CSE 410

May 7, 2020

Last class!

We made it ☺

3rd round of proof reading graded

note @99

0 views

3rd proof reading grading done

Grading of the 3rd round of proof reading is all done. Feedback and scores should be visible on Autolab for everyone now.

proof_reading autolab

edit · good note | 0

Updated Just now by Atri Rudra

Course evaluations

note @93

stop following 5 views

Course evals

Y'all should have received email to fill in the course evals. Please do consider doing so-- your feedback will be very useful for the future offerings of the course. In particular, please do let me know (1) how things could be improved and (2) what worked well so that I can keep doing the latter and improve on the former.

This should be the link for you to go and fill in the evals:

https://sunyub.smartevals.com/

#pin

logistics

edit · good note | 0

Updated 4 days ago by Atri Rudra

Last bonus opportunity deadline TODAY

Alternate name for the course

Options

- View handin history
- View writeup
- Download handout

Due: **May 7th 2020, 5:00 pm**

Last day to handin: **May 7th 2020, 5:00 pm**

☐ I affirm that I have complied with this course's academic integrity policy as defined in the syllabus.

SUBMIT

(∞ submissions left)

Project report due 5pm TODAY

note @96

stop following 4 views

Autolab accepting project reports now

Autolab is now accepting the project report for all groups that is due **Th (May 7) at 5pm**. Please look at the project page for more details on what is expected in the report but it should be a self-contained document and make sure you address all the required part. Also you can add in an appendix if you want and it will not count towards the **10pg limit** (*but there is no guarantee that I'll read the appendix*).

Your groups have been set up on Autolab. However, **please double-check the following:**

- Make sure you can see your group. To do so, follow these steps
  - Go to CSE 410 > Report (under Project)
  - Go to Options > Group options
  - Check to make sure your and your partners' emails are in there.
    - **DO NOT**
      - Click on invite another student to the group or the leave group buttons
- Please submit a dummy PDF well in advance of the deadline and check with ALL your group-mates to make sure they see your submitted PDF in their account.
  - I will grade the last PDF submitted by any of the group members.
  - I recommend that you designate one person in your group as the official submitter but whatever works for your group is fine with me.

project autolab

edit · good note | 0

Updated 3 days ago by Atri Rudra

In case your project is running behind

note @97

0 views

If you need more time on your project

In case your group is not where it wanted to be by the time the project report is due at 5pm this Th (@96), note that you can do more till the deadline to submit your final presentation slides (which is 5pm on Wed, May 13).

In particular, if your group has specific things planned to do between the deadline for the report and the deadline for the final presentation slides make sure you mention those in the "Results" section of your report.

If you have any followup questions, please post them in the comments section below.

logistics project

edit · good note | 0

Updated Just now by Atri Rudra

Final presentation slide due 5pm Wed

Autolab submission should be up over the weekend

Survey part of project starts later today

Submitting the survey

The peer evaluation survey will have to be filled on https://cse.buffalo.edu/teamwork. You will evaluate yourself and your groupmates in all the five categories.

The workflow

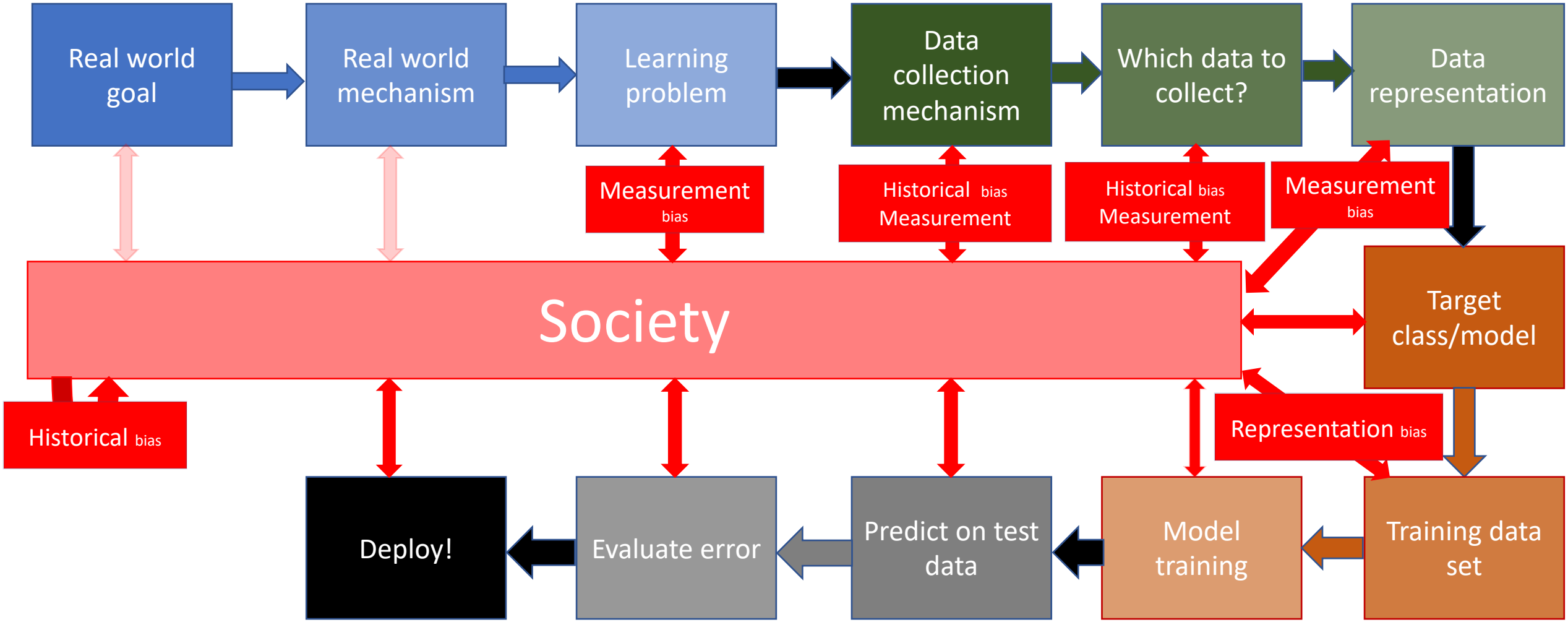
1. At **5:01pm on Thursday, May 7** the website above will be ready for you. You will need to complete the survey by **5pm on Wednesday, May 13**.
2. You will need to enter your UB email and click on a button to generate a verification code.
3. You will have limited time (~10 mins) to enter the verification code into the webpage.
4. You will then fill in the survey: the website will ask you to evaluate yourself and your groupmates in all the five categories above.
5. Your part is done. Atri will use your survey responses and your project score to post your survey scores on Autolab (the scores will be posted on your final presentation submission).
6. If you have not used this system before, please **do a dry run of the system WELL before the final deadline**. Your can re-submit and only your final submission will count (so you can iron out any issues in your trial submissions).

