Technical Journal, University of Engineering and Technology (UET) Taxila, Pakistan *Vol. 26 No. 1-2021*
ISSN:1813-1786 (Print) 2313-7770 (Online)

An Intelligent Analysis of Crime Data using Data Mining Algorithms

S. Saeed[1], M. M. M. Bagram[2], M. M. Iqbal[3]

[1]*Department of Computer Science, Virtual University of Pakistan*
[2]*Department of Business Administration, Allama Iqbal Open University Islamabad, Pakistan*
[3]*Department of Computer Science, University of Engineering and Technology, Taxila, Pakistan*

[3]munwar89@gmil.com

Abstract- Crime is expanding all around the world through many ways, and violent crimes have become a national dilemma. Family members try to ignore, and neighbours do not want to get involved. The victim in such case is usually helpless to understand and to determine the solution of the problem. Also, law enforcement agencies handle each case when someone got affected or when a criminal act has taken place. Data mining turns an essential role in the prediction and the analysis of data. For the crime analysis in various states of the United States, several classification algorithms have been applied to the crime data of the FBI concerning the population of states. Experimental result showed that two classifier algorithms Reduced Error Pruning Tree (REP Tree), and Naïve Bayes produced better results as compared to other algorithms. However, most of the results of Classifier algorithms were same, so we concluded that our dataset had been classified correctly. The outcome from the analysis shows that crime rate is high in populated states. It is observed that rate of property crime and rate of larceny-theft is high amongst all other types of crime rate in populated states of the United States of America (USA).

Keywords- Data Mining, Classification, Violent Crime, Murder, Property Crime, REP Tree, Naïve Bayes.

I. INTRODUCTION

Crime is an unlawful act punishable by a state or other authority. Crime has a direct impact on victims and indirect effects in the broader community. Notably, over the last ten years there has been a massive increase in crimes, e.g. “murder, rape (revised definition), rape (legacy definition), robbery, theft of vehicle, and aggravated assault”. Primary data also show that law enforcement agencies observe overall a 5.3% increase in crime rates [1]. The latest Federal Bureau of Investigation FBI’s Crime Report has also shown that the crime rate increased by 3.4 percent nationwide. In general, we can say that crime is expanding all around the world through many ways and become a national dilemma. Family members try to ignore, and neighbours do not want to get involved. The victim in such case is usually helpless to understand and to determine the solution of the problem. Also, law enforcement agencies handle each case when someone got affected or when a criminal act has taken place. So, there is a strong need to take quick and compact actions to stop or countermeasure the crime. Data mining could be a solution to this problem as data mining techniques have significant influence in some fields such as, crime data analysis, crime pattern analysis, criminal career analysis etc. Else we can say that Data mining acts an important role in the prediction and analysis of data [2].

B. Widner suggested that all crimes in the border region are rising in the northern border-states of Mexico [3-4]. A model of crime and criminal detention was presented using data and methods to estimate the factors that lead to more criminal arrests. Subsequently, measures could be developed to mitigate criminal activity in Mexico. The results of this study show that some actions of the Mexican government have had a positive impact on crime. The detention of prisoners, which is problematic because of prison overcrowding in Mexico, has nonetheless led to criminal activity and reduced the number of criminal arrests [5].

As the FBI is the main department of the Ministry of Justice. Classification technique will be used for data mining. Different classification algorithms will be applied to FBI crime dataset to analyze. All classifier algorithms are available in Weka. Furthermore, the algorithms are compared by the results of the survey, and the investigation and analysis is carried out to analyze crime in different states of the USA with respect to the population of states [5]. The scope of the paper is as follows: Section 2 describes the

literature review; Section 3 discusses the methodological framework; Section 4 deliberates experimental results, analysis, and findings. The last section contains the conclusion and future work.

II. Literature Review

The research classification uses crime records to predict the "crime category" in various states of the United States. This article compares two different algorithms sued for classification: Naive Bayesian and Back Propagation. The results show that Naïve Bayesian performed major calculations [6]. A. Babacura discussed that the classification is the process of creating a model that draws and identifies classes with the goal of being able to use the model to predict the names of the crimes. In this research data mining tool Weka has been used. We will use FBI Crime dataset for data mining as it provides accurate and precise statistics of crime within the United States [7].

K. Janek investigation of crime externality and effects of violent crime on people [8]. In the variance of the region, the results and their evidence show a negative impact of violent crime in society as a whole, through the mechanism of increasing security problems [9]. The study identified the effects of violent crime on the daily activity and on the other hand it also prescribed that targeted action to reduce the fear of violent crime in society can have a positive impact in the future [10].

Sr.#	Author	Year	Problem Definition	Solution	Limitations
01.	B.Widner and et al. [11]	2011	To identify the factors that lead to less arrests of criminals in Mexico.	They used Data and various methods for assessing the factors in Mexico.	They will try to solve legal system problems of Mexico in the future.
02.	A.Babakura and et al. [12]	2014	To predict "Crime category" based on the classification of the criminal data of different states of USA.	By using Naive Bayesian and Back Propagation algorithms, they predicted the "Crime Category".	In Future they will apply different classification algorithms to the criminal records for prediction.
03.	K. Janke and et al. [10]	2016	Companies in most prevalent crime areas have higher variable income.	By exploring crime rates they predicted variability factors.	Impact cost of capital of companies will be analyzed.
04.	J. Wu and et al. [13]	2016	For Complex and changing data descriptive statistical model is used.	They reported pre-processing activities on the survey data.	In Future will analyze historical decision trees for different crimes.
05.	J. Rushwood and et al. [14]	2016	Companies in states where property crime is more common have uncertain profits and higher costs.	Many steps have been taken to minimize property crimes and a corporate capital cost is not wrong.	Different security and management strategies will be used to increase low capital costs.
06.	H. Chi and et al. [15]	2017	Offenses committed by a criminal can help to identify serial crimes.	They presented a crime-based decision support system for problem.	System performance will be further improved.
07.	J.A.Sheikh and et al. [16]	2017	Crime analysis needed an extensive research for crime analysis and mapping.	They combined statistical methods (cluster analysis) and spatial model drawn with Geographic Information System GIS for detailed	Complex (spatial) cluster analysis (block analysis) will be recognized.

				analysis.	
08.	N. Baloian and et al. [17]	2017	It is very complex task to capture the time and risk of certain crimes through computer algorithms.	Through Science fiction they estimated the prediction of future occurrence of crimes.	Integration of results for majority voting function will be performed to minimize the risk.
09.	J.D.A.Fossoul and K.A. Roberts [18]	2017	Authorities have considerable difficulties in combating cross-border criminal groups.	They used forensic science to combat crime with the help of forged documents.	Forensic knowledge generated by this method will be applied to different sectors.
10.	Y. Zhuang and et al. [19]	2018	Crime is a major social problem in the USA that threatens public safety and disrupts the economy.	Hot spots allow police to prevent or respond to accidents more efficiently and effectively.	Include additional features in the Spatio-Temporal Neural Networks STNN model to improve performance.
11	H. Hassani and et al. [20]		Text mining in big data analytics	This paper includes mining of Web and social data for detection.	Actions must be undertaken moment in time to effectively resolve the legal, ethical, and privacy worries in the use of individual data.

