Article

Predicting Spatial Crime Occurrences through an efficient Ensemble-Learning Model

Yasmine Lamari [1], Bartol Freskura [2], Anass Abdessamad [1], Sarah Eichberg [3], and Simon de Bonviller [1,*]

[1] Augurisk, Inc.; United States
[2] Velebit AI LLC; Croatia
[3] The Eichberg Group; United States
* Correspondence: sdebonviller@augurisk.com

Abstract: While the use of crime data has been widely advocated in the literature, its availability is often limited to large urban cities and isolated databases tend not to allow for spatial comparisons. This paper presents an efficient machine learning framework capable of predicting spatial crime occurrences, without using past crime as a predictor, and at a relatively high resolution: the U.S. Census Block Group level. The proposed framework is based on an in-depth multidisciplinary literature review allowing the selection of 188 best-fit crime predictors from socio-economic, demographic, spatial, and environmental data. Such data are published periodically for the entire United States. The selection of the appropriate predictive model was made through a comparative study of different machine learning families of algorithms, including generalized linear models, deep learning, and ensemble learning. The gradient boosting model was found to yield the most accurate predictions for violent crimes, property crimes, motor vehicle thefts, vandalism, and the total count of crimes. Extensive experiments on real-world datasets of crimes reported in 11 U.S. cities demonstrated that the proposed framework achieves an accuracy of 73 and 77% when predicting property crimes and violent crimes, respectively.

Keywords: Crime prediction; Ensemble Learning; Machine Learning; Regression.

1. Introduction

The ability to access reliable, high-resolution crime data has long been advocated by researchers [1]. The analysis of crime data can be useful in many aspects of law enforcement policy. Among other uses, it may help allocate law enforcement resources where they are most needed [2] and adapt law enforcement policies to an ever-changing environment [3].

In the United States, crime data is mainly available through the FBI's Uniform Crime Report program through the Summary Reporting System (SRS), currently transitioning into the National Incident-Based Reporting System (NIBRS). However, the available data is still fragmented and not always directly comparable across the contiguous U.S. In the absence of homogenous data, local crime prediction can provide an additional perspective.

In the field of machine learning (ML), many approaches and models have been defined in relation to crime prediction through methods of classification, clustering, regression, deep learning, and ensemble learning [4, 5]. However, such models face a number of challenges. Among them, many ML models dedicated to crime prediction are exclusively data-driven in their feature selection process: the extensive use of feature engineering and automated feature selection techniques can then limit the out-of-sample reliability of predictions. Besides, the ML models reaching satisfying performances in their predictions tend to use past crime as a determinant of future crime [6–8]. As such data tend to be available only in major urban centers and is often difficult to compare across locations, databases tend to be defined either at an aggregated level (city, county…) or at the local level only (i.e. detailed grid in one city only).

Therefore, offering a prediction with a wide coverage and a high resolution would provide policy makers and individuals with spatial elements of comparison, in addition to the traditional advantages brought by predictive policing [9].

In this paper, we present a ML model able to predict crime counts in all U.S. Census Block Groups, by using data available throughout the entire contiguous U.S. Our model relies on a thorough review of the neighborhood effects literature to identify community correlates of crime.

As a first step, we reviewed different crime theories related to social, economic, and demographic characteristics of a neighborhood, and selected 188 predictors by combining this approach with correlation analysis. These predictors, along with our targets, consisting of crime counts for various crime types between 2014 and 2018, were gathered at the U.S. Census Block Group level for the contiguous U.S. Census Blocks are local areas defined as containing 600 to 3000 people, with a median BG area of about 1.3 km^2. They have been argued to align with residents' perception of their neighborhood, suggesting that they form an appropriate unit of analysis to study neighborhood effects [10]. To build our model, we use the Crime Open Database [11], geodocumenting crimes in 11 U.S. cities between 2014 and 2018, and thereby offering a variety of urban contexts.

Then, since we deal with a regression problem, we studied different ML algorithm families, including Generalized Linear Models (GLMs), Deep Learning, and Ensemble Learning. We maintained the most accurate model for most types of crimes considered, namely: violent crimes, property crimes, motor vehicle theft (MVT), and vandalism. Our model reaches up to 77% accuracy in predicting crime counts for over 13,897 urban Block Groups in the U.S.

In short, the main contributions of this paper are as follows:

- Contribution 1: A spatial crime prediction model using data commonly available throughout the entire continental U.S., thereby enabling spatial comparisons.
- Contribution 2: An efficient data strategy based on an in-depth multidisciplinary literature review on crime and state-of-the-art predictive ML techniques.
- Contribution 3: A concise comparison of the performance of three predictive ML models, namely: Poisson regression, Sequential Neural Network, and Gradient Boosting.
- Contribution 4: A set of extensive experiments on real-world datasets of crimes reported in different U.S. cities, and a detailed discussion of the promising local crime predictions achieved.

The remainder of this paper is structured as follows: Section 2 presents the theoretical background informing neighborhood effects on crime research and some state-of-the-art predictive ML algorithms. Section 3 describes the data strategy followed to produce the input dataset and the proposed predictive method. Section 4 discusses the achieved crime occurrences predictions. Finally, Section 5 concludes and identifies some directions for future research.

2. Background and Related Work

2.1. Theoretical Background

Neighborhood effects is an important concept in geographic, public health and social science research and is concerned with how neighborhood conditions affect social outcomes. The notion can be traced back to University of Chicago sociologists Shaw and McKay [12] who proposed the field's oldest theoretical perspective, social disorganization, positing that neighborhood structures like socioeconomic disadvantage, racial heterogeneity, and residential mobility prevent residents from forming social ties to regulate crime. Shaw and McKay's work heralded a major paradigm shift away from individual-level theories of crime toward ecological models [13].

While social disorganization theory fell out of favor in the 1960s, the approach was revitalized in the 1980s by scholars in the U.S. with a renewed interest in neighborhood dynamics due to rising crime rates and urban decline. These authors updated the framework by addressing criticisms [14], testing and clarifying concepts [15, 16], and expanding causal mechanisms [17–19].

One important extension of social disorganization theory was the concept of collective efficacy [18], which refers to residents' ability to come together to achieve a shared desire for a safe neighborhood [20]. Collective efficacy combines social cohesion, defined as trust and sense of community between neighbors, with informal social control, which refers to residents' ability to regulate community disorder. Subsequent research has repeatedly demonstrated that collective efficacy exerts a strong effect on community crime and violence [21–23].

Routine activities theory is another prominent perspective focused on neighborhood crime. The theory posits that crime is more likely to occur when three factors meet in time and space: a motivated offender, an available target, and the absence of a capable guardian (e.g. an authority figure) [24]. Research in this area is concerned with temporal and spatial effects on crime and focuses on micro-geographies, including "hot spots," such as street segments where crime occurs [25].

Pratt and Cullen [13] assessed routine activity theory and social disorganization theory along with other criminological frameworks in their meta-analysis of macro-level predictors and theories of crime. They found that social disorganization and resource deprivation theory, which links economic inequality with an inability to regulate behavior in accordance with social norms, had the strongest effects on crime. In their systematic, integrative review of the neighborhood effects literature, Brisson & Roll [26] found that social disorganization theory and routine activities theory had the most powerful effects on crime. Based on these findings, we elected to use predictors associated with social disorganization theory in our framework.

Predictors of crime associated with social disorganization theory can be divided into two broad categories: "static" neighborhood conditions that reflect a neighborhood's social structural conditions [27, 26] and "dynamic" neighborhood processes, such as collective efficacy or social cohesion [18, 26, 28, 29]. Single static variables with significant effects on crime include income inequality [30–33]; race/ethnic segregation [34–36]; racial heterogeneity [37–40], residential instability [41], gender [42–45], and age [46–48], all taken into account in our model. Table 1 lists major social structural predictors of crime assessed in prior reviews [26, 49] and a meta-analysis [13] and indicates their effects (positive, negative, unclear) on crime.

