
**Senior Lecturer**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

**Internal Examiner**



**Dr. Md. Saddam Hossain**  
**Assistant Professor**

Department of Computer Science and Engineering  
United International University

**External Examiner**

## DECLARATION

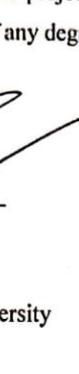
We hereby declare that, this project has been done by us under the supervision of **Zerin Nasrin Tumpa**, Lecturer Department of CSE Daffodil International University. We also declare that neither this project nor any part of this project has been submitted anywhere else for award of any degree or diploma.

Supervised by:



**Zerin Nasrin Tumpa**  
Lecturer  
Department of CSE  
Daffodil International University

Co-Supervised by:



**Dewan Mamun Raza**  
Lecturer  
Department of CSE  
Daffodil International University

Submitted by:



**Saif Mahmud Khan Dourjoy**  
ID: -161-15-7098  
Department of CSE  
Daffodil International University



**Abu Mohammed Golam Rabbani Rafi**  
ID: -161-15-6719  
Department of CSE  
Daffodil International University

© Daffodil International University

ii

## ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to **Zerin Nasrin Tumpa**, Lecturer, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “*Machine Learning*” to carry out this project. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Dr. Syed Akther Hossain, Professor and Head**, Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patients of our parents.

## **ABSTRACT**

Dengue is one of the most common viral fever to the people. This is also known as life threatening disease. Dengue has become more and more evident this year in Bangladesh. It has taken the lives of many in our country. And the number of dengue fever patients is increasing day by day. There are many people is at risk from dengue. Early forecast of dengue can spare individual's life by cautioning them to take legitimate conclusion and care. But it is difficult to say in advance whether this will happen or not. The aim of this piece of research work is to analysis the symptoms of dengue fever and early prediction of the symptoms that can be seen in years ahead. For predicting the symptoms two different machine learning algorithms have been used. Support vector machine (SVM) and random forest classifier algorithm have been used. Finally the accuracy of these two has been evaluated and the confusion matrix has been shown. And then we have talked about the algorithm which is better for our dataset.

## TABLE OF CONTENTS

<b>CONTENTS</b>	<b>PAGES</b>
Board of examiners	i
Declaration	ii
Acknowledgement	iii
Abstract	iv
<b>CHAPTER</b>	
<b>CHAPTER 1: INTRODUCTION</b>	<b>1-5</b>
1.1 Introduction	1
1.2 Motivation	2
1.3 Rationale of the Study	2
1.4 Research Questions	3
1.5 Expected Output	3
1.6 Report Layout	4
<b>CHAPTER 2: BACKGROUND</b>	<b>6-8</b>
2.1 Introduction	6
2.2 Related Works	7
2.3 Research Summary	7
2.4 Scope of the Problem	8
2.5 Challenges	8
<b>CHAPTER 3: RESEARCH METHODOLOGY</b>	<b>9-15</b>
3.1 Introduction	9
3.2 Research Subject and Instrumentation	9
3.3 Data Collection Procedure	10
3.4 Statistical Analysis	11
3.5 Implementation Requirements	14

<b>CHAPTER 4: EXPERIMENTAL RESULTS AND DISCUSSION</b>	<b>16-24</b>
4.1 Introduction	16
4.2 Experimental Results	16
4.3 Descriptive Analysis	23
4.4 Summary	24
<b>CHAPTER 5: SUMMARY, CONCLUSION, RECOMMENDATION AND IMPLICATION FOR FUTURE RESEARCH</b>	<b>25-26</b>
5.1 Summary of the Study	25
5.2 Conclusions	25
5.3 Recommendations	26
5.4 Implication for future scope	26
<b>REFERENCES</b>	<b>27-28</b>
<b>APPENDICES</b>	<b>29</b>

## LIST OF FIGURES

<b>FIGURES</b>	<b>PAGE NO</b>
Figure 1.5.1 Expected Output	4
Figure 2.1.1 Months of highest number of dengue cases in 6 years	6
Figure 3.3.1 Data collection procedure	10
Figure 3.4.1 Gender ratio of patients	11
Figure 3.4.2 Age distribution of patients	12
Figure 3.4.3 Number of occurrence	13
Figure 3.4.4 Number of hospitalized patients	14
Figure 3.5.1 Work methodology	14
Figure 3.5.2 Used library	15
Figure 4.2.1 Result using SVM algorithm	16
Figure 4.2.2 Predicted symptoms for 2020 using SVM	17
Figure 4.2.3 Result using RFC	27
Figure 4.2.4 Predicted symptoms for 2020 using RFC	17
Figure 4.2.5 Accuracy of SVM	18
Figure 4.2.6 Accuracy of RFC	18
Figure 4.2.7 Confusion Matrix using SVM	22
Figure 4.2.8 Prevalent symptoms in bar chart for 2020	23
Figure 4.2.9 Prevalent symptoms in pie chart for 2020	23

## **LIST OF TABLES**

<b>TABLE</b>	<b>PAGE NO</b>
Table 3.4.1 Gender distribution	12
Table 4.2.1 Accuracy of algorithms	18

# CHAPTER 1

## INTRODUCTION

### 1.1 Introduction

Dengue is viewed as one of the most significant arboviral diseases on the planet which is transmitted by Aedes mosquitoes, principally Aedes (Stegomyia) aegypti and furthermore Ae. Alpopictus. . This is also known as life threatening disease nowadays. The disease causes influenza like ailment, and once in a while forms into a possibly deadly inconvenience or extreme dengue, for example, dengue haemorrhagic fever and dengue shock syndrome. Dengue is found in tropical and sub-tropical atmospheres around the world, for the most part in urban and semi-urban territories. Extreme dengue is a main source of genuine disease and passing among kids in some Asian and Latin American nations. There is no particular treatment for dengue but early prediction of dengue can protect person's life and they can drive for taking appropriate finding and care.

Dengue is a mosquito-borne disease found in tropical and sub-tropical locales around the globe. Lately, transmission has expanded prevalently in urban and semi-urban zones and has turned into a significant worldwide general wellbeing concern. Serious dengue was first acknowledged in the 1950s while dengue rises in the Philippines and Thailand [1] . Nowadays, serious dengue influences most Asian and Latin American nations and has turned into a main source of hospitalization and passing among children in these districts.

The occurrence of dengue has developed drastically around the globe in late decades. Over 2.5 billion individuals – over 40% of the total populace – are currently in danger from dengue[1].But in Bangladesh the number of dengue patients have been increasing yearly. Since January 1 till October 5, upwards of 89,616 individuals have been influenced by dengue [2] . So a large number of people has been affected by dengue this year. And many people have died of dengue in 2019. Dengue ailments is realized

by four immovably related diseases of infections known as-DEN-1, DEN-2, DEN-3, and DEN-4 [3] [4] [5] [6] .

The DEN-5 has been presented in 2013 [6] . Nearby are three class characterizations of dengue fever arranges that are- [7]

1. (DF) Dengue fever - In DF stage body temperature increases day by day and patient will feel Headache, Pain in muscle, Itching.
2. (DHF) Dengue Hemorrhagic Fever - In DHF stage body temperature decreases and minor bleeding from patients nose, gums, skins etc.
3. (DSS) Dengue Shock Syndrome - In DSS stage body temperature fluctuating and vomiting accompanied by the flow of minor blood.

Our aim is to develop a model for predicting symptoms of dengue using some Machine Learning Algorithms and increasing the awareness of the people of our country and reducing the number of deaths affected by dengue.