project logistics

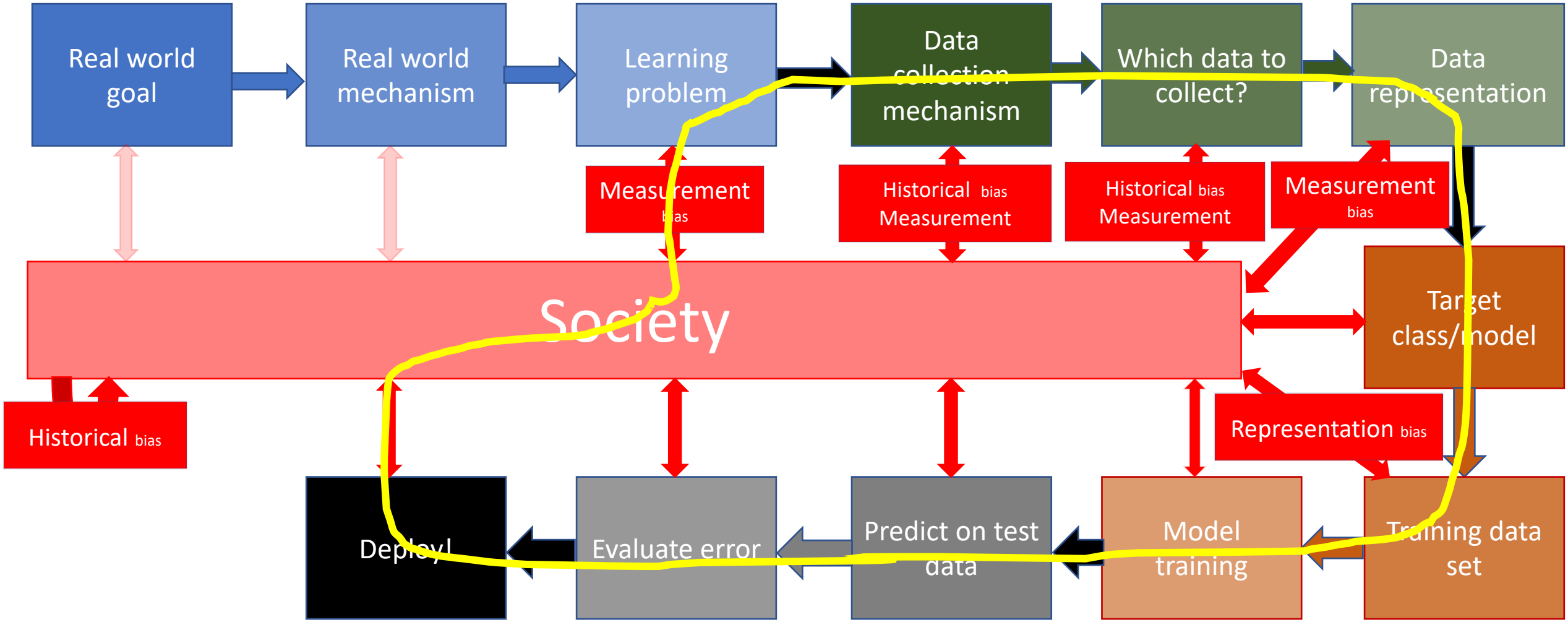
edit · good note | 0

Updated 38 minutes ago by Atri Rudra

What is a feedback loop



The “loop” in feedback loop



News from two weeks back!

Kate Crawford
@katecrawford

Big news: LAPD will end the use of the broken predictive policing system known as PredPol, citing budget concerns under COVID-19. This is thanks in large part to community groups like @stoplapdspying pushing back against its use.

LAPD will end controversial program that aimed to predict where crimes woul...
Chief Moore says, due to financial constraints caused by the pandemic, the LAPD will end a program that predicts where property crimes could occur.
latimes.com

Do feedback loops exist?

How do we “prove” that feedback loops can exist in predictive policing?

