

Toward Causal Machine Learning

Bernhard Schölkopf

Director of the Max Planck Institute for Intelligent Systems
Tübingen Germany

In machine learning, we use data to automatically find dependences in the world, with the goal of predicting future observations. Most machine learning methods build on statistics, but one can also try to go beyond this, assaying causal structures underlying statistical dependences. Can such causal knowledge help prediction in machine learning tasks? We argue that this is indeed the case, due to the fact that causal models are more robust to changes that occur in real world datasets. We discuss the implications of causal models for machine learning tasks such as domain adaptation, transfer learning, and semi-supervised learning. We present an application to the removal of systematic errors for the purpose of exoplanet detection.

Machine learning and "big data" analysis currently mainly focuses on well-studied statistical methods. Some of the causal problems are conceptually harder, however, the causal point of view can provide additional insights that have substantial potential for data analysis.