JAMES MARTENS

PRESENT ADDRESS
664 Bathurst St. Unit 4
Toronto, Ontario, M5S2R3
(416) 875-1543

PERMANENT ADDRESS
124 Hazelwood Ave.
Toronto, Ontario, M4J1K5,
(416) 463-2642

Email: james.martens@gmail.com
Homepage: http://www.cs.toronto.edu/~jmartens/

EDUCATION

2009 - 2015 University of Toronto, PhD. in Computer Science

Advisors: Richard Zemel and Geoffrey E. Hinton
Area: Machine Learning
Average: A+

2007 - 2009 University of Toronto, MSc. in Computer Science

Advisors: Geoffrey E. Hinton and Richard Zemel
Area: Machine Learning
Average: A+

2002 - 2007 University of Waterloo, Honors B.Math (Co-op)

Major areas: Computer Science, Pure Mathematics, and Combinatorics & Optimization
Advisor: Chris Eliasmith
Average: 93%

PROFESSIONAL EXPERIENCE

Jan 2004 - Aug 2004, Apr 2005 - Aug 2005 Developer/Researcher, Z-Tech Inc., Toronto

May 2003 - Aug 2003 Developer/Research Assistant, Descartes Systems Group, Waterloo

Oct 2001 - Aug 2002 Software Developer (part time), Environment Canada, Toronto

Oct 2001 - June 2002 Programmer (part time), University of Toronto, Electrical and Computer Engineering Dept.

HONORS, AWARDS & FELLOWSHIPS

Google PhD Fellowship in Machine Learning	(2012, $70,000 USD)
NSERC Canada Graduate Scholarship PhD (CGS-D)	(2009-2011, $70,000 CAD)
NSERC Canada Graduate Scholarship Masters (CGS-M)	(2007-2009, $35,000 CAD)
Fellowship of Massey College	(2007-2012)
NSERC Undergraduate Research Assistantship Award	(2006)
University of Waterloo National Entrance Scholarship	(2002-2006, $14,000 CAD)
Medalist at the National Canadian Computing Competition	(2002)

PUBLICATIONS

In preparation

1. Beyond Universality: On the Expressive Efficiency of Deep Models
   James Martens

2. A Krylov Subspace Method for Sampling in Very High Dimensional Spaces
   James Martens

In Submission / Preprints

3. A Kronecker-factored Approximate Fisher Matrix for Convolution Layers
   Roger Grosse, **James Martens**
   2016

4. Adding Gradient Noise Improves Learning for Very Deep Networks
   Arvind Neelakantan, Luke Vilnis, Quoc V. Le, Ilya Sutskever, Lukasz Kaiser, Karol Kurach, **James Martens**
   2016

5. On the Expressive Efficiency of Sum Product Networks
   James Martens, Venkatesh Medabalimi
   2015

6. New Insights and Perspectives on the Natural Gradient Method
   James Martens
   2015

Refereed Publications

7. Optimizing Neural Networks with Kronecker-factored Approximate Curvature
   James Martens, Roger Grosse
   In *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, 2015

8. On the Representational Efficiency of Restricted Boltzmann Machines
   James Martens, Arkadev Chattopadhyay, Toniann Pitassi, Richard Zemel
   In *Proceedings of the 27th Annual Conference Neural Information Processing Systems (NIPS)*, 2013

9. On the importance of momentum and initialization in deep learning
   Ilya Sutskever, **James Martens**, George Dahl, and Geoffery Hinton
   In *Proceedings of the 30th International Conference on Machine Learning (ICML)*, 2013

10. Training Deep and Recurrent Neural Networks with Hessian-Free Optimization
    James Martens, Ilya Sutskever
    In *Neural Networks: Tricks of the Trade*, Springer 2012

11. Estimating the Hessian by Back-propagating Curvature
James Martens, Ilya Sutskever, and Kevin Swersky
In *Proceedings of the 29th International Conference on Machine Learning (ICML)*, 2012

12. Learning Recurrent Neural Networks with Hessian-Free Optimization
James Martens, Ilya Sutskever
In *Proceedings of the 28th International Conference on Machine Learning (ICML)*, 2011

13. Generating Text with Recurrent Neural Networks
Ilya Sutskever, **James Martens**, Geoffrey Hinton
In *Proceedings of the 28th International Conference on Machine Learning (ICML)*, 2011

14. Normalization for probabilistic inference with neurons
Chris Eliasmith, **James Martens**
In *Biological Cybernetics*, 2011

15. Deep Learning via Hessian-free Optimization
James Martens
In *Proceedings of the 27th International Conference on Machine Learning (ICML)*, 2010

16. Learning the Linear Dynamical System with ASOS
James Martens
In *Proceedings of the 27th International Conference on Machine Learning (ICML)*, 2010

17. Parallelizable Sampling of Markov Random Fields
James Martens, Ilya Sutskever
In *Proceedings of Artificial Intelligence and Statistics (AISTATS)*, 2010

