Carnegie Mellon University
Language Technologies Institute

Adina Williams
Facebook AI

Adina Williams is a Research Scientist at Facebook AI Research in NYC (started October 2018). Previously, she earned her PhD at New York University in the Department of Linguistics, where she investigated the brain basis of syntactic and semantic processing. Her main research goal is to strengthen the connections between linguistics and cognitive science on the one hand and natural language processing and artificial intelligence on the other. She approaches this process from both directions: she brings linguistic and cognitive scientific insights about human language to bear on training, evaluating, and debiasing NLP systems, and also applies statistical methods and corpus analytic tools from NLP to uncover new quantitative, cross-linguistic facts about particular human languages.

Using Human Perspectives to Understand NLI Models

Given the increasingly prominent role of NLP models in the world, it is crucial that we develop a better understanding of their behavior. In this talk, I present two different angles on this topic that rely on human intuitions about NLI data and NLI model behavior. First, we hand-annotate the development set of the difficult adversarial NLI dataset to better understand which types of reasoning are necessary to get the correct label. We show that the same types of reasoning are responsible for failures and successes of different state-of-the-art models, suggesting we might benefit from focusing our efforts on architectures that can handle those phenomena. Second, we ask whether human-generated explanations of models’ inference decisions align well with how models actually make those decisions. We propose an alignment metric based on human-generated natural language explanations and find that our models are only weakly aligned to humans. Taken together, these projects expose opportunities for using human perspectives to improve our current best models.

Friday, February 12, 2021

2:20 - 3:40 PM EST

Join the meeting on Zoom
Meeting ID 935 3287 1380 Passcode 546823

LTI Colloquium 2020-21
is generously sponsored by abridge