

# **Depression Level Prediction On Bangladeshi Students**

**BY**

**Barnali Adhikari**

**ID: 172-15-9609**

**Syeda Ismatara Era**

**ID: 172-15-10151**

**AND**

**Mohamed Dahir Mohamed**

**ID: 172-15-10089**

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

Supervised By

**Md. Jueal Mia**

Senior Lecturer

Department of CSE

Daffodil International University

Co-Supervised By

**Mr. Abdus Sattar**

Assistant Professor

Department of CSE

Daffodil International University



**DAFFODIL INTERNATIONAL UNIVERSITY**

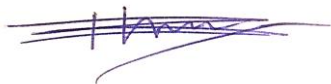
**DHAKA, BANGLADESH**

## **APPROVAL**

This Project/internship titled “**Depression Level Prediction On Bangladeshi Students**”, submitted by \* Barnali Adhikari \*,\*Mohamed Dahir Mohamed\* and \* Syeda Ismatara Era\* ID No: 172-15-9609, 172-15-10089 and 172-15-10151 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 2 June 2021.

## **BOARD OF EXAMINERS**

**Chairman**



---

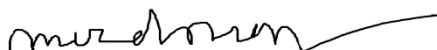
**Dr. Touhid Bhuiyan**

**Professor and Head**

Department of Computer Science and Engineering

Faculty of Science & Information Technology

Daffodil International University



---

**Md. Riazur Rahman**

**Assistant Professor**

Department of Computer Science and Engineering

Faculty of Science & Information Technology

Daffodil International University



---

**Nusrat Jahan**

**Senior Lecturer**

Department of Computer Science and Engineering

Faculty of Science & Information Technology

Daffodil International University

**Internal Examiner**

**Internal Examiner**



**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 **Md. Jueal Mia, Senior Lecturer, Department of CSE** and Co-Supervised **Mr. Abdus Sattar, Assistant 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 the award of any degree or diploma. **Supervised by:**



---

**Md. Jueal Mia**  
Senior Lecturer  
Department of CSE  
Daffodil International University

**Co-Supervised by:**



---

**Mr. Abdus Sattar**  
Assistant Professor  
Department of CSE  
Daffodil International University

**Submitted by:**

**Barnali Adhikari**

*Barnali*

---

ID: 172-15-9609  
Department of CSE  
Daffodil International University

**Syeda Ismatara Era**

*Syeda Ismatara Era*

---

ID: 172-15-10151

Department of CSE

Daffodil International University

**Mohamed Dahir Mohamed**



---

ID: 172-15-10089

Department of CSE

Daffodil International University

Daffodil International University

## ACKNOWLEDGEMENT

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

We grateful and wish our profound indebtedness to **Md. Jueal Mia, Senior Lecturer**, Department of CSE Daffodil International University, Dhaka. Machine Learning and Data Mining interest of our supervisor in the field of “AI” to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior drafts, and correcting them at all stages have made it possible to complete this project.

We would like to express our heartiest gratitude to **Prof. Dr. Touhid Bhuiyan** and Head, Department of CSE, for his kind help to finish our project and also to other faculty members and the staff of the CSE department of Daffodil International University.

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

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

## **ABSTRACT**

Students are the backbone of a nation. But at present there is a very high rate of depression among the students. As a result, many students are suffering from depression and many are choosing the path of suicide. Our research focuses on finding out the best way to find out the causes of their depression and how to keep them free from depression. To this end, we have so far collected 800 data from high school, college and university students and through data leveling we have been able to determine the level of depression among them. And based on this we will take various steps to alleviate the frustration of the students later.

# TABLE OF CONTENTS

<b>CONTENTS</b>	<b>PAGE</b>
Board of examiners	i
Declaration	ii
Acknowledgements	iii
Abstract	iv
Table of contents	v
List of figures	vi
List of tables	vii
<b>CHAPTER 1: INTRODUCTION</b>	<b>1-3</b>
1.1 Introduction	1
1.2 Motivation	1-2
1.3 Rationale of the Study	2
1.4 Research Questions	2
1.5 Expected Outcome	3
1.6 Project Management and Finance	3
1.7 Report Layout	3
<b>CHAPTER 2: BACKGROUND</b>	<b>4-7</b>
2.1 Introduction	4
2.2 Related works	4-6
2.3 Computer analysis and summary	6
2.4 Challenges	7
<b>CHAPTER 3: RESEARCH METHODOLOGY</b>	<b>8-23</b>
3.1 Introduction	8
3.2 Research Subject and Instrumentation	8
3.3 Data Collection Procedure/Dataset Utilized	8-10
3.4 Statistical Analysis	10-14
3.5 Proposed Methodology/Applied Mechanism	15-21
3.6 Implementation Requirements	22-23



<b>CHAPTER 4: EXPERIMENTAL RESULT AND DISCUSSION</b>	<b>24-26</b>
4.1 Experimental Results and Analysis	24-26
4.2 Discussion	26
<b>CHAPTER 5: IMPACT ON SOCIETY</b>	<b>27-28</b>
5.1 Impact on Society	27
5.2 Impact on Environment	27-28
5.3 Ethical Aspects	28
5.4 Sustainability Plan	28
<b>CHAPTER 6: SUMMARY, CONCLUSION, IMPLICATION FOR FUTURE STUDY</b>	<b>29</b>
6.1 Summary of the study	29
6.2 Conclusion	29
6.3 Implication for future study	29
<b>REFERENCES</b>	<b>29-31</b>

## **LIST OF FIGURES**

<b>FIGURES</b>	<b>PAGE NO</b>
Figure3.3.1: Excel file of dataset	9
Figure3.3.2: Numeric Dataset	10
Figure 3.5.1: Selected Machine Algorithm Category	15
Figure 3.5.2: Logistic regression equation	16
Figure 3.6.1: Working procedure using data mining technique	22

## LIST OF TABLES

TABLE	PAGE NO
Table 2.3.1: Comparison of our work and other's work	6
Table 3.4.1. Attribute with their possible values	11-13
Table 3.4.2: Characteristics and Value of all data	14
Table 3.5.1: Performance of Logistic regression	16
Table 3.5.2: Performance of Simple Logistic	17
Table 3.5.3: Performance of SMO	18
Table 3.5.4: Performance of LMT	18
Table 3.5.5: Performance of IBK	19
Table 3.5.6: Performance of J48	20
Table 3.5.7: Performance of k star	21
Table 3.5.8: Performance of Classification Via Regression	21
Table4.1.1: Confusion matrix of all classifier	25
Table 4.1.2: Average Accuracy of all classifier	26

# **CHAPTER 1**

## **INTRODUCTION**

### **1.1 Introduction**

The word “Depression” represents a psychological illness or disorder which is very common nowadays. This declares the negative psychological state of a person; Generally, Depression happens from some common sort of reasons such as frustration, anxiety, social & family problems, medical issues, academic issues, etc. Depression usually affects the daily life of a person negatively, the normal way of a person living gets hampered due to his/her mental feelings or psychological state. This generally hampers a person’s progress in life and daily work. They can not act or concentrate properly as their mental state doesn’t allow them to. A psychological issue like Depression is found common among the students also. This causes the students to not cope up with their academic studies, behavior as well as healthy life; as a result, getting lower grades and apathy in classes, and less communication in the family. Students going through depression are more likely to get less satisfaction with health-related quality of life (HRQOL) than those with no depression symptoms. Depression among students worldwide is now a very widespread problem and very alarming. University students are a unique community of people who are in the midst of a transitional time in which they are attempting to integrate into their new college life as well as the new social world in which they find themselves. During this period, depression is more often seen among some students. Many of them have previous anxiety issues, some have different psychological issues such as communication problems, this is a more commonly found issue that leads most of the students to depression. In our country, there are very few studies found regarding depression among students. We have experimented with eight common data mining classifiers, including Logistic, simple logistic, SMO, IBK, LMT, J48, K-star, and Classification Via Regression, in this report to elaborate on the data we found from the survey we conducted among 800 Bangladeshi students.

