

Monica Munnangi

✉ monicamunnangi23@gmail.com 🏠 khoury/people/monica-munnangi/

RESEARCH INTERESTS

LLMs, Foundation Model Reliability, Robustness & Evaluation, Multi-modality, Retrieval-Augmented Generation (RAG), Clinical NLP, AI Safety & Alignment, Large-scale Pretraining, Human-AI Interaction

EDUCATION

- **Northeastern University** Boston, MA
PhD in Computer Science Sep 2021 - Present
 - **Advisor:** Prof. Saiph Savage
- **University of Massachusetts Amherst** Amherst, MA
Master of Science in Computer Science Sep 2018 - May 2020
 - **Research:** Information Fusion Lab, advised by Prof. Madalina Fiterau
- **Vellore Institute of Technology** Chennai, India
Bachelor of Technology in Computer Science and Engineering Jul 2014 - May 2018

PUBLICATIONS

- **M. Munnangi**, Saiph Savage, **ThReadMed-QA: A Multi-Turn Medical Dialogue Benchmark from Real Patient Questions**, *Under Review* <https://arxiv.org/abs/2603.11281>
- **M. Munnangi***, Swaminathan A.*, Fries J.A*, et al., **FactEHR: A Dataset for Evaluating Factuality in Clinical Notes Using LLMs**, *Accepted to MLHC 2025* <https://arxiv.org/abs/2412.12422>
- AMC Arroyo*, **M. Munnangi***, et al. **Open (Clinical) LLMs are Sensitive to Instruction Phrasings**, *Published at BioNLP (ACL 2024) at Bangkok, Thailand* <https://arxiv.org/abs/2407.09429>
- **M. Munnangi**, et al., **On-the-fly Definition Augmentation of LLMs for Biomedical NER**, *Published at NAACL 2024 at Mexico City, Mexico* <https://arxiv.org/abs/2404.00152>
- Y. Shah, **M. Munnangi**, et al., **Chest Tube Detection on Chest X-Ray Images Using Convolutional Deep Neural Network**, *Poster at European Congress of Radiology, Vienna 2020*

RESEARCH / INDUSTRY EXPERIENCE

- **Northeastern University** Boston, MA
Graduate Research Assistant Sep 2021 - Present
 - **Engineered scalable training and evaluation pipelines** in PyTorch, successfully sharding large foundation models (e.g., 70B parameters) across GPU clusters using DeepSpeed and PyTorch Lightning.
 - **Designing novel evaluation frameworks and multi-turn datasets** to rigorously benchmark LLM factuality, domain adaptation, and real-world reliability.
 - **Evaluated instruction sensitivity** of clinical-tuned models, published at **BioNLP at ACL, 2024**.
- **Stanford University** Palo Alto, CA
Research Intern - Shah Lab May 2024 - Aug 2024
 - **Developed and open-sourced** a comprehensive dataset and entailment-based evaluation framework to detect critical omissions and hallucinations in clinical LLMs.
 - **Built automated pipelines** for large-scale data curation and factuality evaluation, shifting away from rigid metrics (like ROUGE) to atomic-fact decomposition for safe AI deployment.
 - Published FactEHR: dataset for evaluating factual accuracy in clinical LLMs, addressing critical reliability challenges for AI safety applications. Accepted to **Machine Learning for Healthcare, 2025**.
- **Allen Institute for AI** Seattle, WA
Research Intern - Semantic Scholar May 2023 - Aug 2023
 - Enhanced LLM performance through knowledge augmentation techniques in biomedical domains, improving model accuracy and reducing hallucinations.

- Developed ablation studies to understand model behavior and failure modes. **Work published at NAACL 2024**

- **GE Healthcare**

Waukesha, WI

Data Scientist Intern

May 2019 - Aug 2019

- Developed a neural network to identify the presence of a chest tube in an Pneumothorax patient's X-Ray, trained the model on 8000 images and fine tuned on the pre-trained VGG architecture.
- Our results have surpassed the SOTA with 0.95 accuracy and this model is in production now which is helping radiologists prioritize high risk patient cohort using PyTorch framework.

TEACHING / ADVISING EXPERIENCE

- Teaching Assistant for **HCI seminar** and assisted Prof. Saiph Savage at Northeastern University in Spring 2026 semester.
- Teaching Assistant for **Unsupervised Machine Learning and Data Mining** and assisted Prof. Pavlu Virgil at Northeastern University in Spring 2023 semester.
- Teaching Assistant for **Unsupervised Data Mining** and assisted Prof. Pavlu Virgil at Northeastern University in Fall 2022 semester.
- Co-advised a cohort of graduate students for a project titled Naik, A. et al. **Leveraging knowledge distillation for efficient on-device deployment of deep learning models in medical imaging** published in Society for Imaging Informatics in MCMI in Medical Imaging, Nov 2020.
- Teaching Assistant for the course **Database Management Systems** and assisted Prof. Muralidhar A. at Vellore Institute of Technology in the Fall 2017 semester.

ACADEMIC SERVICE AND ACHIEVEMENTS

- Co-chair for Conference on Health, Inference, and Learning 2023 and 2024
- Program Committee member for Human-centered LLMs workshop, ACL 2024
- Reviewer: NeurIPS 2022, AAAI 2022, ML4H 2020, 2021, 2022, 2023, 2024, 2025, MLHC 2025, 2026
- Program Committee member for GenAI4Health workshop, NeurIPS 2025

SKILLS

- Python, PyTorch, C++, Sklearn, Numpy, LangChain, Git, R, distributed training / HPC (e.g., Slurm), Docker, DeepSpeed, vLLM