Literature review: Can machine learning crime prediction reduce crime in England and Wales?

This is an incomplete first draft literature review for formative feedback. Nothing is finalised, but some parts, in regular text *are written in a form similar to how I expect the final version to look for feedback.*

Comments of things "still to complete" are in italics to illustrate the current thinking for this literature review, again for feedback.

Abstract

To be completed at the end….

Crime in England and Wales

In 2015/16, the most recent official data available, costs as a consequence of crime in England and Wales was £40bn, whilst costs in response to crime, such as policing and the criminal justice system, was £13.5bn a total of 16.8m crimes. (Heeks et al, 2018). In that same year there were 4.5m crimes reported to police (Office for National Statistics, 2016). The latest figures show 6.7 million crimes reported to police in (Office for National Statistics, 2023). Whilst for some offences trends in police reported crime "do not provide reliable trends in crime" (Office for National Statistics, 2023: 6), the figures has been included to highlight a general upward trend simply to illustrate that there remains a high number of crimes and therefore a high cost to society. Given the cost constraints on public service, it is therefore asserted that reducing crime in a cost effective way would have a positive social and economic impact on society. The question posed is, "can machine learning help?"

Machine learning models for crime prediction

Machine learning includes unsupervised learning, where patterns in data are learnt and clustered into sets with similar characteristics, and supervised learning where input data are labelled and the characteristics of the data are leant in order to predict the label of a new unlabelled input (Russell & Norvig, 2021). Both unsupervised learning (clustering) and supervised learning (classification) algorithms will be explored for suitability for assisting law enforcement with crime prediction.

Crime is committed by a person or persons, in a place or places. So a machine learning model to predict crime could predict if a person is likely to commit a crime, or if a place is likely to be the location of a crime. Crime is often committed in the same location “because criminals tend to operate in their comfort zone” (Perry et al, 2013: 2) so this review will focus on the second scenario only; is it possible to use machine learning to predict where a crime will take place?

Unsupervised Learning – Clustering

Jain et al (2017) demonstrated clustering of crimes by type and location in New Delhi using K-means. The clusters show the density of a crime type within close proximity to each other. Whilst not a direct predictor of future crime, this could be a useful tool to help to direct police resources to reduce specific criminal activities at specific locations.

Supervised Learning – Classification

Safat et al (2021) tested eight classification algorithms on Chicago and Los Angeles crime datasets, using features common to both datasets, to predict the rates of different crimes in each district with each city. XGBoost performed best on both datasets, but the Chicago dataset gave the most consistent results across all eight algorithms, whilst the Los Angeles dataset gave considerably different results between the different algorithms. This suggests that more work would be needed to conclusively state which algorithm is best for this task. They went on to use Autoregressive Integrated Moving Average (ARIMA) to produce a time series forecasting model to predict crime levels over time for each city, and within district hot spots within each city. The application of such a model could be to assist with policing strategies such as funding and recruitment. However, the outputs looked suspiciously uniform from considerably more varied input data, because ARIMA is "good for short-term forecasting", but "not built for long-term forecasting" (Hayes, 2023)

Hybrid

KNN was used to classify crimes then K-Means, Agglomerative Clustering and DBSCAN compared for clustering crimes of the same category (Sivaranjani et al, 2016). DBSCAN performed best with precision 0.95, recall 0.91 and f-score 0.93. The output assigned each city to a cluster, with a high, medium or low value for each crime type. This is potentially useful for macro level policing policy such as distribution of funds and police specialists, but is not considered a useful crime prediction tool to assist with day-to-day policing because it lacks granularity.

Hajela et al (2020) reverses the approach by clustering first, using K-Means, to identify crime hot spots, and then classify each instance using Naive Bayes and Decision Tree classification algorithms to determine the prevalent crime type at that location. If the location belongs to a hot spot the prediction is high likelihood of the classified crime type happening there. Advantage over Sivaranjani et al (2016) is coordinates are used for the initial clusters so arbitrary city boundaries are not used, making clusters potentially more useful. However the classification of each location results in a single crime class, which might be useful to prioritise policing activities, but would potentially miss other significant predicted criminal activities.

An interesting idea to map crimes using four dimensions; longitude, latitude, time and type of crime was proposed by Stalidis et al (2021). A city is split into a grid and time is split into time windows, so each grid cell will plot crimes within a series of time windows. The grids, output as images, are passed through a CNN to create a feature vector for each grid cell/timespan instance, which are then fed into a LSTM to extract feature vectors. "LSTM is more robust to degradation of gradient than the standard RNN" (Noh, 2021: 10), hence using it in preference to a RNN for time series modelling. The model, named Spatial features First Then Temporal (SFTT), was tested on open datasets of crime for five different cities in North America and compared with state-of-the-art algorithms such as DT, ND, SVM and KNN. SFTT outperformed the other algorithms with consistently excellent results in predicting emerging hotspots of crime.

Additional papers to describe, and once done will look to re-sequence in order to give more of a narative:

- Cichosz (2020):
- Jenga et al (2023):
- Ahishakiye et al (2017):
- Babakura et al (2014):
- Zaidi et al (2020):
- Shamsuddin et al (2017):

- Das & Das (2019):

Data challenges

Might remove this section…

Comment on the availability, or lack thereof, of data for the models in the literature leading to potentially unsafe conclusions.

Ethical considerations

Will describe the relative ethical safety of using crime location data over crime perpetrator data, since the former contains no personal data. However the results of the machine learning models could be used to influence policing which could have implications for the public, such as unnecessary and intimidating police presence in "high risk" areas, or failure to respond to criminal activities in "low risk" areas because resources were diverted elsewhere. It would therefore be important for public acceptance to be able to explain how the models made their recommendations.

