

**CRIME PREDICTION USING CLASSIFICATION RULE MINING**

**BY**

**MD NIMUL ISLAM**

**143-15-4523**

**ABU TALHA FARID**

**143-15-4383**

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

Supervised By

**Narayan Ranjan Chakraborty**

Assistant Professor

Department of Computer Science and Engineering

Faculty of Science & Information Technology

Daffodil International University



**DAFFODIL INTERNATIONAL UNIVERSITY**

**DHAKA, BANGLADESH**

**17 SEPTEMBER 2018**

## **APPROVAL**

This Project/internship titled “**Crime Prediction Using ClassificationRule Mining**”, submitted by Md Nimul Islam, ID No: 143-15-4523 And Abu Talha Farid, ID No: 143-15-4383 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 17 September 2018.

## **BOARD OF EXAMINERS**

---

**Dr. Syed AkhterHossain**  
**Professor and Head**

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

**Chairman**

---

**Dr. SheakRashedHaiderNoori**  
**Associate Professor & Associate Head**

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

**Internal Examiner**

---

**Md. ZahidHasan**  
**Assistant Professor**

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

**Internal Examiner**

---

**Dr. Mohammad ShorifUddin**  
**Professor**

Department of Computer Science and Engineering  
Jahangirnagar University

**External Examiner**

## **DECLARATION**

I hereby declare that, this Research has been done by me under the supervision of **Narayan Ranjan Chakraborty, Assistant Professor, Department of CSE**, Daffodil International University. I also declare that neither this Research nor any part of this Research has been submitted elsewhere for award of any degree or diploma.

### **Supervised by:**

**Narayan Ranjan Chakraborty**

**Assistant Professor**

Department of Computer Science and Engineering

Faculty of Science & Information Technology

Daffodil International University

### **Submitted by:**

**Md Nimul Islam**

ID: 143-15-4523

Department of Computer Science and Engineering

Faculty of Science & Information Technology

Daffodil International University

**Abu Talha Farid**

ID: 143-15-4383

Department of Computer Science and Engineering

Faculty of Science & Information Technology

Daffodil International University

## ACKNOWLEDGEMENT

I have given my efforts to this thesis. However, it would not have been possible without the kind support and help of many individuals. I would like to express my deepest appreciation to all those who provided me the possibility to complete this report.

At first, I express my heartiest thanks and gratefulness to almighty Allah for His divine blessings which allowed me to complete this thesis successfully.

A special gratitude I give to my supervisor, **Narayan Ranjan Chakraborty**, Assistant Professor of CSE department, whose contribution in stimulating suggestions and encouragement helped me to coordinate my thesis especially in writing this report. His endless patience, scholarly guidance, constant and energetic supervision, constructive criticism, valuable advice have made it possible to complete this thesis.

Furthermore, I would also like to acknowledge with much appreciation the crucial role of my department head, **Professor Dr Syed Akhter Hossain**, who provided me with his precious time and kind help to finish this thesis. I also give my deepest thanks to all the faculty members and staff of CSE department of Daffodil International University.

## **ABSTRACT**

Data mining produce workable significant information from a huge amount of discrete data. Data mining is the process of naturally looking huge stores of information to find patterns and prediction that go past basic examination. Data mining also assumes an imperative part as far as prediction and analysis[12]. We are using here three different classification algorithm in Weka. The algorithm we have used here Naive Bayes, Bayes Net, One-R. We have collected data from Dhaka metropolitan police, Chittagong metropolitan police, Bangladesh police, Newspaper and due to sensitive data from a secret agent. The main objective of this paper is to predict place from crime occurred in the last few years. The law enforcement agencies ought to in this way have the capacity to anticipate such increments or abatements or patterns in crime for example, the number of murder, robbery, burglaries or any such crime that may happen in a specific zone in a specific month, year, or any timespan or the general number of crime happening in a nation in a specific year later on, or any other expectation or projection of future crime measurements[7]. So that the law enforcement agency can find out the place easily where the crime rate is high to take proper action, thus reducing the crime rate.

## TABLE OF CONTENTS

<b>CONTENS</b>	<b>PAGE</b>
Board of Examiners	I
Declaration	Ii
Acknowledgements	iii
Abstract	Iv
<b>CHAPTER</b>	
<b>CHAPTER 1 :INTRODUCTION</b>	<b>1-3</b>
1.1 Introduction	1
1.2 Motivation	1
1.3 Rationale of the Study	1
1.4 Research Question	2
1.5 Expected Output	2
1.6 Report Layout	2
<b>CHAPTER 2: BACKGROUND</b>	<b>3-4</b>
2.1 Introduction	3
2.2 Related Works	3
2.3 Research Summary	3
2.4 Scope of the Problem	3
2.5 Challenges	4
<b>CHAPTER 3: RESEARCH METHODOLOGY</b>	<b>5-10</b>
3.1 Introduction	5
3.2 Research Subject and Instrumentation	5
3.2.1 Weka	5
3.2.2 Classification	5
3.3 Data Collection Procedure	6
3.3.1 Generating of Dataset	6
3.3.2 Preprocessing Data	6
3.3.3 Exploring the Data	6
3.4 Statistical Analysis	7

3.4.1 Data Mining	7
3.4.2 Sample Data	8
3.4.3 Test Dataset	9
<b>CHAPTER 4: EXPERIMENTAL RESULTS AND DISCUSSION</b>	<b>11-17</b>
4.1 Introduction	11
4.2 Experimental Result	11
4.2.1 Naïve Bayes	11
4.2.2 Bayes Net	13
4.2.3 ZeroR	15
4.3 Descriptive Analysis	16
4.4 Summary	17
<b>CHAPTER 5: CONCLUSION AND RECOMMENDATION</b>	<b>18</b>
5.1 Summary of the Study	18
5.2 Conclusions	18
5.3 Implication of Further Study	18
<b>REFERENCES</b>	<b>19</b>

