

## SUMMARY

I am a final-year PhD candidate at Mila, University of Montreal, anticipating graduation in the spring or summer of 2025. My research focuses on the fundamental aspects of Large Language Models (LLMs). Specifically, I have explored how data segmentation influences critical functions such as parametric knowledge retrieval and latent multi-hop reasoning.

In addition to investigating these foundational dynamics, my work also seeks to address the inherent limitations of LLMs. For instance, I aim to develop methods that allow these models to better consolidate their knowledge during training, enhancing their utility and impact.

I am deeply committed to advancing the field of machine learning through open-ended research and academic exploration. My goal is to produce impactful, publishable work that contributes meaningfully to the scientific community. I am seeking research-focused roles in industry that align with these values and support my passion for driving innovation.

## EDUCATION

### **PhD in Artificial Intelligence** - *Mila, Université de Montréal*

September 2019 - August 2025

- Thesis Topic: Impact of Data Segmentation on the Understanding and Problem-Solving Capabilities of Large Language Models
- GPA: 4.15/4.3
- Advisors: Sarath Chandar, Alain Tapp

## PUBLICATIONS

### **Consolidating the Knowledge of Large Language Models** - *First Author*

To be submitted

Proposed that LLMs should not only predict the next token during training but also reason about how new information influences their existing knowledge, akin to human information processing.

### **Effect of Data Segmentation and Packing on the Latent Multi-Hop Reasoning Capabilities of Large Language Models** - *First Author*

Under review

Found that duplicating training documents and segmenting & packing them in various ways significantly enhances the performance of LLMs on latent multi-hop reasoning tasks.

### **Do Large Language Models Know How Much They Know?** - *First Author*

EMNLP 2024

Investigated whether LLMs possess an understanding of the span of their knowledge with respect to a given topic.

### **EpiK-Eval: Evaluation for Language Models as Epistemic Models** - *First Author*

EMNLP 2023 (Oral)

Showed that language models tend to struggle more when recalling information spread across multiple training samples compared to when the same information is contained within a single sample. This leads to a higher rate of hallucinations.

### **PatchBlender: A Motion Prior for Video Transformers** - *First Author*

NeurIPS 2022 Workshop

Introduced a learnable pooling function that applies over patch embeddings across the temporal dimension of the latent space of Vision Transformers.

### **Scaling Laws for the Few-Shot Adaptation of Pretrained Image Classifiers** - *First Author*

ICML 2021 Workshop

Showed that the few-shot generalization performance of image classifiers is well approximated by power laws as the pre-training set size increases.

### **Fully Quantized Transformer for Machine Translation** - *First Author*

Findings of EMNLP 2020

First paper to show that Transformers could be quantized to 8-bit without impairing performance.

### **Towards Lossless Encoding of Sentences** - *First Author*

ACL 2019

Proposed a near lossless method for encoding and decoding long sequences of texts into feature rich representations.

## **PROFESSIONAL EXPERIENCE**

### **Research Intern** - *Microsoft Research Montreal*

June 2024 - September 2024

Led a research project on the impact of data segmentation on LLM capabilities, resulting in a paper published at EMNLP.

### **Associate Researcher** - *Huawei Montreal*

January 2019 - December 2019

Tasked with quantizing the Transformer to 8 bits without compromising performance. I successfully achieved this and published a paper at EMNLP.

## **ACADEMIC EXPERIENCE**

### **Teaching Assistant** - *Polytechnique Montreal*

Fall 2021, Fall 2023

Designed and graded homework assignments, provided student assistance, and evaluated exams.

### **Research Mentor - *Mila***

Spring 2024 - Current

Mentored master's students on their research projects by generating ideas, answering questions, planning research schedules, and guiding them in conducting quality research and writing papers.

### **Graduate Application Reviewer - *Mila***

December 2018, December 2023

Evaluated academic credentials and research potential of applicants for Master's and PhD programs, recommending top candidates.

## **AWARDS**

I received an **Excellence Scholarship** for my Bachelor's Degree in Computer Science at Université de Montréal.

## **TECHNICAL SKILLS**

### **AI and Machine Learning Frameworks**

- PyTorch
- Huggingface Transformers & Accelerate
- DeepSpeed
- Numpy

### **Large Scale Training**

I have experience training neural networks up to 60 billion parameters on multi-node compute clusters.

### **Programming Languages**

- Python
- C
- C++
- Java

### **Other**

- Git/Github
- Docker

## **LANGUAGES SPOKEN**

Fluent in both English and French.