# Jean Charle Yaacoub

j.yaacoub@mail.utoronto.ca | (343) 989-1792
www.linkedin.com/in/jean-yaacoub
github.com/jyaacoub


## INTERESTS

I have experience across a wide range of DL topics from conventional CNNs to NLP and recently I am working on a project utilizing Graph NNs and foundation models like ESM for precision oncology.

## EDUCATION

**MSc in Applied Computing** September 2022 – December 2023
University of Toronto, Department of Computer Science
Artificial Intelligence Concentration
Courses:
- CSC2559 Trustworthy ML
- CSC2552 Topics in Computational Social Science
- CSC2231 Visual and Mobile Computing Systems (Winter)
- CSC2545 Advanced Topics in ML – Causal Learning (Winter)

**Bachelor of Computing (Honors)** September 2018 – June 2022
Queen's University of Kingston, Department of Computer Science
Artificial Intelligence Specialization
4.16 GPA

## WORK EXPERIENCE

**Princess Margaret Cancer Centre – UHN** Jan 2024 - Present
(Full-time)
Research Analyst

**Princess Margaret Cancer Centre – UHN** May 2023 – December 2023
(Intern)
Co-op Master's Student
- Worked on project to improve precision oncology using GNNs (see *MutDTA* below).
- Collaborated with others in the lab to help brainstorm ideas and provide feedback during group meetings and 1-1s.

**Vancouver Prostate Centre - UBC Lab** August 2020 – December 2021
(On and off between part-time and full-time)
Undergraduate Academic Assistant
- Helped to improve the performance of the Deep Docking model designed to accelerate the drug discovery process. This included optimizing code to run up to 3x faster and improving the model's architecture to be more accurate.
- Co-led the design and development of a GUI web application that made Deep Docking more accessible.
- Helped review and write papers for submission.

## NOTABLE PROJECTS

**UofT MScAC Thesis Project - MutDTA** May – Dec 2023
*GNNs with Protein Dynamics for Enhanced Drug Targeting - github.com/jyaacoub/MutDTA*
Project looking at using GNNs for binding affinity prediction to improve precision oncology for cancer patients. Researched, designed, and iterated on models under limited resources which required creative solutions to overcome with distributed multi-node computing. Additionally, leveraged existing ESM-2 foundation protein language model to improve prediction results.

**CSC2231 – Visual and Mobile Computing Systems Project** Winter 2023
*Federated Learning Vision Transformers - github.com/jyaacoub/FL-ViT*
A project Looking at optimizing the performance of Vision Transformers under a federated learning environment. Implemented in python using flower library to perform distributed training under simulated and real-world conditions with clients containing varying levels of processing power.

**CSC2559 – Trustworthy ML Project** Fall 2022
*Cross-Domain Attacks in NLP - github.com/jyaacoub/Cross-Domain-Attacks-NLP*
As a team of two we explored the transferability of adversarial examples across problem domains in NLP to get a better understanding of the intuition behind their existence in NLP. I helped review papers and

implement solutions for testing. We found that transferability across domains was weak, signaling that adversarial examples come from high level features of language similar to CV which arise from "non-robust features".

**OpenAI Hackathon for Climate Change** Fall 2022

*Net Zero AI – github.com/jyaacoub/CSR_summarizer* Nov 11-14

Led a team in developing a tool to create meaningful summaries from lengthy Corporate Social Responsibility reports. I helped bring people together, organize and divide up tasks. I was also in charge of the backend programming; utilizing the OpenAI APIs to perform semantic search and summarization via GPT-3 text completion.

**CISC 499 - Undergraduate Project** Winter 2022

*Powerful Puzzling - github.com/QuMuLab/PowerfulPuzzling*

Designed and implemented a novel jigsaw puzzle solver that works with island pieces (group of 2 or more connected pieces). I came up with a matching algorithm that utilized color and shape characteristics to connect the segmented pieces together. I also acted as the team leader, organizing, and tracking tasks that needed to be done via GitHub projects.

**CISC 474 – Reinforcement Learning Project** Fall 2021

*"Playing Atari with Deep Reinforcement Learning" - github.com/jamesgleave/DeepRL-ATARI*

Replicated the results from the 2013 DeepMind paper on Atari games. Oversaw programming the Deep Q-Learning model architectures to be tested using TensorFlow. Also helped design the test environment and utilize transfer learning to speed up training.

**Mayors Innovation Challenge/QHacks** Winter 2020

*Cycle AI - devpost.com/software/cycle-ai* *Feb 1-31*

As a team of 4 we developed Cycle AI, an app that uses ML to classify certain articles of trash/recyclables to incentivize awareness of what a user throws away. Here I oversaw programing the frontend and connecting it to the Computer Vision model in the backend. With this we were the winning team overall at QHacks (hackathon hosted by Queen's University) and got the opportunity to compete in a pitch competition against Kingston's brightest students in the Mayor's Innovation Challenge.

## TECHNICAL SKILLS

**Programming Languages**: Python, JavaScript, Java, C, MATLAB, and Prolog.
**Machine Learning**: PyTorch, pytorch-lightning, torch_geometric, Scikit-learn, and Matplotlib.
**HPC and Distributed Learning**: SLURM, Ray[Tune,Train], Flower
**Web development**: Flask, SQL

## PUBLICATIONS

Yaacoub J.C., Gleave J., Gentile F., et al. "DD-GUI: a graphical user interface for deep learning-accelerated virtual screening of large chemical libraries (Deep Docking)", *Bioinformatics*, Volume **38**, Issue 4, 15 February 2022, 1146–1148. https://doi.org/10.1093/bioinformatics/btab771

Gentile F., Yaacoub J.C., Gleave J., et al. "Artificial Intelligence-Enabled Virtual Screening of Ultra-Large Chemical Libraries with Deep Docking", *Nature Protocols*, **17**, 4 February 2022, 672-697. https://doi.org/10.1038/s41596-021-00659-2

Gentile F., Fernandez M., Ban F., et al. "Automated discovery of noncovalent inhibitors of SARS-CoV-2 main protease by consensus Deep Docking of 40 billion small molecules" *Chem. Sci.,* **12**, 17 November 2021, 15960-15974. https://doi.org/10.1039/D1SC05579H