

THE HONG KONG POLYTECHNIC UNIVERSITY DEPARTMENT OF MANAGEMENT AND MARKETING

Departmental Research Seminar

The Cost of Rigidity – How Class-based Data Collection Affects Machine Learning Performance

By



Prof. Jeffrey Parsons
Memorial University of Newfoundland

Date : 7 May 2025 (Wed)
Time : 10:30 am – 12 noon
Venue : M802, PolyU

Abstract

The design of data collection interfaces can significantly influence the quality of training data for machine learning (ML) systems. While class-based designs—which structure information into predefined categories—are widely used, their rigid classifications risk omitting critical details. This study compares class-based data collection with an instance-based approach, where contributors freely describe observed phenomena, to assess their impact on ML performance. We conducted a controlled experiment in autonomous driving. Participants (N = 260) described driving scenes using class-based or instance-based interfaces. Machine learning models trained on these datasets were evaluated for predictive accuracy. Results show instance-based data outperformed class-based across all models, with higher accuracy and robustness, improving generalization to real-world complexity. In cases where data is crowdsourced to train machine learning models, this work shows the benefits of instance-based data collection design that integrates structured input with open-ended reporting to enhance ML performance.

Prof. Jeffrey Parsons is University Research Professor and Professor of Information Systems in the Faculty of Business Administration at Memorial University of Newfoundland in Canada. His research interests focus on how to better represent human conceptualizations of the world in data and on data repurposing. His work on this and related topics has appeared in journals in several disciplines, including *MISQ*, *ISR*, *Management Science*, *JAIS*, *JMIS*, *ACM ToDS*, *IEEE TKDE*, and *Nature*. Prof. Parsons' research has been recognized a lifetime achievement award from DESRIST, *MISQ* paper of the year, and the INFORMS Design Science Research Award. He is a fellow of the Association for Information Systems, a Distinguished Research Fellow at TU Dresden, a Schoeller Senior Fellow, and an ER Fellow. He currently serves as a Senior Editor at *Information Systems Research*.

All interested are welcome.



THE HONG KONG
POLYTECHNIC UNIVERSITY
香港理工大學

Department of
**MANAGEMENT
& MARKETING**
管理及市場學系