

Semperstr. 111

Bochum 44801

Germany

* 15 April 1989

+49 (169) 6783237

✉ prithagupta.nsit@icloud.com

🌐 leakdetectai.github.io

in pritha-g-32383714

🔗 prithagupta

Google Scholar

Dr. Pritha Gupta

Education

- 2019–2025 **Dr. rer. nat. Topic: Advanced Machine Learning Methods for Information Leakage Detection in Cryptographic Systems**, *Software Innovation Campus Paderborn (SICP)*, Paderborn University, Paderborn, North Rhine-Westphalia, Germany, Grade – magna cum laude
- 2015–2018 **Master of Science in Computer Science**, Paderborn University, Paderborn, North Rhine-Westphalia, Germany, Grade – 1.2
- 2008–2012 **Bachelor of Engineering in Computer Engineering**, Netaji Subhas University of Technology (NSUT), formerly Netaji Subhas Institute of Technology (NSIT), Delhi University, New Delhi, India, Grade – 67.27% (GPA: 7.477; German Scale: 2.2)

Selected Publications

Full publication list available on Google Scholar.

- Archive '2026 Jiawen Wang, **Pritha Gupta**, Ivan Habernal, Eyke Hüllermeier, Xiaoxue Gao, Nancy F. Chen. From ASR to ASP: Evaluating Prompt Attack Vulnerabilities Against Open-Source LLMs. In *arXiv 2026*, Available: <https://arxiv.org/abs/2505.14368>
- Information Sciences **Pritha Gupta**, Marcel Wever, Eyke Hüllermeier. Information Leakage Detection through Approximate Bayes-optimal Prediction. In *Information Sciences, 2025*, '2025 Available: <https://doi.org/10.1016/j.ins.2025.122419>
- IDA '2023 David Schubert, **Pritha Gupta**, Marcel Wever. Meta-Learning for Automated Selection of Anomaly Detectors for Semi-Supervised Datasets. In *International Symposium on Intelligent Data Analysis, 2023*. Available: https://link.springer.com/chapter/10.1007/978-3-031-30047-9_31
- ARES '2023 **Pritha Gupta**, Jan Peter Drees, Eyke Hüllermeier. Automated Side-Channel Attacks using Black-Box Neural Architecture Search. In *18th International Conference on Availability, Reliability and Security, 2023*, Available: <https://dl.acm.org/doi/abs/10.1145/3600160.3600161>

- IJAR '2022 Karlson Pfannschmidt, **Pritha Gupta**, Björn Haddenhorst, Eyke Hüllermeier. Learning Context-Dependent Choice Functions. In *International Journal of Approximate Reasoning*, 2022, Available: <https://www.sciencedirect.com/science/article/pii/S0888613X21001614>
- AISeC '2021 Jan Peter Drees, **Pritha Gupta**, Eyke Hüllermeier, Tibor Jäger, Alexander Konze, Claudia Priesterjahn, Arunselvan Ramaswamy, Juraj Somorovsky. Automated Detection of Side Channels in Cryptographic Protocols: DROWN the ROBOTS! In *Proceedings of the 14th ACM Workshop on Artificial Intelligence and Security*, 2022, Available: <https://dl.acm.org/doi/abs/10.1145/3474369.3486868>
- ICAART '2021 **Pritha Gupta**, Arunselvan Ramaswamy, Jan Peter Drees, Eyke Hüllermeier, Claudia Priesterjahn, Tibor Jäger. Automated Information Leakage Detection: A New Method Combining Machine Learning and Hypothesis Testing with an Application to Side-channel Detection in Cryptographic Protocols. In *Proceedings of the 14th International Conference on Agents and Artificial Intelligence*, 2022, Available: <https://www.scitepress.org/Papers/2022/107930/107930.pdf>
- TFML '2017 Vitalik Melnikov, Eyke Hüllermeier, Daniel Kaimann, Bernd Frick, **Pritha Gupta**. Pairwise versus Point-wise Ranking: A Case Study. In *Theoretical Foundations of Machine Learning (TFML)*, 2017, Available: <http://www.ejournals.eu/Schedae-Informaticae/2016/Volume-25/art/9011>
- IJCA '2015 Gurjashan Singh Pannu, Mohammad Dawud Ansari, **Pritha Gupta**. Design and Implementation of Autonomous Car using Raspberry Pi. In *International Journal of Computer Applications*, 113(9):22–29, 2015. Available: <http://www.ijcaonline.org/archives/volume113/number9/19854-1789>

