

Criminal Identification And Comprehensive Analysis Using Deep Learning: A Review

Laveena Sehgal¹, Dr. P.K. Bharti², Dr. Mahendra Sharma³

¹Ph.D. Scholar (Computer Science & Engineering) Shri Venkateshwara University.

²Professor Shri Venkateshwara University.

³Associate Professor (Information Technology) IIMT College of Engineering, Greater Noida, UP.

Abstract: Deep learning mechanism is a discipline that consists of processes for creation of pattern in very systematic manner so that prediction, forecasting and scaling possible for massive amounts of data. It consists of techniques that are based on the convergence of machine learning and database architecture. Many different disciplines, such as future healthcare, market basket analysis, educational institutions, industrial engineering, and criminal investigation, might benefit from it. The discipline of criminology is one of the most significant for the use of Deep Learning. Criminology is a procedure with the goal of identifying the characteristics of criminal behaviour. Crime analysis entails investigating and identifying crimes as well as the linkages between them. There are many applications that contribute to the welfare of society by assisting law enforcement authorities in lowering crime rates. Crime detection and criminal identification is one of these applications. The Crime Department of India will profit from this technique since it will allow for more accurate crime analysis and prediction. It is the purpose of this research to examine the crime patterns in different Indian states and union territories.

Keywords: Deep Learning, Crime analysis, Crime detection, Crime identification.

I. Introduction

Mining of data is a technique for dealing with large data indexes in order to recognise patterns and build up a group to deal with difficulties via information analysis. The devices that have been used enable for future samples to be accepted. Deep Learning is a method for analysing data from an informational collection in order to transform it into a suitable structure that can be used for further processing [1]. Considering the uniqueness of applying Deep Learning methods to the area of criminal investigation, additional research methodologies are explored in the following parts of this paper. Data clustering and artificial neural networks are two of the approaches [2] that may be used. The first aids in the classification of crimes in an unsupervised way, whilst the second may be used to identify criminal patterns in an unsupervised method.

Clustering is the division of a collection of data or objects into several groups or clusters. Thus, a cluster is formed of data that is like one another and behaves in the same way as a group. It is possible to say that clustering is equivalent to classification, with the main difference being that the classes are not established and chosen in advance, and the data is grouped without the involvement of a supervisor.

Crime Analysis: With the help of machine learning algorithm, using python as core we can predict the type of crime which will occur in a particular area. Building the model will be done using better algorithm

depending upon the accuracy. The K-Nearest Neighbour (KNN) classification and other algorithm will be used for crime prediction.

Significance of Crime analysis in India

- ✓ Support the operations of a police department such as criminal investigation, apprehension, and prosecution, crime prevention and reduction strategies etc.
- ✓ Crime analysts study crime reports, arrests reports, and police calls for service to identify emerging patterns, series, and trends as quickly as possible [3].

B. Crime Detection: In most nations, the police are responsible for the detection of criminal activity, while law enforcement organisations may be tasked with the finding of certain forms of criminal activity (e.g., customs departments may be charged with combating smuggling and related offenses). Crime detection may be divided into three distinct phases: the discovery that a crime has been committed, the identification of a suspect, and the gathering of sufficient evidence to indict the suspect in front of a court of law. Many crimes are found and reported by somebody other than the police. This is a common occurrence (e.g., victims or witnesses) [4].

II. Crime Detection Techniques

DNA Profiling: Deoxyribose Nucleic Acid (DNA) is an organic material present in every living cell, and it is responsible for providing each cell with a genetic blueprint that is nearly entirely unique. DNA may be collected from a broad range of sources present in human beings, including blood, sperm, bone, saliva, and other bodily fluids.

STR Analysis: STR is a microsatellite in the sense and it consists of a unit of two to thirteen nucleotides repeated hundreds of times in a row of the DNA strand. The Federal Bureau of Investigation (FBI) uses a standard set of 13 specific STR regions for CODIS.

Paternity: It is well established that when certain fact is considered as conclusive proof of another fact, the judiciary generally disables the party in disrupting such proof. The only exception occurs when the party is able to show that there was no access to the other party when the conception could have taken place. Whenever paternity is contested, the burden of proof is on the party pleading negative.

