PROFILING GROUPS AT RISK OF RECIDIVISM AND CREATING SOLUTIONS TO REDUCE RECIDIVISM WITHIN EACH GROUP

UNDERSTANDING HOW POORLY DESIGNED ALGORITHMS CAN IMPACT PERCEPTIONS OF PREDICTIVE JUSTICE: A CASE STUDY OF PREDPOL

A Thesis Prospectus
In STS 4500
Presented to
The Faculty of the
School of Engineering and Applied Science
University of Virginia
In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science in Systems Engineering

By
Mohini Gupta

December 5, 2024

Technical Team Members: Zakaria Afi, Sudarshan Atmavilas, Sarah Bedal, Olivia Bernard, Caroline Lee

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

ADVISORS

Kent Wayland, Department of Engineering and Society

Dr. Michael C. Smith and Dr. K. Preston White, Department of Systems and Information Engineering
Dr. Peter Alonzi, School of Data Science

Introduction: How do trends within the justice system impact machine learning algorithms and relevant groups?

Historically, problematic trends and judgments have been present in the justice system, and they have influenced people's perceptions of groups of people and criminal activity. In the past, biased perceptions of criminal activities could only impact the justice system through prejudiced actions from the members of it. However, in the past few decades, the rise of machine learning algorithms has changed the way biases impact the justice system. Police use algorithms to get predictions on who will be committing crimes, where crimes will be committed, and risks that community members may pose (Lum & Isaac, 2016). The algorithms make decisions based on data, and when concerning humans, the data is often biased based on prejudices and discriminatory behaviors from people in the past (Christian, 2020). These biases can be especially problematic in the justice system, as reputations, livelihoods, and futures are at stake for the people involved (Goldkamp, 1987).

These issues have been found within modern-day technologies, with an example being PredPol, a predictive policing technology. PredPol was an algorithm used to predict when and where crimes would occur based on past data (Field, 2020). After its issues came to light, there was a lot of public backlash, and the public started to become more wary of predictive policing (Bhuiyan, 2021). A science, technology, and society (STS) study will be conducted on the factors that led to the failures of PredPol and how those failures influenced society's perception of predictive policing.

Human bias is not the only influencing factor for datasets used by justice algorithms. Recidivism, the act of a convicted criminal reoffending, has also been a factor in influencing datasets. When individuals reoffend, they saturate the dataset. This can cause biases to become

exaggerated in datasets and can also cause the machine learning algorithms to learn and associate behaviors of a small group of people and generalize them to larger groups of people. For a deeper dive into other factors influences criminal justice biases, my capstone project team and I will be working with the Albemarle-Charlottesville Regional Jail and the Charlottesville Police department to find trends in recidivism and ways to reduce it by profiling reoffenders and creating aid plans for each specific group.

Technical Research Problem: Profiling of Groups at Risk of Recidivism and Creating Solutions to Reduce Recidivism within Each Group

Which solutions can be created to reduce recidivism by targeting groups of people with similar characteristics?

In recent years, the jail and police departments servicing the Charlottesville and Albemarle areas have noticed a concerning trend within the population of people getting arrested and going to jail. They realized that there was a large group of high-utilizers, people who would frequently be arrested and come back to jail due to them recommitting crimes after being released.

Why would the jail and police be concerned by this pattern? To understand this, our capstone team spoke with Colonel Martin Kumer, the superintendent of the Albemarle-Charlottesville Regional Jail (ACRJ). There are two aspects making this problem concerning: resource utilization and re-entry effectiveness. In the words of Kumer, in the ACRJ, high-utilizers make up 10% of the jail's population, but they consume about 90% of the jail's resources (Martin Kumer, personal communication, September 20, 2024). From a resource perspective, this makes the population of high-utilizers extremely inefficient to maintain. The

other cause for concern comes from the human outreach side. The ACRJ has a re-entry team whose sole purpose is helping offenders settle back into society and checking on them to make sure they aren't at risk of reoffending (Martin Kumer, personal communication, September 20, 2024). Since high-utilizers reoffend, this means that their needs are no longer being met after release. Factors that put individuals at risk for recidivism are substance abuse, association with "the wrong crowd", problems with mental health, and low income, and unemployment (Yukhnenko et al., 2020). These factors will be used to create profiles of high-utilizers in our study, as well as to find patterns between these factors and recidivism.