## **1.2 Motivation**

Bangladesh has a total of 9913884 students. Students are the backbone of a nation. Students are educated and develop the country, representing the country to the outside world. But if these students drop out prematurely, it is a threat to any country. Youth in Bangladesh face challenges with their studies, careers, relationships, drugs, and family issues, all of which are major or minor contributors to depression. So the motive of our project is to find out the prevalence of depressive symptoms, their associated factors and help to reduce depression among the students of the school, college, and university.

In addition, there are several clear motivations in our research:

- 1) To find out how common depressive symptoms are.
- 2) To determine the prevalence of depression among students.
- 3) To figure out the most effective way to alleviate their depression.

## **1.3 Rationale of the Study**

There has been a lot of research performed on student depression levels. Bangladeshi students, on the other hand, have a much smaller depression research area. In most cases, research has been done on medical students or university students. There is very little study on students in high school and college. But now, for different reasons, the level of depression among all of these tenderhearted students is rising, which is a threat to a country as well as the nation. In our project, we brought together high school, college, and university students to conduct research and use machine learning to alleviate their depression.

## **1.4 Research Questions**

1. Does it predict an actual output by given sample data with your system?
2. Does every algorithm work perfectly (yes/no)?
3. Is depression leveling properly finding out for all students?

We leveled data on 800 students and divided them into three levels and through this, we got an idea about the mental state of the students.

## **1.5 Expected Outcome**

The aim of this study is to determine the level of depression among the students and to provide some solutions to reduce their depression.

## **1.6 Project Management and Finance**

So far, no financial assistance has been needed for this project. However, we are very hopeful about the future of this project.

## **1.7 Report Layout**

We've covered the introduction to depression level detection, inspiration, the study's reasoning, and the thesis's conclusion in this chapter. The report layout is then followed.

In chapter 2, the history of our research subject will be discussed.

In chapter 3, we will go over the research methods we used in our report.

In chapter 4, the Experimental Results and Discussion will be discussed.

In chapter 5, we will discuss the Impact on Society, Environment, and Sustainability.

In chapter 6, we will talk about the conclusion and what we'll do next.

## **CHAPTER 2**

### **BACKGROUND**

#### **2.1 Introduction**

The use of data mining to investigate scientific questions inside educational exploration has grown in popularity in recent decades. In this chapter, we'll talk about data mining as it relates to data classifier research or thesis. Previous related work will be discussed in the first section, the nature of the problems of our study related work will be seen in the second section, and finally, the difficulties we faced in this study will be discussed

#### **2.2 Related Works**

Much work has been done and a lot of research has been done on the effects of depression on students. But in most cases, the research has been done on the students of the university. But we are researching students in schools, colleges, and universities. Since students are the future of the nation, a lot of research has been done so far thinking about them.

The writers of [1] focused on first-year students. They gathered 400 pieces of information during their research. They used a variety of classifiers, but logistic regression proved to be the most efficient. They used 14 variables in their questionnaires and grouped them into four parts. They have 52.3 percent male and 47.8% female candidates, based on their study. They used four classes and found that 50.2 percent of students had mild depression.

The author of [2] conducted a study on stress and depression among Dhaka medical college students. A total of 425 data records were used in this project. In this report, students were split into two groups. The dataset is not listed here, but the results showed that a large percentage of people (66%) recorded no or minor stress, while a lower rate (34%) reported moderate to heavy pressure.

The authors of [3] questioned 523 students, but only 468 had their data examined. The dataset used in this study is not listed, but it included socio-demographic questions and several psychometric instruments. The final results showed that 44 percent of the participants were depressed, while 27.1 percent were nervous. The average overall score is 8.01.

In [4], authors research on mental health problems among Bangladeshi students. The total data they collected is 635, but 590 were retained for final analysis. They use 11 different kinds of questions. In their project, they used the Bangla version of DASS-21. The Cronbach's alphas for the depression was 0.77. The size of the Feature Set of their study is 7. Their best method is logistics regression, and the accuracy is 95%.

The authors of [5] conducted research on college freshmen. They used the (GAIN-SS) to evaluate mental health issues. The correlations of 12-month mental health conditions were estimated using generalized linear modeling. Finally, they discovered that mutually exclusive forms of mental health issues were even less common, with internalizing problems accounting for 14.2 percent (SE=0.56), externalizing problems for 8.6 percent (SE=0.46), drug use for 1.7 percent (SE=0.21), and antisocial problems accounting for 0 percent .

The writers of [6] conducted studies on alexithymia. The sample size was 492 and the feature set was 14. They used version 21 of the (SPSS) software to analyze the results. They used a variety of methods to assess depression, anxiety, and stress. They discovered that alexithymia affects 24.6 percent of the participants. Depression, fear, and stress were all present in 28.5, 38.4, and 22.6 percent of the population, respectively.

The writers of [7] focused on depression. Their sample size was 237, and they used stratified random sampling. Their project's Feature Set is ten features long. Depressive and serious depressive disorder was found to be present in 21.5 percent and 7.6 percent of medical students, respectively.

The writers of [8] focused on depression in college students. They created a standard psychometric instrument in the form of an online questionnaire scheme. They were able to extract 121 features. Their method was used to assess the data of 466 students. They discovered that 25.32 percent of students suffer from depression.

The writers of [9] looked at depression among students at a Nigerian university. Their data had a sample size of 139 people. The precision was 58.2 percent, with mild, moderate, moderately severe, and serious depression being diagnosed in 37.0 percent, 15.7 percent, 3.9 percent, and 1.6 percent, respectively.



## 2.3 Comparative Analysis and Summary

Depression leveling research has been done more with college and university students. But as far as we know, not much research has been done on the students of the school. We're also leveling depression with the help of our project's high school, college, and university students. The relation of our work to other existing work is shown in Table 2.3.1.

Method/Work Done	Object(s) Deal with	Sample Size	Technique used	Size of Feature Set	Classifier	Best Classifier	Accuracy
Our initiatives	Bangladeshi Students School ,college and university Students	800	Machine Learning	29	SMO	SMO	98.94%
Mamun et al. [4]	JU University Students	590	Machine Learning	11	Logistic Regression	Logistic Regression	95%
Islam et al. [1]	First year of JU University Students	400	Machine Learning	14	Logistic Regression	Logistic Regression	95%
Dr. Nazma Sultana [2]	Dhaka Medical College Students	425	Machine Learning	10	NM	NM	NM
Mamun et al. [3]	Pre-university students of JU	468	NM	NM	NM	NM	NM

Table 2.3.1: Comparison of our work and other's work

## **2.4 Challenges**

Collecting data was not an easy task for us at all. Our entire research has been based on this data. So we have been very careful about data collection and have been able to collect 800 data perfectly. The challenges we faced:

### **1. Data collection**

Data collection was the most difficult task because when we took the initiative of data collection the lockdown started for covid-19. So we had to rely online on to collect data through a very difficult google form. Many students did not want to give information because they thought their data would be leaked.

### **2. Leveling**

We collected about 800 data which was a really big challenge for us to level. We have to mark each question based on the answer and find out the total mark. Thus we calculated the total marks of the students for 800 and determined their depression level based on the marks.

## **CHAPTER 3**

### **RESEARCH METHODOLOGY**

#### **3.1 Introduction**

This study describes the procedure of this survey and makes out many strategies that describe which one (study in-home or hostel) is preferred by students. The research draft, Research Topic, and Instrumentation, Data Collection Procedure, and Statistical Analysis will all be examined. This discussion will cover the research project's apparatus, data collection, research subject, pre-processing, processing, and its application. We will use WEKA for data collection, deployment, and other tasks in order to overcome real-world data mining challenges.

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

The most important aspect of any research paper is gathering data. Information is a crucial component of any research project. Finding the ideal data and algorithm for a research project is extremely difficult. We need to read more papers that are relevant to our work. Then we arrive at conclusions such as these:

1. What information would be required?
2. How can you be sure that the information you've gathered is correct?
3. How can we organize the data for our project?
4. What is the best way to mark each piece of information?

