Predictive Analysis Model for Recommending Police Personnel

Gourav Joshi
19MCA0209
Master of Computer Application

School of Information Technology
Vellore Institute of Technology
Vellore, India
gouravjmg123@gmail.com

Vaibhav Bhana
19MCA0002
Master of Computer Application

School of Information Technology
Vellore Institute of Technology
Vellore, India
vaibhav.bhana2019@vitstudent.ac.in

Prashant Dhirendra
19MCA0065
Master of Computer Application

School of Information Technology
Vellore Institute of Technology
Vellore, India
prashant.dhiru@gmail.com

Literature Survey

Manish Gupta, B. Chandra and M.P. Gupta tries to understand the current scenario of Indian Police System and they apply data mining tool and techniques to real data set of Karnataka state and the applied it on data of whole state, Concluded that use of now a days technology will help effectively in decreasing crime rate and also help to find hot spot of crime etc things ("Crime data mining for Indian police information system.", Manish Gupta, B. Chandra and M.P. Gupta) [1].

Author developed a Business Intelligence solution for Personnel planning (Military Staff). Author used data mining concepts and work on how data mining process can carry out, how data from different can preprocessed and represented so that process may become more feasible. Author used Web based Application like OLAP and Google chart tool to represent data extracted ("Development of business intelligence solution for personnel administration.", Keerin, Phimmarin) [2].

They predict Employee Churn by using Data Mining. Since churn analysis done on customers mostly so they did it on data set of Employee of a Company and implemented data Classification techniques such as decision tree, Naive Bayes etc. ("An approach for predicting employee churn by using data mining.", Yiğit, İbrahim Onuralp, and Hamed Shourabizadeh) [3].

In year 2008 authors used Data Mining Technique to make effective use of Human Resource in construction company. They developed data mining framework to extract useful information about relation between personnel and work behavior. This helps in getting the right person for a job ("Data mining to improve human resource in construction company.", Youzheng, Chang, and Guan Ming) [4].

Sangita Gupta and V. Suma used data mining techniques like decision tree for getting right professional according to project requirement in field of Software Engineering so that the quality of software delivered get boosted ("Data mining: A tool for knowledge discovery in human aspect of software engineering.", Gupta, Sangita, and V. Suma)[5].

Compared different regression models like linear regression, support vector regression, decision tree, forest regression and find out which regression model gives better result ("A Comparison of Regression Models for Prediction of Graduate Admissions.", Acharya, Mohan S., Asfia Armaan, and Aneeta S. Antony) [6].

Author of this paper used data-driven approach to predict the short-term crime forecasting. They used deep learning architecture like artificial neural network. Type of ANN (i.e. LSTM, RNN etc.) is decided by using sum of least squared error and virtual leave-one-out test (VLOO) is used to optimize the number of hidden nodes. They were successfully able to train their NN on unstructured police records for Spatiotemporal data. (2018: Zbigniew M. Wawrzyniak, Stanisław Jankowski, Eliza Szczechla, "Data-driven models in machine learning for crime prediction") [7].

Author of this paper was motivated that by the relation between human mobility influencing crime behavior in big cities, this relation can be used to allocate police in to certain spots in the city. It was found that the correlation between police officer and FPA (floating population allocation) is much higher than RPA (resident population allocation), hence the allocation based on the clusters of floating population tends to be more adequate for fighting crime against properties because the distribution of police resources will naturally follow. (2017: Carlos Caminha, Vasco Furtado, "Impact of Human Mobility on Police Allocation") [8].

The authors state that crime records is highly spatial-temporal in nature where the traditional system of criminal records has failed to maintain the desired level of intelligence. To have usable information from time series data, big data and ARIMA (auto regressive integrated moving average) model was proposed. Crime records from National Crime Record Bureau for the state of Haryana(India) was used found that the Generalized Linear Model (GLM) for Crime Site Selection (CSS) using Big Data deliver better results and forecast spatio-temporal crime events with certainty (2018 : Romika Yadav , Savita Kumari Sheoran, "Crime Prediction Using Auto Regression Techniques for Time Series Data")[9].

In this paper the goal of the author is to estimate the staffing levels over different shifts on a time range of days in order to optimize the overall Estimated Time to Restoration (ETR) so that the maximizing crew efficiency can be achieved. The full mathematical model and the constraint programming formulation for the planning problem were presented to the paper. this exposes several decision knobs (node in decision tree) and constraints programming to model complex business constrains into a scheduling problem. (2015: Ali Koc, Amith Singhee, Ashish Sabharwal, Richard Mueller, Gerard Labut, "Optimal spatio-temporal emergency crew planning for a distribution system") [10].

In today's world security is an aspect which is given higher priority by all political and government worldwide and aiming to reduce crime incidence. As data mining is the appropriate field to apply on high volume crime dataset and knowledge gained from data mining approaches will be useful and support police force. So, in this paper crime analysis is done by performing k-means clustering on crime dataset using rapid miner tool [11].

