Robert (Rupert) Wu

Updated June 17, 2024

Web www.cs.toronto.edu/~rupert
Email rupert[at]cs.toronto.edu
Twitter @rhubarbwu
Google Scholar X0M3q_sAAAAJ
GitHub/GitLab rhubarbwu/rhubarbwu
LinkedIn wu-robert

Interests

Deep learning, un/semi-supervised learning, modular neural networks, sequence modeling/LLMs, sparsity, interpretability and safety, high-performance/parallel programming, compilers.

Education

Master of Science (MSc) (Thesis), University of Toronto *Sept 2022 – Jan 2024*
Department of Computer Science – Machine Learning GPA: 3.94

- Supervised by Prof. Vardan Papyan and affiliated with the Vector Institute.
- Reviewed literature and designed architectures in sequence modeling.
- Trained dozens of LLMs on large datasets and analyzed their dynamics, inductive biases, and interpretability.

Bachelor of Science (BSc) (Honours), University of Toronto *Sept 2017 – Apr 2022*
Victoria College (2018-2022), UofT Scarborough (2017-2018) GPA: 3.75

- Specialist in Computer Science (Focus: Artificial Intelligence), Minor in Mathematics

Experience

Machine Learning Intern, Cohere *May – Aug 2022*
Modelling – Embeddings

- Analyzed neural language modelling and information retrieval literature.
- Finetuned LLMs by blending weighted datasets towards multilingual abilities.
- Implemented and executed new tasks in a model evaluation framework.

Undergraduate Researcher, University of Toronto *Jan 2021 – May 2022*
Departments of Computer Science and Mathematics

- Surveyed learning theory techniques applied to neural architecture search (NAS), under Prof. Vardan Papyan (MAT496, Winter 2022).
- Reviewed continual learning methods for image classification to mitigate catastrophic forgetting, under Prof. Florian Shkurti in collaboration with LG AI Research (CSC495, Autumn 2021).
- Developed an OpenAI CLIP-based multimodal application for images/caption similarity search, under Amlan Kar and Prof. Sanja Fidler with compute resources provided by the Vector Institute (CSC494, Winter 2021).

Software Development Engineer (SDE) Intern, Amazon Web Services (AWS) *June – Aug 2020*
AWS Serverless – Simple Queue Service (SQS)

Developer Intern, Interac *May – Aug 2019*
Product & Technology

Software Developer Co-op, Hootsuite *Jan – Apr 2019*
Hootsuite Impact

Publications & Pre-Prints

1. Rishit Dagli, Shivesh Prakash, Robert Wu, & Houman Khosravani (2024). SEE-2-SOUND: Zero-Shot Spatial Environment-to-Spatial Sound. *Under Review.* https://arxiv.org/abs/2406.06612
2. Robert Wu & Vardan Papyan (2024). Linguistic Collapse: Neural Collapse in (Large) Language Models. *Under Review.* https://arxiv.org/abs/2405.17767
3. Nayan Saxena, Robert Wu, & Rohan Jain (2022). Towards One Shot Search Space Poisoning in Neural Architecture Search. *(Student Abstract) Proceedings of the 36th AAAI Conference on Artificial Intelligence.* https://ojs.aaai.org/index.php/AAAI/article/view/21658
4. Robert Wu, Nayan Saxena & Rohan Jain (2022). NeuralArTS: Structuring Neural Architecture Search with Type Theory. *(Student Abstract) Proceedings of the 36th AAAI Conference on Artificial Intelligence* **(Top 20 Finalist, Oral)**. https://ojs.aaai.org/index.php/AAAI/article/view/21679
5. Robert Wu*, Nayan Saxena* & Rohan Jain* (2021). Poisoning the Search Space in Neural Architecture Search. *Workshop on Adversarial Machine Learning, 38th International Conference on Machine Learning, 2021.* https://openreview.net/forum?id=fB3z4GrHCYv

* equal contribution

Professional Service

- Reviewer, Trustworthy Multi-modal Foundation Models and AI Agents (ICML 2024 Workshop) *June 2024*
- Reviewer, Socially Responsible Language Modelling Research (NeurIPS 2023 Workshop) *Oct 2023*

