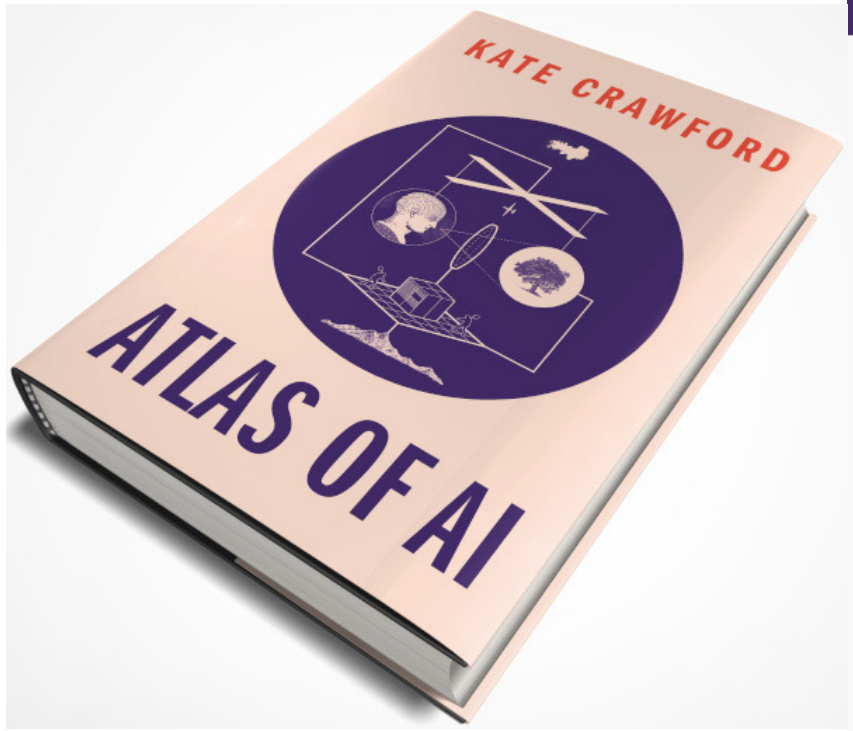
ML and Society

Feb 20, 2023

Passcode: Kate Crawford



KATE CRAWFORD

Kate Crawford is a leading scholar of the social and political implications of artificial intelligence. Over her 20-year career, her work has focused on understanding large-scale data systems, machine learning and AI in the wider contexts of history, politics, labor, and the environment.

Kate is a Research Professor at USC Annenberg, a Senior Principal Researcher at MSR-NYC, and an Honorary Professor at the University of Sydney. She is the inaugural Visiting Chair for AI and Justice at the École Normale Supérieure in Paris, where she co-leads the international working group on the Foundations of Machine Learning. In 2021, she received the Miegunyah Distinguished Visiting Fellowship at the University of Melbourne. She has co-founded multiple interdisciplinary research groups including FATE at MSR, AI Now Institute at NYU, and Knowing Machines at USC. Kate has advised policy makers in the United Nations, the Federal Trade Commission, the European Parliament, the Australian Human Rights Commission, and the White House.

Sit with your team!

Team 1	Afzal	Cole	Navid	Tim	
Team 2	Aishwarya	Herman	Mads	Melvin	
Team 3	Daphkar	Juliana	Ibtida	Monica	
Team 4	Joe	Ken	Vedant	Zach	
Team 5	Chaitanya	Evan	Hitesh	Sushanth	
Team 6	Hannah	Harinee	Gabriella	Suradhya	
Team 7	Alex	Connor	Gopi	Shane	Thanh
Team 8	Aditi	Connor	Jason	Mitali	
Team 9	Botsalano	Niharika	Vedang	Yunmei	
Team 10	Dhiraj	Frank	Kashyap	Michael	

