

Kevin Jordan Swersky

CONTACT INFORMATION 10 King's College Road *E-mail:* kswersky@cs.toronto.edu
Toronto, ON Canada *Website:* <http://www.cs.toronto.edu/~kswersky>
M5S 3G4

RESEARCH INTERESTS My research involves devising strategies for automatically tuning the parameters of machine learning models, in particular deep learning methods. I am primarily interested in the Bayesian approach to this problem as it represents a principled framework that has had considerable empirical success. In the past I've worked on graphical models, inference algorithms, Bayesian methods, time-series problems and unsupervised feature learning.

EDUCATION **University of Toronto**, Toronto, Ontario, Canada

Ph.D. Student, Computer Science, September 2011 to Present

- Adviser: Dr. Richard Zemel
- Area of Study: Machine Learning
- GPA: 4.0/4.0

University of British Columbia, Vancouver, British Columbia, Canada

M.Sc., Computer Science, August 2010

- Thesis Title: *Inductive Principles for Learning Restricted Boltzmann Machines*
- Adviser: Dr. Nando de Freitas
- Area of Study: Machine Learning
- GPA: 92.4% (A+)

University of Alberta, Edmonton, Alberta, Canada

B.Sc., Computing Science, August 2007

- GPA: 3.95/4.0

CONFERENCE PUBLICATIONS Input Warping for Bayesian Optimization of Non-Stationary Functions. Jasper Snoek, Kevin Swersky, Richard Zemel and Ryan P. Adams. *International Conference on Machine Learning (ICML)*, 2014.

Multi-Task Bayesian Optimization. Kevin Swersky, Jasper Snoek and Ryan P. Adams. *Neural Information Processing Systems (NIPS)*, 2013

Learning Fair Representations. Richard Zemel, Yu (Ledell) Wu, Kevin Swersky, Toniann Pitassi and Cynthia Dwork. *International Conference on Machine Learning (ICML)*, 2013

Stochastic k-Neighborhood Selection for Supervised and Unsupervised Learning. Daniel Tarlow, Kevin Swersky, Laurent Charlin, Ilya Sutskever and Richard Zemel. *International Conference on Machine Learning (ICML)*, 2013 .

Probabilistic n-Choose-k models for Classification and Ranking. Kevin Swersky, Daniel Tarlow, Ryan P. Adams, Richard S. Zemel and Brendan J. Frey. *Neural Information Processing Systems (NIPS)*, 2012.

Cardinality Restricted Boltzmann Machines. Kevin Swersky, Daniel Tarlow, Ilya Sutskever, Ruslan Salakhutdinov, Richard S. Zemel and Ryan P. Adams. *Neural Information Processing Systems (NIPS)*, 2012.

Fast Exact Inference for Recursive Cardinality Models. Daniel Tarlow, Kevin Swersky, Richard S. Zemel, Ryan P. Adams and Brendan J. Frey. *Uncertainty in Artificial Intelligence (UAI)*, 2012.

Estimating the Hessian by Backpropagating Curvature. James Martens, Ilya Sutskever and Kevin Swersky. *International Conference on Machine Learning (ICML)*, 2012.

Prediction and Fault Detection of Environmental Signals with Uncharacterised Faults. Michael A. Osborne, Roman Garnett, Kevin Swersky and Nando de Freitas. *AAAI Conference on Artificial Intelligence (AAAI)*, 2012.

On Autoencoders and Score Matching for Energy Based Models. Kevin Swersky, Marc’Aurelio Ranzato, David Buchman, Benjamin M. Marlin and Nando de Freitas. *International Conference on Machine Learning (ICML)*, 2011.

Inductive Principles for Restricted Boltzmann Machine Learning. Benjamin M. Marlin, Kevin Swersky, Bo Chen and Nando de Freitas. *Artificial Intelligence and Statistics (AISTATS)*, 2010.

OTHER
PUBLICATIONS

Learning Unbiased Features. Yujia Li, Kevin Swersky and Richard Zemel. *Neural Information Processing Systems Workshop on Transfer and Multi-Task Learning*. 2014.

Input Warping for Bayesian Optimization of Non-Stationary Functions. Jasper Snoek, Kevin Swersky, Richard Zemel and Ryan P. Adams. *Neural Information Processing Systems Workshop on Bayesian Optimization*. 2013.

