

A Survey on Crime Occurrence Detection and prediction Techniques

Shruti S.Gosavi

Assistant professor, Tilak Maharashtra Vidyapeeth, Pune

Shraddha S. Kavathekar

Master in Computer Engineering, PCCOE, PUNE

Abstract

Due to the increasing rate of crimes there is a need of system that will detect and predict crimes at dynamic time. The aim of this survey is to study Data Mining techniques that will help to detect and predict crimes using association rule mining, k-means clustering, decision trees & naive bayes and Machine learning techniques such as deep neural network and artificial neural network. Findings of this survey were that when the dataset instances have more number of missing values pre-processing becomes a vital task and crime does not occur uniformly across urban landscapes but concentrates in certain areas. So, predicting crime hotspots is a very important task also applying post-processing will help in lowering the rate of crimes.

Keywords

Crime hotspots, criminal profiling, naive bayes, association rule mining, clustering, crime detection, crime prediction.

I. Introduction

Crime Detection is using existing data of crime scene and criminals to extract patterns. By analyzing existing data on past crime, a person can predict when and where new crimes are most likely to occur. The increasing use of computerised systems to track crimes may boost the process of detecting and predicting crimes. Crime analysis is an important aspect in data mining field as there is a huge data at present that needs to be efficiently handled. A solution to this can be proposed using data mining techniques. Automated data collection has fostered the use of data mining for intrusion and crime detection. Indeed, banks, large corporations, insurance companies, casinos, etc. are increasingly mining data about their customers or employees in view of detecting potential intrusion, fraud or even crime.

In this survey the analysis of techniques that can be used for detecting and predicting crime is done to lower the rate of crime. Different techniques are studied such as Association rule mining, Naïve bayes Algorithm, K-Means Clustering, Decision Tree Algorithm, Artificial Neural Network, Deep Neural Network. clustering is Given a set of objects, clustering is the process of class discovery, where the objects are grouped into clusters and the classes are unknown before. Association rule mining aims to find the rules which enable us to predict the occurrence of a specific item based on the occurrences of the other items in the transaction. Frequent patterns are extracted using apriori algorithm to find relation between two crimes, criminals and crime scene. The Naïve Bayes (NB) algorithm is a widely used algorithm for review classification. The main advantages of NB are that they are simple, easy to implement, and comprise better performing algorithms.

Criminal Profiling is an investigative tool used by law enforcement agencies to identify likely suspects (descriptive offender profiling) and analyze patterns that may predict future offenses and victims. analyze criminal records and his background. Areas of concentrated crime are often referred to as hot spots. Researchers and police use the term in many different ways. Regions in map with excessive crime intensity are referred to crime hotspots. Crime Hotspots can be Visualized using Heat Maps or Geo-Plots.

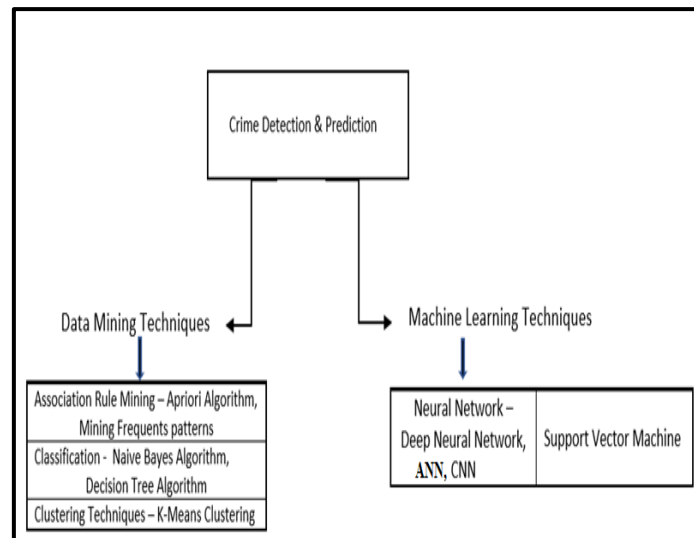


Fig1.1- Classification of Crime Detection & Prediction Techniques

II. Literature Survey

Papers used to study Data Mining Techniques for crime detection are explained below. The techniques are used for clustering crime prone regions, finding patterns in locations of crime, visualising patterns through maps. Crime analysis and prediction using data mining, 2014 IEEE Conference by S. Sathyadevan and S. Gangadharan. Dataset used was Real time data using blogs, social media and used MongoDB to store data. Techniques used were Naive Bayes, Apriori algorithm, Decision Tree. The methodology is to Predict regions that have high probability for crime occurrence and visualize crime prone areas, criminal profiling. The technique was able to predict crime prone regions in India for a particular day and Naive Bayes gives 90% accuracy for their dataset. Strengths are Predict regions that have high probability for crime, visualize crime prone areas. Weakness is predicting the time in which crime is happening.

