International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

CRIME PATTERN PREDICTION USING ML ALGORITHMS

Rakshitha H M[1], Rajesh R[2], Ruchitha K N[3], Vishal V Naik[4], Dr.Arun Nagarle[5]

[1]Rakshitha H M	[2]Rajesh R	[3]Ruchitha K N	[4]Vishal V Naik	[5]Dr.Arun Nagarle
Student, B.E IV Year, I. S. E	Student, B.E IV Year, I. S. E	Student, B.E IV Year, I. S. E	Student, B.E IV Year, I. S. E	Professor & HOD, I. S.E

Department of Information Science and Engineering Vidya Vikas Institute of Engineering and Technology [VVIET]
Mysuru, Karnataka, India

Abstract

Today the crimes are increasing rapidly that too in a high population countries. So we require automated application to detect crime and prevention. The system should stop crimes before it begins. As crimes are on the rise, preventive measures should be taken to stop the occurring of crimes. Finding frequent crimes and related correlation is tiresome and a high-effort task in the current crime sector. machine learning or AI, is an emerging technology to solve this problem. Previously effective unsupervised learning algorithms datasets for procedural crime training and frequent crimes and their relationships are identified [5]. Designed system created to reduce crime and applicable to crime sector. We build an application anywhere it is useful for criminal sectors to reduce crime. Currently, none of the applications in criminal activity and we use effective data science algorithms to predict crimes with better results.

Keywords: *Data Science, Machine Learning, Apriori Algorithm, Apriori TID Algorithm, Crime Types, Patterns*

1.Introduction

there are different categories of crimes such as pickpocketing, murder, kidnapping and chain snatching. Several works related to crime have proposed different techniques to solve crimes that have used too many applications [3]. they are increasing more and more in today's life. All cities and around the world face these crime problems. The type of crime depends on many factors such as location, population, school zone, hospital zones. It is necessary to identify frequent crimes and take frequent measures by creating an application for the crime sector to find frequent crimes in the area and also predict the relationship between different types of crimes. We apply efficient ML algorithms to process crime data and make a prediction. Data Science is a trending technology this helps in solving different types of crimes in all types of fields [1]. In this thesis, data science applied to the crime sector to process old crime data or crime data from previous years and to find common types of crime and their correlations.

machine learning is mainly preferred to train the system using training datasets. Here the dataset named crime dataset is taken, machine learning algorithms used to process crime datasets and extract hidden crime patterns. Unsupervised learning used to train the system with the expected results. unsupervised learning algorithms such as Apriori algorithm, Apriori TID algorithms used to process crime and frequent crime datasets and their predicted relationships. We are developing a web application to predict frequent crimes and crime patterns using ML algorithms. We use an apriori algorithm to predict frequent crimes and their relationships and compare the apriori algorithm with an apriori tid algorithm. Our system identifies crimes and related correlations and provides solutions to reduce crime and increase public vigilance and activity. The system finds correlations between different categories of crimes such as murder, chain snatching, kidnapping, pickpocketing, robbery.

2.Literature Reviews

The venture, titled "Detection of Crime Patterns Using a Simple A priori Algorithm", mainly focused on a crime data set and extracting data based on a specific crime area with the type of crime and the date when these parameters occurred help in creating frequent data curves

aims on the apriori algorithm to process crime data and predict crime patterns. Data science techniques are a powerful subject of data processing. according to Edigar ADERO Crime has been increasing rapidly since the law was first adopted. At the governance level, there is heightened concern over the escalation of crime both internationally and locally in Kenya. In this paper, the researcher proposes the application of associative rule mining to create a model suitable for crime analysis and prevention using the Apriori algorithm to represent the mutual implications between criminal phenomena.

According to Chung-Hsien Most crimes are "undetermined. He discuss the preliminary results of a crime prediction model developed in collaboration with local police departments in the northern United States. He analyzed several classifications to determine which method is best at predicting crime "hot spots" He also looked at adding or creating groups. Here, to display the relationship between crime types uses an unsupervised learning method knowledge in unlabeled data. It is used to solve problems if user gets unlabeled example data. In addition, an association rule can discover interesting co-occurrences of objects in large datasets. At the core of the association rule, the rule consists of two parts.

