The Efficacy of Hotspot Mapping Techniques In Crime Prediction And Optimization

Aakriti Yadav, Nikita Verma, Saumya Rawat, Akshay Kumar, Deepak Bisht, Adarsh Ojha

Dronacharya College of Engineering yadavaakriti23@gmail.com, vermanikita1011@gmail.com, saumya04rawat@gmail.com, kumarishab007@gmail.com, bishtdeepak30016@gmail.com, oadarsh006@gmail.com

ABSTRACT

In hotspot policing, predictive hotspot mapping is significantly vital. Existing approaches, such as kernel density estimation (KDE), does not take the temporal component of crime into account. This paper provides a spatial-temporal paradigm for predicting hotspot mapping and evaluation, based on previous research in related disciplines. In this study optimization is performed structuring the individual data in a way to illustrate the discrete variety of crime zones to ease out the detection for the same. It automatically plots those spots on the map, highlighting hotspots that warrant further investigation. Each kind of map depicts particular crime rate using GIS. The statistics of the figures is spatially analyzed using graphical representation such as histogram and bar graphs and it automatically plots those spots on the map, highlighting hotspots that warrant further investigation.

The degree of precision is contrasted by crime type (murder, rape, theft, burglary and riots) and the mapping approach used to detect crime event clusters (kernel density estimation, spatial ellipses, grid thematic mapping and thematic mapping of Census generated areas).The suggested framework has four important elements that distinguish it from other work in this field:

(1) STKDE that is spatial-temporal kernel density estimation method.

(2) A optimization technique is used called likelihood cross-validation, which helps in choosing the most optimal bandwidths.

(3) To eliminate the false positives in density estimations, a statistical significance test is applied.

(4) To inspect predictive hotspots, prediction accuracy index is used.

KEYWORDS: crime hotspot mapping, crime prediction, spatial analysis, GIS, Hotspot, KDE (Kernel Density Estimation)

Number: 10.14704/nq.2022.20.7.NQ33202 **Neuro Quantology 2022; 20(7):1620-1628**

INTRODUCTION [1][3][4][7][9]

A crime refers to an offence or illegal act which is punishable by the government or other state of authority. One of the biggest challenges in this growing era is the increase in crime rate making it more intense and complicated. Crime has been one of the major public issues affecting the rate of human advancement as well as growth of economy of the country over the past few years. Therefore, to keep a control on the crime happening, analysis of the data of crime taking at particular place. The analysis is done one the basis of geographical location at which the crime took place. Hotspot mapping is one such method which is used to contribute for the same.

The worst part is that the agencies responsible for implementing laws of developing countries are yet to be computerise for productive documentation of records, evaluation of case, easy reference and retrieval, and storage of

information. This in time ahead would be responsible for giving authentic and user-friendly data when needed. Providing proper assistance in predicting the trends and impart the decision support for police agencies will be major roles involved in this. Location, time and process are some spatial qualities of crime. Hence, if this spatial details regarding locations prone to experience crime are easily voluntarily accessible and obtainable, it would count as a great help in contributing in the states successful policing. The problem statement focused by us is : given all crime events (time and location) in a police records management system (RMS), provides rank to a collection of geographic locations in a city according to the likelihood of occurrence of crime. We utilize different types of crime in distinct areas, and our model is applicable to other crime categories and leading indications as well.

There are numerous mapping approaches that can be used to identify and investigate crime patterns. These techniques can be as easy as depicting each crime incident as a point and looking at the geographic distribution of those points.

Both quantitative and qualitative methodologies can be used to analyse crime. In anticipating the future of criminal activity, qualitative tools such as scenario writing or environmental scanning are effective. Meanwhile, a quantitative method is applied to anticipate future crime rates. Furthermore, criminal analysis is a realistic method for analysing and identifying crime statistics.

Utilising functions in geographic information system for various purposes:

Shading administrative areas thematically (demographic regions and territory patrolled by police) or the distribution of crime on a uniform plane in regard to the density of volume of geographic distribution of crime.

GIS is efficient in amplifying the organisational integration. Further on it merge the technology and software data to acquire, evaluate and display all types of spatially connected information.

The purpose of GIS is to analyse data in such a manner that is clearly and quickly circulated in order to help in its assistance to answer questions and solve the arising issues regarding the same. GIS also permits for examining, questioning, inspecting, envision and interpretation of data in various distinct forms such as maps, charts, reports and globes in order to highlight the trends, linkage and patterns.

