

DUMITRU ERHAN

701 1st Ave, Sunnyvale (CA), 94089, USA
+1-330-366-6433 • mail@dumitru.ca

EDUCATION

May 2011

Université de Montréal (UdeM), Canada

Doctor of Philosophy (PhD) in [Computer Science \(Informatique\)](#)

Thesis: Understanding deep architectures and the effect of unsupervised pre-training

Advisor: Yoshua Bengio

Cumulative GPA: 4.3/4.3

December 2006

Université de Montréal

Master of Science (MSc.) in [Computer Science \(Informatique\)](#)

Thesis: Collaborative filtering techniques for drug discovery

Advisor: Yoshua Bengio

Cumulative GPA: 4.2/4.3

June 2004

Jacobs University Bremen, Germany

Bachelor of Science (BSc.) in [Electrical Engineering & Computer Science](#)

Thesis: Combining Echo-State Networks with gradient-descent Recurrent Neural Network learning

Advisor: Herbert Jaeger

Cumulative GPA: 1.85/1.0, Major GPA: 1.84/1.0

WORK EXPERIENCE

March 2011—present Scientist at **Yahoo! Labs, Sunnyvale (CA), USA**

- Solving the problem of *federated search*: automatically learn from click data what type of content to slot where on the search results page. Deciding this in real-time (20ms).
- In charge of all the modelling details: coming up with better models, adding features, removing features, automated feature selection/understanding, using better latent embedding methods, exploration/exploitation trade-offs, unbiased offline evaluation, efficient parallel training (on $O(10^8)$ or more examples), automatic model selection, better measuring and interpreting implicit user feedback, personalization, etc.
- Product is live and serving 100% of Yahoo Search US traffic ($10^7 - 10^8$ queries per day), improving user metrics and revenue considerably.

RESEARCH INTERNSHIPS

Summer 2008, 2009, 2010 **Google Research, Mountain View (CA), USA**

- Learning hierarchical representations for *image annotation*, *speech modeling*, *OCR* and *image/video safe search*.
- Large-scale setting: Efficient implementations of novel learning algorithms in Theano, a Python/C++ machine learning library. Interfacing with MapReduce, BigTable and GPUs. Goal: deep models in production, training on $O(10^9)$ examples.
- Large-scale empirical analyses of deep architectures and the effect of unsupervised pre-training on their performance. Paper at AISTATS'09, subsequent JMLR article.

Summer 2006

Microsoft Research Cambridge, UK

- Research internship on *Social matchmaking in Xbox Live* in the [Machine Learning and Perception](#) Group with Thore Graepel and Ralf Herbrich
- Collaborative filtering algorithms and modifications to the TrueSkill system to predict feedback and its trustworthiness in Xbox Live. Used approximate inference in factor graphs (via Expectation-Propagation) and F#, SQL Server, Visual Studio and MATLAB as tools.

Summer 2004

Max-Planck Institute for Biological Cybernetics, Tübingen, Germany

- Research internship with **Arthur Gretton** in the Department of Empirical Inference.
- Devised new tests of statistical (in)dependence for time series, which were applied for the analysis of simultaneous MRI and electrode time-series recorded in brains of macaques. The entire project was implemented in MATLAB

Summer 2003

Helsinki University of Technology, Helsinki, Finland

- Research internship with **Jaakko Hollmén** in the Pattern Recognition Group of the Department of Computer and Information Science
- Worked on optimization of the Expectation-Maximization algorithm for 0-1 data.

OTHER EXPERIENCE

2010

Startup

- Machine learning consulting to a music recommendation company. Implemented various collaborative filtering techniques and delivered them as prototypes to the company.

Winter 2009, 2010 **Université de Montréal**

- Teaching assistant for two courses **IFT6266** and **IFT3395/6390 Machine Learning**
- Setting up a common **class project**: solve a large-scale character recognition problem and explore the applicability of deep architectures in such a context. Tutorials and support with Machine Learning questions. The project and its results formed an AISTATS paper.
- Organized weekly tutorial and lab sessions (in French/Python), in charge of correcting homework. Was TA for the same course in Winter 2008.

2004 to 2006

Université de Montréal

- Research project on *Collaborative filtering techniques for drug discovery* supervised by **Yoshua Bengio** and **Pierre-Jean L'Heureux**, in collaboration with **AstraZeneca**.
- Applied machine learning algorithms to Quantitative Structure Activity Relationship (QSAR) analysis for drug discovery problems. Implementation in C++/Python and in MATLAB, part of *PLearn*, a machine learning library developed in-house. First author of a journal article

PUBLICATIONS

• *Journals*

Dumitru Erhan, Yoshua Bengio, Aaron Courville, Pierre-Antoine Manzagol, Pascal Vincent, and Samy Bengio. *Why Does Unsupervised Pre-training Help Deep Learning?* Journal of Machine Learning Research, vol. 11(Feb), pp. 625–660, 2010.

