Crime Analysis Through Machine Learning

Suhong Kim (✉), Param Joshi, Parminder Singh Kalsi, and Pooya Taheri
Fraser International College, Simon Fraser University
British Columbia, Canada
suhongk@sfu.ca

***Abstract*—This paper investigates machine-learning-based crime prediction. In this work, Vancouver crime data for the last 15 years is analyzed using two different data-processing approaches. Machine-Learning predictive models, K-nearest-neighbour and boosted decision tree, are implemented and a crime prediction accuracy between 39% to 44% is obtained when predicting crime in Vancouver.**

Keywords—crime prediction, machine learning, data analysis.

I. Introduction

Crime is a socio-economical problem affecting life quality and economic growth [1]. The specifics of how crime is conducted changes depending on the type of society and community. Previous researches in crime prediction have found that factors like education, poverty, employment, and climate affect the crime rate [2]. Vancouver is one of the most populous, ethnically-diverse, and multi-cultural urban cities in Canada. The overall crime rate in Vancouver dropped 1.5% in 2017, but high vehicle break-ins and theft is still an issue [3]. Recently, the Vancouver Police Department (VPD) introduced a crime-predictive model to predict crimes related to property break-ins and, once implemented, the city of Vancouver witnessed a 27% drop in residential break-ins [4]. Crime prediction is a law-enforcement technique that uses data and statistical analysis for the identification of crimes most likely to occur [5]. This field has been subject to continued research in many parts of the world.

Machine learning is the science of having computers make decisions without human intervention. Recently, machine learning has been applied in self-driving cars, speech recognition, web search, and an improved understanding of the human genome. It has also made predicting crime based on referenced data feasible. Classification is a supervised prediction technique which allows for nominal class labels. Classification has been used in many domains including weather forecasting, medical care, finances and banking, homeland security, and business intelligence [6].

Machine-learning-based crime analysis usually involves data collection, classification, pattern identification, prediction, and visualization. Traditional data mining techniques - association analysis, classification and prediction, cluster analysis, and outlier analysis - identify patterns in structured data while newer techniques identify patterns from both structured and unstructured data [7]. The primary objective of this work is to create a prediction model that can accurately predict crime. In our research, two classification algorithms, K-Nearest Neighbour (KNN) and boosted decision tree, were implemented to analyze the VPD crime dataset compiled between 2003 and 2018 with more than 560,000 records. The dataset was processed using two different approaches:

- In the first approach, each neighbourhood and crime category was given a unique number when a certain crime happens in a certain neighbourhood.
- In the second approach, the neighbourhood and the day of the week during which the crime was committed were given a binary number and marked as 1 when the crime happened on that day in that neighbourhood, and 0 otherwise.

The rest of this paper is organized as follows: Section II of this paper provides a brief survey of previous work done on the topic of machine-learning-based crime prediction. In sections III and IV, the data-analysis and machine-learning methodology used in this work are explained, and the results are presented and compared. Conclusions are presented in section V.

II. Related Work

Various researchers have addressed the problems regarding crime control and have proposed different crime-prediction algorithms. The accuracy of prediction depends on the attributes selected and the dataset used as a reference.

In [1], human behavioural data derived from mobile network activity combined with demographic information using real crime data were used to predict crime hotspots in London, UK. In [6], a comparison between two classification algorithms, Decision Tree and Naïve Bayesian, was performed using WEKA, an open-source data mining software, and 10-fold cross-validation. The socio-economic, law-enforcement, and crime datasets for this study were compiled from the 1990 US Census, the 1990 US LEMAS survey, and the 1995 FBI UCR, respectively. The road accident patterns in Ethiopia was studied in [8] considering various circumstantial factors like the driver, weather, car, and road conditions. Three different classification algorithms, KNN, Naïve Bayesian, and Decision tree were used on a dataset of 18,288 accidents. The prediction accuracy for all three algorithms was between 79% to 81%.

A major challenge regarding crime prediction is analyzing large crime datasets accurately and efficiently. Data mining is utilized to find hidden patterns in large crime datasets quickly and efficiently. The increased efficiency and reduced errors in crime data-mining techniques increase the accuracy of crime prediction. A general data-mining framework was developed in [7] based on the experience of the Coplink project, conducted at the University of Arizona. Most research in crime prediction is focused on identifying crime hotspots, which refers to the areas in which the crime rates are above the average level. In [9], authors provided a comparative analysis of Kernel Density Estimation (KDE) and Risk Terrain Modeling (RTM) algorithms for creating hotspot maps and proposed area-specific predictive models using sparse data. In [10], a spatial-temporal model using histogram-based statistical methods, Linear Discriminant Analysis (LDA), and KNN were adopted for crime hotspot prediction. In [11], a crime incidence-scanning algorithm was applied to train Artificial Neural Network (ANN) enhanced by the Gamma test to predict the crime hotspots in Bangladesh. A data-driven machine-learning algorithm based on broken-window theory, spatial analysis, and visualization techniques was used in [12] to analyze drug-related crime data in Taiwan and predict emerging hotspots.

