

# Aton Kamanda

<https://atonkamanda.github.io/> | [atonkamanda@hotmail.com](mailto:atonkamanda@hotmail.com) | <https://github.com/atonkamanda> | 438 543-4133

## EDUCATION

### University of Namur

*Bachelor of computer science with distinction (minor in mathematics)*

Namur

Sept. 2018 – Sept. 2021

### University of Montreal

*Master of artificial intelligence, 3.925/4.3 GPA*

Montreal

Sept. 2021 – Aug. 2023

### University of Montreal

*Ph.D artificial intelligence, advisor: Houari Sahraoui*

Montreal

Sept. 2023 – May 2024

## EXPERIENCE

### Machine learning engineer

March 2024 – Today

*AwakeAI - Mila incubated startup*

*Canada, Montreal*

- Developed a video understanding system based on V-JEPA for real-time activity recognition in nursing homes achieving state-of-the-art performance of 81% on the Toyota Smarthome benchmark.
- Leveraged JEPA's self-supervised learning capabilities to train on over 10,000 hours of unlabeled video data, resulting in robust representations that improved downstream task performance by 5%.
- Engineered a scalable video stream processing pipeline, optimizing the model to conduct real-time inference on 100+ simultaneous video streams using a single GPU server, reducing hardware costs by a factor of 2.6.
- Designed and implemented a comprehensive data pipeline with a feedback loop, enabling continuous model improvement through personalized training data derived from soft predictions and human-annotated labels.

### Machine learning research engineer

May 2023 – February 2024

*VMware*

*Canada, Montreal*

- IVADO partnership with VMware to leverage recent research findings to enhance the large language model (LLM) utilized internally, aimed at bolstering the efficiency of the software engineers for production.
- Some noteworthy improvements include a reduction in inference time, achieving up to 11x increase in inference throughput with performance drops of only 2%, prompt engineering (e.g chain of thought/chain of code), and the implementation of retrieval-augmented generation using CodeLlama, Langchain, and ChromaDB.

### Teacher assistant for a graduate robot learning class

January 2023 – May 2023

*Mila - Montreal institute for learning algorithms*

*Canada, Montreal*

- Course focused on research and composed mainly of PhD students, I have been in charge of creating entirely new assignments with recent research papers, writing automated tests on Gradescope, grading students, and helping students in their research contributions for the final project. (Course website).

### Machine learning intern

June 2021 – September 2021

*SkalUP*

*Belgium, Namur*

- Developed a search engine using CodeBERT to retrieve code snippets based on natural language queries, rank them by cosine similarity, and direct the user to the specific GitHub repo related to the snippet.

## PUBLICATIONS

### CodeUltraFeedback: LLM-as-a-Judge for coding preferences alignment | ACM 2024

March 2024

- Developed CodeUltraFeedback, a comprehensive preference dataset of 10,000 complex instructions aimed at tuning and aligning LLMs to user-defined coding preferences through AI feedback.
- Created CODAL-Bench, a benchmark for assessing LLM alignment with coding preferences, demonstrating that CodeLlama7B-Instruct, aligned through reinforcement learning from AI feedback with direct preference optimization (DPO) using CodeUltraFeedback's data, outperforms 34B LLMs on CODAL-Bench and improves functional correctness on HumanEval+

## TECHNICAL STRENGTHS

**Languages :** Python, Julia, C/C++, R, SQL

**Developer Tools :** Docker, Kubernetes

**Data:** Spark, Hadoop, Pandas, AWS, GCP, Azure, Kafka

**Machine learning:** Pytorch, Jax, TensorFlow, MLflow, LangChain, ChromaDB, NumPy, Gym, Mujoco