

Department of Statistics Archive of Data Science Seminars - Michaelmas Term 2022

Monday 17 October 2022, 2-3pm - Caroline Uhler (MIT)

[Website](#)

This event will take place in the Leverhulme Library (COL 6.15).

Title - From Interventions to Causality using Over-Parameterized Neural Networks.

Abstract - Massive data collection holds the promise of a better understanding of complex phenomena and ultimately, of better decisions. An exciting opportunity in this regard stems from the growing availability of perturbation / intervention data (for example from drug/knockout screens in biology, advertisement, online education, etc.). In order to obtain mechanistic insights from such data, a major challenge is the development of a framework that integrates observational and interventional data and allows causal transportability, i.e., predicting the effect of unseen interventions or transporting the effect of interventions observed in one context to another. I will discuss how over-parameterized neural networks can be used for these problems. In particular, I will characterize the implicit bias of over-parameterized autoencoders and link this to causal transportability in the context of virtual drug screening.

Biography - Caroline Uhler is a Full Professor in the Department of Electrical Engineering and Computer Science and the Institute for Data, Systems, and Society at MIT. In addition, she is a core institute member at the Broad Institute, where she co-directs the Eric and Wendy Schmidt Center. She holds an MSc in mathematics, a BSc in biology, and an MEd all from the University of Zurich. She obtained her PhD in statistics from UC Berkeley in 2011 and then spent three years as an assistant professor at IST Austria before joining MIT in 2015. She is a Simons Investigator, a Sloan Research Fellow, and an elected member of the International Statistical Institute. In addition, she received an NSF Career Award, a Sofja Kovalevskaja Award from the Humboldt Foundation, and a START Award from the Austrian Science Foundation. Her research lies at the intersection of machine learning, statistics, and genomics, with a particular focus on causal inference, representation learning, and gene regulation.

Take a look at [Caroline's slides](#).

.....
Monday 14 November 2022, 2-3pm - Vladimir Vovk (Royal Holloway, University of London)

[Website](#)

This event will take place in the Leverhulme Library (COL 6.15).

Title - Applications of e-values to multiple hypothesis testing

Abstract - In this talk I will review two alternative tools for statistical hypothesis testing, p-values and e-values. Both have been used in the algorithmic theory of randomness for decades (on the log scale and under other names), but only p-values are widely used in non-Bayesian statistics; e-values are related to Bayes factors, especially in the case of a simple null hypothesis. The advantage of e-values is that they are easy to combine. This makes them a convenient and powerful tool for multiple hypothesis testing.

Biography - Vovk started working as a researcher in the Russian Academy of Sciences, then became a Fellow in the Center for Advanced Study in the Behavioral Sciences at Stanford University. He was appointed as a professor of Computer Science at Royal Holloway and Bedford New College, where he currently serves as co-director of the Centre for Machine Learning.

Early in his career, Vovk was heavily involved in the development of the foundations of probability, along with Glenn Shafer. Their work has resulted in a book, *Probability and Finance: It's Only a Game!*, published in 2001, which was subsequently translated into Japanese in 2006 by Masayuki Kumon and edited by Kei Takeuchi. In 2005, he co-invented the Conformal prediction framework with Alexander Gammerman.

Vovk has delivered speeches all around the world. In 2021, he was invited to deliver a series of memorial lectures to Prasanta Chandra Mahalanobis in India. On the 20-year anniversary of The Society for Imprecise Probability (SIPTA) in 2019, he was invited to deliver a talk on "Game-theoretic foundations for imprecise probabilities" in Belgium. In 2016, he delivered a seminar about "Probability-free theory of continuous

martingales" at Imperial College in the UK. In 2014, he delivered a seminar at University of Hawai'i in the USA.

Vovk has written 9 books, more than 280 research papers, and has an estimated h-index of 53. He holds fellowship positions at Stanford University (USA), Arizona State University (USA) and Yandex (Russia).

Take a look at [Vladimir's slides](#)

.....
Monday 21 November 2022, 2-3pm - Anastasia Borovykh (Imperial College London)

[Website](#)

This event will take place in the Leverhulme Library (COL 6.15).

Title - Towards explainable and privacy-preserving machine learning

Abstract - In the last decade, fuelled by drastic increases in computational power and the wide availability of data (i.e. big data), machine learning, and specifically deep neural networks, loosely inspired by neuronal structures in the brain, have been increasingly deployed in the real world. Despite the satisfactory performance achieved in practical applications, these models are generally difficult to analyse and their performance is not always fully understood. This impacts the deployment of neural network models as it directly influences two critical real-world challenges: generalisation - guaranteeing good performance of the model in unseen scenarios and privacy - ensuring the trained model does not give away sensitive information about the datasets it was trained on. In this talk we will first go into more detail on the challenges associated with generalisation and privacy. We will then discuss several recent advancements in defining robust and privacy-preserving machine learning algorithms.

Biography - Anastasia Borovykh is currently an Assistant Professor (lecturer) at the Department of Mathematics at Imperial College London and the Imperial-X initiative. Her group works on computational models to understand and improve information processing in artificial and biological intelligent systems through a combination of tools from stochastic processes, statistical mechanics and

mathematical modeling. They apply this in i) explainable machine learning, ii) privacy-preserving machine learning and iii) neuroscience.

Anastasia also enjoys collaborating with industry on applied problems that use machine learning and optimization in for example finance, smart cities and healthcare.

Prior to her current position, Anastasia was an Assistant Professor at the University of Warwick at the Operations Research department and did postdocs at Imperial College London and CWI Amsterdam. She obtained her PhD cum laude from the University of Bologna as part of a Marie-Curie ITN-EID project. Her MSc was in Quantitative Finance at the VU Amsterdam and my BSc in Applied Mathematics from the Delft University of Technology.

Take a look at [Anastasia's slides](#).