Einladung zum Vortrag

20. Juni 2022, 16.00 Uhr c.t.
Universität Bremen | MZH 4380

Prof. Dr. Jinjun Xiong
(University at Buffalo)

Accelerating AI Systems Innovations to Democratize AI for a better Society

AI has become an increasingly powerful technology force that will transform different aspects of our society, such as transportation, healthcare, education, and scientific discovery. But the diverse layers of software abstractions, hardware heterogeneity, privacy and security concerns, and big-data driven machine learning models have made the development of optimized AI solutions extremely challenging. This results in the business world's expensive investment only on a handful of selective and "profitable" AI solutions, leaving many critical societal needs, such as equitable education and sustainability, much less addressed than deserved. To truly democratize the power of AI for the benefit of the society and humanity, AI systems innovation, driven by "earthly" AI solutions involving big data, AI models, and hybrid cloud computing infrastructures, holds the key to drastically simplify the development of AI solutions and greatly improve AI productivity at a lower cost. This talk will discuss some of my related research efforts in the past decade on developing enterprise-scale high performance AI solutions and innovating related AI systems technologies, in particular, a data-driven reviewer recommendation AI solution, automation tools to identify system performance bottlenecks, and innovative software-hardware co-optimization algorithms for accelerating AI algorithms. I will contextualize these efforts in an ultimate research goal of transforming the current computing paradigm with AI systems innovation to truly democratize AI for a better society.

Biografie

Dr. Jinjun Xiong is currently Empire Innovation Professor with the Department of Computer Science and Engineering at University at Buffalo (UB). Prior to that, he was a Senior Researcher and Program Director for AI and Hybrid Clouds Systems at the IBM Thomas J. Watson Research Center. He co-founded and co-directed the IBM-Illinois Center for Cognitive Computing Systems Research from 2016-2021, the success of which led to the $200M 10-year investment to establish the IBM-Illinois Discovery Accelerator Institute in 2021. His research interests are on across-stack AI systems research, which include AI applications, algorithms, tooling, and computer architectures. Many of his research results have been adopted in IBM's products and tools. He published more than 150 peer-reviewed papers in top AI conferences and systems conferences. His publication won seven Best Paper Awards and eight Nominations for Best Paper Awards. He also won top awards from various international competitions, including the recent champion for the IEEE GraphChallenge on accelerating sparse neural networks, and champions for the DAC'19 Systems Design Contest on designing an object detection neural network for edge FPGA and GPU devices.

Dieser Gast wurde von Rolf Drechsler eingeladen.