Table 1. Direct and Indirect Effects of Variables on Urban Crime [13, 26, 49].

Social Structural Variables	Relationship to Crime
Concentrated Disadvantage	Positive
Unemployment	Unclear, possibly positive
Family Disruption	Positive
Residential Instability	Positive
Racial/Ethnic Heterogeneity	Positive
Segregation	Positive
Income Inequality	Positive
Immigration	Unclear
Gender (Male)	Positive
Age (Younger)	Positive

Multicollinearity among social structural variables is a potential challenge in regression models concerned with causal analysis of crime. This is because of strong links between many of the structural factors associated with crime [50], creating what Wilson [19] referred to as "concentration effects." Concentrated disadvantage or "resource deprivation" [51] is one such index variable that incorporates indicators for income inequality, poverty, racial diversity, educational attainment, residential mobility, unemployment and/or family disruption [50, 52, 53]. Another index variable is family disruption which combines measures of family stability such as nonmarriage, early marriage, early childbearing, parental absenteeism, widowhood, and death [54–56]. While we are aware of multicollinearity issues in crime research, we did not use index variables in our model since collinearity is only an issue for causal inference and not prediction, the purpose of our framework.

Brisson and Roll [26] assessed four dynamic or process variables in their review that tend to interact with static predictors to affect crime. Assessing social cohesion, Brisson and Roll found limited evidence of a relationship between social cohesion and crime in studies on hate crimes [57] and general violence or intimate partner violence [58]. Results were mixed for informal social control, with one study showing a relationship between informal social control and a decline in delinquency rates [59] and another finding effects on anti-Black hate crime [57]. A third study, however, was unable to demonstrate a link between informal social control and general violence and intimate partner violence [58]. Research on social ties, which is a concept closely affiliated with social cohesion that looks at the number of relationships in a community, has demonstrated that effects on crime depend on the type and intensity of relationships and their influence on informal social control [40, 60]. Finally, support for the effect of collective efficacy on crime is robust and the concept is applicable across urban locations. Collective efficacy has been associated with a decline in violent victimization [61]; a decline in homicide [61]; reduced fear of crime [62]; and increased street efficacy [53].

There is a nascent rural crime literature, largely dominated by studies oriented around social disorganization theory [63]. Findings have been inconsistent, with evidence for some aspects of social disorganization but little or no support for others [64]. Consequently, it is difficult to make broad statements about crime patterns but preliminary research indicates that variables like poverty and family disruption affect crime differently in rural communities than in urban areas. For example, research suggests that poverty has no relationship or an inverse relationship with crime [65–68, 63, 69] possibly because community stability produces stronger informal social control [70]. In another example, racial heterogeneity appears to have limited effects on social disorganization in rural settings, given the mixed results of studies. For example, Bouffard and Muftic [65] found no association between ethnic heterogeneity and violent crime, while other scholars have found a positive relationship between variables, including robbery and assault in rural counties [67] and youth violent crime [71]. Table 2 provides an overview of social structural predictors of crime in rural communities.

Table 2. Social Disorganization Variables Effects on Rural Crime [64, 72].

Structural Variables	Relationship to Crime
Poverty, Income, Income Inequality	No relationship or Inverse
Unemployment	Unclear, possibly positive
Family Disruption	Unclear, possibly no relationship or even inverse
Residential Instability	Unclear
Racial/Ethnic Heterogeneity	Unclear

Due to remaining uncertainty about the mechanisms of crime in rural communities, we did not create a separate model for predicting rural crime but applied the same model to rural and urban contexts. Similarly, sparse research into suburban crime [65, 68, 73] meant that we were not able to develop a distinct model to predict crime in suburban settings.

In sum, based on our thorough review of the neighborhood effects literature, we decided to select predictors of urban crime associated with the neighborhood effects perspective, mainly social disorganization theory, to inform our framework. These predictors have been shown to have significant relationships with crime in prior research, and are summarized in Table 3. We subsequently drew on datasets, including the U.S. Census, to select social, economic, and demographic indicators to represent these predictors.

2.2. Related Work: ML and crime prediction

In this section, we review the recent work on spatial crime prediction using different ML techniques, with an emphasis on the methods estimating crime rates or occurrences.

H.W. Kang and H.B. Kang (Kang and Kang, 2017) proposed a deep learning method based on a deep neural network (DNN) for crime occurrences prediction at the US census-tract level. In their data strategy, the authors involved various sources of data, including crime occurrence reports and demographic and climate information. Additionally, they considered environmental context information using image data from Google Street View. In their prediction model, the authors adopted a multimodal data fusion method, in such a way that the DNN is defined with four-layer groups, namely: spatial, temporal, environmental context, and joint feature representation layers. This predictive model produces significant results in terms of accuracy. However, it was trained and tested using only real-world datasets collected from the city of Chicago, Illinois, due to data availability constraints. Thus, it cannot be used uniformly for all US cities.

Based also on the deep learning family of methods, Huang et al. [75] proposed a Recurrent Neural Network (RNN) for predicting spatio-temporal crime occurrences in urban areas. Their method is characterized by detecting dynamic crime patterns using a hierarchical recurrent neural network from hidden representation vectors. These vectors embed spatial, temporal, and categorical signals while preserving the correlations between the crime occurrences and their time slots. This method was trained and evaluated using real-world datasets collected from New York City. In this dataset, crimes are recorded with their respective category, location, and timestamp. However, such a method cannot be uniformly used for all urban areas, since this kind of data is not commonly available for other cities.

A probabilistic model based on the Bayesian paradigm was suggested by [76]. This proposed model was conceived to predict spatial crime rates using demographic and historical crime data. It quantifies the uncertainties in the output predictions and the model parameters using a combination of two Bayesian linear regression models. A first parametric model that takes into account the relationship between crime rate and location-specific factors, and a second nonparametric model that addresses the spatial dependencies. It also handles the inferences on the regression parameters by estimating the posterior probability distribution using the MCMC (Markov Chain Monte Carlo method). Results regarding three types of crime comply with the existing theoretical criminological assumptions. In addition, the proposed model can be generalized to all of Australia, since it uses demographic census data available nearly in all locations.

Besides these efforts, we found that ensemble-learning methods have been the subject of several studies in the literature, and have proven to be effective in the context of spatial crime prediction. This family of ML models draws its strength from the fact that it employs multiple learning algorithms. Each algorithm works on a chunk or on the whole dataset to produce intermediate predictions that are collected and processed in order to obtain the final predictions. Examples of studies relying on ensemble-learning methods include [6, 7, 77].

Alves et al. [6] used a random forest regressor to predict crime in urban areas. Knowing that this ML model is extremely sensitive to its main parameters (the number of trees and the maximum depth of each tree), the authors estimated them using the stratified k-fold cross-validation method, and then, they set them using the grid-search algorithm. Thus, they managed to create a trade-off between bias and variance errors. They have also studied the relationship between crime incidents and urban indicators using various statistical tests and metrics, in order to select the most important explanatory indicators. Their proposed model has been trained and tested using urban indicators data from all Brazilian cities. Experiments showed that it can yield a promising accuracy reaching up to 97% on crime prediction. However, predictions concern only a single type of crime, i.e. homicides, at an aggregated city-level.

