

## Using Big Data, An Extensible System for Forecasting and Analyzing Relations Among Crimes

Subashka Ramesh S.S<sup>a</sup>, Aayush Anshu<sup>b</sup>, Vikash Kumar<sup>c</sup>, Himanshu Kumar<sup>d</sup>

<sup>a,b,c,d</sup>SRM Institute of Science and Technology, Ramapuram, Chennai, India

<sup>a</sup>subashka@gmail.com, <sup>b</sup>aayushanshu31@gmail.com, <sup>c</sup>vikash00mot@gmail.com, <sup>d</sup>himanshushubham555@gmail.com

**Article History:** Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

**Abstract:** While Big Data presents a dilemma for criminal activity analysis, it's even possible to assist them in locating and detecting patterns that can help them crime prevention and investigation. This system intends to draw consideration to current problems in Cyber Crime Investigations – particularly those involving Big Data – as well as potential approaches to combating cybercrime. The proposed system's outcome would help law enforcement and police department agencies a higher comprehension criminal issues and provide insights that will allow them to track operations, forecast the chance of incidents, arrange resources effectively, and optimise policymaking processes. Crime prediction makes use of historical data and, after analysing it, forecasts future crimes based on place and time.

Outdated methods become fewer effective as their ability to deliver needed outcomes in an appropriate and supply-constrained manner deteriorates. To prevent and combat crime, one promising choice for Criminal Inquiries is to use tools for computing based on cutting-edge data analytics. As a result, machine learning and computer modelling should be included in the investigations. One of the solutions is computational, which provides quick and effective data analysis to find small sign in big, unstructured data sets.

**Keywords:** Big Data Analytics, crime prediction, prophet model, SSARIMA

### 1. Introduction

Big Data Analytics (BDA) has recently emerged as a popular method for dissecting information and extracting data and their relationships across a broad range of application domains. Because of consistent urbanization and developing populaces, urban communities assume significant focal parts in our general public. Notwithstanding, such improvements have likewise been joined by an increment in rough violations and mishaps. To handle such issues, sociologists, examiners, and security foundations have given a lot of exertion towards mining likely examples and variables. According to public strategy notwithstanding, there are numerous difficulties in managing a lot of accessible information. As a result, new techniques and innovations have emerged to investigate this data is heterogeneous and comes from a number of sources should be devised. Investigation of such vast data enables us to effectively monitor past events, distinguish similarities from occurrences, send assets, and make timely decisions. This will also help us gain a better understanding of both real-world problems and current events, resulting in improved health, security, and personal satisfaction, in comparison to a rise social and financial growth.

The fast development of computing on a broad scale and information procurement and capacity advancements, from examination and business foundations to different associations and governments, have prompted a colossal number of remarkable extensions/intricacies from information that has been gathered and made freely accessible. It has gotten progressively critical to extricate significant data and accomplish new experiences for understanding examples from such information assets. BDA can successfully address the tests of information that is too fast, too large, and too unstructured to be managed using old-style methods. DBA, as a convincing practice and fast-growing, can assist organizations in making better use of their data and encouraging new freedoms. BDA may also be submitted to assist eager organizations in moving forward with extra powerful actions, satisfied customers, and high benefits. As a result, BDA is becoming progressively important for organizations to address their formative problems.

Information mining is an interdisciplinary, emerging examination territory, and innovative that can fabricate ideal models and procedures in a variety of fields for concluding valuable data and concealed examples from information. It is one of the fundamental strategies of BDA. Information mining is useful for more than just searching new knowledge or marvels, as well as for improving current ones and our interpretation of them. BDA can assistance us differentiate wrongdoing outlines with the aid of certain processes, exist in a specific environment and how they are linked to time. The ramifications of AI and factual methods on wrongdoing or other enormous information applications, for example, auto collisions or time arrangement information, will empower the examination, extraction and comprehension of related examples and patterns, eventually aiding wrongdoing anticipation and the executives.

## 2. Related Works

Within the massive volume of data, data visualisation and mining methods can be used to demonstrate the statistical correlations between various characteristics that have been extracted. To predict inclinations and achieve models with the highest precision are the best, cutting-edge algorithms based on machine learning and deep learning are used.