Several of these mapping techniques have been put through various assessments that have considered their mileage for hotspot. Still, these evaluations were somewhat of a visual depiction of each system and an exercise that estimated its convenience of use. Foremost, the reviews have shown that distinct hotspot mapping methodologies may produce different findings in terms of the size and form of hotspots, as well as their locations. Yet alone none of these reviews has shown that Which of these strategies is the most prominent for predicting where crime trends will occur in the future. It is considered important to examine the differences in the capacity of hotspot mapping methods to anticipate crime patterns spatially. This would help the evaluators to choose the hotspot mapping fashion that is best appropriate for their operation and allows them to decide some level of delicacy in the mileage of the fashion for forecasting future crime patterns.

However, their findings have certain limitations in terms of providing an accurate prediction of crime location. A vast number of research articles have already been published on this subject. As a result, in this paper, we go over each of them in detail and summarise the results. Our goal is to estimate current implementations of crime prediction methods as well as ways to optimize them for future demands.

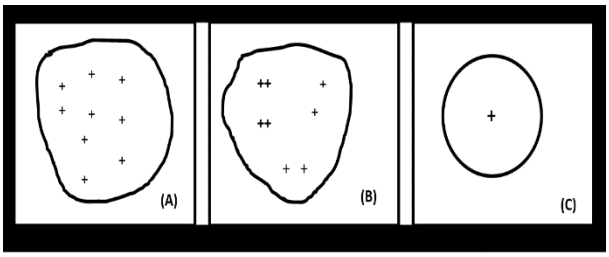
HOTSPOT MAPPING TECHNIQUES:

1. Spatial ellipses [1][3][9]:

STAC i.e., Spatial and Temporal Analysis of Crime became one of the first mapping software application to be widely available to analysts for

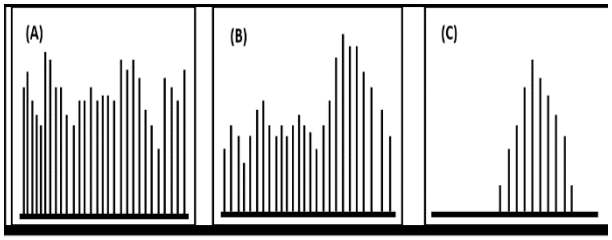
crime mapping. STAC acts as cooperation to users having desktop or GIS mapping potential. It's a tool which scrutinize hotspot areas on the map by determining heavy hot clusters and accordingly it places a 'standardized deviational ellipse' to each one. The nature of the ellipse is indicated by its size & alignment The nature of the ellipse is indicated by its size and alignment.

In North West England, links between disadvantages and crimes were explored by Bowers and Hirschfield (1999) using STAC; in 1998, Martin et al. used STAC during Detroit's infamous Devil's Night period to study reduce rates; Block and Block investigated hotspots around Chicago fast transit stations, whereas Langworthy and Jefferis investigated the impact of the school holiday time on the geographical distribution of crime. STAC was also widely used by Baltimore County to analyse a standard crime priorities (Bloch and Perry, 1993).



Three kinds of spatial hotspots. Each image displays 12 crime taking place points located in the hotspot
(1)dispersed 2)clustered/grouped/congregated (3)Hotpoint
In A the points are distributed all over the hotspot
In B the certain points can be seen clustered together even though they are all over the hotspot
In C all 12 points are assembled together at one point in the middle of the circular hotspot.

STAC is said to have several advantages, including the ability to generate hotspots without relying on established boundaries like police administrative boundaries or Census units, minimal number of parameters are required, and interoperability with most GIS systems. STAC, on the other hand, has been criticised for a number of reasons. To begin with, it's desirable for people who are familiar with how software works. There is minimal guidance for the newbies on how to choose proper parameter values, resulting in uncertainty and increased unpredictability in the consequences. Second, since crime hotspots do not naturally form portable ellipses, they may cause data to be misinterpreted. Eventually, the presentation of the STAC data invalidates any comparison with events that do not fit within the spatial ellipses.



The above shown are three temporal hotspot classes. in each figure, the vertical bars display the expected measured time the crime has been done. The three categories of temporal hotspots are:

1) DISPERSED 2) CONCENTRATED 3) ACUTE

In (A) it shows that there a prevailing risk between the time of day and night even if the graph of crime rises or falls.

