Mohsin Hasan

Tel: 647-571-8737
Email: hasanmohsin19@gmail.com
Github: https://github.com/hasanmohsin
LinkedIn: https://www.linkedin.com/in/mohsin-hasan-421891254/
Website: https://hasanmohsin.github.io/

EDUCATION

University of Montreal, Mila **(Sept 2023 – Expected May 2027)**
PhD. in Computer Science – Supervisor: Yoshua Bengio, Year 3
Topics: RL Fine-tuning, Diffusion and Flow Models, Sampling, Bayesian Inference

University of Waterloo **(September 2021 – August 2023)**
Research Masters in Computer Science – Supervisor: Pascal Poupart
Topics: Bayesian Federated Learning, Calibration
CGPA: **4.0/4.0 (Cumulative Percentage: 100%)**

University of Toronto, St. George Campus **(September 2016 - May 2021)**
Bachelors in Electrical and Computer Engineering (U of T Scholar)
CGPA: **3.99/4.0 (Cumulative Percentage: 95.5%)**
- Rank 2 in Computer Engineering 2nd and 3rd Year
- Rank 1 in Computer Engineering 1st Year

SKILLS

- Technical languages: Python, MATLAB, LaTeX, C++, C, Perl, Java
- Software/Platforms: PyTorch, Tensorflow, Slurm Compute Cluster, Linux, Simulink, Microsoft Office

RESEARCH EXPERIENCE

PhD Research Student – University of Montreal, Mila **(Sept 2023 – Present)**

- Research focuses on scalable methods for reward fine-tuning of generative models (particularly diffusion/flow models) – through both RL training, and inference time methods
- Additionally interested in scalable methods to train samplers for a given energy function (both discrete and continuous)
- Led and collaborated on multiple projects published at conferences such as NeurIPS, ICLR and ICML

Master's Research Student - University of Waterloo, Vector Institute **(Oct 2021 – August 2023)**

- Worked on methods for doing efficient Bayesian inference in a federated setting
- Invented and evaluated a method for aggregating Bayesian posteriors in the predictive space, allowing for federated learning in 1 communication round, with accurate calibration
- Outperforms state of the art methods on heterogeneous data
- Led to work published at AAAI2024 (as first author)

Research Student - University of Toronto (With Will Grathwohl) **(May 2021 – August 2021)**

- Investigated sampling from intractable density with Learned Stein Discrepancy (SD)
- Implemented sampler, compared to baselines on simple distributions and RBMs
- Formulated alternate version of sampler which implicitly minimized KL divergence instead of SD
- Performed better than Stein Variational Gradient Descent in higher dimensions, but below Langevin Sampler

Research Student - University of Toronto, Vector Institute (CS PAIR Lab) **(August 2020 – August 2021)**

- Led research project for designing deep learning algorithm to construct an action graph from an instructional video

- Model took video and transcript text as input, and generated a “visually grounded action graph” with nodes containing critical action steps (as text), and graph edges representing action interdependencies. Entities in action steps are also linked to bounding boxes, tracked over the video.
- Adapted Transformer-based architecture to find joint visual-linguistic representations over the input

PROFESSIONAL EXPERIENCE

Engineering Intern – Qualcomm Markham (Display Team) **(May 2019 – August 2020)**

- Responsible for designing image processing algorithms which enhance image quality (for things such as lighting, and saturation) through the use of deep learning and general statistical techniques
- Developed learning based algorithm for simulating high-dynamic range images by manipulating lighting content in standard dynamic range images

PUBLICATIONS

Adaptive Order Policies for Masked Diffusion **(Jan 2026)**
Mohsin Hasan*, Jama Hussein Mohamud*, Mirco Ravanelli, Yoshua Bengio

- Preprint, workshop version at https://openreview.net/forum?id=0ZBqSpGn4m

Discrete Feynman-Kac Correctors **(Jan 2026)**
Mohsin Hasan, Viktor Ohanesian, Artem Gazizov, Yoshua Bengio, Alán Aspuru-Guzik, Roberto Bondesan, Marta Skreta, Kirill Neklyudov

- Under review, preprint at https://arxiv.org/abs/2601.10403

Outsourced diffusion sampling: Efficient posterior inference in latent spaces of generative models **(Feb 2025)**
Siddarth Venkatraman*, **Mohsin Hasan***, Minsu Kim, Luca Scimeca, Marcin Sendera, Yoshua Bengio, Glen Berseth, Nikolay Malkin

- ICML 2025, as a conference paper https://arxiv.org/abs/2502.06999

Steering masked discrete diffusion models via discrete denoising posterior prediction **(Oct 2024)**
Jarrid Rector-Brooks, **Mohsin Hasan**, Zhangzhi Peng, Zachary Quinn, Chenghao Liu, Sarthak Mittal, Nouha Dziri, Michael Bronstein, Yoshua Bengio, Pranam Chatterjee, Alexander Tong, Avishek Joey Bose

- ICLR 2025, as a conference paper https://arxiv.org/abs/2410.08134

Amortizing intractable inference in diffusion models for vision, language, and control **(May 2024)**
Siddarth Venkatraman*, Moksh Jain*, Luca Scimeca*, Minsu Kim*, Marcin Sendera*, **Mohsin Hasan**, Luke Rowe, Sarthak Mittal, Pablo Lemos, Emmanuel Bengio, Alexandre Adam, Jarrid Rector-Brooks, Yoshua Bengio, Glen Berseth, Nikolay Malkin

- NeurIPS 2024, as a conference paper https://arxiv.org/abs/2405.20971

Estimating Expectations without Sampling: Neural Stein Estimation **(May 2024)**
Mohsin Hasan, Dinghuai Zhang, Cheikh Ahmed, Awa Khouna, Yoshua Bengio

- Appeared at Approximate Inference workshop https://openreview.net/pdf?id=by9cKVIs0F

Calibrated One Round Federated Learning with Bayesian Inference in the Predictive Space **(Feb 2024)**
Mohsin Hasan, Guojun Zhang, Kaiyang Guo, Xi Chen, Pascal Poupart

- AAAI 2024, as a conference paper https://ojs.aaai.org/index.php/AAAI/article/view/29122

TALKS

International Workshop on Trustworthy Federated Learning in conjunction with IJCAI 2022 **(July 2022)**
15 minute presentation (https://federated-learning.org/fl-ijcai-2022/)

ACHIEVEMENTS

- Received USRA NSERC award for undergraduate research **(May 2018)**
- Received over **$ 56,000** in scholarships and awards **(April 2017 - present)**
- One of 8 students to receive First Year Research Fellowship **($ 5000**) **(May 2017)**

RELEVANT COURSEWORK

- Graduate: Dynamical Systems - MAT6215 (**A+**) (University of Montreal)
- Graduate: Probabilistic Graphical Models - IFT6269 (**A+**) (University of Montreal)
- Graduate: Stochastic Processes - AMATH777 (**100%**) (University of Waterloo)
- Graduate: Optimization - CS794 (**100%**) (University of Waterloo)
- Graduate: Reinforcement Learning - CS885 (**100%**) (University of Waterloo)
- Undergraduate: Deep Learning - CSC413 (**96%**) (University of Toronto)

TEACHING ASSISTANT

- Graduate: Representation Learning – IFT6135 (University of Montreal)
- Graduate: Reinforcement Learning - CS885 (University of Waterloo)