According to L. Graham [4] Crime is one of our society's most prevalent and disturbing elements, and preventing it remains a critical task. Crime investigation stands a method of identifying in addition analysing trends and patterns in crime that is done in a systematic manner. In this paper, we analyse Tamil Nadu crime data using different clustering methods from data mining. The crime data comes from India's National Crime Records Bureau (NCRB). It covers crime data from the years 2000 to 2014 for 6 cities: Tirunelveli, Salem, Tiruchirappalli, Coimbatore, Chennai, and Madurai with there are 1760 instances in total, with 9 attributes to represent them. The K-Means clustering, Agglomerative clustering, and Density Dependent Spatial Clustering with Noise (DBSCAN) algorithms are used to cluster crime operations based on such predefined instances, and the consequences of these gathering are contrasted to find the best fitting gathering procedure for crime detection. For interactive and easy understanding, the result of the K-Means clustering algorithm is visualised by means of Google Map. For crime prediction, the K-Nearest Neighbor (KNN) organization is used. Each clustering algorithm's performance is assessed using metrics like accuracy, recall, and measure, and the results are compared. This job aids law enforcement agencies in Tamilnadu in better anticipating and detecting crimes, lowering the crime rate. A period is a collection of numerical data points that are arranged in a certain order. that have been sequentially indexed, listed, or graphed in time order. In most cases, the consecutive information facts in a period are evenly set apart in time, resulting in discrete data.

D. Griffith proposed [2] To assess the level or future crime trend, law implementation authorities as result makers need upcoming data, that can be expected over index crime forecasts. The conversion of information into valuable data is critical. Predicting crime trends can help you make better decisions when it comes to crime prevention. Aside from that, law enforcement authorities need a number of decision-making options based on predicted crime trends. It's required because prediction can be fraught with risk. To get this, you'll need to use a decision-making framework programme. The aim of this paper was to provide a proposed decision-making framework implementation that can include data on crime prevention alternatives. To make assessments, crime predictions and interval forecasting data are used. The level of crime forecasting limit range and crime forecasting values are used to select each decision choice. The limit range is defined by the distribution frequency. Strong, medium, and low are the three stages. The decision options are then made based on this amount. For decision-making procedures, if-then rules are used. The decision has resulted in a number of likely circumstances, which should aid law enforcement authorities in making better crime anticipation decisions. The Crime Prevention Decision Support System is the scheme application for this procedure (CreP-DSS). The Prophet model is an additive model that fits non-linear patterns with annual, weekly, and/or daily seasonality, as well as holiday impacts, to forecast time series data. It's best for time series with robust occasional impacts and various periods of authentic information. Missing information and pattern shifts aren't an issue for the Prophet model, and anomalies are normally dealt with well. The Prophet model is worked to deal with complex time arrangement highlights while additionally having natural boundaries that can be adjusted without knowing the basic model's subtleties. A neural organization is comprised of a specific number of neurons, or hubs in the organization, that are orchestrated in layers and connected to each other across layers.

B. Schneier [1] described, development is producing major financial and communal alterations in cities, as well as posing a number of challenges in terms of city management. Given that bigger cities have advanced crime rates, crime spikes are quickly flattering one of the greatest pressing communal issues in big cities. New tools are empowering police forces to access ever-increasing volumes of crime-related data, which can then be analysed to reveal patterns and trends, resulting in more efficient distribution of police officers across the jurisdiction and more reliable crime reduction. Using spatial analysis and auto-regressive models, this paper presents a framework for automatically identifying high-risk crime regions in metropolitan areas and correctly predicting crime trends in each field. The calculation delivers a spatio-worldly wrongdoing anticipating model, which is comprised of a bunch of wrongdoing thick districts and a bunch of related wrongdoing indicators, every one of which addresses a prescient model for foreseeing the quantity of violations that will happen in its own locale. The proposed strategy accomplishes great precision in spatial and fleeting wrongdoing forecasts throughout moving time skylines, as indicated by the exploratory appraisal, which was directed on genuine information gathered in an enormous zone of Chicago.

S. Morgan [5] proposes Criminological theories enabled crime forecasting, and advances in computing methods/data analytics enhanced forecasting knowledges even further. It assists the numerous police subdivisions in making decisions and developing strategies to deter crime. This paper focuses on using to forecast the annual

crime rate in India, time series models such as Auto-Regressive Integrated Moving Average (ARIMA) and Exponential Smoothing were used. The information comes from the National Crime Record Bureau of India. As part of the modelling process, the data is separated into training data from 1953 to 2008 and test data from 2009 to 2013. According to the model, precision measurements are also significant, and the forecast values are clearly within the 95 percent confidence interval of the test results. As a consequence, a time series model that can be used to predict crime has been developed.

### 3. Existing System

Within the massive volume of data, information perception and mining methods are utilized to exhibit the factual connections between different qualities that have been separated. Bleeding edge AI and profound learning calculations are utilized to conjecture designs and accomplish ideal models with the best exactness. A period is a collection of numerical data points that have been sequentially indexed, listed, or graphed in time order. In most cases, the consecutive statistics facts in a period of time are evenly spaced in time, resulting in discrete information. The Prophet model is an additive model that fits non-linear patterns with annual, weekly, and/or daily seasonality, as well as holiday impacts, to forecast time series data. It's best for time series with robust occasional impacts and various periods of recorded information. Missing data and trend shifts aren't a problem for the Prophet model, and outliers are usually handled well. The Forecaster model is built to handle multifaceted time series landscapes while also having parameters that can be modified without having to know the specifics of the underlying model. A neural network is made up of a certain number of neurons, or nodes, arranged in layers and connected to one another through layers.

$$m(p) = g(p) + s(p) + h(p) + \varepsilon p$$

$$g(p) = (k + a(p) T \delta)t + (m + a(p) T \gamma)$$

$$\{1, \text{if } p \geq s_j,$$

$$a_j(p) =$$

$$\{0, \text{otherwise}$$

Furthermore, this system has the following flaws:

1. There is no uniform framework
2. It is impossible to capture the dynamics within a region
3. It is impossible to accurately predict future crime
4. It's difficult to find accurate patterns
5. Interregional patterns aren't available.

### 4. Proposed System

The proposed application for a framework for making decisions that will include details on crime prevention options. Crime forecasts and interval forecasting data are used to make decision choices. Each decision alternative is selected based on the level of crime forecasting limit spectrum and crime forecasting values. The limit range is defined by the distribution frequency. Strong, medium, and low are the three stages. The decision options are then made based on this amount. For decision-making procedures, if-then rules are used. The decision has resulted in a number of likely circumstances, which should aid law enforcement authorities in making better crime prevention decisions. Techniques for predicting the location of a crime are known as crime prediction techniques. Managing and collecting large amounts of accurate data. On the basis of crime data, crime patterns can be analyzed and avoided.

### 5. Methodology

Prophet is a strategy for expecting time course of action data subject to an additional constituent model where non-direct examples are reasonable with annual, step by step, and step by step routineness, notwithstanding event impacts. It performs most prominent with time game plans that have strong ordinary effects and a couple of times of chronicled data. Prophet is liberal to lost info and changes in the model, and ordinarily handles irregularities well. To start working with the Prophet model, we need to gather the necessary information so the model will investigate and plot a chart showing the different patterns. The four primary advances that should be executed are:

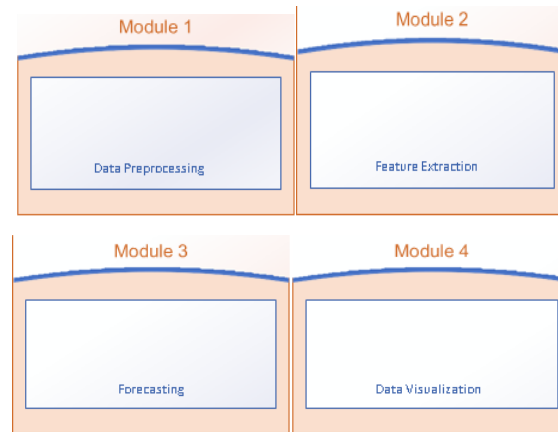


Fig 1.1 Data Modules

**Sub model I: Data Pre-processing**

Information Pre-processing is the process of transforming or encoding data so that it can be easily parsed by the machine. In other words, the algorithm can now readily interpret the data's features.

