

JOYI XUE

✉ joyi.xue@mail.utoronto.ca | [in linkedin.com/in/joyi-x](https://in.linkedin.com/in/joyi-x) | github.com/jjyoi

SKILLS

Languages: Python, SQL, Java, R, C++, JavaScript, HTML5/CSS3

Libraries/Frameworks: Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, TensorFlow, PyTorch, Keras, Pillow, Cohere

Developer Tools: Git, GitHub, VS Code, PyCharm, IntelliJ, JupyterHub, JupyterLab, RStudio, Grav CMS

EDUCATION

University of Toronto

Toronto, ON

Honours Bachelor of Science – Computer Science Specialist and Statistics Major (3.92 cGPA)

Sep. 2024 – Present

- Activities: Google Developer Club Marketing Lead, GenAI Genesis Hackathon Marketing Lead, GenAI Genesis Hackathon Application Review Committee Member, University College Dragon Boat Club
- Scholarships: Bayer Canada Scholarship, Howard Ferguson Scholarship, Louis Savlov (UC'37) Scholarships

EXPERIENCE

Data Analyst Intern

May 2026

Independent Electricity System Operator

Toronto, ON

- Built **Python**, **SQL**, and scripting-based tools to automate data workflows, reducing manual processing effort and improving the reliability of recurring operational analysis.
- **Cleaned, transformed, and migrated data** across multiple data stores, improving dataset structure and consistency for business decision making and testing.
- Led **data validation** and **quality checks** by identifying inconsistencies, documenting issues, and preparing well-structured datasets for analysis and implementation.

Junior Systems Administrator

May 2025 – Sep 2025

University of Toronto

Toronto, ON

- Improved departmental website engagement and usability by maintaining and enhancing web content using **HTML**, **CSS**, **JavaScript**, and **Grav CMS**.
- Improve **UI/UX** for greater engagement; ensure compliance with accessibility and university standards.
- Manage **SharePoint** content and internal communications, including newsletter distribution/design and collaborative platform updates for staff and faculty.

PROJECTS

CNNs for Species Identification – Python, Pandas, Numpy, PyTorch, TensorFlow, Keras, Matplotlib, Pillow, ResNet-18

- Achieved over **90%** species **classification accuracy** with **ResNet-18** transfer learning and a **custom CNN** using **Python**, **TensorFlow**, and **PyTorch**.
- Optimized analysis of **13,000+** images with **Pandas** and **NumPy**, producing clear insights via **Matplotlib**.
- **Top 10 finalist** in 2024 National Big Data Challenge; awarded **AI Excellence Award** for best use of AI (170 teams).

Modeling Emission Effects on Permafrost with Regression – Python, Pandas, NumPy, Matplotlib, Scikit-learn, Seaborn

- Processed and standardized multi-source climate datasets (CO_2 , CH_4 , N_2O , permafrost temperatures) using **Python (pandas, NumPy, scikit-learn)**; engineered features by weighting gases based on heat-trapping effect and applied log transformations for normalization.
- Built **linear regression models** linking greenhouse gas emissions to permafrost change, with statistically significant results and scenario based forecasts.
- **Top 10 finalist** in the 2023 National Big Data Challenge; outperformed 160+ teams.

IntroSpectacle – an AI-powered memory aid – Python, OpenCV, MediaPipe, PyAudio, Cohere, WhisperAI, MongoDB

- Achieved **93%** facial-detection accuracy using **OpenCV** and **MediaPipe**.
- Integrated synchronized audio recording with **PyAudio** to enhance functionality.
- Enabled **real-time AR** by implementing ML-based conversation summarization with **Cohere**.

Data-Driven Insights & Engagement Analysis in Oncology Support Programs – R, RStudio, tidyverse, ggplot2

- Developed an **R** pipeline to analyze **3,000+** cancer-support records using **tidyverse** and **ggplot2**.
- Applied bootstrapping and regression to identify trends, supporting data-driven decisions on outreach.
- Used GitHub for version control and collaborative development.