Alexis J. Ross

alexisr@allenai.org
https://alexisjihyeross.github.io/

Education

Harvard University, Cambridge, Massachusetts — Sept 2016 – May 2020
A.B. with Honors in Computer Science and Philosophy (Joint)

Research Experience

Allen Institute for Artificial Intelligence (AI2), AllenNLP — July 2020 – Present
Predoctoral Researcher
Mentors: Ana Marasović and Matthew E. Peters

Harvard University, Computer Science and Philosophy — Sept 2019 – May 2020
Mentors: Himabindu Lakkaraju and Bernhard Nickel
Senior Thesis: Explaining Blackbox Models with Linear Approximations
Awarded Thomas T. Hoopes Prize for "outstanding scholarly work" by undergraduates.

Microsoft Research — May 2019 – Aug 2019
Research Intern, Project Hanover
Mentor: Tristan Naumann
Project: Machine Reading for Precision Medicine

Harvard University, NLP Group — March 2019 – May 2019
Research Assistant
Mentors: Yonatan Belinkov and Alexander Rush
Project: Probing for Neurons that Encode and Exploit Negation in BERT

Johns Hopkins, Center for Language and Speech Processing — June 2018 – Aug 2018
Undergraduate Researcher, JSALT Workshop
Mentor: Ellie Pavlick
Project: General-Purpose Sentence Representation Learning

Harvard University, Cognitive Neuropsychology Laboratory — Feb 2017 – Aug 2017
Research Assistant
Mentors: Gilles Vannuscorps and Alfonso Caramazza
Project: Object Orientation Perception

Preprints

[1] Tailor: Generating and Perturbing Text with Semantic Controls.
Alexis Ross*, Tongshuang Wu*, Hao Peng, Matthew Peters, Matt Gardner. In submission, 2021.

Publications

[2] Learning Models for Actionable Recourse.
Alexis Ross, Himabindu Lakkaraju, and Osbert Bastani. *NeurIPS, 2021.*

[3] Competency Problems: On Finding and Removing Artifacts in Language Data.
Matt Gardner*, William Merrill*, Jesse Dodge, Matthew E. Peters, **Alexis Ross**, Sameer Singh, Noah Smith. *EMNLP, 2021.*

[4] Explaining NLP Models via Minimal Contrastive Editing (MiCE).
Alexis Ross, Ana Marasović, and Matthew E. Peters. *Findings of ACL 2021.*

[5] How well do NLI models capture verb veridicality?
Alexis Ross and Ellie Pavlick. *EMNLP 2019.* **Oral presentation**.

[6] Probing what different NLP tasks teach machines about function word comprehension.
Najoung Kim, Roma Patel, Adam Poliak, Alex Wang, Patrick Xia, R Thomas McCoy, Ian Tenney, **Alexis Ross**, Tal Linzen, Benjamin Van Durme, Samuel R Bowman, and Ellie Pavlick. **SEM 2019.* **Best Paper Award**.

Teaching Experience

CS 121: Introduction to the Theory of Computation Sept 2018 – Dec 2018
Teaching Fellow, Harvard School of Engineering and Applied Sciences
Instructor: Boaz Barak
Awarded a Certificate of Distinction in Teaching for a student evaluation of 4.8/5.0.