International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 11, Issue 07, July 2022)

Machine Learning for Detecting Criminal Patterns

Nikita Jha[1], Prof. Vishal Paranjape[2], Prof. Saurabh Sharma[3], Prof. Zohaib Hasan[4]
[1,2,3,4]*Global Nature Care Sangathan Group of Institutions, Jabalpur (M.P), India*

Abstract:-- **Data mining is a technique for evaluating data that has been utilised in recent years to examine criminal data that had been previously stored from various sources to uncover patterns and trends in crimes. Additionally, it can be used to automatically inform of crimes and boost efficiency in solving crimes more quickly. There are numerous data mining methods, though. It is important to choose the right data mining techniques in order to boost the effectiveness of crime detection.**

Keywords:-- **ML, SVM, DT, RF**

I. INTRODUCTION

Crime detection and prevention are now major trends in crime and extremely difficult cases to solve. Numerous investigations have uncovered a variety of methods for solving crimes that had several applications. Such research can aid in accelerating the investigation of crimes and aid automated systems in automatically identifying criminals. Additionally, these challenges can be addressed with the aid of the quickly developing technologies. However, crime trends are expanding and changing all the time. The crime statistics that have been previously stored from different sources tend to rise constantly. As a result, managing and analysing vast amounts of data is extremely challenging and complex. Data mining approaches use numerous learning algorithms to extract hidden knowledge from massive amounts of data in order to address the issues outlined above. Data mining is the process of examining data to discover patterns and trends in criminal activity. It can aid in the quicker resolution of crimes as well as the automatic alerting of criminal detection. The non-trivial extraction of implicit, previously unidentified, and potentially relevant information from database data is referred to as data mining, also known as knowledge discovery in databases (KDD). Data mining is a step in the knowledge discovery process, despite the fact that the terms knowledge discovery in databases (or KDD) and data mining are frequently used interchangeably.

II. LITERATURE SURVEY

An Apriori algorithm was used in a straightforward experiment to see if any meaningful patterns could be found in a dataset of bicycle thefts reported to the police. Specific attributes related to modus operandi were included in our model, in contrast to previous research that has concentrated solely on the spatial and temporal aspects of crime pattern recognition using cluster or classification models, in order to uncover potentially useful rules for the distribution of police resources. The results imply that police-recorded data can be used for criminal pattern analysis to identify attributes other than time and place, in this instance modus operandi. [1]

We examined how data mining could be used to discover criminal patterns using clustering algorithms. Our contribution in this case was to structure the task of detecting criminal patterns as a machine learning task and use data mining to assist police detectives in their investigation of crimes. We created a strategy for weighing the significant features after applying an expert-based semi-supervised learning method to identify the significant attributes. The work of crime investigators was made easier by our modelling technique's ability to identify crime patterns from a huge number of crimes. [2]

The project's largest challenge was gathering and staging the data. To allow for the most efficient use of police resources, the scope of crime detection and analysis will soon be expanded to include the generation of crime hotspots. These hotspots will assist in the deployment of police to the areas where crime is most likely to occur during any given time window. The created model will lessen crimes and assist the field of crime detection in many ways, from making arrests of criminals to lessening crimes by doing various necessary actions. [3]

Three findings from our tests demonstrate the benefit of utilising the spatial information present in the crime data set. The success of the straightforward 1NN classifier updated with a location constraint is the first and most noticeable. Finding the situation that was the closest match within the same area instead of the entire city proved to be more successful.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 11, Issue 07, July 2022)

Success and stability of the probability-based Naive Bayes classifier serve as the second indicator. The location constrained 1NN operates on the same fundamental premise as Naive Bayes, namely that events that have occurred at a specific location in the past are likely to occur again. The 24-by-20 grid data is the third finding that suggests spatial awareness. [4]

The application of data mining to law enforcement and intelligence analysis holds the potential to reduce crime-related issues. It is feasible to find relevant data to aid in crime matching, both for single crimes and for series of crimes, using a variety of strategies. In order to predict crime trends, we employ a clustering/classify-based methodology in this paper. The analysis of crime data from databases is done using data mining techniques. The outcomes of this data mining may be used to reduce and even stop crime in the upcoming years. Crime data mining, in our opinion, holds great promise for improving the efficacy and efficiency of criminal and intelligence investigation. [5]

We used data mining techniques to identify the Denial of Service assault in this article. Because it puts the IT resources at peril, this kind of attack is extremely dangerous. By sending phoney messages and making repeated requests, it keeps the server busy. To reduce server performance, the server is overloaded with traffic packets. We have covered cyber security, various forms of cybercrime, clustering, outliers, and pattern recognition in this study article. On the log file, we used the well-known data mining method known as pattern recognition. We established a cutoff point. We consider this to be an attack and notify the administrator if the number of similar requests received at the server exceeds the threshold value. This method makes it simple to recognise a denial of service assault since, in a DoS attack, the hacker or attacker sends identical requests repeatedly to slow down the server. [6]