More recently, Kadar et al. [7] proposed a predictive approach for spatio-temporal crime hotspots predictions in low population density areas. The authors focused mainly on the problem of class imbalance, handled through a repeated under-sampling technique. Indeed, in the learning phase, their predictive model is trained using balanced sub-samples of the input dataset, which are created by randomly selecting the same number of instances from the majority and minority classes. Then, they adopted the random forest classifier as a base learner for predicting crime hotspots after a deep evaluation of other ML models. Results with an input dataset composed of different predictors, such as socio-economic, geographical, temporal, meteorological, and crime variables, showed that this approach outperforms the common baselines in predicting hotspots. However, it is conceived to predict only a single type of crime, burglary incidents.

Another ensemble-learning predictive approach was proposed in [77]. Ingilevich and Ivanov conceived a three-step approach for crime occurrences prediction in a specific urban area. Their approach starts with a clustering step, in which the authors applied the DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm in order to study the spatial patterns of the considered crime types and to remove the noise from the dataset. Then, it is followed by a step of feature selection, in which authors applied the chi-squared test in order to study the relative importance of the features. Finally, in the third step, the authors used the gradient boosting model to predict crime occurrences after a performance comparison of two other models, i.e. the linear regression and the logistic regression. This model was trained and tested using the crime incidents dataset from Saint-Petersburg, Russia. It outperformed the two other models in terms of accuracy for three types of street crimes.

Building on this previous work and on our own efforts, we propose a predictive framework that has been carefully designed to spatially predict crime occurrences at the U.S. Census Block Group level, based on the gradient boosting model.

3. Methodology

3.1. Data strategy

Feature selection for this project was done using several approaches. First, relevant crime predictors were identified using insights from the sociological, geographical, and ML literature, as detailed in the Theoretical Background and Related Works sections. Second, correlations between all variables available from the American Community Survey and our target variables were examined, and variables displaying a correlation over 0.25 were retained. Third, variables were generated for each feature based on neighboring Block Groups' characteristics, to allow for spillover effects. The following sections detail data sources and preprocessing steps used throughout this study.

3.1.1. Data sources

The input dataset of our proposed framework was built from different sources, as listed below:

- Socio-economic and demographic data were extracted from the American Community Survey (ACS) 5-Year Estimates [78]. In the present work, we used the ACS 5-year Estimates collection covering the period 2014-2018 for all U.S. Block Groups.
- Climate data (monthly averages related to wind, rainfall, and temperature) was retrieved from the WorldClim 2 project [79].
- Law enforcement data was collected based on Homeland Infrastructure data related to Local Law Enforcement Agencies in the U.S.
- Crime counts for Violent Crime, Property Crime, and two specific subcases (Vandalism and Motor Vehicle Theft) in the time-period 2014-2018 were extracted and pooled at the U.S. Census Block Group level from the Crime Open Database [11].

3.1.2. Data preprocessing

The feature preprocessing pipeline adopted in our data strategy consists of four steps: preparing the collected data, creating the new features, scaling the features, and de-skewing, as depicted in Figure 1.

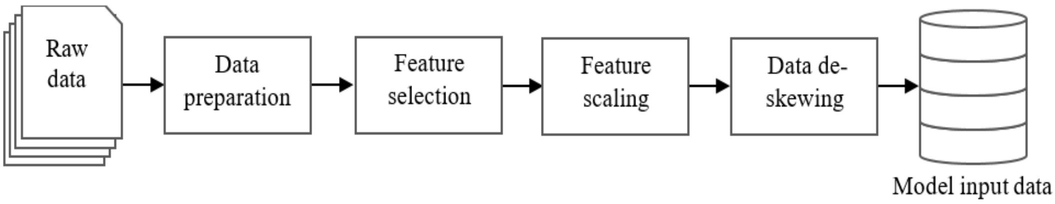


Figure 1. Data preprocessing steps.

First, the collected data was cleaned and formatted. Then, some new features were created by combining the existing features with the goal of adding explicit information. For example, for each socio-economic and demographic variable, a spillover variable was generated using the variable's mean or sum in neighboring Block Groups. In the feature scaling step, a min-max normalization was performed in order to transform all input feature values to the $[0,1]$ range. Finally, a $\log(1+x)$ de-skew function was applied only to variables with a skew score greater than 0.75 (found empirically to be optimal). The skew score was calculated using the skew function from the Scipy [80] library. $\log(1+x)$ de-skewing was also applied to the target variable during the training phase.

The above steps yielded a dataset composed of 13,897 observations where each observation has 188 features. For the sake of clarity, we have aggregated the considered features under 15 themes, as shown in Table 3. We present the mean absolute correlation of features per theme in order to take into account the positive and negative correlations to the count target variable, in addition to the mean of the feature importance per theme. The obtained values are expressed in percentages.

Table 3. Summary of the selected features.

Themes	**Number of attributes**	**Mean absolute correlation (%)**	**Mean feature importance (%)**
Poverty	14	23.57	0.59
Residential instability	4	19.89	0.75
Housing and commuting	14	19.18	0.65
Income	4	18.4	0.68
Population	4	16.95	1.26
Family disruption	10	16.79	0.69
Unemployment	8	11.16	0.66
Gender	2	9.29	0.71
Climate	60	8.99	0.31
Education	36	8.73	0.54
Socio-economic indicators	5	8.67	0.12
Age	10	7.45	0.64
Law enforcement	4	7.37	0.65
Ethnic heterogeneity	12	5.17	0.61
Land area	1	4.47	3.61

Target variables include four types of crime counts and a single variable, which represent a combination of two types of crime counts:

- Target 1: Violent crime occurrences
- Target 2: Property crime occurrences
- Target 3: MVT crime occurrences
- Target 4: Vandalism crime occurrences
- Target 5: Total crime count (Violent + Property)

A brief overview of correlations listed in Table 3 suggests that factors showing the highest correlations with total crime counts are related to static neighborhood conditions as poverty, residential instability, housing and commuting, and income, all clearly identified in the literature as crime determinants [81, 50, 41, 33], along with population and population density. Feature importances reveal that the land area covered by and population in a Block Group have the highest importance, as Block Groups can widely vary in size (with urban Block Groups smaller than rural Block Groups) and population (usually 600 to 3000).

3.2. The proposed method

The considered targets are count variables (the sum of crime type incidents within a fixed zone area, a Block Group, during 5 years) and can be approximated by a Poisson distribution. Thus, we first selected the Poisson regression model, because of its ability to model count data. The considered target variables and the logarithm of its expected values can be modeled by a linear combination of unknown parameters. However, this model assumes that the mean and variance are equal (equi-dispersion). Unfortunately, this assumption is often violated in the observed data [82].

Let y_i be the response variable. We assume that y_i follows a Poisson distribution with mean λ_i defined as a function of covariates x_i. The Poisson probability mass function is given by the equation below:

$$P(y_i) = \frac{e^{-\lambda_i} \ {\lambda_i}^{y_i}}{\lambda_i!} \quad (1)$$

Where: $\lambda_i = E(y_i|x_i)$, and P defines the dimension of the covariates vector incorporated in the model.

We have also examined the possibility of modeling the problem addressed in this paper using Deep Learning methods. The Multilayer perceptron is one of the most widely used class of artificial neural networks (ANN). It is composed of several layers. Each layer contains several, but non-connected perceptrons [83].

The first layer is called the input layer and, in our case, it is composed of 188 units (perceptrons), which correspond to the number of features. Since we are trying to solve a regression problem, the last layer (the output layer) contains only one perceptron. We added two hidden layers. The first layer contains 700 units, and the second includes 25 units. The input units pass their outputs to the units in the first hidden layer. Each of the hidden layer units adds a constant ('bias') to a weighted sum of its inputs, and then calculates an activation function of the result, in our case the *ReLU* activation function:

$$y = \max(0; x) \quad (2)$$

Finally, we have investigated the use of Ensemble Learning methods. We opted for the Gradient Boosting [84] algorithm because it performs well on tasks where the numbers of features and observations are relatively limited and have a small computational footprint. The gradient boosting model produces an ensemble of weak prediction models, typically decisions trees, and it generalizes them by allowing optimization of an arbitrary differentiable loss function, in our case, the *Fair* loss function [85].

