

Hybrid Approach to Crime Prediction using Deep learning

Jazeem Azeez¹

M.Tech Computer Science

Department of Computer Science & Engineering
Hindustan Institute of Technology & Science,
Padur Chennai.

jazeem10@live.com

D. John Aravindhar²

Associate Professor

Department of Computer Science & Engineering
Hindustan Institute of Technology & Science,
Padur, Chennai.

jaravindhar@hindustanuniv.ac.in

ABSTRACT - Prevention is better than Cure. Preventing a crime from occurring is better than investigating what or how the crime had occurred. Just like vaccination is given to a child to prevent disease, in today's world with such higher crime rate and brutal crime happenings, it has become necessary to have a vaccination system that prevents crimes from happening. By vaccinating society against crime it refers to various methods such as educating people, creating awareness, increasing efficiency and proactive policing methods and other deterrent techniques. Inspired by two different existing approaches to crime prediction, the first one presents a visual analytics approach that provides decision makers with a proactive and predictive environment in order to assist them in making effective resource allocation and deployment decisions. Crime incident prediction depends mainly on the historical crime record and various geospatial and demographic information^[1]. Even though it's promising, they do not take into account the rich and rapidly expanding social & web media context that surrounds incidents of interest. Next approach is based on the semantic analysis and natural language processing of Twitter posts via latent Dirichlet allocation, Topic detection & Sentiment analysis^{[3][4]}. But both the techniques face inherent limitations. Crime that happens these days has the following key characteristics such

as crimes repeating in a periodic fashion, crimes occurring as a result of some other activity and occurrence of crimes pre indicated by some other information.

Keywords: Deep Learning, STL, Sentiment Analysis, LDA, Undirected Probabilistic Graph, RNN

1. Introduction

In today's digital world where there is immense data in both online and offline that is related to what happened, happens or what may happen is society. Some of the information may motivate or contribute to the occurrence of a crime and some may push it away from happening. A typical example would be tweets / post in social networks. It is visible how effective or powerful social network posts can be, it can cause social revolutions like in Egypt, Lebanon, India (Regarding Anna Hazare Lokpal Movement) etc. where the data have caused toppled Governments. So how do these posts and tweets relate to and upcoming happenings. Another factor is that recent studies have shown a higher rate of violence against women occurring in locations where there is a higher level of porn subscribers or the users of sexual services is High. This shows that the contributing factors play key roles

It's really shocking that if on observing previous cases of violence in India against women it tends to get higher during 4 quarter of a year i.e. during (October to December) months. Also for the last few years there have been Terrorist attacks during the same time frame too... This shows the pattern of occurrence of a similar crimes repeating periodically at the same time some crimes or actions contributing to the occurrence of another event. From this basic idea of crimes repeating, crimes occurring as a result of some other activity and occurrence of crimes pre indicated by some other information. The discussion of a new algorithm to crime prediction where it considers the three ideas for crime happening. To start with one should understand Crime is Not a Random Event.

In criminology and Crime Studies It is clearly mentioned as crime should not be considered as a random event. That is one of the factors where many algorithms like ARIMA, STL etc. are not effective to deliver accurate explanation of what crime will occur and that's why it is not being used in Law enforcement agency ^{[5][8]}. Crime occurs due to several factors such as motive, situation and reason etc. Where motive is the prime reason for a crime occurrence without which a crime can't be explained. Also there will be periodic factors influencing occurrence of crime which gives the crime a periodic nature. There will be a trending factor that motivates the occurrence of particular crime. Such as increased rate of Prostitution & kidnapping during financial year ending of a nation. During these period mostly man missing cases will be reported and they usually have a trend of occurrences...

