Content

Course Code	Course Name	Semester	Theory	Practice	Lab	Credit	ECTS
ISI 524	Data Science	1	3	0	0	3	6

Prerequisites	
Admission Requirements	

Language of Instruction	English
Course Type	Elective
Course Level	Masters Degree
Objective	This course aims to introduce students to the data mining process. The main objectives of the course include understanding and applying data preparation and preprocessing techniques, various data mining algorithms, and the tools used to evaluate their results. The course focuses on standard approaches related to association rule mining, supervised classification, and unsupervised classification (clustering). Basic statistical knowledge is required to understand mining algorithms and quality evaluation tools. In this way, the course aims to enable students to produce practical solutions in the field of data analysis.
Content	1. An introduction to data mining and predictive analytics 2. Data preprocessing, exploratory data analysis 3. Dimension-reduction methods, univariate statistical analysis 4. Multivariate statistics, preparing to model the data 5. Simple linear regression, multiple regression 6. Model building 7. k-nearest neighbor algorithm, decision trees 8. Logistic regression, naïve bayes and Bayesian networks 9. Midterm exam 10. Model evaluation techniques 11. Graphical evaluation of classification models 12. Hierarchical and k-means clustering, measuring cluster goodness 13. Association rules, ensemble methods 14. Student presentations
References	1. Data Mining - Practical Machine Learning Tools, 2nd edition, Ian H. Witten & Eibe Frank, Morgan Kaufmann, 2005. 2. Neural Networks - A Comprehensive Foundation, 2nd edition, Simon Haykin, Pearson/Prentice Hall,1999. 3. Data Mining: Concepts and Techniques, Jiawei Han & Micheline Kamber, Morgan Kaufmann, 2000. 4. Applied Statistics and Probabilities for Engineers, 4th edition, D.C. Montgomery & G.C. Runger, John Willey & sons, 2006. 5. 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	An introduction to data mining and predictive analytics
2	Data preprocessing, exploratory data analysis
3	Dimension-reduction methods, univariate statistical analysis
4	Multivariate statistics, preparing to model the data
5	Simple linear regression, multiple regression
6	Model building
7	k-nearest neighbor algorithm, decision trees
8	Logistic regression, naïve bayes and Bayesian networks
9	Midterm exam
10	Model evaluation techniques
11	Graphical evaluation of classification models

Week	Weekly Contents
12	Hierarchical and k-means clustering, measuring cluster goodness
13	Association rules, ensemble methods
14	Student presentations