GUÍA DOCENTE
2022 - 2023

FICHA TÉCNICA DE LA ASIGNATURA

Datos de la asignatura	
Nombre completo	Machine Learning I
Código	DOI-MBD-514
Título	Máster en Big Data. Tecnología y Analítica Avanzada/Master in Big Data Technologies and Advanced Analytics
Impartido en	Máster en Big Data. Tec. y Analítica Avanzada/Master in Big Data Technologies and Advanced Analytics [Primer Curso]
Nivel	Master
Cuatrimestre	Semestral
Créditos	6,0 ECTS
Carácter	Obligatoria
Departamento / Área	Departamento de Organización Industrial
Responsable	José Portela
Horario de tutorías	A consultar con el profesor

Datos del profesorado	
Profesor	
Nombre	José Portela González
Departamento / Área	Departamento de Métodos Cuantitativos
Despacho	Santa Cruz de Marcenado 26
Correo electrónico	Jose.Portela@iit.comillas.edu
Profesor	
Nombre	Jaime Pizarroso Gonzalo
Departamento / Área	Departamento de Electrónica, Automática y Comunicaciones
Despacho	Santa Cruz de Marcenado 26
Correo electrónico	jpizarroso@comillas.edu
Teléfono	2732

DATOS ESPECÍFICOS DE LA ASIGNATURA

Contextualización de la asignatura

Aportación al perfil profesional de la titulación

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.

Prerequisitos

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

Competencias - Objetivos

Competencias

Competences[1]	
General competences	
CG1.	Have acquired advanced knowledge and demonstrated, in a research and technological or highly specialized conte detailed and well-founded understanding of the theoretical and practical aspects, as well as of the work methodolog one or more fields of study. *Haber adquirido conocimientos avanzados y demostrado, en un contexto de investigación científica y tecnológica o altam especializado, una comprensión detallada y fundamentada de los aspectos teóricos y prácticos y de la metodología de tra en uno o más campos de estudio.*
CG2.	Know how to apply and integrate their knowledge, understanding, scientific rationale, and problem-solving skills to and imprecisely defined environments, including highly specialized multidisciplinary research and professional contexts. *Saber aplicar e integrar sus conocimientos, la comprensión de estos, su fundamentación científica y sus capacidade resolución de problemas en entornos nuevos y definidos de forma imprecisa, incluyendo contextos de carácter multidiscipl tanto investigadores como profesionales altamente especializados.*
CG3.	Know how to evaluate and select the appropriate scientific theory and the precise methodology of their fields of stud order to formulate judgements based on incomplete or limited information, including, when necessary and pertine discussion on the social or ethical responsibility linked to the solution proposed in each case. *Saber evaluar y seleccionar la teoría científica adecuada y la metodología precisa de sus campos de estudio para form juicios a partir de información incompleta o limitada incluyendo, cuando sea preciso y pertinente, una reflexión sobr responsabilidad social o ética ligada a la solución que se proponga en cada caso.*
CG4.	Be able to predict and control the evolution of complex situations through the development of new and innovative v methodologies adapted to the scientific/research, technological or specific professional field, in general multidisciplir in which they develop their activity. *Ser capaces de predecir y controlar la evolución de situaciones complejas mediante el desarrollo de nuevas e innovad metodologías de trabajo adaptadas al ámbito científico/investigador, tecnológico o profesional concreto, en ger multidisciplinar, en el que se desarrolle su actividad.*
CG5.	Be able to transmit in a clear and unambiguous manner, to specialist and non-specialist audiences, results from scier

	and technological research or state-of-the-art innovation, as well as the most relevant foundations that support them. *Saber transmitir de un modo claro y sin ambigüedades, a un público especializado o no, resultados procedentes d investigación científica y tecnológica o del ámbito de la innovación más avanzada, así como los fundamentos más releva sobre los que se sustentan.*
CG6.	Have developed sufficient autonomy to participate in research projects and scientific or technological collaborations wi their thematic area, in interdisciplinary contexts and, where appropriate, with a high knowledge transfer component. *Haber desarrollado la autonomía suficiente para participar en proyectos de investigación y colaboraciones científic tecnológicas dentro de su ámbito temático, en contextos interdisciplinares y, en su caso, con una alta componente transferencia del conocimiento.*
CG7.	Being able to take responsibility for their own professional development and their specialization in one or more field study. *Ser capaces de asumir la responsabilidad de su propio desarrollo profesional y de su especialización en uno o más campo estudio.*
Specific competences	
CE3.	Be able to design and train systems that learn automatically, mastering both supervised and unsupervised learn techniques. Understand the potential application of these systems in the improvement of industrial processes, relat with clients, etc. *Ser capaces de diseñar y entrenar sistemas que aprendan de manera automática, dominando tanto las técnicas aprendizaje supervisado como no supervisado. Entender el potencial de aplicación de estos sistemas en la mejora de proc industriales, relación con clientes, etc.*

[1] Competences in English are a free translation of the official Spanish version.

Resultados de Aprendizaje

Learning outcomes	
By the end of the course students should:	
RA1.	Understand the basic principles behind machine learning.
RA2.	Have practical experience with the application of the most relevant machine learning algorithms.
RA3.	Have well-formed criteria to choose the most appropriate techniques for a given application.