III. System Architecture

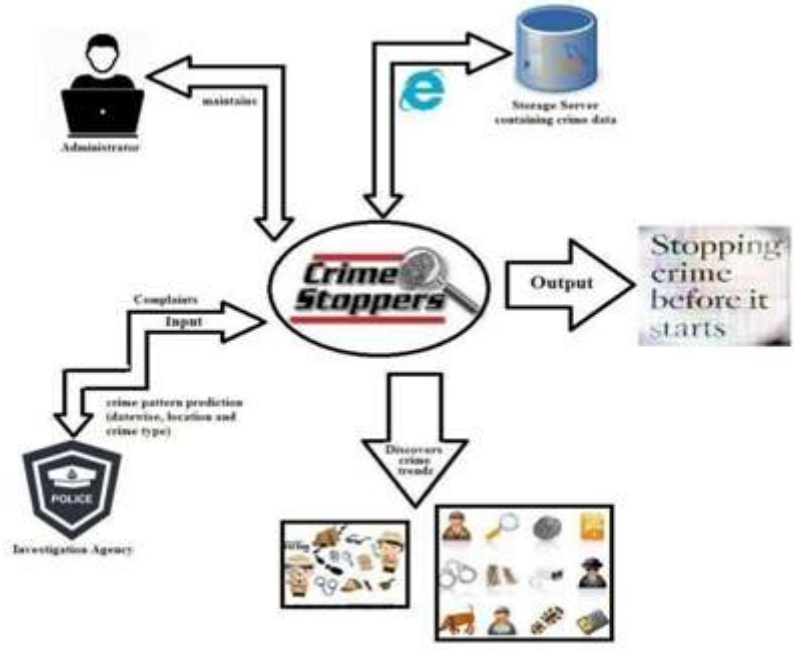

Figure1: System Architecture

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 11, Issue 07, July 2022)

Administrator: Administrator is a one who maintains the entire Crime system. Administrator has the full accessibility ofthe application (Crime system).

Investigation Agencies (Police Stations): Investigation Agencies are registered users. Investigation Agency is a one who receives the services from the application. The key service given by this Crime system is "prediction of Crime Trends" based on the previous crime data.

Public: Public is a one who visits the application, Public has only the limited accessibility.

Modules of the Project: Administrator has the following basic modules

Login Module: In this module, administrator of the application gets login to the application by inputting the credentials such as login id and password which is set in the server.

Manage Cities: In this module administrator manages cities by uploading cities into server. Managing of the cities means adding the new city, editing, updating and deleting the existing cities.

Manage Areas: In this module administrator of the system will add the areas based on the city and can edit, update and delete the existing areas.

Manage Investigation Agencies (IA): In this module, administrator registers the investigation agencies and can edit, update and delete the existing investigation agencies.

*Set Id and Password of IA:*Here administrator sets the unique id and password for each investigation agency. Using these credentials IA access the system.

IV. PROPOSED WORK

Proposed system is applicable in the field of crime. Proposed system includes modeling of crimes for finding suitable algorithms to detect the crime, precise detection, data preparation and transformation, and processing time. Proposed system identifies crime behavior, crime predicting, precise detection, and managing large volumes of data obtained from various sources. Proposed system is an automation for complaints registration, crime pattern prediction based on the previous crime details collected from various sources.

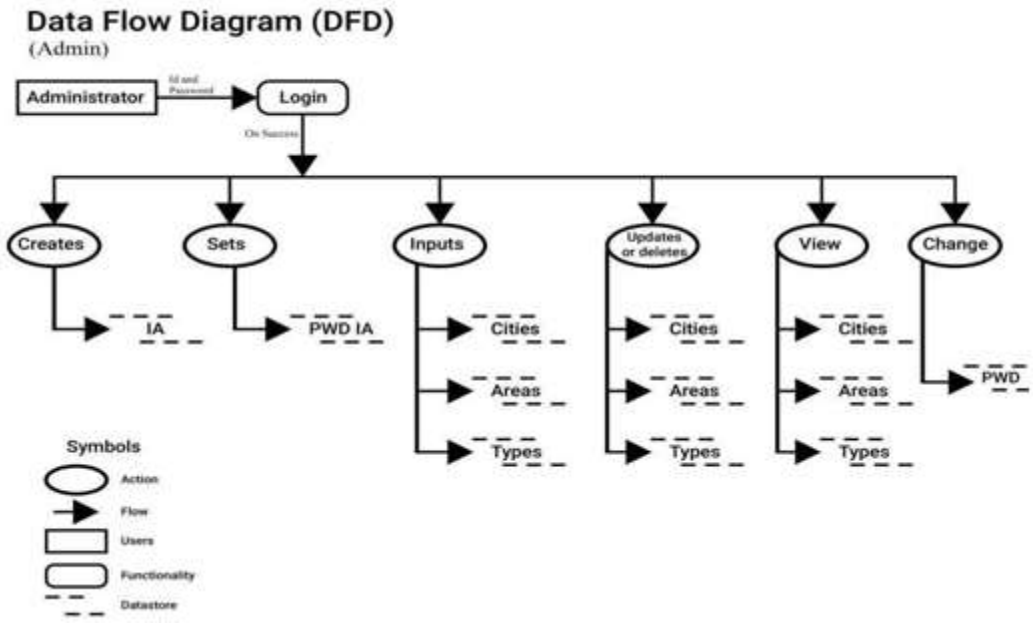

Figure 2: Diagram of Administrator Use case

Apriori algorithm is used to help prune the candidate explored during frequent item set generation to reduce the processing time apriori algorithm needs to scan the all items sets. So it uses a long period of time as well.

