



## BIASED DATA IN MACHINE LEARNING:

### USING POLICE RECORDS TO TRAIN PREDICTIVE POLICING MODELS

---

**Kristian Lum**      HRDAG Lead Statistician  
**Patrick Ball**      HRDAG Director of Research  
**William Isaac**      PhD Candidate, Department of Political Science, MSU

CARE conference  
August 6, 2016

# Machine learning models / algorithms

logistic regression

k-means clustering

random forest

support vector machines

kernel density estimation

reinforcement learning

neural nets

nearest neighbors

generalized linear model

boosting

deep learning

ensemble models

principal components analysis

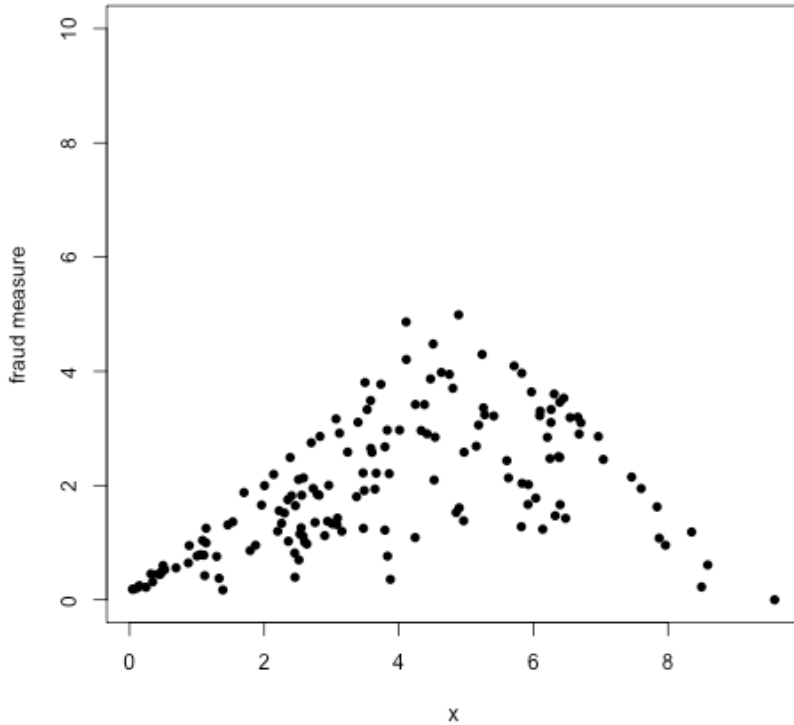
Machine learning models / algorithms

Ways to learn  
patterns &  
structure in data

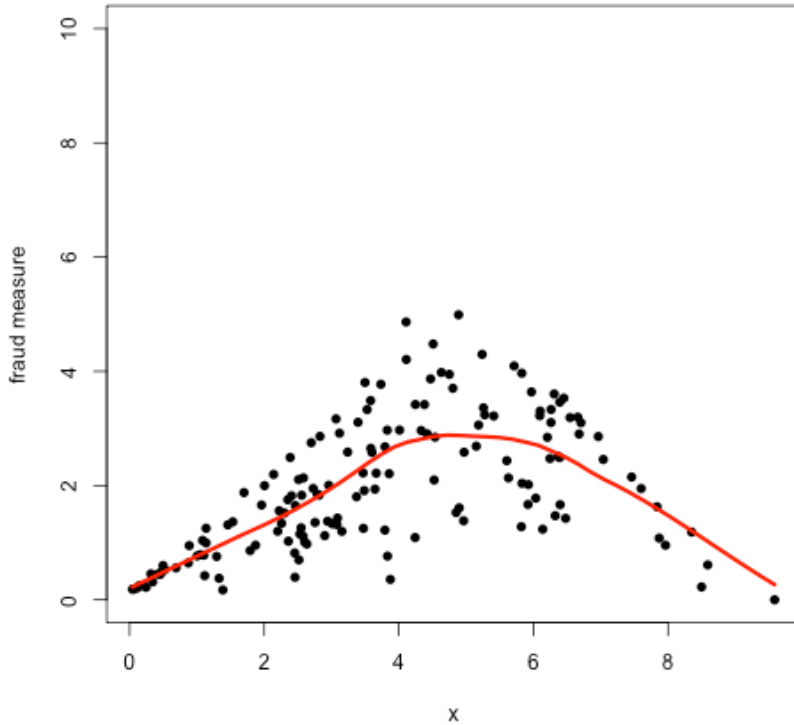
The word cloud includes the following terms: logistic regression, random forest, k-means clustering, support vector machines, kernel density estimation, reinforcement learning, neural nets, nearest neighbors, generalized linear model, boosting, deep learning, ensemble models, and principal components analysis.

