ISSN: 2321-2152

IJMECE

International Journal of modern electronics and communication engineering

E-Mail
editor.ijmece@gmail.com
editor@ijmece.com

www.ijmece.com

ISSN2321-2152
www.ijmece .com
Vol 7, Issue 3, 2019

CRIME RATE PREDICTION & ANALYSIS USING K-MEANS CLUSTERING ALGORITHM

Indira, Parvati Kadli, Manjula. S.D
Asst. Professor, Assco. Professor, Assco. Professor
indira.raj.06@gmail.com, kadli.parvati@gmail.com, manjulasd@pdit.ac.in
Department of CSE, Proudhadevaraya Institute of Technology, Abheraj Baldota Rd, Indiranagar, Hosapete, Karnataka-583225

ABSTRAT:

Every day, the crime rate in India rises. Offender success in committing crimes is now aided by contemporary methods, the impact of social media, and new technologies. Systematically classifying and examining crime trends is the basis of both crime analysis and prediction. While there are a number of clustering techniques available, they do not cover all bases when it comes to crime research and pattern prediction. The K-means algorithm outperforms the others in terms of outcome prediction. The planned study primarily aimed to identify the areas with the highest crime rates and the age groups that are more or less likely to engage in criminal behaviour. To increase efficiency and reduce time complexity, we provide an optimised K-means method.

Keyword: K means, Data set, crime analysis, crime rate.

I. INTRODUCTION

In present scenario criminals are becoming technologically sophisticated in committing crime and one challenge faced by intelligence and law enforcement agencies is difficulty in analyzing large volume of data involved in crime and terrorist activities therefore agencies need to know technique to catch criminal and remain ahead in the eternal race between the criminals and the law enforcement. So appropriate field need to chosen to perform crime analysis and as data mining refers to extracting or mining knowledge from large amounts of data, data mining is used here on high volume crime dataset and knowledge gained from data mining approaches is useful and support police forces. To perform crime analysis appropriate data mining approach need

to be chosen and as clustering is an approach of data mining which groups a set of objects in such a way that object in the same group are more similar than those in other groups and involved various algorithms that differ significantly in their notion of what constitutes a cluster and how to efficiently find them. In this paper k means clustering technique of data mining used to extract useful information from the high volume crime dataset and to interpret the data which assist police in identify and analyze crime patterns to reduce further occurrences of similar incidence and provide information to reduce the crime. In this paper k mean clustering is implemented using open source data mining tool which are analytical tools used for analyzing data .Among the available open source data mining suite such as R, Tanagra ,WEKA ,KNIME ,ORANGE ,Rapid miner.k means clustering is done with the help of rapid miner tool which is an open source statistical and data mining package written in Java with flexible data mining support options. Also for crime analysis dataset used is Crime dataset an offences recorded by the police in England and Wales by offence and police force area from 1990 to 2011-12 .In this paper homicide which is crime committed by human by killing another human is being analyzed .

II.LITERATURE SURVEY

Data Mining Approaches To Criminal Carrer Analysis De Bruin ,J.S.,Cocx,T.K,Kosters,W.A.,Laros,J. and Kok,J.N(2006) Narrative reports and criminal records are stored digitally across individual police departments, enabling the collection of this data to compile a nation-wide database of criminals and the crimes they committed. The compilation of this data through the last years presents new possibilities of analyzing criminal activity through time. Augmenting the traditional, more socially oriented, approach of behavioral study of these criminals and traditional statistics, data mining methods like clustering and prediction enable police forces to get a clearer picture of criminal careers. This allows officers to recognize crucial spots in changing criminal behaviour and deploy resources to prevent these careers from unfolding. Four important factors play a role in the analysis of criminal careers: crime nature, frequency, duration and severity. We describe a tool that extracts these from the database and

creates digital profiles for all offenders. It compares all individuals on these profiles by a new distance measure and clusters them accordingly. This method yields a visual clustering of these criminal careers and enables the identification of classes of criminals. The proposed method allows for several user-defined parameters.

