ETHICAL AND ALGORITHMIC CONSIDERATIONS FOR BIG DATA DECISION-MAKING: PREDICTIVE POLICING CASE STUDY

A Capstone Research Project

Department of Cognitive, Linguistic, & Psychological Sciences

Brown University

In Partial Fulfillment

of the Requirements for Behavioral Decision Sciences

By Camille Fougere

Specialization in Automated Decision-Making in Machines

December, 2020

Ethical and Algorithmic Considerations for Big Data Decision-Making: Predictive Policing Case Study

As governments, businesses, and public policy decision-makers increasingly rely on data science, the ramifications of algorithmic bias and unethical applications of big data increase. This paper reviews considerations for applying big data solutions to decision-making policies through the case study of predictive policing, using a Holt-Winters predictive policing model in conjunction with L.A. crime data from January, 2012, to illustrate the prevalent biases in modeling crime. The O'Neil definition classifies predictive policing as unethical, and the modeling biases suggest that predictive policing is inherently fallible. As such, I conclude that predictive policing is currently an unsuitable application of big data decision-making.

1. INTRODUCTION

Within the age of digital transformation, more of society is outsourcing decision-making to technology, including government agencies, insurance companies, college-ranking publications, and components of our justice system.[1] The stakes of this societal shift are high for the people whose well-being depends on the outcomes of algorithmic decisions. And when biases are present in the data or algorithms used, models can perpetuate systemic inequality.

Predictive policing as an application of big data is extremely controversial relative to the concepts of justice and ethics. Proponents will argue that police are able to more effectively focus their resources. In some scenarios, this directly translates to an increase in crimes stopped and a decrease in crime rates.[2] Furthermore, allowing algorithms to decide what neighborhoods to target purports to exclude the influence of any racial or socioeconomic biases within the police force itself.

Critics of predictive policing argue that these biases are still present. Even if the model is relying on location data, due to the racial stratification of neighborhoods in cities, location can effectively serve as a proxy for race. Though predictive policing software was originally intended to focus on violent, serious crimes, many algorithms were expanded to victimless crimes, as violent crime is difficult to predict. In doing so, those algorithms tend to disproportionately target low-income communities where victimless crimes can be endemic. Lastly, predictive policing software can create self-fulfilling prophecies. If the model predicts higher crime in a given neighborhood, and police are dispatched, it increases the odds of crime being reported there. This creates feedback reinforcing that neighborhood as most problematic.[3] The technology may seem impressive, while severely underperforming and legitimizing inherently biased policies.

This paper first analyzes the ethical context of using big data for policing. I introduce and apply Cathy O'Neil's framework for analyzing the ethics of big data applications. This paper then illustrates potential sources for contextual bias in modeling crime. I identify how crime report data can be susceptible to training data bias, reporting bias, and interpretation bias. To illustrate algorithmic bias, I introduce a model to predict how much crime will occur in each ZIP

---

[1] O'Neill, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York, Penguin Random House LLC.

[2] G. O. Mohler, M. B. Short, Sean Malinowski, Mark Johnson, G. E. Tita, Andrea L. Bertozzi & P. J. Brantingham. (2015). Randomized Controlled Field Trials of Predictive Policing, *Journal of the American Statistical Association*, 110:512, 1399-1411, DOI: 10.1080/01621459.2015.1077710

[3] O'Neill, C., *Weapons of Math Destruction*

code of L.A. during January 2012. I analyze the model's output itself and identify potential sources for algorithmic focus bias and runaway feedback loops. Lastly, I address the limitations of this paper and suggest next steps for related research.

2. ETHICAL CONSIDERATIONS

In Cathy O'Neil's book *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy,* she provides a useful framework for ethically contextualizing different uses of big data. This framework suggests that a problematic application, or a Weapon of Math Destruction (WMD), has three characteristics: it is not transparent, it is easily scaled outside of its intended use, and it has the potential to create significant harm. Predictive policing can be analyzed across these three features. Firstly, many predictive policing software companies, such as PredPol, refuse to release their data or the full version of their code to the public (though they have published one version due to widespread demand).[4] This creates a lack of transparency between the forces that use this technology and the people directly affected by its suggestions. Secondly, predictive policing algorithms are easily scalable; as discussed previously, predictive policing software commissioned to predict only violent crime is now frequently being used to predict nuisance crime. Lastly, predictive policing algorithms have a great potential to create harm in the communities in which they are implemented. Areas predicted to have more crime are therefore more policed, and people living there are more likely to be arrested or enter the prison system.[5] As such, it is appropriate to conclude that predictive policing can be qualified as an unethical application.