J. Wu created significant decision tree models from the Crime Survey. The models cover 18 crimes in 14 years. They reported pre-processing activities on the survey data to enable their use with mining models. They also reported that the results of the earlier analysis of the survey data using decision trees. However, the next big challenge is to analyze historical decision trees for individual crimes and various related crimes. It means that criminologists must analyze very complex models. It will be very difficult to just look at models with human eyes. As a result, data visualization techniques are adopted [13].

J. Brushwood showed that businesses in states with more prevalent real estate (property) crime have lower returns and higher financing costs. The cost of crime in the US economy is considerable. They showed that the potentially high and neglected costs of crime mean that businesses in more criminal sectors incomes are variable and greater financing costs. As a result, the paper provides information on the impact of location and crime on the costs of social capital, which could also affect tax and investment policies [14].

H. Chi presented a crime-based decision support system based on various behavioural characteristics of criminal cases. Its core technology is a similarity-based pairing classification that can be interpreted and easily adjusted. The use of this system would provide strong support to officers in a criminal investigation and would save the law enforcement agency resources. The system can automatically link serial crimes based on a flexible data classification model. However, there are still some limitations in the study. First, the approach can not directly handle textual data. Secondly, the approach only confirms its validity in the context of serial waste. In the future, they will continue to use text extraction techniques such as text classification and thematic modelling of sentences or short notes [15].

Table 2: Analysis of Different Types of Crimes and Techniques

Sr. No	Author	Violent Crime	Murder	Rape	Robbery	Assault Crime	Property Crime	Burglary	Crime Prediction	Decision Support
01.	B.Widner and et al. [11]	YES	YES	YES	YES	NO	NO	NO	NO	YES
02.	A.Babakura and et al. [12]	YES	NO	NO	NO	NO	NO	NO	YES	YES
03.	K. Janke and et al. [10]	YES	YES	NO	YES	NO	YES	YES	NO	YES
04.	J. Wu and et al. [13]	NO	YES	NO	YES	NO	NO	YES	NO	YES
05.	J. Rushwood and et al. [14]	YES	NO	NO	NO	NO	YES	YES	YES	YES
06.	H. Chi and et al. [15]	YES	YES	YES	YES	YES	YES	YES	YES	YES
07.	J.A.Sheikh and et al. [16]	NO	NO	NO	NO	NO	YES	NO	YES	YES
08.	N. Baloian and et al. [17]	YES	NO	NO	YES	YES	NO	YES	NO	NO
09.	Y. Zhuang and et al. [19]	NO	NO	NO	NO	NO	NO	YES	NO	YES
10.	J.D. Alcaraz-Fossoul and K.A. Roberts[18]	NO	NO	NO	NO	NO	YES	NO	NO	YES

J. A. Sheikh have argued that crime exploration has become a common term in recent years, requiring a great deal of research on crime analysis and mapping. The mapping of crime and spatial analysis complements them all and plays an important role in the new intrinsic form of representing and identifying crime and a satisfactory response to the problems of crime. This research combines statistical methods (clustering analysis) and spatial model drawn with GIS from crime reports. The document generated on the various uses of GIS to detect hot spots of crime and promote the development of police investigation strategy [21]. The functional approach of the current crime mapping study can be used successfully to improve the use of GIS in law enforcement agencies. Finally, GIS and mapping can support local and show the complete correlation between crime, victim and perpetrator [16]. N. Baloian argued that science fiction had anticipated the prediction of the future appearance of crime. Prediction is possible. The developed system was tested on historical data, and its performance was found to be acceptable for police use. An interesting discovery is that the performance

of each module is lower than the common performance. In addition, how the results are integrated should probably be changed to a majority vote function to minimize the risk [17]. Y. Zhuang have suggested that crime is a major social problem in the United States that threatens public safety and disrupts the economy. In the study, they proposed the Spatio-Temporal Neural Network (STNN) to predict crime hot spots embedded in spatial information accurately. They have demonstrated that their STNN model which performs better to number of conventional machine approaches. In future work, they plan to include additional features in the STNN model to see if they can improve model performance, ideally to the point where they can be realized in a shorter period of time [19].

III. METHODOLOGY

Figure 1 shows our methodology flow diagram. That is how we proceed in our research. Now we discuss here each phase one by one.

Data Set

The dataset used for analysis is real and authentic. We chose FBI dataset because FBI is the most active law enforcement agency against crime and has more accurate facts and figures. Dataset was in (.xls) format and includes crime statistics of 50 states of USA. Dataset contains total 14 attributes as: "states, city, population, violent crime, murder, rape revised definition, rape legacy definition, robbery, aggravated assault, property crime, larceny-theft, motor vehicle theft, arson and burglary" and 9291 instances.

