

Swetha Pola

spola95@gmail.com | 609.619.1668 | [linkedin.com/in/swethapola](https://www.linkedin.com/in/swethapola) | github.com/swethapola

Data Scientist

Education

Master of Information and Data Science | UC Berkeley

August 2021 (expected)

Selected Coursework: *Machine Learning at Scale, Deep Learning, Statistics, Research Design, Causal Inference*

GPA: 3.9

B.A. Cognitive Science | UC Berkeley

May 2020

Selected Coursework: *Computational Models of Cognition, Neuropsychology, Social Psychology, Linguistics*

GPA: 3.6

Skills

- Python
- R
- SQL
- Hadoop
- Spark
- Cloud Computing
- Machine Learning
- Deep Learning
- Natural Language Processing
- Tableau
- Data Visualization
- Experimentation

Projects

Differential Gender-Based Outcomes of Cardiovascular Disease: BERT & Extracting Insights from Reddit UCSF Cardiology, UC Berkeley School of Information | [Repository](#)

- Developed RNN classifier to predict whether a CVD-related post originated from the r/askWomen or r/askMen subreddit and utilized LIME library to provide interpretability over blackbox model

Assessing the Effectiveness of a Social Norms Based Sexual Violence Prevention Digital Campaign

UC Berkeley PATH To Care Center, UC Berkeley School of Information | [Repository](#)

- Designed experimental study, developed Qualtrics study, cleaned data in Python, built linear regression models in R, and presented statistically conclusive causal results to campus Student Affairs leadership

Experience

Data Science Fellow | [D-Lab at UC Berkeley](#)

September 2020 - Present

- Lead data scientist on American Medical Association research project guiding research design & methods, performing topic modeling, driving large-scale labeling task through MTurk, and developing classification models in Python to conceptualize constructs of racism in medical literature
- Designing and instructing “Introduction to Neural Networks” 24-week curriculum in Python and R for Kaiser Permanente Division of Research Systems
- Mentored 6 undergraduate teams on data science research lifecycles, modeling approaches, and how to use technologies such as Azure ML, Github, and Tableau through the Data Science Discovery Program

Technical Product Manager | [Umba](#)

May 2020 - August 2020

- As sole PM leading a 12-person cross-functional team: set company-wide product vision, roadmapped & drove 12 weeks of product development, and spearheaded rollout of 5 core digital banking features for total active base of 300,000+ users
- Decreased loan default rates by 7% by working closely with senior data scientist on building, tuning, and deploying unsupervised fraud detection model into new user flow
- Led quantitative and qualitative user research via surveys, focus group, multiple stages of usability testing, and guiding customer segmentation project; synthesized findings into product & design decisions

Product & Analyst Intern | [JustAnswer](#)

June 2019 - August 2019

- Launched and validated lean multivariate test, successfully driving free-trialler conversion to subscription membership by engaging users via Concierge, a high touch customer relationship management product
- Developed dashboards in Power BI and analyzed qualitative data to make 4 additional phases of improvement, resulting in 15% lifts in Concierge engagement & 10% lifts in conversions overall