In [13], authors applied reverse-geocoding technique and a density-based clustering algorithm to build a machine-learning model for crime prediction using Open Street Map (OSM) and geospatial data for different categories of crime in the province of Nova Scotia (NS), Canada. A feature-level data-fusion method based on a Deep Neural Network (DNN) trained by spatial-, temporal-, environmental-, and joint-feature representation layers for predicting crimes in the City of Chicago was proposed in [14]. Several crime-prediction methods were reviewed in [15], and Knowledge Discovery in Databases (KDD) techniques, which combine statistical modelling, machine learning, database storage, and AI technologies, was suggested as an effective tool for crime prediction. In [16], a transfer-learning framework that captures temporal-spatial patterns was proposed for leveraging cross-domain urban datasets, meteorological data, points of interests, human mobility data, and complaint data. A fully-probabilistic algorithm based on Bayesian approach was applied in [17] to model the dependency between the offence data and environmental factors such as the demographic characteristics and the spatial location in the state of New South Wales (NSW), Australia. WEKA was used in [18] to conduct a comparative study for measuring the accuracy and effectiveness of linear regression, additive regression, and decision stump algorithms for predicting the crime in the state of Mississippi. In [19], authors presented a survey paper on crime data mining reviewing ANN, decision tree, rule induction, nearest-neighbor method, and genetic algorithm.

An approach based on Auto-Regressive Integrated Moving Average model (ARIMA) was utilized in [20] to design a reliable predictive model for forecasting crime trends in urban areas. In [21], authors proposed a probabilistic model of spatial behavior for known offenders based on a random-walk-based approach to model offender activity in the Metro Vancouver area. The random forest algorithm was used in [22] to quantify the role of urban indicators for crime prediction in Brazil. In [23], prospective method, Dempster-Shafer theory of evidence, and the multi-kernel method were used to develop a crime-prediction solution for Chilean large cities. In [24], three algorithms, KNN, Parzen windows, and Neural Networks, were developed, tested, and compared for predicting the crimes in the city of San Francisco. In [25], Gradient Boosting Machine (GBM) technique was applied in a machine-learning prediction model to find hidden links in criminal networks and the weighted page-rank method was used as an effective strategy to weaken and destroy such networks.

Based on the literature, in this research, the classification algorithms KNN and boosted decision tree were used to analyze the VPD crime dataset.

III. Data Analysis

A. Data Source

The original datasets were obtained from the open data catalog of the city of Vancouver. There are two datasets used for this project: crime and neighbourhood. The crime dataset has been collected by the VPD since 2003 and is updated every Sunday morning. It provides information on the type of crime committed and the time and location of the offence. The neighbourhood dataset contains the boundaries for the city's 22 local areas in the Geographic Information System (GIS). In this project, the crime dataset is used for data analysis and the neighbourhood dataset is used for drawing maps.

B. Preprocessing

The original dataset needs to be preprocessed to fill the empty cells, delete unnecessary columns, and add several relevant features. Fig. 1 shows the original and preprocessed datasets.

TYPE	YEAR	MONTH	DAY	HOUR	MINUTE	HUNDRED_BLOCK	NEIGHBOURHOOD	X	Y
Other Theft	2003	8	8	13	58	10XX ROBSON ST	West End	490998.3	5459018
Theft from Vehicle	2003	8	14	9	30	12XX CHESTNUT ST	Kitsilano	489368.8	5458066
Mischief	2003	4	4	19	15	16XX COMMERCIAL DR	Grandview-Woodlanc	494928.6	5457524
Theft from Vehicle	2003	6	15	22	0	10XX E BROADWAY AVE	Mount Pleasant	494071.3	5456637
Offence Against a Person	2003	10	12			OFFSET TO PROTECT PRIVACY		0	0
Offence Against a Person	2003	10	26			OFFSET TO PROTECT PRIVACY		0	0
Mischief	2003	4	6	23	0	32XX W 39TH AVE	Dunbar-Southlands	487196.8	5453771
Break and Enter Residenti	2003	1	28	18	30	18XX VENABLES ST	Grandview-Woodlanc	495047.8	5458217

(a)

