Andrew Drozdov

mrdrozdov.github.io | adrozdov@cs.umass.edu | mrdrozdov | RsU18o4AAAAJ

Summary

I am a Ph.D. student co-advised by Professors Andrew McCallum and Mohit Iyyer in the College of Information and Computer Sciences at UMass Amherst, and a member of the IESL and NLP research groups. My research interests are in in-context learning and information retrieval, especially retrieval-augmented models. I also work on structured prediction, representation learning, knowledge distillation, question answering, and text generation. I am seeking postdoc positions for 2023.

Education

University of Massachusetts Amherst, *Ph.D. in Computer Science* Sep 2018 - Sep 2023 (exp.)
Co-advised by Andrew McCallum and Mohit Iyyer. Focus: Deep learning methods for natural language processing.

New York University, *M.S. in Computer Science* Sep 2015 - Dec 2016

Cornell University, *M.Eng. in Computer Science* Sep 2013 - Dec 2013
Left early to join Okta full-time.

University of Michigan, *B.S.E. in Computer Science* Sep 2009 - May 2013

Work Experience

Research Assistant, *University of Massachusetts Amherst* Sep 2018 - Present

Student Researcher, *Google - Google Research, w/ Kai Hui & Don Metzler* Apr 2023 - Present

Research Intern & Student Researcher, *Google - Google Research, w/ Xinying Song & Denny Zhou* Summer 2022

Research Intern, *IBM - IBM Research* Summer 2021

Research Intern & Student Researcher, *Google - Google AI Language* Summer 2019

Research Engineer, *eBay - Deep Learning Recommendation Systems* Aug 2017 - Aug 2018

Visiting Scholar, *New York University* Jan 2017 - Jul 2017

Data Engineer, *Datadog* Summer 2015

Software Engineer, *Okta* Jun 2013 - Feb 2015

Selected Publications

Compositional Semantic Parsing with Large Language Models
A. Drozdov, N. Schärli, E. Akyürek, N. Scales, X. Song, X. Chen, O. Bousquet, D. Zhou
ICLR 2022.

You can't pick your neighbors, or can you? When and how to rely on retrieval in the kNN-LM
A. Drozdov, S. Wang, N. Rahimi, A. McCallum, H. Zamani, M. Iyyer
EMNLP 2022 (Findings).

Inducing and Using Alignments for Transition-based AMR Parsing
A. Drozdov, J. Zhou, R. Florian, A. McCallum, T. Naseem, Y. Kim, R. Astudillo
NAACL 2022.

Improved Latent Tree Induction with Distant Supervision
A. Drozdov, Z. Xu, J. Lee, T. O'Gorman, S. Rongali, M. Iyyer, A. McCallum
EMNLP 2021.

Unsupervised Parsing with S-DIORA: Single Tree Encoding for DIORA
A. Drozdov, S. Rongali, Y. Chen, T. O'Gorman, M. Iyyer, A. McCallum
EMNLP 2020.

Unsupervised Labeled Parsing with DIORA
A. Drozdov, P. Verga, Y. Chen, M. Iyyer, A. McCallum
EMNLP 2019 (Short Paper).

Unsupervised Latent Tree Induction with Deep Inside-Outside Recursive Auto-Encoders (DIORA)
A. Drozdov, P. Verga, M. Yadav, M. Iyyer, A. McCallum
NAACL 2019 (Oral).

Emergent Communication in a Multi-Modal, Multi-Step Referential Game
K. Evtimova, **A. Drozdov**, D. Kiela, K. Cho
ICLR 2018.

Do latent tree learning models identify meaningful structure in sentences?
A. Williams, **A. Drozdov**, S. Bowman
TACL 2018.

Professional Service

Reviewing:
AAAI '19, '23; Neurips '19, '20, '21, '22 (Top Reviewer); ICML '20 (Top-33%), '21 (Expert Reviewer), '22, '23; ICLR '22, '23; SIGIR '22 (Secondary Reviewer), '23; CoNLL '20, '21, '22; ACL '21 (Secondary Reviewer); EMNLP '22; ARR

Teaching

UMass Amherst, *Teaching Assistant*
Industry Mentorship Course (CS-696DS) with Andrew McCallum. Spring '22, Spring '23
Advanced Natural Language Processing (CS-685) with Mohit Iyyer. Spring '22

Cornell University, *Teaching Assistant*
Data Science in the Wild (CS-5304) with Giri Iyengar at Cornell Tech. Spring '18

Invited Talks

NYU, Tal Linzen's lab. Unsupervised parsing, success and failures. Spring '22
UMass Amherst, Neural Networks (CS-682) taught by Erik Learned-Miller. Using transformers for NLP. Fall '21
MIT, NLP lab meeting invited by Yoon Kim. Neural alignments for AMR. Fall '21
CMU, Algorithms for NLP (CS-11711) taught by Emma Strubell. Unsupervised parsing with S-DIORA. Fall '20
IBM, NLP reading group, organized by Ramon Astudillo. Unsupervised parsing with DIORA. Spring '20

Research Mentoring

J. Zhao (MS), M. Tulsyan (MS), Y. Kashyap (MS): Confident Student Training for Few-shot Knowledge Distillation.
H. Ananthakrishnan (MS), A. Hattimare (MS), G. Vyas (MS): Improved cross-lingual transfer with data augmentation, Co-mentored with Saleh Sultan (Amazon Alexa).
N. Nizar (MS): Are pre-trained LMs robust to OCR-like noise?
S. Mishra (MS): Combining Chart-based Models for Improved Unsupervised Parsing.
Z. Xu (MS): Improved Latent Tree Induction with Distant Supervision, EMNLP '21.
D. Finkbeiner (MS): Robust Unsupervised Parsing.
S. Suresh (MS): Unsupervised Parsing via Multilingual Span Constraints.
N. Srinivasan (MS), P. Shetty (MS): Document Representation Methods for Tracking Paper Revisions, Co-mentored with Amanda Stent (Bloomberg).
S. Satish (MS), Z. Yao (MS): The Impact of Preprints in the Formation of Novel Ideas, Co-mentored with Boris Veytsman (CZI), EMNLP Workhsop '20.
S. Jalan (MS), S. Gangwar (MS): Semi-Supervised Parsing with Entity Constraints.
Y. Chen (MS): Improved Representation Learning with DIORA.
L. Kantor (BS): Linguistics Honors Thesis, advised by Joe Pater.

Awards

Best Deep Learning Project (Jointly with K. Evtimova) Fall '16
NYU's Center of Data Science Award Ceremony. Award selected by Yann Lecun.
Project Title: Understanding Mutual Information and its Use in InfoGAN

Activities

Data Science Tea, *Co-Organizer* Fall '18, Fall '19
Data Science and Machine Learning Speaker Series

Personal Interests

Outside of research, I like to go hiking and explore museums and art galleries. A long time ago (in high school), I was a competitive runner, setting team mid-distance records and participating in the pentathlon.