

# Karim Zakir

[zakirkarim02@gmail.com](mailto:zakirkarim02@gmail.com) | +1 437 971 3563 | [github.com/karimzakir02](https://github.com/karimzakir02)

---

## SKILLS

---

Programming Languages/Technologies: Python, R, C, Go, Java, React Typescript, SQL

Skills: Software Engineering, Machine Learning, Deep Learning, Data Science, Statistical Analysis, Network Analysis

Relevant Projects on GitHub (link above): Sales-Toronto (Software Engineering / Deployment), Jupyter (Machine Learning)

---

## EDUCATION

---

### University of Toronto (Honors Bachelor of Science) | GPA: 3.84/4.00 | Dean's List Scholar

- Enrolled in the Data Science Specialist Program
  - Relevant Coursework: Introduction to Machine Learning (A+), Data Science I & II (A+), Software Design (A+), Systems Programming (A+), Introduction to Databases (A), Algorithm Design & Analysis (A+), Multivariable Calculus (A)
  - Publications: Gao J, Bonzel C, Hong C, Varghese P, **Zakir K**, Grönsbell J. ssROC: Semi-Supervised ROC Analysis for Reliable and Streamlined Evaluation of Phenotyping Algorithms. [arXiv:2305.01709](https://arxiv.org/abs/2305.01709)
- 

## WORK EXPERIENCE

---

### University of Toronto | Machine Learning Engineer | 09/2023 – 12/2023

- Developed new architecture for semantic segmentation tasks by integrating methods from semi-supervised learning
- Implemented, trained, and tested several architectures, such as Bootstrap Your Own Latent, from scratch in PyTorch
- Created a new dataset by combining PascalVOC and MNIST digits to test semantic models of medium complexity/size
- Added distributed data parallelism to model training, reducing training duration by over 2x
- Conducted extensive literature review in semi-supervised learning and semantic segmentation subsections of deep learning

### Goldman Sachs | Software Engineering Analyst | 06/2023 – 08/2023

- Designed a new workflow and a new set of features for existing privacy governance service used by the whole firm (45K+)
- Communicated with multiple stakeholders from different teams to reach consensus on the most optimal design
- Implemented frontend and the backend for the workflow using React Typescript and Java respectively
- Documented the design decisions for the workflow using Wireframes and UML diagrams
- Conducted extensive end-to-end testing in addition to implementing a comprehensive set of unit tests

### University of Toronto | Machine Learning Researcher | 09/2022 – 04/2023

- Trained semantic and instance segmentation models based on Fully Convolutional and MaskRCNN architectures with PyTorch
- Applied a wide range of edge detection and other visual computing algorithms to OCT images of different objects
- Leveraged high-performance computing nodes to improve the efficiency of OCT denoising pipeline
- Built a data processing pipeline from a custom dataset to allow for most efficient training and validation pipeline

### Recursion Pharmaceuticals | Data Science Intern | 05/2022 – 09/2022

- Led the development of a graph-based tool to introduce a different interpretation of cell perturbation data
- Applied Graph Theory inference algorithms to discover clusters and other significant relationships in 128-dimensional data
- Improved existing workflows by applying the graph tool to known methods and improving their speed by 5-20x
- Validated my findings by communicating with multiple teams, including non-technical members of the company
- Presented my findings to different members of the company, including the executive team, during an internal tech conference