#### **3.3 Data Collection Procedure**

For our project, we have surveyed some questions in which we included age, gender, educational level, etc. Mainly we followed 4 procedures for collecting data. The procedure is:

## 1) Data pre-processing

At first, we collected some raw data and then we process the data for our implementation.

This is the most important part of the research. We collected 800 data from students via a google doc file. Then we convert the CSV file into an excel file.

	A	B	C	D	E	F	
1	Timestamp	আপনার বয়স? (Age?)	আপনার লিঙ্গ? (Gender?)	শিক্ষার স্তর? (Educational Level?)	আমার অসুস্থি লাগে? (I feel trouble)	ইনসানি আমি মননয় থাকি (Nowadays I feel misery)	আমার অসুস্থি স্তর?
2	2020/10/09 8:48:14 PM GMT+6	20	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	মঝামঝি (Moderately)	প্রযোজ্য নয় (Not Applicable)	একোবারেই প্রযোজ্য নয়
3	2020/10/09 8:51:32 PM GMT+6	20	নারী (Female)	বিশ্ববিদ্যালয় (University)	মঝামঝি (Moderately)	মঝামঝি (Moderately)	প্রযোজ্য নয় (Not A
4	2020/10/09 8:57:55 PM GMT+6	20	নারী (Female)	বিশ্ববিদ্যালয় (University)	মঝামঝি (Moderately)	মঝামঝি (Moderately)	প্রযোজ্য নয় (Not A
5	2020/10/09 9:02:27 PM GMT+6	20	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	মঝামঝি (Moderately)	প্রযোজ্য নয় (Not Applicable)	একোবারেই প্রযোজ্য নয়
6	2020/10/09 9:32:52 PM GMT+6	22	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	কিছুটা প্রযোজ্য (A bit applicable)	প্রযোজ্য নয় (Not Applicable)	পূরোপুরি প্রযোজ্য (C
7	2020/10/09 9:33:21 PM GMT+6	23	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	মঝামঝি (Moderately)	মঝামঝি (Moderately)	একোবারেই প্রযোজ্য নয়
8	2020/10/09 9:33:30 PM GMT+6	22	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	কিছুটা প্রযোজ্য (A bit applicable)	মঝামঝি (Moderately)	একোবারেই প্রযোজ্য নয়
9	2020/10/09 9:33:33 PM GMT+6	20	নারী (Female)	বিশ্ববিদ্যালয় (University)	পূরোপুরি প্রযোজ্য (Completely applicable)	পূরোপুরি প্রযোজ্য (Completely applicable)	মঝামঝি (Moder
10	2020/10/09 9:33:33 PM GMT+6	22	নারী (Female)	বিশ্ববিদ্যালয় (University)	মঝামঝি (Moderately)	মঝামঝি (Moderately)	প্রযোজ্য নয় (Not A
11	2020/10/09 9:34:00 PM GMT+6	20	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	প্রযোজ্য নয় (Not Applicable)	কিছুটা প্রযোজ্য (A bit applicable)	মঝামঝি (Moder
12	2020/10/09 9:34:03 PM GMT+6	22	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	প্রযোজ্য নয় (Not Applicable)	মঝামঝি (Moderately)	প্রযোজ্য নয় (Not A
13	2020/10/09 9:34:07 PM GMT+6	23	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	একোবারেই প্রযোজ্য নয় (Totally not applicable)	একোবারেই প্রযোজ্য নয় (Totally not applicable)	একোবারেই প্রযোজ্য নয়
14	2020/10/09 9:34:17 PM GMT+6	20	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	কিছুটা প্রযোজ্য (A bit applicable)	পূরোপুরি প্রযোজ্য (Completely applicable)	কিছুটা প্রযোজ্য (A b
15	2020/10/09 9:34:25 PM GMT+6	22	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	মঝামঝি (Moderately)	কিছুটা প্রযোজ্য (A bit applicable)	একোবারেই প্রযোজ্য নয়
16	2020/10/09 9:34:27 PM GMT+6	23	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	একোবারেই প্রযোজ্য নয় (Totally not applicable)	একোবারেই প্রযোজ্য নয় (Totally not applicable)	একোবারেই প্রযোজ্য নয়
17	2020/10/09 9:34:31 PM GMT+6	21	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	কিছুটা প্রযোজ্য (A bit applicable)	কিছুটা প্রযোজ্য (A bit applicable)	পূরোপুরি প্রযোজ্য (C
18	2020/10/09 9:34:43 PM GMT+6	22	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	কিছুটা প্রযোজ্য (A bit applicable)	প্রযোজ্য নয় (Not Applicable)	প্রযোজ্য নয় (Not A
19	2020/10/09 9:34:45 PM GMT+6	22	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	কিছুটা প্রযোজ্য (A bit applicable)	কিছুটা প্রযোজ্য (A bit applicable)	প্রযোজ্য নয় (Not A
20	2020/10/09 9:34:52 PM GMT+6	21	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	মঝামঝি (Moderately)	মঝামঝি (Moderately)	মঝামঝি (Moder
21	2020/10/09 9:35:04 PM GMT+6	21	নারী (Female)	বিশ্ববিদ্যালয় (University)	পূরোপুরি প্রযোজ্য (Completely applicable)	পূরোপুরি প্রযোজ্য (Completely applicable)	প্রযোজ্য নয় (Not A
22	2020/10/09 9:35:14 PM GMT+6	22	পুরুষ (Male)	বিশ্ববিদ্যালয় (University)	মঝামঝি (Moderately)	মঝামঝি (Moderately)	কিছুটা প্রযোজ্য (A b

Figure3.3.1: Excel file of dataset

## 2) Data Organizing

We used a Microsoft Excel file to organize the information.

## 3) Laveling Data

In this part at first, we count the number of each question of a student then we got a total number. After getting the total number we declare the depression level of the students by using the labeling method, we have 4 levels minimum, mild, moderate, and severe.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	IT	NM	IF	MF	IE	NH	IH	IA	IM	ME	EM	LM	IC	BM	DA	CD	CA	CB	WT	SL	SM
2	2	2	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0	2	1	0	0
3	2	1	0	0	0	4	0	0	0	0	0	0	0	1	1	0	0	1	1	0	3
4	3	1	4	3	2	4	3	0	2	3	3	2	3	2	3	3	2	3	2	1	
5	2	2	0	0	0	1	2	2	2	1	2	0	1	2	3	2	2	2	2	1	
6	3	2	0	1	0	3	1	2	1	3	1	1	1	2	1	1	2	2	2	4	
7	4	4	2	3	1	4	4	2	4	4	4	3	4	4	4	4	4	4	3	1	
8	2	2	1	1	1	2	1	2	2	2	0	1	4	4	3	4	4	4	4	4	
9	1	3	2	1	1	2	1	1	0	0	1	1	0	2	1	1	1	2	2	0	
10	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	1	1	
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
12	3	4	3	2	2	2	2	2	3	4	0	3	1	4	4	4	4	4	4	1	
13	2	3	0	0	0	0	0	0	0	0	0	0	3	3	0	3	4	4	4		
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
15	3	3	4	4	4	4	4	1	4	4	4	4	1	1	2	2	2	2	2	2	
16	3	1	1	1	1	2	1	2	1	1	1	0	1	1	1	1	1	1	1	2	
17	3	3	1	2	1	3	1	2	1	2	1	0	0	3	3	1	2	3	4	0	
18	2	2	2	1	0	2	0	0	2	2	2	0	0	2	0	1	0	2	3	1	
19	4	4	1	0	1	2	3	2	2	2	0	1	2	3	2	2	2	1	2	0	
20	2	2	3	1	0	2	3	2	2	1	1	0	0	1	1	2	3	3	1	3	
21	2	3	3	4	3	4	4	2	1	3	1	1	4	4	3	4	4	4	1	1	
22	3	3	2	0	3	3	3	3	2	3	1	2	2	3	2	3	3	3	2	3	

Figure3.3.2: Numeric Dataset

#### 4) Data Storing

To ensure the data is not lost, we have stored it in Google Drive as Google Sheets. We've also saved it as a CSV format. We'll be able to use them for coding and testing models later.

