

# Fairness in Machine Learning (Recap)



SML201: Introduction to Data Science, Spring 2020

Content from Moritz Hardt, Sam Corbett-Davies,  
Emma Pierson, Avi Feller, Sharad Goel

Michael Guerzhoy

# COMPAS

- “Correctional Offender Management Profiling for Alternative Sanctions”
  - Developed by Northpointe (currently Equivant)
  - Used by *a lot* of probation departments to assess the likelihood of a defendant becoming a recidivist
  - Defendants who are defined as medium or high risk are more likely to be detained before trial
    - (N.B., this is only suggestive of importance)
  - Race is not an input to the algorithm

## COMPAS Probation Risk and Needs Assessment Questionnaire

---

OFFENDER NAME: NYSID: STATUS:  
RACE: SEX: DOB:  
DATE OF ASSESSMENT: MARITAL STATUS:  
SCALE SET: Full COMPAS Assessment v2 AGENCY/COUNTY NAME:

### PART ONE: CRIMINAL HISTORY / RISK ASSESSMENT

#### CURRENT CHARGES

What offenses are covered by the current charges (check all that apply)?

|                             |                 |                  |
|-----------------------------|-----------------|------------------|
| Homicide                    | Arson           | Property/Larceny |
| Assault                     | Weapons         | Fraud            |
| Robbery                     | Drug Sales      | DWI / DWAI       |
| Sex Offense (with force)    | Drug Possession | AUO              |
| Sex Offense (without force) | Burglary        | Other            |

1 Do any of the current offenses involve domestic violence?

Yes No

2 What offense category represents the most serious current charge?

Misdemeanor Non-Assault Felony Assaultive Felony

3 Was there any degree of physical injury to a victim in the current offense?

Yes No

4 Based on your judgment, after reviewing the history of the offender from all known sources of information (PSI, police reports, prior supervision, victim, etc.) does the defendant demonstrate a pattern of violent behavior against people resulting in physical injury?

Yes No

[http://www.northpointeinc.com/downloads/research/D\\_CJS\\_OPCA\\_COMPAS\\_Probation\\_Validity.pdf](http://www.northpointeinc.com/downloads/research/D_CJS_OPCA_COMPAS_Probation_Validity.pdf)

## **COMPAS Probation Risk and Needs Assessment Questionnaire – *Continued***

---

### **PART TWO: NEEDS ASSESSMENT**

#### **A. ASSOCIATES / PEERS**

17 The offender has peers and associates who *(check all that apply)* :

- |                        |                                       |
|------------------------|---------------------------------------|
| Use illegal drugs      | Lead law-abiding lifestyles           |
| Have been arrested     | Are gainfully employed                |
| Have been incarcerated | Are involved in pro-social activities |
| None                   |                                       |

18 What is the gang affiliation status of the offender :

- Current gang membership
- Previous gang membership
- Not a member but associates with gang members
- None

19 Does the offender have a criminal alias, a gang-related or street name?

Yes No

20 Does unstructured idle time contribute to the opportunity for the offender to commit criminal offenses?

Yes Unsure No

21 Does offender report boredom as a contributing factor to his or her criminal behavior?

Yes Unsure No

#### **B. FAMILY**

22 Are the offender 's family or household members able and willing to support a law abiding lifestyle?

Yes Unsure No

23 Is the offender's current household characterized by *(check all that apply)* :

---

## COMPAS Probation Risk and Needs Assessment Questionnaire – *Continued*

### PART THREE: OFFENDER QUESTIONNAIRE

NYSID :

Name :

DOB :

Please look at the following areas and let us know which of them you think will present the greatest problems for you. *Please check one response for each question in the column provided .*

|    | <b>Please answer questions as either No, Yes or Don't Know</b>               | <b>No</b> | <b>Yes</b> | <b>Don't Know</b> |
|----|--|-----------|------------|-------------------|
| 48 | Do you feel you need assistance with finding or maintaining a steady job?    |           |            |                   |
| 49 | Do you feel you need assistance with finding or maintaining a place to live? |           |            |                   |
| 50 | Will money be a problem for you over the next several months?                |           |            |                   |

