Statement of Purpose

Mrigank Raman

mt1170736@iitd.ac.in

My objective is to pursue a Ph.D in Computer Science with a focus on Graph based Machine Learning and it's applications in Natural Language Processing. My past research in Graph based Machine Learning has investigated the robustness of existing Common Sense Question Answering models based on Knowledge Graphs. I have also worked on Visual Question Answering using external knowledge. Being an undergraduate research assistant, I have also explored the fields of Active Sensing and Computer Vision. I am currently working with Professor Xiang Ren of University of Southern California in the field on Graph based Machine Learning. My ultimate goal is to be a Professor at a research institution.

My motivation for research stems from the various research based projects and activities that I have been a part of as an undergraduate. My first research project was based on Energy Efficient Active Remote Sensing under the guidance of Prof. Arpan Chattopadhyay of IIT Delhi. This project was instrumental in providing me an introduction to research. We designed an algorithm for active tracking of a Markov Chain with unknown dynamics. I came up with the idea to treat the model as a Hidden Markov Model and apply the Online version of Expectation Maximization algorithm to estimate the parameters of the Markov Chain. Once we had the dynamics, we used a variant of the Kalman Filter to estimate the state of the Markov Chain. This algorithm also used Gibbs Sampling to constrain the mean number of active sensors which makes our algorithm energy efficient. This experience helped me realize how much I enjoy reading research papers in the field of Machine Learning.

It was the sixth semester when I was introduced to Machine Learning in Graphs. I registered for a graduate course named Special Topics in Database Systems. This research based course was aimed at introducing us to the current state of the art architectures and applications in the field of Machine Learning in Graphs. Unfortunately this course was cut short by COVID-19 but still I had enough time to benchmark and experiment with Struc2Vec, present and critique two research papers[1,2] and read a bunch of other research papers on this topic. During summer 2020, I was selected as a research intern in the Intelligence and Knowledge Discovery (INK) Research Lab at University of Southern California under the guidance of Professor Xiang Ren. After about a month of brainstorming and literature survey we decided the topic of my research which was to investigate the robustness of existing models that leverage KG's for various tasks. Our targeted perturbation using

an RL framework over KGs tries to shed light into how the KG is being leveraged in current systems, and reveal potentials concerns that, even if the system seems to be working well (i.e., improve over non-KG baselines), it is not sensitive to semantic changes of the KG. In particular, such changes might be factually incorrect or ethically harmful. We show our results on the Commonsense Question Answering and Recommendation tasks. I am also working on developing a Question Answering Model that will be robust to such perturbations using a Language Model to generate correct relational evidence between concepts.

I have also worked in the field of Computer Vision on the problem of Domain Generalization. We try to learn domain agnostic features that convert the multi source domain generalization problem to a single source DG problem. We then use a VAE(infinite sampler) to search a nearest neighbour to the target in the source domains. We provide interpretable bounds on the risk of our classifier and demonstrate state-of-the-art numbers on the PACS and VLCS dataset using both Alexnet and Resnet18. We have submitted our paper to the AAAI 2021 conference.

I believe, I can build upon these research experiences and be a perfect fit for the PhD programme in Stanford. I would love to contribute to the ongoing projects in SNAP Research Group headed by Prof Jure Leskovec, mainly the projects based on representation learning in Graphs and Knowledge Graph based Commonsense Reasoning. I think areas such as zero-shot learning, domain generalization and adaptation, adversarial attacks and defenses will be very interesting to explore in Graphs using representation learning. This is a relatively recent field and I believe as a researcher in SNAP research group, I can contribute a lot to the research community.