

Data Mining

Benno Stein Theo Lettmann

Contents

- I. Introduction
- II. Cluster Analysis
- III. Nearest Neighbor Strategies
- IV. Latent Variables Analysis
- V. Association Analysis

Objectives

- ❑ understand and explain the basic concepts of data mining
- ❑ understand formalized concepts and methods and be able to implement them in the form of algorithms
- ❑ sensibly select, adapt, and apply relevant methods
- ❑ being able to educate oneself

Related Fields

1. Statistics [paradigms, models]
2. Mathematics
3. Information Retrieval [methods, algorithms]
4. Knowledge Processing
5. Heuristic Search
6. Decision Support Systems [applications]
7. Business Intelligence
8. Web Technology

Literature

Data Mining:

- ❑ D. Hand, H. Mannila, P. Smyth.
Principles of Data Mining
Bradford, 2001.
- ❑ P.N. Tan, M. Steinbach, V. Kumar.
Introduction to Data Mining
1st edition, Addison Wesley, 2005.
- ❑ I.H. Witten, E. Frank.
Data Mining: Practical Machine Learning Tools and Techniques
3rd edition, Morgan Kaufmann, 2011.

Software

Programming:

- ❑ Eclipse Foundation, Inc., Canada.
Eclipse SDK
Version 4.5. www.eclipse.org/downloads

Statistics:

- ❑ R Development Core Team.
R
Version 3.2. www.r-project.org
- ❑ E. Jones, T. Oliphant, P. Peterson and others.
SciPy
Version 0.16. www.scipy.org
- ❑ J.W. Eaton.
GNU Octave
Version 4.0. www.gnu.org/software/octave