

Divyat Mahajan

Final Year Ph.D. Student
Mila, Université de Montréal

divyatmahajan@gmail.com | divyat.mahajan@mila.quebec
Webpage : divyat09.github.io
Google Scholar, Github, LinkedIn

EDUCATION

- Université de Montréal/Mila** *September '22 - August '26 (Expected)*
Ph.D. in Computer Science GPA: 4.3/4.3
Advisor: Prof. Ioannis Mitliagkas
- Université de Montréal/Mila** *September '21 - August '22*
M.Sc. in Computer Science GPA: 4.3/4.3
Advisor: Prof. Ioannis Mitliagkas
- Indian Institute of Technology Kanpur** *July '14 - June '19*
B.S. in Mathematics and Scientific Computing GPA: 8.6/10
B.Tech in Computer Science and Engineering (Double Major)

RESEARCH INTERESTS

- Primary:** Foundational Models, Out-of-Distribution Generalization, Compositional Generalization
Others: Causal Inference, In-context Learning, World Models, Mechanistic Interpretability

WORK EXPERIENCE

- **Meta Superintelligence Labs (FAIR) - Visiting Researcher** *Montréal*
Advisor: Dr. Kartik Ahuja, Dr. Pascal Vincent, and Dr. David Lopez-Paz *January '24 - January '26*
Topics: Language Model Pretraining, Compositional Generalization
- **Microsoft Research - Research Intern** *Cambridge*
Advisor: Dr. Meyer Scetbon and Dr. Cheng Zhang *June '23 - September '23*
Topics: Amortized Inference, In-context Learning
- **Microsoft Research - Research Fellow** *Bangalore*
Advisor: Dr. Amit Sharma *July '19 - July '21*
Topics: Out-of-distribution Generalization, Trustworthy Machine Learning, Causal Inference
- **Aalto University - Research Intern** *Helsinki*
Advisor: Prof. Samuel Kaski *May '18 - July '18*
Topics: Approximate Bayesian Computation, Bayesian Optimization
- **National University of Singapore - Research Intern** *Singapore*
Advisor: Prof. Wynne Hsu and Prof. Lee Mong Li *May '17 - July '17*
Topics: Recommender Systems

PUBLICATIONS (*: EQUAL CONTRIBUTION)

Preprints

- **Iterative Amortized Inference: Unifying In-Context Learning and Learned Optimizers**
Sarathak Mittal, Divyat Mahajan, Guillaume Lajoie, Mohammad Pezeshki
Under Review

Conference Publications

- **Beyond Multi-Token Prediction: Pretraining LLMs with Future Summaries**
Divyat Mahajan, Sachin Goyal, Badr Youbi Idrissi, Mohammad Pezeshki, Ioannis Mitliagkas, David Lopez-Paz, Kartik Ahuja
Proceedings of the International Conference of Learning Algorithms (ICLR '26)
- **Compositional Risk Minimization**
Divyat Mahajan, Mohammad Pezeshki, Charles Arnal, Ioannis Mitliagkas, Kartik Ahuja, Pascal Vincent
Proceedings of the International Conference on Machine Learning (ICML '25)
- **Amortized Inference of Causal Models via Conditional Fixed-Point Iterations**
Divyat Mahajan*, Jannes Gladrow, Agrin Hilmkil, Cheng Zhang, Meyer Scetbon*
Proceedings of the Transactions on Machine Learning Research (TMLR '25) (J2C Certification), (ICLR '26)

