The Next Horizon of AI Hardware: Silicon That Learns Continually

Dhireesha Kudithipudi

UT San Antonio
E-mail address: dk@utsa.edu

Abstract:

Learning from streaming real-world data remains a core challenge for standard AI systems. Real-world environments generate nonstationary data that arrives incrementally, demanding models that can continually learn new knowledge without forgetting what was learned in the past. Conventional models, however, assume stationary data distributions and depend on high-precision, idealized weights. Although various algorithmic advancements have been proposed to mitigate forgetting, they become inefficient when deployed on emerging non-volatile memory devices like FeFETs and ReRAM whose limited weight precision and intrinsic stochasticity conflict with the assumptions underlying these methods. Overcoming catastrophic forgetting on edge hardware requires an algorithm hardware co design strategy that actively embraces these device non idealities within sub-watt power budget.

Grounded in our team's brain inspired continual learning framework, this talk presents how biological principles, such as metaplasticity, neuromodulation, near memory processing, and spiking neural architectures, can be natively mapped onto custom silicon to achieve continual learning.