In order to have an effective crime prediction technique it becomes necessary to consider various data inputs & each of these data needs to be analyzed and each of them have different characteristics. Analyzing

this wide variety of data is a challenge. And deriving information from them brings another challenge. It is not possible to exactly predict every crime even though but can try to predict from the data which is on hands. Deep learning is a model in which uses multiple computational models to act up on data and generate higher level knowledge representations which can be further analyzed to derive knowledge and information. The algorithm have a deep learning model and uses graphs to represent the knowledge. Typically this can be considered as a complex neural network where one neuron may connect to another if there is a relation between knowledge. Algorithm is based on very simple concept that for any event to happen there should be some contributing factors or events, it's designed with taking in to account of key concept of crime that crime is not a random event. Thereby leading to a better proactive & predictive law enforcement. This algorithm at its core uses various sub component algorithms such as algorithms for graph matching canonicalization, sentiment analysis, Sequence Trend Loess, Topic identification by LDA

“A crime or offence (or criminal offence) is an act harmful not only to some individual or individuals but also to a community, society or the state ("a public wrong")”. Crime prevention is also widely implemented in some countries, through government police and, in many cases, private policing methods such as private security and home defense etc. Law enforcement agencies are working 24x7 to stop or reduce crime to the very extent they can. They focuses on the use of criminal penalties as a means of deterring people from committing crimes and tempo rarely or permanently incapacitating those who have already committed crimes from reoffending. But the very limitation of this approach is that crime penalties have been made so crucial usually the penalties will be made only

after the crime has been made. Another strategy is to prevent crime from happening by increasing policing. But this rise in increase in no of policing routines increases the total burden on society. This leaves us with the fact of having an efficient policing routines.

There are lots approaches to predict crimes but can be broadly classified in to two strategies on the basis of data source. One approach is to use usual statistic techniques and models such as STL ^[1], ARIMA, kernel density distribution, etc... to provide hotspots of crime. Hotspots of crime are those geospatial locations where there is a higher percent of probability for a crime to occur. Many researchers have proposed systems where they have chosen a set of statistical parameter space to predict the probability of crime. But choice of parameters time and location should be made with respect to the data scarcity they face. And it also ignores the vast amount of data possessed by web networks and its statistical information. By referring to web it refer to the vast amount of information possessed by social networks too. There are lot of theories to consider when it comes to predict crime.

The major ones are

- Choice Theory
- Biological Theories
- Psychological Theories
- Sociological Theories
- Conflict Theories
- Integrated Theories
- Victimization Theories

Each of these theory have its own diverse parameter space this give a challenge of variety of data that appears in crime prediction. And there can be arbitrary combination of these theories itself.

The Second major approach in crime prediction is to consider social tweets & posts, the key idea is to understand what public is discussing about. Form these data it has been proved to identify various upcoming crime and event trends such as public outrage and sentiments etc. But in general this approach lacks when and where the crime will happen.

This is not due to the disability algorithms but due to the vast no of noise and unreliable posts, tweets etc... present in social information. This make law enforcement agencies to discard the second approach event thought it describes the type of crime that may happen. Because the enforcement agency working 24x7 want to make the use of resources as efficiently as possible.

It shows numerous no of posts telling or asking for protest but seldom does those turn in to actions. The approach is capable of including diverse parameters and takes account in to various data sources which can be used to predict crimes.

2. Related work

Crime prediction had been a trending research field for a long time. There had been many advancements in recent years. Also the new trending field deep learning to be discussed. The major one among the trending research topic is social posts and data analysis. The following section describes about major related works.

3. Crime Visual Analytics, Mapping & Prediction

Analysts use a variety of techniques to identify and visualize areas of high crime density known as “Hotspots”. Usually a Hotspot indicates a spatial area of relatively high point of interest or a region of high activity represented by underlying model. One of the most coming strategy is to use KDE^[1] or Kernel density estimation. But basically KDE have its own limitations such as lack of contextual information and portability.

The most preferred visual representation for hotspots are choropleth maps. It is a map which uses differences in shading, coloring, or the placing of symbols within predefined areas to indicate the average values of a particular quantity in those areas. This overlay of choropleth map is very useful when it comes to visualizing hotspots but goes in vain when it comes to represent other information along with it.

4. Crime Forecasting

The most common strategy is to use statistical prediction algorithms to predict crime by getting the frequency and parameter of crime. There are lots of research work done in this domain. The major ones include ARIMA, STL generally regression models. These regression models use previous crime records of the space and use them to find the frequency, trend etc. and to generate predictions. Most of the time if proper data had been provided these algorithms work well to give the hotspots. But the general defect being the ignorance or inability to include the vast amount of live data, parameters resulting inaccuracy and not applicable situations. And the fact that crime is not a random

event. But this method enable us to explore the sequence and trends behavior of crime. There by this method can't be ignored^{[6][8]}.

