

Tristan Deleu

6666 St-Urbain

H2S 3H1 Montréal, QC – Canada

<https://tristandeleu.github.io>

Education

2017 - present	Université de Montréal / Mila	Montréal, QC, Canada
2018 - pres.	Ph.D. Computer Science - Artificial Intelligence	
Courses	<i>Advanced Structured Prediction and Optimization, Autonomous vehicles</i>	
Research	<i>Meta-learning, Reinforcement learning, Few-shot learning, Probabilistic Graphical Models, Structure learning, Causality</i>	
2017 - 2018	M.Sc Computer Science - Artificial Intelligence	
Courses	<i>Introduction to Machine Learning, Probabilistic Graphical Models, Representation Learning, Reinforcement Learning</i>	
2011 - 2015	Ecole Normale Supérieure de Cachan	Cachan, France
2013 - 2014	Master 2 - Applied Mathematics (M.Sc) , with highest honors	
	Master <i>Mathematics, Computer Vision and Machine Learning (MVA, ENS Cachan)</i>	
Courses	<i>Introduction to Statistical Learning, Reinforcement Learning, Probabilistic Graphical Models, Simulation-based Learning, Kernel Methods, Convex Optimization, Discrete Optimization, Advanced Learning for Graph & Text Data</i>	
2012 - 2013	Master 1 - Mathematics , with high honors	
	ENS Cachan & Université Paris Diderot VII	
Courses	<i>Functional Analysis, Complex Analysis, Optimization, Stochastic Processes, Statistical Learning, Combinatorics, Advanced Graph Theory, Probabilistic Aspects of Computer Science, Distributed Algorithmic, Game Theory, Advanced Probability, Statistics, Bayesian Statistics</i>	
2011 - 2012	Licence 3 - Mathematics (B.Sc) , with highest honors	
Courses	<i>Measure Theory & Probability, Differential Calculus & Topology, Algebra, Fourier & Hilbert Analysis, Algorithmic, Graph Theory</i>	
2011 - 2012	Licence 3 - Computer Science (B.Sc) , with highest honors	
Courses	<i>Logic & Calculability, Programming, Databases, Complexity, Machine Learning, Advanced Logic, Lambda Calculus, Formal Languages</i>	

Professional Experience

2022	Google Research , Student Researcher	Mountain View, CA, USA
Jun 22 - Aug 22	Project: <i>Combining Zero-shot & Few-shot models</i>	
	Description: <i>Evaluation of zero-shot & few-shot models based on large-scale pre-trained encoders on the Meta-Dataset benchmark. Creation of a more challenging episodic few-shot classification benchmark, based on the Open Images dataset. Study of multiple strategies to combine zero-shot & few-shot models, including an empirically effective regularization method. Design of a bi-level optimization problem to automatically meta-learn the regularization constants.</i>	
2015 - 2017	Snips , Researcher	Paris, France
Oct 16 - Jul 17	Project: <i>Query parsing with Convolutional Neural Networks</i>	
	Description: <i>Design of a Convolutional Neural Network (CNN) and Conditional Random Field-based (CRF) model to parse queries in natural language (intent classification & slot filling) on 10 preset domains (including "Place details", "Restaurant reservation", "Directions", etc...). Collaboration with engineering teams to run inference on mobile devices (see Snips SDK).</i>	

