SA 315 Algorithms in the Criminal Justice System

Bintu Alkassoum [redacted]

Spring 2019

[redacted]

In this paper, I will discuss the implications of algorithm use within the criminal justice system. In doing so, I will consider the following three stages of the criminal justice system: (1) the police system; (2) the courts; and (3) the prison system. Firstly, I argue that the police system uses algorithms as a mechanism of justifying discriminatory policing practices. In police departments, algorithms are used to predict where crime is likely to occur and who is likely to commit a crime in the near future (Joh, 2017). These predictive models are fed crime-related data, often tainted with human biases and assumptions (Joh, 2017). Unsurprisingly, the data that tends to be generated mirrors the data the algorithms were originally given (Joh, 2017). As Joh (2017) put it, "while the adoption of algorithmic decision-making may appear to solve issues about resources, efficiency, and discretion, a closer look at the "raw data" fed to these algorithms reveals some familiar problems (p. 17)." The harm in this is that racial biases, which have long been embedded in the American criminal justice system will perpetuate in a continuous cycle. The only difference being that the mechanisms reinforcing it will be concealed behind a false conception of value-free objectivity.

Afterward, this paper will examine the impact algorithms have had on judicial decision-making as it relates to sentencing. For the last decade, judges have started utilizing algorithms to assist with their decision-making (Piovesan & Ntiri, 2018). This form of algorithm use focuses on providing a risk assessment of offenders for judges to use when determining an appropriate sentence (Piovesan & Ntiri, 2018). Although its applications for judicial decision-making differs from policing, some foundational elements consistent in the nature of algorithms remains the same. For instance, the algorithms operate to assess the probability that an event will occur in the future, facilitating the decision-making of law enforcement and judges. Furthermore, I also look at the underlying implications of risk assessment techniques when it comes to determining bail, sentencing, and parole. The issues here lie in the factors used to determine an individual's risk assessment, which includes age, criminal record, employment history, and place of residence (Kehl, Guo, & Kessler, 2017). These factors appear neutral and impartial but in actuality, they strongly correlate with race and socioeconomic status, hence, placing the impoverished and marginalized at a disadvantage (Kehl, Guo, & Kessler, 2017). These risk assessments are also used in the decision-making process at various levels within the criminal justice system (Monahan & Skeem, 2016). Risk assessments impact an individual's chances of receiving bail, the length of their prison sentence, ability to reduce their prison sentence, and the likelihood of being granted parole (Piovesan & Ntiri, 2018). Therefore, risk assessment strategies are of importance to shed light on the impact of algorithms and big data tools in the criminal justice system.

With the continued advancement of technology, several industries have begun to adopt big data tools to increase efficiency and reduce costs associated with its operation (Joh, 2017). These industries include but are not limited to, finance companies, the health care industry, and the criminal justice system (Joh, 2017). As such, police departments have rapidly developed the use for algorithms to assess who will commit a crime and where these crimes will be committed (Joh, 2017). The goal here is to accurately predict crime prior to a it taking place, reducing the resources police departments will be required to use (Shapiro, 2017). As Shapiro (2017) put it, since algorithmic crime data is able to target crime location precisely, patrols can be sent to a specific city block, as opposed to an entire neighborhood (Shapiro, 2017). In other words, police departments can strategically position officers in areas where criminal activity is likely to take place, increasing efficiency in witnessing a crime and detaining perpetrators.

Furthermore, algorithmic machines are substantially more cost effective than hiring criminologists and criminal analysts to produce information, providing a strong incentive for police departments to adapt to this method of policing (Shapiro, 2017). Since police often indicate that their police departments lack sufficient resources to adequately perform their duties (Morin, Parker, Stepler, & Mercer, 2017), full enforcement conducted by the police is rather unattainable (Joh, 2017). Thus, it is not surprising that several police departments across the United States have already embraced algorithmic machines. For instance, “police departments in Seattle, Los Angeles, and Atlanta, have piloted or adopted predictive policing programs (Joh, 2017, p. 4).” Although algorithm use in police departments appears to be beneficial, I argue that any benefits they bring are significantly outweighed by costs that often fail to be accounted for.

