

CRIME DATA ANALYSIS

¹Ganesh Kukreja, ²Siddharth Mav, ³Anand Vaswani, ⁴Prabhav Pathak, ⁵Dr. Mrs. Greesha Bhatia

¹Student, ²Student, ³Student, ⁴Student, ⁵Deputy HOD,

¹Computer Department

¹Vivekanand Education Society of Institute and Technology, Chembur, India

Abstract: Criminal Analysis is a systematic approach for identifying trends in a given data set and analyzing the trends to come to a conclusion. Given a data set of suspects, our system can give predict which suspect has the most chance of committing the particular crime, which means the police department will know on which suspect the most focus should be driven. This paper present the use of Data Mining to extract previously unknown and useful information from the provided data set. The proposed system identifies through its dataset and the algorithm developed, the most probable culprit. This prediction thus reduces the time complexities involved in narrowing down on the suspect and the culprit. This paper further evaluates the individual criminal statistics that would help the police forces to resolve the crime effectively.

Keywords— Crime, Prediction, Machine Learning model, Suspects.

I. Introduction

In today's world crimes are on the rise, misuse of technology and various socio economic factors have led to global increase of crimes.

Analysis of the crime scenario, the data obtained and determination of probable suspect is a challenge. The existing crime analysis system work on the traditional mechanisms, processes and methods.

The process of criminal identification in India is commonly executed using traditional method. The police stations use a database system to store the criminal information and access the database when criminal information is required.

Crime patterns are changing all the time and growing. The crime data previously stored from sources has a tendency to increase. The management of this data and analysis of the data is very difficult. Thus it can be said that crime data analysis is fundamental to effective prevention of crime. Knowing more information about the criminal will help significantly in the Analysis process [1]. To solve the above mentioned problems, data mining techniques and machine learning algorithms are needed to extract trends from the data [2].

However to improve traditional system, a system has been proposed that operates on utilizing criminal data and its analysis. The analysis basically means storing the criminal data in a particular format, analyzing the stored data to find criminal and deriving conclusions from the analyzed data. This would ensure law enforcement function which involves systematic analysis for identifying and analyzing criminal and its trends in crime and disorder [3].

II. Relevance Of The Proposed Work

The Proposed work is relevant to the Government Agencies Such as Police and other security personals. The Proposed work would Help the Officers to do the analysis on the Criminal and find the probability on the list of the criminal with highest probability to do the crime. This will save the time and can help to solve the case faster.

III. Literature Review

A thorough literature survey was performed to get a better understanding of the topic, models developed and noted their advantages and drawbacks, and highlight the necessary developments [3][4].

In Many of the models developed previously they are just focusing on the Geographic information science which helps them to Visualize the Number of the crime in a particular region on the maps [5] which can't we used for prevention of crime. Or we can say that there is no method/model to find out the probability of the criminal which can greatly help in reducing the crime by focusing on that criminal.

According to the International Association of Crime Analysis (IACA), "Crime analysis is a process in which a set of quantitative and qualitative techniques are used to analyze data useful to the police agencies and their communities. A wide amount of research has been done in the domain of crime data analysis [6].

According to IACA various types included are 'Crime Intelligence Analysis', 'Tactical Crime Analysis', 'Strategic Crime Analysis' and 'Administrative Crime Analysis' [7]. Based on the literature survey conducted it was observed that the existing models operate under GIS system while also needs to be involved in the crime on finding out the probability of suspects [8].

There are several web based crime mapping systems available on the internet however majority of them are custom made for legislative authorities in different countries and those systems are not accessible to parties outside the particular law enforcement authorities [9].

IV. Proposed Methodology

Various steps for performing Crime Data Analysis are :

- Data Collection
- Data Manipulation
- Data Analysis
- Graphical Representation.

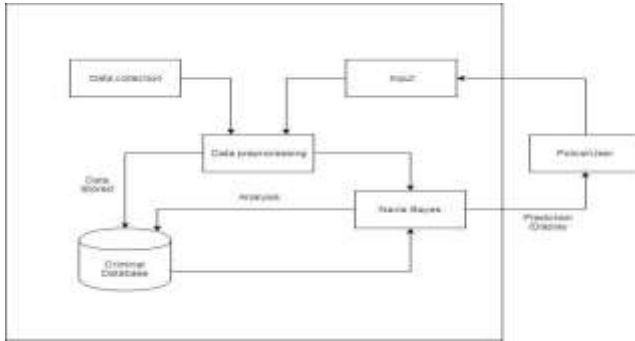


FIGURE:-System representation is overall view of the system.

