Introduction to Data Science

Course Description

Data science emerged as a multi-disciplinary field aimed at creating value from data. This course starts with an overview of data science and related fields and then defines data types and sources. Special focus is put on the assessment of data quality and electronic data processing. The use of data-driven methods has become vital for businesses, and this course outlines how data-driven approaches can be integrated within a business context and how operational decisions can be made using data-driven methods. Finally, this course highlights the importance of statistics and machine learning in the field of data science and gives an overview of relevant methods and approaches.

Contents

1. Introduction to Data Science
   1. Definition of the Term “Data Science”
   2. Data Science and Related Fields
   3. Data Science Activities
2. Data
   1. Data Types and Data Sources
   2. The 5 “V”s of Data
   3. Data Curation and Data Quality
   4. Data Engineering
3. Data Science in Business
   1. Identification of Use Cases
   2. Performance Evaluation
   3. Data-Driven Operational Decisions
   4. Cognitive Biases
4. Statistics
   1. Importance of Statistics for Data Science
   2. Important Statistical Concepts
5. Machine Learning
   1. Role of Machine Learning in Data Science
6. Overview of Machine Learning Approaches