Crime Prediction Model using Deep Neural Networks

Soon Ae Chun
ISI
College of Staten Island City
University of New York, NY, USA
soon.chun@csi.cuny.edu

Venkata Avinash Paturu
I-DSLA & CIMIC Rutgers
University
Newark, NJ, USA

Shengcheng Yuan
I-DSLA Rutgers University
Newark, NJ, USA

Rohit Pathak
I-DSLA & CIMIC Rutgers
University
Newark, NJ, USA

Vijay Atluri
CIMIC & MSIS Department
Rutgers University
Newark, NJ, USA

Nabil R. Adam
I-DSLA & MSIS Department
Rutgers University
Newark, NJ, USA,
adam@adam.rutgers.edu

ABSTRACT

This project investigates the feasibility of using machine learning techniques, specifically neural networks, to make prediction on criminal behavior based on the history of the arrest bookings. The experiment will have to handle imbalanced data frequencies. To combat the challenge, data augmentation and weighted loss function is being developed to extract information from the minority classes. For this project, we have focused on how neural networks can be advantageous in classification of crime prediction. The specific kind of neural network that has been used in the project is a deep fully connected neural network. Fully connected neural networks are suitable for problems where domain knowledge is limited and many to many relations between features are important. As this report shows, machine learning techniques could definitely be of use for classification of criminal behavior, and we recommend exploring the discussed data augmentation and modeling methods more thoroughly to improve on the results and find new patterns.

CCS CONCEPTS

• Applied computing → E-government • Computing methodologies → Neural networks

KEYWORDS

Crime prediction, Machine Learning, Deep Learning, Neural Networks, Crime Trajectory Analysis

dg.o 2019, June 18-20, 2019, Dubai, <, United Arab Emirates

ACM ISBN 978-1-4503-7204-6/19/06.
https://doi.org/10.1145/3325112.3328221

ACM Reference format:

Soon Ae Chun, Venkata Avinash Paturu, Shengcheng Yuan, Rohit Pathak, Vijay Atluri and Nabil R. Adam. 2019. **Crime Prediction Model using Deep Neural Networks. In** ***Proceedings of dg.o 2019: 20th Annual International Conference on Digital Government Research (dg.o 2019), June 18-20, 2019, Dubai, <, United Arab Emirates.*** ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3325112.3328221

1 Introduction

Predicting and forecasting crimes have two important goals. First, to take preventive measures and intelligently allocate law enforcement resources. Second, to assist the criminal justice system to make decisions about individuals. Which will be used for intervention, reform and rehabilitation of criminals while preventing crime from happening the problem we deal with in this work is of great importance for law enforcement agencies. Currently, law enforcement agencies statistically model spatio-temporal patterns of criminal incidents. These statistical models are used to study causality of crimes and predict future criminal incidents [2, 3, 4].

Iqbal et al. [2] used dataset prepared using real data from socio-economic data from US Census, law enforcement data of US LEMAS survey, and crime data from the FBI UCR to predict the crime category for different states of USA. They used Naïve Bayes and Decision Trees for prediction of a multiclass classification with 3 labels Low, Medium and High, achieving an accuracy of 83.95%. Wang et al. [4] apply Twitter-specific linguistic analysis and statistical topic modeling to automatically identify discussion topics across a major city and improves the crime prediction performance over a kernel density estimation. Bogomolov et al. [3] proposed Breiman's Random forests classifier yielding 70% accuracy by combining crime event data of the London borough profile and smart steps data from mobile network activity to predict weather an area in the city will be a crime hotspot or not. Broken windows theory states that visible signs of crime, anti-social behavior, and civil disorder create an urban environment that encourages further crime and disorder, including serious crimes. Kang et al. [1] focused on developing a deep learning model to learn from such extensive multi model data. They employed Alex net to learn from crime location images collected from google street view and fused them with demographics, weather and crime incidence information using deep neural networks for improving crime hotspots prediction.

We apply the machine learning method DNN (Deep Neural Networks) on the individual's criminal charge history to accurately predict the crime and its type at the level of the individual, while previous works mostly focus on the level of a population. Moreover, we study the trend of the criminal charge level developed over the time of every single person, and in a varying number of years. More precisely our goal is to predict (a) whether a person will commit a crime in the near future (n years ahead); (b) which level L of the charge may be (L ∈ 1,2,3, seriousness indicator of a crime); and (c) whether there is any trend in crime progression and correlated crimes.

