EQUITY AND ACCURACY IN HOT-SPOT POLICING

Anupama Santhosh, Devashish Khulbe, Yavuz Sunor & Yuchen Ding
Center for Urban Science + Progress
New York University
{as11566, dk3596, ys3226, yd1402}@nyu.edu

July 24, 2019

ABSTRACT

Several recent studies have demonstrated the efficacy of proactive policing strategies for crime prevention. By predicting emerging geographic hot-spots of violent crime, we can target police patrols and other interventions. However, predictive policing creates moral and ethical concerns, such as fairness and equity, which have been well-documented, yet have not been typically incorporated into the design and evaluation of such systems. In this work we develop machine learning methods to predict hot-spots of crime and present a way to measure equity among those areas. We then adjust the predictions based on the defined equity metric and analyze the trade-offs between accuracy and equity. We also see the performance of the two models based on the racial distribution of the targeted population.

1 Introduction

1.1 Problem statement

Many urban areas are experiencing the lowest levels of violent crimes in decades [1]. Although the decline cannot be attributed to any policy in particular, studies indicate the efficiency of strategies such as predictive policing and hot-spot policing [1]. Predictive policing is the practice of forecasting crime patterns across time and space to inform decision-making for crime prevention. A major example is the identification of crime hot spots, which are micro areas (typically composed of one or more block-long street segments) that have crime densities much higher than the average in a city. Much research has been done on modeling hot spots and crime patterns in general and many cities have adopted these to aid decision making. Crime forecasting methods provide police agencies with data and tools to proactively intervene and reduce the aggregate number of crimes. Despite the positive outcomes of predictive policing, a general critique to these processes is the inherent bias and the inequity in policing. Accuracy and Equity in this context are opposing propositions and a trade-off is believed to be inevitable.

The work presented focus on the following key areas:
(1) Accuracy of predictive hot-spot methods.
(2) Methods to bring about more equity into predictive models.

In order to achieve this, we use open crime data and other publicly available datasets to develop and evaluate machine learning methods with respect to both accuracy and equity. Some of the key questions answered are the following:

• What are the best evaluation metrics for measuring accuracy and equity in predictive policing? Can simple equity metrics, measuring to what extent patrols are concentrated or dispersed across a city, be improved to account for neighborhood demographics, socio-economics, and crime victimization/underlying need for policing?
• Are there trade-offs between policing accuracy and equity, or is it possible to maximize both criteria simultaneously?

1.2 Literature review

While a lot of work has been focused on predictive policing [2, 3, 4, 5] in the past, we focus specifically on two broad classes of crime prediction: Temporal and Spatial. Temporal prediction models use the property of time dependence of crimes to successfully forecast time. Another work use time series model of ARIMA to make short-term forecasting of property crime for one city of China [2].

The social disorganization model focuses on how certain structural characteristics of neighborhoods lead to higher levels of crime and disorder [14]. Past literature in the field of criminology explores relationships between criminal activity and socio-economic variables such as education [17], ethnicity, income level [18]. To minimize the possibility of inducing bias in the model explicitly, racial/ethnic composition, education levels, income levels were not used as neighborhood indicators. In order to capture structural characteristics we looked at other options. Brantingham argues that crimes are the by-product of the built environment [15]. For e.g., shops, offices, government buildings, parks, or bus stops could be crime attractors or crime generators and the probability of crime incidence is higher near these structures. We chose number of police stations, parks/plazas, vacant buildings/plots, schools, health services in the spatial model.

Some literature has emerged around equity and fairness in algorithms in the recent years. Recent works propose new methods for fair classification in terms of demographic parity [6] & [7], which can be leveraged since a lot of discussion regarding fair policing is based on racial disparity and equality.
A paper in Annual Review of Criminology [8] suggest approaches for minimizing potential harm to vulnerable communities while providing an equitable distribution of the benefits of crime prevention across populations within police jurisdiction. Another work [9] proposes fairness by equalizing exposure to certain protected groups during learning. Also, there have been research on quantifying the cost of different types of crimes to the society [10]. This is informative of the fact that Part 1 violent crimes and thefts are most harmful to the society. Thus, we are most interested in predicting these type of crimes in our work.
We also consulted Dylan Fitzpatrick, who is a PhD. student at Carnegie Mellon University and worked extensively on predictive hot-spot policing in the recent past. Also, we have gained some useful insights on fair policing from Divam Jain, who is Principal Engineer at the Center for Policing Equity, an NYC-based organization that uses advanced analytics to diagnose disparities in policing, providing tools that law enforcement agencies can use to improve their relationships with the communities they serve.

2 Data Exploration

Through our prediction model we've aimed at forecasting the weekly crime numbers of Part 1 violent crimes in a city for a given geographical unit. The geographical unit was picked as census tract in this study. Part 1 crimes include murder and non-negligent homicide, rape, robbery, aggravated assault, burglary, motor vehicle theft, larceny-theft, and arson. Apart from these there are shootings and petty thefts. We chose Part 1 crimes as our target as, as per research studies the estimated cost to society of these crimes are the highest and patrolling is most effective for the same. At first, we have analyzed data from three cities for initial exploratory analyses and feature extraction, however eventually we ended up with New York City (NYC) for the modeling. NYC has the most comprehensive and detailed dataset in terms of historical crimes numbers and geographical features that were used in modeling, and also historical arrest numbers that were used in calculation of our equity metrics. Below we see the EDA (Explatory Data Analysis) results for NYC. That said, as we will see further, our equity measure can be applied to any city with sufficient proxies for measuring real crimes and police patrol.