Sep 15 - Oct 16	<p>Project: <i>Memory-augmented Neural Networks applied to Reasoning</i></p> <p>Description: <i>Literature review on memory-augmented networks. Implementation of a Neural Turing Machine (NTM) library in Theano/Lasagne. Reproduction of most of the results on the algorithmic tasks from the original paper, with similar generalization performances. Application of NTMs to the problem of learning matching parentheses (Dyck words). Study of the bAbI dataset and generation of similar unbiased reasoning tasks based on rejection sampling & application of NTMs on these Q&A tasks (graph data).</i></p>	
2014	<p>Snips, <i>Research Internship</i></p>	Paris, France
Apr 14 - Jan 15	<p>Project: <i>Urban context modeling with Dynamic Bayesian Networks</i></p> <p>Supervisors: <i>Dr. Maël Primet, Dr. Rand Hindi</i></p> <p>Description: <i>Theory and design of a hierarchical graphical model to gather contextual data from multiple sources, and infer informations about a user. Design of a 2-stage classifier based on a Random Forest and a Hidden Markov Model to detect the mode of transportation online and on-device from raw accelerometer data.</i></p>	
2013	<p>University College London, <i>Research Internship</i></p>	London, UK
Jun 13 - Aug 13	<p>Project: <i>Hierarchical Representations using Deep Belief Networks</i></p> <p>Supervisor: <i>Dr. David Barber (CSML/UCL)</i></p> <p>Description: <i>Theory of Deep Belief Networks, Gibbs sampling-based learning, and inference algorithms. Design of two autoencoder-based learning algorithms, approaching the state of the art in terms of data reconstruction: the Neural Principal Component Analysis and the Implicit Neural Network. Evaluation of these algorithms on shallow neural networks on the MNIST dataset.</i></p>	
2012	<p>ENS Cachan & Université Paris Sud XI, <i>Research Internship</i></p>	Orsay, France
Feb 12 - Jun 12	<p>Project: <i>Optimal Transport: Theory and Continuity Optimization</i></p> <p>Supervisor: <i>Pr. Filippo Santambrogio (P11)</i></p> <p>Description: <i>Theory of distributions and optimal transport. Design of an evolutionary algorithm for minimizing the oscillation in transport problems, with applications to privacy addressed by the uniform encoding of geolocation data.</i></p>	

Publications

Nov 2022	<p>Mizu Nishikawa-Toomey*, Tristan Deleu*, Jithendaraa Subramanian, Yoshua Bengio, Laurent Charlin, <i>Bayesian learning of Causal Structure and Mechanisms with GFlowNets and Variational Bayes</i> https://arxiv.org/abs/2211.02763</p>
Oct 2022	<p>Jithendaraa Subramanian, Yashas Annadani, Ivaxi Sheth, Nan Rosemary Ke, Tristan Deleu, Stefan Bauer, Derek Nowrouzezahrai, Samira Ebrahimi Kahou, <i>Learning Latent Structural Causal Models</i> https://arxiv.org/abs/2210.13583</p>
Oct 2022	<p>Nikolay Malkin*, Salem Lahlou*, Tristan Deleu*, Xu Ji, Edward Hu, Katie Everett, Dinghui Zhang, Yoshua Bengio, <i>GFlowNets and variational inference</i> https://arxiv.org/abs/2210.00580</p>
Aug 2022	<p>Tristan Deleu*, Akram Erraqabi*, Ramnath Kumar, Yoshua Bengio, <i>Task-Agnostic Initialization in Gradient-Based Meta-Learning</i>, CoLLAs 2022 Workshop Track</p>
Aug 2022	<p>Tristan Deleu, António Góis, Chris Emezue, Mansi Rankawat, Simon Lacoste-Julien, Stefan Bauer, Yoshua Bengio, <i>Bayesian Structure Learning with Generative Flow Networks</i>, UAI 2022 https://arxiv.org/abs/2202.13903</p>
Apr 2022	<p>Tristan Deleu, David Kanaa, Leo Feng, Giancarlo Kerg, Yoshua Bengio, Guillaume Lajoie, Pierre-Luc Bacon, <i>Continuous-Time Meta-Learning with Forward Mode Differentiation</i>, ICLR 2022 (Spotlight) https://arxiv.org/abs/2203.01443</p>
Dec 2021	<p>Ramnath Kumar, Tristan Deleu, Yoshua Bengio, <i>The Effect of Diversity in Meta-Learning</i>, Workshop on Meta-Learning, NeurIPS 2021 https://arxiv.org/abs/2201.11775</p>

