Analytical Software and Frameworks

Course Description

The Analytical Software and Frameworks course provides insight into contemporary software and platform solutions for data analytics in business. The course introduces relevant frameworks and software used in modern data science projects. Commercial and open-source for cloud computing, distributed computing and machine learning, as well as a commercial development platform for in-memory database analytics, are covered. Additional software solutions may be covered by the lecturer as convenient. In particular, in the written assignment, students will be required to apply their technological knowledge to a specific scenario that requires interdisciplinary thinking of how to merge the specifics of a given application domain with the technological options.

Contents

1. Introduction
   1. Software Systems
   2. Frameworks
   3. Distributed Computing
   4. Databases and Data Warehousing
2. Data Storage
   1. Data Clustering
   2. Data Replication
   3. Data Indexing
   4. Data Warehousing
3. Statistical Modeling Frameworks
   1. The R Project for Statistical Computing
   2. The Python Ecosystem
4. Machine Learning and Artificial Intelligence
   1. Overview of Modern Machine-Learning Frameworks
   2. Introduction to TensorFlow and Keras
5. Cloud Computing Platforms and On-Premises Solutions
   1. Advantages and Disadvantages of Cloud, On-Premises, and Edge Solutions
   2. Overview of Cloud Computing Solutions
6. Distributed Computing
   1. Overview of Distributed Computing Approaches
   2. Overview of Streaming Approaches
   3. Other Solutions
7. Database Technologies
   1. Overview of Database Approaches
      1. Row-Based vs. Column-Based
      2. In-Memory DB
      3. Relational DB vs. noSQL
      4. Timeseries DB
   2. Overview of Database Implementations