

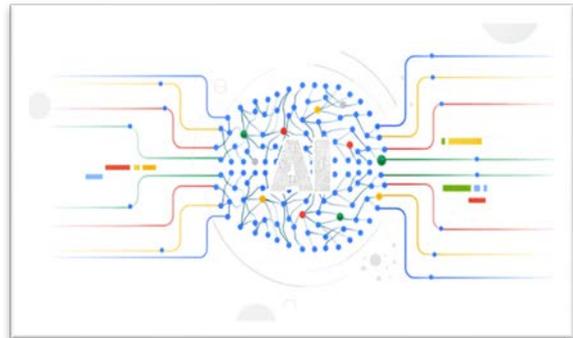
Advanced Topics in Data Management

Summer Term 2021

Lecturers: Prof. Dr. B. Mitschang, Dr. P. Reimann, Dr. H. Schwarz
Contact: Holger Schwarz, holger.schwarz@ipvs.uni-stuttgart.de
Language: English

Description

The comprehensive analysis of large volumes of data is crucial in modern information systems. This is often characterized by terms like big data and advanced analytics. Some of the main driving forces are current trends such as Industrie 4.0 and the Internet of Things (IoT), which result in new challenges for large scale data analysis. In organizations, typically a multitude of machine learning (ML) models are derived from data. These ML models have to be managed in a comprehensive way. Furthermore, the most suitable models have to be selected for a concrete application, e.g., via approaches to model discovery or AutoML. Very often, the data the models were derived from, changes over time (data set shifts in non-stationary environments). Here, major challenges are how to detect and characterize these changes (e.g., different concept drifts), as well as how to consider them in adaptive and incremental machine learning strategies. Moreover, many machine learning models provide some kind of prediction. Here, it is often mandatory for the users to understand why a certain prediction was provided by the model. Finding methods to provide such explanations and thus to help people interpret the models correctly, instead of treating them just as black boxes, is another important challenge in large scale data analysis. In this seminar, we cover important topics in the mentioned fields, for example:



- Explainable artificial intelligence / interpretable machine learning
- Machine learning in non-stationary environments
- Comprehensive management of machine learning models
- Automatic selection of machine learning models
- ...

We address current technologies, concepts, algorithms and system infrastructures for large scale data analysis and focus on the above-mentioned challenges.

Remarks

- Seminar topics will be introduced and assigned to the participants in a first meeting in KW6. Registered participants will be informed about the date by email.
- In this advanced seminar, each student will work on one specific topic. Basic literature will be provided by the advisors. Each student will summarize results in a document of about 20 pages and a 30 minutes presentation.

Prerequisites

Basic knowledge on database systems and information systems, e.g., from the lectures "Modellierung" or „Data Warehouse, Data Mining and OLAP“, is mandatory.