

Sebastian Bordt

Tübingen, Germany | sbordt@posteo.de | +49 177 2674104

Education

University of Tübingen and Max Planck Institute for Intelligent Systems <i>PhD in Computer Science. Advisor: Ulrike von Luxburg</i>	2019 - 2023
Ludwig Maximilian University of Munich <i>Munich Graduate School of Economics (no degree)</i> <i>M. Sc. Economics</i>	2017 - 2019 2017
Technical University of Munich <i>M. Sc. Mathematics</i>	2016
<i>B. Sc. Mathematics</i>	2013

Selected Publications

* Denotes equal contribution

Suraj Srinivas*, **Sebastian Bordt***, Hima Lakkaraju “Which Models have Perceptually-Aligned Gradients? An Explanation via Off-Manifold Robustness”, *NeurIPS*, 2023

Spotlight Presentation

Sebastian Bordt, Harsha Nori, Rich Caruana “Elephants Never Forget: Testing Language Models for Memorization of Tabular Data”, *Table Representation Learning Workshop at NeurIPS*, 2023

Benjamin J Lengerich, **Sebastian Bordt**, Harsha Nori, Mark E Nunnally, Yin Aphinyanaphongs, Manolis Kellis, Rich Caruana “LLMs Understand Glass-Box Models, Discover Surprises, and Suggest Repairs”, *arXiv preprint*, 2023

Sebastian Bordt, Ulrike von Luxburg “From Shapley Values to Generalized Additive Models and back”, *AISTATS*, 2023

Sebastian Bordt, Uddeshya Upadhyay, Zeynep Akata, Ulrike von Luxburg “The Manifold Hypothesis for Gradient-Based Explanations”, *CVPR Workshops*, 2023

Spotlight Presentation

Sebastian Bordt, Michèle Finck, Eric Raidl, Ulrike von Luxburg “Post-Hoc Explanations Fail to Achieve their Purpose in Adversarial Contexts”, *ACM FAccT*, 2022

Sebastian Bordt, Ulrike von Luxburg “A Bandit Model for Human-Machine Decision Making with Private Information and Opacity”, *AISTATS*, 2022

Talks

“Explanations and Regulation”

AITE Conference at the University of Tübingen	2023
Harvard SEAS	2023
IDEAL Institute in Chicago	2023
EML Workshop at the University of Tübingen	2023

“From Shapley Values to Generalized Additive Models and back”

Nice Workshop on Interpretability	2022
-----------------------------------	------

Internships / Workshops

Internship at Microsoft Research, Redmond	2023
Summer Cluster on Interpretable Machine Learning, UC Berkeley	2022
Computer Science, Puzzles and Games, Workshop on Theoretical Computer Science	2016

Teaching

Algorithms and Data Structures	2022
Statistical Machine Learning	2020
Mathematics of Machine Learning	2019

Software

<https://github.com/interpretml/TalkToEBM>
<https://github.com/tml-tuebingen/nshap>

Notable

Scholar at “Studienstiftung des deutschen Volkes“ (German Academic Scholarship Foundation)
Participant in the final round of the German National Mathematics Competition
DAAD Scholarship “A New Passage to India”