Available online at www.sciencedirect.com

ScienceDirect

materialstoday: PROCEEDINGS

AIES 2021

A Survey on Crime Analysis and Prediction

Ashly Thomas[a], Sobhana N V[b] 1

[a]*Research Scholar, Rajiv Gandhi Institute of Technology, Kerala*
[b]*Professor, Rajiv Gandhi Institute of Technology, Kerala*

Abstract

A lot of police forces around the world have adopted mechanisms that use statistical data to guide their decision-making which is coined as predictive policing. As urbanization is increasing day by day, it is highly demanding to keep an eye on the criminal activities of each region so as to reduce the occurrences of unwanted behaviours. Prediction of crimes can be done only using the analysis of the patterns of criminal activities using the past data available with the concerned personals. This mainly makes use of the historic data and analyze them using Deep learning, Statistical Models and Algorithms. This paper makes a study on different approaches used worldwide for the prediction and forecast of crime occurrences. The methods are categorized and their effectiveness in various areas based on the precision and accuracy in their prediction is studied so as to show a light to the existing methodologies and to the need for future developments.

Selection and/or Peer-review under responsibility of International Conference on Artificial Intelligence and Energy Systems (AIES 2021)

Keywords: Deep learning; Neural networks; spatio-temporal; regression.

1. Introduction

In the past years, a lot of police forces around the world have adopted softwares that use statistical data to guide their decision-making which is coined as predictive policing. In this approach the police departments examine statistical historic data to foretell in which geographic areas there's an increased chance of occurrences of criminal activity. This type of data is often employed by law enforcers to efficiently deploy their resources to stop criminal behavior. Predictive policing cannot replace the conventional policing methods rather it enhances these usual practices by applying advanced statistical models and algorithms. Some empirical studies conclude that predictive policing strategies cause a decrease in crime [1]. Let us do as brief study on some of the existing techniques for statistically or probabilistically predicting the occurrences of crimes.

The way crime details are collected in the real world is as follows. Whenever a crime is committed, a police officer most probably visits the crime scene or the information regarding it is taken by telephone, with which they forms the crime report. The details stored may be in a variety of ways depending on the officials, but comprise the following: Time, day and date of the crime; Offence; Location of crime to include post code and Ordnance Survey grid references; Victim information; routine identifies how the crime has been committed. Depending on the crime recording system used by forces the data collected will be a mixture of free text and structured and validated fields. The free text may not even

* Corresponding author. Tel.: +91 9446803437.
*E-mail address:*ashlythomast@rit.ac.in

contain key words or phrases and will have abbreviations, spelling mistakes and, sometimes contradictory information. The quantity and quality of data recorded varies considerably from case to case.

Here comes the need for Artificial Intelligence and Deep learning to analyze thousands of crime reports, classify them and using statistical means predict the time and way a crime can be occurred. We are looking into eight different papers that deal with the above issue and try to identify the gaps in them. This paper takes into account a few commonly used analysis and prediction methods and groups the approaches into the following categories such as Neural Network approaches, Statistical approaches and spatio-temporal approaches and studies the level of accuracy shown by some of the different methods that comes under it.

2. Neural Network Approaches

Deep learning is a man-made intelligence function that copies the operations of the human brain in dealing with data and producing patterns to be used in decision making. Deep learning is a category of machine learning that deals with algorithms stimulated by the composition and function of the brain called Artificial neural networks and has networks able to learn unsupervised from data that's not structured or labeled. It is also referred to as deep neural learning or deep neural network. These approaches usually have the three stages - Pre-processing Stage, Processing Stage, and Post processing Stage.

A. Sarguru et al. [4], uses Artificial Neural Networks along with Recurring LSTM networks instead of traditional and time consuming data mining techniques for the prediction of crimes. Here the dataset is spliced into training, testing data and both are given to the training process. The resulting data are compared with count of crimes and depicted. The work was divided into 5 units. In the data collection step, data sets are collected from websites and converted into a format defined by multiple rows of column data. Next is the data selection step where repetitive or duplicate records are eliminated. In this records who have committed the crimes are trained and eliminated those which have been attempted. Signaling NaNs tend to generate exceptions and hence they must be eliminated or dropped from the table. After these steps the crimes are grouped in terms of monthly frequency. Next step is to create an RNN to forecast as in a typical Back-Propagation Through Time (BPTT), the error flowing backwards tends to fade away. This system is extremely efficient when there are recurrent pattern changes in the records with respect to time because it has a incorporated feed backing loop that allows it to operate as a forecasting engine. LSTM stands for long short term memory networks and are used to diminish the vanishing gradient problem. This dilemma arises when the activation functions in a RNN maps the input into a much small output range. This causes a large sum of inputs to get mapped to a minute collection. As a result, even a big change in the input will only result in a modest change in the resultant output.

