HealthManagement
Promoting Management and Leadership

AI and Real-World Data in Healthcare: A Call for Fairness Assessment

Recent advancements in artificial intelligence (AI), particularly machine learning (ML), have been successfully integrated into healthcare applications, showing promise in risk prediction, disease diagnosis, and sub-phenotyping. The widespread adoption of electronic health record (EHR) systems has provided a wealth of real-world data (RWD) for research, leading to an interest in generating real-world evidence (RWE) to support regulatory decisions. Despite the benefits of AI/ML in analysing RWD for more accurate predictions and efficient identification of disease risk factors, concerns have emerged regarding algorithmic fairness and bias issues, especially towards socioeconomically disadvantaged groups. A recent study published in the Journal of Biomedical Informatics aims to fill this gap by providing an overview of existing literature on fairness assessment, identifying sources of bias, and bias mitigation strategies in ML models using RWD in healthcare.

Addressing Bias in Healthcare: A Call for Comprehensive Fairness Assessment in AI/ML

Algorithmic fairness has become a significant research area, highlighted by examples such as the COMPAS tool used in courts, which exhibited bias against African-American offenders. Similar biases have been observed in healthcare predictions, reflecting underlying societal disparities present in the data used to train ML algorithms. To address these concerns, the AI/ML research community has proposed fairness assessments and mitigation techniques, with some focusing specifically on healthcare applications. However, there remains a need for a comprehensive review of the literature on assessing fairness and mitigating bias when ML techniques are applied to RWD in healthcare. In this comprehensive review, the authors focused on evaluating techniques for ensuring fairness in AI/ML models applied to healthcare using real-world data (RWD). They examined various quantification metrics, publicly available datasets for ML fairness research, and methods for mitigating bias. 35 review articles covering fair ML, fairness assessment, and bias mitigation were analysed, and among them, 11 studies specifically addressed fair ML in healthcare applications using RWD.

Advancing Fair Machine Learning in Healthcare: Challenges and Opportunities

The current state of research in fair machine learning (ML) utilizing real-world data (RWD) has primarily focused on diverse applications and disease domains, demonstrating the potential of fair ML in addressing bias issues within clinical decision support systems. However, there is a need to broaden research efforts to encompass a wider range of health conditions to maximise the impact of fair ML in reducing healthcare inequalities and improving disease outcomes. Several fairness metrics have been employed in RWD studies, with a predominant focus on group fairness metrics such as predictive equality and equalised odds. However, achieving multiple fairness metrics simultaneously remains challenging due to trade-offs between different aspects of fairness. Collaboration with domain and legal experts is crucial to contextualize fairness metrics effectively within healthcare settings. Various techniques for ML bias mitigation have been explored, including pre-processing, in-processing, and post-processing methods. While pre-processing techniques have been widely used, there is a lack of evidence demonstrating their superiority over other approaches. Comprehensive studies comparing the effectiveness of different mitigation strategies are needed to inform the selection of interventions in specific contexts.

Navigating Bias in Multi-Modality Healthcare Data

The emergence of multi-modality data in healthcare research presents both opportunities and challenges for bias mitigation. Current methods primarily focus on single-modality data, and there is a need for innovative techniques to address bias across different modalities, particularly in unstructured data like clinical narratives. Understanding model interpretation through explainable AI (XAI) methods is crucial for identifying the causes of biases and improving transparency in decision-making processes. Enhanced comprehension of model functionalities can lead to more equitable decision-making in healthcare. In addition to mitigating biases in ML algorithms, attention must be paid to data collection issues, including governance and data access. Collaborative initiatives and data-sharing networks can enhance access to representative data samples and mitigate biases in AI/ML models.

Fair machine learning (ML) holds significant promise in enhancing clinical decision support tools to ensure equitable and effective healthcare outcomes. This burgeoning field necessitates increased research attention to developing frameworks to uncover underlying causes of bias in ML algorithms, expanding fair ML research into advanced AI/ML algorithms and broader healthcare data domains, and innovating fair ML methods for multimodality and unstructured data. Systematic investigations into the impact of bias in RWD and analytic methods on study results are also warranted.

Source: [Journal of Biomedical Informatics](#)

Image Credit: [iStock](#)

Published on : Mon, 18 Mar 2024