Sahil **Verma**

THIRD YEAR GRADUATE STUDENT · COMPUTER SCIENCE AND ENGINEERING, UNIVERISTY OF WASHINGTON

(+1)2066366934 | v.sahil1@gmail.com | vsahil.github.io | vsahil | MMBSSXgAAAAJ | sahil-verma-48035264

Education

University of Washington *Seattle, USA*

PH.D., COMPUTER SCIENCE AND ENGINEERING *2019 –*

- GPA at the end of second year: **3.85/4.0**

University of Washington *Seattle, USA*

MASTERS OF SCIENCE, COMPUTER ENGINEERING *2019 - 2021*

- Cumulative Grade Point/ CPI at the end of 6^{th} quarter: **3.85/4.0**

IITK (Indian Institute of Technology, Kanpur) *Kanpur, India*

BACHELOR OF TECHNOLOGY, ELECTRICAL ENGINEERING *2015 - 2019*

- Cumulative Grade Point/ CPI at the end of 8^{th} semester: **8.1/10.0**

DELHI PUBLIC SCHOOL, BOKARO *Bokaro Steel City, India*

ICSE *2015, 2013*

- 12^{th} GRADE | Aggregate **96.0%**
- 10^{th} GRADE | Aggregate **95.2%**

Honors & Awards

2020	**Awarded Nvidia Titan RTX GPU,** For the Best Paper Award at ML-RSA Workshop	*USA*
2020	**Best Paper Award,** ML-RSA @ NeurIPS 2020	*USA*
2018	**Awarded $1500 by ACM SIGPLAN,** Attending PLMW, PLDI	*USA*
2017	**Awarded $1800 by Google India,** Attending FSE	*Germany*
2015	**All India Rank 663,** IIT-JEE Advanced	*India*
2015	**0.1 Percentile,** IIT-JEE Mains	*India*
2015	**KVPY Fellow | All India Rank 205,** IISc Bangalore and Government of India	*Banglore, India*
2015	**Top 1%,** National Standard Examinations in Chemistry	*India*
2015	**Top 1%,** National Standard Examinations in Biology	*India*

Patents

Amortized Generation of Sequential Counterfactual Explanations for Black-box Models *Arthur AI, Washington D.C.*

PROVISIONAL PATENT *November, 2020*

- Provisional patent for our algorithm which is deployed at Arthur AI for generating counterfactual explanations.

Publications

Counterfactual Explanations for Machine Learning: A Review PAPER *Arthur AI*

RESEARCH INTERN *Jun 2020 –*

- Paper won **best paper award** at ML-RSA workshop, NeurIPS 2020.
- We reviewed about 40 papers in counterfactual explainability and evaluated them on desirable properties of a counterfactual.
- We proposed 15 future research directions in this area.

Generating Fast Counterfactual Explanations for Black-box Models Using Reinforcement Learning PAPER *Arthur AI*

RESEARCH INTERN *Jun 2020 – Jan 2021*

- Paper accepted at the XAI workshop, AAAI 2021 and the Recourse workshop, ICML 2021.
- We proposed a novel approach to generate counterfactual explanations which satisfies all desirable properties.
- Our approach is the first to work with black-box model and generate multiple counterfactuals after training once.

Facets of Fairness in Search and Recommendations PAPER — *Univeristy of Washington*

Research Project, Prof. Chirag Shah — *Jan 2020*

- Paper accepted at the Bias workshop, ECIR 2020.
- We collected 25 definitions of fairness in ranking from literature
- We categorized the definitions in 5 major recommendations settings.

Fairness Definitions Explained PAPER — *Univeristy of British Columbia*

Research Project, Prof. Julia Rubin — *August 2017 - Jan 2018*

- Paper accepted at Fairware, ICSE 2018.
- We examined the similarities and differences across all definitions in fairness literature.

NAP: Noise-Based Sensitivity Analysis for Programs PAPER — *Massachusetts Institute of Technology*

Research Project, Prof. Michael Carbin — *May 2018 - August 2018*

- Paper accepted at WAX, 2019.
- We proposed a Noise-based sensitivity analyzer which provides an analysis of each operator and variable in a program.
- We validated NAP's sensitivities by using them to generate mixed-precision approximate programs for a neural network and scientific computing benchmarks.

Bug localization for Heap Programs PAPER — *IIT Kanpur*

Research Project, Prof. Subhajit Roy — *May 2018 - May 2019*

- Paper accepted at FMSD journal for publication in 2021.
- Developed a novel algorithm for bug localization for heap programs.
- Achieved an average speed up of about 225X in repair timings in Wolverine.

Synergistic Debug-Repair for Heap Manipulations PAPER — *IIT Kanpur*

Research Project, Prof. Subhajit Roy — *May 2016 - Feb 2017*

- Paper accepted at ESEC/FSE, 2017.
- Developed interaction of live execution and memory state graphical representation of programs with the repair engine.
- Developed features like hot-patching (runtime repair and insertion of newcode).
- Proposed the idea of synergistic debug and repair of programs in the tool named Wolverine.

Service

Reviewed papers for AAAI 2022, EAAMO 2021, AFCR workshop, NeurIPS 2021, XAIF workshop, ICAIF 2021, Data Mining and Knowledge Journal, 2021

Reviewing responsibility — *October 2021*

Student volunteer for top-tier conference, FSE 2017

Student Volunteer — *September 2017*

Current Projects

Counterfactual explanations for Recommender Systems — *Univeristy of Washington*

Research Project, Prof. Chirag Shah & Prof. John P. Dickerson — *July 2021 - Present*

- Developing the novel concept of counterfactual explanations for recommender systems.
- Developing the algorithm for generating counterfactual explanations.

Fairness in Machine Learning PREPRINT — *Univeristy of Washington*

Research Project, Prof. Michael Ernst & Prof. Rene Just — *Sep 2019 - Present*

- Developed a novel algorithm to identify biased datapoints in a dataset.
- Empirically shows that our techniques leads to zero discrimination levels for all benchmarks.
- Empirically shown to beat many popular previous techniques.