Awards

University of Toronto (UofT)

- Queen Elizabeth II Graduate Scholarship in Science & Technology *May – Dec 2023*
- Dean's List Scholar, Faculty of Arts & Science/UofT Scarborough *Winters 2022, 2021, Summers 2019, 2018*
- Entrance Scholarship ($2000 CAD), UofT Scarborough *Autumn 2017*

Teaching

As a Course Instructor

Course	Title	Co-Instructor(s)	Term
CSC413H1	Neural Networks & Deep Learning	Amir-Massoud Farahmand*, Amanjit Singh Kainth	Winter 2024
CSC207H1*	Software Design		Summer 2023
CSC413H5	Neural Networks & Deep Learning	Florian Shkurti*	Winter 2023
CSC209H5	Software Tools & Systems Programming	Arnold Rosenbloom*, Bahar Aameri	Winter 2023
CSC207H5	Software Design	Sonya Allin*	Autumn 2022

* coordinator

As a Teaching Assistant

Course	Title	Duties	Instructor	Term
CSC2516H1	Neural Networks & Deep Learning	(γ) (η)	Colin Raffel	Autumn 2023
CSC412/2506H1	Probabilistic Learning & Reasoning	(γ) (η)	Murat Erdogdu, David Duvenaud	Winter 2022
CSC413H5	Neural Networks & Deep Learning	(τ) (γ) (η)	Lisa Zhang, Florian Shkurti	Winter 2022
CSC311H5	Introduction to Machine Learning	(τ) (γ) (η)	Anthony Bonner, Lisa Zhang	Autumn 2021
CSC165H1	Mathematics for Computer Science	(λ) (γ)	Francois Pitt, Kirill Serkh	Winter 2020
CSC165H1	Mathematics for Computer Science	(λ) (γ)	David Liu, Jonathan Calver	Winter 2019
CSC/MATA67H3	Discrete Mathematics	(τ) (γ) (η)	Anna Bretscher	Autumn 2019
CSCA08H3	Introduction to Computer Science I	(η) (γ)	Anya Tafliovich	Autumn 2018

(γ) grading (λ) lecture assistance (η) office/lab hours (τ) tutorials

Course Papers/Projects

1. **Custom Expressivity without the Degeneracy** `[CS]/zero-pruning.pdf`
   Robert Wu *(MAT1510)*
2. **Analysis of Heuristics for Neural Architecture Search** `[CS]/nas-heuristics.pdf`
   Robert Wu *(MAT496)* advised by Vardan Papyan
3. **QENAS: Q-Learning for Efficient Neural Architecture Search** `[CS]/q-learning-enas.pdf`
   Robert Wu, Rohan Jain *(CSC498/415)* `[GH]/ENAS-Experiments/tree/qenas`
4. **Multimodal (OpenAI) CLIP Applications (Prototype)** `[GH]/Multimodal-CLIP-Applications`
   Robert Wu *(CSC494)* advised by Amlan Kar/Sanja Fidler `[GH]/CLIP-FAISS-NNs`
5. **Bayesian Filters State Estimation on Directed Graphs** `[CS]/bayesian-filters.pdf`
   Robert Wu, Roland Gao *(CSC412/2506)* `[GH]/Reasonable-Subway-Surfing`
6. **Image Captioning: Nearest-Neighbors versus CNN-LSTM Approaches** `[CS]/image-captioning.pdf`
   Shayan Khalili-Moghaddam*, Jiyu Nam*, Robert Wu* *(CSC413/2516)* `[GH]/Image-Captioning`
7. **Super-Resolution using Deep Learning** `[CS]/super-resolution.pdf`
   Michal Fishkin*, Siddhant Jain*, Robert Wu* *(CSC420)* `[GH]/Super-Resolution-DL`

* equal contribution

`[CS]=www.cs.toronto.edu/~rupert/projects` `[GH]=github.com/rhubarbwu`

Skills & Expertise

- **Knowledge**: deep learning, language modelling, systems/parallel programming (GPUs/TPUs), software engineering.
- **Technology**: Python (PyTorch, JAX, scikit-learn), C/C++ (CUDA, LLVM), Java, Julia, Go, Rust, UNIX/Shell, Slurm, SQL

