

Rex Chao

Backend/ML Systems Engineer | 858-408-5697 | rexfchao27@gmail.com | linkedin.com/in/renxchao

EDUCATION

University of California San Diego

Sep 2021 – Jan 2026 (Expected)

B.S. in Computer Science and Mathematics

Major GPA 3.9/4.0

Relevant Coursework: Advanced Algorithms, Machine Learning, Recommender Systems, Computer Graphics, Databases

Technical Skills: C++, Python, Next.js, PostgreSQL, Docker, GCP, PyTorch, Whisper, LLaMA, RAG

WORK EXPERIENCE

Lunabill - Y Combinator F25 (< 1% acceptance from 20,000+ startups)

Oct 2025 - Present

Software Engineer (Intern - Early Team)

Hybrid

- **Automated call-record auditing** with a 3-way classifier (27k+ insurance calls, 80% accuracy), cutting manual audit spend by **\$1k/mo**.
- **Second engineer hire** at a **Y Combinator F25** startup, a premier U.S. accelerator known for backing companies such as Airbnb, Stripe, and Dropbox.

Formosa Dental Clinic

Aug 2025 - Present

Software Engineer (Contract) | TypeScript, Next.js, PostgreSQL, Vercel

Remote

- Developed and deployed a full **progressive web app (PWA)** for clinic management, digitizing scheduling, billing, patient records, medical inventory (with low stock alerts) previously handled on paper.
- Integrated **36k+ records** and processed **\$200k+ in patient transactions** at a top 5 Taiwan Invisalign clinic.
- Built a secure data pipeline with automated balance reconciliation and audit trails, **reducing administrative workload by 60 min daily** and minimizing revenue-tracking errors.
- Architected **automated backup** infrastructure with **Vercel Cron + Dropbox API**, implementing daily database snapshots, 30-day retention policies, and automated disaster recovery system.

Foxconn iAi

Jun 2025 - Aug 2025

Machine Learning Engineer Intern | Python, PyTorch, Whisper, LLaMA, RAG

Hybrid

- Optimized a real-time factory **voice agent pipeline** using automatic speech recognition (ASR), large language models (LLMs), and text-to-speech (TTS), enabling technicians to query maintenance logs with natural voice commands.
- **Cut ASR cold-start latency by 15%** through multi-threaded model loading and warm-up queries.
- **Cut end-to-end latency by 25%** with streaming optimizations such as sliding-window buffering, adaptive chunk sizing, and voice activity detection (VAD), validated through **A/B testing**.
- Deployed an interactive demo app tailored to factory workflows, earning **90% positive feedback** from internal users.

System Software

Jul 2024 - Sep 2024

Software Developer Intern | Python

Taipei, Taiwan

- Automated and implemented a robust CT scan preprocessing workflow (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

- Built a recommender that suggests Japanese light-novel chapters matched to a learner's level by learning a personal **known-word** profile from quick word marking.
- Implemented **dual-tower** retrieval with transformer **text embeddings** and **user-vocab embeddings**, ranking candidates via **Milvus ANN**.
- Scored difficulty with **unknown-word rate** and **language-model predictability score**, re-ranked by **genre similarity** from prior picks, indexed **60k+** chapters and returns top-10 in **200ms**.

EXTRACURRICULARS

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

Feb 2024 - Present

- Organized **events** and delivered **advanced algorithm workshops** for 30+ members.
- 2024 **ICPC SoCal Regionals** - ranked **#17/88** among university teams in the **world's largest collegiate programming contest** (50k+ participants globally).

Competitive Programming Achievements

Present

- **LeetCode: 2303 rating** (**Top 0.5%** of 17M+ users, largest coding interview platform)
- **Codeforces: 1804 rating** (**Top 7%** globally, leading global competitive programming platform (1M+ users)).