A Review Paper: Using Data Mining Clustering Technique to Predict Criminal Behavior

Dr. Gurpreet Singh
Associate Professor & Head, Department of Computer Sc. & Engg., St. Soldier Institute of Engg. & Tech. Jalandhar (Punjab)
saini3077@gmail.com

Er. Amarjit Singh
M.Tech Scholar, Department of Computer Sc. & Engg., St. Soldier Institute of Engg. & Tech. Jalandhar (Punjab)
amarjitramgarhia66@yahoo.com

Abstract: Crime is defined as illegal and immoral activity. It is a sin against the society that is often prosecuted and punishable by the law. Crime is the behavior disorder that is an incorporated result of community, cost-effective and environment factors. Young people's development occurs within larger Family members, the school, the community, and cultural and sociological forces all influence students as they learn appropriate coping skills and behavior for dealing with their environment. The growing volume of crime data is the reason of concern about the future of national security. Children are the future of nation. The children who had deviant behavior are now in the correction centers will become the main destructors of nation. The main focus is to analyze the inconsistent and incomplete crime data by using data extraction and data preprocessing to make it efficient and accurate.

INTRODUCTION

With the increasing demand of IT and subsequent growth in this sector, the high- dimensional data came into existence. Data Mining plays an important role in analyzing and extracting the useful information. The key information which is extracted from a huge pool of data is useful for decision makers. Data mining is used for forecasting future trends of market and also helps in decision making like in business, science, medical, counter- terrorism, telecommunication etc. Data mining is also used in analyzing, detection, identification and predicting crime. Crime is defined as illegal and immoral activity. It is a sin against the society that is often prosecuted and

punishable by the law. Crime can simply be defined as the violate of social norms of particular society. There can be many different causes of crime and many studies are conducted all around the world to understand and bring down criminal activities. The fight against crime is not a new thing in people and it has, since the concern of society and tried to bring crimes down. Overpopulation, poverty, drugs, family violence, politics, TV violence, depression etc are some causes of crime. Clustering and Classification are the techniques of data mining are generally used methods of analyzing, detection, identification, prediction of crime.

DATA MINING

Data mining is the process of discovering interesting patterns and facts from large amounts of data and also known as knowledge discovery from data [1]. The data sources can databases, the web, data warehousing etc. Data mining converts raw or huge data into interesting and useful information. The given below steps are involved in extracting knowledge from data:

i. **Data Selection:** The necessary data of required field can be collected from various locations. The data related to the analysis task are retrieving from the database.

ii. **Data Preprocessing:** In this stage, the two processes are involved data cleaning and data integration.
a) **Data Cleaning:** In this step, noise data, irrelevant and inconsistent data are distant from the collected data.
b) **Data Integration:** In this step, various data sources of same type are combined.

iii. **Data Transformation:** In this phase, the preferred data is changed into required forms for the mining procedure.

iv. **Data Mining:** It is the important step in which techniques and intelligent methods are applied to extract useful information and data patterns.

v. **Evaluation:** In this step exciting patterns representing facts are recognized.

vi. **Presentation:** This is also an important step, uses visualization and knowledge representing techniques to help users to understand.

Classification

Classification is the most commonly used data mining technique, where a model is constructing to predict class. It is used to classify each item in a set of data into one of predefined set of classes or groups. Classification method uses mathematical techniques such as decision trees, linear programming, neural network and statistics. Classification classifies data items into groups. Types of classification models are Classification by decision tree induction, Bayesian Classification, Neural Networks, Support Vector Machines (SVM), and Classification Based on Associations.

Clustering

Clustering is a data mining technique that makes useful cluster of objects that have similar characteristic. Classification techniques can also be used for distinct groups or classes of object but it becomes costly so clustering can be used as preprocessing technique for selection and classification. Types of clustering methods are Partitioning Methods, Hierarchical

Agglomerative (divisive) method; Density based methods, Grid-based methods, and Model-based methods.

Causes of crime

1. Increase in population is the biggest **cause of crime** and much of the world's worries. Although population increase is related to each and every cause mentioned here, it still needs to be looked at as a cause of crime. The increase of **population** triggers of a dynamo effect in society and this leads to the creation of more people with some form of frustration or resentment towards society as such.

2. **Poverty:**

Economic deprivation or simply poverty is a major **cause of crime** all around the world. People are often driven to great lengths of desperation by poverty and this is a major cause of crime all around the world. The fact that such frustration is created is in itself a very dangerous thing for society on the whole as global inflation has risen significantly over the last few years. Although it does seem that in our world today, the rich get richer and the poor get poorer.

Politics:

Politics is often a **cause of crime**. It is seen that many political associations all around the world have their own mafias running which they use to manipulate and subjugate people. Political power is often misused to take advantage of weaker groups and people and the dissidence that

rises out of such situations often force the **victims** to resort to **crimes**. **Politics** is more related to crime on a much larger and a much heinous level than anything else.

Racism:

Discrimination based on race is a serious issue all around the world. All humans are in a way racist towards some people in some part of the world or another. **Racism** has contributed a lot of unrest to many places all around the world and it's mostly because of one or two idiots that such **crimes** are given birth to. It is a sad fact of the world that we live in that we end up discriminating against something that is the same flesh and blood underneath even though the external appearance and the origins may be different.