Rather than try to explain black box machine learning models such as SVM and ANN, it is preferable to use interpretable models such as DT instead, especially when the results are similar (Rudin, 2019). This provides a much more transparent explanation of the output of the model, and transparency is a key enabler for ethical machine learning models (Walmsley, 2020).

Conclusions

Compare the various results and suggest an optimal direction balancing data availability with ethical considerations and real impact on policing e.g. insights with enough detail to be actionable would be better than more accurate models that produce outputs that are too high level to inform day-to-day policing.

References

Ahishakiye, E., Taremwa, D., Omulo, E.O. & Niyonzima, I. (2017) Crime prediction using decision tree (J48) classification algorithm. *International Journal of Computer and Information Technology*, 6(3): 188-195.

Babakura, A., Sulaiman, M.N. & Yusuf, M.A. (2014) ‘Improved method of classification algorithms for crime prediction’. *International Symposium on Biometrics and Security Technologies.* Kuala Lumpur, Malaysia, 26-27 August. IEEE. 250-255.

Cichosz, P. (2020) Urban crime risk prediction using point of interest data. *ISPRS International Journal of Geo-Information*, 9(7): 459. https://doi.org/10.3390/ijgi9070459

Das, P. & Das, A.K. (2019). ‘Application of Classification Techniques for Prediction and Analysis of Crime in India’. in: Behera, H., Nayak, J., Naik, B. & Abraham, A. (eds) *Computational Intelligence in Data Mining. Advances in Intelligent Systems and Computing*, vol 711. Singapore, Springer. 191-201.

Hajela, G., Chawla, M. & Rasool, A. (2020) A clustering based hotspot identification approach for crime prediction. *Procedia Computer Science*, 167: 1462-1470. https://doi.org/10.1016/j.procs.2020.03.357

Hayes, A. (2023) Autoregressive Integrated Moving Average (ARIMA) Prediction Model. Available from: https://www.investopedia.com/terms/a/autoregressive-integrated-moving-average-arima.asp [Accessed 01 December 2023].

Iqbal, R., Murad, M.A.A., Mustapha, A., Panahy, P.H.S. & Khanahmadliravi, N. (2013) An experimental study of classification algorithms for crime prediction. *Indian Journal of Science and Technology*, 6(3): 4219-4225.

Jain, V., Sharma, Y., Bhatia, A. & Arora, V. (2017) Crime prediction using K-means algorithm. *GRD Journals-Global Research and Development Journal for Engineering*, 2(5): 206-209.

Jenga, K., Catal, C. & Kar, G. (2023) Machine learning in crime prediction. *Journal of Ambient Intelligence and Humanized Computing*, 14(3): 2887-2913.

Keeks, M., Reed, S., Tafsiri, M., & Prince, S. (2018) *The economic and social costs of crime: Second edition*. London: Home Office

Khan, M., Ali, A. and Alharbi, Y., 2022. Predicting and preventing crime: a crime prediction model using san francisco crime data by classification techniques. *Complexity*, 2022: 1-13. https://doi.org/10.1155/2022/4830411

Noh, S.H. (2021) Analysis of gradient vanishing of RNNs and performance comparison. *Information*, 12(11): 442.

Office for National Statistics (2023) *Crime in England and Wales: year ending June 2023*. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingjune2023/pdf [Accessed 28 November 2023].

Office for National Statistics (2016) *Crime in England and Wales: year ending Mar 2016*. Available from:

https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingmar2016/pdf [Accessed 28 November 2023].

Perry, W.L., McInnis, B., Price, C.C., Smith, S. & Hollywood, J.S. (2013) Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations. Santa Monica: RAND Corporation.

Rudin, C. (2019) Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature machine intelligence*, 1(5): 206-215.

Russell, S. & Norvig, P. (2021) *Artificial Intelligence: A Modern Approach*. 4th ed. Harlow: Pearson Education Limited.

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. https://doi.org/10.1109/ACCESS.2021.3078117

Shamsuddin, N.H.M., Ali, N.A. & Alwee, R. (2017) ‘An overview on crime prediction methods’. *6th ICT International Student Project Conference*. Johor, Malaysia. 23-24 May. IEEE. 1-5.

Sivaranjani, S., Sivakumari, S. & Aasha, M. (2016) ‘Crime prediction and forecasting in Tamilnadu using clustering approaches’. *2016 International Conference on Emerging Technological Trends.* Kollam, India, 21-22 October. IEEE. 1-6.

Stalidis, P., Semertzidis, T. & Daras, P. (2021) Examining deep learning architectures for crime classification and prediction. *Forecasting*, 3(4): 741-762. https://doi.org/10.3390/forecast3040046

Walmsley, J. (2021) Artificial intelligence and the value of transparency. *AI & Society*, 36(2): 585-595.

Zaidi, N.A.S., Mustapha, A., Mostafa, S.A. & Razali, M.N. (2020) ‘A classification approach for crime prediction’. *Applied Computing to Support Industry: Innovation and Technology.* Ramadi, Iraq,15–16 September. Switzerland: Springer International Publishing. 68-78.

Not yet used – will be deleted unless used:

Ali, S., Alvi, S.A. & Rehman, A.U. (2019) ‘The Usual Suspects: Machine Learning Based Predictive Policing for Criminal Identification’, *13th International Conference on Open Source Systems and Technologies*. Lahore, Pakistan, 17-19 December. IEEE. 1-6.

Zamri, N.F.M., Tahir, N.M., Ali, M.S.A.M., Ashar, N.D.K. & Al-Misreb, A.A. (2021 Mini-review of street crime prediction and classification methods. *J. Kejuruter*, 33: 391-401.