Crime Data Mining for Indian Police Information System, Manish Gupta1*, B.Chandra1 and M. P. Gupta1,2007 .There has been an enormous increase in the crime in the recent past. Crime deterrence has become an upheaval task. The cops in their role to catch criminals are required to remain convincingly ahead in the eternal race between law breakers and law enforcers. One of the key concerns of the law enforcers is how to enhance investigative effectiveness of the police. 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 paper highlights the existing systems used by Indian police as e-governance initiatives and also proposes an interactive query based interface as crime analysis tool to assist police in their activities. 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. The effectiveness of the proposed interface has been illustrated on Indian crime records. An interactive interface as crime analysis tool has been designed for this purpose.

III.EXISTING SYSTEM

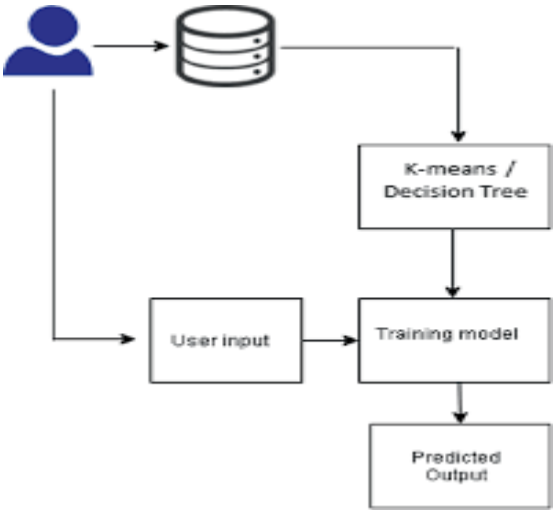
Crime analysis tool is developed using various distinct data mining methods. It supports the police officers for investigating crimes [6]. Implementing a clustering algorithm on crime datasets enables analysis of crimes [7]. It makes identification and analysis of various criminality trends over the years through their conclusion. The random initial starting points produced by K-means which gives results in the form of cluster that helps in reaching the local optima [8]. So to overcome this problem, the partitioned data along with the data axis with the highest variance for assigning the initial centroid for K-Means clustering was applied. So it is observed that the proposed technique uses a lesser number of iteration thereby reducing the clustering time. Using

ISSN2321-2152
www.ijmece .com
Vol 7, Issue 3, 2019

merge sort, K-means algorithm can be improved for clustering the Hidden Markov Model (HMM

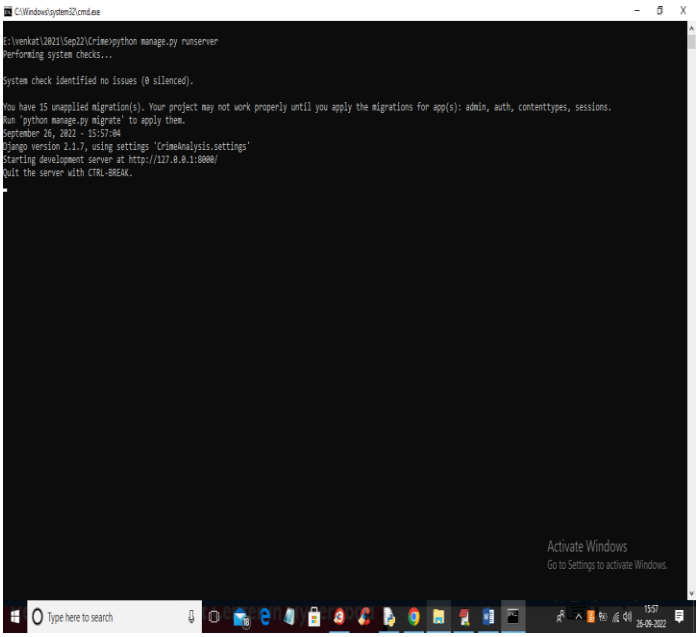
IV.PROPOSED SYSTEM

We are working on Spyder for implementation. Here we use a Spyder 3.7 version. Spyder is an integrated development environment for systematic programming in Python. Here we implemented different packages like matplotlib,numpy,sklearn, pandas,etc. Which helps to plot elbow graph and data frame table using a K-means clustering algorithm? Dataset is collected from Kaggle datasets and import datasets into Spyder in CSV format as shown in Fig 1. We perform normalization for finding the accurate number of clusters (k) using the elbow method. The elbow method performs k-means clustering on the obtained dataset for a range of values of k (2-15) and calculates the SSE. A line chart of the SSE is plotted for each value of k



