Rajat Kulshreshtha

in LinkedIn   𝕘 Google Scholar   🌐 Website   ✉ rk.kuls@gmail.com   📞 +1-732-397-3890

Education

Carnegie Mellon University	Pittsburgh, PA
Master of Language Technologies (Research Master's), Language Technologies Institute	2018
Indian Institute of Technology, Guwahati	Guwahati, India
B. Tech., Electronics & Electrical Engineering (with a Minor in Computer Science)	2015

Work Experience

Amazon AGI Foundations   Aug 2022 – Present
Applied Scientist

- Conduct LLM research as part of the core modeling team focusing on multimodal capabilities with real-time user experience for Amazon's Nova family of LLMs.

- Focus on post-training data and inference optimization for Nova Micro, Lite, Pro models
  - Led speculative decoding and quantization efforts achieving ~4.5x speedup while preserving model performance
  - Post-training data curation for long-tail recognition enabling ~1% WERR improvement on rarewords

- Led the design and implementation of robust evaluation and release pipelines for on-device language models, enabling scalable benchmarking and deployment.

Telling.ai   Oct 2018 – Jul 2022
Founding Engineer - CMU Startup

- Led development of cloud-based, speech-first AI systems for longitudinal monitoring of respiratory diseases such as COPD using voice and breath analysis.

- Built and maintained production-grade pipelines for data preprocessing, model training, and real-time inference, ensuring safety and reliability for clinical use.

- Developed models, published peer-reviewed studies, and co-invented patents that were crucial in securing an NSF SBIR grant and series A funding.

Goldman Sachs, Bangalore   Jun – Nov 2015
Senior Analyst, Securities Division

- Implemented automated surveillance of orders using Monte Carlo simulations on exchange data.

Marooner Technologies   Apr 2014 - May 2015
Cofounder - IIT Guwahati startup

- Developed a virtual trial room software that generates target apparel images on the user.

Research Experience

Carnegie Mellon University   Aug 2016 – Aug 2018
Graduate Research Assistant – Advisors: Prof. Bhiksha Raj, Prof. Florian Metze

- Conducted research on adversarial robustness in ASR systems, designing and evaluating targeted perturbations to expose model vulnerabilities and inform safer deployment strategies.

- Developed adaptive speech and conversation models to study early language acquisition in infants.

Research Intern – Advisors: Prof. Bhiksha Raj, Prof. Rita Singh   Dec 2013; May – Jul 2014

- Built an information mining system for Polly, a voice-based service reaching low-literate populations in India, delivering developmental content through simple telephonic interactions.

- Analyzed large-scale call data to identify usage patterns, track network growth, and inform content design for maximum impact.
- Created a novel method using Bayesian inference and semantic analysis to find the safest travel route based on modeling of crime distribution from police records and news feeds.

Hanyang University, South Korea   *May – Jul 2013*
Research Intern – Advisor: Prof. Frank Rhee

- Implemented iterative versions of fuzzy clustering algorithms to improve classification accuracy and computational efficiency for large-scale datasets.

Patents

- Systems and methods for analyzing and monitoring lung function using voice and breath sound samples for respiratory care. US Patent 11,937,911 published 2024.
- Methods and systems for voice profiling as a service. US Patent App. US20210020191A1

Publications

- LLM Powered After Visit Summary (Poster, selected for Panel Discussion). American Academy of Ophthalmology, 2024. [live demo]
- Voice based monitoring of COPD, CHEST Journal, 2021.
- Voice based screening & monitoring of chronic respiratory conditions, CHEST Journal, 2021.
- N Ryant et al. *Enhancement and Analysis of Conversational Speech: JSALT 2017*, ICASSP 2018.
- A Ravichander*, S Rijhwani*, R Kulshreshtha*, C Nagpal, T Baltrušaitis, LP Morency. *Preserving Intermediate Objectives: One Simple Trick to Improve Learning for Hierarchical Models*, arXiv 2017.
- AA Raza, R Kulshreshtha, S Gella, S Blagsvedt, M Chandrasekaran, B Raj, R Rosenfeld. *Viral Spread via Entertainment and Voice-Messaging Among Telephone Users in India*, ICTD 2016.
- R Kulshreshtha, V Sharma, P Singh, N Agrawal, A Kumar. *Analyzing Newspaper Crime Reports for Identification of Safe Transit Paths*, HLT-NAACL 2015.
- V Sharma, R Kulshreshtha, A Kumar, N Agrawal, P Singh. *Image summarization using Topic Modeling*, IEEE ICSIPA 2015.

Academic Achievements

- Represented India at the International Junior Science Olympiad 2008 held in South Korea, and won a Bronze Medal.
- National Topper, Panini National Linguistics Olympiad, India 2014.
- Awarded Gandhian Young Technological Innovation Award, 2014.
- Graduate Student Fellowship at Carnegie Mellon University 2016-2017.
- Awarded Institute Merit Scholarship for topping the department, IIT Guwahati 2012.
- Won Best Project Award at the Carnegie Mellon IPTSE Winter School 2013.
- Kishore Vaigyanik Protsahan Yojana (KVPY) Scholar.
- National Talent Search Examination (NTSE) Scholar.
- Rank 24 (Team trojans), ACM-ICPC, India, 2012.

Mentorship and Community

- Invited Speaker - STEM Symposium 2020 at Fox Chapel Area High School, Pittsburgh.
- Attended Jelinek Summer Workshop on Speech and Language Technology (JSALT) 2017.
- Teaching Assistant for Deep Learning (2018) and NLP courses (2017) at Carnegie Mellon University
- Volunteer - Greater Pittsburgh Community Food Bank 2024-2025