Polygraph test: The psychological stress evaluator, the polygraph or any “lie detection” device cannot detect deception per se. rather, such machine record reaction to a given situation, most commonly a questions and answer session. Polygraph test have been traced back to the time of Lombroso, who experimented with the machine that measured blood pressure and pulse rate to examine the truthfulness of the statement and honesty of suspected persons in the crime.

Fingerprints: Fingerprint identification become without question the most trusted type of forensic evidence in criminal trials. Judges and juries have shown themselves readily willing to believe that no two fingerprints are alike and that fingerprint examiners can reliably trace “latent prints”, fragments of fingerprints patterns found at the crime scene, to one and only one person. [5]

III. Methods and Remarks of Criminal Identification

Criminal identification is an applied science that involves the study of facts that are then used to inform criminal trials. A complete criminal investigation can

include searching, interviews, interrogations, evidence collection and preservation, and various methods of investigation.^[1] Modern-day criminal investigations commonly employ many modern scientific techniques known collectively as forensic science [6].

Methods of criminal identification

- Height
- Head Length
- Head Breadth
- Arm Span
- Sitting Height
- Left Middle Finger Length
- Left Little Finger Length
- Left Foot Length
- Right Ear Length
- Cheek Width

These measurements were organised into a complex indexing system using specially trained clerks. When trying to identify someone, these clerks would navigate through the filing system by searching smaller and smaller sub-categories of the measurements, and produce ‘anthropometric cards’ of all the people who matched the measurements. Because the odds of two people having exactly the same eleven measurements were low, process of elimination based on facial description could be applied to the small number of possible people [7].

Fingerprinting: Fingerprints have been used as signature since ancient times. In ancient Babylon, for example, fingerprints were used on clay tablets for business transactions, and in ancient China thumb prints have been found on clay seals. It is unknown, however, whether or not these civilisations recognised the unique characteristics of fingerprints [8].

IV. Related Work

They (J. Agarwal, R. Nagpal, and R. Sehgal) conducted an analysis of crime and analysed murder crime, taking into consideration the corresponding year, and found that the trend is downward between the years 1990 and 2011. In order to extract relevant information from the crime dataset, they have employed the k-means clustering approach, which has been implemented using the RapidMiner tool, which is a robust and full package with configurable support options. [9].

In their study, Priyanka Gera and Dr. Rajan Vohra employed a linear regression to predict the incidence of crimes in Delhi (India). They examine a dataset spanning the previous 59 years in order to forecast the incidence of various crimes such as murder, burglary, robbery, and so on. Their work will be beneficial to the local police stations in terms of decision-making and criminal monitoring, among other things. Following training, the systems will be able to estimate data values for the following fifteen years. The system is trained by performing linear regression on data from the previous year. This will result in a formula as well as a squared correlation.

L. Ding and colleagues created an integrated system named PrepSearch, which stands for preparation search. It has been integrated by combining two distinct sorts of visualisation tools: those that provide a spatial picture of crimes and those that allow for visualisation of social networks. A supplied description of a crime, including the location and kind of crime, as well as the physical description of suspects (personal characteristics), will be used to programme the system.

Bogahawatte and Adikari conducted research that resulted in the development of an intelligent criminal identification system (ICIS), which has the capacity to differentiate a criminal based on observations gathered from the crime scene for a certain class of crimes. When identifying a criminal, the system makes advantage of existing evidence in the case by using a clustering technique to partition crime data into subsets, and the Nave Bayesian classification has been utilised for identifying likely suspects in crime episodes, among other things. ICIS has made advantage of the communication capabilities of a multi-agent system to increase the efficiency with which it is able to identify potential suspects. User interface, managed bean, multi agent system, and database are the components of ICIS that are described in detail in this document. The Oracle Database is being utilised for the database implementation, while the Java platform is being used for the identification of crime patterns implementation.