To continue, algorithms generate predictions based on the information they have been given—or their input (Joh, 2017; Lum & Isaac, 2016; Shapiro, 2017). Algorithms cannot produce outputs in the absence of inputs. In the realm of policing, algorithms used to predict criminality are fed inputs which come from pre-existing flawed data. Thus, if the data is flawed, the algorithmic data produced will also be flawed, as it is produced based on the initial data it has been programmed to follow. To illustrate this point, consider the following analogy. Let us say a human being is taught to speak a language incorrectly, it can be assumed that unless they are given correct information pertaining to the proper way of speaking that language, they will continue to speak that language incorrectly. From this, it can be reasonably drawn that if you do not have the necessary information to correct your actions, you will continue to perform that action incorrectly. This reflects the issue with algorithm use in determining instances of future crime. Since predictive models are designed to learn and reproduce patterns in data, if the data itself is flawed, this will render the models either ineffective or a contributing force to discriminatory policing (Lum & Isaac, 2016). While I do agree with Lum and Isaac (2016), the concern here is that the individuals who determine algorithm use in police departments are unlikely to view discriminatory models as such. In other words, models tend to reinforce pre-conceived biases already held by police officers (Joh, 2017). Hence, its predictions often confirm police officers' views, providing little reason for them to question the models' effectiveness (Joh, 2017). As Lum and Isaac (2016) put it, "selection bias meets confirmation bias (p. 6)." As a result, police officers become passive consumers (Joh, 2017), using algorithms as justification for the perpetuation of discriminatory policing practices.

Within the United States (U.S.), this idea renders significant importance when examining the police system. This is the case because race has long been embedded in the U.S. criminal justice system. To a significant degree, race relations have been reflected in the policing practices employed by police departments. For instance, stop and frisk tactics permitted law enforcement the ability to stop and pat down a person under mere suspicion that the individual may be about to or in the process of committing a crime (Fangman, 2013). During the time period between January 2004 and June 2012, the New York Police Department (NYPD) conducted over 4.4 million stops (Fangman, 2013). The issue here is that the individuals that were stopped were 52 percent Black, 31 percent Hispanic, and 10 percent White (Fangman, 2013). Yet, As of July 1, 2018, the American population was 13.4 Black, 18.1 percent Hispanic, and 60.7 percent White (United States Census Bureau, 2018). This demonstrates a stark racial bias underlying the operations of the American police system. That is to say, when using algorithms to assess crime, the potential for these models to reflect similar racial discrepancies should be an area of serious concern.

Furthermore, in the U.S, race is a strong indicator of geographic location and socioeconomic factors (Shapiro, 2017). As such, it is implausible for algorithms to fully control for racial bias (Shapiro, 2017). For example, Lum and Isaac (2016) conducted a study comparing the results of an illicit drug use public survey and algorithm-generated predictions of drug-related crimes (Joh, 2017). These predictions were made on the basis of demographic characteristics, such as sex, household income, the neighborhood of residence, and age (Lum & Isaac, 2016). Their findings indicated that while drug crimes occur everywhere, drug arrests were limited to particular neighborhoods (Lum & Isaac, 2016). Specifically, neighborhoods with a high concentration of non-white and low-income residents (Lum & Isaac, 2016). Their findings also revealed that targeted areas were already over-represented in police databases, signifying that predictive models are merely regurgitating information consistent with their initial input data (Lum & Isaac, 2016). Lum and Isaac (2016) state that by "filtering [the] decision-making process through sophisticated software that few people understand lends unwarranted legitimacy to biased policing strategies (p. 13)."

This point is significant in that it brings about issues surrounding the lack of transparency of algorithm systems. They not only maintain the racially bias nature of the police system but do so behind a guise of pure objectivity. Thus, creating a "black box" algorithm, in which a lack of transparency leads to the inability to question the algorithm's legitimacy (Shapiro, 2017). This can have a detrimental impact on the livelihood of individuals who find themselves victim to algorithm use for crime prevention. As stated above, algorithms are used by police departments to (1) pinpoint the location of future criminal activity; and (2) to predict individuals who are likely to commit a crime (Joh, 2017). With respect to the latter, algorithms have been used to compile "heat lists," which name off individuals who have been categorized as high risk for violent crime perpetuation (Joh, 2017). With this information, law enforcement officers issue warnings to the subjects of the "heat list," letting them know that they are on the police department's radar (Joh, 2017).

