Course syllabus

Memory Technology for Machine Learning
Minnesteknologi för maskininlärning

EITP25, 7.5 credits, A (Second Cycle)

Valid for: 2024/25
Faculty: Faculty of Engineering LTH
Decided by: PLED C/D
Date of Decision: 2024-04-16
Effective: 2024-05-08

General Information

Main field: Nanoscience **Depth of study relative to the degree requirements:** Second cycle, in-depth level of the course cannot be classified
Elective for: E4-is, F4, F4-fel, MNAV1, MSOC2, N4-hn
Language of instruction: The course will be given in English

Aim

The purpose of this course is to give an in depth understanding for the physics of common memory device technologies with focus on non-volatile memories. Furthermore, the course covers how these memory devices can be integrated to create neuromorphic hardware for applications in machine learning and artificial intelligence. Finally, the course gives an introduction to the architectures and algorithms that are used in machine learning, to give a basic understanding for the needs that memory devices and their connections need to fulfil.

Learning outcomes

Knowledge and understanding
For a passing grade the student must

- Explain in general the memory hierarchy of a modern computer and in detail how its memory components function.

- Explain the physical processes that determine the functionality of common non-volatile memory types such as RRAM, PCM, FeRAM and STT-MRAM.
- Understand in general the limitations and benefits of the various memory technologies treated in the course.
- Understand in general how the integration of memory devices into neural network circuits can be realized and the benefits and limitations of these approaches.
- Explain in detail how learning occurs in a spiking neural network and how memristors can be used in such networks.
- Explain in general how to perform training of a neural network with the help of backpropagation and the gradient descent algorithm.

Competences and skills
For a passing grade the student must

- perform measurement and analysis of the current-voltage characteristics of a RRAM device.
- perform measurements and from a measured polarization-field diagram be able to extract important parameters for a ferroelectric capacitor.
- be able to give suggestions on how speed, reliability and energy consumption can be improved in the memory device technologies treated in the course.
- design and train a spiking neural network to perform image recognition.

Judgement and approach
For a passing grade the student must

- realize the need for energy efficient and scalable neuromorphic hardware for machine learning and AI.
- Evaluate the applicability of a given memory technology for a range of application areas with respect to the pros and cons of the technology.

Contents

Memory devices of the computer: SRAM, DRAM, NAND

Non-volatile memory devices: The memristor. Resistive memories (RRAM), phase change memories (PCM), ferroelectric memories (FeRAM), magnetic memories (MRAM).

Integration of memory devices: 3D stacking for scalability, crossbar architecture

Neural network architectures: Fully connected networks, convolutional networks, recursive networks. spiking networks.

Machine learning methods: Backpropagation, gradient descent, STDP.

Examination details

Grading scale: TH - (U, 3, 4, 5) - (Fail, Three, Four, Five)
Assessment: Written exam, project work, lab work with report.

The examiner, in consultation with Disability Support Services, may deviate from the regular form of examination in order to provide a permanently disabled student with a form of examination equivalent to that of a student without a disability.

Modules
Code: 0119. **Name:** Written Exam.
Credits: 4.0. **Grading scale:** TH - (U, 3, 4, 5). **Assessment:** Exam **The module includes:** Written examination that will cover all the topics of the course.
Code: 0219. **Name:** Lab Exercise.
Credits: 1.0. **Grading scale:** UG - (U, G). **Assessment:** Passing grade upon participation in the lab

exercise and an approved lab report. **The module includes:** Practical lab exercise with subsequent lab report.

Code: 0319. **Name:** Project Assignment.
Credits: 2.5. **Grading scale:** UG - (U, G). **Assessment:** The level of understanding reflected in the written report as well as a control of the written program code will determine whether a passing grade has been achieved. **The module includes:** The project consists of designing and training a neural network to recognize at least 90% of the images from a predefined database. The project should be executed alone or in groups of two persons. The project should result in a written report and working MATLAB code.

Admission

Assumed prior knowledge: Basic knowledge in solid state physics.
The number of participants is limited to: No

Reading list

- An Chen, James Hutchby, Victor Zhirnov, George Bourianoff: Emerging Nanoelectronic Devices. Wiley, 2014, ISBN: 9781118958254.

Contact

Course coordinator: Mattias Borg, mattias.borg@eit.lth.se
Course homepage: https://www.eit.lth.se/course/eitp25

Further information

The course uses Canvas for communication, hand-ins and study materials.