The following are common drawbacks of real-world data: Incompleteness, noise, and inconsistency are all words that come to mind when I think of inconsistency. As a result, these data must be preprocessed in order to make them appropriate for analysis, and this pre-processing entails the following tasks:

1. Data reduction
2. Data discretization
3. Data integration

$$x(o) = \sigma (Ax \cdot [ho-1, xo] + bx)$$

$$io = \sigma (Ai \cdot [ho-1, xo] + bi)$$

Missing values and outliers in both crime and criminal data sets have absolutely different meanings than in many other data sets. For example unknown criminal address is different from that for an employee because crime analysis tends to discover and find out the unknowns, whereas in employment data set these unknowns can be cancelled to prepare data for mining algorithms, in addition to that outliers in crime and criminal data sets represent some types of knowledge needed to be explored and processed and needed to be focused on, for example increasing a specific crime “Theft” for example in 5% annually is expressed as a normal situation between adults whereas 0.005% increase of this type of crimes between children can be of great meaning, this percent may be expressed as outliers in applications other than crime analysis.

S.No	Year	Crime Incidence	Crime Rate
1	2006	1878293	167.7
2	2007	1989673	175.1
3	2008	2093379	181.5
4	2009	2121345	181.4
5	2010	2224831	187.6
6	2011	2325575	192.2
7	2012	2387188	196.7
8	2013	2647722	215.5
9	2014	2851563	229.2
10	2015	2949400	234.2
11	2016	2975711	233.6
12	2017	3062579	237.7
13	2018	3132955	236.7
14	2019	3225701	241.2

Table 1.1 IPC Crimes over the years 2006 – 2019

The table determines the successful attacks in India from the year 2006 to 2019. The successful crime attacks gradually increase from 167.7% in 2006 to 241.2% in 2019. In between these years, the crime rate peaked in the year 2017 and decreased the following year. Notwithstanding, later on the pace of criminal assaults may differ and from the diagram, there is a chance of the crime percentages expanding soon. With the ascent of crime

percentages across different urban areas on the planet and different families getting influenced by such interruptions, a prescient model to forestall such offense later on must be made by examining the information from over a significant time span record. Making credits and ordering them dependent on the comparable characteristics will show when and where the wrongdoing has occurred. From this, we can distinguish the regular time, area and different parts of wrongdoing that is occurring in a specific district. Bad behavior event assumption is basic to bad behavior expectation in the overall population by aiding the law execution workplaces to structure ideal watch techniques. Reduction of bad behavior will benefit society from different viewpoints. It will assemble the open prosperity and decay the monetary mishap. Regardless, bad behavior event assumption is a troublesome task. Various factors are material to the probability that a particular sort of bad behavior event will occur in a locale in the near future. This consolidates economics, the spread of different kinds of organizations, bad behavior history, human adaptability, and so on criminal allegation. Grouping is a technique for certainty assessment which is utilized to remove and anticipate upcoming patterns in data dependent on closeness trials. In future work, this calculation can be ad libbed to find crooks extra productively. From the examination on Forecasting of yearly Crime include in India: A contextual investigation [5]. It had an inclination, it gives a methodology to stop wrongdoing, by finding the example and patterns in wrongdoing should be possible by embracing a wrongdoing study [6]. Despite the fact that it had various benefits still This paper clarifies the thought hypothetically we should almost investigate the idea basically.

### Sub model II: Feature Extraction

Aside from the conventional transformed and non-transformed signal characteristics and texture, feature extraction approaches include structural and graph descriptors.

One of the dimensionality reduction techniques is feature selection. It's used to get rid of features that aren't needed. It also increases the accuracy of classification. Unlike feature extraction methods, feature selection techniques use the original set to create a new generalized feature set. Due to some objective functions and important criteria, this subset of features chosen provides the best performance.

$$C_o = \tanh(AC \cdot [h_{o-1}, x_o] + bC)$$

### Sub model III: Forecasting

$$C_o = f_o * C_{o-1} + i_o * \sim C$$

where  $x_o$  is a sigmoid feature that specifies if the previous state can be retained,  $C_{o-1}$  is the old cell state,  $C_o$  is the modified cell state,  $A_x$ ,  $A_i$ , and  $AC$  are the previous values in each layer,  $h_{o-1}$  and  $x_o$  are the input values,  $b_x$ ,  $b_i$ , and  $b_c$  are constant values,  $i_o$  specify which value will be used to change the state, and  $C_o$  stands for the current candidate values.