The purpose of the "heat list" is to deter future crime from happening (Joh, 2017). However, the process responsible for producing the list and its utilization in police departments are unknown to the public—and the subjects of the lists (Lum & Isaac, 2016). As an outcome, concerns have been made regarding the potential contradiction between the "black box" algorithm in policing and individual rights against unlawful search and seizure and unreasonable suspicion (Lum & Isaac, 2016). Take the case of Robert McDaniel. In 2013, Mr. McDaniel ended up on the Chicago Police Department's (CPD) "heat list," subsequently receiving a visit from Chicago law enforcement, warning him of this (Lum & Isaac, 2016). Mr. McDaniel, a 22-year-old African American male living in Chicago's South Side had no violent criminal history and no recent issues with the law (Lum & Isaac, 2016). Taking this into account, one must consider the potential rationale behind the selection process of the predictive model used by the CPD. In my view, it appears as if Mr. McDaniel was selected, at least in part, based on factors more or less out of his control. For example, his place of residence—a violent, impoverished, and predominantly African-American area of Chicago.

With that being said, Mr. McDaniel was left with little explanation as to the mechanisms used to render him a subject of the “heat list.” Without transparency, disputing against the algorithm is not possible. And in the absence of a clear explanation of this process, it is difficult to assess whether law enforcement is operating under reasonable suspicion in these circumstances. The notion of reasonable suspicion, functioning under the Fourth Amendment, was established as a safeguard for unlawful policing (Goel, Perelman, Shroff, & Sklansky, 2017). An assessment of reasonableness need not “inquire into the motivations of individual officers;” and should instead employ objective measures (Goel, Perelman, Shroff, & Sklansky, 2017, p. 231). However, Goel et al. (2017) argue that objective measures cannot test racial fairness with respect to reasonable suspicion. As such, reasonable suspicion, as it relates to algorithm use should be re-examined to consider issues surrounding fairness in predictive policing. In my view, an algorithm that labels individuals as high risk for crime perpetuation should not—on its own—render sufficient grounds for reasonable suspicion.

My point here is that reasonable suspicion provides justification for the actions of law enforcement. First, police departments use predictive models to determine high-risk areas and individuals for crime perpetuation. Second, police use algorithmic data to justify increasing police presence in these areas. As a result, it can be assumed that with an increase in law enforcement, arrests will also increase. This is not necessarily due to a higher rate of criminal activity, but simply due to greater police presence (Lum & Isaac, 2016). As a result, crime data will reflect the increasing number of arrests, perpetuating the notion that criminal activity in these areas has risen. However, this claim is unfounded because it rests on the assumption that an increase in arrests is indicative of an increase in crime—which is not always the case. This also creates data in which targeted areas will appear to have A) an increase in crime; and B) a high rate of crime relative to other areas. This is because more law enforcement in one area detracts from law enforcement in other areas, making it less likely arrests will occur in areas with less policing. Thus, the algorithm will receive misleading data, breeding a negative feedback loop. As a result, disadvantaged groups will continue to be targeted by law enforcement, under a false sense of objectivity in their policing mechanisms.

However, this notion extends to other aspects of the criminal justice system. As such, this paper will now assess the impact of algorithms in the judicial decision-making process. According to Piovesan and Ntiri (2018), Artificial Intelligence (AI) has the ability to assist in judicial decision-making by identifying trends based on precedent, decreasing the length of time it takes the court to resolve matters. This has the ability to increase the efficiency of the court system while decreasing trial delay and relieving backlog. However, I argue this only benefits the criminal justice system, so long as it does not have a contributing factor outweighing the harm. In recent years, judges and prosecutors have increased their reliance on automated machines to produce risk assessments on offenders for the purposes of bail, sentencing and parole (Piovesan & Ntiri, 2018). Risk assessments measure an offenders' likelihood to re-offend based on characteristics including age, prior criminal record, employment history, and social networks (Shapiro, 2017; Piovesan & Ntiri, 2018). As mentioned, big data tools lack the transparency necessary for individuals to understand the results the system generates (Piovesan & Ntiri, 2018). Consequently, risk assessments are often difficult to challenge, much like the predictive models used by law enforcement. This is because any underlying racial and socio-economic biases are undisclosed in the absence of legal regulation (Piovesan & Ntiri, 2018).