(B)even though there is general level risk but the crime rate is significantly high in a particular evening time period.

(C)level of crime is severe in early evening time period only. Each bar represents one hour running left to right from 00:00-00:59 to 23:00- 23:59 hours or more simply hourly from midnight to midnight

2. Thematic mapping [4][5]:

Geographic border thematic mapping, often known as choropleth mapping, is a popular means of displaying crime patterns (Home Office, 2001). Boundary Thematic mapping zones are frequently arbitrarily determined for this form of mapping. They can be used for administrative or political purposes, such as police stations, census blocks, and so on. Districts or wards Offenses can be arranged together as points on a map. Areas that can be divided into geographical units be coloured based on the number of people in them crimes that fall within their jurisdiction According to Williamson et al., these maps are simple to create and interpret with minute technical skills. Furthermore, by 'zeroing in' on specific regions, the user can rapidly determine which regions have a high frequency of crime, allowing for more in-depth investigation of the issue. Furthermore, Census regions can be easily coupled with the help of other data references, they can use data like population to calculate a crime rate, enhancing their analytical adaptability.

Thematic shading can deceive map readers into detecting the existence of the greatest concentrations of crime due to the varying sizes and shapes of most geographical boundaries (Eck et al., 2005).

As a result, this may not be able to detect patterns across and inside border region of geographical divisions. The problem of modifiable area units (MAUP Openshaw, 1984), like any map based on established geographical boundaries, introduces additional complications.

Boundary changes themselves can have a direct impact on the models displayed on the map. Thematic mapping of border areas resumes to be widely used, from analysing different volumes of single and repeated thefts in different locations. Comparing car theft in connection to land use as well as for the research and presentation of crime trends and audits in all administrative zones of the partnership.

Thematic mapping of border areas continues to be widely applied. used to compare different volumes of single and repeat thefts over census tracts of a study area, to compare theft related to land usage in Overland Park, Kansas as well as criminal analysis and presentation and audit patterns throughout the partnership administrative regions (Ministry of Home Affairs, 2005; Chainey, 2001). MapInfo's basic procedures were used to do this. The choice of thematic coverage method and parameters for establishing crime hotspot thresholds are discussed here, as they relate to both KDE and grid thematic mapping. Mapping grids by subject: MapInfo (and many other GIS applications) requires a lot of work.

3. Grid thematic mapping [1][9]:

To eliminate the limitations of thematic mapping grid thematic mapping was introduced which can be drawn in GIS as a layer over the research area with uniform quadrants (or grids) that are thematically shaded. Hence, all areas used for topic shading are constant and of equal size, making hotspot identification easier and faster. This method was used as a part of GIS-based database application configured to recognize vulnerable neighbourhoods and has since been targeted for enhancement. In North Carolina, LeBeau used this approach to map emergency call volume and violent crime per square mile. This method has various drawbacks, such as the usage of grids, which limits how hotspots may be presented. Therefore, the spatial details of the space within and above each square are lost due to the criminal event subject to a particular quartile, which can lead to misunderstandings by map users. Furthermore, several people have commented on how grid cell Cut affects the 'blocky' quality of this approach. By reducing the size of each tile, the method reduces the precision of the thematic map, making it seem "spotted," and may not give meaningful information about where criminals congregate. Ultimately, subject-mapped grids suffer from similar MAUP problems.

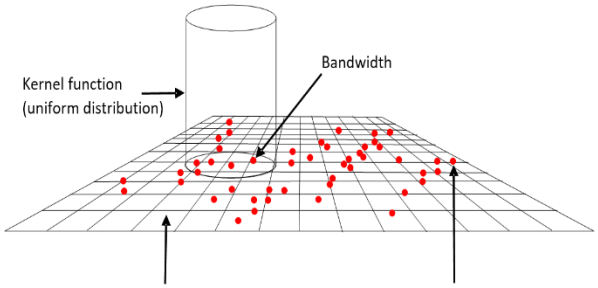
Thematic map consists of a target specific and a statistical map based on spatial variation.

- The correct value is not displayed.
- Provides background information about spatial patterns.
- It might be tough to tell the difference between two vivid colours.
- Variants in the map are hidden
- More data is needed for accuracy
- Symbol sizes can be biased. Density and,
- Scales are difficult to show on a map.