A. Babakura, N. Sulaiman, and M. Yusuf suggested a better approach of classification algorithms for crime prediction in, which was published in the journal Science. A comparison of Nave Bayesian and Back Propagation (BP) classification algorithms for predicting crime category in different states of the United States was conducted. The model is developed on the training data in the first step phase, and the model is implemented in the second phase of the process. When comparing classification algorithms, performance metrics such as Accuracy, Precision, and Recall are utilised to evaluate their effectiveness. When BP is employed as a classifier, the accuracy and recall stay the same as before.

Chung-Hsien Yu, Chung- The authors of this work focussed on developing a forecasting model in collaboration with the police department of a city in the United States' Northeast region. After extracting the dataset from the original crime record, the technique goes on to analyse the dataset, which comprises information such as the crime location, time, and other crime-related aspects. For the purpose of crime forecasting, this strategy makes use of a categorization technique. After evaluating several different classification approaches, the optimal classification method is selected.

Chung-Hsien Yu, Chung-Hsien Yu, The authors of this work focussed on developing a forecasting model in collaboration with the police department of a city in the United States' Northeast region. After extracting the dataset from the original crime record, the technique goes on to analyse the dataset, which comprises information such as the crime location, time, and other crime-related aspects. For the purpose of crime forecasting, this strategy makes use of a categorization technique. After evaluating several different classification approaches, the optimal classification method is selected.

Arunima & Raju Gopal, Arunima & Raju Gopal, The authors of the research make use of data mining approaches such as clustering and classification to achieve their results. In India, a technology is being created to analyse crime information, which will aid in the automation of the investigative process there. The tool that has been created is quite beneficial in detecting offenders and, as a result, aids in the speeding up of the investigation.

Crime Detection Using Text Recognition and Face Recognition is a paper published by Bachhety et al. (2018). Text and face recognition processes for criminal identification are being investigated in this study with the goal of providing a better result, both in terms of speed and accuracy, than currently available traditional approaches. In order to avert such a catastrophe, we have devised a very effective criminal detection system that makes use of text and face recognition technology. For toll tax collection, parking system management, airport management, and border crossing management, such systems will be very successful. Text recognition may be used to identify a vehicle's licence plate, and facial recognition can be used to identify a suspect in a crime. The quality, lighting, and perspective of the picture collected are all important factors in the system's efficiency.

In 2015, Tayal and colleagues published Crime detection and criminal identification in India using data mining approaches. They portray these occurrences with the help of 35 specified criminal qualities. Accessibility to the criminal database is protected using security measures. Each of the other four modules is beneficial for a different kind of crime detection, criminal identification and prediction, and crime verification, for example. K-means clustering is used to examine crime detection since it repeatedly develops two crime groups that are based on comparable characteristics of the crimes that are being investigated. The use of a Google map enhances the depiction of k-means. Using KNN classification, it is possible to examine criminal identification and prediction.

The survey on crime analysis and prediction using data mining methods, conducted by David et al. (2017). The primary goal of this study is to conduct a survey on the supervised learning and unsupervised learning approaches that have been used in the field of criminal identification in recent years. This work includes a survey on the analysis and prediction of crime using various data mining approaches, which was conducted by the author.

Hussein et al. (2019), Improvement of Criminal Identification by Smart Optimization Method, Journal of Criminal Identification. Using the smart approach, the decision tree algorithm is utilised to categorise the data and optimise its structure in this work, as well as other works. This approach is used on two datasets: criminals from Iraq and criminals from India. The suggested technique has as its purpose the identification of offenders via the use of a mining method based on smart search. This contribution contributes to the acquisition of better results than those supplied by classic mining techniques by regulating the size of the tree by lowering leaf size. This is accomplished by managing the size of the tree by decreasing leaf size.

Monitoring the Impact of Economic Crisis on Crime in India Using Machine Learning is a paper published by Mittal et al. (2018). For the purposes of this work, machine learning methods have been used to crime data in order to determine the economic variables that influence crime in India. Experimental work has been carried out using a variety of machine learning algorithms, including Decision Trees, Random Forest, Linear Regression, and Neural Networks, with the unemployment rate and gross domestic product (GDP) as the independent input variables and theft, burglary, and robbery as the target variables, among others. In addition, the performance of the regression model has been examined in detail using the correlation coefficient r , the coefficient of determination R^2 , the absolute mean error, and the accuracy, among other metrics.