\*\*Which model you pick defines *how* you learn the patterns and the types of patterns we can learn.

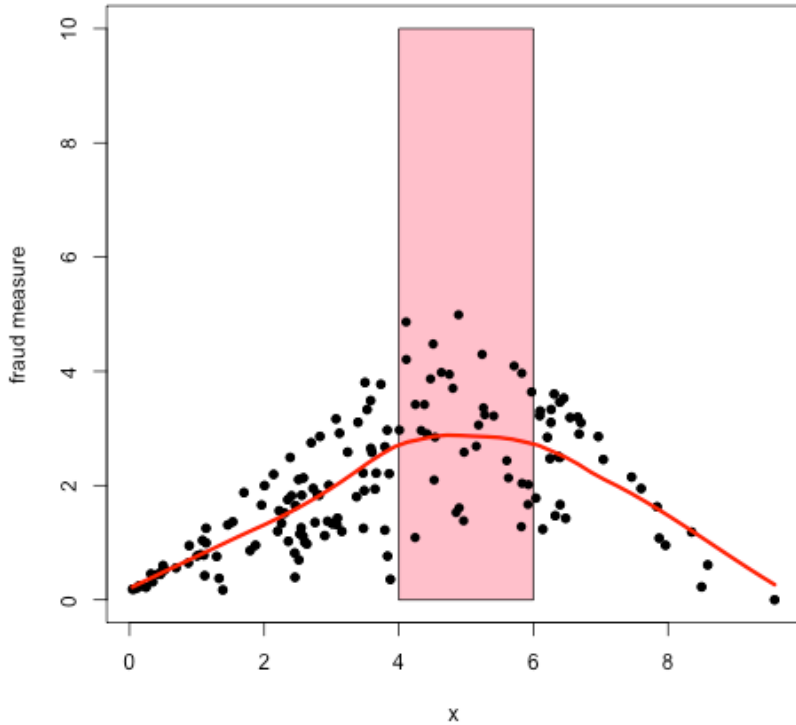
# Machine learning example



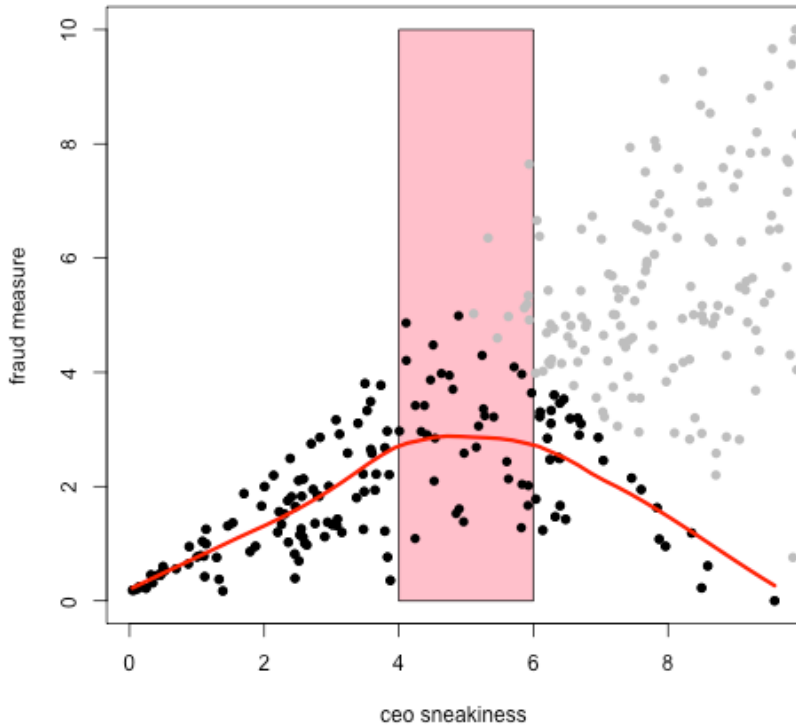
# Machine learning example



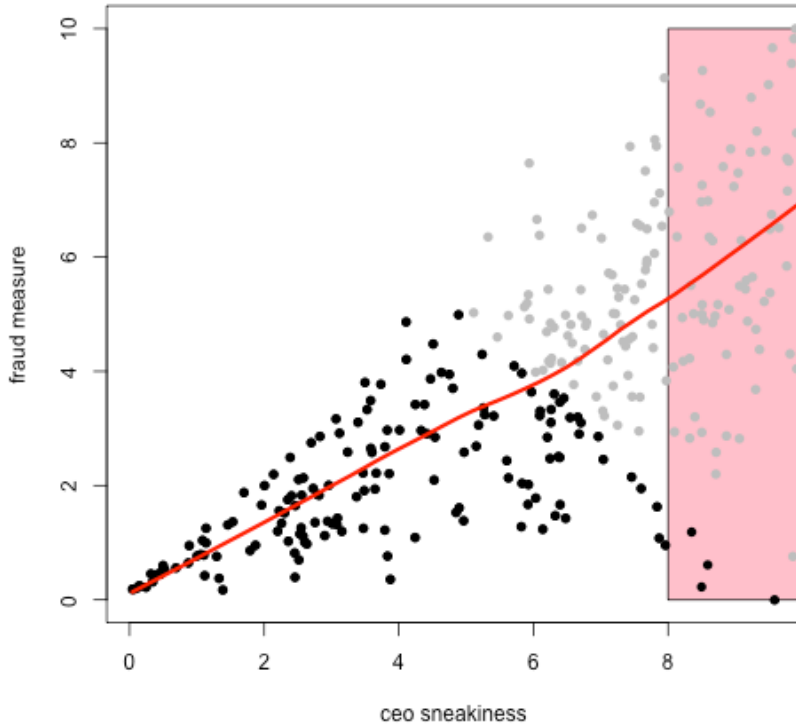
# Machine learning example



# Machine learning example



# Machine learning example





# What is your data measuring?

- Unless you have a complete census or a random sample (with known sampling weights) from the population of things you are trying to measure...
  - Data collected about Z is **not** a measurement of Z; It is a measurement of a complex interaction between Z and your ability to observe or record Z.
- Examples:
  - Data on reported fraud is measuring the interaction between a firm's committing fraud and {the willingness of employees at that firm to be whistleblowers, regulatory agencies' institutional priorities in the year fraud was committed, etc.}

# Predictions

- Predictions using a model trained on  $Z$  are predictions about the future value of  $Z$ :
- In this case:
  - Models trained on observed fraud data make predictions about where fraud will be *observed* in the future.
  - Models will help you find the types of fraud you would have found anyway.

# What is Predictive Policing

Predictive policing uses police records to learn patterns in the occurrence of crime.