2.1 New York City

The crime data for New York City is extracted from NYC OpenData using NYPD Complaint Dataset. This dataset includes all valid felony, misdemeanor, and violent crimes reported to the New York City Police Department (NYPD). For the analysis purpose, data from 2008 to 2017 (last 10 years) is used with time, crime type and location features. Crime complaints which involve multiple offenses are classified according to the most serious offense.
We also utilized the NYPD Arrest Incident Level Data for the equity metric. Each record represents an arrest effected in NYC by the NYPD and includes information about the type of crime, the location and time of enforcement. Arrests which involve multiple charges are classified according to the top charge. The footnotes along with the arrest data indicate possible errors in the geo-coding of the data. No corrections or adjustments were done.
There happen to be two possible bias in the study: First, the general statistical bias caused by the the fact that the exact crime numbers might not have reported or recorded properly. Second, the bias stemmed from the specific crime dataset which possibly reflects an inherent correlation between crimes and certain demographic characteristics (age, sex, education, poverty). All analyses presented assumes that the data used is the ground truth and no corrections or

adjusting for errors were undertaken.
For geospatial analysis, Facilities database and census tract shapefile provided by Department of City Planning are used. Census tract file updated in 2010 was used which may differ from earlier census tract boundaries. There happen to be two possible bias in the study: First, the general statistical bias caused by the the fact that the exact crime numbers might not have reported or recorded properly. Second, the bias stemmed from the specific crime dataset which possibly reflects an inherent correlation between crimes and certain demographic characteristics (age, sex, education, poverty). All analyses presented assumes that the data used is the ground truth and no corrections or adjusting for errors were undertaken.

For exploratory data analysis (EDA), we checked the general distribution of the number of Part1 crimes occurred in 2017 and looked whether we can see a temporal pattern in terms of both seasonal and weekly manners. When we look into days of the week, we see a pattern especially a difference between weekdays and weekends. Crimes like Assault, Murder and Rape were observed more on weekends, while Burglary and Robbery more on weekdays. New York has a kind of temporal crime patterns showing seasonality like other cities in our scope. One can infer that the crime numbers are increasing in Summer season for each year which is understandable because people spend more time outside and are more prone to be victimized. The crime distribution is also visualized spatially in census tract granularity. Here, we put the spatial distribution for 2017.
The corresponding graphs and results for all cities (NYC, Portland, Chicago) can be found in Appendix at the end of this report.

3 Methodology

3.1 Predictive Modeling

3.1.1 LSTM Model

We observe a clear periodicity for the all historical crime data-sets of three cities in data exploration. We thus decide to apply time-series analysis for our modeling. Considering LSTM has worked very well in similar scenarios (weather forecasts, stock market predictions etc.) and based on our experience with it in different projects, we focused specifically on LSTM.

For our first version of the model, we only used one layer with 32 units, and we picked our activation function as "ReLU". The data-set used in this model was the aggregated crime numbers of whole NYC based on last 10 years. The model worked pretty well in terms of a good fitted line and RMSE (We got an RMSE of 115 for the data with a mean of 2000 and std of 215). This inspired us to go on with alternative LSTM structures for more specific data-set rather than whole city's aggregated numbers (Figure 1).

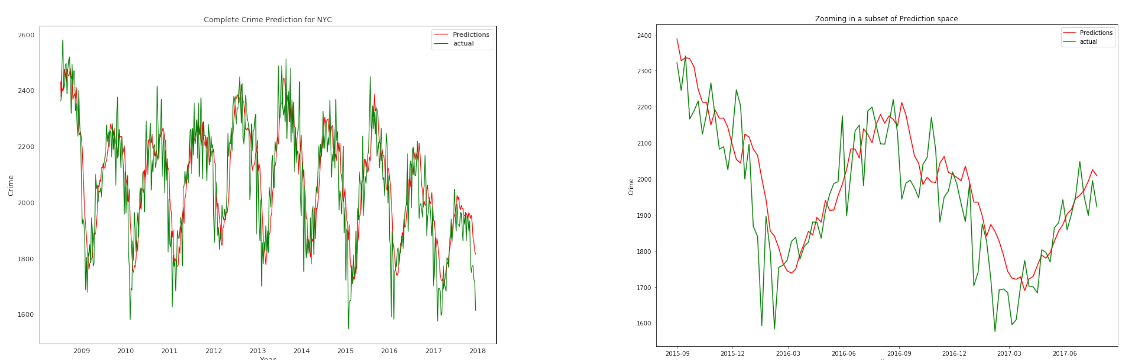

Figure 1: Complete prediction of whole NYC last 10 years and zooming for the last two years

Before delving into modeling part extensively, we notice that an aggregated prediction for whole census tracts would not make much sense in terms of individual tracts, as they don't have the same characteristics and trends. On the other hand, modeling them one by one or using a small set of them would also create problems due to data scarcity reasons. Thus, we came with the idea of applying time-series clustering for census-tracts, and applying separate models for each of those clusters. We have applied the elbow method to get the most optimized cluster numbers. Accordingly, the number of clusters has chosen as five for this analysis (Figure 2).