Affiliations

Vector Institute (Apr 2023–, Jan–May 2021), Cohere For AI (community) (June 2022–), ML Collective (community) (Sept 2021–), UTMIST (Aug 2019–)

Student Leadership

President, Computer Science Graduate [Students' Benevolent] Society (CSG[SB]S), UofT *Sept 2022 - Oct 2023*

UofT Machine Intelligence Student Team (UTMIST)

- Director, Automated Machine Learning Group (AutoMLG) *Sept 2021 – Mar 2022*
  - Founded a research group within UTMIST that explored AutoML at the intersection of combinatorial optimisation, type theory and neural architecture search (NAS).
  - Facilitated open collaboration and built a culture of open, cross-institutional research collaboration among researchers of diverse and non-traditional backgrounds.
  - Advised by Alex Adam and Chuan-Yung Tsai (Vector Institute), and the ML Collective community.
- Co-President, with Yixuan (Richard) Xu *July 2020 – July 2021*
- Vice-President, Engineering/Infrastructure *Jan 2020 – Jan 2021*
- Web Developer *Aug 2019 – Dec 2019*

Computer Science Student Union (CSSU), UofT

- President (Interim) *Apr – May 2022, Oct – Dec 2021*
- Vice-President *May 2021 – May 2022*
- First-Year Orientation: Leader (Summer 2020), Organizer (Summers 2019, 2018) *May 2018 – Sept 2020*

Student Councillor, Victoria College/University *Oct 2021 – Apr 2022*

Graduate Coursework

Code	Title	Instructor	Term
CSC2537/STA2555H	Information Visualization	Fanny Chevalier	Autumn 2023
CSC2125H	Blockchain Technology & Engineering	Fan Long	Winter 2023
CSC2231H	Visual & Mobile Computing Systems	Nandita Vijayakumar	Winter 2023
CSC2224H	Parallel Computer Architecture & Programming	Gennady Pekhimenko	Autumn 2022
MAT1510H	Deep Learning: Theory & Data Science	Vardan Papyan	Autumn 2022

Advanced Undergraduate Coursework

Code	Title	Instructor/Supervisor	Term
CSC488/2107H1	Compilers & Interpreters	Fan Long	Winter 2022
MAT496H1	Reading: Mathematics of Deep Learning	Vardan Papyan	Winter 2022
VIC493H1	Vic Capstone Research Colloquium	Emanuel Istrate	Year 2021-2022
CSC485/2501H1	Computational Linguistics	Gerald Penn	Autumn 2021
CSC495H1	Project: Continual Learning	Florian Shkurti	Autumn 2021
CSC498/475H5	Topics: Introduction to Reinforcement Learning	Animesh Garg	Autumn 2021
CSC384H1	Introduction to Artificial Intelligence	Sonya Allin	Summer 2021
CSC412/2506H1	Probabilistic Learning & Reasoning	Jesse Bettencourt	Winter 2021
CSC413/2516H1	Neural Networks & Deep Learning	Jimmy L. Ba, Bo Wang	Winter 2021
CSCD70H3	Compiler Optimizations	Gennady Pekhimenko	Autumn 2020
CSC494H1	Project: Multimodal CLIP Applications	Sanja Fidler, Amlan Kar	Winter 2021
CSCC11H3	Introduction to Machine Learning & Data Mining	David J. Fleet, Bryan Chan	Autumn 2020
CSC369H1	Operating Systems	Karen Reid	Autumn 2020
CSC420H1	Introduction to Image Understanding	Babak Taati, Morteza Rezanejad	Autumn 2020
HPS391H1	History of Mathematics from 1700 to the Present	Sylvia Nickerson	Winter 2020
CSC324H1	Principles of Programming Languages	David Liu	Autumn 2019
CSC373H1	Algorithm Design, Analysis, & Complexity	Koushik Pal	Summer 2019
CSC300H1	Computers & Society	Mathew Zaleski, Ishtiaque Ahmed	Winter 2019
CSC336H1	Numerical Methods	Kenneth R. Jackson	Autumn 2018