# What is Predictive Policing

Predictive policing uses police records to learn patterns in the **occurrence of crime.**

**police records**

# What is Predictive Policing

Predictive policing uses police records to learn patterns in the **occurrence of crime.**

**police records**

Using these patterns, the computer then predicts the most likely locations of future crimes.

# What is Predictive Policing

Predictive policing uses police records to learn patterns in the **occurrence of crime.**

**police records**

Using these patterns, the computer then predicts the most likely locations **of future crimes.**

**where crime will  
be detected in the future.**

# What is Predictive Policing

Predictive policing uses police records to learn patterns in the **occurrence of crime.**

**police records**

Using these patterns, the computer then predicts the most likely locations **of future crimes.**

**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.

# What is Predictive Policing?

Predictive policing uses police records to learn patterns in the ~~occurrence of crime.~~

police records

Using these patterns, the computer then predicts the most likely locations ~~of future crimes.~~

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?



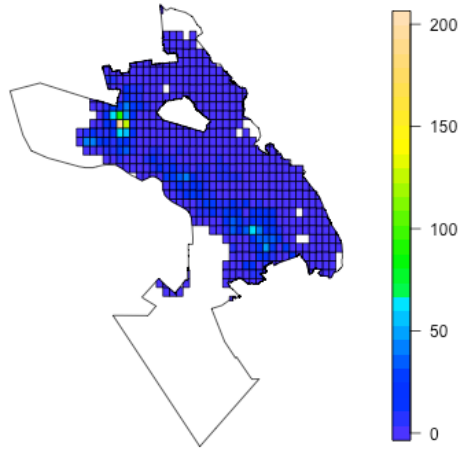
# Why do we think police records are not representative?

