



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume: 6      Issue: IV      Month of publication: April 2018**

**DOI: <http://doi.org/10.22214/ijraset.2018.4468>**

**[www.ijraset.com](http://www.ijraset.com)**

**Call:  08813907089**

**E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)**

# “Performance Evaluation of Cluster Based Techniques for Zoning of Crime Info”

Ms. Ashwini Gharde<sup>1</sup>, Mrs. Ashwini Yerlekar<sup>2</sup>

<sup>1</sup>M.Tech Student, RG CER, Nagpur Maharashtra, India

<sup>2</sup>Asst. Prof, Department of Computer Science and Engg. RG CER, Nagpur Maharashtra, India

**Abstract:** *There has been an enormous increase in the crime in the recent past. Crime deterrence has become an upheaval task. There is need for user interactive interfaces based on current technologies to give them the much needed edge and fulfil the new emerging responsibilities of the police. The proposed interface is used to extract useful information from the vast crime database maintained by National Crime Record Bureau (NCRB) and find crime hot spots using crime data mining techniques such as clustering etc. In India, the regional location has a powerful impact on criminal activity. The crime profiling and zoning can be modeled with utilization of data mining. In this paper, we make cluster analysis by using k-means cluster algorithm on criminal dataset of India. The cluster input is used to create custom India map with the cluster zones of states. In this paper we are also providing comparative analysis of crime rates in different states of India using KNN and Hybrid algorithm. The custom maps displays an overall crime profiles of states which helps police and law enforcement department to take additional preventive measures to combat against the crime and plan advanced investigation strategies.*

**Index Terms:** *Data mining, crime profiling, clustering, k-means, KNN classifier, Hybrid classifier*

## I. INTRODUCTION

India is a vast country with more than one billion populations, and has a police force of 1.5 million. Police is a critical component of civil administration in India. It has created its own executive apparatus to discharge assigned responsibilities. Indian constitution assigns responsibility for maintaining law and order to the states and territories, and almost all routine policing, including apprehension of criminals, is carried out by state-level police forces. The police functioning have remained a constant area of governmental concern and efforts to improve it upon further and further. Way back in 1986, the Government of India created National Crime Record Bureau (NCRB). In this paper we have used the dataset of NCRB for determining the crime rates in India and also for determining the high rated crime zone. NCRB provided the data according to population of male, female and juvenile in different states and their crime rates according to their respective population. Here we are using NCRB dataset of year 2015. This dataset provide the details about the crimes in different cities. Dataset is very vast and have many crimes but in this paper we have used the basic three crimes i.e. Murder, Rape and human trafficking. These three crimes are highly occurring crimes in India so we have chosen these crimes.

Sr.no	Crime Name in Adult	Formula
1.	Rate of murder in male	Rate=No. of murder/population of male
2.	Rate of Attempt to Murder in male	Rate=No. of Attempt to murder in male/population of male
3.	Rate of rape in Female	Rate=No. of rape/ population of Female
4.	Rate of Gang Rape in female	Rate=No. of gang rape / population of Female

- 1) Murder has 3 basic sub categories such as: Not amounting to murder, Murder and attempt to murder.
- 2) Rape has 3 basic sub categories such as : Gang Rape, Commit to rape and other kinds of rape
- 3) Human trafficking has 2 sub categories such as: Immoral trafficking and Procuration of girls.

We have used above 3 categories and their sub categories for determining the crime rates in Male, Female and Juvenile in different states.