Concern about national security has increased after the 26/11 Mumbai attack. In this paper we look at the use of missing value and clustering algorithm for a data mining approach to help predict the crimes patterns and fast up the process of solving crime. We will concentrate on MV algorithm and Apriori algorithm with some enhancements to aid in the process of filling the missing value and identification of crime patterns. We applied these techniques to real crime data. We also use semi-supervised learning technique in this paper for knowledge discovery from the crime records and to help increase the predictive accuracy [12].

Data mining is the computer-assisted process of digging through and analyzing enormous sets of data and then extracting the meaning of the data and it is the process of analyzing data from different perspectives and summarizing it into useful information. Data mining plays an important role in terms of prediction and analysis. Crimes are a social nuisance and cost our society dearly in several ways. Crime investigation has very significant role of police system in any country. Clustering is the task of grouping a set of objects in such a way that objects in the same group are more similar to each other than to those in other groups. This paper presents detailed study on clustering techniques and its role on crime applications. This study also helps crime branch for better prediction and classification of crimes [13].

This research paper highlights the importance of data mining technology to design proactive services to reduce crime incidences in the police stations jurisdiction. Crime investigation has very significant role of police system in any country. Almost all police stations use the CIPA system to store and retrieve the crimes and criminal data and subsequent reporting. It became useful for getting the criminal information but it does not help for the purpose of designing an action to prevent the crime. It has become a major challenge for police system to detect and prevent crimes and criminals. There is no any kind of information is available before happening of such criminal acts and it result into increasing crime rate. The presented paper highlights the use of data mining techniques for effective investigation of crimes [14].

Many police departments all around the world lack of good and efficient crime recording and analysis systems. The vast geographical diversity and the complexity of crime patterns have made the analyzing and recording of crime data even difficult. According to the Sri Lankan police department, they face these problems for many years. This paper presents an intelligent crime analysis and recording system designed to overcome problems that appear mainly in the Sri Lankan police department. The proposed system is a GIS based system which comprises of data mining techniques such as Hotspot detection, Crime clock, Crime comparison, Crime pattern visualization, Outbreaks detection and the Nearest police station detection. Salient features of the proposed system include a rich environment for crime data analysis and a simplified environment for location-based data analysis. It facilitates the identification of various types of crimes in detail and assists the police personals to control and prevent such incident efficiently. The SL-SecureNet was tested for about 1000 crime records. The test results indicated that it functions in an efficient and reliable manner [15].

Data Mining is the procedure which includes evaluating and examining large pre-existing databases in order to generate new information which may be essential to the organization. The extraction of new information is predicted using the existing datasets. Many approaches for analysis and prediction in data mining had been performed. But many few efforts have made in the criminology field. Many few have taken efforts for comparing the information all these approaches produce. The police stations and other similar criminal justice agencies hold many large databases of information which can be used to predict or analyze the criminal movements and criminal activity involvement in the society. The criminals can also be predicted based on the crime data. The main aim of this work is to perform a survey on the supervised learning and unsupervised learning techniques that has been applied towards criminal identification. This paper presents the survey on the Crime analysis and crime prediction using several Data Mining techniques [16].

This research paper highlights the importance of data mining technology in crime and criminal investigation to reduce crime incidences in Pune Rural Police Stations. Crime investigation has very significant role of police system in any country. Pune Rural Police Stations Crime and criminal record is stored and retrieved using CIPA and CCIS and it became useful for getting the criminal information but it does not help for the purpose of designing an action to prevent the crime, it has become a major challenge for police system to detect and prevent crimes and criminals. There is no any kind of information is available before happening of such criminal acts and it result into increasing crime rate. The presented paper highlights the use of Data Mining Time series technique to forecast crime in a specific area and seasonal crime pattern which can be useful for crime prevention [17].

Summary Table

Author	Title	Year	Finding
Manish Gupta, B. Chandra and M.P. Gupta	Crime data mining for Indian police information system.	2008	Data Mining techniques helps in decreasing crime rates.
Keerin, Phimmarin	Development of business intelligence solution for personnel administration.	2016	Data Visualization using web-based application like Olap and Google chart tool.
Yiğit, İbrahim Onuralp, and Hamed Shourabizadeh	An approach for predicting employee churn by using data mining.	2017	We can also use data mining concept in employee churn just like in customer churn.
Youzheng, Chang, and Guan Ming	Data mining to improve human resource in construction company.	2008	One more field where we can use data mining. In recruitment of workers.
Gupta, Sangita, and V. Suma	Data mining: A tool for knowledge discovery in human aspect of software engineering	2015	Selection of employee or professional in IT for particular project.
Acharya, Mohan S., Asfia Armaan, and Aneeta S. Antony	A Comparison of Regression Models for Prediction of Graduate Admissions	2019	Every model works better than other in different cases.
Zbigniew M. Wawrzyniak, Stanisław Jankowski, Eliza Szczechla	Data-driven models in machine learning for crime prediction	2018	virtual leave-one-out test is used to find the optimal number of hidden nodes for Spatiotemporal data
Carlos Caminha, Vasco Furtado	Impact of Human Mobility on Police Allocation	2017	correlation between FPA is more than that of RPA with crime rate.
Romika Yadav, Savita Kumari Sheoran	Crime Prediction Using Auto Regression Techniques for Time Series Data	2018	for time series data, big data and ARIMA model used for Crime Site Selection.
Ali Koc, Amith Singhee, Ashish Sabharwal, Richard Mueller, Gerard Labut	Optimal spatio-temporal emergency crew planning for a distribution system	2015	constraints programming to model complex business constrains into a scheduling problem
Jyoti Agarwal	Crime analysis using k-means clustering	2013	crime analysis is done by performing k-means clustering on crime dataset using rapid miner tool.
Malathi. A	An enhanced algorithm to predict a future crime using data mining	2011	semi-supervised learning technique for knowledge discovery from the crime records and to help increase the predictive accuracy.
A.Bharathi	A survey on crime data analysis of data mining using clustering techniques	2014	clustering techniques and its role on crime applications.
Hanmant N. Renushe	Data mining practices for effective investigation of crime	2012	CIPA system to store and retrieve the crimes and criminal data and subsequent reporting.

