Applications of Machine Learning to Pretrial Sentencing: A Literature Review
(Technical Paper)

The Social Construction of Digital Justice

(STS Paper)

A **Thesis Prospectus Submitted to the**

Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements of the Degree
Bachelor of Science, School of Engineering

Jalen Gregory
Fall, 2020

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.

Signature _Jalen Gregory_ Date 05/11/2021
     Jalen Gregory

Approved _Jundong Li_ Date 04/29/2021
     Jundong Li, Department of Computer Science

Approved _____ Date _____
     Travis Elliott, Department of Engineering and Society

Introduction

In the foreword to his book *Amusing Ourselves to Death*, Neil Postman (1985) takes time to shine light on the juxtaposition between the prophetic visions of two the 20th century's greatest science fiction novels and their authors, Orwell's *1984* and Huxley's *Brave New World*. He summarizes this difference in saying, "In short, Orwell feared that what we hate will ruin is. Huxley feared that what we love will ruin us". In the information age, there is nothing as ubiquitous as data, the processing and analysis of information that we love being refined for the purpose of rendering the uncertainty that we hate a thing of the past. Machine learning has made it possible to take massive amounts of information and distill them into assumptions, if not outright assertions, and this has made data more powerful than ever. The question to ask is whether the insights gained from these machine learning algorithms are truly reflective of the world we aim to simulate and predict, or if they are simply reinforcing the beliefs of the engineers that build and entities that fund their creation.

The technical project in this thesis aims to answer that question in the context of criminal justice algorithms and predictive policing. There has been backlash over the ethical legitimacy as well as the general efficacy of these models. In my technical project, I will undergo a review of literature on the subject of using machine learning algorithms in criminal justice settings in order to find what experts believe to be the core concepts to be studied as these practices are further researched and developed.

In the STS research paper, I will attempt to understand the creation and deployment of these predictive policing algorithms through the lens of the Social Construction of Technology. Software is man-made, as is the law, and at the cross section lies a subject fraught with controversy. The creation, deployment, and results of the use of these algorithms affect and are

viewed differently by the varying social classes living on the other end of the equation. As Pinch and Bijker (2012) put it in *The Social Constructions of Facts and Artifacts*, "…both science and technology are socially constructed cultures and bring to bear whatever cultural resources are appropriate for the purposes at hand". Given the lasting effects of an encounter with the criminal justice system, it is worth asking exactly what cultural resources we are bringing to bear in using technology to solve this problem.

Technical Project: Machine Learning Algorithms in Criminal Justice, a Literature Review

Project Background

Algorithms help run the everyday world, even when people are not aware of their presence. In its base form, an algorithm is essentially a set of instructions that allow a set of inputs to produce an expected output. A popular example given to students by teachers and professors is the algorithm of making a peanut butter and jelly sandwich. This process is used to illustrate to students the steps that they inherently skip or take for granted when approaching a solution to a problem, steps that a computer is not able to skip. For years, even the most complex algorithms required expert knowledge to both transform into inputs and to assess the meaning of the outputs. In his proposal for a federal administration to regulate algorithms, Andrew Tutt (2016) describes how algorithms as large and widely used as Google's PageRank and the Deep Blue algorithm created by IBM to play chess against world champions were programmed with inputs largely tuned by human hand. Machine learning, the process by which algorithms refine themselves, has changed the goal of algorithmic programming from one of training a machine to be good at a task to one of training a machine to learn what it means to be good at a task and how to get better (Tutt, 2016). Handing the reins over to the algorithm itself makes it difficult if not

outright impossible to trace the "thinking" that led to any given decision, obfuscating the reasoning behind what could be a crucial decision.