### 3.4 Statistical Analysis

A survey was used to collect the data for this study. The survey consists of 29 questions in total. Personal and academic considerations are also included in these 29 questions. Here, 29 variables are used to try to mark all of the questions. The number of independent variables in the dataset is 29, so the number of dependent variables is 1. Table 3.4.1 lists all of the variables and their potential values. This work necessitates the use of over 800 individual documents. A total of 80% of information is used in the preparation of the classifier, with the remaining 20% being used for research. We gathered information from 800 students. We graded depression and divided it into four categories using 29 questions. We have gathered some basic information for the students in addition to the 29 questions, which is shown in Table 3.4.2.

Variable	Full-form	Variable type	Possible value
IFT	I feel trouble	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
NIFM	Nowadays I feel misery	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
IFHNF	I feel like I have no future	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
MCIWF	My condition is getting worse day by day in the future	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)

ILE	I have lost everything	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
NLIH	I think, nowadays life is very hard	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
IFH	Nowadays I think, I am a complete failure as a human	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
DFHA	I don't find any happiness anywhere	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
FLM	I feel very low for myself	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
MCLE	My confidence has become low in everything	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)

EFSM	I feel everyone feel sympathy for me	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
LISM	Life is meaningless	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
IFC	Mostly I feel cry	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
IFBM	I feel bored mostly	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)

DFEA	I don't feel the eagerness in anything	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
CNTMD	Nowadays, I can not think and make any decision	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
CNPAA	Nowadays I can not pay attention to anything	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
CNKTB	I can not keep things in my mind like before	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
WATM	I feel weak and tired mostly	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
ISLN	I sleep less now	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)

ISMN	I sleep more now	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
IFIM	I feel irritable mostly	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
IFLH	I feel less hungry now	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
IFMH	I feel more hungry now	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
ILW	I have lost weight [not to lose weight intentionally]	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
PWSD	I feel like the pace of my work has slowed down	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
CNLA	I can't laugh anymore even if something funny happens	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
CNPSW	I can't participate in social work like before	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
FEPW	Nowadays, it has become difficult to focus on education and professional work as before	Independent	Totally not applicable(0), Not Applicable(1), Moderately(2), A bit applicable(3), Completely applicable(4)
DL	Depression Level	Dependent	No depression(0-30), Minimum(30-100),Mild(101-114),Moderate(115-123)

Table 3.4.1. Attribute with their possible values



## Characteristics and Value

Gender	
Female	294
Male	506
Age in years	
17	30
18	30
19	43
20	120
21	135
22	160
23	133
24	63
25	38
26	16
27	6
28	4
29	2
30	2
31	1
Educational Status	
School	41
College	64
University	695

Table 3.4.2: Characteristics and Value of all data

### 3.5 Proposed Methodology

Weka has a significant portion of classification algorithms. One of the benefits of using the Weka platform to solve our machine learning problems is the large number of machine learning algorithms available. Classification is a popular data processing activity that has its origins in machine learning. Data is divided into two or more categories, or classes, in the "Classification Issue." The method of creating a model of classes from a collection of records containing class levels is known as classification. The aim of the Decision Tree Algorithm is to figure out how the attributes-vector behaves in a variety of situations. The classes for the newly created instances are also determined using the training instances. The principles for predicting the target variable are created by this algorithm. The essential distribution of data can be understood with the aid of a tree classification algorithm.

We worked on about 11 classifications at Weka and came up with the following eight best classifications. IBK, J48, K-star, Classification Via Regression, Logistic, Simple Logistic, SMO, LMT (Logistic Model Tree). Our classification is much more accurate than the others.

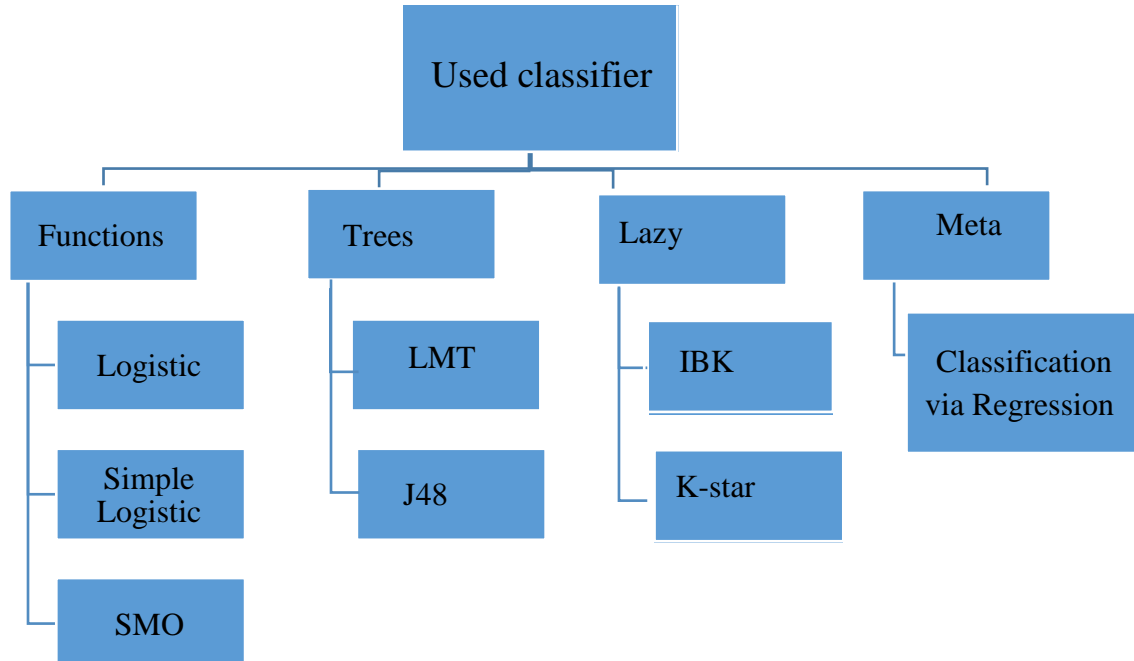


Fig 3.5.1: Selected Machine Algorithm Category

## Logistic

For binary classification, logistic regression is one of the most widely used machine learning algorithms. This is due to the fact that it is a simple algorithm that works well on a wide variety of problems. The dependent variable in logistic regression is binary, and the independent variables are frequently continuous or binary. A categorical variable's performance is predicted using logistic regression. As an outcome, the result must be a discrete or categorical value. Let's look at the logistic equation, which is at the heart of the logistic regression methodology. By estimating probabilities using the logistic equation, logistic regression may also calculate the relationship between a single dependent variable and one or more independent variables [10] [11]. This is how the logistic function is defined:

$$p = \frac{1}{1 + e^{-(b_0 + b_1 X_1 + b_2 X_2 + \dots + b_k X_k)}}$$

Figure 3.5.2: Logistic regression equation

	Accuracy	Error rate	Sensitivity	Specificity	Precision	FPR	FNR
No depression	98.00%	2.00%	95.58%	98.40%	90.75%	1.60%	4.42%
min	94.48%	5.52%	94.48%	94.48%	98.71%	5.52%	5.52%
mild	97.36%	2.64%	86.67%	97.78%	60.46%	2.22%	13.33%
moderate	98.87%	1.13%	100%	98.86%	18.18%	1.14%	0.00%

Table 3.5.1: Performance of Logistic regression

## Simple Logistic

When you have one nominal variable with two values and one measurement variable, you can use Simple Logistic [12]. The nominal variable is the experimental variable, while the measurement variable is the nominal variable. Simple logistic regression is similar to rectilinear regression, except the dependent variable is a nominal value rather than a test. One aim is to see how the probability of getting a specific nominal variable value is related to the measurement variable; the other goal is to estimate the probability of getting a specific nominal variable value provided the measurement variable.