| S. No. | Paper Title | Author/Year | Technique | Conclusion |
|---------------|--------------------|--------------------|------------------|-------------------|
|---------------|--------------------|--------------------|------------------|-------------------|

| | | | | |
|---|--|----------------------|--|--|
| 1 | Crime detection and criminal identification in India using data mining techniques | Tayal et al., (2015) | k-means clustering, KNN classification | Crime verification of our results is done using WEKA [®] . This verifies an accuracy of 93.62 and 93.99 % in the formation of two crime clusters using selected crime attributes. |
| 2 | Crime against women: Analysis and prediction using data mining techniques | Kaur et al., (2019) | Data Mining predictive techniques | discusses the various data mining techniques, used for analysing and predicting crime or violence against women. |
| 3 | Behavioural analysis of crime against women using a graph-based clustering approach. | Das & Das (2017) | graph based clustering approach, clustering algorithm | The study evaluates the efficacy of Infomax clustering algorithm for detecting communities of states and union territories in India based on crimes. |
| 4 | Mitigating vulnerabilities through forecasting and crime trend analysis | Orong et al. (2018) | k-means clustering algorithm, Autoregressive Integrated Moving Average (ARIMA) model | Results showed that 3 of the indexed crime data were in the first group and five are in the second group. Moreover, rape, cattle rustling, physical injury, robbery, and theft showed an increasing pattern based on the forecasted data from 2015 to 2020 |
| 5 | Crime pattern detection, analysis & prediction. | Yadav et al. (2017) | regression model, supervised, semi-supervised and unsupervised learning technique | The system is trained by feeding previous year's record of crimes taken from legitimate online portal of India listing various crimes such as murder, kidnapping and abduction, dacoits, robbery, burglary, rape and other such crimes |

V. Crime Detection and Identification using Deep Learning

In recent years, Deep Learning has become a popular approach for analysing crime data that has been previously saved from multiple sources in order to uncover patterns and trends in criminal activity. Such research may aid in the speeding up of the criminal investigation process as well as the automated detection of offenders by electronic systems. The relationship between crime detection and Deep Learning methods is shown in the diagram below. Finally, the success of Crime Detection using Deep Learning is dependent on the selection of data, the selection of algorithms, and the selection of the platform on which the results are to be shown.