1) Ancestor that is on the left side or called left side (LHS).

2) The consequence which is placed on the right side or is called the right side (RHS).

The form of the general rule of association is LHS! RHS, where LHS and RHS are disjoint itemset. If an LHS item set occurs, a RHS item set is likely to occur. Important statistical measures, support and trust measures are to be used together to effectively discover association rules

3.Problem statement

Crime is increasing rapidly these days and we need a system that detects crimes and stops crimes before they start. Crime prevention and detection is becoming an important trend in crime and very challenging to detect and solve crimes. Determining crime patterns is a tough job in today's world to reduce crime and take preventive measures to prevent crimes. Finding common crimes and relationships between crime types is a challenging task in today's crime sector. There is no automation for crime prediction.

4.Methadology

4.1 Proposed System

The proposed system is applicable in the field of crime. It includes crime modeling to find suitable algorithms for crime detection, Frequent graph, data preparation and transformation, and processing time of the apriori algorithms and based on the comparison with apriori tid algorithm, it is producing the graph. The proposed system identifies crime behavior, crime prediction, accurate detection and managing of large volumes of data obtained from various sources.It is an automation for registering complaints, predicting crime patterns based on previous crime details collected from various sources.

4.2 Functional Requirements

 Mainly 3 actors have access to the system: the administrator, investigative agencies and the public.

 The administrator is the registrar of investigative agencies.

 The administrator sets a unique user ID and password for each investigative agency.

 The system predicts crime patterns based on past crime data.

 The investigation office is registers complaints.

 Admin manages crime types and areas related to the specified city.

 Crime trend prediction is relay on date, location and type of crime.

4.3 Non Functional Requirements

Usability- our application will be useful to police stations, investigation agencies and public where system predicts the crime patterns which will be useful to reduce

crime rates. A Generic application which can be used my n number of users as the system is browser based application.

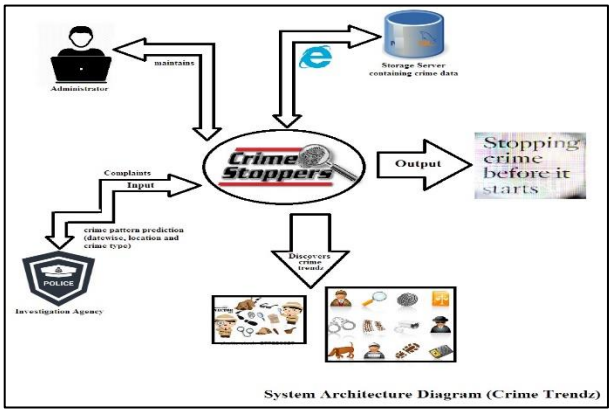
Reliable - our application provides the services according the users satisfaction and interest, and designed as per users requirements and more user friendly, so the application is more reliable compare to other crime based applications.

Maintainability - as we update the software regularly it will be easy to maintain it. Application is designed for modifications and enhancements can be done easily.

Efficiency - The application provides the efficient results for crime pattern prediction.

Re-usability – The system is a web based application, once the user creates an account; user can access the system multiple times.

4.4 System Design



System Architecture Diagram (Crime Trendz)

4.5 FRONT END USED: ASP.NET

ASP.NET could be the predominant web application system created by Microsoft. It enables users to create dynamic websites, applications and web administrations. ASP.NET provides a programming demo, extensive libraries, and facilities for building web applications using .NET dialects such as C# and Visual Essential. Are the key features of ASP.NET is its ability to create distributed web applications through the Model-View-Controller (MVC) structural design or the more modern Model-View-View Show (MVVM) design with systems like ASP.NETCore.ASP.NET also support the creation of Web APIs using innovations such as ASP.NET Web API and ASP.NET Center MVC that are commonly used for creating relaxation services. By integrating with the .NET System and the Visual Studio IDE, ASP.NET offers a rich environment for enhancements to create powerful, versatile, and secure web applications. It gives preferences like confirmation, authorization, session administration, caching, and numerous others out of the box, making it a favored choice for numerous designers and organizations.