Rage students in Green. ML+Soc students in black

Team meetings

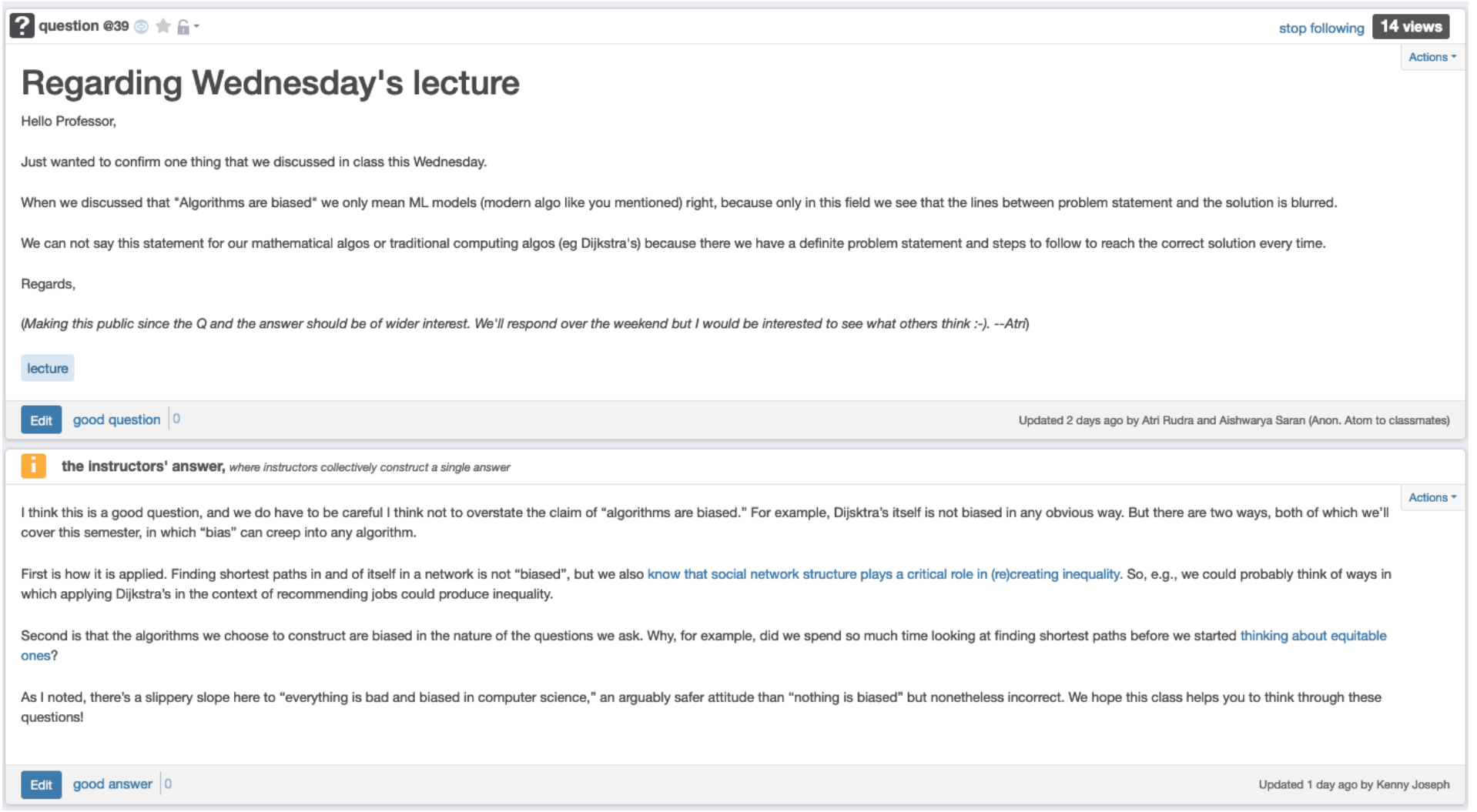
In class tomorrow (Mon, Feb 20), after the initial attendance and checkin, Kenny and I will meet with y'all. Some logistics:

- We will meet with each group for **6 mins**
- We will meet with the groups whose submission we graded.
- Each of us will meet the groups in the following order (Kennny and I will meet with two groups in parallel at any time):
    - Kenny
        - Team 9
        - Team 2
        - Team 3
        - Team 4
        - Team 8
    - Atri
        - Team 5
        - Team 7
        - Team 10
        - Team 1
        - Team 6

Looking forward to some great discussions tomorrow!

Break!

Are algorithms racist?



question @39 stop following 14 views

Actions

Regarding Wednesday's lecture

Hello Professor,

Just wanted to confirm one thing that we discussed in class this Wednesday.

When we discussed that "Algorithms are biased" we only mean ML models (modern algo like you mentioned) right, because only in this field we see that the lines between problem statement and the solution is blurred.

We can not say this statement for our mathematical algos or traditional computing algos (eg Dijkstra's) because there we have a definite problem statement and steps to follow to reach the correct solution every time.

