# Varun Babbar

+1 (919) 951-5082 | vb97@duke.edu | LinkedIn | Google Scholar | varunbabbar.github.io

## Summary

PhD student at Duke University, with 4+ years of AI research and engineering experience across industry and academia. Skilled in PyTorch, Scikit-learn, and TensorFlow, with expertise in deep learning, statistical modeling, and discrete optimization. First author publications at ICML (**Spotlight**), IJCAI (**Oral**), and JMLR amongst other conference and workshop papers.

## Education

**Duke University** — *Durham, NC, USA*
*PhD, Computer Science* — *August 2023 – Current*
Advisor: *Cynthia Rudin*

**University of Cambridge** — *Cambridge, United Kingdom*
BA + MEng in Information and Computer Engineering — *2018-2022*
Advisors: *Adrian Weller (MBE), Umang Bhatt* — *1st Class (BA), Distinction (MEng) - **Top 5% of class***

## Industry Experience

**AI Applied Scientist Intern at Zillow** — *New York, United States*
*May 2025 – Aug 2025*

- Working on developing novel recommendation algorithms for homes. Internship just started.

**AI Applied Scientist at JP Morgan Chase and Co** — *London, United Kingdom*
Managed by Dr Sean Moran — *Sept 2022 - July 2023*

- Developed ML algorithms for source code understanding, using large language models and novel deep learning architectures.
- Won an internal generative AI hackathon, published a paper at ICSE'24, developed a unit test quality assessment tool that saved the firm ~ $2 million, contributed to open source, and secured a patent.

**AI Applied Scientist Intern at JP Morgan Chase and Co** — *London, United Kingdom*
Managed by Dr Sean Moran — *June 2021 - Sept 2021*

- Developed a federated deep learning algorithm robust to distribution shifts in client datasets. Outcome: a patent and a paper accepted at SPIE Medical Imaging 2022.

## Research Experience

**Duke Interpretable Machine Learning Lab** — Aug 2023 – Current
**Supervisor**: Dr Cynthia Rudin

- Developed algorithms for finding $\epsilon$-optimal, sparse, and interpretable ML models 100× faster than SOTA. Spotlight paper at ICML 2025 **(top** 2.6%)+ another paper under review at NeurIPS 2025.
- Developed an interpretable toolbox to explain distribution shifted data. Recently accepted at Journal of Machine Learning Research (JMLR).

**University of Cambridge: Department of Engineering** — Sept 2021 – July 2022
**Supervisors**: Dr Adrian Weller (OBE), Dr Umang Bhatt

- Developed an uncertainty quantification method with theoretical guarantees that provides confidence intervals with human preference alignment.
- Won a department prize for my thesis. Oral paper at IJCAI 2022 (**top** 3%) and workshop papers at ICML and NeurIPS 2022.

**University of Cambridge: Department of Computer Science** June 2020 – Sept 2020
**Supervisor**: Dr Rafal Mantiuk.

- Constructed and tested novel human-perceptual loss functions for image to image translation tasks (super-resolution, de-blurring, de-noising), with *state of the art performance.*
- Co-author of paper accepted at IEEE Winter Conference on Applications of Computer Vision (WACV '22).

## Awards and Honours

**YCombinator AI Startup School - only** 2000 **people selected worldwide** *June 2025*
**Graduate School Fellowship @ Duke University** *Aug 2023*
**JP Morgan Innovation Week AI Hackathon** *(Runner Up)* *June 2023*
**MEng Project Prize** *(For a top ranked MEng Project @ Cambridge)* *July 2022*
**Ruth Hendry Year Prize** *(For exemplary academic performance)* *June 2022*
**Foundation Scholar (2x)** *(For 1st class performance in exams)* *Oct 2021, Oct 2022*
**The James and Jean Bennet Prize** *(For exemplary academic performance)* *June 2022*
**Hawkes' Trust Award** *(For exemplary academic performance)* *March 2021*
**HackCambridge, Oxford Hackathon, HackBrunel** *(Winner)* *Jan 2022, Nov 2018, Oct 2019*
**World Topper in IB School Exams** *($100^{th}$ percentile out of ~170000 students)* *May 2018*