Unlike previous approaches, we focus on the rehabilitation and intervention in troubled cities to support recurring offenders in leading a path away from crime. It is planned to support the State Department of Police or judicial system wherein they can examine a criminal history and focus on criminals who may be predicted to be a threat to society. The prediction model can be used for intervention, reform and rehabilitation of criminals while preventing crime from happening, i.e. encountering the criminals at an early stage to avoid them to commit more severe crimes.

2 Data Set

The problem we are trying to solve is to accurately predict the crime reoccurrences based on the personal criminal charge history. Data collection is critical for the accurate prediction of crime reoccurrences. The dataset at our disposal is the booking history of criminals from 1997 to 2017. The raw data has 5 features per record with a unique PersonID to link multiple records together (NCIC Crime Code, gender, age, race, booking date). There are 16,841 unique people with 63,133 records of arrest history. There are 42 unique crime types with 3 levels of seriousness. We plan to create a model which predicts the possible crime level in the next 5 years. In order to create the labels for supervised learning, we used 5-year windows to look ahead for each crime record and pick the most serious level (e.g. level 1, 2, 3 or 0 for NoCrime) as its label for the predicted crime. The problem with this dataset is that each unique individual has his own timeline of sequence of crimes. Some people have days between 2 successive arrests and some have years. We need to find as effective method to create feature vectors for analysis while keeping the information intact.

3 Crime Trajectory Analyses

The ordinary correlation analysis does not reveal the directionality of the crime progression. To find out the crime progression which may shed some lights on the prediction of crimes that occur, we used a graph-based progression analyses, one with pairwise progression between two crimes (Figure 1), and the other the temporal progression from each crime type (Figure 2).

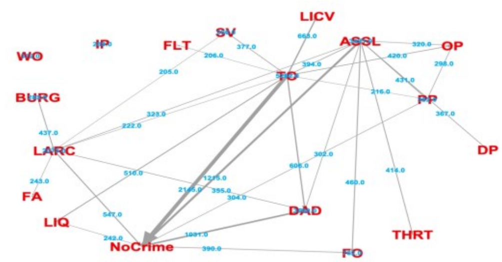

Figure 1: Pairwise Crime Progression Analysis

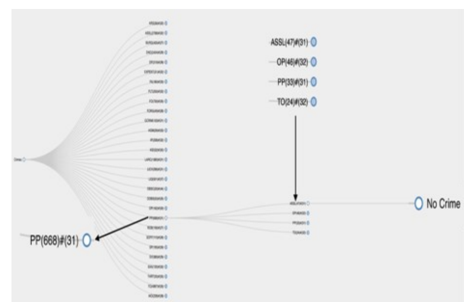

Figure 2: Crime trajectory over time

These tree-based crime transition analyses help us to understand how offenders move from one crime to another over time and the average of criminals involved in the progression. In Figure 2 , each crime node shows the frequency of crime and average age of the people who committed the crime. Figure 2 shows that there are 668/16841 people had first crime as *Public Peace disturbance(PP) with average age of 31*, out of these 668 people 47 moved to *Assault* as second crime. None of the 47 committed any other crime further.

4 Crime Trajectory Analyses

In order for an individual to commit a crime within 5 years, given a set of booking history, we developed a Deep Neural Network to build the prediction model. We have a crime booking data ranging from year 1997 to 2017. The dataset is highly imbalanced 42.2% of the population has just one crime in their record, 60% of the population has less than or equal to 2 crimes. The data is not explicitly labeled whether someone is going to commit a crime. We have used a window method to create labels for each individual data record. Given a year with a crime x, the window of 5 year look ahead is created and used the worst crime committed within that 5 year window is considered as the label.

For instance, for a record at 2012, we look at the window of next 5 years (from 2013 to 2017) and select the worst one as its label, i.e. either No_crime or Level1, Level 2 or Level3. If there is a level 1 or level 2 crime within the five year window after 2012, then it is labeled as Level 2, meaning that he is likely to commit the Level 2 crime within 5 years. The crime history data up to 2012 is the input training data x, and its label from the next 5 years as its label. Thus, for each booking record, we generated its annual crime history data (h1, h2, h3...) from 1997 up to 2012, and its 5 year prediction labels (y1, y2, y3, ...) for each history record. We used an expanding window with a stride length of 1 year starting from first booking year to last year for 5-year future label. i.e 2012 of each person to create aggregate feature vectors at every year along with label for next 5 years.