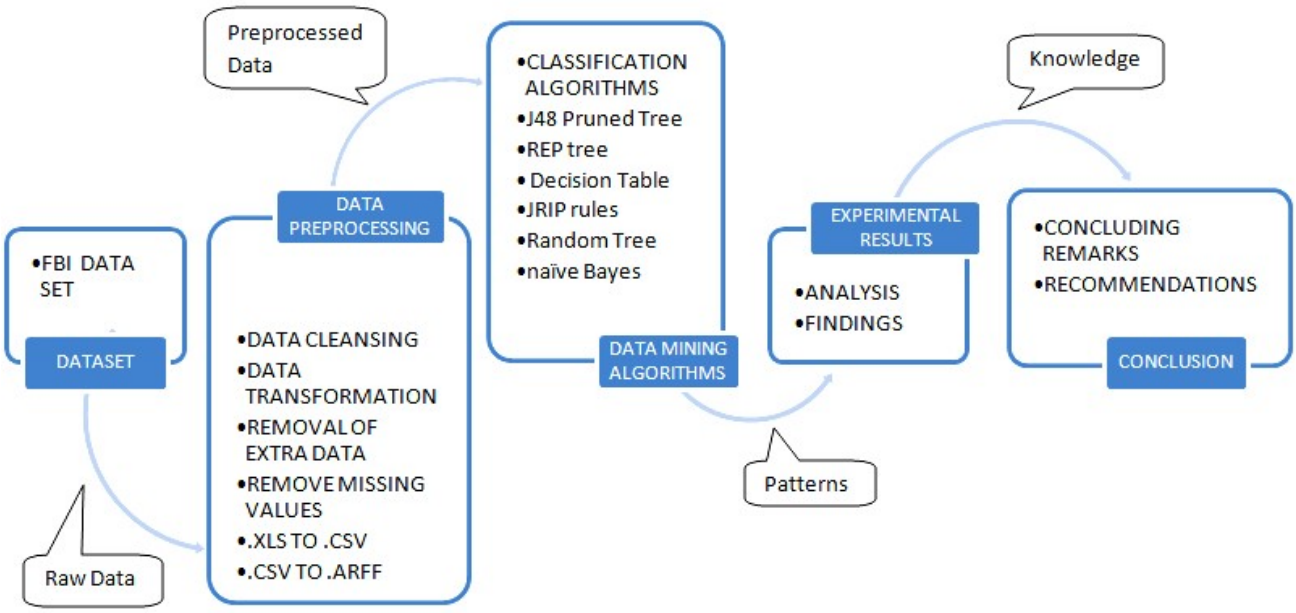


Figure 1: Flow Diagram of Methodology for Proposed Solution

Data Preprocessing and Cleansing

As our dataset was so huge so we reduce the size of our dataset by removing extra values. As we want to analyze crime at significant state level, so we remove extra states and cities attribute. Further by categorizing aggravated assault, burglary, and arson etc. under violent crime, we removed these attributes also. Resulted database contains 850 Instances and 9 Attributes as: "States, Population, Violent Crime, Murder, Rape, Robbery, Property Crime, Larceny Theft, Motor Vehicle Theft". Then we replaced missing values and saved file in (.csv) format. After that, we loaded (.csv) file in Weka tool for further pre-processing. In Weka, applied two filters on our dataset "Replace missing values" to replace remaining missing values and "Numeric to nominal" to convert the numeric type of values into nominal type, respectively. Furthermore, in data transformation, we convert the file format from (.csv) to (.arff), which is detected and run smoothly on Weka.

Data Mining Algorithms

In many areas like homeland security, weather forecasting, health care, medical, financial, business intelligence etc classification got a lot of success.

Data mining has several techniques e.g. Clustering, Classification etc. for mining useful patterns and data. Clustering is the process of grouping data objects in such a way that the objects in one group are similar with respect to the other group objects [22]. On the other hand, Classification is a class prediction technique which predicts class labels which are nominal. That is why we used Classification technique as it is most important and utilized a lot. These are the classification algorithms we used for analysis (J48 Pruned Tree, REP tree, Decision Table, JRIP rules, Random Tree, naïve Bayes) to produce competitive results. Classification has been applied through Weka tool as all algorithms are available in Weka.

Environmental Setup

We used Weka tool for the analysis and comparison of different classification algorithms. It is used as a data mining tool which contains different types of algorithms. We choose classification technique and compared six classification algorithms based on "time taken to build model, correctly classified instances, incorrectly classified instances, Root mean squared error, TP rate (True Positive), FP rate (False Positive), Precision, Recall and F-Measure values". For evaluation purpose, a test mode 10-Fold (Cross-Validation) has been used.

IV. Experiment Results, Analysis and Findings

The Table below is showing different values of "Time taken to build model, correctly classified instances, incorrectly classified instances, Root mean squared error" for classifier algorithms. Two classifier algorithms (REP Tree, Naïve Bayes) are producing better results in less time as compared to other algorithms. The graphical representation of the table has been shown in Figure 2. Most of the results using different algorithms are approximately the same.

Table 3: Time Comparison of Different Algorithms

Sr. No	Classifier Model	Time taken to build model	Correctly Classified Instances	Incorrectly Classified Instances	Root mean squared error
1	J48 pruned tree	O.03 s	85.6115 %	14.3885 %	0.1033
2	REP Tree	O.02 s	86.3309 %	13.6691 %	0.1131
3	Decision Table	0.06 s	86.3409 %	13.6591 %	0.116
4	JRIP rules	0.03 s	86.3409 %	13.6591 %	0.1131
5	Random Tree	0.03 s	85.6115 %	14.3885 %	0.12
6	Naive Bayes	O.02 s	91.3709 %	8.628 %	0.0163

Naïve Bayes Classifier

Naïve Bayes classifiers perform the best performance over the data set. The detail of the result is given below:

Table 4: Result summary of Naïve Bayes

1	Correctly Classified Instances	3279	91.3709 %
2	Incorrectly Classified Instances	308	8.628 %
3	Kappa statistic	0.9101	
4	Mean absolute error	0.081	
5	Root mean squared error	0.0163	
6	Relative absolute error	0.011 %	
7	Total Number of Instances	3277	
8	Ignored Class Unknown Instances	5	

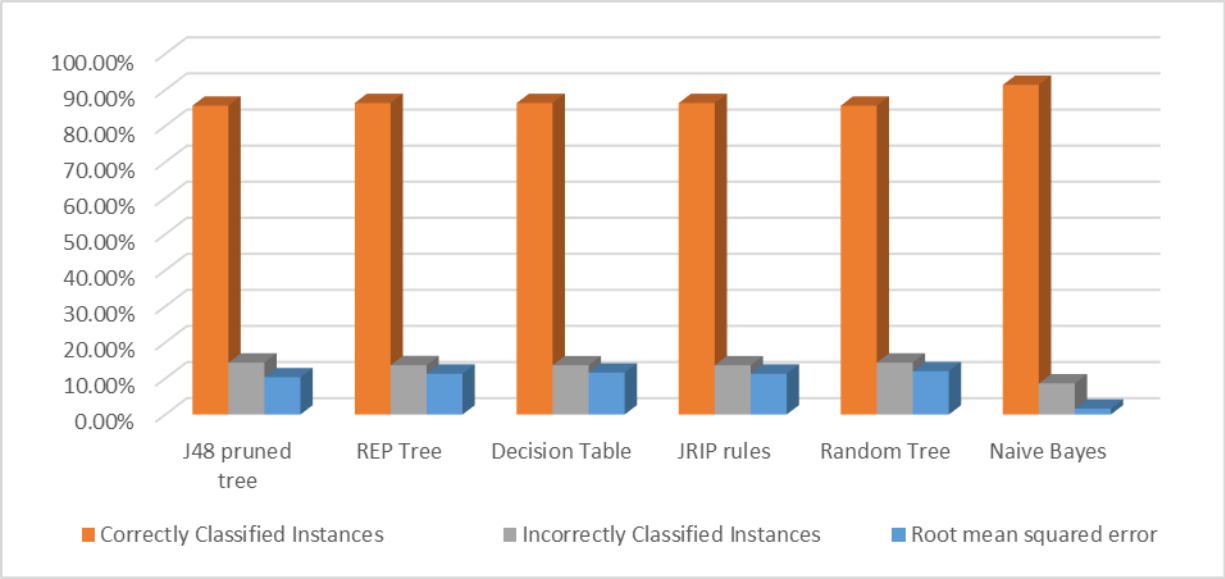


Figure 2: Time Graph of Different Algorithms

Detailed Accuracy (Weighted Average)

In Table detailed accuracy in term of a weighted average of TP rate (True Positive), FP rate (False Positive), Precision, Recall and F-Measure has been shown. Two classifier algorithms (REP Tree, Naïve Bayes) are producing better results again. So, we can say that both classifier algorithms (REP Tree, Naïve Bayes) are better as compared to other algorithms for our scenario.