- Nov 2021 Yoshua Bengio*, Salem Lahlou*, **Tristan Deleu***, Edward J. Hu, Mo Tiwari, Emmanuel Bengio, *GFlowNet Foundations*
<https://arxiv.org/abs/2111.09266>
- Apr 2021 Yoshua Bengio, Prateek Gupta, Tegan Maharaj, Nasim Rahaman, Martin Weiss, **Tristan Deleu**, Eilif Muller, Meng Qu, Victor Schmidt, Pierre-Luc St-Charles, Hannah Alsdurf, et al., *Predicting Infectiousness for Proactive Contact Tracing*, ICLR 2021 (**Spotlight**)
<https://arxiv.org/abs/2010.12536>
- Feb 2021 **Tristan Deleu**, Yoshua Bengio, *Structured Sparsity Inducing Adaptive Optimizers for Deep Learning*
<https://arxiv.org/abs/2102.03869>
- May 2020 Hannah Alsdurf, Edmond Belliveau, Yoshua Bengio, **Tristan Deleu**, Prateek Gupta, Daphne Ippolito, Richard Janda, Max Jarvie, Tyler Kolody, Sekoul Krastev, Tegan Maharaj, et al., *COVI White Paper*
<https://arxiv.org/abs/2005.08502>
- Feb 2020 Bhairav Mehta, **Tristan Deleu**, Sharath Chandra Raparthy, Chris J. Pal, Liam Paull, *Curriculum in Gradient-Based Meta-Reinforcement Learning*
<https://arxiv.org/abs/2002.07956>
- Oct 2019 Mandana Samiei, Tobias Würfl, **Tristan Deleu**, Martin Weiss, Francis Dutil, Thomas Fevens, Geneviève Boucher, Sebastien Lemieux, Joseph Paul Cohen, *The TCGA Meta-Dataset Clinical Benchmark*
<https://arxiv.org/abs/1910.08636>
- Sep 2019 **Tristan Deleu**, Tobias Würfl, Mandana Samiei, Joseph Paul Cohen, Yoshua Bengio, *Torchmeta: A Meta-Learning library for PyTorch*
<https://arxiv.org/abs/1909.06576>
- Jun 2019 Shagun Sodhani, Anirudh Goyal, **Tristan Deleu**, Yoshua Bengio, Sergey Levine, Jian Tang, *Learning Powerful Policies by Using Consistent Dynamics Model*, RLDM 2019, Montreal, Canada
<https://arxiv.org/abs/1906.04355>
- Jun 2019 Sébastien Lachapelle, Philippe Brouillard, **Tristan Deleu**, Simon Lacoste-Julien, *Gradient-Based Neural DAG Learning*, ICLR 2020, Addis Ababa, Ethiopia
<https://arxiv.org/abs/1906.02226>
- Jan 2019 Yoshua Bengio, **Tristan Deleu**, Nasim Rahaman, Rosemary Ke, Sébastien Lachapelle, Olexa Bilaniuk, Anirudh Goyal, Christopher Pal, *A Meta-Transfer Objective for Learning to Disentangle Causal Mechanisms*, ICLR 2020, Addis Ababa, Ethiopia
<https://arxiv.org/abs/1901.10912>
- Dec 2018 **Tristan Deleu**, Yoshua Bengio, *The effects of negative adaptation in Model-Agnostic Meta-Learning*, Workshop on Meta-Learning, NeurIPS 2018, Montreal, Canada
<https://arxiv.org/abs/1812.02159>
- Jul 2018 **Tristan Deleu**, Simon Guiroy, Seyedarian Hosseini, *On the reproducibility of gradient-based Meta-Reinforcement Learning baselines*, Reproducibility in Machine Learning Workshop, ICML 2018, Stockholm, Sweden
- Dec 2016 **Tristan Deleu**, Joseph Dureau, *Learning Operations on a Stack with Neural Turing Machines*, 1st Workshop on Neural Abstract Machines & Program Induction (NAMPI), NIPS 2016, Barcelona, Spain
<https://arxiv.org/abs/1612.00827>

Awards & Scholarships

- 2021 **Sony Research Award Program**
Project: *The learning dynamics of meta-learning*
- 2019 **Antidote scholarship 2019**
Druide informatique
- Oct 2019 **Best in Show award – Global PyTorch Summer Hackathon 2019**
1st place for the project Torchmeta

Teaching

- Fall 2022 **Université de Montréal – Probabilistic Graphical Models**
Teaching assistant
- Fall 2021 **Université de Montréal – Probabilistic Graphical Models**
Teaching assistant
- Winter 2021 **Université de Montréal – Representation Learning**
Teaching assistant

Talks

- Oct 2019 **PyTorch Developer Conference** San Francisco, US
Title: *Torchmeta: A Meta-Learning library for PyTorch*
Description: *Presentation of the library Torchmeta as part of the Best in Show award at the Global PyTorch Summer Hackathon.*