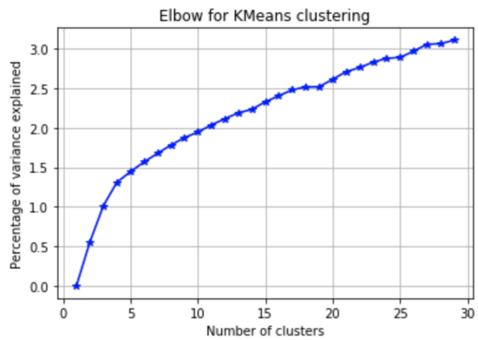

Figure 2: Explained Variance. The number of clusters chosen as 5

Afterwards, we have applied k-means time-series clustering on more than 2000 census tracts over the whole city. As a result, we obtained five different time-series each showed particular patterns. Here, some of the clusters start with high crime numbers and represent decreasing patterns, while others start with low numbers and represent increasing patterns.

We wanted to apply the same architecture for each cluster because of compatibility reasons and also that they have the very similar periodicity only with different trends. First, we tried the same architecture as we used for the whole aggregated data previously. However, it didn't result well and led us to search different architectures. At his point, we analyzed our cluster datasets in terms of their time horizons to decide on specific train-test splitting and looking back window-size. We picked 7-year(2009-2016) of train and 1.5-year(2016-2018) of test data with a 26 weeks window-size.

For a second version of LSTM modeling, we used three layers having 50 units each with three Dropouts(0.2) after each one. We haven't use a non-linear activation function this time, and changed the optimizer from 'ReLU' to 'RMSprop'. It worked better than the previous, but for some clusters(especially the ones don't follow a very distinctive pattern) didn't give the sufficient results.

Finally, for a third version, we simplified our architecture to only one LSTM layer having 100 units and two output Dense layers with two non-linearities, 'ReLU' and 'sigmoid' respectively. We also added a Dropout(0.5) in between LSTM and Dense layers. This model resulted the best, so we kept and used it as our final model. The losses can be seen below for five(5) different clusters(Figure 3).

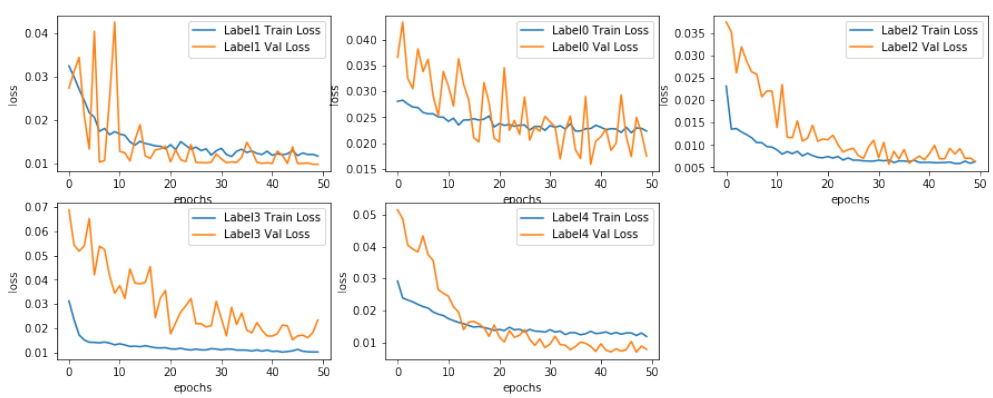

Figure 3: Train and validation losses for the five clusters

Limitations of LSTM modeling LSTM works well if there is a huge amount of sequential data. So, one limitation in this study is the historical crime data cannot be found for a period of more than 10 year. This is not short but also not very long for using 26 weeks look-back window. Another issue is, there are tens of different possible architectures in terms of layer/unit size, optimizer/activation function, Dropouts/batch-norm etc. On one hand, this provides many well-running modeling options; but on the other, finding the optimum model out of them is too difficult. And finally, training time and computational complexity increase dramatically with each new small touch to the model.

3.1.2 Gaussian Process Model

We use Gaussian Process (GP) as another model for prediction. For spatial data sets, GP is considered to be one of the most accurate models as it takes into account the spatial dependencies in the data. We also took into account the temporal dependencies by the past numbers as features for an area. For prediction of weekly crime numbers of a census tract, we consider the crime numbers of its neighbouring census tracts as features. This takes into account the criminal activity in surrounding areas of the concerned tract. We use the weekly data from 2008 to 2016 as training and 2017 as the test year. After training with cross validation, the model optimizes for Radial Basis Function (RBF) kernel with average length of 5 times steps (weeks), the Exponential Kernel with average periodicity of 52 and the White Noise kernel with length 0.1. For the 10 years of prediction (2008-2017), the model obtains RMSE value of 334.

Limitations of GP modeling One of the limitations for GP is data sparsity, which is prevalent for weekly crime numbers. Another possible issue with GP modeling is the long training time for the data, and given that we predict for around each of the 2100 census tracts.