	Accuracy	Error rate	Sensitivity	Specificity	Precision	FPR	FNR
No depression	98.36%	1.64%	94.69%	98.98%	93.85%	1.02%	5.31%
min	96.73%	3.27%	97.69%	92.42%	98.30%	7.58%	2.31%
mild	98.36%	1.64%	80.00%	99.08%	78.00%	0.92%	20.00%
moderate	99.88%	0.12%	100%	99.75%	66.67%	0.25%	0%

Table 3.5.2: Performance of Simple Logistic

## SMO

Sequential Minimal Optimization (SMO) [13] is a simple algorithm that solves the Support-Vector Machine Quadratic Programming problem easily and without the need for any additional matrix storage or numerical QP optimization steps. The SMO algorithm is used to decompose a multi-variable optimization problem into a series of sub-problems, each of which optimizes an objective function of a small number of variables, usually only one, while all other variables are treated as constants that remain unchanged within the sub-problem. The coordinate descent algorithm, for example, is a decomposition method that solves a drag in a multi-dimensional space by breaking it down into a series of sub-problems, each of which is solved in a one-dimensional space.

	Accuracy	Error rate	Sensitivity	Specificity	Precision	FPR	FNR
No depression	99.00%	1.00%	94.69%	99.71%	98.16%	0.29%	5.31%
min	98.12%	1.88%	99.24%	93.11%	98.47%	6.89%	0.76%
mild	98.88%	1.12%	86.66%	99.35%	8.38%	0.65%	13.34%
moderate	99.75%	0.25%	0%	100%	0%	0%	100%

Table 3.5.3: Performance of SMO

## LMT

LMT (Linear Model Tree) [14] is a classification model with a supervised training algorithm that combines rectilinear regression and decision trees. LMTs are based on the earlier concept of a model tree: a decision tree with linear regression models at its leaves that produces a piecewise rectilinear regression model, while ordinary decision trees with constants generate a piecewise-constant model. The LogitBoost algorithm is used in the logistic version to provide an LR model at each node in the tree, which is then split using the C4.5 principle. Each Logit Boost invocation is warmed up by the parent node's leads. The tree is finally trimmed.

	Accuracy	Error rate	Sensitivity	Specificity	Precision	FPR	FNR
No depression	98.37%	1.63%	94.69%	98.98%	93.86%	1.02%	5.31%
min	97.08%	2.92%	97.70%	84.38%	99.22%	15.62%	2.30%
mild	98.37%	1.63%	80.00%	99.08%	77.42%	0.92%	20.00%
moderate	99.75%	0.25%	100%	99.75%	50.00%	0.25%	0.00%

Table 3.5.4: Performance of LMT

## IBK

Instance-based learning [15] refers to systems that learn machine learning training examples from memory and then generalize to new instances using a similarity metric. It's called instance-based because the hypotheses are built from the training data. ALFA, Niagara Mohawk Power Service of Central NY State, and Clark's (1989) method of geologic prospect appraisal are some of the specialized Domains that had to use IBK to produce classification predictions using only unique instances. In Machine Learning, K-Nearest Neighbors (KNN) algorithms are used to solve regression and classification problems. Weka calls the KNN classifier IBK, and its algorithms use data to identify new data points based on similarity measures.

	Accuracy	Error rate	Sensitivity	Specificity	Precision	FPR	FNR
No depression	96.86%	3.14%	89.38%	98.09%	88.59%	1.90%	10.62%
min	93.97%	6.03%	95.09%	88.97%	97.48%	11.03%	4.91%
mild	96.86%	3.14%	80.00%	97.52%	55.81%	2.48%	20.00%
moderate	99.75%	0.25%	100%	99.75%	50.00%	0.25%	0.00%

Table 3.5.5: Performance of IBK

## J48

J48 is an ID3 extension. Accounting for missing values, decision tree pruning, continuous attribute value ranges, rule derivation, and other features of J48 [16] are included. J48 is an open-source Java implementation of the C4.5 algorithm used in the WEKA data processing tool. The WEKA tool offers a number of choices for tree pruning. Pruning is often used as a method for préising only in case of possible over-fitting. Other algorithms perform recursive classification until every single leaf is pure, implying that the information classification should be as perfect as possible. The algorithm considers all possible tests for splitting the data set and chooses the one with the simplest knowledge gain. Since the number of distinct values of the attribute is taken into account, one test with as many outcomes is performed for each discrete attribute. Binary tests concerning each distinct

value of the attributes are considered for each continuous attribute. The training data set belonging to the node in considerations is sorted for the values of the continuous attribute, and thus the entropy gains of the binary cut supported each distinct value are measured in one scan of the sorted data to collect the entropy gain of these binary tests competently. For each continuous attribute, the process is repeated.

	Accuracy	Error rate	Sensitivity	Specificity	Precision	FPR	FNR
No depression	92.97%	7.03%	76.11%	95.76%	74.78%	4.24%	23.89%
min	89.84%	10.16%	93.56%	73.10%	93.99%	26.90%	6.44%
mild	95.89%	4.11%	48.57%	98.04%	53.13%	1.96%	51.43%
moderate	99.62%	0.38%	0.00%	99.88%	0.00%	0.12%	100.00%

Table 3.5.6: Performance of J48

The Correctly Classified Instances in our data interpreted is: 713  $\rightarrow$  89.4605 %

### **K-star**

K\* is an instance-based classifier [17], which means that the category of a test instance is based on the category of training instances that are nearly identical to it, as defined by a similarity function. The K-Nearest-Neighbor classifier is equivalent to K\*. The K-star algorithm employs an entropic metric, which is based on the probability of remodeling one instance into another by selecting random transformations from all possible options. Using entropy as a meter for an instance distance is very beneficial and information theory helps in computing the space between the instances. The complexity of a change of one instance into another is the space between instances.

	Accuracy	Error rate	Sensitivity	Specificity	Precision	FPR	FNR
No depression	96.74%	3.26%	88.50%	98.10%	88.50%	1.90%	11.50%
min	93.60%	6.40%	94.78%	88.28%	97.32%	11.72%	5.21%
mild	96.61%	3.39%	80.00%	97.26%	53.33%	2.74%	20.00%
moderate	99.75%	0.25%	100%	99.75%	50.00%	0.25%	0.00%

Table 3.5.7: Performance of k star

## Classification Via Regression

When only two suitable models/techniques, namely logistic regression and rectilinear regression, are used, classification is often done by regression. However, before we get into ClassificationViaRegression, it's crucial to understand the distinction between regression and classification models. Regression models may predict a variety of variables, such as the amount of rainfall or the intensity of sunlight. They can also predict probabilities, such as the likelihood of a dog appearing in a frame. By applying a preference rule - for example, if the likelihood is 50% or higher, decide it's a dog - a probability-predicting regression model can be used as part of a classifier. Logistic Regression is a regression algorithm that predicts probabilities. In the machine learning literature, it's generally referred to as a classification process and Linear Regression, which is the most basic form of regression. It's a method in which the variable remains constant throughout.

	Accuracy	Error rate	Sensitivity	Specificity	Precision	FPR	FNR
No depression	96.50%	3.50%	91.25%	97.25%	83.20%	2.75%	8.75%
min	94.35%	5.65%	98.16%	77.24%	95.10%	22.76%	1.84%
mild	97.62%	2.38%	53.33%	99.35%	76.19%	0.65%	46.65%
moderate	99.75%	0.25%	0.00%	100%	0.00%	0.00%	100%

Table 3.5.8: Performance of ClassificationViaRegression



### 3.6 Implementation Procedure

The primary goal of this project is to determine the extent of depression among students.

We attempted to determine the number of students in our country who are depressed. We took several measures to complete this project, which are outlined below.