A fact of the usability is that these algorithms results in prediction that tell where to look for or at which time to look for a crime. But it doesn't tell anything about the crime and its nature. This is particularly hard when the fact that crime occurs due to a combination of contributing factors.

5. Use of Web & Social Media for Prediction

Recent advances in text mining, web trends and social media have prompted a surge in interest towards using the vast amount of data for various predictive purposes. The key examples being use of twitter to predict Election results stock market trends, box office results, product sentiments etc.^{[3][11]}.

Performing Sentiment Analysis on various topics can give public's perspective and intend. For example web statistics of clicks to pornography sites have been correlated to sex crimes such as crime against women. In a given particular location. Social tweets can be analyzed sentiment of users and their geographic locations can be mapped to show the public mood in general and sentiment for a given topic. Higher the sentiment polarity there is a higher chance of occurrence for related events^[11].

6. Deep Learning

Deep learning are composed of multiple layers of non-linear operations such as neural net or in complicated propositional formulae re-using many sub-formulae.

Deep Belief Network have been recently been proposed to tackle many difficult problems. The key focus of artificial intelligence have been giving the capability for computers to model our world. Crime data is used by the system to understand the various relationship between crimes and there by model the world of various crimes^[12].

As shown in the diagram below the deep learning system will use various algorithms and transform the raw information in higher representations and so on...

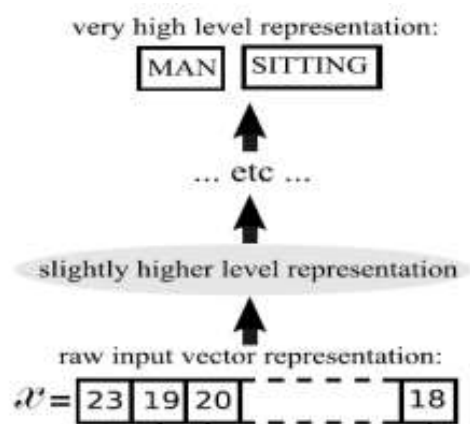


Figure 1 System Representation

7. Experimental setup

At Core of this system is a minimalistic custom built graph data store framework. The Graph database model which serves our need to handle variety of data stream and to maintain the complex modeling and relationship between them. It is here in the graph data base where the complex prediction algorithm runs. The key reason for a custom built graph data store framework is that the algorithm works on the component of 3 basic dimension of physics such as space time energy. Where by at any given time the graph data base can be analyzed in any three of these dimensions. Following diagram shows the component architecture of the system

Input: Time frame to which predictions needs to be made. Required Dataset & information

Output: Visual representation of hotspots created by considering multiple data sources, along with type & general details of crime that occur filtered out with respect to probability

The key feature of the system is that it learns about patterns of events and this knowledge can be reused of across different time and space. The experimental setup have chosen Thiruvananthapuram Kerala India as the space location and a dataset of last 5 years had been collected and built. The Graph Loader Component will load the Graph on to the graph data store and will be analyzed for groups of events.

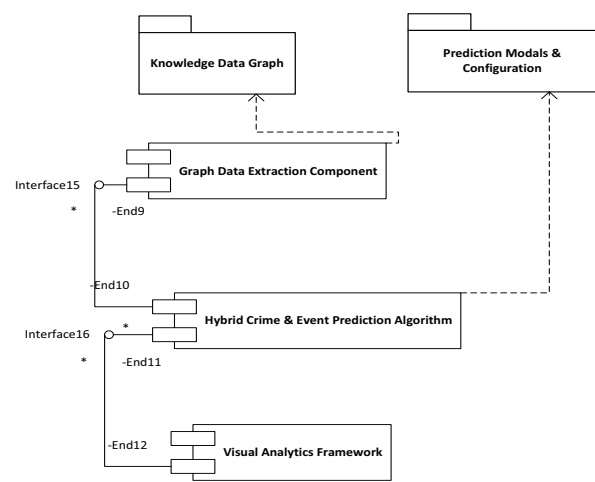


Figure 2 Architecture

Algorithm Description

The algorithm consist of three stages

- (1) Preprocessing stage
- (2) Processing to generate events
- (3) Post Processing.