## LIST OF FIGURES

<b>FIGURES</b>	<b>PAGE NO</b>
Figure 3.1: Data Preprocessing	8
Figure 3.2: Training Data	9
Figure 3.3: Test Data	10
Figure 4.1: Naïve Bayes Training Set Result	11
Figure 4.2: Visualization of Crime Rate	12
Figure 4.3: Bayes Net Test Dataset Result	13
Figure 4.4: Zero R Test Dataset Result	13
Figure 4.5: Procedure for Prediction of Crime Rate	14



## LIST OF TABLE

<b>TABLE</b>	<b>PAGE NO</b>
Table 4.1: Table Shows The Percentage of The Instance	12
Table 4.2: Performance Comparison between Different Algorithms	15

# CHAPTER 1

## INTRODUCTION

### 1.1 Introduction

Crime rates in our society are growing up rapidly day by day. Crime will play a negative impact on our society or inside a country when it happens regularly[12]. Crime becomes the dangerous issue for any country or society crime destroy a country's natural environment and destabilizes a country. Now in this modern era criminal are offered a different type of crime in different ways. At present criminals are using a lot of technology to offered crime[7]. So to prevent crime we have to move forward. This research is about to help the law enforcement agency by analyzing the crime related data that occurs last few years.

### 1.2 Motivation

Crime is increasing day by day at the alarming rate[5]. It is not possible to suppress crime accordingly it we look at the crimes that committed in our country in the last few years it will be seen that that crime rate is in cruising highly day after day. The general public is suffering from this but the manpower on the resource of our law enforcement agency is insufficient. It is needed to help the law enforcement agency to reduce it. To help local law enforcement agencies so that they would be able to more aware about criminal activities[7]. On the other hand, there are several related works with this research but in Bangladesh, no one has done yet.

### 1.3 Rationale of the Study

It is clear that in order to improve the development of a developing country like Bangladesh the crime committed inside a country should be reduced within the faster time. Crime pattern analysis and detection can only help the law enforcement agencies to find out the crime rate and place [6]. To reduce the amount of crime that committed in the past few years, collecting the data about the type of crimes amount and area by using different data mining classification algorithm like Naïve Bayes, Bayes Net, One-R to find

out the place where the crime rate is so high and which type of crime is occurring there.[11]

#### **1.4 Research Question**

Question of this work are:

- How crime threats be identified on initial stage.
- When algorithms can provide best outcome for this research work
- What are the parameters to help find out the criminal activity responsible for high rate?
- How to find out the area where crime occurs in high rate.

#### **1.5 Expected Outcome**

We will try to find out the place where the crime rate is high and also try to find out the type of crime.

#### **1.6 Report Layout**

The following description is given to understand which section covers which themes furthermore, their relevant discussions. In Chapter One we have discussed the motivation for choosing this research, The rationale of the Study, which question we have faced to do this research, what would be the expected outcome, we have also discussed the introduction here. In Chapter Two we have discussed the related work which has been done with this research, research summary, the scope of the problem and which challenges we have faced to do this research. In Chapter Three we have discussed the methodology and procedure, what is the research Subject and Instrumentation also discussed the data preprocessing process, statistical analysis has been shown with some figure, requirements for implementation. In Chapter Four experimental result, Descriptive Analysis described with some figure and table. In Chapter Five the summary of the project, conclusion and future scope of this research has been discussed here as well.

## CHAPTER 2

### BACKGROUND

#### 2.1 Introduction

This project is about crime and the prediction of crime. The main work of this project is to find out the rate of committed crime last few years.

#### 2.2 Related Works

- To determine crime pattern between different geographical location [5].
- In which pattern crimes are occurring in a particular area the interim/interval of occurring crimes [7].
- To help the user some future predictions would make by training crime data sets and to visualize the output in a simple order to perceive for the user [12].
- Use of clustering algorithm for data mining obtainment to detect the crime patterns and excite the method of solving crimes [6].
- Optimization of outlier operation parameter using G A [4].

#### 2.3 Research Summary

We have noticed that several research has already done internationally. But nobody has done any research in Bangladesh. They have done their research using classification and clustering method on the data mining approach. They used a big dataset where some data were training data and some were training test data. We also used training data and test data but our data was not so much because there was no available data.

#### 2.4 Scope of the Problem

To complete this research we have faced many problems. The main problem we faced that nobody has done any research on this problem. So it was difficult for us to do this research.

## **2.5 Challenges**

The challenges we faced for the research are:

- To collect crime data from law enforcement agency
- Because of having sensitive data there was not enough data available
- Nobody agreed to provide data about a crime to us.

## CHAPTER 3

### RESEARCH METHODOLOGY

#### 3.1 Introduction

It has been already mentioned in the last chapter that the crime rate in Bangladesh has been increasing day by day. So it's needed to find out a solution to reduce the crime rate. So that the law enforcement agency can find a way to solve it.

#### 3.2 Research Subject and Instrumentation

##### 3.2.1 Weka

Weka contains a collection of visualization tools and algorithm. Weka is used for analyzing data and predictive modeling. Weka is also used for data preprocessing, clustering, classification, and feature selection of research in education [5]. It is easy to use. Weka is built in Java language. ARFF format, CSV format, libSVM'S format and Cu.5's format are supported in Weka [11]. By Weka, we can predict on the assumption of that data which is available and usable [8].

The advantage of using Weka

- Free availability under the GNU General public license [13].
- Weka is implemented in Java language and it easily runs in any modern computing platform [13].
- It is a technique of widespread gathering of data pre-processing and modeling [13].

