



Big Data Analytics and Crime Patterns Detection and Prevention

¹ Mayuri M. Menkudle, ² Rachana S. Potpelwar

¹ Student, M Tech (IT), ² Assistant Professor, ME (IT)

Dept of Information Technology,

Shri Guru Gobind Singhji Institute of Engineering and Technology
, Nanded-431606 (MH), India

Abstract: Analysis of crime data aims to the extraction of patterns of crimes and trends of crimes from previously stored data of crimes. Analysis of crime data plays an important role to prevent crimes that occur in the country and simultaneously that increase the security of people. The paper aims to seek out a frequent occurrence of crime patterns with knowledge discovery and its prediction. We have used machine learning techniques on criminal records for knowledge discovery and to assist in increasing the predictive accuracy of the crime. This work is going to be helpful to the local police stations in crime suppression.

Index Terms - Big data, LSTM model, prophet model, machine learning.

I Introduction

This project is estimated on big data. It seeks to unify the hitherto fragmented discourse on what contains large data, what dimensions define the dimensions and other features of bulk data, and what tools and technologies are available to utilize the potential of bulk data. Crime prevention and detection is becoming a crucial trend in crime and it can be really challenging to uncover crime. Numerous studies have found different methods to unravel crimes that most applications do not. Such studies can help speed up the process of solving crime and computerized systems can help identify criminals automatically. In addition, rapidly evolving technologies can help solve such problems. However, patterns of crimes are always get changing and increasing. Data of crime previously get from various sources tends to gradually expand. Consequently, management and analysis with heavy data are very difficult and sophisticated. To unravel the problems mentioned earlier, data processing methods uses various learning algorithms to mining hidden knowledge from a large volume of data. Data processing is data analysis methods to search for patterns and trends in crime. It helps to resolve crimes more quickly and also helps to automatically alert criminal identities. The area commits the crime where criminal activities are done and crimes committed by criminals, and produces and leaves traces of crime, and testimonies within a specific range of your time and place. Crime scene analysis is the starting line of a crime investigation, so it is very important for the success of case resolution. Traditional crime scene analysis is based on knowledge gained from onsite surveillance. However, it is well integrated into the private experiences of the police, which are significantly fragmented and lacking in systematic and scientific scrutiny. Therefore, it is imperative to introduce new theories and methods in crime scene analysis. Crime rates are constantly rising and therefore the pattern of crime is constantly changing. As a result, it is difficult to describe behaviors in criminal models. This paper explains crime prevention in order to social development. Its aim is to provide a comprehensive review of the theory and research related to crime prevention in society and to implement various data analysis algorithms that address the relationship between crime and its model. Information for the project is collected from legal government sources.

One challenge for a developing team is to formulate important expectations. A common consequence of this requirement is that an assessment must be made, i.e., "crime takes place somewhere in the city" cannot be innocently imagined. In other words, the expected world should be too small for police to be able to patrol with greater priority than unselected areas. Another consequence is that a clear estimate is of little value. For example, it is uninteresting for the national police to find that theft is taking place in an outdoor mall, especially since mall thefts can be easily assessed and prevented by mall guards. Therefore, in achieving a successful new system and giving the user confidence that the new method will work and be more accurately efficient, it can be considered as the most critical level.

II LITERATURE SURVEY

There are so many researchers who are working on crime pattern analysis and prediction. The most common crime pattern detection, future prediction of crime is done by various machine learning, data mining techniques such as apriority algorithm, association mining techniques, LSTM, acp approach etc. with maximum accuracy. Below is the table that describes the summary of different techniques used in different papers for detecting the crime pattern.