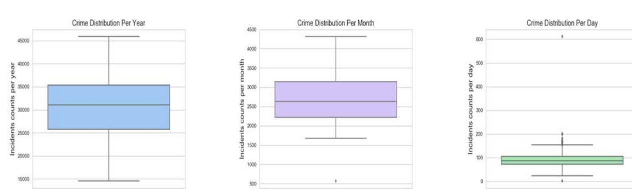
TYPE	YEAR	MONTH	DAY	HOUR	HUNDRED_BLOCK	NEIGHBOURHOOD	X	Y	DATE	WEE
other theft	2003	8	8	13	10XX ROBSON ST	west end	490998.3	5459018	2003-08-08	Frid
theft from vehicle	2003	8	14	9	12XX CHESTNUT ST	kitsilano	489368.8	5458066	2003-08-14	Thu
mischief	2003	4	4	19	16XX COMMERCIAL DR	grandview-woodlai	494928.6	5457524	2003-04-04	Frid
theft from vehicle	2003	6	15	22	10XX E BROADWAY AVE	mount pleasant	494071.3	5456637	2003-06-15	Sun
offence against a pe	2003	10	12	0	OFFSET TO PROTECT PRI\	n/a	0	0	2003-10-12	Sun
offence against a pe	2003	10	26	0	OFFSET TO PROTECT PRI\	n/a	0	0	2003-10-26	Sun
mischief	2003	4	6	23	32XX W 39TH AVE	dunbar-southlands	487196.8	5453771	2003-04-06	Sun
break and enter res	2003	1	28	18	18XX VENABLES ST	grandview-woodlai	495047.8	5458217	2003-01-28	Tue

(b)

Fig. 1. The snapshot of the (a) original and (b) preprocessed datasets.

C. Statistical Analysis

The distribution of the crime dataset described in Fig. 2 is based on year, month, and day. In Vancouver, the average number of crime incidents is around 31624 per year, 2720 per month, and 90 per day. The dataset tends to show a normal distribution as the time intervals lengthen. However, the graph of each day has an abnormal max value of 650 incidents, which is suspected as an outlier - and turns out to indicate the Stanley-Cup riot on June 15, 2011.



(a) (b) (c)

Fig. 2. Distribution of crimes per (a) year, (b) month, (c) day

D. *Trend Analysis*

According to Fig. 3, the overall trend shows that the average number of crimes per month decreased from 2003 to 2013 but increased in 2016, and again fell slightly to about 3000 incidents per year in 2018.

According to the time-featured heat-map graphs, shown in Fig. 4, the summer season and the middle of each month are the most dangerous. In addition, there are more crimes on Fridays, Saturdays, and during the evenings. Since all the empty data cells were filled with zero, the heat map showed the highest numbers near zero hours, which should be ignored.

The greatest number of incidents was recorded in the category of theft from vehicles, followed by mischief. However, theft from vehicles has significantly declined in recent years, but other incidents of theft have increased. Fig. 5 shows the number and trends of each crime.

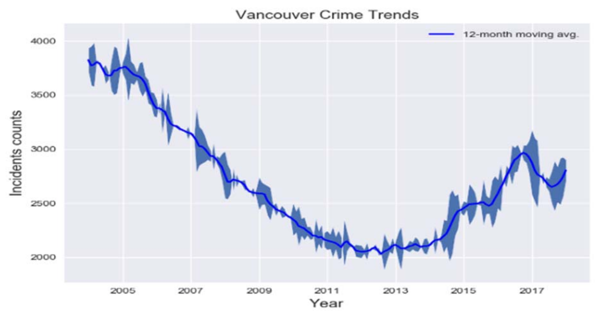
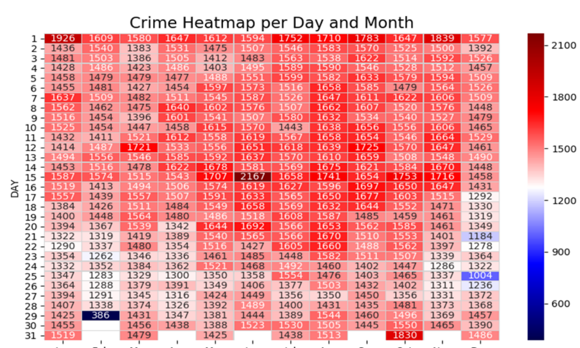
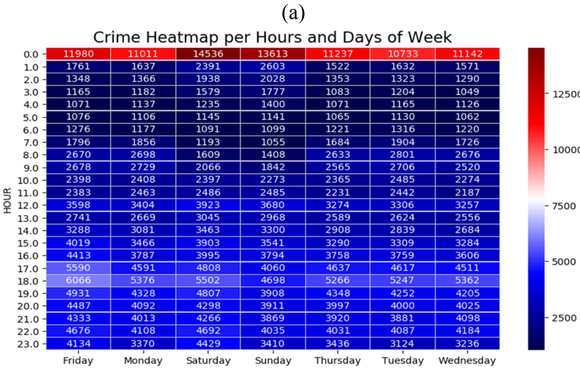


