

Nicolas Le Roux

Researcher

✉ nicolas@le-roux.name

🌐 <http://nicolas.le-roux.name>

Professional experience

2017–Now **Research scientist**, *Google Brain Montreal*, Canada.

- Large-scale optimization
- Variance reduction methods
- Reinforcement learning

2017–Now **Adjunct professor**, *McGill*, Canada.

2012–2017 **Research lead**, *Criteo*, France.

- Created and grew the Paris research team
 - Defined the interactions with the other teams in the R&D
 - Defined the recruiting process and recruited researchers
 - Mentored researchers
 - Led weekly meetings with remote research teams to favor communication
 - Led the quarterly discussions on projects and deliverables
 - Scientific point of contact for the rest of the company
 - Organized a workshop to increase awareness of Criteo's scientific problems
- Ensured proper interactions between the business and research teams
 - Defined the short-term and long-term scientific roadmaps
 - Led projects yielding additional revenue of several million dollars per year (distributed learning, improved product recommendation, feature selection)
 - Communicated about the research achievements to the rest of the company

2010–2012 **Postdoc at Inria**, *École Normale Supérieure de Paris*, France.

- Convex optimization
- Metric learning

Summer 2010 **Invited researcher at the Courant Institute**, *New-York University*, USA.

2008–2010 **Postdoc at Microsoft Research**, *Cambridge*, United Kingdom.

- Large-scale optimization
- Generative model of images

2004–2008 **PhD at the LISA lab**, *University of Montreal*, Canada.

- Theoretical and practical aspects of neural networks
- Large-scale optimization

Education and awards

2018 **Lagrange prize in continuous optimization**, *SIAM*.

2010 **Excellence scholarship (declined)**, *CIFAR*, Canada.

2008–2010 **Microsoft Research Fellowship**, *Darwin College*, Cambridge, UK.

2004–2008 **PhD in machine learning**, *University of Montreal*, Canada.

2002–2003 **MSc. in mathematics, vision and learning**, *ENS Cachan*.

2000–2003 **MSc. in applied mathematics**, *École Centrale Paris*.

Miscellaneous

Cycling 3800 miles from Vancouver to the Arctic Ocean: <http://www.arctic2007.org>
1500 kilometres in New Zealand

Research duties

- Reviewer **NeurIPS, ICML, ICLR, JMLR, PNAS, CVPR, Neural Computation.**
- Area chair **ICML, ICLR, NeurIPS.**
- Associate editor **TPAMI.**
- Organizer **Montreal AI Symposium, 2018.**
- Organizer **Deep Learning Workshop, NIPS 2011.**
- Creator **Machine learning in the real world yearly Workshop, Criteo.**
- Organizer **Machine Learning reading group, MSR Cambridge.**

Teaching

- 2005 **Learning algorithms (TA), MSc., University of Montreal.**
- 2009 **Optimization, MSc., Gatsby Computational Neuroscience Unit, 3 hours.**
- 2010–2016 **Neural networks and optimization, MSc., ENS Cachan, 3 hours.**
- 2012 **Introduction to machine learning, MSc., ENS Ulm, 9 hours.**
- 2015 **Neural networks and optimization, MSc., Télécom ParisTech, 3 hours.**
- 2016 **ML for the industry, Machine Learning Summer School, Cádiz, 3 hours.**
- 2017 **Optimization, Machine Learning Summer School, Montreal, 1.5 hours.**
- 2018 **Optimization, Machine Learning Winter School, Montreal, 1.5 hours.**

Patents

- 2012 **Data processing using restricted Boltzmann machines.**
N. Le Roux, J. Winn and J. Shotton
US Patent 8,239,336
- 2012 **Image processing using masked restricted Boltzmann machines.**
N. Le Roux, J. Winn, J. Shotton and N. Heess
US Patent 8,229,221

Publications

Zafarali Ahmed, Nicolas Le Roux, Mohammad Norouzi, and Dale Schuurmans. Understanding the impact of entropy in policy learning. In *International Conference on Machine Learning*, 2019.

Guillaume Alain, Nicolas Le Roux, and Pierre-Antoine Manzagol. Negative eigenvalues of the hessian in deep neural networks. In *International Conference on Learning Representations Workshop*, 2019.

