Joseph Yu

647-528-6378 | joseph.yu2004@gmail.com | linkedin.com/in/joseph-yskyuu/ | github.com/YoshikuYuu

Education

University of Toronto Sept. 2022 – May 2026
Honours Bachelor of Science — Computer Science, Bioinformatics, & Mathematics *Toronto, ON*

- **cGPA**: 3.96/4.0
- **Relevant Coursework**: Java & Software Design, C & Systems Programming, Data Structures, SQL & Databases, Algorithms, OS & Concurrency, Machine Learning, Multivariable Calculus, Linear Algebra, Statistics

Experience

Software Engineer May 2025 – Aug 2025
Verily (formerly Google Life Sciences) *Toronto, ON*

- Designed and launched LLM-powered tools using **Go** for **translating between natural language logic/description and FHIR data model objects/expressions**, enabling **non-engineers** to efficiently express clinical workflows in a computable format

LLM Interpretability & Alignment Researcher Aug. 2025 – Present
University of Toronto Machine Intelligence Student Team *Toronto, ON*

- Isolated vector representations of political sycophancy behavior in LLMs and demonstrated that modifying LLM activations along these vectors affect LLM sycophancy

Machine Learning Researcher Mar. 2024 – May 2025
SickKids (PGCRL) — Yuen Lab *Toronto, ON*

- Improved understanding of genetic disease predictors by designing, training, and fine-tuning **three** different **PyTorch CNNs** with CUDA integration, achieving robust performance on over eight distinct biological datasets
- Streamlined bioinformatics workflows by building a genome data pre-processing pipeline using Bash and command-line utilities
- Processed, analyzed, and visualized CNN training and evaluation data using pandas and matplotlib Python libraries for research presentations

Skills

Languages: Python, Go, C/C++, Java SQL (PostgreSQL), Bash, HTML/CSS, JavaScript
Tools/Frameworks: PyTorch, Linux, Flask, MongoDB, Google Cloud Platform, Git, Unity, sklearn, pandas/polars, numpy, matplotlib, pytest, JUnit 5, Conda/Mamba

Projects

RespiraCheck: COVID-19 Detection with CNNs on Cough Audio | *PyTorch, Pandas*

- Built ML pipeline for fine-tuning CNN models **e.g. ResNet18, EfficientNet** on cough audio spectrogram data for COVID-19 detection, attaining a **test accuracy of 79% and F1 score of 0.76**
- Employed a broad range of ML techniques to address class imbalance and avoid under/overfitting, including data augmentation, hyperparameter optimization, weight decay, and random weighted sampling

Atlas Adventures | *Java, MongoDB, Git*

- Created an educational geography quiz app using **Java**, BingMaps API, and **MongoDB**, helping users learn geography through interactive gameplay
- Wrote unit, integration, and end-to-end tests for the application using JUnit5, covering **6000+** lines of code to ensure code quality and reliability
- Collaborated effectively in a team development environment using **Git** for version control, code reviews, and merge conflict resolutions