Fig. 3. Moving average of crimes per month

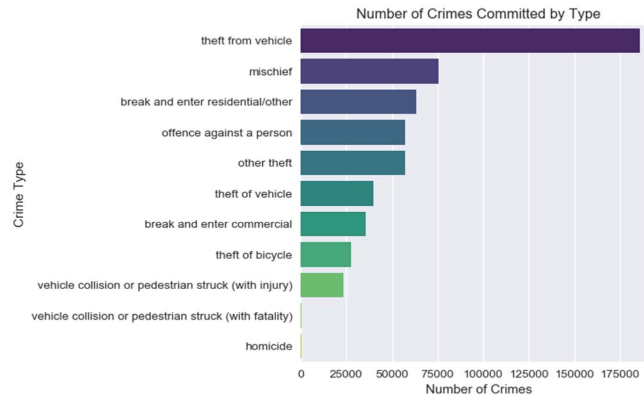


(a)

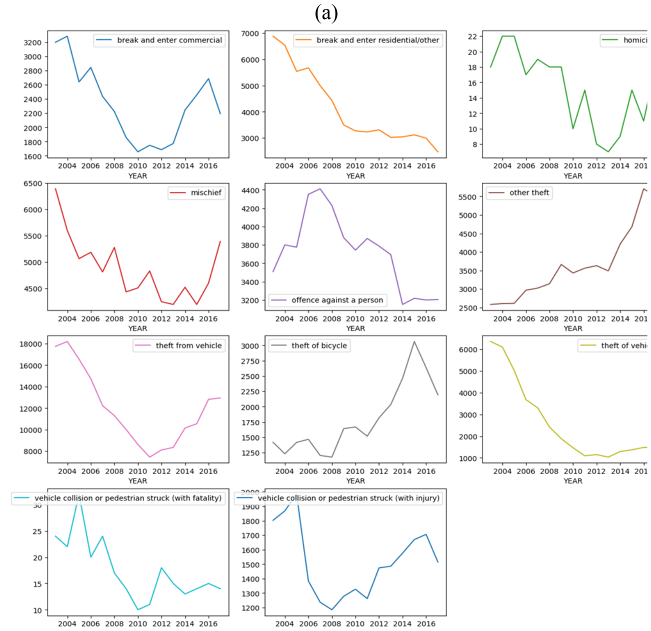


(b)

Fig. 4. The crime heatmap (a) per days and month, and (b) per hours and days of week



(a)



(b)

Fig. 5. (a) Number and (b) trend of crimes committed by type

E. *Geographical Analysis*

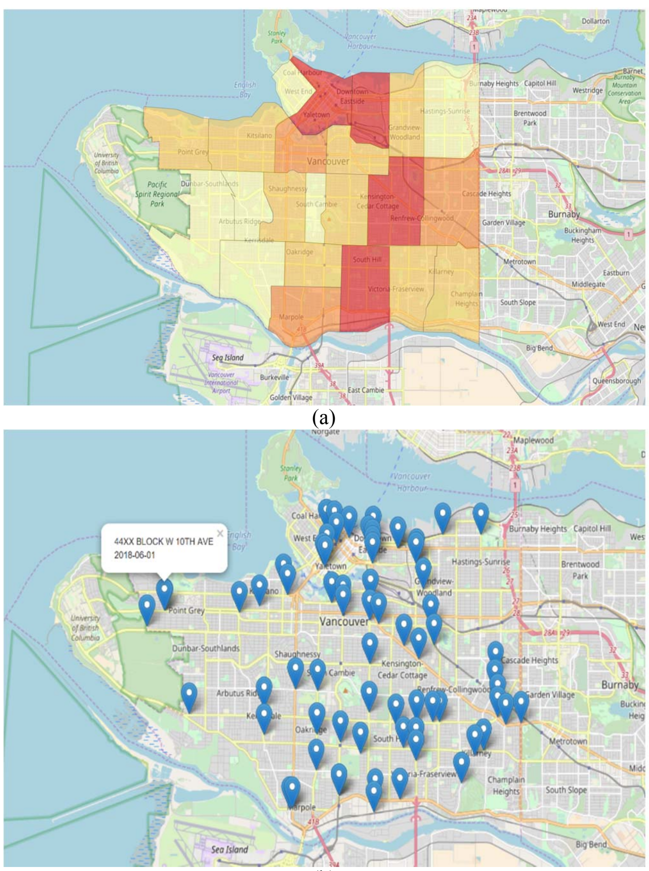
There are various techniques to map hotspots, but among them, the choropleth mapping is widely used to describe the geographic information of crime incidents [26]. Choropleth map represents the proportion of statistical measurements or its density with shaded colors. This makes it easy to recognize where crime incidents are more condensed, which gives insights into criminal behavior. Geographic Information System (GIS) has been used as a powerful analytical tool for crime mapping. It shows the locations of crime series with various geographic information on one map, which helps police officers to make decisions for operational and tactical purposes [27].

