University of California, Berkeley

Algorithmic Efficiency in the United States Criminal Justice System:

Assessing Bias, Improvement, and Proposing Solutions

Abdul Choudhry

College Writing R4B: SEM 014, 22548

Tara Hottman

15 May 2020

Table of Contents

Abstract

In the United States criminal justice system, judges and other justice system officials rely on algorithms to make important decisions that are severely consequential and negatively impact not only the lives of defendants, victims, their families, but also low-income communities of color across the nation. If these algorithms continue to be used in the justice system without improved testing, then these communities will become even more susceptible to racial prejudice, which they already largely face. In this study, I focus on the current widespread algorithms used in the justice system such as predictive policing and risk assessment tools. More specifically, I aim to discuss how a popular predictive policing company named PredPol uses its algorithm as a method of predicting crime in crime hotspots in the US such as Oakland and Chicago. I also emphasize the use of risk assessment tools in the justice system such as the LSI-R and COMPAS that negatively affect the lives of minority offenders and currently incarcerated individuals. I intend to study the bias that exists in these algorithms, the inaccurate data and poor data collection processes that jurisdictions using these algorithms nationwide practice, the negative effects that they bring to low-income communities of color, and how we can more importantly improve the efficiency of these algorithms with some potential solutions. By emphasizing the bias that exists in these algorithms and the context of their consequences, I aim to show readers that policy and decision makers must quickly propose some new solutions in order to prevent the continued proliferation of racial bias towards low-income communities of color nationwide.

Algorithmic Efficiency in the United States Criminal Justice System:

Assessing Bias, Improvement, and Proposing Solutions

Introduction

Within the past decade, the levels of mass incarceration in the criminal justice system has exponentially increased across both state and federal prisons as well as juvenile facilities, causing the United States to have the highest documented incarceration rate in the world. Among many issues, this increasing rate that disproportionately targets people of color is mainly driven by poor data collection and racial, non-human automated decision-making biases caused by widespread algorithms in the criminal justice system such as predictive policing and risk assessment tools. In this context, these algorithms are evidence-based methods or software that aim to predict the future behavior of defendants and incarcerated people and are intended to make decision making processes such as sentencing, bail determination, and parole more efficient. Predictive policing refers to the use of these methods by law enforcement in order to predict crime while risk assessment tools are data-driven methods in the form of questionnaires or software that intend to assess the likelihood of a defendant going back to prison for a similar offense, which is a concept known as recidivism. Judges and other justice system officials rely on these algorithms to make important decisions that are severely consequential and negatively impact not only the lives of defendants, victims, their families, but also low-income communities of color across the nation. If these algorithms continue to be used in the justice system without improved testing, then these communities will become even more susceptible to racial prejudice, which they already largely face. Widespread algorithms used in the criminal justice system such as predictive policing and risk assessment tools must further be improved and tested by their creators, or used in alternative methods within the justice system if testing fails to improve in order to prevent the proliferation of racial bias and harm towards incarcerated individuals, defendants, and communities of color across the nation.

Inaccurate Data and Bias in Predictive Policing Algorithms

Due to patterns of poor data collection and usage, a commonly used algorithm in the criminal justice system that unfairly targets minority groups and proliferates racial bias through inaccuracy is predictive policing. Predictive policing consists of using analytical and quantitative techniques in order to identify likely targets for police intervention and ultimately prevent and predict crime. This algorithm uses data on location, times, and nature of past crimes within the respective cities that use it so that law enforcement can gain insight as to where and at what times potential crime can possibly occur. Although intended to prevent crime, a majority of the predictive policing algorithms currently used by police departments in the United States proliferate racial profiling towards minorities in an already biased nation by often collecting inaccurate and large sets of data. Decades of criminology research has shown that police databases are not a complete and accurate census of all criminal offenses and do not constitute a representative random sample (Lum and Isaac 15). A representative random sample of all criminal offenses is crucial in ensuring that no preconceived biases exist in police databases such that each offense has an equal chance of being chosen in order to not skew the algorithm's output. This same logic can be applied to police data relating to an offender's ethnicity or race, which could be linked to their criminal offense. Since the patterns in the data is what drives predictive policing, biased data and improper usage of sampling will yield biased results. Subsequently, these biased results will cause police officers to consider other factors such as race and ethnicity in determining which individuals to detain and which neighborhoods to patrol, even though they should be focusing on other factors such as the level of danger or harm the crime brings. As a result, historical police records will overrepresent those individuals and ethnic groups such that crimes that occur in locations where policing is more prevalent are more likely to appear in the historical data records that algorithms like predictive policing use.