Simulation results

Theoretical modeling results

A simulation result

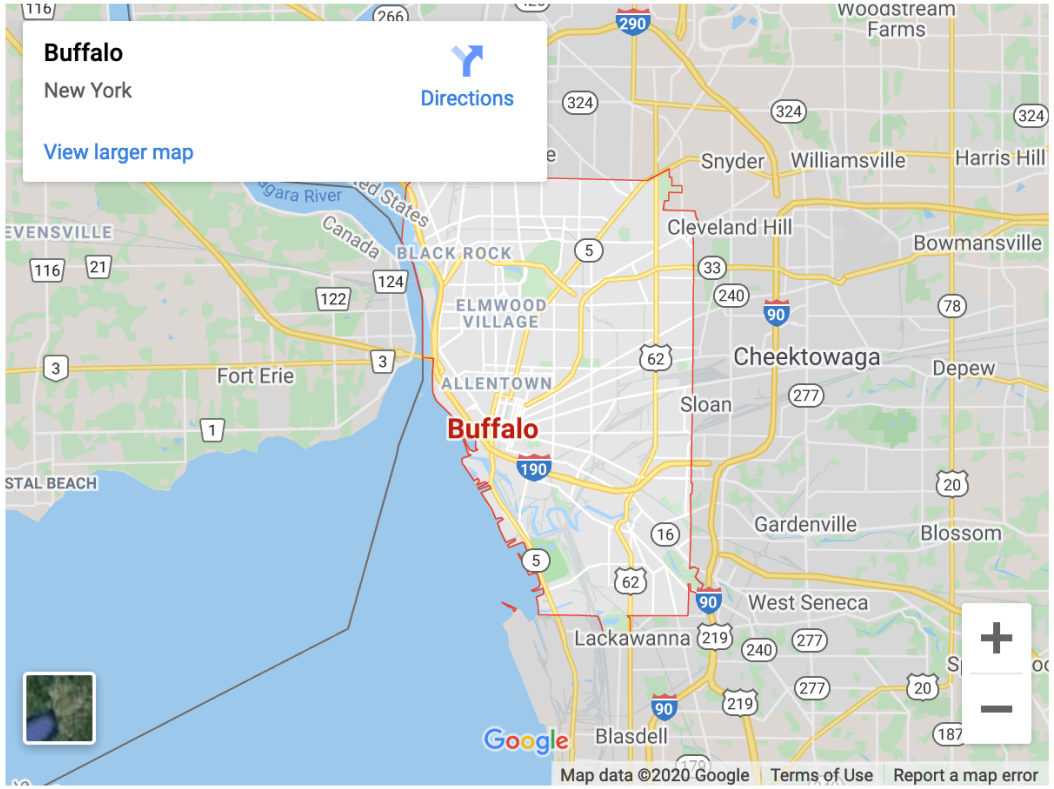
IN DETAIL

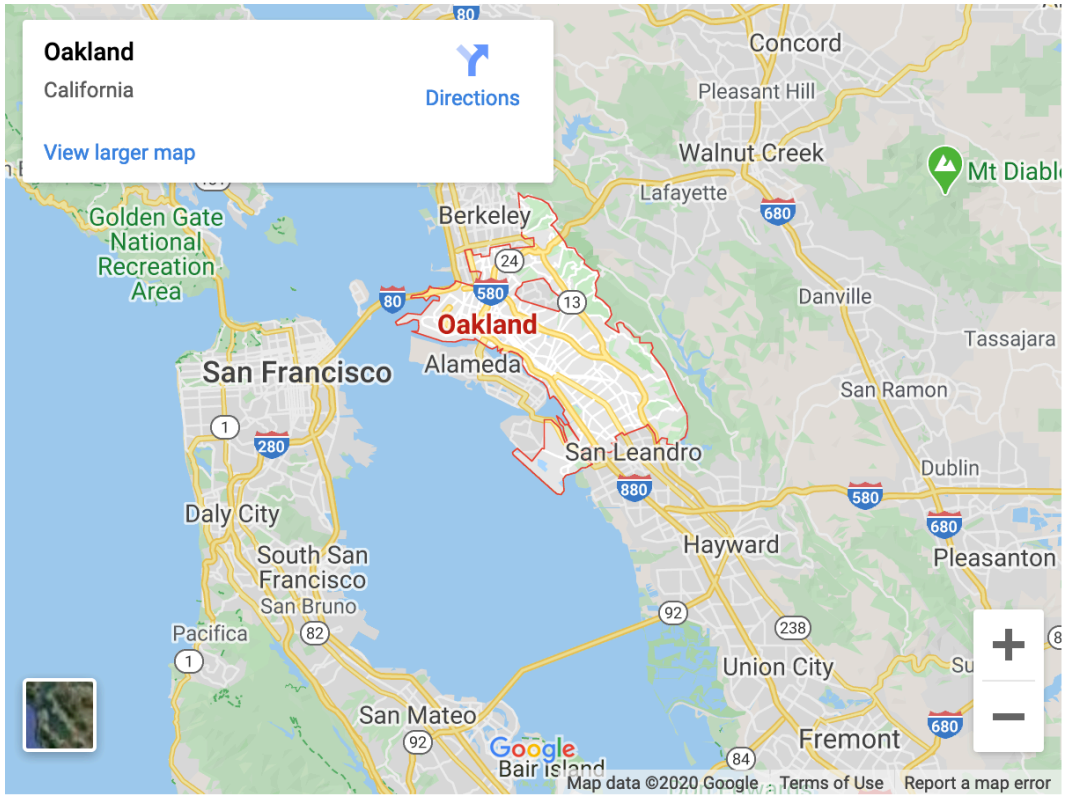
To predict and serve?

Predictive policing systems are used increasingly by law enforcement to try to prevent crime before it occurs. But what happens when these systems are trained using biased data? **Kristian Lum** and **William Isaac** consider the evidence – and the social consequences

How would you run this simulation?

Step 1: Pick a city





Run predpol on historical crime data and see if there is indeed a feedback loop

Step 2: Pick a crime type

Drug related crimes

For the rest of the section **let us assume you have access to drug crimes arrests for say year 2019 in Buffalo**. This is a realistic assumption since many cities in the US male crime incidents data public: e.g. here is crime incidents reported in Buffalo.

Access to predpol?

What about access to predpol?

We will **assume that you have access to the predpol algorithm**. It turns out that there is enough information about the algorithm (see e.g. overview from predpol that this is not an unreasonable assumption.

One of the video above briefly talked about how predpol works but here is a quick summary:

1. The geographical areas are divided into "grids" and all the crime information is aggregated in each cell.
2. predpol (as per its founders and via Lum and Isaac) "only three data points in making predictions: past type of crime, place of crime and time of crime. It uses no personal information about individuals or groups of individuals, eliminating any personal liberties and profiling concerns."

Questions we want answered

Initial State Question

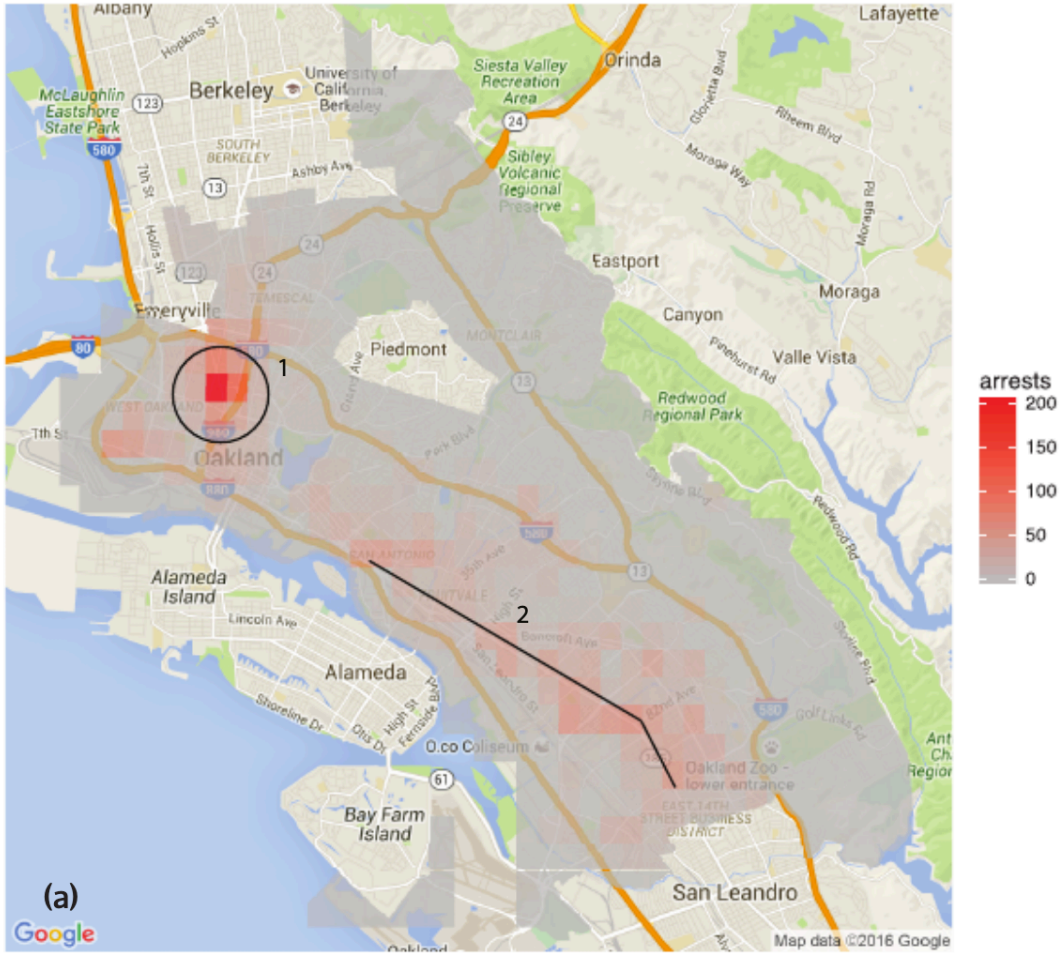
Here the question basically is the following. If we were to use predpol to say decide on where potential drug crimes were to happen on January 1, 2019 in Buffalo, then how does predpol's prediction compare to "actual" drug crimes on the same day.