As the first step towards geographical analysis, the neighborhood boundary dataset was transformed from the Universal Transverse Mercator (UTM) into World Geodetic System 1984 (WGS84), known as latitude and longitude. Python was used to draw the map as it has various supportive libraries for visualizing geographic data such as PySal, GeoPandas, Folium, and Shapely.

To represent the crime hotspots on Vancouver's city map, the crime incidents were counted for each neighbourhood. Table 1 describes the top 10 crime-dense neighbourhoods within a 30-day time frame, linked to the choropleth map. Fig. 6 represents the hotspot map and point clusters of incidents occurred in a 30-day period in the city of Vancouver.

Table 1. Top-10 crime-dense neighbourhoods

Map ID	Name	Density (per square miles)
CBD	Downtown	3.938709
SUN	Sunset	3.150967
KC	Kensington-Cedar Cottage	2.363225
STR	Strathcona	2.363225
RC	Renfrew-Collingwood	1.969355
MARP	Marpole	1.575484
FAIR	Fairview	1.575484
MP	Mount Pleasant	1.181613
KITS	Kitsilano	1.181613
OAK	Oakridge	1.181613



(a)

(b)

Fig. 6. The city of Vancouver's (a) hotspot map, and (b) incident point clusters in a 30-day time

IV. Machine Learning

Machine learning is the part of artificial intelligence which deals with statistical methods and gives computers the ability to learn from past experiences [28]. Machine learning can be divided into categories such as supervised, unsupervised, and reinforcement learning. This study uses supervised learning due to the nature of required input data and output targets.

Supervised learning can be categorized into classification and regression. Classification is predicting a discrete class label, while regression is the task of predicting a continuous quantity. This work attempts to predict the types of crime in a particular location. Therefore, the objective of this study is the classification of crime. There are many algorithms that can be used for the classification such as K-Nearest Neighbour (KNN), Support Vector Machine (SVM), Naïve Bayesian, Decision Tree, and Ensemble Methods. Every algorithm has its own advantage and disadvantage in terms of complexity, accuracy, and training time and can provide different results from a single dataset.

In this work, KNN and decision-tree algorithms were used to train our model. KNN is one of the simplest classification algorithms. It assigns a sample z to class A if the most values nearer to z are from class A, otherwise, it assigns the sample to class B. In KNN, the following formula is used to calculate the probability of the test sample belonging to category C_i:

$$P(x \in C_i) = \frac{\sum_{j \in C_i} {n_j}/{d_j}}{\sum_{l=1}^{k} {n_l}/{d_l}} \tag{1}$$

where n_j and n_l are the number of elements represented by each training dataset sample, and d_j and d_l are the distance between the test and the corresponding training sample calculated based on Euclidean norm [24]. KNN stores all available objects and classifies new objects based on the similarity measure by searching for the nearest neighbor of the input values [29].

On the other hand, decision trees can deal better with large datasets that have many layers with different nodes [18]. Decision-tree classifiers can provide a better balance of flexibility and accuracy while limiting the number of possible decision points. Decision-tree-classification model forms a tree structure from the dataset by dividing the dataset into smaller pieces. At each step in the algorithm, decision tree chooses a feature that best splits the data with the help of two functions: Gini impurity and information gain. Gini impurity measures the probability of classifying a random sample incorrectly:

$$I_G(p) = \sum_{i=1}^{k} p_i(1 - p_i) \tag{2}$$

Information gain helps to decide which feature to split next. Information gain can be calculated using entropy defined as:

$$H(T) = I_E = -\sum_{i=1}^{k} p_i \, log_2(p_i) \tag{3}$$

where p_i represents the percentage of each feature being present in the child node after a split [30].

Before applying the algorithms, the data must be altered into a usable form. We implemented two different approaches for data processing and compared the results.