4.6 BACK END USED: C# and MySQL

C# (pronounced 'C Sharp') is a modern computer programming language developed by Microsoft.
C# is a completely object-oriented language like Java and is a primary component-oriented dialect. It was outlined to support key features of the .NET system, unused Microsoft's advanced level for creating component-based programming arrangements. It is simple, efficient, profitable and type-safe derived from the well-known C and C++ dialects. the fact that it has its place in the C/C++ family, it can be an absolutely object-oriented, high-end dialect suitable for creating web applications. C# is designed for building strong, reliable, and resilient components to work with real-world applications.

The main strengths of C# are:

- A completely new language derived from C/C++ family.
- Simplifies and modernizes C++.
- the only component oriented language available today.
- the only language designed for the .NET framework.
- It is concise and modern language.
- Combines the best features of many commonly used languages: productivity Visual Basic, the power of C++ and the elegance of Java.
- It embodies today's interest in simplicity, productivity and robustness.
- Becomes the language of choice for .NET programming.
- The main work of the .NET framework is hardcoded by C#.

MySQL is an open-source social database management system framework (RDBMS) that is widely used for building customizable, high-performance databases. is created distributed and supported by Prophet Enterprise. MySQL is known for its unwavering quality, ease of use and compatibility with different phases and programming dialects.

Here are some key points of MySQL:

• MySQL is based on the social presentation of information,
where information is organized into tables with columns
and columns. Powers SQL (Structured Query Language)
for interrogation, review and supervision of social databases.

• MySQL is an open-source program, which means it is open access for download, use and distribution
under the GNU Common Open Permit (GPL). This
is open to engineers and organizations of all sizes.
Cross-Platform Compatibility:

• MySQL is accessible for different frameworks, including Linux, Windows, macOS, and UNIX. It can be
shipped on-premise, in the cloud or half-breed situation.

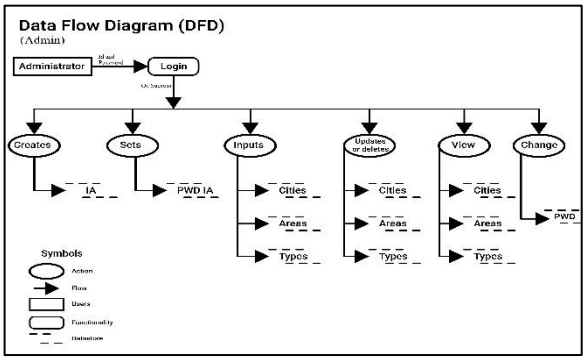
• MySQL is designed to handle large volumes information and high exchange rates. Enhances highlights
such as ordering, caching, replication, and allocation optimize performance and adaptability.

• MySQL provides built-in highlighting to ensure high pitch availability and resistance to guilt like master-slave replication, multi-master replication and clustering arrangements such as MySQL Gather Replication and MySQL InnoDB Cluster. MySQL includes security

• MySQL includes a large and dynamic community engineers, clients and donors who provide support, documentation, instructional exercises and extensions to improve its usefulness.

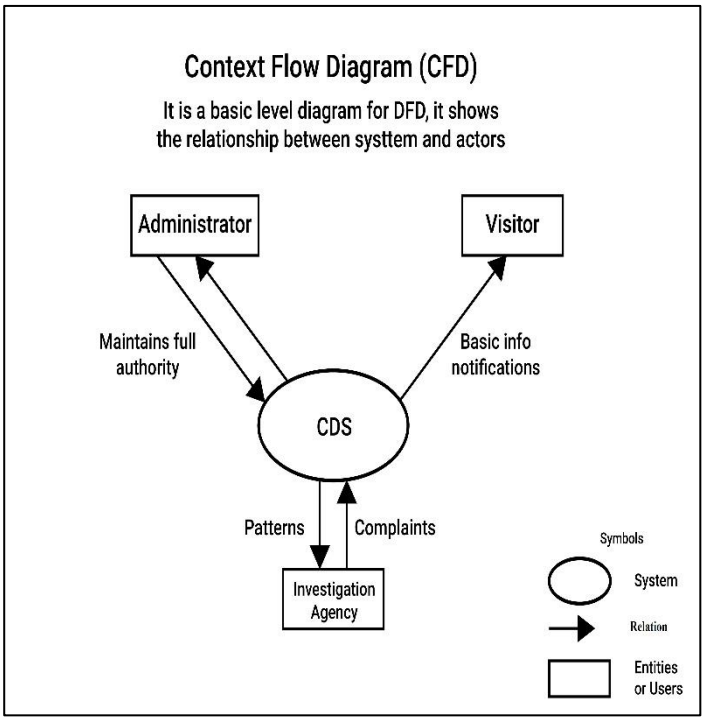
5.Data flow diagram:

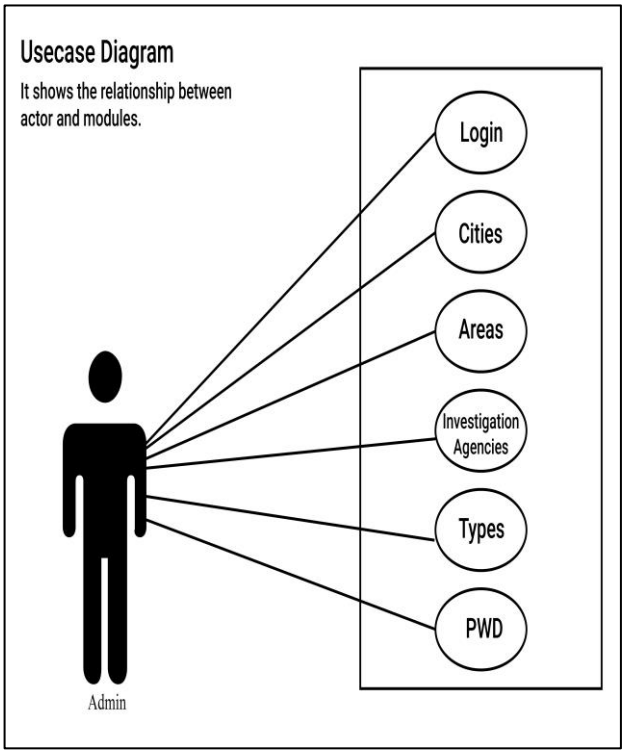
A Data flow diagram (DFD) is a graphical representation of the "flow" of information through a data frame. DFD is also used to visualize information manipulation (organized plan). On a DFD, information flows from an external source to an internal information store sink through an internal handle. The DFD provides no data nearly on the timing of the forms, only on whether the forms will work in arrangement or in parallel. It is therefore very different from a flow diagram, which appears as a flow of control through a calculation, allowing the user to decide what operations will be performed, in what order, and under what circumstances, but not what kinds of information will be input and output from the frame.



6.Control flow diagram:

It is common to draw an information flow diagram at the context level, showing the interaction between the framework and external operators that act as information sources and information sinks. In a setup diagram (also known as a "Level DFD"), the system's intelligence with the outside world is demonstrated absolutely in terms of information flows across framework boundaries. The settings table appears as the entire framework as a single node and gives no clue as to its internal organization. This context-level DFD is then "triggered" to provide a Level 1 DFD that shows several details of the displayed image. It shows how the framework is divided into subsystems (forms), each of which negotiates one or more information flows to or from an external specialist and which together provide the overall usefulness of the framework as a whole. It also distinguishes the internal storage of information that must be displayed for the framework to work, and the flow of information between different parts of the framework appears.





7.Use Case Diagram:

Unified Modeling Language (UML) is a type of behavior diagram defined and created from analysis. Its task is to provide a graphical overview of the functions provided by the system. The main purpose of a use case diagram is to show which system functions are performed for which actors. It does not show interaction between the actor. this interaction is necessary for a coherent description of the desired behavior, the boundaries of the system or use case should be re-examined. Alternatively, it is part of the assumptions used when using interaction between actors.

8. Current work

proposed system detects criminal behavior, predicts crime, accurately analyzes and manages largeamounts of information obtained from various sources. It uses data science technology "association rules" to predict the relationship between different types of crimes.