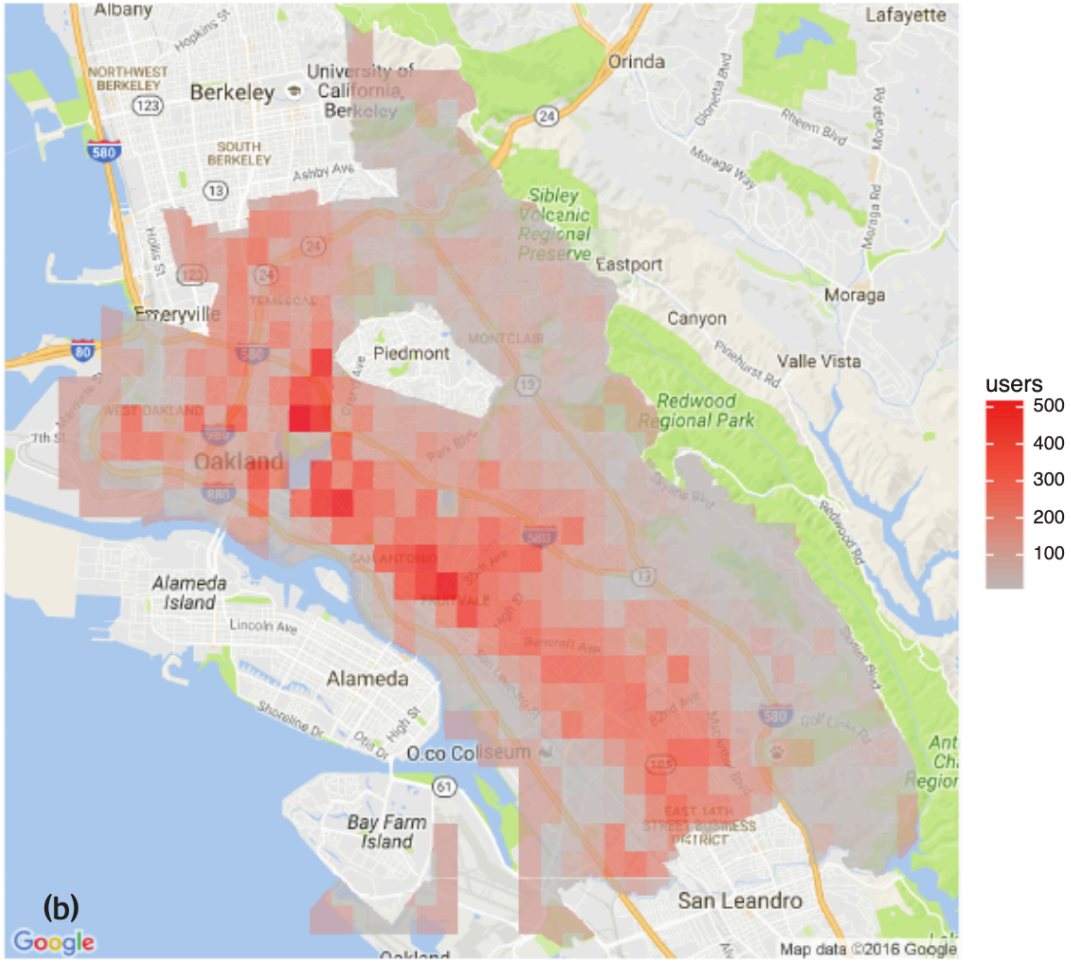
`Note` The use of quotes in "actual" is on purpose-- we will come to this shortly.

Feedback Loop State Question

Here the question is actually to test whether predictive policing could lead to feedback loop. The idea is to simulate the use of predpol in Buffalo from January 1, 2019 to December 31, 2019 in Buffalo and use the observed drug crime data for a given day to generate predpol's prediction for the next day. The goal here is to check if the the odds of location being "targeted" by predpol goes up (say when compared to the reported 2019 drug crimes as a baseline).