##### 3.2.2 Classification

Data mining classification is a process that makes out in a gathering of targeted classes. The Classification used to correctly predict the destination class for each case in the data. Classification is a technique which is used to predict the relevance among the group of data instance [11]. There are different types of classification method in data mining. These methods are Bayesian-networks, decision tree, induction, case-based reasoning,

fuzzy logic, genetic algorithm and k-nearest neighbor. In this paper, we have used some classification algorithm like naïve-Bayes. Bayes net and ZeroR [11]. These classification methods can be used in Weka. Additionally the output of using these algorithm are likened and converted into the graphical representation also accomplished[11].

### **3.3 Data Collection Procedure**

#### **3.3.1 Generating of Data set**

To generate a dataset it can be done by getting crime data from the different organization. We have collected data from the website of Dhaka metropolitan police, Bangladesh police and from a secret agency.

#### **3.3.2 Preprocessing Data**

Data pre-processing is a technique by which we can get understandable data from raw data[8]. Genuine information is frequently fragmented, conflicting as well as ailing in specific practices or inclines and is probably going to contain numerous blunders. Information preprocessing is a demonstrated technique for settling such issues. Data pre-processing gets ready crude information. First of all, we have collected data from the different place as monthly, yearly. We have also collected from different metropolitan police in Bangladesh as district wise and police station wise. All the data were separated. Then we arranged the data accordingly to the year. And we categorized the crime in different types like dacoity, robbery, murder etc.

#### **3.3.3 Exploring the data**

Finally using some algorithm technique we have analyzed those data what have been collected to get the better result.

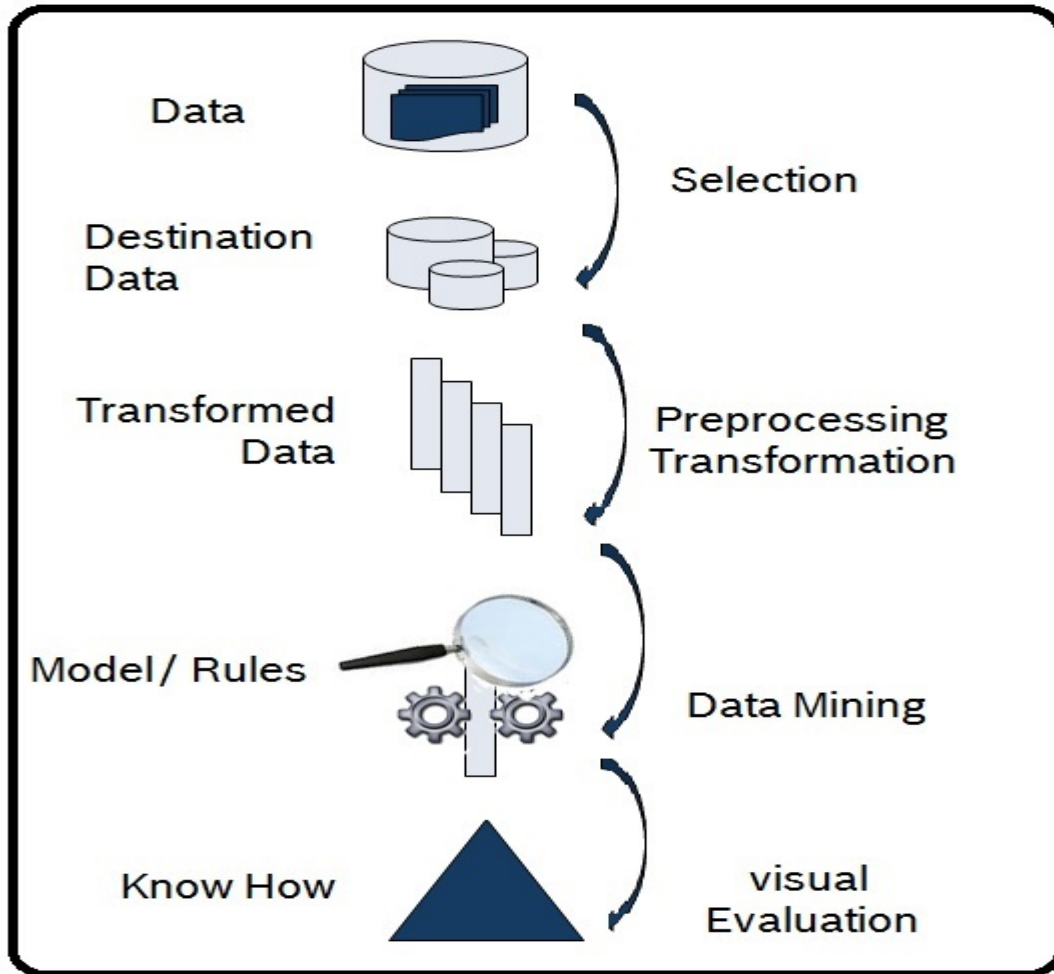


Figure 3.1: Data preprocessing

### 3.4 Statistical Analysis

#### 3.4.1 Data mining

Data Mining is used to find out the anomalies, patterns, and correlation using huge data to predict some outcome[8]. We can also find out new information from a lot of data. The would be new and useful. Data mining includes six basic classes of errands [8].

- Anomaly detection
- Association rule learning
- Clustering
- Classification



- Regression
- Summarization

We used classification in this research for prediction of crime.

There are different types of data mining algorithm[10]. Some of them are

- Pattern recognition
- Bayesian network
- Neural network
- Classification tree
- Naive Bayes

We used in our thesis paper naïve Bayes classification algorithm for the better result.

