DETECTION OF MAJOR SYMPTOMS OF COVID 19 DELTA VARIANT USING MACHINE LEARNING TECHNIQUE

BY

UJJAL RAZ BONGSHI
ID:203-25-003

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

Supervised By

Dr. S. M. Aminul Haque
Associate Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

JANUARY 2021

APPROVAL

This Project/Thesis titled **"Detection of Major Symptoms of COVID 19 Delta Variant Using Machine Learning Technique**" , submitted by Ujjal Raz Bongshi, ID No: 203-25-003 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 M.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 19-01-2022.

BOARD OF EXAMINERS

Chairman

Dr. Touhid Bhuiyan
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Dr. Fizar Ahmed
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

Naznin Sultana
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

External Examiner

Dr. Mohammad Shorif Uddin
Professor
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

We hereby declare that this project has been done by us under the supervision of **Dr. S. M. Aminul Haque, Associate Professor, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

Dr. S. M. Aminul Haque
Associate Professor
Department of CSE
Daffodil International University

Submitted by:

(Ujjal Raz Bongshi)
ID: 203-25-003
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

Foremost, I am grateful to Allah for the good health and welling that were necessary to complete this thesis.

We would like to express our sincere gratitude to our supervisor **Dr. S. M. Aminul Haque, Associate** Professor, Department of Computer Science and Engineering, Daffodil International University, Dhaka, for the continuous support of our project study, for his patience, motivation, enthusiasm and immense knowledge. His guidance helped me in all the time of working and writing of this report. I could not imagine having a better supervisor and mentor for my thesis.

Beside my supervisor, I would like to thank project committee, for their encouragement, insightful comments and hard questions.

I would like to thank other faculty members and the staffs of Computer Science and Engineering Department of Daffodil International University, for their kind help to finish my thesis.

And specially I would like to thank those patients who gave their valuable information as my thesis data.

Last but not least, I would like to thank my family specially my parents, for giving birth and supporting me spiritually throughout my life.

ABSTRACT

This research is intended as a guide for researchers who want to take part in future research work and current people who want to know the major symptoms of the Covid 19 Delta variant. Machine learning in the health sector helps to detect, study and diagnose different diseases. The primary purpose of this work is to find out the major symptoms of the COVID 19 Delta variant. This research has been conducted, and data has been taken from 1169 Covid patients who were affected with the different variant of the Covid 19. After gathering the data, different classifier algorithms were applied. RandomTree algorithm provides the best result, which tells the major symptoms of Delta variant. By knowing this variant's symptoms, one will be able to discriminate one person if he is affected by Covid 19 Delta variant or not. The analysis is done with an intention so that machine can save time to discriminate between persons of different vaiant and Covid positive people. It can help the researcher find the vaccine and make any required medicine that will save the day. There are other possibilities and research sectors that can help to fight and find a cure for this virus.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLE

CHAPTER 1

INTRODUCTION

1.1 Introduction

Life is very much valuable than any other thing. In recent times COVID 19 has been pandemic and has taken a lot of innocent lives. COVID-19 or Coronavirus disease 2019 is one of the contagious diseases that is caused by severe acute respiratory syndrome coronavirus 2 (Or SARS-CoV-2). It has spread worldwide and changed its form several times [1].

It started with the variant Alpha with several symptoms, including fever, cough, fatigue, breathing problems, and etc. Then it changed its form, and symptoms are also changed. No one can say that if one has fever and cough, therefore he is affected with Coronavirus. Or specifically with any variant of Coronavirus [2].

Different variants of COVID 19 are casing risks at different levels. In a new environment, the virus is also mutating and evolving. In Bangladesh, this virus spread in March 2020, and the first 3 cases were reported on month of March 8 in IEDCR [3]. Since that day, the virus has spread throughout the nation rapidly. Bangladesh is the 2nd most affected country by this virus in South Asia.

The first variant of COVID was Alpha; it has taken away a lot of life, and its next version, Beta, was also very similar to Alpha [4]. There was neither any treatment nor any medication, and only one way was available ‘Wait’. Suffer & wait for death and stay safe & wait for the vaccine.

Delta variant has already made up more an 90% of all infections[5]. Public Health England said that the data shows a great risk for patients who are admitted in contrast to alpha and beta variants. More study is required to know the patterns of this virus [6]. The latest wave of Delta variant of SARS COV 2 mutates to new variants that are emerging as ongoing transmission persists [7].

The B.1.1.7 (Alpha), P.1 (Gamma), B.1.351(Beta), and B.1.617.2 (Delta) variants are a concern, and more mutations make the virus less responsive to treatments [8].

After the waves of COVID 19, people have an extremely cautious vibe. A small cough is thought of as an infection or COVID 19 infection [9]. But in reality, it might be a general cough.

This project works with the data of COVID 19 delta-affected patients and finds the pattern of symptoms for delta variants. This pattern will help one to significantly tell if he is affected by the delta variant or not.

This research will also help one to find the proper treatment for a specific type of variant. Besides, there is a very low amount of research regarding COVID 19 in Bangladesh. Different variants respond differently in different countries or in different temperatures. It will help one to understand what effects delta variant bring on the victim's body. Then one will be able to use this data in his future works.

1.2 Motivation

Corona Virus has taken a lot of lives, and its effects are not over yet. Many people have lost their relatives and close ones. Over two years, this virus didn’t disappear, in fact, people have found the only vaccine to slow it down. No cure is still invented. Due to mutation, the virus is taking different shapes. To find any proper cure, we must understand the virus first, what its symptoms are and how to fight it [10]. In this research, I have worked with the symptoms of Delta variants.

Moreover, there is very little research that took place regarding COVID 19 and Delta variant. If no one moves forward to take down the virus, then who will? Besides, saving a life is very much important. From 2019 to 2022, a lot of data has been recorded. If we utilize this data it will be easier to find a breakthrough for this virus.

Machine learning can help find the information that will help us fight this syndrome and discover more about this virus.

Every reason that is talked about above is the motivation behind this research.

1.3 Research Objectives

Every research has some specific objectives. Without the objective, the candidate might get lost from his track of research. While starting the research, the candidate must have a clear idea of the objective of the research. On measuring the variable, such as identifying or describing them, is generally focused on the research objective.

The research objectives are given below:

- To find the major symptoms of the COVID 19 Delta variant.
- When to know if one is affected by COVID 19 Delta variant.

1.4 Research Questions

When I started this project, there were several questions that arose in mind, which were project-related. Among those questions, there were some major sectors, with which are included in the research. These sectors are “occupation”, “hour spent outside home”, “Fever”, “Shortness of breath”, “Sore throat”, “Anosmia”, “Loss of smell”, “Fatigue” and “other activities”.

1.5 Expected Outcome

In this project expected outcome is to find specific symptoms of COVID 19 Dela variant patients. We will also know which group of people are mostly affected by this variant and what are the other criteria.

1.6 Report Layout

This research paper consists of five chapters. Introduction, Backgroun, Research Methodology, Experimental Results and Discussion, and Conclusion and Future Work.

Chapter 1: Introduction; Introduction, Motivation, Research Objectives, Research Questions, Expected Outcome, Report Layout.

Chapter 2: Background; Introduction, Related Works, Scope of the problem, Challenges.

Chapter 3: Research Methodology; Introduction, Research Subject and Instrumentation, Data Collection Procedure, Statistical Analysis.

Chapter 4: Experimental Results and Discussion; Introduction, Experimental Results, Results of Other Classifiers, Descriptive Analysis.

Chapter 5: Conclusion and Future Work; Conclusion, Future Work.

CHAPTER 2

BACKGROUND

2.1 Introduction

Health is considered as wealth. Though you have everything but if you are not well, you won't be able to enjoy everything you have therefore good health is a must. In 2019 COVID 19 started creating havoc in Wuhan in China. From then, it spread to the whole world and was announced a pandemic [11].

Whenever COVID 19 spreads through an area, it mutates and changes its form according to the weather and environment of the area. Therefore, the symptoms and conditions of patients are different. These reasons make it very difficult to find any cure for this virus.

In cold regions, people suffer from different symptoms, and in the warm region, people suffer from different symptoms [12]. The date rate also varies in different regions. The first variation was Alpha that took many lives until people became a little bit cautious. At that time, there was nothing to fight back and stop the virus from spreading. The only way was to stay at home and stay safe [13].

Later on, by researchers found the way to detect COVID 19. There were several ways. But detecting is not any cure. Every country was locked down due to pandemic [14]. Then the next version of Corona was found and named "Beta". Its symptoms were similar to the Alpha version, but its death rate was higher than the Alpha version [15].

With time some more mutations happened, and the virus changed its form. These variations include Alpha, Beta, Gamma, Delta, Lambda, Mu, and even new variation is introduced after a period.

Every variation has its own specific criteria and specific symptoms, though they have common symptoms. All these variations should be analyzed along with their symptoms, without research, it is impossible to find any real cure for this virus. Till now, after the Beta variation, Delta has spread very much, and its death rate is very high [16].

In Bangladesh, it has taken many lives and spread almost everywhere. It has some variations in its symptoms, including taste loss, smell loss, conjunctivitis, malaise, etc. It is seen that symptoms are different in terms of different persons [17].

This research works with the symptoms of the Delta variant and finds the major symptoms of the COVID 19 Delta variant.

2.2 Related Works

In the recent past, there were not enough related works are done. But there were a few related works are done, and some of those works helped a lot.

Related works:

1. "COVID-19 Detection using Artificial Intelligence
   - by Samy S. Abu-Naser, Fatima M. Salman, Eman Alajrami, Belal A. M. Ashqar, Bassem S. Abu-Nasser, Department of Information Technology, Faculty of Engineering and Information Technology, Al-Azhar University, Gaza, Palestine.
   This article focuses on the ways of detection of Covid 19 using Artificial intelligence. It is aimed to construct a model by adopting deep learning so that Pneumonia of Covid 19 can be detected using high-resolution X rays and relieve the working pressures of radiologists. At here, artificial intelligence is used on X-ray images of non Covid and covid patients. Then findings are published as the result of the work.
2. "COVID-19 Prediction and Detection Using Deep Learning"
   - by Moutaz Alazab, Abdelwadood Mesleh, Albara Awajan, Salah Alhyari, Ajith Abraham, Vansh Jatana, Faculty of Artificial Intelligence, Al-Balqa Applied University, Al-salt, Jordan, Amman, Jordan, Machine Intelligence Research (MIR) Labs, Auburn, WA, US, School of IT, SRM University, Kattankulathur, India.
   This paper works for detecting the Covid 19 using the deep learning. In this research, augmented chest x-ray images of healthy people and Covid patients were used. From 1000 X-ray images, AI has been trained, and then this AI is used for identifying the Covid and healthy x-ray images.

3. “Detection of COVID-19 Infection from Routine Blood Exams with Machine Learning: A Feasibility Study”- by Davide Brinati, Andrea Campaigner, Davide Ferrari, Massimo Locatelli, Giuseppe Banfi, Federico Cabitza.

   In this research, two machine learning models have been developed to discriminate between healthy and Covid patients. The accuracy of the model ranges between 82%-85% and sensitivity between 92%-95%. It takes the blood exam results and applies machine learning techniques to find the result, if it is Covid or not.

4. “Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study”- by Prof Nanshan Chen MD, Prof Min Zhou MD, Xuan Dong PhD, Prof Jieming Qu MD, Fengyun Gong MD, Yang Han PhD, Prof Yang Qiu PhD, Jingli Wang MD, Ying Liu MD, Yuan Wei MD, Jia'anXia MD, Ting Yu MD, Prof Xinxin Zhang MD, Prof Li Zhang MD

   This research works with 99 cases and investigates with the characteristics they got from this cases. The sample was taken from the Wuhan, and all the patient were suffering from Covid and pneumonia both. It reveals many characteristics of Covid virus for that time in Wuhan. This research also shows blood element values and other symptoms that a Covid patient feels.

5. “Detection and analysis of nucleic acid in various biological samples of COVID-19 patients”- by Jianguo Wu, Jiasheng Liu, Shijun Li, Zhiyang Peng, Zhe Xiao, Xufeng Wang, Ruicheng Yan, Jianfei Luo.

