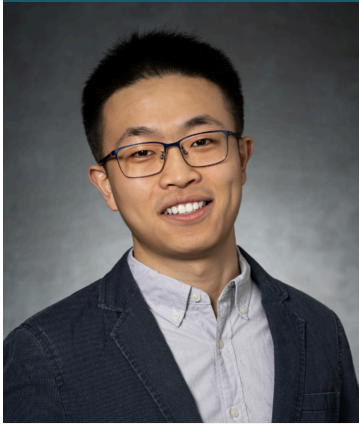


THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Department of Information Systems, Business Statistics and Operations Management

IS SEMINAR ANNOUNCEMENT



Human-Algorithmic Bias: Source, Evolution, and Impact

by

Mr. Xiyang HU
Carnegie Mellon University

DATE	10 November 2023 (Friday)
TIME	10:10 am - 11:40 am
VENUE	4/F ISOM Meeting Room (Room 4047), LSK Business Building

ABSTRACT

Prior work on human-algorithmic bias has seen difficulty in empirically identifying the underlying mechanisms of bias, because in a typical "one-time" decision-making scenario, different mechanisms tend to generate the same patterns of observable decisions. In this study, leveraging a unique repeat decision-making setting in a high-stakes micro-lending context, we aim to uncover the underlying source, evolution dynamics, and associated impacts of bias. We first develop and estimate a structural econometric model of the decision dynamics to understand the source and evolution of potential bias in human evaluators in microloan granting. We find that both preference-based bias and belief-based bias are present in human evaluators' decisions and are in favor of female applicants. Through counterfactual simulations, we quantify the effects of the two types of bias on both fairness and profits. The results show that the elimination of either of the two biases improves the fairness in financial resource allocation, as well as the platform profits. The profit improvement mainly stems from the increase in the approval probability for male borrowers, especially those who would eventually pay back loans. Furthermore, to examine how human biases evolve when being inherited by machine learning (ML) algorithms, we then train a set of state-of-the-art ML algorithms for default risk prediction on both real-world datasets with human biases encoded within and counterfactual datasets with human biases partially or fully removed. By comparing the decision outcomes in different counterfactual settings, we find that even fairness-unaware ML algorithms can reduce bias present in human loan-granting decisions. Interestingly, while removing both types of human biases from the training data can further improve ML fairness, the fairness-enhancing effects vary significantly between new and repeat applicants. Based on our findings, we discuss how to reduce decision bias most effectively in a human-machine learning pipeline.

BIOGRAPHY

Xiyang Hu is a Ph.D. candidate at Carnegie Mellon University. His research focuses on Human-Centered Machine Learning, from both technical and empirical aspects. He utilizes econometric methods to analyze the impact of AI and human-machine interactions. Meanwhile, he develops accessible and user-friendly tools, as well as interpretable, fair, and efficient algorithms, promoting responsible AI use and wide technology access. His papers have been published at top AI/ML conferences and journals, and are under review at top business journals. His research impact extends beyond academia, with industry firms such as Amazon, Alibaba, and Morgan Stanley implementing his work to revolutionize AI applications. Additionally, Xiyang is a recipient of the prestigious Amazon Research Award as a co-PI.