

EDUCATION

- PhD in Computer Science, University of Toronto** **2019 - Present**
Advisors: Roger Grosse, Geoffrey Hinton
- Bachelor of Applied Science, University of Toronto** **2014 - 2019**
Division of Engineering Science – Robotics (+ professional experience year at NVIDIA)
Academic Standing: First in 2019 Class in Engineering Science, CGPA: 3.98 (high honours)

RESEARCH EXPERIENCE

- Exploring Length Generalization in Large Language Models** **2021 - 2022**
Toronto, Google Research, Blueshift Team (large language models, reasoning, length generalization)
Advisor: Behnam Neyshabur
- Demonstrated that while finetuning large language models show significant length generalization deficiencies, combining chain-of-thought and few-shot-learning capabilities yield significant improvements.
- Minerva: A Language Model that Solves Quantitative Reasoning Problems** **2021-2022**
Toronto, Google Research, Blueshift Team (large language models, reasoning, transformers)
Advisors: Behnam Neyshabur, Guy Gur-Ari, Vedant Misra
- Contributed to a project that demonstrated that large language models trained on large amounts of technical content are capable of solving highly complex quantitative reasoning problems.
- Upwards Generalization with Path-Independent Equilibrium Models** **2021-2022**
Toronto, University of Toronto, Vector Institute (equilibrium models, test-time methods, upward generalization)
Advisors: Roger Grosse, Zico Kolter
- Identified a condition – named path independence – that allows equilibrium models to generalize to problem instances that are more difficult than those observed during training.
- Learning to Elect** **2020-2021**
Toronto, University of Toronto, Vector Institute (social choice, equivariant architectures, OOD generalization)
Advisors: Nisarg Shah
- Proposed a steerable learning framework that can discover novel voting rules based on custom loss functions.
 - Demonstrated that the learned rules remain effective on out-of-distribution elections.
- Prover-Verifier Networks** **2019 - Present**
Toronto, University of Toronto, Vector Institute (interactive proof systems, smooth games, AI Safety)
- Developed a training framework that encourages neural networks to find solutions that are easily verifiable. This is done by framing learning as a game between a powerful prover network, and a constrained verifier network.
- Object-Centric Source Separation** **2019 - 2020**
Toronto, University of Toronto, Vector Institute (source separation, capsule nets, audio processing)
Advisors: Geoffrey Hinton, Roger Grosse, Sageev Oore
- Investigating the effectiveness of object-centric inductive biases in blind source separation.
 - Demonstrated that dynamical routing capsule networks can successfully separate highly overlapping spectrograms of multiple notes and instruments. Currently scaling up to larger scale tasks and novel architectures.
- Sorting Out Lipschitz Function Approximation** **2018 - 2019**
Toronto, University of Toronto, Vector Institute (Lipschitz networks, adversarial robustness, optimal transport)
Advisor: Roger Grosse
- Developed a novel neural network architecture with a controllable Lipschitz constant using a novel GroupSort nonlinearity.
 - Proved GroupSort networks are universal Lipschitz function approximators. Empirically demonstrated that GroupSort networks improve performance on tasks including Wasserstein distance estimation and adversarial robustness.
- Preventing Gradient Attenuation in Lipschitz Convolutional Networks** **2019**
Toronto, University of Toronto, Vector Institute (Lipschitz networks, gradient preservation, provable robustness)
Advisors: Roger Grosse, Jörn Jacobsen
- Extended our earlier analysis of Lipschitz networks on convolutional networks.
 - Proposed an expressive parametrization of orthogonal convolutions, analyzed related optimization challenges. The resulting model exhibits state-of-the-art provable robustness for image classification tasks.
- TimbreTron: A WaveNet(CycleGAN(CQT))) Pipeline for Musical Timbre Transfer** **2018, 2019**
Toronto, University of Toronto, Vector Institute (GANs, style transfer, audio processing)
Advisors: Roger Grosse, Sageev Oore
- Worked on building a GAN and WaveNet-based deep learning-based pipeline that can perform musical timbre transfer.
- Metastatic Brain Tumor Segmentation** **2016**
Toronto, University of Toronto, Biophysics Department (medical image processing, segmentation, convnets)
Advisor: Anne Martel
- Built a 3D brain tumor segmentation pipeline using convolutional neural networks.
- Effect of Motor Preparation on Auditory Evoked Potentials** **2015**
Toronto, University of Toronto, Biophysics Department (auditory neuroscience, MEG imaging, signal processing)
Advisor: Bernhard Ross
- Researched the effect of preparation in sound making on brain responses with magneto-encephalogram (MEG) imaging.