## **1.2 Motivation**

One of the most consuming subject or episode of this current year is the monstrous breakout Of the Dengue fever. Dengue Fever exhibits a huge and developing weight of infection to endemic nations, where kids are at specific hazard. Around the world, no viable enemy of viral treatment has been recognized, hence vector control is key for malady avoidance. In such case our rationale in the postulation was to –

1. Make Participants mindful of dengue side effects.
2. Ensure transmission and hatchlings destruction techniques.
3. Eradicating Misconceptions about the conduct of the *Aedes aegypti* mosquito and disarray with other mosquito-borne infections affected safeguard practice.
4. Reduce Community-wide resistance and consolidate individuals to battle together Against Dengue.

5. Ensuring individual activities just as to guarantee steps taken by the Government and wellbeing focuses.

### **1.3 Rationale of the Study**

The effect of our exploration on dengue has a long haul impact on the. Our exploration

territories are-

1. Develop an extensive early cautioning and observation framework that has prescient capacity for pandemic dengue and advantages from powerful instruments for lab conclusion and vector checking.
2. Understand and foresee the danger of worldwide spread of dengue inside the setting of parameters of vectorial limit, worldwide portability, and environmental change.
3. Enable individuals to comprehend the indications and make strides as right on time as conceivable Against dengue.

### **1.4 Research Questions**

While doing the research on Dengue fever some questions arose in my mind. And the questions are-

1. What are the reasons of occurrence of the disease?
2. What are the symptoms of dengue fever?
3. What are the available treatment for dengue?
4. What should an individual do if I suspect I have dengue?
5. Where do the mosquitoes breed?
6. What can be done to reduce the risk of dengue?
7. What type of personal and household protection must be taken to prevent dengue?

### **1.5 Expected Output**

In this research developing a model for predicting the dengue fever symptoms is being tried using some different Algorithms of Machine Learning. And also our aim is to make people more conscious and make people more aware about dengue fever and it's symptoms so that people can go for a proper diagnosis and treatment. Our expected output is a list of some symptoms which can be seen in the next years. From that a sample of our expected output is given below

```
['Headache', 'Body_pain', 'Aversion', 'Feeling_dizzy']
```

Figure 1.5.1 Expected Output

## **1.6 Report Layout**

### **Chapter 1: Introduction**

In this part, we've talked about the motivation of ours for doing this research and also the rationale of the study, research questions and the expected outcome of the research.

### **Chapter 2: Background**

In this chapter, we will discuss about the related work of our research and comparative studies between our developed model and existing related model as well as the problems and challenges that we've faced.

### **Chapter 3: Research Methodology**

In this chapter, we will discuss about our research methodology that means we will talk about how we did our research and about the terminology.

### **Chapter 4: Experimental results and Discussion**

In this chapter, we will discuss about our research experimental results and outputs. We will talk about which results we have found by implementing the algorithms of machine learning over our collected data.

## **Chapter 5: Summary, Conclusion, Recommendation and Implication for Future Research**

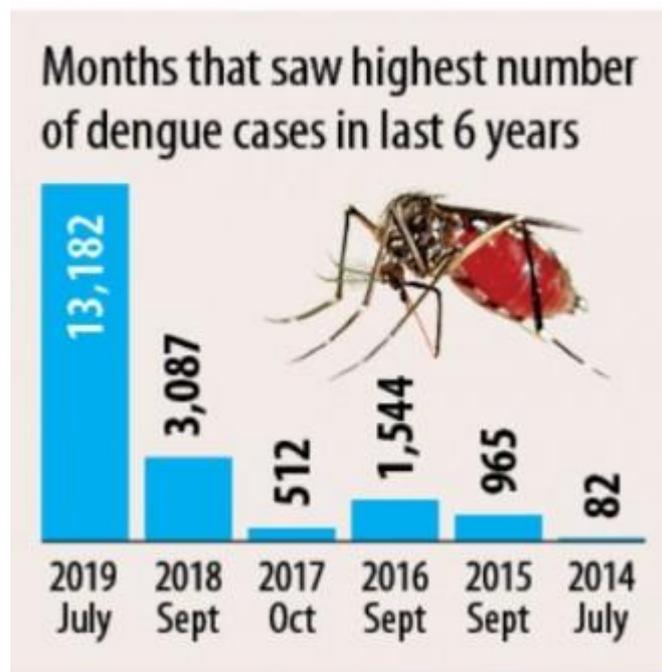
In this chapter, we will discuss about summary, conclusion, recommendation and scope for further development of our research

## CHAPTER 2

### BACKGROUND

#### 2.1 Introduction

Machine Learning empowers IT frameworks to perceive designs based on existing calculations and informational collections and to create sufficient arrangement ideas. Its algorithms are very powerful for predicting any outcome over the given data or information. That's why we have used machine learning algorithms for prediction. Dengue has become as a threat for our life. It took too many lives in this year from our Bangladesh. Not only for our country it has become a threat for the whole world. WHO at present gauges there might be 50–100 million dengue contaminations worldwide consistently [1]. In Bangladesh in the month of July of 2019 has reported the highest dengue patients.



**Source-** Dengue outbreak: 1,649 hospitalized in 24 hours (The Daily Star) [8]

Figure: 2.1.1 Months of Highest Number of Dengue Cases in 6 Years

## 2.2 Related Works

One of the most consuming subject or episode of this current year is the monstrous breakout Of the Dengue fever. Dengue Fever exhibits a huge and developing weight of infection to endemic nations, where kids are at specific hazard. Around the world, no viable enemy of viral treatment has been recognized, hence vector control is key for malady avoidance.

In order to predict the dengue fever Iqbal N., et al has used the dataset which is made with clinical data, Genes, climate factors. In their research using this dataset and Tree based model, Neural based model, Evolutionary based classifier, Ensemble classifier algorithms they have predicted whether or not dengue will occur [5]. There are many research for predicting dengue fever but Jain R., et al made a dataset using meteorological data, lag variable of surveillance of the disease, clinical data, socioeconomic data and used algorithms were Generalized Additive Models, R-squared for prediction [7]. For predicting the dengue fever Guo P., et al used Demographic data, clinical data, laboratory data, dengue cases, Meteorological data, Search and query data as data and Support vector machine (SVR), gradient boosted regression model (GBM), step-down linear regression model, negative binomial regression (NBM), LASSO model, GAM model, RMSE as algorithm [8]. On the other hand Rahim N. F., et al predicted dengue fever using Real data on dengue and weather data as data to make dataset and Decision Tree, Support Vector Machine, Random Forest, CN2, Naïve Bayes as algorithm for prediction [9]. For the purpose of predicting dengue fever Sanjudevi R., et al made a dataset using dengue cases and used Support Vector Machine (SVM), SMO, Decision Tree algorithms for detection and prediction [10]. Muhilthini P., et al with the intention of predicting dengue fever used Gradient Boosting Regression (GBR), Mean Square Error algorithms over Temperature, Population, Rain Fall, Urbanization, Dengue Cases data [11]. With the objective to predict dengue fever Choi Y., et al took help from Environment, Temperature, Rain Fall, Humidity data and Generalizes Linear Model, Negative Binomial Regression (NBR) algorithms [12].

### **2.3 Research Summary**

Developing a model for doing the prediction of dengue fever symptoms is being tried which can be seen in future years using some different and effective algorithms of machine learning. There are some related works which were done using different types of algorithms over their different types of collected data and information. Someone has used clinical data, someone has used weather data and another one has used environmental data to predict dengue using different techniques. A study by [13] in Cambodia, by [14] in Taiwan concluded that weather is an effective predictor for dengue cases and outbreak. In this research the symptoms or syndromes are being used in our data set to predict dengue using some machine learning technique like Support Vector Machine (SVM), Random Forest algorithm. After that the cross validation will be used for differentiate the accuracy rate of different Machine Learning Algorithms. And finally will be used the Evaluation Metrics to evaluate the performance of the Algorithms.