M. Feng et al. [5], apply BDA to criminal data where inquisitive data analysis is conducted for visualization and trends forecast along with some data mining and deep learning techniques. The predictive results demonstrate that the Prophet model and Keras stateful LSTM attain better outcomes than the neural network models, in which the best size of the training data is found to be three years and it has a performance accuracy of 80% proved by them.

N. Esquivel et al. [6], proposes a Convolutional Neural Network (CNN) is a Long-Short Term Memory (LSTM) network (CLSTM-NN) to predict the existence of crime events over the town of Baltimore (USA). The proposed method is in a position to require under consideration spatial and temporal correlations present within the past data to enhance the future forecast. It makes use of Convolutional neural networks which are a kind of neural networks that are frequently used for automatic image and data identification tasks and it consists of a number of layers where each convolution layer normally takes into consideration different steps like convolution, pooling and activation operations.

Ulises M. et al. [7], proposes an approach based on deep learning for the categorization of crime incidents of public safety using predictive analysis. The predictive model is implemented on a neural network Long STM (LSTM), trained with a small set of attributes, enabling the forecast of the group tag in the validation stage, with a**n** elevated percentage of prediction precision. The prediction of the group tag is done by implementing stages such as pre-processing, categorization of the data, and by implementing a predictive model. This model shows an accuracy of 87.84%.

P. Stalidis et al. [10], studies on crime classification and prediction using deep learning architectures and examines the usefulness of deep learning algorithms in this realm. With 5 openly obtainable datasets, it is exhibited that the deep learning-based methods constantly do superior than the existing best-performing methods. Convolution Neural Networks contains many convolutional layers, which are characterized by an input map, a group of filters and biases that generate an output map. In the case of crime maps, by cumulating the crime incidents ‘xi’ to cells for every event type and time-span, and produces an ordered compilation of incident maps for a complete time period, with a defined height, width and

channels. LSTMs, are a variation of RNNs that are better suitable for lengthy sequences because they are not suffering from the issue of vanishing gradient effect. A blend of CNNs and RNNs can be together used for a spatiotemporal forecasting model. In SFTT method incident map is taken from a CNN sub model, constructing a feature vector time period that encodes the spatial allocation of incidents feature space that is less significant compared to the original. For the spatial feature extraction, the architectures based on VGGNet, ResNet and FastMask can be employed. A more complicated CNN has been proved to be more efficient in extracting spatial information in image categorization than VGGNet and ResNet, as the extra residual layers seek to resolve the degradation dilemma that was encountered when the convolutional networks becomes too deep. The temporal feature extraction is obtained by extracting the temporal correlations series of vectors holding the spatial features by using an LSTM sub model.

3. Statistical Approaches

For analyzing the records of crime to uncover the information about diverse perspectives of a crime event being happened, data mining techniques have been suggested [15] e.g. association rule mining, classification, regression varieties and clustering[20]. These methods are been widely used along with other models like LSTMs and RNNs to improve their efficiency and been establish that they outperform those models which are not using statistical methods.

Anusha K et al. [9], finds out the different questions of the statistics and attempt to answer those questions by means of the machine learning based crime forecasting solutions. The Predictive model gives you an idea that the prophet model handles outliers properly and also is robust to omitted or absent data and shifts the fashion. Data is accumulated from various sources like news sites, blogs, social media websites, RSS feeds etc., over a span of years. For every crime, 12 featured traits are incorporated explicitly Type, Year, Month, Day, Hour, Minute, Hundred_Block, Neighborhood, X, Y, Latitude, and Longitude. A sequence of preprocessing steps are performed for data conditioning and formatting such as the policies for entering dates, time, coordinates, etc. Considering the geographic nature of the crime incidents, an interactive map based on Google map was added for data visualization, in which the crime occurrences are clustered. Deep learning algorithms and time series methods were utilized for prediction. The Prophet model is a method for predicting time series data with respect to an additive model where non-linear trends and works most excellent with time series that have strong seasonal effects and a number of seasons of historical data. The Prophet model decomposes time series into three main components- the trend, seasonality, and holidays. To calculate the fault for this model the Mean Absolute Percentage Error is used (MAPE) and the error calculated is 12.64 which provide an accuracy of 87.36 making it a good choice for crime analysis and prediction process.