The use of data in policing is not new but it has seen novel uses if not an outright renaissance in the machine learning age. Named one of the best inventions of 2011 by Time magazine, predictive policing is "the usage of mathematical, predictive analytics, and other analytical techniques in law enforcement to identify potential criminal activity" (Predictive policing, 2020). Criminal justice entities face many of the same dilemmas faced in the private sector caused by issues of supply and demand in any number of areas. When police department budgets are stretched too thin or they are simply overrun with work, decisions must be made as to where to allocate their officers and how to approach the citizens in those areas. Judges in districts with overcrowded prisons often look for alternative solutions for those who they believe are less likely to recommit crimes but may not have all the information at hand to make individualized decisions in every case. For those not so lucky as to be offered the alternative, the judges are still cognizant of a public that expects them to be fair in their judgment and not deliver overly harsh punishments based on information not necessarily relevant to the facts of the case. Machine learning can be and is used to guide all these decisions, but by its very nature makes the decisions reached difficult to predict and understand. The pros and cons of "big data policing" are varied, and that is part of what I hope to uncover in my literature review.

Literature Review Method

In my literature review, I intend to survey scholarly articles on the practice of using data analytics in criminal justice to answer a few questions that immediately come to mind for me and hopefully to uncover those being focused on currently by those undergoing the research. These sources will include advocates for the use of these algorithms such as Dr. George Mohler, the

Chief Data Scientist and co-founder of PredPol, one of the most recognizable names in the predictive policing space as well as those such as Danielle Ensign et. al (2017) who published "Runaway Feedback Loops in Predictive Policing", in which they created their own small version of PredPol and showed how it would inherently lead to biased results. Another important source I believe will be *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*, part of a research series by members of the RAND corporation. This was the result of a funding effort by the National Institute of Justice to help police understand the tools at their disposal and the benefits and pitfalls of each of them (Perry et. al, 2013). In researching the way that these algorithms are already in use, it will be important to understand to what extent even the police departments are aware of the effects of the software they pay for.

The questions I have on the use of machine learning in criminal justice settings deal with the relationship behind their results and the path taken to get there. In other words, how is the effectiveness of the algorithms being fine tuned in relation to the ethical dilemma of using the algorithms in the first place. Answering this will first require establishing what it means for an algorithm to be successful vs. unsuccessful, something that is clear in some cases and less so in others. Is there agreement between software engineers, criminal justice officials, and the public on the goal in using these algorithms? After establishing what success means, the next step will be to determine in what ways the implementation of machine learning is achieving and failing to achieve the expected goal. As shown in the difference in findings between Mohler and Ensign et. al. in looking at similarly designed systems but reaching different conclusions, the output given by methods already in use are subject to interpretation. Lastly, I hope to understand the ethical concerns faced in the use of these algorithms and whether they are being considered or whether they are being shoved aside in an "ends justifies the means" sense. In undergoing this literature

review, I hope to gain a better understanding of the present and future of both the use of data and the criminal justice system.

STS Research Paper: The Social Construction of Digital Justice

Introduction

The goal of this STS Research Paper will be to study how the use of machine learning algorithms in criminal justice came about, what their implementation means to the varying social classes that encounter them, and how that will determine their use going forward. In their current form, most implementations of these algorithms are closed off to the public or "black-box", meaning what is going on under the hood is only available to those writing the code, leaving everyone else to base their opinions of them on their output. AI Now Institute co-founder Meredith Whittaker worries that that closed systems cut off from public judgment are dangerous, and "went so far as to recommend that no public agencies responsible for such matters as criminal justice, health care, welfare, and education should use black box AI systems" (Rieland, 2018). On the other side of the spectrum are those like UVA Law alum and Fourth Circuit judge James Wilkinson III, who defended the practice of predictive policing in the face of what his fellow judges considered a clear 4th amendment violation in the case of US v. Curry (2020). He lamented that it is the part of the country that historically has the least protection whom "ends up shouldering the weightier burden when our branch of government oversteps its proper role". The spectrum of opinions on the use of algorithms makes the social construction of technology an ideal way to study their use and implementation as they are open to widely different interpretations based on the vantage point they are observed from.