Raiders of the Lost Architecture: Kernels for Bayesian Optimization in Conditional Parameter Spaces. Kevin Swersky, David Duvenaud, Jasper Snoek, Frank Hutter and Michael Osborne. *Neural Information Processing Systems Workshop on Bayesian Optimization*. 2013.

Efficient Feature Learning using Perturb-and-MAP. Ke Li, Kevin Swersky and Richard Zemel. *Neural Information Processing Systems Workshop on Perturbations, Optimization and Statistics*. 2013.

A Machine Learning Approach to Pattern Detection and Prediction for Environmental Monitoring and Water Sustainability. Michael Osborne, Roman Garnett, Kevin Swersky and Nando de Freitas. *ICML 2011 Workshop on Machine Learning for Global Challenges*, 2011.

Inductive Principles for Learning Restricted Boltzmann Machines. Kevin Swersky. *Master’s thesis, University of British Columbia*, 2010.

Sparsity Priors and Boosting for Learning Localized Distributed Feature Representations. Bo Chen, Kevin Swersky, Benjamin M. Marlin and Nando de Freitas. *Technical Report TR-2010-04. University of British Columbia, Department of Computer Science*, 2010.

A Tutorial on Stochastic Approximation Algorithms for Training Restricted Boltzmann Machines and Deep Belief Nets. Kevin Swersky, Bo Chen, Benjamin M. Marlin, and Nando de Freitas. *Information Theory and Applications (ITA) Workshop*, 2010.

COMMUNITY
SERVICE

Workshop Organizer

- NIPS 2014 Workshop on Bayesian Optimization in Academia and Industry.

Referee Service

- Neural Information Processing Systems (NIPS)
- Uncertainty in Artificial Intelligence (UAI)
- International Conference on Machine Learning (ICML)
- Artificial Intelligence and Statistics (AISTATS)
- Journal of Machine Learning Research (JMLR)
- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)

AWARDS

University of Toronto

- NSERC PGS-D Postgraduate Scholarship, 2012–2014

University of British Columbia

- Graduate Teaching Assistant Award, 2009
- NSERC PGS-M Postgraduate Scholarship, 2007–2009

University of Alberta

- Gold Medal in Computing Science, 2007
- Sony Computer Science Scholarship, 2006
- Louise Mckinney Post-Secondary Scholarship, 2006
- Louise Mckinney Post-Secondary Scholarship, 2005
- Jason Lang Scholarship, 2004

PROFESSIONAL
EXPERIENCE

Google Inc., Mountain View, California, United States

Software Engineering Intern

June 2012 to August 2012

- Host: Dr. Georg Heigold
- Worked on probabilistic graphical models for speech recognition.

Aquatic Informatics, Vancouver, British Columbia, Canada

Research and Development Engineer

September 2010 to June 2011

- Developed algorithms and software for forecasting, missing data imputation, anomaly detection and fault diagnosis in water quality time-series data.

Institute for Reconstructive Sciences in Medicine, Edmonton, Alberta, Canada

Student Researcher

May 2005 to August 2007

- Developed a database program in C# to manage patients and assist with medical research. This system is being expanded for usage by similar surgical groups worldwide.
- Converted MRI scans of a human tongue into a 3D model for the University of Calgary's CAVeMan Virtual Human Project.

TEACHING
EXPERIENCE

University of Toronto, Toronto, Ontario, Canada

Teaching Assistant

September 2011 to Present

- CSC 412/2506: Probabilistic Graphical Models
 - January 2015 to April 2015
 - Gave tutorials, developed and marked assignments.
- CSC 2515: Introduction to Machine Learning
 - September 2013 to December 2013
 - Gave tutorials, developed and marked assignments, held office hours.
- Neural Networks for Machine Learning
 - September 2012 to December 2012
 - Coursera online course with approximately 35,000 students enrolled.
 - Created quizzes and programming assignments, monitored discussion boards.

- CSC 411: Machine Learning and Data Mining
 - September 2011 to December 2011
 - Gave tutorials, designed and marked homework assignments.

University of British Columbia, Vancouver, British Columbia, Canada

Teaching Assistant

September 2007 to December 2008

- CPSC 111: Introduction to Computation
 - September 2008 to December 2008
 - Graded exams, ran weekly labs.
- CPSC 260: Object-Oriented Program Design
 - September 2007 to May 2008
 - Graded assignments and exams, ran weekly labs, held weekly office hours.