Therefore, offenders do not know how their risk assessment was calculated or used in the decision-making process, creating significant barriers for successful appeal (Kehl, Guo, & Kessler, 2017).

To illustrate the aforementioned issue, take the *Wisconsin v. Loomis* case. In *Wisconsin v. Loomis*, Eric Loomis, the defendant, argued that Judge Scott Horne's use of The Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) to determine his prison sentence infringed on his due process rights (Piovesan & Ntiri, 2018). He argued that the use of COMPAS was discriminatory on the basis of race and gender, and violated his ability to accurately challenge his case (Piovesan & Ntiri, 2018). The Supreme Court concluded that his rights were not violated because COMPAS was only one element used in the decision-making process (Piovesan & Ntiri, 2018). The concern here is that there should be mechanisms in place to assess the weight to which algorithmic systems impact judges' decision-making. Although seemingly ideal, implementing an effective strategy to assess judicial discretion when considering risk assessments would be considerably difficult. If the information was collected from judges through self-reporting, it would be rather uninformative due to bias. Another strategy would be to compare their sentencing patterns prior to and after adopting risk assessment tools. This may be effective if enough data is collected. Similarly, the Sentencing Guidelines and the Sentencing Reform Act were enacted to reduce variation in sentencing among judges (Christin, Rosenblat, & Boyd, 2015). The idea here is to reduce the possibility of judges applying algorithmic data to varying degrees when sentencing offenders (Christin, Rosenblat, & Boyd, 2015).

Furthermore, Behavioural Economists, Amos Tversky, and Daniel Kahneman have demonstrated that individuals make decisions based on the initial evidence they are given—regardless of its strength (Christin, Rosenblat, & Boyd, 2015). As a result, judges may subconsciously increase their sentences according to the algorithm (Christin, Rosenblat, & Boyd, 2015). On the other hand, at times this applies on both sides; when algorithms produce results consistent with lighter sentences, judges could potentially give out shorter sentences than they otherwise would have (Christin, Rosenblat, & Boyd, 2015). In contrast, Christin et al. (2015) also purport that judges are more likely to give out harsher sentences in the presence of a high-risk assessment, than in the opposite situation. Nonetheless, risk assessments leave much to be revealed about their influence on judicial decision-making.

Moreover, algorithms tend to emphasize incapacitation as the central justification for punishment (Christin, Rosenblat, & Boyd, 2015). “Algorithms privilege a view of justice based on estimating the “risk” posed by the offender when deciding on a sentence designed to incapacitate dangerous individuals (Christin, Rosenblat, & Boyd, 2015, p. 7).” In other words, an individual’s potential risk directly equates to the severity of their sentence. This rationale also emphasizes punishment based on the potential for the occurrence of future crime (Kehl, Guo, & Kessler, 2017). In doing so, sentencing is no longer focused on merely the facts of the crime and the defendant’s criminal history (Kehl, Guo, & Kessler, 2017). Instead, it also considers an offender’s risk assessment score, which is comprised of factors that should be irrelevant to prison sentencing.

For instance, risk assessments account for an individual's employment history, among other factors (Piovesan & Ntiri, 2018). While one could argue that a consistent work history decreases an individual's chances of recidivism, it also severely penalizes the socio-economically disadvantaged. To put it differently, strong employment history reflected in an offender's risk assessment is favorable, in that it provides strong support for successful re-integration back into society following release. This is positive and I do not discredit using employment history when it renders a positive outcome. However, a poor employment history does not merely reveal the opposite of what a strong employment history reveals. What I mean is that the absence of employment does not necessarily demonstrate a higher chance of recidivism in the same way strong employment shows a lower chance of recidivism. In my view, it shows a lack of opportunity and provides valuable information that the criminal justice system should use to lower this individual's chances of re-offending. Instead, risk assessments emphasize utilizing these measures as cause and effect precursors to assess the likelihood of re-offending, which is not entirely accurate. As the use of algorithmic systems continues to grow, one should consider the impact objectivity has on the criminal justice system and the groups most vulnerable to it. As Pasquale and Cashwell (2018) argue, relying on algorithms in the judicial decision-making process can lead to judges merely converting certain inputs into outputs—systematizing the judiciary.

