

Academic Year: ( 2018 / 2019 )

Review date: 07/05/2018 18:47:21

Department assigned to the subject: Statistics Department

Coordinating teacher: DELGADO GOMEZ, DAVID

Type: Basic Core ECTS Credits : 6.0

Year : 1 Semester : 1

Branch of knowledge: Engineering and Architecture

## OBJECTIVES

At the end of the course students will be able:

To understand the importance of data science in today's knowledge society.

Use data visualization techniques to understand the problems faced by a data scientist and to report the results obtained.

To know when to use a supervised or an unsupervised data analysis technique.

To know the main data analysis techniques and applications where they have been used successfully.

To know the main problems a data scientist may encounter and how to deal with them

To know the different actual data analysis tools.

## DESCRIPTION OF CONTENTS: PROGRAMME

1. The Importance of Data Science
2. Understanding the Data: Case Studies of Exploratory Data Analysis and Visualization Techniques I
3. Understanding the Data: Case Studies of Exploratory Data Analysis and Visualization Techniques II
4. Importance of a good design of the experiment and choice of performance measures: precision, sensitivity, specificity, ROC curves. Over-fitting.
5. Introduction to unsupervised techniques: case studies of clustering I
6. Case studies of clustering II
7. Introduction to unsupervised classification: case studies on decision trees and random forests.
8. Case studies on data reduction techniques (Principal Component Analysis, Independent Component Analysis, Fisher Discriminant Analysis).
9. Introduction to Regression: Case Studies of Linear Regression.
10. Case studies of Logistic Regression.
11. Case studies on probabilistic models.
12. Introduction to the state of the art: case studies on Support vector machines.
13. Case studies on Deep Learning.

## LEARNING ACTIVITIES AND METHODOLOGY

The course is taught in 14 theoretical lessons and 14 practical lessons.

In the theoretical classes, the main concepts of the subject will be explained. These concepts will be further elaborated in the practical classes in which various computer-based data analyses will be carried out.

The students will also have a collective tutorial where they will have the opportunity to resolve any doubts they may have about the theoretical and practical classes or about the assignments they have to carry out.

## ASSESSMENT SYSTEM

**% end-of-term-examination/test:** 50

**% of continuous assessment (assignments, laboratory, practicals...):** 50

The final grade of the course will be obtained by averaging the grades of the continuous evaluation (50%) and the final exam(50%).

The continuous evaluation will consist of two assignments in which the student will have to apply the knowledge learned during the course.

#### BASIC BIBLIOGRAPHY

- Peter Bruce & Andrew Bruce Practical Statistics for data scientists:50 essential concepts, O'really, 2018