Atharva Kulkarni

Ph.D. in Computer Science | University of Southern California

🌐 athrvkk.github.io @ atharva.kulkarni@usc.edu github.com/athrvkk 📣 Google Scholar in linkedin.com/in/athrvkk
🐦 x.com/athrvkk 👤 athrvkk.bsky.social

Research Interests: Theoretical & Empirical Understanding of LLMs | Robust, Generalizable, & Trustworthy NLP

Education

Present
Aug 2024
	University of Southern California | Thomas Lord Dept. of Computer Science
	Doctor of Philosophy (Ph.D.) in Computer Science | GPA: 3.85/4.00
	  ❯ Advisor: Prof. Swabha Swayamdipta
	  ❯ Viterbi School of Engineering Fellowship (AY 2024-25)
		Los Angeles, CA

May 2025 **Jan 2025**	**University of California, Berkeley | Simons Institute for the Theory of Computing** Visiting Ph.D. Student | Program: Special Year on LLMs and Transformers (Part 2) [🌐]	**Berkeley, CA**

May 2024
Aug 2022
	Carnegie Mellon University | Language Technologies Institute
	Master of Science (M.S.) in Language Technologies | GPA: 4.04/4.00
	  ❯ Advisors: Prof. Graham Neubig & Prof. Barnabás Póczos
	  ❯ Collaborators: Prof. Aditi Raghunathan, Prof. Ameet Talwalkar, Prof. Emma Strubell
	  ❯ CMU LTI Research Fellowship (AY 2023-24)
		Pittsburgh, PA

May 2020 **Aug 2016**	**Savitribai Phule Pune University | PVG's College of Engineering and Technology** Bachelor of Engineering (B.E.) in Computer Science | GPA: 9.1/10	**Pune, India**

Experience

Aug 2024 **May 2024**	**Apple Inc. | Siri & Information Intelligence (SII)** *Machine Learning Research Intern* | *Mentors: Yuan Zhang & Hong Yu* *Collaborators: Joel Moniz, Xiou Ge, Bo-Hsiang Tseng, & Dhivya Piraviperumal* ❯ Evaluating evaluation metrics for hallucination detection in LLM-generated texts (under review ACL'25).	**Seattle, WA**
Aug 2023 **May 2023**	**Apple Inc. | Siri & Information Intelligence (SII)** *Machine Learning Research Intern* | *Mentors: Shruti Bhargava & Hong Yu* *Collaborators: Bo-Hsiang Tseng, Joel Moniz, & Dhivya Piraviperumal* ❯ Synthetic data generation using LLMs for dialog state tracking. (*EACL'24*)	**Seattle, WA**

July 2022
March 2021
	IIIT Delhi | Laboratory for Computational Social Systems (LCS2) [🌐]
	Predoctoral Researcher | *Advisor: Prof. Tanmoy Chakraborty*
	  ❯ Hate speech detection under subpopulation and covariate shifts. (*SIGKDD'23*)
	  ❯ Semantic role labeling of multimodal memes using optimal transport. (*EACL'23*)
	  ❯ Claim span identification using structural constraints and domain-specific rationales. (*EMNLP'22*)
	  ❯ Sarcasm explanation generation in mult-imodal multi-party dialogues. (*ACL'22*)
		Delhi, India

July 2022 **Dec 2021**	**IIT Bombay | Cognitive & Behavioural Neuroscience Laboratory [🌐]** *Research Collaborator* | *Advisor: Prof. Rashmi Gupta* ❯ Evaluating LLM performance in diagnosing comorbid mental health conditions, identifying their key limitations, and creating a novel dataset to address these gaps. (*EMNLP'24*)	**(Remote) Powai, India**

Selected Publications (* denotes equal contribution)

[11] **Evaluating Evaluation Metrics – The Mirage of Hallucination Detection**
<u>Atharva Kulkarni</u>, Yuan Zhang, Joel Moniz, Xiou Ge, Bo-Hsiang Tseng, ..., Swabha Swayamdipta, Hong Yu.
Annual Meeting of the Association for Computational Linguistics   **[Under review ACL'25]**

[10] **Still Not Quite There! Evaluating Large Language Models for Comorbid Mental Health Diagnosis** [PDF | code]
Amey Hengle*, <u>Atharva Kulkarni</u>*, Shantanu Patankar, ..., Rashmi Gupta
Conference on Empirical Methods in Natural Language Processing   **[EMNLP'24]**

[9] **Multitask Learning Can Improve Worst-Group Outcomes** [PDF | code]
<u>Atharva Kulkarni</u>*, Lucio Dery*, Amrith Setlur, Aditi Raghunathan, Ameet Talwalkar, Graham Neubig
Transactions of Machine Learning Research   **[TMLR'24]**

[8] **SynthDST: Synthetic Data is All You Need for Few-Shot Dialog State Tracking** [PDF | code]
Atharva Kulkarni, Bo-Hsiang Tseng, Joel Moniz, Dhivya Piraviperumal, Hong Yu, Shruti Bhargava.
Conference of the European Chapter of the Association for Computational Linguistics **[EACL'24]**

