Christos Tsirigotis

Ph.D. Student

Montréal, QC, Canada | +1 514 679 8945 | tsirif@gmail.com
Google Scholar | GitHub | LinkedIn

Profile

Ph.D. student who focuses in making multimodal and retrieval models work better and more efficiently. I tackle this from two angles: 1. improving robustness, retrieval, and multi-modal tokenization, via novel representation learning algorithms (logit adjustment, graded relevance loss), and 2. optimizing the underlying systems (Triton kernels, large-batch gradient caching for contrastive learning) to make training feasible at scale in multi-node clusters. Publications at **NeurIPS, ICLR, ICML, COLM**. Strong at data curation, throughput/cost optimization, memory efficiency, and reproducibility.

Experience

ServiceNow Research — Visiting Researcher Aug 2022 – Oct 2024
Montréal, Canada

- Drove research on robust generalization, developing a **logit adjustment** method that improved **worst-group accuracy** on compositional OOD benchmarks and standardized robust classification evaluation protocols.
- Scaled dense-retrieval fine-tuning on **LLM-graded relevance** data across multi-node clusters (64 H100s); achieved **higher nDCG@10** with **reduced annotation costs** compared to standard data and InfoNCE/DPR-based training pipelines.

HEC Montréal – NextAI — Scientist-in-Residence Apr 2022 – Aug 2022
Montréal, Canada

- Mentored 6 AI startups on MLOps, model and product fit, **data pipelines**, and **scaling evaluation**.

Université de Montréal — Teaching Assistant (IFT6135) Jan 2021 – Apr 2021
Montréal, Canada
Representation Learning (Prof. Aaron Courville).

Mila, Université de Montréal — Research & SWE Intern Oct 2017 – May 2018
Montréal, Canada

- Contributed to the design of **Oríon**, an open-source distributed experimentation and HPO software supervised by Frédéric Bastien. Open-source release on GitHub. Official documentation website.
- Research on SGD generalization (*A Walk with SGD* – supervised by Yoshua Bengio).

Theano (Google Summer of Code) — Software Developer May 2016 – Aug 2016

- Implemented **multi-GPU collectives** and distributed optimizers for the first deep learning software framework with differentiable backpropagation support.

Pandora Robotics, Aristotle University of Thessaloniki — Software Engineer Jan 2014 – Oct 2015
Thessaloniki, Greece

- Served as **Software Lead (2015)**, developing the autonomy stack for a UGV that won **2nd place** at the 2015 RoboCup Rescue League. Key points: sensor data fusion and mapping, reinforcement learning (RL) based kinodynamic modeling, novelty detection and exploration. Open-source software stack on Github.

Selected Publications

BiXSE: Improving Dense Retrieval via Probabilistic Graded Relevance Distillation COLM 2025
Tsirigotis C., Adlakha V., Monteiro J., Courville A., Taslakian P.

- A pointwise ranking loss that outperforms InfoNCE approaches by **+3-10% in nDCG@10** via leveraging LLM-graded relevance data to train, not only filter. Decrease in annotation costs. Code release on GitHub.

FLAM: Frame-Wise Language-Audio Modeling ICML 2025
Wu Y., *Tsirigotis C.*, Chen K., Huang C.Z.A., Courville A., Nieto O., Seetharaman P., Salamon J.

- **State-of-the-art open-vocabulary sound-event detection; impact at Adobe**. My core contribution was to propose and develop a novel *logit adjustment* training protocol for multi-modal contrastive learners, key for **+16-23% improvements in AUROC** for open-set detection. Official project webpage.

Group Robust Classification Without Any Group Information NeurIPS 2023
Tsirigotis C., Monteiro J., Rodriguez P., Vazquez D., Courville A.

- Derived a novel *logit adjustment* modeling technique for training on datasets with spurious correlations; improvements in compositional generalization. Code release on GitHub.

Simplicial Embeddings in Self-Supervised Learning and Downstream Classification ICLR 2023
Lavoie S., *Tsirigotis C.*, Schwarzer M., Vani A., Noukhovitch M., Kawaguchi K., Courville A.

- Oral - top 25%
- Constrain SSL features onto a product of simplices (SEM). Induce structures sparsity, **improves +4-5% accuracy** in downstream image classification tasks.

Full list: Google Scholar

Education

PhD in Computer Science (Expected: **April 2026**) — Université de Montréal, Montréal, Canada
Research: Scalable learning for zero-shot robust generalization – Supervisor: Aaron Courville – Affiliation: Mila

Diploma in Electrical & Computer Engineering — Aristotle University of Thessaloniki, Greece
GPA: 8.96/10 (Rank 4/205) – Thesis: Stabilizing GAN Training Without Gradient Penalties – Supervisor: Pericles A. Mitkas

Engineering Skills

Distributed training: PyTorch; DDP, FSDP, ZeRO; grad checkpointing and caching; NCCL; mixed precision; SLURM and Kubernetes; multi-node clusters; object storage.
Systems and performance: CUDA; Triton; nsys and ncu; checkpoint, IO, memory optimization.
Retrieval: FAISS; IVF; HNSW; OPQ/PQ; multi-vector retrieval; re-ranking; large-batch contrastive training.
Frameworks: PyTorch (primary), JAX, NumPy / SciPy; W&B.
Data curation: Large-scale synthetic data generation; similarity search and LLM-as-a-judge pipelines for denoising, filtering, and decontamination; WebDataset; Arrow and Parquet; active learning.
Hyperparameter optimization: Scaling-laws-aware zeroth-order search; batch-size and schedulers tuning; data mixtures; compute-constrained Chinchilla-style tradeoffs.

Awards

Mitacs Accelerate Scholarship (2024).
2nd place, Best-In-Class Autonomy – RoboCup Rescue League (2015).

Languages

Greek (native); English (fluent; TOEFL iBT 113/120); French (intermediate); Italian (elementary).