18. Novel Lead Configurations for Robust Bio-Impedance Acquisition
Joel Ironstone, Milan Graovac, **James Martens**, Martin Rozee, K.C. Smith
In *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2007

Non-refereed Presentation and Posters

19. A neurologically plausible implementation of statistical inference applied to random dot motion
James Martens, Chris Eliasmith
Poster, Computational Neuroscience (CNS-2007), 2007

20. Challenges for biological bayes: Solving normalization
Chris Eliasmith, **James Martens**
Presentation, COSYNE Workshop on Statistical Inference in the Brain, 2007

21. A biologically realistic model of statistical inference applied to random dot motion
James Martens, Chris Eliasmith
Poster, COSYNE, 2007

THESES

Second-order Optimization for Neural Networks
James Martens
Ph.D. Thesis, Dept. of Computer Science, University of Toronto

A New Algorithm for Rapid Parameter Learning in Linear Dynamical Systems
James Martens
Master's Thesis, Dept. of Computer Science, University of Toronto

PATENTS

Weighted gradient method and system for diagnosing disease
Graovac M., **Martens J.**, Pavlovic, Z., Ironstone, J.

Jul, 13 2006: US 20060151815

TEACHING ASSISTANTSHIPS

CSC 2515, Graduate Introduction to Machine Learning, taught by Richard Zemel (2010)

CSC 236, Introduction to the Theory of Computation, taught by Francois Pitt (2009)

CSC 236, Introduction to Theory of Computation, taught by Azadeh Farzan (2009)

CSC 336, Numerical Methods, taught by Hossein ZivariPiran (2008)

CSC 411, Machine Learning and Data Mining, taught by Aaron Hertzmann (2007)

REVIEWING

Neural Information Processing (NIPS 2015)	2015 (7)
International Conference on Machine Learning (ICML 2015)	2015 (6)
Neural Computation (NECO)	2014 (1)
Nature Communications	2014 (1)
Neural Information Processing (NIPS 2014)	2014 (5)
Journal of Machine Learning Research (JMLR)	2014 (1)
International Conference on Learning Representations (ICLR 2014)	2014 (5)
International Conference on Machine Learning (ICML 2014)	2014 (7)
NIPS 2013 Deep Learning Workshop	2013 (2)
Neural Information Processing (NIPS 2013)	2013 (5)
International Conference on Learning Representations (ICLR 2013)	2013 (4)
International Conference on Machine Learning (ICML 2013)	2013 (7)
Neural Computation (NECO)	2013 (1)
Neural Information Processing (NIPS 2012)	2012 (6)
International Conference on Machine Learning (ICML 2012)	2012 (5)
IEEE Transactions on Neural Networks	2011 (1)
Neural Networks	2011 (1)
Journal of Machine Learning Research (JMLR)	2010 (1)

TALKS

Optimizing Neural Networks with Kronecker-factored Approximate Curvature (invited talk)
Maluuba, Waterloo, Canada (2015)

Optimizing Neural Networks with Kronecker-factored Approximate Curvature (invited talk)
Amazon, Seattle, USA (2015)

Optimizing Neural Networks with Kronecker-factored Approximate Curvature (paper talk)
ICML 2015, Lille, Fance (2015)

Optimizing Neural Networks with Kronecker-factored Approximate Curvature (seminar)
Machine Learning Group, University of Toronto (2015)

The Effect of Depth on Representational Efficiency in Neural Networks and Certain Probabilistic Models (invited talk)

SAS, Raleigh, USA (2014)

The Effect of Depth on Representational Efficiency in Neural Networks and Certain Probabilistic Models (invited talk)
Skytree, San Jose, USA (2014)

The Effect of Depth on Representational Efficiency in Neural Networks and Certain Probabilistic Models (invited talk)
Deepmind, London, UK (2013)

The Effect of Depth on Representational Efficiency in Neural Networks and Certain Probabilistic Models (invited talk)
ICML 2013 Workshop on Deep Learning for Audio, Speech, and Language Processing, Atlanta, USA (2013)

Estimating the Hessian by Back-propagating Curvature (paper talk)
ICML 2012, Edinburgh, UK (2012)

An Approximate Newton Optimization Algorithm for Deep and Temporal Neural Networks (invited talk)
Numerical Analysis Group, University of Toronto (2011)

Learning Recurrent Neural Networks with Hessian-Free Optimization (paper talk)
ICML 2011, Bellevue, USA (2011)

Deep Learning via Hessian-free optimization (invited tutorial)
The CIFAR Summer School, Toronto, Canada (2010)

Deep Learning via Hessian-free Optimization (paper talk)
ICML 2010, Haifa, Israel (2010)

Learning the Linear Dynamical System with ASOS (paper talk)
ICML 2010, Haifa, Israel (2010)

Deep Learning via Hessian-free Optimization (seminar)
Machine Learning Group, University of Toronto (2010)