

Data Analysis 3

A course introduction

Jan Platoš

September 16, 2019

Department of Computer Science
Faculty of Electrical Engineering and Computer Science
VŠB - Technical University of Ostrava



Course Content

- Associative Pattern and Rule Mining
- Clustering algorithms and principles
 - Representative based clustering
 - Hierarchical clustering
 - Density based clustering
 - Clustering validation
- Anomaly detection
- Classification
 - Feature selection
 - Probabilistic Classification
 - Decision trees
 - Rule-Based Classification
 - Linear discriminant analysis
 - Support Vector Machine
- Regression

- Aggarwal C.C. (2015), Data Mining: The Textbook, Springer.
- Bramer, M. (2013). Principles of data mining. Springer.
- Leskovec, J., Rajaraman, A., Ullman, J. D. (2014). Mining of massive datasets. Cambridge University Press.
- Witten, I. H., Frank, E. (2011). Data Mining: Practical machine learning tools and techniques [3rd Ed.]. Morgan Kaufmann.
- Zaki, M. J., Meira Jr, W. (2014). Data Mining and Analysis: Fundamental Concepts and Algorithms. Cambridge University Press.

Course Evaluation - Full-time students

- There is 30 points from exercises (10x3):
 - up to 3 points for finishing task (small implementation/analytical tasks) at the exercise or at home (1-2 hours of work).
- Small presentation for 20 points.
 - Presentation related to the data analysis or implementation task below. (1-2h)
- Method implementation and results comparison for 25 points.
 - Evaluation of the implemented method on the standard dataset and comparison of the results with standard implementation (2-4h).
- Dataset analysis for 25 points.
 - A report (PDF) about analysis of the selected dataset with standard algorithm (2-4h).

Questions?