Implementing Data Mining Algorithms for the Analysis of Crime Trends

Prof. Maria Gonzalez [1], Dr. Javier Martinez [2] & Ms. Lucia Rivera[*][3]

[1]Department of Civil Engineering, Universidad Nacional Autónoma de México (UNAM), Mexico City, Mexico

[2]Research Associate, Institute of Structural Engineering, Universidad Politécnica de Madrid, Spain

[3]Graduate Student, Department of Environmental Engineering, Universidad de los Andes, Santiago, Chile

ABSTRACT

Crimes are a social disturbance and cost the general public to a great degree from numerous points of view. Any examination that can help in breaking down and solving crime faster pays for itself. Crime data mining has the capacity of extricating helpful data and concealed examples from the substantial wrongdoing informational indexes. The crime data mining challenges are getting to be fortifying open doors for the coming years. Since the writing of crime information mining has expanded energetically as of late, it winds up obligatory to build up a diagram of the cutting edge. This orderly survey centers around crime data mining procedures and innovations utilized as a part of past investigations. The current work is grouped into various classifications and is introduced utilizing perceptions. This paper additionally demonstrates a few difficulties identified with crime data research.

KEYWORDS: Crime data mining; crime data analysis; systematic review; systematic study.   .

I.   INTRODUCTION

Wrongdoing is one of the predominant and concerning feature in any general public. The expansion in wrongdoing rates is one of the reason for the alert. Law requirement offices, knowledge organizations and police keep up wrongdoing databases[1,2]. The wrongdoing information be investigated to pick up bits of knowledge and to remove learning from it. A few investigations have found number of strategies to break down the wrongdoing information [3]. The crime data analysis can provide the crime statistics of a region, country or world [4]. The law requirement organizations can take better choices for success of the subjects by understanding the different parameters that impact the violations [5].

A. Crime data mining information

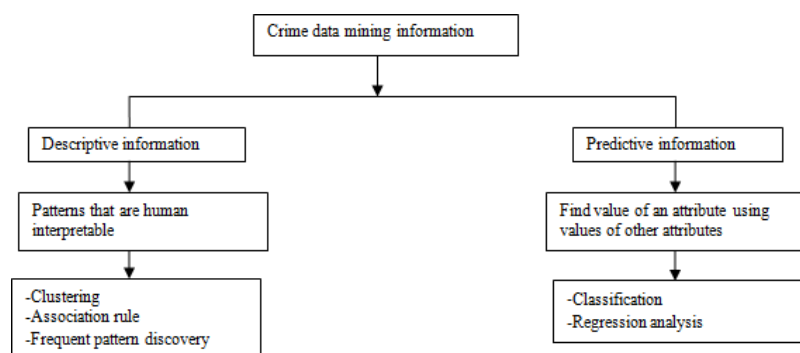Crime data mining information can be of different types as shown in the Figure 1.

Figure 1. Crime data mining information types

B. Motivations and Objectives

This investigation is completed keeping in mind the end goal to investigate the wrongdoing information mining systems, challenges and to know the advancements which are pertinent. This investigation is obligatory to make it conceivable to know which classes of wrongdoing information mining innovations, procedures and difficulties have been canvassed in past research and distinguishes holes.

This examination goes for efficiently assessing the wrongdoing information mining methods, issues and difficulties and advancements utilized as a part of existing investigations. The outcomes may help the scientist to get a review of the status of wrongdoing information mining and feature the exploration holes

The information mining benchmark is grounded by orders, for example, machine learning [4], manmade brainpower [5], likelihood [6] and insights [7]. The orders distinguished among the papers looked into are condensed in Table 1. Table 1 resources the orders specified for the papers looked into.

This paper is structured as follows. Section II describes the research methodology used in this study. Section III gives the classification of crime data mining and analysis papers considering the following criteria: (1) techniques used, (2) technologies used and (3) challenges addressed. Section IV presents the sources of crime data and type of crime. Section V presents the summary of researches in crime data mining. Section VI discusses the paper. Section VII presents the conclusion and future directions

II.   Related Work

The research methodology is composed of three stages. The first stage involves the research of works related to crime data mining and analysis. The second stage is concerned with establishing a classification scheme described in Section III. The third stage involves the presentation of summary of researches in crime data mining and analysis and the report of detailed literature review.

(1)   RQ.1: What are the different techniques for the crime data mining and analysis?

To answer this question Section III describes the frequency of publications corresponding to the various techniques in this area.

(2)   RQ.2: What are the technologies used in crime data mining and analysis?

To order to be able to answer this question, Table I in Section III describes the technologies used in crime data mining and analysis.

(3)   RQ.3: What are the various challenges involved in crime data mining and analysis?

