Data Science

Course Description

This course provides the framework to create value from data. After an introduction, the course covers how to identify suitable use cases and evaluate the performance of data-driven methods. In an interdisciplinary approach, the requirements from a specific application domain need to be understood and transferred to the technological understanding to identify the objectives and value proposition of a data science project. The course covers techniques for the technical processing of data and then introduces advanced mathematical techniques and selected methods from artificial intelligence that are used to analyze data and make predictions.

Contents

1. Introduction to Data Science
   1. Overview of Data Science
   2. Terms and Definitions
   3. Applications and Notable Examples
   4. Sources of Data
   5. Structured, Unstructured, and Streaming
   6. Typical Data Sources and their Data Types
   7. The Four “V”s of Data: Volume, Variety, Velocity, and Veracity
   8. Introduction to Probability Theory
   9. What Are Probabilities and Probability Distributions
   10. Introduction to Bayesian Statistics
   11. Relation to Data Science: Prediction as a Probability
2. Use Cases and Performance Evaluation
   1. Identification of Use Cases for Data Science
   2. Identifying Data Science Use Cases
   3. From Prediction to Decision: Generating Value from Data Science
   4. Evaluation of Predictions
   5. Overview of Relevant Metrics
   6. Business-Centric Evaluation: The Role of KPIs
   7. Cognitive Biases and Decision-Making Fallacies
3. Pre-Processing of Data
   1. Transmission of Data
   2. Data Quality and Cleansing of Data
   3. Transformation of Data (Normalization, Aggregation)
   4. Reduction of Data Dimensionality
   5. Data Visualization
4. Processing of Data
   1. Stages of Data Processing
   2. Methods and Types of Data Processing
   3. Output Formats of Processed Data
5. Selected Mathematical Techniques
   1. Linear Regression
   2. Principal Component Analysis
   3. Clustering
   4. Time-Series Forecasting
   5. Overview of Further Approaches
6. Selected Artificial Intelligence Techniques
   1. Support Vector Machines
7. Neural Networks and Deep Learning
8. Feed-Forward Networks
9. Recurrent Networks and Memory Cells
10. Convolutional Networks
11. Reinforcement Learning
12. Overview of Further Approaches