4. Kernel density estimation (KDE) [1][7]:

KDE is considered as one of the most appropriate interpolation methods for displaying crime statistics. Because of its increasing accessibility, the Intuit accuracy of hotspot detection and visual appeal of the map generated, this has become a more popular method. A user-specified search radius is used to aggregate point data (offences)representing the volume or density of crime occurrences throughout the desired area over a continuous surface is calculated. There's no need to stick to geometric shapes like ellipses in order to create a smooth surface map that shows the fluctuations in point/crime density across the research region. KDE requires two parameters which has to be entered by the user that are bandwidth (search radius) and the cell size. Yet, despite numerous beneficial, there is no general guideline for establishing them and when they should be executed.

KDE's use is now well documented, with popular crime mapping literature displaying several examples. KDE does not come without flaws, Eck et al. (2005) point out that deciding which thematic range to employ is still difficult since agencies fail to analyse the map's validity or statistical consistency, instead falling prey to its 'visual appeal.'



This has a significant impact on how hotspots are discovered, as well as the variety of maps created from the data. According to certain concerns, minimal amounts of data might possibly mislead the map reader. Nonetheless, KDE method is popular right now, not because itsthe most visually appealing, but also because it can locate hotspots using a statistically relevant methodology.

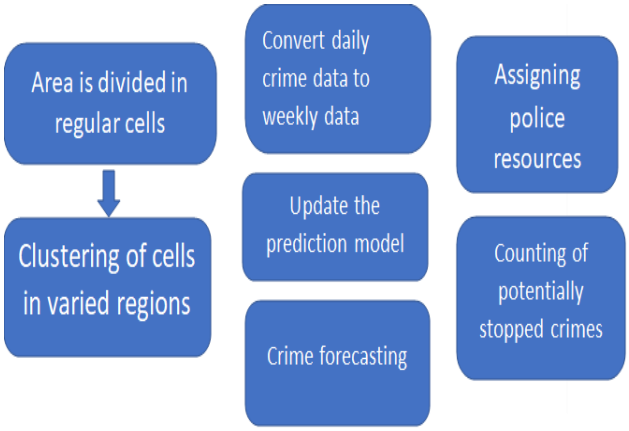
RESEARCH OBJECTIVES:

There were two key goals for this study:

- In order to find the accuracy of hotspot maps in predicting crime trends that differ from other forms of crime.
- Performing analysis that helps to compare the differences in the capability of hotspot mapping approaches in order to determine crimes in future.

CRIME DATA AND METHODOLOGY [11][12]:

Methodological concerns for this study focused on how to best investigate a variety of mapping methodologies with data that was adequate for this analysis, utilizing a consistent (to allow for comparisons across results) and representative of what practitioners would do. MapInfo Professional mapping software was used to create hotspots and analyze their predictive potential, which included the usage of the 'Hotspot Detective' add-on application (Ratcliffe 2002). The Crime Stat program was also utilized since it includes a recent updated version of STAC, which was one of the mapping techniques used to construct hotspot maps in this study.

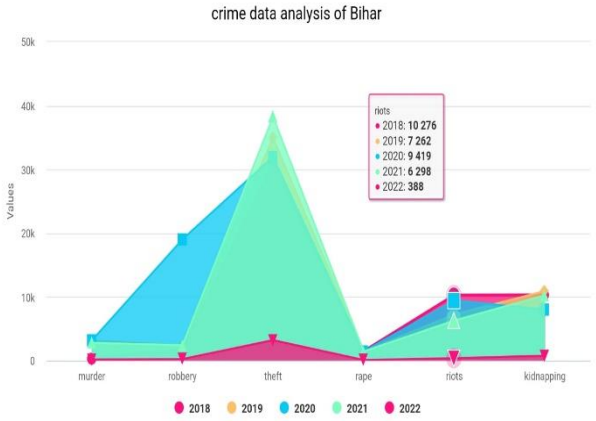


- *Study area of Bihar:*

the study area of Bihar state is located in eastern part of India consisting of 9 divisions 38 districts, municipal corporations, 49 Nagar Parishads and 80 Nagar Panchayats for administrative purposes this area consists of vast urban topography including 5 major stations (Patna, Patliputra, Bhagalpur, Katihar, Gaya junction), popular purchase areas.
Its ancient history and rich culture attracts a mob of natives as well as tourists. This massive attraction of tourists gives rise to crimes like theft, burglary etc. Due to lack of education in some areas, people tend to opt for unfair means of livelihood which adds to growing crime rates. Table 1 lists indicates each type of crime and the figures of number of crimes taken place in Bihar in 5 years(2018-2022)