   This research paper shows that real-time fluorescence quantitative reverse transcriptase-polymerase chain reaction and viral gene sequencing are the gold standard for diagnosing the Covid 19. The positive rate of nucleic acid detection in covid patients is low; however, there are few reports on the clinical application of nCov 2019 nucleic acid detection.

6. “Real-time RT-PCR in COVID-19 detection: issues affecting the results”- by Alireza Tahamatan, Abdollah Ardebili.

   This research is done to find if Covid 19 can be detected using real-time RT-PCR. At here detection process and the its performance has been described. This method has successfully been able to minimize the chance of false-positive results.

7. “Knowledge, attitudes, and fear of COVID-19 during the Rapid Rise Period in Bangladesh”- by Mohammad Anwar Hossain, Md. Iqbal Kabir Jahid, K. M Amran Hossain , Lori Maria Walton, Zakir Uddin, Md. Obaidul Haque, Md. Feroz Kabir, S. M. Yasir Arafat, Mohamed Sakel, Rafey Faruqui, Zahid Hossain.

   This paper shows the result regarding the knowledge, attitudes, and fear that are nurtured among the people due to Covid 19. It shows what people are doing wrong and why the infection is increasing in a rapid way. It also talks about putting a stop to this virus.

2.3 Scope of the Problem

In every research, we all face some problems. It’s very much difficult to cope with those problems. The working procedure of us might not match with the future or with the past. This is one of the biggest problems. In our project, we faced several problems.

- Collecting the thoughts of specialists on the title of this thesis work was essential and challenging work.
- It was tough to set the attributes for the survey question. It took a lot of time to select the correct attributes.
- After setting the survey question, it needed to be verified. We needed to know if the questions were correct to find the best result.
- Data collection was very tough in this pandemic situation. Everybody had a fear of getting affected by the virus.

In this thesis work, all the problem or challenges was solved with the help of the supervisor, co-supervisor, doctors, and research papers. Every problem was like a challenge. The solution to those problems made this work more perfect.

2.4 Challenges

Every important works and hard works always there have some challenges. Our projects also were some challenges. Those challenges make our works interesting and enjoyable. From those challenges, we can be learned many new things.

First of all, it was challenging to set the questionnaire set. But it was managed and then verified by Dr. Musa Ibrahim Taslim, and Dr. Arif Aziz. Without them it was unthinkable to set the questionnaire.

The biggest challenge was collecting the data. The data is very much rare, but with the help of the supervisor, co-supervisor and some of the doctors, data was successfully managed. At here, data was collected from Ad-din hospital, Dhaka National Hospital, and Mitford hospital. Data was collected by doing survey.

Then I had to convert those data into digital form, and it was done with google forms. Collecting data and then converting those was quite challenging.

At last, I want to say that I enjoyed those challenges very much.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

The nature of the research question should formulate the type of research design. The process is used to collect data and information to make decisions. The research methodology might include surveys, publication research, interviews, and other kinds of techniques for research, including both of the historical and present information.

This research was concerned with finding out the major symptoms of Coronavirus Delta variation. So that one can easily discriminate if he is affected by the Delta variation of Coronavirus or not. It will help researchers in the future for finding the cure and analyze new symptoms for new variations. It will also help to study about the Corona virus Delta variation.

The research was designed for multiple approach. Attributes and data were gathered from different sources by using different data collection tools. A physical interview was conducted to collect the attributes from specialists. A questionnaire in Google form was used to collect data from patients, doctors, and people who are cured. The research work was conducted in the Ad-din hospital, Dhaka National Hospital, and Mitford hospital in Dhaka, Bangladesh. The number of collected data is 1169.

3.2 Research Subject and Instrumentation

In this research, the main subject was to find out the major symptoms of the Covid 19 Delta variant. At here there are some sub-subjects, and they are:

Choosing Attributes Table: In making question, several research papers, different websites, and different person's suggestions are used.

Collecting pure data: Raw data is collected from the hospital patients and the record of the hospital. Some data is also collected using Google form.

Data Pre-processing: At here, Weka's built-in pre-processing technique as a pre-processing instrument for data pre-processing is used.

Data Mining: I used Weka's default data mining algorithms to mine from the data.

3.3 Data Collection Procedure

First sample attributes were collected from doctors and patients. Based on those sample attributes, the main attributes were prepared. The final attribute table or Questionnaire to collect data is given below.

Table 3.3.1: Questionnaire to collect data

No.	Attribute	Attribute Value
1.	What is your age?	5-18, 19-35, 36-55, 56-100
2.	What is your gender?	Male, Female, Others
3.	What is your occupation?	Public service, Healthcare, Business or work in crowd, study, other
4.	How many hours do you spend outside your home?	1,2,3,4,5,6,7,8,9,10 hour
5.	Do you have or had a confirmed case of Covid 19?	Yes, No
6.	Fever > 100F in the last 14 days of infection? If yes, then tell the duration.	No, Yes 1-7 days, Yes 7-14 days
7.	Shortness of breaths in the last 14 days of infection? If yes, then tell the duration.	No, Yes 1-7 days, Yes 7-14 days
8.	Dry cough in the last 14 days of infection? If yes, then tell the duration.	No, Yes 1-7 days, Yes 7-14 days
9.	Sore Throat in the last 14 days of infection? If yes, then tell the duration.	No, Yes 1-7 days, Yes 7-14 days

10	Cough with sputum in the last 14 days of infection? If yes, then tell the duration.	No, Yes 1-7 days, Yes 7-14 days No, Yes 1-7 days, Yes 7-14 days
11	Pneumonia in the last 14 days of infection? If yes, then tell the duration.	No, Yes 1-7 days, Yes 7-14 days
12	Diarrhea in the last 14 days of infection? If yes, then tell the duration.	No, Yes 1-7 days, Yes 7-14 days
13	Do you have Anosmia or taste loss? If yes, then tell the duration.	No, Yes 1-7 days, Yes 7-14 days
14	Do you have Fatigue in last 14 days of infection? If yes, then tell the duration.	No, Yes 1-7 days, Yes 7-14 days
15	Do you have Loss of Smell in last 14 days of infection? If yes, then tell the duration.	No, Yes 1-7 days, Yes 7-14 days
16	Do you have Conjunctivitis in last 14 days of infection? If yes, then tell the duration.	No, Yes 1-7 days, Yes 7-14 days
17	Do you have Malaise in last 14 days of infection? If yes, then tell the duration.	No, Yes 1-7 days, Yes 7-14 days
18	Do you have Lung disease or COPD? If yes, then tell the duration.	No, Yes for more than 1 years, Yes for more than 5 years, Yes for more than 10 years, Yes for more than 20 years
19	Do you smoke? If yes, then tell the duration.	No, Yes for more than 1 years, Yes for more than 5 years, Yes for more than 10 years, Yes for more than 20 years
20	Which variant of Covid Case it was?	Alpha, Beta, Gamma, Delta, Lambda, Mu, I do not know

Based on these questionnaires raw data was collected from patients and from the records of hospital. Some of the data was collected using Google forms

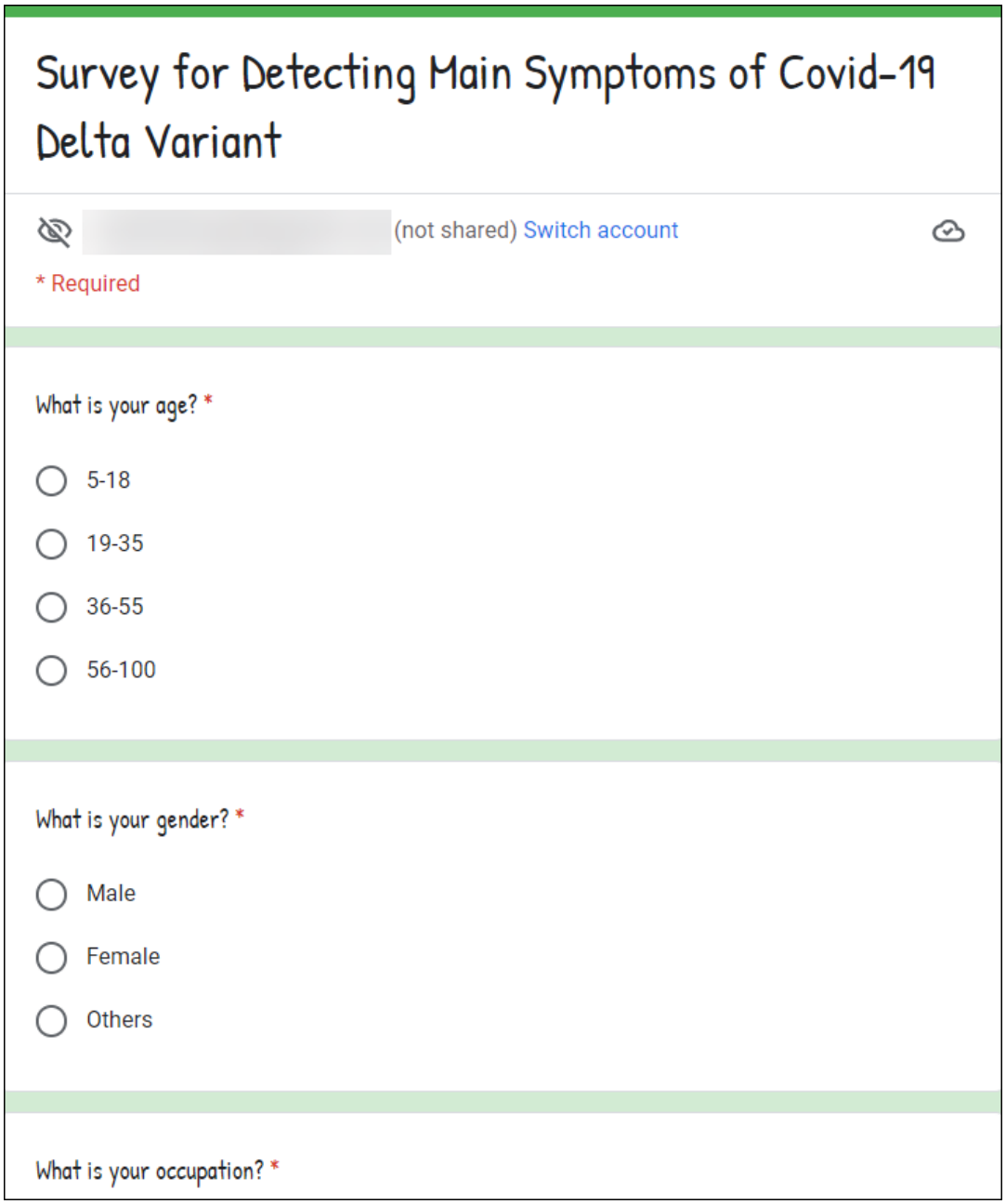


Figure 3.3.1: Google form for collecting data

3.4 Statistical Analysis

From patients, doctors, and other respondents, responses were recorded.

Responses of the Questionnaires

1. **What is your age?**

   Among all the respondents, it was like 63.7% were of age 19-35, 28.3% were of age 36-55, 7.3% were of age 56-100, and 0.7% were of age 5-18%. Therefore it can be said that most of the victims of this variant are mostly from the age of 19-55.

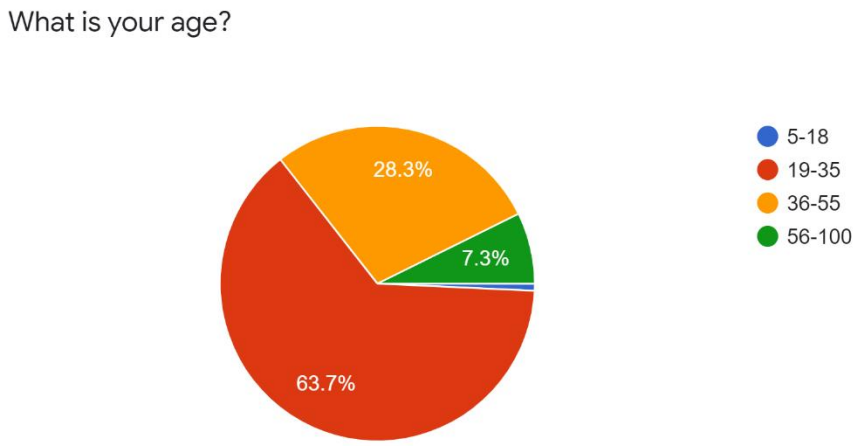


Figure 3.4.1: What is your age

2. **What is your gender?**

   From all respondents, it is found that 73.5% of the patients were male and the rest of the 26.5 were female. Males are more prone to get infected with this virus.

