uc3m Universidad Carlos III de Madrid

Data Analitics

Academic Year: (2022 / 2023) Review date: 28-04-2022

Department assigned to the subject: Statistics Department Coordinating teacher: NOGALES MARTIN, FRANCISCO JAVIER

Type: Electives ECTS Credits: 3.0

Year: 4 Semester: 1

REQUIREMENTS (SUBJECTS THAT ARE ASSUMED TO BE KNOWN)

Linear algebra
Probability and Data Analysis
Introduction to Statistical Modeling

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. Introduction
 - 1.1 Basics of multivariate data analysis and statistical learning
 - 1.2 Supervised vs. unsupervised learning
 - 1.3 Data visualization techniques
- 2. Supervised Learning: Regression
 - 2.1 Linear regression
 - 2.2 Linear model selection and regularization
 - 2.3 Cross-validation on regression problems
- 3. Supervised learning
 - 3.1 Logistic regression
 - 3.2 Bayes classifier
 - 3.3 Linear discriminant analysis and k-Nearest neighbor classifier
 - 3.5 Random Forests
 - 3.6 Support vector machines
- 4. Unsupervised Learning and Dimensionality Reduction Techniques
 - 4.1 Clustering methods: k-means and hierarchical clustering
 - 4.2 Principal component analysis
 - 4.3 Multidimensional scaling
 - 4.4 ISOMAP and locally-linear embedding

LEARNING ACTIVITIES AND METHODOLOGY

Theory (3 ECTS), Practice (3 ECTS).

50% lectures with teaching materials available on the Web. The other 50% practical sessions (computer labs).

ASSESSMENT SYSTEM

The assessment will be made by weighting the continuous evaluation (50%) and the final exam (50%), with a minimum grade of 5 points over 10 in each assessment activity.

% end-of-term-examination:	50
% of continuous assessment (assigments, laboratory, practicals):	50