Step 1:- Data Collection

In data collection step we are collecting data from different websites like blogs, news sites, social media, etc [10] . The data information collected comprises of name, address etc. And this collected data is stored into database for further process. Crime data is an structured data therefore the no of field, content, and size of the document will be same due to which it will greatly improve the efficiency .

Step 2:- Data Manipulation

After Data Collection we are performing Data manipulation where we are going to add the information of the criminal as the new entry which will then automatically will be stored in the database and will be used for future purpose as well as we can delete the data stored in it [11]. And Foremost we can sort & search the criminal data according the parameters or in the alphabetical order [12].

Step 3:- Data Analysis

For data analysis Naive Bayes which is a supervised learning method as well as a statistical method for classification is utilized. Naïve Bayes Algorithm is a most efficient probabilistic classifier [12]. Naive Bayes is used when we have to calculate the probability of event which depends on several independent conditions. It predicts a class values based on set attributes given as input [13][14].

The general formula of Naïve Bayes algorithm is given by: [15]

$$P(c | x) = \frac{P(x | c)P(c)}{P(x)}$$

Likelihood
Class Prior Probability
Posterior Probability
Predictor Prior Probability

$$P(c | X) = P(x_1 | c) \times P(x_2 | c) \times \dots \times P(x_n | c) \times P(c)$$

For the proposed system, Naive Bayes algorithm takes 3 inputs for each round, 2 out of 3 inputs will remain constant that is Crime Intensity, Crime type and the third input would vary according to gender, age, region, motive, method, modus operandi in different rounds and it will calculate the probability contribution of all input parameters. After that, the average of all these values

would be calculated and the system will recommend the suspect to concentrate first which will have maximum unit value after calculating the probability contribution of all parameters.

$$P(\text{Murder}/\text{Male,CI}) = \frac{P(\text{Murder})P(\text{Male,CI}/\text{Murder})}{P(\text{Male,CI})}$$

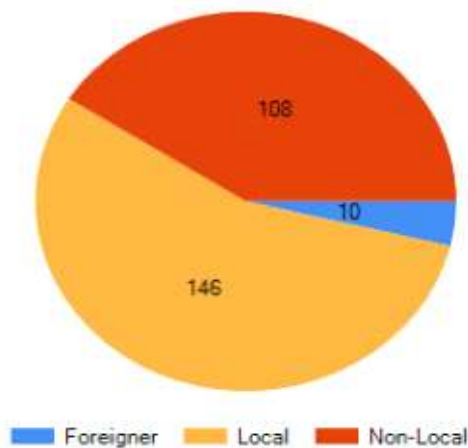
Parameters	Count	Total Dataset count	Probable Contribution
P(Murder)	05	21	0.23
P(Male,CI)	06	21	0.28
P(Male,CI/ Muder)	02	06	0.33
Total Contribution			0.2727

Table 1:Probability of suspect based on gender

Step 4:- Graphical Representation

In graphical representation plot of individual criminal statistics which is represented with respect to its parameters wise contribution. These are represented as:-

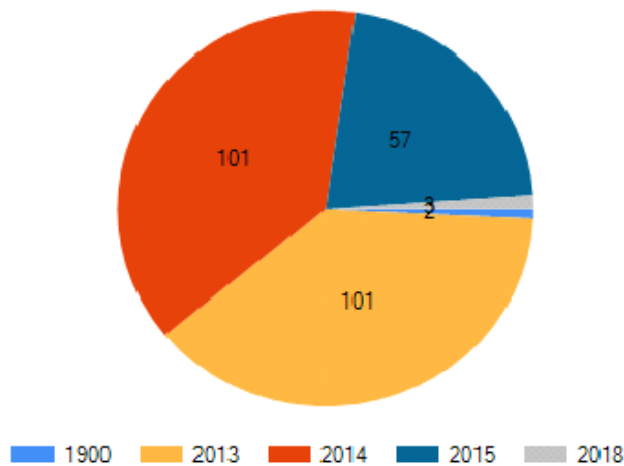
- Region wise representation.
- Year representation.
- Gender wise representation.



