DEPARTMENT OF COMPUTER SCIENCE

TEXAS TECH
Whitacre College of Engineering

Explainable AI for Cybersecurity

Prabhat Mishra, Ph.D.
University of Florida
Tuesday, February 4, 2025
4:30 p.m.
Zoom: https://texastech.zoom.us/my/stas.tiomkin
Meeting ID: 544 589 4586

Abstract: This talk will provide a comprehensive overview of security attacks as well as detection techniques using explainable AI. First, I will outline a wide variety of software and hardware security threats and vulnerabilities. Next, I will introduce explainable AI algorithms to interpret machine learning behaviors in a human-understandable way. I will discuss state-of-the-art attack detection methods using explainable AI. I will also cover how to enable hardware acceleration of explainable AI models for real-time vulnerability detection. Finally, I will discuss the security threats toward machine learning models, and effective countermeasures to design robust AI models.

Bio: Prabhat Mishra is a Professor in the Department of Computer and Information Science and Engineering at the University of Florida. His research interests include embedded systems, hardware security, energy-aware computing, system-on-chip validation, trustworthy AI, and quantum computing. He has published 9 books, 35 book chapters, 26 patents, and more than 250 research articles in premier international journals and conferences. He is an IEEE Fellow, a Fellow of the American Association for the Advancement of Science, and an ACM Distinguished Scientist.