Invited Talks

- 02 Feb, 2026 **LLM Security and Prompt Injection Risks**, *AI Safety & Ethics Workshop, organized in collaboration with The Debugging Society (TDS), ASN, and IEEE NSUT, Netaji Subhas University of Technology (NSUT), Delhi, India*, **Invited Speaker**
advisor Prof. Dr. Ivan Habernal and Prof. Dr.-Ing. Shampa Chakraverty
- 19 Jul, 2026 **AI-Powered Security Testing for Crypto Wallets and Financial Systems**, *Internet Engineering Task Force (IETF) 126, ACM/IRTF Applied Networking Research Workshop, Vienna, Austria*, **Invited Speaker**
advisor Prof. Dr.-Ing. Shampa Chakraverty and Manusheel Gupta
- 02 Aug, 2026 **Architecting Artificial Intelligence — Data Annotation and Ground Truth for Machine Learning**, *Expert Consultancy Sessions under the IndiaAI Project — in collaboration with the Ministry of Electronics and Information Technology (MeitY), Government of India; NIELIT, India*, **Keynote Speaker**
advisor Prof. Dr.-Ing. Shampa Chakraverty and Manusheel Gupta
- 02 Aug, 2026 **Python-Based Data Annotation Pipelines — Automating Human-in-the-Loop Labeling**, *Expert Consultancy Sessions under the IndiaAI Project — in collaboration with the Ministry of Electronics and Information Technology (MeitY), Government of India; NIELIT, India*, **Keynote Speaker**
advisor Prof. Dr.-Ing. Shampa Chakraverty and Manusheel Gupta

- 21 Aug, 2026 **Open-Source AI, Security and Digital Public Goods**, *Devcon 8 India — Ethereum Community Hub on Open AI Public Goods and Digital Public Goods*; organized with *Ethereum*, India, **Keynote Speaker**
advisor Prof. Dr.-Ing. Shampa Chakraverty and Manusheel Gupta

Teaching

- Summer **Attacks on Large Language Models**, *Privacy-Preserving Natural Language Processing*, *Trustworthy Human Language Technologies (TrustHLT)*, *Ruhr University Bochum (RUB)*, Bochum, North Rhine-Westphalia, Germany, as **Guest Lecturer**
Semester, 2025
advisor Prof. Dr. Ivan Habernal
- Summer **Seminar Recent Advances in Machine Learning**, *Software Innovation Campus Paderborn (SICP)*, *Paderborn University*, Paderborn, North Rhine-Westphalia, Germany, as **Mentor and Co-organizer**
Semester, 2022
advisor Prof. Dr. Eyke Hüllermeier
- Summer **Foundations of Machine Learning**, *Software Innovation Campus Paderborn (SICP)*, *Paderborn University*, Paderborn, North Rhine-Westphalia, Germany, as **Tutor and Co-organizer**
Semester, 2021
advisor Prof. Dr. Arunselvan Ramaswamy