3.1.3 Random Forest Model

We chose a Random Forest regressor (RF) as it has shown good prediction quality over high-dimensional heterogeneous feature spaces. Due to their non-parametric nature, they make no assumption about the data and can work with many, correlated features, while also requiring little preparation of the data [19]. In the random forest model we used both geographic features temporal features.

Geographic features: Traditional methods use structured data like population, race/ethnicity, income levels but we decided to not induce any bias explicitly. Inspired by the Broken Windows theory which infers that untended property attracts vandals, even among people who would not ordinarily consider committing crimes[21], and Crime Prevention through Environmental Design [22, 23] which suggests that appropriate environmental design can increase the perceived likelihood of detection and apprehension, the single most deterrent to crime, we used as predictors geographical information like existence of Parks or vacant plots/buildings, Police stations, Transportation/Health facilities. Violence has been associated with residential instability of neighborhoods[20]. These features describes the characteristics of a neighborhood and can be seen as crime attractors or crime deterrents. We obtained information about these predictors from the Facilities database provided by the Department of Planning. Vacant Plots/Buildings suggest amount of disordering in a neighborhood while other predictors suggest orderliness in census tracts. These features are measured by the counts of the mentioned venues in the census tracts.

Temporal features: We aggregated past crime data for each census tract. We had years worth of data but decided to group it as crimes that happened in the past week, two weeks, one month, three months, one year, three years and five years. Weekly predictions were made for one year and the results analyzed. Internally, the regressors optimize the mean squared error (MSE) of the dependent variable. RMSE value for the Random Forest model is 34.

Limitations of RF Modeling process Census tract is a rather large unit for police intervention but in terms of data, offers more stability than other geographical units. Random Forests models offer less interpretability but in general performs better.As with any other supervised model, the causal effects cannot be inferred.

Model Comparison Comparing the LSTM, GP and the RF models, we notice that RF performs better in terms of RMSE value. Also looking at the plot of number of crimes with the top predicted census tracts (Figure 4, 50 top hotspots shown for clarity), the RF predictions are better and are close to the true number of crimes in the city. It can also be inferred from the graph that RF predictions captures more crimes as we intervene on top crime hotspots. Also worth noticing is the training time for the models, for which the RF model takes lowest time among the three.

Model used	RMSE value
LSTM	115
Gaussian Process	334
Random Forests	34

Table 1: RMSE values for 2017 predictions

Since RF performs better in predictions, we use it further for the predictions with equity metric.

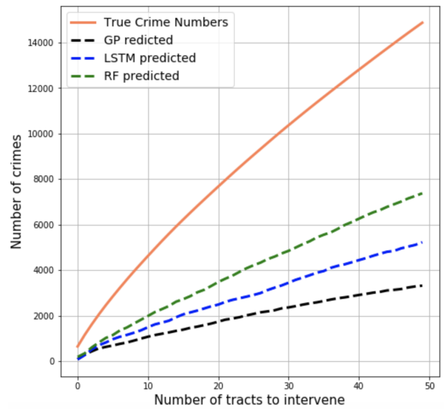

Figure 4: Crimes captured with different models as we intervene on top hotspots

3.2 Equity Measure

Equity in our analysis is based on the premise that all members of the community, regardless of their race, ethnicity or economic status receives the same policing services. We have defined equity in terms of policing metric that balances the demand and supply of policing in a neighborhood. We have defined the metric as a proportion of estimated need of policing to the actual deployment on ground. Both these metrics cannot be directly measured but can be indirectly assessed through proxy measures. As a proxy for demand for policing, we consider the metric "Level of disorder" in neighborhoods measured by Vacant Buildings/Plots, 311 complaints, Garbage disposal problems etc. But from further discussions we decided to measure the Part 1 Violent Crimes as a proxy for policing demand with the assumption that all violent crimes are captured by the police department and gets reflected in the crime data.To assess the actual deployment of police forces we considered using drug/narcotics related crimes but historical data suggests that some locations might be more prone to drug related crimes. Instead we chose total number of arrests within a census tract as proxy considering that most of these instances happen during police patrolling.

$$PolicingRatio = \frac{PolicingDemand}{PolicingSupply} = \frac{\sum ViolentCrimes}{\sum Arrests}$$

Number of Part 1 Violent Crimes and Number of Arrests made in the past one year (2016 for our data) for each census tract. Both proxies were standardized and normalized for comparison. The metric can be used as a measure of over-policing or under-policing with a value of more than one indicating under-policing.

4 Results

We present our results in three subsections. The first section shows our best (Random Forest) model's predictive accuracy in terms of a fitted line and the lift curve (our intervening accuracy) without taking into account the equity metric. Second shows how the lift curve changes after incorporating the equity measure. And finally in the third section, we look the racial distribution of the hot-spot areas before and after equity measure to get a sense of how our equity measure works in terms of fairness.

4.1 Predictions without Equity measure

Here, we are interpreting our predicted crime numbers using two different ways. First is the fitted curve for the corresponding time-series in test data (2017). Second is the lift curve which depicts a comparison between the number of captured crimes by actual and predicted Top-50 census tracts. The lift curve is developed by putting the actual crime numbers of hot-spot (Top50) areas and adding predicted crime numbers for the same hot-spot areas if there is a match with actual ones. The comparison is made by weekly, but the numbers aggregated yearly afterwards. The curve plays a role representing the potential police intervening accuracy of our suggestion.

