

# Nicolas Le Roux

Research lead

☎ +33 6 64 81 70 48

✉ [nicolas@le-roux.name](mailto:nicolas@le-roux.name)

🌐 <http://nicolas.le-roux.name>

## Professional experience

2012–Present **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 candidate at the LISA lab**, *University of Montreal*, Canada.

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

## Education and awards

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.

- PhD thesis elected amongst the top 5% of the discipline

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

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

## Miscellaneous

Programming MATLAB, Python

Cycling 3800 miles from Vancouver to the Arctic Ocean: <http://www.arctic2007.org>

1500 kilometres in New Zealand

---

## Research duties

- Reviewer **NIPS, ICML, ICLR, JMLR, PNAS, CVPR, Neural Computation.**  
Organizer **Deep Learning Workshop, NIPS 2011.**  
Area chair **ICML 2015, ICLR 2015, ICLR 2016.**  
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.  
2010–2015 **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.

---

## 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

- Numerical Optimization **Minimizing Finite Sums with the Stochastic Average Gradient.**  
M. Schmidt, **N. Le Roux** and F. Bach  
arXiv:1309.2388.

**Fast Convergence of Stochastic Gradient Descent under a Strong Growth Condition.**

M. Schmidt and **N. Le Roux**  
arXiv:1308.6370.

**A Stochastic Gradient Method with an Exponential Convergence Rate for Strongly-Convex Optimization with Finite Training Sets.**

**N. Le Roux**, M. Schmidt and F. Bach  
*NIPS 25*, 2012 (oral presentation).

**Convergence Rates of Inexact Proximal-Gradient Methods for Convex Optimization.**

M. Schmidt, **N. Le Roux** and F. Bach  
*NIPS*, 2011.

**Improving first and second-order methods by modeling uncertainty.**

**N. Le Roux**, Y. Bengio and A. Fitzgibbon  
Book chapter, *Optimization for Machine Learning*, MIT Press, Cambridge, MA, USA, 2011  
Edited by S. Sra, S. Nowozin and S.J. Wright.

**A fast natural Newton method.**

**N. Le Roux** and A. Fitzgibbon  
*ICML*, 2010.

**Topmoumoute Online Natural Gradient Algorithm.**

**N. Le Roux**, P.A. Manzagol and Y. Bengio  
*NIPS 20*, 2008.

Computer **Ask the locals: multi-way local pooling for image recognition.**

Vision Y.-L. Boureau, **N. Le Roux**, F. Bach, J. Ponce and Y. LeCun  
*ICCV*, 2011.

**Weakly Supervised Learning of Foreground-Background Segmentation using Masked RBMs.**

N. Heess, **N. Le Roux** and J. Winn  
*ICANN*, 2011.

**Learning a generative model of images by factoring appearance and shape.**

**N. Le Roux**, N. Heess, J. Shotton and J. Winn  
*Neural Computation*, March 2011, Vol. 23, No. 3, Pages 593-650.

**Learning the 2-D Topology of Images.**

**N. Le Roux**, Y. Bengio, P. Lamblin, M. Joliveau and B. Kegl  
*NIPS 20*, 2008.

Neural **Deep belief networks are Compact Universal Approximators.**

Networks **N. Le Roux** and Y. Bengio  
*Neural Computation*, August 2010, Vol. 22, No. 8, Pages 2192-2207.

**Avancées théoriques sur la représentation and l'optimisation des réseaux de neurones.**

**N. Le Roux**  
PhD thesis, University of Montreal, 2008.

**Representational Power of Restricted Boltzmann Machines and deep belief networks.**

**N. Le Roux** and Y. Bengio  
*Neural Computation*, June 2008, Vol. 20, No. 6, Pages 1631-1649.

**Continuous Neural Networks.**

**N. Le Roux** and Y. Bengio  
*AISTATS 11*, 2007.

**Convex Neural Networks.**

Y. Bengio, **N. Le Roux**, P. Vincent, O. Delalleau and P. Marcotte  
*NIPS 18*, 2006.

Representation **Local Component Analysis.**

Learning **N. Le Roux** and F. Bach  
*ICLR*, 2013.

**A latent factor model for highly multi-relational data.**

R. Jenatton, **N. Le Roux**, A. Bordes and G. Obozinski  
*NIPS 25*, 2012.

**Spectral Dimensionality Reduction.**

Y. Bengio, O. Delalleau, **N. Le Roux**, J.-F. Paiement, P. Vincent and M. Ouimet  
Book chapter, *Feature Extraction, Foundations and Applications*, Physica-Verlag,  
Springer, 2006  
Edited by I. Guyon and al.

**The Curse of Highly Variable Functions for Local Kernel Machines.**

Y. Bengio, O. Delalleau, and **N. Le Roux**  
*NIPS 18*, 2006.

**Out-of-Sample Extensions for LLE, Isomap, MDS, Eigenmaps, and Spectral Clustering.**

Y. Bengio, J.-F. Paiement, P. Vincent, O. Delalleau, **N. Le Roux** and M. Ouimet  
*NIPS 16*, 2004.

**Learning Eigenfunctions Links Spectral Embedding and Kernel PCA.**

Y. Bengio, O. Delalleau, **N. Le Roux**, J.-F. Paiement, M. Ouimet and P. Vincent  
*Neural Computation*, October 2004, Vol. 16, No. 10, Pages 2197-2219.

Semi- **Label propagation and quadratic criterion.**

Supervised Y. Bengio, O. Delalleau and **N. Le Roux**  
Learning Book chapter, *Semi-supervised learning*, MIT Press, Cambridge, MA, USA, 2006  
Edited by O. Chapelle, B. Schölkopf and A. Zien.

**Large-scale algorithms.**

O. Delalleau Y. Bengio, and **N. Le Roux**  
Book chapter, *Semi-supervised learning*, MIT Press, Cambridge, MA, USA, 2006  
Edited by O. Chapelle, B. Schölkopf and A. Zien.

**Efficient Non-Parametric Function Induction in Semi-Supervised Learning.**

O. Delalleau, Y. Bengio, and **N. Le Roux**  
*AISTATS 10*, 2005.