The primary objective of our capstone team is to aid high-utilizers with re-entry. We have two areas of interest for improving recidivism: diverting high-utilizers pre-arrest and aiding high-utilizers out of jail. For the cases of both arrests and jail high-utilizers, we plan to profile high-utilizers to create groups of them with common traits. Then, by conducting analysis on arrest details within those groups, we will identify patterns that differ from those of the population of offenders who do not go through recidivism.

For the steps detailed above, we will get data on jail stays and arrests from the ACRJ and Charlottesville and Albemarle police departments respectively. To profile high-utilizers, we will first use basic sorting and filtering techniques, and then we will use statistical software packages, such as Minitab, to ensure each group is statistically different from each other and from the general jail population. To identify basic trends in arrests, we will create different graphs to identify trends between different variables. For a deeper dive into trends, we will use machine learning and regression analysis techniques to understand how different combinations of factors can cause different outcomes within groups of people, and specific models that we plan on using for trend analysis are linear regression models and logistic regression models.

The insights and recommendations generated from our analysis will be presented to the stakeholders to help them make informed decisions. We will work in conjunction with the ACRJ and police to develop feasible solutions to help each of the different groups of high-utilizers. For solutions generation, input will be needed from our stakeholders to create solutions within their scope. Our project will not have an immediate impact on high-utilizers, but our hope is that once solutions are rolled out by the ACRJ and police, their quality of life will improve, causing them to be less likely to undergo recidivism.

STS Research Problem: Understanding How Poorly Designed Algorithms Can Impact Perceptions of Predictive Justice: A Case Study of PredPol

How are society members and structures impacted by decisions made with criminal justice algorithms?

PredPol was a predictive policing software that now goes by Geolitica (*How We Assessed the Accuracy of Predictive Policing Software*, 2023). PredPol was developed by researchers at the University of California, Los Angeles (UCLA) in collaboration with the Los Angeles Police Department (LAPD) (Field, 2020). It was developed to identify individuals who might be linked to violent crime to stop crime before it occurred (Field, 2020). However, when analyses were run on PredPol results, it was found that police were most often directed to low-income and marginalized communities (*How We Assessed the Accuracy of Predictive Policing Software*, 2023). Due to public backlash, the LAPD stopped using PredPol in 2020 (Field, 2020).

To understand how PredPol affected people's attitude towards the use of predictive algorithms in criminal justice, first a general understanding must be created of the relationship between algorithms and the justice system. What are factors that make criminal justice

algorithms biased? Why do people care about decisions made by those algorithms? Who should be held responsible for making sure fair outcomes are produced with the use of algorithms? In the past couple of decades, criminal justice algorithms have become more prevalent as a way of predicting crime and helping judges make verdicts (Sousa et al., 2024, Chapter 14, p. 282). Though predictive methods can be helpful, problems emerge when data is biased (Floridi, 2023). When criminal justice algorithms come up with a decision based on biased data, there are severe negative impacts for groups involved. In the case of PredPol, residents unfairly faced increased police surveillance and suspicion due to the biased predictions.

Understanding Machine Learning in the Justice System

Criminal justice algorithms have the potential for good. Studies have found that judges would be more likely to issue favorable sentences after having lunch, which is a source of bias that a machine learning algorithm would not have (Taylor, 2023). Though some sources of bias are missing from the algorithms, other more problematic sources cannot be replaced so easily.