In order to investigate the effect of police-recorded data on predictive policing models, two researchers named Lum and Isaac applied PredPol’s algorithm to the drug crime records in Oakland using its software. This drug crime records data was taken from drug arrest logs by the Oakland Police Department (OPD) in 2010. The two researchers used this data to create a visualization showing a map of Oakland containing points where drug arrests occurred in 2010 and the estimated number of drug users based on data from the 2011 National Survey on Drug Use and Health. According to Lum and Isaac, PredPol’s algorithm uses seismographic models to produce predictions of crime rates across a given city using only previously recorded crimes and the areas with the highest predicted crime rates are flagged as “hotspots” on the map generated (Lum and Isaac 18). When comparing the data visualizations created by both applying PredPol’s algorithm on the 2010 OPD data and the actual OPD data itself, Lum and Isaac found that those same flagged locations that PredPol’s algorithm predicted as being crime hotspots were also over-represented in the OPD data visualization. This led to the conclusion that rather than correcting for the apparent biases in the police data, the model reinforces these biases (Lum and Isaac 18). This implies that historical crime data, which in this case was Oakland’s drug crime records, does not accurately predict future criminal activity. From the original OPD dataset, drug use by race in Oakland is evenly distributed amongst white, black, and other groups; however, using PredPol's software, black individuals are targeted by police for drug use at a much higher rate than white individuals because of the geographic locations that the software predicts these crimes occur, which consists of predominantly low-income black communities. In an online NYU law review by Rashida Richardson et al. titled *Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice,* Richardson explains to readers that PredPol claims that its algorithm includes drug related offenses, which according to Richardson have known and well-documented racial disparities, and traffic citations data to remove officer bias and generate predictions based on officer discretion (Richardson, et al. 199-200). However, as Lum and Isaac's research shows, data based on drug related offenses yields racial and biased outputs from this exact same algorithm. This shows that in a broader context, predictive policing is as much as predicting the police as it is predicting actual crime since it predicts where arrests have occurred instead of actual violent crimes such as murder. In this case, the OPD 2010 drug arrest data is driving the PredPol algorithm to be inaccurate, which implies that this algorithm is causing harm towards low-income communities of color in Oakland and must be further improved through testing.

The Use of Risk Assessment Tools

Risk assessment tools used in the criminal justice system, in addition to widespread predictive policing algorithms, negatively affect the lives of minority offenders and currently incarcerated individuals. In a legal context, recidivism risk is the risk that a formerly incarcerated individual is rearrested for a similar offense and also a component of many jurisdictions' efforts to reduce the mass incarceration rate in the U.S. One of the most common pretrial risk assessments is the Level of Service Inventory -- Revised (LSI-R) that intends to classify an offender’s risk of re-offending through a series of questions. According to Cathy O’ Neil’s *Weapons of Math Destruction*, the questions posed in this assessment tool are unjust since they include circumstances of a criminal’s birth and upbringing, which subsequently include his or her family and friends. She argues that these details should not be relevant to a criminal case or to the sentencing (O’ Neil 26). This questionnaire is biased because it targets minority groups who are offenders and aims to increase their recidivism score by asking questions that require personal information such as upbringing, which may not have been of the best quality in the offender’s past. O’ Neil argues that although the LSI-R benefits some individuals by landing them lighter sentences, it is unfair since “the questionnaire judges the prisoner by details that would not be admissible in court” (O’ Neil 29). Similarly, Edward Lempinen discusses how researchers at Dartmouth University raise concerns in the accuracy of tools like the LSI-R in a criminal justice framework. The research results showed that “both the people and the algorithm were accurate slightly less than two-thirds of the time” (Lempinen) and that these risk assessments are indeed an issue for individuals on trial who represent a minority group and come from low-income socioeconomic backgrounds. In his online article titled “The Use and Abuse of the LSI-R In Parole Evaluations Challenging So-Called ‘Objective’ Testing”, attorney Eric Marcy emphasizes that the LSI-R is biased towards current incarcerated individuals since it does not account for rehab programming, work performance, maturation, education, and being a model inmate (Marcy). Although it was intended to make the justice system more efficient, the LSI-R brings harm to minority offenders and currently incarcerated individuals by essentially keeping them in the prison industrial complex and away from their families where they may face physical threats, violence, and psychological harm such as depression.

