

# Shachi Deshpande

PHD CANDIDATE · DEPARTMENT OF COMPUTER SCIENCE  
Cornell Tech and Cornell University, New York, NY 10044

## Education

---

### Cornell University

PHD IN COMPUTER SCIENCE

- Advisor: Prof. Volodymyr Kuleshov

New York

Present

### Cornell University

MS IN COMPUTER SCIENCE

- Advisor: Prof. Volodymyr Kuleshov

New York

Aug 2022

### Indian Institute of Technology, Bombay

BTECH WITH HONORS IN COMPUTER SCIENCE

- Advisor: Prof. Sudarshan

Mumbai

Aug 2018

## Research Interests

---

My research interests span probabilistic models, uncertainty estimation and causal machine learning. I work on improving the reliability of predictive uncertainties in the context of deep learning, causal inference and sequential decision-making. I also work on deep generative models for causal inference. I am interested in the applications of my research to the identification of causal genetic variants for traits in humans (Genome-Wide Association Studies) and in accurate quantification of uncertainties in genetic risk to develop a disease (Polygenic Risk Scores).

## Publications

---

### Probabilistic Conformal Prediction Improves Clinical Decision-Making.

**Shachi Deshpande**, Kamilė Stankevičiūtė, April Wei, Volodymyr Kuleshov. (In Preparation)

### Online Calibrated Regression for Adversarially Robust Forecasting.

Volodymyr Kuleshov and **Shachi Deshpande**. (In Submission)

### Calibrated Propensity Scores for Causal Effect Estimation.

**Shachi Deshpande** and Volodymyr Kuleshov.

*Uncertainty in Artificial Intelligence (UAI)*, 2024

Also appeared at *Spurious Correlations, Invariance and Stability (SCIS) Workshop*, ICML 2023

### Online Calibrated and Conformal Prediction Improves Bayesian Optimization.

**Shachi Deshpande**, Charles Marx and Volodymyr Kuleshov.

*International Conference on Artificial Intelligence and Statistics (AISTATS)* 2024

### Deep Multi-Modal Structural Equations For Causal Effect Estimation With Unstructured Proxies.

**Shachi Deshpande**, Kaiwen Wang, Dhruv Sreenivas, Zheng Li and Volodymyr Kuleshov.

*Advances in Neural Information Processing Systems (NeurIPS)* 2022

Also appeared at *Women in Machine Learning Workshop (WiML)*, NeurIPS 2022

### Calibrated and Sharp Uncertainties in Deep Learning via Simple Density Estimation.

Volodymyr Kuleshov and **Shachi Deshpande**.

*International Conference on Machine Learning (ICML)* 2022. (Spotlight)

Also appeared at *Distribution-Free Uncertainty Quantification (DFUQ) Workshop*, ICML 2022.

### New Genome Similarity Measures based on Conserved Gene Adjacencies.

Daniel Doerr, Luis Antonio B. Kowada, Eloi Araujo, **Shachi Deshpande**, Simone Dantas, Bernard M.E. Moret, and Jens Stoye.

*Journal of Computational Biology* 2017.

## Work Experience

---

### MULTIMODAL GENERATIVE CAUSAL MODELS FOR IMPROVING AGRICULTURAL YIELD MAY 2023 - AUG 2023

Swati Sharma, Angels de Luis Balaguer, Emre Kiciman and Ranveer Chandra | Microsoft Research, Redmond

We built deep generative causal models to answer causal queries from farmers as a part of the FarmBeats project. The conversational system aids decision making and improves the agricultural yield by estimating causal effects. The causal model incorporates multimodal covariates like satellite images of agricultural farms and historical weather time-series.

### LEARNING SPATIAL RELATIONSHIPS FOR ROBOTIC MANIPULATION JUNE 2022-AUG 2022

Dr. Chaitanya Mitash | Research at Amazon Robotics

We design vision transformers to perform spatial reasoning between objects in cluttered scene images, aimed at improving scene understanding and enabling robotic arm manipulation of new objects.

## Research

---

### PROBABILISTIC CONFORMAL PREDICTION FOR CLINICAL DECISION-MAKING NOV 2023-CURRENT

Prof. Volodymyr Kuleshov, Prof. April Wei | Cornell University

We design a unified framework to enforce probabilistic conformal and calibrated prediction of genetic risk to develop a disease. With reliable predictive PRS uncertainties for populations with distribution shifts, we demonstrate an improved ability to perform clinical decision-making on several semi-simulated and real datasets derived from the UK Biobank.

### CALIBRATION OF LEARNED PROPENSITY SCORE MODELS FOR CAUSAL INFERENCE NOV 2022-MAY 2023

Prof. Volodymyr Kuleshov | Cornell University

We propose probabilistic calibration to improve the effectiveness of propensity score models in causal effect estimation. Calibration reduces the variance and error bounds on causal effect estimates, thus allowing the use of simple propensity models that improve the computational throughput of Genome-Wide Association Studies by more than two-fold.

