# Sahil Verma

Paul G. Allen School of Computer Science and Engineering
University of Washington, Seattle, WA, USA
Work Email:
Personal Email:

Website
Google Scholar
vsahil@cs.washington.edu
v.sahil1@gmail.com

## Research Interests

My research is broadly focused on Trustworthy ML or Responsible AI, specifically focused on fairness, explainability, and robustness of ML.

## Education

| | |
|---|---|
| Sept 2019 - Present | PhD in Computer Science<br>**University of Washington, Seattle**<br>Advisors: Prof. Chirag Shah and Prof. John Dickerson |
| July 2015 - July 2019 | BTech in Electrical Engineering<br>**Indian Institute of Technology Kanpur (IIT Kanpur)**<br>Advisor: Prof. Subhajit Roy |

## Honors and Awards

| | | |
|---|---|---|
| 2020 | Best Paper Award and Nvidia Titan RTX GPU | ML-RSA Workshop at NeurIPS |
| 2019 | Allen School Fellowship | Paul G. Allen School, UW |
| 2018 | Student Travel Award of $1500 | ACM SIGPLAN |
| 2017 | Student Travel Award of $1800 | Google India |
| 2015 | All India Rank 663 | IITJEE Advanced Exam |
| 2015 | KVPY Fellow with All India Rank 205 | IISc Bangalore |

## Publications

Post-Hoc Attribute-Based Explanations for Recommender Systems
Sahil Verma, Anurag Beniwal, Narayanan Sadagopan, Arjun Seshadri
*TEA Workshop at NeurIPS 2022*

Amortized Generation of Sequential Counterfactual Explanations for Black-box Models
Sahil Verma, Keegan Hines, John P Dickerson
*AAAI 2022*

Counterfactual Explanations for Machine Learning: A Review
Sahil Verma, John Dickerson, Keegan Hines
*ML-RSA Workshop at NeurIPS 2020* **(Best Paper Award) (Citations 200+)**

Removing biased data to improve fairness and accuracy
Sahil Verma, Michael Ernst, Rene Just

Fairness Definitions Explained
Sahil Verma and Julia Rubin
*FairWare Workshop at ICSE 2017* **(Citations 700+)**

Facets of Fairness in Search and Recommendations
Sahil Verma, Ruoyuan Gao, Chirag Shah

*Algorithmic Bias Workshop at ECIR 2020*

ShapeFlow: Dynamic Shape Interpreter for TensorFlow
Sahil Verma and Zhendong Su

Debug-Localize-Repair: A Symbiotic Construction for Heap Manipulations
Sahil Verma and Subhajit Roy
*FMSD Journal 2021*

Synergistic Debug-Repair for Heap Manipulations
Sahil Verma and Subhajit Roy
*ESEC/FSE 2017*

## Patents

Amortized Generation of Sequential Counterfactual Explanations for Black-box Models
Sahil Verma, Keegan Hines, John P Dickerson
*U.S. Patent Application No.: 17/520,069*

## Work Experience

| | |
|---|---|
| June 2022 - Sept 2022 | Research Intern at Amazon, USA.<br>Developed novel post-hoc explainability technique for recommender systems. |
| June 2020 - Sept 2021 | Research Fellow at Arthur AI, USA.<br>Developed industry deployable ML explainability algorithm. |
| June 2019 - Sept 2019 | Research Intern at ETH Zurich, Switzerland.<br>Developed tensor shape incompatibility bugs detection in TensorFlow. |
| May 2018 - Aug 2018 | Research Intern at CSAIL, MIT, USA.<br>Developed tool for automating floating bit allocation in programs. |
| May 2017 - Aug 2017 | Research Intern at NUS, Singapore.<br>Developed tool to convert CSP programs into C code. |

## Professional Responsibilities

- **Reviewed research papers for:**
  - Workshops: AFCR 2021 (3), AFCP 2022 (7), HCAI 2022 (4).
  - Conferences: EAAMO 2021 (1), XAIF 2021 (3), NeurIPS 2022 (1), AAAI 2022 (1), ICML 2022 (1), FAccT 2022 (1), AIES 2022 (1), XAIF 2022 (2), AAAI 2023 (4).
  - Journals: IEEE Transactions on Artificial Intelligence (1), Data Mining and Knowledge Discovery (1), International Journal of Data Science and Analytics (1), Journal of Decision Systems (1), Computer and Operations Research (1), Machine Learning (2).
- **Student Volunteer** at ESEC/FSE 2017.

## Coursework

| | | |
|---|---|---|
| Computer Vision | Deep Learning | Fairness in Machine Learning |
| Machine Learning | Convex Optimization | Reinforcement Learning |
| Probability and Statistics | Linear Algebra | Reasoning for Software |

## Teaching Experience

**Teaching Assistant:** Machine Learning (CSEP546), Introduction to Machine Learning (CSE 416)