### **3.4.2 Sample Data**

In statistics and quantitative research methodology, a sample data is a set of data which is collected from a statistical population by a defined procedure. The elements of a sample are known as sample points, sampling units or observations. The sample usually represents a subset of manageable size.

```

C:\Users\nimul_31e0jb\OneDrive\Desktop\All\project 27.07.18\police data.aff - Notepad++
File Edit Search View Encoding Language Settings Tools Macro Run Plugins Window ?
police data.aff
3 @ATTRIBUTE unit-name {BMP, Barisal-Range, CMP, Chittagong-Range, DMP, Dhaka-Range, FMP, Khulna-Range, RMP, Railway-Range, Rajshahi-Range, Rangpur-Range, SMP, Sylhet-Range}
4 @ATTRIBUTE year REAL
5 @ATTRIBUTE dacoity REAL
6 @ATTRIBUTE robbery REAL
7 @ATTRIBUTE murder REAL
8 @ATTRIBUTE speedy-trial REAL
9 @ATTRIBUTE riot REAL
10 @ATTRIBUTE woman-and-child-repression REAL
11 @ATTRIBUTE kidnapping REAL
12 @ATTRIBUTE police-assault REAL
13 @ATTRIBUTE burglary REAL
14 @ATTRIBUTE theft REAL
15 @ATTRIBUTE other cases REAL
16 @ATTRIBUTE arms-act REAL
17 @ATTRIBUTE explosive REAL
18 @ATTRIBUTE narcotics REAL
19 @ATTRIBUTE smuggling REAL
20 @ATTRIBUTE total-cases REAL
21
22 @DATA
23 DMP,2013,47,220,245,363,3,1370,139,155,555,1915,7228,518,82,10535,144,23519
24 DMP,2014,47,265,262,332,7,1611,176,104,650,2190,6219,225,104,7063,222,19417
25 DMP,2015,45,205,239,226,26,1550,146,118,642,1711,5795,263,195,8365,201,19727
26 DMP,2016,4,20,48,43,1,402,24,16,126,332,1236,55,17,2633,20,4977
27 DMP,2017,20,103,218,202,7,1779,85,132,554,1197,5315,157,108,13638,115,23630
28 DMP,2018,11,45,115,117,1,903,33,65,292,648,2672,79,72,8111,72,13236
29 CMP,2013,16,108,94,31,7,455,37,31,123,314,1831,51,0,866,99,4063
30 CMP,2014,5,63,120,52,17,387,53,31,97,310,1457,121,10,2488,56,5267
31 CMP,2015,5,61,105,67,25,357,51,30,128,262,1497,101,40,2090,43,4962
32 CMP,2016,1,5,21,13,2,89,3,3,22,53,412,47,1,967,10,1649
33 CMP,2017,2,71,70,80,0,350,7,21,92,267,1547,231,2,4260,44,7044
34 CMP,2018,6,34,37,42,0,194,3,15,43,105,739,135,3,2310,24,3690
35 FMP,2013,3,9,29,25,0,153,11,4,65,91,551,19,2,792,13,1767
36 FMP,2014,2,29,22,17,0,206,11,3,44,101,522,34,2,599,10,1602
37 FMP,2015,9,15,30,19,0,146,7,3,49,95,574,21,13,659,2,1642
38 FMP,2016,1,2,9,9,0,40,1,2,11,27,135,4,3,213,1,458
39 FMP,2017,6,9,25,19,0,248,4,6,35,50,443,25,4,1706,32,2612
Normal text file length: 6,443 lines: 107 Ln:3 Col:2 Sel:0|0 Unix (LF) UTF-8 INS

```

Figure 3.2: Training Data

### 3.4.3 Test Dataset

We can divide data set into the training dataset and testing dataset. It is important for evaluating a model in data mining. When we divide a dataset into two parts most of them are used for training dataset, and least is used for test data set for testing the result.

```

* C:\Users\nimul_31e0jib\OneDrive\Desktop\All\project 27.07.18\police data.aff - Notepad++
File Edit Search View Encoding Language Settings Tools Macro Run Plugins Window ?
police data.aff
3 @ATTRIBUTE unit-name {BMP, Barisal-Range, CMP, Chittagong-Range, DMP, Dhaka-Range, EMP, Khulna-Range, RMP, Railway-Range, Rajshahi-Range, Rangpur-Range, SM
4 @ATTRIBUTE year REAL
5 @ATTRIBUTE dacoity REAL
6 @ATTRIBUTE robbery REAL
7 @ATTRIBUTE murder REAL
8 @ATTRIBUTE speedy-trial REAL
9 @ATTRIBUTE riot REAL
10 @ATTRIBUTE woman-and-child-repression REAL
11 @ATTRIBUTE kidnapping REAL
12 @ATTRIBUTE police-assault REAL
13 @ATTRIBUTE burglary REAL
14 @ATTRIBUTE theft REAL
15 @ATTRIBUTE other-cases REAL
16 @ATTRIBUTE arms-act REAL
17 @ATTRIBUTE explosive REAL
18 @ATTRIBUTE narcotics REAL
19 @ATTRIBUTE smuggling REAL
20 @ATTRIBUTE total-cases REAL
21
22 @DATA
23
24 ?,2018,11,45,115,117,1,903,33,65,292,648,2672,79,72,8111,72,13236
25 ?,2017,2,71,70,80,0,350,7,21,92,267,1547,231,2,4260,44,7044
26 ?,2016,1,2,9,9,0,40,1,2,11,27,135,4,3,213,1,458
27 ?,2015,0,19,20,12,4,131,8,18,39,66,463,23,30,733,179,1745
28 ?,2014,7,16,15,3,0,94,3,9,35,48,375,16,1,759,5,1386
29 ?,2013,162,199,1153,362,7,4272,171,71,643,1477,19966,309,30,4459,993,34274
30 ?,2017,64,74,626,114,0,1889,77,112,360,786,11676,581,16,12555,190,29120
31 ?,2014,80,125,520,277,15,2351,77,53,310,585,10247,381,152,2935,1766,19874
32 ?,2015,31,47,175,58,1,1497,53,20,156,266,6766,26,22,5797,27,14942
33 ?,2013,54,130,757,276,29,4095,105,35,442,1442,20654,134,10,3614,2509,34286
34 ?,2018,10,5,173,9,0,829,7,6,39,164,3945,37,9,6386,451,12070
35 ?.2014.1.8.33.3.0.6.0.0.1.89.170.10.3.395.110.829
Normal text file length: 1,396 lines: 37 Ln: 25 Col: 23 Sel: 0|0 Unix (LF) UTF-8 INS

