Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Special Issue 4, April : 2024

UNVEILING CRIME HOTSPOTS

Nalluri Charishma1, Palakitti Yamini Siva Parvathi2, Marthula Mounika3, Kolli Hema4
1,2,3,4 Department of Computer Science and Engineering, Vignan's Institute of Engineering for Women, Visakhapatnam, Andhra Pradesh, India

ABSTRACT
This study is to identify and predict the criminal hotspot in crime data. Crime is one of the society's most pressing issues and preventing it is critical. This research explores the application of ensemble learning algorithms in machine learning to enhance the accuracy of crime hotspot prediction. By combining multiple base models, such as decision trees, random forests the ensemble model aims to provide a more robust and reliable prediction framework. The study employs various spatial and temporal features, including historical crime data, and environmental factors. Through the ensemble learning approach, the model demonstrates improved performance in identifying crime hotspots, offering valuable insights for proactive crime prevention strategies and resource allocation in urban areas.
Keywords: Random Forest Classifier, Ensemble Learning Approach ,Crime Hotspots, Decision tree.

INTRODUCTION

In this project, we employ the power of machine learning to pinpoint areas with elevated crime rates. Random Forest algorithm enables us to analyse vast amounts of data efficiently, uncovering patterns and trends to identify crime hotspots accurately. By leveraging this advanced technique, we aim to assist law enforcement agencies and urban planners in developing targeted strategies for crime prevention and resource allocation. It takes the historical data of public property crime from 2015 to 2018 from a section of India as research data to assess the predictive power of machine learning algorithm. Results based on the historical crime data alone suggest that the random forest performs well.Therefore, future crime prediction should take advantage of both historical crime data and covariates associated with criminological theories. Not all machine learning algorithms are equally effective in crime prediction.

LITERATURE SURVEY

1) A survey of data mining techniques for analyzing crime patterns AUTHORS: U. Thongsatapornwatana

In recent years the data mining is data analyzing techniques that used to analyze crime data previously stored from various sources to find patterns and trends in crimes. In additional, it can be applied to increase efficiency in solving the crimes faster and also can be applied to automatically notify the crimes. However, there are many data mining techniques. In order to increase efficiency of crime detection, it is necessary to select the data mining techniques suitably. This paper reviews the literatures on various data mining applications, especially applications that applied to solve the crimes. Survey also throws light on research gaps and challenges of crime data mining. In additional to that, this paper provides insight about the data mining for finding the patterns and trends in crime to be used appropriately and to be a help for beginners in the research of crime data mining.

2) Risk terrain modeling: Brokering criminological theory and GIS methods for crime forecasting AUTHORS: J. M. Caplan, L. W. Kennedy, and J. Miller

The research presented here has two key objectives. The first is to apply risk terrain modeling (RTM) to forecast the crime of shootings. The risk terrain maps that were produced from RTM use a range of

contextual information relevant to the opportunity structure of shootings to estimate risks of future shootings as they are distributed throughout a geography. The second objective was to test the predictive power of the risk terrain maps over two six-month time periods, and to compare them against the predictive ability of retrospective hot spot maps. Results suggest that risk terrains provide a statistically significant forecast of future shootings across a range of cut points and are substantially more accurate than retrospective hot spot mapping. In addition, risk

terrain maps produce information that can be operationalized by police administrators easily and efficiently, such as for directing police patrols to coalesced high-risk areas.

3) Using geographically weighted regression to explore local crime patterns AUTHORS: M. Cahill and G. Mulligan.
The present research examines a structural model of violent crime in Portland, Oregon, exploring spatial patterns of both crime and its covariates. Using standard structural measures drawn from an opportunity framework, the study provides results from a global ordinary least squares model, assumed to fit for all locations within the study area. Geographically weighted regression (GWR) is then introduced as an alternative to such traditional approaches to modeling crime. The GWR procedure estimates a local model, producing a set of mappable parameter estimates and t-values of significance that vary over space. Several structural measures are found to have relationships with crime that vary significantly with location. Results indicate that a mixed model— with both spatially varying and fixed parameters—may provide the most accurate model of crime. The present study demonstrates the utility of GWR for exploring local processes that drive crime levels and examining misspecification of a global model of urban violence.

4) Language usage on Twitter predicts crime rates AUTHORS: A. Almehmadi, Z. Joudaki, and R. Jalali
Social networks 1 produce enormous quantity of data. Twitter, a microblogging network, consists of over 230 million active users posting over 500 million tweets every day. We propose to analyze public data from Twitter to predict crime rates. Crime rates have increased in the past recent years. Although crime stoppers are utilizing various technics to reduce crime rates, none of the previous approaches targeted utilizing the language usage (offensive vs. non-offensive) in Tweets as a source of information to predict crime rates. In this paper, we hypothesize that analyzing the language usage in tweets is a valid measure to predict crime rates in cities. Tweets were collected for a period of 3 months in the Houston and New York City by locking the collection by geographic longitude and latitude. Further, tweets regarding crime events in the two cities were collected for verification of the validity of the prediction algorithm. We utilized Support Vector Machine (SVM) classifier to create a model of prediction of crime rates based on tweets. Finally, we report the validity of prediction algorithm in predicting crime rates in cities.

