Tokey Tahmid

Tennessee, USA | +1(865)438-2188 | ttahmid@icl.utk.edu
linkedin.com/in/ttahmid | https://tokey-tahmid.github.io/

QUALIFICATION SUMMARY

Passionate AI engineer developing performance API for AI hardware and software at the Innovative Computing Laboratory.

SKILLS

Programming: • Expert in Python, C, C++, C#, CUDA, Fortran, Javascript (Node.js), and Java • Intermediate in Assembly, PHP, Rust
AI Expertise: Machine Learning, Deep Learning, Reinforcement Learning, Large Language Models, Neuromorphic Computing, Parallel Computing, Distributed Computing, High Performance Computing, Mixed Precision, Benchmarking, and Performance Analysis
Backend & DevOps: • API Design & Development • Databases (RDBMS, PostgreSQL, MySQL, MongoDB) • Cloud Platforms (AWS, GCP, Azure) • CI/CD · Docker · Kubernetes • Web Frameworks (Express.js, Flask, and FastAPI) • Performance Optimization / Profiling
Software Frameworks: TensorFlow, Horovod, CMAKE, MPI, PyTorch, Scikit-learn, Scipy, Score-P, mpiP, Tau, Perf, PAPI, CodeCarbon
Technical Skills: • Excellent in problem solving and analytical thinking • Experienced with manipulating and analyzing complex large data sets for AI problems • Experienced with writing technical reports, documentation, papers, posters, and presentations

WORK EXPERIENCE

Innovative Computing Laboratory (ICL) | Research Associate | Tennessee, USA | Feb 2025 – Present
- Developing Performance Application Programming Interface (PAPI) software support for ever-growing AI chips and accelerators including **Intel Habana Gaudi** on the MINCER project with **Dr. Heike Jagode** as the PI

National Renewable Energy Laboratory (NREL) | Graduate Intern | Colorado, USA | May 2024 – Aug 2024
- Demonstrated performance results (speedup of ~2.05X and 80.75% more energy efficiency) for AI applications as an intern for the "Low Precision and Efficient Programming Languages for Sustainable AI" role, under the mentorship of **Dr. Weslley Da Silva Pereira**, published two papers

Innovative Computing Laboratory (ICL) | Graduate Research Assistant | Tennessee, USA | Jan 2023 – Jan 2025
- Developed SABATH, a Benchmarking Infrastructure (5 models with GPU support and ~7.1 TB of datasets) for evaluating scientific surrogate AI applications, under the supervision of **Dr. Piotr Luszczek** and **Dr. Mark Gates,** published a paper

TENNLab - Neuromorphic | Graduate Research Assistant | Tennessee, USA | May 2023 – Dec 2024
- Developed SpikeRL, a scalable and energy-efficient framework (≈22% better and ≈39% more efficient), under the supervision of **Dr. Catherine Schuman**, published, and presented a paper

Chowkosh Limited | Research Assistant | Dhaka, Bangladesh | Sep 2021 – Mar 2022
- Developed a Deep Learning based Android Malware Detection model for zero day detection

Cye Retail Tech Ltd | Full Stack JavaScript Developer Intern | Dhaka, Bangladesh | May 2022 - Jul 2022
- Learned JavaScript for frontend and backend web and android applications

EDUCATION

Master of Science in Computer Science | University of Tennessee, Knoxville | CGPA: 3.66 | Jan 2023 - Dec 2024
Bachelor of Science in Computer Science | BRAC University | CGPA: 3.27 | Jan 2017 - Dec 2021

PUBLICATIONS

[1] PAPI Support for Specialized AI Architectures. SC25 (PDSW'25).
[2] SpikeRL: A Scalable and Energy-efficient Framework for Deep Spiking Reinforcement Learning. ICONS2025.
[3] Energy-Efficient Computing for Scalable and Sustainable AI. University of Tennessee 2024.
[4] Towards Scalable and Efficient Spiking Reinforcement Learning for Continuous Control Tasks. ICONS2024.
[5] Low Precision for Lower Energy Consumption. ASCR Energy Efficient Workshop 2024.
[6] Low Precision and Efficient Programming Languages for Sustainable AI: Final Report for the Summer Project of 2024. NREL 2024.
[7] Towards the FAIR Asset Tracking Across Models, Datasets, and Performance Evaluation Scenarios. HPEC2023.
[8] Character animation using reinforcement learning and imitation learning algorithms. ICIEV and icIVPR 2021.

ACADEMIC PROJECTS

LLM Chatbot | University of Tennessee, Knoxville | Jan 2024 – May 2024
- Developed a Chatbot for International Students using Google's Gemini API

Operating System Development | University of Tennessee, Knoxville | Aug 2023 – Dec 2023
- Built a fully functional Operating System from scratch using core C/C++

Jailbreak GPT Project | University of Tennessee, Knoxville | Aug 2023 – Dec 2023
- Analyzed Jailbreak-llms dataset (adversarial prompts) for Large Language Models (LLMs) to identify the unethical use of LLMs