Table 5: Detailed Accuracy Table in term of Weighted Average

Sr. No	TP Rate	FP Rate	Precision	Recall	F-Measure	Classifier Model
01	0.856	0.137	0.842	0.856	0.849	J48 pruned tree
02	0.863	0.863	0.745	0.863	0.800	REP Tree
03	0.863	0.863	0.745	0.863	0.800	Decision Table
04	0.863	0.863	0.745	0.863	0.800	JRIP rules
05	0.856	0.092	0.849	0.856	0.853	Random Tree
06	0.913 %	8.73 %	0.90	0.96	0.941	Naive Bayes

The graphical representation of the table has been shown in Figure 3. As Most of the results of Classifier algorithms are same here also so we can conclude that our dataset has been classified correctly.

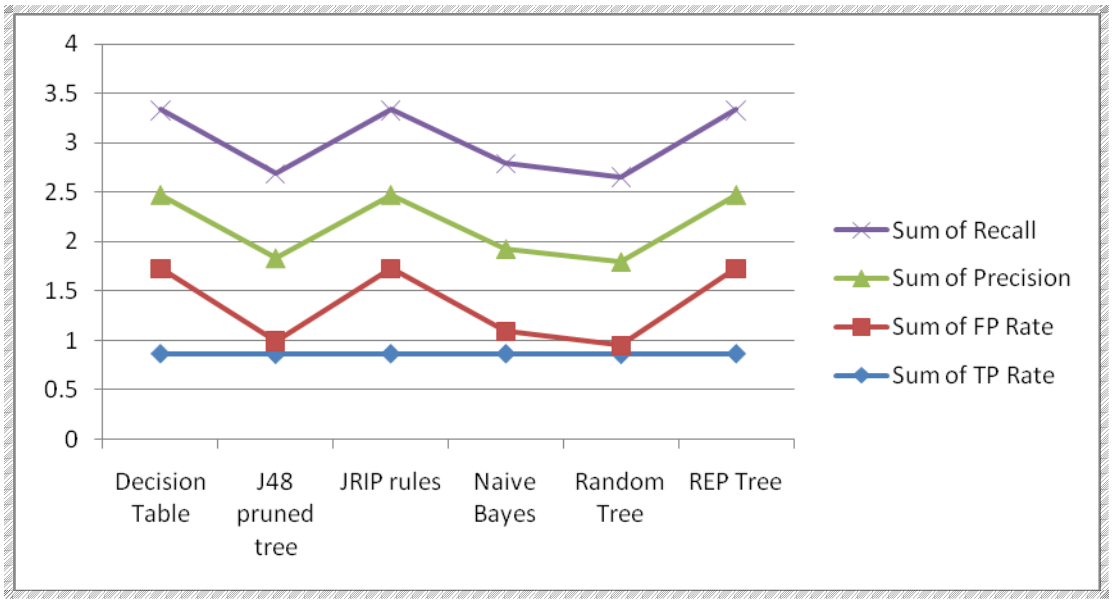


Figure 3: Graph of Detailed Accuracy Table

In Table 1, the Naïve Bayes Algorithm results have been compared. We found results in two literatures hence we compared their results with our results to get a better understanding. Results of literature [12] are not so good as they used dataset having 5000 instances and 45 attributes. Results of literature [19] are very good because of a dataset having 1078 Instances and 23 attributes. Our dataset contain 850 instances and nine attributes, as we mentioned earlier. Our results are quite reasonable and quite near to literature [19] and far good from literature [12]. So, by this comparison, we can get that our Naïve Bayes classification algorithm also produced better, consistent results and classified data correctly.

Table 6: Comparison of Naive Bayes Algorithm

Algorithm (Naïve Bayes)	**Accuracy (Correctly classified instances)**	**Incorrectly Classified Instances**	**Precision**	**Recall**	**F-Measure**
Results in Literature [10]	86.3309 %	13.6691 %	0.836	0.863	0.849
Results in Literature [12]	74.3 %	25.7 %	0.79	0.66	0.719
Results in Literature [19]	90.2207 %	9.7793 %	0.958	0.93	0.937
Proposed Solution	91.37 %	8.73 %	0.90	0.96	0.941

Below in Figure 4, the graphical representation of comparative results from Table 1 has been shown. It clearly indicates that our Naïve Bayes classification algorithm produced reasonable results in comparison to literature results.

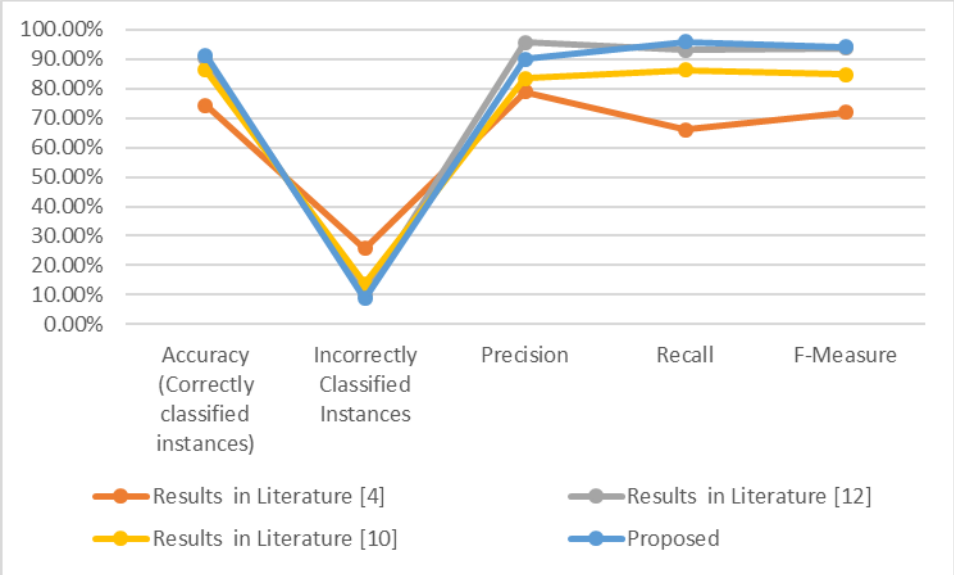


Figure 4: Precision, Recall, F-Measure Comparison of naive Bayes Algorithm

Figure 5 is showing total population of different states. Population is a key factor in our analysis as we want to know the crime rate in states with respect to population. In simple words, we want to conclude that either crime has a direct relation with the population or not. Mean crime is occurring in more populated states or in low populated states. According to the graph, California is at; first, Texas is at second, New York is at the third position in population.