Here the system learns from historical crime cases and data using multiple component algorithms and creates a knowledge graph as a higher level representation, which will be then used create predictions on the basis of projections of various factors generated and other fitting algorithms will be used over this to find time and space of crime events.

Preprocessing stage the data from various resources will be cleaned and transformed. These transformed data will be stored on to graph database. The user will provide the input space-time frame to which the prediction have to be made. The system will automatically take a buffer zone and analyze for all possible event combination form the zone.

In preprocessing stage the main task is to collect and convert various data or information, where by the system focuses on xml & text parsing to understand the type of information and its relations, a typical type of semantic & topic relationship discovery is performed. Topic or information classification is backed by LDA algorithm.

In processing stag the generated event combination will analyzed to find out the possible production configuration; suppose an event A occurs and is followed by B then C. the System will form out a production of event ABC_E . The system will scan for the most suitable productions and compute its combinations. The observed sequence will be analyzed for matches will previous. Sequence matching is done in the same way as DNA match is performed .the graph will be converted to a canonical representation and cross matched with respect to time space and energy. End result of

this stage is a set of locational hotspots with set of possible events. This part signifies deep learning architectures.

In post processing stage the generated event have to be filtered across for interesting events. This is done by using several output stage threshold based filters. For each event the contributing factors will be different and there by the methods to compute the probability also will differ. But this diversity has been handled by the Graph Data structure. One of the advantage of Using Graph database its schema last feature. This particularly aid for the complex environment of the program. This post processing is required are there will be noise or unwanted labels be predicted from the graph of information.

8. Result

This section will analyses the out of the algorithm and its performance characteristics. The following Chart shows randomly select 10 types of crime and the probabilistic anomalies given out by the algorithm for a given geo spatial cell. This will be recomputed for all the geo spatial cells in the prediction zone. Form this it is able to generate the type of crime and its time space constraints of the crime that may happen.

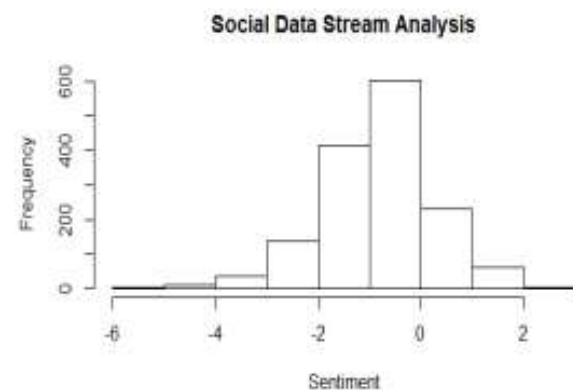


Figure 3 Frequency distribution of Sentiment

Social data streams are analyzed to find out sentiment level score of the public and there reaction to events happen. This is achieved by using a combination of LDA and sentiment analysis. The LDA gives out various topics as points of interest and the system runs sentiment analysis on them.

