Andrew Drozdov

914.574.1221 ◇ adrozdov@cs.umass.edu ◇ mrdrozdov.com ◇ github.com/mrdrozdov

Education

UMass-Amherst, *Ph.D. in Computer Science* Sep '18 - Dec '23 (Expected)
Co-advised by Andrew McCallum and Mohit Iyyer. Deep learning methods for information extraction and natural language processing.

New York University, *M.S. in Computer Science* Sep '15 - Dec '16

Cornell University, *M.Eng. in Computer Science* Sep '13 - Dec '13
Left to join Okta full-time.

University of Michigan, *B.S.E. in Computer Science* Sep '09 - May '13

Experience

Google Research, Brain Team, *PhD Research Intern* Summer '22

IBM, *PhD Research Intern* Summer '21

Google Research, Language, *PhD Research Intern* Summer '19

eBay, *Research Engineer* Aug '17 - Aug '18

New York University, *Research Assistant* Jan '17 - Jul '17

Datadog, *Data Engineer* Summer '15

Okta, *Software Engineer* Jun '13 - Feb '15

Research Papers

Compositional Semantic Parsing with Large Language Models
A. Drozdov, N. Schärli, E. Akyürek, N. Scales, X. Song, X. Chen, O. Bousquet, D. Zhou. We achieve state-of-the-art results on semantic parsing through dynamic least-to-most prompting. In submission.

You can't pick your neighbors, or can you? Improving kNN-LM
A. Drozdov, S. Wang, N. Rahimi, A. McCallum, H. Zamani, M. Iyyer
We extend kNN-LM using an adaptive interpolation coefficient. In submission.

Inducing and Using Alignments for Transition-based AMR Parsing
A. Drozdov, J. Zhou, R. Florian, A. McCallum, T. Naseem, Y. Kim, R. Astudillo
We use a neural model to learn alignments for AMR parsing. NAACL 2022 (Poster).

Improved Latent Tree Induction with Distant Supervision
A. Drozdov, Z. Xu, J. Lee, T. O'Gorman, S. Rongali, M. Iyyer, A. McCallum
We demonstrate the effectiveness of distant supervision via entity-based span constraints for unsupervised constituency parsing with DIORA. EMNLP 2021 (Poster).

Unsupervised Parsing with S-DIORA: Single Tree Encoding for DIORA
A. Drozdov, S. Rongali, Y. Chen, T. O'Gorman, M. Iyyer, A. McCallum
By adding a beam to each cell in DIORA's chart data structure, we show substantial improvements in unsupervised constituency parsing. EMNLP 2020 (Poster).

Deep Inside-Outside Recursive Autoencoders (Labeled Parsing)
A. Drozdov, P. Verga, Y. Chen, M. Iyyer, A. McCallum
We extend DIORA with latent codes and achieve state-of-the-art results when inducing labels for constituency parse trees. EMNLP 2019 (Short).

Deep Inside-Outside Recursive Autoencoders (Unlabeled Parsing)
A. Drozdov, P. Verga, M. Yadav, M. Iyyer, A. McCallum
Using dynamic programming and neural networks, we learn structured representation of text that are efficient for reconstructing sentences. Our method achieves various state-of-the-art results in unsupervised grammar induction. NAACL 2019 (Oral).

Research Papers (cont.)

Emergent Language in a Multi-Modal, Multi-Step Referential Game
K. Evtimova, **A. Drozdov**, D. Kiela, K. Cho
When training cooperative agents to classify an image in a zero-shot setting, we observe patterns in their messages and conversation length. ICLR 2018 (Poster).

Do latent tree models identify meaningful structure in sentences?
A. Williams, **A. Drozdov**, S. Bowman
We analyze four models for grammar induction with neural networks. Although they perform well on their respective semantic tasks, it remains unclear whether the generated syntax is optimal or linguistically justified. TACL 2018.

Service

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

Teaching

UMass Amherst, *Teaching Assistant*
Industry Mentorship Course (CS-696DS) with Andrew McCallum. Spring '22
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 labeled and unlabeled parsing with DIORA. Spring '20

Advising

UMass Amherst
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