INDUSTRY RESEARCH EXPERIENCE

Student Researcher at Google Blueshift Team

2021 - 2022

Toronto, Google Research, Blueshift Team

(large language models, reasoning, generalization)

Manager: Behnam Neyshabur

- Lead a project that demonstrated severe deficiencies in the length generalization capabilities of large language models and showed that making use of few-shot learning and scratchpad techniques significantly improve performance.
- Contributed to the effort that led to Minerva, a language model trained on technical material that can solve highly nontrivial quantitative reasoning problems.

Professional Experience Year at NVIDIA

2017 - 2018

Toronto, NVIDIA Corporation

(generative modelling, RL, semi-supervised learning)

Manager: Gavriel State

- Applied novel learning techniques on a diverse set of areas, including robot simulation, game design and animation.
- Proposed a semi-supervised method to error-correct time series predictions, which led to a patent application.
- Contributed to a project that focused on using randomized synthetic data to train object detection and segmentation systems that operate on complex real world data.

PUBLICATIONS AND PATENTS

[Exploring Length Generalization in Large Language Models](#)

NeurIPS 2022

C Anil, Y Wu, A Andreassen, A Lewkowycz, V Misra, V Ramasesh, A Slone, G Gur-Ari, E Dyer, B Neyshabur

[Solving Quantitative Reasoning Problems with Language Models](#)

NeurIPS 2022

A Lewkowycz, A Andreassen, D Dohan, E Dyer, H Michalewski, V Ramasesh, A Slone, C Anil, I Schlag, T Gutman-Solo, Y Wu, B Neyshabur, G Gur-Ari, V Misra

[Path Independent Equilibrium Models Can Better Exploit Test-Time Computation](#)

NeurIPS 2022

C Anil, A Pokle, K Liang, J Treutlein, Y Wu, S Bail, Z Kolter, R Grosse

[Learning to Elect](#)

NeurIPS 2021

C Anil, Xuchan Bao

Preventing Gradient Attenuation in Lipschitz Constrained Convolutional Networks

NeurIPS 2019

Q Li, S Haque, C Anil, J Lucas, RB Grosse, JH Jacobsen

Sorting out Lipschitz Function Approximation

ICML 2019

C Anil, J Lucas, RB Grosse

Timbretron: A Wavenet (cycleGAN (CQT (Audio))) Pipeline for Musical Timbre Transfer

ICLR 2018

S Huang, Q Li, C Anil, X Bao, S Oore, RB Grosse

Training Deep Networks with Synthetic Data: Bridging the Reality Gap by Domain Randomization

CVPR 2018

J Tremblay, A Prakash, D Acuna, M Brophy, V Jampani, C Anil, T Hong To, E Cameracci, S Boochoon, S Birchfield

(workshop)

Refining Labeling of Time-Associated Data

Patent 2019

C Anil (US Patent App. 16/153,430)

Generation of Synthetic Images for Training a Neural Network Model

Patent 2019

J Tremblay, A Prakash, M A Brophy, V Jampani, C Anil, S Thomas Birchfield, T Hong To, D Jesus Acuna Marrero (US Patent App. 16/256,820)

SELECT HONOURS AND AWARDS

Award – NSERC Alexander Graham Bell Canada Graduate Scholarship-Doctoral (CGS D)

2020-2023

Award- W.S. Wilson Medal (first in graduating class in Engineering Science)

2019

Award- NSERC Undergraduate Summer Research Award (USRA)

2016

Award- NSERC ACN Create Undergraduate Research Award

2015

Award- High School Valedictorian - 100% Academic Merit Scholarship

2014

Physics Olympiads – Top 25 in Turkey (2013).

2014

Philosophy Olympiads - Invited to join High School International Philosophy Olympiads

2013

COMMUNITY SERVICE

Conference Reviews: Consistently review for ICML, ICLR and NeurIPS (top 10% reviewer at NeurIPS2021 and ICML2022)

2019 – Present

EXTRACURRICULAR SKILLS

Music: Proficient piano player and composer.

- Composed original scores for two motion pictures (Siyah-Beyaz and The Smell of Money)
- Composed the original scores for the short film The Teacup (2016), which won the Best Music and Sound Design Award in the 2016 Dingle International Film Festival in the student category.

Tennis and Scuba Diving: Former licensed tennis player and scuba diver.