

Rishabh Maheshwary

rf.rishabh@gmail.com | +91 – 8427119320 | Google Scholar | Website | GitHub | LinkedIn

Summary

My research encompass training, alignment, and benchmarking of language models. With the growing integration of AI into our daily lives, I am passionate about developing safe and robust AI systems.

Work Experience

- **ServiceNow Research - Senior Applied Scientist** **Hyderabad, India**
My work focuses on enhancing LLMs in post-training phase through data curation, alignment, tool-augmented reasoning, and benchmarking. *Aug 2023 – Present*
 - **Facebook AI Research - AI Resident** **California, U.S.**
My research was focused on designing intelligent and reliable systems having joint understanding of vision and language modalities. *Nov 2021 – Dec 2022*
 - **Verisk AI – Research Intern** **Hyderabad, India**
I worked on joint language and vision understanding of multimodal content and semantic understanding of natural language documents. *May 2021 – Oct 2021*
 - **Google Summer of Code – Software Developer Intern** **Remote**
I developed an application enabling users to report incidents thus facilitating nearby assistance and communication. *Apr 2018 – Sept 2018*
-

Publications

1. **INCLUDE: Evaluating Multilingual Language Understanding with Regional Knowledge.**
Core Contributor. International Conference on Learning Representations (ICLR 2025), Singapore.
2. Srishti Gureja*, Lester James V. Miranda*, Shayekh Bin Islam*, **Rishabh Maheshwary***, Drishti Sharma, Gusti Winata, Nathan Lambert, Sebastian Ruder, Sara Hooker, Marzieh Fadaee. **M-RewardBench: Evaluating Reward Models in Multilingual Settings.**
Association for Computational Linguistics (ACL 2025), Vienna, Austria.
3. **Rishabh Maheshwary**, Vikas Yadav, Hoang Nguyen, Khyati Mahajan, Sathwik Tejaswi Madhusudhan. **M2Lingual: Enhancing Multilingual, Multi-Turn Instruction Alignment in LLMs.**
North American Chapter of the Association for Computational Linguistics (NAACL 2025), Albuquerque, New Mexico.
4. Pulkit Pattnaik*, **Rishabh Maheshwary***, Kelechi Ogueji, Vikas Yadav, Sathwik Tejaswi Madhusudhan. **Enhancing Alignment using Curriculum Learning & Ranked Preferences.**
Findings of Empirical Methods in Natural Language Processing (EMNLP 2024), Miami, Florida.
5. Corentin Dancette, Spencer Whitehead, **Rishabh Maheshwary**, Ramakrishna Vedantam, Stefan Scherer, Xinlei Chen, Matthieu Cord, Marcus Rohrbach. **Improving Selective Visual Question Answering by Learning from Your Peers.**
IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2023), Vancouver, Canada.

6. Vivek Kumar, Rishabh Maheshwary, Vikram Pudi. **Practice Makes a Solver Perfect: Data Augmentation Methods for Math Word Problem Solvers.**
North American Chapter of the Association for Computational Linguistics (NAACL 2022), Seattle, Washington.
7. Rishabh Maheshwary*, Saket Maheshwary*, Vikram Pudi. **A Strong Baseline for Query Efficient Attacks in a Black Box Setting.**
Empirical Methods in Natural Language Processing (EMNLP 2021), Punta Cana, Dominican Republic.
8. Rishabh Maheshwary*, Vivek Kumar*, Vikram Pudi. **Adversarial Examples for Evaluating Math Word Problem Solvers.**
Findings of EMNLP 2021, Punta Cana, Dominican Republic.
9. Rishabh Maheshwary, Saket Maheshwary, Vikram Pudi. **Generating Natural Language Attacks in a Hard Label Black Box Setting.**
Association for the Advancement of Artificial Intelligence (AAAI 2021), Vancouver, Canada.
10. Rishabh Maheshwary, Saket Maheshwary, Vikram Pudi. **A Context Aware Approach for Generating Natural Language Attacks.**
Association for the Advancement of Artificial Intelligence (AAAI 2021), Vancouver, Canada.

* Equal Contribution

Education

- **International Institute of Information Technology, Hyderabad** **Hyderabad, India**
MS by Research in Computer Science and Engineering | CGPA: 8.7/10 *2019 – 2021*
 - **University Institute of Engineering and Technology, Panjab University** **Chandigarh, India**
BTech in Computer Science and Engineering | CGPA: 8.4/10 *2015 – 2019*
-

Major Projects

- **Information Extraction from Form like Documents** *2021*
The aim is to design an intelligent reading system that is expected to respond to ad-hoc requests for information, expressed in natural language questions by human users.
 - **MultiHop Question Answering** *2019*
The aim is to answer questions which require reasoning over multiple supporting documents.
 - **Deep Learning for detecting Hate Speech Tweets** *2018*
The aim is to identify abusive language, flag offensive content using natural language processing.
-

Miscellaneous

- Actively reviewing for ACL, TMLR, ICLR, ECCV, NeurIPS and CoNLL. *2021 - Present*
 - **Google Summer of Code** and **Google CodeIn** mentor. *2018*
 - Ranked 1st out of 100+ teams in **CODETRIX** (National level coding contest). *2017*
 - Ranked 3rd out of 100+ teams in **CODE-IT** (National level coding contest). *2016*
-

Programming Languages and Technologies

- Python, PyTorch, C++, C, Shell, Git
- NLP, Multimodal vision & language, Reinforcement learning