- Graph 1:-

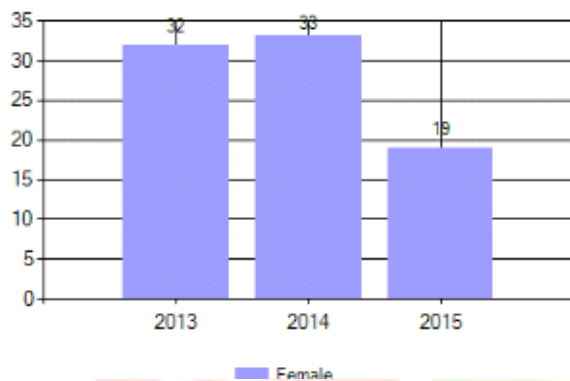
Region wise representation

This graph represents the region-wise distribution of criminals such as local, non-local, foreigner i.e it shows how many criminals are present in a particular region



Graph 2:-Year wise representation.

It represents the concentration of crimes according to different years till the year it is stored in database.



Graph 3:-Gender wise representation.

It represents the count of female criminals year wise.

V. Conclusion

Looking at the difficulties posed by the security forces towards handling the various crimes, the proposed work was carried out. The system aims to apply Machine Learning algorithm (ML) to help in on the suspect and thus aid the police force and other security personals. The proposed system is capable to enhance the accuracy, performance, speed of predicting the crime

VI. References

- 1) KS: Author",Definition and Types of Crime Analysis"International Association of Criminalists. (2014). Definition and types of crime analysis [White Paper 2014-02].
https://www.iaca.net/Publications/Whitepapers/iacawp_2014_02_definition_types_crime_analysis.pdf
- 2).Ubon Thongsatapornwatana, "A survey of data mining techniques for analyzing crime patterns", *Defence Technology (ACDT) 2016 Second Asian Conference on*, pp. 123-128, 2016.
- 3).S. Sathyadevan, M. Devan, and S. Surya Gangadharan, "Crime analysis and prediction using data mining," in *Networks Soft Computing (ICNSC), 2014 First International Conference on*, Aug 2014, pp. 406– 412.
- 4)..Isuru Jayaweera, Chamath Sajeewa, Sampath Liyanage, Tharindu Wijewardane, Indika Perera "Crime Analytics: Analysis of Crimes Through Newspaper Articles,"*Moratuwa Engineering Research Conference (MERCOn), 2015*
- 5).D.Usha, Dr.K.Rameshkumar,"A Complete Survey on application of Frequent Pattern Mining and Association Rule Mining on Crime Pattern Mining,"*Volume 3, No.4, April 2014 International Journal of Advances in Computer Science and Technology*
- 6)*International Association of Crime Analysts - Wikipedia*
https://en.wikipedia.org/wiki/International_Association_of_Crime_Analysts.
- 7).P. Thongthai and S. Srisuk,"An Analysis of Data Mining Applications in Crime Domain ," *Computer and Information Technology Workshops, 2008. CIT Workshops 2008. IEEE 8th International Conference on Computer and Information Technology Workshops*

- 8).Chung-Hsien Yu 1 , Max W. Ward1 , Melissa Morabito 2 , and Wei Ding1,"Crime Forecasting Using Data Mining Techniques ."
2011 11th IEEE International Conference on Data Mining Workshops
- 9).Manish Gupta, B. Chandra and M. P. Gupta ,"Crime Data Mining for Indian Police Information System."
Indian Institute of Technology Delhi,Hauz Khas, New Delhi. India 110 016.
- 10).Revathy Krishnamurthy,J. Sathesh Kumar", Survey of data mining techniques on crime data analysis,"Vol 01, Issue 02, December 2012 International Journal of Data Mining Techniques and Applications ISSN: 2278-2419
- 11).Shyam Varan Nath ,"Crime Pattern Detection Using Data Mining ," 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT 2006 Workshops)(WI-IATW 06)
- 12).Dr. A.Bharathi,R. Shilpa,"A Survey on Crime Data Analysis of Data Mining Using Clustering Techniques,"Volume 2, Issue 8, August 2014 International Journal of Advance Research in Computer Science and Management Studies ISSN: 232 7782
- 13).George He Shuxian, Varun Ponnuru, Eric Chow,"Crime Data Analysis using Data Mining and Data Mining and Business Analytics,"

https://sis.smu.edu.sg/sites/default/files/sis/pdf/is424_2010_crime_data_analysis.pdf

14) Naive Bayesian

http://www.saedsayad.com/naive_bayesian.htm

15) Characteristics of Structured Data

https://www.webopedia.com/TERM/S/structured_data.html