Improving The Efficiency of Current Algorithms

In order to maximize fairness and minimize harm in the justice system, predictive policing algorithms and risk assessment tools must further be improved. One way that this fairness can be reached is by providing transparency about how these algorithms have been developed, the assumptions made in their implementation, and how frequently they are assessed and updated to the individuals that are mainly affected by their use. In an online article by Danielle Kehl and other researchers titled “Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing”, this ethical obligation of algorithmic transparency is emphasized to readers as Kehl persuades readers that law professors including Danielle Keats Citron have developed and advocated for a process known as “technological due process” (Kehl 32). This process aims to ensure that the many people negatively affected by these algorithms, primarily people of color, are able to challenge the decisions made by the algorithms and have the “right to inspect, correct, and dispute inaccurate data and to know the sources (furnishers) of the data” (Kehl 33). Transparency allows for local and state government policymakers to determine which companies to include or remove in using their developed tools or software since some companies do not publicly share how exactly their algorithms are implemented. According to Stephanie Wykstra’s online article titled “Philosopher's Corner: What Is ‘Fair’? Algorithms in Criminal Justice”, in 2016, a defendant in Wisconsin who pleaded guilty to eluding the police and operating a vehicle without its owner’s consent sued the state, claiming that reliance on Newpointe’s COMPAS risk assessment algorithm violated his rights to due process, in part because he was not able to see the algorithm’s code (Wykstra). Northpointe refused to share the code and the Wisconsin Supreme Court eventually ruled against the defendant, which shows how crucial transparency really is in improving the efficiency of current algorithms used in the justice system.

In addition to improving criminal justice algorithms through open transparency, the people who design, implement, and employ them must improve their models by implementing modern machine learning techniques such as classification and forecasting. Since most criminal justice outcomes tend to be categorical, classification in criminal justice algorithms allows assigning particular crime categories to individuals that have already been arrested. Classifying these individuals by select features like crime outcomes, crime history, and age will increase the algorithm's predictions since these features will guide subsequent outcomes in the future such as arrest. According to Richard Berk's *Criminal Justice Forecasts of Risk: a Machine Learning Approach*, classification using least squares linear regression and decision boundaries subsets the data that is being used in the algorithm to create a new class which is then used in the forecasting process to predict multiple outcomes (Berk 28-29). For example, in predicting the recidivism risk of offenders, if features such as age and crime history instead of ethnicity and socioeconomic status are used to predict violent outcomes, then depending on how the algorithm is coded, the data points either above or below the decision boundary will forecast if the individual is likely to commit more crime after being released from prison. This decision boundary regression process allows the model's predictions to be more accurate and efficient by reducing bias variance error, uncertainty, and using features that will better predict future crime.

Jeff Larson and other researchers bring to attention that Newpointe's COMPAS algorithm was only correct in its predictions of violent recidivism twenty percent of the time because their analysis showed that "black defendants who did not recidivate over a two-year period were nearly twice as likely to be misclassified as higher risk compared to their white counterparts" (Larson, et al.). In developing such a widespread tool like the COMPAS, the developers at Newpointe should have properly used more advanced machine learning techniques such as classification and forecasting to better improve their model. In response to the aforementioned ProPublica article, in an online study by William Dieterich and others from Newpointe's research department it is argued that ProPublica focused on "classification statistics that did not take into account the different rates of recidivism for blacks and whites" (Dieterich, et.al 1). However, regardless, Newpointe did not apply the highest level of technological prediction tools at the time like linear regression through decision boundaries to improve the prediction of defendant recidivism. The level of detail that is required in the implementation of these widespread algorithms must be taken seriously and private companies like Newpointe are to be held accountable for the lack of these implementations.

To further improve the overall efficiency of algorithms used in the criminal justice system, it is essential that these algorithms are worked on collaboratively through open source code instead of being deployed privately within companies such as Newpointe. Open source collaboration encourages algorithmic transparency and lowers cost as well as competition within the criminal justice industry. This not only reduces issues of error and bias, but it also allows for leading developers and researchers in multiple fields such as machine learning, social science, criminal justice, law, and more to come together and propose much more efficient algorithms than a private company. Although these companies may be able to improve their algorithms through testing and introducing new statistical techniques, this does not compare to the efficiency that leaders in their respective fields will bring in improving the algorithms.

Additionally, according to Phillip D. Waggoner and Alec Macmillen's paper titled "Pursuing Open-Source Development of Predictive Algorithms: The Case of Criminal Sentencing Algorithms", proprietary algorithms like COMPAS are much more expensive to deploy compared to potential open source algorithms. Leading researchers can work with other researchers in different industries and use free, open source statistical software to create models that perform just as well or if not better than commercially used models (Waggoner and Macmillen 4). Adopting an open source framework for these algorithms will also allow for anyone in the general public that is interested in algorithmic models in the justice system to view the code and analyze it for themselves as a greater educational opportunity. Through a different view, some criticisms that this open source framework can create include ethical privacy concerns about the defendants and their families; however, this issue can be accounted for if all the defendant names and personal information is replaced with ID numbers that lose track of their identity. In the long run, the models that open source algorithms predict must be more accurate and evade legal loopholes of data privacy because they will constantly be worked on by a group of well established developers and researchers that are better equipped in answering these issues collectively rather than a private company whose experience is limited. This implies that these biased algorithms should not completely be removed from the justice system, but rather improved through transparent, collaborative, and statistical methods.