Ref no	Author	Paper	Classification
[1]	A. Gandomi and M. Haider	"Beyond the hype: Big data concepts, methods, and analytics,"	Big data analytics used for text, audio, video analysis.
[2]	U. Thongsatapornwatana	"A survey of data mining techniques for analyzing crime patterns,"	Association rule mining is used for analyzing crime patterns.
[3]	S. Wang, X. Wang	"Parallel crime scene analysis based on ACP approach,"	Parallel execution approach and computational experiments artificial societies, are used for crime analysis.
[4]	S. Yadav, A. Yadav	"Crime pattern detection, analysis & prediction"	K-mean, naïve Bayes, apriori algorithm used for crime prediction and suppression.
[5]	N. Baloian, C. E. Bassaletti	"Crime prediction using patterns and context"	Multi kernels, Dempster-Shafer theory Applications used for crime prediction
[6]	X. Zhao and J. Tang	Spatio-temporal method are used for mining historical data.	Spatio-temporal pattern used for crime prediction.
[7]	S. Wu, J. Male	"Spatial temporal campus crime pattern mining from historical alert messages,"	Spatio-temporal method are used for mining historical data.
[8]	K. R. S. Vineeth, T. Pradhan	"A novel approach for intelligent crime pattern discovery and prediction,"	FP max, KDD is used or crime pattern discovery and prediction.

Table I Related work

III METHODOLOGY

We use different models for identification and classification in this paper. As we saw in above table no. of different methods are used in different paper for classification of data and improve the result. In this paper we use following methods as shown in fig.1 .

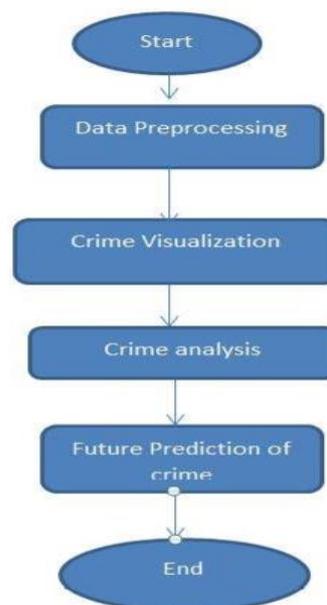


Fig 1 System Design flow

3.1 Data Pre-processing

The First ever step is to take the input crime data from the crime dataset and improve the input data quality which uses filter to contain unwanted noise. Crime dataset consist of various featured attributes such as Incident Numb, Dates, and Day of Week etc. we have to preprocess the data by removing unwanted features from dataset such as incident Numb, coordinate. By imputing random values on missing values by taking their mean value. Dataset consist of timestamp then we can deduce it in various attributes such as year, month, day, hour 9 and minute.

3.2 Crime Visualization

We can visualize the crime by using various visualization techniques such as pie chart using various colors. By visualizing crime we can get idea of various types of crimes, with their rate of crimes occur in various area.

3.3 Crime Analysis

Crime analysis is one of the most important steps in crime pattern detection. In which we take previous crime data and analyses that data which considered the which type of crime is occur with how much amount of rate.

3.4 Future Prediction of Crime

After analyzing the crime data the next step is prediction of crime in future. in this prediction of crime after analyzing the crime we predict the amount of rate of crime in future. For prediction of crime we use two machine learning techniques such as LSTM model and Prophet Model. Lstm model is a long-short term memory model which is used to predict the future crime. And prophet model is used to predict the seasonality, trend, holiday model.

3.4.1 LSTM Method

Long Short-Term Memory (LSTM) is used in the field of deep learning. LSTM is a neuronal, artificial network. There are two types of feed forward and feedback-word neural network. LSTM has relations with feedback word. LSTM processes the entire data sequence, rather than single data points. For example, LSTM applies to tasks such as connected handwriting recognition, unsegmented detection and speech recognition in network traffic or IDSs (intrusion detection systems). LSTMs have been developed to deal with the question of vanishing gradients that can be encountered while training conventional RNNs. Relative insensitivity to gap length is an advantage of LSTM over RNNs, hidden Markov models and other sequence learning methods in numerous applications

3.4.2 PROPHET MODEL

Prophet is a tool for estimating time series data based on an additive model that combines nonlinear patterns with annual, weekly, and regular seasonality and holiday impacts. It works best with a time series which has clear seasonal effects and historical data over multiple seasons. The protagonist approach at its core is an additive regression model composed of four main components: a linear or logistic growth curve trend pixies. By selecting change points from the data the protagonist detects changes in trends automatically prophet forecasting process is shown in figure 2.