### DEEP MULTIMODAL STRUCTURAL EQUATIONS FOR CAUSAL INFERENCE AUG 2021-MAY 2022

Prof. Volodymyr Kuleshov | Cornell University

We incorporate multi-modal, unstructured information in modern datasets within the framework of causal inference using deep structural equations. We propose novel generative architectures and inference algorithms that scale to multi-modal setups with missing data. We create multimodal causal inference benchmarks for evaluation of causal effect estimation.

### CALIBRATING UNCERTAINTIES FOR BAYESIAN OPTIMIZATION JAN 2021-MAY 2021

Prof. Volodymyr Kuleshov | Cornell University

We propose a simple algorithm to calibrate the uncertainty of posterior distributions over the objective function as part of the Bayesian optimization process. We show that by improving the uncertainty estimates of the posterior distribution with calibration, Bayesian optimization makes better decisions and arrives at the global optimum in fewer steps.

### STREAMING QUERY OPTIMIZATION AUG 2017-MAY 2018

Prof. S Sudarshan | Undergraduate Thesis, IIT Bombay

We proposed dynamic tuple routing policies to optimize join computation in streaming query application within a distributed computing environment. We designed parallel query execution policies and demonstrated reduction in latency of computation without compromising the application throughput in simulated star join datasets.

### DERIVATIVE CLOUDS AUG 2017-MAY 2018

Prof. Umesh Bellur | R & D Project, IIT Bombay

We performed an empirical study of the impact CPU and memory overcommitment on application performance over a range of derivative setups and virtualization frameworks. We determined efficient Virtual Machine cluster configurations corresponding to application requirements through empirical validation of our models.

### MAXIMUM LIKELIHOOD ESTIMATION OF PHYLOGENETIC TREES JUNE 2016-AUG 2016

Prof. Bernard Moret | Summer@EPFL Scholar, EPFL, Switzerland

We inferred the evolutionary history of plant genomes using sequential genetic data using maximum likelihood estimation. We demonstrated significant reduction in the computational requirements of phylogenetic (evolutionary) tree reconstruction for plant species.

## Awards and Fellowships

---

|      |  |
|------|--|
| 2022 | <b>Doctoral Grant for Grace Hopper Celebration</b> , Cornell University      |
| 2022 | <b>NeurIPS Travel Grant</b> , Women in ML                                    |
| 2018 | <b>Cornell University Fellowship</b> , Cornell University                    |
| 2013 | <b>INSPIRE Scholarship</b> , Dept of Science and Technology, Govt of India   |
| 2013 | <b>Dhirubhai Ambani Scholarship</b> , Reliance Foundation                    |
| 2008 | <b>National Talent Search Examination Scholarship</b> , NCERT, Govt of India |

## Seminars and Talks

---

|           |   |
|-----------|---|
| Oct 2023  | <b>Causality Discussion Group</b> , TU Darmstadt, Germany (Virtual) |
| Oct 2023  | <b>Microsoft Research</b> , Cambridge, UK (Virtual)                 |
| Sept 2022 | <b>Machine Learning Reading Group</b> , Cornell Tech, NY, USA       |
| Aug 2022  | <b>Amazon Robotics</b> , Boston, MA, USA                            |

## Teaching Experience

---

|                           |  |
|---------------------------|--|
| Spring 2021               | <b>Deep Probabilistic and Generative Models</b> , Teaching Assistant |
| Fall 2020                 | <b>Applied Machine Learning</b> , Teaching Assistant                 |
| Summer 2020               | <b>Introduction to Programming Using Python</b> , Teaching Assistant |
| Spring 2020               | <b>Interactive Computer Graphics</b> , Teaching Assistant            |
| Fall 2017,<br>Spring 2018 | <b>Computer Programming and Utilization</b> , Teaching Assistant     |

## Key Coursework

---

Causal Machine Learning, Reinforcement Learning, Decision Theory, Topics in Machine Learning and Natural Language Processing (Seminars), Emerging Cloud Technologies, Model Checking, Computer Graphics, Information Retrieval, Probability Theory, Stochastic Processes

## Outreach & Professional Development

---

### SERVICE AND OUTREACH

|           |  |
|-----------|--|
| 2021-2022 | <b>Computer Science Graduate Organization, Cornell Tech</b> , Vice-President     |
| 2021-2022 | <b>Dept of Computer Science, Cornell</b> , Mentor for incoming Graduate students |
| 2019-2020 | <b>CS Graduate Admissions Committee, Cornell</b> , Reviewer                      |
| 2017-2018 | <b>Dept of Computer Science, IIT Bombay</b> , Academic mentor                    |

### REVIEWER

AISTATS 2022, 2023  
ICML 2023  
NeurIPS 2021, 2022, 2023

### EXTRACURRICULAR

|  |             |
|--|-------------|
| Coordinator for internship recruitment of undergraduate students at IIT Bombay | 2016-2017   |
| Panelist for BitStream Newsletter at IIT Bombay                                | 2015-2016   |
| Organizer of programming competition CodeBlitz at Techfest, IIT Bombay         | Winter 2014 |