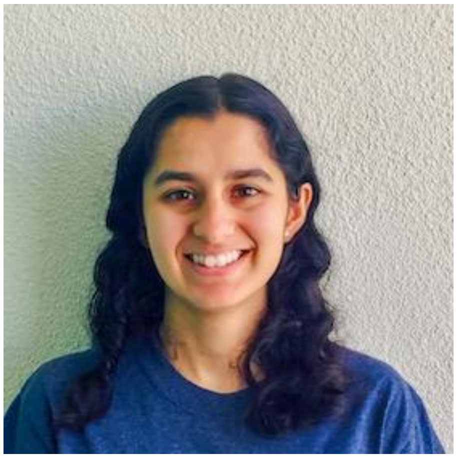
COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF MICHIGAN

DISSERTATION DEFENSE



Puja Trivedi

Improving Graph Representation Learning with Augmentations, Uncertainty Quantification and Large Language Model Guidance

Thursday, November 7, 2024
11:00am – 1:00pm
Virtual – Zoom
Password: pujasroom

ABSTRACT: Expressive graph representation learning is important to many high-impact applications as structured data across many domains can be naturally represented using graphs. While the advent of graph neural networks has led to considerable success on a variety of graph-based tasks, there remains room for improvement when the quality of GNN representations is determined with respect to not only in-distribution task performance but also other desiderata, such as generalization under distribution shift, trust-worthiness, or robustness. To this end, this thesis is broadly interested in understanding and improving the quality of graph representations {beyond accuracy} and makes contributions to several of the aforementioned desiderata.

The first part of this thesis considers the generalization and expressiveness of graph representations learnt using contrastive learning, as expressivity is often a prerequisite for other model desiderata. In particular, we study the role of augmentations, identifying several weaknesses of popular, generic graph augmentations strategies and derive a corresponding generalization bound. In the second part, we focus on improving the reliability of uncertainty estimates when performing predictive tasks, as expressivity alone is not sufficient for safe model deployment. To this end, we propose a lightweight training protocol that improves estimate quality, even under challenging distribution shift settings. The final part of this thesis considers strategies for leveraging large language models to further enhance GNN performance, particularly by supporting capabilities that are not realizable from only the GNN or structured data alone. Specifically, we propose a framework for LLM guided text-attributed graph clustering that uses the LLM to disambiguate uncertain nodes, and improves zero-shot clustering capabilities for disconnected nodes. Finally, we study the robustness of joint LLM and GNN models to structural and text-based adversarial attacks.

CHAIR: Prof. Danai Koutra