**Case Study - Predictive Policing** (Electronic submission due 2/27 at 11:59 pm)

Your task: Working together as a group, completely answer the four questions asked in this case study. Use readings from class to inform your answers. Each answer should be approximately 200 words in length.

1. Priming question: **Predictive policing systems identify places and people who are at greatest risk of crime. This is a type of decision that has traditionally been made by humans that is now being “outsourced” to computer systems. Are there decisions that you would only ever trust a human police officer (and never trust an algorithm alone) to make? If you would only trust a human officer with certain decisions, what do you think is special about human decision-making? If you *would* trust an algorithm with any decision, how accurate would it need to be in order for you to trust it?**

Part1 Increasing racial bias

As we saw in lecture, place-based predictive police algorithms could amplify biases that are present in historical crime data. For instance, if police have historically made more arrests in black parts of town because of racial bias, then black individuals might be more likely to reside in an area predicted to be a crime “hot spot” by a predictive policing algorithm. If police send more patrol cars to places that the algorithm has labeled hot spots, then black residents will receive more attention, leading to even more arrests in that area. The end result would be a “runaway feedback loop”, in which the predictions made by an algorithm are continuously reconfirmed.

* “If we consider that institutional racism in this country is an ongoing unnatural disaster, then crime prediction algorithms should more accurately be called crime *production* algorithms. The danger with New Jim Code predictions is the way in which self-fulfilling prophecies enact what they predict, giving the allure of accuracy.” - Ruha Benjamin, *Race After Technology*
* *“*Hiring algorithms only receive feedback on people who were hired, predictive policing algorithms only observe crime in neighborhoods they patrol, and so on. Decisions made by the system influence the data that is fed to it in the future. For example, once a decision has been made to patrol a certain neighborhood, crime discovered in that neighborhood will be fed into the training apparatus for the next round of decision-making.” - Ensign et al., “Runaway Feedback Loops in Predictive Policing”[[1]](#footnote-1)

1. **Is the bias in predictive policing described above an example of preexisting bias, technical bias, or emergent bias? Maybe it’s more than one? Do you think that a predictive policing algorithm could escape the “feedback loop” problem? If so, how? If not, why not?**

Part 2 Introducing racial bias

Boonin argues that even if crime data is not influenced by racial bias in the way police choose to enforce the law, a predictive policing algorithm might *still* introduce racial bias in policing. On pages 40 – 43 of his manuscript, Boonin imagines an example in which Black residents of a city live in more densely populated areas, while White residents live in less densely populated areas. Even though the crime rates are the same in both areas (crimes per resident), there is more crime per square mile in the more densely populated parts of town. This is simply because there are more people per square mile. Due to the difference in crime per square mile in black and white parts of town, a place-based predictive policing system identifies more high-risk areas in the black parts of town. The result is that the Black population of the city receives more attention from police.

1. **Explain why this scenario might be unfair or unethical. Then give what you think is the best defense of using the algorithm in spite of its disproportionate effects (assume that all data is true and that the algorithm itself is procedurally fair). Which side is more convincing?**

Part 3 Which bias is worse: Human or machine?

As discussed above, algorithms that are trained to identify patterns in historical data are likely to replicate biases present in that data. For instance, if police arrest Black people at a higher rate due to racial animus, then predominantly Black areas of the city will be more likely to be labelled as historically “high crime”.

Keeping this in mind, read the following passage very carefully. In a 2021 report entitled *Garbage In, Gospel Out*,[[2]](#footnote-2) the National Association of Criminal Defense Lawyers writes,

*Police departments must not utilize data-driven policing technologiesbecause they are ineffective; lack scientific validity; create,* ***replicate and exacerbate “self-perpetuating cycles of bias”; deeply entrench existing inequities in the system****; hyper-criminalize individuals, families, and communities of color; and divert resources and funds from communities that should be allocated towards social services and community-led public safety initiatives.*

We might think that many of those concerns are not about predictive policing specifically, but of policing generally. This raises a question whether using predictive policing makes things worse than they were before the introduction of algorithms in policing.

1. **If predictive policing algorithms are bad because they preserve pre-existing bias, would their use be any worse than relying on humans who are probably also biased? Why or why not? Is there a special ethical problem with using a data-driven method that reaches the same racially biased conclusions as people would? Could predictive policing algorithms improve on human biases? How?**

1. <http://proceedings.mlr.press/v81/ensign18a/ensign18a.pdf> [↑](#footnote-ref-1)
2. <https://www.nacdl.org/getattachment/eb6a04b2-4887-4a46-a708-dbdaade82125/garbage-in-gospel-out-how-data-driven-policing-technologies-entrench-historic-racism-and-tech-wash-bias-in-the-criminal-legal-system-09142021.pdf> [↑](#footnote-ref-2)