STS Framework: Social Construction of Technology

As explained by Pinch and Bijker (2012), it is important to realize that "all knowledge and all knowledge claims are sought in the domain of the social world rather than the natural world". While it is largely understood that the social beliefs of the time determine the laws, it is less widely understood that technology is susceptible to this same process. Algorithms, presented as infallible as math equations, can give the impression that they are unbiased, but "if biased data is used to train these predictive models, the models will reproduce and in some cases amplify those same biases" (Lum & Isaac, 2016). Knowing this, it becomes critical that a study of the relevant social groups is undertaken to ensure that all their concerns are met. Unfortunately, as often is the case, the creation of the algorithms is left in the hands of a select few who may or may not have those concerns in mind.

At the forefront of the groups with an interest in criminal justice algorithms are the criminal justice officials who pay for their use. In the wake of mass protests in the summer of 2020, NYPD commissioner Dermot Shea, partially in response to calls for budget cuts, stated the need for his department of "better utilizing data, intelligence and all the technology at our disposal…That means for the NYPD's part, we'll redouble our precision-policing efforts" (Heaven, 2020). He is not nearly alone in this belief. It was former LAPD commissioner Charlie Beck whose staunch support for the algorithms that led to them becoming more widely accepted across the nation. It was his belief that they were invaluable as they shift "law enforcement from focusing on what happened to focusing on what will happen and how to effectively deploy resources in front of crime, thereby changing outcomes" (Pearsall, 2010). As stated before, other public officials find benefits in their use as well. For example, in 2014, judges in Virginia using risk assessment tools trained by machine learning to determine the likelihood of defendants

committing future crimes sent half of their defendants to alternatives to prison, slowing prison population growth to 5% between 2005 and 2016 from 31% in the 90's (Angwin et. al, 2016).

Outside of the criminal justice system, reaction to the creation of these algorithms is less positive. Experts on artificial intelligence, such as those at the aforementioned AI Now Institute, have come out and lambasted their use, criticizing the championing of these algorithms as little more than "tech-washing, where a veneer of objectivity covers mechanisms that perpetuate inequities in society" (Heaven, 2020). In a letter titled "Abolish the #TechToPrisonPipeline", 2435 experts across different fields lent their signature for the rescinding of publication of a research article that claimed to be able to predict criminality based on facial recognition on the basis that attempts to predict crime through technology do far more harm than good to the societies they aim to protect (Coalition for Critical Technology, 2020). As technology becomes more complex and fewer people have the knowledge to assess the full impact of its adoption, for so many academic researchers in different areas pertaining to machine learning to come to agreement that these algorithms are potentially as toxic as the crimes they aim to prevent must be relevant. The other main social group left in the equation is the citizenry that fall under the watchful eye of criminal justice algorithms. Some believe that they have nothing to hide, while others subscribe to the Edward Snowden belief that giving up rights because one has nothing to hide is akin to giving up freedom of speech because one has nothing to say. Many are not even aware of the implementation of software in the criminal justice system to the extent that it already has been. An LA resident, Hamid Khan, was so alarmed by his discovery of these algorithms that he has dedicated a large part of his life to fighting their adoption and raising awareness to fellow citizens. He demanded an audit of the algorithm used by PredPol but was rejected due to being told that the system was too complicated to determine pathways that led to

the outputs given (Heaven, 2020). With the groups divided so fundamentally on the use of the algorithms, the next step of bringing them together to a conclusion seems difficult to achieve.

The malleable nature of the law combined with the constantly transforming algorithms in use to assess them renders attempts at closure unlikely. The closure mechanism most reflective of the situation in my opinion is redefinition of the problem. As our societal attitude toward criminal justice ebbs and flows the clashes between the laws we create to govern and the software we create to help deal with the consequences will shift. In assessing criminal justice algorithms in relation to the wider sociopolitical milieu, it could be said that they are emblematic of our current moment: technology leaned on to solve problems that the engineers that create them are unable to solve themselves. There is a technological hubris in their use that is representative of our nation as a whole.