|    | <b>How difficult will it be for you to...</b>                       | <b>Not Difficult</b> | <b>Somewhat Difficult</b> | <b>Very Difficult</b> |
|----|---|----------------------|---------------------------|-----------------------|
| 51 | manage your money?  |                      |                           |                       |
| 52 | keep a job once you have found one or if you currently have one?    |                      |                           |                       |
| 53 | find or keep a steady place to live?                                |                      |                           |                       |
| 54 | have enough money to get by?  |                      |                           |                       |
| 55 | find or keep people that you can trust?                             |                      |                           |                       |
| 56 | find or keep friends who will be a good influence on you?           |                      |                           |                       |
| 57 | avoid risky situations?   |                      |                           |                       |
| 58 | learn to control your temper?                                       |                      |                           |                       |
| 59 | find things that interest you?                                      |                      |                           |                       |
| 60 | learn better skills to get or keep a job?                           |                      |                           |                       |
| 61 | find a safe place to live where you won't be hassled or threatened? |                      |                           |                       |
| 62 | get along with people?  |                      |                           |                       |

# COMPAS Probation Risk Assessment

Offender: **Joe Sample**

DOB: **2/2/1950**

Gender: **Male**

Screening Date: **9/13/2007**

Screeener: **Hellem, Dan**

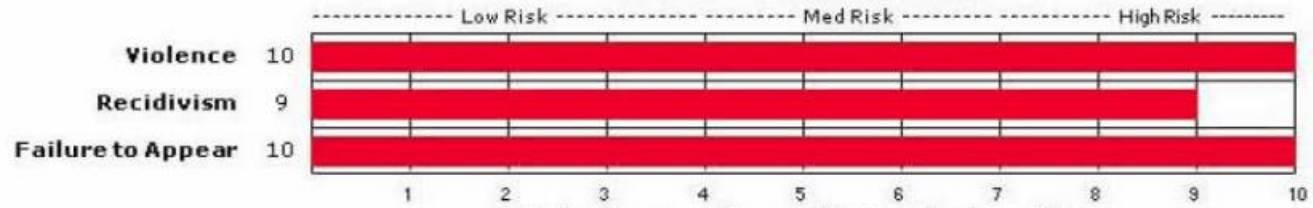
Ethnicity: **Native A**

Scale Set: **DMB-PSI**

Case: **009943**

Marital Status: **Single**

## Overall Risk Potential



## Criminogenic and Needs Profile



# Dressel and Farid (2018)

- A logistic regression model with age, sex, number of juvenile misdemeanors, number of juvenile felonies, number of prior (nonjuvenile) crimes, crime degree, and crime charge as the predictors did as well as COMPAS (around 65% correct rate)
- Untrained humans, presented with the same data, guessed correctly 62% of the time
  - If we had several people vote, they got the outcome correct 67% of the time
  - Both 62% and 67% are very close to 65%
- Consistent with what you saw in the project

# Fairness

- False positive parity
  - Probability of a person who won't recidivate staying in jail is the same regardless of the sensitive characteristic
- Accuracy parity
  - Everyone is classified correctly at the same rate regardless of sensitive characteristics
- Demographic parity
  - The same proportion of people is predicted to recidivate, regardless of demographics
- Calibration
  - "6/10" means the same thing regardless of race
  - $P(\text{recid} | \text{pred}=0)$  and  $P(\text{recid} | \text{pred}=1)$  are the same regardless of race
    - Equivalent to both PPV and NPV parity holding
- In general, can only mathematically have one



# Making the classifier fair

- We built a classifier that we found did not satisfy false positive parity
- Adjusted thresholds to make it fair

$$\begin{cases} \text{Yes if } y > 0.5, \text{ demographic} = A \\ \text{Yes if } y > 0.55, \text{ demographic} = B \end{cases}$$

- Increasing the threshold
  - Decreases the number of times we say “yes”
  - Will generally decrease the false positive rate, and increase the false negative rate
  - Moving the threshold away from 0.5 decreases accuracy (usually)

# A classifier is only as good as the data

- We are predicting re-arrests
- If the arrests themselves are biased, the classifier will be biased

- “Part of the field [of fairness in machine learning] is about answering the question "How do we make sure no one ever uses logistic regression to sentence or convict people" or something equally problematic” – George D.
- Counterpoint: are people better?
- Important to be aware of fairness and bias issues