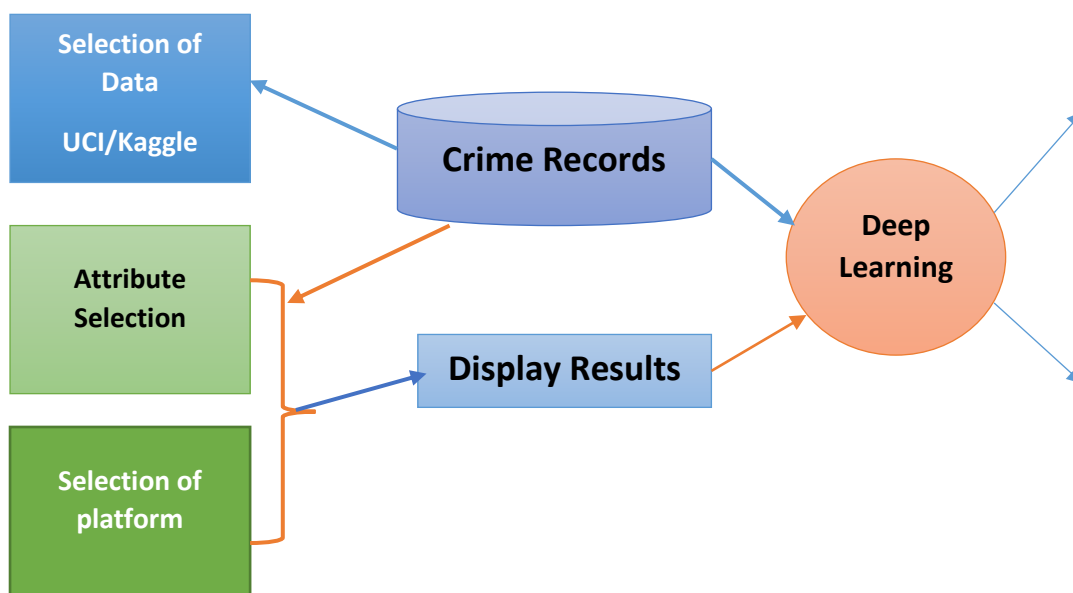


Fig 1: Crime Detection using Deep Learning

5.1 Criminal Identification using Deep Learning

In a criminal investigation, a criminal identification system is used to identify the suspect of a crime based on numerous characteristics and evidence gathered at the crime scene. To confirm the correctness of the results, we will first create a sample dataset, which will then be used to assess the J48, Nave Bayes, and Deep Learning methods against it. Their performance will be compared in terms of classification accuracy, correctly classified instances, true positive rate, false positive rate, precision, recall, and other factors. The algorithm with the best performance will then be selected for use in real-world crime and criminal data Learning scenarios, among other things. So, in summary, we are designing a criminal identification system that classifies sample crimes and criminal data in reference to numerous features before generating classification rules based on those classes and attributes. These categorization standards will be utilised to narrow down the list of potential suspects in the investigation. Following that, the suspect's information will be collected from the database. [24]

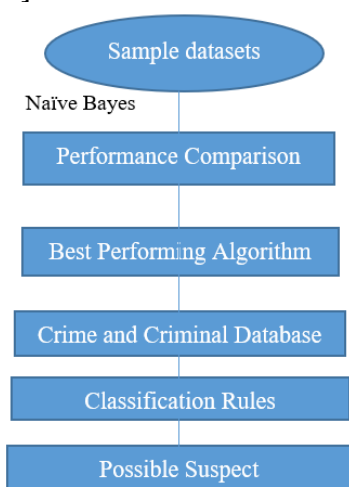


Fig 2: Illustrate the crime detection process using Naïve Bayes

IV. Conclusion and Future Work

We undertake crime analysis by examining crime murder and plotting it about year, and we get to the conclusion that homicide has decreased from 1990 to 2011. This research is concerned with crime analysis by applying clustering algorithm on crime dataset using fast miner tool. It is simple to establish a crime pattern over time based on the clustered data, which may be utilised to develop preventative measures for the foreseeable future. We may infer from the above research that Nave Bayes is the most efficient algorithm since it has the best accuracy among the three algorithms tested. In addition, Nave Bayes takes less time to categorise the data than other classification methods. Traditional criminal investigative techniques need a large amount of competent manpower as well as a large amount of documentation. There is a dearth of technological application in sensitive domains such as criminal investigation. As a result, criminal investigation has become a time-consuming endeavour. It is the process of obtaining meaningful information or knowledge from massive data sets that is known as Deep Learning. While a large quantity of information is gathered over the course of a criminal investigation, only the most valuable information is necessary for analysis. As a result, Deep Learning may be employed to accomplish this goal. The choice of a certain Deep Learning approach has a significant impact on the results that are achieved. This is the primary reason for comparing performance and selecting the best-performing Deep Learning method, as well as other factors. As a result, Deep Learning techniques may be used to crime and criminal data sources, allowing for the identification of potential criminal suspects as well as the tracking of their actions. The police department is experiencing difficulties with inter-agency coordination. A fingerprint recognition system, which may be installed at the police station level, can help to solve this issue.

