



IST Seminar Series Presents:

Knowledge-Guided Machine Learning for Scientific Discovery: Challenges and Opportunities

Dr. Xiaowei Jia

Associate Professor, Department of Computer Science
University of Pittsburgh

Monday, April 27, 2026 | 10:30-11:30AM | UH at Sugar Land, AMG 300

<https://teams.microsoft.com/meet/216474889214755?p=nqxcBPxFHWPqxnLFgi>

Meeting ID: 216 474 889 214 755 Passcode: E3Jm73VQ

Abstract: Data science and machine learning (ML) models, which have found tremendous success in several commercial applications where large-scale data is available, e.g., computer vision and natural language processing, has met with limited success in scientific domains. Despite their extensive use, these models have several well-known limitations due to incomplete or inaccurate representations of the physical processes being modeled. Given rapid data growth due to advances in sensor technologies, there is a tremendous opportunity to systematically advance modeling in these domains by using machine learning methods. However, capturing this opportunity is contingent on a paradigm shift in data-intensive scientific discovery since the “black box” use of ML often leads to serious false discoveries in scientific applications. Because the hypothesis space of scientific applications is often complex and exponentially large, an uninformed data-driven search can easily select a highly complex model that is neither generalizable nor physically interpretable, resulting in the discovery of spurious relationships, predictors, and patterns. This problem becomes worse when there is a scarcity of labeled samples, which is quite common in science and engineering domains.

My work aims to build the foundations of knowledge-guided machine learning (KGML) by exploring several ways of bringing scientific knowledge and machine learning models together. In particular, we discuss gaps and opportunities in scientific discovery and show the effectiveness of KGML in multiple applications of great societal and scientific relevance.

Biosketch: Xiaowei Jia is an Associate Professor in the Department of Computer Science at the University of Pittsburgh. He got his Ph.D. degree from the University of Minnesota. His research interests include knowledge-guided data science and spatio-temporal data mining for real-world applications of great societal relevance. A major highlight of his research is a general paradigm Knowledge-guided Machine Learning for combining machine learning and physics-based modeling approaches. He is leading multiple projects funded by NSF, NASA, and USGS. He is also the recipient of the NSF CAREER Award, NASA Early Career Investigator Award, the University of Minnesota Best Dissertation Award, and best paper awards from ICDM, SDM, ASONAM, and BIBE.