UNIVERSITY of GUELPH

COLLEGE *of* ENGINEERING AND PHYSICAL SCIENCES

SCHOOL OF COMPUTER SCIENCE

MSc Defence

Tuesday May 7, 2024, at 1PM, online via Zoom (Remote)

Fatemeh Haghighi

*Evaluating End-Users' Perspectives Toward
Explainability in Artificial Intelligence*

Chair: Dr. Fei Song
Advisor: Dr. Rozita Dara
Advisory: Dr. David Flatla
Non-Advisory: Dr. Charlie Obimbo

Abstract:

The rapid evolution of Artificial Intelligence (AI) algorithms in decision-making has profoundly impacted various fields, offering promising advancements. As these advances have been made, explainability has gained prominence, particularly in areas with direct human impact. As a result, Explainable Artificial Intelligence (XAI) has emerged as an essential pursuit, aiming to provide understandable explanations for the decisions of AI systems. Although there has been extensive research in XAI, there is a gap in the literature regarding end-users' perspectives on black-box and white-box AI decision-making algorithms regarding AI explanations and prediction abilities. This thesis addresses the gap by investigating users' perspectives on AI decision-making and XAI, focusing on balancing accuracy and interpretability in AI algorithms for different applications.

Two experiments were conducted: Experiment I, which involved developing and evaluating AI algorithms and explanation methods, and Experiment II, which conducted user interviews to assess their perceptions of AI and XAI. This study contributes to a deeper understanding of users' perspectives on XAI methodologies and the accuracy and explainability aspects of AI algorithms. This research will bridge the gap between human needs and AI development, enabling the development of a more user-friendly and transparent AI system.