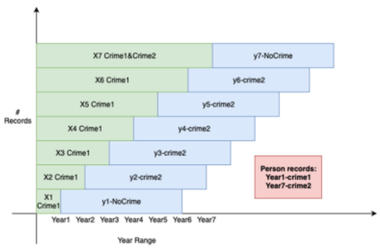

Figure 3: 5 year look ahead window for labeling

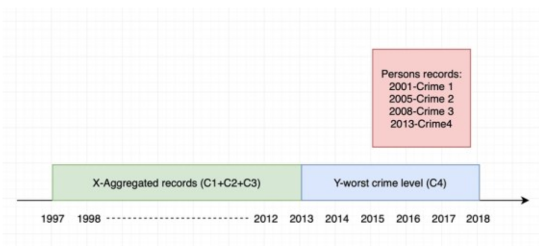

Figure 4: Last year fixed 5 year as labeling

Figure 3 shows the window labeling mechanism with two crime records in Year1 and second in Year7. We are going to start at Year1 and take all the features as our X1-in this case its crime1 to create h1. Look ahead next five years and create label y1= NoCrime. Next, we move to Year2 and take the aggregated features of Year1 and Year2 as h2 and look ahead to Year3-Year7 and create label y2 = crime2. In the same manner we continue until year 7 to create pooled feature vectors at each year. Our experiments contrast its results with the prediction model based on the last 5 year only (2013-2017) as labels, as shown in Figure 4.

We used a Deep Neural Network, using fully connected convolution layers to build the prediction model for multi-label classification (crime levels 1, 2, 3 and No crime). The network was implemented with TensorFlow (TF), an open source API developed by Google mainly for Machine Learning and Deep Learning, with drop out layers and with L2 Regularization to avoid overfitting. The network used a loss function of cross entropy function, a learning rate of 0.005, gradient descent as optimizer and the softmax for multi-label output.

The experimental results show that the fixed window model (see Figure 4) achieved 99.7% accuracy to predict whether one will commit a crime or not given all the history, and 94% accuracy in predicting the level of crimes (1, 2 or 3) with F1 score 54.3%. In other other hand, the 5 year look-ahead window for each year with history provided 82% for predicting the crime or no crime, but 80% for crime level prediction with F1 score of 34%. These were also compared with the baseline SVM binary classification models to predict level 1, 2, 3 crimes with accuracy level of 92.5%, 91.5% and 97.6% with F1 scores 51%, 37%, and 87% respectively. The results using the fixed window model is better for crime levels 1 or 2, while the conventional svm model is better in predicting the large class crime level 3.

5 Conclusion

We presented a prediction learning model with Dynamic Window based data modeling to predict whether a specific person will commit a new crime in the future years within n window time. For multi-class crime predictions, we found the data pooling method of looking at all the possible historical years per person performs the best.

6 ACKNOWLEDGMENTS

This project is partially funded by NSF CNS 1624503 and CNS 1747728, and partially supported by the National Research Foundation of Korea Grant from the Korean Government (NRF-2017S1A3A2066084). The booking dataset was provided by Kovid Inc.

REFERENCES

[1] Kang, H. and Kang, H. (2017). Prediction of crime occurrence from multi-modal data using deep learning. *PLOS ONE*, 12(4), p.e0176244.
[2] Iqbal R, Azmi Murad MA, Mustapha A, Hassany Shariat Panahy P, Khanahmadliravi N. (2013) An experimental study of classification algorithms for crime prediction. Indian Jour- nal of Science and Technology, 6(3):4219–25.
[3] Bogomolov A, Lepri B, Staiano J, Oliver N, Pianesi F, Pentland A (2014) Once upon a crime: towards crime prediction from demographics and mobile data. In: Proceedings of the 16th international conference on multimodal interaction. ACM, New York, pp 427-434
[4] Wang X, Gerber MS., and Brown, DE.(2012) Auto- matic crime prediction using events extracted from twitter posts. In Social Computing, Behavioral-Cultural Modeling and Prediction, pages 231–238. Springer