Achieving Neural and Quantum Supremacy in AI/ML

Debanjan Bhowmik
Assistant Professor,
Department of Electrical Engineering,
Indian Institute of Technology Bombay
Conventional Implementation of AI/ML

Device (conventional silicon MOSFET)

System (digital/von Neumann architecture)

Develops some functionalities of the brain (sometimes better than the brain)

Some functionalities are missing

High power requirement
AI/ML: A Common Neural Network

Input layer

- $x_1$
- $x_2$

Hidden layer

- $z_1$
- $z_2$
- $z_3$

Output layer

- $y_1$
- $y_2$
- $y_3$

Non-Linear Activation Function (Neuron)

- $f$

Synapses

Weight Update

- $w_{1,1}$
- $w_{1,2}$
- $v_{1,1}$
- $v_{1,3}$

Summation

- $\sum$

Total Output

- $Y_1$
- $Y_2$
- $Y_3$

Non-Linear Activation Function (Neuron)

- $f$

Output layer

- $Y_4$
- $Y_5$
- $Y_6$

Synapses

Weight Update

- $w_{2,1}$
- $w_{2,2}$
- $v_{2,1}$
- $v_{2,3}$

Summation

- $\sum$

Total Output

- $Y_7$
- $Y_8$
- $Y_9$
Conventional Implementation of AI/ML

Vector-Matrix Multiplication (VMM), Non-linear activation function (Neuron), Weight update:

Instructions implemented on the CPU

Synaptic weight values and all instructions stored in the memory
Advantage: Highly Generalizable

Theory: Universal Turing Machine (can implement most algorithms)

Implementation: Von Neumann architecture
Advantage: Easily Implementable

Silicon-transistor-based processor

Volatile and non-volatile memory (NVM) banks
Disadvantage: Less Parallelism/ Low Speed

Way to overcome:

More parallelization (GPU and TPU)
Disadvantage: High Power Consumption

- **GPU**
  - Kilowatts and Megawatts of Power!!!

- **Supercomputer (IBM Blue Gene)**
  - Kilowatts and Megawatts of Power!!!
Edge AI/ tiny ML: Where Low Power Is Required

State of the art:
Learning, Inference -> cloud

Target:
Learning -> cloud, Inference -> local
Learning, Inference -> local

R. Sanchez Imora and Antonio Skarmeta, IEEE Circuits and Systems Magazine, 2020
Neuromorphic Implementation of AI/ML:
Inspired by the brain, aims at low power

Our brain only consumes 20 W of power.

It can do many things which modern AI/ML algorithms can’t.
Overview of Group’s Research: Neuromorphic and Quantum Implementation of AI/ML

- Conventional transistor
- Charge-based NVM
- Spin devices

Devices: New functionalities

Circuits: New architectures

Oscillator

Quantum

Novel Algorithms
Spintronic domain-wall-device-based neuromorphic computing:


Spintronic-oscillator-based neuromorphic computing:


Quantum Machine Learning:
