Syllabus
2020 - 2021

TECHNICAL SHEET OF THE SUBJECT

Data of the subject	
Subject name	Machine Learning
Subject code	E000009337
Quarter	Semestral
Credits	6,0 ECTS
Type	Obligatoria
Department	Department of Industrial Organization
Course overview	The purpose of this course is to provide students with a fundamental understanding and an extensive practical experience of how to extract knowledge from an apparently unstructured set of data. By the end of the course, students will: § Understand the basic principles behind machine learning. § Have practical experience with the most relevant machine learning algorithms. § Have well-form criteria to choose the most appropriate techniques for a given application.

Teacher Information	
Teacher	
Name	Mario Castro Ponce
Department	Department of Telematics and Computer Sciencies
Office	Alberto Aguilera 25 [D-411]
EMail	marioc@iit.comillas.edu
Phone	4224

SPECIFIC DATA OF THE SUBJECT

Contextualization of the subject

Prerequisites

Students willing to take this course should be familiar with linear algebra, basic probability and statistics, and undergraduate-level programming. Previous experience with the R programming language is also desired although not strictly required.

Competencies - Objectives

THEMATIC BLOCKS AND CONTENTS

Contents - Thematic Blocks

Contents

Unit 1. Introduction

1.1 Data mining & machine learning
1.2 The learning process
1.3 Smart industry levers and drivers
1.4 Types of machine learning

Unit 2. Classification methods

2.1 The classification problem
2.2 Logistic regression
2.3 Discriminant analysis
2.4 K-nearest neighbors
2.5 Decision trees
2.6 Support vector machines
2.7 Multilayer perceptrons for classification

Unit 3. Regression methods

3.1 The regression problem
3.2 Linear regression. Model selection and regularization
3.3 Polynomial regression
3.4 Splines
3.5 Generalized additive models
3.6 Multilayer perceptrons for regression
3.7 Radial basis function networks

Unit 4. Time series forecasting

4.1 Stochastic processes
4.2 Exponential smoothing
4.3 Decomposition methods
4.4 ARIMA models
4.5 Dynamic regression models

Unit 5. Unsupervised learning

5.1 Probability density estimation
5.2 Dimensionality reduction methods
5.3 Clustering and vector quantization
5.4 Self-organizing feature maps

TEACHING METHODOLOGY

General methodological aspects of the subject

EVALUATION AND CRITERIA

Ratings

Regular assessment

Theory will account for 50%, of which:

- Midterm:15%
- Final exam:35%

Lab will account for the remaining 50%
In order to pass the course, the weighted average mark must be greater or equal to 5 out of 10 points, and the mark of the final exam must be greater or equal to 4 out of 10 points. Otherwise, the final grade will be the lower of the two marks.

Retake

Lab marks will be preserved. In addition, all students will take a final exam. The resulting grade will be computed as follows:
Final exam: 50%
Lab practices: 50%
As in the regular assessment period, in order to pass the course, the weighted average mark must be greater or equal to 5 out of 10 points, and the mark of the final exam must be greater or equal to 4 out of 10 points. Otherwise, the final grade will be the lower of the two marks.

BIBLIOGRAPHY AND RESOURCES

Basic Bibliography

-Slides prepared by the lecturer (available in Moodlerooms).
-G. James, D. Witten, T. Hastie, and R. Tibshirani, An Introduction to Statistical Learning with Applications in R, Springer, 2013.