3. CONTEXTUAL BIAS SOURCES

Before delving into algorithmic bias, it is important to note contextual biases related to crime data and model interpretation. Training data bias occurs when the data used to train a model is inconsistent with the data the model hopes to predict.[6] Police report data is a notoriously inaccurate representation of underlying crime patterns. One factor of this is reporting bias. Certain demographic groups are more likely to report crime than others. Rates of report are correlated with an individual's trust in police, which vary greatly across racial, gender, and socioeconomic groups.[7] Furthermore, police report data represents what crimes were discovered or reported, not what crimes occurred. Using police report data to train a model will serve primarily to make a model excellent at predicting future police reports.

Secondly, interpretation bias may play a role in predictive policing. Interpretation bias happens when the users of a model misinterpret the purpose or functioning of the model[8]; in this case, if police forces interpret the output of a predictive policing model to be where crimes are likely to occur, instead of where crimes are likely to be reported.

---

[4] Lum, K. (2016). Predictive Policing: Bias In, Bias Out. *Data and Society.* https://datasociety.net/library/predictive-policing-bias-in-bias-out/
[5] O'Neill, C., *Weapons of Math Destruction*
[6] Danks, David & London, Alex. (2017). Algorithmic Bias in Autonomous Systems. 4691-4697. 10.24963/ijcai.2017/654.
[7] Lum, K., *Predictive Policing: Bias In, Bias Out*
[8] Danks, David & London, Alex, *Algorithmic Bias in Autonomous Systems,* 4691-4697

4. METHODOLOGY

1.1 DATA:

The data selected for this model is Los Angeles crime data for January, 2012. I selected Los Angeles as an appropriate city to analyze because PredPol has been implemented there since 2008. The year 2012 falls within the timeframe of PredPol's implementation and is accessible on L.A.'s Open Data site[9].

1.2 MODEL:

To select a model to create predictions on a rolling basis, I analyzed the different components of the crimes reported over January (see Figure 1). Because there was a noticeable trend and seasonality, I implemented a Holt-Winters model with root mean squared error (RMSE) hovering around 1.6 (see Figure 2).

Figure 1

	Date	Prediction RMSE
0	2012-01-03	1.427834
1	2012-01-04	1.616448
2	2012-01-05	1.441326
3	2012-01-06	1.515511
4	2012-01-07	1.393533
5	2012-01-08	1.681014
6	2012-01-09	1.713325
7	2012-01-10	1.663653
8	2012-01-11	1.472325
9	2012-01-12	1.752418
10	2012-01-13	1.515511
11	2012-01-14	1.456908
12	2012-01-15	1.212835
13	2012-01-16	1.746887
14	2012-01-17	1.331989
15	2012-01-18	1.949359
16	2012-01-19	1.962553
17	2012-01-20	1.372542
18	2012-01-21	1.231312
19	2012-01-22	1.365473
20	2012-01-23	1.730187
21	2012-01-24	1.553352
22	2012-01-25	1.690581
23	2012-01-26	1.450250
24	2012-01-27	1.530338
25	2012-01-28	1.545023
26	2012-01-29	1.183216
27	2012-01-30	1.540842
28	2012-01-31	1.526117

[9] LAPD Open Data. (2020). Crime Data From 2010 to 2019. Retrieved from https://data.lacity.org/Public-Safety/Crime-Data-from-2010-to-2019/63jg-8b9z