Proposed Solutions and Future Implementation

Including improving the efficiency of algorithms, the issue of biased algorithmic usage in the justice system can be solved by using new algorithms in alternative methods within the justice system instead of completely removing them. The algorithms that the justice system currently uses to predict sentencing or parole determination for defendants as well as recidivism risk scores can instead be used in determining which defendants need supplemental resources to better cope with life after prison or which defendants require confinement in a maximum security prison based on their criminal record, as a few examples. Algorithms used in these two suggested areas within the justice system will be less biased and more effective then, for example, risk assessment tools being used for pretrial servicing in a bail context, parole determination, or sentencing because the data needed to develop these models is not as general and does not include a large sample size. In this case, unbiased inputs will not yield biased results. For example, in order to create a model that predicts which defendants in a state prison need supplemental resources to better cope with life after prison, assuming they are guilty of the crime, data such as mental health records, cases of psychological trauma, and family dependency or conditions are all valid features that can be used as the training dataset for the model. The negative consequences that biases in current predictive policing and risk assessment algorithms bring far outweigh the negative consequences that potential biases in algorithms that model life after prison bring. The context and subsequent consequences immediately shift from individuals not being eligible for parole, acquiring prolonged sentences, and not being able to apply for jobs compared to individuals not receiving the resources they need to cope with life after prison. This implies that current algorithms in the justice system such as predictive policing and risk assessment tools should be replaced with new algorithms that are used within the justice system for different purposes if these current algorithms do not improve.

Conclusion

To conclude, judges and other justice system officials rely on widespread algorithms used in the criminal justice system such as predictive policing and risk assessment tools like the LSI-R and COMPAS to make important decisions that are severely consequential and negatively impact not only the lives of defendants, victims, their families, but also low-income communities of color across the nation. These algorithms contain racial bias targeted towards communities of color and must further be improved and tested by their creators, or used in alternative methods within the justice system if testing fails to improve, in order to prevent the proliferation of racial bias and harm towards these communities. If the data that is being inputted into these algorithms was to be cleaned and initially looked over by humans instead of direct software implementation, then inaccuracy, racial profiling, and police brutality would decrease. It is in the best interest of policy and decision makers that decide to use these algorithms within their jurisdictions to deeply consider and assess the potential solutions aforementioned sooner rather than later due to the current overall state of the prison industrial complex. With the current presence of an ongoing global pandemic, these implementations are not likely to occur any time soon; however, raising awareness through public discourse and virtual means can motivate government officials to act on this issue seriously and in a timely manner in order to provide transparency, fairness, accountability, and more importantly justice for all.

Annotated Bibliography

Richardson, Rashida, et al. "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice." *Nyulawreview.org*, vol. 94, no.192, 2019, pp.206-210, www.nyulawreview.org/wp-content/uploads/2019/04/NYULawReview-94-Richardson-Schultz-Crawford.pdf.

The above New York University law review written by Rashida Richardson and other scholars discusses the analysis of thirteen jurisdictions that have used or developed predictive policing tools that negatively impact the lives of minority groups in the United States. The law review emphasizes to readers that multiple law enforcement agencies across the nation are increasingly using predictive policing algorithms to forecast criminal activity; however, most of the data that is produced by jurisdictions that use these algorithms is inaccurate and systematically biased towards people of color which leads to unlawful practices by police departments across the nation. Richardson mainly argues that this "dirty data" implemented into predictive policing algorithms is the primary cause for the unlawful and biased police practices that exist in the United States today. Through three main case studies in Chicago, New Orleans, and Maricopa County, Richardson gathers strong evidence that shows the predictive policing algorithms in each region were using dirty data, which leads to the overall conclusion that it is highly likely for a predictive policing algorithm to contain biased data that primarily negatively target people of color.

I believe that this law review serves to be a great source for my final research paper because it emphasizes the relationship of inaccurate data implementation in a widespread algorithm used in the United States criminal justice system known as predictive policing, and the negative effects that this algorithm has towards people of color in particular. In my paper I plan to discuss the relationship between inaccurate "dirty" data and the harm that it brings to people of color specifically through predictive policing. I found that the first case study that Richardson and other scholars focus on, which talks about the case in Chicago where public outcry and city-wide protests followed the fatal shooting of a young black teenager named Laquan McDonald through unlawful police practices, is important to include because it supports my argument and also complements some other sources I plan to use which also discuss Chicago's police practices. For example, another one of my sources discusses the year long investigation by the Department of Justice of McDonald's case that concluded that poor data collection did in fact lead to unlawful conduct.

