School of Engineering and Applied Science Harvard University, Cambridge, MA Phone: (215) 317-8089 srush@seas.harvard.edu http://nlp.seas.harvard.edu @harvardnlp

## ALEXANDER M. RUSH

Appointment

Harvard University School of Engineering and Applied Sciences 2015-. Assistant Professor of Computer Science

Facebook Artificial Intelligence Research Lab 2015. Post-Doctoral Fellowship

Advisor: Yann LeCunn

Education

Massachusetts Institute of Technology

2009-2014. Ph.D, Computer Science.

Advisor: Michael Collins

Dissertation: Relaxation Methods for Natural Language Decoding.

Columbia University

2011-2014. Visiting Scholar, Department of Computer Science.

Harvard University

2007. B.A., Computer Science. (Magna Cum Laude With Highest Honors.)

Grants and Awards

- 2012 Best Paper Award, North American Association of Computational Linguistics.
- 2010 Best Paper Award, Empirical Methods in Natural Language Processing.
- 2009 Graduate Research Fellow, National Science Foundation.
- 2006 Outstanding Undergraduate Award Finalist, Computing Research Association.

Recent Publications

Yoon Kim, Sam Wiseman, Alexander M. Rush. Deep Latent-Variable Models for Natural Language. EMNLP 2018 (Tutorial)

Yuntian Deng\*, Yoon Kim\*, Justin Chiu, Demi Guo, Alexander M. Rush. Latent Alignment and Variational Attention. NIPS 2018

Sam Wiseman, Alexander Rush. Learning Neural Templates for Text Generation. EMNLP 2018

- Sebastian Gehrmann, Yuntian Deng, Alexander Rush. Bottom-Up Abstractive Summarization. EMNLP 2018
- Luke Melas-Kyriazi, George Han, Alexander Rush. Training for Diversity in Image Paragraph Captioning. EMNLP 2018 (Short)
- Luong Hoang, Sam Wiseman, Alexander Rush. Entity Tracking Improves Cloze-style Reading Comprehension. EMNLP 2018 (Short)
- Hendrik Strobelt, Sebastian Gehrmann, Michael Behrisch, Adam Perer, Hanspeter Pfister, Alexander M. Rush. Seq2Seq-Vis: A Visual Debugging Tool for Sequence-to-Sequence Models . VAST 2018, EMNLP-BlackBox 2018 (Best Paper Honorable Mention)
- Alexander M. Rush. The Annotated Transformer. ACL NLP-OSS 2018
- Jean Senellart, Dakun Zhang, Bo Wang, Guillaume Klein, J.P. Ramatchandirin, Josep Crego, Alexander M. Rush. OpenNMT System Description for WNMT 2018: 800 words/sec on a single-core CPU. WNMT 2018 (First-Place CPU Speed/Memory)
- Yoon Kim, Sam Wiseman, Andrew C. Miller, David Sontag, Alexander M. Rush. Semi-Amortized Variational Autoencoders. ICML 2018
- Alexander M. Rush and Slav Petrov, Vine Pruning for Efficient Multi-Pass Dependency Parsing. Proceedings of NAACL 2012. [Best Paper Award]
- Terry Koo, Alexander M. Rush, Michael Collins, Tommi Jaakkola, and David Sontag. *Dual Decomposition for Parsing with Non-Projective Head Automata*. Proceedings of EMNLP 2010. [Best Paper Award]

## Patents

- Techniques for discriminative dependency parsing (Google). Slav Petrov, Alexander M. Rush, 2015.
- Efficient parsing with structured prediction cascades (Google). Slav Petrov, Alexander M. Rush, 2013
- Determining user affinity towards applications on a social networking website (Facebook., Thomas S. Whitnah, Alexander M. Rush, Ding Zhou, Ruchi Sangvhi, 2010.

Industry

Lead Engineer (Platform Team), Facebook, June 2007 – August 2009, Palo Alto, CA.

Deep Latent-Variable Models for Natural Language - Inference in deep latent variable models.

Variational Attention - fast training for latent var. attention (e.g. hard attention) with accuracy like soft attention

 $\rm Seq 2 Seq \ Vis \$  - Interaction with seq 2 seq translation dynamics.

OpenNMT - Open-source neural machine translation system.