Data Analytics in IC4.0

Academic Year: (2020 / 2021) **Review date: 06-07-2020**

Department assigned to the subject: Statistics Department

Coordinating teacher: GUERRERO LOZANO, VANESA

Type: Compulsory ECTS Credits : 3.0

Year : 1 Semester : 1

OBJECTIVES

BASIC COMPETENCES:
- CB6: Possess and understand knowledge that provides a basis or opportunity to be original in the development and / or application of ideas, often in a research context
- CB7: That students know how to apply the knowledge acquired and their ability to solve problems in new or unfamiliar environments within broader (or multidisciplinary) contexts related to their area of study
- CB10: That students have the learning skills that allow them to continue studying in a way that will be largely self-directed or autonomous

GENERAL COMPETENCES:
- CG3: Capacity to develop basic distributed applications for the transport, storage and management of information.
- CG5: Capacity for basic analysis of the requirements for information management and treatment of large volumes of data.
- CG6: Capacity to adapt to changes in requirements associated with new products, new specifications and environments.

SPECIFIC COMPETENCES:
- CE8: Ability to apply augmented reality technology, in the context of Connected Industry 4.0
- CE9: Ability to identify computer security requirements in connected industry environments
- CE10: Programmatic data processing capabilities in solving particular problems of the connected industry

LEARNING RESULTS:
- Knowledge and use of data visualization techniques and tools.
- Understanding and practical use of regression and classification models (supervised learning).
- Understanding and practical use of clustering and dimensionality reduction models (unsupervised learning).

DESCRIPTION OF CONTENTS: PROGRAMME

1. Introduction
1.1 Basics of Multivariate Data Analysis
1.2 Introduction to Statistical Learning
1.3 Supervised vs. Unsupervised Learning
1.4 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
2.4 Extensions
3. Supervised Learning: Classification
3.1 Logistic Regression
3.2 Bayes classifier
3.3 Linear Discriminant Analysis
3.4 k-Nearest Neighbor classifier

3.5 Random Forests
3.6 Support Vector Machines
3.7 Cross-Validation on Classification problems
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

LEARNING ACTIVITIES:
- Theoretical and practical lessons using the statistical language R.
- Team work
- Individual work of the student

METHODOLOGY:
- Theoretical lessons, with support material available on the Web, to present and develop the main concepts of the course. Teachers with provide students with supplementary material.
- Critical reading of documents provided by the teachers: newspaper articles, reports, manuals and / or academic papers, either for later discussion in class, either to expand and consolidate the knowledge of the subject.
- Resolution of practical cases, problems, etc. proposed by the teacher individually or in groups.
- Preparation of projects individually or in group.

TUTORING SESSIONS:
- Weekly individual tutoring sessions
- Group tutorials might be possible

ASSESSMENT SYSTEM

100% of the assessment consists of individual and/or group projects carried out during the course.

For those students who do not follow the continuous assessment process, the grade in the extraordinary call is obtained through an exam consisting of both theoretical questions and practical problems.

% end-of-term-examination:	0
% of continuous assessment (assigments, laboratory, practicals…):	100

BASIC BIBLIOGRAPHY

- G. James, D. Witten, T. Hastie and R. Tibshirani An Introduction to Statistical Learning, Springer, 2013
- H. Wickham ggplot2. Elegant Graphics for Data Analysis, Springer, 2016
- T. Hastie, R. Tibshirani and J. Friedman The Elements of Statistical Learning, Springer, 2009
- T. Hastie, R. Tibshirani and M. Wainwright Statistical Learning with Sparsity, CRC Press, 2015