6.5930/1

Hardware Architectures for Deep Learning

DLH-24 Program Committee

March 20, 2024

Joel Emer and Vivienne Sze

Massachusetts Institute of Technology
Electrical Engineering & Computer Science

Paper Review Schedule and Deliverables

- Paper Review (10% of grade)

- We will be forming a program committee of the (fictional) *Deep Learning Hardware (DLH)* Conference 2024
  - Learn how to read papers (critique and extract information)
  - Use concepts from class to analyze and gain deep understanding of paper
  - Enable wider coverage of papers
  - Gain insight of how paper decisions are made

Paper Review Schedule and Deliverables

- April 1 – **Two papers** assigned (check for undetected conflicts, e.g., advisor or past co-authors)
  - One paper on dataflow/mapping and one paper on sparsity
- April 19 – Reviews due 11:59PM (Phase I)
  - Note: reviews must be submitted on time
  - Do not change the review after the deadline as it will be reflected as a late submission!
- April 26 – Online Paper discussion plus decision complete for **one** of the papers (Phase II)
  - You will be assigned to participate in the discussion of either the dataflow/mapping or sparsity paper

HotCRP Website

Overall merit *

- ○ **1.** Should have been rejected
- ○ **2.** Just above the bar
- ○ **3.** Average
- ○ **4.** Award quality

Paper summary

Summarize paper in your own words (i.e., do not copy and paste from paper). and use the common terminology you learned in class (e.g., activations and filter weights not neurons and synapses). What is the problem that this paper is addressing? What was their solution? And how is it different from existing solutions?

Markdown styling and LaTeX math supported · Preview

Paper strengths (hidden from authors)

Write two or three sentences describing the strengths of the paper. Please focus on the technical content and insights. Evaluate the paper in terms of innovation and impact (i.e., focus on an idea or principle not that the results were 10x faster than prior papers) . Has this work been done before (at the time of publication)? Does the technical contribution address an important challenge? Does the paper provide important insights?

Markdown styling and LaTeX math supported · Preview

https://mit-dlh24.hotcrp.com/

Submitted Reviews Should Capture

- Paper Summary (In your own words!)
  - Motivation of Paper: What is the problem? Why is it important? What are the current solutions (i.e., previous work)?
  - Give insight on contributions of proposed techniques in paper. Describe in terms of concept discussed in class (e.g., stationarity of the dataflow)
  - Discuss key results in paper and how do they support claimed contributions
  - What are the main takeaways from this paper
- Strengths of paper
- Weakness of paper
- Innovation
  - Describe key insights based on concepts from class

Submitted Reviews Should Capture

- Quality of Evaluation
  - Do the results backup the claimed benefits? What baseline was used? Are the experiments unbiased? Are tradeoffs and overheads evaluated?
  - See https://gernot-heiser.org/benchmarking-crimes.html
- Possible extensions to the paper
  - Not just a list of additional techniques (e.g., combine with X, Y, Z). Explain how various approaches could address specific short comings of existing work.
- Questions for Authors
  - What parts of the paper was not clear? What would you ask authors?
- Comments for professors

Other Considerations

- Graded based on ability to give intuition on technique, interpretations of graphs and data, etc.
  - Use of insights from class to analyze papers
    - e.g., even if authors don't explicitly state dataflow, can you identify which dataflow is used and why? What are the strengths and weaknesses of this?
  - Refer to "Evaluating DNN" lecture on performing a comprehensive evaluation
  - Critique not about writing and figures but about quality of technology
- Depth better than breadth (particularly summary)
  - Extensions not just a list of other work to add on top, but insight in terms of ***how*** to improve weakness; ***why*** approach would be a better on top of this
- **We encourage you to draw on your expertise from outside class**
  - **Diverse background: Devices, Algorithms, Architecture – we want to learn from your expertise**