Biswas A. A. et al. [11], examines and forecasts the crimes in Bangladesh by means of linear regression, polynomial regression, and random forest regression which are used to predict the trends and patterns of crimes. The dataset contains documentation about different crime types like robbery, kidnapping, murder, women & child repression, theft, burglary, arms act, explosive, narcotics, and smuggling happened at Bangladesh. Initially, training of regression models on the training data for the prediction of crime is carried out on the test data with different regression models. In this 94% of collected records are made use of for the training purpose and 6% of data are used for testing purpose. Linear regression did the prediction of the value of the dependent variable from the independent variable. Thus, this architecture finds out a linear relation between dependent and independent variables. In the polynomial regression the same function uses polynomial arithmetic for calculation. Random Forest regression employs Ensemble learning method combined with a variety of base models in order to generate one predictive model that is most favorable and more powerful. From the investigational results, it is found that Polynomial Regression is much better than the others with an EVS of 0.95, MAE of 2832 and R2 Score 0.95. Hence, this approach is found to be one of the effective prediction techniques.

R. Yaday et al. [18], explored the Auto Regression Techniques to precisely predict the crime with lowest error for such time series data by recognizing the correlation among crime attributes. In this method, the standard deviation of every variable is determined to discover the most differing variable. The deviating variable predicted using autoregressive model is the key aspect of crime data set. Additionally clustering technique is applied to analogous crime trends, those have high probability of crime site selection by criminals. Also have been used ARIMA model in which a function of past value is considered. When forecasting past of the series, series values that are not yet observed are scrutinized.

W. Safat et al.[19] applied different machine learning algorithms, to be precise, the logistic regression, support vector machine (SVM), Naïve Bayes, k-nearest neighbors (KNN), decision tree, multilayer perceptron (MLP), random forest, and eXtreme Gradient Boosting (XGBoost), and analysis by long-short term memory (LSTM) and autoregressive integrated moving average (ARIMA) model to predict the crime data in a better manner. The performance of LSTM for

time series analysis was rationally sufficient in the basis of magnitude of root mean square error (RMSE) and mean absolute error (MAE), on both the analyzed data sets. Exploratory data analysis foretells more than 35 crime types and recommends a annual decline in Chicago crime rate, and a insignificant boost in Los Angeles crime rate. They have taken two big datasets namely Chicago and Los Angeles acquired from open access data portals. The bootstrap random sampling method is used here to produce statistically considerable data to create precise crime predictions. The RMSEs of the predicted crime rate for Chicago and Los Angeles were 31.8 and 24.65 and MAE was 29.8 and 20.83 correspondingly.

Mary Shermila et al.[21] detects crime patterns from implications gathered from the crime sight and predicts the description of the executor who is possibly assumed to have committed the crime. This work has two main aspects: Crime Analysis and Prediction of perpetrator personality. The Crime Analysis stage discover of unsolved crimes, and investigates the power factors like year, month, and weapon on these unsolved crimes. The prediction stage calculates approximately the features of the perpetrators like, their age, sex and relationship with the victim. These predictions are made using the proofs acquired from the crime scene. This system predicts the description of the crime executor by means of algorithms like, Multilinear Regression, KNeighbors Classifier and Neural Networks. Year, Month, Crime Type, Crime Solved, Victim Gender, Victim Age, Victim Race, Victim Count and Weapon are selected as input features for the scheme. The features such as Perpetrator Age, Perpetrator Sex and Relationship of the criminal with the victim are selected as the target variable that is to be predicted by the system.

4. Spatio-Temporal Approaches

This is a sophisticated forecasting methodology, which have the capacity for making use of all the accessible information collected from diverse locations. An ever-growing quantity of information offers a valuable chance for researchers and practitioners to forecast crime events and facilitate police departments to build up simpler approaches for avoiding occurrences of crime. As a result, mixed big datasets and spatiotemporal prediction of crime [12, 13] turn out to be very important. This method is generally implemented in smart cities and studies give you the idea that occurrences of criminal events are not evenly distributed within a city [16]. Because of the imbalanced distribution of crimes, it can be considered as a locality specific trait as some places can show evidences of greater threat of crime to be committed than some other regions [17].