### **2.4 Scope of the Problem**

In every research there can be a scope of problem. So there is no exception in our work too. Our collected data should be appropriate and our data set should be smooth to do work with. We need to train our collected data properly while working with algorithms. If we can't do this properly our output won't be appropriate and won't be similar to the real life scenario. If we want to reduce the problem we should work properly and carefully.

### **2.5 Challenges**

No task can be accomplished without challenges. While doing any work, challenges could come in any steps of the work. In our research the main challenge we had to face while collecting data from the dengue patients. That means in our research surveying data from patients and collecting the information is the main challenge. Collection of proper and appropriate data is needed from patients for building a

smooth dataset. So we can say that the main challenge is that the collected data should be more appropriate for our research for finding the better output.

## **CHAPTER 3**

### **RESEARCH METHODOLOGY**

#### **3.1 Introduction**

This section gives a framework of research techniques that were followed in the investigation. It gives data on the members, that is, the criteria for incorporation in the examination, who the participants were and how they were tested. We depict the examination structure that was picked with the end goal of this investigation and the purposes behind this decision. The instrument that was utilized for information accumulation is likewise depicted and the methodology that were pursued to do this examination are included. We additionally talk about the strategies used to break down the information. In conclusion, implementation requirement that were followed in the process are likewise talked about.

#### **3.2 Research Subject and Instrumentation**

Our research for prediction of dengue fever symptoms. We are doing this using some different machine learning algorithms. So our research subject is - A Comparative Study on Prediction of Dengue Fever Using Machine Learning Algorithm.

In research of human subjects, a survey is a rundown of questions planned for separating explicit information from a specific gathering of individuals. Surveys are

directed by telephone, mail, by means of the web, and now and again eye to eye . Surveys are utilized to expand learning in fields, for example, social research and demography. For our research structural questionnaires was first prepared, it was pre-tested. After finalization it was used as research instrument. So in our research work surveys questionnaires are the main instrument which can be said as research instrument. After surveying the collected data has been generalized for using some preprocessing techniques. After that the preprocessed data has been prepared for our research work. Than the prepare data has been used into machine learning (ML) algorithms for predicting our desired dengue symptoms .

### 3.3 Data Collection Procedure

Data collection is one of the most important tasks in our research work. Data collection or information accumulation is the way toward social event and estimating data on factors of enthusiasm, in a built up precise design that empowers one to respond to expressed research questions, test speculations, and assess results. The information accumulation part of research is basic to all fields of study including physical and sociologies, humanities, business, and so forth. While strategies differ by discipline, the accentuation on guaranteeing exact and fair accumulation continues as before. It is one of the toughest things that we have faced. There are many ways to collect information and data but We have collected our data in two different ways. They are-

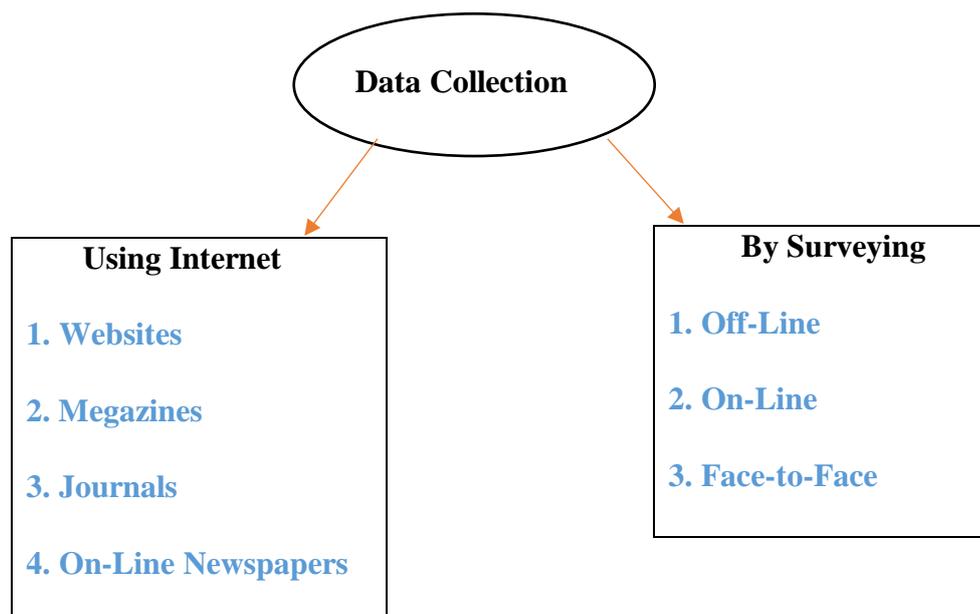


Figure 3.3.1 Data Collection Procedure

### Surveys:

We have surveyed over more than 500 hundreds patients for collecting data and finding Information. We made a question paper including the questions that are necessary for our research work. The question paper was consist of with questions like – Age, Gender, How many times they have been infected with dengue fever, What were the symptoms they faced, Duration of staying at the hospital, Duration of recovery and some other questions. By collecting their data we fill it into our dataset.

### Internet:

We have taken further support from internet like websites, magazines, journals, on-line newspapers for enriching our dataset and to make bigger dataset.

### 3.4 Statistical Analysis

After surveying we have found some analytical statistic data by which some differentiation can be seen and these are shown in below-

#### Gender Ration:

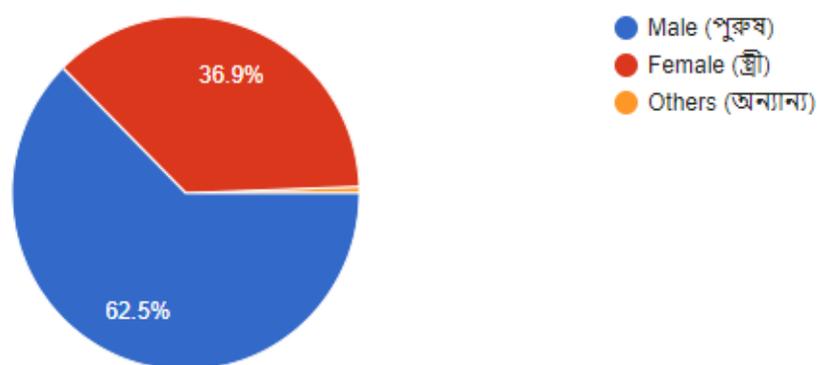


Figure 3.4.1 Gender Ratio of Patients

According to this figure we can see that male are in danger zone more than female. The pace of dengue fever is higher among men, however more ladies are among the individuals who have died from the mosquito-borne viral ailment in Bangladesh. [16]. After our surveying we have found that 62.5% male, 36.9% female and less than 1% other got attacked by dengue this year in our country. This number is too high comparing to the previous years. Many people of them already have died of dengue and female are more in number. It will be more clarify when it will be shown in table format.

Table 3.4.1 Gender Distribution

Gender	Percentage %
Male	62.5%
Female	36.9%
Other	Less than 1%

### Age Distribution

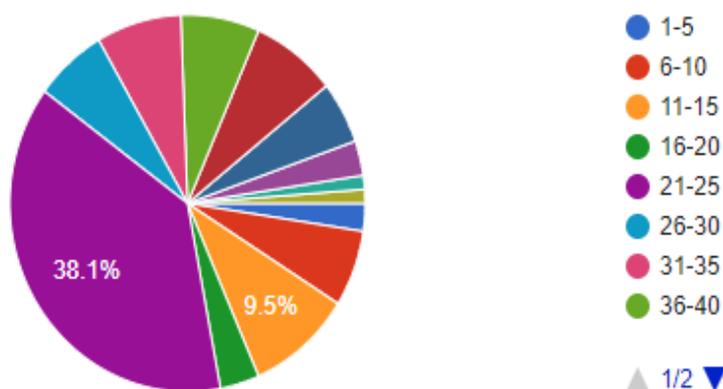


Figure 3.4.2 Age Distribution of Patients

From this figure it can be said that age among 21-25 are in a great danger. 38.1% from age 21- 25 have attacked by dengue in our country according to our survey. Age among 6-10, 11-15 and 36-40 are also in great danger.