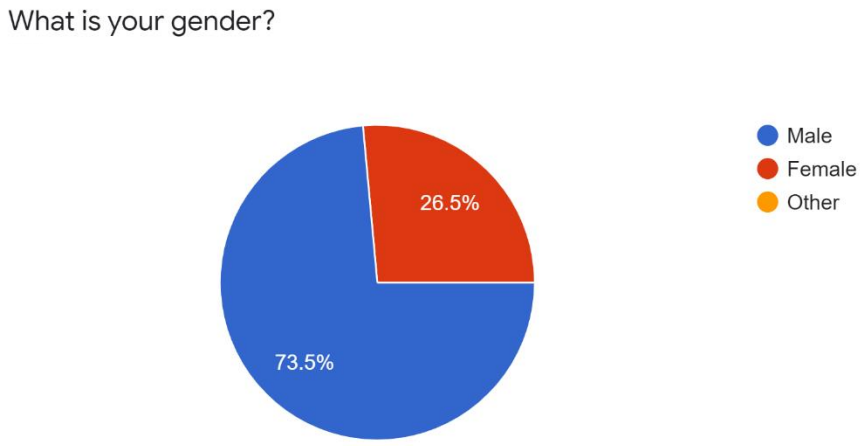


Figure 3.4.2: What is your gender

3. What is your occupation?

In the questionnaire, there were five types of occupation. Among these, 42.8% of persons were businessman or the person who works in the crowd, 23.2% are those persons who are related in other types of occupations, 22.9% are students, 10.1% are those who are in public services, rest are in healthcare.

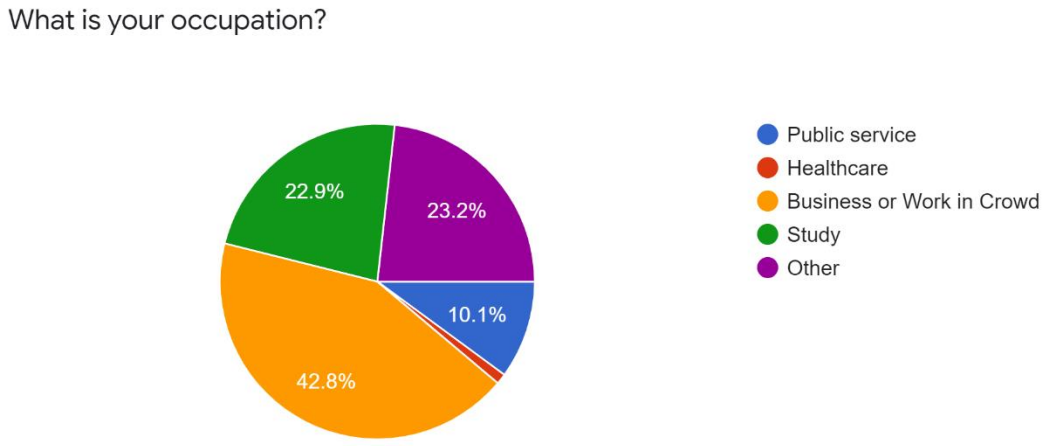


Figure 3.4.3: What is your occupation

4. **How many hours do you spend outside your home?**

From the survey it is found that most of the persons who are related to the occupation that requires to go outside. At here it is found that 26.9% person spends 10 hours, 23.4% spend 9 hours, 20.4% spend 8 hours at outside the home.

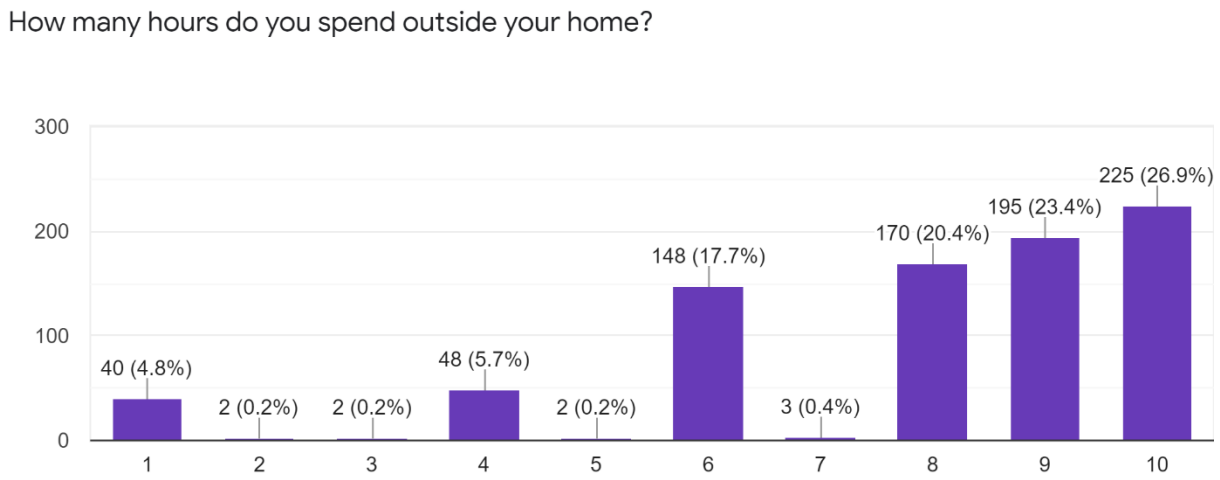


Figure 3.4.4: How many hours do you spend outside your home

5. **Do you have or had a confirmed case of Covid 19?**

All the respondents have all had a confirmed case of Covid 19.

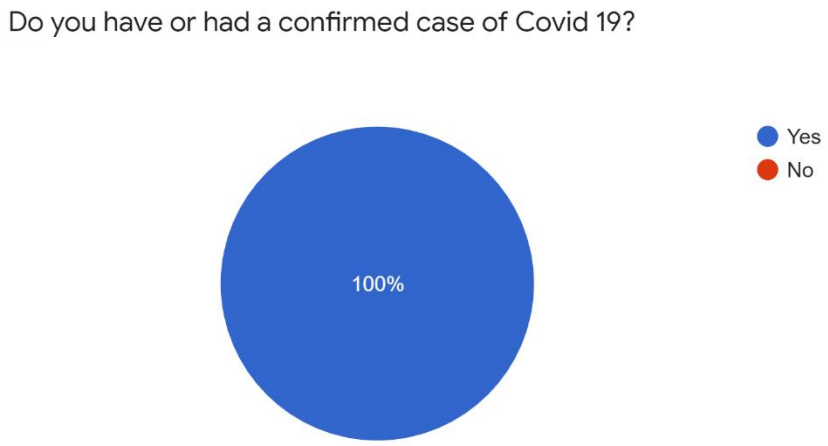


Figure 3.4.5: Do you have or had a confirmed case of Covid 19

6. **Fever > 100F in the last 14 days of infection? If yes, then tell the duration.**

   Here, 99.5% of respondents had a fever of 7-14 days, and the rest had a fever of 1-7 days.

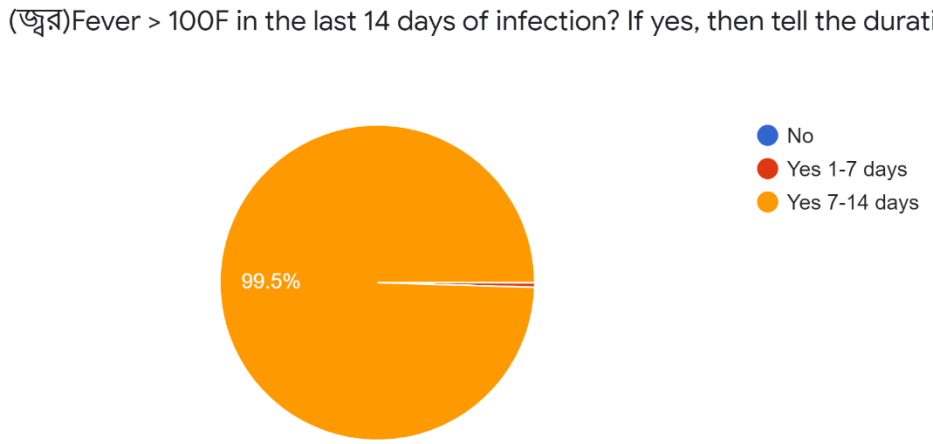


Figure 3.4.6: Fever > 100F in the last 14 days of infection

7. **(শ্বাসকষ্ট)Shortness of breaths in the last 14 days of infection? If yes, then tell the duration.**

   Among all the respondents, 96.3% had shortness of breath for a period of 7-14 days, and 2.3% had for 1-7 days.

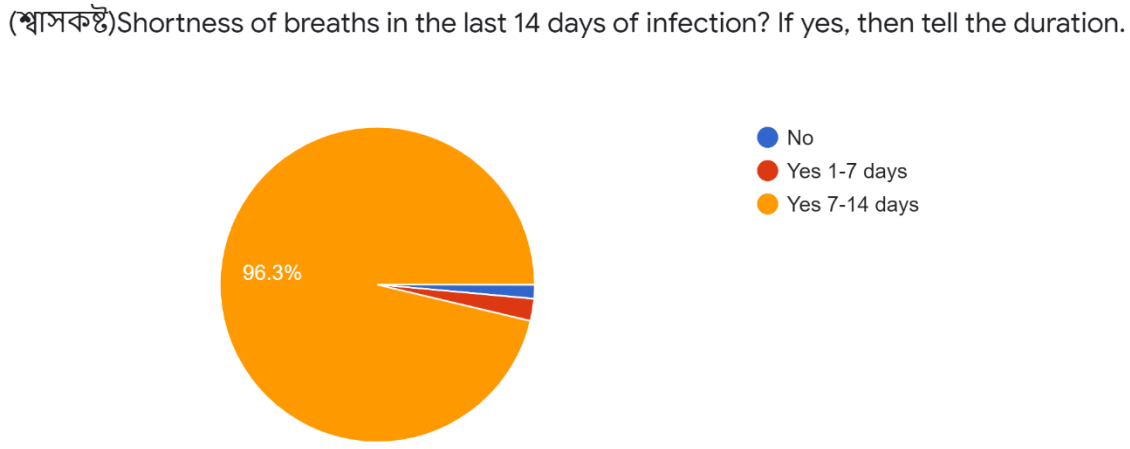


Figure 3.4.7: Shortness of breaths in the last 14 days of infection

8. **(শুকনা কাশি)Dry cough in the last 14 days of infection? If yes, then tell the duration.**

90.1% of the respondents had a dry cough for 7-14 days, and 9.6% had for 1-7 days. The rest of the 0.3% didn't have any cough.

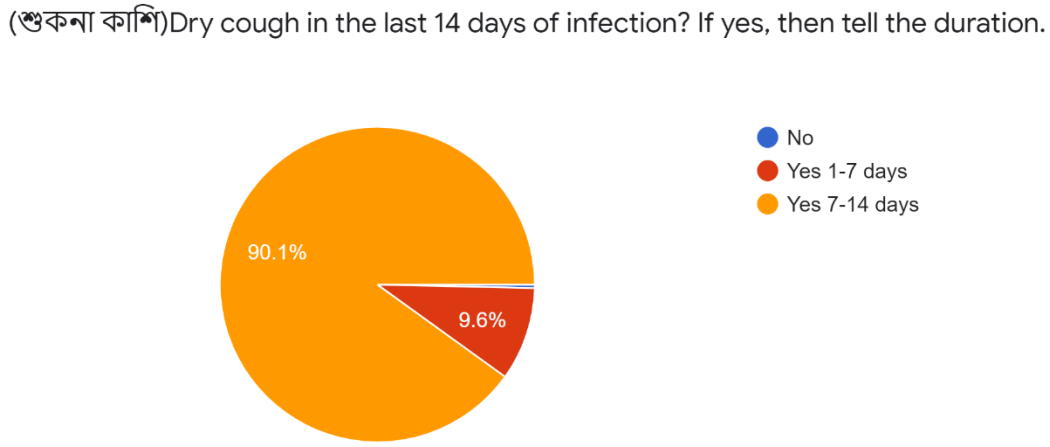


Figure 3.4.8: Dry cough in the last 14 days of infection

9. **(গলা ব্যথা)Sore Throat in the last 14 days of infection? If yes, then tell the duration.**

There were 96.2% of people had a sore throat for 7-14 days, and 3.5% had a sore throat for 1-7 days.