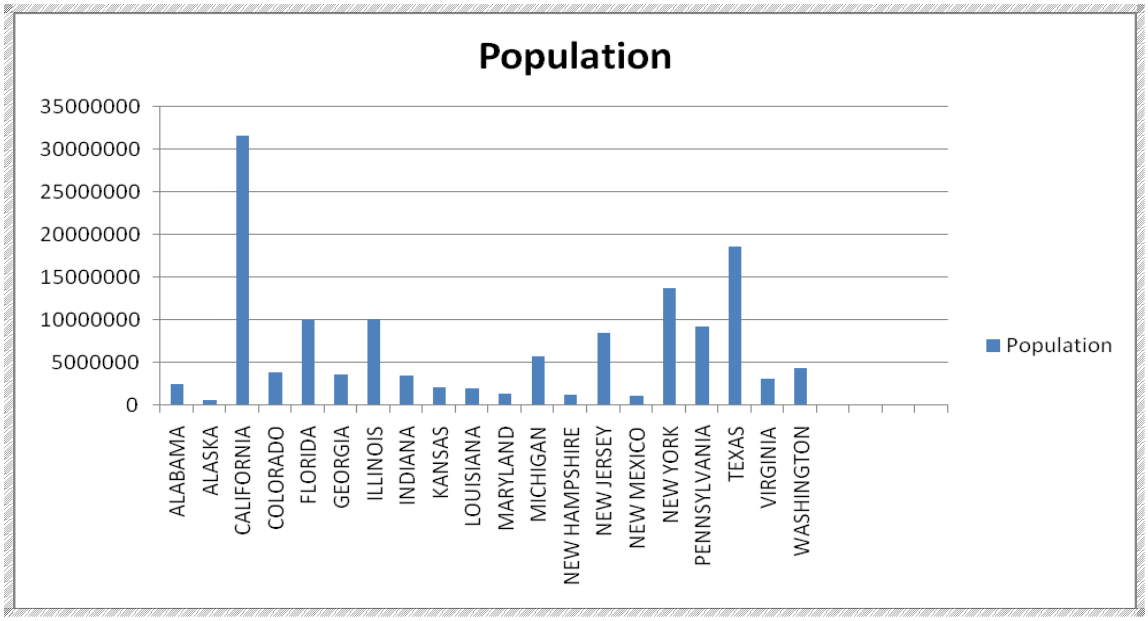


Figure 5: States Population Graph

Figure 6 is showing comparisons between total no of murders and total no of violent crimes in different states with respect to population. By the given graph, we can easily interpret that violent crime has a high rate as compared to murder. Murder rate is almost equal in every state, but violent crime is at the peak in California, then reduced a little in Texas, New York and Florida respectively and so on. The point to

remember is that although the murder rate is the same in every state, but the violent crime rate is high, specifically in highly populated states.

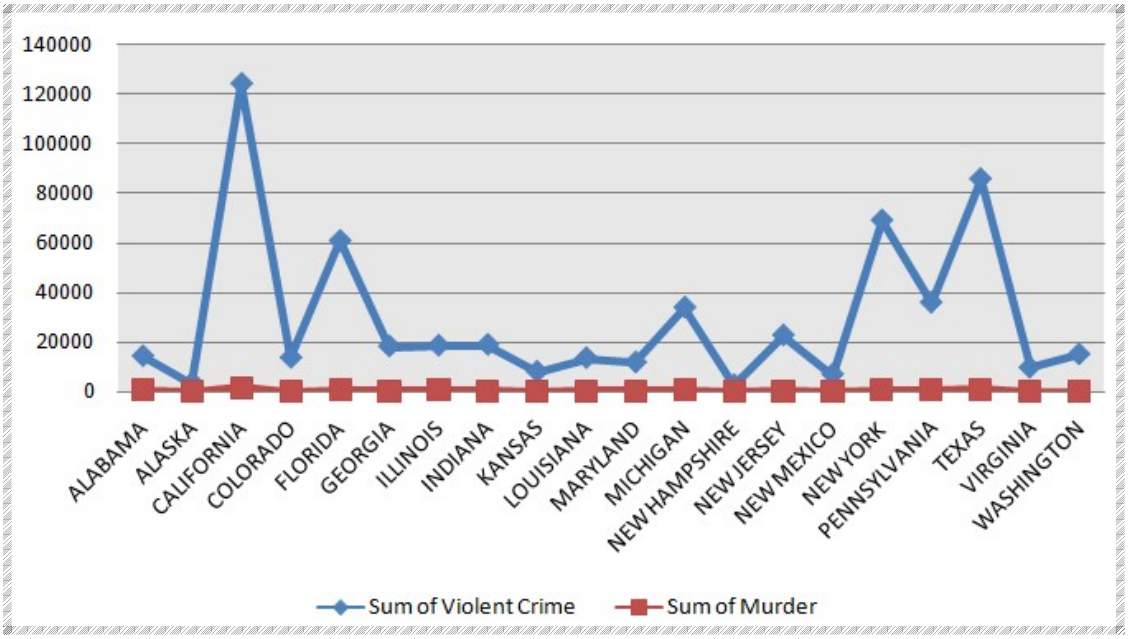


Figure 6: Graph of Total Murders and Violent Crimes in Different States

Figure 7 is showing comparison between total no of rape cases and total no of robbery cases in different states with respect to population. By the given graph, we can easily interpret that total no of robbery cases are high as compare to total no of rape cases. Total no of rape cases has few fluctuations in three states, but total no of robbery cases are high specifically in Texas, New York, California and Florida respectively and so on. Again the point to remember is that although the rate of rape crime is almost the same in every state but the rate of robbery crime is high specifically in highly populated states.