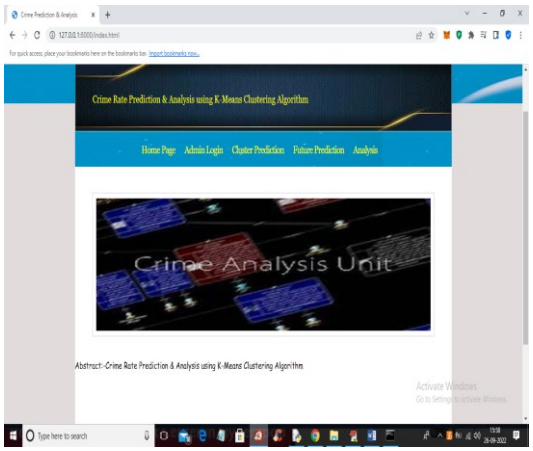
V.RESULTS ANALYSIS

Double click on run.bat file to start python server and get below screen

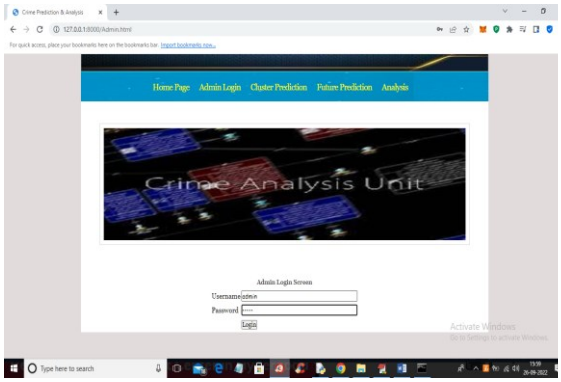


In above screen python server started and now open browser and enter URL as 'http://127.0.0.1:8000/index.html' and press enter key to get below page

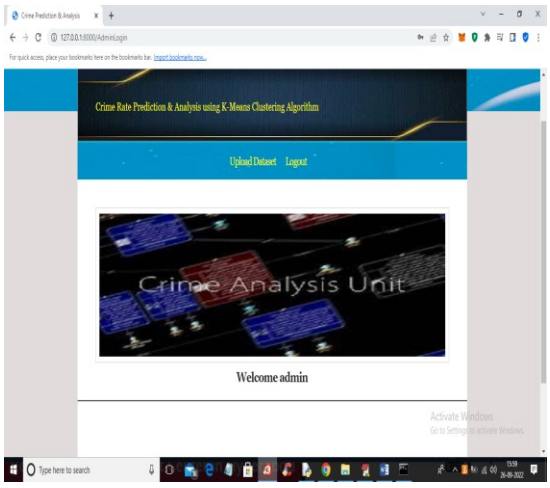
ISSN2321-2152
www.ijmece .com
Vol 7, Issue 3, 2019



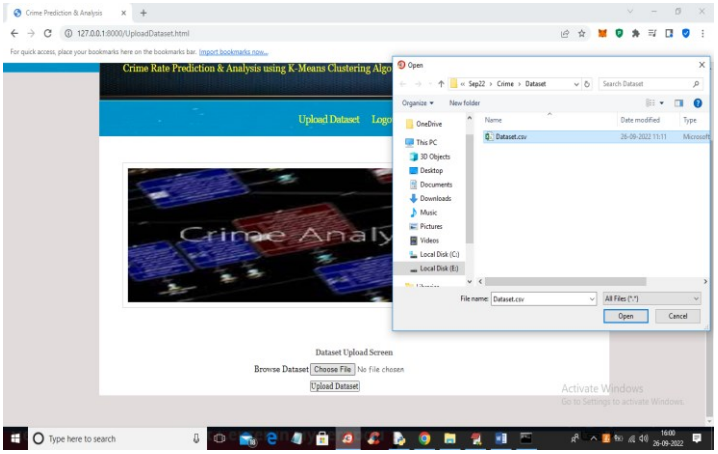
In above screen click on 'Admin Login' link to get below login page