Reference proposed the improved apriori algorithm by using the compressed data set algorithm for association rule mining to reduce the amount of time needed to read data from the database. This a novel algorithm will delete the transaction that not contained in interesting item set.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435(Online) Volume 11, Issue 07, July 2022)

Apriori designed and developed by has improved apriori algorithm to find the effective associationrule and to reduce the amount of processing time. Additionally, there are several techniques that have been developed in order to analyze associations between two item sets effectively such as mutual information concept association bundle, audio watermarking etc.

Association rule mining is used to uncover closely related item sets in transactions for deciding business policies. Apriori algorithm is widely adopted is association rule mining for generating closely related item sets. Traditional apriori algorithm is space and time consuming since it requires repeated scanning of whole transaction database. In this paper we propose improved apriori algorithm based on compressed transaction database. Transaction database is compressed based on the consequence of interest.

Apriori Algorithm

STEP 1: Scan the opinion data set and determine the support(s) of each item.

STEP 2: Generate L1 (Frequent one item set).

STEP 3: Use Lk-1, join Lk-1 to generate the set of candidate k - item set.

STEP 4: Scan the candidate k item set and generate the support of each candidate k – item set.

STEP 5: Add to frequent item set, until C=Null Set.

STEP 6: For each item in the frequent item set generate all non empty subsets.

STEP 7: For each non empty subset determine the confidence. If confidence is greater than or equal to this specifiedconfidence .Then add to Strong Association Rule.

Snap Shots

Figure 3: Login page

Figure 4: Output page

V. CONCLUSION

Crime is defined by a constant rise and change over time. The difficulties of comprehending criminal behaviour, crime prediction, accurate detection, and managing substantial amounts of data gathered from numerous sources have arisen as a result of crime's changing and growing prevalence. These problems have been addressed via research interests. The accuracy of crime detection is still lacking despite these studies. The difficulties in the field of criminal detection result from this. The difficulties include modelling criminal activity in order to identify appropriate algorithms for crime detection, accurate detection, data preparation and transformation, and processing time.

REFERENCES

[1] Peng Chen, Justin Kurland,"Time place and Nods operandi:A simple Apriori Algorithm Experiment for Crime PatternDetection", 9th International Conference on Information, Intelligence, Systems and Applications (IISA),2018.

[2] ,Shyam Varan Nath , "Crime pattern Detection Using Data Mining", International conference on web intelligence and intelligent agent technology,International Conference on Current Trends towards Converging Technologies,2018

[3] Sunil yadav, Meet tmbadia, Amit yadav,rohit vishwakarma and nikhilesh yadav,"Crime pattern detection,analysis and prediction" ,international conference on electronics communication and aerospace technology ICECA 2017

[4] Chung-Hsien Yu1, Max W. Ward1, Melissa Morabito2, and Wei Ding11Department of Computer Science, 2Department of Sociology,"Crime Forecasting Using Data Mining Techniques", University of Massachusetts Boston, 100 Morrissey Blvd., Boston, MA 02125.

[5] S.Yamuna,N.Sudha Bhuvaneswari,"Datamining Techniques to Analyze and Predict Crimes" The International Journal of Engineering And Science (IJES)||Volume|| 1 ||Issue|| 2 ||Pages|| 243-247 ||2012||ISSN: 2319 – 1813 ISBN: 2319 – 1805www.theijes.com The IJES Page 243 Datamining Tech.

[6] Mohiuddin Ali Khan1, Sateesh Kumar Pradhan2, Huda Fatima ,"Applying Data Mining Techniques in Cyber Crimes" 2nd International Conference on Anti-Cyber Crimes (ICACC)978-1-5090-5814-3/17 2017 IEEE

[7] K. Chitra Lekh Dr. S. Prakasam,"Data Mining Techniques in Detecting and Predicting Cyber Crimes in Banking Sector" International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS-2017)

[8] Nidhi Tomar (Research Scholar) Amit Kumar Manjhvar (Asst.prof),"An Improved Optimized Clustering Technique For Crime Detection" Symposium on Colossal Data Analysis and Networking (CDAN),2016.

[9] , Lalitha Saroja Thota, Suresh Babu Changalasetty,"Cluster based Zoning of Crime Info", 2nd International Conferenceon Anti-Cyber Crimes (ICACC)978-1- 5090-5814-3/17 2017 IEEE