A. Approach 1:

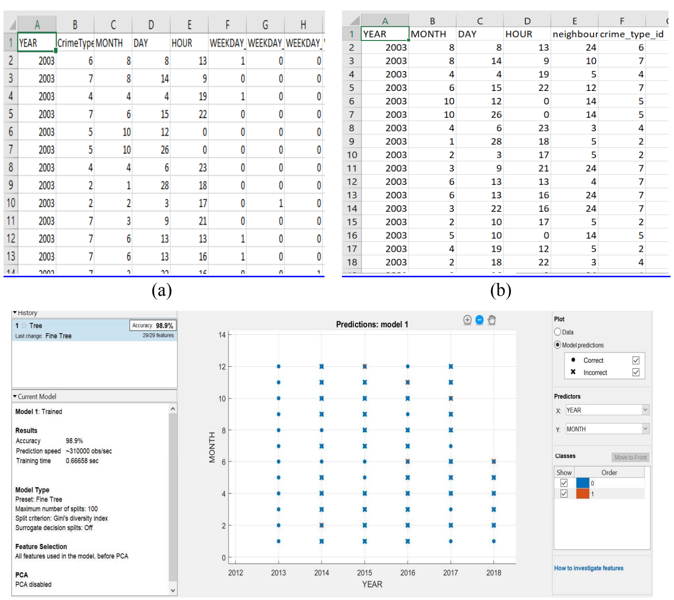
In the first approach, all categorical variables are converted into binary variables 0 and 1. All the neighbourhoods and days were made into features. Only the correct variable is assigned to "1" and all other variables are assigned to "0". Basically, all the variables with "0" are dummy variables. This gives the algorithm more variables to train and prevents the data from skewing to one side. Experiments with skewed data give 98.9% fake accuracy which is not reliable.

B. Approach 2:

In the second approach, categorical variables are converted into numerical variables with unique IDs. All the crime types and neighbourhoods have different IDs. For example, vehicle theft is given the crime type ID 10. This approach is inspired by the work done on Kaggle [31]-[32].

For both approaches, the same algorithms are applied with the same parameters and same validation processes. 5-fold cross-validation is applied to evaluate the classifier algorithm's validity. Crime type is chosen as the target to train the algorithm.

Cross-validation prevents the overfitting problem and ensures that the prediction model has a satisfactory performance on new unseen data. The processed dataset using approaches 1 and 2, and the skewed data from approach 1 are depicted in Fig. 7.



	A	B	C	D	E	F	G	H
1	YEAR	CrimeType	MONTH	DAY	HOUR	WEEKDAY_	WEEKDAY_	WEEKDAY_
2	2003	6	8	8	13	1	0	0
3	2003	7	8	14	9	0	0	0
4	2003	4	4	4	19	1	0	0
5	2003	7	6	15	22	0	0	0
6	2003	5	10	12	0	0	0	0
7	2003	5	10	26	0	0	0	0
8	2003	4	4	6	23	0	0	0
9	2003	2	1	28	18	0	0	0
10	2003	2	2	3	17	0	1	0
11	2003	7	3	9	21	0	0	0
12	2003	7	6	13	13	1	0	0
13	2003	7	6	13	16	1	0	0

(a)

	A	B	C	D	E	F
1	YEAR	MONTH	DAY	HOUR	neighbour	crime_type_id
2	2003	8	8	13	24	6
3	2003	8	14	9	10	7
4	2003	4	4	19	5	4
5	2003	6	15	22	12	7
6	2003	10	12	0	14	5
7	2003	10	26	0	14	5
8	2003	4	6	23	3	4
9	2003	1	28	18	5	2
10	2003	2	3	17	5	2
11	2003	3	9	21	24	7
12	2003	6	13	13	4	7
13	2003	6	13	16	24	7
14	2003	3	22	16	24	7
15	2003	2	10	17	5	2
16	2003	5	10	0	14	5
17	2003	4	19	12	5	2
18	2003	2	18	22	3	4

(b)

(c)

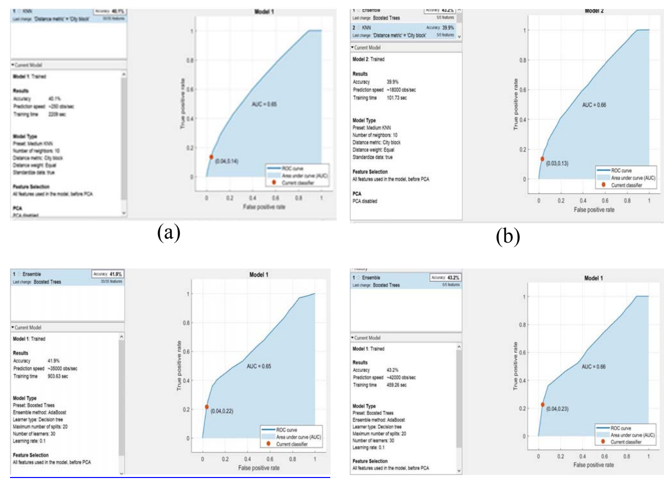
Fig. 7. Processed dataset (a) approach 1, (b) approach 2, and (c) result of skewed data.