As the model was trained on the $\log(1+x)$ transformed targets, we use the inverse $e^x - 1$ on the model predictions when inferencing in order to get proper crime count values.

The dataset is randomly split into train and test sets using a 80:20 ratio respectively. To find optimal model hyperparameters we employ the cross-validation strategy on the train set (n_folds = 6) along with grid search for the hyperparameter space search. The cross-validation chooses the optimal hyperparameters according to the lowest negative mean absolute error score.

We used the LightGBM gradient boosting algorithm implementation and the following optimal hyperparameters found using grid search:

Table 4. The optimal hyperparameters set using the grid search algorithm.

Parameters	Values
learning_rate	0.005
reg_lambda	0.01
bagging_fraction	1
num_leaves	128
max_bin	512
max_depth	7
num_iterations	5000
feature_selection	0.5
objective	Fair
seed	1337

Hyperparameter tuning was done on the total crime count target variable and the same optimal hyperparameters were used to train models for the remaining four target variables. In the end, each target variable has a dedicated gradient boosting model.

4. Results and Discussion

4.1. Experimental settings

All operations related to the training and the test of the three models, i.e. Gradient Boosting, neural network, and Poisson regressor, were conducted on a computer having a processor Intel (R) Core (TM) i5 of 2.40 GHz and eight Giga bytes of RAM.

The proposed framework was implemented using Python 3.7, installed on a virtual environment of the package manager Anaconda. For the Gradient Boosting model implementation, we used the Light GBM library. For the Poisson model implementation, we used the Scikit-learn package. And, for the neural network model implementation, we used the Keras library based on the TensorFlow backend.

4.2. Evaluation metrics

In order to assess the quality of the predictions obtained with our proposed framework, we relied on the most commonly used evaluation metrics for regression problems, namely: the Mean Absolute Error (MAE) and the Root Mean Squared Error (RMSE).

$$MAE = \frac{\sum_{i=1}^{n} |r_i - \hat{r}_i|}{n} \quad (3)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (r_i - \hat{r}_i)^2}{n}} \quad (4)$$

Where r_i denotes the ground truth target value for the i-th data point, $\hat{r}_i$ denotes the predicted target value for the i-th data point, and n is the total number of data points.

Additionally, we used a third metric to quantify the percentage of how close the predictions are against the ground truth: the MAE divided by the mean of target values. This metric, what we call accuracy in this paper, is defined as follows:

$$AC_p = 1 - \left(\sum_{i=1}^{n} |r_i - \hat{r}_i| \Big/ \sum_{i=1}^{n} r_i \right) \quad (5)$$

4.3. Experiment Results

Table 5 shows the performances of three ML models, namely: Poisson regression, Deep learning, and Gradient boosting. We applied these models for each crime type, in addition to the total count of crimes, using the same input dataset and in the same conditions. Then, we measured their accuracy using the MAE and RMSE described above.

The gradient boosting model outperforms the other models in all the evaluated types of crime. It should be noted, however, that the deep learning model also yields performances close to the gradient boosting results.

Table 5. Comparison of the performance of three predictive ML models in terms of MAE and RMSE.

Crime types	Metrics	Models		
		Poisson regression	Deep learning	Gradient boosting
Count	MAE	181.94	130.69	123.24
	RMSE	439.35	331.14	318.28
Violent	MAE	76.41	52.48	49.87
	RMSE	175.70	132.39	132.37
Property	MAE	114.34	86.61	79.13
	RMSE	309.25	246.30	230.73
MVT	MAE	15.54	9.35	8.70
	RMSE	37.64	23.28	23.81
Vandalism	MAE	28.56	20.18	18.54
	RMSE	56.25	39.04	38.19

In order to further evaluate the performance of these predictive models, we selected a random set of 1000 observations from the input dataset, and then we compared the predicted crime occurrences of each type of crime, in addition to the total count of crime occurrences, against the ground truth, as depicted in Figure 2.