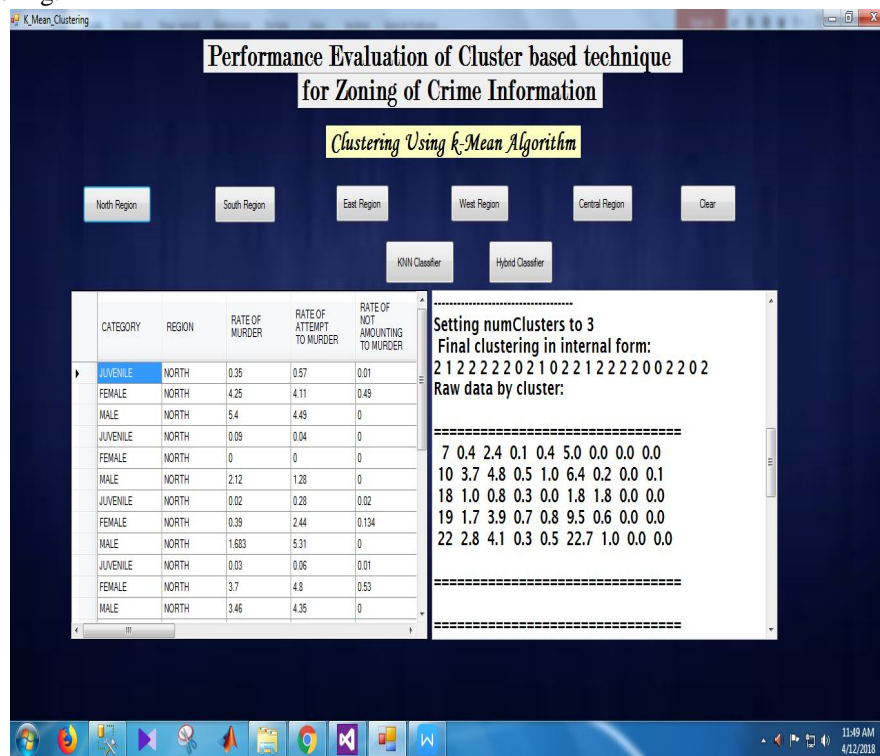
Sr.no	Crime Name in juvenile	Formula
1.	Rate of murder in juvenile	Rate=No. of murder/population of juvenile
2.	Rate of rape in juvenile	Rate=No. of rape/ population of juvenile
3.	Rate of immoral trafficking	Rate=No. of immoral trafficking in juvenile/ population of juvenile

**A. Background**

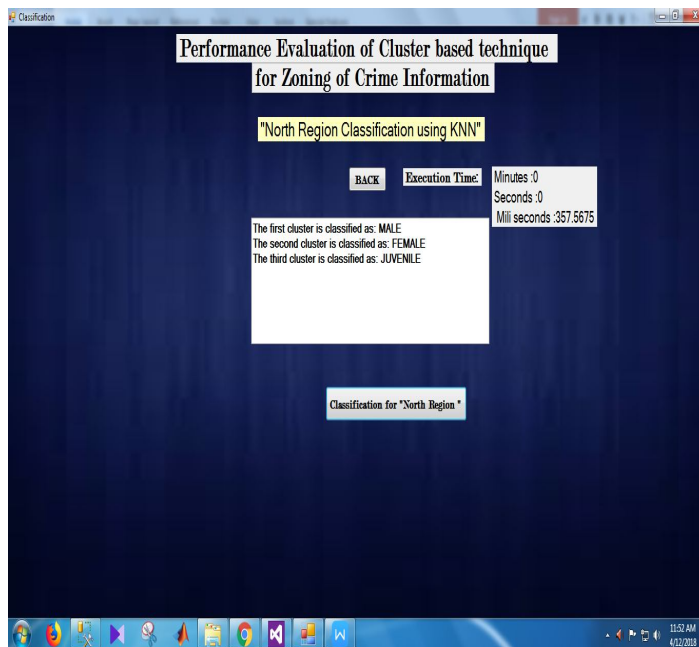
Now the collected dataset need to be preprocessed for finding the clustering in different regions according to the population of India in different genders of adult and juvenile. For preprocessing of dataset we find the crime rates according to the population of Male, Female and Juvenile. In preprocessing we have used formula for finding the rates. These formulas are displayed in below tables: Please not that other subcategories of murder, rape and human trafficking are also considered. These preprocessed data is maintained in excel sheet and then this data is sent to the database for finding the results.

