UNIVERSIDAD COMPLUTENSE MADRID

Facultad de Informática

# ANUNCIO DE CONFERENCIA

POSGRADO

# AI Hardware for Real-Time Machine Learning

Seda Ogrenci-Memik
Northwestern University

Facultad de Informática
meet.google.com/yxx-btyd-qzi - miércoles 17 de marzo de 2021 - 17:00
*Entrada libre hasta completar el aforo*

**Resumen:**

Computing systems for AI workloads have evolved towards data-center clusters of GPUs and TPUs, with architectures optimized for performing linear algebra and tunable for variable precision. As new AI paradigms emerge, more distinct divergence between hardware architectures for powering AI and other workloads are observed. GPU manufacturers are developing different architectures and chipsets for the HPC/supercomputing, cloud, edge computing, and robotics domains. FPGA vendors are also joining this ecosystem (e.g., Intel FPGAs deployed within Microsoft Azure). Moving forward, many industries and services ranging from cloud computing to consumer electronics are making hardware-accelerated AI a prominent component in their portfolio. In this talk, some examples of AI hardware architectures and available silicon technologies will be presented. The concept of co-design will be discussed. This makes the unique needs of an application domain transparent to the hardware design process. Finally, an overview of design automation tool flows will be presented to gain an understanding of how to support a high productivity framework for domain experts to design and deploy AI hardware.

**Sobre Seda Ogrenci-Memik:**

Seda Ogrenci-Memik (IEEE Senior Member in 2005) received the B.S. degree in electrical and electronic engineering from Bogazii University, Istanbul, Turkey, and the Ph.D. degree in computer science from the University of California at Los Angeles, Los Angeles, CA, USA. She is currently a Professor with the Electrical Engineering and Computer Science Department, Northwestern University, Evanston, IL, USA. Furthermore, she is the Director of Computer Engineering Division at ECE Department. Her research interests include embedded and reconfigurable computing, HLS, thermal-aware design automation, and thermal management for microprocessor systems. She has served as a technical program committee member, an organizing committee member, and the track chair of sev- eral conferences, including ICCAD, DATE, FPL, GLSVLSI, and ISVLSI. She received the National Science Foundation Early Career Development (CAREER) Award in 2006. She is currently serving on the Editorial Board of the IEEE Transactions on Very Large Scale Integration. In April 2021 she will be joining Google during her sabbatical leave.