

Big Data and Security

Jeffrey Borowitz, PhD

Lecturer

Sam Nunn School of International Affairs

AI For High Stakes Decisions with Multiple Goals

AI Based Decisionmaking

- There are winners and losers with any high-stakes decisionmaking
- But how can AI based decisions be better?

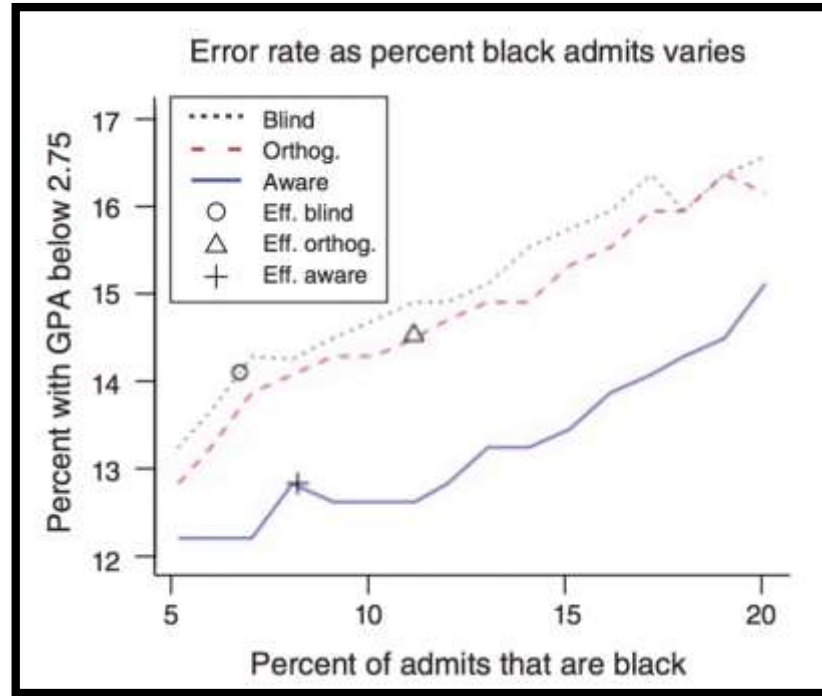
Education Example from Kleinberg et al

- College admissions is a great prediction problem
 - Leaving aside the bigger issues, we could say we want students to just “not get super bad grades” as an admission goal
- Two concerns:
 - “Efficiency” - how many students can we admit who will pass?
 - “Equity” - how many students who are admitted are from minority groups?
- How can we combine equity and efficiency?
 - Should we consider race?
 - Should we NOT consider race?
 - Should we “orthogonalize” with respect to race?
- Paper: Algorithmic Fairness

Education Example: Key Theorem

- Theorem
 - Either planner should use an “optimal” predictor to score all applicants, including using race
 - The equitable planner should choose the highest ranking applicants from each group, for a given preference towards equality.
- Key idea:
 - To get the best “efficiency”, you want to pick students within each group who will do the best
 - You can change “equity” by taking different numbers from each group

Results: Improved Performance and Minority Admissions



Implications of Model Errors: Criminal Justice Decisions

- We want the criminal justice system to be fair to individuals and also protect the public from crime
- A great recent strand of work in economics analyzes this:
 - Kleinberg et al: Human Decisions and Machine Predictions
- What is fairness?

Kleinberg et al (2018)

- Frame a criminal justice problem in terms of a prediction problem (here, decision to release before trial)
- Fit the prediction problem with a model
- Analyze the “outcomes”
 - Turns out this is hard because we don’t know if people in jail would have committed crimes
 - Use random assignment of judges with different leniency to different cases
- We can use people let out by lenient judges to see what might have happened if less lenient judges had let out “similar” people
- Results are impressive!

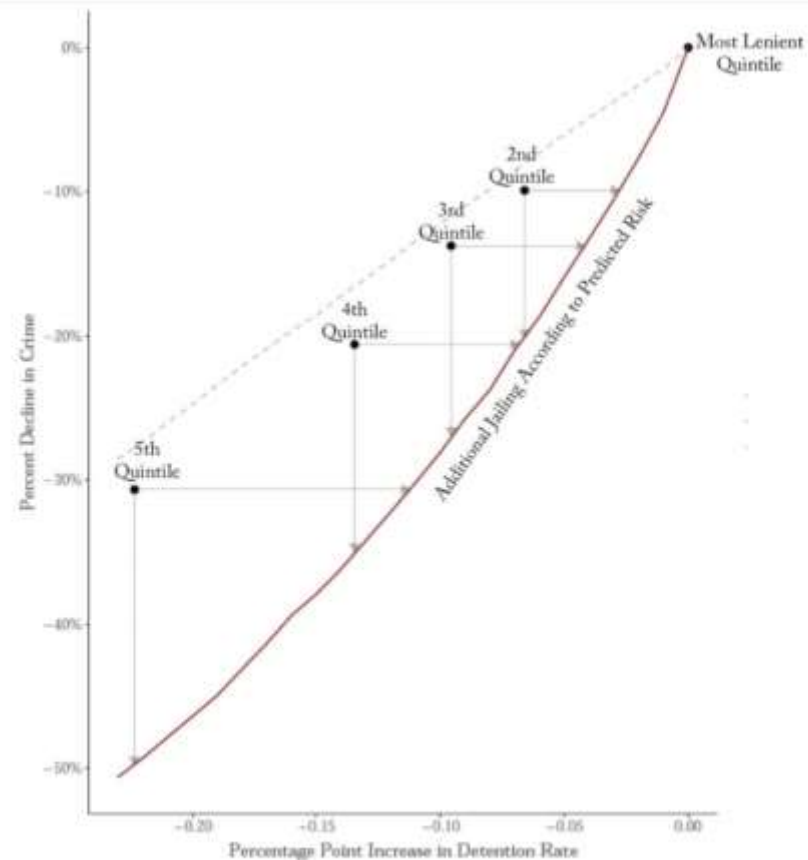


Figure 5: Comparing Impact of Detaining in order of Predicted Risk To What Judges Achieve

Conclusion: What's not to Like?

- This sounds like with careful work we can have better outcomes on a range of metrics
- But we have to make changes to the process
 - People will not be making as many decisions
 - Predictions feel more rigid.

Lesson Summary

- It's possible to simultaneously improve multiple dimensions of goals with good AI based models
- We saw this in
 - College admissions
 - Pretrial detention