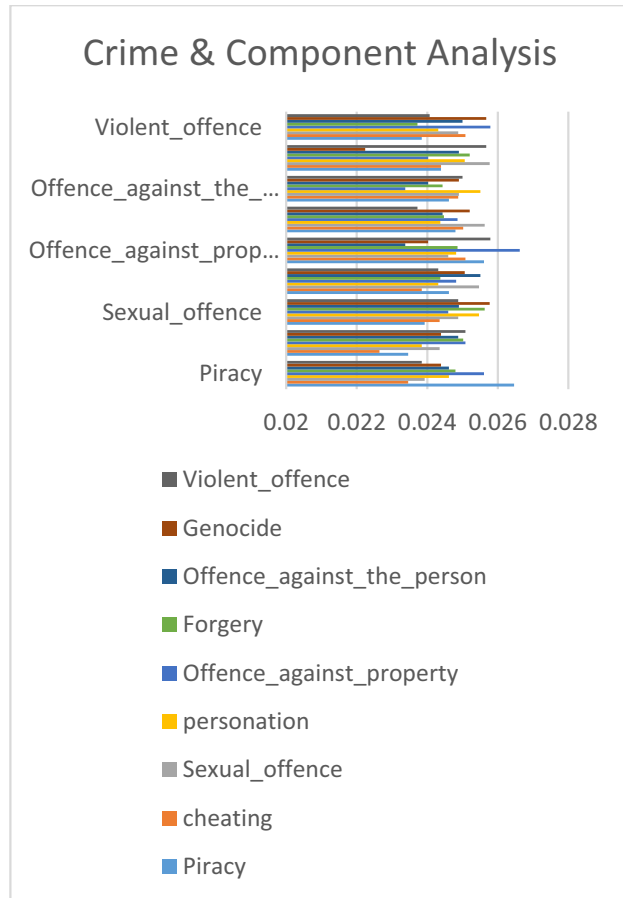


Figure 3 Component Analysis

For example recent researches have shown that areas where porn relate tweets made are of places where here is higher rates of crimes against women. Also from web traffic logs it is clear that geographic locations and the corresponding web traffic is related. A related semantic analysis can be done to get more relevant information. Also web various web statistics are used with respect to the locations.

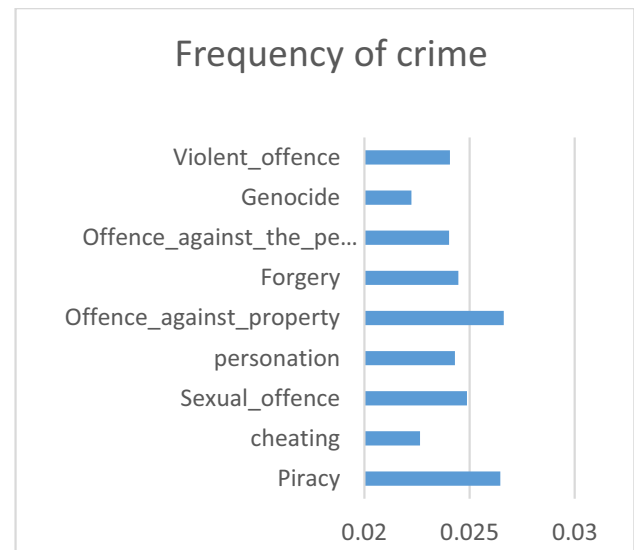


Figure 4 Frequency of Crime

The above result shows the Crime frequency with respect to crimes that happened in a Geo Spatial Zone. From the test run it is clear that Crimes with respect property and Piracy are the topmost crimes in that particular zone. Following is a corresponding intermediate dump table of correlational values.

X/Y	Piracy	cheating	Sexual_of/personati	Offence_ Forgery	Offence_ Genocide	Violent_offence
Piracy	0.02546	0.023457	0.023924	0.024613	0.025608	0.024796
cheating	0.023457	0.022645	0.02435	0.023842	0.02508	0.025019
Sexual_of	0.023924	0.02435	0.024877	0.025466	0.024593	0.025628
personati	0.024613	0.023842	0.025466	0.024309	0.024816	0.02437
Offence_	0.025608	0.02508	0.024593	0.024816	0.026622	0.024857
Forgery	0.024796	0.025019	0.025628	0.02437	0.024857	0.024471
Offence_	0.024613	0.024877	0.024898	0.025506	0.023376	0.024431
Genocide	0.02439	0.02439	0.02577	0.02506	0.024025	0.025202
Violent_	0.023842	0.02508	0.024877	0.024309	0.02579	0.023721

Figure 5 Probability Matrix

A sample visualisation of internal events graph along with probability of there relations is show in following image. this simply is the representation of reated events the and their respective probability graph.

