

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 specifically focused on robustness and interpretability in ML. I am available from **June 12th to October 1st 2024** for an internship.

EDUCATION

SEPT 2019 - Present	PhD in Computer Science University of Washington, Seattle Advisors: Jeff Bilmes and Chirag Shah
JULY 2015 - JULY 2019	BTech in Electrical Engineering Indian Institute of Technology Kanpur (IIT Kanpur) Advisor: Subhajit Roy

HONORS AND AWARDS

2024	Conference Travel Grant of \$1000	SATML 2024
2023	Best Oral Presentation Award	BUGS Workshop at NeurIPS
2022	Best Student Paper Award	TEA Workshop at NeurIPS
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

[Effective Backdoor Mitigation Depends on the Pre-training Objective](#)

SAHIL VERMA, GANTAVYA BHATT, SOUMYE SINGHAL, ARNAV DAS, CHIRAG SHAH,
JOHN P DICKERSON, JEFF BILMES
Best Oral Paper at *BUGS@NeurIPS 2023* and under submission at *CVPR 2023*

[RecRec: Algorithmic Recourse for Recommender Systems](#)

SAHIL VERMA, ASHUDEEP SINGH, VARICH BOONSANONG, JOHN P DICKERSON, CHIRAG SHAH
CIKM 2023 (Short Paper)

[Post-Hoc Attribute-Based Explanations for Recommender Systems](#)

SAHIL VERMA, CHIRAG SHAH, JOHN P. DICKERSON, ANURAG BENIWAL,
NARAYANAN SADAGOPAN, ARJUN SESHADRI
TEA Workshop at NeurIPS 2022 (Best Student Paper Award)

[Amortized Generation of Sequential Counterfactual Explanations for Black-box Models](#)

SAHIL VERMA, KEEGAN HINES, JOHN P DICKERSON
AAAI 2022

[Counterfactual Explanations and Algorithmic Recourses for Machine Learning: A Review](#)

SAHIL VERMA, VARICH BOONSANONG, MINH HOANG, KEEGAN HINES, JOHN P. DICKERSON, CHIRAG SHAH
ML-RSA Workshop at NeurIPS 2020 (Best Paper Award) (500+ Citations)

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 (1100+ Citations)

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

PROFESSIONAL RESPONSIBILITIES

- Reviewed research papers for:

- *Workshops*: AFCR 2021 (3), AFCEP 2022 (7), HCAI 2022 (4), RO-FoMo 2023 (2), BUGS 2023 (2).
- *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 (5).
- *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), IEEE Transactions on Knowledge and Data Engineering (1)

- Student Volunteer at ESEC/FSE 2017.

COURSEWORK

Computer Vision
Machine Learning
Probability and Statistics

Deep Learning
Convex Optimization
Linear Algebra

Fairness in Machine Learning
Reinforcement Learning
Reasoning for Software