

Yann N. Dauphin

CONTACT INFORMATION

E-mail: yann@dauphin.io

Webpage: dauphin.io

RESEARCH INTERESTS

deep learning, machine learning, natural language processing, computer vision

PROFESSIONAL EXPERIENCE

Google Research, *Research Scientist* **January 2019 to now**

Facebook AI Research, *Research Scientist* **May 2015 to December 2018**

Developed novel and influential regularization and training methods as well as new architectures for natural language processing. Some of my work there has been featured in the media, such as in [Wired](#).

Microsoft Research, *Research Intern* **May 2013 to August 2013**

Implemented and developed deep learning algorithms for spoken language understanding, focusing on the problem of semantic parsing.

Google, *Research Intern* **May 2012 to August 2012**

Implemented and developed new deep learning algorithms for automatic speech recognition that scale to billions of examples.

SteerAds, *Research Engineer* **January 2011 to May 2012**

Implemented all learning algorithms for intelligent online ad placement and deployed in production serving over 10,000 requests per second.

EDUCATION

Université de Montréal, Montréal, Quebec Canada

Ph.D., Computer Science, 2015

- Dissertation Topic: “Advances in scaling deep learning algorithms”
- Advisor: [Professor Yoshua Bengio](#)
- Area of Study: Machine Learning

École Polytechnique de Montréal, Montréal, Quebec Canada

B.Eng., Computer Engineering, 2010

HONORS AND AWARDS

IEEE SPS '20 Best Paper Award

ACL '18 Best Paper Honourable Mention

NeurIPS '11 Best Student Paper Award Honourable Mention

Pascal2 Best UTLC Paper Award

Winner of the [Unsupervised and Transfer Learning Challenge: Phase 2](#).

Winner of the [Emotion Recognition in the Wild Challenge](#).

PUBLICATIONS

Dauphin, Y. and Cubuk, E.D., 2020, September. Deconstructing the Regularization of BatchNorm. In International Conference on Learning Representations.

Song, J., Dauphin, Y., Auli, M. and Ma, T., 2020, August. Robust and on-the-fly dataset denoising for image classification. In European Conference on Computer Vision (pp. 556-572). Springer,

Cham.

Dauphin, Y.N. and Schoenholz, S., 2019. MetaInit: Initializing learning by learning to initialize. *Advances in Neural Information Processing Systems*, 32, pp.12645-12657.

Yee, K., Dauphin, Y.N. and Auli, M., 2019, January. Simple and Effective Noisy Channel Modeling for Neural Machine Translation. In *EMNLP/IJCNLP* (1).

Lowe, R., Foerster, J., Boureau, Y.L., Pineau, J. and Dauphin, Y., 2019, May. On the Pitfalls of Measuring Emergent Communication. In *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems* (pp. 693-701).

Fan, A., Lewis, M. and Dauphin, Y., 2019, July. Strategies for Structuring Story Generation. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 2650-2660).

Wu, F., Fan, A., Baevski, A., Dauphin, Y. and Auli, M., 2018, September. Pay Less Attention with Lightweight and Dynamic Convolutions. In *International Conference on Learning Representations*.

Zhang, H., Dauphin, Y.N. and Ma, T., Fixup Initialization: Residual Learning without Normalization. In *International Conference on Learning Representations*.

Fan, A., Lewis, M., & Dauphin, Y. (2018). Hierarchical Neural Story Generation. In *Proceedings of the Association for Computational Linguistics 2018*.

Zhang, H., Cisse, M., Dauphin, Y. N., & Lopez-Paz, D. (2018). mixup: Beyond Empirical Risk Minimization.

Lewis, M., Yarats, D., Dauphin, Y., Parikh, D., & Batra, D. (2017). Deal or No Deal? End-to-End Learning of Negotiation Dialogues. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (pp. 2443-2453).

Gehring, J., Auli, M., Grangier, D., Yarats, D., & Dauphin, Y. N. (2017, July). Convolutional Sequence to Sequence Learning. In *International Conference on Machine Learning* (pp. 1243-1252).

Cisse, M., Bojanowski, P., Grave, E., Dauphin, Y., Usunier, N. (2017, July). Parseval Networks: Improving Robustness to Adversarial Examples. In *International Conference on Machine Learning* (pp. 854-863).

Dauphin, Y. N., Fan, A., Auli, M., & Grangier, D. (2017, July). Language Modeling with Gated Convolutional Networks. In *International Conference on Machine Learning* (pp. 933-941).