- Decades (maybe centuries) of criminological research
- Data from the National Crime Victimization Survey (NCVS) indicates that reporting rates vary substantially by demographic group, meaning that some crimes are more likely than others to be reported to police based upon who was victimized.
- Crimes that are committed in areas that are highly patrolled by police are more likely to be discovered by police than those committed in less patrolled areas.
- While white and black populations use marijuana at similar rates, blacks are arrested for marijuana possession at a rate several times that of whites.\*

\*<https://www.washingtonpost.com/news/wonk/wp/2013/06/04/the-blackwhite-marijuana-arrest-gap-in-nine-charts/>

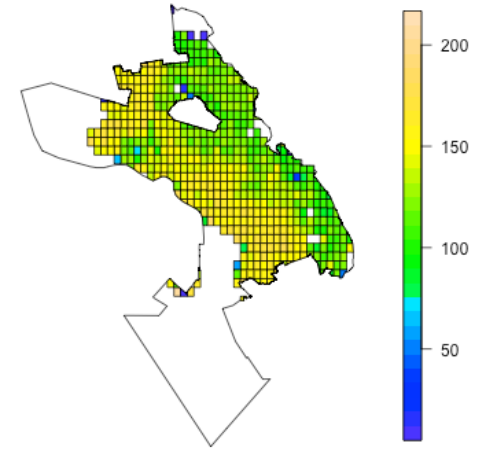
# Drug Crimes in Oakland, CA

Drug Crimes In Police Database



All drug crimes in Oakland PD  
data from 2009-2011

Drug Use Rates Per 1K People



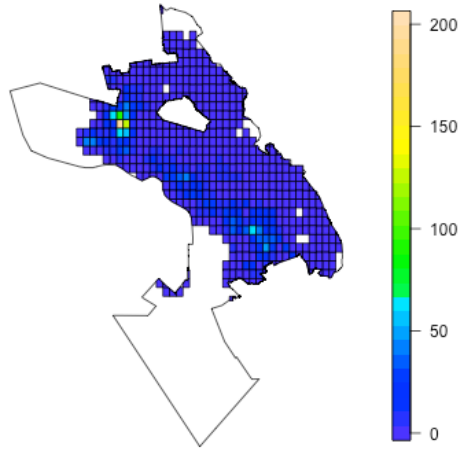
Estimate of drug use based on  
NSDUH [1] and detailed US  
Census data [2]

[1] United States. Office of Applied Studies. Substance Abuse and Mental Health Archives. National Survey on Drug Use and Health.

[2] 2010 RTI U.S. Synthetic Population Ver. 1.0 RTI International. May, 2014. URL: <https://www.epimodels.org/midas/pubsyntdata1.do>

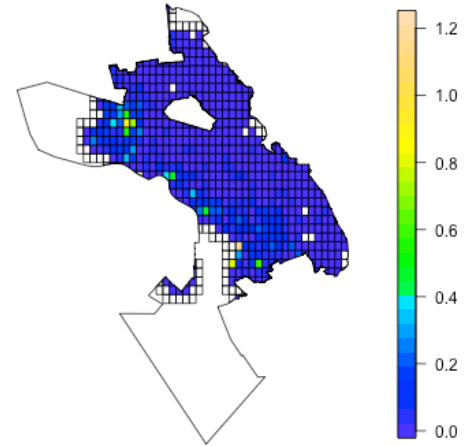
# Drug Crimes in Oakland, CA

Drug Crimes In Police Database



All drug crimes in Oakland PD  
data from 2009-2011

Drug Arrests to Drug Users Ratio



Ratio

[1] United States. Office of Applied Studies. Substance Abuse and Mental Health Archives. National Survey on Drug Use and Health.

