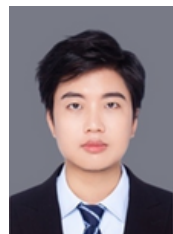## Session Chairs: Penglin Dai, Southwest Jiaotong University

Dr. Dai is currently an Associate Professor at the School of Computing and Artificial Intelligence, Southwest Jiaotong University (SWJTU), in Chengdu. He has authored and co-authored more than 70 papers at the top international conferences and journals such as AAAI, INFOCOM, TMC, TSC, TITS, TSMC.

Xincao Xu, University of Electronic Science and Technology of China

Dr. Xu is currently an Associate Researcher at the Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China (UESTC). From 2023 to 2025, he was a Postdoctoral Research Fellow with the Shenzhen Institute for Advanced Study, UESTC. He has authored and co-authored more than 15 papers.

Brief Description of the Session Thematic:

Edge Intelligence (EI) represents a transformative computing paradigm that brings artificial intelligence and machine learning capabilities directly to the source of data generation, i.e., the network edge. By deploying and executing AI algorithms on edge devices, such as IoT sensors, smartphones, and autonomous vehicles, EI overcomes the critical limitations of traditional cloud-centric AI, including high latency, bandwidth constraints, and data privacy concerns. This shift enables real-time data processing, instantaneous decision-making, and personalized user experiences with enhanced security. However, unlocking the full potential of EI presents a complex set of challenges, including: a) Algorithmic and Model Innovation. Developing lightweight and efficient AI models that can operate under the severe resource constraints (computation, memory, power) of edge devices; b) Hardware and System Heterogeneity. Designing and managing AI workloads across a diverse and fragmented ecosystem of edge hardware with varying capabilities; c) Data Governance and Security. Ensuring robust data privacy and security in decentralized environments, especially in multi-stakeholder scenarios like federated learning; and d) Reliability and Manageability. Guaranteeing the robustness and scalability of distributed intelligent systems, especially when facing unreliable network conditions and the need for seamless model deployment and updates.

Related topics for this special session (but not limited to) :

• Lightweight, efficient, and robust machine learning models for the edge (e.g., model compression, quantization, pruning, neural architecture search).

• Theoretical limits and performance bounds of edge computing.

• Novel learning paradigms for the edge (e.g., federated learning, split learning, continual learning).

• Hardware accelerators and energy-efficient processors for AI on edge devices.

• Compilers, runtimes, and middleware for deploying AI models on heterogeneous edge hardware.

• Edge-Cloud collaborative intelligence and computing orchestration.

• Collaborative architectures for large and small models (e.g., collaborative inference between cloud-based large models and edge-based small models). • Operating system support for edge intelligence. • Privacy-preserving machine learning on the edge.

• Security and trustworthiness of distributed edge AI systems.

• Data management and analytics at the network edge. • Adversarial attacks and defenses for edge AI models.

• Edge AI in Industrial IoT (IIoT), smart manufacturing, and robotics.

• Intelligent applications for smart homes, smart cities, and environmental monitoring.

• On-device AI for autonomous vehicles, drones, and transportation systems.

• Edge intelligence in healthcare, wearables, and mobile computing.

• Real-time augmented reality (AR) and virtual reality (VR) powered by the edge.

• Metrics, benchmarks, and datasets for evaluating edge AI systems.

• Performance analysis and optimization of edge AI applications.