Figure 2

5. RESULTS

1.1 ANALYZING ALGORITHMIC FOCUS BIAS:

As introduced before, there are often discrepancies between the geographic makeup of Part 1 crimes and Part 2 crimes. Part 1, or violent crimes, include murder, arson, and assault. Part 2 are victimless, or "nuisance" crimes, including vagrancy, minor drug offenses, and panhandling. To analyze the relationship between the two types, I ran a correlation test on the total crime in each ZIP code in L.A. in January, 2012 (see Figure 3).

```
CORRELATION FOR # PART 1 and # PART 2 CRIMES REPORTED:
__________________________________________________________
Correlation Coefficient btw # PART 1 and # PART 2 CRIMES REPORTED: 0.9375599920574694
P-Value for Correlation btw # PART 1 and # PART 2 CRIMES REPORTED: 3.344385327092054e-76
Correlation btw # PART 1 and # PART 2 CRIMES REPORTED: Very high significant
```

Figure 3

There is a strong positive correlation between the amount of Part 1 crimes and the amount of Part 2 crimes in a ZIP code. However, I wanted to analyze outliers to this correlation, ZIP codes with a proportionally high number of Part 1 crimes and a proportionally low number of Part 2 crimes and vice versa (Figure 4). These outliers could be useful in testing the theory that Part 1 and Part 2 crimes occur in different demographics. I used a GeoJSON from L.A. Times[10] to visualize the quantity of Part 1 and Part 2 crimes reported for January.

	ZIPCODE	# PART 1 CRIMES REPORTED	# PART 2 CRIMES REPORTED		PROPORTIONAL DIFF
0	91401	182.0	170.0		1.094595
1	91331	198.0	179.0		1.030146
2	90057	117.0	118.0	High Part 2, Low Part 1	0.939189
3	90008	122.0	120.0		0.897089
4	91402	101.0	103.0		0.841476
...	...	...	...		...
159	90731	165.0	91.0		-0.713617
160	90045	113.0	52.0		-0.767672
161	90037	180.0	95.0	High Part 1, Low Part 2	-0.893971
162	90064	161.0	71.0		-1.177235
163	91367	108.0	25.0		-1.401247

164 rows × 4 columns

Figure 4

[10] Los Angeles Times. (2012). ZIP Code Tabulation Areas (2012). Retrieved from http://boundaries.latimes.com/set/zip-code-tabulation-areas-2012/

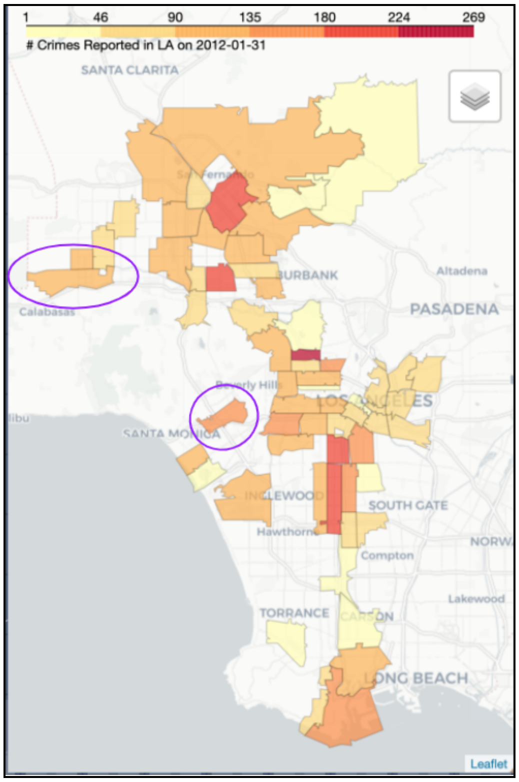

Figure 5 (Part 1 Crimes)

