

Content

Course Code	Course Name	Semester	Theory	Practice	Lab	Credit	ECTS
INF 511	Data Science	1	3	0	0	3	6

Prerequisites	
Admission Requirements	

Language of Instruction	English
Course Type	Elective
Course Level	Masters Degree
Objective	This class aims at introducing the data mining process to students. This includes the description of data preparation and preprocessing, of various data mining algorithms and of the tools available to assess their results. The class focuses on standard approaches regarding association rules mining, supervised classification and unsupervised classification (clustering). Basic statistical knowledge is necessary to understand the mining algorithms and the quality assessment tools.
Content	<ul style="list-style-type: none">- data pre-processing- supervised classification- clustering- complex data mining- results validation and quality assessment
References	<ul style="list-style-type: none">• Data Mining - Practical Machine Learning Tools, 2nd edition, Ian H. Witten & Eibe Frank, Morgan Kaufmann, 2005.• Neural Networks - A Comprehensive Foundation, 2nd edition, Simon Haykin, Pearson/Prentice Hall, 1999.• Data Mining: Concepts and Techniques, Jiawei Han & Micheline Kamber, Morgan Kaufmann, 2000.• Applied Statistics and Probabilities for Engineers, 4th edition, D.C. Montgomery & G.C. Runger, John Wiley & sons, 2006.• The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd edition, T. Hastie, R. Tibshirani & J. Friedman, Springer, 2009.

Theory Topics

Week	Weekly Contents
1	Introduction
2	Data preparation
3	Association rules and a priori algorithm
4	FP-trees and complex rules
5	Decision trees and naïve Bayes classifier
6	Statistical regression and Bayesian networks
7	Neural networks and other classifiers
8	Quality assessment on classification results
9	Classifier comparison
10	Distance and partitioning
11	Hierarchical clustering methods
12	Clustering with grids and density

Week	Weekly Contents
13	Model-based processing
14	Outliers detection