Thesis Supervisor

- 2026 **Enhanced Membership Inference Attacks via Statistical Score Aggregation in Large Language Models**, *Trustworthy Human Language Technologies (TrustHLT)*, *Research Center for Trustworthy Data Science and Security (RC Trust)*, *Ruhr University Bochum (RUB)*, Bochum, North Rhine-Westphalia, Germany, Master's Thesis Supervision
- 2025 **Implementation of Combined Hyperband Search Strategies in Keras Tuner**, *Chair of Artificial Intelligence and Machine Learning (AIML)*, *Ludwig Maximilian University of Munich (LMU Munich)*, Munich, Bavaria, Germany, Master's Thesis Supervision
- 2025 **Is Your Prompt Safe? Learning Prompt Injection Attacks against Open-Source Large Language Models**, *Chair of Artificial Intelligence and Machine Learning (AIML)*, *Ludwig Maximilian University of Munich (LMU Munich)*, Munich, Bavaria, Germany, Master's Thesis Supervision
- 2025 **Assessing the Security of TabPFN through Membership Inference Attacks**, *Chair of Artificial Intelligence and Machine Learning (AIML)*, *Ludwig Maximilian University of Munich (LMU Munich)*, Munich, Bavaria, Germany, Master's Thesis Supervision
- 2023 **Multi-Label Deep Learning-based Side-Channel Analysis using Different Encodings**, *Chair of Artificial Intelligence and Machine Learning (AIML)*, *Ludwig Maximilian University of Munich (LMU Munich)*, Munich, Bavaria, Germany, Bachelor's Thesis Supervision
- 2022 **Automated Side-Channel Attacks using Black-Box Neural Architecture Search**, *Software Innovation Campus Paderborn (SICP)*, *Paderborn University*, Paderborn, North Rhine-Westphalia, Germany, Master's Thesis Supervision

2020 **Application of Anomaly Detection in Autonomous Systems**, *Software Innovation Campus Paderborn (SICP)*, Paderborn University, Paderborn, North Rhine-Westphalia, Germany, Master's Thesis Supervision

2020 **ASOD: A Python Toolbox for the Automatic Selection of Outlier Detection Techniques**, *Software Innovation Campus Paderborn (SICP)*, Paderborn University, Paderborn, North Rhine-Westphalia, Germany, Master's Thesis Supervision

Academic Experience

Feb 2025 – **Postdoctoral Researcher Topic: Adversarial Attacks on Large Language Models**, *Trustworthy Human Language Technologies (TrustHLT)*, Ruhr University Bochum (RUB), Bochum, North Rhine-Westphalia, Germany

Oct 2026

Technologies: Python, Hugging Face, LaTeX, Jupyter, Ollama, vLLM

Achievements:

- Investigating LLM privacy and security, focusing on prompt injection, membership inference, model inversion/reconstruction attacks, including privacy-preserving fine-tuning methods.
- Developed and benchmarked prompt injection pipelines for open-source LLMs.
- Delivered invited lectures on adversarial attacks on LLMs and mitigation strategies.
- Member of the Research Center for Trustworthy Data Science and Security (RC Trust).

Feb 2019 – **Research Associate (Wissenschaftliche Mitarbeiterin)**, *Software Innovation Campus Paderborn (SICP)*, Paderborn University, Paderborn, North Rhine-Westphalia, Germany

Jan 2025

Technologies: Python, PyTorch, TensorFlow, Keras, Scikit-learn, AutoML, LaTeX, Git, Docker

Achievements:

- Developed ML and information-theoretic methods to quantify information leakage in cryptographic implementations.
- Applied AutoML (HPO, NAS) to automate attack generation and vulnerability discovery in AES-based systems.
- Performed ML-assisted analysis of TLS implementations and cryptographic protocols.
- Coordinated with academic and industry partners and produced technical reports and documentation.

Jan 2016 – **Student Research Assistant**, *Intelligent Systems and Machine Learning Group (ISML)*, Paderborn University, Paderborn, North Rhine-Westphalia, Germany

Nov 2018

Technologies: Python, MATLAB, Java, Android

Achievements:

- Researched choice theory and preference learning for recommender systems.
- Developed context-sensitive neural ranking models for preference learning.
- Collected tablet purchase preferences through a self-developed Android application.

