Aakanksha Naik

✉ aakankshanaik19@gmail.com • 🌐 www.cs.cmu.edu/~anaik/

Research Interests

Transfer Learning, Learning with Limited Supervision, NLP for Healthcare, Information Extraction (especially for Expert Domains), Robust Model Evaluation

Education

Carnegie Mellon University - School of Computer Science   **Pittsburgh, PA**
Doctor of Philosophy in Language and Information Technologies (PhD)   *May 2022 (expected)*

- **Thesis:** Adapting to the Long Tail in Language Understanding
- **Advisor:** Prof Carolyn Rose
- QPA: 4.0/4.0

Carnegie Mellon University - School of Computer Science   **Pittsburgh, PA**
Masters in Language Technologies, (MLT)   *Aug 2018*

- **Advisor:** Prof Carolyn Rose
- QPA: 4.0/4.0

Birla Institute of Technology and Science (BITS) Pilani   **Rajasthan, India**
Bachelor of Engineering, BE (Hons.) in Computer Science   *Jul 2016*

- **Undergraduate Thesis:** Speeding Up Coreference Resolution
- Cumulative GPA: 9.68/10

Experience

National Institutes of Health Clinical Center RMD   **Pittsburgh, PA**
Machine Learning Engineer II, contracted by FedPoint Systems LLC   *Oct 2020 - present*

- **Manager:** Elizabeth Rasch, Julia Porcino
- **Project:** Extraction and Temporal Ordering of Events from Clinical Records

Allen Institute for Artificial Intelligence   **Seattle, WA**
Research Intern, Semantic Scholar   *June 2021 - Oct 2021*

- **Manager:** Tom Hope, Lucy Lu Wang, Sergey Feldman
- **Project:** Improving Clinical Outcome Prediction Models Using Medical Literature

Facebook   **Menlo Park, CA**
Research Intern, Assistant Reasoning Team   *May 2019 - Aug 2019*

- **Manager:** Pararth Shah
- **Project:** Zero-shot Dialog State Tracking Using Pretrained Language Models

Indian Institute of Science (IISc)   **Bangalore, India**
Undergraduate Thesis Student, Machine and Language Learning (MALL) Lab   *Dec 2015 - Jun 2016*

- **Advisor:** Prof. Partha Talukdar
- **Project:** Speeding up Coreference Resolution

Institute for Creative Technologies   **Playa Vista, CA**
Visiting Undergraduate Researcher   *May 2015 - Jul 2015*

- **Advisor:** Prof. Anton Leuski
- **Project:** Improving Language Understanding for Virtual Humans using Query Expansion

Homi Bhabha Center for Science Education   **Mumbai, India**
Summer Intern   *May 2014 - Jul 2014*

- **Advisor:** Prof G. Nagarjuna
- **Project:** Integrating Mailman Mailing Lists With Gnowledge Web Forum

Publications

1. **Naik, Aakanksha**, Jill Lehman, Carolyn Rose, Adapting Event Extractors to Medical Data: Bridging the Covariate Shift, *EACL 2021*
2. **Naik, Aakanksha**, Carolyn Rose, Adapting to the Long Tail in Language Understanding, *Adapt-NLP Workshop, EACL 2021 (non-archival)*
3. Yoder, Michael, Sopan Khosla, Qinlan Shen, **Aakanksha Naik**, Huiming Jin, Hariharan Muralidharan, Carolyn Rose, FanfictionNLP: A Text Processing Pipeline for Fanfiction, *Workshop on Narrative Understanding, NAACL 2021*
4. **Naik, Aakanksha**, Carolyn Rose, Towards Open Domain Event Trigger Identification using Adversarial Domain Adaptation, *ACL 2020*
5. Ravichander, Abhilasha*, **Aakanksha Naik***, Carolyn Rose, Eduard Hovy, EQUATE: A Benchmark Evaluation Platform for Quantitative Reasoning in Natural Language Inference, *CoNLL 2019*
6. **Naik, Aakanksha**, Luke Breitfeller, Carolyn Rose, TDDiscourse: A Dataset for Discourse-Level Temporal Ordering of Events, *SIGDIAL 2019*
7. **Naik, Aakanksha***, Abhilasha Ravichander*, Carolyn Rose, Eduard Hovy, Exploring Numeracy in Word Embeddings, *ACL 2019*
8. Yan, Xinru, **Aakanksha Naik**, Yohan Jo, Carolyn Rose, Using Functional Schemas to Understand Social Media Narratives, *Storytelling Workshop, ACL 2019*
9. **Naik, Aakanksha***, Abhilasha Ravichander*, Norman Sadeh, Carolyn Rose, and Graham Neubig, Stress Test Evaluation for Natural Language Inference, *COLING 2018* (Area Chair Favorite Paper)
10. **Naik, Aakanksha**, Chris Bogart, and Carolyn Rose, Extracting Personal Medical Events for User Timeline Construction using Minimal Supervision, *BioNLP 2017*
11. Chandu, Khyathi, **Aakanksha Naik**, Aditya Chandrasekar, Zi Yang, Niloy Gupta, and Eric Nyberg, Tackling Biomedical Text Summarization: OAQA at BioASQ 5B, *BioNLP 2017*
12. Lohmueller, Lisa, **Aakanksha Naik**, Luke Breitfeller, Colleen McIlvennan, Manreet Kanwar, Srinivas Murali, Carolyn Rose, James Antaki, Factors Affecting Healthcare Engagement by Patients with Severe Heart Failure: An Exploratory Analysis, *Medical Decision Making Policy & Practice*
13. **Naik, Aakanksha R.**, Lisa C. Lohmueller, Carolyn P. Rose, and James F. Antaki, Improving Decision Support Tools for LVAD Recipients via Computational Analyses of Online Support Group Discussions, *Journal of Cardiac Failure*
14. **Naik, Aakanksha R.**, Lisa C. Lohmueller, Carolyn P. Rose, and James F. Antaki, A Computational Analyses of Online Support Group Discussions: Identifying Coping Strategies in LVAD Patients, *Society for Medical Decision Making* (Awarded Lee B. Lusted Student Prize)