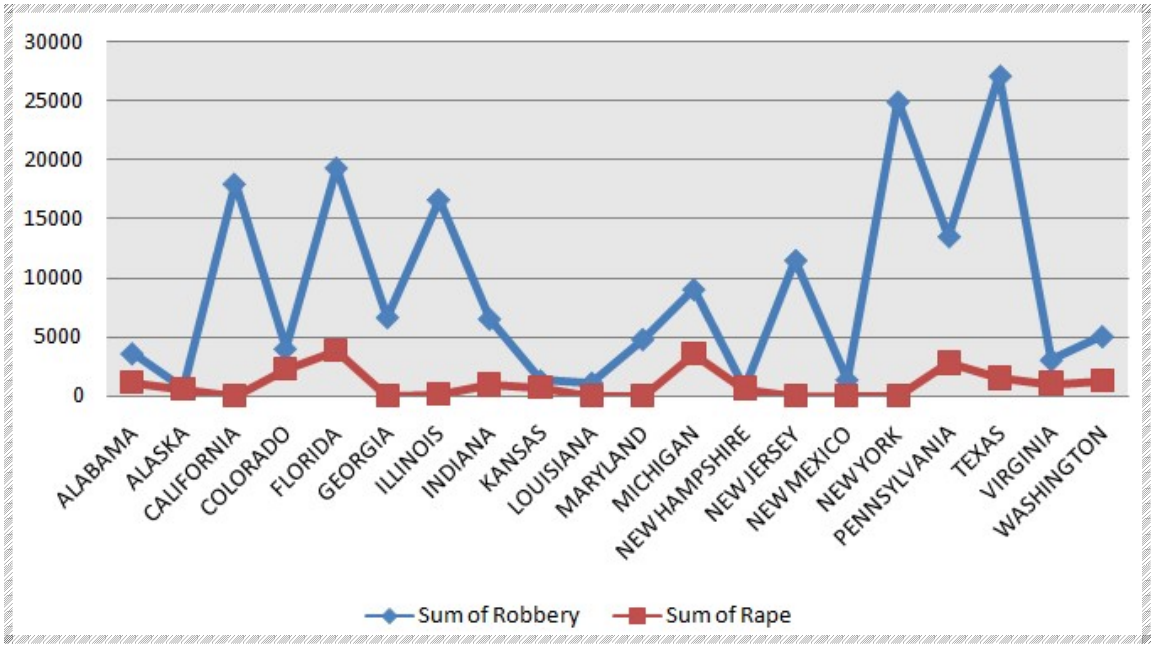


Figure 7 : Graph of Total Robberies and Rape in Different States

Figure 8 is showing comparison between total no of property crimes, total no of larceny-theft and total no of motor vehicle theft in different states with respect to population. By the given graph, we can easily interpret that both crimes total no of property crime and total no of larceny-theft has high rates as compared to total no of motor vehicle theft. Rate of motor vehicle theft crime is almost equal in every state except California state in which it is high. But the rate of both crimes total no of property crime and total no of larceny-theft are at peak in California, then reduced a little in Texas, New York and Florida respectively and so on. The point to remember again is that although the rate of motor vehicle theft crime is almost same in every state except California state but rate of both crimes total no of property crime and total no of larceny-theft are high specifically in highly populated states.

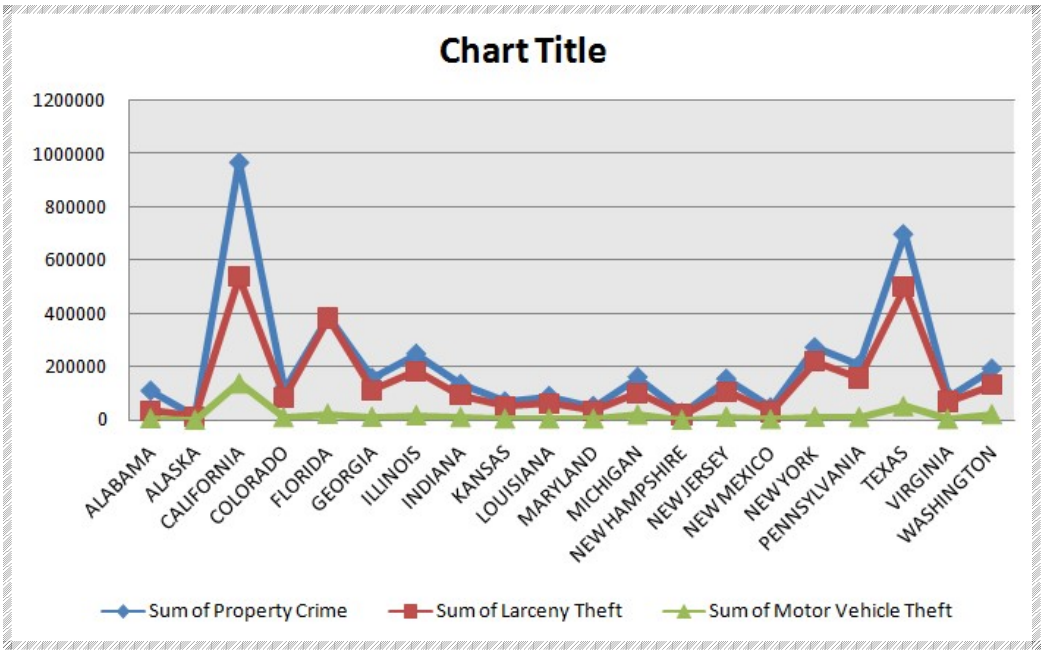


Figure 8: Graph of Total Property Crimes, Larceny Theft and Motor Vehicles Theft in Different States