**B. Cluster Formation**

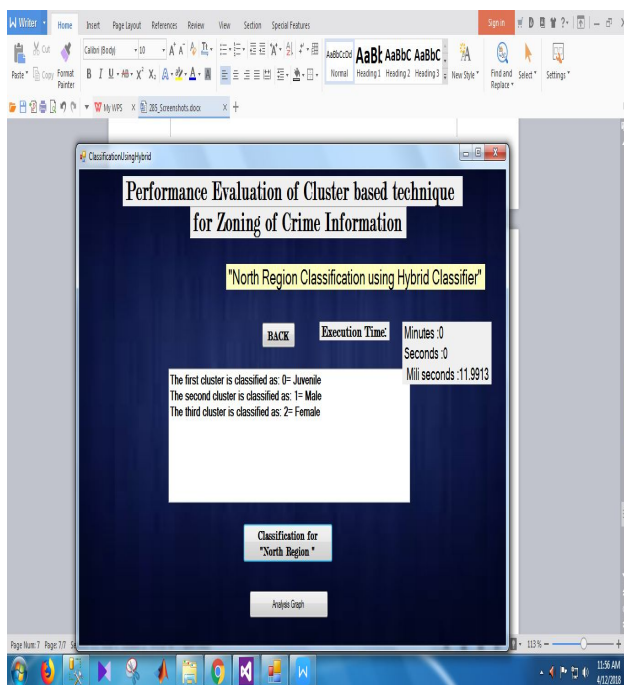
Once the rate has been calculated our preprocessed data is categories by crime rates of Male, Female and juvenile in different states and also for different regions. And finally the data has been preprocessed by allocating every region with crime rates for male female and juvenile category. Now the output of preprocessed dataset is used for clustering using k-means algorithm. This k-means algorithm uses the Euclidean distance for finding the centroid and after finding the centroid gives the clusters. After applying the K-means algorithm we get the clustering of different regions in male, female and juvenile population. Following figure shows the output of K-means clustering.



After this we need to classify the clusters i.e clusters belongs to which category i.e. male, female and juvenile in a particular region. For the classification we have used the K-NN classifier. This K-NN classifier is based on the nearest neighbor distance. This classifier classifies the clusters after applying K-means algorithm. Following fig shows the results of KNN classifier (For North Region) which have identified all three clusters. First as Male, Second as Female and third as Juvenile.



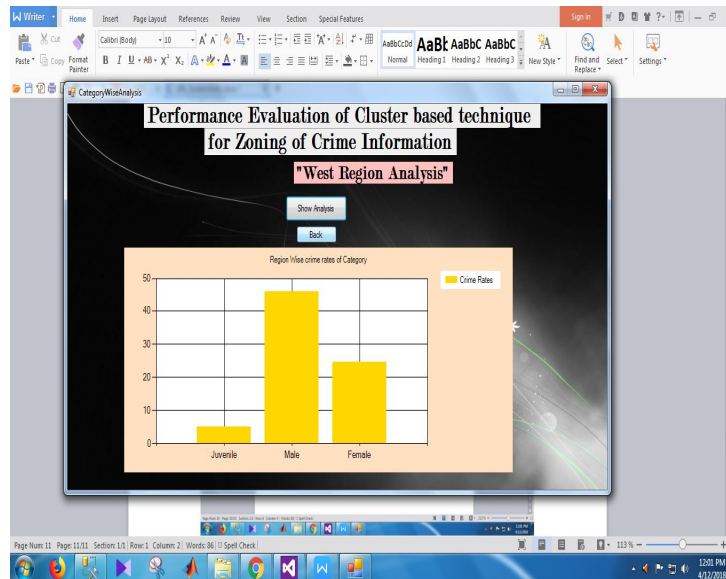
Now for proposed work we have used the Hybrid classifier which is based on all the classifier algorithm i.e. K-NN and SVM. But the execution time for the classification of Hybrid algorithm is better than the K-NN algorithm. Following fig shows the output of Hybrid classifier that is used for classification for getting the better result.



### C. Identification of High risk crime zone

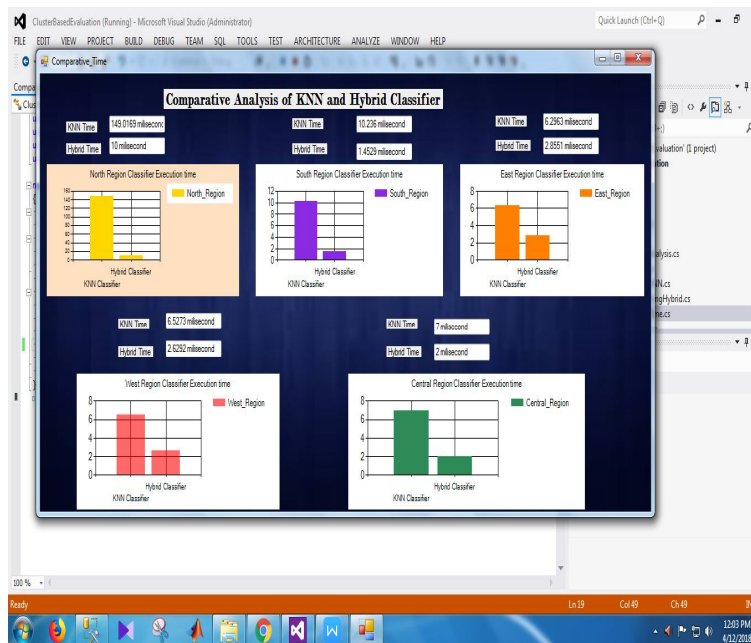
For the final output we have given the analysis graph for determining the highly rated crime regions. The analysis graph is shown according to different regions and in this graph it gives the information about the highly involved population in crime on the basis of male, female and juvenile. Following figure shows the analysis graph of west region. In west region male is highly involved in crime.





