

COMPUTER SCIENCE SEMINAR

Addressing Biases for Robust, Generalizable AI

Swabha Swayamdipta
Allen Institute for AI

Abstract: Artificial Intelligence has made unprecedented progress in the past decade. However, there still remains a large gap between the decision-making capabilities of humans and machines. In this talk, I will investigate two factors to explain why. First, I will discuss the presence of undesirable biases in datasets, which ultimately hurt generalization, regardless of dataset size. I will then present bias mitigation algorithms that boost the ability of AI models to generalize to unseen data. Second, I will explore task-specific prior knowledge which aid robust generalization, but are often ignored when training modern AI architectures on large amounts of data. In particular, I will show how linguistic structure can provide useful biases for inferring shallow semantics, which help in natural language understanding. I will conclude with a discussion of how this framework of dataset and model biases could play a critical role even in the societal impact of AI, going forward.

Wednesday, February 10, 2021, 1:00 pm
<https://emory.zoom.us/j/95438087188>

COMPUTER SCIENCE
EMORY UNIVERSITY