Crimes involving burglaries, drugs, and other types of crimes can be projected using patterns of violence in a city or society in order to find resources in critical areas or to well deter crime in the future or to deal with crime as it happens. For many years, crime forecasting has been based on FBI reports released each year. These reports, on the other hand, provide only a sliver of the crimes that occur every year, or even every day, across the country. By allowing politicians and police units to screen instances of corruption or police response tactics, any people community will be better positioned to deal with crimes and police outreach on their own patios. The midway-assembled and regularly revised criminality detecting instruments and online data would assist officials in exchanging attitudes and creating stronger early warning frameworks to protect more people in their networks. Crime detection can help with avoiding repeat offences in a given area by identifying previous wrongdoing patterns or separating the more common types of wrongdoing in a given field. It can also send data to multiple offices rather than just one, allowing for improved asset detection across a range of networks across the globe. This is the advantage of incorporated information while as yet assembling information from every individual office to hold the special data and wrongdoing designs found locally. Advancements now accessible give a more secure stockpiling territory, simpler or quicker admittance to information, and the capacity to think about extensive crime data from the littlest towns to the biggest urban areas to acquire a superior comprehension of how and where wrongdoing happens the nation over. This information will also be used to strengthen our areas and prevent breaches from happening in the first place. Officials can more easily resolve the problems of their own networks through education, countermeasures, and enhanced response strategies. Because of these mechanisms, networks may be set up with early alert systems to protect their citizens, as well as continuous surveillance within the police department to deter these wrongdoings. This also allows it possible for police forces to distribute their personnel to protect horror zones, allowing authorities to concentrate on these areas more quickly.

### Sub model IV: Data Visualization

Graphs can reveal patterns and reveal connections between people, times, and places. Graph visualisation is a crucial tool for analysing and comprehending graph data at a large scale.

Because our original dataset included a District property, we can use it in our visualisation. Law enforcement will be able to see and understand complex connections in their data thanks to the proposed system's link analysis tools. The tools help to break down data silos, increase investigations, uncover more information, and make operations more effective.

The proposed technologies, in particular, aid in overcoming the following data challenges that modern policing faces:

**Volume** – the proposed link analysis tools perform admirably even when visualising large datasets like those found in policing.

**Complexity** – the node-link model is a simple way to understand complicated data connections.

**Urgency** – the proposed software is based on current web technology, which means it can be used on any device that has a web browser.

**Formula**

$$X(t) = N(t) + C(t) + Z(t) + \epsilon_0$$

$$C(t) = \sum_{n=1}^N \left( a_n \cos\left(\frac{2\pi nt}{L}\right) + b_n \sin\left(\frac{2\pi nt}{L}\right) \right)$$

**X(t)**: logistic growth curve for demonstrating non-intermittent changes in time arrangement

**C(t)**: recurring differences (e.g., yearly or weekly seasonality)

**Z(t)**: impacts of occasions (user given) with unpredictable schedules

**ε0**: error term account for any uncommon variation that cannot fit in the model

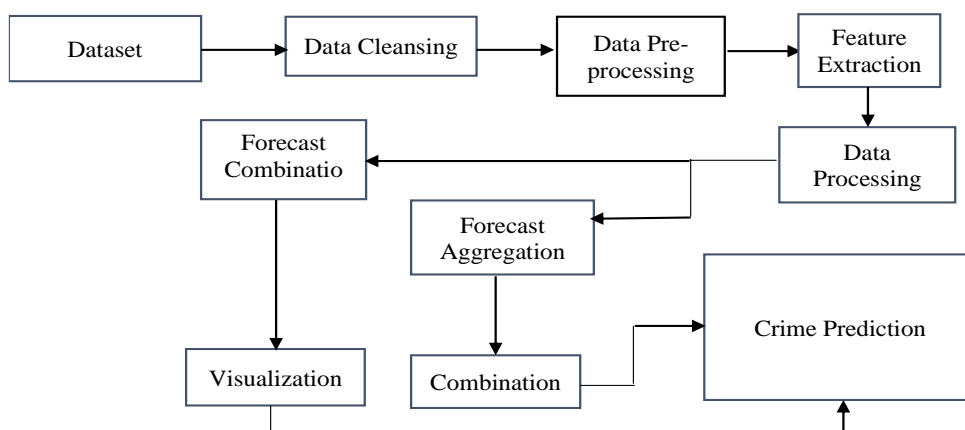
To fit and figure the impacts of regularity, prophet relied upon Fourier arrangement to give a reasonable model. Seasonal effects C(t) is approximated by the following formula:

L is the period yearly data is 365 and weekly data is 7) Parameters [a1, b1, ..., aN, bN] should be assessed for an offered N to model seasonality.

The fundamental Hardware and Software required are:

- Input Data: Eclipse IDE, Sci-learn
- Tools for simulation: Python and
- Documentation: NumPy

**6. System Architecture**



**Fig 1.2** Architecture Diagram

From the above outline, the wrongdoing informational index is gathered and a four-section activity called information mining occurs. Information mining alludes to the act of investigating gigantic previous datasets to create new insights. At first, the information goes through an interaction called information purging or information cleaning that recognizes and remedies bad or erroneous records. After the information is adjusted and remedied a technique called information preprocessing happens which changes over the crude information into a helpful and proficient configuration. At that point an interaction where the crude information is diminished to a more sensible gathering of preparing is alluded to as highlight extraction. When the information is gathered, a strategy called information preparing happens where tasks of the information are executed by the PC to recover, change or arrange data.

The underneath flowchart depicts the cycle of the prophet model. At first, the information which is consolidated from pattern, irregularity and occasions are determined and the information is gotten. The main technique that happens is include choice where the client naturally or physically chooses those highlights which will give the most to the expectation variable or yield. When the highlights are chosen the information goes through a four-venture measure for example demonstrating, estimate assessment, surface issues and outwardly examining the figures. After the cycle is finished, the informational indexes are examined to sum up the principal qualities regularly with visual techniques.

### 7. Result And Conclusion

This module gives law requirement heads and strategy producers with objective research on the adequacy of prescient examination in diminishing wrongdoing and to give proposals for those assessing whether to put time and assets into a prescient policing program. Law authorization pioneers look for powerful wrongdoing control methodologies, since, supposing that wrongdoing can be decreased, at that point local area security and flexibility are improved. This theory additionally affirms that diminishing wrongdoing improves local area flexibility, consequently improving country security. By adequately fighting wrongdoing in their networks, police bosses will have the chance to redistribute assets towards other country security needs. Actualizing another wrongdoing control system requires a pledge to plan and actualize the system, put cash in reasonable gear and programming, train staff, screen and assess results, and change the methodology as fitting. A particularly enormous venture ought not be embraced without realizing that the work on being put resources into has been approved and will create the normal outcome. As of now, prescient investigation and prescient policing programs have not been broadly concentrated from a level headed, logical point of view. While starting encounters by the offices that have either completely executed or explored different avenues regarding prescient policing appear to be positive, prescient policing's effect on wrongdoing still can't seem to be authoritatively decided.

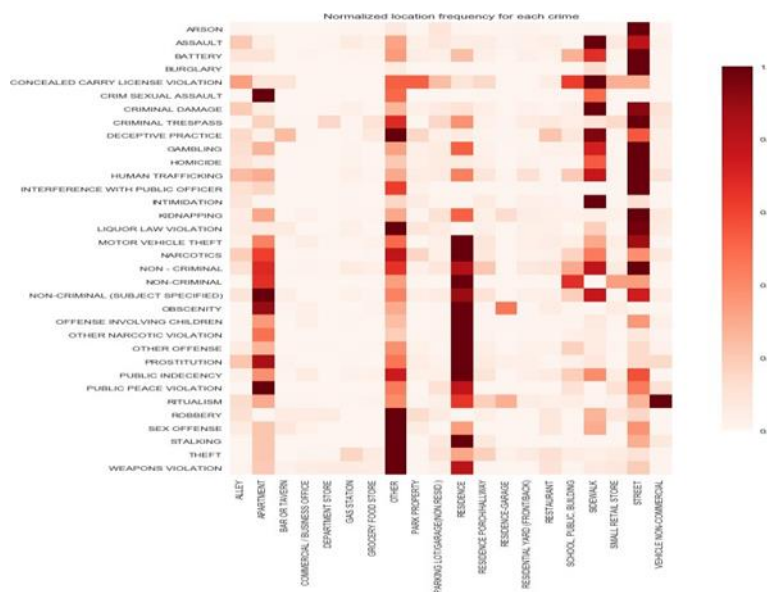


Fig 1.4 Heat Graph (Normalized Location Frequency for Each Crime)