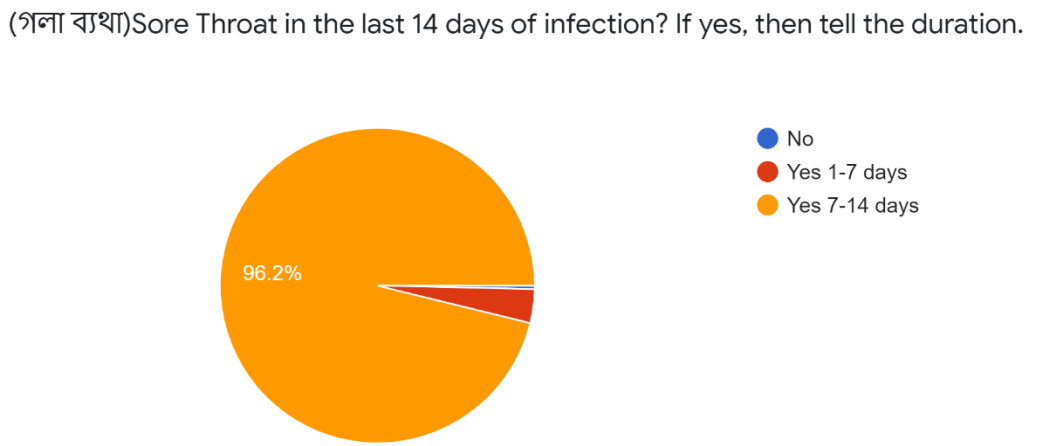


Figure 3.4.9: Sore throat in the last 14 days of infection

10. **(থু থু বা কফ এর সাথে কাশি)Cough with sputum in the last 14 days of infection? If yes, then tell the duration.**

Cough with sputum wasn't seen in 21.6% of people. Only 33.2% of people had cough with sputum for 1-7 days, and the rest of 45.3% had cough with sputum for 7-14 days.

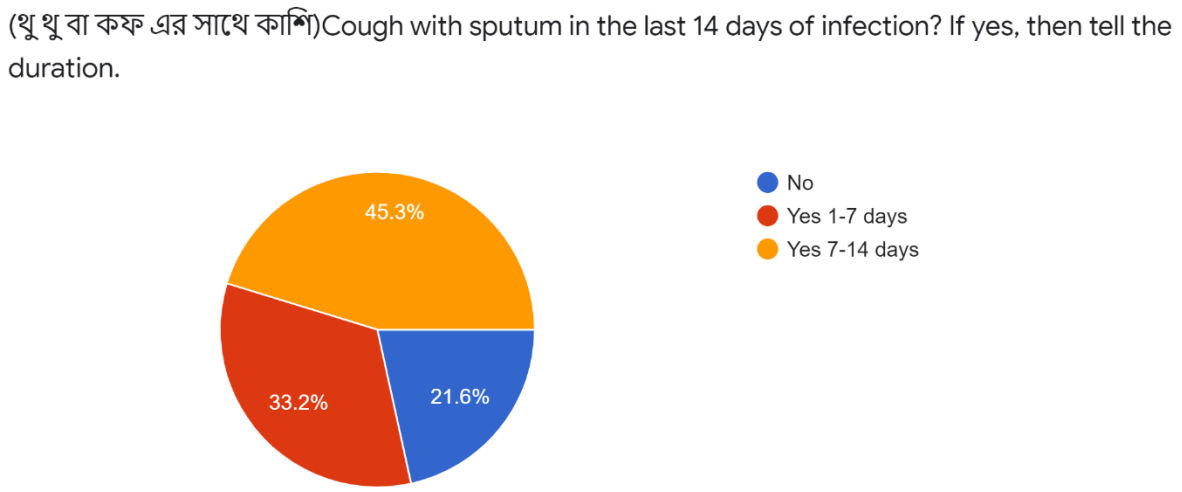


Figure 3.4.10: Cough with sputum in the last 14 days of infection

11. **(নিউমনিয়া)Pneumonia in the last 14 days of infection? If yes, then tell the duration.**

Among respondents 92.2% people didn't had pneumonia, but 7.2% people had.

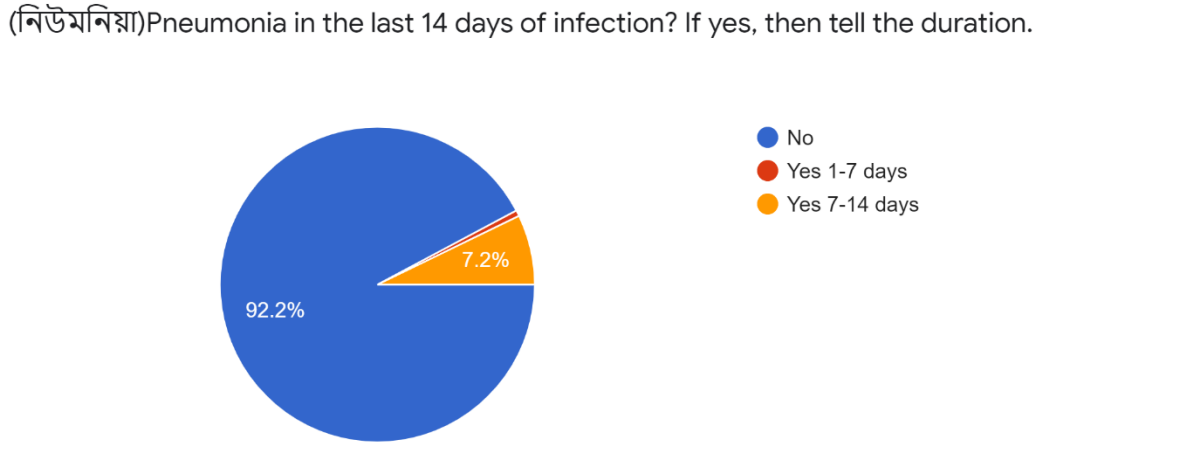


Figure 3.4.11: Pneumonia in the last 14 days of infection

12. **(পাতলা পায়খানা)Diarrhea in the last 14 days of infection? If yes, then tell the duration.**

It is seen that 89.7% of people didn't had diarrhea, but 9.2% of people had it for 1-7 days. A small portion of 1.1% had diarrhea for 7-14 days.

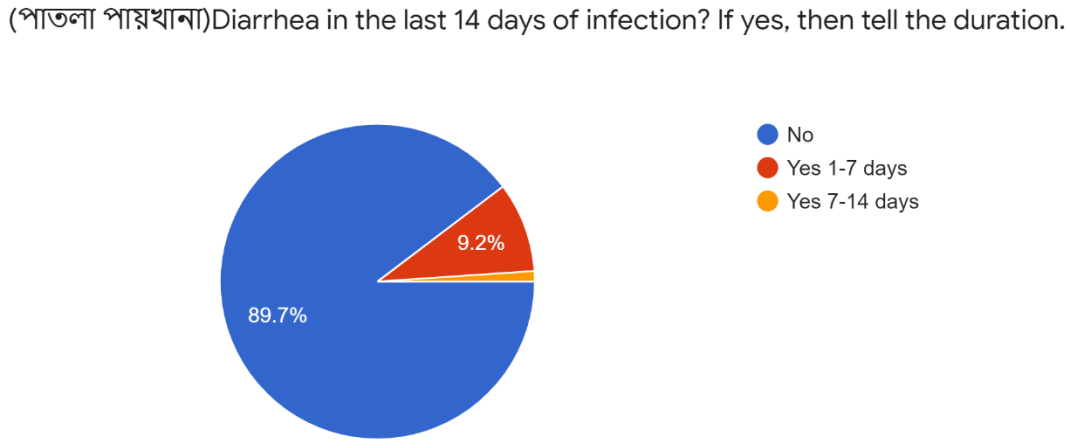


Figure 3.4.12: Diarrhea in the last 14 days of infection

13. **Do you have Anosmia or taste loss(স্বাদবিহীন মুখ)? If yes, then tell the duration.**

From these respondents, 47.2% of people had anosmia for 7-14 days, 28.6% people had anosmia for 1-7 days, and 24.2% people didn't have anosmia.

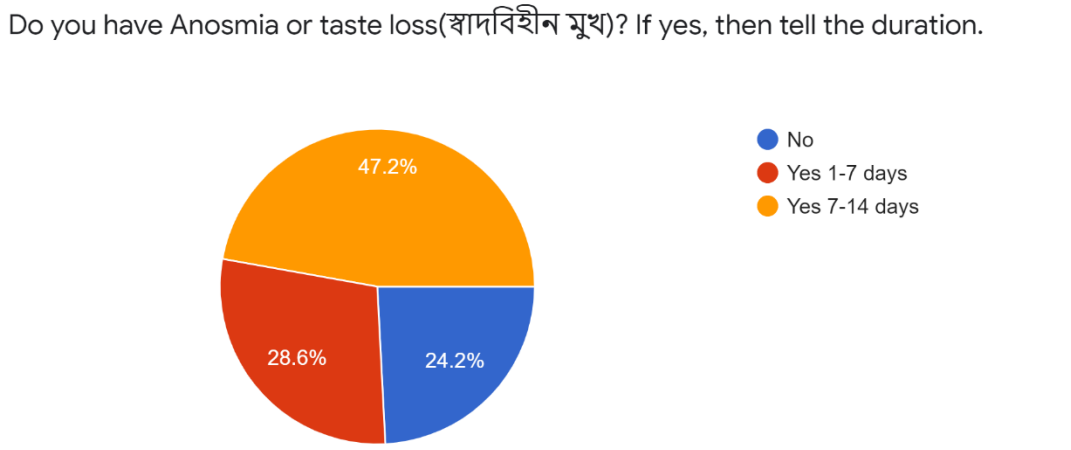


Figure 3.4.13: Do you have anosmia or taste loss

14. Do you have Fatigue(ক্লান্তি, অবস) in last 14 days of infection? If yes, then tell the duration.

Among the respondents, 41.6% had fatigue for 7-14 days, 30.4% had it for 1-7 days, and 28% didn't had fatigue.

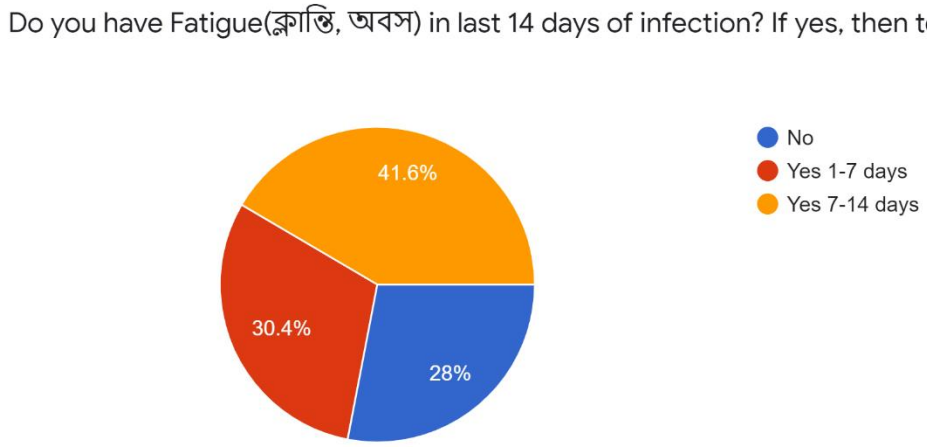


Figure 3.4.14: Do you have fatigue in the last 14 days of infection

15. Do you have Loss of Smell(ঘ্রাণ শক্তি হ্রাস) in last 14 days of infection? If yes, then tell the duration.

Due to Covid, 78.4% had suffered from loss of smell for 7-14 days, 14.9% for 1-7 days and the rest of 6.7% didn't suffer.