Crime type	2018	2019	2020	2021	2022
Murder	293	3138	3149	2799	180
Robbery	1734	2399	19022	2421	247
Theft	30915	34970	31971	38277	3211
Rape	1475	1450	1438	1439	102
Riots	10276	7262	9419	6298	388
kidnapping	10310	10925	8004	10254	767

Table 1

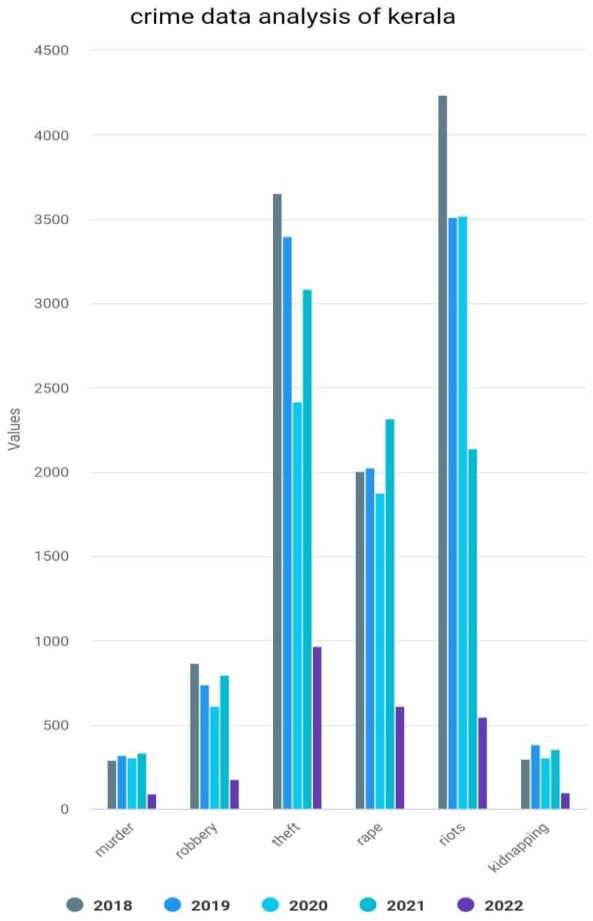


- *Study area of Kerala:*

If we talk about the quality & welfare of life, Kerela is considered as one of the most escalating states having highest literacy rates, high life expectancy & low child mortality. The date of crime is merged together to give various crime types like: residential theft and burglary (individual property), crime robbery in streets, vehicle or personal Vehicle theft. The merging of these crimes is grouped together and steadily evaluated by police and practitioners focused on reducing crime. Hence the indications regarding research are easily attainable and would be immediately converted in policing and crime reduction practice. Table 2 lists indicates each type of crime and the figures of number of crimes taken place in Kerala in 5 years (2018-2022).

Crime type	2018	2019	2020	2021	2022
Murder	292	323	306	337	92
Robbery	867	741	610	793	177
Theft	3651	3401	2418	3086	968
Rape	2005	2023	1880	2318	614
Riots	4236	3514	3520	2142	546
kidnapping	297	386	307	358	101

Table 2



PREDICTION ACCURACY INDEX (PAI) [2][6][8][10]:

- *Artificial neural networks:*

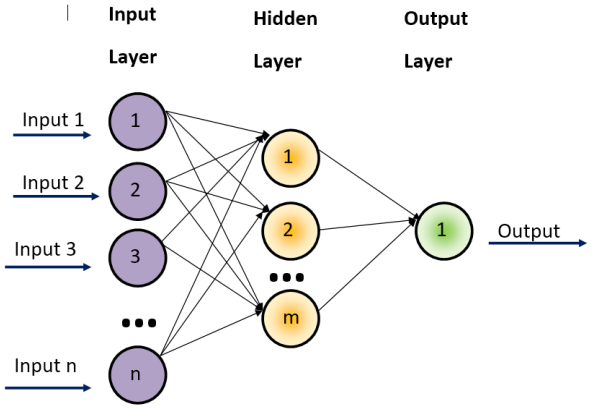
Artificial neural networks i.e., ANNs are computational systems comprising of a huge number of processing units in which data is processed, carried and passed on. the simplest way to put forward the representation of neural networks is a structure of nerve cell, neurons. Just like neurons, in ANN every unit is connected to one another and these neurons are organized in layers. There are three main segregations in a whole neural network: the input layer, output layer and an intermediate between the two is the hidden layer.