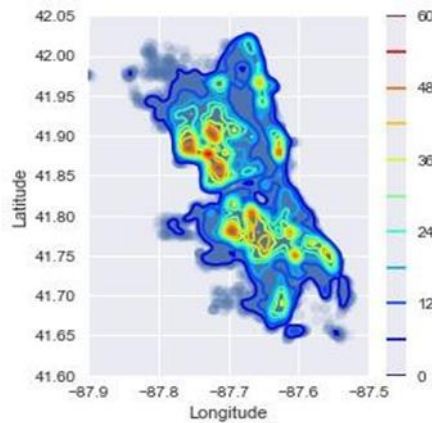
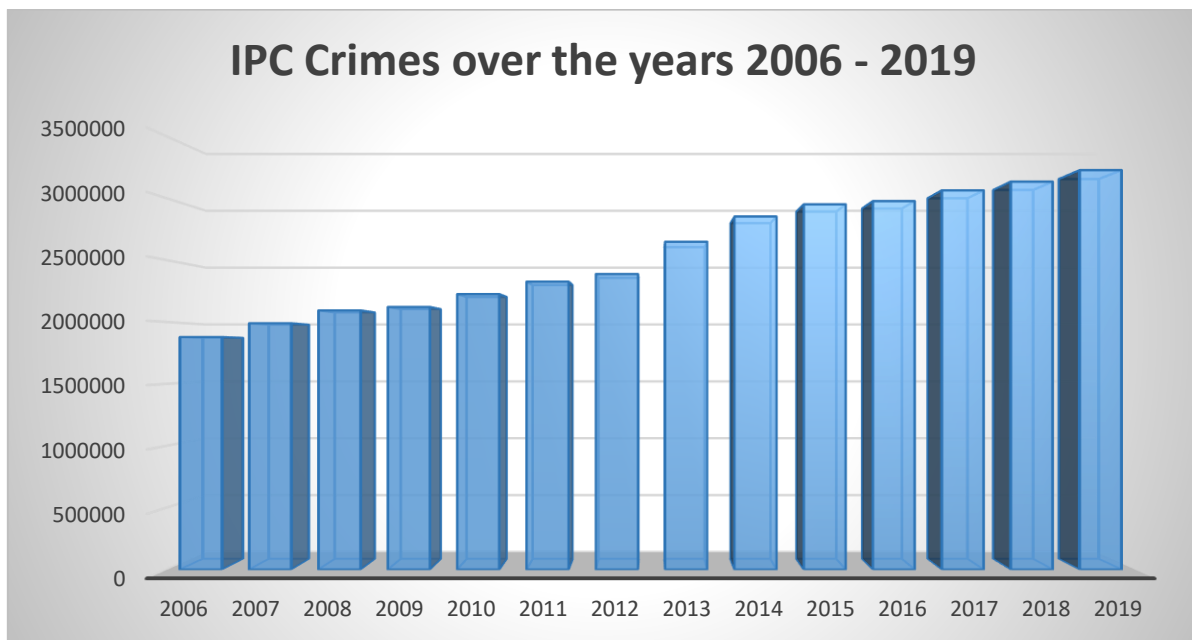


Fig 1.5 Heat Map of the Area



Graph 1.1 Year vs Crime Incidence

**8. Future Work**

In the upcoming, the idea is to whole our ongoing stage aimed at nonexclusive massive information examination, which will be suitable for preparing a variety of data formats for a variety of applications. To uncover more likely examples and patterns within these datasets, we plan to combine multivariate visualisation, diagram mining strategies, and fine-grained spatial analysis. Furthermore, we intend to conduct more rational contextual analyses in order to assess the adequacy and adaptability of the similar models in our framework.