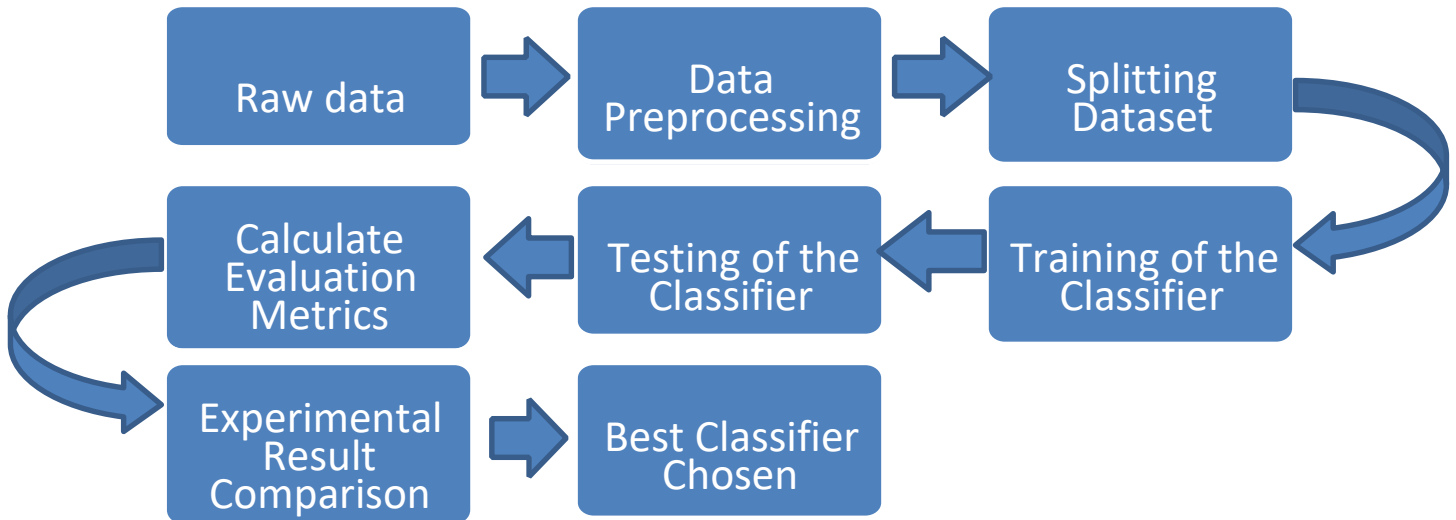


Figure 3.6.1: Working procedure using data mining technique.

First and foremost, we've built a 29-question questionnaire form. Then, using this questionnaire, we gathered information from a large number of people. Then, in order to feed this data into the classifier, we used some preprocessing techniques. One query is leveled using only one element. As a result, a total of 29 variables are used to level all of the questions. The numbers 0–4 are then used to level the answer to the actual query. For example, the variable/attribute “I feel irritable mostly(FI)” in our dataset has five possible outcomes: Totally not applicable(0), Not Applicable(1), Moderately(2), a touch applicable(3), and Completely applicable(4) (4). After preprocessing, our prepared data is partitioned into the training and testing set. For training purposes, 80 percent of the data from the entire data collection was used. The remaining 20% of the total data set was used for testing. This division is done at random. Following that, the classifiers were trained using the training data. We used test data to predict current professional status after training the classifiers. We've estimated a few of the performance assessment metrics right here. We found the best classifier to estimate during this context using these metrics. Using Eqs. (1–7), several performance metrics in percentage are calculated based on the confusion matrix created by the classifier.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (1)$$

$$Error\ Rate = \frac{FP+FN}{TP+TN+FP+FN} \times 100\% \quad (2)$$

$$Sensitivity = \frac{TP}{TP+FN} \times 100\% \quad (3)$$

$$Specificity = \frac{TN}{FP+TN} \times 100\% \quad (4)$$

$$Percision = \frac{TP}{TP+FP} \times 100\% \quad (5)$$

$$FPR = \frac{FP}{FP+TN} \times 100\% \quad (6)$$

$$FNR = \frac{FN}{TP+FN} \times 100\% \quad (7)$$

## Chapter 4

### Experimental Results and Discussion

#### 4.1 Experimental Results & Analysis

The classifier generated a 4\*4 confusion matrix because this is primarily a multiclass problem. Table 4.1.1 shows the generated matrix for each of the classifiers. Accuracy, Error rate, Sensitivity, Specificity, FPR, FNR, and Precision are computed from the above confusion matrix to test this job. Tables 3.5.1 to 3.5.8 show the results of many performance assessment measurements. Person precision, error rate, FPR, FNR, and other values for each class can be found in tables 3.5.1 to 3.5.8. The classifier's accuracy for no depression, minor depression, mild depression, and severe depression is 98.99 percent, 98.11 percent, 98.37 percent, and 99.87 percent, respectively, which is the highest of all classifiers. Our data were also 99.73 percent accurate, with 14.03 percent no depression, 81.70 percent minimal depression, 3.75 percent mild depression, and 0.25 percent moderate depression.

## Confusion Matrix

Model	Matrix						Model	Matrix					
Logistic	Actual	Predicted					Simple Logistic	Actual	Predicted				
			A	B	C	D				A	B	C	D
		A	108	5	0	0			A	107	6	0	0
		B	11	616	17	8			B	7	637	7	1
		C	0	3	26	1			C	0	5	24	1
		D	0	0	0	2			D	0	0	0	2
Model	Matrix						Model	Matrix					
SMO	Actual	Predicted					LMT	Actual	Predicted				
			A	B	C	D				A	B	C	D
		A	107	6	0	0			A	107	6	0	0
		B	2	647	3	0			B	7	637	7	1
		C	0	4	26	0			C	0	5	24	1
		D	0	0	2	0			D	0	0	0	2
Model	Matrix						Model	Matrix					
IBK	Actual	Predicted					J48	Actual	Predicted				
			A	B	C	D				A	B	C	D
		A	101	12	0	0			A	86	27	0	0
		B	13	620	19	0			B	29	610	13	0
		C	0	4	24	2			C	0	12	17	1
		D	0	0	0	2			D	0	0	2	0
Model	Matrix						Model	Matrix					
Kstar	Actual	Predicted					Classification Via Regression	Actual	Predicted				
			A	B	C	D				A	B	C	D
		A	100	13	0	0			A	94	19	0	0
		B	13	618	21	0			B	9	640	3	0
		C	0	4	24	2			C	0	14	16	0
		D	0	0	0	2			D	0	0	2	0

Table 4.1.1: Confusion matrix of all classifier

## Average Accuracy

Classifier	Accuracy	Errorrate	Sensitivity	Specificity	Precision	FPR	FNR
Logistic	97.18%	02.82%	94.18%	97.38%	67.02%	02.62%	05.82%
Simple Logistic	98.33%	01.67%	93.00%	97.56%	84.13%	02.44%	07.00%
SMO	98.94%	01.06%	70.00%	98.00%	51.00%	02.00%	30.00%
LMT	98.39 %	01.61%	93.10%	95.55%	80.12%	04.45%	06.90%
IBK	96.86%	03.14%	91.00%	96.08%	73.00%	03.92%	09.00%
J48	94.58%	05.42%	54.56%	91.70%	55.47%	08.30%	45.44%
K star	96.68%	03.32%	91.00%	95.85%	72.00%	04.15%	09.00%
Classification Via Regression	97.05 %	02.95%	60.69%	93.46%	63.62%	06.54%	39.31%

Table 4.1.2: Average Accuracy of all classifier

## 4.2 Discussion

We know that the level of depression among students has increased many times. After analyzing the data of 600 students, it is seen that the accuracy of the minimum level is the highest. This means that all these students have a minimum level of depression. And we got the lowest accuracy at the moderate level which means that only 0.25% of the students have moderate level depression.

