

Bishwamittra Ghosh

Research Interests

My research applies formal methods to machine learning, with a focus on advancing the foundational understanding of large language models and improving the trustworthiness of machine learning systems through enhanced fairness and interpretability.

Appointment

- Since 04.24 **Postdoctoral Researcher**, *Max Planck Institute for Software Systems*, Germany.
Advisor: Dr. Krishna P. Gummadi
- 06.23 – 03.24 **Scientist**, *Institute of High Performance Computing (IHPC)*, A*STAR, Singapore.
- 01.18 – 06.23 **Graduate Research Assistant**, *National University of Singapore (NUS)*, Singapore.
CS2030: Programming Methodology II, 2018 – 2021
CS3243: Introduction to Artificial Intelligence, 2020
- 10.17 – 01.18 **Lecturer**, *Department of CSE, United International University (UIU)*, Bangladesh.

Education

- 01.18 – 09.23 **PhD in Computer Science**, *National University of Singapore (NUS)*, Singapore.
Thesis: Interpretability and Fairness in Machine Learning: A Formal Methods Approach
Advisor: Dr. Kuldeep S. Meel
- 02.13 – 09.17 **BSc in Computer Science and Engineering**, *Bangladesh University of Engineering and Technology (BUET)*, Bangladesh, CGPA: 3.83 out of 4 (with Honours).
Thesis: The Flexible Socio Spatial Group Queries
Advisor: Dr. Mohammed Eunus Ali

Publications (sorted by publication year)

Conference Papers

- ICLR-25 Logical Consistency of Large Language Models in Fact-checking
[Bishwamittra Ghosh](#), Sarah Hasan, Naheed Anjum Arafat, Arijit Khan
- AAAI-25 Active Fourier Auditor for Estimating Distributional Properties of ML Models
Ayoub Ajarra, [Bishwamittra Ghosh](#), Debabrota Basu
- WSDM-25 Towards Reliable Latent Knowledge Estimation in LLMs: Zero-Prompt Many-Shot Based Factual Knowledge Extraction
Qinyuan Wu, Mohammad Aflah Khan, Soumi Das, Vedant Nanda, [Bishwamittra Ghosh](#), Camila Kolling, Till Speicher, Laurent Bindschaedler, Krishna P. Gummadi, Evimaria Terzi
- CAI-24 Split Learning of Multi-Modal Medical Image Classification
[Bishwamittra Ghosh](#), Yuan Wang, Huazhu Fu, Wei Qingsong, Yong Liu, Rick Goh
- FAccT-23 “How Biased are Your Features?”: Computing Fairness Influence Functions with Global Sensitivity Analysis
[Bishwamittra Ghosh](#), Debabrota Basu, Kuldeep S. Meel
- VLDB-23 Neighborhood-based Hypergraph Core Decomposition
Naheed Anjum Arafat, Arijit Khan, Arpit Kumar Rai, [Bishwamittra Ghosh](#)
- AAAI-22 Algorithmic Fairness Verification with Graphical Models
[Bishwamittra Ghosh](#), Debabrota Basu, Kuldeep S. Meel
- AAAI-21 Justicia: A Stochastic SAT Approach to Formally Verify Fairness
[Bishwamittra Ghosh](#), Debabrota Basu, Kuldeep S. Meel
- AAAI-20 A MaxSAT-based Framework for Group Testing
Lorenzo Ciampiconi, [Bishwamittra Ghosh](#), Jonathan Scarlett, Kuldeep S. Meel

- ECAI-20 Classification Rules in Relaxed Logical Form
Bishwamittra Ghosh, Dmitry Malioutov, Kuldeep S. Meel
- AIES-19 IMLI: An Incremental Framework for MaxSAT-Based Learning of Interpretable Classification Rules
Bishwamittra Ghosh, Kuldeep S. Meel
- VLDB-18 The Flexible Socio Spatial Group Queries
Bishwamittra Ghosh, Mohammed Eunus Ali, Farhana M. Choudhury, Sajid Hasan Apon, Timos Sellis, Jianxin Li

Journal Papers

- JAIR-22 Efficient Learning of Interpretable Classification Rules
Bishwamittra Ghosh, Dmitry Malioutov, Kuldeep S. Meel
- TSAS-21 Social-Spatial Group Queries with Keywords
 Sajid Hasan Apon, Mohammed Eunus Ali, Bishwamittra Ghosh, Timos Sellis

Professional Experiences

Tutorials

- IJCAI-23 Auditing Bias of Machine Learning Algorithms: Tools and Overview

Talks

- Invited Talks Trustworthy Machine Learning: A Formal Methods Approach
 ○ Microsoft Research (24), University of Michigan, Dearborn (23), MPI-SWS, Germany (22), INRIA, France (22), IHPC, Singapore (22)
- Conference Talks ICLR (25), FAccT (23), AAAI (22, 21, 20), ECAI (20), CP (19), VLDB (19), IJCAI (19), AIES (19)

Program Committee

AAAI (25)

Reviewer

NeurIPS, ICML, AAAI, TMLR, KDD (**outstanding reviewer**), JAIR, AIES, CAV, TNNLS, TMI, TKDE

Research Visit

- 03.23 – 05.23 Max Planck Institute For Software Systems, Saarbrücken, Germany
 11.21 – 02.22 INRIA, Lille-Nord, France
 06.21 – 08.21 Goldman Sachs, Singapore
 03.20 – 06.20 Max Planck Institute For Software Systems, Kaiserslautern, Germany

Awards and Honors

- 2026 Invited to Dagstuhl seminar on User-Aligned Assessment of AI Systems
 2025 Invited to Dagstuhl seminar on Interactions in Constraint Optimization
 2021 Moblilex Scholarship, Université de Lille, France
 2018 – 2023 NUS Research Scholarship, Singapore
 2013 – 2017 University Deans List, BUET, Bangladesh
 2011, 2012 National Math Olympiad winner, Bangladesh
 2012 National Chemistry Olympiad winner, Bangladesh

References

- **Krishna P. Gummadi**
Scientific Director
MPI-SWS, Germany
- **Kuldeep S. Meel**
Associate Professor
Department of Computer Science
University of Toronto, Canada
- **Debabrota Basu**
Faculty (equivalent of Associate Professor)
Scool Team
INRIA at Université de Lille and CNRS, France