### Number of Occurrence

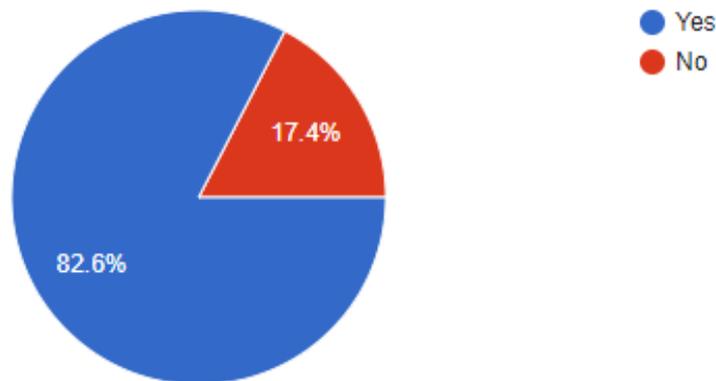


Figure 3.4.3 Number of Occurrence

In our country 82.6% people has attacked by dengue as a first time in their life. And remaining 17.4% people who has already attacked by dengue. There have been many fatalities this year.

### Number of Hospitalized

Here we can see that 36.2% patients didn't have to go to the hospital for getting admitted but On the other hand the remaining 63.8% people went to hospital for getting admitted for taking treatment and stayed for different days. We have talked about our statistical data which have been found by doing survey. Here is a bad news for us that dengue fever is getting too much dangerous day after day. It needs to be prevented, the damage from the dengue.

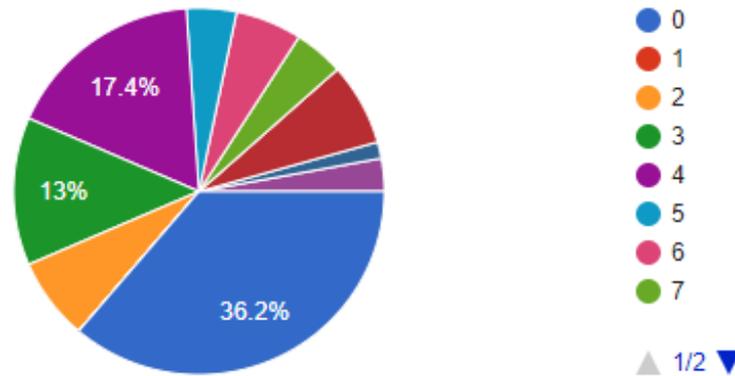


Figure 3.4.4 Number of Hospitalized Patients

### 3.5 Implementation Requirements

For this research dengue fever symptoms are being predicted for reducing the future damage of a patient. Early dengue symptoms prediction can save a patient life and they can go for better and early diagnosis. First of all data have been collected by us by surveying and using internet. Collected data have been prepared and cleaned using data preprocessing techniques. Then data have been generalized finally a output has been generated by algorithm when it has been given to the algorithm for prediction.

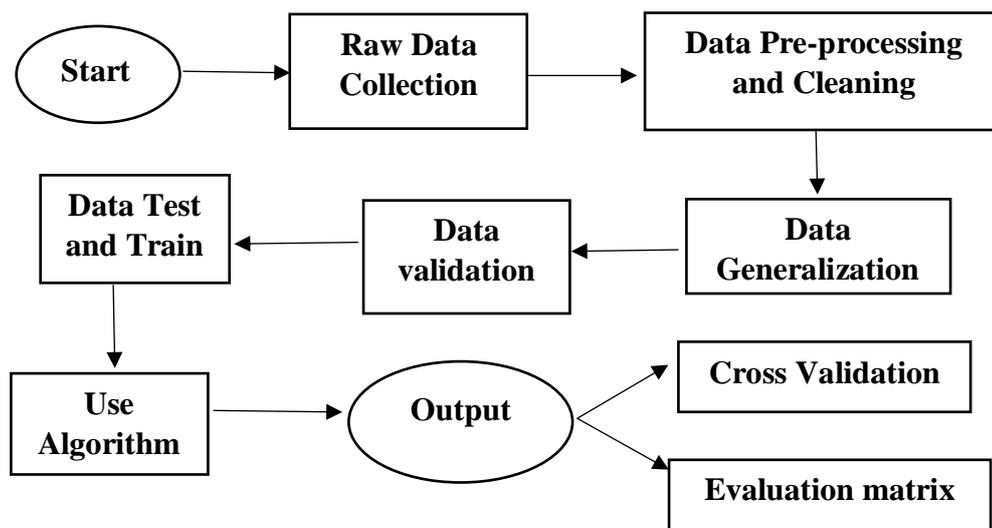


Figure 3.5.1 Work Methodology

Using this methodology our expected output will be predicted. When it comes to algorithm we must say that two algorithm have been used for prediction of dengue fever symptoms. The algorithms are Support Vector Machine(SVM) and Random Forest Algorithm. We are using two algorithm for differentiate the accuracy rate of these two and finally cross validation will be found and confusion matrix also to evaluate the efficiency of the algorithm. For doing these all tasks we need to import some library function in our IDE. Jupyter notebook has been used as our main working platform for this dengue prediction. Which libraries have been used is given below-

```
import pandas as pd
import numpy as ny
from sklearn.model_selection import train_test_split
from sklearn.model_selection import cross_val_score
from sklearn.svm import SVC
from sklearn.ensemble import RandomForestClassifier
import matplotlib.pyplot as plt
from sklearn.metrics import confusion_matrix
import seaborn as sb import seaborn as sb
```

Figure 3.5.2 Used Library

We have used pandas, numpy library in our project. Pandas is used for data frame. Sklearn.model\_selection is used for importing the models or algorithm which we want to use and also used for importing train\_test\_split function for train and test the dataset.

## CHAPTER 4

### EXPERIMENTAL RESULTS AND DISCUSSION

#### 4.1 Introduction

The prescient intensity of the determining model-surveyed by Support Vector Machine (SVM) and Random Forest Algorithm are improved by including the new parameters for dengue fever indications expectation. The restrictive utilization of looking over and gathering of information from numerous areas is the model's primary expectation control. Female mosquitoes and seasons were emphatically corresponded with dengue cases. Hence, the dengue rate patterns gave by this model will help the enhancement of dengue anticipation. The present work shows the significant jobs of female mosquito contamination rates from the past season and individuals will be increasingly cautious with the dengue fever side effects in dengue episodes. Fusing these two factors in the model essentially improves the prescient intensity of dengue hemorrhagic fever predicting models.

Our model is able to predict the symptoms of dengue fever which can be seen in the next years. Early prediction is very important in this factor and using that people can go for better treatment in early stage before losing their lives. And our government can also take proper steps for reducing the damage.

#### 4.2 Experimental Results

Some results have been found after our experiment and that is discussed below.

##### Output of SVM:

```
[1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0]
```

Figure 4.2.1 Result using SVM Algorithm

Using SVM algorithm in our dataset we have predicted that seven symptoms can be seen in the year of 2020. Using the models which we have developed we can predict the dengue symptoms of any next years.

```
SVM :  
['Headache', 'Pain_back_of_the_eye', 'Body_pain', 'Aversion', 'Joint_pain', 'Feeling_dizzy', 'Platelet_depletion']
```

Figure 4.2.2 Predicted Symptoms for 2020 using SVM

### **Output of Random Forest:**

```
[1, 1, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0]
```

Figure 4.2.3 Result using RFC

Using Random Forest Classifier algorithm in our dataset we have predicted that seven symptoms can be seen in the year of 2020 which is similar to the result of Support Vector Machine.

```
RFC :  
['Headache', 'Pain_back_of_the_eye', 'Body_pain', 'Aversion', 'Joint_pain', 'Feeling_dizzy', 'Platelet_depletion']
```

Figure 4.2.4 Predicted Symptoms for 2020 using RFC

We can see here that using both algorithm like Support Vector Machine (SVM) and Random forest classifier (RFC) similar results have been found.