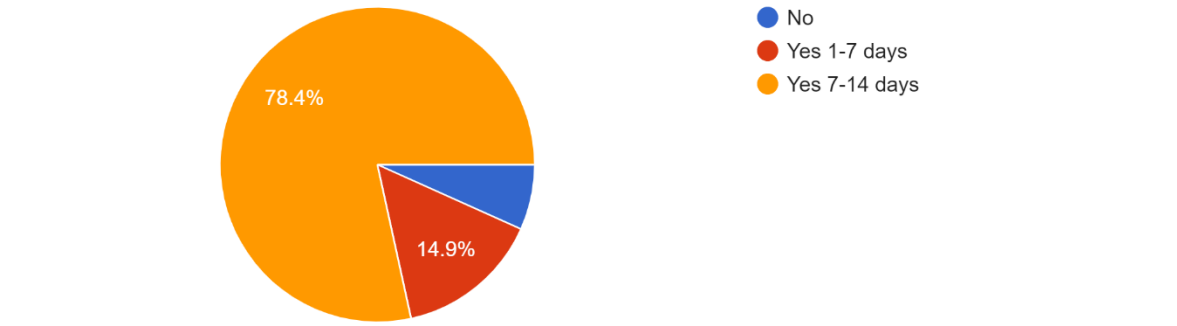


Figure 3.4.15: Do you have loss of smell in last 14 days of infection

16. Do you have Conjunctivitis(চোখ উঠা) in last 14 days of infection? If yes, then tell the duration.

Almost anyone didn't suffer from conjunctivitis. The percentage of these people is 98.6%. Rest of 1.5% of people suffered from conjunctivitis.

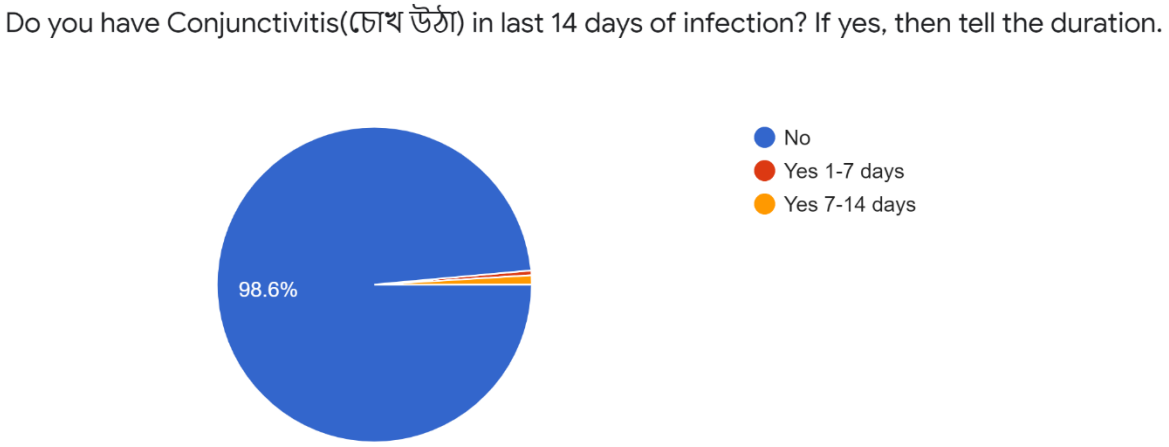


Figure 3.4.16: Do you have conjunctivitis in last 14 days of infection

17. Do you have Malaise(অস্থিরতা) in last 14 days of infection? If yes, then tell the duration.

In the earlier variant of Covid, people had malaise, but in this variant, 57.1% of people didn't have malaise. 29.5% people had malaise for 1-7 days, and 13.4% of people had malaise for 7-14 days.

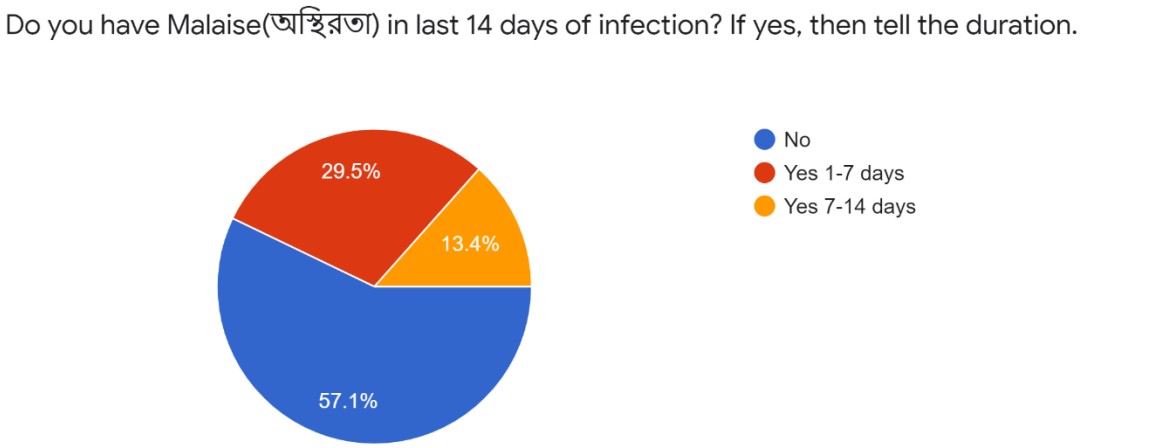


Figure 3.4.17: Do you have malaise in last 14 days of infection

18. Do you have Lung disease or COPD (শ্বাসতন্ত্র জনিত রোগ)? If yes, then tell the duration.

Most of the patients didn't had lung disease or COPD, and its percentage is 96.4%. From the rest of the respondents 2.8% had COPD for more than 1 year, and 0.8% had COPD for more than 5 years.

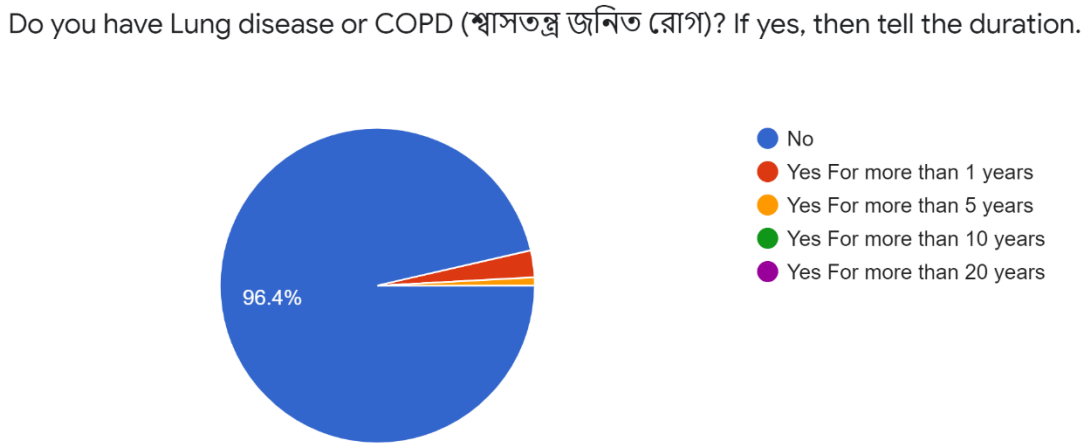


Figure 3.4.18: Do you have lung disease or COPD

19. Do you smoke? If yes, then tell the duration.

57.5% of people don't smoke, yet they were infected. 33.7% of people smoke for more than 5 years, 7.8% of people smoke for more than 1 years, and rest don't smoke.

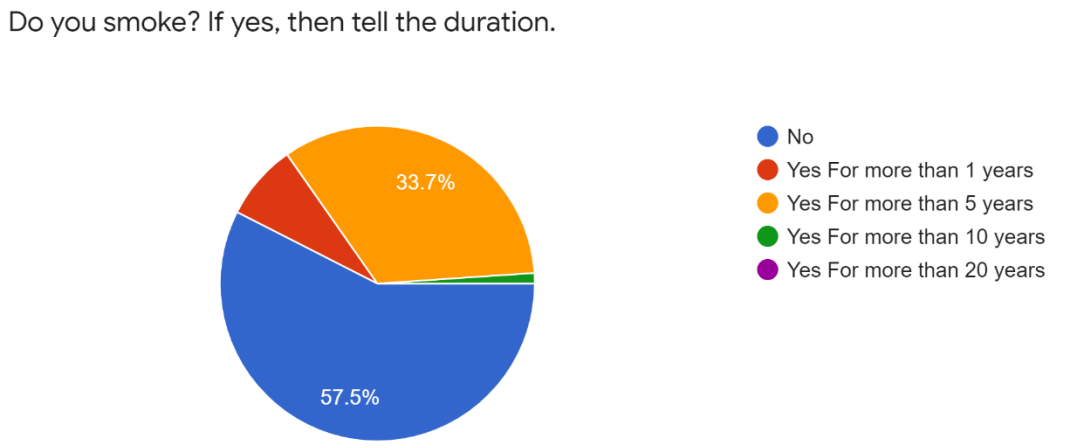


Figure 3.4.19: Do you smoke

20. Which variant of Covid case it was?

48% of people in data were infected with the Covid 19 Delta variant. 22.8% were with Beta and 28.8% were with the Alpha variant.

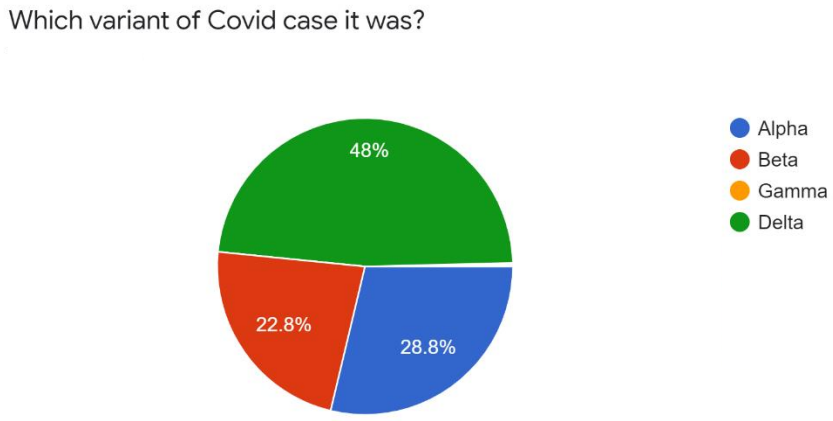


Figure 3.4.20: Which variant of covid case it was

CHAPTER 4

EXPERIMENTAL RESULTS AND DISCUSSION

4.1 Introduction

This study contains results that is determined with different algorithm techniques. In this research, popular algorithm techniques like Naïve Bayes, RandomTree, and REPTree. RandomTree provides the best result with accuracy of 99.4012%.

Random Tree is a supervised Classifier algorithm; it is an ensemble learning algorithm that generates various individual learners. This classifier employs a bagging idea for producing a random set of data for the construction of a decision tree. In typical tree each of the node is split by using the best split among all the variables. In a random forest, every node is split using the best one among the subset of the predicators randomly chosen at that node [18].

This algorithm can deal with both regression and classification problems. RandomTree is a collection of tree predictors that is known as forest. This classification first takes the input feature vector, and classifies it with every tree in the forest. Then outputs the class label that is received with the votes of majority. For regression, the classifier response is average of responses over all the trees in the forest [19].

4.2 Experimental Results

Best experimental result is found on applying Random Tree classifier algorithm with the following classify attributes.

