LESIA SEMENOVA

lsemenova@microsoft.com | https://lesiasemenova.github.io/

My ***theoretical*** work creates a foundation for the existence of accurate simpler, including interpretable, machine learning models. I introduced a new simplicity measure of a learning problem and proposed first methods to measure the Rashomon set (the set of equally well-performing models), which enabled the shift in the machine learning community towards model multiplicity and underspecification. My ***applied*** work with immunologists has led to understanding of how cannabis affects the immune system of people with HIV. Student teams that I've coached have won the ASA Data Challenge Expo (twice) and placed third in a competition on a scholarly document processing.

RESEARCH INTERESTS

Responsible and Trustworthy AI, Interpretability, Machine Learning, Human-Centered Design, AI in Healthcare, Reinforcement Learning, Reasoning,

EMPLOYMENT HISTORY

Postdoctoral Researcher, Microsoft Research, NYC, USA — July 2024 –Present
Machine Learning Team

Research Intern, Pinterest Labs, Palo Alto, CA, USA — Summer 2020, Summer 2021
Applied Science Team

Software Engineer, Samsung R&D Institute, Kyiv, Ukraine — 2012 –2014
Interaction Lab, Augmented Reality Team

EDUCATION

Duke University, Durham, NC, USA
Ph.D. in Computer Science — 2016 – 2024
Advisors: Cynthia Rudin, Ronald Parr

Taras Shevchenko National University of Kyiv, Kyiv, Ukraine
MS in Applied Mathematics — 2012 – 2014
BS in Applied Mathematics — 2008 – 2012

PUBLICATIONS

Google Scholar
(* denotes equal contributions)

14 Zachery Boner*, Harry Chen*, **Lesia Semenova***, Ronald Parr, Cynthia Rudin. Using Noise to Infer Aspects of Simplicity Without Learning. *Neural Information Processing Systems (NeurIPS), 2024.*

13 Cynthia Rudin, Chudi Zhong, **Lesia Semenova**, Margo Seltzer, Ronald Parr, Jiachang Liu, Srikar Katta, Jon Donnelly, Harry Chen, Zachery Boner. Amazing Things Come From Having Many Good Models. *International Conference on Machine Learning (ICML), 2024 (Spotlight).*

12 Siong Thye Goh*, **Lesia Semenova***, Cynthia Rudin. Sparse density trees and lists: an interpretable alternative to high-dimensional histograms. *INFORMS Journal on Data Science (IJDS), 2024.*

11 **Lesia Semenova**, Yingfan Wang, Shane Falcinelli, Nancie Archin, Alicia D Cooper-Volkheimer, David M Margolis, Nilu Goonetilleke, David M Murdoch, Cynthia D Rudin, Edward P Browne. Machine learning approaches identify immunologic signatures of total and intact HIV DNA during long-term antiretroviral therapy**.** *eLIFE, 2024*

10 **Lesia Semenova**, Harry Chen, Ronald Parr, Cynthia Rudin. A path to simpler models starts with noise. *Neural Information Processing Systems (NeurIPS), 2023.*

9 Shane D Falcinelli, Alicia Volkheimer, **Lesia Semenova**, Ethan Wu, Alexander Richardson, Manickam Ashokkumar, David M Margolis, Nancie M Archin, Cynthia D Rudin, David Murdoch, Edward P Browne. Impact of cannabis use on immune cell populations and the viral reservoir in people with HIV on suppressive antiretroviral therapy. *The Journal of Infectious Disease (JID), 2023.*

8 **Lesia Semenova**, Cynthia Rudin, Ronald Parr. On the existence of simpler machine learning models. *ACM Conference on Fairness, Accountability, and Transparency (FAccT), 2022.*

7 Cynthia Rudin, Chaofan Chen, Zhi Chen, Haiyang Huang, **Lesia Semenova**, Chudi Zhong. Interpretable machine learning: fundamental principles and 10 grand challenges. *Statistics Surveys, 2022.*

WORKSHOPS

6 Ronald Parr, Cynthia Rudin, Harry Chen, Zachery Boner, Michal Moshkovitz, **Lesia Semenova.** Transition Noise Facilitates Interpretability. *Workshop on Interpretable Policies in Reinforcement Learning (InterpPol) @ RLC - 2024 (Oral).*

5 Dennis Tang, Frank Willard, Ronan Tegerdine, Luke Triplett, Jon Donnelly, Luke Moffett, **Lesia Semenova**, Alina Jade Barnett, Jin Jing, Cynthia Rudin, Brandon Westover. ProtoEEGNet: An interpretable approach for detecting interictal epileptiform discharges. *Medical Imaging meets NeurIPS workshop, 2023 (Oral).*

4 Gaurav Rajesh Parikh, Jenny Huang, Albert Sun, **Lesia Semenova**, Cynthia Rudin. Moving towards a more equal world, one ride at a time: studying public transportation initiatives using interpretable causal inference**.** *NeurIPS Workshop on Causality for Real-world Impact, 2022.*