Dumitru Erhan, Pierre-Jean L'Heureux, Yoshua Bengio, and Shi Yi Yue. *Collaborative filtering on a family of biological targets*. Journal of Chemical Information and Modeling, volume 46, number 2, pages 626–635, 2006.

James Bergstra, Norman Casagrande, **Dumitru Erhan**, Douglas Eck, and Balazs Kegl. *Aggregate features and AdaBoost for music classification*. Machine Learning, volume 65, issue 2-3, December 2006.

• *Conference proceedings*

Yoshua Bengio, Frédéric Bastien, Arnaud Bergeron, Nicolas Boulanger-Lewandowski, Thomas Breuel, Youssouf Chherawala, Moustapha Cisse, Myriam Côté, **Dumitru Erhan**, Jeremy Eustache, Xavier Glorot, Xavier Muller, Sylvain Pannetier Lebeuf, Razvan Pascanu, Salah Rifai, Francois Savard and Guillaume Sicard, *Deep Learners Benefit More from Out-of-Distribution Examples*, in: *JMLR W&CP: Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics (AISTATS 2011)*, Fort Lauderdale, FL, USA, 2011.

Dumitru Erhan, Pierre-Antoine Manzagol, Yoshua Bengio, Samy Bengio, and Pascal Vincent. *The Difficulty of Training Deep Architectures and the effect of Unsupervised Pre-Training*. In *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics (AISTATS 2009)*, Clearwater (FL), USA, 2009.

Hugo Larochelle, **Dumitru Erhan**, and Pascal Vincent. *Deep Learning using Robust Interdependent Codes*. In *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics (AISTATS 2009)*, Clearwater (Florida), USA, 2009.

Hugo Larochelle, **Dumitru Erhan**, and Yoshua Bengio. *Zero-data Learning of New Tasks*. In the *Proceedings of the Twenty Third AAAI conference on Artificial Intelligence*, Main Technical Track (AAAI), Chicago (IL), USA, 2008.

Hugo Larochelle, **Dumitru Erhan**, Aaron Courville, James Bergstra and Yoshua Bengio. *An Empirical Evaluation of Deep Architectures on Problems with Many Factors of Variation*. In *Proceedings of 24th International Conference on Machine Learning, ICML 2007*, Corvallis, OR, USA.

- *Posters and presentations*

Olivier Chapelle and **Dumitru Erhan**. *Improved preconditioner for hessian free optimization*. Oral presentation (by Olivier Chapelle) at the NIPS 2011 Workshop on Deep Learning and Unsupervised Feature Learning. Sierra Nevada, Spain.

Dumitru Erhan, Aaron Courville, Yoshua Bengio, Pierre-Antoine Manzagol, Pascal Vincent and Samy Bengio. *Why does Unsupervised Pre-training Help Deep Discriminant Learning?* Oral presentation at the NIPS 2009 Workshop on “The Generative & Discriminative Learning Interface”, Whistler, Canada.

Dumitru Erhan, Yoshua Bengio, Aaron Courville, and Pascal Vincent. *Visualizing Higher Layer Features of a Deep Network*. Spotlight presentation and poster at the ICML 2009 Workshop on “Learning Feature Hierarchies”, Montreal, Canada.

Hugo Larochelle, **Dumitru Erhan**, and Yoshua Bengio. Generalization to a zero-data task: an empirical study. Oral presentation (by Hugo Larochelle) at the Learning Workshop (“Snowbird”), San Juan, Puerto Rico, 2007.

Presented the work on *Collaborative filtering on a family of biological targets* at the 230th ACS National Meeting, in Washington, DC (2005) and at the 7th International Conference on Chemical Structures, Holland (2005).

COMPUTER SKILLS

- Experience with Python, C/C++, MATLAB, R, F#, SQL, L^AT_EX.
- Worked with: Hadoop, Pig, Oozie, Hive, HBase, BigTable.
- Also: the GNU toolchain (gcc, make, etc), awk, sed.

LANGUAGE SKILLS

- Romanian: Mother tongue; English, French, Russian: Fluent; German: Intermediate

AWARDS AND HONORS

2009–2010	End of PhD studies scholarship, UdeM (\$12,000)
2004–2008	Scholarship covering the difference in tuition fees for foreign students, FESP, UdeM (\$5,500/year)
2004, 2007, 2008	Scholarship, Department of Computer Science and Operations Research, UdeM (varying)
2001–2004	Undergraduate tuition scholarship, Jacobs University Bremen (€13,500/year)

SERVICE

- Reviewer for International Conference on Machine Learning (ICML) 2012 & 2009, International Joint Conference on Artificial Intelligence (IJCAI) 2011, Neural Information Processing Systems (NIPS) 2009, NIPS Workshop on Deep Learning (2010), NIPS Workshop on Deep Learning and Unsupervised Feature Learning (2011). Ad-hoc reviewing for Journal of Machine Learning Research (JMLR)