Nov 2014 – **Research Associate**, *Department of Computer Science and Engineering, Thapar University*, Patiala, India

Feb 2015

Technologies: C/C++, Linux, Java, Android, Python, OpenCV

Achievements:

- Designed a Raspberry Pi-based autonomous car prototype with road detection.
- Developed a WiFi-controlled Android application for remote navigation.
- **Project Demonstration:** <https://youtu.be/NwgXK1Fjv-Q>

Industrial Experience

Jun 2012 – **Software Engineer**, *Samsung India Electronics Pvt. Ltd. (SRI Noida and SDD Division)*, Noida, India

Technologies: Java, Android, Android Framework, C/C++, XML, Perforce

Achievements:

- Designed and implemented Android OS framework features for Samsung smartphones.
- Developed and customized Android applications and TouchWiz Home components, including My Magazine for Galaxy Note 3 and Easy Launcher for Galaxy S4.
- Conducted feasibility studies and proposed product enhancements, with focus on performance, multi-resolution support, and localization.
- Coordinated with Samsung HQ in South Korea and European test teams; mentored interns.
- SRI Noida is part of Samsung Research.

Aug 2015 – **Student Employee**, *Wincor Nixdorf (now Diebold Nixdorf)*, Paderborn, North Rhine-Westphalia, Germany

Technologies: Java, Vaadin, Maven

Achievements:

- Built a Vaadin-based web application for statistical evaluation of ATM fleet performance.
- Implemented tools in Java for analysis and diagnostics of ATM issues.

Skills and Languages

Languages Python, Java, C/C++, Android, Git, Docker, LaTeX, Maven

ML/AI PyTorch, TensorFlow, scikit-learn, AutoML, NLP, Hugging Face, Weights & Biases (W&B), MLflow

Scholarship Owl Deutschlandstipendium (Sept'2016 - Sept'2018) by Studienfonds Owl

Languages English (C1), German (A2), Hindi (Native), Marwari (Native), Punjabi (Intermediate)

Interests Running, Bouldering, KraftTraining, Bollywood Dance

Academic References

- **Prof. Dr. Eyke Hüllermeier**, Chair of Artificial Intelligence and Machine Learning (AIML), Ludwig Maximilian University of Munich (LMU Munich), eyke@ifi.lmu.de
- **Prof. Dr. Ivan Habernal**, Trustworthy Human Language Technologies (TrustHLT), Ruhr University Bochum (RUB), ivan.habernal@ruhr-uni-bochum.de
Member of the Research Center for Trustworthy Data Science and Security (RC Trust)
- **Prof. Dr.-Ing. Juraj Somorovsky**, Institute of Computer Science, Paderborn University, Deputy Head of Institute and Head of Systems Security, juraj.somorovsky@uni-paderborn.de
- **Assoc. Prof. Dr. Arunselvan Ramaswamy**, Department of Industrial Engineering and Operations Research, Indian Institute of Technology Bombay (IIT Bombay), arunselvan.ramaswamy@kau.se, arunselvan.r@gmail.com
- **Prof. Dr. Daniel Beverungen**, Software Innovation Campus Paderborn (SICP), Paderborn University, Speaker of the SICP Steering Committee and Chairman of the SI-Lab Board, daniel.beverungen@uni-paderborn.de
- **Prof. Dr. Shampa Chakraverty**, Dean Student Welfare, Faculty of Design and Computer Engineering, Netaji Subhas University of Technology (NSUT), Delhi, shampa@nsut.ac.in

- **Prof. Dr. Bijendra Kumar**, Department of Computer Engineering, Netaji Subhas University of Technology (NSUT), Delhi, bizender@gmail.com

Industrial References

- **Manusheel Gupta (Manu Gupta)**, Entrepreneur, manusheel.gupta@gmail.com
- **Prof. Dr. Stefan Sauer**, Software Innovation Campus Paderborn (SICP), Paderborn University, sauer@uni-paderborn.de