Lum, Kristian, and William Isaac. "To predict and serve?" *Royal Statistical Society*, John Wiley & Sons, Ltd, 7 Oct. 2016, rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1740-9713.2016.00960.x

Kristian Lum and William Isaac's online article titled "To predict and serve?" discusses the effects and social consequences of what happens when a predictive policing algorithm used by law enforcement is trained using bias data. The two researchers argue that police databases are not a complete and accurate census of all criminal offenses and do not constitute a representative random sample. One main cause for this inaccuracy is that police officers consider race and ethnicity in determining which individuals to detain and which neighborhoods to patrol. As a result, historical police records will overrepresent those individuals and ethnic groups such that crimes that occur in locations where policing is more prevalent are more likely to appear in the historical data records that algorithms such as predictive policing use, according to the article. The case study that the two researchers focus on is applying the PredPol predictive policing algorithm to Oakland's drug crime arrest records taken from the Oakland Police Department back in 2010. Lum and Isaac found that rather than correcting for the apparent biases in the police data, the PredPol model reinforces these biases which implies that by using PredPol's algorithm, black individuals are targeted by police for drug use at a much higher rate than white individuals solely due to inaccuracy. Towards the end, the researchers conclude that when the costs of policing are disproportionate to the level of crime, this amounts to discriminatory policy. They state that although predictive policing reinforces the same biases that police have historically held, filtering this decision-making process through sophisticated software leads to legitimacy issues regarding police practices.

I strongly believe that Lum and Isaac's online article will support my claim that poor data collection leads to biased predictive policing algorithms that will eventually lead to either physical or emotional harm towards people of color in the U.S. It would make sense for me to dedicate this source for a paragraph solely emphasizing the negative effects that an inaccurate dataset brings to the implementation of a predictive policing algorithm because the researchers do a great job of showing readers these effects through applying PredPol's algorithm on Oakland Police Department's inaccurate drug arrest records. This source will definitely support my thesis statement and help develop my argument because I am providing an example of an application that researchers have recently done related to the topic. Depending on how many sources I want to dedicate specifically for the usage of the predictive policing algorithm, I think that introducing this source in my first body paragraph will help organize my paper since it follows the following flow I want my body paragraphs regarding predictive policing to be in: describe data inaccuracy, show how data inaccuracy affects the predictive policing algorithm, show physical or emotional harm caused by the inaccurate data-driven algorithm.

Marcy, Eric. “The Use And Abuse Of The LSI-R In Parole Evaluations Challenging So-Called ‘Objective’ Testing”. *Wilentz, Goldman & Spitzer, P.A.*, 2015, www.wilentz.com/perspectives/criminal-law/2015-04-15-the-use-and-abuse-of-the-lsi-r-in-parole-evaluations-challenging-so-called-objective-testing.

Eric Marcy’s online article titled “The Use And Abuse Of The LSI-R In Parole Evaluations Challenging So-Called ‘Objective’ Testing” discusses how a common risk assessment tool in criminal justice system known as the Level of Service Inventory -- Revised (LSI-R) that intends to classify an offender’s risk of re-offending through a series of questions is biased towards current incarcerated individuals since it does not account for rehab programming, work performance, maturation, education, and being a model inmate. Marcy argues that the LSI-R’s training manual itself cautions administrators about its use as a forensic tool in the wrong context or improper administration and also mentions that the test is frequently administered based upon an incomplete record of the inmate’s history or false information about the inmate that has not been verified. He concludes by saying that the use of the LSI-R in determining suitability for parole is highly suspect and should be challenged especially if the factors generated to score the test involve factual circumstances that go back many decades.

I believe that Marcy’s article will be good for me to use and combine with other class readings in my final paper including O’ Neils *Weapon of Math Destruction* and Edward Lempinen’s online article “Algorithms are better than people in predicting recidivism, study says.” Since I hope to address the issues that current algorithms in the criminal justice system cause, including recidivism risk and not only predictive policing, this article does a great job of explaining the negative implications that assessment tools such as the LSI-R bring to people of color. I think that this source will be a great complement to O' Neils discussion of the LSI-R being an unjust assessment tool because according to her the questions posed in the LSI-R should not include circumstances of a criminal's birth and upbringing, which subsequently include his or her family and friends, which is similar to Marcy's idea that the LSI-R should not be used for determining suitability for parole.

James, Nathan. “Risk and Needs Assessment in the Federal Prison System.” 2018, fas.org/sgp/crs/misc/R44087.pdf.

In an online Congressional Research service article by Nathan James titled “Risk and Needs Assessment in the Federal Prison System”, it is discussed that as levels of mass incarcerations in the U.S. continue to rise, awareness regarding reforms to the criminal justice system including improving the system’s ability to rehabilitate incarcerated offenders by better assessing their risk for recidivism is also increasing. James suggests that the most commonly used assessment instruments can, with a moderate level of accuracy, predict who is at risk for violent recidivism. The article emphasizes some questions and issues that policymakers may contemplate if Congress decides to implement a risk and needs assessment system in federal prisons, such as the following: Is there the potential for bias in the use of risk and needs assessment? Should risk assessment be incorporated into sentencing?