Francis Bach, Rodolphe Jenatton, Julien Mairal, Guillaume Obozinski, Martin Andersen, Joachim Dahl, Zhang Liu, Lieven Vandenberghe, Dimitri Bertsekas, Anatoli Juditsky, et al. Optimization for machine learning, 2011.

Marc G Bellemare, Will Dabney, Robert Dadashi, Adrien Ali Taiga, Pablo Samuel Castro, Nicolas Le Roux, Dale Schuurmans, Tor Lattimore, and Clare Lyle. A geometric perspective on optimal representations for reinforcement learning. *arXiv preprint arXiv:1901.11530*, 2019.

Marc G Bellemare, Nicolas Le Roux, Pablo Samuel Castro, and Subhodeep Moitra. Distributional reinforcement learning with linear function approximation. In *International Conference on Artificial Intelligence and Statistics*, 2019.

Y Bengio, O Delalleau, and NL Roux. A discussion of semi-supervised learning and transductive, 2006.

Y Bengio, O Delalleau, and NL Roux. Semi-supervised learning, chapter label propagation and quadratic criterion, 2007.

Yoshua Bengio, Olivier Delalleau, and Nicolas Le Roux. The curse of dimensionality for local kernel machines. *Techn. Rep.*, 1258, 2005.

Yoshua Bengio, Olivier Delalleau, and Nicolas Le Roux. The curse of highly variable functions for local kernel machines. In *NIPS*, volume 18, page 107. MIT; 1998, 2006.

Yoshua Bengio, Olivier Delalleau, and Nicolas Le Roux. Label propagation and quadratic criterion. *Semi-supervised learning*, pages 193–216, 2006.

Yoshua Bengio, Olivier Delalleau, Nicolas Le Roux, Jean-François Paiement, Pascal Vincent, and Marie Ouimet. Learning eigenfunctions links spectral embedding and kernel pca. *Neural Computation*, 16(10):2197–2219, 2004.

Yoshua Bengio, Olivier Delalleau, Nicolas Le Roux, Jean-François Paiement, Pascal Vincent, and Marie Ouimet. Spectral dimensionality reduction. In *Feature Extraction*, pages 519–550. Springer Berlin Heidelberg, 2006.

Yoshua Bengio, Nicolas Le Roux, Pascal Vincent, Olivier Delalleau, and Patrice Marcotte. Convex neural networks. In *NIPS*, pages 123–130, 2005.

Yoshua Bengio, Jean-François Paiement, Pascal Vincent, Olivier Delalleau, Nicolas Le Roux, and Marie Ouimet. Out-of-sample extensions for lle, isomap, mds, eigenmaps, and spectral clustering. In *NIPS*, volume 16, pages 177–184. Cambridge, MA, USA: MIT Press, 2004.

- Yoshua Bengio, Pascal Vincent, Jean-François Paiement, Olivier Delalleau, Marie Ouimet, and Nicolas Le Roux. *Spectral clustering and kernel PCA are learning eigenfunctions*, volume 1239. CIRANO, 2003.
- Y-Lan Boureau, Nicolas Le Roux, Francis Bach, Jean Ponce, and Yann LeCun. Ask the locals: multi-way local pooling for image recognition. In *ICCV*, 2011.
- Clément Calauzènes and Nicolas Le Roux. Distributed saga: Maintaining linear convergence rate with limited communication. *arXiv preprint arXiv:1705.10405*, 2017.
- Robert Dadashi, Adrien Ali Taïga, Nicolas Le Roux, Dale Schuurmans, and Marc G Bellemare. The value function polytope in reinforcement learning. In *International Conference on Machine Learning*, 2019.
- Oliver Delalleau, Yoshua Bengio, and Nicolas Le Roux. Large-scale algorithms. *Semi-supervised learning*, pages 333–341, 2006.
- Olivier Delalleau, Yoshua Bengio, and Nicolas Le Roux. Efficient non-parametric function induction in semi-supervised learning. In *AISTATS*, volume 27, page 100, 2005.
- Robert M Gower, Nicolas Le Roux, and Francis Bach. Tracking the gradients using the hessian: A new look at variance reducing stochastic methods. In *International Conference on Artificial Intelligence and Statistics*, 2018.
- Nicolas Heess, Nicolas Le Roux, and John Winn. Weakly supervised learning of foreground-background segmentation using masked rbms. In *International Conference on Artificial Neural Networks*, pages 9–16. Springer Berlin Heidelberg, 2011.
- Rodolphe Jenatton, Nicolas Le Roux, Antoine Bordes, and Guillaume R Obozinski. A latent factor model for highly multi-relational data. In *NIPS*, pages 3167–3175, 2012.
- Nicolas Le Roux. *Avancées théoriques sur la représentation et l’optimisation des réseaux de neurones*. 2008.
- Nicolas Le Roux. Using gradient descent for optimization and learning. Technical report, Technical report, University of Cambridge, 2009.
- Nicolas Le Roux. Efficient iterative policy optimization. *arXiv preprint arXiv:1612.08967*, 2016.
- Nicolas Le Roux. Tighter bounds lead to improved classifiers. In *International Conference on Learning Representations*, 2017.
- Nicolas Le Roux. Anytime tail averaging. *arXiv preprint arXiv:1902.05083*, 2019.
- Nicolas Le Roux, Reza Babanezhad, and Pierre-Antoine Manzagol. Online variance-reducing optimization. In *International Conference on Learning Representations Workshop*, 2018.
- Nicolas Le Roux and Francis Bach. Local component analysis. In *International Conference on Learning Representations*, 2013.