C. K-Nearest Neighbour (KNN)

KNN was applied in both approaches with the same parameters, and the accuracies and training time was compared. For approach 1, KNN's accuracy was 40.1% and training time is 2209 seconds, while for approach 2 it turned out to be 39.9% accurate and took 101.73 seconds to train.

D. Boosted Decision Tree

We applied boosted decision tree algorithm in both approaches and compared the results. For both approaches, we used the Adaptive Boosting (AdaBoost) ensemble method and learner-type decision tree. AdaBoost is a meta-algorithm that combines several weak learners to improve a weak classifier. The maximum number of splits was 20. Accuracy and training time for approach 1 was 41.9% 903.63 seconds, respectively, while approach 2 was 43.2% accurate with 459.26 sec training time. The results from both methods (KNN and boosted decision tree) are shown in Fig. 8 for both approaches.



(a) (b) (c) (d)

Fig. 8. KNN results for approaches 1 and 2 (a, b) and boosted decision tree results for approaches 1 and 2 (c, d)

V. Conclusion

In this research, Vancouver crime data for the last 15 years was used in two different dataset approaches. Machine Learning predictive models KNN and boosted decision tree were used to obtain crime-prediction accuracy between 39% to 44%. The accuracy, complexity, and training time of algorithms were slightly different for different approaches and algorithms. The prediction accuracy can be improved by tuning both the algorithm and the data for specific applications. Although this model has low accuracy as a prediction model, it provides a preliminary framework for further analyses.

Acknowledgment

This paper is based on the capstone design project completed at the Fraser International College (FIC) as part of the pre-master's program in product realization. The authors would like to thank Ms. Amy Yeung, Dr. Alberta Seah, Ms. Sandra Kimber, and Ms. Sharla Reid for their valuable support.

References

[1] A. Bogomolov, B. Lepri, J. Staiano, N. Oliver, F. Pianesi, and A. Pentland, "Once upon a crime: towards crime prediction from demographics and mobile data," *Proc. of the 16th Intl. Conf. on Multimodal Interaction*, pp. 427-434, 2014.

[2] H. Adel, M. Salheen, and R. Mahmoud, "Crime in relation to urban design. Case study: the greater Cairo region," *Ain Shams Eng. J.*, vol. 7, no. 3, pp. 925-938, 2016.

[3] "Overall crime rate in Vancouver went down in 2017, VPD says," *CBC News*, Feb. 15, 2018. [Online] Available: https://www.cbc.ca/news/canada/british-columbia/crime-rate-vancouver-2017-1.4537831. [Accessed: 09- Aug- 2018].

[4] J. Kerr, "Vancouver police go high tech to predict and prevent crime before it happens," *Vancouver Courier*, July 23, 2017. [Online] Available: https://www.vancourier.com/news/vancouver-police-go-high-tech-to-predict-and-prevent-crime-before-it-happens-1.21295288. [Accessed: 09- Aug- 2018]

[5] J. Han, *Data mining: concepts and techniques*, Morgan Kaufmann, 2012.

[6] R. Iqbal, M. A. A. Murad, A. Mustapha, P. H. Shariat Panahy, and N. Khanahmadliravi, "An experimental study of classification algorithms for crime prediction," *Indian J. of Sci. and Technol.*, vol. 6, no. 3, pp. 4219-4225, Mar. 2013.

[7] H. Chen, W. Chung, J. J. Xu, G. Wang, Y. Qin, and M. Chau, "Crime data mining: a general framework and some examples," *IEEE Computer*, vol. 37, no. 4, pp. 50-56, Apr. 2004.

[8] T. Beshah and S. Hill, "Mining road traffic accident data to improve safety: role of road-related factors on accident severity in Ethiopia," *Proc. of Artificial Intell. for Develop. (AID 2010)*, pp. 14-19, 2010.

[9] M. Al Boni and M. S. Gerber, "Area-specific crime prediction models," *15th IEEE Intl. Conf. on Mach. Learn. and Appl.*, Anaheim, CA, USA, Dec. 2016.

[10] Q. Zhang, P. Yuan, Q. Zhou, and Z. Yang, "Mixed spatial-temporal characteristics based crime hot spots prediction," *IEEE 20th Intl. Conf. on Comput. Supported Cooperative Work in Des. (CSCWD)*, Nanchang, China, May 2016.

[11] N. Mahmud, K. Ibn Zinnah, Y. Ar Rahman, and N. Ahmed, "CRIMECAST: a crime prediction and strategy direction service," *IEEE 19th Intl. Conf. on Comput. and Inform. Technol.*, Dhaka, Bangladesh, Dec. 2016.