As shown in the Figure 5, the predicted crime numbers for the whole tracts give us a good fit. On the other hand, the lift curve says it's not easy to predict the exact hot-spot areas for each separate week even though the numbers can be predicted. Still, the model does a good job to guess around half of the tracts(with the assumption of getting the numbers well).

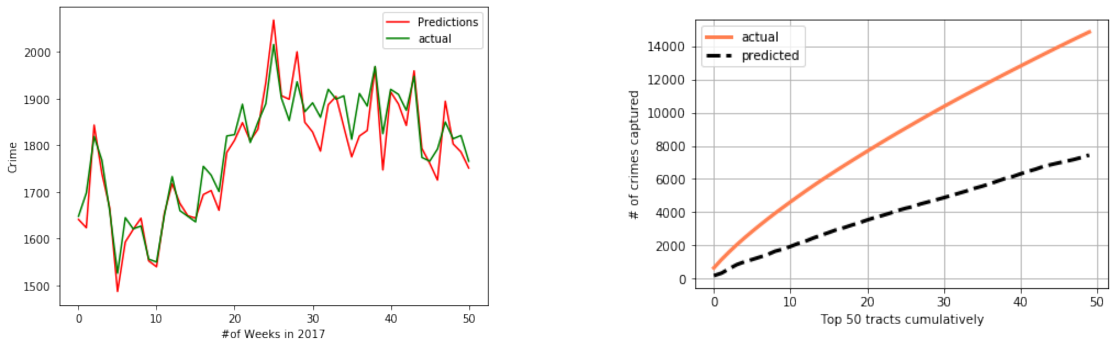

Figure 5: Actual vs. pred crimes in 2017 (left) and number of crimes captured for Top tracts (right)

4.2 Predictions with Equity measure

Here, we are providing a comparison between our lift curves before using equity measure and after using it. We see in Figure 6 that the black dashed curve is the same curve above represents the sole result of the predictive model. And the green curve represents the result of our model with equity measure. We get this updated curve simply by multiplying the predicted numbers with our equity measure which is a sign for over/under-policing. The reason the updated curve lifted and got closer to the actual curve is that we've increased the predicted crime numbers by multiplying them with the equity measure. So, we can infer that the general pattern for hot-spot areas matched by our models tend to be under-policed(policing ratio > 1).

Finally, we wanted to see the Top50 hot-spot areas in 2017 on a map with a comparison among actual areas, sole predictions and predictions after equity measure. Figure 7 shows the hot-spot areas do not change much from one map to other as we compare real and prediction without equity. We see a concentration around midtown and downtown areas, also some around northern part of the city and near JFK airport. However, it can be seen that some hot-spots change after we adjust for equity metric.

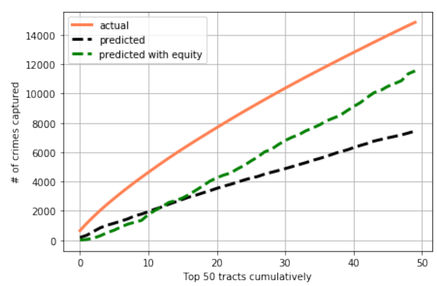

Figure 6: Comparison of lift curves among real, predicted and predicted with equity

Figure 7: Spatial distribution of top 50 hot-spots with true values (left), predicted (middle) and after adjusting for equity metric (right)

4.3 Racial distribution of predictions

We compared the predictions of the RF models (with and without incorporation of equity measure) with respect to the demographics they are targeting. We plot the distribution of races as we intervene on the top hotspots measured by the prediction results from the two models (Figure 8). We observe that for the unconstrained model (without equity), the top (around 10) hotspots have higher percentage of black people and lower proportion of whites among them.

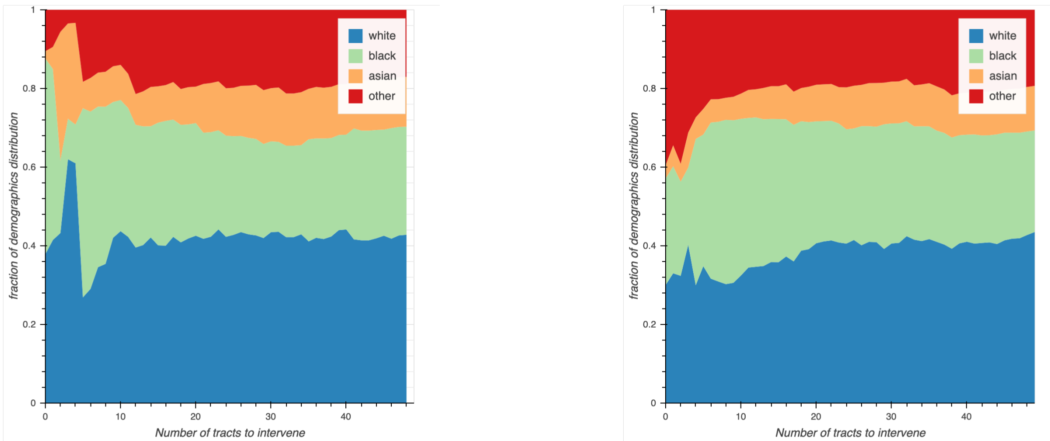

Figure 8: Demographics distribution of RF model with (right) and without (left) equity measure

Looking at the top 50 tracts, there is clearly a larger percentage of white people. Whereas for the predictions after adjusting for equity metric, the top 10 hotspots see a decrease in black population and overall it can be seen that the proportion of blacks is more compared to the unconstrained model.