Thesis

In developing an understanding of the way that social movements and interpretative flexibility have an impact on the creation of smart algorithms, I plan to survey beliefs from different members of the relevant social groups to better learn their stance. By doing this, I hope to shed light on whether the algorithms are being employed in the interests of the majority, or if there is a smaller group with greater power that benefits from their use.

Conclusion

Data is information, information is knowledge, and knowledge is power. The benefits of algorithms that train themselves to be better at solving problems that plague humans than humans themselves are vast, but not without detriments. The technical project aims to extract from a range of papers the key concepts and questions that have been and still need to be

explored as machine learning expands further into the criminal justice field. The STS Research Paper aims to assess the impact of these algorithms on the public at large and how they are interpreted by the different groups that they affect using SCOT as the STS framework. Together, they should serve as a guide to the present and future of data analytics in criminal justice and the affect it will have in our country. Martin Luther King Jr. emphasized in his speech *The Three Evils* in 1967: "When scientific power outruns moral power, we end up with guided missiles and misguided men...It is this moral lag in our thing-oriented society that blinds us to the human reality around us and encourages us in the greed and the exploitation which creates the sector of poverty in the midst of wealth" (Doalmsavid 2013). It could be said that our scientific power has long outrun our spiritual power, and that algorithms are now that guided missile, taking in the information we give them as coordinates and constantly being adjusted until they hit the targets that we have programmed them to. Predictive policing, if it is to be trusted at all, must be more than a mechanism by which we identify the weakest and most in need of our society and monitor them further. If we as software engineers are going to be guiding missiles, they should be aimed at the foundation of the problems of our society, not at the results of them.

References

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine Bias. *ProPublica.* https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Coalition for Critical Technology. (2020, June 23). Abolish the #TechToPrisonPipeline. *Medium.* https://medium.com/@CoalitionForCriticalTechnology/abolish-the-techtoprisonpipeline-9b5b14366b16

Doalmsavid, E. J. (2013, September 14). Martin Luther King The Three Evils of Society[Video]. YouTube. https://youtu.be/j8d-IYSM-08

Ensign, D., Friedler, S., Neville, S., Scheidegger, C., & Venkatasubramanian, S. (2017, December 22). *Runaway Feedback Loops in Predictive Policing.* https://arxiv.org/pdf/1706.09847.pdf

Heaven, W. D. (2020, July 17). Predictive policing algorithms are racist. They need to be dismantled. *MIT Technology Review.* https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/

Lum, K. & Isaac, W. (2016). To predict and serve? *Significance, Vol. 13 (5)*, 14-19. https://doi.org/10.1111/j.1740-9713.2016.00960.x

Pearsall, B. (2010). Predictive Policing: The Future of Law Enforcement? *National Institute of Justice Journal, Vol. 1* (266), 16-19. https://www.ncjrs.gov/pdffiles1/nij/230409.pdf

Perry, W., McInnis, B., Price, C., Smith, S., & Hollywood, J. (2013). Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations. *RAND Corporation Research Reports.* https://doi.org/10.7249/RR233

Pinch, T., & Bijker, W. (2012). The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other. In Douglas D. (Author) & Bijker W., Hughes T., & Pinch T. (Eds.), *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology* (pp. 11-44). Cambridge, Massachusetts; London, England: The MIT Press.

Postman, N. (1985). *Amusing Ourselves to Death.* Viking Penguin.

Predictive policing. (2020, November 29). In *Wikipedia.* https://en.wikipedia.org/wiki/Predictive_policing

Rieland, R. (2018, March 5). Artificial Intelligence is Now Used to Predict Crime. But Is It Biased? *Smithsonian Mag.* https://www.smithsonianmag.com/innovation/artificial-intelligence-is-now-used-predict-crime-is-it-biased-180968337/

Tutt, A. (2016). An FDA for Algorithms. *Administrative Law Review, Vol. 69 (1),* 83-124.

United States v. Curry, No. 18-4233 (4th Cir. 2020). https://law.justia.com/cases/federal/appellate-courts/ca4/18-4233/18-4233-2020-07-16.html