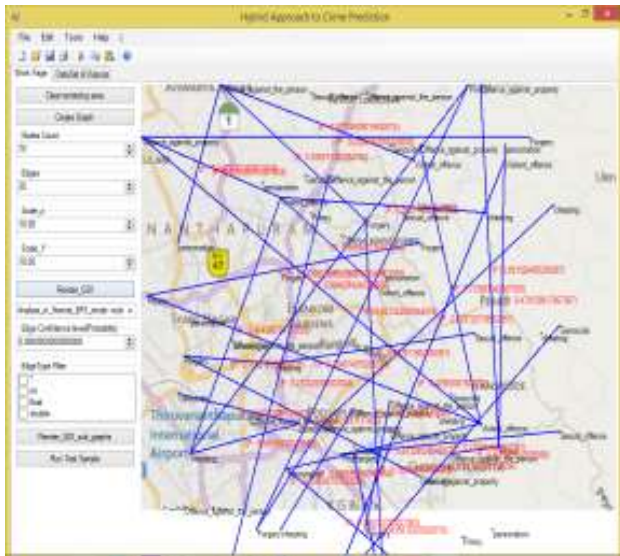


Figure 6 Crime and its relations across various events

PERFORMANCE ANALYSIS

To start with this algorithm have a minimum time and space complexity of $O(n^2)$. In order to manage such a huge space requirement and complexity large graph based databases are used. The average Training time required by the algorithm for 100000 events have been observed around 53 min in average as show in chart below.

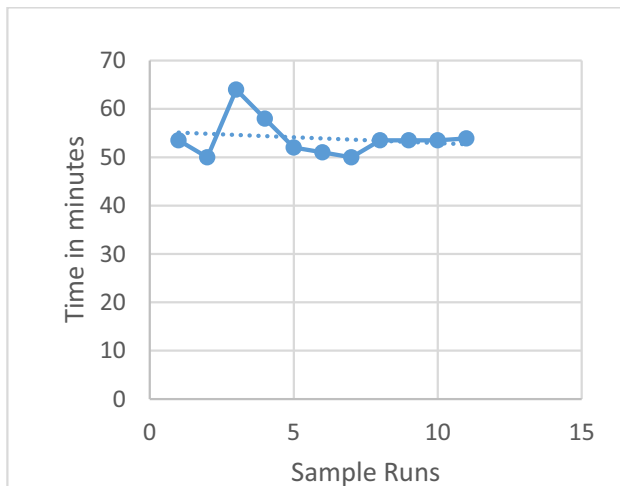


Figure 7 Time complexity

The Time complexity is measure with respect to the no of iterations required to find interdependency of events

where each events should be cross checked with respect to every other events. Shown in the chart below

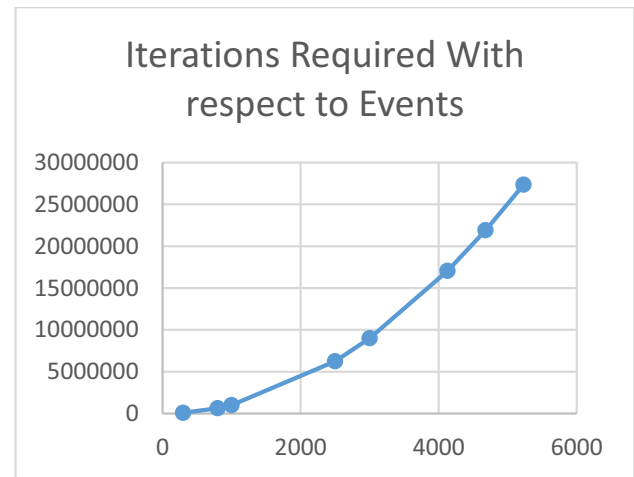


Figure 8 No of iteration required with respect to events

From the above chart & table it can infer that the process is both CPU intensive and memory intensive. The Test bed used was AMD x64 based 4cpu true core machine running at 3.6 GHZ with 16GB ram running at 2133MHZ. As said above the algorithms performance characteristics the output of an algorithms is its results and the quality of algorithm.