H. Aitelbour et al. [8], uses historical data and convert it into spatiotemporal data based on the types of crimes and forecasts with precision the possibility of occurrence of crimes in a city. It is implemented by plotting the spatiotemporal points with the collected past data. This scheme has a number of layers and it takes into consideration two types of data, historical crime data (HIST DATA) and site information (LOC DATA) for performing crime prediction. The first layer handles data preparation in which it takes input and it does the process in 2 steps rather implemented as two layers: cleaning layer and transformation layer. The outcomes of those layers are been stored using the Storage layer and this accumulated data will be used by the feature selection layer that permits in selecting the most appropriate fields. Then it is separated into model training data and post-training testing data. The most vital layer in this model is transformation that contains the process of calculating a risk vector for every crime included in the considered dataset. These vectors play a part within the learning to predict, at each point in slice s, a vector of size n. Subsequent to the transformation phase, the dataset will comprise the prediction fields describing time and position in addition to the decision vector that enclose the risk values by crime type. Transformation was carried on the dates, times, day, month, year, and slice number. The slice number is obtained by dividing a day into four six-hour session.

X. Ye et al. [14], uses deep inception-residual networks (DIRNet) to carry out specific and precise, theft-related offense forecast based on non-emergency service request data. It sketches the employment of inception units containing asymmetrical convolution layers to draw low-level spatiotemporal dependencies concealed in crime dealings and grievance reports. This had considered 825,064 incidents for 3 theft-related types: burglary, grand theft, and petty larceny. Every incident studied had the properties of offense type, geographic coordinates, and time. Each grievance report included coordinates, complaint type, and hence the time at which the complaint was accepted and registered. An image registration deep learning network (DIRNet) is used for the prediction with an architecture involving 4 layers along with an inception and Residual Unit. This kind of an asymmetric neural structure can communicate the wide-ranging trends in a number of local spatiotemporal scopes while constructing cheaper computations for prediction and analysis with much less parameters.

S. K. Dash et al. [23], analyzes Chicago city crime data merged with a few other social information resources by means of network analytic procedures to predict criminal actions for succeeding year. This prediction models not only

guess the total amount of crimes for the entire Chicago city, rather they foretell the quantity of crimes for all sorts of crimes and that is done for different areas in the City of Chicago. This combines different types of social and historical data in a network that facilitates to uncover associations between different communities within the city and the identified relationships are integrated in the prediction model. To discover the social composition, communities are assembled and computed the similarity matrix between the communities which was used to take out two most similar communities for every reference community. To forecast the quantity of crimes in a given community for a certain year and month, three sorts of regression models, specifically, polynomial regression, support vector regression, and auto-regressive model are used. The effectiveness of complete network and data fusion in the prediction is instantly examined when the polynomial regression is applied.

U. M. Butt et al. [24], proposes an absolutely distinctive AI (AI) based approach empowering the authorities to raised envisage the risks and to support them recognizing the highly-reported crime sectors yielding a superior expectedness of crime hot-spots in a smart city. To this end, it foremost inspects the Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) to identify the hot-spots that have threat of crimes to be committed. For crime prediction, Seasonal Auto-Regressive Integrated Moving Average (SARIMA) is utilized in each intense crime section to predict the quantity of crimes in future with spatial and temporal data collected. It uses a 80:20 ratio where 80% of the labeled data is used for training purpose and 20% data was used for testing purpose. The projected approach does better than other methods with a Mean Absolute Error (MAE) of 11.47. The model consists of the following steps. After the data is acquired, as it is in raw form, in Step 1 it is getting converted into a processed form. Secondly the crime hot-spots are recognized by applying clustering algorithm. In third step 3, splitting of crime data for each hot-spot is performed. Lastly, forecasting algorithm is used to discover the amount of crimes that will be taking place for each hot-spot.

