

# Crimerate Prediction using Datamining

Milind Nayak<sup>1</sup>, Vikas Yadav<sup>2</sup>, Yamini Patil<sup>3</sup>

<sup>1,2</sup>Student of Information Technology Department, KC College of Engineering, Thane, Maharashtra, India

<sup>3</sup>Faculty of Information Technology Department, KC College of Engineering, Thane, Maharashtra, India

**Abstract**—Crime is a social stimulus and in many ways our society suffers. Crime is one of the most important and traumatic factors of our society, and prevention is an important function. Any research that will help you solve crimes will soon be paid. This system was influenced by the Criminal Record of the Indian Online Portal last year, which listed various crimes, such as murder, kidnapping, robbery, rape and other crimes. Crime analysis is a systematic method of investigating and investigating crime models and trends. This job helps local police station to suppress crime. With the rise of computerized systems, criminal data analysts can help law enforcement officials to speed up the crime-solving process. Use data mining concept. Due to increasing crime rates over the years, we have to deal with a large number of criminal data stored on the warehouse, which is difficult to manually analyze, and now one day, criminals are making technological advances, so advanced technology is required to keep police. They are. The main purpose of examining the algorithms and skills to identify criminals in this article.

**Keywords**—Data Mining, Predictive analysis, YouTube, twitter, IMDb.

## I. INTRODUCTION

Crime is a violation of humanity, often punishable by law. Criminology is a study of crime, interdisciplinary science that investigates and investigates crime and criminal performance data. Criminal activity is now high and the police department is responsible for controlling and reducing criminal activity. Crime prospects and criminal identity elements are the major issues facing the police department. Crime rates are increasing every day in many countries. Today, with such high crime rates and cruel offenses, there must be some protective measures to prevent such crimes. Here, we introduce a system that reduces crime rates. Crime data must be entered into the system. We have introduced data mining algorithms to guess the crime. This system assumes the crimes that occur in society. Analyze the criminal data stored on the database. The data mining algorithm receives information from the database. Administrators enter into a system that requires criminal details to be forecast. Administrators can view criminal history data. The fate of crime events is mainly based on historical criminal records and current records and demographic information.

### Data Mining

The process is to find different models and algorithms in large data sets to predict the results. Data mining is a multi-functional sub-field of information technology, whose goal is

to extract information from the database and convert it into a coherent structure for further use. Previously, we've seen a lot of data each year. Only 90% of the digital world is chaotic data accounts. Larger data analytics help us move uncomfortable and repetitive sounds from large data sets. And understand more important data to do more and help speedup the development of Enlightenment Decisions. There square measure 2 types of information analysis which will be used for extracting models describing necessary categories or to predict future information trends. These are as follows:

- Classification
- Prediction

The classification model assumes the classification label and the future model assumes the constant value function. For example, the classification model that categorizes bank loan applications as safe or risk and designs the forecasting pattern to predict the customer's monthly budget based on consumer income and career.

### Classification

It is a data processing technique that assigns data to a category to assist with more accurate predictions and analysis. For larger data like the world today, more data is being produced. Database with multiple types of data - Terabytes are a trillion bytes data. Think about Facebook itself in just a day and create 600TB new data. To process this data, some automated modes are separated to develop useful information and classification is one of them.

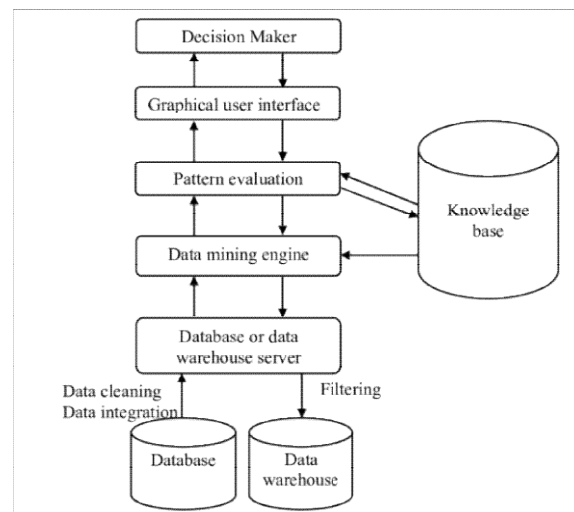


Fig1. Big data Architecture.

