Content

Course Code	Course Name	Semester	Theory	Practice	Lab	Credit	ECTS
IT 533	Data Science and Applications	2	4	0	0	3	8

Prerequisites	
Admission Requirements	

Language of Instruction	English
Course Type	Compulsory
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	W1: Introduction, overview W2: Descriptive Statistics W3: Data Preprocessing W4: Inferential Statistics and its preprocessing tools W5: Code Application 1 W6: Regression W7: Classification1 W8: Classification2 W9: Clustering1, 2 W10: Code Application 2 W11: Project Presentations
References	• 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 Willey & 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, overview
2	Descriptive Statistics
3	Data Preprocessing
4	Inferential Statistics and its preprocessing tools
5	Code Application 1
6	Regression
7	Classification1
8	Classification2
9	Clustering1,2
10	Code Application 2
11	Project Presentations