Dylan Dai

dylanydai.github.io dylan.dai@uwaterloo.ca linkedin.com/in/dylanydai github.com/dylanydai

EDUCATION

University of Waterloo *Expected May 2027*
Bachelor of Computer Science (Honours) Waterloo, Ontario
Relevant coursework: Functional Programming, Compilers, Calculus, Linear Algebra, Computer Organization & Design, Algorithms

Canadian Computing Olympiad | Bronze Medalist
- Placed **14th out of 10,000+ participants** in national computing competition with **less than a year** of programming experience
- Solved algorithmic problems using tools such as segment trees, single source shortest path, max flow, and disjoint set union

SKILLS

Languages: Python, C++, C, SQL, Bash
Frameworks & Tools: AsyncIO, Pandas, Weights & Biases, PyTorch, NumPy, Git, PostgreSQL, MongoDB, GCP, Cursor
Interests: Fashion, Digital art, Tetris, Puzzle games, Rubik's cube (11s average), Clarinet (National band festival winner)

WORK EXPERIENCE

Stealth *September 2025 – December 2025*
Software Engineer Intern San Fransisco, CA
- Seed stage startup with sufficient seed funding backed by general catalyst and sequoia to speed up pharmaceutical market research
- Collaborating with researchers to build data pipelines to parse medical surveying data for synthetic data creation and backtesting
- Building synthetic data creation pipelines by training machine learning models and backtesting via existing data

Cohere *May 2025 – August 2025*
Software Engineer Intern San Fransisco, CA
- Refactored and optimized all company-wide large language model inference queries by implementing **batched query processing**, cutting query runtime by **35%** and **saving $100,000+ per month** in external large language model query costs
- Built a tool to display all of Cohere's large language model benchmarks, and runtime statistics, reducing manual effort by **30%**
- Implemented cost tracking and aggregation for **all** company-wide large language model queries to provide cost transparency and business metrics for managers, saving the company **20%** in operational costs and **40% time** in manual work
- Improved database efficiency for storing AI model queries by implementing decision trees, reducing server load by **90%**

Cohere *September 2024 – December 2024*
Data Engineer Intern San Fransisco, CA
- Oversaw and managed large-scale code datasets used to train a state of the art machine learning model **Command-A**
- Built a web-scraper to parse competitive programming websites to extract over **1,000** questions for an LLM training dataset
- Designed and solved advanced data structure and algorithm problems to train and evaluate Cohere's LLM models
- Optimized and reviewed over **700** coding test entries for evaluating the quality and accuracy of LLM-generated code

PROJECTS

AI Dataset Undersampling tool | Diversity Equity Inclusion Winner
- Won best **DEI AI Hack** from **620+ participants** by building a tool to diversify and analyze AI model training datasets
- Built a web-based data processing platform to identify and filter similar datapoints using a transformer and **k-means clustering** to segment and diversify datasets by **30%**, as measured by the average pairwise cosine similarity between all vectorized datapoints
- Added **bias identification** by querying AI Models to pinpoint outstanding attributes in results after data segmentation
- Implemented data visualization by implementing **vector compression** to render segmented data points

Exercise Assistant | DataBricks Winner
- Won "Best use of DataBricks" from **340+ participants** by developing a web-based physiotherapy game with gyroscope input
- Stores and recreates exercises by translating movement into position-storing vectors from fetching controller gyroscope data
- Incorporated live feedback via a AI model voice assistant to enhance user engagement and retention using LangChain
- Automated deployment to GCP with Terraform, supporting secure and persistent storage of user health data and exercises

Music Tracking Game | MATLAB Winner
- Won **Best use of MATLAB** from **200+ participants** by developing a musical accuracy tracking game
- Compared live audio to in-game music using two vectors of amplitudes by implementing cross-correlation and lag adjustment