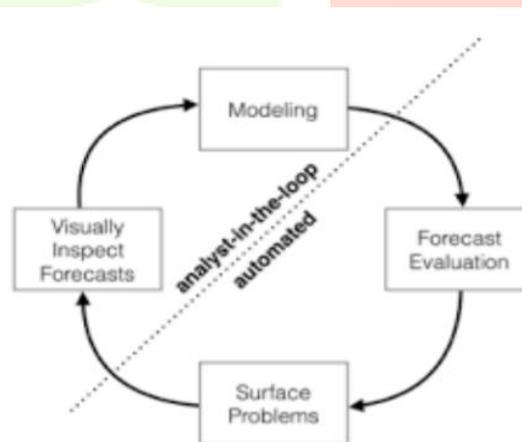


Fig: 2 Prophet forecasting process

IV RESULTS

In this Philadelphia city it shows the prediction of crimes for next two years based on previous two years. From the year 2017 to 2019 it shows the same rate of crimes and based on that result by using LSTM methodology. we predict the crime rate of next two years i.e. 2019 and 2020. Figure 3 shows the Lstm model for Philadelphia city for crime prediction. Figure 4 and figure 5 shows future crime prediction of Philadelphia city using prophet model for holiday and seasonality.

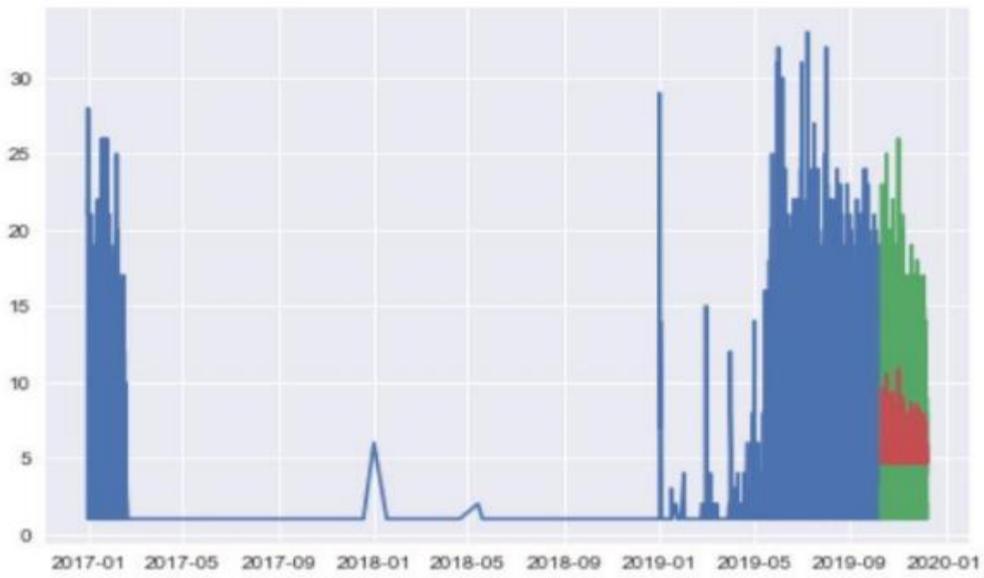


Fig: 3 LSTM model for Philadelphia

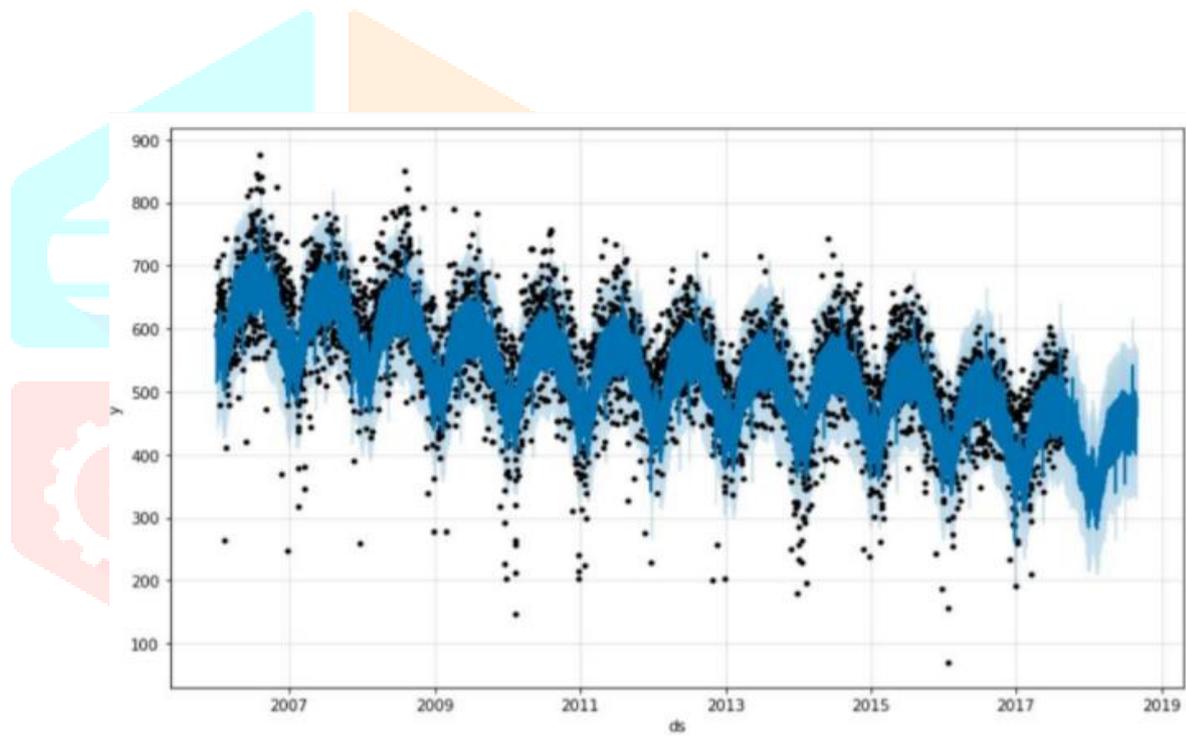


Fig: 4 prophet model of holiday for Philadelphia

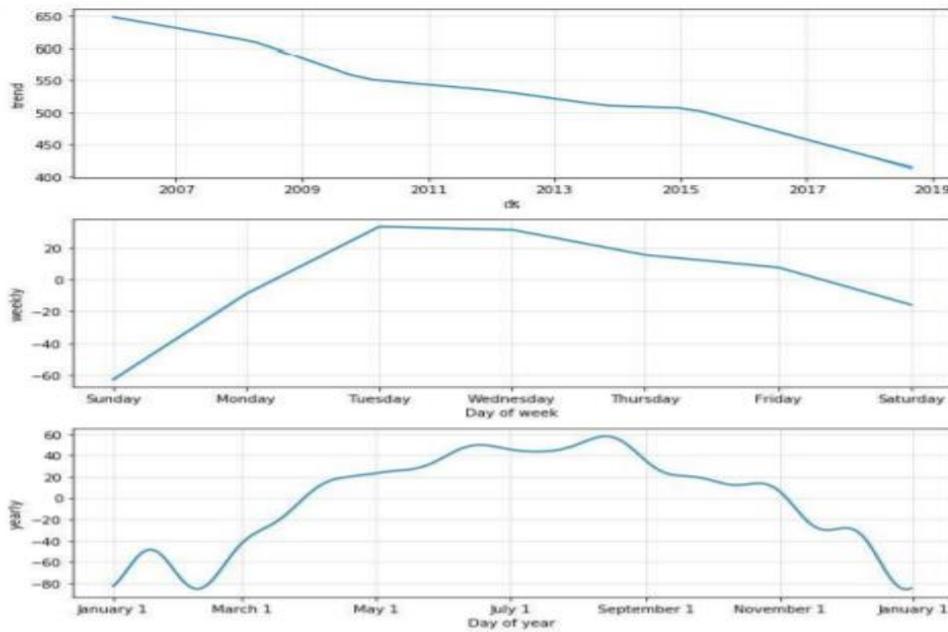


Fig: 5 prophet model of seasonality for Philadelphia

V CONCLUSIONS

The biggest challenge of the project is data collection and data staging. The further scope extension of crime detection and analysis is to get crime hotspots, which will help deploy police in crime prone areas for any window of your time, allowing you to make the best use of police resources. The developed system reduces crime and helps the field of crime detection in some ways, from arresting criminals to reducing crime by completing various necessary actions. Crime is classified, it changes over time and constantly increases. Changing and increasing crime can lead to problems such as understanding criminal behavior, assessing