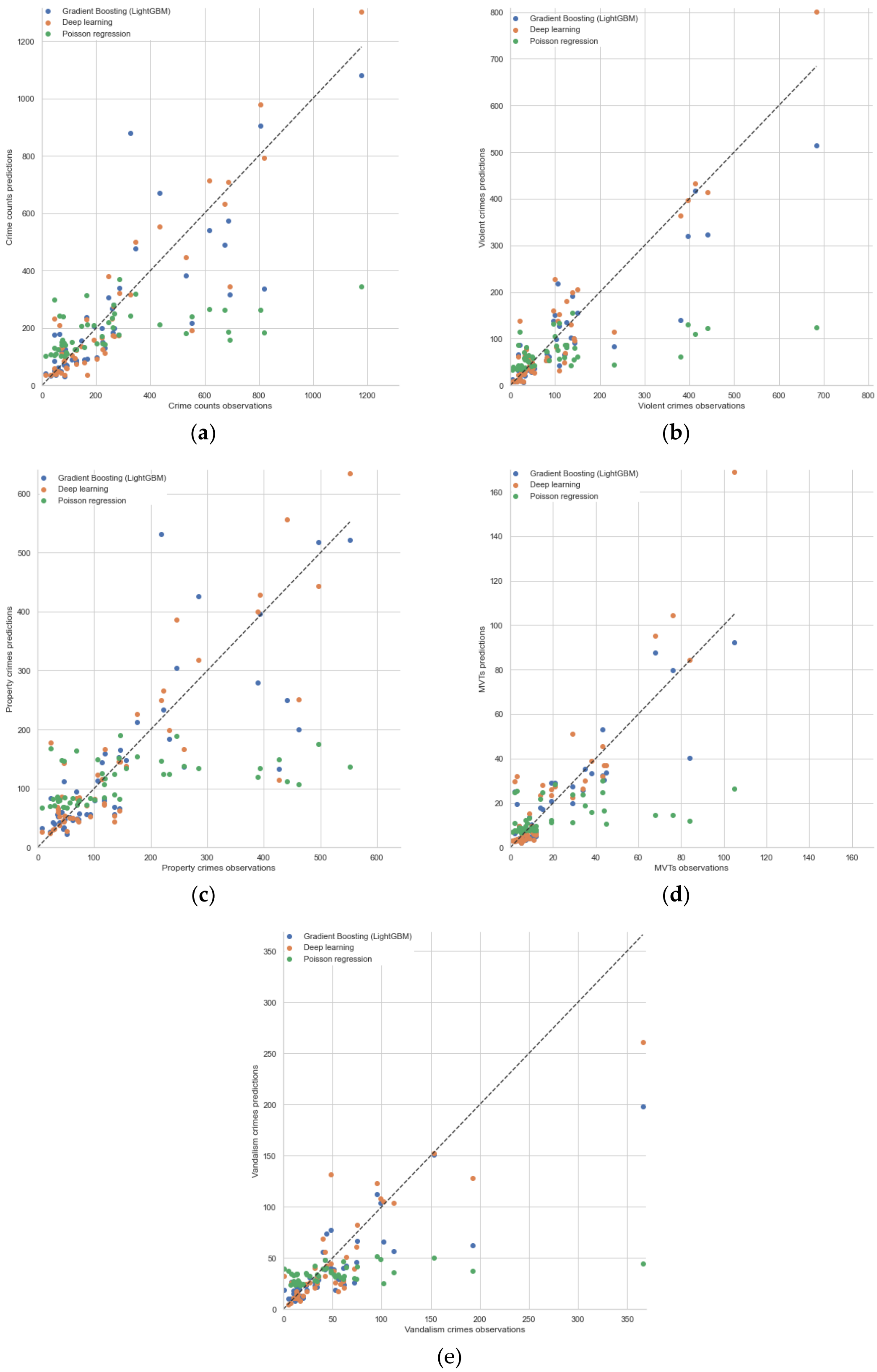
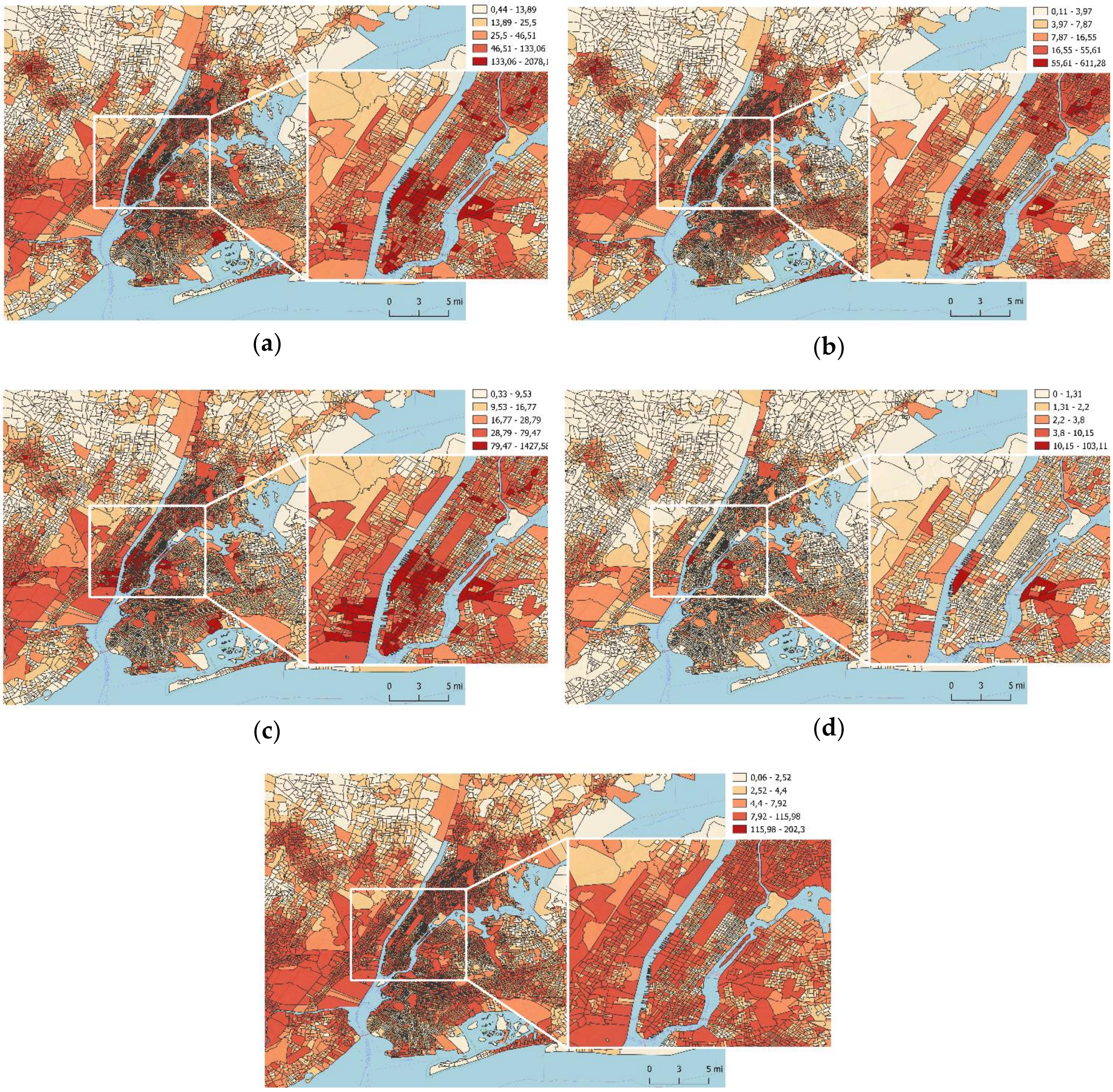


Figure 2. Comparison of the predicted occurrences of crimes against the ground truth using three different models. (a) Total crime count: predictions vs. real observations; (b) Violent crimes: predictions vs. real observations; (c) Property crimes: predictions vs. real observations; (d) MVT: predictions vs. real observations; (e) Vandalism: predictions vs. real observations.

As stated before, our framework is able to provide predicted crime occurrences for all Block Groups in the contiguous U.S. The learning phase was performed on 188 identified features, used to predict crime occurrences for 11 U.S. cities across 13,897 Block Groups and for 5 years (2014-2018). The model was then used to produce predictions for crime occurrences for the same period and all U.S. Block Groups. For the sake of clarity, Figure 3 represents our findings for one year using map visualizations of the New York City area, with a focus on Manhattan.



(**a**) (**b**) (**c**) (**d**) (**e**)

Figure 3. Map visualizations of yearly-predicted crime occurrences in New York City[1]. (a) Predicted total crime (count) occurrences; (b) Predicted violent crime occurrences; (c) Predicted property crime occurrences; (d) Predicted MVT occurrences. (e) Predicted vandalism acts.

4.4. *Discussion*

Our approach generates Mean Absolute Errors (MAE) between 36 (vandalism) and 41% (property crime) of the targets' means, suggesting accuracies between 59% and 64% in our ability to

[1] Categories used to generate maps (from light to dark) correspond to the first quartile, second quartile, third quartile, fourth quartile (excluding the 2 highest centiles), and the 2 highest centiles, of crime count predictions, respectively. Basemap obtained from OpenStreetMap, and U.S. Census Block Groups delimitations were extracted from the Tiger Census Shapefiles.

predict the exact count of crimes occurring in a Block Group between 2014 and 2018. This performance can appear moderate in comparison to studies using aggregated data (city, county, state) and past crimes as features, that can reach up to 97% accuracy [6]. However, we believe it to be remarkable given that (1) we predict crime at a higher resolution (Census Block Groups) and (2) our approach does not use past crimes as a predictor. Our approach has the advantage of only using features available throughout the entire U.S. Its results can thus provide elements of comparison to policy makers at the national level, including in urban environments where crime data is scarce. Furthermore, our tests reveal that predicting whether an observation lies within one of the categories displayed in Figure 3 instead of the exact crime count can increase our accuracy to 75% when predicting the total count of crimes, 77% for violent crimes, 73% for property crimes, 77% for motor vehicle thefts, and 77% for vandalism acts.

Analyzing the importance of selected features in the decision process can add perspective to our results. In the context of a tree-based algorithm, feature importance can be calculated by the sum of all improvements over all internal nodes where this feature is used [86, cited by 6]. The resulting feature importance, as calculated by the LightGBM regressor within the Python SciKitlearn library [87], sums to 100 (across all features used) and provides a way to describe a feature's relative importance in generating the final prediction.

A number of features were found to be particularly important in predicting crime using this approach. Population and population density are the most important predictors. Such factors are followed by the total area covered by the Block Group, which can vary significantly (with larger Block Groups located in rural areas). The median age comes third, followed by the distance to the nearest local law enforcement agency. However, those features collectively explain less than 11% of the total feature importance (with the 10 most important, involving additional factors related to social mobility and education, explaining 17% of the total importance). These results highlight the complexity of crime as a social phenomenon, as an important number of features in our framework significantly improve our ability to predict crime occurrences.

Besides, in many instances, spillover features (i.e. features describing attributes of the neighboring Block Groups) were found as more important than original features (describing attribute of a single Block Group). This is further illustrated by an important spatial autocorrelation in crimes predicted. If we consider total crime throughout the U.S., the Moran's I (i.e. the correlation between crime in a Block Group and the average crime predicted in neighboring Block Groups) predicted by our approach is around 0.7 nationwide, and the existence of clusters is particularly clear in the case of violent crime, vandalism, and motor vehicle theft (see Figures 3: (b), (d) and (e) for the case of New York).

Finally, a number of limitations should be stated. First, due to the methodological framework used, we can identify features of importance but not their impact (positive or negative) on crime in our model. Second, our approach is based on more than 180 features gathered from multiple different sources. Therefore, it involves a significant amount of work in terms of data processing. Third and nonetheless, our accuracy could be improved by adding additional types of features to the analysis. In this perspective, considering point of interests (involving a significant amount of social interactions), such as bus stops [2], malls, bars, churches, or schools [77], additional factors related to street lights [74] and/or social networks data [88] are additional approaches that can be complementary to our analysis. Considering ambient population instead of residential population [89] is also a promising perspective for future research. Finally, our model is trained on various urban contexts, meaning that it does not necessarily capture crime dynamics in rural settings. Consequently, predictions relative to rural areas are more uncertain than their urban counterparts.

5. Conclusions

In this paper, we proposed a ML framework able to provide predictions for spatial crime occurrences across all U.S. Census Block Groups in the contiguous U.S. Our findings from a set of extensive experiments on real-world datasets of crimes reported in 11 U.S. cities demonstrate that the proposed framework yields accurate predictions for the different crime types considered, i.e. violent crimes, property crimes, motor vehicle thefts, vandalism acts, and total count of crime occurrences. For these crimes types, our ability to predict whether crime count in a Block Group belongs to the first, second, third, or fourth quartile or the 2 highest centiles range between 73 and 77%.

We believe that our findings could be further enhanced if we consider selected additional features, such as social networks data, sites involving significant amounts of social interactions (malls, bars, churches, schools, etc.), land use, and streetlights, among others. Another track to explore in future research could be a deep dive into the subject of rural crime: although many factors defining rural areas (such as lower population density) have indeed be taken into account by our model, differing societal frameworks might justify the use of a separate model in the future.

Author Contributions: Conceptualization, Simon de Bonviller and Sarah Eichberg; Methodology, Anass Abdessamad and Bartol Freskura; Software, Anass Abdessamad and Bartol Freskura; Validation, Simon de Bonviller, Anass Abdessamad and Bartol Freskura; Formal Analysis, Yasmine Lamari, Bartol Freskura and Anass Abdessamad; Investigation, Bartol Freskura and Anass Abdessamad; Resources, Simon de Bonviller, Yasmine Lamari, Anass Abdessamad, Sarah Eichberg, and Bartol Freskura; Data Curation, Yasmine Lamari, Anass Abdessamad and Simon de Bonviller; Writing-Original Draft Preparation, Yasmine Lamari, Simon de Bonviller, Anass Abdessamad, Sarah Eichberg, and Bartol Freskura; Writing-Review & Editing, Yasmine Lamari, Simon de Bonviller, Anass Abdessamad, and Sarah Eichberg; Visualization, Yasmine Lamari and Anass Abdessamad; Supervision, Simon de Bonviller and Yasmine Lamari; Project Administration, Simon de Bonviller and Yasmine Lamari; Funding Acquisition, Simon de Bonviller, Anass Abdessamad and Yasmine Lamari.

All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by Augurisk in the context of a crime risk assessment project for commercial purposes.

Conflicts of Interest: Simon de Bonviller has been involved in Augurisk as a lead scientist. Yasmine Lamari and Anass Abdessamad have been involved in Augurisk as data scientists. Sarah Eichberg and Bartol Freskura have been involved in Augurisk as external consultants. The funder (Augurisk) has been involved in (1) the research goal of the study (providing crime predictions in the contiguous United States), and (2) the decision to publish the results of this study.

References

1. Clancey, G. Are We Still 'Flying Blind?' Crime Data and Local Crime Prevention in New South Wales. *Current Issues in Criminal Justice*, **2011**, *22* (3), 491–500. https://doi.org/10.1080/10345329.2011.12035901.
2. Cichosz, P. Urban Crime Risk Prediction Using Point of Interest Data. *ISPRS International Journal of Geo-Information*, **2020**, *9* (7), 459. https://doi.org/10.3390/ijgi9070459.
3. Inayatullah, S. The Futures of Policing: Going beyond the Thin Blue Line. *Futures*, **2013**, *49*, 1–8. https://doi.org/10.1016/j.futures.2013.01.007.
4. Almaw, A.; Kadam, K. Survey Paper on Crime Prediction Using Ensemble Approach. *International Journal of Pure and Applied Mathematics*, **2018**, *118* (8), 133–139.
5. Prabakaran, S.; Mitra, S. Survey of Analysis of Crime Detection Techniques Using Data Mining and Machine Learning. *J. Phys.: Conf. Ser.*, **2018**, *1000*, 012046. https://doi.org/10.1088/1742-6596/1000/1/012046.
6. Alves, L. G. A.; Ribeiro, H. V.; Rodrigues, F. A. Crime Prediction through Urban Metrics and Statistical Learning. *Physica A: Statistical Mechanics and its Applications*, **2018**, *505*, 435–443. https://doi.org/10.1016/j.physa.2018.03.084.
7. Kadar, C.; Maculan, R.; Feuerriegel, S. Public Decision Support for Low Population Density Areas: An Imbalance-Aware Hyper-Ensemble for Spatio-Temporal Crime Prediction. *Decision Support Systems*, **2019**, *119*, 107–117. https://doi.org/10.1016/j.dss.2019.03.001.
8. Lin, Y.-L.; Yen, M.-F.; Yu, L.-C. Grid-Based Crime Prediction Using Geographical Features. *ISPRS International Journal of Geo-Information*, **2018**, *7* (8), 298. https://doi.org/10.3390/ijgi7080298.
9. Meijer, A.; Wessels, M. Predictive Policing: Review of Benefits and Drawbacks. *International Journal of Public Administration*, **2019**, *42* (12), 1031–1039. https://doi.org/10.1080/01900692.2019.1575664.
10. Konkel, R. H.; Ratkowski, D.; Tapp, S. N. The Effects of Physical, Social, and Housing Disorder on Neighborhood Crime: A Contemporary Test of Broken Windows Theory. *ISPRS International Journal of Geo-Information*, **2019**, *8* (12), 583. https://doi.org/10.3390/ijgi8120583.