## **Chapter-05**

### **Impact on Society, Environment and Sustainability**

#### **5.1 Impact on Society**

The most critical component of society is students. Common mental health issues are considered public health concerns, and they can lead to risky behaviors among students, such as suicide in extreme cases. Students' mental health problems have a negative impact on society. When a student suffers from an unexpected psychological problem, he can't focus on his studies. This unbearable disease impacts their normal regular life. The student becomes unconscious in every sphere of his life. When a student suffers from a psychological problem in his life he goes far away from his family, friends as well as society. Then, a student can't do his duty for society properly. Our survey on mental health issue among students shows that 81.70% of students among 800 students of Bangladesh including school, college, and the university is suffering from minimum level depression which is a negative sign for the future of those students as well as for society. If most of the students suffer from a mental health issue of society then the whole society will suffer from this problem and face an unexpected situation like suicide.

#### **5.2 Impact on Environment**

Nowadays depression, anxiety, stress has become very common words among students. Most of the students are suffering at least a minimum level of depression in their life. We did our survey among 800 students from different places in Bangladesh. From our data, we got that, at least about 99.73% of students are suffering from depression. Most students age 20 to 23 are suffering from depression. Teenage students are on the second level of suffering from mental health issues. The psychological problem of students impacts their family, classmates and friends. When a student suffers from a mental health issue, it impacts the environment of their family as well as their educational institutions and society. Some students do not have the psychological resilience to cope with the university world, which can become unbearable and bring unexpected difficulties (such as living apart from family and friends for the first time) (Saeed et al. 2018; Taneja et al. 2018).<sup>1</sup>

### **5.3 Ethical Aspects**

Based on the statistics of our survey and analysis, it can be clearly stated that the depression level among the students is very minimum, 81.70% of the students who participated in the survey have a minimum depression level which is very usual in the current social condition of the World. And a remarkable output came out in the moderate depression section which is 0.25% (lowest among the five choices). So, we can say that majority of the students are going through good mental health and such a condition is helping them to concentrate on their area of interest and having a good life. As we discovered, this type of study and survey should be undertaken on a regular basis to track progress in students' mental health so that appropriate measures can be taken to create a better future.

### **5.4 Sustainability Plan**

The students require a healthier environment in every aspect, in education, educational premises, playing areas, as well as at their home. This will help them to grow up with wider knowledge and good mental health. In many surveys, it is seen that students who get to study in institutions with a wide range of extra-curricular activities available in their premises grow up with much better knowledge and a wide range of skills. This helps a nation to grow strong and become much more stable. A good practice of behavior among the parents and their children helps the students most; they tend to do engage themselves in much more creative works and also helps in their study. Also, when parents give their children the freedom to explore their area of interest, the child or student can develop very fast and the success rate becomes very high.

## **Chapter-06**

### **Summary, Conclusion, Recommendation and Implication for Future Research**

#### **6.1 Summary of the study**

Our aim was to look into the prevalence of depression, anxiety, and stress among Bangladeshi students, as well as the related risk factors, and to provide them with a treatment plan to help them overcome this unsettling disease. Depression has become a very common mental health issue for people. Nowadays it has become an unbearable and unexpected mental health issue among the students of Bangladesh. This study was conducted to get a ratio of depressed students in Bangladesh. For the proper study, we surveyed the students of Bangladesh in different places. We got 800 data of school, college, and university students. From the survey, we found that 99.73% of students are suffering from depression. Among them, 81.70% of students are suffering from at least a minimum level of depression. Which is a negative sign for Bangladeshi students. Besides, we also got that, 14.03% of the students are consisting of no depression level. Which is also a positive sign of our study.

#### **6.2 Conclusions**

The aim of this project is to reduce the number of depressed students in Bangladesh. This is accomplished using a variety of data mining techniques. To accomplish this work, we collected 800 data of students and we got four levels of depressed students. In order to train and evaluate the classifier, the dataset was used. We calculated a significant amount of output confusion matrix to test the working classifier. We found accuracy among all the data through the classifier. We also found depressed students according to gender, age, and educational status.

#### **6.3 Implication for Further Study**

For the implication of further study, we will work for the solution to get rid of the MH issue among students. We will also work for an application by which a student can test their depression level and get a solution of it if he is in depression.



## References

- [1] Saiful Islam, Rejina Akter, Tajuddin Sikder, Mark D. Griffiths. "Prevalence and Factors Associated with Depression and Anxiety Among First-Year University Students in Bangladesh: A Cross-Sectional Study". *International Journal of Mental Health and Addiction* (2020). <https://doi.org/10.1007/s11469-020-00242-y>.
- [2] Dr. Nazma Sultana. "Stress and Depression among undergraduate Medical Students of Bangladesh". *Bangladesh Journal of Medical Education*, ISSN: 2306-0654 (2011). Vol.-02, Issue-01.
- [3] Bhuiyan, Md Aminul H., Mark D. Griffiths, and Mohammed A. Mamun. "Depression literacy among Bangladeshi pre-university students: differences based on gender, educational attainment, depression, and anxiety." *Asian journal of psychiatry* 50 (2020).
- [4] Mamun, Mohammed A., Md Sharif Hossain, and Mark D. Griffiths. "Mental health problems and associated predictors among Bangladeshi students." *International Journal of Mental Health and Addiction* (2019): 1-15.
- [5] Bruffaerts, Ronny, Philippe Mortier, Glenn Kiekens, Randy P. Auerbach, Pim Cuijpers, Koen Demyttenaere, Jennifer G. Green, Matthew K. Nock, and Ronald C. Kessler. "Mental health problems in college freshmen: Prevalence and academic functioning." *Journal of affective disorders* 225 (2018): 97-103.
- [6] Hamaideh, Shaher H. "Alexithymia among Jordanian university students: Its prevalence and correlates with depression, anxiety, stress, and demographics." *Perspectives in psychiatric care* 54, no. 2 (2017): 274-280.
- [7] Sidana, Surbhi, Jugal Kishore, Vidya Ghosh, Divyansh Gulati, R. C. Jiloha, and Tanu Anand. "Prevalence of depression in students of a medical college in New Delhi: a cross-sectional study." *The Australasian medical journal* 5, no. 5 (2012): 247.
- [8] Mei, Guang, Weisheng Xu, Li Li, Zhen Zhao, Hao Li, Wenqing Liu, and Yueming Jiao. "The role of campus data in representing depression among college students: exploratory research." *JMIR mental health* 7, no. 1 (2020): e12503.
- [9] Dabana, Aisha, and Abdulrazaq A. Gobir. "Depression among students of a Nigerian University: Prevalence and academic correlates." *Archives of Medicine and Surgery* 3, no. 1 (2018): 6.
- [10] 15 types of regression in data science. Available online: <https://www.listendata.com/2018/03/regression-analysis.html#Logistic-Regression>, [Last accessed on March 28, 2021]
- [11] Logistic Regression in Machine Learning. Available online: <https://www.javatpoint.com/logistic-regression-in-machine-learning>, [Last accessed on March 12, 2021]
- [12] Simple Logistic Regression. Available online: [https://stats.libretexts.org/Bookshelves/Applied\\_Statistics/Book%3A\\_Biological\\_Statistics\\_\(McDonald\)/05%3A\\_Tests\\_for\\_Multiple\\_Measurement\\_Variables/5.06%3A\\_Simple\\_Logistic\\_Regression](https://stats.libretexts.org/Bookshelves/Applied_Statistics/Book%3A_Biological_Statistics_(McDonald)/05%3A_Tests_for_Multiple_Measurement_Variables/5.06%3A_Simple_Logistic_Regression), [Last accessed on March 28, 2021]

[13] Platt, John. "Sequential minimal optimization: A fast algorithm for training support vector machines." (1998).

[14]Landwehr, Niels, Mark Hall, and Eibe Frank. "Logistic model trees." *Machine learning* 59, no. 1-2 (2005): 161-205.

[15]Aha, David W., Dennis Kibler, and Marc K. Albert. "Instance-based learning algorithms." *Machine learning* 6, no. 1 (1991): 37-66.