```

Figure 3.3: Test Data

## CHAPTER 4

### EXPERIMENTAL RESULTS AND DISCUSSION

#### 4.1 Introduction

The research has been done to reduce the number of crime rates that committed in Bangladesh. We have collected data about crimes that committed the last few years in the different area and try to find out which crime is occurring so much and which area it's rate is high.

#### 4.2 Experimental Results

##### 4.2.1 Naïve Bayes

Naive Bayes is a classifier technique. This strategy depends on freedom presumption between indicators. Bayes theorem used to computing the back likelihood  $P(c/x)$  from  $P(c)$ ,  $p(x)$  and  $p(c)$ . In this classifier expect that the result of the estimation of a predictor(x) or a specific class(c) is an autonomous of the estimation of indicator. This is called class restrictive freedom.  $p(c/x) = (p(x/c)p(c))/p(x)$

We can define it as,

Posterior = ((prior)(likelihood))/evidence

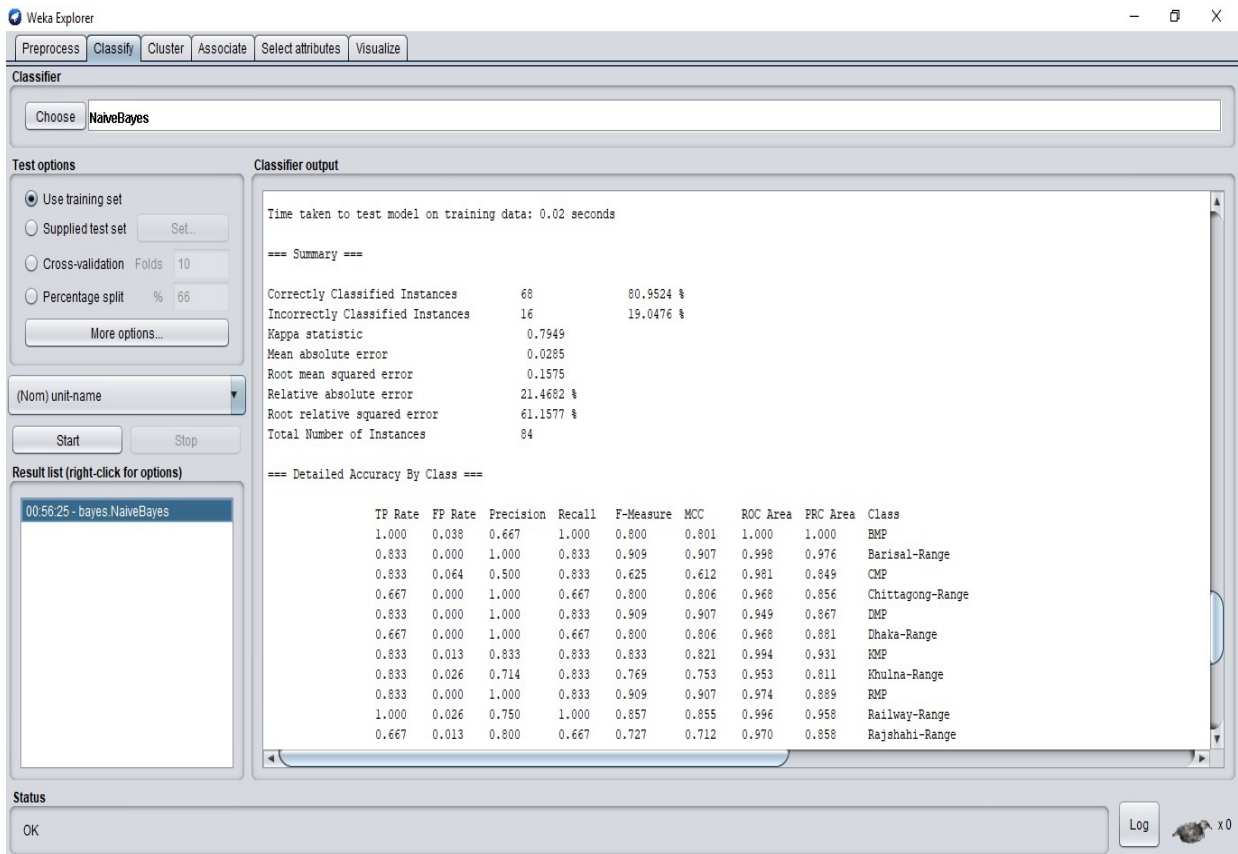


Figure 4.1: Naïve Bayes training set result

For displaying the model we used here Naïve Bayes classification algorithm in Weka. The model represents the relation, attributes, an instance which is involved in the execution section.