Crime data analysis using data mining techniques to improve crime prevention IJCA by D. Z. ZUBI and A. MAHMMUD. The dataset was Taken manually from Libyan police department. Data mining algorithms used are Association rule mining, k-means clustering. The Methodology is to extract crime patterns and clustering them to classify crime records. The Technique was able to extract Relation between criminal age and no of crimes is high. Clusters are made using location and crime type for analysis. Strength is Analysing crime patterns become easy due to clusters. Weakness is cannot predict for multi-modal data or high dimensional data.

Crime pattern detection, analysis and prediction IEEE Conference 2017 by Sunil Yadav, Meet Timbadia, Ajit Yadav, Rohit Vishwakarma and Nikhilesh Yadav. The dataset was taken from Online portal of India (2001-2014). Data mining algorithms used are Apriori, k-means clustering, naïve Bayes, correlation and regression. The methodology is to Create clusters and mining frequent data, classification and finding correlation, regression from dataset. Detect crime for different states. The technique was able to identify Correlation between states and rate of crime which is 0.98. Regression shows that out of 10 cases only 3

are convicted of the charges. Strength is Crime prediction is done based on states and age group with respect to dates. Weakness is cannot predict crime hotspots with respect to time, handling insufficient data is difficult.

And Machine Learning Techniques include same additional making use of images for crime detection and prediction are as shown below. Prediction of Crime Occurrence from multi modal data using deep learning Research Article, 2017. The dataset used was American Fact-Finder 2014, weather data, Google street view images. The Techniques used are DNN, Pearson correlation -coefficient analysis, SoftMax classifier. The methodology is to predict to Crime occurrence by fusing multi-modal data, using criminal activity records in certain areas to predict the occurrence of crimes. Accuracy is 84.25 for multi modal data fusion using DNN, also it can Efficiently fuse multi-modal data with environmental context information. The paper Predicts occurrences of crime using past criminal activity. Strength is It works well on high dimensional and multi-model data. Weakness is the DNN –Based crime occurrence and prediction cannot be applied on insufficient data.

CRIMECAST: A Crime Prediction and Strategy Direction Service IEEE,2016 by Mahmud-Zinnah. The dataset is prepared, probability of crime i in location s is calculated, hotspot detection is done, & using ANN crime-cast is implemented. Calculating probability factor of each crime for location. Calculating probability factor of each crime for every location using ANN. The system proposed a crime prediction model CRIMECAST simulating in statistical and ANN implementation. Strengths is the ANN implementation gives us more precise prediction than mathematical implementation.

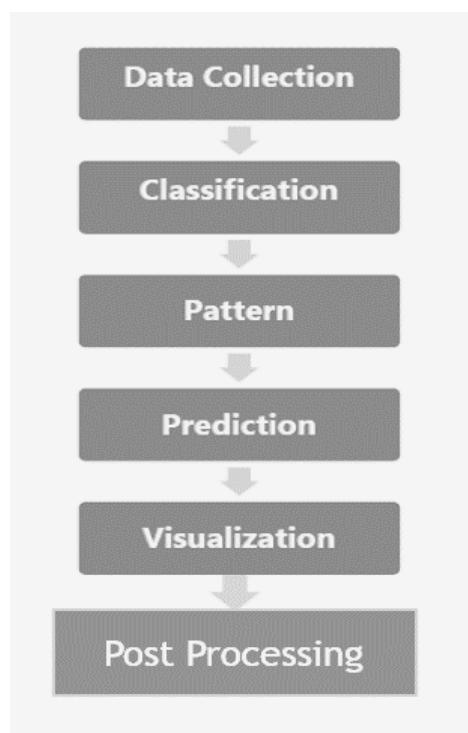


Fig 1.2 Crime Prediction workflow

The model for crime prediction and detection may change by making use of either the data mining techniques or machine learning techniques. As per the literature survey the crime detection process works as follows:

1. **Creating a Map Grid –**

The area to be predicted is manually defined. The crime prone areas can be graphically represented using a heat map which indicates level of activity, usually darker colours to indicate low activity and brighter colours to indicate high activity. The main responsibility

of the data visualization is to create images, diagrams, or animations to provide data summarization.

2. **Data Collection & Pre-processing –**

The fields that will be useful have to be extracted which are the date when crime was committed, type of crime, address of the committed crime including the latitude and longitude as well as criminal information.

3. **Model for Crime Prediction –**

Predictive policing starts with collecting large amounts of data on past crime and co-relating them with present crimes. Analyze Data: Software looks for patterns and correlations in past crime data. Predictive Maps: Algorithm predicts where and when a crime is likely to happen in future. Increased surveillance: Based on the possibility of the crime, redeploy resources during certain period to prevent crime.

4. **Statistical Analysis –**

The results generated by the neural network are to be then analysed using the Sum of Squares Error (SSE) of the output values and the target values, also accuracy needs to be checked.

5. **Visualizations of Results -**

The output of the crime prediction model is probability of crime & non-crime occurrences. Outcomes of predicted data is visualized using heat maps or geographic plots.

6. **Postprocessing –**

Increased surveillance: Based on the possibility of the crime, redeploy resources during certain period to prevent crime.

III. **Conclusion**

This survey reviews the literatures on various techniques and applications that are applied to solve the crimes. Thus, for certain regions crime hotspots need to be predicted to identify which areas are more prone to crimes and the type of crime. This survey concludes that Crime does not occur uniformly across urban landscapes but concentrates in certain areas. The data set to be processed is huge so pre-processing and handling missing values becomes an important task. For predicting future crimes from large amount of data the Artificial Neural Network and Deep Neural Network can be used for Detecting Crime as well as Predicting Crime using past crime datasets. When data is supervised, data mining algorithms can be preferred, and when data is multi-modal, huge & unsupervised the deep learning techniques can be used.

IV. **References**

- [1] S. V. Nath, "Crime Pattern Detection Using Data Mining," in *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, Washington, DC, USA, 2006*, pp. 41–44.
- [2] S. Nath, "Crime Pattern Detection Using Data Mining," *Proc. 2006 IEEE/WIC/ACM Int. Conf. Web Intell. Intell. Agent Technol. WI-IAT 2006 Work-IATW06 0-7695-2749-306 2006* © 2006, Jan. 2007.
- [3] M. A and S. Santhosh Baboo, "An Enhanced Algorithm to Predict a Future Crime using Data Mining," *Int. J. Comput. Appl.*, vol. 21, no. 1, pp. 1–6, May 2011.
- [4] J. Azeez and D. J. Aravindhar, "Hybrid approach to crime prediction using deep learning," in *2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2015*, pp. 1701–1710.
- [5] U. Thongsatapornwatana, "A Survey of Data Mining Techniques for Analyzing Crime Patterns," *Def. Technol. ACDT 2016 Second Asian Conf. On, A Survey of Data Mining Techniques for Analyzing Crime Patterns*.

- [7] P. Thongtae and S. Srisuk, "An Analysis of Data Mining Applications in Crime Domain," in *2008 IEEE 8th International Conference on Computer and Information Technology Workshops*, 2008, pp. 122–126.
- [8] S. Sathyadevan and S. Gangadharan, "Crime Analysis and Prediction Using Data Mining," *First Int. Conf. New. Soft Comput.*, 2014.
- [9] D. Z. ZUBI and A. MAHMMUD, "Crime Data Analysis Using Data Mining Techniques to Improve Crimes Prevention," *Int. J. Comput.*, vol. 8, 2014.
- [10] "Crime Hot-Spots Prediction Using Support Vector Machine - IEEE Conference Publication." <https://ieeexplore.ieee.org/document/1618468/>.
- [11] K. Sukhija, S. N. Singh, and J. Kumar, "Spatial visualization approach for detecting criminal hotspots: An analysis of total cognizable crimes in the state of Haryana," in *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT)*, 2017, pp. 1060–1066.
- [12] "Crimecast: A crime prediction and strategy direction service - Semantic Scholar." Available: [/paper/Crimecast%3A-A-crime-prediction-and-strategy-service-Mahmud-Zinnah/89f4dc20e2c1e713c03f2eac9855411d57fc9b4b](https://arxiv.org/abs/1808.08441).
- [13] M. A. Tayebi, M. Ester, U. Glässer, and P. L. Brantingham, "CRIMETRACER: Activity space-based crime location prediction," in *2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014)*, 2014, pp. 472–480.
- [14] "Prediction of crime occurrence from multi-modal data using deep learning." <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0176244>.
- [15] S. Sivaranjani, S. Sivakumari, and M. Aasha, "Crime prediction and forecasting in Tamilnadu using clustering approaches," in *2016 International Conference on Emerging Technological Trends (ICETT)*, 2016, pp. 1–6.
- [16] Sunil Yadav, Meet Timbadia, Ajit Yadav, Rohit Vishwakarma and Nikhilesh Yadav, "Crime Pattern Detection, Analysis & Prediction", in *2017 International Conference on Electronics, Communication and Aerospace Technology ICECA 2017*