Feiyang Zheng

TEL: +86-13660868810 Email: zhengfy2021@mail.sustech.edu.cn
GitHub: https://mr-jeffery.github.io
Address: No. 1088 Xueyuan Blvd., Nanshan District, Shenzhen, Guangdong, China
Post Code: 518055

EDUCATION BACKGROUND

Southern University of Science and Technology (SUSTech) September 2021-June 2025

- B.Sc. in Mathematics and Applied Mathematics *(expected)*
- GPA: 3.45/ 4.0

Summer School: North Carolina State University Summer 2022

- Project: Solving American Option Pricing by Newton's Method

STANDARDIZED TEST

TOEFL: **108** (Reading: 30, Listening: 30, Speaking: 23, Writing: 25) September 16, 2023

GRE: **324** (Verbal: 159, Quantitative: 165, Analytical Writing: 3.0) November 15, 2024

PROFESSIONAL SKILLS

Programming Languages: Proficient in Python, C/C++, Java, Shell, MATLAB
Frameworks/Libraries: PyTorch, DeepSpeed, MONAI, OpenMP, SIMD, CUDA
Profiling Tools: ITAC, nvProf, Torch Profiler, hipprof

COMPETITIONS & RESEARCH EXPERIENCE

5th Priority Research Application Competition July 2024-November 2024

- Optimized the inference throughput of a LLaMA-3 8B language model by restructuring the prefilling stage. This included reordering prompts based on input length to maximize token batch utilization, reducing the number of prefilling rounds, and enhancing performance by 25%.
- Use Torch compiler to minimize PyTorch overhead through operator fusion, improving computational efficiency.
- Reduced CPU-GPU data transfer latency by implementing pinned memory for sampling, streamlining decoding processes.
- Received 7th place and Third Prize.

ISC 2024: International Supercomputing Competition 2024 February 2024-May 2024

- Responsible for fluid dynamics simulations using Neko, with a compilation of Fortran benchmarks using mpiifort via hpcx.
- Utilized nsys for in-depth MPI performance analysis.
- Implemented and managed parallel computing operations on a single-node system equipped with eight GPUs.

11th ASC Student Supercomputer Competition January 2024-April 2024

- Led optimization of LLaMa-v2-70B large language model inference by restructuring LLaMA.cpp and enhanced token throughput by approximately 20%.

- Awarded ***Second Prize*** at ASC24.

10th ASC Student Supercomputer Competition January 2022-May 2023

- Optimized DeepMD challenge using TensorFlow's built-in Profiler and improved performance by 10-50% through adjustments in GPU thread mode and TensorFlow Graph.
- Awarded the ***First Prize*** and ***the Super Team Award*** at ASC23.

Semantic Segmentation for Autodriving September 2023-December 2023

- Developed an advanced Attention U-net deep learning model utilizing PyTorch and MONAI frameworks, specifically engineered to run on a single NVIDIA A100 GPU.
- Optimized the model performance through testing and tuning of network parameters.

6th Asia-Pacific HPC-AI Competition June 2023-November 2023

- Led the management and optimization of inference tasks for the BLOOM large language model, ensuring robust performance through strategic environment configurations using Spack.
- Implemented a two-machine, eight-GPU inference system using DeepSpeed, which significantly enhanced computational throughput and model scalability.
- Awarded the ***Second Prize*** at the 6th Asia-Pacific HPC-AI Competition.

HONORS AND AWARDS

7th place and Third Prize, 5th Priority Research Application Competition November 2024

Second Prize, 11th ASC Student Supercomputer Competition (ASC24) April 2024
First Prize, 10th ASC Student Supercomputer Competition (ASC23) May 2023
Super Team Award, 10th ASC Student Supercomputer Competition (ASC23) May 2023
Second Prize, 6th Asia-Pacific HPC-AI Competition November 2023
Runner-Up, 1st Student Supercomputing Competition, SUSTech March 2023
Third Prize, Outstanding Student Award, 2021-2022 Academic Year, SUSTech October 2022
Second Prize, First Classical Chinese Recitation Contest, SUSTech September 2022
Second Prize, Freshmen Swimming Competition, Men's 50m Breaststroke, SUSTech September 2021

EXTRACURRICULAR ACTIVITIES

Volunteer, Conducted health screenings for students at the Shenzhen Bay Port.
President, the Oak Tree Bartending Club, SUSTech 2022-Present

- Led recruitment efforts for new members and conducted cocktail-making workshops to educate new members on the history and techniques of cocktail preparation.
- Organized various events, including whiskey-tasting sessions and external collaboration activities.

Member, the Baking Club, SUSTech 2021- Present
Board Member, the Science Fiction Association, SUSTech 2021-Present
Member, the Latin Language Club, SUSTech 2021- Present

Hobbies: Swimming, Cycling, Bartending