Spencer R. Szabados

Contact Information

Address : 1410-812 King St W
Kitchener, ON, Canada

Web : https://spencerszabados.github.io/
Email : spencerszabados@gmail.com
sszabados@uwaterloo.ca

About

Currently a Machine Learning Engineer at hum.ai, applying diffusion models to remote satellite sensing. I completed a Master's in Computer Science at the University of Waterloo with Yao-liang Yu, focusing on generative models, and an undergrad in Applied Math and Computer Science at the University of Manitoba, where I worked on data depth in the GADA Lab.

I enjoy tackling challenging problems—especially in computational math, applied stats, and numerical methods. Outside of work, I'm an avid reader, mostly on topics relating to human decision making, finance, and sci-fi.

Education and Certificates

Jan. 2023 - Jan. 2024	Master's of Mathematics in Computer Science, **University of Waterloo**, Canada. Specializing in Machine Learning. Supervised by Yaoliang Yu. Grade: 95.4%
Feb. 2022	Bachelor's of Science in Mathematics, **University of Manitoba**, Canada. Specialized in Applied Mathematics and Theoretical Computer Science. Graduated with Distinction.
June. 2020	*Certificate* in Front-End Web Development, **Udemy**, Instructed by Joseph Delgadillo and Nick Germaine. Course detailed the use of html5, css, JavaScript, and jQuery in the design and deployment of custom dynamic websites.
June. 2016	Highschool Diploma, **Institut Collégial Vincent Massey Collegiate**, Canada. Graduated with Honors.

Awards[1]

June. 2024	Directed Reading Program Bursary. University of Waterloo, *Awarded for serving as graduate mentor.*	($1000)
July. 2024	Vector Institute Research Bursary. Vector Institute, *Awarded to full-time graduate students at a Canadian educational institution in good academic standing who are supervised by a Vector Faculty Member.*	($6000)
Jan. 2023	Domestic Mathematics Award. University of Waterloo, *Awarded to full-time Mathematics graduate students in high academic standing.*	($4,000)
Jan. 2023	Graduate Scholarship. University of Waterloo, *Awarded to full-time graduate students in a Master's program with a minimum first-class (80%) cumulative average in their current program or over the last two full-time academic years.*	($4,000)
Jan. 2022	Faculty of Science Undergraduate Research Bursary. University of Manitoba, *Awarded to undergraduate students to fund their research while working under the supervision of individual professors. Awarded to me by Stephane Durocher.*	($10,000)
Oct. 2021	Philosophia Mathematica Prize in Applied Mathematics. University of Manitoba, *Awarded to best applied mathematics student in either second or third year.*	($3850)
Sept. 2016	Faculty of Science Entrance Bursary.	($1000)

[1] Values listed in cumulative amounts, opposed to repeating entries.

Publications

2024 Lu, H.*, **Szabados, S.***, and Yu, Y. (2024). Diffusion Models with Group Equivariance. ICML2024 SPIGM Workshop. openreview.net
This paper focuses on structure-preserving diffusion models (SPDM), a subset of diffusion processes. We complement existing sufficient conditions for constructing SPDM by proving complementary necessary ones. Within this framework, we propose a method of preserving the alignment between endpoint couplings in bridge models to design a novel structure-preserving bridge model.

2024 (Preprint.) Lu, H.*, **Szabados, S.***, and Yu, Y. (2024). Structure Preserving Diffusion Models. arxiv.org
This paper introduces structure-preserving diffusion models (SPDM), a novel approach for learning distributions with inherent symmetries. Theoretical conditions are developed to ensure that diffusion process preserves symmetry, enabling equivariant data sampling trajectories.

2023 Lu, H., Lu, Y., Jiang, D., **Szabados, S.**, Sun, S., and Yu, Y. (2023). CM-GAN: Stabilizing GAN Training with Consistency Models. ICML2023 SPIGM Workshop. openreview.net
To overcome GAN training instability we replace the standard generator architecture with a pre-trained Consistency Diffusion Model. In this way, we provide a method to combine the main strengths of diffusion and GAN models while mitigating both their major drawbacks.

2022 Durocher, S., and **Szabados, S.*** (2022). Curve Stabbing Depth: Data Depth for Plane Curves. 34th Canadian Conference on Computational Geometry (CCCG 2022). cccg.ca
A novel depth measure for plane curves and functional data is proposed from a computational geometry perspective, focusing on the development and analysis of an exact algorithm for computing the depth of arrangements of polylines.

