Sweta Karlekar

Email: sweta.karlekar@gmail.com Phone: (703) 732-7048 LinkedIn: swetakarlekar Website: https://sweta.dev

Education

Columbia University Aug 2023-present
Research Advisor: Dr. David Blei
Ph.D. Computer Science
Research Specialization in Machine Learning, Bayesian Statistics, Causal Inference, and NLP
Fellowships: Provost's Diversity Fellowship
Relevant Course Work: Probabilistic Graphical Models and Machine Learning, Computation and the Brain, NLP

University of North Carolina at Chapel Hill Aug 2016-May 2020
Research Advisor: Dr. Mohit Bansal
B.S. Computer Science, Entrepreneurship Minor -- GPA: 3.87/4.00, Major GPA: 3.94/4.00
Scholarships: STEM Diversity Scholarship, Chancellor's Science Scholarship (Full Ride, Academic Merit)
Honor Societies: Phi Beta Kappa Honor Society, Sigma Xi Scientific Research Honor Society, Honors Carolina
Relevant PhD Course Work: Deep Learning and Natural Language Processing, Conversational Models in AI

Refereed Publications

Jesson, A., Beltran-Velez, N., Chu, Q., **Karlekar, S.**, Kossen, J., Gal, Y., Cunningham J. & Blei, D. (2024). *Estimating the Hallucination Rate of Generative AI.* In NeurIPS 2024, Vancouver, Canada.
Karlekar, S., Bansal, M. (2018, Nov). *SafeCity: Understanding Diverse Forms of Sexual Harassment Personal Stories.* In EMNLP 2018, Brussels, Belgium .
Karlekar, S., Niu, T., Bansal, M. (2018, June). *Detecting Linguistic Characteristics of Alzheimer's Dementia by Interpreting Neural Models.* In NAACL 2018, New Orleans.
Karlekar, S., Bansal, M. (2018, June). *#MeToo: Neural Detection and Explanation of Language in Personal Abuse Stories.* In NAACL-WiNLP 2018, New Orleans.
Becker, S.J., Daughtry, C.S.T., Jain, J., & **Karlekar, S.** (2015, December). *Developing a Method to Mask Trees in Commercial Multispectral Imagery*. Poster presented at the American Geophysical Union Fall Meeting, San Francisco, CA.

Awards and Honors

Columbia Provost's Diversity Fellowship ($8,000) Sept 2023
CRA Outstanding Undergraduate Research Award Runner-Up Jan 2020
Ernest H. Abernethy Prize for Student Research Publication – Chancellor's Award Apr 2019
Phi Beta Kappa Honor Society Apr 2019 – present
EMNLP 2018 Student Scholarship for Travel, Lodging, and Conference Registration Sep 2018
Grace Hopper 2018 UNC Chapel Hill Scholarship May 2018
1st Place Undergrad Math & Computer Science Poster - National Sigma Xi Research Conference Nov 2017
STEM Diversity Scholarship – Full Scholarship (Tuition/Room/Board), Academic Merit Jun 2016 – May 2020
Chancellor's Science Scholars – 10k/yr Scholarship, Academic Merit Jun 2016 – May 2020

Relevant Work Experience

Capital One Machine Learning Data Science PhD Intern, *Generative AI, Encoding Models* May 2024 – Aug 2024

- Evaluated and benchmarked multiple text encoder models for effectiveness in encoding call transcripts, improving model selection for various summarization and similarity search tasks
- Developed a scalable evaluation framework using Kubeflow Pipelines, automating model performance testing and facilitating efficient experimentation

Meta Applied ML Research Engineer, *Gen AI Safety*, *Bayesian Modeling, & Causal Inference* Oct 2020 – Jul 2023

- Tech lead for GenAI text-to-image Safety; work included coordinating efforts of multiple Integrity and responsible AI teams & using prompt engineering, fine-tuning, red-teaming and other techniques to create safer products
- Founded and acted as tech lead for sub-team pillar of Causal Inference: responsibilities included sourcing, building, and supporting cross-functional (XFN) collaborations; roadmapping; driving architectural discussions; hosting and mentoring Ph.D. research interns; and motivating research directions
- Was lead engineer on multiple Bayesian + Causal Inference projects from prototyping to productionization
- Gained experience with various Bayesian models/techniques including Bayesian regressions, Gaussian Processes, conjugate priors, Bayesian Structural Time Series (BSTS), Bayesian contextual multi-armed bandits, etc., and with causal inference approaches like Difference-in-Difference, Regression Discontinuity, Instrumental Variables, Propensity Scores, Bayesian hypothesis testing, and missing data de-biasing, etc.

Facebook Machine Learning Engineering Intern, *Anomaly Detection* May – Aug 2019

- Designed and built an anomaly detection model to remove mismatched candidate recommendations as part of the Recruiting org; Implemented a voting ensemble method to combine performances of various models, including regressions, sparse NNs, and decision trees

Yelp Applied Machine Learning Intern, *Survival Analysis* Jan – May 2019

- Developed WTTE-RNN (Weibull Time to Event RNN) models to produce survival curves and predict business and advertiser retention
- Explored and created feature sets of various advertiser signals to deploy ML models at scale

Google AI Research Mentoring Program, *Deep Learning & Interpretability* Sep 2018 – Sep 2019

- Nominated by WiNLP workshop at NAACL to collaborate and be mentored by Google Brain researchers
- Mentored by Dr. Been Kim on interpretability and democratizing deep learning models

Walt Disney Company Emerging Technologies Intern, *Natural Language Topic Modelling* Aug – Dec 2018

- Leveraged Oracle ConText packages and SQL developer toolkits to extract salient themes from large amounts of SMS data
- Developed methods using regular expressions to automate responses for over 20% of incoming SMS messages
- Created models using Google's BigQuery ML to predict demographic information of users who visit Disney websites to better serve ads

MITRE Corp. Deep Learning and AI Intern, *Computer Vision and Generative Models* May – Aug 2018

- Utilized generative adversarial networks, specifically SD-GANs and PG-GANs, to generate training images for classifiers to detect objects in satellite imagery (Python)
- Performed literature reviews on the current state of neural computer vision and gave briefings on CNNs and CapsuleNets (https://github.com/swkarlekar/summaries)

UNC Chapel Hill Research Assistant, *Deep Learning and Natural Language Processing* Aug 2017 – May 2018

- First-authored three papers working in Dr. Mohit Bansal's Deep Learning + NLP Lab in the UNC CS Dept.
- Projects: (1) Used recurrent neural networks (RNNs) + natural language processing (NLP) techniques to identify the linguistic characteristics of early signs of Alzheimer's and dementia, (2) Used various neural models such as CNNs, RNNs, and CNN-RNN hybrids to perform text classification and visualization on circumstances of domestic abuse