Fairness, Accountability, & ransparency A few (Counter)-Examples from predictive models in Criminal Justice

Kristian Lum Department of Computer and Information Science, University of Pennsylvania

Fairness Predictive Policing

Unfair policing practices of the passed through a model to cause unfair practices in the future.

What is predictive policing?

Predictive policing uses police records to learn patterns in the occurrence of crime. Using these patterns, the computer then predicts the most likely locations of future crimes.

police records

where crime will be detected in the future.

Additional police are then dispatched to the locations with the highest predicted rate of crime, thus preventing the crime that would have occurred from occurring or catching criminals in the act.

finding crime you wouldn't have found otherwise?

Drug Crimes in Oakland



Data provided by openoakland.org

Drug Crimes in Oakland





Drug Crimes in Oakland

2010 Oakland PD Drug Crime Reports Albany Lafayette Siesta Valley Orinda NORTHWEST University of Recreation Berkeley 🗢 Cali (123) Area BERKELE' Ashby A (13) Siblev lcanic Regional reserve Fastpor Moraga Canyon Records Piedmont Valle Vista Redwood 200 gional Park 150 100 50 Alameda 0 Island Lincoln Av Alameda O.co Coliseum (61) Luther Bay Farm .Ir Rea Shore San Leandro kland mational Google Airport Man data ©2016 Goode



Image Copyright, 2013, Weldon Cooper Center for Public Service, Rector and Visitors of the University of Virginia (Dustin A. Cable, creator)

Demo on Oakland Data

other

other



Lum, Kristian, and William Isaac. "To predict and serve?." Significance 13.5 (2016): 14-19.

Accountability

Risk Assessment

Lack of accountability allows for "overbooking", which can have large effects on the outputs of pre-trial risk assessment models

Pre-Trial Risk Assessment



Decision-Making Framework

| | NCA 1 | NCA 2 | NCA 3 | NCA 4 | NCA 5 | NCA 6 |
|----------|----------|-----------------|-----------------|----------------------------|----------------------------|----------------------------|
| FTA 1 | OR - NAS | OR - NAS | | | | |
| FTA 2 | OR - NAS | OR - NAS | OR - NAS | OR - Minimum | SFPDP - ACM | |
| FTA 3 | | OR - NAS | OR - Minimum | SFPDP - ACM | SFPDP - ACM | Release Not Recommended |
| FTA 4 | | OR - Minimum | SFPDP - ACM | SFPDP - ACM | Release Not Recommended | Release Not Recommended |
| FTA | | SFPDP - ACM | SFPDP - ACM | SFPDP - ACM* | Release Not Recommended | Release Not Recommended |
| | / | 7.01 | 1.011 | Recommended | | |
| FTA 6 | | | | Release Not Recommended | Release Not Recommended | Release Not Recommended |

* Release Not Recommended if any booked offense is a felony or violent misdemeanor per PSA Violent Offenses List; SFPDP – ACM if booked offense(s) are non-violent misdemeanors.

The Plan



Results

| Charges | exclusions | bump-ups | nvca | rec |
|------------|------------|----------|--------|-------|
| Conviction | 8.7 | 9.1 | 9.3 | 2.5 |
| Booking | 29.4 | 23.7 | 20.2 | 3 |
| Difference | 20.7 * | 14.6 * | 10.8 * | 0.5 * |

Table 1: Percent of cases with exclusions, bump-ups, nvca flags, and the average recommendation by input charges.

| exclusions | bump-ups | nvca | rec |
|------------|----------|------|------|
| 20.9 | 17.0 | 10.9 | 27.4 |

Table 2: Percent of cases for which each PSA component was higher under the booking charges than under the conviction charges.