**Title:** Exploring Trustworthy AI: Fundamentals, Challenges, and Outlooks

**Abstract:** Artificial Intelligence (AI) algorithms including Deep Neural Networks (DNN) and Large Language Models (LLM) have achieved tremendous success in the last decade. However, recent research investigations show that AI systems may lead to hallucinations and/or be vulnerable to small perturbations or malicious attacks, making them less trustworthy to be applied in many real scenarios. This talk will first address fundamentals and theories of AI including its history, milestones, and typical models such as DNN and LLM. Highlights will be also put on the challenges and ethics of AI as well as possible future research directions.

**Bio:** Kaizhu Huang is Tenured Professor and Director of Digital Innovation Research Center, at Duke Kunshan University (DKU). He was Professor and Associate Dean of Research, School of Advanced Technology, Xi’an Jiaotong-Liverpool University before joining DKU. Prof. Huang obtained his PhD degree from Chinese University of Hong Kong (CUHK) in 2004. He worked in Fujitsu Research Centre, CUHK, University of Bristol, National Laboratory of Pattern Recognition, Chinese Academy of Sciences from 2004 to 2012. He published more than 280 international papers including 140+ journal papers. He was the recipient of the 2011 Asia Pacific Neural Network Society Young Researcher Award. He also received the best (or runner-up) paper or book awards about 10 times in major AI and machine learning conferences. In particular, he received the 2024 IEEE ICDM 10 Years Highest-Impact Paper Award. He acts as Editor-in-Chief of Elsevier CSSI and serves as associated editors/advisory board members in 6 international journals and book series (e.g. Pattern Recognition Journal, and Neural Network Journal). He was invited as a keynote/tutorial speaker to more than 50 international conferences or workshops.