Regards,

(*Making this public since the Q and the answer should be of wider interest. We'll respond over the weekend but I would be interested to see what others think :-). --Atri*)

lecture

Edit good question 0

Updated 2 days ago by Atri Rudra and Aishwarya Saran (Anon. Atom to classmates)

the instructors' answer, *where instructors collectively construct a single answer*

Actions

I think this is a good question, and we do have to be careful I think not to overstate the claim of "algorithms are biased." For example, Dijsktra's itself is not biased in any obvious way. But there are two ways, both of which we'll cover this semester, in which "bias" can creep into any algorithm.

First is how it is applied. Finding shortest paths in and of itself in a network is not "biased", but we also know that social network structure plays a critical role in (re)creating inequality. So, e.g., we could probably think of ways in which applying Dijkstra's in the context of recommending jobs could produce inequality.

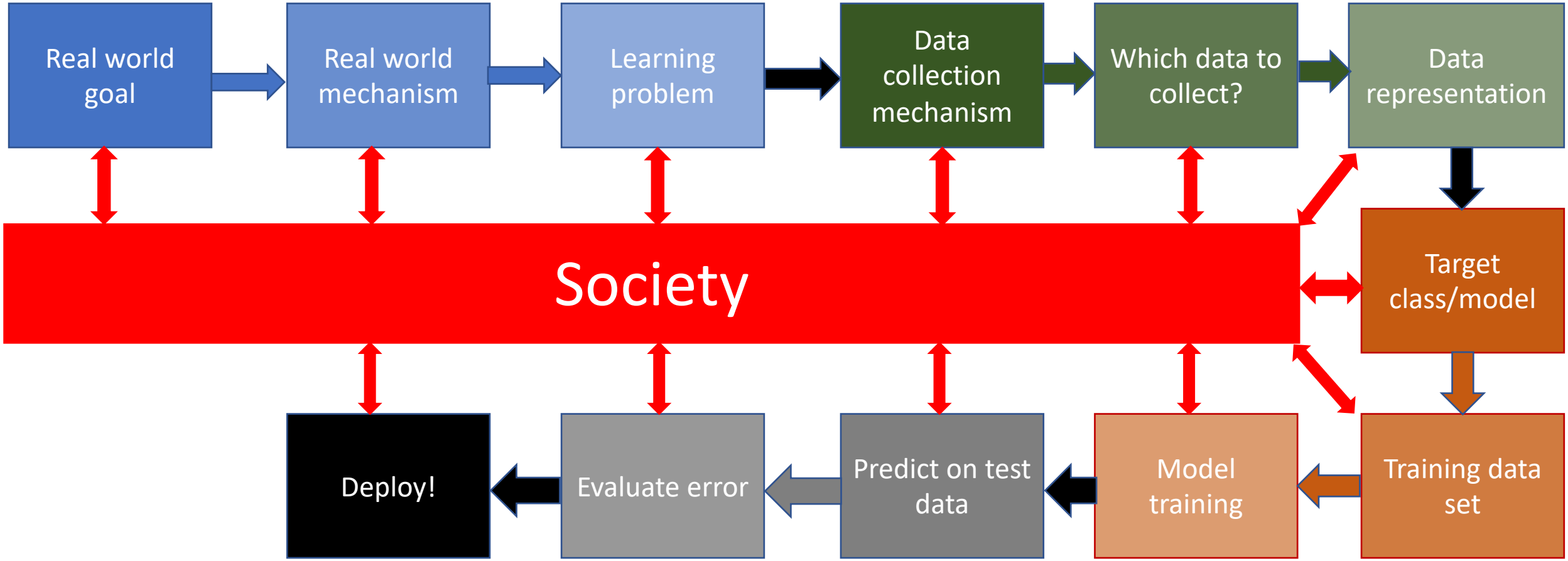
Second is that the algorithms we choose to construct are biased in the nature of the questions we ask. Why, for example, did we spend so much time looking at finding shortest paths before we started thinking about equitable ones?

As I noted, there's a slippery slope here to "everything is bad and biased in computer science," an arguably safer attitude than "nothing is biased" but nonetheless incorrect. We hope this class helps you to think through these questions!