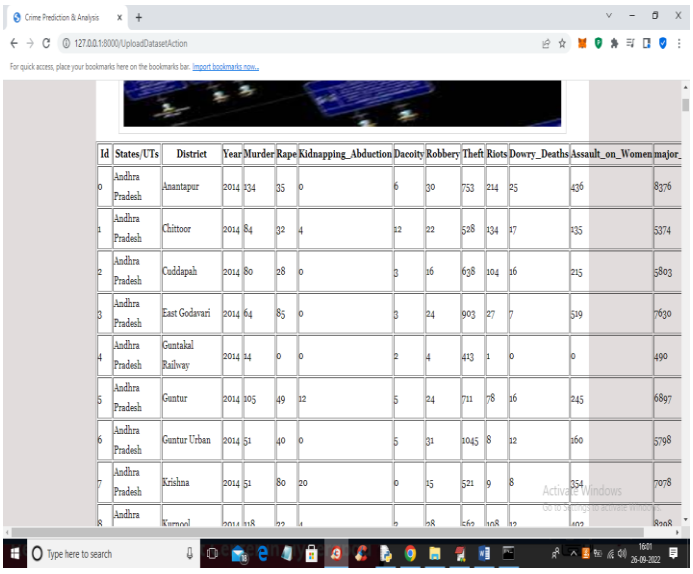
In above screen admin is login and after login will get below screen



In above screen admin can click on 'Upload Dataset' link to upload dataset and then click submit button to load dataset and then train it with machine learning algorithms



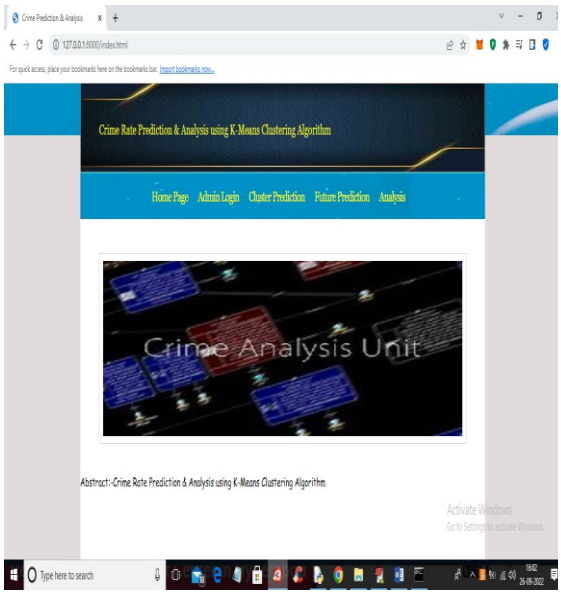
In above screen selecting and upload dataset and then click on 'Open' and 'Upload Dataset' button to load and complete training process and get below output



