

Research Discussion Assignment 3

Bikash Bhowmik

18-Jun-2025

Recommender systems, which platforms like Netflix, Amazon, and some social media use, aim to personalize content and improve user experience. However, these systems often depend on historical user behavior and data. This reliance can introduce and escalate existing human biases. For example, if women have historically engaged less with tech job ads, a recommender system might stop displaying such ads to them entirely, which reinforces gender stereotypes. Similarly, biased training data can result in uneven visibility or representation. For instance, music platforms may over-recommend mainstream artists to minority users, or algorithms might show certain ads disproportionately to specific racial or gender groups.

These issues mainly arise from feedback loops, biased data inputs, and unfair targeting based on sensitive user traits. When left unchecked, recommender systems can worsen inequalities in areas like job opportunities, financial products, and media exposure. Real-world cases illustrate this, such as YouTube's algorithm promoting polarizing content to boost engagement, or Netflix customizing thumbnails based on inferred racial preferences. Both situations raise ethical concerns about profiling and discrimination.

Despite these problems, recommender systems can become fairer if designers focus on equity. Approaches like Equality of Opportunity and adversarial debiasing help ensure that model predictions treat all user groups fairly. Using diverse and balanced datasets, along with clear evaluation metrics, improves fairness and reduces bias. By adopting responsible development practices, recommender systems can not only minimize harm but also actively address the biases present in historical and societal data.