Input layer: it is the very beginning of ANN which comprises of neurons that allow the initial data into the network. No logistic operations or processing of information takes place here.

Hidden layer: this is required in order to separate data non-linearly. The hidden layer

identifies the characteristic features from input data and uses these observations to correlate between the input which has now been given and the correct output.

Output layer: it is responsible to bring out the information that has been processed by the foregoing layers. The connections are termed as edges and these edges with neurons together makes up 'weight'. The weight determines strength of the connection between the units. a neural network is firstly provided with a huge amount of data, and these are known as training data sets. The thorough analyses of these training data sets make ANN able to produce outputs. Each layer is constituted of units, the input is made available to all the units of the input layer and it's then weighed and similarly passed on to the units of next layer. There are multiple hidden layers where processing takes place subsequently. The units i.e., the neurons present in each layer calculates the linear combination of weights and inputs in order to produce a result. Activation function is present in hidden layer which helps the network learn complex patterns in the data. The activation function is responsible for forward propagation in network. The neurons work in the light of weight, bias and their activation functions. The errors calculated while comparing the analyses of training data sets and current data are minimized in order to produce the most appropriate output, this is possible with the assistance of activation function.



- *SVM:*

SVMs are one of the most popular and efficient classification as well as regression methods. It is a non-linear model which functions on the principle of classification. If we consider a certain problem, the training data set will be classified and labelled as positive sample and negative sample. The system analyses the results by using this labelled samples in order to build a classifying strategy by applying the algorithm. Since the classification done here is positive and negative, it's binary classification. The SVM segregates the data in positive and negative samples by constructing a hyperplane. The algorithm works in such a way that there is maximum margin between the objects being separated and the hyperplane. Data mapping is done such that n-dimension space of the input space is resolved into higher dimensional space and kernel function is responsible for this unique functionality of SVMs.

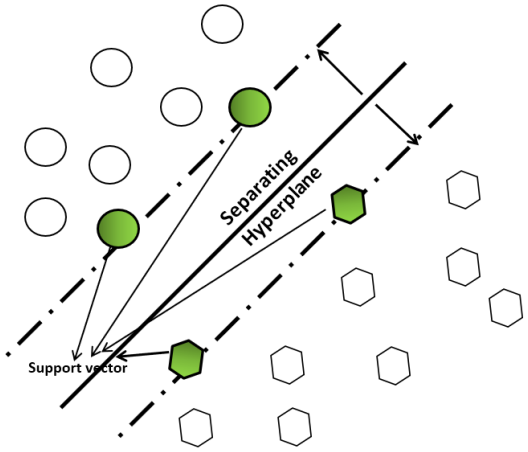


Figure 2: Classification hyperplane of SVM

- *Multivariate Time Series:*

Time series analysis is one of the most common algorithms that has been used in machine learning techniques. Time series data refers to a chronological and consistent stream of data over certain cycle of time. Analysis of time series data leads to better forecasting and prediction outcomes which is the main concern of artificial intelligence nowadays. It establishes a correlation between data within a course of time and how it variates. For time series analysis, there are two types of data sets:

Univariate Time Series- it involves time with one more other variable;

Multivariate time series- it involves multiple other factors apart from time which tends to influence the results. Multivariate time series is a statistical tool which examine time dependent data in order to make predictions such that these predictions are made in light of previous variations in data. Such a series is formed when there is dependence on more than one time-dependent variable and the dependency is on previous values of both, the variable itself and other variables too.

In prediction models there is a huge number of data sets being provided, comprising of both known and unknown parameters. This gives rise to complications when non-linearities are introduced. The jump made from univariate non-linear models to multivariate models is straightforward. but in practical testing, there is availability of small number observations accessible to the predictors, which makes the best way for implementation of nonlinear multivariate models vague and results in no definite conclusions.