Employment

Industry

Jan. 2025 - Present	Machine Learning Engineer - (Contract, Full-time, Hybrid) **hum.ai** (prev. Coastal Carbon) (url), Waterloo Promoted from Graduate Intern. Designing and deploying latent diffusion models for remote sensing of multi-spectral satellite data. Work includes building highly distributed pipelines for multi node training, performing time series interpolation, implementing cross sensor fusion techniques.
Aug. 2024 - Jan. 2025	Machine Learning Graduate Intern - (Contract, Part-time, Hybrid) **Coastal Carbon** (url), Waterloo Focused on design, benchmarking, and implementation of latent diffusion models for satellite data analysis.

Academic

Apr.-Sept. 2024	Teaching Assistant CS480 - Introduction to Machine Learning, **University of Waterloo**, Waterloo. Introduction to neural networks, kernel based models, support vector machines, etc...
Jan.-Apr. 2024 Sep.-Aug. 2023	Teaching Assistant CS335 - Computational Methods in Fiance, **University of Waterloo**, Waterloo. Covering the design and analysis of algorithms for simulating financial stock pricing, the theory surrounding stochastic differential equations, Black-Scholes risk free model, etc...
Jan.-Apr. 2023	Teaching Assistant CS330 - Infrastructure Management, **University of Waterloo**, Waterloo. Design decisions for computing infrastructure and networking, how to price compare distributor options, etc...
May.-Aug. 2022 Jan.-Apr. 2022	Undergraduate Research Assistantship, **University of Manitoba**, Winnipeg. Assistantship extended for an additional term. Due to my academic standing, I received a bursary for a collaborative research project with Dr.Durocher, working on the definition and analysis of a novel data depth measure for curve and functional data.

Jan.-Apr. 2022	Teaching Assistant CS3170 - Advanced Design of Algorithms, **University of Manitoba**, Winnipeg. Covering the design and analysis of advanced algorithms and data structures; e.g., Amortized analysis, Fibonacci heaps, Randomized data structures and expected running time analysis, Skip lists, etc...

Volunteer Experience

Academic

May.-Aug. 2024	Directed Reading Program Mentor (DRP), **University of Waterloo**, Waterloo. The woman in mathematics directed reading program is an initiative that aims to provide an opportunity for undergraduate students (mentees), who identify as women or underrepresented gender identities, to work with graduate students on a mathematical project. During this time I mentored two undergraduates on machine learning methods (latent space models) and measures of statistical depth (and combinatorial data depth) for out-of-distribution diction (anomaly detection).
Jan. 2022 - Jan. 2023	Lab organizer (GADA Lab), **University of Manitoba**, Winnipeg. Organizer for the Geometry Algorithms and Distributed Algorithms open problem solving sessions. These meetings were a platform for new students and existing lab members to collaborate on open problems in the fields of Computational Geometry, Graph Algorithms, Distributed Algorithms, etc...

Skills

Proficient in modern machine learning tools and infrastructure, with a strong focus on generative modeling, distributed training, and scientific computing:

Languages :	Python *(proficient)*, C *(comfortable)*, Julia, Java.
Machine Learning Frameworks :	PyTorch *(2+ years)*, PyTorch Lightning, Torch Compile, JAX, TensorFlow. Experience includes building and training VAEs, Normalizing Flows, Latent Diffusion Models, GANs. Recently implemented a Fourier-space VAE.
Distributed Computing :	Deployed large-scale training jobs (10M–500M+ parameters) across multi-node AWS EC2 (H100) clusters using Ray Train, Slurm, and PyTorch Distributed.
Experiment Tracking :	Heavy user of Weights & Biases (WandB) for experiment management, hyperparameter sweeps, and logging.
Data Handling & Visualization :	Skilled in data exploration and quality assessment with quantile plots, high-dimensional clustering, and statistical summaries.
Web & API Tools :	FastAPI (via Ray Serve), html5, css, JavaScript, Hugo.
Technologies :	Docker *(used extensively across ML workflows)*, Git, Ray, Slurm, AWS EC2.

Data Science

Experience performing statistical data analysis and exploratory visualization in Python. Regularly use pandas, scipy, and custom tools for dataset assessment, with a focus on identifying outliers, visualizing structure, and guiding preprocessing. Familiar with clustering, dimensionality reduction, and empirical validation of sample statistics.