GENERAL INFORMATION

Course information	
Name	Data Mining and Machine Learning for Smart Systems
Code	DEA-OPT-614
Degree	MII, MIT
Year	2º
Semester	Fall
ECTS credits	3 ECTS
Type	Elective
Department	Electronics, Control Engineering and Communications
Area	
Coordinator	Antonio Muñoz

Lecturer	
Name	Antonio Muñoz
Department	Electronics, Control Engineering and Communications
Area	
Office	D-513
e-mail	amunoz@comillas.edu
Phone	915 406 147
Office hours	Ask for an appointment by email.
Lecturer	
Name	
Department	
Area	
Room	
e-mail	
Tutorial timetable	

DETAILED INFORMATION

Contextualization of the course

Contribution to the professional profile of the degree

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 Data Mining and Machine Learning.
- Have practical experience with the most relevant Data Mining and Machine Learning algorithms.
- Have well-formed criteria to choose the most appropriate technique for a given application.

Prerequisites

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

CONTENTS

Contents

COMPETENCES AND LEARNING OUTCOMES

Competences and Learning Outcomes

Competences

General Competences

CG3. The capability of adapting to new theories, methods and changing engineering situations based on a sound technical training.

CG4. The capability of solving problems with personal initiative, efficient decision making, critical reasoning and transmitting technical information in the engineering world.

CG5. The capability of conducting measurements, calculations, assessments, studies, reports, planning, etc.

CG10. The ability to work in a multilingual and multidisciplinary environment.

Basic Competences

Specific Competences

Learning outcomes

RA1. The student understands the basic principles behind Data Mining and Machine Learning.

RA2. The student has a practical experience with the application of the most relevant Data Mining and Machine Learning algorithms.

RA3. The student has well-formed criteria to choose the most appropriate technique for a given application.

TEACHING METHODOLOGY

General methodological aspects

Each session will combine theory and practice. The teacher will explain the basics of the subject and will go in depth in the more important issues with illustrative examples. The students will be grouped in pairs in order to put in practice the proposed methods and techniques using software tools in a collaborative way.

In-class activities

1. **Lectures and problem-solving sessions (13 hours)**: The lecturer will introduce the fundamental concepts of each chapter, 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.
2. **Lab sessions (13 hours):** Under the instructor's supervision, students, divided in small groups, will apply the concepts and techniques covered in the lectures to real problems and will become familiar with the practical application of the most relevant algorithms using software tools and libraries.
3. **Assesment (4 hours)**

Off-class activities

1. **Personal study** of the course material and resolution of the proposed exercises (30 hours)
2. **Lab session** preparation, analysis of results and reporting (30 hours).

ASSESSMENT AND GRADING CRITERIA

Assessment activities	Grading criteria	Share
Mid-term 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. There will be an intra-group evaluation method to differentiate among team members.	50%

GRADING AND COURSE RULES

Grading

Regular assessment

- **Theory** will account for 50%, of which:
  - Mid-term: 15%
  - Final exam: 35%

- **Lab** will account for the remaining 50%

In order to pass the course, the mark of the final exam must be greater or equal to 4 out of 10 points.

Retakes

Lab practice 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 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.

Course rules

- Class attendance is mandatory according to Article 93 of the General Regulations (Reglamento General) of Comillas Pontifical University and Article 6 of the Academic Rules (Normas Académicas) of the ICAI School of Engineering. Not 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 the regular assessment period.

  - Regarding laboratory, absence to more than 15% of the sessions can result in losing the right to take the final exam 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 disciplinary procedure will follow (cf. Article 168 of the General Regulations (Reglamento General) of Comillas Pontifical University).

