

Justin Du

📧 dujstn — in du-justin — ✉ justin.du@mail.utoronto.ca — 🌐 justindu.ca — ☎ 7788621655

Education

University of Toronto, St. George

HBSc. Computer Science, Co-Op Program (ASIP) — GPA: 3.94 / 4.00 (Dean's List Scholar)

2023 - 2027

Toronto, ON

Technical Skills

Languages: Python, SQL, Bash, JavaScript/TypeScript, HTML/CSS, R, C/C++, JSX/TSX

Libraries: matplotlib, sklearn, pandas, transformers, opencv, pytorch, openai, requests, numpy, asyncio, fastmcp

Frameworks & Platforms: FastAPI, Linux, SageMaker, Azure, Google Cloud, React, Angular, Flask, Expo, PyTest

Development Tools: Docker, JIRA, Confluence, OpenShift, Stonebranch, GitHub, PostgreSQL, Hugging Face

Experience

Royal Bank of Canada (RBC)

Data Scientist — Co-Op

May 2025 - present

Toronto, ON

- Built and deployed **Whisper**-powered speech input server, load balancing **10,000+** monthly users globally.
- Finetuned **Qwen3-0.6B** for search autocomplete service using ingestion pipeline processing **225,000+** files.
- Redesigned **RAG** model risk analysis agent to support concurrency with **async FastAPI** response streaming.
- Created **scale-friendly** architecture for data storage by replacing local caches with **vector stores** via **ada-002**.
- Deployed **ETL** pipelines for employee data in **17 countries** with errorless refreshes of **1M+ PostgreSQL** rows.
- Assisted in algorithm improvements in models for business handling **~80,000,000** transactions month-to-month.
- Leveraged **CI/CD** pipelines to deploy via Github Actions, **OpenShift**, and VMs with **Stonebranch** automation.
- Designed **MCP** tooling to auto-triage and resolve tickets within service inboxes for North American departments.

U of T Machine Intelligence Student Team — 🌐 *Website*

Sep 2024 – present

Director, Industry Collaborations

Jun 2025 – present

- Project management with **Lawbrokr Inc.** to develop real-time marketing analysis & recommendation agents.
- Project management with **BuildingAssets AI Inc.** to develop automated improving energy auditing efficiency.
- Coordinated **Agile**-like workflows to manage cross-collaborating subteams and deliver market-ready product.
- Organized & conducted recruitment to fill **75+** development roles, receiving **300+** applicants in under 2 weeks.

Machine Learning Developer, Knockri Inc.— 🌐 *Website*

Jun 2025 – present

- Improved training set labeling efficiency by **over 500%** with LLM-as-a-Judge and **prompt engineering** flow.
- **Eliminated data collection costs** by synthesizing training data with finetuned LLM for transcript generation.

Projects

Stack: Course Multimedia Indexer — 🌐 *Github*

May 2024 - Mar 2025

- Outperformed SOTA models by **24%** in mIoU on *SPaSe* benchmark with custom finetuned Mask R-CNN model.
- Designed **React Native** app with Expo for search and automatic compilation and indexing of course multimedia.

TapIt: NFC-Enabled Social Media Sharing (Hack the North 2024) — 🌐 *Github*

Sep 2024

- Designed tooling for social media networking & sharing with **Ndef**-enabled NFC tags via react-native-nfc-manager.
- Produced **Expo + React Native** mobile app in under 1 day, enabling quick profile transfer from phones to tags.

Intersect: LLM Notes Summarizer — 🌐 *Website* — 🌐 *Github*

Sep 2022 - May 2023

- Finetuned **BART** model for study notes summarization sourced from aggregated work taken by students.
- Achieved **>10%** improvement in ROUGE-2 against **Pegasus** in bullet-point generation (prior to GPT-3.5 (ChatGPT)).
- Received over **80%** satisfaction rate for **React + HF Inference** app among test users from high school cohort.

Awards

Youth Science Canada

2021, 2023

2-time bronze medalist at the **Canada-Wide Science Fair** and top **0.1%** of projects in the country.