Faculty Name:

Richard Zemel

Faculty Email:

zemel@cs.columbia.edu

Lab:

Controllable Machine Learning Lab

Project Title:

Uncertainty Quantification in Frontier AI Modelss

Description:

Multimodal AI has the potential to revolutionize healthcare. Combining vision and language

capabilities, these systems can analyze visual data (e.g., medical imaging) and linguistic

information (e.g., patient records), to offer a comprehensive and nuanced assessment of patient health. However, while modern large vision and language models (LVLMs) like GPT-4 and LlaVa have impressive accuracy on a wide range of benchmarks, they are prone to errors that can be difficult to predict and interpret. To enable reliable LVLM deployment, we must produce robust techniques for identifying when a model is uncertain and enable a practitioner to calibrate their confidence to the model's prediction when making critical decisions.

The aim of the summer project is to develop an algorithm to quantify the uncertainty of large multimodal medical AI systems. Nikita will work with instruction-tuned VLMs, including LLaVa-Med, which are trained to give their answers in a conversational fashion. She will draw on existing literature on uncertainty quantification for LLMs and vision models to develop a method for scoring the uncertainty of individual predictions made by these models. Although we focus on the safe usage of multimodal models in healthcare, such a method can also be applied to other risk-sensitive domains, for example autonomous vehicles.

Location of Research:

On Site

# of hrs/week:

40

Department/Program:

Computer Science

Eligibility:

MS

To apply, please contact:

Richard Zemel

zemel@cs.columbia.edu