- **Path-specific effects for pulse-oximetry guided decisions in critical care**
Kevin Zhang, Yonghan Jung, [Divyat Mahajan](#), Karthikeyan Shanmugam, Shalmali Joshi
Proceedings of the Neural Information Processing Systems (NeurIPS '25)
- **Empirical Analysis of Model Selection for Heterogeneous Causal Effect Estimation**
[Divyat Mahajan](#), Ioannis Mitliagkas, Brady Neal, Vasilis Syrgkanis
Proceedings of the International Conference of Learning Algorithms (ICLR '24) (**Spotlight**)
- **Additive Decoders for Latent Variables Identification and Cartesian-Product Extrapolation**
Sébastien Lachapelle*, [Divyat Mahajan](#)*, Ioannis Mitliagkas, Simon Lacoste-Julien
Proceedings of the Neural Information Processing Systems (NeurIPS '23) (**Oral**)
- **Interventional Causal Representation Learning**
Kartik Ahuja, [Divyat Mahajan](#), Yixin Wang, Yoshua Bengio
Proceedings of the International Conference on Machine Learning (ICML '23) (**Oral**)
- **Synergies Between Disentanglement and Sparsity: A Multi-Task Learning Perspective**
Sebastien Lachapelle*, Tristan Deleu*, [Divyat Mahajan](#), Ioannis Mitliagkas, Yoshua Bengio, Simon Lacoste-Julien, Quentin Bertrand
Proceedings of the International Conference on Machine Learning (ICML '23)
- **Towards efficient representation identification in supervised learning**
Kartik Ahuja*, [Divyat Mahajan](#)*, Vasilis Syrgkanis, Ioannis Mitliagkas
Proceedings of the Conference on Causal Learning and Reasoning (CLear '22)
- **Domain Generalization using Causal Matching**
[Divyat Mahajan](#), Shruti Tople, Amit Sharma
Proceedings of the International Conference on Machine Learning (ICML '21) (**Oral**)
- **Split-Treatment Analysis to Rank Heterogeneous Causal Effects for Prospective Interventions**
Yanbo Xu, [Divyat Mahajan](#), Liz Manrao, Amit Sharma, Emre Kiciman
Proceedings of the ACM International Conference on Web Search and Data Mining (WSDM '21) (**Oral**)
- **Towards Unifying Feature Attribution and Counterfactual Explanations: Different Means to Same End**
Ramaravind Mothilal, [Divyat Mahajan](#), Chenhao Tan, Amit Sharma
Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '21)
- **A Generative Framework for Zero-Shot Learning with Adversarial Domain Adaptation**
Varun Khare*, [Divyat Mahajan](#)*, Homanga Bharadhwaj, Vinay Verma, Piyush Rai
Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV '20)

Workshop Papers

- **Learning to Defer for Causal Discovery with Imperfect Experts**
Oscar Clivio, [Divyat Mahajan](#), Perouz Taslakian, Sara Magliacane, Ioannis Mitliagkas, Valentina Zantedeschi, Alexandre Drouin
Workshop on Reasoning and Planning for Large Language Models (ICLR '25)
- **Evaluating Interventional Reasoning Capabilities of Large Language Models**
Tejas Kasetty, [Divyat Mahajan](#), Gintare Karolina Dziugaite, Alexandre Drouin, Dhanya Sridhar
Workshop on Causality and Large Models (NeurIPS '24)
- **The Connection between Out-of-Distribution Generalization and Privacy of ML Models**
[Divyat Mahajan](#), Shruti Tople, Amit Sharma
Workshop on Privacy Preserving Machine Learning (NeurIPS '20)
- **Preserving Causal Constraints in Counterfactual Explanations for Machine Learning Classifiers**
[Divyat Mahajan](#), Chenhao Tan, Amit Sharma
Workshop on Machine Learning and Causal Inference for Improved Decision Making (NeurIPS '19) (**Oral**)

RESEARCH TALKS

- **Compositional Risk Minimization**

Imperial College London

King's College London

May' 25

May' 25

- **Latent Identification and Extrapolation with Additive Decoders**

Microsoft Research India

January' 24

- **Latent Identification via Interventions**

Microsoft Research Cambridge

July' 23

- **Latent Identification in Multi-Task Learning**

Microsoft Research India

August '22

CausalML Reading Group, ServiceNow Research

November '22

Causality Discussion Group, <https://www.matej-zecevic.de/cdg/> [Recording]

November '23

ACADEMIC SERVICE

- **Reviewer:** ICML 2022-2025, NeurIPS 2021-2025, ICLR 2024-2026, AISTATS 2024, TMLR 2025

