Statistical Learning

Academic Year: (2023 / 2024) **Review date: 29-05-2023**

Department assigned to the subject: Statistics Department

Coordinating teacher: DELGADO GOMEZ, DAVID

Type: Compulsory ECTS Credits : 3.0

Year : 1 Semester : 2

OBJECTIVES

Become familiar with different analytical tools, based on data, to make business decisions

Capacity to develop skills to analyze and find relationships between many variables/features

Know how to evaluate supervised-learning models

Develop skills to classify observations based on probabilistic learning and machine learning tools

Handle the R language for statistical-learning tools

DESCRIPTION OF CONTENTS: PROGRAMME

1. Principal Component Analysis (PCA)
2. Multivariate Normal Distribution
3. Discriminant Analysis
4. Supervised Learning: k-Nearest Neighbors, Decision Trees, and Random Forests
5. Bias-Variance Tradeoff and Cross-Validation
6. Support Vector Machines (SVM)
7. Unsupervised Learning: K-means and Expectation-Maximization (EM) algorithm for Gaussian Mixture Models

LEARNING ACTIVITIES AND METHODOLOGY

Lectures (50% of the sessions): the contents of the course will be introduced, explained and illustrated with examples. Teaching materials will be provided on Aula Global.

Computer Labs (50% of the sessions): Examples and cases studies with the R language.

ASSESSMENT SYSTEM

% end-of-term-examination:	40
% of continuous assessment (assigments, laboratory, practicals…):	60

Group assignments and presentations in class (60%)
Final test (40%)

Extraordinary evaluation similar to the ordinary evaluation.

BASIC BIBLIOGRAPHY

- G. James, D. Witten, T. Hastie and R. Tibshirani An Introduction to Statistical Learning with Applications in R, Springer, 2013

- Kevin P. Murphy Machine Learning: A Probabilistic Perspective, The MIT Press, 2012

- Machine Learning with R Brett Lantz, Packt Publishing, 2015