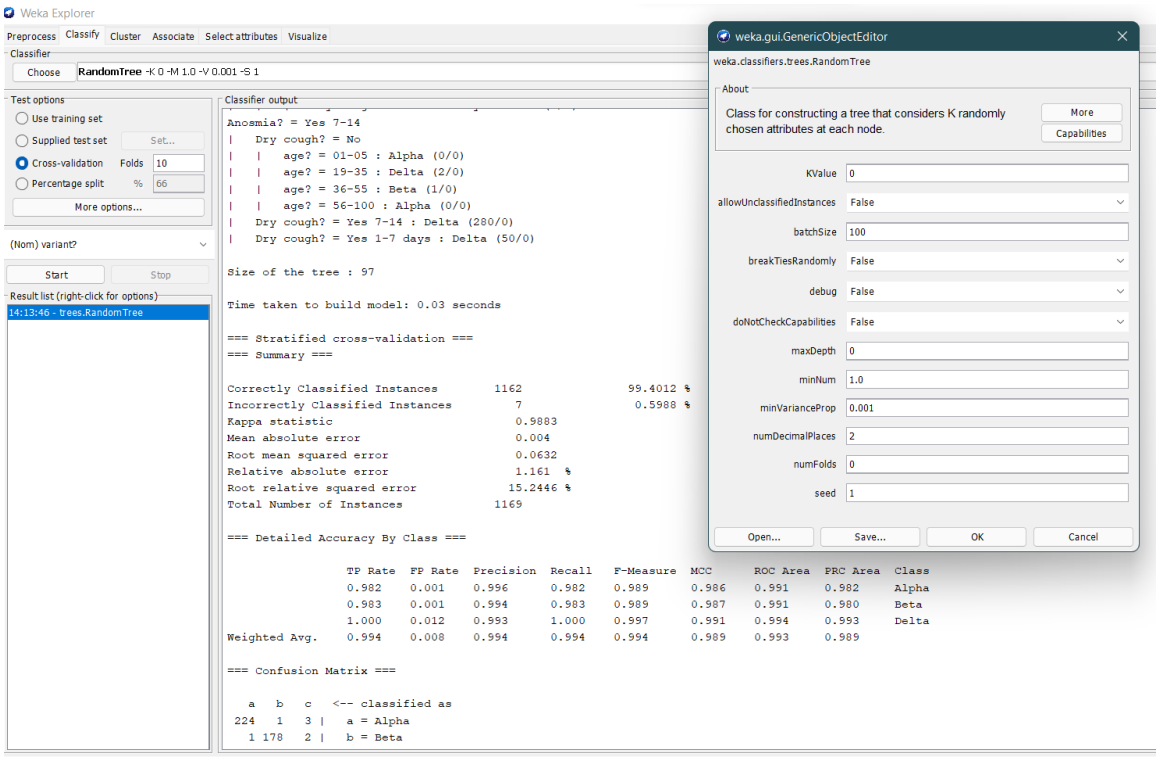


Figure 4.2.1: Attributes of classifier

At here KValue is 0, batchsize is 100, maxDepth 0, numFolds 0 and seed is 1. The test mode is 10-fold cross-validation.

```
=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1162               99.4012 %
Incorrectly Classified Instances         7                0.5988 %
Kappa statistic                          0.9883
Mean absolute error                      0.004
Root mean squared error                  0.0632
Relative absolute error                  1.161  %
Root relative squared error             15.2446 %
Total Number of Instances             1169
```

Figure 4.2.2: Stratified cross-validation

At here using RandomTree classifier 1162 instances were correctly classified and 7 instances were incorrectly classified. The percentage of correctly classified instances is 99.4012% and incorrectly classified instances is 0.5988%. The number of total instances are 1169.

```
=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.982    0.001    0.996      0.982    0.989      0.986    0.991     0.982     Alpha
                 0.983    0.001    0.994      0.983    0.989      0.987    0.991     0.980     Beta
                 1.000    0.012    0.993      1.000    0.997      0.991    0.994     0.993     Delta
Weighted Avg.    0.994    0.008    0.994      0.994    0.994      0.989    0.993     0.989
```

Figure 4.2.3: Detailed Accuracy By Class

The true positive rate for Alpha, Beta, and Delta variant are 0.982, 0.983, and 1.000. The false positive rate for Alpha, Beta, and Delta variant are 0.001, 0.001, and 0.012. The precision rate for Alpha, Beta and Delta variant are 0.996, 0.994, and 0.993.

```
=== Confusion Matrix ===

   a   b   c   <-- classified as
 224   1   3 |   a = Alpha
   1 178   2 |   b = Beta
   0   0 760 |   c = Delta
```

Figure 4.2.4: Confusion Matrix Values

In the confusion matrix we can see that for 228 patients of Alpha variant 224 are correctly predicted, 1 was predicted as beta and 3 was predicted as Delta.

For 181 patients of Beta variant, 178 was predicted correctly, 1 was predicted as alpha, and 2 was predicted as delta.

For 760 patients of Delta variant, 760 was predicted correctly and non was predicted incorrectly.

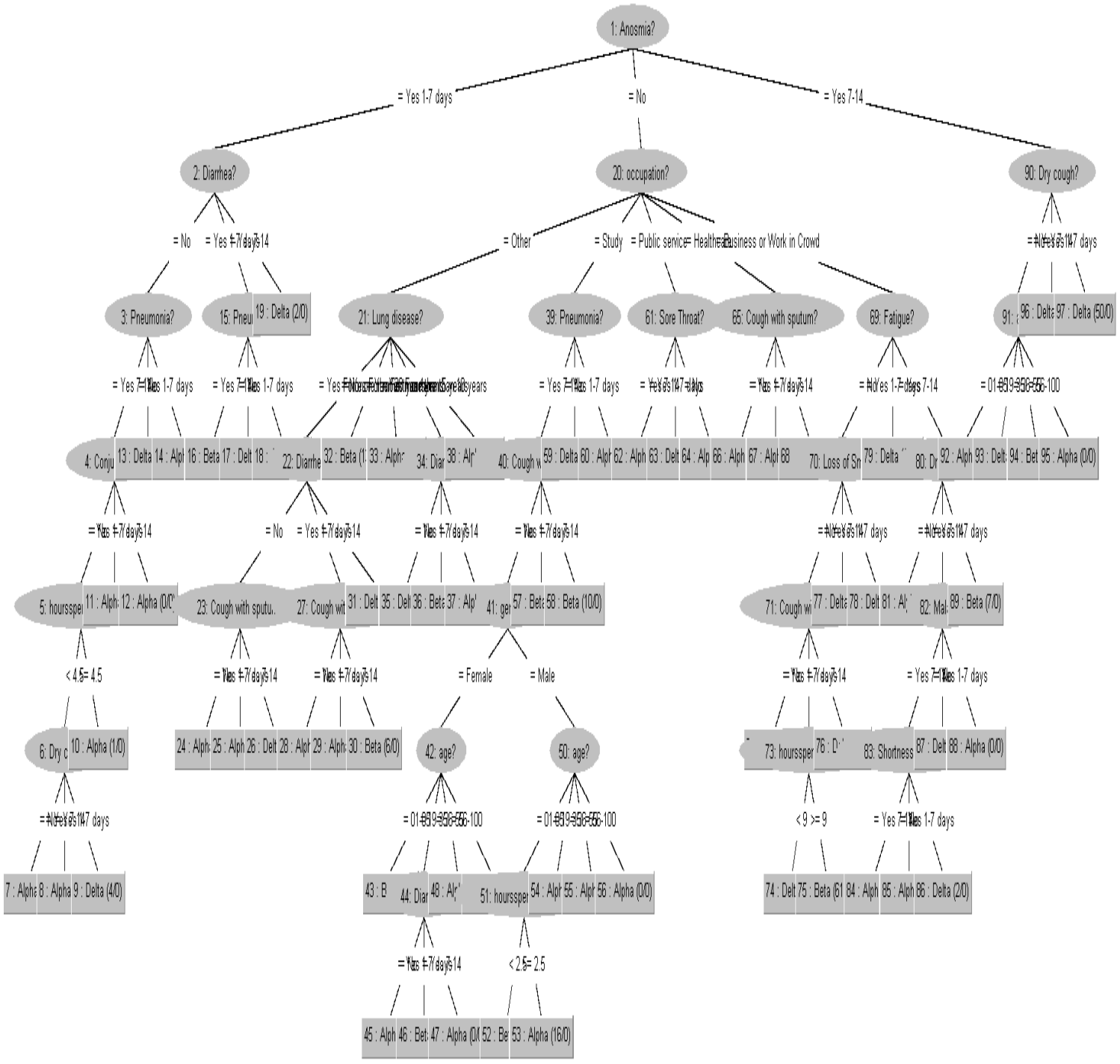


Figure 4.2.5: Tree generation for the classifier

4.3 Model Diagram

Anosmia? = Yes 1-7 days
| Diarrhea? = No
| | Pneumonia? = Yes 7-14
| | | Conjunctivitis? = No
| | | | hoursspend outside? < 4.5
| | | | | Dry cough? = No : Alpha (1/0)
| | | | | Dry cough? = Yes 7-14 : Alpha (0/0)
| | | | | Dry cough? = Yes 1-7 days : Delta (4/0)
| | | | hoursspend outside? >= 4.5 : Alpha (1/0)
| | | Conjunctivitis? = Yes 1-7 days : Alpha (11/0)
| | | Conjunctivitis? = Yes 7-14 : Alpha (0/0)
| | Pneumonia? = No : Delta (218/0)
| | Pneumonia? = Yes 1-7 days : Alpha (0/0)
| Diarrhea? = Yes 1-7 days
| | Pneumonia? = Yes 7-14 : Beta (35/0)
| | Pneumonia? = No : Delta (2/0)
| | Pneumonia? = Yes 1-7 days : Alpha (0/0)
| Diarrhea? = Yes 7-14 : Delta (2/0)
| | Lung disease? = No
| | | Diarrhea? = No
| | | | Cough with sputum? = No : Alpha (81/0)
| | | | Cough with sputum? = Yes 1-7 days : Alpha (0/0)
| | | | Cough with sputum? = Yes 7-14 : Delta (2/0)
| | | Diarrhea? = Yes 1-7 days
| | | | Cough with sputum? = No : Alpha (1/0)
| | | | Cough with sputum? = Yes 1-7 days : Alpha (0/0)
| | | | Cough with sputum? = Yes 7-14 : Beta (6/0)
| | | Diarrhea? = Yes 7-14 : Delta (2/0)
| | Lung disease? = Yes For more than 20 years : Beta (13/0)
| | Lung disease? = Yes For more than 1years : Alpha (0/0)
| | Lung disease? = Yes For more than 5 years
| | | Diarrhea? = No : Delta (2/0)
| | | Diarrhea? = Yes 1-7 days : Beta (14/0)
| | | Diarrhea? = Yes 7-14 : Alpha (0/0)

Anosmia? = No
| occupation? = Other

| | Pneumonia? = Yes 7-14
| | | Cough with sputum? = No
| | | | gender? = Female
| | | | | age? = 01-05 : Beta (5/0)
| | | | | age? = 19-35
| | | | | | Diarrhea? = No : Alpha (1/0)
| | | | | | Diarrhea? = Yes 1-7 days : Beta (5/0)
| | | | | | Diarrhea? = Yes 7-14 : Alpha (0/0)
| | | | | age? = 36-55 : Alpha (0/0)
| | | | | age? = 56-100 : Alpha (0/0)
| | | | gender? = Male
| | | | | age? = 01-05
| | | | | | hoursspend outside? < 2.5 : Beta (1/0)
| | | | | | hoursspend outside? >= 2.5 : Alpha (16/0)
| | | | | age? = 19-35 : Alpha (1/0)
| | | | | age? = 36-55 : Alpha (0/0)
| | | | | age? = 56-100 : Alpha (0/0)
| | | Cough with sputum? = Yes 1-7 days : Beta (10/0)
| | | Cough with sputum? = Yes 7-14 : Beta (10/0)

| | Lung disease? = Yes For more than 10 years : Alpha (0/0)
| occupation? = Study

| | Sore Throat? = Yes 7-14 : Alpha (1/0)
| | Sore Throat? = Yes 1-7 days : Delta (4/0)
| | Sore Throat? = No : Alpha (0/0)
| occupation? = Healthcare
| | Cough with sputum? = No : Alpha (42/0)
| | Cough with sputum? = Yes 1-7 days : Alpha (0/0)
| | Cough with sputum? = Yes 7-14 : Beta (13/0)
| occupation? = Business or Work in Crowd
| | Fatigue? = No
| | | Loss of Smell? = No
| | | | Cough with sputum? = No : Alpha (41/0)
| | | | Cough with sputum? = Yes 1-7 days
| | | | | hoursspend outside? < 9 : Delta (2/0)
| | | | | hoursspend outside? >= 9 : Beta (61/0)
| | | | Cough with sputum? = Yes 7-14 : Delta (2/0)

| | Pneumonia? = No : Delta (142/0)
| | Pneumonia? = Yes 1-7 days : Alpha (0/0)
| occupation? = Public service
| | | Loss of Smell? = Yes 7-14 : Delta (22/0)
| | | Loss of Smell? = Yes 1-7 days : Delta (6/0)
| | Fatigue? = Yes 1-7 days : Delta (12/0)
| | Fatigue? = Yes 7-14
| | | Dry cough? = No : Alpha (0/0)
| | | Dry cough? = Yes 7-14
| | | | Malaise? = Yes 7-14
| | | | | Shortness of breaths? = Yes 7-14 : Alpha (31/0)
| | | | | Shortness of breaths? = No : Alpha (0/0)
| | | | | Shortness of breaths? = Yes 1-7 days : Delta (2/0)
| | | | Malaise? = No : Delta (4/0)
| | | | Malaise? = Yes 1-7 days : Alpha (0/0)
| | | Dry cough? = Yes 1-7 days : Beta (7/0)
Anosmia? = Yes 7-14
| Dry cough? = No
| | age? = 01-05 : Alpha (0/0)
| | age? = 19-35 : Delta (2/0)
| | age? = 36-55 : Beta (1/0)
| | age? = 56-100 : Alpha (0/0)
| Dry cough? = Yes 7-14 : Delta (280/0)

| Dry cough? = Yes 1-7 days : Delta
(50/0)

At here the size of the tree is 97.

