**Pre-Course Reflection**

As I read through the “About the Book” and “Introduction” sections in *Fairness and*

*Machine Learning*, I learned quite a lot about inequality in data science and biases that are inherent in machine learning algorithms. This is a topic that has come up fairly often in the past few years, as recently, there have been a large number of cases where minorities have faced prejudice due to factors they possess that they have no control over but are frowned upon. This has caused a lot of people to face scrutiny when the only thing that made them stand out was, for example, a physical trait they had, such as their skin color.

The “About the Book” section focused quite a bit on the moral obligations of those who create and use machine learning algorithms and the value of debating whether such an algorithm is morally unjustifiable to use. It’s not just about creating solutions that make the technologies *seem* less reprehensible; it’s about discussing how the algorithm impacts those it relates to and whether or not its uses are morally questionable.

This is seen in *Fairness and Machine Learning* as the author writes, “A narrow framing of machine learning ethics might be…a way to focus on technical interventions while sidestepping deeper questions about power and accountability…For example, mitigating racial disparities in the accuracy of face recognition systems, while valuable, is no substitute for a debate about whether such systems should be deployed in public spaces and what sort of oversight we should put into place” (Barocas et al. 8-9).

The “Introduction” section of *Fairness and Machine Learning* focused more on the technical aspects of *why* such prejudice is inherent in machine learning algorithms. Due to demographic disparities, there is often a bias against minority groups. In terms of data, information about a minority group is often skewed due to a combination of minority groups having a smaller sample size in comparison to majority groups (which leads to the data being skewed towards the majority groups) as well as the small amount of data about minority groups (which often leads to a situation where the data is inaccurate or not accurate enough for the vast majority of members of the minority group). Thus, machine-learning algorithms get data from this system and refine their models to include these prejudices, eventually creating a machine-learning system that propagates the bias. This is a major reason why it is important to debate whether using such systems for significant and impactful decisions is morally justifiable.

As of right now, I am familiar with the Python programming language and have some experience using it for personal projects as well as competitively. However, I do not have any experience with the R programming language. My goal for this class is to learn how to program with R and learn to use both R and Python in the context of data science. Using these programming languages, I hope to learn how to manipulate data and data structures efficiently as well as learn how to create visualizations of data that are effective and easy to understand in a given scenario. In addition, I wish to familiarize myself with any notable strategies you may have that will help me better develop myself as a learner in the field of data science and perhaps computer science. I am extremely excited to join this class and meet passionate, like-minded individuals who I can relate with and who can aid me in advancing my knowledge of this field.