BLOQUES TEMÁTICOS Y CONTENIDOS

Contenidos – Bloques Temáticos

Contents

Theory and laboratory

Unit 1. Introduction

1. Data mining & machine learning
2. The learning process
3. Smart industry levers and drivers
4. Types of machine learning

Unit 2. Classification methods

1. The classification problem
2. Logistic regression
3. Discriminant analysis
4. K-nearest neighbors
5. Decision trees
6. Support vector machines
7. Multilayer perceptrons for classification

Unit 3. Regression methods

1. The regression problem
2. Linear regression. Model selection and regularization
3. Polynomial regression
4. Splines
5. Generalized additive models
6. Multilayer perceptrons for regression
7. Radial basis function networks

Unit 4. Time series forecasting

1. Stochastic processes
2. Exponential smoothing
3. Decomposition methods
4. ARIMA models
5. Dynamic regression models

Unit 5. Unsupervised learning

1. Probability density estimation
2. Dimensionality reduction methods
3. Clustering and vector quantization
4. Self-organizing feature maps

METODOLOGÍA DOCENTE

Aspectos metodológicos generales de la asignatura

Metodología Presencial: Actividades

In-class activities	Competences
• **Lectures:** The lecturer will introduce the fundamental concepts of each unit, along with some practical recommendations, and will go through worked examples to support the explanation. Active participation will be encouraged by raising open questions to foster discussion and by proposing short application exercises to be solved in class either on paper or using a software package.	CG1, CG3, CG7, CE3
• **Lab sessions:** Under the instructor's supervision, students, divided in small groups, will apply the concepts and techniques covered in the lectures and will become familiar with the practical application of the most relevant algorithms using software tools and libraries.	CG1, CG2, CG3, CG4, CG5, C CG7, CE3
• **Tutoring** for groups or individual students will be organized upon request.	–

Metodología No presencial: Actividades

Out-of-class activities	Competences
• Personal study of the course material and resolution of the proposed exercises.	CG1, CG3, CG7, CE3
• Lab session preparation, analysis of the results and report writing.	CG1, CG2, CG3, CG4, CG5, C

RESUMEN HORAS DE TRABAJO DEL ALUMNO

STUDENT WORK-TIME SUMMARY		
IN-CLASS HOURS		
Lectures	Lab sessions	Assessment
28	28	4
OUT-OF-CLASS HOURS		
Self-study	Lab preparation and reporting	

60	60
ECTS credits:	**6 (180 hours)**

EVALUACIÓN Y CRITERIOS DE CALIFICACIÓN

Assessment activities	Grading criteria	Weight
Midterm exam	▪ Understanding of the theoretical concepts. ▪ Application of these concepts to problem-solving. ▪ Critical analysis of numerical exercises' results.	15%
Final exam	▪ Understanding of the theoretical concepts. ▪ Application of these concepts to problem-solving. ▪ Critical analysis of numerical exercises' results.	35%
Lab sessions and reports	▪ Application of theoretical concepts to real problem-solving. ▪ Ability to use and develop data mining and machine learning software. ▪ Attitude and effort: Initiative and proactive work will be encouraged ▪ Written communication skills.	50%

Calificaciones

GRADING AND COURSE RULES

Grading

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 fin 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

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 the two marks.

Course rules

- Class attendance is mandatory according to Article 93 of the General Regulations (Reglamento General) of Comill Pontifical University and Article 6 of the Academic Rules (Normas Académicas) of the ICAI School of Engineering. N complying with this requirement may have the following consequences:
  - Students who fail to attend more than 15% of the lectures may be denied the right to take the final exam during t regular assessment period.
  - Regarding laboratory, absence to more than 15% of the sessions can result in losing the right to take the final exa of the regular assessment period and the retake. Missed sessions must be made up for credit.
- Students who commit an irregularity in any graded activity will receive a mark of zero in the activity and disciplina procedure will follow (cf. Article 168 of the General Regulations (Reglamento General) of Comillas Pontifical University).

PLAN DE TRABAJO Y CRONOGRAMA

Actividades	Fecha de realización	Fecha de entrega
WORK PLAN AND SCHEDULE[1]		

In and out-of-class activities	Date/Periodicity	Deadline
Midterm exam	Session 14	–
Final exam	Exam period	–
Lectures and lab sessions	Twice a week	–
Review and self-study of the concepts covered in the lectures	After each lesson	–
Lab preparation	Before every lab session	–
Lab report writing	–	One week after the end of each session

[1] A detailed work plan of the subject can be found in the course summary sheet (see following page). Nevertheless, this schedule is tentative and may vary to accommodate the rhythm of the class.		

BIBLIOGRAFÍA Y RECURSOS

Bibliografía Básica

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*, Springe Second edition, 2021.

Bibliografía Complementaria

Complementary bibliography

- M. Kuhn and K. Johnson, *Applied Predictive Modeling*, Springer, 2013.
- T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning. Data Mining, Inference and Prediction*, 2nd E Springer, 2009.
- E. Alpaydin, *Introduction to Machine Learning*, 3rd Ed., MIT Press, 2014.
- S. Marsland, *Machine Learning: An Algorithmic Perspective*, 2nd Ed., Chapman & Hall/CRC Machine Learning & Patte Recognition, 2015.
- T. Mitchell, *Machine Learning*, McGraw-Hill, 1997.
- R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd Ed., Wiley-Interscience, 2000.
- C. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2007.
- S. Haykin, *Neural Networks. A Comprehensive Foundation*, 2nd Ed., Pearson, 1999.
- W. Wei, *Time Series Analysis. Univariate and Multivariate Methods*, 2nd Ed., Addison-Wesley, 2006.