

Adam Ibrahim

Interests

Mathematics:

Topology, Algebra,
Measure Theory,
Number Theory

Computer Science:

Machine Learning,
Algorithms and
Optimisation, Quantum
Computing,
Cryptography

Physics:

Theoretical Physics,
Computational Physics

Programming Skills

Python, C/C++, C#,
Unity, MATLAB, R,
LaTeX

Languages

Native/bilingual

proficiency:

French, Arabic, English

Limited working

proficiency:

Spanish

Elementary

proficiency:

German (currently
learning)

Extra-curriculars

Gym, Rock climbing,
Motorcycling (track),
Scuba diving,
Piano, Guitar

Work Experience

06/24 - 09/24

Head of LLM team. Also led the infrastructure team to set up and maintain clusters. [H](#)

01/24 - 05/24

Consulting on pretraining and finetuning foundation models (mixture-of-experts, LLMs, and image generation). [Zyphra](#)

05/23 - 12/23

Mentoring/supervising confidential research project on foundation models for scene understanding as external consultant. [Bosch](#)

05/23 - 12/23

Mentoring/supervising two confidential research projects on image generation and self-supervised learning for computer vision as external consultant. [Optina Diagnostics](#)

05/23 - 11/23

Mentoring/supervising confidential research project on Large Language Models as external consultant. [Staples](#)

06/23 - 09/23

Consulting for the development of Large Language Models for code. [Blackbox AI](#)

05/22 - 12/22

Mentoring/supervising confidential research project on reinforcement learning and exploration as external consultant. [AMD](#)

02/22 - 06/22

Mentoring/supervising confidential research project on speech recognition as external consultant. [Microsoft](#)

05/21 - 12/21

Mentoring/supervising confidential research project on multimodal machine learning in hardware-constrained environments as external consultant. [Apple, Cupertino](#)

Summer 2016

Design and optimisation of computer vision and signal processing algorithms for the Analysis Studio software (C# / C++). Design of computer vision algorithms for a healthcare related NDA project. [Anasys Instruments](#)

Education

2018 - 2024

Doctor of Philosophy

Machine Learning, under the supervision of Pr. Ioannis Mitliagkas and Irina Rish. GPA: 4.30/4.30. [Mila, Université de Montréal, Canada](#)

2015 - 2018

Master of Science

Computer Science. Areas of focus: Machine Learning, Human-Computer Interaction, Computer Vision, Cryptography. GPA: 4.0/4.0. [University of California, Santa Barbara, USA](#)

Relevant graduate courses: CS 595I Advanced Machine Learning seminar, CS 290I Deep Learning, CS 292F Foundations of Data Science, MATH 260J Foundations of Machine Learning, ECE 210A Matrix Analysis, CS 240A Parallel Computing.

2012 - 2015

Bachelor of Science

Joint Honours Mathematics and Physics. First-class honours. [McGill University, Montreal, Canada](#)

Relevant graduate courses: General Relativity, Introduction to String Theory, Group Theory, Topics in Topology, Quantum Field Theory, Very Early Universe.

2011 - 2012	MPSI Preparatory school. Main subjects: Mathematics, Physics, Engineering, specialisation Computer Science.	Collège Stanislas, Paris, France
2011	Baccalauréat Scientifique Bilingual French/English Secondary School. Passed all of the Cambridge English proficiency exams, including the Certificate of Proficiency in English.	Lycée Saint-Charles, Orléans, France

Research Experience

06/24 - 09/24	Head of LLM team. Also led the infrastructure team to set up and maintain clusters.	H
09/18 - 06/24	Research in optimisation, out-of-distribution generalisation and adversarial machine learning, and continual learning. Ongoing work on Large Language Models as part of the 2023 INCITE Allocation program by the US Department of Energy. Interests in reinforcement learning.	Mila, Université de Montréal
01/24 - 05/24	Consulting on pretraining and finetuning foundation models (mixture-of-experts, LLMs, and image generation).	Zyphra
05/23 - 12/23	Mentoring/supervising confidential research project on foundation models for scene understanding as external consultant.	Bosch
05/23 - 12/23	Mentoring/supervising two confidential research projects on image generation and self-supervised learning for computer vision as external consultant.	Optina Diagnostics
05/23 - 11/23	Mentoring/supervising confidential research project on Large Language Models as external consultant.	Staples
06/23 - 09/23	Consulting for the development of Large Language Models for code.	Blackbox AI
05/22 - 12/22	Mentoring/supervising confidential research project on reinforcement learning and exploration as external consultant.	AMD
02/22 - 06/22	Mentoring/supervising confidential research project on speech recognition as external consultant.	Microsoft
05/21 - 12/21	Mentoring/supervising confidential research project on multimodal machine learning in hardware-constrained environments as external consultant.	Apple, Cupertino
03/16 - 06/18	Research in recommender systems and in particular the perception of recommendations in Augmented Reality. Language learning project in Augmented Reality using machine learning to recognise objects in the environment in order to provide situated and personalised learning. Design of computer vision algorithms based on deep learning to allow users of Augmented Reality devices to select objects in the environment. Designed and conducted user studies to test the potential of Augmented Reality as a vocabulary learning medium.	Four Eyes lab, UC Santa Barbara
09/14 - 04/15	THz photons trapped in dynamical optically-pumped cavities in silicon materials. Worked in a group to test numerically the consistency of the results with the theory using an FDTD algorithm.	David Cooke group, McGill University
05/14 - 04/15	Investigating whether string theoretical monodromy inflation can be uplifted to a de Sitter universe.	Keshav Dasgupta, McGill University

Probing for cosmic string wakes signatures in the CMB using Canny's algorithm and analytical methods. Attended the weekly research meetings of the cosmology group until 2015.