- Nicolas Le Roux and Yoshua Bengio. Continuous neural networks. In *International Conference on Artificial Intelligence and Statistics*, pages 404–411, 2007.
- Nicolas Le Roux and Yoshua Bengio. Representational power of restricted boltzmann machines and deep belief networks. *Neural Computation*, 20(6):1631–1649, 2008.
- Nicolas Le Roux and Yoshua Bengio. Deep belief networks are compact universal approximators. *Neural computation*, 22(8):2192–2207, 2010.
- Nicolas Le Roux, Yoshua Bengio, and Andrew Fitzgibbon. Improving first and second-order methods by modeling uncertainty. *Optimization for Machine Learning*, pages 403–429, 2011.
- Nicolas Le Roux, Yoshua Bengio, Pascal Lamblin, Marc Joliveau, and Balázs Kégl. Learning the 2-d topology of images. In *Advances in Neural Information Processing Systems*, pages 841–848, 2008.
- Nicolas Le Roux and Andrew W Fitzgibbon. A fast natural newton method. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, pages 623–630, 2010.
- Nicolas Le Roux, Nicolas Heess, Jamie Shotton, and John Winn. Learning a generative model of images by factoring appearance and shape. *Neural Computation*, 23(3):593–650, 2011.
- Nicolas Le Roux, Hugo Larochelle, and Yoshua Bengio. Discriminative training of rbms using bhattacharyya distance. In *Learning Workshop, Cliff Lodge, Snowbird, Utah*, 2008.
- Nicolas Le Roux, Pierre-Antoine Manzagol, and Yoshua Bengio. Topmoumoute online natural gradient algorithm. In *NIPS*, pages 849–856. MIT Press, 2008.
- Nicolas Le Roux, Mark Schmidt, and Francis Bach. A stochastic gradient method with an exponential convergence rate for finite training sets. In *NIPS*, pages 2663–2671. Curran Associates Inc., 2012.
- Romain Lerallut, Diane Gasselin, and Nicolas Le Roux. Large-scale real-time product recommendation at criteo. In *Proceedings of the 9th ACM Conference on Recommender Systems*, pages 232–232. ACM, 2015.
- Thomas Nedelec, Nicolas Le Roux, and Vianney Perchet. A comparative study of counterfactual estimators. In *NIPS Workshop on Causal Inference and Machine Learning for Intelligent Decision Making*, 2017.
- Mark Schmidt and Nicolas Le Roux. Fast convergence of stochastic gradient descent under a strong growth condition. *arXiv preprint arXiv:1308.6370*, 2013.
- Mark Schmidt, Nicolas Le Roux, and Francis Bach. Convergence rates of inexact proximal-gradient methods for convex optimization. In *NIPS*, 2011.
- Mark Schmidt, Nicolas Le Roux, and Francis Bach. Minimizing finite sums with the stochastic average gradient. *Mathematical Programming*, 162(1-2):83–112, 2017.