

Chen Lei

Incoming M.Eng. Student in AI · Rust Developer · Practical AI Builder

dev@somnus.wiki · GitHub · Blog · Nanjing, China

Profile

Incoming M.Eng. student in Artificial Intelligence at the Institute of Computing Technology, Chinese Academy of Sciences, with a B.Eng. in Intelligent Manufacturing Engineering from Southwest Jiaotong University. Interested in retrieval-augmented generation, embeddings and reranking, parameter-efficient fine-tuning, model efficiency, and Rust tooling for machine learning. Experienced in computer vision research and e-commerce data engineering.

Education

Institute of Computing Technology, CAS / UCAS 2026 — Present

M.Eng. in Artificial Intelligence · Recommended Admission

Focus: high-speed vision, multimodal information processing, and practical AI systems.

Southwest Jiaotong University 2022 — 2026

B.Eng. in Intelligent Manufacturing Engineering

Thesis: Optimization and lightweight deployment of deep-learning-based binocular stereo matching.

Experience

AI / Data Engineering Intern · E-commerce Company 2026

- Built a multi-store merchant data aggregation and operational analytics platform.
- Developed a WeChat mini program for daily sales collection and reporting.
- Designed workflows for collecting, cleaning, storing, and scheduling e-commerce operational data.
- Explored LLM-based normalization of product names, models, and SKUs.

Selected Projects

Model Weight Parser Rust · Ratatui · Safetensors

A Rust tool for parsing Safetensors model weights and inspecting tensor names, data types, and shapes through an interactive terminal interface.

RAFT-Stereo Optimization and Deployment PyTorch · Stereo Matching · FP16

Evaluated refinement, edge-loss, and partial-unfreezing strategies on Middlebury 2014 and delivered an FP16 inference path that reduced average latency from 209.80 ms to 173.59 ms while keeping accuracy effectively stable.

Superlabel Python · Data Aggregation · E-commerce Analytics

An operational analytics platform that consolidates merchant data from multiple sources and provides unified querying, metric organization, and reporting workflows.

Technical Skills

Languages	Rust, Python, TypeScript, JavaScript
AI	PyTorch, RAG, embeddings, reranking, LoRA, Safetensors

**Engineering
Environment**

Git, Jujutsu, Docker, PostgreSQL, GitHub Actions, Cloudflare
CachyOS, Linux, Zed, Bash