# SI CHEN

Suwanee | 4703344900 | sichen.emory@gmail.com | LinkedIn

## SUMMARY

AI Systems Engineer and Computer Scientist experienced in deploying scalable AI solutions from research to production. Demonstrated success in accelerating simulation performance and improving prediction accuracy through innovative ML techniques. Skilled in experimental design and operationalizing AI pipelines, with a focus on optimizing throughput, quality, and cost-efficiency in manufacturing operations. Committed to translating complex challenges into measurable outcomes.

## SKILLS

- **Full Stack**: React, TypeScript, JavaScript, Figma, RESTful APIs, NodeJS
- **Database**: MySQL, MongoDB, Cassandra
- **DevOps & Tools**: Git, Docker, GitHub Actions, Linux
- **Machine Learning**: Python, Pytorch, Linux Shell Scripting, AWS, GCP

## EXPERIENCE

**Photon Forces Studio** **Dec 2024 - Present**

*Software Development Engineer*

- Architected and implemented a scalable full-stack educational web application leveraging React, Node.js, and MongoDB.
- Engineered RESTful APIs for customer registration, AI-assisted venue booking, and group notifications to ensure robust integrations.
- Automated interactive form-filling and scheduling with a bot-guided flow to enhance operational efficiency.

**National Center for Atmospheric Research (NCAR)** **May 2023 - Dec 2023**

*Research Intern*

- Containerized HPC simulation applications using Docker and Singularity, reducing build time by 60%.
- Established automated CI/CD pipelines with GitHub Actions to improve engineering efficiency and deployment reliability.
- Deployed containerized workloads across CPU and GPU nodes on supercomputers using Spack for dependency management.
- Collaborated with research and operations teams to architect scalable simulation workflows.

**Bytedance** **May 2021 - Aug 2021**

*Research Intern*

- Developed an NLP-powered chatbot for automated log diagnosis and root cause analysis, reducing support workload by 30%.
- Teamed with architecture groups to design infrastructure monitoring and automated error feedback tools.
- Executed RESTful API integration tests using Postman within internal deep learning platforms.

**NetApp** **May 2020 - Aug 2020**

*Research Intern*

- Modeled performance for ONTAP data management software using queuing theory, improving CPU efficiency by 20%.
- Designed modular analytics pipelines in Jupyter to detect system bottlenecks and support predictive tuning.
- Documented and implemented scalable profiling workflows for workload signature detection and reuse.

**Emory University** **Aug 2017 - Dec 2024**

*Graduate Research Assistant*

- Achieved a 127× speedup in model training by developing a cross-architecture HPC prediction system with a Meta-learning model and Gem5 simulator.
- Enhanced the SimPoint architecture simulation acceleration framework using advanced clustering, realizing a 5× speedup in simulation time while maintaining accuracy.
- Utilized feature selection analysis to identify critical hardware performance events with Perf, reducing data collection time by 95%.
- Developed a gradient-boosting classification model for storage provisioning using time-series-based feature extraction to identify concurrent I/O workloads.

## EDUCATION

**Emory University** **Dec 2024**

*Ph.D., Computer Science*

**Huazhong University of Science and Technology**

*M.S., Electrical Engineering*

**Huazhong University of Science and Technology**

*B.S., Electrical Engineering*