11. Ashby, M. P. J. Studying Crime and Place with the Crime Open Database: Social and Behavioural Scienes. **2018**. https://doi.org/10.1163/24523666-00401007.
12. Shaw, C. R.; McKay, H. D. *Juvenile Delinquency and Urban Areas*; Juvenile delinquency and urban areas; University of Chicago Press: Chicago, IL, US, 1942; pp xxxii, 451.
13. Pratt, T. C.; Cullen, F. T. Assessing Macro-Level Predictors and Theories of Crime: A Meta-Analysis. *Crime and Justice*, **2005**, *32*, 373–450. https://doi.org/10.1086/655357.
14. Bursik, R. J. Social Disorganization and Theories of Crime and Delinquency: Problems and Prospects. *Criminology*, **1988**, *26* (4), 519–552. https://doi.org/10.1111/j.1745-9125.1988.tb00854.x.
15. Kornhauser, R. R. Social Sources of Delinquency: An Appraisal of Analytic Models. **1978**.
16. Sampson, R.; Groves, W. B. Community Structure and Crime: Testing Social-Disorganization Theory. *American Journal of Sociology*, **1989**. https://doi.org/10.1086/229068.
17. Bursik, R. J. J.; Grasmick, H. G. Economic Deprivation and Neighborhood Crime Rates, 1960-1980. *Law & Soc'y Rev.*, **1993**, *27*, 263.
18. Sampson, R. J.; Raudenbush, S. W.; Earls, F. Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy. *Science*, **1997**, *277* (5328), 918–924. https://doi.org/10.1126/science.277.5328.918.
19. Wilson, W. J. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*; University of Chicago Press: Chicago, 1987.
20. Cole, S. J. Social and Physical Neighbourhood Effects and Crime: Bringing Domains Together Through Collective Efficacy Theory. *Social Sciences*, **2019**, *8* (5), 147.
21. Browning, C. R. The Span of Collective Efficacy: Extending Social Disorganization Theory to Partner Violence. *Journal of Marriage and Family*, **2002**, *64* (4), 833–850.
22. Morenoff, J. D.; Sampson, R. J.; Raudenbush, S. W. Neighborhood Inequality, Collective Efficacy, and the Spatial Dynamics of Urban Violence. *Criminology*, **2001**, *39* (3), 517–558. https://doi.org/10.1111/j.1745-9125.2001.tb00932.x.
23. Sampson, R. J.; Wikström, P.-O. H. The Social Order of Violence in Chicago and Stockholm Neighborhoods: A Comparative Inquiry. In *Order, Conflict, and Violence*; Shapiro, I., Kalyvas, S. N., Masoud, T., Eds.; Cambridge University Press: Cambridge, 2008; pp 97–119. https://doi.org/10.1017/CBO9780511755903.006.
24. Cohen, L. E.; Felson, M. Social Change and Crime Rate Trends: A Routine Activity Approach. *American Sociological Review*, **1979**, *44* (4), 588–608. https://doi.org/10.2307/2094589.
25. Weisburd, D.; Groff, E. R.; Yang, S.-M. *The Criminology of Place: Street Segments and Our Understanding of the Crime Problem*; Oxford University Press, 2012.
26. Brisson, D.; Roll, S. The Effect of Neighborhood on Crime and Safety: A Review of the Evidence. *null*, **2012**, *9* (4), 333–350. https://doi.org/10.1080/15433714.2010.525407.
27. Furstenberg, F. F.; Cook, T. D.; Eccles, J.; Elder, G. H.; Sameroff, A. *Managing To Make It: Urban Families and Adolescent Success. Studies on Successful Adolescent Development*; University of Chicago Press, 5801 South Ellis Avenue, Chicago, IL 60637; Fax: 773-702-9756; 1999.
28. Coleman, J. S. Social Capital in the Creation of Human Capital. *American Journal of Sociology*, **1988**, *94*, S95–S120.
29. Putnam, R. D. *Bowling Alone: The Collapse and Revival of American Community*; Bowling alone: The collapse and revival of American community; Touchstone Books/Simon & Schuster: New York, NY, US, 2000; p 541. https://doi.org/10.1145/358916.361990.
30. Chiu, W. H.; Madden, P. Burglary and Income Inequality. *Journal of Public Economics*, **1998**, *69* (1), 123–141. https://doi.org/10.1016/S0047-2727(97)00096-0.
31. Hsieh, C.-C.; Pugh, M. D. Poverty, Income Inequality, and Violent Crime: A Meta-Analysis of Recent Aggregate Data Studies. *Criminal Justice Review*, **1993**, *18* (2), 182–202. https://doi.org/10.1177/073401689301800203.
32. Kelly, M. Inequality and Crime. *The Review of Economics and Statistics*, **2000**, *82* (4), 530–539. https://doi.org/10.1162/003465300559028.
33. Weatherburn, D. *What Causes Crime?*; NSW Bureau of Crime Statistics and Research Sydney, 2001.
34. Feldmeyer, B. The Effects of Racial/Ethnic Segregation on Latino and Black Homicide. *The Sociological Quarterly*, **2010**, *51* (4), 600–623.
35. Krivo, L. J.; Peterson, R. D.; Kuhl, D. C. Segregation, Racial Structure, and Neighborhood Violent Crime. *American journal of Sociology*, **2009**, *114* (6), 1765–1802.

36. Peterson, R. D.; Krivo, L. J. *Divergent Social Worlds: Neighborhood Crime and the Racial-Spatial Divide*; Russell Sage Foundation, 2010.
37. Balkwell, J. W. Ethnic Inequality and the Rate of Homicide. *Social Forces*, **1990**, *69* (1), 53–70.
38. Blau, P. M.; Golden, R. M. Metropolitan Structure and Criminal Violence. *Sociological Quarterly*, **1986**, *27* (1), 15–26.
39. Kubrin, C. Racial Heterogeneity and Crime: Measuring Static and Dynamic Effects. *Research in Community Sociology*, **2000**, *10*, 189–219.
40. Warner, B. D.; Rountree, P. W. Local Social Ties in a Community and Crime Model: Questioning the Systemic Nature of Informal Social Control. *Social problems*, **1997**, *44* (4), 520–536.
41. Schieman, S. Residential Stability and the Social Impact of Neighborhood Disadvantage: A Study of Gender-and Race-Contingent Effects. *Social Forces*, **2005**, *83* (3), 1031–1064.
42. Burton Jr, V. S.; Cullen, F. T.; Evans, T. D.; Alarid, L. F.; Dunaway, R. G. Gender, Self-Control, and Crime. *Journal of research in crime and delinquency*, **1998**, *35* (2), 123–147.
43. Carrabine, E.; Iganski, P.; South, N.; Lee, M.; Plummer, K.; Turton, J.; Iganski, P.; South, N.; Lee, M.; Plummer, K.; et al. *Criminology : A Sociological Introduction*; Routledge, 2004. https://doi.org/10.4324/9780203642955.
44. Chrisler, J. C.; McCreary, D. R. *Handbook of Gender Research in Psychology*; Springer, 2010; Vol. 1.
45. Rowe, D. C.; Vazsonyi, A. T.; Flannery, D. J. Sex Differences in Crime: Do Means and within-Sex Variation Have Similar Causes? *Journal of research in Crime and Delinquency*, **1995**, *32* (1), 84–100.
46. Hirschi, T.; Gottfredson, M. Age and the Explanation of Crime. *American journal of sociology*, **1983**, *89* (3), 552–584.
47. Farrington, D. P. Childhood Aggression and Adult Violence: Early Precursors and Later-Life Outcomes. *The development and treatment of childhood aggression*, **1991**, *5*, 29.
48. Flanagan, T. J.; Maguire, K. *Sourcebook of Criminal Justice Statistics - 1989*; Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics, 1990.
49. Sampson, R. J.; Morenoff, J. D.; Gannon-Rowley, T. Assessing "Neighborhood Effects": Social Processes and New Directions in Research. *Annual review of sociology*, **2002**, *28* (1), 443–478.
50. Land, K. C.; McCall, P. L.; Cohen, L. E. Structural Covariates of Homicide Rates: Are There Any Invariances across Time and Social Space? *American journal of sociology*, **1990**, *95* (4), 922–963.
51. Messner, S. F.; Rosenfeld, R.; Baumer, E. P. Dimensions of Social Capital and Rates of Criminal Homicide. *American Sociological Review*, **2004**, *69* (6), 882–903.
52. Lo, C. C.; Zhong, H. Linking Crime Rates to Relationship Factors: The Use of Gender-Specific Data. *Journal of Criminal Justice*, **2006**, *34* (3), 317–329.
53. Sharkey, P. T. Navigating Dangerous Streets: The Sources and Consequences of Street Efficacy. *American Sociological Review*, **2006**, *71* (5), 826–846.
54. McLanahan, S.; Bumpass, L. Intergenerational Consequences of Family Disruption. *American Journal of Sociology*, **1988**, *94* (1), 130–152.
55. Messner, S. F.; Sampson, R. J. The Sex Ratio, Family Disruption, and Rates of Violent Crime: The Paradox of Demographic Structure. *Social Forces*, **1991**, *69* (3), 693–713.
56. Sampson, R. J. Neighborhood Family Structure and the Risk of Personal Victimization. In *The social ecology of crime*; Springer, 1986; pp 25–46.
57. Lyons, C. J. Community (Dis) Organization and Racially Motivated Crime. *American Journal of Sociology*, **2007**, *113* (3), 815–863.
58. Frye, V. The Informal Social Control of Intimate Partner Violence against Women: Exploring Personal Attitudes and Perceived Neighborhood Social Cohesion. *Journal of Community Psychology*, **2007**, *35* (8), 1001–1018.
59. Cantillon, D. Community Social Organization, Parents, and Peers as Mediators of Perceived Neighborhood Block Characteristics on Delinquent and Prosocial Activities. *American journal of community psychology*, **2006**, *37* (1–2), 111–127.
60. Bellair, P. E. Social Interaction and Community Crime: Examining the Importance of Neighbor Networks. *Criminology*, **1997**, *35* (4), 677–704.
61. Browning, C. R.; Dietz, R. D.; Feinberg, S. L. The Paradox of Social Organization: Networks, Collective Efficacy, and Violent Crime in Urban Neighborhoods. *Social Forces*, **2004**, *83* (2), 503–534.