TABLE 4.1: Table Shows The Percentage of The Instance

Testing options	Correctly classified instance	Incorrectly classified instance
Training set	80.9524%	19.0476%

The final result of test data set is given below

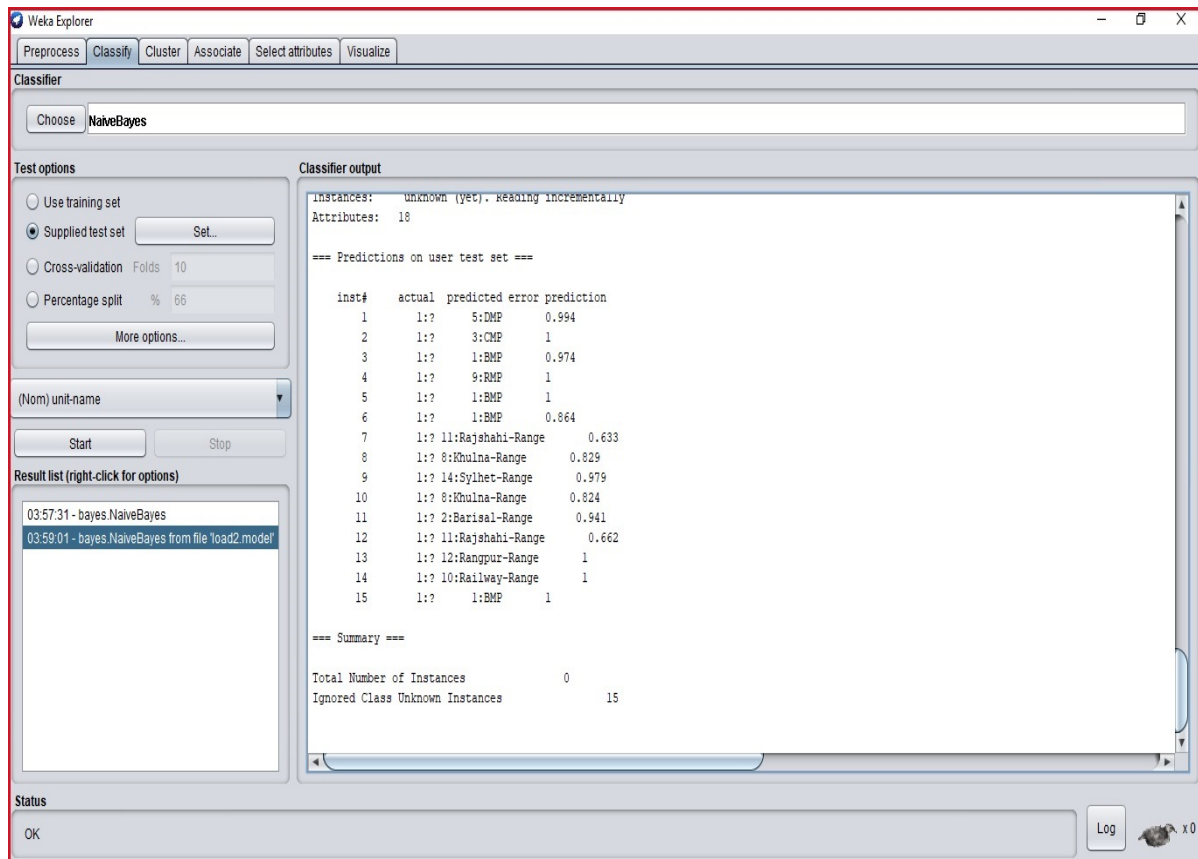


Figure 4.2: Visualization of crime rate

From the result, we can predict the place or unit name from the crime rate.

### 4.2.2 Bayes net

Bayes net otherwise called Bayesian system Bayes organize .Bayes net are probabilistic graphical model. We can utilized it to fabricate a model from dataset. We can utilize this for indicator, identification of abnormality reason, and expectation of time arrangement .Bayesian systems is utilized to speak to an arrangement of irregular examination and under their contingent acquiescence by a coordinated non-cyclic diagram.

