Geometry and Topology in Data Science (Spring 2024)

Organization

- **The Organizers**: Alvaro Diaz, Marzieh Eidi, Celia Hacker, Guillermo Restrepo, Daniel Spitz, Diaaeldin Taha, Francesca Tombari
- **MPI Seminar Page**: Mathematical Methods in Data Science
- **Contact**: To contact the organizers, email the lab at lab [at] mis [dot] mpg [dot] de.
- **Mailing List**: To stay informed of Lab activities, including this group's meetings, join the Lab mailing list.

Schedule

Week	Date	Time	Location	Speaker	Title
Week 19	Fri 10.05	09:00–10:30	E1 05	Justin Curry	"To Predict is NOT to Explain"
Week 20	Fri 17.05	09:00–10:30	—	—	CAG Conference - **NO MEETING**
Week 21	Fri 24.05	09:00–10:30	G3 10	Celia Hacker	TBA
		TBA	TBA	Francesca Tombari	TDA Course 01/06
Week 22	Fri 31.05	09:00–10:30	G3 10	Jeff Philips	The Geometry of Kernel Methods and Kernel Range Spaces
		TBA	G3 10	Francesca Tombari	TDA Course 02/06
Week 23	Fri 07.06	09:00–10:30	E1 05	Simon Telen	TBA
		—	—	Francesca Tombari	**NO COURSE**
Week 24	Fri 14.06	09:00–10:30	G3 10	Parvaneh Joharinad	TBA
		TBA	G3 10	Francesca Tombari	TDA Course 03/06
Week 25	Fri 21.06	09:00–10:30	E1 05	Bei Wang	Topology-Preserving Data Compression
		TBA	TBA	Francesca Tombari	TDA Course 04/06
Week 26	Fri 28.06	—	—	—	ScaDS Summer School - **NO MEETING**
		TBA	TBA	Francesca Tombari	TDA Course 05/06
Week 27	Fri 05.07	09:00–10:30	E1 05	Marzieh Eidi	TBA
		TBA	TBA	Francesca Tombari	TDA Course 06/06
Week 28	Fri 12.07	09:00–10:30	E1 05	Karel Devriendt	TBA
Week 29	Fri 19.07	09:00–10:30	E2 10	SPEAKER TBA	TBA
Week 30	Fri 26.07	09:00–10:30	E1 05	SPEAKER TBA	TBA

Abstracts

Week 19

Speaker: Justin Curry (University of Albany, USA)

Coordinates: Fri 10.05, 9–10:30 AM, MiS E1 05

Title: "To Predict is NOT to Explain"

Abstract: Modern day neural networks are amazing prediction machines, but to get at explanations one has to understand higher order relations between data as they fiber over their predictions. In this talk I will connect the urgent questions of modern data science with the distinguished history of applied topology by considering simple geometric examples and probing them with increasingly complicated tools. Ideas from dynamics, stratification theory and sheaf theory will be introduced in a loose and intuitive fashion to trace future directions for research.

Week 20

Week 21

Week 22

Speaker: Jeff Philips (University of Utah, USA)

Jeff Phillips is a Professor in the School of Computing at the University of Utah. He founded the Utah Center for Data Science, and directs the Data Science academic program there. He works on geometric data analysis, algorithms for big data, and how these intersect with data science. His book, Mathematical Foundations for Data Analysis, was published by Springer-Nature in 2021.

Coordinates: Fri 31.05, 9–10:30 AM, MiS G3 10

Title: The Geometry of Kernel Methods and Kernel Range Spaces

Abstract: I will start by overviewing kernel methods in machine learning, and how the simple kernel trick allows one to effortlessly turn intuitive linear methods into non-linear ones. While these methods can seem mysterious, I'll try to give insight into the geometry that arises, especially in kernel SVM. This will lead into kernel range spaces, which describes all the ways one can inspection a data set with a kernel. From there I will discuss approximation of these with coresets, and just approximating the spaces themselves which leads to surprising results in high dimensions.

Week 23

Week 24

Week 25

Speaker: Bei Wang (University of Utah)

Coordinates: Fri 21.06, 9–10:30 AM, MiS E1 05

Title: Topology-Preserving Data Compression

Abstract: Existing error-bounded lossy compression techniques control the pointwise error during compression to guarantee the integrity of the decompressed data. However, they typically do not explicitly preserve the topological features in data. When performing post hoc analysis with decompressed data using topological methods, preserving topology in the compression process to obtain topologically consistent and correct scientific insights is desirable. In this talk, we will discuss a couple of lossy compression methods that preserve the topological features in 2D and 3D scalar fields. Specifically, we aim to preserve the types and locations of local extrema as well as the level set relations among critical points captured by contour trees in the decompressed data. This talk is based on joint works with Lin Yan, Xin Liang, Hanqi Guo, and Nathan Gorski.

Week 26

Week 27

Week 28

Week 29

Week 30