### Prediction

In data mining, it is used to strictly identify information points on the outline of contact information. It's nothing for future events, but the variables used are unknown. The forecast will result in the connection between the factors you identify and the elements you want to predict for future reference. For example, the forecasting model in the data mining area is used by the businessman, who speculates how much a particular customer can pay for the full purchase and is therefore planning a forthcoming sales volume. The predictions in data processing are considered a numerical prophecy. Typically, multidimensional analysis is used for the future. For data futures, some algorithms are used, such as C4.5, k-devices, support vector machines, upper, EM, PageRank, adburst, KNN, navigate and cart.

## II. LITERATURE REVIEW

In paper [1] We have studied data mining use to identify patterns of crime models using clustering techniques. This means developing crime model detection as a machine learning task, which uses data mining to support police detectives to address crime issues. We recognize the key features; We use expert-based semi-monitoring learning techniques and have developed a plan for inserting key features. Our sample technology identifies the crime patterns of a large number of crimes and facilitates the detection of criminal detectives.

In paper [2] Crime Analysis is a method of identifying and analyzing crime patterns and trends. With the rise of computerized systems, criminal data analysts can help law enforcement officials to speed up the crime-solving process. Using the data mining concept, we can review useful information that we did not previously know from data we did not create. Predictive pursuits use analytics and forecasting tools to identify crime, and it has been found to be very effective in managing the same operation.

In paper [3] The crime rate is increasing and crime patterns are constantly changing. Therefore, it is difficult to explain the behavior of a criminal model. This article explains how social development can lead to crime prevention. Its goal is to conduct a comprehensive review of the different data analysis algorithms of social crime prevention and research to address the connection between crime and its type. The project data comes from legitimate government sources. Data is converted to .csv format, which is pre-placing data.

In paper [4] Mugdha Sharma and others. The advanced ID3 algorithm is proposed to present the importance of significance-attribute on characteristics with fewer values but more importance than the values and low importance characteristics and the attributes with higher values to solve classification errors to choose. Experimental data analysis shows that the advanced ID3 algorithm has more reasonable and more effective classification rules. In this Z-crime device,

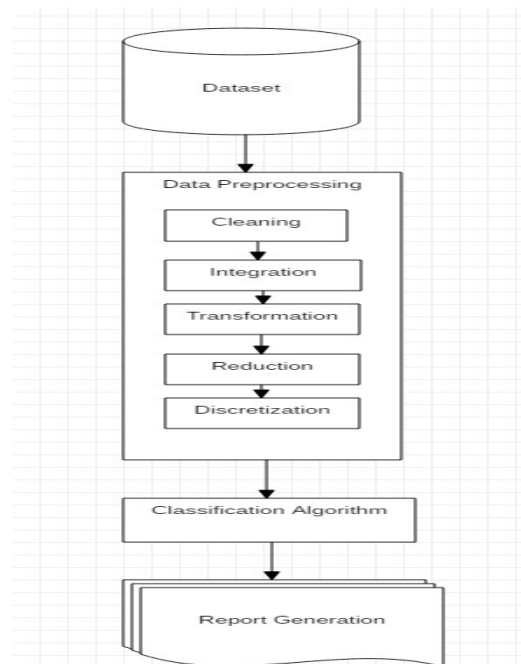
it is also suggested to analyze criminal activities through email communication.

