**Book Review**
**Berry, C. A., & Marshall, B. H. (Eds.). (2024)**
***Mitigating Bias in Machine Learning***
**McGraw Hill**
**Kindle Edition**

.

# Teaching Justice: Why *Mitigating Bias in Machine Learning* Is an Essential Textbook for the Future of AI

**Gift Iyioku**

**Stanford University**

As a Nigerian PhD student at Stanford University studying political science, policy, immigration, and technology, I found *Mitigating Bias in Machine Learning* to be a timely and valuable contribution to the literature. Edited by Dr. Carlotta A. Berry and Dr. Brandeis Hill Marshall, two leading Black American women in STEM, the book addresses a longstanding gap in the field by offering a comprehensive and socially responsive guide to understanding and mitigating bias in artificial intelligence systems. Most of all, these two technologists offer a clear, technologically sound set of instructions for responsible model design and maintenance, accompanied by helpful charts and figures.

In a lucid, pedagogical style, Berry and Marshall present directives:

> We need a set of metrics aligned with responsible practices, avoiding damages. These metrics and the performance of the model should be constantly monitored. Build an action plan to identify and respond to failures and damages when they happen. Always try to collect feedback from the population that uses the product built. (Berry & Marshall, 2024, p. 12)

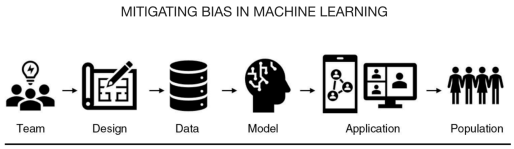


**FIGURE 1.3** Best practices to build a fairer application.
*Source:* Own authorship.

Structured as a collection of topical chapters authored by contributors ranging from full professors to graduate students, the volume reflects a diversity of lived experience and scholarly expertise. Most contributors come from historically marginalized backgrounds, which enriches the perspectives offered. Each chapter focuses on a specific domain, such as healthcare, social media, or natural language processing and includes case studies as well as practical strategies for addressing algorithmic bias. The book engages with both systemic and technical dimensions of bias and presents readers with ethical frameworks alongside actionable tools. It is written with accessibility in mind, without sacrificing technical rigor, making it suitable for interdisciplinary audiences.

The authors underscore that algorithmic bias is not merely a technical issue but one with ethical, political, and social consequences, particularly for communities disproportionately affected by surveillance

and exclusion. The book's treatment of these issues is grounded in real-world implications and informed by a strong commitment to equity.

Educationally, *Mitigating Bias in Machine Learning* is a flexible resource. It introduces foundational concepts in approachable language for secondary and early post-secondary students, while also offering depth and guidance suitable for graduate-level study. For professionals in industry, policy, or public service, it provides practical mitigation strategies and prompts critical reflection on the relationship between innovation and equity. It is well-suited for inclusion in curricula across computer science, engineering ethics, and interdisciplinary AI studies, as well as for use in training programs focused on responsible technology development.

Several chapters stand out for their relevance to current debates. For example, the discussion on large language models addresses the challenges posed by generative AI, while another chapter on hate speech detection explores how sociocultural context influences the interpretation of harmful content. As someone engaged in research on the use of algorithmic systems in immigration and asylum processes, I found the book particularly instructive in framing how ethics must be embedded in system design and how affected communities should be involved from the outset.

*Mitigating Bias in Machine Learning* deserves a broad readership. Its insights and tools have applications far beyond academia, from classrooms and research labs to policy workshops and civic technology initiatives. By foregrounding both ethical responsibility and practical implementation, the book makes a compelling case for a more inclusive and accountable approach to artificial intelligence.