To answer this question pie chart in Section III describes the various challenges and issues involved. The research is initialized with these queries and then follows the steps described.

A. Search Strategy and Screening
1. Sources of Information

Violations in India are stirred up at a disturbing rate, and lawbreakers are picking strange exercises to confer them. Daily papers, Web online journals, and so forth are everyday loaded with var ious wrongdoing occurrences. A portion of the perplexed wrongdoings that happened in India in most recent few years are specified here.12 A teacher was pounded the life out of by his own understudies in Ujjain, Madhya Pradesh. A group of nine cab drivers from Gurgaon, Haryana, ransacked and slaughtered no less than 35 individuals in the wake of offering them lift. Rowdy horde stripped and attacked a young lady in full general visibility at the Gateway of India, Mumbai, on the New Year eve. Days after awful Nithari killing, 4 disintegrated assemblages of kids were recuperated from surrendered godown in Punjab. Sexually ambushed young ladies in the Kashmir valley are as yet attempting to adapt up to injury. These occurrences uncover how violations are turning into a developing scourge in India and have turned into a Some capable variables that win in India for sheer increment in wrongdoings are neediness, relocation, joblessness, disappointment, starvation, lack of education, debasement, nepotism, swelling, and so on. Effect of such wrongdoings is that today individuals living in India now center their eyes toward wrongdoing examination organizations and security offices to check and control violations. Right now, physical examination by organizations has the likelihood to disregard and disregard the strong wrongdoing highlights.

III.   SOURCES OF CRIME DATA AND CRIME TYPES

In order to discover the right insights and successful investigation, it is necessary to recognize available data sources of crime and the various types of crime[23].

A.   Data sources for crime data mining and analysis
1.   Police reports

Police reports such as FIRs (First Information Report) contain information about the crime, complainant and suspect.
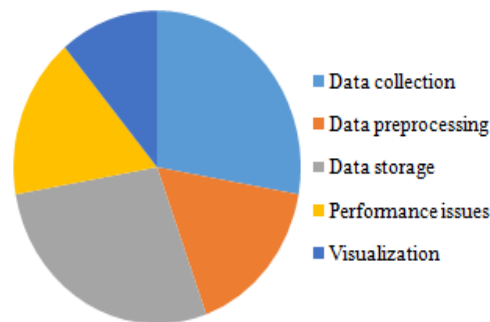
*Figure 2. Challenges addressed by research paper*s

FIRs are written by police staff on paper and have unstructured data format. These are one of the reliable sources for collecting crime data[].

2. Previous investigation files
a) In the wake of recognizing a formerly indicted suspect, the police request past examination records of the proposed suspect. These records come in content, photograph, video, CCTV video documents. financial balance, financial records, telephone call, email send-get records, measurable reports, witness and casualty explanations and legal counselor proclamations [25,26].
b) Intelligence reports
    Intelligence agencies maintain information about the criminals. Intelligence agencies of India are Research and Analysis (RAW), Intelligence Bureau (IB), Narcotics Control Bureau (NCB) [27,28].
c) Open source intelligence findings
    Open source intelligence findings are extracted from the web, search engines, social networking sites (Facebook, Twitter, LinkedIn). This information is in unstructured data format [29,30].
d) Police arrest records
    When police officers arrest suspects then their arrest records are maintained. These records are mainly in relational format or text format [31].

A. Type of crimes

Table V presents different crime categories classified by various law-enforcement agencies. Some type of crime, such as arson and traffic violations concern police at the city, state and national level [32,33]. Other crime types such as terrorism, cybercrimes are investigated by local agencies as well as national and international agencies [34]

### IV.   METHODOLOGY

Table V. Crime types and law enforcements

Crime type	Local law enforcement	National and enforcement
Traffic violations	Speeding, causing harm or damage in a crash amid affected by medications or liquor, attempt at manslaughter, neglectful driving	------------- ------------- ------
Sex crime	Tyke attack, sexual manhandle, assault, rape, kid explicit entertainment, prostitution	Human traffi cking, prosti tution, pornography
Theft	Robbery, burglary, motor vehicle theft	Fraud, corruption, trafficking in stolen software, music, movies
Arson	Intentionally setting fires to damage property	----------------- ---------------
Gang/drug offenses	Having, appropriating and offering illicit drugs	Drug trafficking, individuals carrying
Violent crime	Murder, hate crime, armed robbery	Terrorism, bombings
Cybercrime	Web extortion, fake sites, unlawful web based betting and exchanging, digital robbery, organize interruption and hacking, charge card and propel expense misrepresentation	

V.   PROBLEM FORMULATION AND PROPOSED SOLUTION

To do so, we applied the classification on the dataset. In which we used the decision tree algorithm and apriori algorithm that identified the data and classify the data accordingly.