4.4 Result of Other Classifiers

There are also other algorithms that are used in this data set. These classifiers are Naïve Bayes, and REPTree.

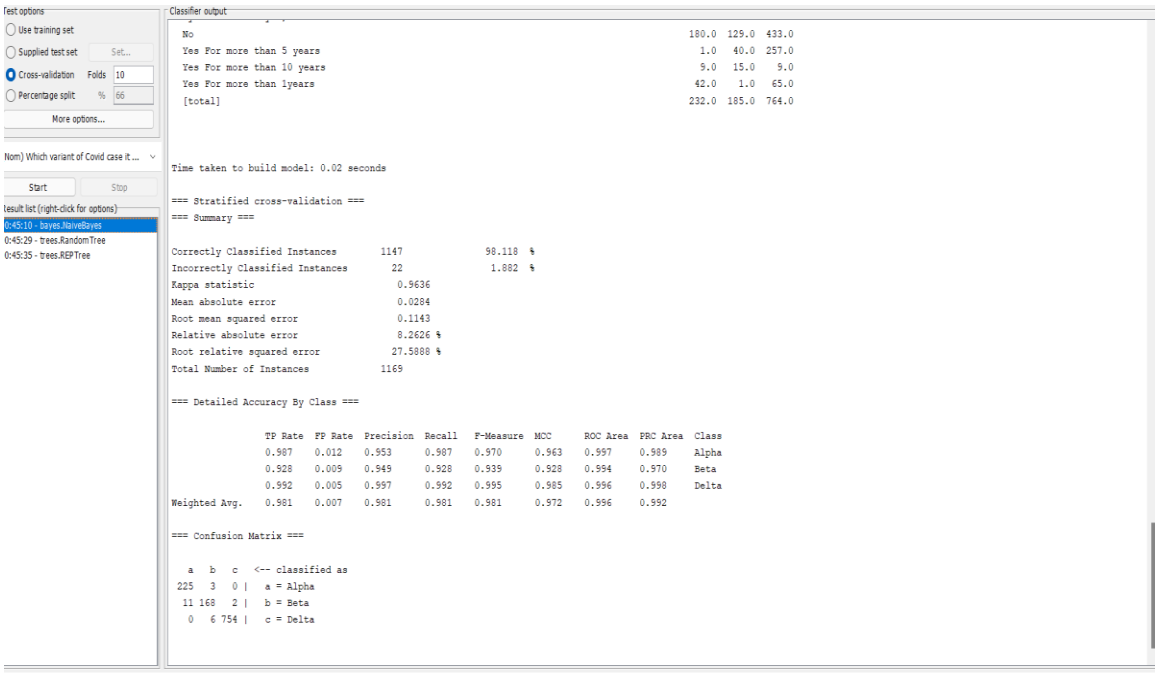


Figure 4.4.1: Result in Naïve Bayes classifier

The correctly classified instances in Naïve Bayes classifier is 98.118%.

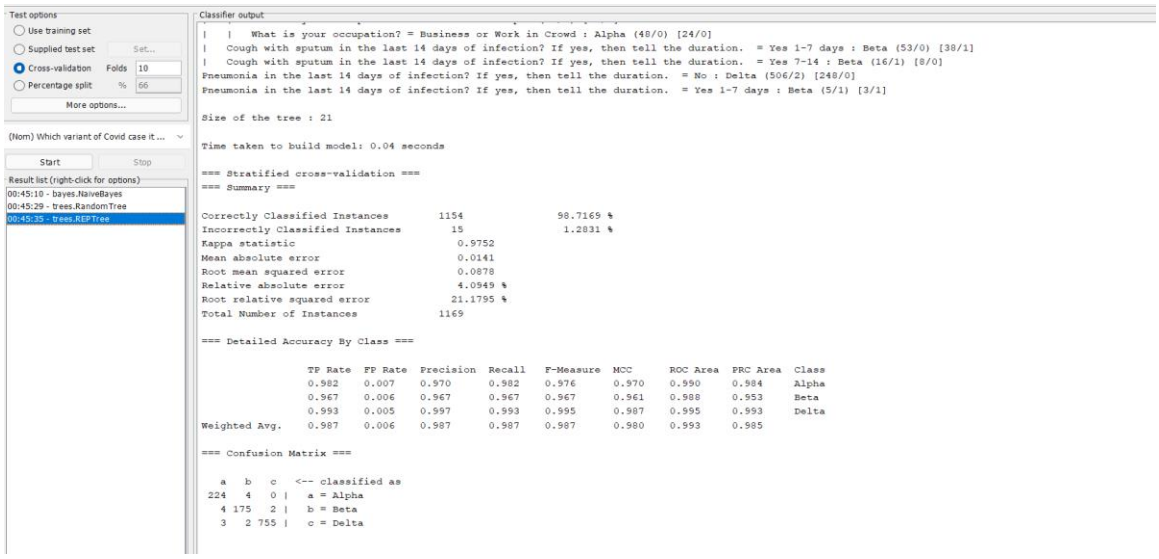


Figure 4.4.2: Result in REPTree classifier

The correctly classified instances in REPTree classifier is 98.7169%.

The tree generation of REPTree classifier is below:

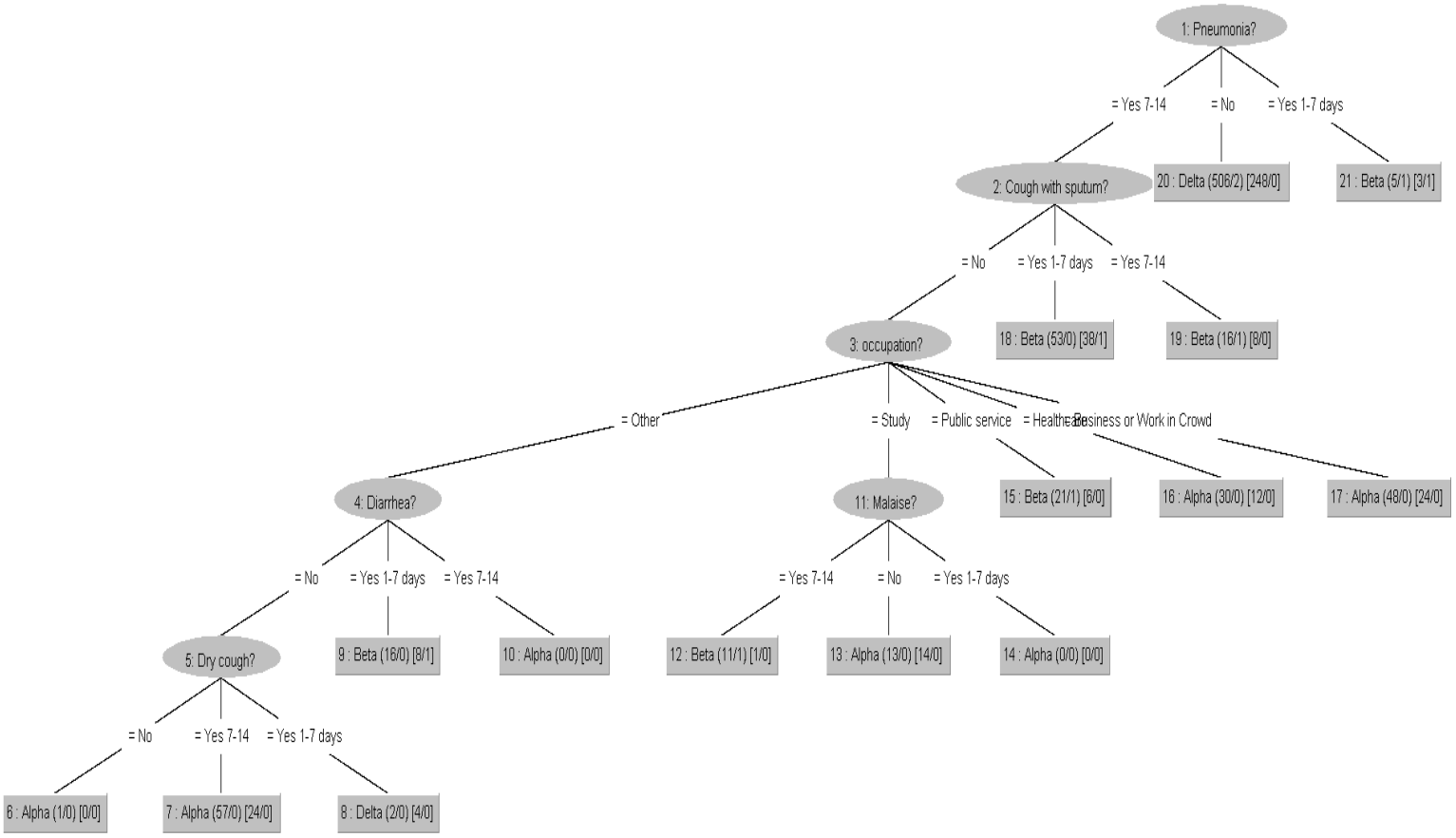


Figure 4.4.3: Tree generation for REPTree classifier

4.5 Descriptive Analysis

In this research, the best result is found using RandomTree classifier with 10 folds. The experimental result matches with the expected result. In real life if we match this data, we can discriminate Delta variant with other variants of Covid 19.

According to the result of the analysis:

a. **Duration of anosmia**: Some of the people seems to have anosmia but most of the patients didn't had anosmia.

b. **Age of affected people**: It is seen that most of the person who are affected by the Delta variant is people who are from age 19-35. These are mostly adult persons who stay outside of the home most of the time.
c. **Gender of affected people**: It is seen that rather than women, men are more likely to get infected with the Delta variant of the Corona virus.
d. **Occupation of the people**: People who work or stays outside are more likely to get affected by Covid 19 Delta variant.
e. **Hour's people spend outside who are affected by delta variant**: Different people stay different hours at outside the home. From the result, it is seen that people who stay more than 10 hours are more likely to get affected, and on average, this time is above 7 hours.
f. **Confirmed case in people**: According to the report, all the respondents have or had a confirmed case of Covid 19.
g. **Duration of fever**: From the report, it is seen that those who were affected with the delta variant had a fever of 7-14 days.
h. **Duration of shortness of breath**: The report also shows that people had shortness of breath for 7-14 days.
i. **Duration of dry cough**: The duration of dry cough is 7-14 days.
j. **Duration of sore throat**: Most of the respondents had a sore throat for 7-14 days.
k. **Duration of cough with sputum**: Most of the respondents had cough with sputum for 1-7 days.
l. **Duration of pneumonia**: Most of the respondents didn't had Pneumonia.
m. **Duration of diarrhea**: Most of the respondents didn't had Diarrhea.
n. **Duration of fatigue**: During the infection period, fatigue was found for 1-7 days.
o. **Duration of loss of smell**: Other variants didn't have this symptom, but in Delta variant, people suffer from loss of smell. According to the analysis, the period is 7-14 days.
p. **Duration of conjunctivitis**: Some of the cases in Delta variant has been seen who suffer from conjunctivitis and problems in the eyes. But in most of the cases, Conjunctivitis is not found.

q. **Duration of malaise**: In other variants, people would suffer from shortness of breath and become malaise. In this variant report, some people seem to suffer from malaise, but most of them didn't suffer.

r. **Duration of lung disease or COPD**: Most of the respondents didn't have any lung disease or COPD. That means it can be more fatal for those people who have lung disease or COPD.

s. **Duration of smoking**: It is seen in the result that people were not a smoker.