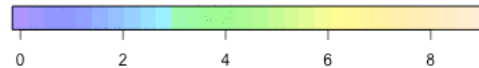
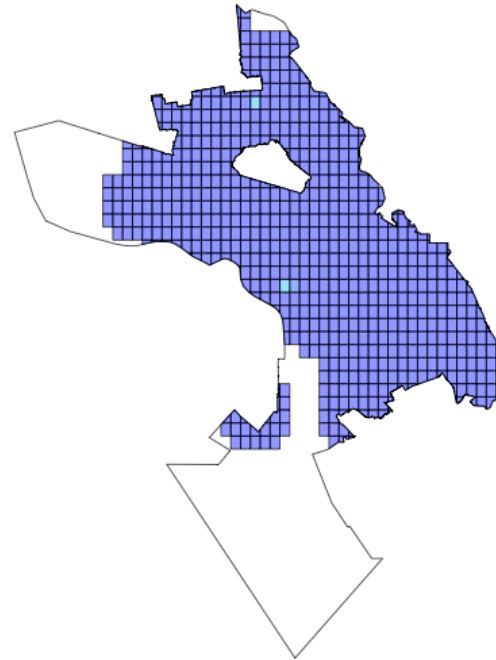
[2] 2010 RTI U.S. Synthetic Population Ver. 1.0 RTI International. May, 2014. URL: <https://www.epimodels.org/midas/pubsyntdata1.do>

# Applying predictive policing

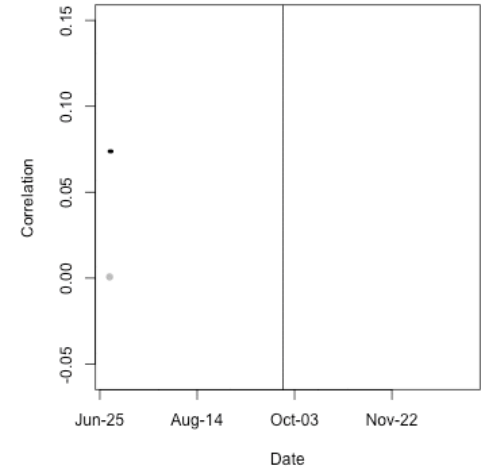
- The algorithm directs additional policing to the over-policed areas.
- Biases that previously existed in the police database are confirmed.
- Minority and low-income communities receive a disproportionate amount of targeted policing.