**References**

1. B. Schneier, "Tracking the owner of kickass torrents," <https://www.schneier.com/blog/archives/2016/07/tracking-the-ow.html>, July 2017, accessed: 05.07.2017.
2. D. Griffith, "How to investigate cybercrime," <http://www.policemag.com/channel/technology/articles/2003/11/how-to-investigate-cybercrime.aspx>, November 2003, accessed: 07.11.2017.
3. Guarino, "Digital forensics as a big data challenge," in ISSE 2013 Securing Electronic Business Processes. Springer, 2013, pp. 197–203.
4. L. Graham, "Cybercrime costs the global economy \$450 billion: Ceo," <https://www.cnbc.com/2017/02/07/cybercrimecosts-the-global-economy-450-billion-ceo.html>, February 2017, accessed: 07.11.2017.



5. S. Morgan, "Cyber crime costs projected to reach \$2 trillion by 2019," <https://www.forbes.com/sites/stevemorgan/2016/01/17/cyber-crime-costsprojected-to-reach-2-trillion-by-2019/>, January 2017, accessed: 07.11.2017.
6. K. Chitra Lekha and Dr. S. Prakasam, "An analysis of finding the Influencing Factors of supporting for the "GiveitUp" LPG Subsidy for the Government using Data mining Techniques", *IJCA*, Vol. 143, Issue 5, June 2016, pp.34-39.
7. Dr. R. Jayabrabu, Dr. V. Saravanan and Dr. J. Jebamalar Tamilselvi, "A Framework for Fraud Detection System in Automated Data mining using Intelligent agent for better Decision making process", *Science Direct*, March 2014.
8. Subashka Ramesh S.S, Kartik Singh Rathore, Ritik Raj, Kumar Vatsalya, MridulaVatsa "Integrated Malware Analysis Using Markov Based Model in Machine Learning" *International Journal of Engineering and Advanced Technology (IJEAT)* ISSN: 2249-8958, Volume-8 Issue-4, April 2019, pp 219-222.
9. Muhammad Arif, Khabaib Amjad Alam and Mehdi Hussain, "Crime Mining : A Comprehensive Survey", *International Journal of u- and eService, Science and Technology*, Vol. 8, Issue 2, 2015, pp. 357-364.
10. Dr. Zakaria Suliman Zubi and Ayman Altaher Mahmmud, "Crime Data Analysis using Data mining Techniques to Improve Crimes Prevention", *International Journal of Computers*, Vol. 8, 2014, pp. 39-45.
11. Tushar Sonaqwanev, Shirin Shaikh, Shaista Shaikh, Rahul Shinde and Asif Sayyad, "Crime Pattern Analysis, Visualization and Prediction using Data Mining", *IJARIE*, Vol. 1, Issue 4, 2015, pp.681-686.
12. Akshay Kumar Singh, Neha Prasad, Nohil Narkhede and Siddharth Mehta, "Crime: Classification and Pattern Prediction", *IARJSET*, Vol. 3, Issue 2, February 2016, pp. 41-43.
13. K. Chitra Lekha and Dr. S. Prakasam, "Performance Assessment of Different Classification Techniques", *CiiT International Journal of Data mining and Knowledge Engineering*, Vol. 9, Issue 1, January 2017, pp. 20-23.
14. D. J. Karn, *Policing and Crime Reduction, The evidence and its implications for practice: The Police Foundation*, 2013.
15. R. B. Santos, "Effectiveness of Police in Reducing Crime and the Role of Crime Analysis," in *Crime Analysis with Crime Mapping*, ed: Sage Publications, Inc., 2012.
16. Atul bamrara, Gajendra Singh and Mamta Bhati, "Cyber Attacks and Defense Strategies in India: An Emprical Assessment of Banking Sector", *IJCC*, Vol.7, Issue 7, January-June 2013, pp. 49-61.
17. K. Tayal, V. Ravi. "Fuzzy association rule mining using binary particle swarm optimization: applications to cyber fraud analytics", *IEEE International Conference on Computational Intelligence & Computing Research (ICCIC)*, pp. 1-5, 2015.
18. Chauhan, S. Sehgal. "A review: crime analysis using data mining techniques and algorithms". 2017 *International Conference on Computing, Communication and Automation (ICCCA)*, pp. 21-25, 2017