Edit good answer 0

Updated 1 day ago by Kenny Joseph

ML pipeline


Real world goal
Real world mechanism
Learning problem
Data collection mechanism
Which data to collect?
Data representation
Society
Target class/model
Deploy!
Evaluate error
Predict on test data
Model training
Training data set

COMPAS

COMPAS (software)

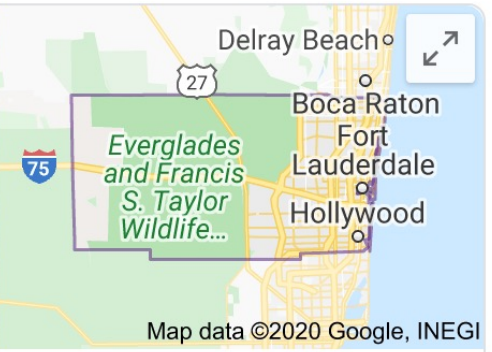
From Wikipedia, the free encyclopedia

COMPAS, an acronym for Correctional Offender Management Profiling for Alternative Sanctions, is a case manag
Equivant) used by U.S. courts to assess the likelihood of a defendant becoming a recidivist.[1][2]

COMPAS has been used by the U.S. states of New York, Wisconsin, California, Florida's Broward County, and oth

Contents [hide]

Risk Assessment [edit]



Broward County

County in Florida

Broward County is a county in southeastern Florida, US. According to a 2018 census report, the county had a population of 1,951,260, making it the second-most populous county in the state of Florida and the 17th-most populous county in the United States. The county seat is Fort Lauderdale. Wikipedia

Incorporated cities: 24

Population: 1.936 million (2017)