Following the decisions held by the court, serious offenders are sent to federal prison to carry out their prison sentence. Risk assessment allows for the structuring of risk-reduction strategies and earned time credit (Monahan & Skeem, 2016). Prison staff uses risk assessments to develop case plans oriented to reduce low-risk offenders' likelihood of re-offending (Monahan

& Skeem, 2016). Those who are able to successfully follow their established plan can earn up to a 33% reduction off their prison sentence (Monahan & Skeem, 2016). In support of this, Former Attorney General Eric Holder (2014) claimed “data can help design paths for federal inmates to lower these risk assessments, and earn their way towards a reduced sentence, based on participation in programs that research shows can dramatically improve the odds of successful re-entry (as cited by Monahan & Skeem, 2016, 497).”

To conclude, this paper sought to examine the implications of algorithms on different aspects of the criminal justice system. This paper began by analyzing the use of algorithms for the purposes of crime prevention. Through algorithms, police systems assist in maintaining discriminatory policing through the legitimization of algorithmic data (Joh, 2017). To prevent this from happening, considerations must be made as to whether these tools promote fairness and impartiality or simply worsen and underlying social issues. In doing so, law enforcement should aim to provide greater transparency regarding the operation and use of algorithmic systems (Shapiro, 2017). Aligning with Shapiro (2017), government agencies should play a substantive role in regulating the use of algorithms for crime control. This can be done through funding studies to analyze the harmful impact of predictive policing and by also taking steps to alter legislation in light of these policing strategies (Shapiro, 2017). As for risk assessment systems, it renders two main constitutional concerns: its effects on an individual’s right to due process and the potential for variables to violate equal protection (Kehl, Guo, & Kessler, 2017). This is because risk assessments operate with minimal transparency; and as such, concerns have been raised regarding whether defendants have the chance to properly refute, supplement, or explain the information that contributed to their sentencing (Kehl, Guo, & Kessler, 2017).

All in all, this paper aimed to shed light on the emergence of algorithms and its intersection with the criminal justice system. Algorithms have the ability to exacerbate racial and socio-economic issues already plaguing the criminal justice system and therefore warrant close attention and scrutiny. In the future, importance should be placed on designing legislation outlining the appropriate uses of algorithmic machines. Emphasis should also be made regarding transparency in terms of the data it generates and its application to decision-making.

Bibliography

Christin, A., Rosenblat, A., & Boyd, D. (2015, 10 27). Court and Predictive Algorithms . *Data & Civil Rights: A New Era of Policing and Justice* , 1-11.

Fangman, J. L. (2013). Stop the "Stop and Frisk?" How Floyd v. City of New York Will Limit the Power of law Enforcement Across the Nation. *Public Interest Law Reporter, 19*(1), 50-54.

Goel, S., Perelman, M., Shroff, R., & Sklansky, D. A. (2017). Combatting Police Discrimination in the Age of Big Data . *New Criminal Law Review* , 181-232.

Joh, E. E. (2017). Feeding The Machine: Policing, Crime Data, & Algorithms. *William & Mary Bill of Rights Journal* , 1-17.

Kehl, D., Guo, P., & Kessler, S. (2017). Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing. *Responsive Communities Initiative, Berkman Klein Center for Internet & Society*, 2-36.

Lum, K., & Isaac, W. (2016). To predict and serve? *Significance, 13*(5), 1-18.

Monahan, J., & Skeem, J. L. (2016). Risk Assessment in Criminal Sentencing. *Annual Review of Clinical Psychology*, 489-513.

Morin, R., Parker, K., Stepler, R., & Mercer, A. (2017, January 11). *Inside America's Police Departments* . Retrieved from Pew Research Center: Social & Demographic Trends : https://www.pewsocialtrends.org/2017/01/11/inside-americas-police-departments/

Pasquale, F., & Cashwell, G. (2018). Prediction, Persuasion, and the Jurisprudence of Behaviourism . *University of Toronto Law Journal , 68*(1), 63-81.

Piovesan, C., & Ntiri, V. (2018, Spring). Adjudication by Algorithm. *The Advocates' Journal* , pp. 42-45.

Shapiro, A. (2017, January 25). Reform Predictive Policing. *Springer Nature* , pp. 1-10.

United States Census Bureau. (2018). *Quick Facts: United States.* Retrieved from United States Census Bureau : https://www.census.gov/quickfacts/fact/table/US/PST045218