M.A.P. Chamikara	SL-SecureNet: intelligent policing using data mining techniques	2012	proposed system is a GIS based system which comprises of data mining techniques such as Hotspot detection, Crime clock, Crime comparison, Crime pattern visualization, Outbreaks detection
H. Benjamin Fredrick	survey on crime analysis and prediction using data mining techniques	2017	the survey on the Crime analysis and crime prediction using several Data Mining techniques.
Milind J. Joshi	Importance of data mining time series technique in crime and criminal investigation: A case study of pune rural police stations	2011	Data Mining Time series technique to forecast crime in a specific area and seasonal crime pattern which can be useful for crime prevention.

Reference

[1] Gupta, Manish, B. Chandra, and M. P. Gupta. "Crime data mining for Indian police information system." *Computer society of India 40.1 (2008): 388-397.*

[2] Keerin, Phimmarin. "Development of business intelligence solution for personnel administration.*" 2016 Second Asian Conference on Defence Technology (ACDT). IEEE, 2016.*

[3] Yiğit, İbrahim Onuralp, and Hamed Shourabizadeh. "An approach for predicting employee churn by using data mining.*" 2017 International Artificial Intelligence and Data Processing Symposium (IDAP). IEEE, 2017.*

[4] Youzheng, Chang, and Guan Ming. "Data mining to improve human resource in construction company." *2008 International Seminar on Business and Information Management. Vol. 1. IEEE, 2008.*

[5] Gupta, Sangita, and V. Suma. "Data mining: A tool for knowledge discovery in human aspect of software engineering." *2015 2nd International Conference on Electronics and Communication Systems (ICECS). IEEE, 2015.*

[6] Acharya, Mohan S., Asfia Armaan, and Aneeta S. Antony. "A Comparison of Regression Models for Prediction of Graduate Admissions." *2019 International Conference on Computational Intelligence in Data Science (ICCIDS). IEEE, 2019*

[7] Zbigniew M. Wawrzyniak, Stanisław Jankowski, Eliza Szczechla, "Data-driven models in machine learning for crime prediction" *Data-driven models in machine learning for crime prediction. 2018 26th International Conference on Systems Engineering (ICSEng). doi:10.1109/icseng.2018.8638230*

[8] Caminha, C., & Furtado, V. (2017) " Impact of human mobility on police allocation." *2017 IEEE International Conference on Intelligence and Security Informatics (ISI). doi:10.1109/isi.2017.8004886*

[9] Romika Yadav, Savita Kumari Sheoran, "Crime Prediction Using Auto Regression Techniques for Time Series Data"*3rd International Conference and Workshops on Recent Advances and Innovations in Engineering, 22-25 November 2018*

[10] Ali Koc, Amith Singhee, Ashish Sabharwal, Richard Mueller, Gerard Labut, "Optimal spatio-temporal emergency crew planning for a distribution system*" 2015 IEEE Power & Energy Society Genera.*

[11] Agarwal, Jyoti, Renuka Nagpal, and Rajni Sehgal. "Crime analysis using k-means clustering." *International Journal of Computer Applications* 83.4 (2013).

[12] Baboo, S. Santhosh. "An enhanced algorithm to predict a future crime using data mining." *International Journal of Computer Applications* 975 (2011): 8887.

[13] Bharathi, A. Shilpa, and R. Shilpa. "A survey on crime data analysis of data mining using clustering techniques." *International Journal of Advance Research in Computer Science and Management Studies* 2.8 (2014): 9-13.

[14] Renushe, Hanmant N., Prasanna R. Rasal, and Abhijit S. Desai. "Data mining practices for effective investigation of crime." *International Journal of Computer Applications in Technology* 3 (2012): 865-870.

[15] Chamikara, M. A. P., et al. "SL-SecureNet: intelligent policing using data mining techniques." *International Journal of Soft Computing and Engineering* 2.1 (2012): 175-180.

[16] David, H., and A. Suruliandi. "survey on crime analysis and prediction using data mining techniques." *ICTACT journal on soft computing* 7.3 (2017).

[17] Reuse, H., Milind J. Joshi, and R. Rascal. "Importance of data mining time series technique in crime and criminal investigation: A case study of pune rural police stations." *International Journal of Computer Applications* 30.9 (2011).