5 Conclusion

5.1 Policy Implications

Through our work, we predicted weekly numbers of Part 1 violent crimes at a census tract level. We explored three different methodologies for crime prediction: An LSTM model incorporating the seasonality of crimes, a Gaussian Process model taking into account spatial correlation and a Random Forest model which uses the geographic features of the neighborhood as predictors. Our results show that the Random Forest model performs better in capturing more crimes by intervening specific tracts. The better performance of the Random Forest model corroborates the Broken Windows theory and the concept of Crime Prevention through Environmental Design (CPTED). Our novel approach of implementing LSTMs is motivated by the distinctive periodicity in crime patterns for each cities we analyzed.

An extension of our weekly predictions can be used to assess short term risk and aid in deployment of patrolling. Our model specifically would tell top census tracts to intervene in a given week. The number of such census tracts to be intervened can be decided based on the available resources. The efficacy of these results would depend on choice of intervention and ground deployment. Solutions currently in use utilize information on recent crime and does not incorporate geographic features of the area expanding the scope of our model to street crimes.

5.2 Discussion on Modelling

The models we chose deals well with heterogeneous data sources. By using tree-based models the need to pre-process the features to remove multicollinearity can be avoided. In future works, more geographic features can be added to the model. Even though we haven't directly used any indicators of protected groups (like race/ethnicity, gender),associations with historical crime numbers can bring in bias. The model has achieved better predictive accuracy, but the underlying causal effect cannot be deduced from the analysis. So the work only gives empirical evidence that incorporating geographical features gives higher accuracy.

5.3 Discussion on Equity

We have defined equity in a novel way using the policing ratio. We have compared the predictions of our machine learning models based on accuracy, equity and targeted racial distribution. In terms of accuracy, the predictions adjusted

for equity metric capture more crime as compared to unconstrained predictions. This is a little surprising as one would expect unconstrained model to be more accurate as it focus on maximizing accuracy. The better performance of adjusted predictions may be as a result of two reasons: 1) The top captured hot-spots change after we adjust for equity and/or 2) The crime numbers in the same captured are increased by some amount.
Looking at the racial distribution of the two models, we can infer that the equity metric does not necessarily account for racial parity for policing as except for the first few top hot-spots, the racial distribution does not change across models. We can't conclusively comment on the accuracy-equity trade-off from our results without further investigations and more robust equity metric.

5.4 Future Work

Further investigations are needed to validate our approach and empirical results and the robustness of the predictors used. There is still room for improvement for guessing the correct hot-spot sets for a given week and implementation of the measure. Also the models can be enhanced using dynamic data sources to capture movements within the city. More spatial features of the geography can be incorporated. Finally, an interactive dashboard with weekly predictions and top hot-spots to intervene can directly aid in decision making.

References

[1] Robin S. Engel, John E. Eck Effectiveness v. Equity in Policing: Is a Tradeoff Inevitable? In Ideas in American Policing; 2015.

[2] Chen P, Yuan H, Shu X. Forecasting Crime Using the ARIMA Model In Proceedings of the 5th IEEE International Conference on Fuzzy Systems and Knowledge Discovery. vol. 5; 2008. p. 627–630.

[3] Gorr W, Olligschlaeger A, Thompson Y. Short-term forecasting of crime In International Journal of Forecasting. 2003;19(4):579–594.

[4] Wang P, Mathieu R, Ke J, Cai HJ. Predicting Criminal Recidivism with Support Vector Machine. In Proceedings of the 2010 IEEE International Conference on Management and Service Science; 2010. p. 1–9.

[5] Mohler GO, Short MB, Brantingham PJ, Schoenberg FP, Tita GE. Self-Exciting Point Process Modeling of Crime. In Journal of the American Statistical Association, 2011;106(493):100–108.

[6] Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez Rodriguez, Krishna P. Gummadi. Fairness Constraints: Mechanisms for Fair Classification In *Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on*, pages 417–422. IEEE, 2014.

[7] Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez Rodriguez, Krishna P. Gummadi. Fairness Beyond Disparate Treatment & Disparate Impact: Learning Classification without Disparate Mistreatment In 20th International Conference on Artificial Intelligence and Statistics (AISTATS), Fort Lauderdale, FL, April 2017.

[8] Dylan J. Fitzpatrick, Wilpen L. Gorr and Daniel B. Neill Keeping Score: Predictive Analytics in Policing In Annual Review of Criminology 2019

[9] Ashudeep Singh, Thorsten Joachims Fairness of Exposure in Rankings In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, London, UK, 2018

[10] Kathryn E. McCollister, Michael T. French, Hai Fang The cost of crime to society: New crime-specific estimates for policy and program evaluation Drug and Alcohol Dependence Volume 108, Issues 1–2, 1 April 2010, Pages 98-109

[11] Chammah, Maurice, and Mark Hansen. Policing the Future. The Marshall Project, The Marshall Project, 3 Feb 2016, www.themarshallproject.org/2016/02/03/policing-the-future.

