Statistical Machine Learning

Instructor: D.KARLIS

Course Code: 61232
Course Type: Compulsory of Course Group 2
Course Level: Graduate (MSc)
Year of Study: A'
Semester: 2nd
ECTS: 7,5
Language: English

Course Description

A range of statistical and machine learning methods will be described for supervised and unsupervised learning problems. Unsupervised learning: clustering (hierarchical, partition clustering, k-means and its variants, model-based clustering), data reduction methods. Model Assessment and Selection. Supervised learning: Methods of Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), k-nn, decision trees, random forests, SVM, naïve Bayes and others. Cross-validation methods. Statistics for big data problems, new approaches. Regularizations. Statistical methods for networks. Smoothing approaches in regression.

Prerequisites

Statistical Inference

Target Learning Outcomes

Upon completion of the course, students will have the knowledge and the skills

- to implement statistical methods aiming to deal with the problem of data dimension reduction,
- to apply classification models/algorithms and access their performance
- to apply clustering and access its performance
- to be familiar with new methodologies developed to deal with big data.

Recommended Bibliography

- T. Hastie, R. Tibshirani and R. Friedman (2009) Elements of Statistical Learning, Springer.
- Witten, J, Hastie, T. and Tibshirani, R. (2011) Introduction to Statistical Learning with applications in R, Springer

- C. Giraud (2015). Introduction to High-Dimensional Statistics. Philadelphia: Chapman and Hall/CRC.
- E. D. Kolaczyk (2014) Statistical Analysis of Network Data with R. Springer

Teaching and Learning Activities

Course lasts 12 3-hours lectures (one each week). Every week there will be exercises as homework (some to be submitted). There will be also a team project.

Assessment and Grading Methods

The final grade is the weighted average of the final examination grade (80%) and the assignment/projects (20%).