Example: location "vijaynagar", based on the previous crime data we can predict the differenttypes of crime that takes place in future.

Output – Crime patterns with different crimetypes in "vijaynagar" for upcoming days.

1. robbery, murder related to suicide
2. child abusing related to murder
3. chain snatching, robbery related to murder
4. pick pocket, chain snatching related torobbery and murder

Datasets

	A
1	Crimes
2	Murder, Kidnap, Chain Snatching, Robbery, Half Murder
3	Kidnap, Chain Snatching, PickPocket, Neck Hanging
4	Murder, Single Car Chain Snatching, Half Murder
5	Murder, Kidnap, Chain Snatching, Hit n Run
6	Hit n run, Drunk n Drive, Chain Snatching, Theft
7	Murder, Kidnap, Chain Snatching, Neck Hanging, Man Handling
8	Murder, Chain Snatching, Robbery
9	Drunk n Driven, Kidnap, Chain Snatching, Theft
10	Kidnap, Hit n run, Half Murder
11	Kidnap, Chain Snatching, Robbery, Half Murder
12	Murder, Kidnap, Chain Snatching, Robbery, Half Murder
13	Chain Snatching, Kidnap,Hit n run, Neck Hanging, Man Handling
14	Single Car Accident, Chain Snatching, Hit n Run
15	Bad Light, Kidnap, Chain Snatching, Hit n Run, Neck Hanging
16	Kidnap, Chain Snatching, Neck Hanging, Hit n Run
17	Murder, Single Car Chain Snatching, Half Murder
18	Murder, Kidnap, Chain Snatching, Hit n Run
19	Hit n run, Kidnap, Chain Snatching, Neck Hanging
20	Bad Light, Kidnap, Chain Snatching, Hit n Run, Neck Hanging

Fig 1: Crime Datasets

Unsupervised Learning

A descriptive model is used for tasks that benefits from the insight gained by summarizing data in new and interesting ways. The unsupervised learning technique has no predefined labels. The goal is to explore the data and find some structure within it. Unsupervised learning works well on transactional data.

In our project Association learning algorithms used such as "Apriori *Algorithm and Apriori TID Algorithm.* **These algorithms preferred as algorithms supports small datasets and also large datasets.**

9.Process Flow

Step 1: Crime data need to be collected from data sources like kaggle.com, dataworld.com, data.gov.in, githib.com etc...

Step 2: here crime data is preprocessed, where weremove unwanted data and fetch the required data. Unwanted data such as crime no, serial no etc. will be removed.

Step 3: required data is inputted to algorithms. We use efficient algorithms such as Apriori algorithm and apriori tid algorithm to process the data.

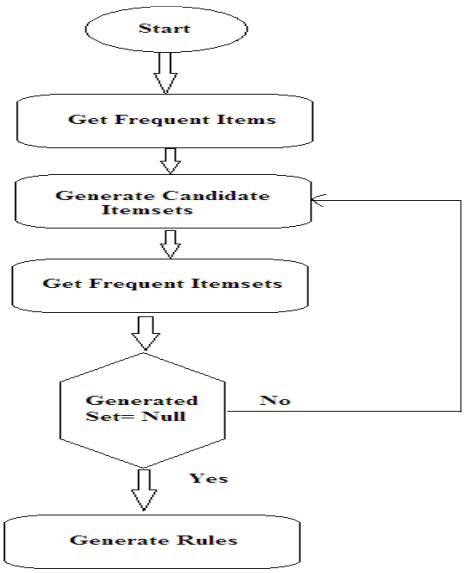
Step 4: after processing, frequent crimes will be extracted and displayed.

Step 5: Relationship between different types of crimes will be extracted and displayed**.**

Step 6: both algorithms results compared and efficient algorithm will be chosen.

Step 7. Final outputs displayed on GUI.

Step 8: Comparative analysis of algorithms displayed using graphs.