### Accuracy of Algorithms:

Table 4.2.1 Accuracy of Algorithms

Algorithm	Accuracy
Support Vector machine	0.7898307549470339
Random Forest Classifier	0.7898307549470339

Here it is clear that the accuracy that have been found it is similar in both Support Vector

Machine and Random Forest classifier.

### Accuracy of SVM:

SVM Score:  
0.7898307549470339

Figure 4.2.5 Accuracy of SVM

### Accuracy of RFC:

Random Forest Classifier Score:  
0.7898307549470339

Figure 4.2.6 Accuracy of RFC

### Confusion Matrix:

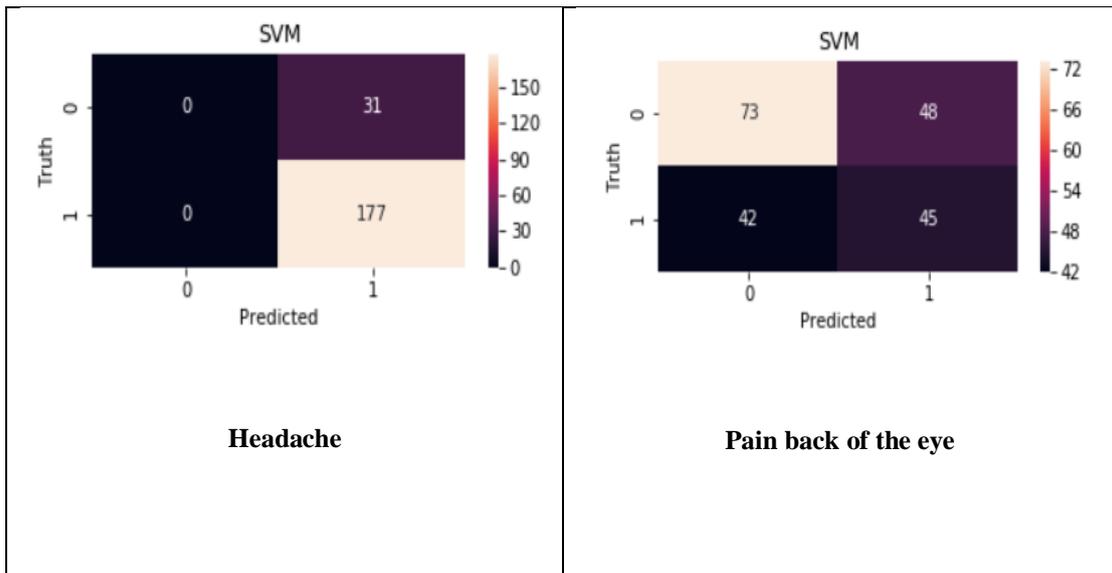
A confusion matrix is a counter that is regularly used to portray the exhibition of an order model on a lot of test information for which the genuine qualities are recognized. It permits the perception of the exhibition of a calculation. Fifteen confusion matrix have been found for our model.

**Formulas-**

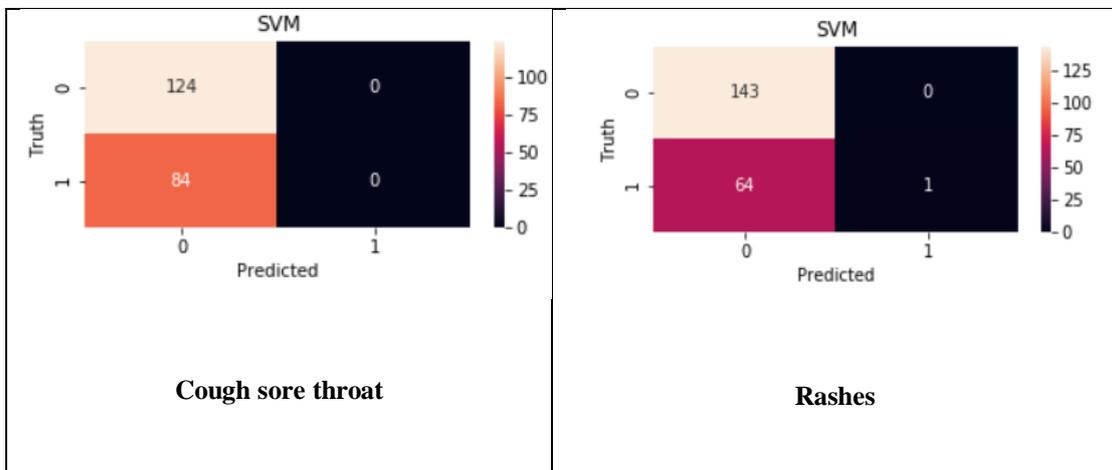
Recall- [ TP / Actual Yes ]

Precision- [ TP / Predicted Yes ]

Accuracy- [ ( TP + TN ) / Total ]

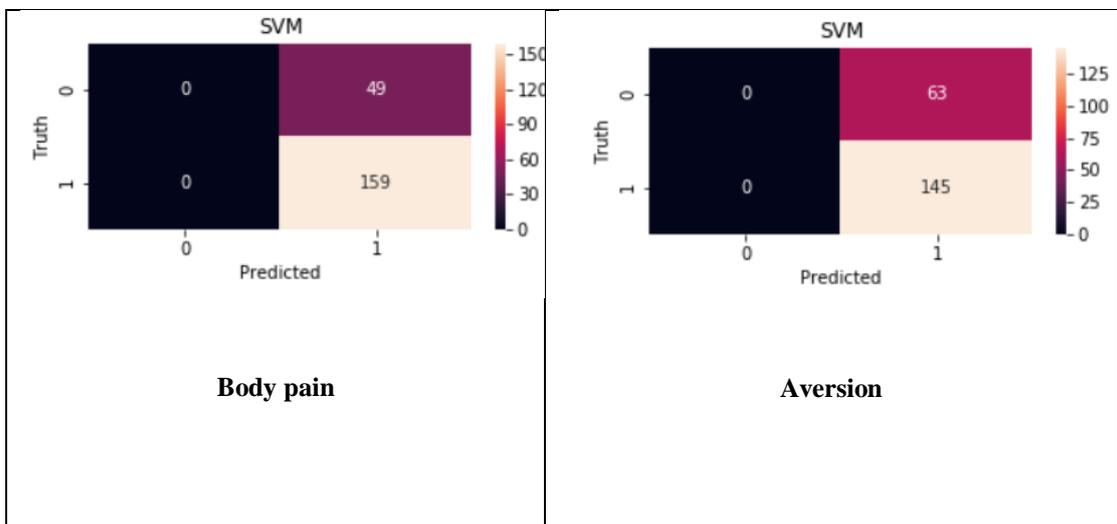


	Recall	Precision	Accuracy
<b>Headache</b>	1.00	0.85	0.85
<b>Pain back of the eye</b>	0.52	0.48	0.57

