CRIME DETECTION USING DATA MINING TECHNIQUES

Md. Sumon Rony , Sagor Chandra Bakchy and Hadisur Rahman

Department of Computer Engineering, Varendra University, Rajshahi, Bangladesh.

ABSTRACT

As crime rates keep spiraling each day, new challenges are faced by law enforcement agencies. They have to keep their on the lookout for any signs criminal activity. The law enforcement agencies should therefore be able to predict such increase or decrees or trends in crime. Such as theft, Killing. Crime that may occur in a particular area in a particular month, year, any timespan. Data mining is defined as a process of discovering hidden valuable knowledge by analyzing large amounts of data, which is stored in databases or data warehouse, using various data mining techniques such as machine learning, artificial intelligence, statistical. Many algorithms for data mining approach to help detect the crimes patterns. Data Collection, Data Preprocessing Phase, Data Filtering, Linier Regression. Wekasoft are used for collection of data analyzing. Visualization finally get results. The advantage of using this tool is that clustering will be performed automatically.

KEYWORDS

Crime patterns, Data mining, Decision tree, Wekasoft, Prediction

1. INTRODUCTION

Data mining is a process of useful information and patterns from huge data. It is also called as knowledge discovery process, knowledge mining from data, knowledge extraction or data /pattern analysis. This term originally referred to the algorithmic step in the data mining process, which initially was known as the Knowledge Discovery in Databases (KDD) process. Data mining is a process that takes data as input and outputs knowledge.

2. METHODOLOGY AND ALGORITHM

The following formatting rules must be followed strictly. This (.doc) document may be used as

2.1. Data Filtering

An abstract instance filter that assumes instances form time-series data and performs some merging of attribute values in the current instance with attribute values of some previous (or future) instance. An instance filter that adds a new attribute to the dataset [4].

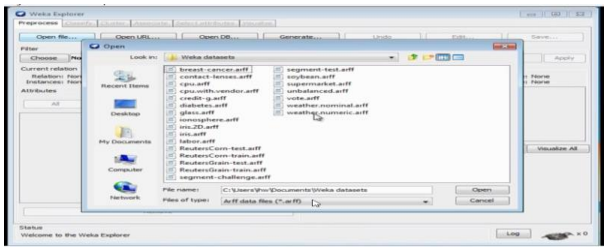


Fig.1. Data Filtering

2.2. Cluster Algorithm

We will show a simple clustering example here. Let us take an oversimplified case of crime record. A crime data analyst or detective will use a report based on this data sorted in different orders, usually the first sort will be on the most important characteristic based on the detective's experience.

Table 1. Collection of data

Crime Type	Suspect Race	Suspect Sex	Suspect Age	Victim Age	Weapon
Robbery	B	M	Middle	Elderly	Knife
Robbery	W	M	Young	Middle	Bat
Robbery	B	M	?	Elderly	Knife
Robbery	B	F	Middle	Young	Piston
Murder	W	M	Young	Middle	Knife

We look at table 1 with a simple example of crime list. The type of crime is robbery and it will be the most important attribute [4]. The rows 1 and 3 show a simple crime pattern where suspect description matches and victim profile is also similar.

2.3. Linier Regression

Numeric prediction is called regression**.** Regression is a data mining function that predicts a number. Age, weight, distance, temperature, income, or sales could all be predicted using regression techniques. For example, a regression model could be used to predict children's height, given their age, weight, and other factors [5]. The simplest and oldest form of regression is linear regression used to estimate a relationship between two variables.

