

# Rajeev Verma

## PhD Candidate at the University of Amsterdam

📍 Amsterdam 🌐 [rajeevv.github.io](https://rajeevv.github.io)

@rajeev.ee15@gmail.com 🐙 [github.com/rajeevv](https://github.com/rajeevv)

### Research Interests

AI safety, imprecise probabilities, human-AI collaboration, learning to defer, decision theory, safe statistics

## 🎓 EDUCATION

### PhD in Machine Learning

AMLab, University of Amsterdam (UvA)

Jan 2023 – present

Amsterdam, Netherlands

> Advisors: **Eric Nalisnick** (Johns Hopkins University), **Christian A. Naesseth** (UvA)

> Research on AI safety, decision-making, safe statistics, and learning to defer.

### Master of Science in Artificial Intelligence

University of Amsterdam (UvA)

Sep 2020 – Sep 2022

Amsterdam, Netherlands

> Thesis: *On the Calibration of Learning to Defer Systems*.

> **Nominated** for the Amsterdam AI thesis award and UvA-wide thesis award. Covered by **UvA News** and featured on [amsterdamsciencepark.nl](https://amsterdamsciencepark.nl).

### Bachelor of Technology in Electrical Engineering

Indian Institute of Technology Patna

Aug 2015 – Aug 2019

Patna, India

> Thesis: *Knowledge Graph Representation Learning Based Drug Informatics*.

> **Nominated** for the best thesis award at the institute. Undergraduate research at the AI-NLP-ML lab on NLP, and at Nanyang Technological University on random-walk based graph representation learning.

## 📖 PUBLICATIONS

<sup>1</sup>equal contribution

### Selected Conference Articles

1. Abbavaram Gowtham Reddy, **Rajeev Verma**, Celia Rubio-Madrigal, Krikamol Muandet<sup>1</sup>, Rebekka Burkholz<sup>1</sup>. Boosting for Predictive Sufficiency. *International Conference on Learning Representations*, 2026.
2. Alexander Timans<sup>1</sup>, **Rajeev Verma**<sup>1</sup>, Eric Nalisnick, Christian A. Naesseth. On Continuous Monitoring of Risk Violations under Unknown Shift. *Uncertainty in Artificial Intelligence (UAI)*, 2025.
3. **Rajeev Verma**, Volker Fischer, Eric Nalisnick. On Calibration in Multi-Distribution Learning. *ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 2025.
4. Dharmesh Tailor, Aditya Patra, **Rajeev Verma**, Putra Manggala, Eric Nalisnick. Learning to Defer to a Population: A Meta-Learning Approach. *Conference on Artificial Intelligence and Statistics (AISTATS)*, 2024. *Oral, Student paper award (top 1%)*.
5. **Rajeev Verma**<sup>1</sup>, Daniel Barrejón<sup>1</sup>, Eric Nalisnick. Learning to Defer to Multiple Experts: Consistent Surrogate Losses, Confidence Calibration, and Conformal Ensembles. *Conference on Artificial Intelligence and Statistics (AISTATS)*, 2023.
6. **Rajeev Verma**, Eric Nalisnick. Calibrated Learning to Defer with One-vs-All Classifiers. *International Conference on Machine Learning (ICML)*, 2022.

### Selected Workshop Articles

1. Alejandro Monroy Muñoz, **Rajeev Verma**<sup>1</sup>, Alexander Timans<sup>1</sup>. Detecting Object Tracking Failure via Sequential Hypothesis Testing. *WACV Workshop: Real World Surveillance: Applications and Challenges*, 2026. *Oral; as co-advisor*
2. **Rajeev Verma**<sup>1</sup>, Rabanus Derr<sup>1</sup>, Christian A. Naesseth, Volker Fischer, Eric Nalisnick. So What are Good Imprecise Forecasts? *EurIPS Workshop: Epistemic Intelligence in Machine Learning*, 2025. *work in progress*

3. Anurag Singh, Julian Rodemann, **Rajeev Verma**, Siu Lun Chau, Krikamol Muandet. Incentive Aware AI Regulation. *EurIPS Workshop: Beyond Regulation, Private Governance & Oversight Mechanisms for AI*, 2025, *Oral; work in progress*
4. Jakub Podolak, **Rajeev Verma**. Read Your Own Mind: Reasoning Helps Surface Self-Confidence Signals in LLMs. *EMNLP Workshop: Uncertainty-Aware NLP*, 2025. *as advisor*
5. **Rajeev Verma**, Volker Fischer, Eric Nalisnick. On the Calibration of Conditional-Value-at-Risk. *ICML Workshop: Next Generation of AI Safety*, 2024.

## 🔧 EXPERIENCE

### Visiting Researcher

Rational Intelligence Lab *CISPA Helmholtz Center for Information Security*

Summer 2025  
Saarbrücken, Germany

- Visiting and collaborating with **Krikamol Muandet** at the **Rational Intelligence lab** investigating problems on AI regulation and regulatory markets.

### Research Assistant

AI-NLP-ML Lab *Indian Institute of Technology Patna*

July 2018 – Sep 2021  
Patna, India

- Worked on natural language processing problems on scholarly data, resulting in publications at JCDL, ACL, and other venues.

### Software Design Engineer

Telestream

Aug 2019 – July 2020  
Bengaluru, India

- Developed video deinterlacing algorithms and image processing pipelines for video quality monitoring of commercial content.

## ⚙️ MISCELLANEOUS

### Awards and Honors

- Amsterdam AI Thesis award (as advisor; for the thesis on studying fairness in two-sided markets), Outstanding Student Paper Award (as co-author; AISTATS 2024), NeurIPS Top Reviewer (2023), ICML Participation Grant (2022), Microsoft Research Travel Award (2019)

### Talks and Presentations

- Invited talk: "On Continuous Monitoring of Risk Violations under Unknown Shift" (CISPA). *Slides*.
- Oral presentation: "On Calibration in Multi-Distribution Learning" (FAccT 2025).
- Invited talk: "On Calibration in Multi-Distribution Learning" (*2nd Workshop on Learning Under Weakly Structured Information*, Tübingen AI Center). *Slides*.
- Talk: "On the Calibration of Systems that Learn to Defer to Experts" (*EPIC Research group*, Swansea University).
- Talk: "On the Calibration of Systems that Learn to Defer to Experts" (ICAI). *Slides*.

### Technical Skills

- Methods: Bayesian inference, conformal prediction, calibration, statistical learning theory, safe statistics (e-values) and game-theoretic probability

### Reviewing

- NeurIPS 2023-2025; ICLR 2023, 2025; ICML 2023-2025; UAI 2024-2025; ACL 2021, 2025

### Teaching and Advising

- Teaching Assistant for Human-in-the-machine learning (2023), Deep Learning 2 (2024), and Machine Learning 2 (2025) at Master AI program, UvA
- Supervised three master's projects:
  1. *Equity by Design: Fairness-Driven Recommendations in Two-Sided Markets*. (*Awarded the Amsterdam AI Thesis award*),
  2. *On Reliable Confidence Scoring for LLMs: Domain Shifts and Test-Time Compute* (*published at the EMNLP 2025 workshop on UncertainNLP*),
  3. *Detecting Object Tracking Failure via Sequential Testing*