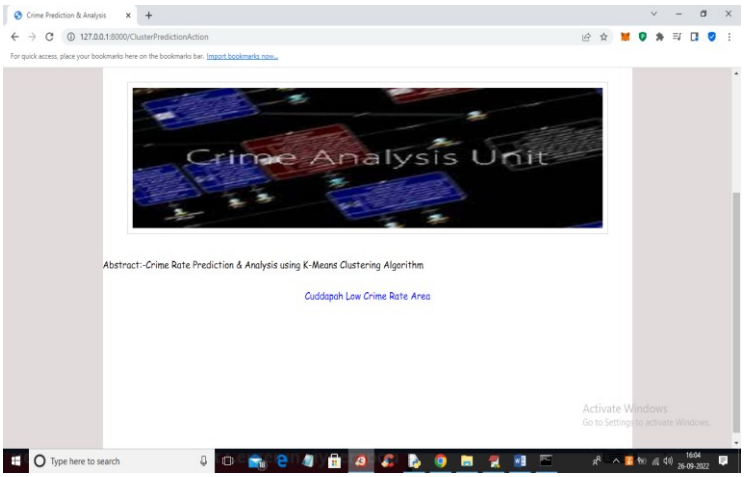
Id	States/UTs	District	Year	Murder	Rape	Kidnapping_Abduction	Dacoity	Robbery	Theft	Riots	Dowry_Deaths	Assault_on_Women	major_
0	Andhra Pradesh	Anantapur	2014	334	35	0	6	30	753	214	25	436	8376
1	Andhra Pradesh	Chittoor	2014	84	32	4	12	22	528	134	17	335	5374
2	Andhra Pradesh	Cuddapah	2014	80	28	0	3	16	638	104	16	215	5803
3	Andhra Pradesh	East Godavari	2014	64	85	0	3	24	903	27	7	519	7630
4	Andhra Pradesh	Guntakal Railway	2014	34	0	0	2	4	413	1	0	0	490
5	Andhra Pradesh	Guntur	2014	305	49	12	5	24	711	78	16	245	6897
6	Andhra Pradesh	Guntur Urban	2014	51	40	0	5	31	1045	8	12	160	5798
7	Andhra Pradesh	Krishna	2014	51	80	20	0	15	521	9	8	354	7078
8	Andhra	Kurnool	2014	108	22	4	2	28	562	108	12	402	8968

In above screen training is completed and then we got all dataset details and now click on 'Logout' link to get below screen

ISSN2321-2152
www.ijmece .com
Vol 7, Issue 3, 2019

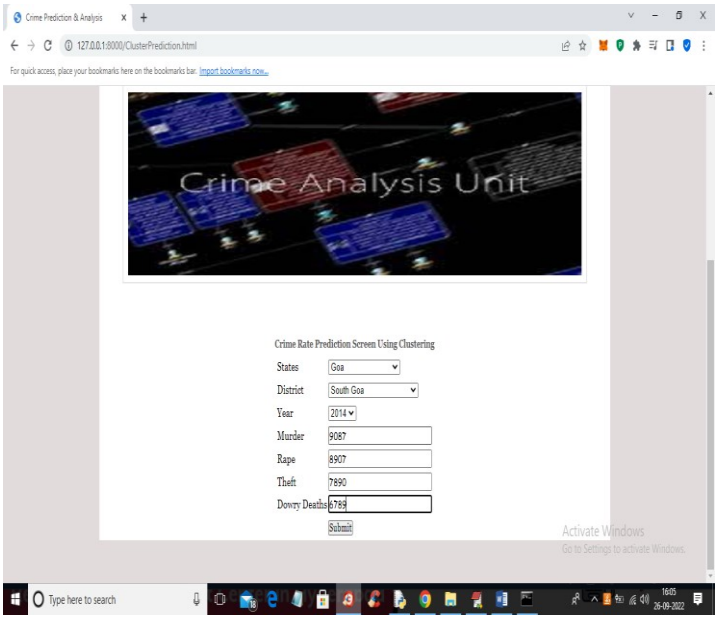


In above screen click on 'Cluster Prediction' link to get below screen



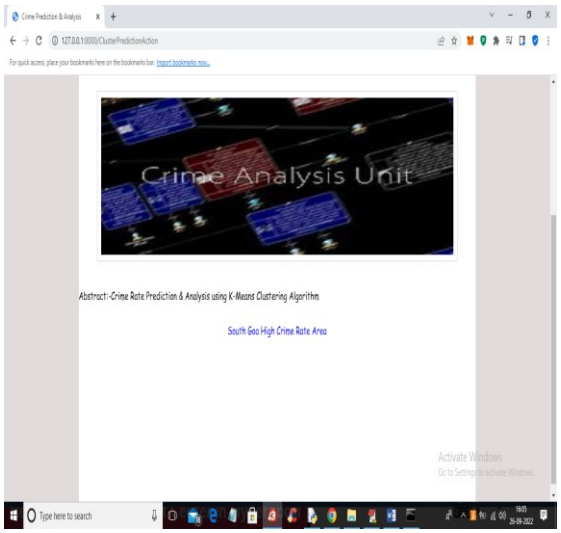
In above screen in blue colour we got output as 'Cuddapah is the Low Crime Area' and similarly we can test any other state