Flow of the Algorithm

STRONG ASSOCIATION RULE:

1. { Murder } -> { Pickpocket }
2. { Kidnap Pickpocket } -> { Robbery }
3. { Robbery Pickpocket } -> { Kidnap }
4. { Robbery } -> { Kidnap }
5. { Kidnap } -> { Robbery }

10.Results

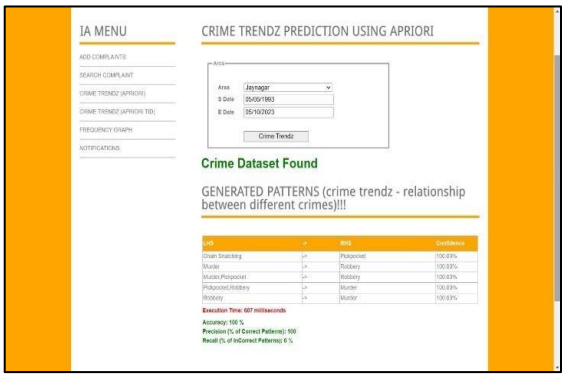


Fig 1: Apriori Patterns with Accuracy

Fig 2: Apriori TID Patterns

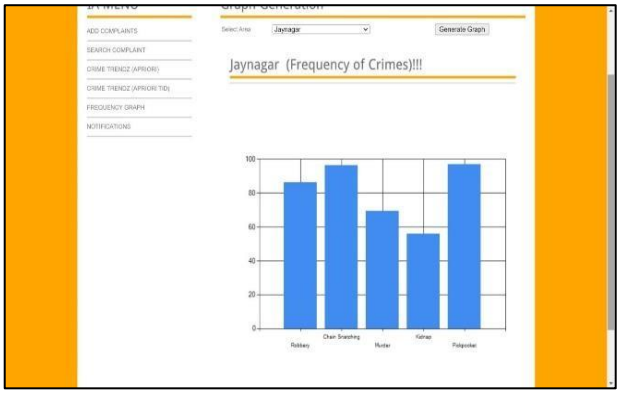


Fig 3: Crimes Frequency Plot

11.Conclusion

Identifying the different types of crimes that may take place in the coming days plays a vital role in today's crime or investigation sector. To reduce crime in the city, it is necessary to take preventive measures to avoid crimes. The proposed system predicts frequent crimes and their interrelation in advance, so that the police or investigation department can take necessary measures to reduce crime. We are creating a real-time application where it is useful for the crime sectors to reduce the crime rates.

12.References

[1]Pushpalatha R., Ramesh B, “A CFRO Taylor Based Deep Learning Model for Service Level Agreement - Aware VM Migration and Workload Prediction - Power Allowed Power Model in Cloud Computing”

[2]R Pushpalatha, B Ramesh - Advanced Concepts in VLSI, Signal Processing and Power Electronics, "Merging Neural Network and Genetic Algorithm for Effective Data Center Workload Prediction"

[3]M Chandini, R Pushpalatha, R Boraia, “A Brief Study on Load Prediction in Cloud Environment”

[4]S. Sathyadevan, M. Devan and S. Surya Gangadharan, "Crime Analysis and Prediction Using Data Mining" in Networks and Soft Computing (ICNSC), 2014 1st International Conference, Aug 2014, pp. 406-412.

[5]S. Kaza, T. Wang, H. Gowda, and H. Chen, “Target

vehicle identification for border security using mutual information," in Smart Transportation Systems, 2005. Proceedings. 2005 IEEE, September 2005, pp. 1141–1146.

[6]W. Huang, M. Krneta, L. Lin, and J. Wu, "Association Package - A New Paradigm for Association Analysis," in Data Mining Workshops, 2006. ICDM Workshops 2006. IEEE Sixth International Conference on, Dec. 2006, pp. 601–605 .

[7]N. Sasaki, R. Nishimura, and Y. Suzuki, "Audiowatermarking based on association analysis," 2006 8th Inter Conference on, vol. November 4, 2006.

[8] A. Ben Ayed, M. Ben Halima, and A. Alimi, "A Survey on Clustering Methods: Towards a Fuzzy clustering for big data" 2014 6th Inter Conference on, August 2014, pp. 331-336.

[9] A. Thammano and P. Kesisung, "Improved k-means algorithm for solving classification problems," in Machine Learning.