There are two major entry points for biases to enter machine learning results: data collection and the review process (Travaini et al., 2022). Since algorithms are created from data collected over a long period of time, historical biases within people and in the main players of the justice system have caused data being fed into algorithms to have those biases present within it. Biases still exist in modern day policing and are commonly caused by implicit preconceived notions about groups of people (Spencer et al., 2016). Algorithms cannot learn outside of the data that is given to them, so if the data being fed into it is biased, it will also be biased (Sousa et al., 2024, Chapter 14, p. 283). As the users of policing algorithms, responsibility for checking those decisions for biases falls upon the police (Dignum, 2019).

Society's opinions on the usage of decision-making algorithms have been mixed. Police departments tend to have more favorable opinions, as they believe the algorithms could help reduce discrimination and increase workplace efficiency (Bell et al., 2021). Views among community members, especially those of marginalized groups, tend to be more mixed, as they are the ones most negatively impacted by faulty or biased decisions (Lum & Isaac, 2016). Due to controversies caused by predictive algorithms, members of the Senate have also begun questioning the ethics of using predictive methods (Wyden et al., 2024). Senate members have concerns about how marginalized groups would be affected (Wyden et al., 2021). The opinions and perspectives of different groups on predictive policing will be used as evidence for the STS research.

Methodology

The main part of the research will be evidence-based to understand how society affected the successes and failures of PredPol and to gain insight to changes being made in light of PredPol's failures. To understand public opinions on predictive policing and how opinions have changed after the release of PredPol, statements from public figures, police departments, and scholarly sources will be examined to find out how opinions have changed over time. Two letters written by Senate members, Letter to DOJ on Predictive Policing Tech and Letter to DOJ - Predictive Policing & Title VI, will be used to understand changes in lawmakers' opinions. A 2024 bill for legislation against the use of predictive algorithms in criminal justice will also be used to understand current opinions of lawmakers (Justice in Forensic Algorithms Act, 2024). Reports from UCLA on PredPol will be used to analyze how academic and police opinions on PredPol changed from when it was implemented to the time that the LAPD stopped using it

(Field, 2020). Currently, there is no literature that ties in specific factors that led to changes in public opinion about PredPol, so the aim of the research is to understand how society reacted to specific failures of PredPol.

The Effect Predictive Algorithms and the Justice System have on Each Other

To further analyze the relationship between society and PredPol, the value-neutrality thesis and the concept of technological momentum will be applied. The value-neutrality thesis states that technologies are unbiased, and biases come from the users and creators. Machine learning is created and trained through human behaviors and human data, so by understanding human behaviors, machine learning processes can be better understood. PredPol is a technology, so it inherently cannot produce biased decisions, so to understand its failures, the behaviors of its creators and the police must be understood. Technological momentum is the idea that a technology's influence increases as it becomes more embedded in society. Machine learning algorithms have been around for a while, and people have started to become more reliant on it. This reliance has also come into play with law enforcement, as they have been trying to implement predictive policing for over a decade (Vargas, 2023). By understanding factors that made algorithms important for police, we can better understand how the relation between the two entities introduced failures into the PredPol system.

Conclusion

Through the STS research, I aim to understand the failures of PredPol and the effects of those failures on users of the justice system, so that future creators of predictive algorithms are

aware of and more mindful of algorithmic biases. Predictive policing algorithms have the potential to be helpful in making communities safer, but as they are now, there are too many problems without solutions. The technical research will uncover reasons behind reincarceration to find out how the law can aid people who reoffend. By understanding recidivism trends, sources of biases in the justice system can be uncovered. Once lawmakers and decision-making groups understand the significance of reducing those biases, they can start developing next steps to reduce biases.

References

Bell, K., Hong, J., McKeown, N., & Voss, C. (2021a, December). *A new direction for machine learning in criminal law*. Human Centered Artificial Intelligence. https://hai.stanford.edu/sites/default/files/2021-12/Policy Brief - A New Direction for Machine Learning in Criminal Law.pdf

Bhuiyan, J. (2021, November 8). *LAPD ended predictive policing programs amid public outcry. A new effort shares many of their flaws*. The Guardian; The Guardian. https://www.theguardian.com/us-news/2021/nov/07/lapd-predictive-policing-surveillance-reform

Christian, B. (2020). *The alignment problem: Machine Learning and human values* (1st ed.). W.W. Norton & Company.