In above screen select state and district name and then enter details of crime and then press 'Submit' button to get below output

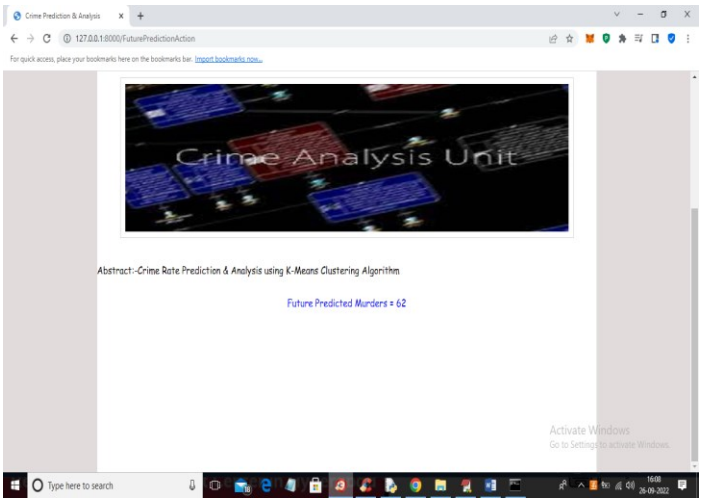


In above screen I entered some other state and crime rate and press button to get below output

ISSN2321-2152
www.ijmece .com
Vol 7, Issue 3, 2019

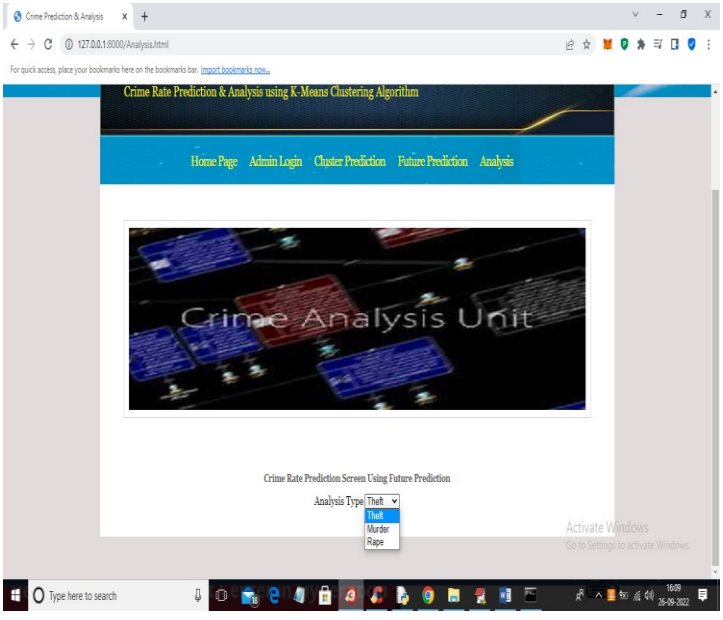


In above screen in blue colour text we got output as 'Goa is High Crime area' and now click on 'Future Prediction' link to get below screen



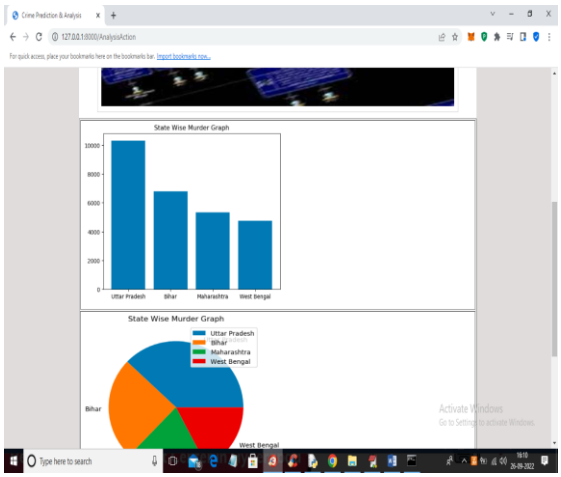
In above screen future predicted Murders for Gujarat state and Ahmadabad district is 62 and similarly you can select option and get future prediction and now click on 'Analysis' link to get below screen