C.Catlett et al. [16], presents a predictive approach that supports spatial analysis and auto-regressive models to involuntarily identify high-risk crime sectors in urban areas and to unfailingly predict crime trends in every province. The algorithm produces a spatio-temporal crime forecasting model, created of crime-dense regions with connected crime predictors, all of them representing a predictive model for approximating number of crimes liable to take place in its associated sector. The experimental evaluation was carried out on two real-world datasets gathered inside the cities of Chicago and NY City. The input file of the investigation is that the set of gathered crime data to be processed and analyzed. The methodology consists of recognition of crime dense areas from the collected dataset. Next step does the spatial data splitting of the actual crime data, based on the clustering model. The third step is intended to extract a selected crime prediction model for every crime dense region (or the foremost representative regions), and scrutinize the crime data split during the previous step. When calculated about 40% of the crime events were detected correctly in the whole area.

5. Observations

From the above study, we have come to the following conclusions. An analysis is been done by considering the quality of dataset used, the variety of crimes considered or evaluation and the accuracy of the predicted output when compared with the actual scenarios. The comparison of different works that comes under Neural Network, Spatio-temporal and Statistical approaches are given in Table 1, Table 2 and Table 3 respectively. A graphical representation of the same is been depicted in the Fig 1 as three separate graphs plotting performance of Neural Network, Spatio-temporal and Statistical approaches.

Table 1: Details of different Neural Network approaches

Work	Methodology/Language Used	Variety of Crimes/ Dataset considered	Accuracy
[4]	RLSTM - Python	All Crimes	-
[5]	Keras LSTM - Python	All Crimes	80%
[6]	CLSTM - Python	City	86%
[7]	LSTM - Python	Smart City	87.84%
[10]	LSTM - Python	All crimes	76.48 %

Table 2: Details of different Spatio-Temporal approaches

Work	Methodology/Language Used	Variety of Crimes/ Dataset considered	Accuracy
[8]	Spatio temporal - Python	All Crimes	81%
[14]	Spatio-temporal- Python	All Crimes	71%
[23]	Spatio temporal - Python	City	-

[24]	Spatio- temporal - Python	Smart City	(MAE) of 11.47
[16]	Spatio temporal - Python	All crimes	(MAE) of 10.56

Table 3: Details of different Statistical approaches

Work	Methodology/Language Used	Variety of Crimes/ Dataset Considered	Accuracy
[9]	Prophet - Python	All Crimes	87.36 %
[11]	Regression - Python	Selected crimes in Bangladesh	95 %
[18]	Standard Deviation- R Programming	All Crimes	High
[19]	*ARIMA model* - Python	Chicago datasets Los Angeles datasets	94% 88%
[21]	K-neighbours - Python	All Crimes	60% - Age, 96% - Sex 97% Relationship