crime, accurate identification, and maintaining large volumes of knowledge gained from a variety of sources. Researchers have interests in sought to unravel these issues. However, these investigations are gaps in the accuracy of crime detection. This leads to challenges in the field of crime detection

VI ACKNOWLEDGEMENT

I take this opportunity to express a deep sense of gratitude towards my guide Prof. R .S. Potpelwar, for providing excellent guidance, encouragement and inspiration throughout the project work. Without her invaluable guidance, this work would never have been a successful one. I would also like to thank all my classmates for their valuable suggestions and helpful discussions.

VII REFERENCES

- [1] A. Gandomi and M. Haider, "Beyond the hype: Big data concepts, methods, and analytics," *Int. J. Inf. Manage.*, vol. 35, no. 2, pp. 137_144, Apr. 2015.
- [2] U. Thongsatopornwatana, "A survey of data mining techniques for analyzing crime patterns," in *Proc. 2nd Asian Conf. Defence Technol.*, Chiang Mai, Thailand, 2016, pp. 123_128.
- [3] S.Wang, X.Wang, P. Ye, Y. Yuan, S. Liu, and F.-Y.Wang, "Parallel crime scene analysis based on ACP approach," *IEEE Trans. Computat. Social Syst.*, vol. 5, no. 1, pp. 244_255, Mar. 2018.
- [4] S. Yadav, A. Yadav, R. Vishwakarma, N. Yadav, and M. Timbadia, "A survey of data mining techniques for analyzing crime patterns," in *Proc. 2nd Asian Conf. Defence Technol.*, Chiang Mai, Thailand, 2016, pp. 123_128.
