**Data Bias:**

**Prejudice in Recidivism Model for Prison Sentence**

Lavanya Govindarajan

Department of Mathematics and Computer Science, Webster University

CSDA 5130 07 S2 2023: Social and Ethical Issues in Analytics

Prof. Oubonvanh Douangkeomany

March 26, 2023

**Introduction**

In most industries today, the word "data" has become something of a catchphrase. Data is what fuels the industries' daily operations, growth, and sustainability. Ultimately, it all comes down to running a few queries against the data and interpreting the findings to help run the business. Clive Humby (2006) famously said, "Data is the new oil." Like oil, data is valuable, and it can be broken down and processed to gain insights. As the dependency on data increases, so does the potential for data privacy breaches and other ethical issues surrounding data, right from the data collection phase. Data bias is one of the ethical issues that caught my attention when reading the first two chapters of the book "Weapons of Math Destruction". The author argues about how recidivism models that were developed to reduce prejudice in prison sentencing still contain data bias wrapped within a blanket of technology.

**Synthesis**

The judiciary system is regarded as the most powerful sector of a nation, and it is believed to operate solely in accordance with the law and not under any political influences. Along with power comes the responsibility to uphold justice and remain impartial. Having stated that, the author discusses an incident that took place in a trial in Texas, when the jury opted to sentence the convicted murderer to death based on the race factor. The accused was an African American man named Duane Buck; as per a study conducted by a psychologist, the recidivism rates in Texas prisons were high for people of Buck’s race. The author states that:

According to the American Civil Liberties Union, sentences imposed on black men in the federal system are nearly 20 percent longer than those for whites convicted of similar crimes. And though they make up only 13 percent of the population, blacks fill up 40 percent of America’s prison cells. (O’Neil, 2016, p. 27)

The Texas attorney general identified some incidents that were similar. To eradicate the human bias in sentencing, recidivism models were developed. The LSI-R was one of the most popular models; it featured a questionnaire to be filled out by the prisoners, and the model was then updated using their responses. Based on how many points the convicts score, they would be rated as high, medium, or low risks. To return to the questionnaire for a moment, most of the questions are both surprising and alarming. The inquiries include the circumstances surrounding their birth, their family history, whether they had any police confrontations, their prior convictions, whether their relatives or friends have any criminal records, and so on, which are quite unfair.

**Conclusion**

It is a human tendency to often believe computerized models to be accurate, error-free, and unbiased. Although this is true to some extent, we must also remember that these models were created and trained by human beings. Is it fair to expect the algorithm to be unbiased when it's fed with biased data? Going back to the recidivism example, the questions in the questionnaire of the LSI-R model had more to do with the convict's personal life and social background while paying little to no attention to the case itself or how the convict was involved in the crime committed. When erroneous information is introduced into the models, the results will be deceptive. And in the end, the society that relies on these results will suffer the consequences. It is not only about gathering data but also collecting accurate, relevant, and specific data while strictly adhering to moral standards.

Reference

O’Neil, C. (2016). *Weapons of Math Destruction: How Big Data increases inequality and threatens democracy.* Crown.

Mavuduru, A. (2020, December 11). Is Data Really the New Oil in the 21st century? *Medium*. **https://towardsdatascience.com/is-data-really-the-new-oil-in-the-21st-century-17d014811b88**