WORK PLAN AND SCHEDULE[1]

In and out-of-class activities	Date/Periodicity	Deadline
• Mid-term exam	Week 8	-
• Final exam	Last week	-
• Lectures + Lab sessions	Weekly	-
• Review and self-study of the concepts covered in the lectures	Weekly	-
• Lab preparation and reporting	Weekly	One week after the end of each lab session

STUDENT WORK TIME SUMMARY			
IN_CLASS HOURS			
Lectures	**Lab sessions**	**Assessment**	
13	13	4	
OFF_CLASS HOURS			
Self-study	**Lab preparation and reporting**		
30	30		
		ECTS credits:	3 (90 hours)

BIBLIOGRAPHY

Basic

- Notes prepared by the lecturer (available in Moodle).
- E. Alpaydin (2014), *Introduction to Machine Learning*, 3d Ed., MIT Press, ISBN-13 978-0-262-02818-9

Complementary

- C. Bishop (2007). *Pattern Recognition and Machine Learning*. Springer. ISBN-13: 978-0387310732
- R. Duda, P. Hart & D. Stork (2000). *Pattern Classification*. 2nd Ed. Wiley-Interscience. ISBN-13: 978-0471056690
- T. Hastie, R. Tibshirani & J. Friedman (2009). 2nd Ed. *The Elements of Statistical Learning. Data Mining, Inference and Prediction*. Springer. ISBN-13: 978-0387848570
- S. Marsland (2015), *Machine Learning: An Algorithmic Perspective*, 2nd Ed., Chapman & Hall/Crc Machine Learning & Pattern Recognition. ISBN-13: 978-1-4665-8333-7
- T. Mitchell (1997). *Machine Learning.* McGraw-Hill. ISBN-13: : 978-0070428072

[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.

	IN-CLASS ACTIVITIES				OUT-OF-CLASS ACTIVITIES				LEARNING OUTCOMES
Week	h/w	LECTURE & PROBLEM SOLVING	LAB	ASSESMENT	h/w	SELF-STUDY	LAB PREPARATION AND REPORTING	OTHER ACTIVITIES	Learning Outcomes
1	2	Introduction (1h)	Introduction to IDAT (1h)		4	Review and self-study (2h)	Practice with IDAT stastistical software (2h)		RA1, RA2, RA3
2	2	Classification I (1h)	Lab Practice 1 (1h)		4	Review and self-study (2h)	Lab preparation and report writing (2h)		RA1, RA2, RA3
3	2	Classification II (1h)	Lab Practice 2 (1h)		4	Review and self-study (2h)	Lab preparation and report writing (2h)		RA1, RA2, RA3
4	2	Regression I (1h)	Lab Practice 3 (1h)		4	Review and self-study (2h)	Lab preparation and report writing (2h)		RA1, RA2, RA3
5	2	Regression II (1h)	Lab Practice 4 (1h)		4	Review and self-study (2h)	Lab preparation and report writing (2h)		RA1, RA2, RA3
6	2	Forecasting I (1h)	Lab Practice 5 (1h)		4	Review and self-study (2h)	Lab preparation and report writing (2h)		RA1, RA2, RA3
7	2	Forecasting II (1h)	Lab Practice 6 (1h)		4	Review and self-study (2h)	Lab preparation and report writing (2h)		RA1, RA2, RA3
8	2	Forecasting III (1h)		Mid term exam (1h)	4	Review and self-study (2h)	Lab preparation and report writing (2h)		RA1, RA2, RA3
9	2	Density estimation (1h)	Lab Practice 7 (1h)		4	Review and self-study (2h)	Lab preparation and report writing (2h)		RA1, RA2, RA3
10	2	Clustering (1h)	Lab Practice 8 (1h)		4	Review and self-study (2h)	Lab preparation and report writing (2h)		RA1, RA2, RA3
11	2	Self Organising Maps (1h)	Lab Practice 9 (1h)		4	Review and self-study (2h)	Lab preparation and report writing (2h)		RA1, RA2, RA3
12	2	Genetic Algorithms I (1h)	Lab Practice 10 (1h)		4	Review and self-study (2h)	Lab preparation and report writing (2h)		RA1, RA2, RA3
13	2	Genetic Algorithms II (1h)	Lab Practice 11 (1h)		4	Review and self-study (2h)	Lab preparation and report writing (2h)		RA1, RA2, RA3
14	2			Final exam (3h)	4	Review and self-study (4h)	Report writing (4h)		RA1, RA2, RA3
15	2				4				