- [5] N. Baloian, C. E. Bassaletti, M. Fernández, O. Figueroa, P. Fuentes, R. Manasevich, M. Orchard, S. Peña_el, J. A. Pino, and M. Vergara, "Crime prediction using patterns and context," in *Proc. 21st IEEE Int. Conf. Comput. Supported Cooperat. Work Design*, Wellington, New Zealand, Apr. 2017, pp. 2_9.
- [6] X. Zhao and J. Tang, "Exploring transfer learning for crime prediction," in *Proc. IEEE Int. Conf. Data Mining Workshops*, New Orleans, LA, USA, Nov. 2017, pp. 1158_1159.
- [7] S. Wu, J. Male, and E. Dragut, "Spatial-temporal campus crime pattern mining from historical alert messages," in *Proc. Int. Conf. Comput., Netw. Commun.*, Santa Clara, CA, USA, 2017, pp. 778_782. 29
- [8] K. R. S. Vineeth, T. Pradhan, and A. Pandey, "A novel approach for intelligent crime pattern discovery and prediction," in *Proc. Int. Conf. Adv. Commun. Control Comput. Technol.*, Ramanathapuram, India, 2016, pp. 531_538.
- [9] A. Joshi, A. S. Sabitha, and T. Choudhury, "Crime analysis using K-means clustering," in *Proc. 3rd Int. Conf. Comput. Intell. Netw.*, Odisha, India, 2017, pp. 33_39.
- [10] Z. Zhao, S. Tu, J. Shi, and R. Rao, "Time-weighted LSTM model with redened labeling for stock trend prediction," in *Proc. IEEE 29th Int. Conf. Tools Artif. Intell. (ICTAI)*, Boston, MA, USA, Nov. 2017, pp. 1210_1217.
- [11] H. Hassani, X. Huang, M. Ghodsi, and E. S. Silva, "A review of data mining applications in crime," *Stat. Anal. Data Mining, ASA Data Sci. J.*, vol. 9, no. 3, pp. 139_154, Apr. 2016.