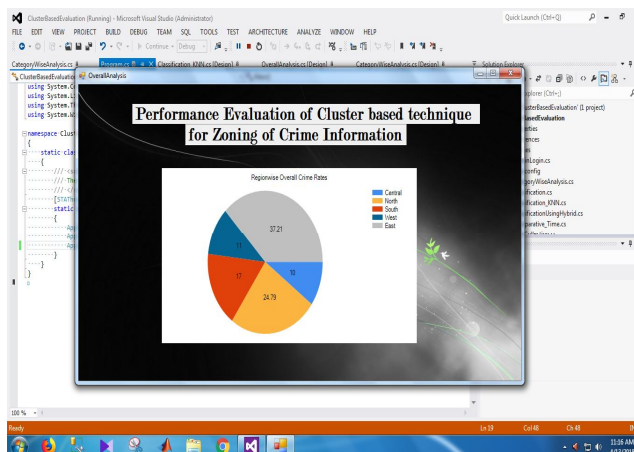
#### D. Comparative Analysis and Performance Evaluation

We have compared KNN and hybrid classifier and their performance evaluation is calculated by the parameter execution time. The result shows that the hybrid classifier has less execution time than the KNN classifier and both the classifier gives the same result. Following fig shows that comparative analysis of both the classifier and their performance evaluation is calculated depending on the execution time.



## II. CONCLUSION

In this paper, we have concluded that the east region is the highly crime rated zone and this information is helpful for police to analyze the crime rates in a region. And it is also helpful to give information about which gender population mostly involved in the crime. This result is calculated on the basis of k-means clustering method, KNN classifier and Hybrid classifier on the NCRB dataset. Following figure shows the final analysis of the all crime zones.



## REFERENCES

- [1] A Review of Different Clustering Techniques in Criminal Profiling, Adeyiga J.A,Bello A.O,Volume 6, Issue 4, April 2016
- [2] A Short Review, Bio-house Journal of Computer Science, Bio-house volume M Shiva Prasad, Asst. Professor. Issue 1, Jan-Feb 2016
- [3] A Survey of Data Mining Techniques for Analyzing Crime Patterns, Ubon Thongsata pornwatana Defence Technology Institute Nonthaburi, 2016 IEEE
- [4] Analysis and Prediction of Crimes by Clustering and Classification RasoulKiani ,Siamak Mahdavi, International Journal of Advanced Research in Artificial Intelligence,vol 4, No.8, 2015
- [5] Crime Analysis and Prediction Using Data Mining Shiju Sathyadevan, Devan M.S, 2014 IEE
- [6] Cluster Analysis for Anomaly Detection in Accounting Data :An Audit Approach, The International Journal of Digital Accounting Researce Sutapat Thiprungsri Miklos Vasarhelyi, Vol. 11, 2011
- [7] A Proposed Framework for Analyzing Crime Data Set Using Decision Tree and Simple K-Means Mining Algorithms KadhimB. Swadi al-Janabi, Journal of Kufa for Mathematics and Computer, Vol.1, No.3, may 2011
- [8] Crime Pattern Detection Using Data Mining Shyam Varan Nath Florida Atlantic university/ Oracle Corporation. Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology
- [9] Cluster Analysis for Reducing City Crime Rates Adel Ali Alkhaibari, Long Island University, Brooklyn, NY, *Student Member, IEEE*.
- [10] Mining Crime Data by Using New Similarity Measure Guangzhu Yu1, Shihuang Shao1, Bing Luo



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)