

Mathieu Duchesneau – Ingénieur IA & Logiciel

Québec, Canada – (514) 246-6064

 duchesneau.mathieu@gmail.com

 mathieuduchesneau.ca

 [LinkedIn](#)

 [GitHub](#)

RÉSUMÉ

Ingénieur en intelligence artificielle et en développement logiciel avec plus de 10 ans d'expérience dans la création de systèmes d'apprentissage automatique et d'infrastructures informatiques qui les entraînent et les exécutent.

Titulaire d'un doctorat de Mila (Université de Montréal) avec des travaux originaux introduisant la première méthode neuronale permettant une **récupération exacte de données en $O(1)$ à grande échelle** grâce à des encodeurs binaires conçus pour un adressage direct dans des tables de hachage.

Co-fondateur d'une start-up ayant déployé un système d'arbitrage d'électricité en temps réel automatisé avec des modèles IA exécutés sur une infrastructure Linux locale sur mesure.

Pour une présentation plus interactive, j'héberge moi-même un grand modèle de langage affiné (fine-tuned) sur mathieuduchesneau.ca, capable de répondre à des questions sur mon parcours. Je maintiens moi-même l'ensemble de la plateforme, de l'infrastructure Kubernetes jusqu'au front-end.

EXPÉRIENCE PROFESSIONNELLE & FORMATION

09/2017 – 01/2025

Doctorat en Intelligence Artificielle
Mila, DIRO, Université de Montréal

Montréal, Canada

- Thèse: [Learning Equivalence Hash Functions](#)
 - Conçu le premier objectif d'entraînement permettant à des réseaux de neurones d'encoder des éléments provenant de milliards de classes en représentations binaires uniques à chaque classe.
 - Développé un prototype capable d'effectuer des recherches complètes en **moins de 100 ms** dans une base de données contenant un milliard d'images, en utilisant un seul GPU pour l'encodage et un CPU standard pour la récupération via une table de hachage.
- Grande maîtrise des modèles de vision par ordinateur et de traitement du langage naturel.
- Gestion de pipelines d'expérimentations à grande échelle sur les grappes de calcul de Mila et de l'Alliance de recherche numérique du Canada via **Slurm**, ayant exécuté des centaines de tâches (jobs) d'entraînement.

08/2017 – 06/2019	Administrateur Système Linux & Ingénieur IA <i>Énergie Manifold Inc. (co-fondateur)</i>	Mascouche, Canada
	<ul style="list-style-type: none"> • Construit un système entièrement automatisé d'arbitrage d'électricité fonctionnant entre NYISO (US-NY) et IESO (CA-ON). • Conçu et implémenté des classifieurs ML pour décider entre « IESO → NYISO », « NYISO → IESO » ou « aucune transaction », à partir des prix du réseau en temps réel ainsi que de plus de 40 variables dérivées. • Déployé et maintenu une infrastructure Linux locale pour l'ingestion de données en temps réel, l'inférence des modèles et l'intégration complète des API avec les deux opérateurs de réseau. 	
09/2016 – 04/2022	Auxiliaire d'Enseignement <i>DIRO, Université de Montréal</i>	Montréal, Canada
	<ul style="list-style-type: none"> • Encadrement du cours Introduction à l'informatique théorique durant quatre sessions, incluant la correction des travaux et des présentations sur les automates finis, les langages context-free, les machines de Turing, les classes de complexité (telles que P et NP) et la décidabilité. • Encadrement du cours Introduction à la science des données, incluant l'évaluation des travaux et des présentations sur la classification, la régression, le clustering, les intervalles de confiance et les tests d'hypothèses. 	
09/2012 – 04/2015	Baccalauréat en Mathématiques & Informatique <i>DIRO, Université de Montréal</i>	Montréal, Canada
09/2010 – 05/2011	D.E.C. en Sciences de la Nature <i>Cégep de Lanaudière à Terrebonne</i>	Terrebonne, Canada

COMPÉTENCES	INGÉNIEUR IA	INGÉNIEUR LOGICIEL
	<ul style="list-style-type: none"> • Python • PyTorch • MLOps • MLflow • Hugging Face • Llama.cpp • Slurm • Vision par ordinateur • Réseaux convolutifs 	<ul style="list-style-type: none"> • Bash • FastAPI • Jinja2 • CSS • JavaScript • SQL • CI/CD • GitOps • DevOps

- Traitement du langage naturel
- Réseaux récurrents
- Grands modèles de langage
- Transformers
- Réglage fin (fine-tuning)
- LoRA / QLoRA
- Calcul différentiel et intégral
- Algèbre linéaire
- Probabilités & statistiques
- Docker
- Kubernetes
- GitHub Actions
- FluxCD
- Nginx
- Traefik
- Prometheus
- Grafana
- Structures de données
- Complexité algorithmique

PROJETS

10/2025 – Présent

Site Web Auto-Hébergé avec Assistant LLM

Montréal, Canada

- Déploiement full-stack d'une **plateforme Kubernetes** utilisant FastAPI, Traefik, Prometheus et Grafana pour le service et la surveillance.
- Mise en place d'un **pipeline CI/CD** avec GitHub Actions et FluxCD.
- Affinage (QLoRA) et hébergement d'un **LLM personnalisé** destiné à répondre à des questions sur mon travail et mon parcours.
- Intégration d'MLflow pour le **suivi continu d'expérimentations** avec la promotion automatique de nouveaux modèles SOTA en production.
- Le modèle est servi en direct sur mathieuduchesneau.ca, illustrant le cycle complet de déploiement IA, de l'entraînement à la production.

LANGUES

Français (courant)

Anglais (courant)

LIENS

- Thèse: <https://hdl.handle.net/1866/40944>
- Site web: <https://mathieuduchesneau.ca>
- LinkedIn: <https://www.linkedin.com/in/mathieu-duchesneau-047a0b307>
- GitHub: <https://github.com/duchesneumathieu>