

# Md Rifat Arefin

Montreal, Canada

✉ rifat.arefin515@gmail.com • 🌐 [rarefin.github.io/](https://rarefin.github.io/)

## Intro

I am a PhD student at Mila and University of Montreal, supervised by Irina Rish. I had been fortunate to work with Yoshua Bengio and Yann LeCun. I study how architectural and optimization-induced inductive biases shape the geometry of representations in foundation models, aiming to improve generalization and multi-step reasoning. I develop optimization strategies and work on architectural modifications that regulate signal propagation to learn better representations for Foundation Models.

## Work Experience

### **Research Scientist Intern:**

Meta; California, US

August 2025 to January 2026

Working on LLMs pretraining dynamics to improve and design better optimization algorithms.

### **Research Intern:**

Samsung; Montreal, Canada

February 2025 to July 2025

Working on improving reasoning capabilities in large language models (LLMs) by exploring architectural modifications and training strategies to enhance logical consistency and problem-solving efficiency.

### **Applied Research Scientist Intern:**

Amazon; Seattle, USA

October 2024 to January 2025

Worked on diffusion-based generative models for controllable image generation, focusing on improving model conditioning techniques to enhance fidelity, diversity, and user-driven customization.

### **Visiting Researcher (part-time):**

ServiceNow; Montreal, Canada

May 2024 to September 2024

Conducting research on improving reasoning in large language models (LLMs), focusing on architectural modifications and training techniques to enhance structured problem-solving abilities.

### **Research Intern (part-time):**

Recursion Pharmaceutical; Montreal, Canada

June 2023 to December 2023

Worked on a project related to learning representations from images of human cells that have been manipulated by CRISPR/Cas9-mediated gene knockouts, compounds, or other reagents to predict trillions of relationships across biology and chemistry.

### **Visiting Researcher:**

Eberhard Karls University of Tübingen; Tübingen, Germany

January 2021 to June 2021

Worked on a project related to speech separation conditioning on speaker style, which was published

in Interspeech 2021. Focused on improving deep learning models for audio processing and source separation.

**Deep Learning R&D Engineer:**

UpStride; Paris, France

February 2020 to November 2020

Explored alternative algebraic structures such as complex numbers and quaternions for deep learning architectures, aiming to improve efficiency and expressiveness of neural networks.

**Research Intern:**

MILA - Quebec AI Institute; Montreal, Canada

August 2018 to February 2019 (7 months)

Conducted research under the supervision of Professor Yoshua Bengio, focusing on a satellite image super-resolution competition organized by the European Space Agency. Secured 1st runner-up position in the PROBA-V Super Resolution competition. Competition website: <https://kelvins.esa.int/proba-v-super-resolution/leaderboard/results> - Team "rarefin".

**Internship:**

Orion Informatics Limited; Dhaka, Bangladesh

January 2016 to June 2016 (6 months)

Completed a six-month internship as part of my Bachelor's degree requirements, working on software development projects, gaining hands-on experience in programming, and contributing to the development of web applications.

## Education

---

**Ph.D. Student**

*MILA - Quebec AI Institute & University of Montreal,*  
CGPA 4.3 (out of 4.3)

**Probable Graduation Year**

*December 2026*

**Masters in Computer Science & Engineering**

*Department of Computer Science & Engineering, Kyung Hee University,*  
CGPA 4.3 (out of 4.3)

**Passing Year**

*February, 2019*

**B. Sc. in Software Engineering**

*Institute of Information Technology, University of Dhaka,*  
CGPA 3.89 (out of 4)

**Passing Year**

*February, 2017*

## Achievements

---

- IVADO PhD Excellence Scholarships
- 1st Runner up in PROBA-V Super Resolution Competition organized by European Space Agency
- IEEE BDS best paper award in International conference on Computer and information Technology 2015
- Obtained Govt. scholarship in class V, VIII and X standard
- Obtained President Scholarship during the study of Masters

## Mentorship and Reviewer Experience

---

- Reviewer of NeurIPS 2024. Reviewer of ICML, ICLR, NeurIPS 2025
- Served as reviewer of AISTATS 2024
- Served as reviewer of ICLR 2024 Workshop on Reliable and Responsible Foundation Models

- Reviewer of ICML 2024 Workshop on Foundation Models in the Wild
- Internship Mentor of Mila professional Masters Student (May 2024 to December 2024)

## Research Work

---

### Publication:

- [1] Oscar Skean, **Md Rifat Arefin**, Dan Zhao, Niket Nikul Patel, Jalal Naghiyev, Yann LeCun, Ravid Schwartz-Ziv. Layer by Layer: Uncovering Hidden Representations in Language Models. (**Accepted at, ICML 2025, with Spotlight**)
- [2] **Md Rifat Arefin**, Gopesh Subbaraj, Nicolas Gontier, Yann LeCun, Irina Rish, Ravid Schwartz-Ziv, Christopher Pal. Seq-VCR: Preventing Collapse in Intermediate Transformer Representations for Enhanced Reasoning. (Published at, ICLR 2025)
- [3] **Md Rifat Arefin**, Zhang Y, Baratin A, Locatello F, Rish I, Liu D, Kawaguchi K. Unsupervised Concept Discovery Mitigates Spurious Correlations. arXiv e-prints. 2024 Feb:arXiv-2402. (Published at, ICML 2024)
- [4] Hu, Shell Xu, **Md Rifat Arefin**, Viet-Nhat Nguyen, Alish Dipani, Xaq Pitkow, and Andreas Savas Tolias. "AvaTr: One-Shot Speaker Extraction with Transformers." arXiv preprint arXiv:2105.00609 (2021).
- [5] Ostapenko, Oleksiy, Timothee Lesort, Pau Rodríguez, **Md Rifat Arefin**, Arthur Douillard, Irina Rish, and Laurent Charlin. "Foundational Models for Continual Learning: An Empirical Study of Latent Replay." arXiv preprint arXiv:2205.00329 (2022).
- [6] **Md Rifat Arefin**, Vincent Michalski, Pierre-Luc St-Charles, Alfredo Kalaitzis, Sookkyung Kim, Samira E. Kahou, and Yoshua Bengio. "Multi-Image Super-Resolution for Remote Sensing Using Deep Recurrent Networks." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 206-207. 2020.
- [7] Deudon, Michel, Alfredo Kalaitzis, Israel Goytom, **Md Rifat Arefin**, Zhichao Lin, Kris Sankaran, Vincent Michalski, Samira E. Kahou, Julien Cornebise, and Yoshua Bengio. "HighRes-net: Recursive Fusion for Multi-Frame Super-Resolution of Satellite Imagery." arXiv preprint arXiv:2002.06460 (2020).
- [8] **Md Rifat Arefin**, Farkhod Makhmudkhujaev, Oksam Chae, and Jaemyun Kim. "Background Subtraction based on Fusion ofColor and Local Patterns." 2018 Asian Conference on Computer Vision (ACCV), 2-6 December, 2018, Perth, WA, Australia.
- [9] **Md Rifat Arefin**, Farkhod Makhmudkhujaev, Oksam Chae, and Jaemyun Kim. "Aggregating CNN and HOG Features for Real-Time Distracted Driver Detection." IEEE International Conference on Consumer Electronics - January 11-13, 2019, Las Vegas, Nevada, USA.