CSDS 500 and ECSE 500
Fall 2021 Colloquium

11:30AM to 12:30PM
Tuesday, November 23, 2021
Virtual
Zoom Webinar ID: 940 7438 8634
Passcode: 357363

"Robust and Explainable Machine Learning Algorithms for Social Good"

Abstract: Deep Neural Networks (DNNs) are complex nonlinear functions parameterized by model weights to map inputs to outputs that have attracted much attention in machine learning community due to its state-of-the-art performance on various tasks. Despite the initial success, robustness and explainability of a complex DNN remain an open problem, hindering its wide deployment in safety and security-critical domains. In this talk, I will introduce a new algorithm, Attentive Multitask Collaborative Filtering (AMCF), to tackle the interpretability problem of recommender systems by integrating a feature mapping strategy into the recommender systems and novel metrics to evaluate the quality of explanation, and a new gradient based DNN interpretation algorithm, Adversarial Gradient Integration (AGI), which utilizes backpropagation and adversarial effects via exploiting adversarial examples. I will also introduce a new loss function as a drop-in replacement for standard cross-entropy loss to improve DNN's adversarial robustness.

Dongxiao Zhu
Wayne State University

Bio: **Dongxiao Zhu** is currently an Associate Professor of the Department of Computer Science at Wayne State University. He earned PhD from University of Michigan, Ann Arbor in 2006 and his current research interest lies in developing robust, explainable, and fair machine learning algorithms with applications to public health and safety, clinical medicine, cybersecurity, and human mobility for social good. Dr. Zhu is a leading expert in AI and machine learning with applications to health, social and urban computing and he is the founding director of Wayne AI Research Initiative (http://ai.wayne.edu/) and the Director of Trustworthy AI research lab (https://dongxiaozhu.github.io/) at Wayne State University.

This is to certify that ________________________________ attended this seminar. Certified by ________________________________.
Certificates of attendance and other evidence of CPD activity should be retained by the attendee for auditing purposes.

CASE SCHOOL
OF ENGINEERING
CASE WESTERN RESERVE
UNIVERSITY