

# **Data Mining and Knowledge Discovery**

Course level: Master

Course code: MLDM DMKD ECTS Credits: 4.00

**Course instructors:** Baptiste Jeudy, Fabrice Muhlenbach (UJM, Saint- Etienne)

# Education period (Dates): 2<sup>nd</sup> semester

# Language of instruction: English

### Aim and learning outcomes:

This course presents an advance study of some data mining algorithms and techniques. The necessary theoretical background is also provided.

# Topics to be taught (may be modified)~17h:

- Itemset and association rule mining: principles, APRIORI and ECLAT algorithms
- Constraint-based pattern mining: types of constraints
- Condensed Representations: formal concepts, MDL heuristic patterns (KRIMP algorithm)
- Other patterns: sequence, stream and graph mining

### Practical Laboratory Sessions and Tutorials~13h:

- 1. Survey of data mining softwares
- 2. Introduction to R
- 3. Basics of data mining with R (data description, clustering, classification, overfitting problem)
- 4. Data mining process (association rule mining, mining frequent patterns)
- 5. Efficient data mining with R, R and SQL, R and NoSQL
- 6. Data Mining applications (recommender systems, crime analysis, football mining with R)

### Teaching methods: Lectures and lab classes.

Form(s) of Assessment: written exam (70%), practical work (30%)

### Literature and study materials:

Basic textbooks:

- Introduction to Data Mining, Pang-Ning Tan, Michael Steinbach, Vipin Kumar.
- Data Mining: Concepts and Techniques, Jiawei Han and Micheline Kamber. 2nd ed. The Morgan Kaufmann Series in Data Management Systems

### Additional information:

Baptiste Jeudy, Fabrice Muhlenbach. University Jean Monnet, Saint- Etienne E-mail: {fabrice.muhlenbach,baptiste.jeudy}@univ-st-etienne.fr

### Home page: http://mldm.univ-st-etienne.fr

e-mail: master.MLDM@univ-st-etienne.fr