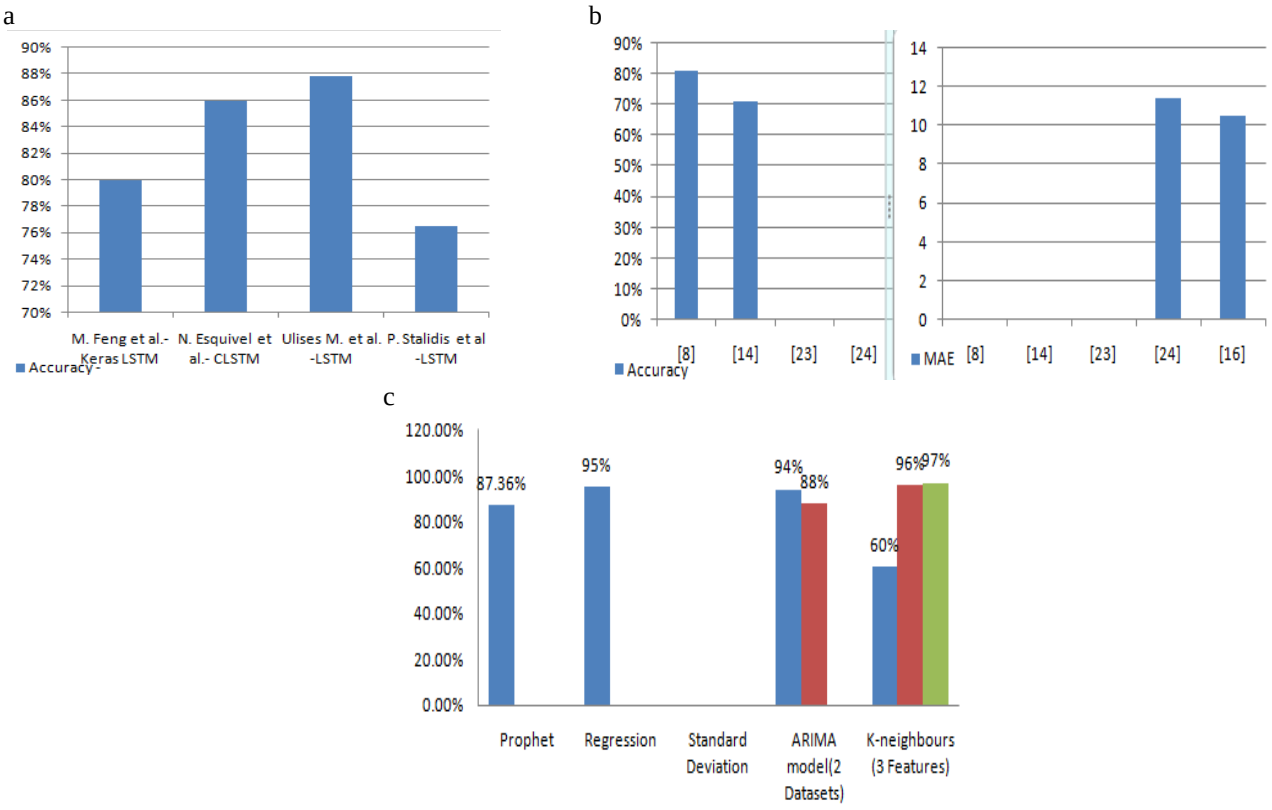

Fig. 1. Performance of (a) Neural Network (b) Spatio-temporal and (c) Statistical approaches.

Conclusion

This analysis mainly considered the performance of some of the models that are used for the prediction and analysis of crime occurrences based on the datasets having historical or past data of criminal activities happened in an area over the span of a fixed time period. It is found that normal neural networks or CNNs or the data mining techniques are not providing accurate results more than a threshold when used separately. But can be used combining with some other models. From the above study we can conclude that using LSTM and regression modelling we can attain better prediction results which matches to the actual scenarios to an average extend of 90% accuracy. Time series models can also be

compared to LSTMs for getting better results. In many of these referred methods combination of more than one model is done for improving the accuracy of prediction.

References

[1] Albert Meijer & Martijn Wessels (2019): Predictive Policing: Review of Benefits and Drawbacks, International Journal of Public Administration, DOI:10.1080/01900692.2019.1575664

[2] Camacho-Collados, M., & Liberatore, F. (2015). A decision support system for predictive police patrolling. Decision Support Systems, 75, 25–37. doi:10.1016/j.dss.2015.04.012

[3] Grover V., Adderley R., Bramer M. (2007) Review of Current Crime Prediction Techniques. In: Ellis R., Allen T., Tuson A. (eds) Applications and Innovations in Intelligent Systems XIV. SGAI 2006. Springer, London. https://doi.org/10.1007/978-1-84628-666-7_19

[4] A. Sarguru, A. Krishnan, A.C. Shantha Sheela, Predicitve analysis of crime data using deep learning, Int. J. Pure Appl. Math. 118 (20) (2018) 4023– 4031.

[5] M. Feng *et al*., "Big Data Analytics and Mining for Effective Visualization and Trends Forecasting of Crime Data," in *IEEE Access*, vol. 7, pp. 106111-106123, 2019, doi: 10.1109/ACCESS.2019.2930410.

[6] N. Esquivel, O. Nicolis, B. Peralta and J. Mateu, Spatio-Temporal Prediction of Baltimore Crime Events Using CLSTM Neural Networks, in *IEEE Access*, vol. 8, pp. 209101-209112, 2020, doi: 10.1109/ACCESS.2020.3036715.

[7] Ulises M. Ramirez-Alcocer, Edgar Tello-Leal, Jonathan A. Mata-Torres, "Predicting Incidents of Crime through LSTM Neural Networks in Smart City Domain", *SMART 2019 : The Eighth International Conference on Smart Cities, Systems, Devices and Technologies, 2019*

[8] Hicham Aitelbour, Soumia Ounacer, Yassine Elghomari, Houda Jihal, Mohamed Azzouazi, "A crime prediction model based on spatial and temporal data", Periodicals of Engineering and Natural Sciences, ISSN 2303-4521, Vol. 6, No. 2, December 2018, pp.360-364

[9] Anusha K, Dr M.K Jayanti Kannan, "Big Data Analytics for Forecasting of Crime Data", International Research Journal of Engineering and Technology (IRJET), Volume: 07 Issue: 06, June 2020

[10] P. Stalidis, T. Semertzidis, P. Daras**, "**Examining deep learning architectures for crime classification and prediction", CoRR (2018), pp. 1-12. abs/1812.00602

[11] A. A. Biswas and S. Basak, "Forecasting the Trends and Patterns of Crime in Bangladesh using Machine Learning Model," *2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT)*, Jaipur, India, 2019, pp. 114-118, doi: 10.1109/ICCT46177.2019.8969031.