References

1. <https://iopscience.iop.org/article/10.1088/1742-6596/1000/1/012046/pdf>.
2. <https://pdf.sciencedirectassets.com/280203/1-s2.0-S1877050911X00029/1-s2.0>.
3. https://en.wikipedia.org/wiki/Crime_analysis.
4. <https://www.britannica.com/topic/crime-law/Detection-of-crime>.
5. <https://www.biocoreopen.org/ijcf/Innovative-Forensic-Techniques-In-Crime-Detection-An-Appraisal.php>.
6. https://en.wikipedia.org/wiki/Criminal_investigation.
7. <https://www.police.govt.nz/about-us/history/museum/exhibitions/suspicious-looking-19th-century-mug-shots/other-methods-criminal-identification>.
8. <https://www.police.govt.nz/about-us/history/museum/exhibitions/suspicious-looking-19th-century-mug-shots/other-methods-criminal-identification>.
9. J. Agarwal, R. Nagpal, and R. Sehgal, —Crime analysis using k-means clustering, | International Journal of Computer Applications, Vol. 83 – No4, December 2013.
10. P. Gera, and R. Vohra, —Predicting Future Trends in City Crime Using Linear Regression, | IJCSMS (International Journal of Computer Science & Management Studies) Vol. 14, Issue 07 Publishing Month: July 2014.
11. L. Ding et al., —Perp-Search: an integrated crime detection system, | 2009 IEEE 161-163 ISI 2009, June 8-11, 2009, Richardson, TX, USA.
12. K. Bogahawatte, and S. Adikari, —Intelligent criminal identification system, | IEEE 2013 The 8th International Conference on Computer Science & Education (ICCSE 2013) April 26-28, 2013. Colombo, Sri Lanka.

13. A. Babakura, N. Sulaiman, and M. Yusuf, —Improved method of classification algorithms for crime prediction, | International Symposium on Biometrics and Security Technologies (ISBAST) IEEE 2014. [15] S. Sathyadevan, and S. Gangadharan, —Crime analysis and prediction using data mining, IEEE 2014.
14. Chung-Hsien Yu, Max Ward, Melissa Morabito, Wei Ding, “Crime Forecasting Using Data Mining Techniques”, 2011 IEEE 11th International Conference on Data Mining Workshops, 2011, DOI: 10.1109/ICDMW.2011.56.
15. Chung-Hsien Yu, Max Ward, Melissa Morabito, Wei Ding, “Crime Forecasting Using Data Mining Techniques”, 2011 IEEE 11th International Conference on Data Mining Workshops, 2011, DOI: 10.1109/ICDMW.2011.56
16. Arunima Kumar, Raju Gopal, “Data mining-based crime investigation systems: Taxonomy and relevance”, 2015 Global Conference on Communication Technologies (GCCT) – 2015, DOI: 10.1109/GCCT.2015.7342782
17. Bachhety, S., Singhal, R., & Rawat, K. (2018), Crime Detection Using Text Recognition and Face Recognition. *International Journal of Pure and Applied Mathematics*, 119(15), 2797-2807.
18. Tayal, D.K., Jain, A., & Arora, S. (2015), Crime detection and criminal identification in India using data mining techniques. *AI & Soc*, 3(3), 117–127.
19. David, H.B.F., & Suruliandi, A. (2017). Survey on crime analysis and prediction using data mining techniques. *ICTACT journal on soft computing*, 7(3), 2229-6959.
20. Hussein, R.R.A., & Croock, M.S. (2019), Improvement of Criminal Identification by Smart Optimization Method. *MATEC Web of Conferences*, 2(8).
21. Mittal, M., Goyal, L.M., & Sethi, J. (2018), Monitoring the Impact of Economic Crisis on Crime in India Using Machine Learning. *Comput Econ*
22. Jayapratha et al. (2018), real time criminology detection and criminal identification (RCDCI). *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 7(5), 2278 – 1323.
23. Tayal, D. K., Jain, A., Arora, S., Agarwal, S., Gupta, T., & Tyagi, N. (2015). Crime detection and criminal identification in India using data mining techniques. *AI & society*, 30(1), 117-127.
24. Kaur, B., Ahuja, L., & Kumar, V. (2019, February). Crime against women: Analysis and prediction using data mining techniques. In 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon) (pp. 194-196). IEEE.
25. Das, P., & Das, A. K. (2017, January). Behavioural analysis of crime against women using a graph-based clustering approach. In 2017 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-6). IEEE.
26. Orong, M. Y., Sison, A. M., & Hernandez, A. A. (2018, May). Mitigating vulnerabilities through forecasting and crime trend analysis. In 2018 5th International Conference on Business and Industrial Research (ICBIR) (pp. 57-62). IEEE.
27. Yadav, S., Timbadia, M., Yadav, A., Vishwakarma, R., & Yadav, N. (2017, April). Crime pattern detection, analysis & prediction. In 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA) (Vol. 1, pp. 225-230). IEEE.