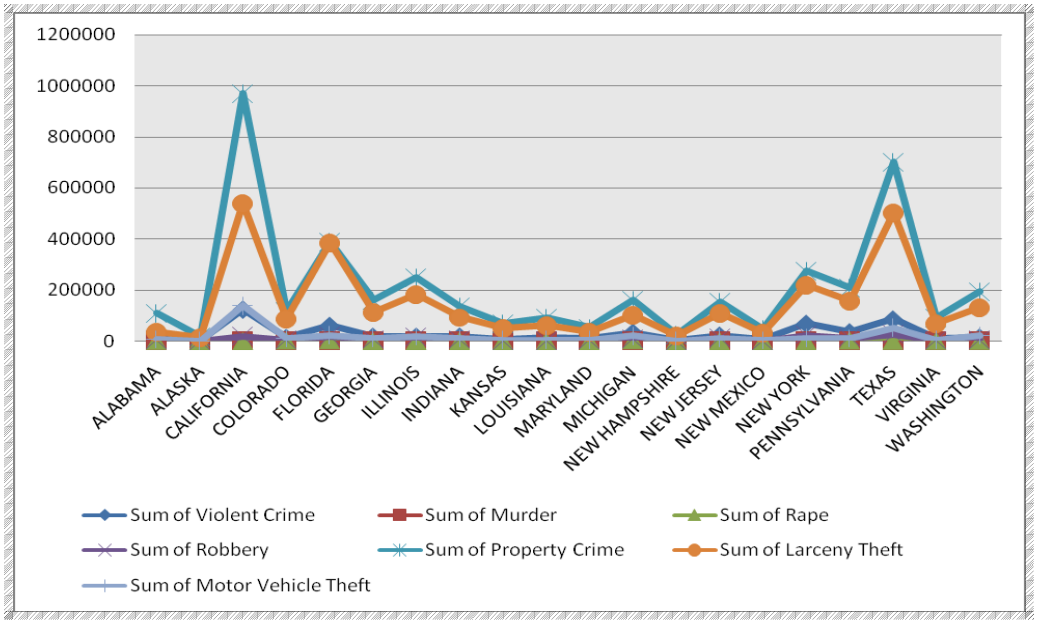


Figure 9: Graph of All Crimes Committed in Different States

The Figure 9 is also providing an overview of all crimes rate in different states with respect to population. Rate of property crime and the rate of larceny-theft is high amongst all other types of crime rate in populated states. Analysis result also showing that the crime rate is high in populated states. It could have two main reasons that either man force of law enforcement agency is less in those states or either they are not performing their duties so well.

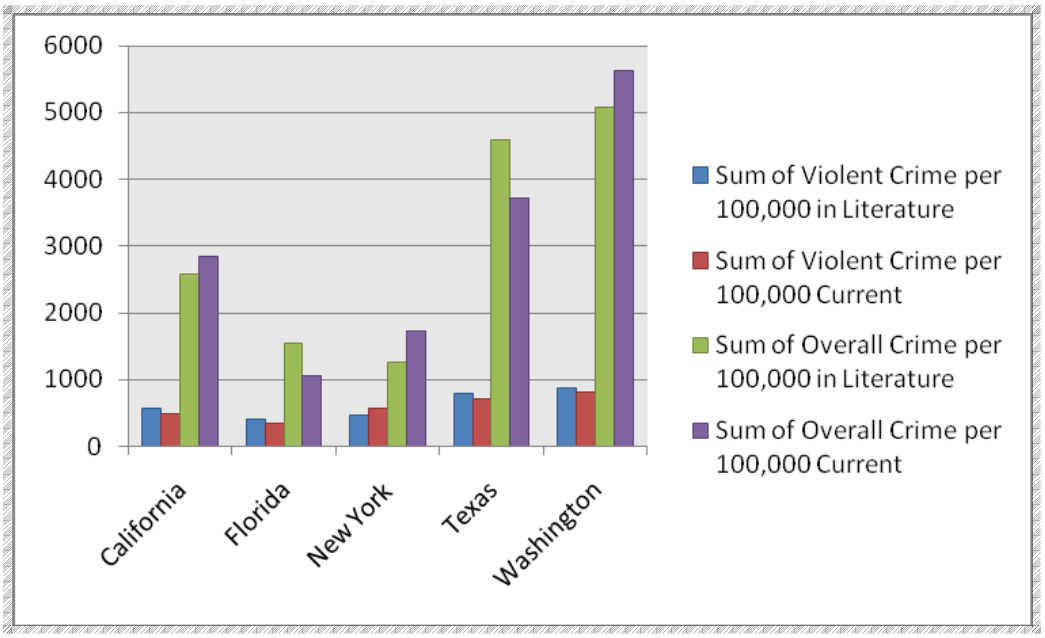


Figure 10: Result Comparison of violent crime and overall crime

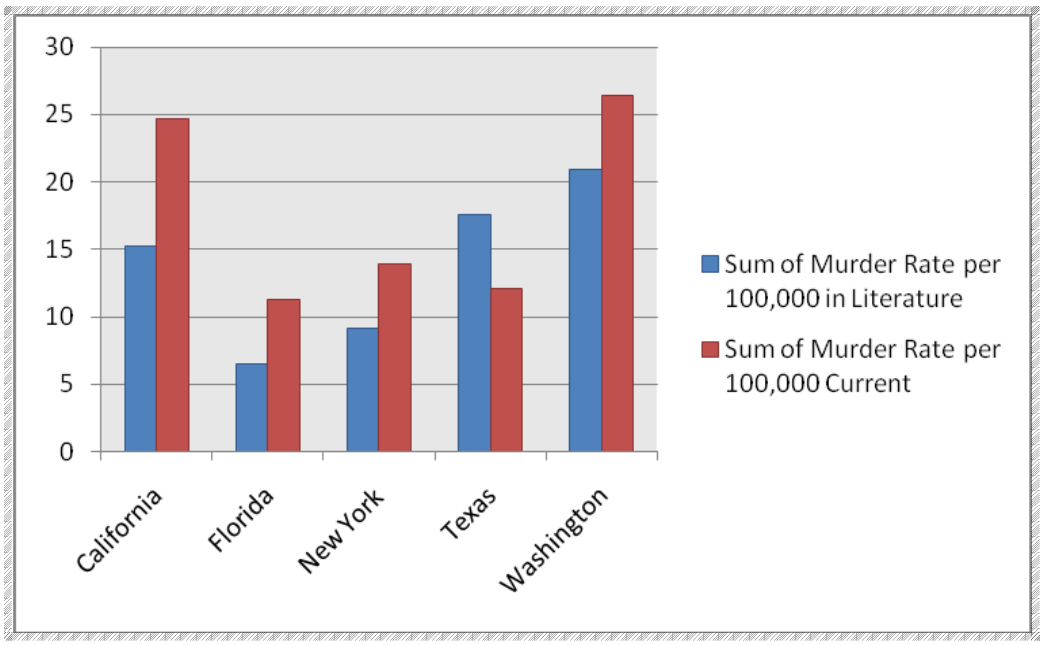


Figure 11: Result Comparison of murder crime

Figure 10 and Figure 11 are showing comparisons of our results with literature results. We found stats of overall crime rate per 100,000 people, the violent crime rate per 100,000 people and the murder rate per 100,000 people in literature. Statistics are of five common states found in literature and in our analysis

also. We compared those stats with our stats to show a comparison of how much our results are deviating from previous results. By the results, there are slight differences between literature and in our results. Finally, by results, we can conclude that our results are accurate and reasonable as there is no major difference between our and literature stats.