## Conference, Industry, and Academic Presentations

**International Conference in Machine Learning (ICML 2025)**, Vancouver, Canada July 2025
**New York University**, Responsible Data Science Course April 2025
**Responsible AI Symposium, Duke University** Feb 2025
**Duke University, Guest Lecture**, MEng in AI Program Sept 2024, Feb 2025
**Duke University, Guest Lecture**, Theory and Algorithms in Machine Learning Sept 2024
**JP Morgan, Innovation Week Speaker Series**, London, UK June 2023
**NeurIPS 2022 Workshop on Human-in-the-Loop Learning**, Virtual Dec 2022
**JP Morgan, Applied Innovation in AI Seminar**, London, UK Nov 2022
**International Joint Conference on AI (IJCAI 2022)**, Vienna, Austria July 2022
**ICML Workshop on Distribution Free Uncertainty Quantification**, Virtual July 2022
**University of Cambridge**, Department of Engineering July 2022

## Service

**Research Mentorship:** Tracy Sun, MS in CS @ Duke Oct 2024-April 2025
Michael Thomas, Allen Yao, Chris Li - CS+Math Undergrad @ Duke Jan 2025-Current
**Program Committee:** Interpretable AI Workshop @ NeurIPS Dec 2024
**Teaching Assistant:** CS671: Theory and Algorithms in Machine Learning Fall 2024
**Reviewer:** ECAI 2023, WANT@ICML 2024, TheWebConf2023, Journal of the ACM

## Discrete Optimization

[DiscOpt1] **Varun Babbar***, Hayden McTavish*, Cynthia Rudin, and Margo Seltzer. Near-Optimal Decision Trees in a SPLIT Second. In *Proceedings of the 42nd International Conference on Machine Learning (ICML 2025)*. **Spotlight**—top 2.6% of submissions.

## Uncertainty Quantification

[Uncertainty1] **Varun Babbar**, Umang Bhatt, and Adrian Weller. On the Utility of Prediction Sets in Human-AI Teams. In Lud De Raedt, editor, *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence, IJCAI-22*, pages 2457–2463, 7 2022. Main Track (**Oral:** Top 3% of submissions). Also accepted as a workshop paper at ICML 2022.

[Uncertainty2] **Varun Babbar**, Umang Bhatt, Miri Zilka, and Adrian Weller. Conformal Prediction for Resource Prioritisation in Predicting Rare and Dangerous Outcomes, 2022. Human in the Loop Learning (HiLL) Workshop at NeurIPS 2022 (**Top 15%** of Accepted Papers).

## Deep Learning and Computer Vision

[DLPapers1] **Varun Babbar***, Zhicheng Guo*, and Cynthia Rudin. What is Different Between these Datasets?, 2025. *Journal of Machine Learning Research (JMLR)*.

[DLPapers2] Aamir Mustafa, Aliaksei Mikhailiuk, Dan Andrei Iliescu, **Varun Babbar**, and Rafał K. Mantiuk. Training a Task-Specific Image Reconstruction Loss. In *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 21–30, 2022.

[DLPapers3] Antonios Georgiadis*, **Varun Babbar***, Fran Silavong, Sean J. Moran, and Rob Otter. ST-FL: Style Transfer Preprocessing In Federated Learning For COVID-19 Segmentation. In *Medical Imaging 2022: Imaging Informatics for Healthcare, Research, and Applications*. SPIE, April 2022.

[DLPapers4] Agathe Lherondelle, **Varun Babbar**, Yash Satsangi, Fran Silavong, Shaltiel Eloul, and Sean Moran. Topical: Automatic Repository Tagging using Attention on Hybrid Code Embeddings. In *Proceedings of the 1st IEEE/ACM Workshop on Software Engineering Challenges in Financial Firms*, FinanSE '24, page 23–30. Association for Computing Machinery, 2024.

[DLPapers5] Xiaoying Zhi, **Varun Babbar**, Rundong Liu, Pheobe Sun, Fran Silavong, Ruibo Shi, and Sean Moran. Learning a consensus sub-network with polarization regularization and one pass training. In *Arxiv*, 2025.

## Patents

[Patents1] Antonios Georgiadis, **Varun Babbar**, Fanny Silavong, Sean Moran, and Rob Otter. Systems and Methods for Noise-Agnostic Federated Learning, 2023. Patent No. US20230058972A1.

[Patents2] Peter Maciver, **Varun Babbar**, and Sean Moran. Systems and Methods for Automated Application and Platform Generation, 2025. Patent No. US20250005530.