Targeted

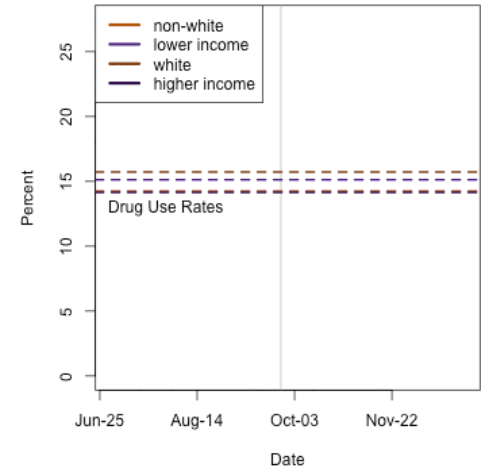
2011-Jun-30



Correlation Between % Low Income & Arrest Rate



% Experiencing Targeted Policing

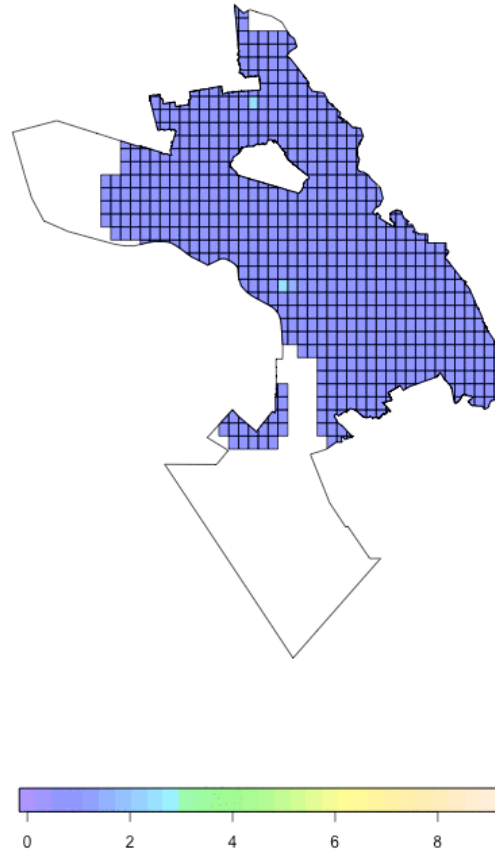


Targeted

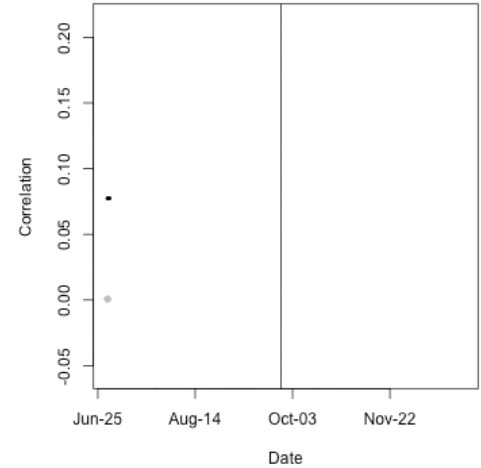
2011-Jun-30

# Simulating an alternative scenario

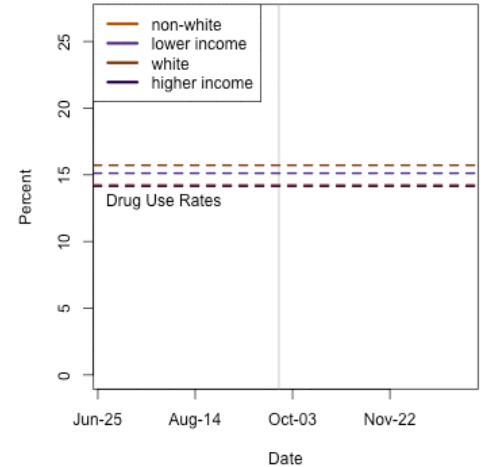
- We assume that sending targeted policing to an area results in a slight increase in the number of drug crimes discovered by police.
- Biases in the data are *amplified*.
- Minority and low-income communities receive a disproportionate amount of targeted policing.



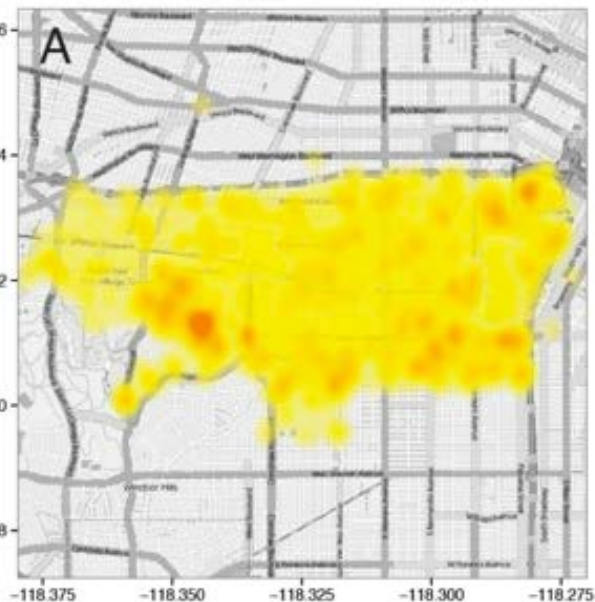
Correlation Between % Low Income & Arrest Rate