Gehring, J., Auli, M., Grangier, D., & Dauphin, Y. (2017). A Convolutional Encoder Model for Neural Machine Translation. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (Vol. 1, pp. 123-135).

Mesnil, G., Dauphin, Y., Yao, K., Bengio, Y., Deng, L., Hakkani-Tur, D., ... Zweig, G. (2015). Using recurrent neural networks for slot filling in spoken language understanding. In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(3), 530-539.

Dauphin, Y. N., Grangier, D. (2015). Predicting distributions with Linearizing Belief Networks. *arXiv preprint arXiv:1511.05622*.

Ebrahimi Kahou, S., Michalski, V., Konda, K., Memisevic, R., Pal, C. (2015, November). Recurrent neural networks for emotion recognition in video. In: *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (pp. 467-474). ACM.

Dauphin, Y., de Vries, H., Bengio, Y. (2015). Equilibrated adaptive learning rates for non-convex optimization. In: *Advances in Neural Information Processing Systems* (pp. 1504-1512).

Dauphin, Y. N., Pascanu, R., Gulcehre, C., Cho, K., Ganguli, S., Bengio, Y. (2014). Identifying and attacking the saddle point problem in high-dimensional non-convex optimization. In: *Advances in neural information processing systems*.

Y. Dauphin, G. Tur, D. Hakkani-Tur, L. Heck. Zero-Shot Learning and Clustering for Semantic Ut-

terance Classification. In: Proceedings of the International Conference on Learning Representations (ICLR 2014).

Y. Dauphin, Y. Bengio. Stochastic Ratio Matching of RBMs for Sparse High-Dimensional Inputs. In: Advances in Neural Information Processing Systems 26 (NeurIPS 2013).

S. Kanou, et al. Combining modality specific deep neural networks for emotion recognition in video. In: Proceedings of the 15th ACM on International conference on multimodal interaction (ICMI 2013).

Y. Bengio, G. Mesnil, Y. Dauphin, S. Rifai. Better Mixing via Deep Representations. In: Proceedings of the 30th International Conference on Machine Learning (ICML 2013).

S. Rifai, Y. Bengio, Y. Dauphin, P. Vincent. A Generative Process for Sampling Contractive Auto-Encoders. In: Proceedings of the 29th International Conference on Machine Learning (ICML 2012).

S. Rifai, Y. Dauphin, P. Vincent, Y. Bengio, X. Muller. The Manifold Tangent Classifier. In: Advances in Neural Information Processing Systems (NeurIPS 2011). **Invited as plenary talk (1.4% of submissions).**

Y. Dauphin, X. Glorot, Y. Bengio. Large-Scale Learning of Embeddings with Reconstruction Sampling. In: Proceedings of the 28th International Conference on Machine Learning (ICML 2011).

S. Rifai, G. Mesnil, P. Vincent, X. Muller, Y. Bengio, Y. Dauphin, X. Glorot. Higher Order Contractive Auto-Encoder. In: Proceedings of the European Conference on Machine Learning (ECML 2011).

G. Mesnil, Y. Dauphin, X. Glorot, S. Rifai, Y. Bengio, et al. Unsupervised and Transfer Learning Challenge: a Deep Learning approach. In: Journal of Machine Learning Workshop and Conference Papers (JMLR W&CP 2011).

See full list on my [Google Scholar page](#).

TEACHING
EXPERIENCE

École Polytechnique de Montréal, Montréal, Quebec Canada

Teaching Assistant

September 2007 to September 2009

SERVICE

Program co-Chair: NeurIPS 2021

Area Chair: ICML 2017, ICML 2018, NeurIPS 2018, ICLR 2019 ICML 2019, NeurIPS 2019, ICLR 2020, NeurIPS 2020, AAI 2020, ICLR 2021

Reviewer: International Conference on Machine Learning, Neural Information Processing Systems, Journal of Machine Learning Research, International Conference on Representation Learning

Contributor to several open-source software projects, including:

- [Scikit Learn](#), one of the top machine learning libraries in Python
- [Theano](#), Theano was a Python library for fast numerical operations that inspired [Tensorflow](#).
- [GNU CLisp](#), one of the top Lisp compilers
- [JGAP](#), a popular library for implementing Genetic Algorithms
- [SWTBot](#), a UI testing tool