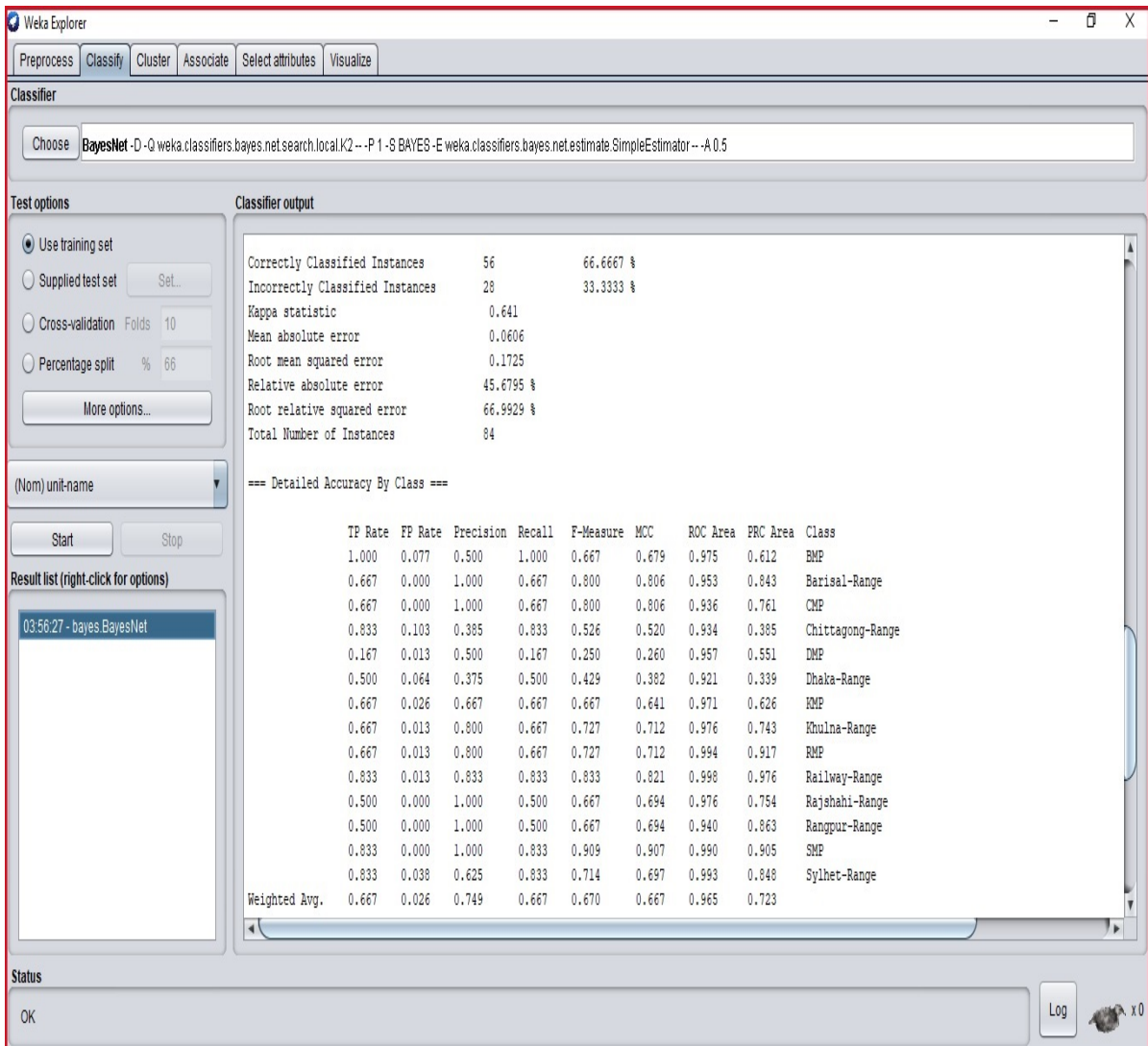


Figure 4.3: Bayes Net Test Data Result

We can see here that when we use Bayes net classification algorithm the correctly instance is 66.66% and the incorrectly instance is 33.33%.

### 4.2.3 ZeroR

Zero are is a basic characterization technique that relies upon the goal and neglect all indicators. Zero classifier basically estimate the lion's share class .Zero R has no consistency .So we can utilize this classifier for deciding gauge achievement for other order strategy .In this calculation we need to build a table and have select most regular esteem.

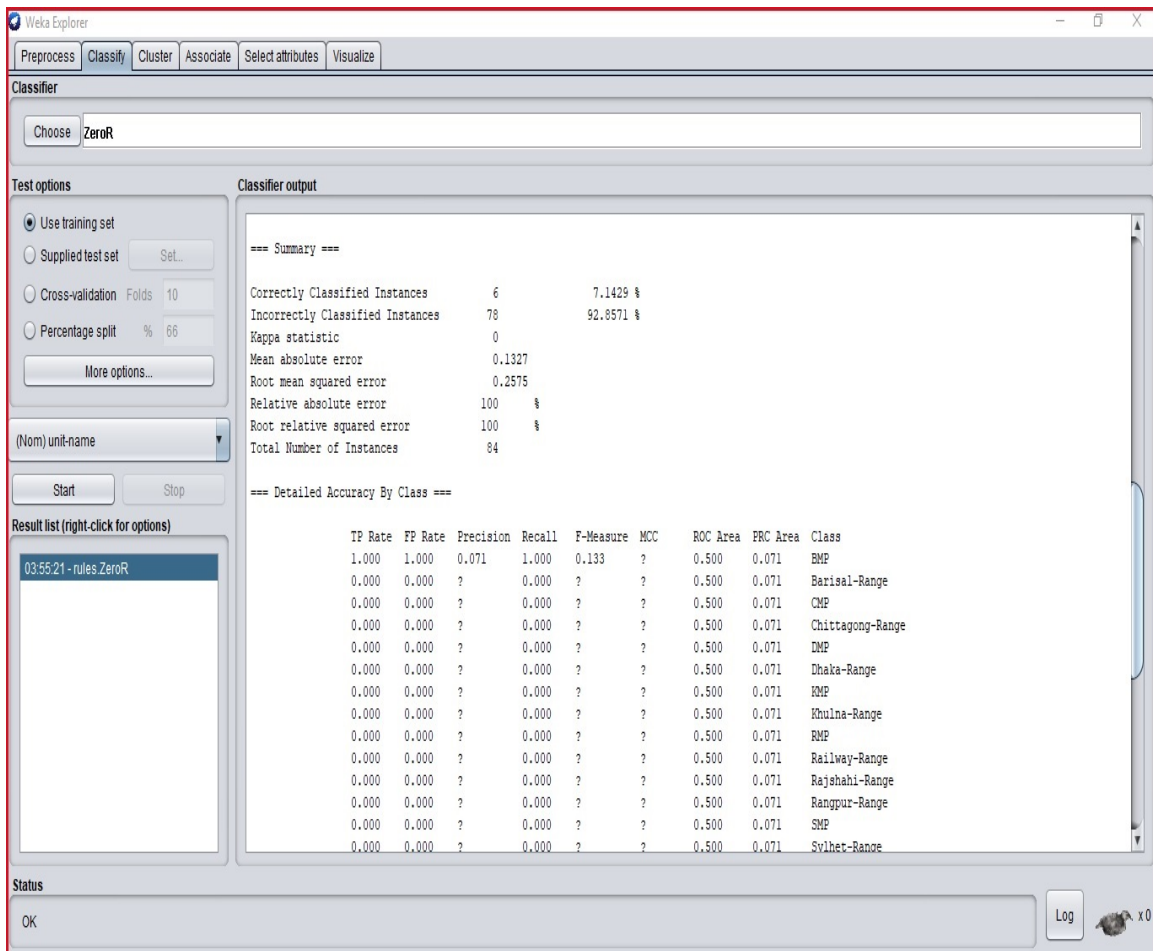


Figure 4.4: ZeroR Test Data Result

We can see here that when we use Bayes net classification algorithm the correctly instance is 7.14% and the incorrectly instance is 92.85%.



