

Cover Letter

I am Rishabh Maheshwary, interested in working at the intersection of research and engineering on NLP, MultiModal vision & language, and scalability, alignment, and safety of large language models. As AI systems are being increasingly used in day to day tasks, I am passionate about creating AI systems that are not only intelligent but also safe, robust, and that align with human goals and values. My work focuses on improving the robustness, reliability and safety of large language models.

I am currently working on aligning language models with human values using reinforcement learning from human feedback. I recently completed my AI Residency from [Facebook AI Research](#), where my research was focused on designing intelligent systems having joint understanding of vision and language modalities [\[CVPR '23\]](#). I worked with [Marcus Rohrbach](#), to design visual question answering models that generate reliable predictions by reading text in images, reasoning over the text and other visual content to help visually impaired people in day to day tasks.

I did my Masters from [International Institute of Information Technology, Hyderabad \(IIIT-H\)](#), where I worked on the task of evaluating the generalization and robustness of various NLP models in a real world setting. My first work [\[AAAI '21\]](#) proposes a new adversarial attack strategy that attacks NLP models in a highly constrained setting. In my second work [\[AAAI '21\]](#), I improved the quality of existing adversarial attack methods using masked language models. I extended my research by another work [\[EMNLP '21\]](#) which proposes a query efficient attack mechanism. Besides, I worked on the problem of evaluating Math Word Problem solvers for their robustness and natural language understanding ability [\[Findings of EMNLP '21\]](#) and [\[NAACL' 22\]](#).

Previously, I was also a research intern at [Verisk AI](#) where I worked on the task of information extraction from form-like documents. During my undergraduate years at [UIET, Panjab University, Chandigarh](#), I did [Google Summer Of Code](#) Internship in [AOSSIE](#) organization and worked as a Software Developer Intern at [MSB Docs](#). I am an active contributor to the [TextAttack](#) and [MMF](#) frameworks. I am proficient with Python and its full ML stack, and have hands-on experience with deep learning frameworks like Pytorch and Tensorflow as well.

Relevant Links [GitHub](#) | [GitLab](#) | [LinkedIn](#) | [Google Scholar](#) | [Website](#)

Contact Information

Email: rf.rishabh@gmail.com, rishabh.maheshwary@research.iiit.ac.in

Phone: +91-8427119320, +1-650-609-1067