Decision Tree Algorithm Pseudocode

1. Place the best characteristic of the dataset at the base of the tree.
2. Split the preparation set into subsets. Subsets ought to be made such that every subset contains information with a similar incentive for a property.
3. Repeat stage 1 and stage 2 on every subset until the point when you discover leaf hubs in all the branches of the tree.
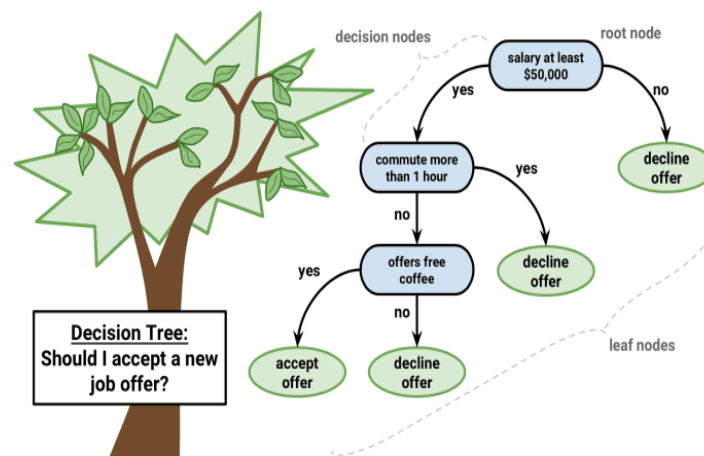
Fig 3: Decision Tree

VI.   EXPERIMENTS RESULTS

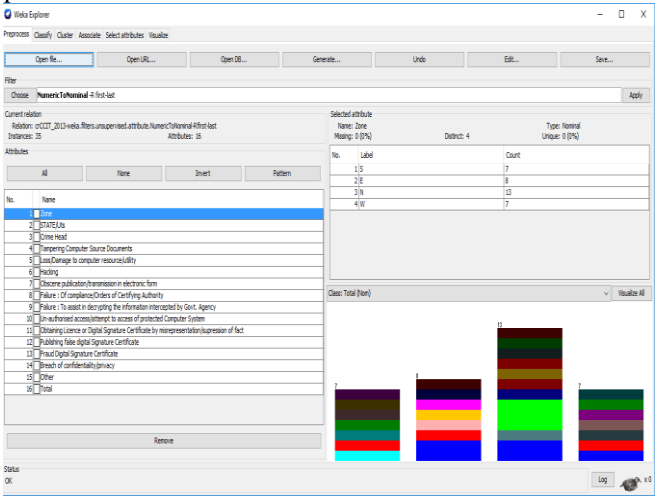To select a file we use preprocess tab

Figure 4 : Preprocess Tab

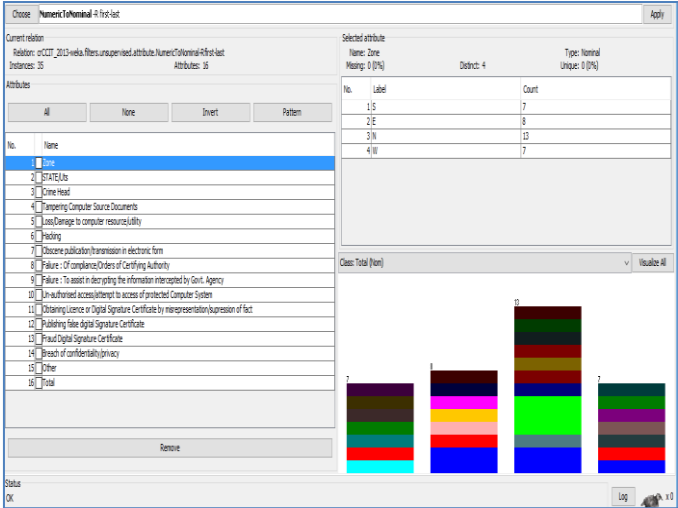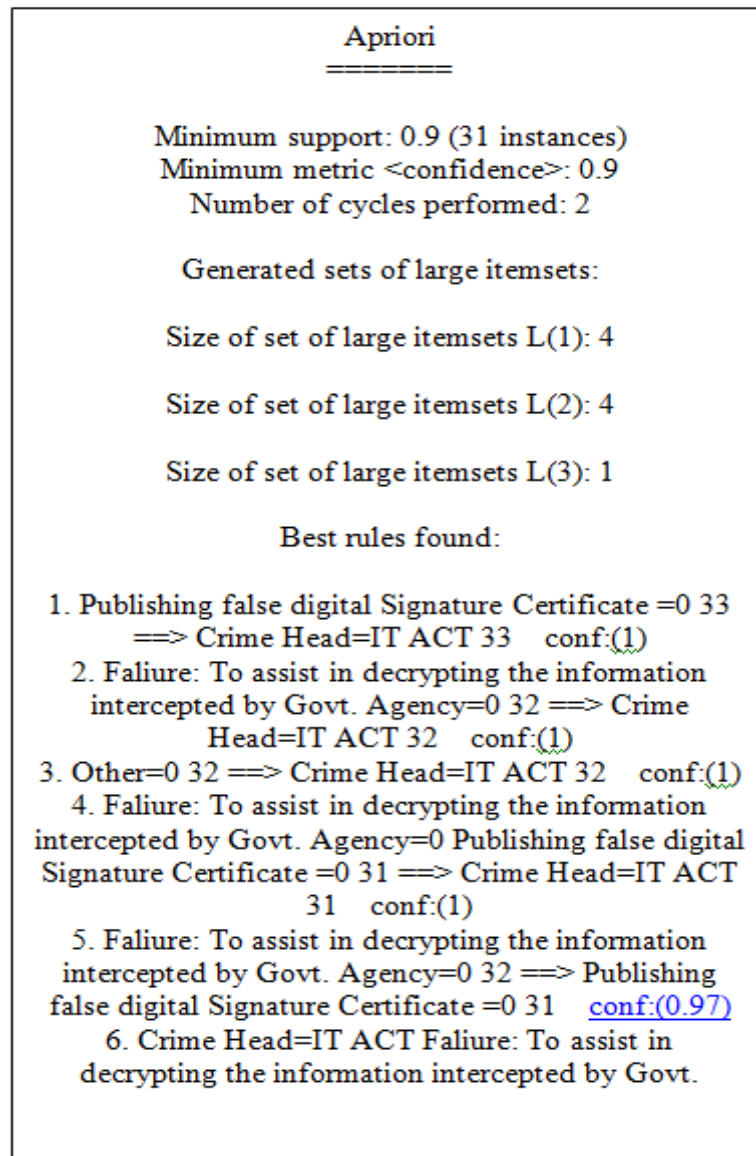No we have all the list of attributes and their values as shown below