In above screen I selected state, district and then select crime as 'Murder' and then press 'Submit' button to get below output



In above screen select the type of analysis and press button to get below graphs

ISSN2321-2152
www.ijmece .com
Vol 7, Issue 3, 2019



VI.CONCLUSION

Here, we conduct crime analysis by taking homicide into account and plotting it with respect to year. We conclude that homicide is decreasing from 1990 to 2011 by implementing a clustering algorithm on the crime dataset using the rapid miner tool.Crime trends over years may be easily seen from the clustered findings, which can then be utilised to develop future prevention strategies.

VII.FUTURE SCOPE

We are optimistic about the future of crime data mining as a tool to improve intelligence and criminal analysis because to the excellent outcomes. The development of visual and intuitive intelligence and criminal investigation approaches for crime pattern is possible. We may use additional data mining methods, such classification, since we have already used the clustering methodology for crime analysis. Enterprise survey, poverty, assistance effectiveness, and other datasets are all within our analytical capabilities.

VIII.REFERANCES

[1] De Bruin, J.S., Cocx, T.K, Kosters, W.A.,Laros,J. and Kok,J.N (2006) Data mining approaches to criminal carrer analysis ,"in Proceedings of the Sixth International Conference on Data Mining (ICDM"06) ,Pp. 171-177

[2] Manish Gupta1*, B.Chandra1 and M. P. Gupta1,2007 Crime Data Mining for Indian Police Information System

[3] Nazlena Mohamad Ali1, Masnizah Mohd2, Hyowon Lee3, Alan F. Smeaton3, Fabio Crestani4 and Shahrul Azman Mohd Noah2 ,2010 Visual Interactive Malaysia Crime News Retrieval System

[4] Sutapat Thirprungsri Rutgers University .USA ,2011 Cluster Analysis of Anomaly Detection in Accounting Data : An Audit Approach 1

[5] A.Malathi ,Dr.S.Santhosh Baboo. D.G. Vaishnav College,Chennai ,2011 Algorithmic Crime Prediction Model Based on the Analysis of Crime Clusters.

[6] Malathi.A 1, Dr. S.Santhosh Baboo 2 and Anbarasi . A 31 Assistant professor, Department of Computer Science Govt Arts College Coimbatore, India. 2 Readers, Department of Computer science, D.G. Vaishnav Collge, Chennai, India, 2011 An intelligent Analysis of a city Crime Data Using Data Mining

[7] Malathi A; Santhosh Baboo , S, 2011 An Enhanced Algorithm to Predict a Future Crime using Data Mining

[8] Kadhim B.Swadi al-Janabi. Department of Computer Science. Faculty of Mathematics and Computer Science University of Kufa/Iraq, 2011 A Proposed Framework for Analyzing Crime DataSet using Decision Tree and Simple K-means Mining Algorithms.

[9] Aravindan Mahendiran, Michael Shuffett, Sathappan Muthiah, Rimy Malla, Gaoqiang Zhang,2011 Forecasting Crime Incidents using Cluster Analysis and Bayesian Belief Networks

[10] Sutapat Thiprungsri,2012 Cluster Analysis for Anomaly Detection in Accounting Data: An Audit Approach1

[11] K. Zakir Hussain, M. Durairaj and G. Rabia Jahani Farzana, 2012 Application of Data Mining Techniques for Analyzing Violent Criminal Behavior by Simulation Model

[12] https://www.gov.uk/government/publications/offences-recorded-by-the-police-in-england-and-wales-by-offence-and-police-force-area-1990-to-2011-12