

THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Department of Information Systems, Business Statistics and Operations Management

IS SEMINAR ANNOUNCEMENT



Walk a Mile in Someone Else's Mindset: Cognitive Empathy Priming to Improve Labeling of Controversial Statements

by

Prof. Il-Horn HANN
University of Maryland

DATE	12 December 2024 (Thursday)
TIME	10:30 am - 12:00 noon
VENUE	4/F Meeting Room (Room 4047), LSK Business Building

ABSTRACT

Human-annotated labels are used for both training and evaluating machine learning models in subjective domains. As we show in this paper, the quality of these labels is poor when there is high disagreement among human annotators. We term this the controversial labeling problem. Through three related studies in the context of identifying sexist content using crowdsourced labelers, we establish the prevalence and implications of the controversial labeling problem. We propose Cognitive Empathy Priming as a strategy to improve labeling and demonstrate that this scalable, low-cost treatment increases annotators' ability to take others' perspectives and thereby improves the accuracy in a subjective labeling task. We find that state-of-the-art open source Large Language Models (LLMs) have increased performance when trained using empathy-primed labels.

BIOGRAPHY

Il-Horn Hann, Professor of Information Systems, teaches the the core MBA in the full-time and flex-time program at the Robert H. Smith School of Business at the University of Maryland. His research interest is broadly in the intersection of digital innovations and markets, which currently includes the gaming ecosystem, online incivility, and AI. Il-Horn's papers have been published in Management Science, Journal of MIS, and Communication of the ACM. He served on the editorial board of Information Systems Research and on the editorial board of Management Science. Il-Horn received his MA and PhD from the Wharton School of the University of Pennsylvania. He is the co-director of the school's Center for Digital Innovation, Technology and Strategy (DIGITS).