Figure 5: List of attributes and their values

```
                    Apriori
                    ======

        Minimum support: 0.9 (31 instances)
        Minimum metric <confidence>: 0.9
        Number of cycles performed: 2

        Generated sets of large itemsets:

        Size of set of large itemsets L(1): 4

        Size of set of large itemsets L(2): 4

        Size of set of large itemsets L(3): 1

                Best rules found:

1. Publishing false digital Signature Certificate =0 33
        ==> Crime Head=IT ACT 33    conf:(1)
   2. Faliure: To assist in decrypting the information
    intercepted by Govt. Agency=0 32 ==> Crime
            Head=IT ACT 32    conf:(1)
3. Other=0 32 ==> Crime Head=IT ACT 32    conf:(1)
   4. Faliure: To assist in decrypting the information
intercepted by Govt. Agency=0 Publishing false digital
Signature Certificate =0 31 ==> Crime Head=IT ACT
                31    conf:(1)
   5. Faliure: To assist in decrypting the information
    intercepted by Govt. Agency=0 32 ==> Publishing
false digital Signature Certificate =0 31    conf:(0.97)
      6. Crime Head=IT ACT Faliure: To assist in
    decrypting the information intercepted by Govt.
```

Figure 2 Result of Apriori Algorithm

VII.   CONCLUSION

Crime Data information mining and examination is a dynamic territory of research. The consequences of this investigation may help new potential clients in understanding the scope of accessible wrongdoing information mining methods and advances. Crime information mining can be utilized to give entire crime measurements of a specific locale or territory that gives advantage to the general public by striking the administration and law implementation organizations to comprehend the different causes that expansion the crime rates. The administration and law requirement offices can take better choices for better living of the nationals that normally mean parcel of lives

VIII.   REFERENCES

[1]   Ubon Thongsatapornwatana, "A Survey of Data Mining Techniques for Analyzing Crime Patterns," in 2016 Second Asian Conference on Defence Technology (ACDT), IEEE, 2017.

[2]   Antoine Bagula, Sonia Berman Omowunmi Isafiadea, "A Revised Frequent Pattern Model for Crime Situation Recognition Based on Floor-Ceil Quartile Function," Procedia Computer Science, vol. 55, Elsevier, pp. 251-260, 2015.

[3]   Wingyan Chung,JenniferJie Xu,Gang Wang, Yi Qin,Michael Chau Hsinchun Chen, "Crime Data Mining: A General Framework and Some Examples," Computer, IEEE, pp. 50- 56, April 2014.

[4]   S. Srisuk P. Thongtae, "An Analysis of Data Mining Applications in Crime Domain," IEEE 8th International Conference on Computer and Information Technology Workshops, IEEE, pp. 122-126, 2014