Yifan Qin

yqin3@nd.edu | (574) 401-9049 | [Website](#) | [Google Scholar](#) | [LinkedIn](#)

Education

University of Notre Dame — **Notre Dame, IN**
Ph.D. candidate in Computer Science and Engineering — *2022–present*

Huazhong University of Science and Technology — **Wuhan, China**
Master in Software Engineering, machine learning track — *2018–2021*
Outstanding Graduates

Huazhong University of Science and Technology — **Wuhan, China**
Bachelor in Electronic Science and Technology — *2013–2017*

Experience

Research Assistant — **Notre Dame, USA**
University of Notre Dame — *2022–present*

- Developed robust and efficient training and inference methods for machine learning models, including large language models (LLMs), by introducing a negative feedback–based training strategy and a lightweight parameter-sharing mechanism to reduce inference cost while maintaining accuracy under noise and system constraints.
- Analyzed prediction uncertainty to characterize model reliability under noisy and constrained inference conditions, enabling robust and risk-aware deployment.
- Published multiple first-author papers in top-tier conferences and journals on efficient and robust machine learning.

Research Internship — **Hong Kong**
AI Chip Center for Emerging Smart Systems(ACCESS) — *2024.05–2024.07*

- Developed an efficient ML pipeline for real-time ventricular arrhythmia detection under strict latency and energy constraints.
- Applied quantization and pruning techniques to reduce inference cost while preserving detection accuracy in a deployment-oriented setting.
- Validated the approach in a realistic end-to-end system, including real-time inference and monitoring, demonstrating reliable performance for healthcare applications.

Research Assistant — **Wuhan, China**
Huazhong University of Science and Technology — *2018–2022*

- Conducted research on low-bit and quantization-aware learning methods for CNNs, improving robustness and accuracy–efficiency trade-offs under non-ideal inference conditions.
- Published two papers in journals, including one featured as a journal back-cover article.

Selected Publications

- **Yifan Qin**, Zheyu Yan, et al., "NeFT: Negative Feedback Training to Improve Robustness of Compute-In-Memory DNN Accelerators", in **TCAD'25**.
- **Yifan Qin**, Zheyu Yan, et al., "TSB: Tiny Shared Block for Efficient DNN Deployment on NVCIM Accelerators", in **ICCAD'24**.
- Zheyu Yan, **Yifan Qin**, et al., "Improving realistic worst-case performance of NVCiM DNN accelerators through training with right-censored gaussian noise", in **ICCAD'23(Best paper)**.

Award

- ICCAD'23 Best Paper Award
- ICCAD'24 Best Paper Award Candidate
- National Second Prize, China Undergraduate Mathematical Modeling Contest

Skills

Technical: Python (PyTorch, TensorFlow, JAX, NumPy, Pandas), C++; Linux, Git, LaTeX

ML & Systems: Model training and inference; robustness and uncertainty analysis; quantization and low-precision inference; inference optimization under runtime and resource constraints