Justin Du

 dujstn — du-justin — justin.du@mail.utoronto.ca — justindu.ca — +1(778)862-1655

Education

University of Toronto, St. George — 2023 - 2027
HBSc. Computer Science, Co-Op Program (ASIP) — GPA: 3.94 / 4.00 (Dean's List Scholar) — *Toronto, Ontario*

Experience

Machine Learning Developer — Sep. 2024 - present
Knockri — *Toronto, ON*

- **Improved dataset labeling efficiency** by **over 500%** by designing a prompt engineering workflow with Transformer models via Azure AI Foundry, reducing costs for data preprocessing in model training.
- **Eliminated data collection costs** by synthesizing training data; tuned LLM hyperparamters to achieve text generation replicating the characteristics of existing data, expanding available assets to train classification models.

Projects

Stack: Course Multimedia Indexer — *Github* — *Figma* — May 2024 - present

- Outperformed SOTA models by **24%** for the *SPaSe* benchmark dataset with a new Mask R-CNN model designed to segment presentation slide elements (e.g. title, links, text, diagrams).
- Desiging **React Native** app with Expo that automatically compiles and indexes uploaded course multimedia, providing students with centralized and efficient topic lookup when studying.

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

- Implemented the `react-native-nfc-manager` library to write social media links to any `Ndef`-enabled NFC tag, as well as handle data read by the device.
- Brought an **Expo - React Native** app to production in under 32 hours, enabling users to to load mutiple custom profiles and write to an available NFC tag.

Interrsect: LLM Notes Summarizer — *Website* — *Github* — Sep 2022 - May 2023

- Developed BART model based on **bart-large-cnn**, finetuned on **arxiv-summarization**, to generate improved study notes using aggregated work by students.
- Achieved a **ROUGE-2** score of **15.13**, expressing **strong shortform text generation**.
- Designed **React.js** end user interface hosted on Google Firebase, with data managed through **Firestore Database** and model connected via Hugging Face Inferencing.

Heliios: Solar Panel Generation Forecaster — *Github* — Sep 2020 - Apr 2022

- Trained a **K-neighbors regressor** from scratch to predict, using metereological data, power generation of solar panels **anywhere in Canada**.
- Reduced the error margin of model predictions from **10%** to **<3%** through iterative **finetuning** process with **K-fold cross validation** using **GridSearchCV**.

Awards

Youth Science Canada — 2021, 2023
2-time bronze medalist at the **Canada-Wide Science Fair**, placing in the top **0.1%** of projects in the country.

Technical Skills

Lanugages: Python, Java, JavaScript, Assembly, TypeScript, R, C++, HTML/CSS
Frameworks & Platforms: React.js, React Native, Angular, Flask, Expo, Material UI, PyTest, JUnit
Development Tools: Atlassian JIRA + Confluence, Microsoft Azure, Google Firebase, Hugging Face, Git, Github
Libraries & Misc.: matplotlib, sklearn, pandas, transformers, opencv, pytorch, Maven