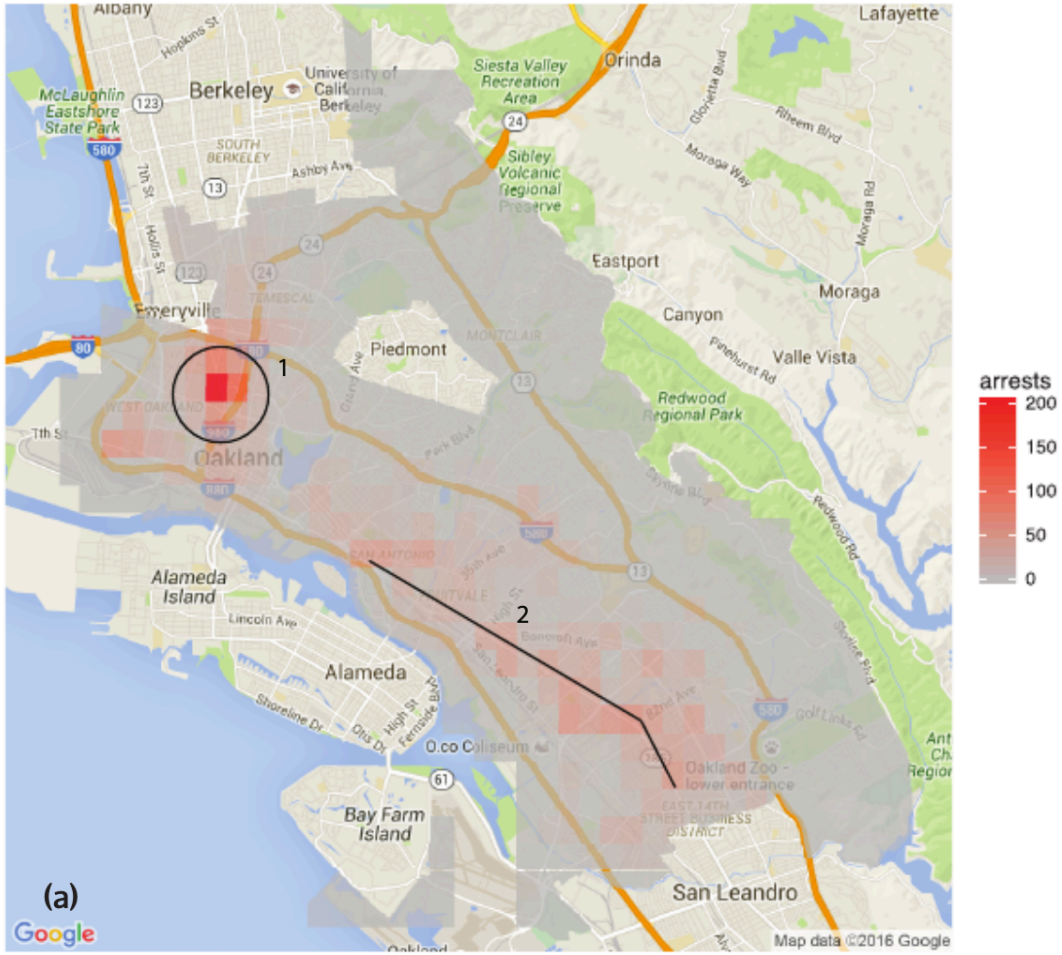
Lum and Isaac results: arrests vs ground truth