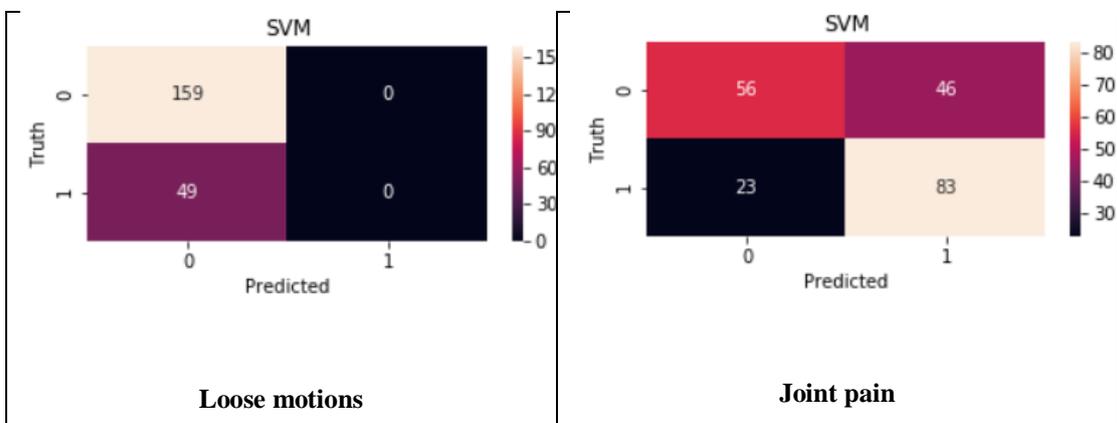


--	--

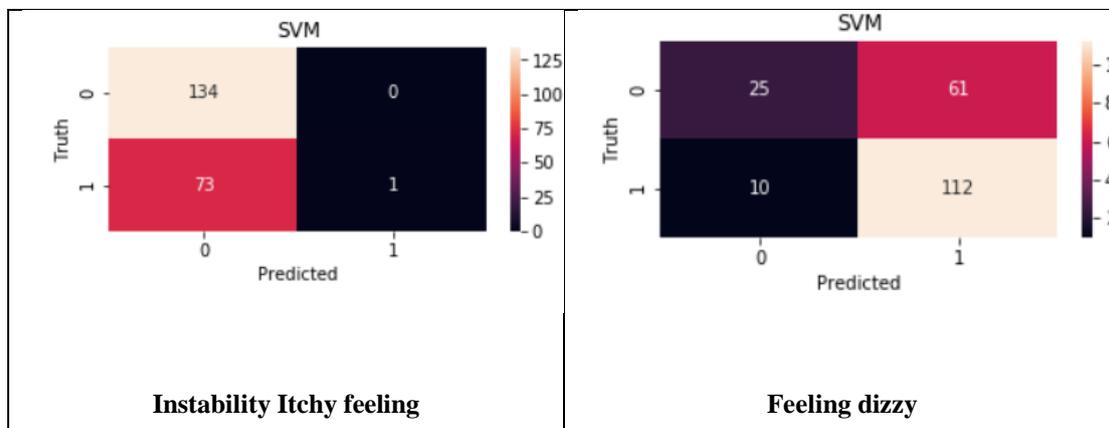
	<b>Recall</b>	<b>Precision</b>	<b>Accuracy</b>
<b>Cough sore throat</b>	1.00	0.00	0.60
<b>Rashes</b>	0.02	1.00	0.69



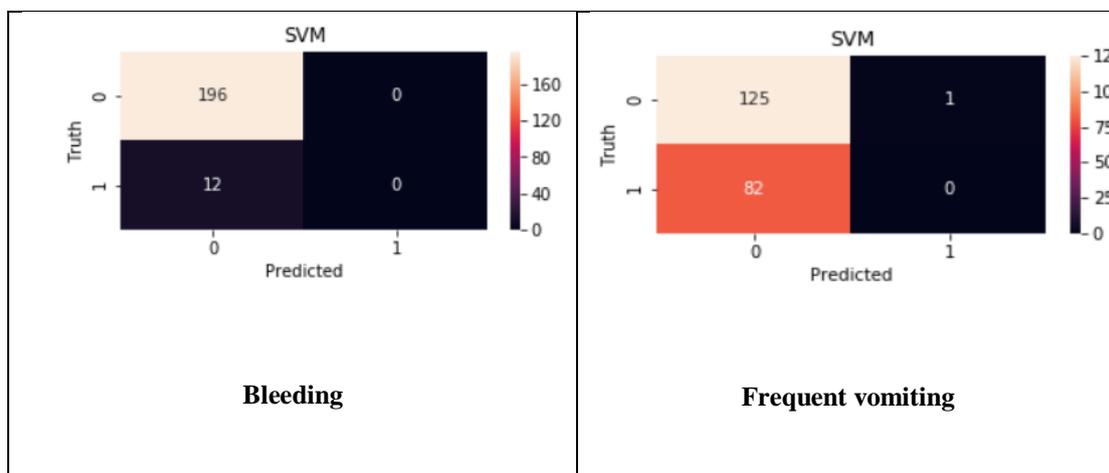
	<b>Recall</b>	<b>Precision</b>	<b>Accuracy</b>
<b>Body pain</b>	1.00	0.76	0.76
<b>Aversion</b>	1.00	0.70	0.70



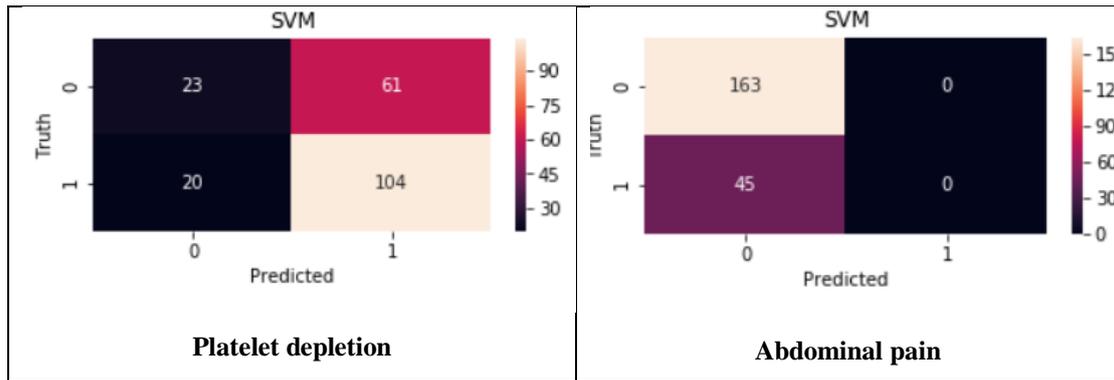
	<b>Recall</b>	<b>Precision</b>	<b>Accuracy</b>
<b>Loose Motions</b>	0.00	0.00	0.76
<b>Joint Pain</b>	0.78	0.64	0.67



