

Annie Cai

✉ ajcai@uwaterloo.ca | 🏠 anniecai.com | 🐙 github.com/aanxniece | 🌐 in/anniecai2004

Education

University of Waterloo

B.ASc in Systems Design Engineering - GPA: 4.00/4.00, 95%

Waterloo, ON

2022 - 2026

- *Awards:* 3x Term Dean's List, Faculty of Engineering Scholarship, President's Distinction
- *Relevant Courses:* Data Structures, Algorithms, Computer Networks, Digital Systems, AI Algorithms

Skills

Languages: Python, TypeScript, JavaScript, C++, Java, SQL, GraphQL, HTML, CSS

Technologies: React, Flask, Node.js, Next.js, Express.js, Vue.js, Django

Tools: Git, PostgreSQL, AWS, Kubernetes, Docker, Redis, Pinecone, LangChain, Snowflake, CircleCI

Work Experience

Newfront (YC W18) 🔗

Software Engineer Intern

San Francisco, CA

May. 2024 - Aug. 2024

- Refactored synchronous I/O-bound tasks within **FastAPI** endpoints to asynchronous operations, enabling concurrent task execution and **reducing latency and server response times by 3x**
- Improved user interaction speeds by developing a real-time **Websockets** messaging system, with user authentication to support instant bidirectional communication for **10000+ users**
- Proposed and led text chunking and hybrid search experimentations with **Python**, **Pinecone** and **LangChain**, resulting in **11% reduction** in query response times and improved retrieval accuracy
- Conducted A/B testing of various combinations of vector embeddings and LLM models. Identified configurations that **reduced computational load by 28%** while maintaining output quality

Royal Bank of Canada 🔗

Software Engineer Intern

Toronto, ON

Sep. 2023 - Dec. 2023

- Architected an ETL pipeline in **Python** and **Django** to extract and serialize **100,000+ data points**
- Achieved a **40% cost reduction** in storage expenses by **migrating 8GB+** of data to **AWS S3 buckets**
- Scaled database retrieval job during high load periods using **Redis** and **Python**, leveraging resource tuning and caching strategies to **lower job time from 5 minutes to a few seconds**
- Built full-stack aggregated metric widgets and graphs using **Vue.js**, **Typescript**, and **RESTful APIs**, facilitating quicker, data-driven decision-making

Home Trust Company 🔗

Software Developer Intern

Toronto, ON

Jan. 2023 - Apr. 2023

- Transferred **over a million** rows of client banking data from **Azure Blob Storage** to **Azure SQL** by architecting an end-to-end ETL pipeline, utilized parallel processing to achieve **63% faster data migration** while ensuring data consistency
- Re-built CI pipeline by dockerizing services and optimizing execution with YAML scripting

Projects

UW Blueprint 🔗

Leading a cross-functional team of 9 designers and developers to build a pet volunteering platform for a non-profit, serving 2000+ users. Previously developed a recruitment tool using React, GraphQL, and Express.js. Automated a 2 hour manual process by developing a recruiter-candidate matching algorithm.

Spotifeel 🔗

Dynamic Spotify playlist based on audio emotion recognition. Trained scikit-learn neural network using RAVDESS data set and leveraged data augmentation and feature engineering to achieve 94.77% accuracy in recognizing sentiment.