Projects

- Oct 2019 **Model-Agnostic Meta-Learning**
Link: <https://github.com/tristandeleu/pytorch-maml>
Description: *Implementation of the “Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks” (MAML) paper in PyTorch and Torchmeta. Applications to few-shot supervised learning tasks.*
- Sep 2019 **Torchmeta: A Meta-Learning library for PyTorch**
Link: <https://github.com/tristandeleu/pytorch-meta>
Description: *A collection of extensions and data-loaders for few-shot learning & meta-learning in PyTorch. It provides a unified for a wide range of standard benchmarks, including Omniglot, Mini-ImageNet, Tiered-ImageNet, CIFAR-FS, Fewshot-CIFAR100.*
- May 2018 **Neural Discrete Representation Learning**
Link: <https://github.com/ritheshkumar95/pytorch-vqvae>
Description: *Implementation of the “Neural Discrete Representation Learning” (VQ-VAE) paper in PyTorch. Applications to computer vision problems (MNIST, CIFAR10 and Imagenet).*
- Apr 2018 **Meta-Learning for Reinforcement Learning**
Link: <https://github.com/tristandeleu/pytorch-maml-rl>
Description: *Implementation of the “Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks” (MAML) paper in PyTorch. Applications to Bandits, Tabular Markov Decision Processes and Continuous Control problems.*
- Dec 2017 **Memory Architectures in Recurrent Neural Network Language Models**
Link: <https://github.com/arianhosseini/MemArchs-in-RNNLM>
Description: *Implementation of the “Memory Architectures in Recurrent Neural Network Language Models” paper in PyTorch, as part of the ICLR 2018 Reproducibility Challenge.*
- Dec 2017 **Disentanglement in Variational Autoencoders & β -VAEs**
Link: <https://github.com/nithin127/nest-vae>
Description: *Study of the disentanglement properties of VAEs and β -VAEs, qualitatively and quantitatively with a novel disentanglement metric.*
- May 2016 **One-shot Learning with Memory-Augmented Neural Networks**
Link: <https://github.com/tristandeleu/ntm-one-shot>
Description: *Implementation of the “One-shot Learning with Memory-Augmented Neural Networks” paper in Theano, with experiments on the Omniglot dataset.*

- Feb 2016 **NTM-Lasagne: Neural Turing Machines in Theano/Lasagne**
 Link: <https://github.com/snipsco/ntm-lasagne>
 Description: *Implementation of the “Neural Turing Machines” paper in Lasagne/Theano as a Lasagne layer. Source and pre-trained models to reproduce the original experiments on algorithmic tasks (Copy, Repeat Copy, Associative Recall, Dyck words). Visualization of the heads & animations of the memory with Matplotlib.*
- Dec 2015 **Neural GPU**
 Link: <https://github.com/tristandeleu/neural-gpu>
 Description: *Implementation of the “Neural GPU” paper in Theano, with experiment on the duplication task. Animation of inference with Matplotlib.*
- Mar 2015 **Switching Kalman Filters**
 Link: <https://github.com/tristandeleu/switching-kalman-filter>
 Description: *Implementation of the Switching Kalman Filter (filtering and smoothing algorithms) in Python. Detection of stop points from GPS trajectories (AXA Driver Telematics Analysis dataset) using a Continuous Wiener Process Acceleration/Brownian Motion SKF.*
- Jan 2014 **Bandit algorithm for online multi-class prediction**
 Description: *Study of a multi-class prediction problem with a bandit feedback in the fully adversarial setting. Theory & implementation in Matlab of two efficient algorithms minimizing the regret over time: the Neutron and the Banditron.*
- Dec 2013 **Topic modeling**
 Description: *Review of multiple existing probabilistic topic modeling algorithms. Theory & implementation in Matlab of the Latent Dirichlet Allocation, with variational inference and Gibbs sampling. Empirical comparison of these two inference algorithms on the NIPS full papers dataset.*
- May 2012 **Active learning-based Web search engine**
 Description: *Implementation in Python of a web crawler with user-feedback to focus on a given query. Design of a semi-supervised classification algorithm based on Random Forests.*

Languages & Skills

French	Native
English	Fluent (Oral & Written)
Computer Skills	JAX, L ^A T _E X, Python, PyTorch, Rust
Github	https://github.com/tristandeleu/
Mila	Meta-Learning Reading Group organizer