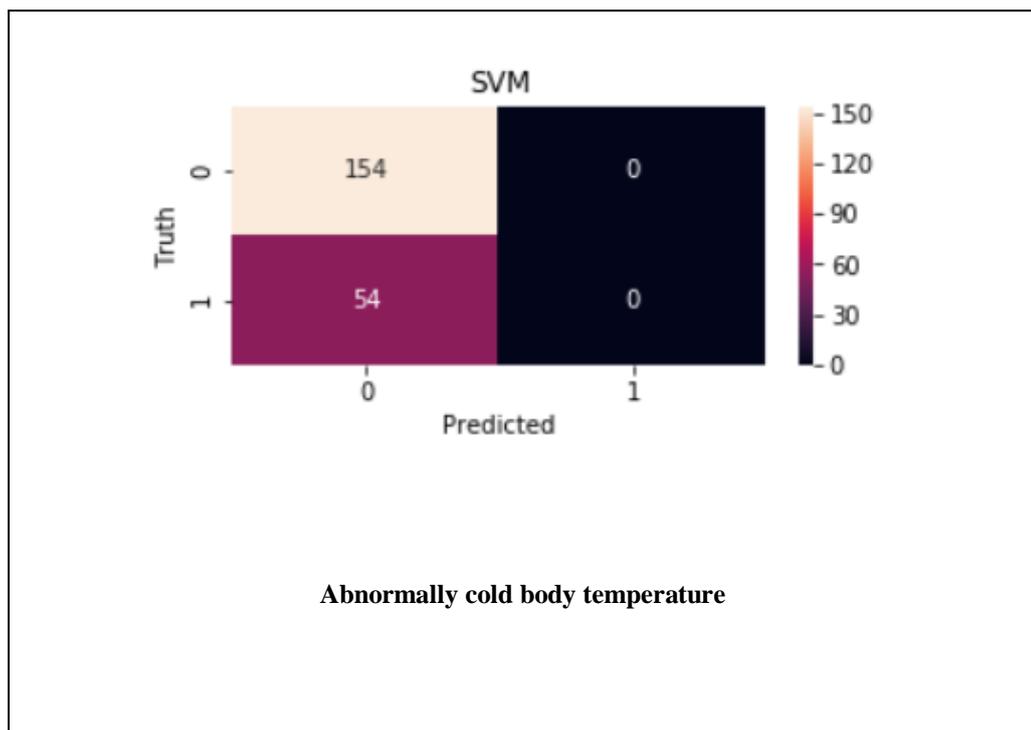
	<b>Recall</b>	<b>Precision</b>	<b>Accuracy</b>
<b>Instability Itchy feeling</b>	0.01	1.00	0.65
<b>Feeling Dizzy</b>	0.92	0.65	0.66



	<b>Recall</b>	<b>Precision</b>	<b>Accuracy</b>
<b>Bleeding</b>	0.00	0.00	0.94
<b>Frequent Vomiting</b>	0.00	0.00	0.60



	<b>Recall</b>	<b>Precision</b>	<b>Accuracy</b>
<b>Platelet Depletion</b>	0.84	0.63	0.61
<b>Abdominal Pain</b>	0.00	0.00	0.78



	<b>Recall</b>	<b>Precision</b>	<b>Accuracy</b>
<b>Abnormally Cold Body Temperature</b>	0.00	0.00	0.74

Figure 4.2.7 Confusion Matrix using SVM

### 4.3 Descriptive Analysis

Which outbreak will be more prevalent in 2020 among patients have been found in our experiment over our dataset. This prediction is for the year of 2020.

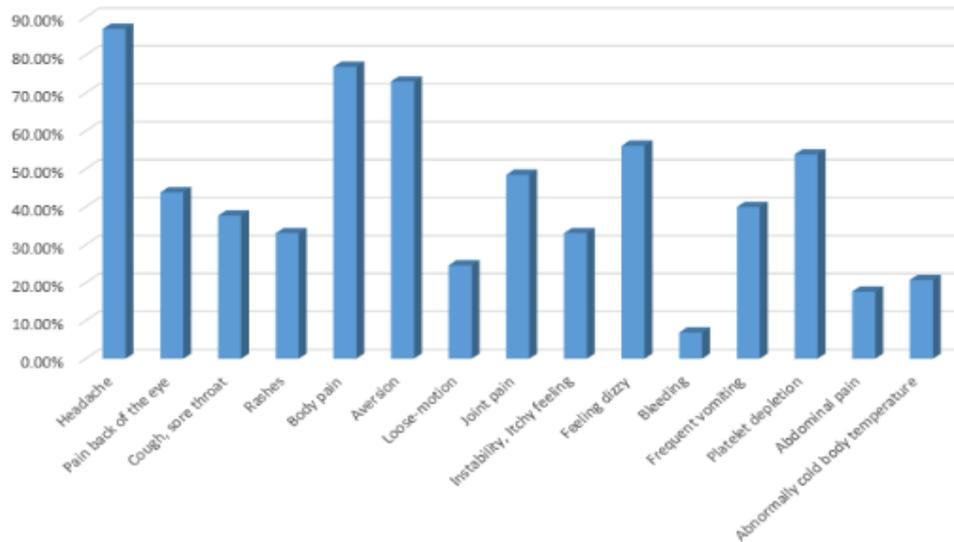


Figure 4.2.8 Prevalent Symptoms in bar chart for 2020

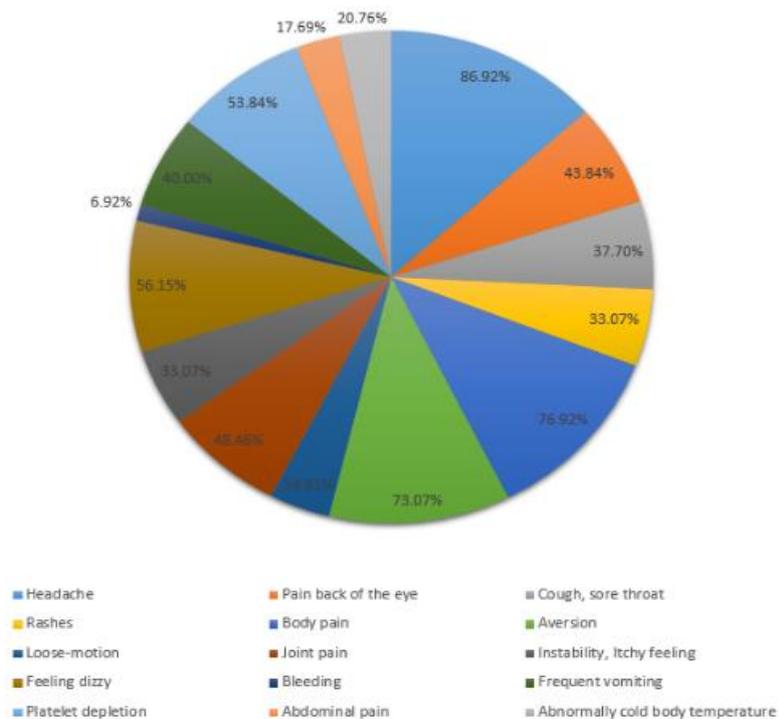


Figure 4.2.9 Prevalent Symptoms in Pie Chart for 2020

According to the figure headache will be the most prevalent symptoms among patients in our predicted year. The percentage of headache will be 86.92% which is a huge number of percentage. After that body pain will be the second in prevalent symptoms and it will be 76.92. After getting the results of our predicted values it can be said that Headache, Body pain, Aversion, Feeling dizzy, Joint pain, platelet depletion, Pain back of the eye will be the most seen symptoms in the year of 2020.

#### **4.4 Summary**

In this experiment and research some dengue fever symptoms have been found which can be seen in the next year using our Support Vector machine (SVM) and Random Forest Classifier (RFC). In our research over our dataset the Support Vector Machine and Random Forest Classifier have given the same output of dengue symptoms as well as the similar accuracy and which is 0.7898307549470339. The symptoms which have been predicted by these two algorithm for next years are Headache, Body pain, Aversion, Feeling dizzy, Joint pain, platelet depletion, Pain back of the eye. Confusion matrix has been found also and we have describe it on the part of experimental results. By predicting future dengue symptoms it can be possible to reduce the damage of our manpower. And we have to ensure the better treatment for the patients.

## CHAPTER 5

### SUMMARY, CONCLUSIONS, RECOMMENDATIONS AND IMPICATION FOR FUTURE RESEARCH

#### 5.1 Summary of the Study

Dengue is one of the most common known viral diseases in recent days. It has taken a very serious effect in this year in our country, Bangladesh. Many people have died of dengue this year. So, we need to make people more careful and aware about dengue fever symptoms so that they can go for a quick and better diagnosis and care. How can we make people careful about dengue fever this question arose when we first thought about this research. Finally, we have come to a conclusion that we could work with dengue fever symptoms and that would be so much beneficiary for people. So a model for predicting dengue fever symptoms using some different machine learning algorithm has been developed. Support Vector Machine (SVM) model, Random Forest Classifier have been used for prediction the dengue symptoms over our collected data and our dataset. Data collection was the main challenge for our research work.

