

Chien Liu

🌐 chienliu.com **in** linkedin.com/in/chien-liu ✉️ chienliu6001@gmail.com ☎️ +49 1573 838 6049
📍 Berlin, Germany

PROFESSIONAL SUMMARY

MLOps Engineer with 3+ years of experience, developing and optimizing scalable ML pipelines. I genuinely enjoy the process of investigating and solving complex problems for my team. While at Aignostics, I led a pipeline migration that cut model training time by **50%**. At NIO, I saved my team over **10 hours** of manual QA effort per week by designing a full-stack leaderboard app to automate evaluations for speech-to-text models. My core expertise includes Python, PyTorch, and TensorFlow, along with a deep focus on cloud platforms, containerization & orchestration, ML lifecycle management, and CI/CD best practices.

WORK EXPERIENCE

MLOps Engineer | Aignostics, Berlin Jul 2024 – Feb 2026

- Reduced model training time and computational cost by 50% via cloud infrastructure optimization and system design improvements.
- Built internal tooling for cost-efficient GPU access on GCP, giving data scientists self-service infrastructure for model training and saving 10+ hours/day in team development time.
- Accelerated CI/CD workflows by reducing Docker image build time by 90% while preserving all functionalities.
- Managed team access control and IAM roles using Terraform, enforcing infrastructure-as-code practices for secure, auditable cloud resource provisioning on GCP.

Machine Learning Researcher | National Tsing-Hua University, Taiwan Oct 2018 – Jun 2020

- Conducted fundamental research on publications at CVPR and ICML by designing model architectures and training strategies.
- Developed scalable cloud-based ML pipelines, reducing model training time from weeks to days.

Software Engineer Intern | NIO, Berlin Dec 2023 – May 2024

- Built an internal ML model performance tooling platform (Airflow + full-stack dashboard) enabling engineers to track, compare, and debug ASR model benchmarks weekly, eliminating 10 hours of manual QA per week.
- Improved speech-to-text accuracy by 35% relative to Microsoft's solution by redesigning the model training strategy.

Software Engineer Intern | Teraki, Berlin Apr 2021 – Sep 2021

- Accelerated the radar data pipeline by 10x by rewriting the bottleneck from Python to C++.

EDUCATION

M.Sc., Computational Science and Engineering(GPA 1.7) Sep 2020 – July 2023

University of Rostock, Germany

- Designed a data-driven anomaly detection system for PostgreSQL on Kubernetes, improving data quality monitoring and event detection accuracy, completed as Master's thesis with **IBM**.

SKILLS

MLOps

- Cloud Platform (AWS, GCP)
- Infrastructure as Code (Terraform, IAM)
- Databases (PostgreSQL, Cloudflare D1, Vectorize)
- Data Visualization (React, FastAPI, Grafana)
- CI/CD (GitHub Actions, GitLab CI/CD)
- Container Ecosystem (Kubernetes, Docker)
- ETL (Airflow, n8n)
- ML Monitoring (MLflow, Weights & Biases)
- System Monitoring (Prometheus, Grafana)
- Open Source AI Models (HuggingFace, ONNX)

Project Management

- Stakeholder Collaboration
- Cross-functional Collaboration
- Time Management
- Risk Management

Programming Languages

- Python
- C++
- JavaScript

SIDE PROJECT

- **Chien's Chatbot** — RAG-based chatbot on Cloudflare; combines a relational database (D1) for source-of-truth text with a vector database (Vectorize) for semantic search, orchestrated via serverless Workflows.
- **E-commerce Website** — A full-stack web app utilizing Next.js, Django, Redis, and PostgreSQL.
- **AI Face Generation** — A command-line tool published on PyPI for generating synthetic digital faces.