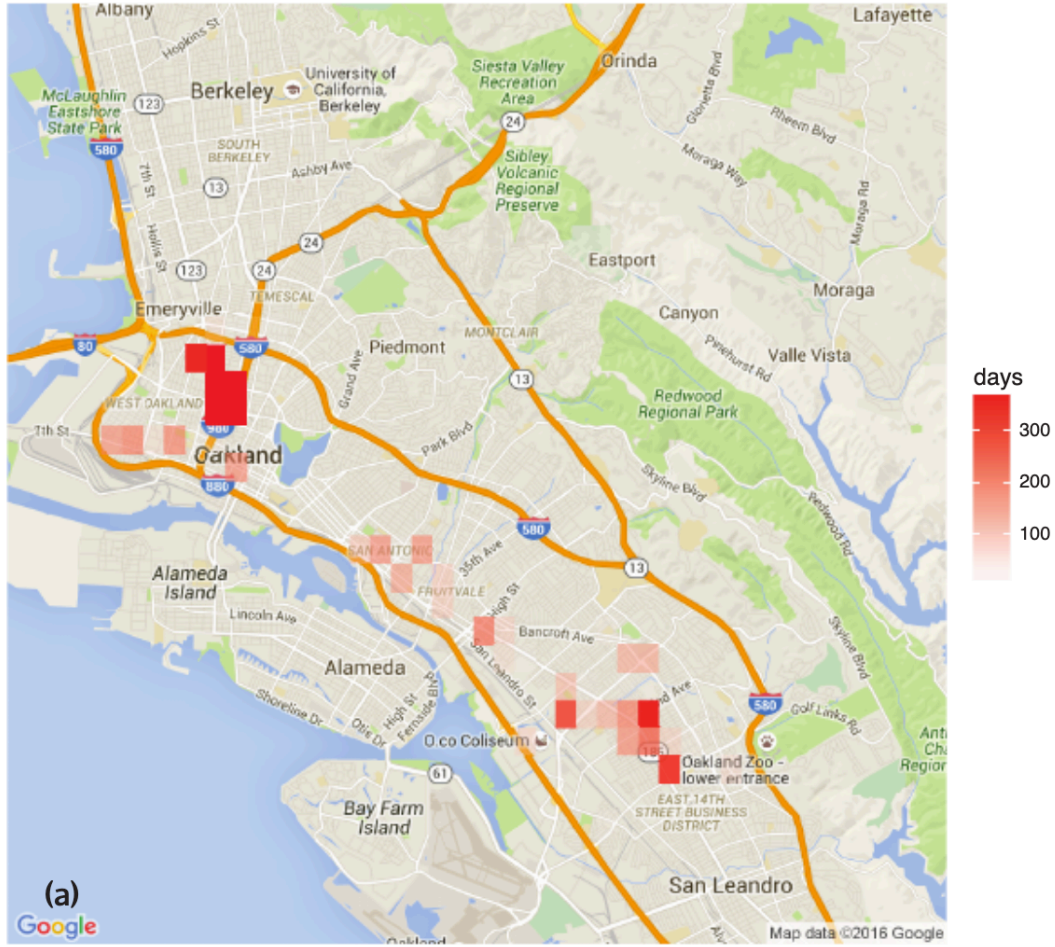




Figures taken from the Lum and Isaac paper

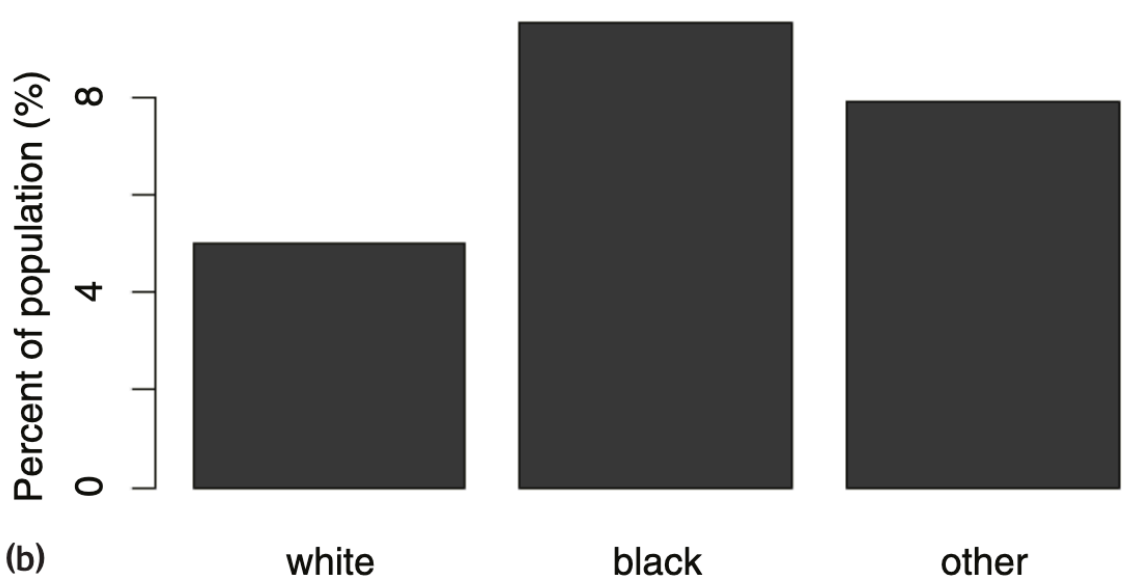
Lum and Isaac result: predpol output

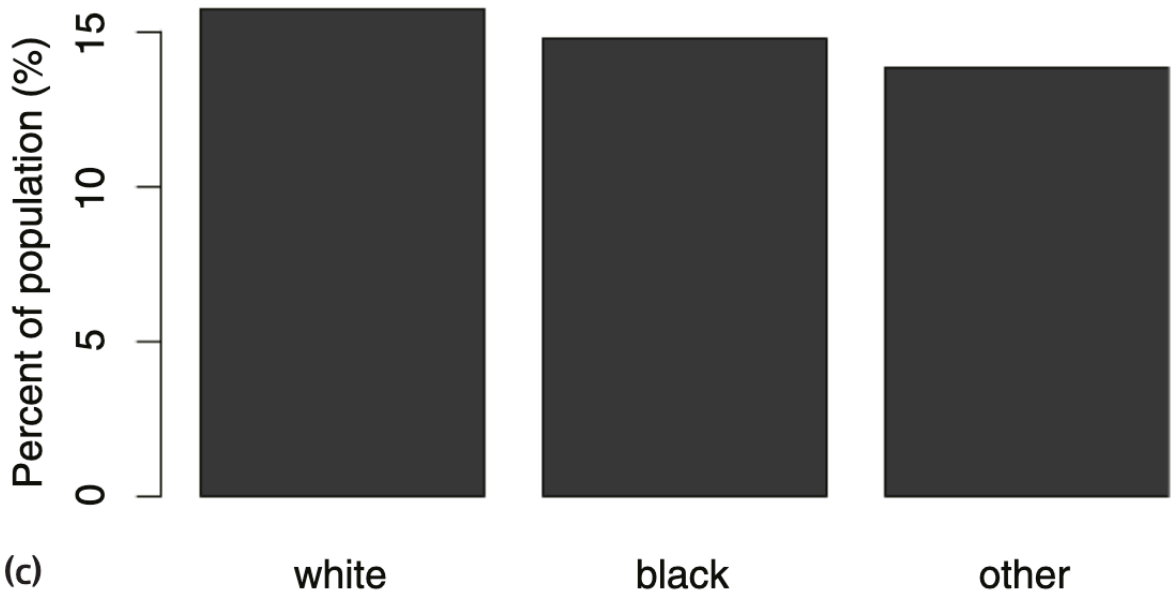




Figures taken from the Lum and Isaac paper

Lum and Isaac result





Figures taken from the Lum and Isaac paper

Back to feedback loop state question

Feedback Loop State Question

Here the question is actually to test whether predictive policing could lead to feedback loop. The idea is to simulate the use of predpol in Buffalo from January 1, 2019 to December 31, 2019 in Buffalo and use the observed drug crime data for a given day to generate predpol's prediction for the next day. The goal here is to check if the the odds of location being "targeted" by predpol goes up (say when compared to the reported 2019 drug crimes as a baseline).

Exercise (using predpol only on historical data)

Consider the following simulation. From January 1, 2019 to December 3 e use the **original** drug crime data for a given day to generate predpol's prediction for the next day.

Do you expect the above simulation result to s up (when compared to the reported 2019 drug crimes)? If so, why? If not, why not?

Click here for the answer

Discuss!

Since the effect of predpol prediction is not fed back into predpol, we should expect the odds of predpol targeting an area to track the historical reported drug crimes.

Now really back to the Q

Exercise (feedback loop state question)

How would you go about incorporating the effect of predpol's predictions into reported drug crimes? I.e. how would you incorporate the "feedback"?

Click here for how Lum and Isaac answered this question

One simple way to do this to increase the num police be sent. E.g., one could bump up the reported drug crimes in a area by $x\%$ (for some value of x) in all areas were pe geted by predpol, use the historical reported drug crime data. The idea here is that if more police are sent to an area then more crimes wil de. This updated crimes data is then fed into predpol for get predictions for next day.

Discuss!

Lum and Isaac result (20% increase)

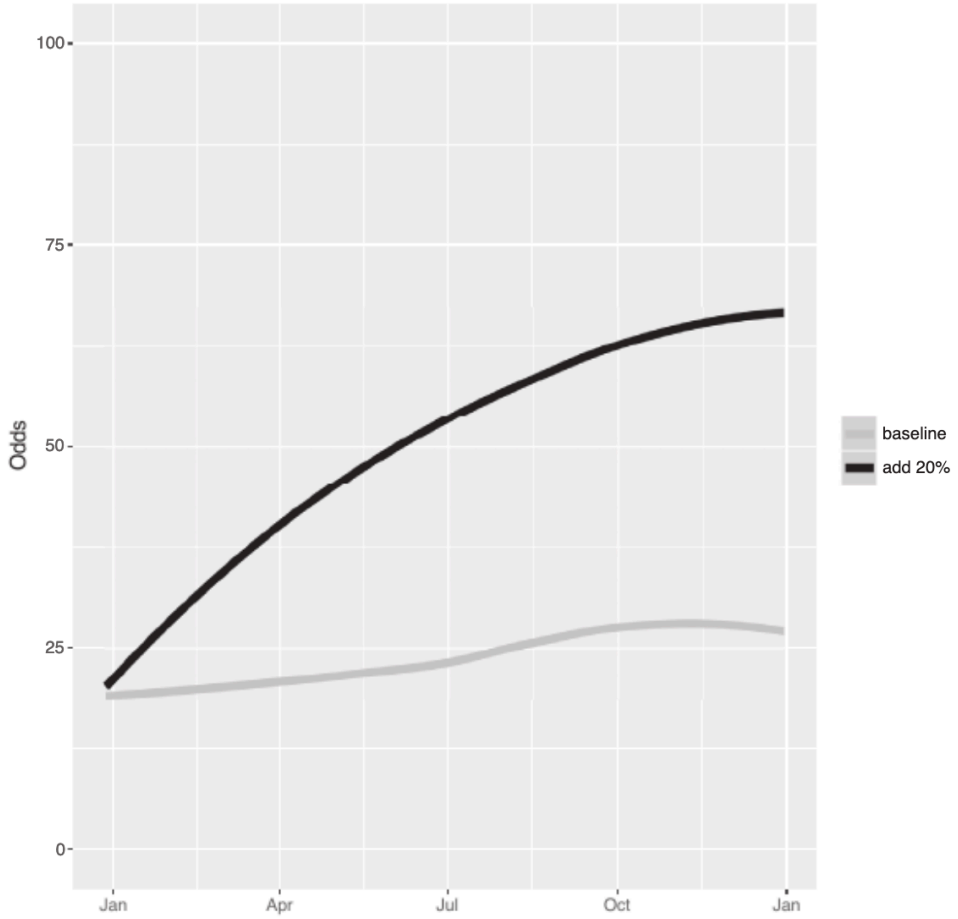


FIGURE 3 Predicted odds of crime in locations targeted by PredPol algorithm, relative to non-targeted locations. 'Baseline' is original Oakland police data. 'Add 20%' simulates the effect of additional crimes being observed in targeted locations

Figure taken from the Lum and Isaac paper

Any weakness of Lum and Isaac result?

