Abstract

Oberwolfach Workshop:

Machine Learning for Science: Mathematics at the Interface of Data-driven and Mechanistic Modelling

Dates:

11 Jun - 17 Jun 2023 (Code: 2324)

Organizers:

Neil Lawrence, Cambridge UK
Jessica Montgomery, Cambridge UK
Bernhard Schölkopf, Tübingen

Science advances our understanding of the world around us by collecting data and using mathematical and statistical tools to extract insights. This century a major revolution has occurred in terms of the quantity and accessibility of data. Today, automated data analysis methods – machine learning, for example – offer new tools for scientists to interrogate these complex datasets. Such methods are often collectively referred to in the mainstream media as Artificial Intelligence. These techniques have enabled advances in many fields, from understanding patterns of protein folding, to identifying new astronomical phenomena. The promise of these methods suggests that automated data analysis could revolutionise our approach to science, generating new hypotheses, automating experimental processes, and developing novel theories.

Despite this promise, the complexity of modern data sets, and limitations in our theoretical understanding of these data are limiting our rate of progress. We have made impressive advances in its ability to identify patterns in data and make predictions about the future state of a system, but we are poorly equipped to examine why such patterns arise or what implications they have for the system being studied. This understanding is vital for the progress of science. Researchers need analytical tools that not only reveal what the dynamics of a system are, but that also provide information about the cause-effect relationships underlying those dynamics, how these relationships influence the system's response to perturbation, and how that understanding should be succinctly summarised to inform human operators.

These limitations need to be overcome through advances in our foundational mathematical understanding of systems that learn from data that bridge those systems to classical modelling approaches that relate to the physical interpretability of the system.

Mathematics is the toolset we require to overcome these limitations. Innovations in the mathematics of machine learning can help researchers establish the causal structures that underpin their data and model complex systems in ways that allow them to test different hypotheses about how a system works. This workshop will explore how such mathematical innovations can help produce machine learning tools that can be deployed in support of scientific discovery, creating new interfaces between physical and data-driven modelling approaches.