Publications

- 2024 **Zyda: A 1.3 T Dataset for Open Language Modeling** [Technical report \(arXiv\)](#)
Yury Tokpanov, Beren Millidge, Paolo Glorioso, Jonathan Pilault, Adam Ibrahim, James Whittington, Quentin Anthony
- 2024 **Why Has Predicting Downstream Capabilities of Frontier AI Models with Scale Remained Elusive?** [Under review NeurIPS 2024, presented at ICML 2024 workshops](#)
Rylan Schaeffer, Hailey Schoelkopf, Brando Miranda, Gabriel Mukobi, Varun Madan, Adam Ibrahim, Herbie Bradley, Stella Biderman, Sanmi Koyejo
- 2024 **Zamba: A Compact 7B SSM Hybrid Model** [Technical report \(arXiv\)](#)
Paolo Glorioso, Quentin Anthony, Yury Tokpanov, James Whittington, Jonathan Pilault, Adam Ibrahim, Beren Millidge.
- 2024 **Simple and Scalable Strategies to Continually Pre-train Large Language Models** [TMLR](#)
Adam Ibrahim*, Benjamin Thérien*, Kshitij Gupta*, Mats Leon Richter, Quentin Gregory Anthony, Timothée Lesort, Eugene Belilovsky, Irina Rish.
A * denotes equal contribution.
- 2023 **Continual Pre-Training of Large Language Models: How to Re-warm Your Model?** [ESFoMo ICML 2023, ENLSP NeurIPS 2023](#)
Kshitij Gupta*, Benjamin Thérien*, Adam Ibrahim*, Mats Leon Richter, Quentin Gregory Anthony, Eugene Belilovsky, Timothée Lesort, Irina Rish.
A * denotes equal contribution.
- 2022 **Towards Out-of-Distribution Adversarial Robustness** [arXiv preprint](#)
Adam Ibrahim, Charles Guille-Escuret, Ioannis Mitliagkas, Irina Rish, David Krueger, Pouya Bashivan.
- 2022 **Learning Robust Kernel Ensembles with Kernel Average Pooling** [arXiv preprint](#)
Pouya Bashivan, Adam Ibrahim, Amirozhan Dehghani, Yifei Ren
- 2022 **Gradient Descent Is Optimal Under Lower Restricted Secant Inequality And Upper Error Bound** [NeurIPS 2022](#)
Charles Guille-Escuret, Adam Ibrahim, Baptiste Goujaud, Ioannis Mitliagkas.
- 2022 **Towards Generalisable Robustness: A Domain Generalisation Approach** [ICML 2022 AdvML](#)
Adam Ibrahim, Charles Guille-Escuret, Ioannis Mitliagkas, Irina Rish, David Krueger, Pouya Bashivan.
- 2021 **Adversarial Feature Desensitization** [NeurIPS 2021](#)
Pouya Bashivan, Reza Bayat, Adam Ibrahim, Kartik Ahuja, Mojtaba Faramarzi, Touraj Laleh, Blake Richards, Irina Rish.
- 2020 **Linear Lower Bounds and Conditioning of Differentiable Games** [ICML 2020](#)
[Also presented at MAIS 2019, DeepMath 2019 and NeurIPS 2019 SGO workshop](#)
Adam Ibrahim, Waïss Azizian, Gauthier Gidel, Ioannis Mitliagkas.
- 2019 **User Perception of Situated Product Recommendations in Augmented Reality** [International Journal of Semantic Computing 13 \(03\)](#)
Brandon Huynh, Adam Ibrahim, Yun Suk Chang, Tobias Höllerer, John O'Donovan.

- 2018 **ARbis Pictus: A Study of Vocabulary Learning with Augmented Reality** [ISMAR 2018](#)
[Also published as a journal paper in IEEE transactions on visualization and computer graphics 24 \(11\)](#)
Adam Ibrahim, Brandon Huynh, Jonathan Downey, Tobias Höllerer, Dorothy Chun, John O'Donovan.
- 2018 **A Study of Situated Product Recommendations in Augmented Reality** [AIVR 2018](#)
Brandon Huynh, Adam Ibrahim, Yun Suk Chang, Tobias Höllerer, John O'Donovan.

