

Rex Chao

☎ 858-408-5697 | ✉ rexchao.careers@gmail.com | [in linkedin.com/in/rexchao](https://www.linkedin.com/in/rexchao) | github.com/seraph27 | seraph.tw

EDUCATION

University of California San Diego

Sep 2021 – Dec 2025 (Expected)

B.S. in Computer Science and Mathematics | Major GPA: 3.9/4.0

La Jolla, CA

- **Coursework:** Design & Analysis of Algorithms · Recommender Systems & Web Mining · Machine Learning Algorithms · Advanced Data Structures · Data Science in Practice · Intro to Stochastic Processes · Intro to Numerical Analysis

EXPERIENCE

Lunabill — Y Combinator F25

Sep 2025 – Present

Founding Software Engineer

San Francisco, CA

- Built real-time observability dashboard processing **5,000+ calls/wk** across **\$25M+** in insurance claims, automating call status tracking and re-dial workflows that eliminated **200+ hrs/mo** of manual rescheduling and triage work.
- Designed and deployed LLM classification pipeline integrated as core infrastructure processing every inbound call transcript, analyzing **55k+** transcripts to automate human audits, cutting operational costs by **\$24k/yr**.
- Consolidated 3 separate applications into unified codebase with shared component library and single CI/CD pipeline, reducing code duplication by 20%.

Foxconn iAi

Jul 2025 – Sep 2025

Machine Learning Engineer Intern

Milwaukee, WI

- Optimized real-time factory voice-agent pipeline (Whisper ASR + LLaMA + RAG) enabling technicians to query maintenance history and equipment logs via natural language instead of manual searches.
- Cut ASR cold-start latency by **15%** via multi-threaded model loading and warm-ups and reduced end-to-end latency by **25%** with streaming tweaks (sliding windows, adaptive chunking, VAD) validated by A/B tests.
- Deployed demo in factory environment for internal testing with **12 technicians**, achieving **77%** query success rate and collecting feedback that validated core use cases for production consideration.

Formosa Dental Clinic

May 2025 – Jul 2025

Software Engineer

Taipei, Taiwan

- Engineered a full-stack PWA for operations—scheduling, billing, patient records, and inventory with low-stock alerts.
- Migrated **36k+** patient records, and built a system that has processed **\$1M+ in transactions**.
- Built a secure data pipeline with automated balance reconciliation and audit trails to ensure integrity and reliability.
- Implemented automated backups (Vercel Cron + Dropbox API): daily snapshots, 30-day retention, and disaster recovery.

Systex Software

Jul 2024 – Sep 2024

Software Developer Intern

Taipei, Taiwan

- Automated a CT scan preprocessing workflow for a research project on liver tumor detection via computer vision (DICOM ingestion → inference → annotation → export), saving 100+ hrs/week of manual work.

PROJECTS

Tsukiyomi — AI Japanese Light-Novel Recommender | Python, PyTorch, Hugging Face, SentencePiece, Milvus

- Learns a personal known-word profile from quick word marking to recommend chapters matched to a learner's level.
- Implemented dual-tower retrieval (text embeddings + user-vocab embeddings) with Milvus ANN; re-ranks via unknown-word rate, LM predictability, and genre similarity.
- Indexed over 60k+ chapters and returns top-10 prediction in **200ms**.

LEADERSHIP & ACHIEVEMENTS

Officer, Competitive Programming Club @ UCSD (Jane Street Sponsored)

Feb 2024 – Present

- Organized and presented advanced algorithm workshops to over 30+ members; revamped club website.
- ICPC SoCal Regionals 2024: #17/88.

Competitive Programming Achievements

Present

- LeetCode: 2303 rating (Top 0.5% of 17M+ users); Codeforces: 1804 rating (Top 7% globally).

TECHNICAL SKILLS

Stack: C++, Python, TypeScript, SQL; PyTorch, React/Next.js; PostgreSQL; Docker, GitHub Actions, Vercel, GCP, Linux.