In the context of my research paper, a particular area in this article that I would focus on would be the concerns about bias in risk and needs assessment. In this section of the article, James talks about how research by investigative journalists and data scientists with ProPublica, a nonprofit based in New York city, on risk classifications generated by the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) has increased the debate about racial disparities in risk and needs assessment. According to this research, I would try and complement this source with Eric Macy’s article regarding the LSI-R since they both portray similar claims that support my thesis but provide two different risk assessment algorithms that are still used in the criminal justice system. It would be interesting to also compare this article with Lempinen's "Algorithms are better than people in predicting recidivism, study says." and perhaps combine all three sources into a single body paragraph or find another source that also discusses the COMPAS in depth and break two sources for each body paragraph. Although I am not entirely positive that I will end up using this source, I do think that I can further build my overall argument and persuade readers that algorithms in the criminal justice system such as predictive policing and risk assessment tools are one main reason for an increase in racial bias and harm towards minority groups.

Kehl, Danielle, Priscilla Guo, and Samuel Kessler. "Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing." *DASH Home,* Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School, 25 July 2017, https://dash.harvard.edu/bitstream/handle/1/33746041/2017-07_responsivecommunities_2.pdf

The article above focuses on the use of risk assessment tools in the criminal justice system in the sentencing process which raises legal and ethical questions of accountability, fairness, and transparency. Kehl and other scholars aim to provide local policymakers with some key considerations and questions for further research that can help them decide if these algorithms should be implemented in their local legal systems, more specifically in a sentencing context. The authors make clear and argue that the use of risk assessment software for sentencing poses greater challenges and bias than the use of these tools in pretrial risk assessments by including the potential for constitutional challenges under the Due Process and Equal Protection clauses of the Fourteenth Amendment. Lastly, the article summarizes the challenges that these algorithms make for law and policymakers and offers solutions of best possible practices to make sure that these tools are accountable, fair, and transparent for all individuals in the criminal justice system.

I believe that this article will serve to be a great source for my research paper because it raises awareness towards the ethical questions regarding the use of risk assessment tools in the criminal justice system. Since my paper aims to highlight the challenges caused by algorithms including risk assessment tools for people of color, this article will give more depth as to how these tools actually create issues for people of color in sentencing proceedings which will greatly complement my discussion of O' Neil's ideas regarding the inherent bias in the LSI-R. I can also choose to complement this source with James Nathan's "Risk and Needs Assessment in the Federal Prison System" since Nathan talks about how risk classifications generated by the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) has increased the debate about racial disparities in risk and needs assessment and this article similarly discusses the causes of these disparities. Lastly, I think it would be a good idea to use this source in addition to other new sources in a separate paragraph discussing some proposed solutions to how algorithms in the criminal justice system can improve since this article addresses policymakers about best possible practices for these kinds of systems.

Wykstra, Stephanie. "Philosopher's Corner: What Is 'Fair'? Algorithms in Criminal Justice." *Issues in Science and Technology*, 4 Oct. 2018, issues.org/perspective-philosophers-corner-what-is-fair-algorithms-in-criminal-justice/.

Wykstra's online article discusses the fairness involved with algorithms used in the criminal justice system. She introduces the article by mentioning how these algorithms have been used in some kind of decision making process since the 1920s and in the modern day they are gaining widespread usage especially in the context of pretrial decision making. She argues with evidence from ProPublic journalists that the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) risk assessment tool is unfair after analyzing the data obtained from a jurisdiction in Florida that uses the algorithm and concludes that it is racially biased. The main reason as to why COMPAS is deemed unfair and this conclusion was reached is because there are different base rates of rearrest fed into the algorithm for different racial groups. Wykstra then argues that achieving this kind of fairness creates more unfairness. She makes the claim that since black defendants are arrested at higher rates and their criminal history is largely fed into risk assessment tools like the COMPAS, a larger percentage of these black defendants are assigned to higher risk scores than white defendants. Later in the article, she draws views from computer scientists and researchers that have actually worked on the development of these assessment tools and concludes that due to the lack of accountability, unfairness, and transparency of these algorithms, they should be limited in usage and regularly be audited by independent researchers to assess racial inequality.