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Conclusion

General cough and cold are not Covid 19. Covid 19 and its variants have their patterns and specific symptoms. It is seen that all the variants of Covid 19 have some common symptoms, including fever, shortness of breath, fatigue, etc. Anyone affected with Covid 19 can have these symptoms, but it is almost impossible to confirm if it is Covid or not and what is the variant without any test. This data analysis will help one to find the major symptoms of Covid 19 Delta variant, and one will be able to easily tell it.

The Alpha and Beta variants of Covid 19 had the same symptoms, but the effects were different and very much severe. In the Beta variant, more people died than Alpha. But due to lacking of test, people were not sure if he was affected by Corona. They didn't had any proper list of symptoms that could tell if he is infected or not and what the variant is. Everyone should know about the variants of Covid 19 and the symptoms of its variants. So that one person can easily identify the variants and take proper cautions. In this way, many lives could be saved.

In this research, it is shown that what are the major symptoms of the Covid 19 Delta variant. People had several symptoms who were affected with the Delta variant. After collecting the data, algorithms are applied. Then major symptoms along with the duration are found. These symptoms include fever, shortness of breath, dry cough, sore throat, fatigue, loss of smell, cough with sputum, etc.

5.2 Limitation of the Study

As the study has benefits, but there are some limitations too. This study was conducted in Bangladesh and in summer. Therefore, the results might not match exactly for other countries. The machine might detect the the wrong variation for the patient or not detect at all.

5.3 Future Work

In this way, one can find the major symptoms for other variants of Covid 19. The latest variant of Covid 19 is Omicron. It is reported that vaccines work on a good percentage of people of all the old variants. Some people still get infected with the old variant of the Covid, and others are getting infected with the Omicron variant.

Finding the symptoms of Omicron and other variants will help the researcher to find the vaccine that will be effective for this and all the future variants.

REFERENCE

[1] Fatima M. Salman, Samy S. Abu-Naser, Eman Alajrami, "COVID-19 Detection using Artificial Intelligence," IJAER ISSN2643-9085, vol. 4, Issue 3, pp 18-25, March 2020.

[2] Moutaz Alazab, Albara Awajan, Abdelwadood Mesleh, "COVID-19 Prediction and Detection Using Deep Learning," IJCISIM ISSN 2150-7988, vol. 12, pp168-181, May 2020.

[3] Brinati, D., Campagner, A., Ferrari, D. et al. "Detection of COVID-19 Infection from Routine Blood Exams with Machine Learning: A Feasibility Study,". J Med Syst 44, July 2020.

[4] Nanshan Chen, Min Zhou, Xuan Dong, "Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study," The Lancet ISSN 0140-6736, vol 395, Issue 10223, pp 507-513, January 2020.

[5] Alireza Tahamatan, Abdollah Ardebili, "Real-time RT-PCR in COVID-19 detection: issues affecting the results," Taylor & Francis, vol 20, Issue 5, May 2020.

[6] Jianguo Wu, Jiasheng Liu, Shijun Li, "Detection and analysis of nucleic acid in various biological samples of COVID-19 patients," travel Med Infect Dis, April 2020.

[7] Mohammad Anwar Hossain, Md. Iqbal Kabir Jahid, K. M Amran Hossain, "Knowledge, attitudes, and fear of COVID-19 during the Rapid Rise Period in Bangladesh" Plos One, September 2020.

[8] Learn about Conclusion, available at << https://www.deseret.com/coronavirus/2021/10/24/22652369/delta-variant-signs-symptoms-you-have-it >>, last accessed on 30-11-2021 at 2:12pm.

[9] Kai Kupferschmidt, Meredith Wadman, "Delta variant triggers new phase in the pandemic," vol 372, Issue 6549, pp 1375-1376, 25 June 2021.

[10] Adrian O'Dowd, "Covid-19: Cases of delta variant rise by 79%, but rate of growth slows," BMJ 2021;373:n1596, 21 June 2021.

[11] Del Rio C, Malani PN, Omer SB, "Confronting the Delta Variant of SARS-CoV-2," JAMA. 2021;vol 326, Issue 11, pp 1001–1002, 18 August 2021

[12] 5 Things To Know About the Delta Variant, available at << https://www.haptonomie-angers.fr/wp-content/uploads/2021/07/5-choses-a-savoir-sur-la-variante-Delta-_-Actualites-_-Yale-Medicine.pdf >>, last accessed on 05-01-2022 at 10:24pm.

[13] Jeromie Wesly Vivian Thangaraj, Pragya Yadav, CP Girish Kumar, "Predominance of delta variant among the COVID-19 vaccinated and unvaccinated individuals, India, May 2021" Doi: https://doi.org/10.1016/j.jinf.2021.08.006, 05 August, 2021.

[14] Alizon Samuel, Haim-Boukobza Stéphanie, Foulongne Vincent, Verdurme Laura, Trombert-Paolantoni Sabine, Lecorche Emmanuel, Roquebert Bénédicte, Sofonea Mircea T, "Rapid spread of the SARS-CoV-2 Delta variant in some French regions," June 2021. Euro Surveill. 2021;26(28):pii=2100573. https://doi.org/10.2807/1560-7917.ES.2021.26.28.2100573, 15 July 2021

[15] Zhang M., Xiao J., Deng A., Zhang Y., Zhuang Y., Hu T., "Transmission dynamics of an outbreak of the COVID-19 Delta variant B. 1.617. 2—Guangdong Province, China," China CDC Wkly, pp. 3(27): 584–586. 2 July 2021.

[16] Ying Liu, Joacim Rocklöv, "The reproductive number of the Delta variant of SARS-CoV-2 is far higher compared to the ancestral SARS-CoV-2 virus," Journal of Travel Medicine, Volume 28, Issue 7, taab124, 9 August 2021,

[17] Farinholt, T., Doddapaneni, H., Qin, X. et al, "Transmission event of SARS-CoV-2 delta variant reveals multiple vaccine breakthrough infections," BMC Med 19, 255 (2021), https://doi.org/10.1186/s12916-021-02103-4, 1 October 2021

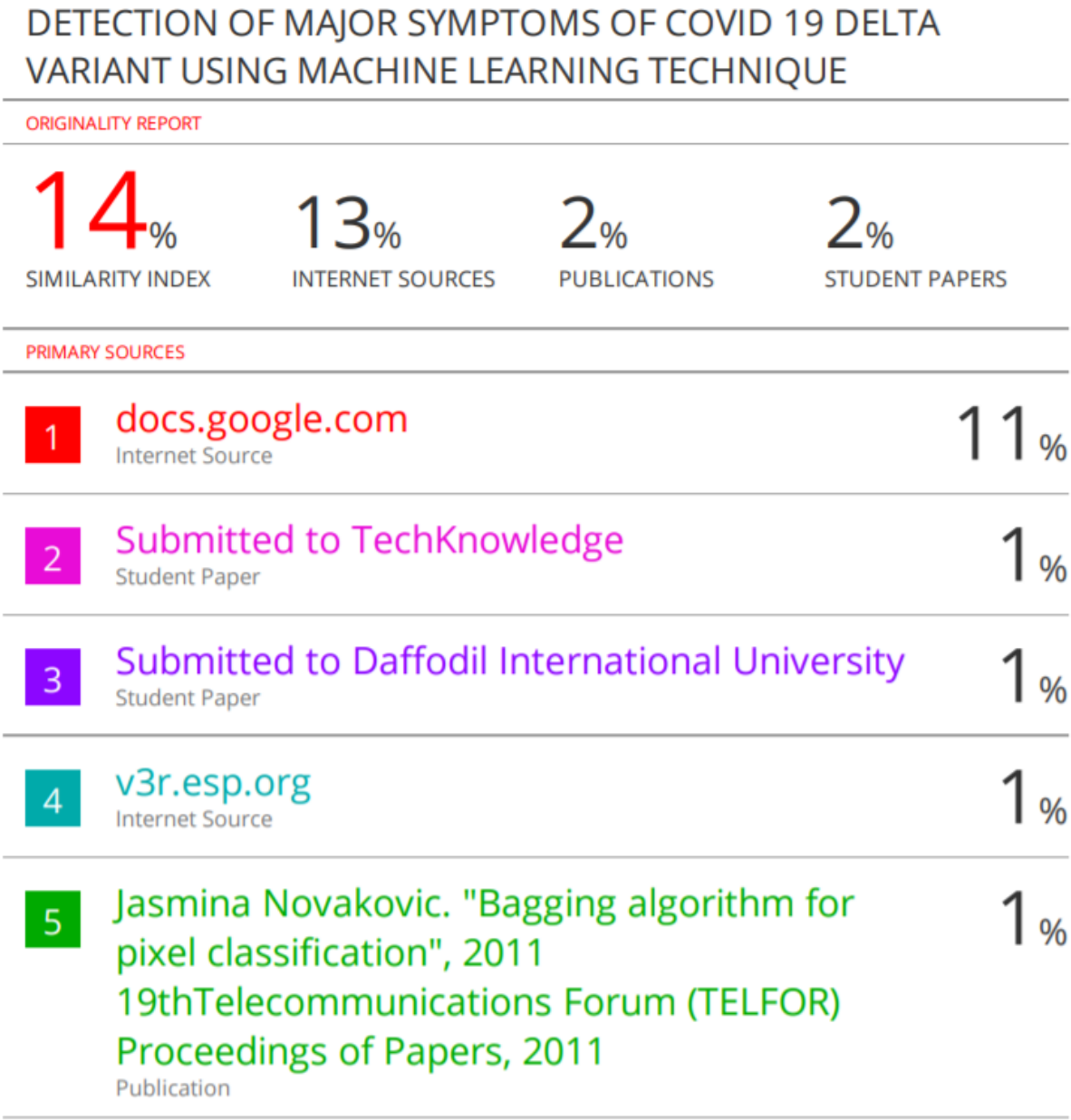
[18] Pragya D Yadav, PhD, Gajanan N Sapkal, PhD, DM, "Neutralization of Beta and Delta variant with sera of COVID-19 recovered cases and vaccinees of inactivated COVID-19 vaccine BBV152/Covaxin," Journal of Travel Medicine, Volume 28, Issue 7, October 2021, taab104, https://doi.org/10.1093/jtm/taab104, 6 July 2021

[19] SushilKumar Kalmegh, "Analysis of WEKA Data Mining Algorithm REPTree, Simple Cart and Random Tree for Classification of Indian News", IJISET, Vol 2 Issue 2, ISSN 2348-7968, February 2015

APPENDIX

1. Link of data collection questionnaire.
   https://forms.gle/BmZQXSMx9bs5Sm1a9
2. Link of Google site.
   https://sites.google.com/diu.edu.bd/covid-19-delta-symptoms

PLAGIARISM REPORT



DETECTION OF MAJOR SYMPTOMS OF COVID 19 DELTA VARIANT USING MACHINE LEARNING TECHNIQUE

ORIGINALITY REPORT

14%	13%	2%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	docs.google.com Internet Source	11%
2	Submitted to TechKnowledge Student Paper	1%
3	Submitted to Daffodil International University Student Paper	1%
4	v3r.esp.org Internet Source	1%
5	Jasmina Novakovic. "Bagging algorithm for pixel classification", 2011 19thTelecommunications Forum (TELFOR) Proceedings of Papers, 2011 Publication	1%