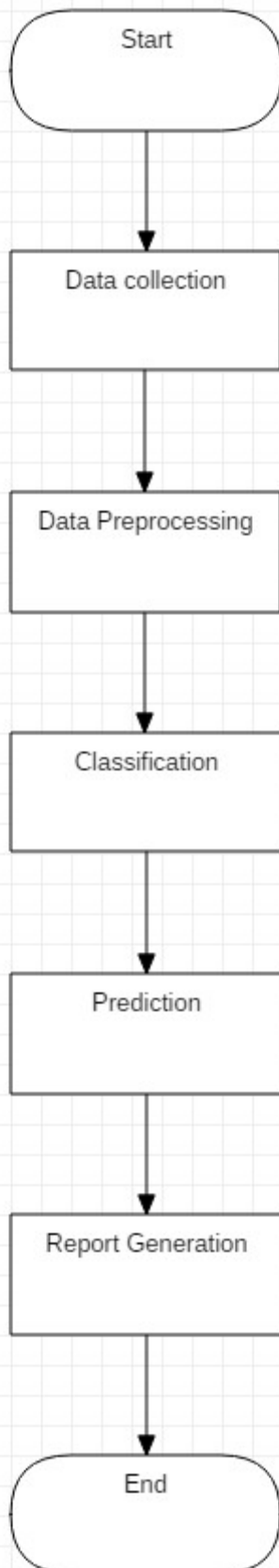
In paper [5] Sushant Bharathi and others. Hidden link algorithm has been proposed to find hidden links to network of collaborators that can potentially show potential future criminal partners and different networks of the real network. This paper also analyzes the node's focal point. This analysis explains the importance of network nodes. It is used to find powerful people, the power of people and the role of people in the network. This paper is a predictive method of criminal analysis, which helps prevent crime before crime, analyzes the network of Indian companions and predicts the network of existing criminals.

In paper [6] ShijuSathyadevan et al. Apriori algorithm has been proposed to identify crime trends and patterns. This algorithm is used to determine the association rules that highlight the general trend in the database. This paper also refers to the naive Bayesian algorithm to create a model by training Crime data. After the test, the results show the accuracy of the naive base algorithm is 90%.

In paper [7] Prashant K. Khobragadeand so on. The proposed forensic tool kit provides 4.0 remote data surveys and visual analytics. In remote data, surveys include analysis of process information, service information, driver information, network tools, and network information. The tool produces files and analyzes the data. This tool is used to analyze the victim's system of attacks. Physical and logical reminiscences are analysed with the help of criminal investigations.

## III. DIAGRAM





#### IV. CONCLUSION

In this report, we define criminal detection as a dynamic and emerging field of research in real-world environments designed to prevent crime rates. Data mining plays an important role in law enforcement agencies for crime detection and prevention of crime. Classification techniques are used for crime futures. A good-looking model helps criminal investigations by reducing crime and implementing various necessary steps to reduce criminals.

#### REFERENCES

- [1]. Dr. Zakaria Suliman Zubi , Ayman Altaher Mahmud, Using Datamining Techniques to Analyse Crime patterns in the Libyan National Crime Data, Recent advances in image, audio and signal processing. ISBN: 978-960-474- 350-6
- [2]. Shiju Sathyadeven, Deven M.S, Surya Gangadhara. S, Crime Analysis and prediction using data mining, IEEE, 2014
- [3]. Ubon Thansatpornwatana, A Survey of Data Mining Techniques for Analyzing Crime.
- [4]. Mugdha Sharma. "Z - CRIME: A Data Mining Tool for the Detection of Suspicious CriminalActivities Based on Decision Tree", International Conference On Data Mining and Intellegent
- [5]. HarshTaneja1, Anupam Dewan2, Vineet Bhardhwaj3, "Pre-Release Success Quotient Prediction of Movies", International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064 Index Copernicus Value (2013): 6.14 | Impact Factor (2015): 6.391
- [6]. Babita M. Jangid1 , Chaitali K. Jadhav2 , Swati M. Dhokate3 , Grish M.Jadhav4, Prof. G.M.Bhandari5," Survey on Movies Popularity Prediction System Using Social Media Feature", (An ISO 3297: 2007 Certified Organization) Vol. 4, Issue 9, September 2016.
- [7]. Prashant K. Khobragade and Latesh G. Malik. "Data Generation and Analysis for Digital Forensic Application using Data mining", Fourth International Conference on Communication Systems and Network Technologies, 2014.