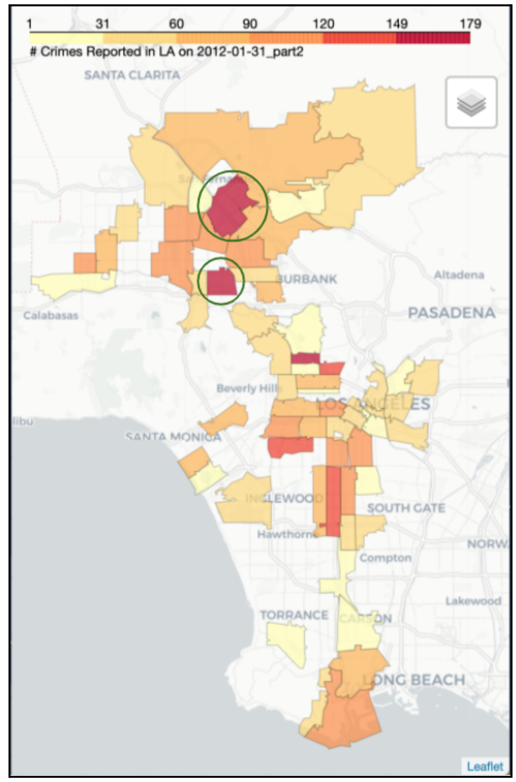

Figure 6 (Part 2 Crimes)

The two top ZIP codes with proportionally high Part 1 crimes but proportionally low Part 2 crimes are circled in Figure 5. They correspond to Rancho Park and Woodland Hills. The median household income in Rancho Park in 2015 was $84,573[11], and in Woodland Hills, was $76,548[12]. Rancho Park is 61% white[13], and Woodland Hills is 68% white[14]. By contrast, the top two ZIP codes with proportionally low Part 1 crimes and proportionally high Part 2 crimes and proportionally high Part 2 crimes, circled in Figure 6, correspond to Van Nuys and Pacoima. The median household income in Van Nuys in 2015 was $46,420[15] and in Pacoima was $51,477[16]. In addition, Van Nuys is 46% Black and 43% Latinx, and Pacoima is 88% Latinx.[17] This supports the argument that even though location-based predictive policing intends to avoid racial and socioeconomic bias, including Part 2 data may disproportionately target low-income neighborhoods with primarily Black and Latinx residents.

---

[11] United States Zip Codes.org. (2015). ZIP Code 90064. Retrieved from https://www.unitedstateszipcodes.org/90064/

[12] United States Zip Codes.org. (2015). ZIP Code 91367. Retrieved from https://www.unitedstateszipcodes.org/91367/

[13] Los Angeles Almanac. City of Los Angeles Population by Zip Code & Race, 2010 Census. Retrieved from http://www.laalmanac.com/population/po24la_zip.php

[14] Los Angeles Almanac, *City of Los Angeles Population by Zip Code & Race, 2010 Census*

[15] United States Zip Codes.org. (2015). ZIP Code 91401. Retrieved from https://www.unitedstateszipcodes.org/91401/

[16] United States Zip Codes.org. (2015) ZIP Code 91331. Retrieved from https://www.unitedstateszipcodes.org/91331/

[17] Los Angeles Almanac, *City of Los Angeles Population by Zip Code & Race, 2010 Census*

1.2 ANALYZING FEEDBACK LOOPS:

In order to illustrate how feedback loops can be created by predictive policing, I simulated what would happen when more police forces are sent to particular ZIP codes. To do so, I drew from Kristian Lum and William Isaac's research and modeling project *To Predict and Serve?*, which ran a similar simulation on different data/geography and also tried to represent increased reports in crime in the regions police were targeting[18]. I ran the model week by week to make predictions on which ZIP codes would have the most crime. In the control trial, the only data used was historical crime reports (see Figure 7). In another trial, I selected the ZIP codes predicted to have the most crime in the following week, and multiplied the amount of crimes that were actually reported by 20% as suggested by Lum and Isaac (to represent a higher police presence leading to more discovered crimes) (see Figure 8). To understand how these two models differed, I analyzed the distribution of crimes predicted across the ZIP codes. Figure 9 shows the distribution of crimes predicted over ZIP code based only on actual reports, and Figure 10 shows the distribution of crimes predicted with adding in simulated "discovered" crimes where applicable. Both Figure 9 and Figure 10 also include relevant statistics to show the tendency and spread of these two distributions.

