COURSE DESCRIPTION – ACADEMIC YEAR 2017/2018

Course title	**Data Mining**
Course code	72113
Scientific sector	INF/01
Degree	Master in Computer Science (LM-18)
Semester	1
Year	1
Credits	8
Modular	No

Total lecturing hours	48
Total lab hours	24
Total exercise hours	--
Attendance	Not compulsory, but strongly recommended.
Prerequisites	Students should have a good knowledge of probability and algorithms and suitable programming skills for projects.
Course page	https://www.inf.unibz.it/~mkacimi/teaching.shtml

Specific educational objectives	The course belongs to the type "caratterizzanti – discipline informatiche" in the curriculum "Data and Knowledge Engineering". It is designed to give a general overview of scientific contents about Data Mining. This course aims at developing a deep understanding of the strengths and the limitations of a wide range of data mining techniques to be able to identify their use cases and important applications. Given a data-mining problem, students will be able to define what are the data sources to be exploited, the mining tasks to be performed, and the algorithms that need to be used to solve the problem. To this end, students will have the possibility to actively participate in data mining projects to perform extensive experiments on real datasets and potentially propose extensions to existing algorithms. A useful takeaway from the course will be the ability to identify the knowledge to be mined from data and employ mining algorithms using easy-to-use software and cases.

Lecturer	Mouna Kacimi El Hassani
Contact	Piazza Domenicani 3, Room 2.02, Mouna.Kacimi@unibz.it
Scientific sector of lecturer	INF/01
Teaching language	English
Office hours	To be arrange beforehand by email or during the lecture.
Lecturing Assistant (if any)	--
Contact LA	--
Office hours LA	--
List of topics	• Data types, quality and pre-processing • Data exploration • Classification • Association analysis • Clustering • Error estimation • Data mining applications

	• Data mining tools
Teaching format	Lectures and labs with theoretical and practical exercises. Group projects.

Learning outcomes	Knowledge and understanding • Know the main techniques for extracting information (associations, trends, dependencies, forecasts) from structured and unstructured data. • Understand the methods of mathematics and statistics which are of support to information technology and its applications. Applying knowledge and understanding • Be able to identify new application requirements and business opportunities in the field of systems based on data and knowledge. • Be able to design and execute experimental analyses on information systems or their components. Making judgments • Be able to independently select the documentation required to keep abreast of the frequent technological innovations in the field by using a wide variety of documentary sources: books, web, magazines. • Be able to identify reasonable work goals and estimate the resources required to achieve the objectives. Communication skills • Be able to structure and prepare scientific and technical documentation describing project activities. • Be able to coordinate the work of a project team and to interact positively with members of the group. Ability to Learn • Be able to autonomously extend the knowledge acquired during the study course by reading and understanding scientific and technical documentation. • Be able to independently keep up to date with developments in the most important areas of Data Mining.

Assessment	• **Final written exam**: with verification questions and problem solving tests • **Project**: consists in two parts (1) acquiring the basic data mining knowledge via assignments and (2) applying or adapting existing techniques to solve a concrete problem chosen by the student. The project results are a written report, implementation of the used techniques, and a presentation. • **Mid-term exam (optional)**: with verification questions and problem solving tests
Assessment language	English
Evaluation criteria and criteria for awarding marks	**Evaluation Criteria** • Written final exam: 50% of the mark • Project: 50% of the mark • The project marks are valid for the three exam sessions.

	The project is a prerequisite for attending the written exam **Criteria for awarding marks** • Exam: correctness and clarity of answers, the ability to adequately solve data mining problems and to understand how to choose the right technique. • Project: ability to apply data mining tools and algorithm in a real world problem, creativity, and ability to work in team.

Required readings	• Jiawei Han and Micheline Kamber, "Data Mining: Concepts and Techniques", Second Edition, 2006
Supplementary readings	Additional sources will be announced during the course.
Software used	• Weka and Java