Monica Munnangi

CONTACT INFORMATION	2228F, Floor 22 177 Huntington, Boston MA - 02115 USA	Profile: khoury.northeastern.edu/people/monica-munnangi/ Webpage: monicamunnangi.github.io/ E-mail: monicamunnangi23@gmail.com

RESEARCH INTERESTS
: Machine Learning, Clinical Natural Language Processing for Healthcare, Retrieval for Large Language Models, Retrieval Augmented Generation

EDUCATION

Khoury College of Computer Sciences, Northeastern University — Boston, MA
Doctor of Philosophy in Computer Science — Sep 2021 - Present

CICS, University of Massachusetts Amherst — Amherst, MA
Masters of Science in Computer Science — Sep 2018 - May 2020

Vellore Institute of Technology — Chennai, IN
Bachelor of Science in Computer Science and Engineering — Aug 2014 - May 2018

PUBLICATIONS

M. Munnangi et al. **On-the-fly Definition Augmentation of LLMs for Biomedical NER**, *To appear 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 EXPERIENCE

Zak Lab, Harvard Medical School — Boston, MA
Lead - Matthew McDermott — Sep 2023 - Present
- Working on retrieval augmented generation for clinical task prediction.

Semantic Scholar, Allen Institute for AI — Seattle, WA
Research Intern, Ph.D. — May 2023 - Sep 2023
- Working on few shot domain adaptation with LLMs in biomedical and scientific domain. Benchmarking IE results on these datasets with state-of-the-art methods for inference with LLMs.
- Working on methods to improve inference performance of LLMs in knowledge intensive domains.

Clinical NLP Lab, Khoury College of Computer Sciences — Boston, MA
Advisor - Byron Wallace and Silvio Amir — Sep 2021 - May 2026
- Working on zero shot and few shot predictions of large language models in the clinical domain.
- Research interests lie in the areas of clinical natural language processing, multi-modal learning, learning from limited labeled data.

Krishnaswamy Lab, Yale School of Medicine — New Haven, CT
Advisor - Smita Krishnaswamy — Jul 2020 - Jun 2021
- Worked on classification and regression problems with recurrent neural networks on time series data of ICU patients and visualizing the patterns in data with sophisticated techniques.
- Worked on a natural language processing model to classify patient physician communication and to improve message triage.

Information Fusion Lab, University of Massachusetts Amherst — Amherst, MA
Advisor - Madalina Fiterau — Feb 2020 - May 2020
- Implemented a novel forecasting framework which utilizes a CNN to extract features from a patient's brain MRIs which we then fused with patient data and use RNN to track progression.

- Showed that the inclusion of these customised/patient-specific features increases the F1-score of 0.4644, with recall at 0.4974 and precision of 0.4355 of forecasting the disease stages.

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.
- Developed a solution using VGG architecture to separate obstructing radiopaque objects in a chest X Ray image from non-obstructing radiopaque objects with 14,000 images and achieved an accuracy of 0.89 to help reduce the number of redundant X-Rays.

Quantiphi Solutions, University of Massachusetts Amherst Amherst, MA
CS 696DS - Independent Study Jan 2019 - May 2019

- Used time series ICU data of over 40,000 patients and computed baselines, logistic regression and random forests to predict the onset of Sepsis as early as six hours.
- Benchmarked and compared our model results to baselines such as random forests, regression and validated the results where we have achieved an F1 score of 0.82.

TEACHING/ ADVISING EXPERIENCE

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.

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.

PROJECTS

Semi-supervised Named Entity Recognition for Clinical data UMass Amherst
CS 685 - Advanced NLP Mar 2020 - Apr 2020

- The aim of the project was to make annotations for named entity recognition faster by using semi-supervised learning techniques exclusively for clinical data.

Auto Generation of Image Captions for Medical Images UMass Amherst
CS 682 - Neural Networks Oct 2019 - Nov 2019

- We worked on automatic image captioning for medical images, used the IU chest X-Ray images which have 3965 unique patient reports and images. We have achieved a 0.168020 BLEU-1 score for the dataset.

PROFESSIONAL EXPERIENCE

DoctorC (Simplify Wellness Pvt. Ltd) Hyderabad, IN
Software Developer Intern Jan 2018 - Apr 2018

- Enhanced user experience on iOS with development of key interface changes using Xcode and React Native which improved usability by 25% and worked on a REST API service.

Autochat.io Hyderabad, IN
Software Developer Intern Sep 2017 - Dec 2017

- Created an English learning bot using telegram API, helps correct syntactic and semantic errors.
- Developed and deployed chat bots for E-commerce applications which improved user interface.

Pixelvide Solutions Pvt. Ltd Hyderabad, IN
Software Developer Intern May 2017 - Jul 2017

- Designed, wire-framed, prototyped and developed corporate website using HTML5, CSS and JavaScript which is currently in production.

ACADEMIC SERVICE
- Communications chair (Organizing Committee) Conference on Health, Inference and Learning (CHIL), 2024
- Program Committee at Human-centered LLMs workshop, ACL 2024
- Logistics co-chair (Organizing Committee) for CHIL, 2023
- Reviewer: ML4H 2020, 2021, 2022, 2023
- Program Committee at User-centered Natural Language Processing Workshop, WWW 2022
- Student reviewer at Northeastern University's CS PhD Admissions Committee 2022

AWARDS AND GRANTS
- Student Grant for NeurIPS 2020 and EMNLP 2020
- Central Board of Secondary Education Excellence award for outstanding performance (AISSE).
- City topper, Science Olympiad Foundation - National Science Olympiad 2012

SKILLS AND TOOLS
- **Languages** : Python, R, JavaScript, HTML, CSS, SQL, LaTeX
- **Libraries and Frameworks** : TensorFlow, PyTorch, Sklearn, Numpy, Pandas, ReactNative, Angular
- **Applications and Tools** : Docker, DataMiner, Jira

VOLUNTARY WORK
- Volunteer at NAACL 2022
- Volunteer at the Un-workshop in Woman in ML (WiML) at ICML, 2020 and NeurIPS, 2020.
- Part of an event at DESIRE Society, Hyderabad - served children affected with HIV/AIDS.
- Lead Volunteer of student led organization - Orange Leaf, Hyderabad

LEADERSHIP EXPERIENCE
- Student representative for the School of Computer Science and Engineering, VIT University.
- Publicity and marketing head, responsible for managing the online and offline marketing events at VIT Chennai with over 5000 participants from more than 30 universities.
- School and literacy captain, responsible for managing the cohort of school cabinet, conduct and manage the events conducted in school for the academic year (2011- 2012).