TV violence:

TV violence has gone up to staggering levels and it does not help when people are influenced and try to emulate such acts of **violence**. TV violence is a major **cause of crime** especially among younger people that are unable to differentiate between fiction and reality. Since TV has become such an integral part of people's lives now days, it is important to draw clear lines between what is real and what is not.

Family conditions:

There are a lot of things that go on in families that often **cause** people to get into a life of crime. Here again there are a lot of different conditions that lead a person into crime. Abuse during formative years from **family** members and other such acts also instigate a person into a life of **crime**. People who are neglected by their families and do not get the love and attention that they desire also get into criminal activities. Family violence and other issues are also related to crime in many ways.

Drugs:

Drugs are a bane, no matter how we look at them. A person addicted to **drugs** is unable to support their addiction and more often than not they end up in a life of crime to fuel their habits. It is not unknown anywhere in the world that a **drug** addict ends up committing **crimes** to raise money for their habits. Besides that there are also a large number of people that are involved in the drug trade. Though these people may not really be drug users themselves, they often lure others into **drugs** and **crimes**.

Victim of unfair rulings and the correction system (prison system):

Victims of unfair or incorrect rulings from court often **cause** people to enter a life of **crime**. It often happens that a person is a **victim** of chance and happens to fall into **crimes**. Besides that, people are also often falsely accused of committing crimes which ends up in a **court** conviction. Jails or prisons often make worse criminals out of people because of the conditions that exist there. Corrections anywhere does not involve major rehabilitation for criminals and more often than not they end up being thrown in overpopulated jails full of people who are either victims or perpetrators of crimes more serious than their own. The declassification of people in prisons is also a major cause of **crime** creation.
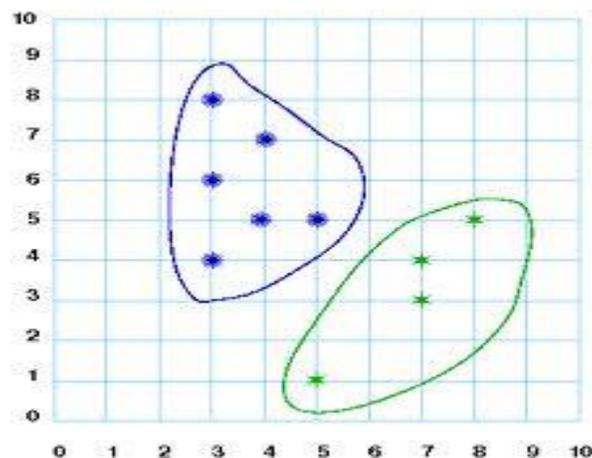

<u>**Factors Identified as Contributing to Violent Crime**</u>.

1. Gangs 82 %
2. Juveniles / youth crime 80 %
3. Economy / poverty / unemployment 74 %
4. Impulsive violence/disrespect issues 74 %
5. Release of offenders from correctional institutions 69 %

6. Drugs-Cocaine 67 %

7. Poor parenting 63 %

8. Increased availability of guns 55 %

9. Reduced cooperation from witnesses / victims (not in top 10 last year) 37 %

10. Educational system-increasing dropout rates (not in top 10 last year) 36 %

Clustering

Cluster analysis divides data into meaningful or useful groups (clusters). If meaningful clusters are the goal, then the resulting clusters should capture the "natural" structure of the data. For example, cluster analysis has been used to group related documents for browsing, to find genes and proteins that have similar functionality, and to provide a grouping of spatial locations prone to earthquakes. However, in other cases, cluster analysis is only a useful starting point for other purposes, e.g., data compression or efficiently finding the nearest neighbors of points. Whether for understanding or utility, cluster analysis has long been used in a wide variety of fields: psychology and other social sciences, biology, statistics, pattern recognition, information retrieval, machine learning, and data mining.

Conclusion

Crime is defined as illegal and immoral activity. It is a sin against the society that is often prosecuted and punishable by the law. Young people's development occurs within larger Family, the society, and educational and sociological factors manipulate students as they learn, coping skills and behavior for dealing with their environment. The growing volume of crime data is the reason of concern about the future of national security. Children are the future of nation. The

children who distract from right path are now in the correction centers will become the main destructors of nation.

The result has been showed the age within 12 to 18 indulged in crime intentionally due to his/her age and crime type.

References

1. Wang Lijun, Rege Manjeet, Dong Ming, Member, IEEE, and Yongsheng Ding, Senior Member, IEEE, "Low-Rank Kernel Matrix Factorization for Large-Scale Evolutionary Clustering" , *IEEE Transactions on Knowledge and Data Engineering,* vol. 24, no. 6, june 2012.

2. Vongsingthong Suwimon, Wisitpongphan Nawaporn, "classification of university students behave.or in sharing information on facebook*", IEEE/International Joint Conference on Computer Science and Software Engineering,* 2014.

3. Tayal Kumar Devendra , Jain Arti , Arora Surbhi ,AgarwalSurbhi, Gupta Tushar, Tyagi Nikhil, "Crime detection and criminal identification in India using data mining Techniques*", Springer*,march/april 2014.

4. Wei Wei, Jinjiu Li, Longbing Cao, Yuming Ou, Jiahang Chen , "Effective detection of sophisticated online banking fraud on extremely imbalanced data", *Springer Science Bussiness Media, LLC 2012.*

5. Mansingh Gunjan, Rao Lila, Osei-Bryson Kweku-Muata,"Profling internet banking users: A Knowledge discovery in data mining process model process model based approach", *Springer Science Business Media New York 2013*.