5) Self-organised critical hot spots of criminal activity AUTHORS: H. Berestycki and J.-P. Nadal
In this paper1 we introduce a family of models to describe the spatio-temporal dynamics of criminal activity. It is argued here that with a minimal set of mechanisms corresponding to elements that are basic in the study of crime, one can observe the formation of hot spots. By analysing the simplest versions of our model, we exhibit a self-organised critical state of illegal activities that we propose to call a warm spot or a tepid milieu2 depending on the context. It is characterised by a positive level of illegal or uncivil activity that maintains itself without exploding, in contrast with genuine hot spots where localised high level or peaks are being formed. Within our framework, we further investigate optimal policy issues under

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Special Issue 4, April : 2024

the constraint of limited resources in law enforcement and deterrence. We also introduce extensions of our model that take into account repeated victimisation effects, local and long range interactions, and briefly discuss some of the resulting effects such as hysteresis phenomena.

EXISTING SYSTEM

The random forest and spatiotemporal KDE(kernal density estimation) method, found that the random forest algorithm is more efficient than the traditional spatiotemporal KDE method in the smaller time scale and grid space unit.Compared with the traditional Network-time KDE method, the diffusion network approach significantly increased the prediction accuracy. The ability of machine learning algorithm in processing non-linear relational data has been confirmed in many fields, including crime prediction. It has a faster training speed, can handle very high-dimensional data, and can also extract the characteristics of the data.

Disadvantages

Existing system has no research that at the same time

1. Evaluates its efficiency against a traditional hotspot policing approach implemented by the police
2. Provides a clear breakdown of the processing steps involved to implement such a predictive system. Small police departments, which often have more worrying demands for violence, may not be able to provide more efficient tools. If they want to build a prediction system, it can cost even more than buying one and they can take much time to build.

PROPOSED SCHEME

Crime hotspot prediction aims to predict the likely location of future crime events and hotspots where the future events would concentrate. In this project, random forest algorithm is used for crime prediction.
The randomness of random forest is reflected in two aspects: one is to randomly select the training sample set by using bagging algorithm; the other is to randomly select the split attribute set.The objective would be to train a model for prediction. The training would be done using the training data set which will be validated using the test dataset. Building the model will be done using better algorithm depending upon the accuracy. The Random Forest will be used for crime prediction. Visualization of dataset is done to analyze the crimes which may have occurred in the country

ADVANTAGES

- The purpose of this work is to improve our previously proposed prediction framework through alternative crime mapping and feature engineering approaches, and provide an open-source implementation that police analysts can use to deploy more effective predictive policing.
- This work helps the law enforcement agencies to predict and detect crimes in India with improved accuracy and thus reduces the crime rate.

SYSTEM BLOCK DIAGRAM

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Special Issue 4, April : 2024

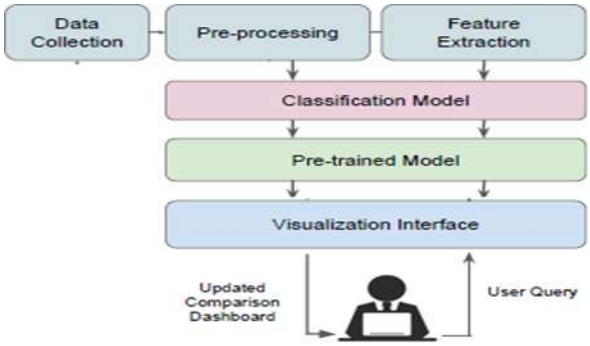

Fig1: System Architecture

RESULT ANALYSIS

Fig1: Login Form

Fig2: Data Upload

Fig 3:Predicting Theft Crime

Fig4: Predicting Murder Crime

Fig5: Predicting Rape Crime

Fig 6: Predicting Total Crime

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Special Issue 4, April : 2024