I believe that Wyktra's article fits in well with my research topic because it adds a new perspective, supports my claim, and complements my other sources. This article supports my claim that risk assessment tools such as the LSI-R and COMPAS are unfair and proliferate racial bias through the perspectives of the very researchers and computer scientists that implemented these tools, and I may want to quote their thoughts on this issue. I think that I will also use this article in a separate paragraph discussing the solutions to these types of algorithms being used in the criminal justice system since Wykstra proposes a lot of solutions that I agree with. To be used in the proposed solutions paragraph, I would complement this article with Danielle Kehl's "Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing" since both Kehl and Wykstra propose similar solutions to how risk assessment tools in particular can be improved and regularly automated to check for racial inequalities. Excluding the solutions paragraph, I will most likely complement this article with Nathan and Marcy's articles discussing the impacts of different risk assessment tools like the LSI-R and COMPAS.

Ferguson, Andrew G. *Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. New York University Press, 2020, *jstor.org*, pp. 107-130, www.jstor.org/stable/j.ctt1pwtb27.3?refreqid=excelsior%3A7870a3a7909973cbd608e1de5b7179aa&seq=6#metadata_info_tab_contents.

Andrew Ferguson's *Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement* is a great book that studies how big data can be used by law enforcement across the nation to reduce privacy. Big data systematically extracts data from large datasets that are too large for traditional data-processing tools and Ferguson argues that big data collects data that contains preconceived biases, outdated, and unreliable. He looks at racial bias in big data and researches whether the system is biased, data is biased, or if the data is correct and discusses the constitutionality of using big data as probable cause. Specifically in chapter six of his book titled "How we police: Data mining digital haystacks", he argues that some algorithms used by law enforcement are no better than the data they process since they try to smooth out the anomalies in the data, where these anomalies correspond with minority groups. Ferguson also includes mathematical reasoning as to why these algorithms may be inherently biased. He argues that in the context of criminal justice, a machine learning model's only ability to learn is to sort through historical data to determine a correlation. He concludes this chapter by emphasizing to readers that data collection, data mining, and data analysis shape both police investigation and policing practice across the nation and this field of work requires great level of technical expertise and significant investment in technology.

I strongly believe that Ferguson's book will deepen my claim that due to patterns of poor data collection and usage, commonly used algorithms in the criminal justice system such as predictive policing unfairly target minority groups and proliferate racial bias. I think it would be best for me to compare the main arguments and points of chapter six of this source with Lum and Isaac's online article titled "To predict and serve?" because these two researchers also argue that the data that is being collected to run the predictive policing software that technology companies such as PredPol create, and many police departments nationwide rely on, is a by-product of biased police activity. It would make the most sense for me to complement Ferguson's discussion with this online article in a paragraph dedicated solely for emphasizing the negative effects that poor data collection and usage actually has on current widespread algorithms such as predictive policing. Additionally, I plan to paraphrase Ferguson's claim that mathematical reasoning and machine learning are also some causes for why these algorithms are inaccurate in predicting what they are intended to since it will only support my overall thesis that these algorithms negatively affect the lives of minority groups and may cause physical or emotional harm.

Lansing, Sharon. "New York State COMPAS-Probation Risk and Need Assessment Study: Examining the Recidivism Scale's Effectiveness and Predictive Accuracy." 2012, www.criminaljustice.ny.gov/crimnet/ojsa/opca/compas_probation_report_2012.pdf.

Lansing's online article presents the findings from a study which examined the effectiveness and predictive accuracy of the New York State COMPAS-Probation Recidivism scale. This scale predicts the likelihood of rearrest for any felony or misdemeanor offense over a two year follow-up period for offenders under probation supervision. In addition to examining the effectiveness and predictive accuracy of the scale, the study also examined the prevalence of COMPAS-Probation risk among the probationers to see if factors were correlated with the likelihood of rearrest. According to Lansing, with respect to the offender age at assessment, the study found that the recidivism scale underestimated the likelihood of rearrest for offenders 16 to 18 years old and subsequently overestimated the likelihood of rearrest of older individuals. The graph presented in the research shows that through probabilistic regression at any given age, the percent rearrested curve predicted by the COMPAS algorithm seemed to be close to the expected rearrest percentages curve given any age. Lansing concludes that young adult offenders, because of their youth, comprise a special population with needs not fully addressed by COMPAS. She also mentions that the original dataset that was fed into COMPAS may have created discrepancies for the predicted curve which may have led to an inaccurate prediction.

I am currently on the fence with this source and will potentially use it in my final research paper. From reading chapter six of Craft of Research, I found it important to find a source related to my research topic that does not directly support my claim so I think it may be interesting to include Lansing’s analysis of New York’s COMPAS case study with my other sources that support that risk assessment tools like the COMPAS are biased towards people of color. This article does not fully support my claim that algorithms used in the criminal justice system negatively affect people of color because from Lansing’s analysis, New York’s COMPAS algorithm is technically accurate in predicting recidivism for younger individuals because it underestimates and is fitted pretty closely to the actual recidivism curve. In a paragraph specifically dedicated to the COMPAS algorithm, I may or may not compare this source with Wykstra’s analysis of the COMPAS used in Florida and her argument that it is in fact racially biased towards people of color.