V. Conclusion and Future Work

The analysis of the data collection of the FBI crime provides the insight for identification of the patterns. For this purpose, Weka which is the primary data mining tool is used. First, we choose the classification technique and applied six classifier algorithms, namely (J48 Pruned Tree, REP tree, Decision Table, JRIP rules, Random Tree, naïve Bayes) on our dataset to find competitive results. We interpreted Algorithm's result based on different evaluation parameters and in precise accuracy (weighted average) form. We performed an analysis of different crimes in different states based on population. Experimental result showed that two classifier algorithms (REP Tree, Naïve Bayes) produced better results as compared to other algorithms. However, most of the results of Classifier algorithms were same, so we concluded that our dataset had been classified correctly. Analysis result showed that the crime rate is high in populated states. It could have two main reasons that either man force of law enforcement agency is less in those states or they are not performing their duties so well. The outcome from the analysis shows that rate of property crime and rate of larceny-theft is high amongst different crime rates are high in populated states. Rate of property crime and the rate of larceny-theft is also high amongst all other types of crime rate in populated states. For evaluation purpose, a test mode 10-Fold (Cross-Validation) has been used. Agencies responsible for Law enforcement, can take great advantage from this analysis, by using this type of machine learning algorithms like "Naïve Bayesian or REP Tree". They can further make future strategies based on our analysis to counter measure crime before its occurrence, and they can point out specific crime areas as well. In general, the primary purpose of our analysis was to help law enforcement agencies to analyze, predict and counter measure the crime more accurately and precisely. For future research, we have a plan to use of crime datasets of previous 10-15 years, to perform compact analysis and to predict future patterns. Furthermore, we can analyze crime at the city level based on which intelligent investigation techniques can be developed for crime pattern and prediction. In future the trained model is used for the analysis of crime data in Pakistan.

References

[1] Alkhaibari, A.A. and P.-T. Chung. *Cluster analysis for reducing city crime rates*. in *Systems, Applications and Technology Conference (LISAT), 2017 IEEE Long Island*. 2017. IEEE.

[2] Camacho, D., et al., *The four dimensions of social network analysis: An overview of research methods, applications, and software tools.* Information Fusion, 2020. 63: p. 88-120.

[3] Slack, J. and D.E. Martínez, *Postremoval Geographies: Immigration Enforcement and Organized Crime on the US–Mexico Border.* Annals of the American Association of Geographers, 2020: p. 1-17.

[4] Orraca-Romano, P.P. and E.D. Vargas-Valle, *Drug-related violence and the decline in the number of Mexican cross-border workers.* Review of Development Economics, 2020. 24(2): p. 485-502.

[5] Malathi, A., S. Babboo, and A. Anbarasi. *An intelligent analysis of a city crime data using data mining*. in *International conference information electronic engineering*. 2011.

[6] Yao, J. and Y. Ye, *The effect of image recognition traffic prediction method under deep learning and naive Bayes algorithm on freeway traffic safety.* Image and Vision Computing, 2020. 103: p. 103971.

[7] Wheelus, C. and X. Zhu, *IoT Network Security: Threats, Risks, and a Data-Driven Defense Framework.* IoT, 2020. 1(2): p. 259-285.

[8] Dogar, I.A., et al., *Prevalence and risk factors for depression and anxiety in hospitalized cardiac patients in Pakistan.* Psychiatry (Edgmont), 2008. 5(2): p. 38.

[9] Monuteaux, M.C., et al., *Firearm ownership and violent crime in the US: an ecologic study.* American journal of preventive medicine, 2015. 49(2): p. 207-214.

[10] Janke, K., C. Propper, and M.A. Shields, *Assaults, murders and walkers: The impact of violent crime on physical activity.* J Health Econ, 2016. 47: p. 34-49.

[11] Widner, B., M.L. Reyes-Loya, and C.E. Enomoto, *Crimes and violence in Mexico: Evidence from panel data.* The Social Science Journal, 2011. 48(4): p. 604-611.

[12] Babakura, A., M.N. Sulaiman, and M.A. Yusuf. *Improved method of classification algorithms for crime prediction*. in *Biometrics*

and Security Technologies (ISBAST), 2014 International Symposium on. 2014. IEEE.

[13] Wu, J., et al. *Mining the crime survey to support crime profiling*. in *Computational Intelligence for Multimedia Understanding (IWCIM), 2016 International Workshop on*. 2016. IEEE.

[14] Brushwood, J., et al., *Property crime, earnings variability, and the cost of capital.* Journal of Corporate Finance, 2016. 40: p. 142-173.

[15] Chi, H., et al., *A decision support system for detecting serial crimes.* Knowledge-Based Systems, 2017. 123: p. 88-101.

[16] Sheikh, J., et al. *IST: Role of GIS in crime mapping and analysis*. in *Communication Technologies (ComTech), 2017 International Conference on*. 2017. IEEE.

[17] Baloian, N., et al. *Crime prediction using patterns and context*. in *Computer Supported Cooperative Work in Design (CSCWD), 2017 IEEE 21st International Conference on*. 2017. IEEE.

[18] De Alcaraz-Fossoul, J. and K.A. Roberts, *Forensic intelligence applied to questioned document analysis: A model and its application against organized crime.* Sci Justice, 2017. 57(4): p. 314-320.

[19] Zhuang, Y., et al., *Crime Hot Spot Forecasting: A Recurrent Model with Spatial and Temporal Information.* 2017: p. 143-150.

[20] Hassani, H., et al., *Text mining in big data analytics.* Big Data and Cognitive Computing, 2020. 4(1): p. 1.

[21] Onnoom, B., et al. *An ontology framework for recommendation about a crime scene investigation*. in *Communications and Information Technologies (ISCIT), 2014 14th International Symposium on*. 2014. IEEE.

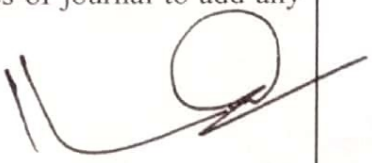
[22] Bharathi, A. and R. Shilpa, *A Survey on Crime Data Analysis of Data Mining Using Clustering Techniques.* International Journal of Advance Research in Computer Science and Management Studies, 2014. 2(8): p. 9-13.

Paper Titled: An Intelligent Analysis of Crime Data using Data Mining Algorithms

Certificate

The subject article has not been published or submitted to or accepted for publication in any form, in any other journal or conference etc. Subject article is our own original contribution.

We also guarantee that the authorship of this article will *not* be contested by anyone whose name(s) is/are not listed by us here and we will not request the authorities of journal to add any more author(s) after submission.



Corresponding Author

Authorship and Contribution Declaration

Sr.#	Author-s Full Name	Contribution to Paper	Signature
1	Mr. Shakeel Saeed (Main/principal Author)	Theme and idea of topic which includes representation theory & methodology developed regarding	 Signature by the Corresponding author on Behalf of Co-Authors
2	Dr. Muhammad Majid Mahmood Bagram (2nd Author)	Involved in collection of data & working software used to collect character Table	
3	Dr. Muhammad Munwar Iqbal (3rd Author)	Literature review of group theory and Contribution in theory of projective Symplectic group	