Alireza Mousavi-Hosseini

mousavi@cs.toronto.edu | mousavih.github.io | Google Scholar | Github

Research Interests

Foundations of LLMs/Deep Learning, Generative Modeling, High-Dimensional Statistics

Education

University of Toronto Sept. 2021 – Present
Ph.D. in Computer Science
- GPA 4.0/4.0

Sharif University of Technology Sept. 2017 – July 2021
B.Sc. in Computer Engineering
- GPA 19.76/20 (equivalent to 4.0/4.0)

Research Experience

Research Internship Apr. 2025 - Sept. 2025
Apple ML Research Paris, France
Supervisor: Marco Cuturi

Graduate Student Researcher Sept. 2021 - Present
Vector Institute Toronto, Canada

Visiting PhD Student Sept. 2023 - Nov. 2023
École Polytechnique Fédérale de Lausanne (EPFL) Lausanne, Switzerland
Supervisor: Lénaïc Chizat
- Kernel learning via mean-field Langevin dynamics.

Research Internship July 2020 - Dec. 2020
IST Austria Vienna, Austria
Supervisor: Dan Alistarh
- Using second-order information for neural network weight quantization.

Research Internship July 2019 - Sept. 2019
École Polytechnique Fédérale de Lausanne (EPFL) Lausanne, Switzerland
Supervisor: Christoph Koch
- Designing an RNN-based system for efficient approximation of real-world simulation behavior.

Honors and Awards

- Ontario Graduate Scholarship (CAD 15,000) 2025-2026
- ICLR Notable Reviewer 2025
- Borealis AI Fellowship (CAD 10,000) 2024-2025
- Mary H. Beatty Fellowship (CAD 10,000) 2024-2025
- Department of Computer Science 50th Anniversary Graduate Scholarship, University of Toronto 2023-2025
- C.C. Gotlieb (Kelly) Graduate Fellowship, University of Toronto 2021-2023
- Vector Institute Research Grant 2021-2026
- Graduated in the top 4% of Computer Engineering class of 2021, Sharif University of Technology 2021
- International Physics Olympiad (IPhO) Silver Medalist 2017
- National Physics Olympiad Gold Medalist 2016

Publications

- **Alireza Mousavi-Hosseini**, Clayton Sanford, Denny Wu, Murat A. Erdogdu. "When Do Transformers Outperform Feedforward and Recurrent Networks? A Statistical Perspective." To appear in **NeurIPS**, *Advances in Neural Information Processing Systems*, 2025.
  [**Oral Presentation** at *ICML 2025 workshop on high-dimensional learning dynamics.*]

- KC Tsiolis, **Alireza Mousavi-Hosseini**, Murat A. Erdogdu. "From Information to Generative Exponent: Learning Rate Induces Phase Transitions in SGD". To appear in **NeurIPS**, *Advances in Neural Information Processing Systems*, 2025.
- **Alireza Mousavi-Hosseini**, Adel Javanmard, Murat A. Erdogdu. "Robust Feature Learning for Multi-Index Models in High Dimensions." **ICLR**, *Proceedings of the Thirteenth International Conference on Learning Representations*, 2025.
- **Alireza Mousavi-Hosseini**, Denny Wu, Murat A. Erdogdu. "Learning Multi-Index Models with Mean-Field Neural Networks." **ICLR**, *Proceedings of the Thirteenth International Conference on Learning Representations*, 2025.
- Guillaume Wang*, **Alireza Mousavi-Hosseini***, Lénaïc Chizat. "Mean-Field Langevin Dynamics for Signed Measures via a Bilevel Approach." **NeurIPS**, *Advances in Neural Information Processing Systems*, 2024. **(Spotlight)**
- Ye He, **Alireza Mousavi-Hosseini**, Krishnakumar Balasubramanian, Murat A. Erdogdu. "A Separation in Heavy-Tailed Sampling: Gaussian vs. Stable Oracles for Proximal Samplers." **NeurIPS**, *Advances in Neural Information Processing Systems*, 2024.
- **Alireza Mousavi-Hosseini**, Denny Wu, Taiji Suzuki, Murat A. Erdogdu. "Gradient-Based Feature Learning under Structured Data." **NeurIPS**, *Advances in Neural Information Processing Systems*, 2023.
- **Alireza Mousavi-Hosseini***, Tyler Farghly*, Ye He, Krishnakumar Balasubramanian, Murat A. Erdogdu. "Towards a Complete Analysis of Langevin Monte Carlo: Beyond Poincaré Inequality." **COLT**, *Proceedings of the Thirty Sixth Conference on Learning Theory*, 2023.
- **Alireza Mousavi-Hosseini**, Sejun Park, Manuela Girotti, Ioannis Mitliagkas, and Murat A. Erdogdu. "Neural Networks Efficiently Learn Low-Dimensional Representations with SGD." **ICLR**, *Proceedings of the Eleventh International Conference on Learning Representations*, 2023. (**Spotlight**)