Puente, Mark. “LAPD Pioneered Predicting Crime with Data. Many Police Don't Think It Works.” *Los Angeles Times*, Los Angeles Times, 3 July 2019, www.latimes.com/local/lanow/la-me-lapd-precision-policing-data-20190703-story.html.

Mark Puente’s online article discusses how the Los Angeles Police Department agrees that their predictive policing algorithm targets crimes in areas that are never committed and may be inaccurate. Since the software that drives this algorithm forecasts risk, much of its efficacy is in crimes never committed due to poor data collection that can be cleaned or verified by a human before relying on the algorithm’s predicted output. The data that is driving this algorithm in Los Angeles is negatively affecting the communities of color and minorities that live there because it is blindly suggesting police officers to enforce the law in areas where crime never occurred. Puente argues that the use of complicated algorithms for data-driven programs like PredPol make it difficult to evaluate their overall effectiveness in targeting crime. He mentions that the Stop LAPD Spying Coalition has long criticized the department’s data tools, since they argue that information collected is inherently biased because poor, black, and brown communities are more prone to being targeted by these tools. Finally, he concludes with the words of John Hollywood, a senior researcher and professor at Rand Corp., who said that these predictive algorithms are just mapping tools and do little to decrease crime.

I believe that Puente’s article serves to be a good source for my research paper because it supports my claim that predictive policing algorithms are inaccurate in actually predicting crimes and lead to racial biases that target minority groups, in this case in LA. If I end up using this source for my paper, which I am still unsure of, I would develop and complement this source with Lum and Isaac’s article “To Predict and Serve?” because the two researchers in that article found that while using PredPol’s predictive policing algorithm on Oakland’s drug arrest records, rather than correcting for the apparent biases in the police data, the PredPol model reinforces these biases which implies that by using PredPol’s algorithm, black individuals are targeted by police for drug use at a much higher rate than white individuals solely due to inaccuracy. In connection with Puente’s LAPD article, this implies that LAPD’s predictive policing algorithm would also be inaccurate because it is developed and implemented by the same company, PredPol.

Works Cited

Berk, Richard. *Criminal Justice Forecasts of Risk a Machine Learning Approach*. Springer New York, 2012, pp. 28-29.

Dieterich, William, et al. "ProPublica-Commentary-Final-070616." *DocumentCloud*, Northpointe Inc., 8 July 2016, www.documentcloud.org/documents/2998391-ProPublica-Commentary-Final-070616.html

Kehl, Danielle, Priscilla Guo, and Samuel Kessler. "Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing." *DASH Home,* Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School, 25 July 2017, https://dash.harvard.edu/bitstream/handle/1/33746041/2017-07_responsivecommunities_2.pdf

Larson, Jeff, et al. "How We Analyzed the COMPAS Recidivism Algorithm." *ProPublica*, 23 May 2016, www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm.

Lempinen, Edward. "Algorithms Are Better than People in Predicting Recidivism, Study Says." *Berkeley News*, 19 Feb. 2020, news.berkeley.edu/2020/02/14/algorithms-are-better-than-people-in-predicting-recidivis m-study-says/?utm_source=Berkeleyan&utm_campaign=1a5a33f584-berkeleyan&utm_medium=email&utm_term=0_99ee3800d7-1a5a33f584-388623173.

Lum, Kristian, and William Isaac. "To predict and serve?" *Royal Statistical Society*, John Wiley & Sons, Ltd, 7 Oct. 2016, rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1740-9713.2016.00960.x

Marcy, Eric. "The Use And Abuse Of The LSI-R In Parole Evaluations Challenging So-Called 'Objective' Testing". *Wilentz, Goldman & Spitzer, P.A.*, 2015, www.wilentz.com/perspectives/criminal-law/2015-04-15-the-use-and-abuse-of-the-lsi-r-in-parole-evaluations-challenging-so-called-objective-testing.

O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Penguin Books, 2018.

Richardson, Rashida, et al. "Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice." *Nyulawreview.org*, vol. 94, no.192, 2019, pp.206-210, www.nyulawreview.org/wp-content/uploads/2019/04/NYULawReview-94-Richardson-Schultz-Crawford.pdf.

Waggoner, Phillip D, and Alec Macmillen. "Pursuing Open-Source Development of Predictive Algorithms: The Case of Criminal Sentencing Algorithms." *Pdwaggoner.github.io*, 2018, pdwaggoner.github.io/Research/Manuscript.pdf.

Wykstra, Stephanie. “Philosopher's Corner: What Is ‘Fair’? Algorithms in Criminal Justice.” *Issues in Science and Technology*, 4 Oct. 2018, issues.org/perspective-philosophers-corner-what-is-fair-algorithms-in-criminal-justice/.