Exercise (potential issues with the Lum and Isaac result)

Can you think of some weakness(es) in the Lum and Isaac result. In particular, if you were tasked with the job of poking holes in their setup how would you go about it? For example, can you think of ways that their conclusion might not be general enough?

Click here for one potential drawback

One potential issue is that the results are based on a synthetic population (a vell as synthetic ground truth). While it is apparent that at the aggregate level, their population and ground truth are accurate (since they are based on solid surveys), on argue that perhaps the feedback look observed is due to the specific synthetic population that was chosen. I.e. there *could* have been some weird co-incidence i cted synthetic population that led to the occurrence of the feedback loop then it being some general phenomenon.

Discuss!

We would hasten to state that the **above is a very weak ar** l only works on essentially aggregate counts of crime incident the composition of the synthetic population should not matter. However, n a way to is very weak objection.

Theoretical model

Mathematically prove that a feedback loop **always** exist

Have to make (many) simplifying assumptions

Proceedings of Machine Learning Research 81:1–12, 2018 Conference on Fairness, Accountability, and Transparency

Runaway Feedback Loops in Predictive Policing*

Danielle Ensign DANIPHYE@GMAIL.COM
University of Utah

Sorelle A. Friedler SORELLE@CS.HAVERFORD.EDU
Haverford College

Scott Neville DROP.SCOTT.N@GMAIL.COM
University of Utah

Carlos Scheidegger CSCHEID@CSCHEID.NET
University of Arizona

Suresh Venkatasubramanian† SURESH@CS.UTAH.EDU
University of Utah

Basic setup

Basic setup

We will assume that there are only two regions we are interested in-- E (for East side in Buffalo and W (for (Upper) West side in Buffalo).

We will assume there is only **one cop** tasked with patrolling E and W. Further, given that only one cop is patrolling an area we **assume that in one day the cop can only discover (at most) one crime** (there is paperwork etc. involved and with one cop patrolling these big areas the chances of them catching multiple crimes are pretty-much non-existent.

These are indeed **NOT realistic** assumptions. But we make these assumptions since it makes proving things in this model easier. The way to look at this model is that this is a *first attempt* and indeed followup modeling would be needed to make the model more realistic.

Assigning the cop to a region

Assumption 1 (assigning the cop to a region)

We will assume that the **cop can only go to one region each day**. (If you are OK with the assumption of there being only one cop, this is not that unrealistic!)

Further, we will assume that the **cop will visit s region with probability directly proportional to how many crimes have been reported till date in the given region**.

Notation alert

For any given day t, we will use $n_E^{(t)}$ and $n_W^{(t)}$ to denote the number of observed crimes in E and W respectively from day 0 to day t.

In particular, the cop will visit E with probability

$$\frac{n_E^{(t)}}{n_E^{(t)} + n_W^{(t)}}$$

and will visit W with probability

$$\frac{n_W^{(t)}}{n_E^{(t)} + n_W^{(t)}}.$$

Unequal crime rates

Assumption 2 (unequal crime rates)

We will assume that the E and W have **different crime rates (which are unknown)**.

Notation alert

The crime rate for E is λ_E and the crime rate for W is λ_W.

In particular, we are assuming

$$\lambda_E \neq \lambda_W.$$

How crimes are observed

Assumption 3 (observed crime rate = actual crime rate)

We will **assume that crime is discovered by the cop in either region EXACTLY at the same as its actual crime rate**.

NOT a realistic assumption

As we have seen before this is not a realistic assumption.

However, as we will see shortly even under this very generous (to predpol) assumption, we will show the existence of a feedback loop, which only makes the argument for dangers of feedback loop stronger.

In particular, if the cop goes to E, then they will observe one crime with probability λ_E and no crimes with probability $1 - \lambda_E$. Similarly if the cop goes to W, then they will observe one crime with probability λ_W and no crimes with probability $1 - \lambda_W$.

The overall process..

The evolution of the number of observed crimes

- The process starts with initial values $n_E^{(0)}$ and $n_W^{(0)}$.

- For $t = 1, 2, \ldots$

```
//The process repeats "forever"
```

  - With probability $\frac{n_E}{n_E^{(t)} + n_W^{(t)}}$ do:

```
//Cop visits E
```

    - With probability λ_E set $n_E^{(t+1)} = n_E^{(t)} + 1$
    - Else with probability $1 - \lambda_E$ set $n_E^{(t+1)} = n_E^{(t)}$

  - Otherwise with probability $\frac{n_W^{(t)}}{n_E^{(t)} + n_W^{(t)}}$ do:

```
//Cop visits W
```

    - With probability λ_W set $n_W^{(t+1)} = n_W^{(t)} + 1$
    - Else with probability $1 - \lambda_W$ set $n_W^{(t+1)} = n_W^{(t)}$

Wait, what was the question again?

Exercise

To make things concrete assume that $n_E^{(0)} = n_W^{(0)} = 100$ and $\lambda_E = 10.5\%$ and $\lambda_W = 11\%$. What would you consider to be a manifestation of feedback loop as the process above runs?

`Hint`: Think about how the ratios $\frac{n_E^{(t)}}{n_E^{(t)}+n_W^{(t)}}$ and $\frac{n_W^{(t)}}{n_E^{(t)}+n_W^{(t)}}$ evolve as t grows larger. (Side question: why are these ratios something worth monitoring?)

First note that the ratios $\frac{n_E^{(t)}}{n_E^{(t)}+n_W^{(t)}}$ and $\frac{n_W^{(t)}}{n_E^{(t)}+n_W^{(t)}}$ [illegible] E and W on day t respectively. Ideally, one would like these probabilities to be proportional to λ_E and λ_W respectively.