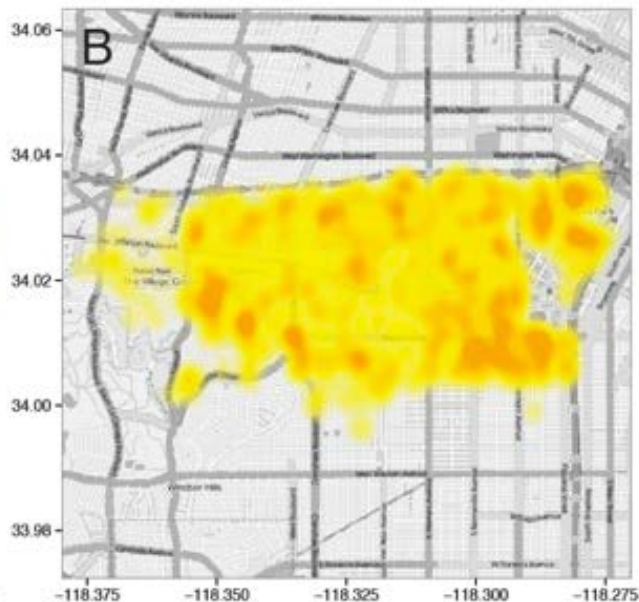
% Experiencing Targeted Policing



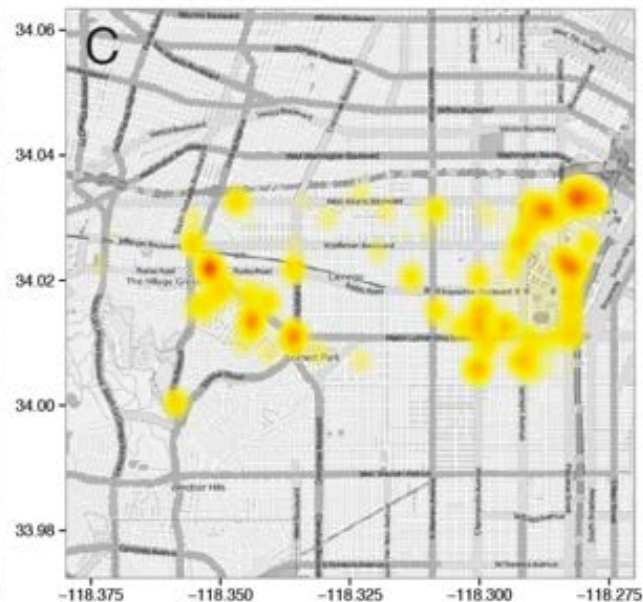
# A real example



training data



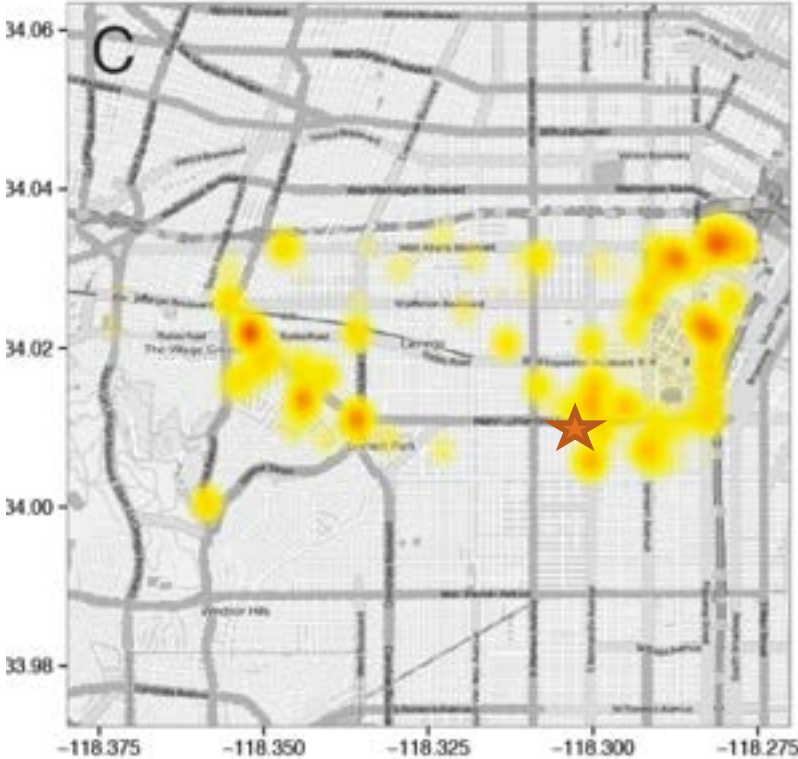
human predictions



machine learning predictions

Figure from Mohler et. al, JASA 2015

# A real example



machine learning predictions

# Summary

- Crime is everywhere, and police data on crime is not representative of all crime.
- Biased data leads to biased predictions in any machine learning algorithm— predictive policing included.
- Predictive policing reproduces the biases that already exist in the police data in a much more precise and targeted way.
- If increased targeted policing leads to an increase in the number of arrests, targeted policing will magnify pre-existing biases.
- Because of the biases in the existing police data, predictive policing will disproportionately affect historically over-policed communities.
- There is no way to assess whether this is happening using police data— needs human judgment to find such issues.



Thanks!

[kl@hrdag.org](mailto:kl@hrdag.org)

@klDivergence