[16]Saravanan, N., and V. Gayathri. "Performance and classification evaluation of J48 algorithm and Kendall's based J48 algorithm (KNJ48)." *International Journal of Computer Trends and Technology (IJCTT)* 59, no. 2 (2018): 73-80.

[17]Mahmood, Deeman Y., and Mohammed A. Hussein. "Intrusion detection system based on K-star classifier and feature set reduction." *International Organization of Scientific Research Journal of Computer Engineering (IOSR-JCE)* Vol 15, no. 5 (2013): 107-112.

**Thank You!**  
**Best of Luck**

## Depression Level Prediction on Bangladeshi Students

### ORIGINALITY REPORT

16%	13%	9%	10%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

### PRIMARY SOURCES

1	<a href="https://dspace.daffodilvarsity.edu.bd:8080">dspace.daffodilvarsity.edu.bd:8080</a> Internet Source	3%
2	Submitted to Griffith College Dublin Student Paper	1%
3	<a href="https://en.wikipedia.org">en.wikipedia.org</a> Internet Source	1%
4	<a href="https://link.springer.com">link.springer.com</a> Internet Source	1%
5	<a href="https://pdfs.semanticscholar.org">pdfs.semanticscholar.org</a> Internet Source	1%
6	Submitted to The University of the South Pacific Student Paper	<1%
7	<a href="https://mafiadoc.com">mafiadoc.com</a> Internet Source	<1%
8	A. B. M. Shawkat Ali. "Chapter 20 Performance Analysis of Statistical Classifier SMO with other Data Mining Classifiers",	<1%

Springer Science and Business Media LLC,  
2003

Publication

- 
- |           |  |                |
|-----------|--|----------------|
| <b>9</b>  | Md. Saiful Islam, Most. Zannatul Ferdous, Marc N. Potenza. "Panic and generalized anxiety during the COVID-19 pandemic among Bangladeshi people: An online pilot survey early in the outbreak", Journal of Affective Disorders, 2020 | <b>&lt;1 %</b> |
| <hr/>     |  |                |
| <b>10</b> | <b>irep.ntu.ac.uk</b><br>Internet Source   | <b>&lt;1 %</b> |
| <hr/>     |  |                |
| <b>11</b> | Naoual Nassiri, Abdelhak Lakhouaja, Violetta Cavalli-Sforza. "Arabic L2 readability assessment: Dimensionality reduction study", Journal of King Saud University - Computer and Information Sciences, 2021                           | <b>&lt;1 %</b> |
| <hr/>     |  |                |
| <b>12</b> | <b>www.ijert.org</b><br>Internet Source  | <b>&lt;1 %</b> |
| <hr/>     |  |                |
| <b>13</b> | <b>Submitted to Ain Shams University</b><br>Student Paper  | <b>&lt;1 %</b> |
| <hr/>     |  |                |
| <b>14</b> | <b>Submitted to California State University, Fresno</b><br>Student Paper   | <b>&lt;1 %</b> |
| <hr/>     |  |                |
| <b>15</b> | <b>Submitted to Middlesex University</b><br>Student Paper  | <b>&lt;1 %</b> |
-

16	<a href="https://bpspsychub.onlinelibrary.wiley.com">bpspsychub.onlinelibrary.wiley.com</a> Internet Source	<1 %
17	Submitted to Northcentral Student Paper	<1 %
18	Submitted to KIPP NYC College Prep LLC Student Paper	<1 %
19	Submitted to American InterContinental University Student Paper	<1 %
20	<a href="http://www.larafrediani.it">www.larafrediani.it</a> Internet Source	<1 %
21	Submitted to Central Queensland University Student Paper	<1 %
22	Abu Sayeed, Md Nazmul Hassan, Md Hafizur Rahman, Samer El Hayek et al. "Facebook addiction associated with internet activity, depression and behavioral factors among university students of Bangladesh: A cross-sectional study", Children and Youth Services Review, 2020 Publication	<1 %
23	Submitted to Green University Of Bangladesh Student Paper	<1 %
24	<a href="https://arxiv.org">arxiv.org</a> Internet Source	<1 %

25	<a href="http://www.tandfonline.com">www.tandfonline.com</a> Internet Source	<1 %
26	Hrishikesh Bharadwaj Chakrapani, Smruti Chouraisa, Anish Saha, J. N. Swathi. "Predicting Performance Analysis of System Configurations to Contrast Feature Selection Methods", 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020 Publication	<1 %
27	Submitted to East Forsyth High School Student Paper	<1 %
28	Olusola Ayandele, Olugbenga A Popoola, Tolulope O Oladiji. "Addictive use of smartphone, depression and anxiety among female undergraduates in Nigeria: a cross-sectional study", Journal of Health Research, 2020 Publication	<1 %
29	Submitted to University of Liverpool Student Paper	<1 %
30	Submitted to KYUNG HEE UNIVERSITY Student Paper	<1 %
31	Submitted to University of London External System Student Paper	<1 %

---

32	Rachida Ihya, Abdelwahed Namir, Sanaa El Filali, Mohammed Ait Daoud, Fatima Zahra Guerss. "J48 algorithms of machine learning for predicting user's the acceptance of an E-orientation systems", Proceedings of the 4th International Conference on Smart City Applications - SCA '19, 2019	<1 %
----	---	------

Publication

---

33	<a href="http://www.listendata.com">www.listendata.com</a>	<1 %
----	--	------

Internet Source

---

34	A. Srivastava, Eui-Hong Sam Han, V. Singh, V. Kumar. "Parallel formulations of decision-tree classification algorithms", Proceedings. 1998 International Conference on Parallel Processing (Cat. No.98EX205), 1998	<1 %
----	--	------

Publication

---

35	Anamika Ahmed, Raihan Sultana, Md Tahmidur Rahman Ullas, Mariyam Begom, Md. Muzahidul Islam Rahi, Md. Ashrafur Alam. "A Machine Learning Approach to detect Depression and Anxiety using Supervised Learning", 2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), 2020	<1 %
----	--	------

Publication

---

36	<a href="http://researchbank.rmit.edu.au">researchbank.rmit.edu.au</a>	<1 %
----	--	------

---

Internet Source



37	<a href="http://www.hindawi.com">www.hindawi.com</a> Internet Source	<1 %
38	<a href="http://www.mdpi.com">www.mdpi.com</a> Internet Source	<1 %
39	<a href="http://www.preprints.org">www.preprints.org</a> Internet Source	<1 %
40	Hong Qiao, Yan-Guo Wang, Bo Zhang. "A simple decomposition algorithm for support vector machines with polynomial-time convergence", Pattern Recognition, 2007 Publication	<1 %
41	Submitted to University of Abertay Dundee Student Paper	<1 %
42	<a href="http://www.appraisalinstitute.org">www.appraisalinstitute.org</a> Internet Source	<1 %
43	Cheng, P.S.. "GC-MS quantification of ketamine, norketamine, and dehydronorketamine in urine specimens and comparative study using ELISA as the preliminary test methodology", Journal of Chromatography B, 20070601 Publication	<1 %
44	M. Rasheduzzaman, Firoj Mamun, Md. Omar Faruk, Ismail Hosen, Mohammed A. Mamun. "Depression in Bangladeshi university students: The role of sociodemographic,	<1 %

personal, and familial psychopathological factors", Perspectives in Psychiatric Care, 2021

Publication

---

45	<a href="#">idoc.pub</a> Internet Source	<1 %
46	<a href="#">ir.kabarak.ac.ke</a> Internet Source	<1 %
47	<a href="#">www.atlantis-press.com</a> Internet Source	<1 %
48	Md. Saiful Islam, Rejina Akter, Md. Tajuddin Sikder, Mark D. Griffiths. "Weight-Related Status and Associated Predictors with Psychological Well-being among First-Year University Students in Bangladesh: A Pilot Study", International Journal of Mental Health and Addiction, 2020 Publication	<1 %

---

---

Exclude quotes Off

Exclude matches Off

Exclude bibliography Off