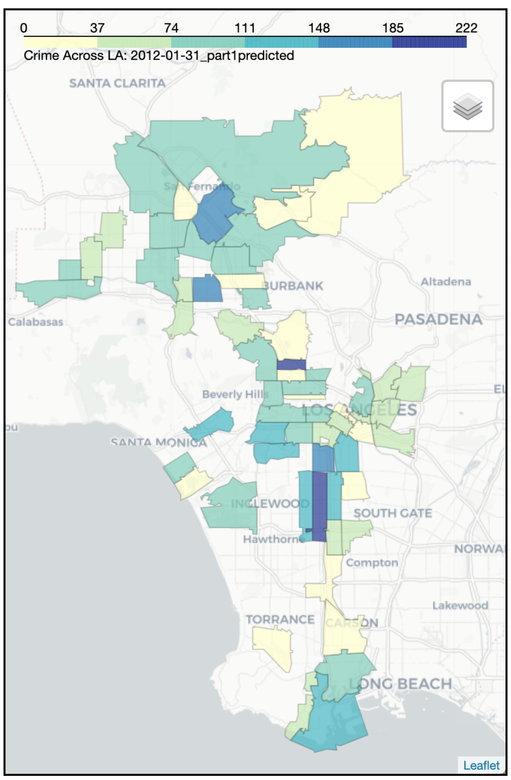

Figure 7 (using only actual reports)

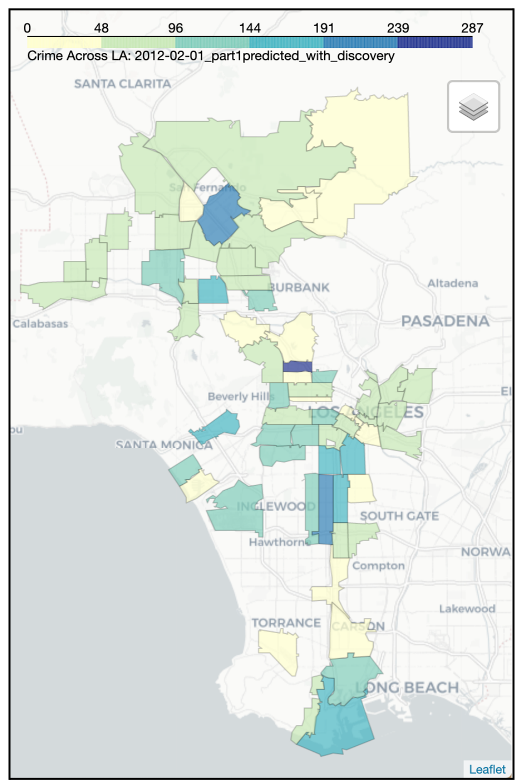

Figure 8 (with simulated discovery report)

[18] Lum, K., & Isaac, W. (2016). To predict and serve? *Significance*, 13(5), 14– 19.

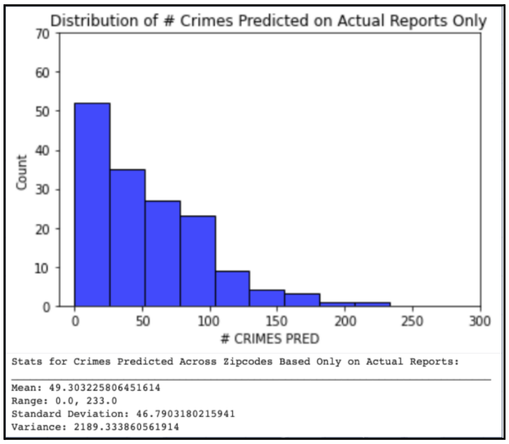

Figure 9 (using only actual reports)

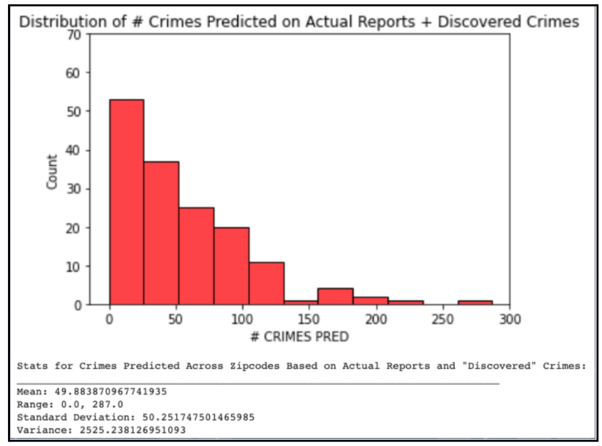

Figure 10 (with simulated discovery reports)

Figure 9 and Figure 10 illustrate the change in distribution of crimes predicted across ZIP codes when discovery is introduced and feedback loops reinforce that particular ZIP codes have more crime. After including simulated discovery data, the model predicted a significant increase for the number of crimes predicted for the top ZIP codes (from 222 to 287 crimes). In addition, the distribution after simulating discovery is more spread out, and suggests that the model ended up focusing disproportionately on few locations and was susceptible to bias from reinforcement loops.