[12] Butt, U.M.; Letchmunan, S.; Hassan, F.H.; Ali, M.; Baqir, A.; Sherazi, H.H.R. Spatio-Temporal Crime HotSpot Detection and Prediction: A Systematic Literature Review. IEEE Access 2020, 8, 166553–166574.

[13] Hossain, S.; Abtahee, A.; Kashem, I.; Hoque, M.M.; Sarker, I.H. Crime Prediction Using Spatio-Temporal Data. arXiv 2020, arXiv:2003.09322. preprint.

[14] Ye, X.; Duan, L.; Peng, Q. Spatiotemporal Prediction of Theft Risk with Deep Inception-Residual Networks. Smart Cities 2021, 4, 204–216. https://doi.org/10.3390/smartcities4010013

[15] Tapia-McClung, R. Exploring the Use of a Spatio-Temporal City Dashboard to Study Criminal Incidence: A Case Study for the Mexican State of Aguascalientes. Sustainability 2020, 12, 2199.

[16] Catlett, C.; Cesario, E.; Talia, D.; Vinci, A. Spatio-temporal crime predictions in smart cities: A data-driven approach and experiments. Pervasive and Mobile Computing 2019, 53, 62–74.

[17] Wortley, R.K.; Mazerolle, L.A. Environmental Criminology and Crime Analysis; Vol. 6, 2016; pp. 1–294.

[18] Yadav, R., & Kumari Sheoran, S. (2018). *Crime Prediction Using Auto Regression Techniques for Time Series Data. 2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE).* doi:10.1109/icraie.2018.8710407

[19] Safat, W., Asghar, S., & Gillani, S. A. (2021). *Empirical Analysis for Crime Prediction and Forecasting Using Machine Learning and Deep Learning Techniques. IEEE Access, 9, 70080–70094.* doi:10.1109/access.2021.3078117

[20] Agarwal, S., Yadav, L., & Thakur, M. K. (2018). *Crime Prediction Based on Statistical Models. 2018 Eleventh International Conference on Contemporary Computing (IC3).* doi:10.1109/ic3.2018.8530548

[21] Mary Shermila, A., Bellarmine, A. B., & Santiago, N. (2018). *Crime Data Analysis and Prediction of Perpetrator Identity Using Machine Learning Approach. 2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI).* doi:10.1109/icoei.2018.8553904

[22] Shukla, A., Katal, A., Raghuvanshi, S., & Sharma, S. (2021). Criminal Combat: Crime Analysis and Prediction Using Machine Learning. 2021 International Conference on Intelligent Technologies (CONIT). doi:10.1109/conit51480.2021.9498397

[23] S. K. Dash, I. Safro and R. S. Srinivasamurthy, "Spatio-temporal prediction of crimes using network analytic approach," *2018 IEEE International Conference on Big Data (Big Data)*, 2018, pp. 1912-1917, doi: 10.1109/BigData.2018.8622041.
[24] U. M. Butt *et al*., "Spatio-Temporal Crime Predictions by Leveraging Artificial Intelligence for Citizens Security in Smart Cities," in *IEEE Access*, vol. 9, pp. 47516-47529, 2021, doi: 10.1109/ACCESS.2021.3068306.

materialstoday: PROCEEDINGS

Volume 58, Part 1, 2022, Pages 310-315

A survey on crime analysis and prediction

Ashly Thomas, N.V. Sobhana

https://doi.org/10.1016/j.matpr.2022.02.170

Abstract

A lot of police forces around the world have adopted mechanisms that use statistical data to guide their decision-making which is coined as predictive policing. As urbanization is increasing day by day, it is highly demanding to keep an eye on the criminal activities of each region so as to reduce the occurrences of unwanted behaviours. Prediction of crimes can be done only using the analysis of the patterns of criminal activities using the past data available with the concerned