$$X=W_0 +W_1A_1 + W_2A_2 + \cdots + W_kA_k$$

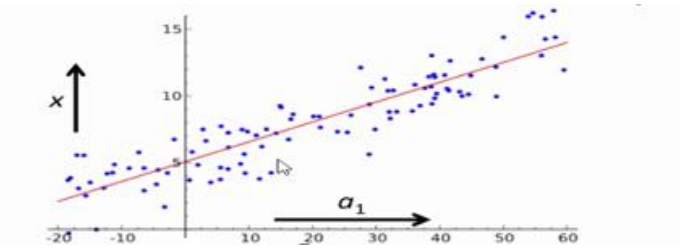


Fig.2. Regression Process

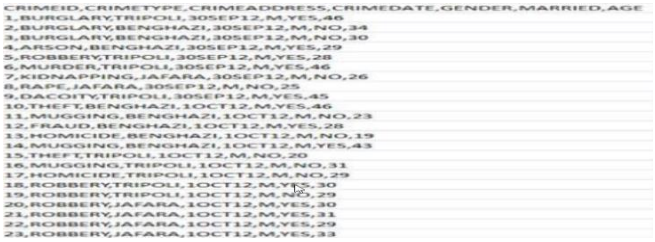
2.4. Terrorist-Behavior

Predicting terrorist attacks by group networks is an important but difficult issue in intelligence and security informatics. [14] According to our analysis, almost half (44 percent) of all terrorists examined lived within 30 miles of their targets. (See spatial analysis "Distance from Terrorist Residence to Target (All Groups).") When the types of terrorist groups are examined separately, however, the finding are so much different.

3. IMPLEMENTATION

3.1. Data Collection

It includes criminal offenses and crime incidents in the city and county. The exploratory factor analysis was also used to reduce the large number of related variables to a more efficient number to avoid redundancy. Since crime is increasing at an alarming rate globally it is important to control it. [7] In order to reduce crime rates we need to study the crime rates of various places of a country. In this research all the states and union territories crime rate (Fig) is studied in detail for different types of crimes.



```
CRIMEID,CRIMETYPE,CRIMEADDRESS,CRIMEDATE,GENDER,MARRIED,AGE
1,BURGLARY,TRIPOLI,30SEP12,M,YES,46
2,BURGLARY,BENGHAZI,30SEP12,M,NO,34
3,BURGLARY,BENGHAZI,30SEP12,M,NO,30
4,ARSON,BENGHAZI,30SEP12,M,YES,29
5,ROBBERY,TRIPOLI,30SEP12,M,YES,28
6,MURDER,TRIPOLI,30SEP12,M,YES,46
7,KIDNAPPING,JAFARA,30SEP12,M,NO,26
8,RAPE,JAFARA,30SEP12,M,NO,25
9,DACOITY,TRIPOLI,30SEP12,M,YES,45
10,THEFT,BENGHAZI,1OCT12,M,YES,46
11,MUGGING,BENGHAZI,1OCT12,M,NO,23
12,FRAUD,BENGHAZI,1OCT12,M,YES,28
13,HOMICIDE,BENGHAZI,1OCT12,M,NO,19
14,MUGGING,BENGHAZI,1OCT12,M,YES,43
15,THEFT,TRIPOLI,1OCT12,M,NO,20
16,MUGGING,TRIPOLI,1OCT12,M,NO,31
17,HOMICIDE,TRIPOLI,1OCT12,M,NO,29
18,ROBBERY,TRIPOLI,1OCT12,M,YES,30
19,ROBBERY,TRIPOLI,1OCT12,M,NO,29
20,ROBBERY,JAFARA,1OCT12,M,YES,30
21,ROBBERY,JAFARA,1OCT12,M,YES,31
22,ROBBERY,JAFARA,1OCT12,M,YES,29
23,ROBBERY,JAFARA,1OCT12,M,YES,33
```

Fig.3. Data Collection

3.2. Compare a Crime Period Chart

Comparing crime statistics between different jurisdictions can be misleading without sufficient knowledge about local data and methods. In this article, direct comparisons between jurisdictions have been avoided and the focus is instead on developments over time.[8] Such comparisons rely on the assumption that the characteristics of the national recording systems, etc. are fairly constant over time.