CONCLUSION:

In order to recognize geographical patterns of crime, hotspot mapping is utilized by many crime reduction and policing professionals. The Predictive Accuracy Index was created in this study as a metric to compare the predictive power of crime mapping approaches and to evaluate if typical hotspot mapping techniques are the tools used by investigators and if their capacity to anticipate crime patterns is distinct. This study also looked at whether hotspot maps for different types of crime have different prediction ability. In comparison to the most frequent oppositional hotspot mapping approaches, such as common deviation ellipses, thematic mapping of border regions, and thematic grid mapping, this article found that KDE is also effective at forecasting future geographic crime hot spots. Many practitioners make the assertion that the value of hotspot maps is the same regardless of the kind of crime. When this finding was taken into account, it revealed a connection linking the consistency of scope and occasion for crime to be committed and the required application of backdated data in predicting spatial patterns of future. Crimes in street possibilities were found to be more balanced than the possibilities of other crimes taking, and thus data which is recollected and backdated is assed suitable for street crime as compared to crime types where possibility of this taking place are variable. The results of this study propose that assessing the efficiency and effectiveness (or volume) of computer files and carefully picking out the parameters will improve the strength of hotspot mapping method to detect crime patterns in the home and, after all, have influenced the event of recent crime prediction mapping techniques.

REFERENCES:

1. The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime (Lisa Tompson, University College London, Spencer Chainey, University College London): Article in Security Journal, February 2008.

2. The Continuum of Classification Fuzziness in Thematic Mapping (Glles M. Foody).

3. Marked point process hotspot maps for homicide and gun crime prediction in Chicago: George Mohler, Department of Mathematics and Computer Science, Santa Clara University

4. Thematic Map Comparison (Foody, Giles M.) American for Photogrammetry and Remote sensing

5. Technical Problems on Thematic Mapping (Willis R. Heath (1964) Technical Problems on Thematic Mapping, The Cartographic Journal)

6. Time Series: Economic Forecasting (J.H. Stock, in International Encyclopedia of the Social & Behavioral Sciences, 2001)

7. Effectiveness of Support Vector Machine for hot-spots prediction (Keivan Kianmehr, Department of Computer Science, University of Calgary , Calgary, Alberta, Canada)

8. Applied Machine Learning in Social Sciences: Neural Networks and Crime Prediction by Ricardo Francisco Reier Forradellas *ORCID, Sergio Luis Náñez AlonsoORCID, Javier Jorge-VazquezORCID and Marcela Laura Rodriguez Department of Economics-DEKIS Research Group, Catholic University of Ávila, 05005 Ávila, Spain

9. An Overview on Crime Prediction Methods by Nurul Hazwani Mohd Shamsuddin, Nor Azizah Ali, Razana Alwee Faculty of Computing, Universiti Teknologi Malaysia, Johor, Malaysia

10. Artificial neural networks: https://en.wikipedia.org/wiki/Artificial_neural_network

11. https://keralapolice.gov.in/crime-statistics/ipc-cases

12. http://police.bihar.gov.in/menuhome/crime-in-bihar.html

ACKNOWLEDGMENT:

To begin with, We would like to express our heartfelt appreciation to Mr. Naveen Rathi, our professor and mentor, for allowing us the opportunity to work on the review article at Dronacharya College Of Engineering and for providing us with unrivalled assistance, which enabled us to complete the outline on time. Prof. Anchal Bhutani, our English teacher, took a personal interest in our job and helped us through it, providing all required material and ensuring that we had a crystal-clear comprehension of it.

Furthermore, we owe a debt of gratitude to our adoring parents and siblings for instilling motivation and ensuring stability. Last but not least, we'd want to express our gratitude to everyone who assisted us, whether directly or indirectly in the direction of project completion.

AUTHOR INFORMATION:

- **Aakriti Yadav,** Undergraduate student, Department of Computer Science Engineering and Information Technology, Dronacharya College of Engineering, Gurgaon

- **Nikita Verma,** Undergraduate student, Department of Computer science Engineering and Internet of things, Dronacharya College of Engineering,Gurgaon

- **Saumya Rawat,** Undergraduate student, Department of Computer Science Engineering and Information Technology, Dronacharya College of Engineering, Gurgaon

- **Akshay Kumar,** Undergraduate student, Department of Computer Science Engineering and Information Technology, Dronacharya College of Engineering, Gurgaon

- **Deepak Bisht,** Undergraduate student, Department of Computer Science Engineering and Information Technology, Dronacharya College of Engineering

- **Adarsh Ojha,** Undergraduate student, Department of Computer Science Engineering and Information Technology, Dronacharya College of Engineering,Gurgaon