Mayor: Mark D. Bogen

PRO PUBLICA

Donate

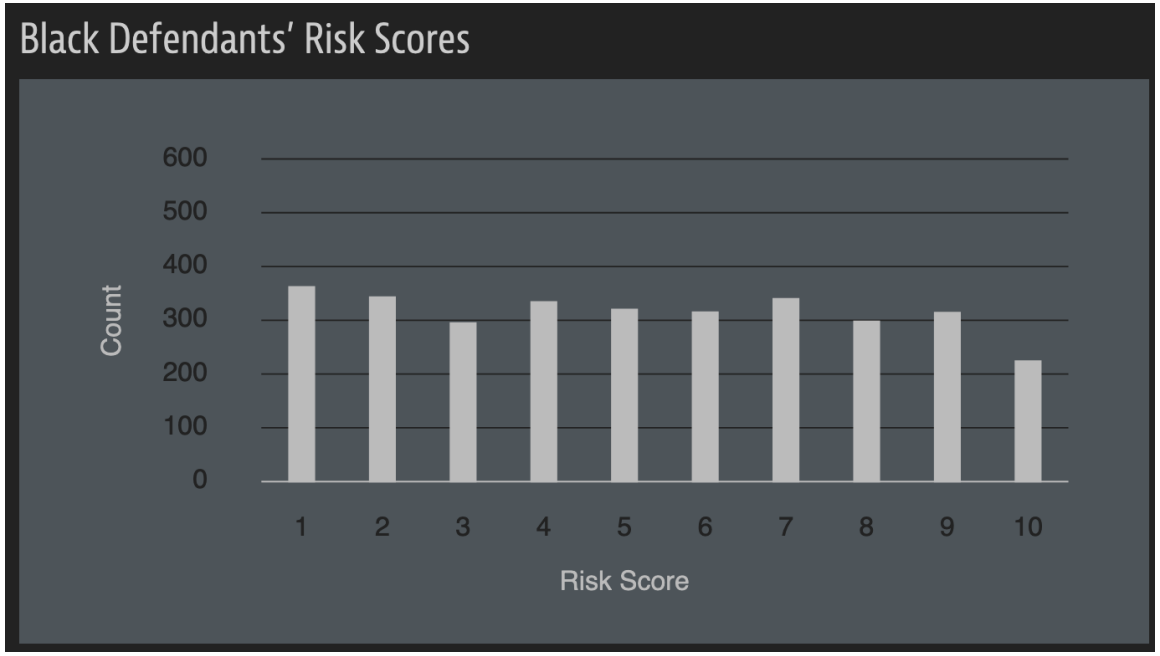
Machine Bias

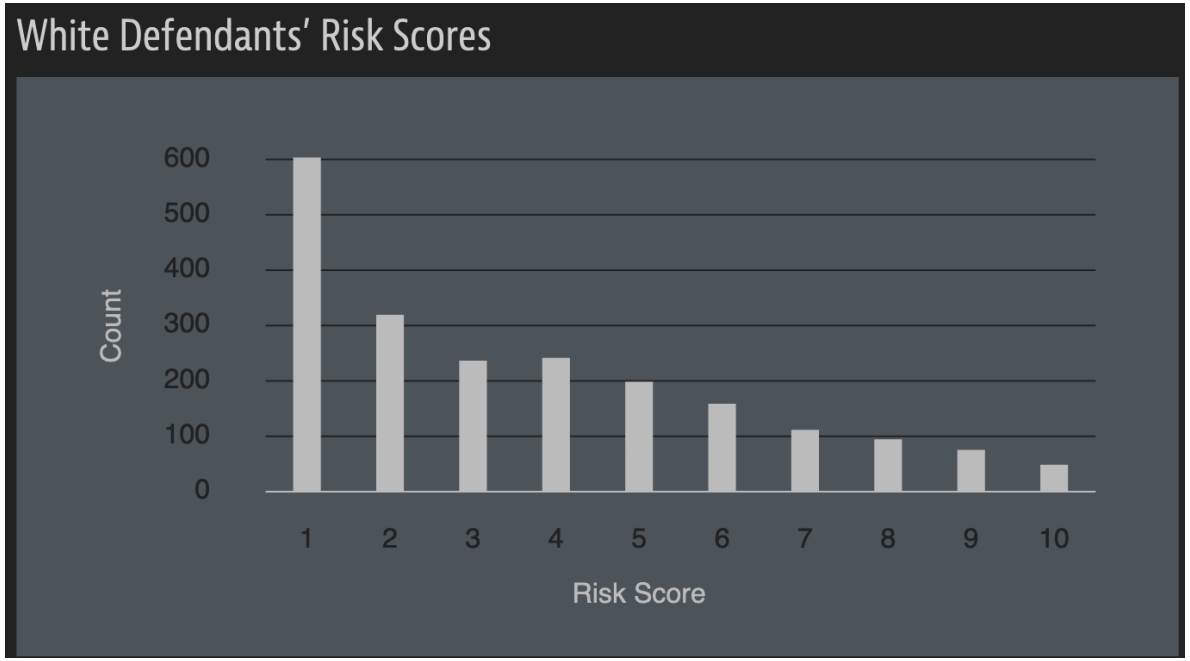
There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

A sample of their result

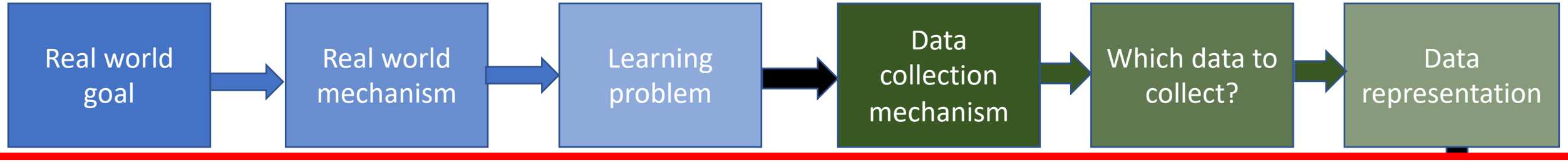




False Positives, False Negatives, and False Analyses: A Rejoinder to "Machine Bias: There's Software Used Across the Country to Predict Future Criminals. And It's Biased Against Blacks."

Anthony W. Flores
California State University, Bakersfield
Kristin Bechtel
Crime and Justice Institute at CRJ
Christopher T. Lowenkamp
Administrative Office of the United States Courts
Probation and Pretrial Services Office

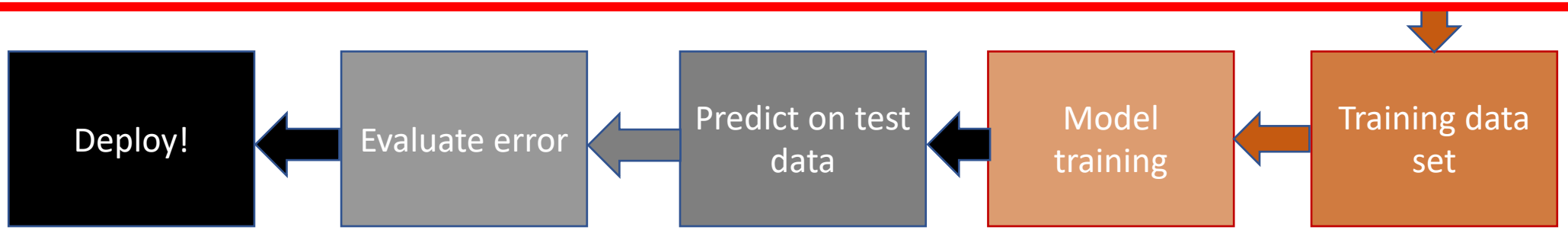
A walkthrough



The Problem

Imagine a situation where the creator of COMPAS had access to the COMPAS dataset. In particular, you are in the team that wants to predict recidivism based on the COMPAS dataset. How would you go about doing it?

Well, let's just walk through the ML pipeline to see how you would go about doing this.



Real world goal

Real world goal

Real world goal

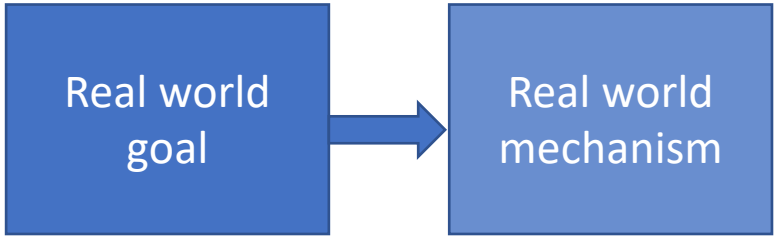
Reduce crime in society.

The Problem

Imagine a situation where the creator of COMPAS had access to the COMPAS dataset. In particular, you are in the team that wants to predict recidivism based on the COMPAS dataset. How would you go about doing it?

Well, let's just walk through the ML pipeline to see how you would go about doing this.

Real world mechanism



Real world mechanism

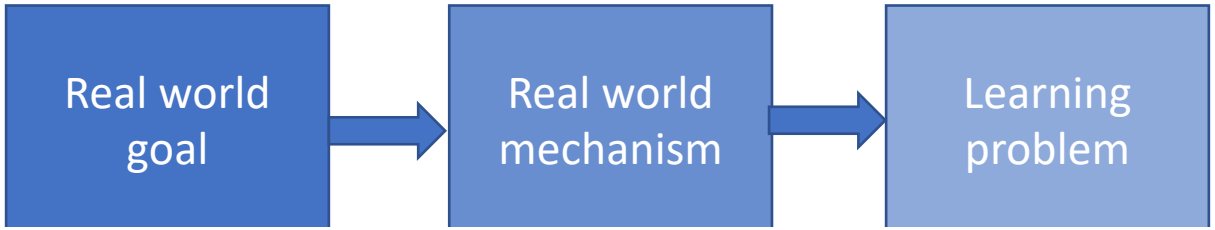
Based on some studies (or not!), your superiors decided that repeat offenders contribute most to crime. This in turn they decided would mean that if one could identify who would commit a crime again in the future, then one could use this information when making judgment on the current crime. Thus, they decided they wanted a system that can identify folks who will re-offend in the future and then promptly handed off the problem to your group to solve it.

The Problem

Imagine a situation where the creator of COMPAS had access to the COMPAS dataset. In particular, you are in the team that wants to predict recidivism based on the COMPAS dataset. How would you go about doing it?

Well, let's just walk through the ML pipeline to see how you would go about doing this.

Learning problem



Learning problem

Your group decides on the simplest learning problem: given a defendant *predict* if they will re-offend or not (in other words you are doing *binary classification* (binary because you are "labeling" defendants as either going too re-offend or not going to re-offend and you are doing classification because you are putting people into the two bins-- i.e. giving them a binary label and hence assigning them a "class."

There is another related option (which is what COMPAS: instead of assigning defendants to two scores: they assign a score from 1 (being least likely to re-offend) to 10 (most likely to defends). This range of score (rather than a binary classification) could potentially be more useful to the end user of your system.