Preprints

1. **Naik, Aakanksha**, Jill Lehman, Carolyn Rose, Adapting to the Long Tail: A Meta-Analysis of Transfer Learning Research for Language Understanding Tasks (*under review*)
2. **Naik, Aakanksha**, Sravanthi Parasa, Sergey Feldman, Lucy Lu Wang, Tom Hope, Literature-Augmented Clinical Outcome Prediction (*under review*)
3. Breitfeller, Luke, **Aakanksha Naik**, Carolyn Rose, STAGE: Tool for Automated Extraction of Semantic Time Cues to Enrich Neural Temporal Ordering Models

Academic Honors and Awards

- National Institutes of Health (NIH) pre-doctoral fellowship, 2020-2022
- ISCA Travel Grant to attend SIGDIAL, 2019

- Area Chair Favorite Paper at COLING, 2018
- Lee B Lusted Student Prize from Society for Medical Decision Making, 2017
- Fellowship in Digital Health, Center for Machine Learning and Health, Carnegie Mellon University, 2017
- Viterbi-India Scholarship, 2015
- BITS-Pilani Merit Scholarship, 2012-2016
- Dhirubhai Ambani Scholarship for academic excellence, 2012-2016

Professional Service

Organization
- Co-organizer, Pattern-based Approaches to NLP in the Age of Deep Learning (PAN-DL) COLING 2022
- Academic Inclusion Co-chair, Diversity and Inclusion Committee NAACL 2022
- Academic Inclusion Co-chair, Diversity and Inclusion Committee ACL 2020
- Mentorship Co-chair ACL 2019
- Member, Academic Inclusion Group, Diversity and Inclusion Committee NAACL 2019

Reviewing
- **Conferences:** ACL Rolling Review, ACL 2020-21, EMNLP 2020 (**Outstanding Reviewer**), 2021, NAACL 2019, EACL 2021, COLING 2020
- **Workshops:** ACL-SRW 2020-21, WiNLP 2019-21, LOUHI 2020-21, LTI-SRS 2018-19, SMDM, ESMDM, NASSLLI (2018)
- **Journals:** NEJLT

Invited Talks

- **Adapting Information Extractors to Long Tail Domains**
  - Salesforce Research, Feb 2022
  - BITS Pilani Alumni Research Talks (ART), Feb 2022
- **Adapting to Long Tail Domains: A Case Study in Clinical Information Extraction**
  - Amazon AWS AI, Jan 2022
  - CURAI External Speaker Series, Jan 2022
  - Microsoft Research, Nov 2021
  - Allen Institute for Artificial Intelligence, Nov 2021
- **Event Extraction from Medical Text: A Glimpse into the Semantics of the Long Tail**
  - National Institutes of Health Epidemiology and Biostatistics Section, May 2020

Teaching

- Teaching Assistant, Machine Learning for Structured Data (Masters-level) Fall 2019
- Teaching Assistant, Introduction to Machine Learning (Masters-level) Fall 2018

Personal

- **Language Skills:** English, Hindi, Gujarati, Marathi (Native/Bilingual), Konkani (Limited)
- Interested in painting and sketching. **Artwork:** https://www.cs.cmu.edu/~anaik/art.html.
- Volunteered for Nirmaan, a student-run not-for-profit organization at BITS-Pilani for three years, tutoring underprivileged primary school kids in mathematics and science.
- Studied Hindi and Sanskrit literature. Reached "Parichay" level in national exams conducted by the Rashtrabhasha Hindi Parishad and "Madhyama" level in exams conducted by the "Bruhad Gujarati Sanskrit Parishad" (with "First class" in all levels).