*Equal Contribution.

Preprints

- **Alireza Mousavi-Hosseini***, Stephen Zhang*, Michal Klein, Marco Cuturi. "Flow Matching with Semidiscrete Couplings." *arXiv preprint arXiv:2509.25519*, 2025.
- Stephen Zhang*, **Alireza Mousavi-Hosseini***, Michal Klein, Marco Cuturi. "On Fitting Flow Models with Large Sinkhorn Couplings." *arXiv preprint arXiv:2506.05526*, 2025.
- Jivan Waber, **Alireza Mousavi-Hosseini**, Murat A. Erdogdu. "Fundamental Limits of Learning Single-Index Models under Structured Data." In preparation, 2025.
  [**Oral Presentation** at *ICML 2025 workshop on high-dimensional learning dynamics.*]

Technical Skills

Python, C++, R, Java, PyTorch, JAX, Tensorflow/Keras, Numpy/Scipy/Scikit-Learn, Jupyter Notebook, Git, Slurm

Invited Talks

Learning and Optimization with Mean-Field Langevin Dynamics.
Mila - Quebec AI Institute. November 2024

Robustness and Feature Learning in Neural Networks.
Vector Institute. November 2024

Gradient-Based Feature Learning under Structured Data.
Foundations of Learning and AI Research (FLAIR) Seminar, EPFL. October 2023

Gradient-Based Feature Learning of Neural Networks.
Institute of Applied Mathematics, UBC. June 2023

Neural Networks Efficiently Learn Low-Dimensional Representations with SGD.
Mila - Quebec AI Institute. October 2022

Teaching Experience

Teaching Assistant at the University of Toronto Sept. 2021 - Present

Statistical Methods for Machine Learning II (STA 414/2104), Introduction to Machine Learning (CSC 311), Probabilistic Learning and Reasoning (CSC 412/2506), Foundations of Computer Science I (CSC 110).

Teaching Assistant at Sharif University of Technology Sept. 2019 - Dec. 2020

Machine Learning, Probability and Statistics, Data Structures and Algorithms, Computer Networks

Physics Olympiad Teacher Nov. 2016 - Jan. 2018

Allameh Helli High School

Academic Service

Journal Reviewer

Journal of Machine Learning Research **(JMLR)**, SIAM Journal on Mathematics of Data Science **(SIMODS)**, Transactions on Machine Learning Theory **(TMLR)**

Conference Reviewer

Neural Information Processing Systems **(NeurIPS)**, International Conference on Machine Learning **(ICML)**, International Conference on Learning Representations **(ICLR)**, Conference on Learning Theory **(COLT)**, International Conference on Artifiticial Intelligence and Statistics **(AISTATS)**

Departmental Service

Graduate Application Assistance Program (GAAP) Mentor November 2024
Department of Computer Science, University of Toronto

Coaching in Excel to AI for Black & Indigenous Students October 2024
Vector Institute

Graduate Applications Triager December 2023
Department of Computer Science, University of Toronto