

Academic Year: (2018 / 2019)

Review date: 18-09-2018

Department assigned to the subject: Statistics Department

Coordinating teacher: GALEANO SAN MIGUEL, PEDRO

Type: Compulsory ECTS Credits : 6.0

Year : 2 Semester : 1

REQUIREMENTS (SUBJECTS THAT ARE ASSUMED TO BE KNOWN)

Linear algebra
 Probability and Data Analysis
 Introduction to Statistical Modeling

OBJECTIVES

CB2: That students know how to apply their knowledge to their work or vocation in a professional manner and possess the skills that are usually demonstrated through the elaboration and defense of arguments and the resolution of problems within their area of study.

CB3: That students have the ability to gather and interpret relevant data (usually within their area of study) to make judgments that include a reflection on relevant issues of social, scientific or ethical nature.

CE3: Ability to correctly identify classification problems corresponding to certain objectives and data and to use the basic results of the multivariate analysis as a basic foundation of the methods of classification, clustering and reduction of the dimension.

CG1: Knowledge and adequate skills to analyze and synthesize basic problems related to engineering and data science, solve them and communicate them efficiently.

CG4: Ability to solve technological, computer, mathematical and statistical problems that may arise in engineering and data science.

CG5: Ability to solve mathematically formulated problems applied to diverse subjects, using numerical algorithms and computational techniques.

CG6: Ability to synthesize the conclusions obtained from the analyzes carried out and present them clearly and convincingly, both in writing and orally.

RA1: Have acquired advanced knowledge and demonstrated an understanding of the theoretical and practical aspects and work methodology in the field of science and data engineering with a depth that reaches the forefront of knowledge.

RA2: Power, through arguments or procedures developed and supported by them, apply their knowledge, understanding of these and their abilities to solve problems in complex or professional and specialized work environments that require the use of creative and innovative ideas.

RA3: Have the ability to collect and interpret data and information on which to base their conclusions including, when necessary and relevant, reflection on social, scientific or ethical issues in the field of their field of study.

RA4: Being able to cope with complex situations or requiring the development of new solutions in the academic, work or professional field within their field of study.

RA5: Know how to communicate clearly and accurately to all types of audiences (specialized or not), knowledge, methodologies, ideas, problems and solutions within the scope of their field of study.

RA6: Be able to identify their own training needs in their field of study and work or professional environment and to organize their own learning with a high degree of autonomy in all types of contexts (structured or not).

DESCRIPTION OF CONTENTS: PROGRAMME

1. Introduction to the statistical learning.
2. Evaluation of learning methods:
 - a. Cross-validation.
 - b. Bootstrap.
3. Distance-based methods:
 - a. k nearest neighbors.
 - b. k means.
 - c. Hierarchical clustering.

4. Model-based methods:
 - a. Mixture distributions.
 - b. The EM algorithm.
 - c. Model selection.
5. Discriminant analysis:
 - a. The Bayes rule.
 - b. Linear discriminant analysis.
 - c. Quadratic discriminant analysis.
 - d. Extensions.
6. Dimension reduction:
 - a. Principal component analysis.
 - b. Factor analysis.

LEARNING ACTIVITIES AND METHODOLOGY

AF1: THEORETICAL-PRACTICAL CLASSES. During the classes, the teacher will present the knowledge that students should acquire. The students will receive class notes and will have basic texts of reference to facilitate the follow-up of the classes and the development of the subsequent work. Exercises, practical problems on the part of the student will be solved and workshops and evaluation tests will be held to acquire the necessary skills.

AF3: INDIVIDUAL OR GROUP WORK OF THE STUDENT.

AF9: FINAL EXAM. In which the knowledge, skills and abilities acquired throughout the course will be assessed globally.

MD1: CLASS THEORY. Exhibitions in the teacher's class with support of computer and audiovisual media, in which the main concepts of the subject are developed and materials and bibliography are provided to complement the students' learning.

MD2: PRACTICES. Resolution of practical cases, problems, etc. raised by the teacher individually or in groups.

MD3: TUTORIALS. Individualized assistance (individual tutorials) or group (collective tutorials) to students by the teacher.

ASSESSMENT SYSTEM

SE1: FINAL EXAMINATION In which the knowledge, skills and abilities acquired throughout the course will be assessed globally.

SE2: CONTINUOUS EVALUATION. In it, work, presentations, debates, exhibitions in class, exercises, practices and work in the workshops throughout the course will be evaluated.

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

BASIC BIBLIOGRAPHY

- Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani An Introduction to Statistical Learning, Springer, 2013

ADDITIONAL BIBLIOGRAPHY

- Bradley Efron, Trevor Hastie Computer Age Statistical Inference, Cambridge University Press, 2016
- Trevor Hastie, Robert Tibshirani, Jerome Friedman The Elements of Statistical Learning, Springer, 2009
- Trevor Hastie, Robert Tibshirani, Martin Wainwright Statistical Learning With Sparsity, CRC Press, 2015