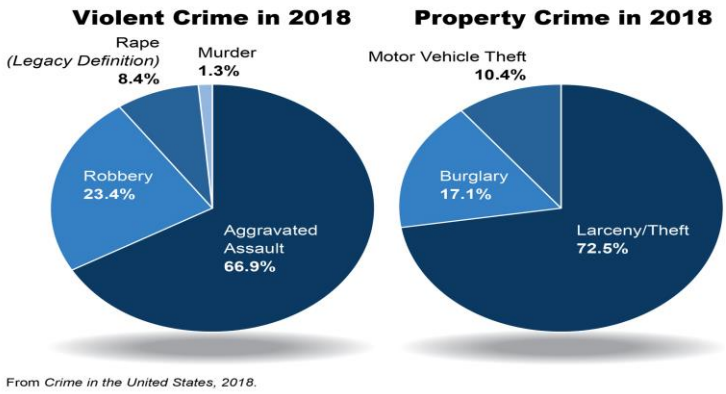


Fig.4. Comparative to Crime Period

Out of 23,883 total crime incidents recorded from January to November 2018, City Police Office recorded the highest number of crime incident at 14,328 or 59.99%, followed del Norte Police Provincial Office at 5,053 or 21.15%, del Sur Police Provincial Office at 3,783 or 15.83% while Police Provincial Office recorded the lowest crime incidents at 1,719 or 7.19%.

Data mining is a new discipline lying at the interface of statistics, database technology, pattern recognition, machine learning, and other areas. We looked at the use of data mining for identifying crime patterns crime pattern using the clustering techniques. Our modeling technique was able to identify the crime patterns from a large number of crimes making the job for crime detectives easier. Detected some feature for criminal identify system. Such as non-educated person, record all crimes, identify any person with criminal background in any event or group, Voice analysis, non-secure area etc.

4. FUTURE WORK

As a future extension of our work, we plan to apply more classification models to increase crime prediction accuracy and to enhance the overall performance. It is also a helpful extension for our study to consider the income information for neighbourhoods in order to see if there are relationships between neighbourhood's income level and their crime rate. Moreover, we intend to analyses Los Angeles demographics information with its crime findings. Furthermore, we want to study other crimes datasets from new cities along with their demographics datasets. Last but not least, we hope by publishing this paper starting a trend of crimes prediction, which can help law enforcements and keep our community safer for everyone.

REFERENCES

[1] Shakir Khan, Dr. Arun Sharma, Abu Sarwar Zamani, Ali Akhtar, "Data Mining For Security Purpose & Its Solitude Suggestions", University of Warsaw, Polna 40 00-635 Warsaw Poland International Journal of Engineering Research & Technology, Vol. 1, ISSUE 7, August 2012 ISSN 2277-8616.

[2] Mrs. Bharati M. Ramageri, "Data Mimning Techniques and Applications", "International Confeerence – 2008" , Modern Institute of Information Technology and Research, Yamunanagar, Nigdi Pune, Maharashtra, India-411044. Vol. 1 No. 4 301-305.

[3] N.Sudha Bhuvaneswari, S.Yamuna, "Datamining Techniques to Analyze and Predict Crimes", Dr.G.R.Damodaran College of science Coimbatore, Avinashi Road, Civil Aerodrome Post, Peelamedu, Coimbatore, Tamil Nadu 641014, India, The International Journal of Engineering And Science (IJES), ISSN: 2319 – 1813 ISBN: 2319 – 1805, Vol. 2-2016.

[4] Viraj Samant , Sindhu Nair , Vineet Pande , "Crime Detection using Data Mining ",D. J. Sanghvi College of Engg , Mumbai, Maharashtra, India, Published by : International Journal of Engineering Research & Technology (IJERT) http://www.ijert.org, ISSN: 2278-0181, Vol. 5 Issue 01, January-2016 .

[5] Levine, N., 2002. CrimeStat: A Spatial Statistic Program for the Analysis of Crime Incident Locations (v 2.0). Ned Levine & Associates, Houston, TX, and the National Institute of Justice, Washington, DC. May 2002.

[6] D. E. Brown, "The regional crime analysis program (RECAP): A framework for mining data to catch criminals," International Proceedings of IEEE International Conference on Systems and Cybernectics, pp. 2848- 2853, 1998.

[7] M. Strano, "A neural network applied to criminal psychological profiling: An Italian initiative,"International Journal of Offender Therapy and Comparative Criminology, vol. 48, no. 4, pp. 495-503,2004.

AUTHORS

Md. Sumon Rony
Bangladesh.
Cell: +880 1744352472
Email: sumoncse2395@gmail.com
Short Biography

Sagor Chandra Bakchy
Bangladesh
Cell- +8801776919609
Email: sagorchandro.10@gmail.com

Hadisur Rahman
Bangladesh
Cell- +880174969697
Email: hudacse6@gmail.com