Ishita Gopal

Tel: +1 (925) 394-8793 Mail: ishitagopal@gmail.com Web: ishitagopal.github.io

I have 5 years of experience in translating abstract questions into concrete data problems. I leverage diverse data types (text, networks, images) and use cutting-edge techniques in machine learning, statistics, and experimentation to generate actionable insights.

SKILLS

Tools and Languages: Python, R, SQL, Git, AWS, Azure
Data Science & Statistics: Machine learning, natural language processing (NLP), deep learning, network analysis, inferential statistics, hypothesis testing, causal inference, time–series, panel-data models

WORK EXPERIENCE

Data Science Intern, Aware HQ **Summer 2022**

- Developed and deployed a credit card detection model with AWS SageMaker to flag sensitive data sharing in digital workspaces to be used in Aware's product.
- Used deep learning (EfficientNets) for transfer learning on hand-labeled data, utilized data augmentation techniques to reduce overfitting, improved model performance and achieved a 90% accuracy rate.

Doctoral Research Scientist, Center for Social Data Analytics (C-SoDA) **Aug 2018 – Current**
Advancing computational methods for social research. 3 publications, 4 working papers, 8 conference presentations.

- Led a team of 6 to analyze complex behavioral data of 4K policymakers using network simulation models on high performance computing. Found gender and race reliably predict cross-state connections.

- Delivered a high-impact journal article on when policymakers respond to public health crisis. Developed NLP pipelines. Fine-tuned large language models to identify COVID-19 discussions in 1M+ tweets of US legislators (F1 of 85%). Used hierarchical models and showed partisan divide moderated response to crisis severity.

- Collected protest, repression, and demographic data from 50 districts. Used machine learning and topic models to identify protest discourse in Telegram group chats. Modeled civic engagement's impact on government repression and censorship using panel regressions.

- Published results from 2 survey experiments (9000 participants) to show age and digital literacy limits the generalizability of experiments conducted on MTurk.

- Designed an experiment (1000 subjects) to test if peer influence affects legislators' support for environmental policies. Accounted for latent network effects and identified peers using backbone extraction methods. Identified relevant treatment policies using zero-shot text classifiers.

- Creating instructional materials (book, chapters, and tutorials) on computational methods for practitioners to collect and analyze unstructured text and network data.

Economist, The Energy & Resources Institute **Aug 2016 – Aug 2018**

- Developed time series (ARIMA) models for electricity demand forecasting.
- Used scenario modeling to forecast the impact of renewable uptake on coal capacity growth in India. Results provided policy assessment support to the Indian Government.

EDUCATION

Ph.D. Social Data Analytics & Political Science, *Pennsylvania State University* Expected 2023
(Awards: Princeton University Dissertation Scholar, C-SoDA Predoctoral Fellow)

M.Sc. Economics, *University of Warwick* 2015

B.A. (Hons) Economics, *Miranda House* 2014