6. DISCUSSION

There are a few key limitations in this research to highlight. Firstly, I built a model based off of the guidelines of a sliding window, smoothed averaging type. Companies such as PredPol are likely to have more complicated algorithms that would've been beyond the scope of this project to recreate. Secondly, many predictive policing models divide geography by 150 by 150 sq. ft. sections, whereas this model relies on ZIP codes. Lastly, the police report data is derived from the digitization of written police reports, and there could be inaccuracies due to translation.

This ethical and model-based analysis illustrated why predictive policing can be a flawed application of big data. In addition, I clarified which specific biases could play a role in both the context and implementation of a predictive policing algorithm, and how those biases could affect the well-being of people living in the policed communities. However, there is a promising trend away from predictive policing in the recent years across the U.S., primarily due to activism and legal work of civil rights and civil liberties organizations[19]. In November 2019, the LAPD reported that it would be changing its use of the technology due to an inability to determine how effective PredPol was at reducing crime, and in April of this year, they stopped using PredPol. My hope is that this simulation provides a framework for contextualizing the use of big data decision-making across society. Further work could use different versions of predictive algorithms (such as directly recreating PredPol[20] algorithm) or different cities/timelines to assess the prevalence of these patterns.

---

[19] Lau, T. (2020). Predictive Policing Explained. *Brennan Center for Justice.* https://www.brennancenter.org/our-work/research-reports/predictive-policing-explained

[20] Lum, K., & Isaac, W., *To predict and serve?,* 14– 19.

BIBLIOGRAPHY:

Danks, David & London, Alex. (2017). Algorithmic Bias in Autonomous Systems. 4691-4697. 10.24963/ijcai.2017/654.

G. O. Mohler, M. B. Short, Sean Malinowski, Mark Johnson, G. E. Tita, Andrea L. Bertozzi & P. J. Brantingham. (2015). Randomized Controlled Field Trials of Predictive Policing, *Journal of the American Statistical Association*, 110:512, 1399-1411, DOI: 10.1080/01621459.2015.1077710

Lau, T. (2020). Predictive Policing Explained. *Brennan Center for Justice.* https://www.brennancenter.org/our-work/research-reports/predictive-policing-explained

LAPD Open Data. (2020). Crime Data From 2010 to 2019. Retrieved from https://data.lacity.org/Public-Safety/Crime-Data-from-2010-to-2019/63jg-8b9z

Los Angeles Almanac. City of Los Angeles Population by Zip Code & Race, 2010 Census. Retrieved from http://www.laalmanac.com/population/po24la_zip.php

Los Angeles Times. (2012). ZIP Code Tabulation Areas (2012). Retrieved from http://boundaries.latimes.com/set/zip-code-tabulation-areas-2012/

Lum, K., & Isaac, W. (2016). To predict and serve? *Significance*, 13(5), 14– 19.

Lum, K. (2016). Predictive Policing: Bias In, Bias Out. *Data and Society.* https://datasociety.net/library/predictive-policing-bias-in-bias-out

O'Neill, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York, *Penguin Random House LLC*.

United States Zip Codes.org. (2015). ZIP Code 90064. Retrieved from https://www.unitedstateszipcodes.org/90064/

United States Zip Codes.org. (2015). ZIP Code 91367. Retrieved from https://www.unitedstateszipcodes.org/91367/

United States Zip Codes.org. (2015) ZIP Code 91331. Retrieved from https://www.unitedstateszipcodes.org/91331/

United States Zip Codes.org. (2015). ZIP Code 91401. Retrieved from https://www.unitedstateszipcodes.org/91401/