### 4.3 Descriptive Analysis

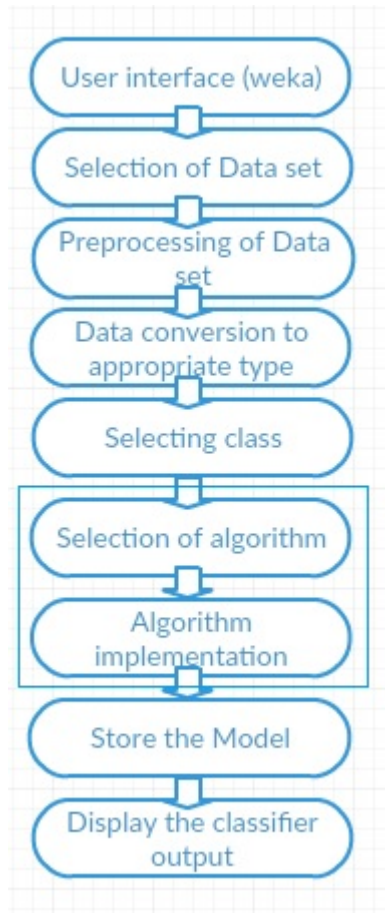


Figure 4.5: Procedure for Prediction of crime rate

We have used here different type of Classification algorithm to compare the performance by using the test data set [11]. We can see the correctly classified instance and mean absolute error from the table.

TABLE 4.2: Performance Comparison between Different Algorithms

Algorithm	Correctly classified instances	Mean absolute error
Naïve Bayes	80.95%	0.0285
Bayes net	66.66%	0.0606
One R	7.14%	0.1327

The table shows the correctly classified instances using the different type of classification algorithm. From the table, it is clear that the correctly classified instance of Naïve Bayes

classifier algorithm is maximum than other algorithms. The minimum correctly classified instances is 7.14%. So we use here Naïve Bayes classification algorithm for a better result to find out the place where crime rate so high.

From the result, we can predict the place or unit name from the crime rate.

#### **4.4 Summary**

By using the different type of algorithm it is clear that the Naïve Bayes classification algorithm will be better than using other algorithms because it gives 80.95% accuracy.

## **CHAPTER 5**

### **CONCLUSION AND RECOMMENDATION**

#### **5.1 Summary of the study**

The research work is done to predict the place where the crime rate is high. We needed the crime related data to accomplish the research work. We have collected data from the different place as metropolitan wise, district wise and police station wise that occurs in the different year. Then we preprocessed the data. We used 75% as train dataset and 25% as the test dataset.

#### **5.2 Conclusion**

This research work represents a new framework for classification and predicting crime rate based on real data. In this research we tried to use different data mining techniques, algorithms and models to find out the better result for predicting. It will help the law enforcement agency to reduce the crime rate that occurs in the different place.

#### **5.3 Implication for Further Study**

In the future, we will try to build a system which will be able to predict the crime rate which may occur in the future based on committed crime in the last few years.

## REFERENCES

- [1] Dhaka Metropolitan police website ,<https://dmp.gov.bd/>, Last accessed on 25.07.2018
- [2] Chittagong Metropolitan Police Website, <http://cmp.gov.bd/> ,Last accessed on 26.07.2018
- [3] Bangladesh police website, [https://www.police.gov.bd/en/crime\\_statistic/year/2015](https://www.police.gov.bd/en/crime_statistic/year/2015),  
Last accessed on 27.07.2018
- [4] Rasoul Kiani , Siamak Mahdavi and Amin Keshavarzi Analysis and Prediction of Crimes by Clustering and Classification, (IJARAI) International Journal of Advanced Research in Artificial Intelligence, Vol. 4, No.8, 2015
- [5] Shyam Varan Nath, Crime Pattern Detection Using Data Mining
- [6] Neha Gohar Khan and Prof.V.B.Bhagat(Mate) Effective Data Mining Approach For Crime-Terrorpattern Detection Using Clustering Algorithm Technique, International Journal of Engineering Research & Technology (IJERT), Vol. 2 Issue 4, April – 2013
- [7] Vineet Pande , Viraj Samant and Sindhu Nair, Crime Detection using Data Mining, International Journal of Engineering Research & Technology (IJERT) <http://www.ijert.org> ISSN: 2278-0181[paper1]
- [8] [https://en.wikipedia.org/wiki/Data\\_mining](https://en.wikipedia.org/wiki/Data_mining)Last accessed on 27.07.2018
- [9 ] <https://docs.microsoft.com/en-us/sql/analysis-services/data-mining/testing-and-validation-data-mining?view=sql-server-2017>Last accessed on 25.07.2018]
- [10] [https://www.ibm.com/support/knowledgecenter/en/SSEPGG\\_9.7.0/com.ibm.im.overview.doc/c\\_naive\\_bayes\\_classification.html](https://www.ibm.com/support/knowledgecenter/en/SSEPGG_9.7.0/com.ibm.im.overview.doc/c_naive_bayes_classification.html)Last accessed on 23.07.2018
- [11] S.Muthuselvan and Dr. k. Soma Sundaram(2016),Prediction of breast cancer Using classification Rule Mining Technique in Blood Test Dataset.ISBN:978-1-5090-2552-7
12. Dr. A.Bharathi, R. Shilpa,A Survey on Crime Data Analysis of Data Mining Using Clustering Techniques.[last accessed on 25.07.2018
- 13.[https://en.wikipedia.org/wiki/Weka\\_\(machine\\_learning\)](https://en.wikipedia.org/wiki/Weka_(machine_learning)) Last accessed on 25.07.2018]