#### 5.2 Conclusions

The present work shows the significant jobs of female mosquito disease rates from the past season in dengue episodes. Joining this factor in the model essentially improves the predict intensity of dengue hemorrhagic fever determining models. Dengue disease has risen as a significant wellbeing worry in Southeast Asia particularly Bangladesh and till a particular antibody or an antiviral operator become accessible, we need to depend on early distinguishing proof of hazard variables related with creating intricacies of DF; simply like the admonition includes just as the ones recognized in this research. A model for identification of dengue symptoms has been developed so that they can take quick and proper treatment for perfect and early overcome.

### **5.3 Recommendations**

Machine Learning (ML) is the logical investigation of calculations and factual models that PC frameworks use to play out a particular assignment without utilizing unequivocal directions, depending on examples and surmising. It is viewed as a subset of artificial intelligence. Machine Learning calculations assemble a numerical model dependent on test information, known as training data, so as to settle on expectations or choices without being expressly modified to play out the work. Two different machine learning algorithms have been used to develop our model for predicting dengue fever symptoms and for showing which is better in prediction over our collected dataset. The three algorithms are- Support vector machine which is also known as SVM, Random forest algorithm. While doing our research work over our dataset similar accuracy rate has been found from these different algorithms. Though it can vary from dataset to dataset of the accuracy of algorithms but for our dataset the accuracy rate is similar for the used algorithms. When we increase or decrease our test value the accuracy of Support Vector Machine gets higher than Random forest. So we have a recommendation for you to use or try Support Vector Machine in predicting research work.

### **5.4 Implication for Further Study**

In this research work a model for predicting dengue fever symptoms has been developed using different machine learning algorithms. But in future we have a different plan with it. Like, in future we will predict the pattern of the dengue symptoms using ML algorithms. I think it will be also more beneficiary for people to find the actual reason behind the dengue fever and they will take the proper diagnosis and treatment.

## References

- [1] "Dengue," South-East Asia Regional Office, [Online]. Available: [http://www.searo.who.int/entity/vector\\_borne\\_tropical\\_diseases/data/data\\_factsheet/en/](http://www.searo.who.int/entity/vector_borne_tropical_diseases/data/data_factsheet/en/) [Accessed 2 Nov. 2019].
- [2] "DGHS: 269 more hospitalized for dengue in 24hrs," Dhaka Tribune, 2019. [Online]. Available: <https://www.dhakatribune.com/bangladesh/2019/10/05/dghs-269-more-hospitalized-for-dengue-in-24hrs>. [Accessed 2 Nov. 2019].
- [3] F. Mahmud, "Why is Bangladesh grappling with a record dengue fever outbreak?," Aljazeera.com, 7 8 2019. [Online]. Available: <https://www.aljazeera.com/news/2019/08/bangladesh-grappling-record-dengue-fever-outbreak-190807091210915.html>. [Accessed 2 Nov. 2019].
- [4] "Dengue and severe dengue," Who.int, 15 4 2019. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/dengue-and-severe-dengue>. [Accessed 2 Nov. 2019].
- [5] "Bangladesh: Dengue | IAMAT," Iamat.org, 24 10 2019. [Online]. Available: <https://www.iamat.org/country/bangladesh/risk/dengue>. [Accessed 2 Nov. 2019].
- [6] N. I. M. Iqbal, "Machine learning for dengue outbreak prediction: An outlook," *International Journal of Advanced Research in Computer Science*, vol. 8, pp. 93-102, 2017. [Accessed 2 Nov. 2019].
- [7] S. R. S. Hadinegoro, "The revised WHO dengue case classification: does the system need to be modified?," *Paediatrics and International Child Health*, vol. 32, no. sup1, pp. 33-38, 2012. [Accessed 2 Nov. 2019].
- [8] S. Report, "Dengue outbreak: 1,649 hospitalised in 24 hours," The Daily Star, [Online]. Available: <https://www.thedailystar.net/country/dengue-fever-outbreak-1649-hospitalised-in-24-hours-1781182>. [Accessed 2 Nov. 2019].
- [9] S. S. S. I. H. P. Raghvendra Jain, "Prediction of dengue outbreaks based on disease surveillance, meteorological and socio-economic data," *BMC Infectious Diseases*, vol. 19, no. 1, 2019. [Accessed 2 Nov. 2019].

- [10] T. L. Q. Z. L. W. J. X. Q. Z. G. L. Z. L. J. H. Y. Z. W. M. Pi Guo, "Developing a dengue forecast model using machine learning: A case study in China," *PLOS Neglected Tropical Diseases*, vol. 11, no. 10, p. e0005973, 2017. [Accessed 2 Nov. 2019].
- [11] S. T. A. I. Z. A. N. F. Rahim, "Dengue fatality prediction using data mining," *Journal of Fundamental and Applied Sciences*, vol. 9, no. 6S, pp. 671-683, 2018. [Accessed 2 Nov. 2019].
- [12] D. S. R. Sanjudevi, "Dengue Fever Prediction using Classification Techniques," *International Research Journal of Engineering and Technology (IRJET)*, vol. 06, no. 02, pp. 558-563, 2019. [Accessed 2 Nov. 2019].
- [13] B. M. S. L. S. S. P. Muhilthini, "Dengue Possibility Forecasting Model using Machine Learning Algorithms," *International Research Journal of Engineering and Technology*, vol. 05, no. 03, pp. 1661-1665, 2018. [Accessed 2 Nov. 2019].
- [15] H.-R. G. S.-C. L. C.-Y. L. H.-J. S. Pei-Chih Wu, "Weather as an effective predictor for occurrence of dengue fever in Taiwan," *Acta Tropica*, vol. 103, no. 1, pp. 50-57, 2007. [Accessed 2 Nov. 2019].
- [16] "Women dying more from dengue but disease afflicting more men in Bangladesh," *Bdnews24.com*, 20 8 2019. [Online]. Available: <https://bdnews24.com/health/2019/08/20/women-dying-more-from-dengue-but-disease-afflicting-more-men-in-bangladesh>. [Accessed 2 Nov. 2019].

## **APPENDICES**

### **APPENDIX**

We have put a lot of effort doing this research. While developing the prediction model of dengue fever symptoms, we have learnt lots of things of which we did not have a proper knowledge before and we are still learning. Now we know how to use technology for saving human lives. We have used different algorithms of machine learning in this research for predicting dengue fever that will help us to take early treatment before having any damages from dengue. We are hopeful that the knowledge and experience we have gained while doing this research work will help us improve a lot in our future work resulting in better output.

## A comparative study of Dengue

### ORIGINALITY REPORT

<b>19%</b>	<b>8%</b>	<b>6%</b>	<b>18%</b>
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

### PRIMARY SOURCES

<b>1</b>	<b>Submitted to Pittsburg State University</b> Student Paper	<b>2%</b>
<b>2</b>	<b>Submitted to Westcliff University</b> Student Paper	<b>1%</b>
<b>3</b>	<b>Submitted to University of Surrey</b> Student Paper	<b>1%</b>
<b>4</b>	<b>Submitted to Central Queensland University</b> Student Paper	<b>1%</b>
<b>5</b>	<b>Submitted to University of Westminster</b> Student Paper	<b>1%</b>
<b>6</b>	<b>Submitted to The Norwood Morialta High School</b> Student Paper	<b>1%</b>
<b>7</b>	<b>Submitted to Fiji National University</b> Student Paper	<b>1%</b>
<b>8</b>	<b>Submitted to California State University, Fresno</b> Student Paper	<b>1%</b>
<b>9</b>	<b>Submitted to Far Eastern University – Nicanor Reyes Medical Foundation</b>	<b>1%</b>