

DAVID THORSLEY

Phone: (206) 972-3263 | E-mail: david.thorsley@gmail.com

LinkedIn: <https://www.linkedin.com/in/david-thorsley/> | Personal Page: <https://torsli.github.io>

SUMMARY

- Experienced researcher and software engineer in the deep learning, machine learning, hardware/software co-design, and optimization spaces
- Passionate about exploring cutting-edge research and building tools that make innovative ideas accessible to all

EXPERIENCE

- **Deep Learning Senior Staff Software Engineer, Samsung Semiconductor Inc., San José, CA, 07/19 – 02/24.**
 - Led the algorithms group within the Samsung Advanced Institute of Technology Neural Processor Lab, handling both technical leadership and people management
 - Researched and developed methods for efficient deep learning, including quantization, pruning, and neural architecture search
 - Researched efficient deep learning methods for multiple applications, including computer vision, natural language processing, recommendation systems, and 3D scene understanding
 - Collaborated closely with hardware architecture researchers to evaluate proposed deep learning accelerator architectures for task accuracy, throughput, and related metrics
 - Used commercial software tools and developed custom software tools to profile AI and HPC workloads and develop cycle-accurate hardware simulations
 - Managed external university research collaborations and performed joint research projects with academia
 - Prepared multiple research publications and patent applications
- **Senior Deep Learning Engineer, Mythic AI, Redwood City, CA, 09/17 – 07/19.**
 - Early member of the software team for an artificial intelligence chip startup developing a novel deep neural network (DNN) accelerator for low-power inference on the edge
 - Responsible for the software aspects of DNN hardware-software co-design, i.e., understanding how hardware mixed-signal design choices affect DNN accuracy & performance
 - Collaborated with the analog & digital hardware teams to develop custom software models and simulations of hardware artifacts
 - Designed and prototyped the software for an *accuracy estimator* tool that models the effects of quantization, noise, and other hardware artifacts in Keras
 - Designed, developed, and supported software for a *retraining* tool that allows users to train or retrain quantized DNNs that are hardened with respect to noise and highly performant on Mythic hardware
 - Used these tools to train the first DNN to run on Mythic prototype hardware
 - Collaborated with our internal AI research team on computer vision problems such as object detection and semantic segmentation
- **Senior Software Engineer, Quantcast, San Francisco, CA, 05/16 – 09/17.**
 - Developed and supported the programmatic brand advertising product pipeline, including: user experience in configuring ad campaigns; development of machine learning models to reach desired audiences; and monitoring campaign performance
 - Developed a new reporting system for maintaining records of data source quality and tracking performance and revenue of our product lines

EXPERIENCE (CONT'D)

- **Analytics Development Engineer, Teradata Aster, San Carlos, CA, 08/14 ~ 05/16.**
 - Implemented advanced distributed algorithms in the Aster SQL-MapReduce framework for the Aster Analytics software package. These algorithms covered many areas of statistics, data science, and machine learning, including clustering, text analysis, time-series analysis, etc.
 - Designed APIs for the analytics team to automate the parsing and validation of user input, leading to massively increased velocity for development engineers and quality engineers and increased consistency in the user experience
- **Mathematical Modeler, Immunetrics, Pittsburgh, PA, 06/13 ~ 07/14.**
 - Predicted the outcomes of clinical trials using biomathematical modeling, machine learning and stochastic optimization approaches on large databases of virtual patients
- **Research Scientist II, Biotechnology High Performance Computing Software Applications Institute, Telemedicine and Advanced Technology Research Center, U.S. Army Medical Research and Materiel Command, Frederick, MD, 09/10 ~ 04/13.** (Employed under contract from the Henry M. Jackson Foundation for the Advancement of Military Medicine.)
 - Developed individualized biomathematical time-series models of the effect of sleep loss on the human stress response system using stochastic optimization
 - Developed novel statistics to assess between-individual variability in cognitive performance impairment caused by sleep loss

EDUCATION

- Research Associate (Postdoctoral Position), Electrical Engineering, University of Washington, Seattle, WA, 10/06 ~ 09/10.
- Ph.D., Electrical Engineering: Systems, University of Michigan, Ann Arbor, MI, 09/00 ~ 05/06.
- M.S., Electrical Engineering: Systems, University of Michigan, Ann Arbor, MI, 09/00 ~ 12/02.
- B.E.Sc., Electrical Engineering, University of Western Ontario (Western University Canada), London, ON, 09/96 ~ 06/00.

PROFESSIONAL SKILLS

- Programming languages: Python, Java, C/C++
- Deep learning frameworks: Pytorch, Keras, TensorFlow
- Machine learning/Data science: numpy, scipy, scikit-learn, pandas, R, SQL, MATLAB, Mathematica
- Data engineering: Docker, AWS
- Software development practices: agile development, scrum, Jira/Confluence, git

RECENT PUBLICATIONS AND PREPRINTS

- L. Li, D. Thorsley and J. Hassoun. "SaiT: Sparse Vision Transformers through Adaptive Token Pruning." arXiv:2210.05832, October 2022.
- S. Kim*, S. Shen*, D. Thorsley*, A. Gholami, W. Kwon, J. Hassoun, and K. Keutzer, "Learned Token Pruning for Transformers." In *KDD '22: Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, August 2022. (*equal contribution)
- J. Fang, A. Shafiee, H. Abdel-Aziz, D. Thorsley, G. Georgiadis, and J.H. Hassoun, "Post-training Piecewise Linear Quantization for Deep Neural Networks." In *Proceedings on the 2020 European Conference on Computer Vision*, August 2020.
- Other publications and patent applications are listed on my webpage: <https://torsli.github.io>