[12] Muhammad Bilal Zafar, Isabel Valera, Manuel Gomez Rodriguez, Krishna P. Gummadi, Adrian Weller. From Parity to Preference-based Notions of Fairness in Classification 31st Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, December 2017.

[13] Lingxue Zhu, Nikolay Laptev. Deep and Confident Prediction for Time Series at Uber. arXiv:1709.01907v1 [stat.ML] 6 Sep 2017, https://arxiv.org/pdf/1709.01907.pdf

[14] John R. Hipp. Block, Tract, and Levels of Aggregation: Neighborhood Structure and Crime and Disorder as a Case in Point In American Sociological Review Vol. 72, No. 5 (Oct., 2007), pp. 659-680 (22 pages)

[15] Brantingham, P.P., Brantingham, P.P. Criminality of place. European Journal on Criminal Policy and Research 3, 5–26 (1995)

[16] Lauritsen JL, Rezey ML, Heimer K. Violence and Economic Conditions in the United States, 1973-2011: Gender, Race, and Ethnicity Patterns in the National Crime Victimization Survey. In Journal of Contemporary Criminal Justice. 2014;30(1):7–28.

[17] I. Ehrlich. On the relation between education and crime. In Education, Income, and Human Behavior, NBER 1975;(p. 313 - 338).

[18] E. B. Patterson. Poverty, income inequality, and community crime rates. In Criminology, 29(4):755–776, 1991.

[19] Kadar, Cristina, and Irena Pletikosa. Mining large-scale human mobility data for long-term crime prediction. In EPJ Data Science, 7.1 (2018): 26.

[20] Sampson, R.J., Rauenbush, S.W., Earls, F. Neighborhoods and violent crime: A multilevel study of collective efficacy In Science 277, 918–924 (1997)

[21] Wilson JQ, Kelling GL. Broken windows. Critical Issues in Policing: Contemporary Readings. In NCJRS, 1982; p. 395–407.

[22] Casteel C, Peek-Asa C. Effectiveness of crime prevention through environmental design (CPTED) in reducing robberies. In American Journal of Preventive Medicine. 2000;18(4):99–115. pmid:10793286

[23] Cozens PM, Saville G, Hillier D. Crime prevention through environmental design (CPTED): a review and modern bibliography In Property Management. 2005;23(5):328–356.

A Appendix

A.1 Team Collaboration

Anupama - Analyzed and conducted exploratory analysis for the city of Chicago. Prepared data and implemented the Random Forest modeling. Conducted literature review and collaborated on writing the report. Worked on implementing the website.

Devashish - Analyzed and conducted exploratory analysis for the city of Portland and New York. Prepared data and implemented the Gaussian Processes modeling. Visualized the results and collaborated on writing the report.

Yavuz - Analyzed and conducted exploratory analysis for the city of New York. Prepared data and implemented the LSTM modeling. Came up with alternative methods to present results in an analytical and interpretable way. Collaborated on writing the report.

Yuchen - Studied on the web-page with relevant details of the project.

A.2 EDA results

A.2.1 EDA results for Chicago

The crime data for the city of Chicago is extracted from the Chicago Police Department's CLEAR(Citizen Law Enforcement Analysis and Reporting) system. For the analysis we used filtered the data of all reported crimes from the period Jan 2014 to present. The dataset has information on where(x,y coordinates, latitude, longitude), when(date), and which(primary type, description) crime was reported.
Yearly crime statistics follow a general pattern. There's a rise in the trend at the beginning of the year and peak during June/July and then a rapid drop.
The patterns of each of the Part 1 crimes were looked at with a rolling sum graph. Although the general trend is that crimes are decreasing, crimes like sexual assault or assault otherwise and theft is going up. By numbers the most occurring crimes are theft and damage.
It can be seen that crimes follow a temporal pattern by the time of the day. For the modeling we are ignoring the temporal variations of crime types and focusing more on the historical crime rates and other factors such as geographic places of interest(aggregate number of bars, police stations, vacant plots etc.)

A.2.2 EDA results for Portland

For Portland, Oregon, we checked the crime data provided by the National Institute of Justice (NIJ) which is available for the 2012 to 2017. There are 99 different types of case descriptions based on the type of offences. On a more high level, the data is categorized into 4 classes - 'Burglary', 'Motor Vehicle Theft', 'Street Crimes' and 'Others'. Relevant features from the data include day of the crime, location coordinates and the census tract in which the crime took place.
Initial data exploration included filtering out the Part 1 crimes from the data which are to be used as target variables. Considering data for 3 years (2012-2015), these amount for 29.45% of the total crimes. We aggregated the relevant offences monthly to get an idea about the yearly cycles and overall trend for the period.
We can see a cyclical component where the number reaches a low during the winter months and high during the summer months. Also the general trend is slightly increasing as time progresses from 2013 to end of 2015. Next, to extract temporal features, we create a simple heatmap of standardized number of crimes for each crime type seen against the day of the week. Here we observe that certain types of offences are prevalent during the weekends and other during the weekdays.
To get a sense of the major hotspots of part 1 crimes, we aggregated the offences on census tract level for all three years and visualized through a map. Also, in order to see if there is any weekly trends, we visualized aggregating the crimes weekly for each month, but no significant trends were seen.