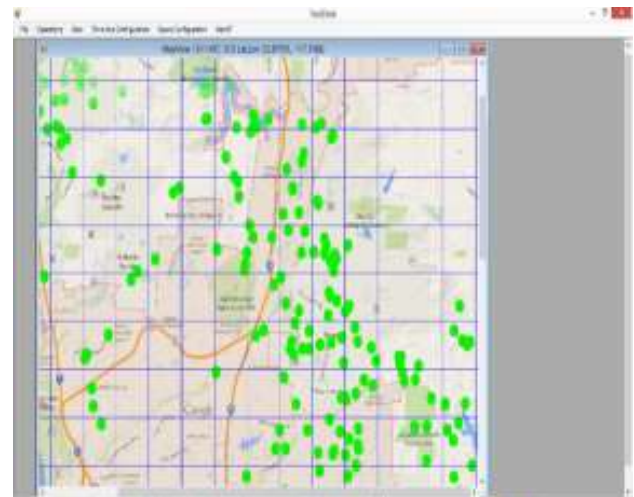


Figure 9 Heat Map Plotted over prediction zone

With respect to each heat map location and latitude & longitude the deep learning system is able to provide

near accurate probability chart of crime. With a sorted view on confidence levels.



Figure 10 Probable Crimes Listing for Given Location

We have used H₂O framework to test out the prediction and the records say that the crimes scenario predicted came true. What this implies is that not only the law enforcement agencies can locate the most probable location of crime with this system of combined knowledge and deep learning we are able to predict the most probable crime that may happen at a given time space frame.

9. Conclusion

A hybrid approach to crime event prediction Using Deep Learning, covers the limitations of both of the existing strategies that have discovered by doing literature survey. Also found that it is possible to predict crime by focusing patterns and trends from various contributing factors.

Each of these techniques have its own limitations and advantages, it is obvious because to predict properly should consider all the possible factors that affect that prediction. This gives rise to another challenge of dealing with wide variety of data. Also another challenge that should be addressed is the volume of web data &

its velocity. It also considers how to deal with wide variety of data that can be used for predicting crime, for this develop a hybrid predicting algorithm that basically deals with both the strategies for prediction of crime. From this it is clear that it's possible to map most probable crime events & its geo spatial & temporal details.

10. Future work

The approach is to crime prediction using deep learning system even though gives a tremendous edge over existing approaches but the computational complexity and achieving accuracy with more fine grained detail is an open research topic. Acquiring data, information required to make such a prediction is a great challenge which is same as maintaining the resources required to make such a prediction for a researcher. This opens up further the requirement of further advancements that needs to be achieved in the field of deep learning modules & its optimization for event prediction

11. References

- [1]. Abish Malik (2013) IEEE Transactions on Visualization and Computer Graphics “Proactive Spatiotemporal Resource Allocation and Predictive Visual Analytics for Community Policing and Law Enforcement”
- [2]. Hua Liu 2003 International Institute of Forecasters Published by Elsevier B.V
- [3]. Keivan Kianmehr Department of Computer Science University of Calgary 2006 IEEE: “Crime Hot-Spots Prediction Using Support Vector Machine”
- [4]. Kira Radinsky Technion–Israel Institute of Technology Haifa, Israel 2013 ACM 978-1-4503-

1869-3/13/02: “Mining the Web to Predict Future Events”

[5]. Nafees Ur Rehman 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining: “OLAPing Social Media: The case of Twitter”

[6]. Patricia Morreale Department of Computer Science Kean University Union 2013 27th International Conference on Advanced Information Networking and Applications Workshops: “Data Mining and Analysis of Large Scale Time Series Network Data”

[7]. Peng Chen Fifth International Conference on Fuzzy Systems and Knowledge Discovery Forecasting: “Crime Using the ARIMA Model”

[8]. Shyam Varan Nath Florida Atlantic University/ Oracle Corporation 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology: “Crime Pattern Detection Using Data Mining”

[9]. Sitaram Asur Social Computing Lab HP Labs Palo Alto, California 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology: “Predicting the Future With Social Media”

[10]. Spyros E. Polykalas 2013 27th International Conference on Advanced Information Networking and Applications Workshops: “A general purpose model for future prediction based on web search data (Predicting Greek and Spanish elections)”

[11]. Xiaofeng Wang 2012 ACM International Conference on Advances in Social Networks Analysis: “Automatic Crime Prediction using Events Extracted from Twitter Posts”

[12]. Dong Yu ; Microsoft Res., Redmond, WA, USA ; Li Deng, Signal Processing Magazine, IEEE (Volume:28 , Issue: 1) : “Deep Learning and Its Applications to Signal and Information Processing”