Talks

- 10/09/24 **Empowering AI - The Essential Infrastructure** [Earlybird VC and OVNI Capital](#)
- 05/23/24 **Improving the Efficiency of Large Language Model Pretraining** [Meta](#)
- 02/07/23 **Towards Out-of-Distribution Adversarial Robustness** [Microsoft Research](#)
- 07/05/19 **Linear Lower Bounds and Conditioning of Differentiable Games** [Montreal MLOpt](#)
- 05/11/17 **Motivating Convolutional Neural Networks** [Microsoft Station Q](#)

Awards

- Bourse en Intelligence Artificielle (IA) des ESP** [Université de Montréal](#)
- Natural Sciences and Engineering Research Council of Canada**
- Trottier-Lavigne Physics Department Award**
- Edward Beatty Scholarship in Mathematics**
- John V Galley Scholarship in Mathematics**
- Dean's Honour List**

Events Organised

- 12/15/2023 **6th Neural Scaling Laws Workshop** [NeurIPS 2023](#)
Workshop colocated with NeurIPS 2023. Co-organised with Irina Rish, Julia Bossmann, and the CERC-AAI team.
Link: <https://sites.google.com/mila.quebec/6thnslw-no/home>
- 07/28/2023 **Emergent Behaviours and Phase Transitions in Deep Learning** [ICML 2023](#)
Workshop colocated with ICML 2023. Co-organised with Irina Rish, Guillaume Dumas, Mohammad Pezeshki, Pascal J. Tikeng Notsawo, Hattie Zhou, Gabriela Moisescu-Pareja, Ethan Caballero, Yi Ren, Eric Michaud.
Link: <https://sites.google.com/mila.quebec/5thnslw>
- 12/02/2022 **4th Neural Scaling Laws Workshop** [NeurIPS 2022](#)
Unofficial workshop held during NeurIPS on Friday 2nd, 2022. Co-organised with Irina Rish. You are encouraged to attend or reach out for more information!
Link: <https://sites.google.com/mila.quebec/4thnslw/home>
- 2019-2021 **Deep Learning Theory / Out-of-Distribution Generalisation Reading Group** [Mila](#)

Teaching Assistant Experience

Spring 2018	CS 178 Introduction to Cryptography	UC Santa Barbara, USA
Winter 2018	CS 130B Data Structures and Algorithms II	UC Santa Barbara, USA
Fall 2017	CS 174A Fundamentals of Database Systems	UC Santa Barbara, USA
Summer 2017	CS 16 Problem Solving with Computers 1	UC Santa Barbara, USA
Spring 2017	CS 165B Machine Learning	UC Santa Barbara, USA
Winter 2017	CS 181B Introduction to Computer Vision	UC Santa Barbara, USA
Fall 2016	CS 40 Foundations of Computer Science	UC Santa Barbara, USA
Summer 2016	CS 16 Problem Solving with Computers 1	UC Santa Barbara, USA
Spring 2016	CS 24 Problem Solving with Computers 2	UC Santa Barbara, USA
Winter 2016	CS 16 Problem Solving with Computers 1	UC Santa Barbara, USA
Fall 2015	CS 16 Problem Solving with Computers 1	UC Santa Barbara, USA
2014	MATH 381 Complex Variables and Transforms for Engineers	McGill University
2013 & 2014	MATH 249 Honours Complex Variables	McGill University

Conferences Attended

05/24	2024 ICLR International Conference on Learning Representations
12/23	2023 NeurIPS Neural Information Processing Systems Presentation of <i>Continual Pre-Training of Large Language Models: How to Re-warm Your Model?</i> at workshops.
07/23	2023 ICML International Conference on Machine Learning Presentation of <i>Continual Pre-Training of Large Language Models: How to Re-warm Your Model?</i> at the Efficient Systems for Foundation Models workshop and <i>Towards Out-of-Distribution Adversarial Robustness</i> at the New Frontiers in Adversarial Machine Learning workshop.
12/22	2022 NeurIPS Neural Information Processing Systems Poster presentation of <i>Gradient Descent Is Optimal Under Lower Restricted Secant Inequality And Upper Error Bound</i> .
08/22	2022 CoLLAs Conference on Lifelong Learning Agents
07/22	2022 ICML International Conference on Machine Learning Poster presentation of <i>Towards Generalisable Robustness: A Domain Generalisation Approach</i> at the New Frontiers in Adversarial Machine Learning workshop.
20-22 COVID	Attended the virtual NeurIPS, ICML and ICLR conferences.
07/20	2020 ICML International Conference on Machine Learning Poster presentation of <i>Linear Lower Bounds and Conditioning of Differentiable Games</i> .
10/19	2019 DeepMath Conference on the Mathematical Theory of Deep Neural Networks Poster presentation of <i>Linear Lower Bounds and Conditioning of Differentiable Games</i> .
09/19	2019 MAIS Montreal AI Symposium Poster presentation of <i>Linear Lower Bounds and Conditioning of Differentiable Games</i> .
12/18	2018 NeurIPS Neural Information Processing Systems
10/18	2018 ISMAR International Symposium on Mixed and Augmented Reality Oral presentation of <i>ARbis Pictus: A Study of Vocabulary Learning with Augmented Reality</i> .
12/17	2017 NeurIPS Neural Information Processing Systems

10/17

2017 South California Machine Learning Symposium (USC)

03/17

2017 IEEE VR 3DUI Conference on Virtual Reality and 3D User Interfaces