Enterprise Platform and Integration Concepts: Research Group of Prof. Dr. Hasso Plattner

Master Project Winter Term 2018/19

Design and Implementation of a Causal Inference Pipeline

Motivation

The questions that motivate most data analysis are not associational but causal in nature. What are the causes and what are the effects of events under observation? Nevertheless, the statistical methods commonly used today to answer those questions are of associational nature. But “Correlation does not imply causation!”, and misinterpretation often results in incorrect deduction. In the recent years, causality has grown from a nebulous concept into a mathematical theory that allows for causal reasoning.

Based on this development, this master project has the intention to engineer a modular IT system that enables the application of machine learning techniques in the context of causal inference. Therefore, you will build and combine the components needed to prepare, analyze and process diverse datasets to detect causal dependencies. As part of the journey, you will have the opportunity to deepen your knowledge about tools for data science, improve machine learning skills and influence the architecture of an end-to-end pipeline for causal inference.

Project Goals

1. Learn basic concepts of Causal Inference
2. Research & understand existing tools for data scientists
3. Design & implement a pipeline for Causal Inference
4. Optimize your implementation for underlying hardware

Technology & Skills

The core of the work will be based upon existing implementations and previously developed extensions (e.g., using In-Memory Database Management Systems or GPU-Acceleration) of the PC-algorithm.
Prior understanding of the fundamentals of machine learning techniques (e.g., having attended the lectures Causal Inference – Theory and Applications, Big Data Analytics or equivalent) is expected as well as knowledge in one of the following areas (C++, R, Python, In-Memory Databases, GPU acceleration).

Contact

You are welcome to visit us in the “Villa” or reach out to one of the following contacts:

Dr. Matthias Uflacker, Johannes Hügle, Christopher Schmidt