Dignum, V. (2019). *Responsible artificial intelligence how to develop and use AI in a responsible way virginia dignum aut* (1st ed.). Springer.

Field, S. (2020, June 23). *Predictive Policing: What Is It And Should It Be Used In 2020?* UC Institute for Prediction Technology; UCLA. https://predictiontechnology.ucla.edu/predictive-policing-what-is-it-and-should-it-be-used-in-2020/

Floridi, L. (2023). The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities. Oxford University Press.

Goldkamp, J. S. (1987). Prediction in criminal justice policy development. *Crime and Justice*, *9*.

> https://doi.org/10.1086/449133

How We Assessed the Accuracy of Predictive Policing Software. (2023). Policingequity.org;

> Policing Equity. https://policingequity.org/newsroom/press/how-we-assessed-the-
> accuracy-of-predictive-policing-software

Justice in Forensic Algorithms Act, H.R. 7394, 118th Congress. (2024).

> https://www.congress.gov/bill/118th-congress/house-bill/7394/text

Lum, K., & Isaac, W. (2016). To predict and serve? *Significance*, *13*(5), 14–19.

> https://doi.org/10.1111/j.1740-9713.2016.00960.x

Sousa Antunes, H., Freitas, P. M., Oliveira, A., Martins Pereira, C., Vaz de Sequeira, E., &

> Barreto Xavier, L. (Eds.). (2024). Multidisciplinary Perspectives on Artificial Intelligence
> and the Law, *Judicial Decision-Making in the Age of Artificial Intelligence Law* (pp. 281-
> 297). *Governance and Technology Series*. https://doi.org/10.1007/978-3-031-41264-6

Spencer, K. B., Charbonneau, A. K., & Glaser, J. (2016). Implicit Bias and Policing. *Social and*

> *Personality Psychology Compass*, *10*(1), 50–63. https://doi.org/10.1111/spc3.12210

Taylor, I. (2023). Justice by Algorithm: The Limits of AI in Criminal Sentencing. *Criminal*

> *Justice Ethics*, 42(3), 193–213. https://doi.org/10.1080/0731129X.2023.2275967

Travaini, G. V., Pacchioni, F., Bellumore, S., Bosia, M., & De Micco, F. (2022). Machine

> Learning and criminal justice: A systematic review of advanced methodology for

recidivism risk prediction. *International Journal of Environmental Research and Public Health*, *19*(17). https://doi.org/10.3390/ijerph191710594

Vargas, A. (2023, August 29). *Reporting on the long shadow of the LAPD's data-driven policing programs | USC Center for Health Journalism*. Centerforhealthjournalism.org. https://centerforhealthjournalism.org/our-work/insights/reporting-long-shadow-lapds-data-driven-policing-programs

Wyden, R., Clark, Y., Merkley, J., & Padilla, A. 118th Congress. (January 24, 2024). ["Letter to DOJ - Predictive Policing & Title VI". This is a letter written by Senate members to the DOJ with questions about how they plan to uphold their policies with the introduction of predictive policing softwares.] https://www.documentcloud.org/documents/24389851-letter-to-doj-predictive-policing-title-vi-jan-24-2024/

Wyden, R., Clark, Y., Markey, E., Warren, E., Merkley, J., Padilla, A., Warnock, R., & Lee, S. 117th Congress. (April 15, 2021). ["Letter to DOJ on Predictive Policing Tech". This is a letter written by Senate members to the DOJ with questions about the monitoring, oversight, and budgeting for predictive policing algorithms.] https://www.wyden.senate.gov/imo/media/doc/041521%20Wyden%20Led%20Letter%20to%20DOJ%20on%20Predictive%20Policing%20Tech.pdf

Yukhnenko, D., Blackwood, N., & Fazel, S. (2020). Risk factors for recidivism in individuals receiving community sentences: a systematic review and meta-analysis. *CNS spectrums*, 25(2), 252–263. https://doi.org/10.1017/S1092852919001056