CleaR 2024-2025, ML RC 2021-2022, MAIS 2022, IEEE SMDS 2021

STUDENT MENTORSHIP

- **Yukti Mukhija** (Ph.D. at Mila, University of Montreal)

Finite Sample Analysis of Compositional Risk Minimization

2025-Present

- **Tejas Kasetty** (M.Sc. at Mila, University of Montreal)

Evaluating Interventional Reasoning Capabilities of Large Language Models

2023-24

SOFTWARE

- **RobustDG - Microsoft** [Github] [Commit History]

Core developer for Microsoft's open-source framework for building robust ML models that generalize to unseen domains

- **DiCE - InterpretML** [Github] [Commit History]

Collaborating on the InterpretML's open-source framework to support feasible counterfactual explanation approaches

AWARDS & ACHIEVEMENTS

- **Oral/Spotlight Conference Publication**

WSDM 2021, ICML 2021, ICML 2023, NeurIPS 2023, ICLR 2024

- **Outstanding/Top Reviewer Award**

ML RC 2021, ML RC 2022, ICML 2022, NeurIPS 2022, NeurIPS 2024

- **Excellence Scholarship, University of Montreal (\$5,000)**

Awarded for the session 2025-2026 for exceptional academic performance in the Ph.D. program

- **FRQNT Doctorate Scholarship (\$100,000)**

Rank #1 among all applicants in my category. Competition 2024-2025. [DOI]

- **Fast Track Scholarship, University of Montreal (\$14,000)**

Awarded for exceptional academic performance upon transition to Ph.D. from M.Sc. program in Fall 2022.

- **Amii (Upper Bound) Talent Bursary (\$1,500 × 3)**

Awarded to attend AI Week 2022 and Upper Bound 2023, 2024 organized by Amii

- **Quebec Ministry of Higher Education International Students Scholarship (\$4,000)**

Awarded for the session 2021-2022 during the M.Sc. program

- **NeurIPS Travel Assistance (\$1,000)**

Awarded to present my work at NeurIPS 2019

- **Academic Excellence Award, IIT Kanpur**

Awarded for the session 2018-2019 for exceptional academic performance in the undergraduate program

TEACHING AND LEADERSHIP

- (2025) Co-organizing Mila Tea Talks, institution-wide academic talks.
- (2024) Guest lecturer for two classes of the course IFT 6168 (Causal Inference & ML) at the Université de Montréal
- (2023-24) Member of student recruitment committee for screening the M.Sc./Ph.D. applications at Mila
- (2020-21) Co-organizing the Machine Learning Reading Group at Microsoft Research India
- (2018) Project mentor for the course Machine Learning Techniques (CS771A) at IIT Kanpur
- (2018) Mentored 5 students for a project on Recommender Systems under Association of Computing Activities, IITK
- (2017-18) Managed a team of 5 members to publish newsletter Alpha under Statmatics, mathematics society of IITK
- (2016-17) Volunteered in Blood Connect, NGO working to provide a solution for the shortage of blood in India
- (2015-16) Volunteered in National Service Scheme at IIT Kanpur to provide better education to underprivileged children

TECHNICAL SKILLS

Programming Languages Python, C/C++, Bash

Software and Utilities Git, Slurm, L^AT_EX, PyTorch, Hugging Face, vLLM, Sklearn, ELFI, DoWhy

RELEVANT COURSEWORK (*: RECEIVED BEST GRADE)

Machine Learning	Machine Learning*, Representation Learning*, Computer Vision* Causal Inference*, Probabilistic Graphical Models* Theoretical Deep Learning*, Probabilistic Machine Learning*
Statistics	Probability and Statistics*, Applied Stochastic Process, Statistical Inference
Mathematics	Calculus*, Linear Algebra & Differential Equations*, Numerical Computation* Real Analysis, Complex Analysis, Abstract Algebra, Topology*
Algorithms & Theory	Data Structures and Algorithms, Advanced Algorithms, Quantum Computing