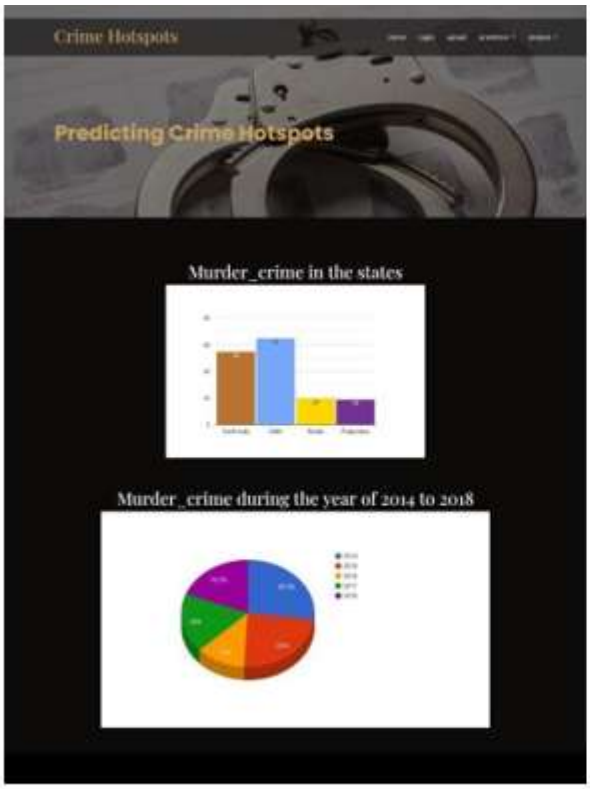

Fig 8.Data Analysis of Murder Crime

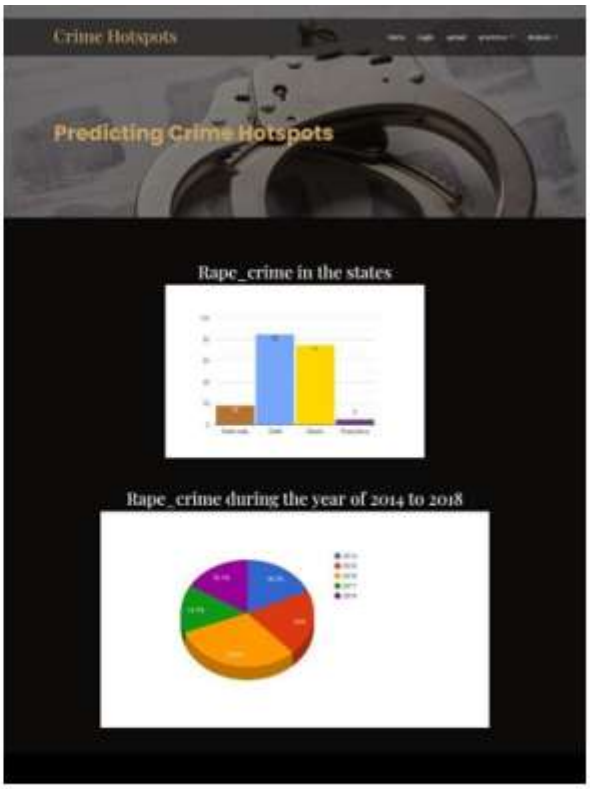

Fig 7: Data Analysis of Rape Crime

CONCLUSION

With the help of machine learning technology, it has become easy to find out relation and patterns among various data's. The work in this project mainly revolves around predicting the type of crime which may happen if we know the location of where it has occurred. Using the concept of machine learning we have built a model using training data set that have undergone data cleaning and data transformation. The model predicts the type of crime with Good Accuracy. Data visualization helps in analysis of data set. The graphs include bar, pie, line and scatter graphs each having its own characteristics. We generated many graphs and found interesting statistics that helped in understanding Indian crimes datasets that can help in capturing the factors that can help in keeping society safe.

REFERENCES

[1] U. Thongsatapornwatana, ``A survey of data mining techniques for analyzing crime patterns, in Proc. 2nd Asian Conf. Defence Technol. (ACDT), Jan. 2016,pp. 123128.

[2] J. M. Caplan, L. W. Kennedy, and J. Miller, ``Risk terrain modeling: Brokering criminological theory and GIS methods for crime forecasting, Justice Quart., vol.28, no. 2, pp. 360381, Apr. 2011.

[3] M. Cahill and G. Mulligan, ``Using geographically weighted regression to explore local crime patterns, Social Sci. Comput. Rev., vol. 25, no. 2, pp. 174193,May 2007.

[4] A. Almehmadi, Z. Joudaki, and R. Jalali, ``Language usage on Twitter predicts crime rates; in Proc. 10th Int. Conf. Secur. Inf. Netw. (SIN), 2017, pp. 307310.

[5] H. Berestycki and J.-P. Nadal, ``Self-organised critical hot spots of criminal activity,; Eur. J. Appl. Math., vol. 21, nos. 45, pp. 371399, Oct. 2010.

[6] K. C. Baumgartner, S. Ferrari, and C. G. Salfati, ``Bayesian network modelling of offender behavior for criminal proling,&; in Proc. 44 th IEEE Conf. Decis. Control, Eur. Control Conf. (CDC-ECC), Dec. 2005,
pp. 27022709.

[7] W. Gorr and R. Harries, ``Introduction to crime forecasting; Int. J. Fore-casting, vol. 19, no. 4, pp. 551555, Oct. 2003.