Since λ_E and λ_W are very close to each other, a feedback lo[illegible] steady state the probability that the cop goes to E is not close to the probability that they go to W. For example, an extreme case of feedba[illegible] surely goes to only one region.

Discuss!

The existence of feedback loop

Feedback loop always exists

It can be shown that if $\lambda_E > \lambda_W$, then the cop will go to E with probability 1 and if $\lambda_W > \lambda_E$, then the cop will go to W with probability 1. In other words, if there is any imbalance (however slight) between the crime rates in E and W, eventually the cop will only go to the region with the higher crime rate.

Exercise

Can you argue the above claim?

Discuss!

Here is the intuitive argument. Assume that $\lambda_E > \lambda_W$ ($> \lambda_E$ is essentially the same), then assuming the cop goes to a specific region, the number of observed crimes in E will be more t ses this "advantage" is re-inforced till cop only ends up going to E.

Back to Pranshu's Question

A potential fix

In the above model, Ensign, Friedler, Neville, Scheidegger and Venkatasubramanian suggest the following fix (which they can mathematically prove that it works) is based roughly on the following idea. If the cop visits a specific region most of the time, then it should not be a surprise if they discover a crime in the region and in such a case they should "discount" the crime discovery by not recording such discovery mo[illegible]f the time. On the other hand, if the cop visits region infrequently and they discover a crime, they have learned something "new" and hence would record such crime disco[illegible]ost of the time.

Discuss!

Passphrase: Suresh Venkatasubramanian

I'm a professor in the School of Computing at the University of Utah. My research interests are a random walk through the theoretical and applied aspects of data science, including computational geometry, sublinear algorithms, clustering, and kernel methods.

I'm currently very interested in the social ramifications of automated decision making. I'm a founding member of the FAccT conference. I'm also a member of the board at the ACLU of Utah and a member of the Computing Community Consortium Council.

News

- Jan 27-30: At the ACM Conference on Fairness, Accountability and Transparency
- Jan 9-10: At Data Science across the Undergraduate Curriculum
- Dec 12-14: At Social Responsibility of Algorithms, 2019

For more, see my news page

Research

I'm interested in the problem of *algorithmic fairness*: ensuring that in a world of automated decision-making, decisions that get made about us and for us are fair, accountable and transparent.

This is the culmination of a series of research explorations that started with algorithms and computational geometry, lifted to high dimensional geometry and sublinear algorithms, with a random sampling of work in clustering and kernel methods.

For more, see my publications

Teaching

Spring 2020

- Clustering

Fall 2019

- Ethics of Data Science

Spring 2019

- Advanced Algorithms

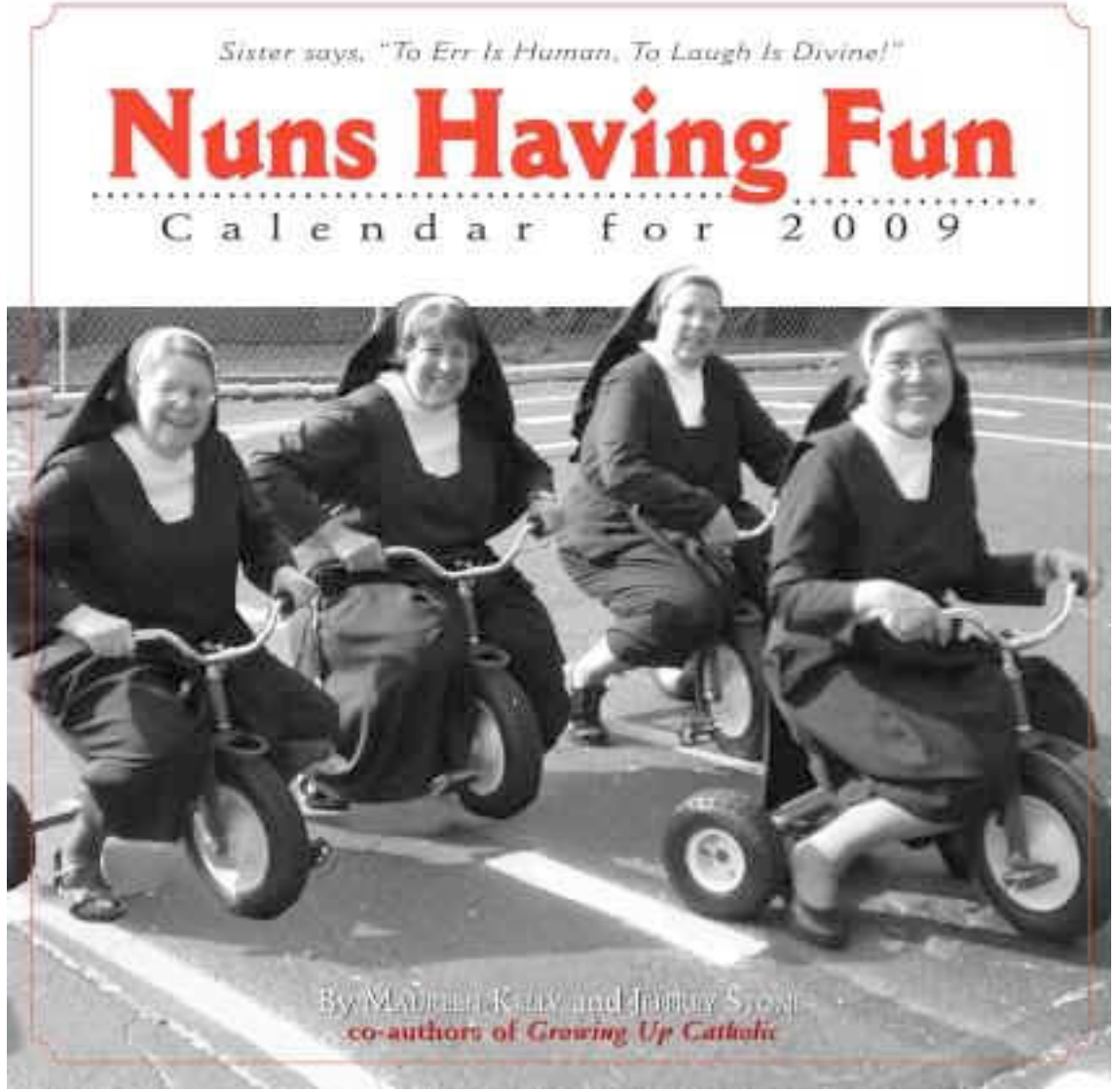
For more, see my teaching page.

Let’s reflect back

Humans have been making decisions for centuries….

Why is replacing human decision making by ML pipeline problematic?

Hopefully the class was fun for y'all!



Thanks!

Except of course the final project presentation