62. Gibson, C. L.; Zhao, J.; Lovrich, N. P.; Gaffney, M. J. Social Integration, Individual Perceptions of Collective Efficacy, and Fear of Crime in Three Cities. *Justice quarterly*, **2002**, *19* (3), 537–564.
63. Wells, L. E.; Weisheit, R. A. Patterns of Rural and Urban Crime: A County-Level Comparison. *Criminal Justice Review*, **2004**, *29* (1), 1–22.
64. Kaylen, M. T.; Pridemore, W. A. Social Disorganization and Crime in Rural Communities: The First Direct Test of the Systemic Model. *British Journal of Criminology*, **2013**, *53* (5), 905–923.
65. Bouffard, L. A.; Muftić, L. R. The" Rural Mystique": Social Disorganization and Violence beyond Urban Communities. *Western Criminology Review*, **2006**, *7* (3).
66. Li, Y.-Y. Social Structure and Informal Social Control in Rural Communities. **2011**.
67. Petee, T. A.; Kowalski, G. S. Modeling Rural Violent Crime Rates: A Test of Social Disorganization Theory. *Sociological Focus*, **1993**, *26* (1), 87–89.
68. Osgood, D. W.; Chambers, J. M. Social Disorganization Outside the Metropolis: An Analysis of Rural Youth Violence. *Criminology*, **2000**, *38* (1), 81–116.
69. Wells, L. E.; Weisheit, R. A. Explaining Crime in Metropolitan and Non-Metropolitan Communities. **2012**.
70. Barnett, C.; Mencken, F. C. Social Disorganization Theory and the Contextual Nature of Crime in Nonmetropolitan Counties. *Rural sociology*, **2002**, *67* (3), 372–393.
71. Osgood, D. W.; Chambers, J. M. Community Correlates of Rural Youth Violence. *Juvenile Justice Bulletin*, **2003**.
72. Ward, K. C.; Kirchner, E. E.; Thompson, A. J. Social Disorganization and Rural/Urban Crime Rates: A County Level Comparison of Contributing Factors. **2018**.
73. Kaylen, M.; Pridemore, W. A.; Roche, S. P. The Impact of Changing Demographic Composition on Aggravated Assault Victimization during the Great American Crime Decline: A Counterfactual Analysis of Rates in Urban, Suburban, and Rural Areas. *Criminal justice review*, **2017**, *42* (3), 291–314.
74. Kang, H.-W.; Kang, H.-B. Prediction of Crime Occurrence from Multi-Modal Data Using Deep Learning. *PLOS ONE*, **2017**, *12* (4), e0176244. https://doi.org/10.1371/journal.pone.0176244.
75. Huang, C.; Zhang, J.; Zheng, Y.; Chawla, N. V. DeepCrime: Attentive Hierarchical Recurrent Networks for Crime Prediction. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*; CIKM '18; Association for Computing Machinery: Torino, Italy, 2018; pp 1423–1432. https://doi.org/10.1145/3269206.3271793.
76. Marchant, R.; Haan, S.; Clancey, G.; Cripps, S. Applying Machine Learning to Criminology: Semi-Parametric Spatial-Demographic Bayesian Regression. *Security Informatics*, **2018**, *7* (1), 1. https://doi.org/10.1186/s13388-018-0030-x.
77. Ingilevich, V.; Ivanov, S. Crime Rate Prediction in the Urban Environment Using Social Factors. *Procedia Computer Science*, **2018**, *136*, 472–478. https://doi.org/10.1016/j.procs.2018.08.261.
78. US Census Bureau. 2014-2018 ACS 5-year Estimates https://www.census.gov/programs-surveys/acs/technical-documentation/table-and-geography-changes/2018/5-year.html (accessed Aug 18, 2020).
79. Fick, S. E.; Hijmans, R. J. WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas. *International Journal of Climatology*, **2017**, *37* (12), 4302–4315. https://doi.org/10.1002/joc.5086.
80. Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0--Fundamental Algorithms for Scientific Computing in Python. *Nat Methods*, **2020**, *17* (3), 261–272. https://doi.org/10.1038/s41592-019-0686-2.
81. Armitage, R.; Monchuk, L.; Rogerson, M. It Looks Good, but What Is It like to Live There? Exploring the Impact of Innovative Housing Design on Crime. *European Journal on Criminal Policy and Research*, **2011**, *17* (1), 29–54.
82. Mouatassim, Y.; Ezzahid, E. H. Poisson Regression and Zero-Inflated Poisson Regression: Application to Private Health Insurance Data. *Eur. Actuar. J.*, **2012**, *2* (2), 187–204. https://doi.org/10.1007/s13385-012-0056-2.
83. Fallah, N.; Gu, H.; Mohammad, K.; Seyyedsalehi, S. A.; Nourijelyani, K.; Eshraghian, M. R. Nonlinear Poisson Regression Using Neural Networks: A Simulation Study. *Neural Comput & Applic*, **2009**, *18* (8), 939. https://doi.org/10.1007/s00521-009-0277-8.
84. Friedman, J. H. Greedy Function Approximation: A Gradient Boosting Machine. *The Annals of Statistics*, **2001**, *29* (5), 1189–1232.

85. Zhang, Z. *Parameter Estimation Techniques: A Tutorial with Application to Conic Fitting*; Research Report RR-2676; INRIA: Sophia Antipolis, France, 1995; pp 59–76.
86. Breiman, L.; Friedman, J.; Stone, C. J.; Olshen, R. A. *Classification and Regression Trees*; Routledge & CRC Press, 1984.
87. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research*, **2011**, *12* (85), 2825–2830.
88. Bogomolov, A.; Lepri, B.; Staiano, J.; Oliver, N.; Pianesi, F.; Pentland, A. Once Upon a Crime: Towards Crime Prediction from Demographics and Mobile Data. In *Proceedings of the 16th International Conference on Multimodal Interaction*; ICMI '14; Association for Computing Machinery: New York, NY, USA, 2014; pp 427–434. https://doi.org/10.1145/2663204.2663254.
89. He, L.; Páez, A.; Jiao, J.; An, P.; Lu, C.; Mao, W.; Long, D. Ambient Population and Larceny-Theft: A Spatial Analysis Using Mobile Phone Data. *ISPRS International Journal of Geo-Information*, **2020**, *9* (6), 342. https://doi.org/10.3390/ijgi9060342.