[12] Y. L. Lin, L. C. Yu, and T. Y. Chen, "Using machine learning to assist crime prevention," *IEEE 6th Intl. Congr. on Advanced Appl. Inform. (IIAI-AAI)*, Hamamatsu, Japan, Jul. 2017.

[13] F. K. Bappee, A. S. Júnior, and S. Matwin, "Predicting crime using spatial features," *Can. AI 2018: Advances in Artificial Intel.-Lecture Notes in Comput. Sci.*, vol. 10832, pp. 367-373, Springer, Mar. 2018.

[14] H. W. Kang, H. B. Kang, "Prediction of crime occurrence from multi-modal data using deep learning," *PLoS ONE*, vol. 12, no. 4, Apr. 2017.

[15] V. Grover, R. Adderley, and M. Bramer, "Review of current crime prediction techniques," *Intl. Conf. on Innovative Techn. and Appl. of Artificial Intel.*, pp. 233-237, Springer, London, 2007.

[16] X. Zhao and J. Tang, "Exploring transfer learning for crime prediction," *IEEE Intl. Conf. on Data Mining Workshop (ICDMW)*, New Orleans, LA, USA, Nov. 2017.

[17] R. Marchant, S. Haan, G. Clancey, and S. Cripps, "Applying machine learning to criminology: semi□parametric spatial□demographic Bayesian regression," *Security Inform.*, vol. 7, no. 1, Dec. 2018.

[18] L. McClendon and N. Meghanathan, "Using machine learning algorithms to analyze crime data," *Mach. Learn. and Appl.: an Intl. J. (MLAIJ)*, vol.2, no.1, Mar. 2015.

[19] S. Prabakaran and S. Mitra, "Survey of analysis of crime detection techniques using data mining and machine learning," *Nat. Conf. on Math. Techn. and its Appl. (NCMTA 2018), IOP J. of Physics: Conf. Series*, vol. 1000, 2018.

[20] E. Cesario, C. Catlett, and D. Talia, "Forecasting crimes using autoregressive models," *IEEE 14th Intl. Conf. on Dependable, Auton. and Secure Comput.*, Auckland, New Zealand, Aug. 2016.

[21] M. A. Tayebi, U. Glässer, and P. L. Brantingham, "Learning where to inspect: location learning for crime prediction," *IEEE Intl. Conf. on Intel. and Security Inform. (ISI)*, Baltimore, MD, USA, May 2015,.

[22] L. G. A. Alves, H. V. Ribeiro, and F. A. Rodrigues, "Crime prediction through urban metrics and statistical learning," *Physica A*, vol. 505, pp. 435-443, 2018.

[23] N. Baloian, E. Bassaletti, M. Fernández, O. Figueroa, P. Fuentes. R. Manasevich, M. Orchard. S. Peñafiel, J. A. Pino, and M. Vergara, "Crime prediction using patterns and context," *IEEE 21st Intl. Conf. on Comput. Supported Cooperative Work in Des. (CSCWD)*, Wellington, New Zealand, Apr. 2017.

[24] M. V. Barnadas, *Machine learning applied to crime prediction*, Thesis, Universitat Politècnica de Catalunya, Barcelona, Spain, Sep. 2016.

[25] E. Budur, S. Lee, and V. S. Kong, "Structural analysis of criminal network and predicting hidden links using machine learning," *arXiv:1507.05739*, Sep. 2015.

[26] G. Zhou, J. Lin and W. Zheng, "A web-based geographical information system for crime mapping and decision support," *IEEE Intl. Conf. on Comput. Problem-Solving (ICCP)*, Leshan, China, Oct. 2012.

[27] S. Chainey and J. Ratcliffe, *GIS and crime mapping*, John Wiley & Sons, Incorporated, 2015.

[28] M. Stamp, *Introduction to machine learning with applications in information security*, Chapman and Hall/CRC, New York, 2017.

[29] S. Sivaranjani, S. Sivakumari, and M. Aasha, "Crime prediction and forecasting in Tamilnadu using clustering approaches," *IEEE Intl. Conf. on Emerg. Technol. Trends (ICETT)*, Kollam, India, Oct. 2016.

[30] N. Shama, *A machine learning approach to predict crime using time and location data*, B.Sc. Thesis, BRAC University, 2017.

[31] W. Osaku, "EDA of crime in Vancouver (2003-2017)," *Kaggle*, 2018. [Online] Available: https://www.kaggle.com/wosaku/eda-of-crime-in-vancouver-2003-2017. [Accessed: 09- Aug- 2018]

[32] Nicolas, "Analysis Strategy," *Kaggle*, 2018. [Online] Available: https://www.kaggle.com/nickpaps/analysis-strategy [Accessed: 09-Aug-2018]