[7] **The Student Becomes the Master: Matching GPT3 on Scientific Factual Error Correction** [PDF | code]
Dhananjay Ashok, Atharva Kulkarni, Hai Pham, Barnabás Póczos.
Findings of the Association for Computational Linguisticss **[Findings EMNLP'23]**

[6] **Counting the Bugs in ChatGPT's Wugs: A Multilingual Investigation into the Morphological Capabilities of a Large Language Model** [PDF | code]
Leonie Weissweiler, Valentin Hofmann, ..., Atharva Kulkarni, ..., Hinrich Schuetze, Kemal Oflazer, David Mortensen.
Conference on Empirical Methods in Natural Language Processing **[EMNLP'23]**

[5] **Learning & Reasoning Multifaceted Longitudinal Data for Poverty Estimates and Livelihood Capabilities of Lagged Regions in Rural India** [PDF] [AI for Social Good Award]
Atharva Kulkarni, Raya Das, Ravi Srivastava, Tanmoy Chakraborty
International Joint Conference on Artificial Intelligence **[IJCAI'23]**

[4] **Revisiting Hate Speech Benchmarks: From Data Curation to System Deployment** [PDF | code]
Atharva Kulkarni*, Sarah Masud*, Vikram Goyal, Tanmoy Chakraborty
ACM SIGKDD Conference on Knowledge Discovery and Data Mining **[KDD'23]**

[3] **Characterizing Entities in Harmful Memes: Who is Hero, Villain, Victim?** [PDF | code]
Shivam Sharma*, Atharva Kulkarni*, . . ., Preslav Nakov, Md. Shad Akhtar, Tanmoy Chakraborty
Conference of the European Chapter of the Association for Computational Linguistics **[EACL'23]**

[2] **Empowering the Fact-checkers! Automatic Identification of Claim Spans on Twitter** [PDF | code]
Megha Sundriyal*, Atharva Kulkarni*, Vaibhav Pulastya, Md. Shad Akhtar, Tanmoy Chakraborty
Conference on Empirical Methods in Natural Language Processing **[EMNLP'22]**

[1] **When did you become so smart, oh wise one?! Sarcasm Explanation in Multi-modal Multi-party Dialogues** [PDF | code]
Shivani Kumar*, Atharva Kulkarni*, Md. Shad Akhtar, Tanmoy Chakraborty
Annual Meeting of the Association for Computational Linguistics **[ACL'22]**

Academic Service

Conference Reviewer	COLM'25, ARR'25, TMLR'24, ARR'24, EMNLP'23, ACL'23, EMNLP'22, ICON'22
Workshop Reviewer	WiNLP'24, SSLTheoryPractice NeurIPS'23, R0-FoMo NeurIPS'23
Conference Volunteer	EACL'24, EACL'23, ACL'21, EACL'21

Teaching & Mentorship Roles

Mentor – CS Undergraduate Mentorship Program | **USC** Fall'24
> Mentoring USC undergraduate students for starting out in research (Mentee – Tshephang Makubate & Grace Li).

Teaching Assistant – Advanced NLP | **CMU** Fall'23
> TA for 11-711 Advanced NLP taught by Prof. Daniel Fried. Designed assignments and guided teams for course projects.

Teaching Assistant – Introduction to NLP | **CMU** Spring'23
> TA for 11-611 Intro to NLP taught by Prof. David Mortensen. Designed assignments and guided teams for course projects.

Mentor – AI Undergraduate Mentorship Program | **CMU** Fall'22
> Mentored CMU undergraduate students for starting out in research (Mentee – Yi (Nicole) Xiang).

Talks

SynthDST: Synthetic Data is All You Need for Few-Shot Dialog State Tracking
> NiklasOPF AI podcast [video link] Nov 2024 – YouTube
> EACL'24 Oral Presentation March 2024 – St. Julian's, Malta
> Siri & Information Intelligence (SII), Apple. Inc Aug 2024 – Seattle, WA

When did you become so smart, oh wise one?! Sarcasm Explanation in Multi-modal Multi-party Dialogues
> Student Research Symposium, CMU LTI Aug 2022 – Pittsburgh, PA

Honours & Awards

1. USC Viterbi Fellowship (2024-25) – Awarded to selected incoming PhD student at USC, covering full tuition & stipend.

2. CMU LTI Research Fellowship (2023-24) – Awarded to selected MLT students at CMU, covering full tuition & stipend.

3. Best Paper – AI for Social Good | IJCAI'23 [⚲] – Poverty estimation using neural networks for rural Indian setting.

4. Shared Task Winner | WASSA @ EACL'21 [⚲] – Empathy & distress detection in reaction to news stories.

5. Shared Task Winner | TechDoFication @ ICON'20 [⚲] – Coarse-grained technical domain identification of documents for indigenous Indian languages.

6. Bachelor's Thesis Award | ZS ASPIRE 2020 [⚲] – 2nd place amongst 33,000 teams & cash prize of 2,00,000 INR for Bachelor's Thesis project.

References

Prof. Swabha Swayamdipta   *Assistant Professor, University of Southern California, USA* [🌐, @]

Prof. Graham Neubig   *Associate Professor, Carnegie Mellon University, USA* [🌐, @]

Prof. Barnabás Póczos   *Associate Professor, Carnegie Mellon University, USA* [🌐, @]