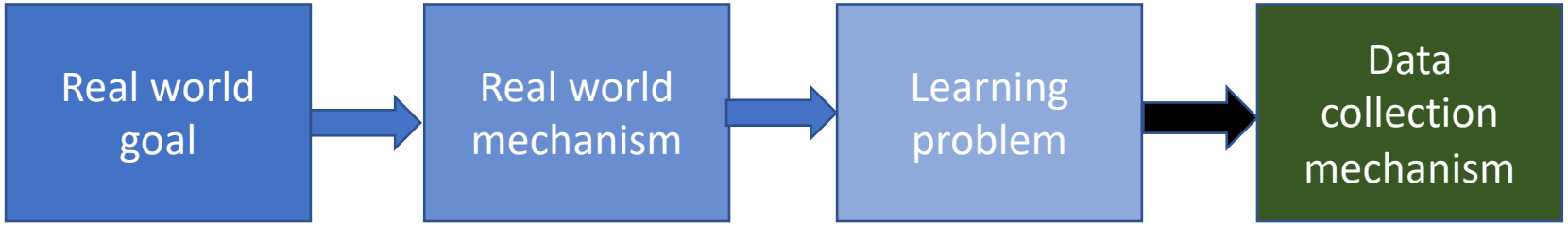
However, for our discussion (and indeed for most of the rest of the course), we will focus on binary classification.

The Problem

Imagine a situation where the creator of COMPAS had access to the COMPAS dataset. In particular, you are in the team that wants to predict recidivism based on the COMPAS dataset. How would you go about doing it?

Well, let's just walk through the ML pipeline to see how you would go about doing this.

Data collection mechanism



Data collection mechanism

Your group decides to use the COMPAS dataset.

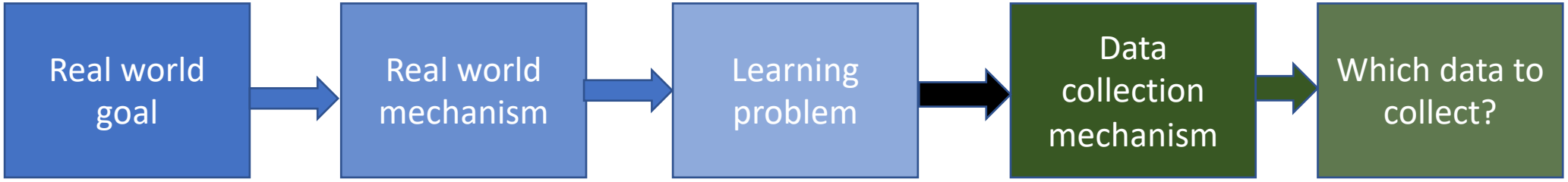
However, it is a useful exercise to recall what mechanism ProPublica used to collect the data (see the accompanying article to the main ProPublica article for details). In short, they used the existing public records law to get some data and generated the rest of the data was generated via a public government website. An important point to note this is a very *labor intensive process* and it's not like writing a script to log certain information about a system (though that also can work as in Hal Duame III's blog post on the machine learning pipeline. In other words, generating data can be *expensive* (if not directly in terms of money then in person-hours).

The Problem

Imagine a situation where the creator of COMPAS had access to the COMPAS dataset. In particular, you are in the team that wants to predict recidivism based on the COMPAS dataset. How would you go about doing it?

Well, let's just walk through the ML pipeline to see how you would go about doing this.

Which data to collect?



Which data to collect?

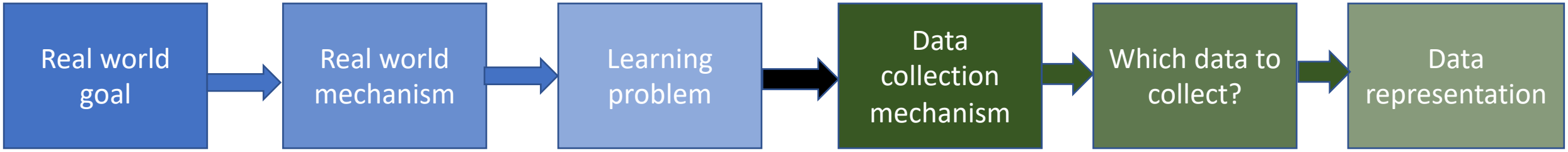
Your group decides to use whatever data the COMPAS dataset has.

However, it's worth it to note that in the ProPublica data collection, they could only collect data that was public and so your group does not have access to data that is not in the public domain that could be relevant to solve your learning problem. See the next callout for a pertinent example.

Measuring crime

We would now like to highlight one unavoidable (and potentially huge) issue with measuring/collecting data on when a crime was committed. For example, ideally in your group's problem you would like to figure out when someone re-offends: i.e. commits a crime again. However, public/police records can only show when someone was *arrested for a crime*. Keep this distinction in mind-- we will come back to this later on in the course (especially when we talk about feedback loops).

Data representation



Data representation

Since your group is using the COMPAS dataset, the data representation is also given to you.