— Won 2022 American Statistical Association Data Challenge Expo Student Competition

3 Alex Oesterling, Angikar Ghosal, Haoyang Yu, Rui Xin, Yasa Baig, **Lesia Semenova**, Cynthia Rudin. Multitask learning for citation purpose classification. *Second Workshop on Scholarly Document Processing, ACL, 2021.*

— Won third place in 3C Shared Task Competition

PREPRINTS

2 Chloe Qinyu Zhu, Muhang Tian, **Lesia Semenova**, Jiachang Liu, Jack Xu, Joseph Scarpa, Cynthia Rudin. Fast and interpretable mortality risk scores for critical care patients. *Submitted to Journal of the American Medical Informatics Association (JAMIA).*

1 Allan Guo, Eric Song, Gaurav Rajesh Parikh, Harry Chen, **Lesia Semenova**, Cynthia Rudin. Weed and violence: The impact of marijuana legalization on crime in California. *To be submitted to Harvard Data Science Review (HDSR).*

— Won 2023 American Statistical Association Data Challenge Expo Student Competition

INVITED TALKS

Joint Statistics Meetings (JSM)	2025
18th International Joint Conference on Computational and Financial Econometrics (CFE) and Computational and Methodological Statistics (CMStatistics)	2024
INFORMS Annual Meeting	2024
Theory of Interpretable AI Seminar	2024
The 25th International Symposium on Mathematical Programming	2024
Johns Hopkins University, Applied Physics Lab, AI/ML seminar	2024
Conference on Information Sciences and Systems (CISS)	2023
JSM, Near-Optimization Topic-Contributed session	2021
INFORMS Annual Meeting	2020

SELECTED AWARDS AND HONORS

Rising Stars in Computational and Data Science, University of Texas at Austin	2024
SAMSI Fellowship	2018
Duke CS Department Fellowship	2016 – 2018
Scholarship of Mayor of Kyiv	2012 – 2014
Scholarship of Academic Council of Taras Shevchenko National University	2011 – 2012
Scholarship of President of Ukraine	2008 – 2009

ADVISING AND MENTORING

Zach Boner, PhD Student, Duke University
Harry Chen, undergraduate, Duke University
Dennis Tang, undergraduate, Duke University (now research intern at UNC)
Chloe Zhu, undergraduate, Duke University (now PhD student at Duke)
Muhang (Tony) Tian, undergraduate, Duke University (now PhD student at NYU)
Flora Shi, undergraduate, Duke University (now PhD student at MIT)

Mentor in the RAI for Ukraine academic research program:
Bogdan Turbal, undergraduate, Taras Shevchenko National University of Kyiv
Iryna Voitsitska, undergraduate, Ukrainian Catholic University

Mykola Vysotskyi, undergraduate, Ukrainian Catholic University

Co-instructor for Duke Data Science teams that participated in various Data Science competitions (9 teams total, 29 Duke undergraduate students)

TEACHING

Certificate in College Teaching, Duke Graduate School
Formal pedagogical training in the college teaching.

Institute of Advanced Study, Teaching Assistant
2022 Program for women and mathematics: "The Mathematics of Machine Learning"
Terng Lecture Course on Interpretable Machine Learning — May 2022

Duke University, Teaching Assistant
CS474, Data Science Competition — SP23, SP22, SP21
CS571, Probabilistic Machine Learning — SP18
— TA Award Honorable Mention
CS101, Introduction to Computer Science — SP17

SERVICE

Organizer: Interpretable AI: Past, Present and Future Workshop @ NeurIPS 2024
Reviewer:
NeurIPS: 2024
ICLR: 2025
AISTATS: 2025
Journals: JMLR, Patterns, ACM Journal on Responsible Computing
Workshops: DeepMath (2022, 2023), International Workshop on Advances in Interpretable Machine Learning and Artificial Intelligence (2021, 2022, 2023)

Volunteer procurement manager at "Razom for Ukraine" nonprofit — 2022-2023
Co-lead the "Artificial Intelligence for Art and Fun" capstone event as a part of Duke's FEMMES+ (Females and Allies Excelling More in Math, Engineering, and Science) outreach program to introduce young female students (4th-6th grade) to math, science, and engineering — 2021
Student Assistant at NSF-sponsored Seamless/Seamful Human Technology Interaction (HTI) Workshop — May 2021
Co-organized the discussion series "Controversial Topics in Precision Medicine and Learning" as a collaboration between SAMSI and Duke Computer Science — 2019
Graduate student committee member for the Faculty Search and Prospective Student Visit for the Department of Computer Science, Duke University — 2016–2019
Conference volunteer — KDD 2017, ICML 2019
Alumni of ComSciCon Triangle (science communication workshop for graduate students) — 2018
Volunteer tutor at Study Zone at King County Library System (provided homework help to K-12 students) — 2015
Head of Scientific Association of Students and Postgraduates of Taras Shevchenko National University of Kyiv — 2012 – 2013