A.2.3 EDA graphs for all cities

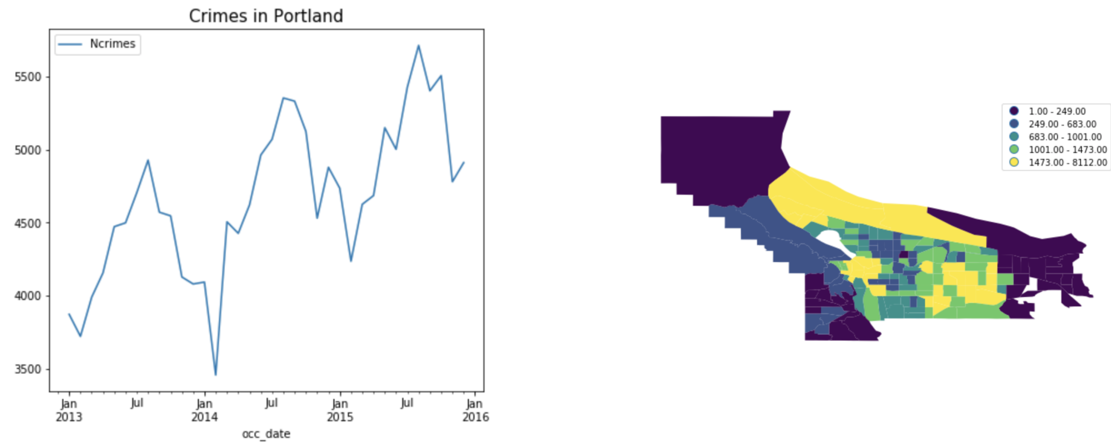

Figure 9: Yearly trend of Part 1 crimes in Portland and their Spatial distribution for 2017

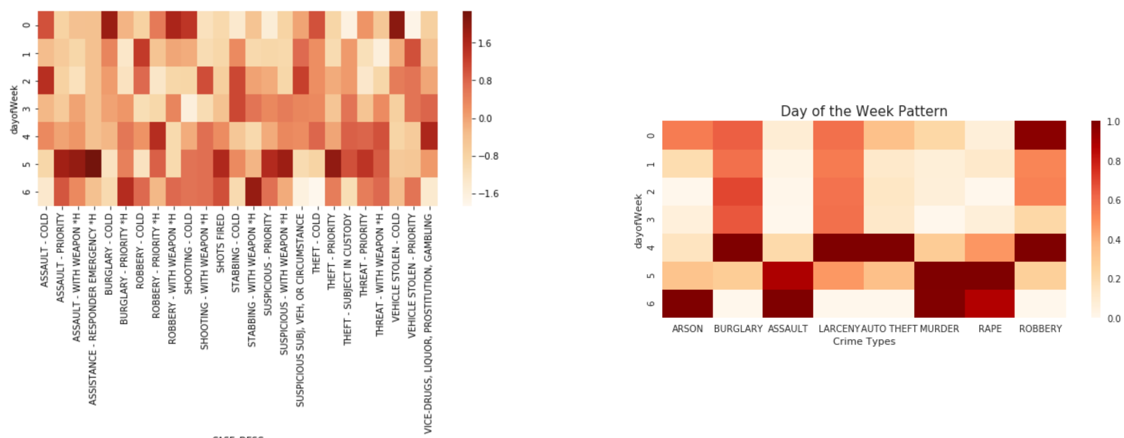

Figure 10: Day of week patterns of Crimes

Figure 11: Time series of Part 1 crimes in New York and their Spatial distribution for 2018

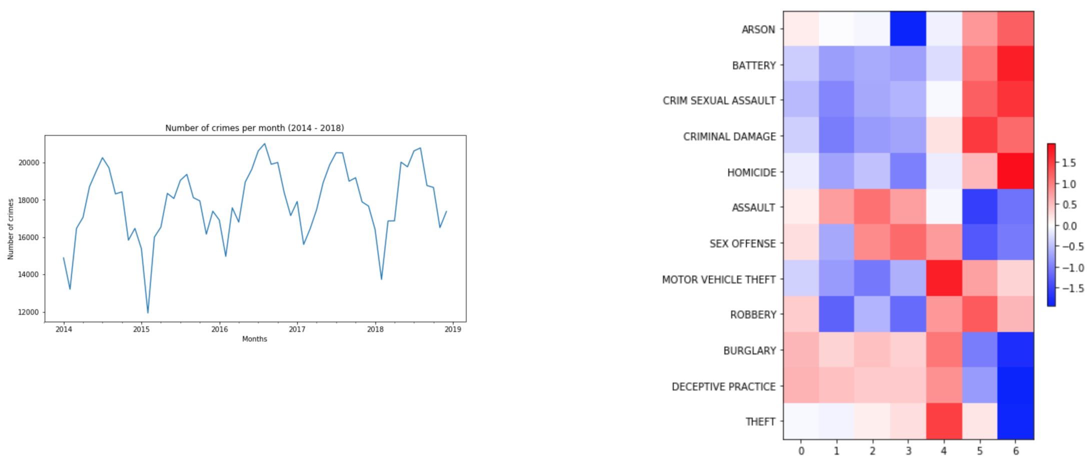

Figure 12: Time series of Part 1 crimes in Chicago and their Weekly distributions