UAB

Universitat Autònoma
de Barcelona

2022/2023

Fundamentals of Machine Learning

Code: 106572
ECTS Credits: 6

Degree	Type	Year	Semester
2504392 Artificial Intelligence	OB	2	1

Contact

Name: Dimosthenis Karatzas

Email: dimosthenis.karatzas@uab.cat

External teachers

Oguz Mulayim

Use of Languages

Principal working language: english (eng)

Some groups entirely in English: Yes

Some groups entirely in Catalan: No

Some groups entirely in Spanish: No

Prerequisites

To have completed the first-year subjects of Fundamentals of Programming, Fundamentals of Mathematics, Probability and Statistics, and Data Engineering.

Objectives and Contextualisation

The objective of this subject is to provide comprehensive training in a wide range of fundamental machine learning techniques and algorithms that an engineer specializing in artificial intelligence must master.

Students will learn, implement and use a range of techniques covering both supervised and unsupervised learning. They will derive their own implementations, as well as explore existing libraries, and apply this new knowledge to solving practical problems.

Competences

- Develop critical thinking to analyse alternatives and proposals, both one's own and those of others, in a well-founded and argued manner.
- Identify, understand and apply the fundamental concepts and techniques of knowledge representation, reasoning and computational learning for the solution of artificial intelligence problems.
- Introduce changes to methods and processes in the field of knowledge in order to provide innovative responses to society's needs and demands.
- Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
- Work cooperatively to achieve common objectives, assuming own responsibility and respecting the role of the different members of the team.

Learning Outcomes

1. Develop critical thinking to analyse alternatives and proposals, both one's own and those of others, in a well-founded and argued manner.
2. Formulate the most appropriate cost function for a specific problem according to the characteristics of the experimental data and to the requirements/restrictions of the problem.
3. Propose new methods or informed alternative solutions.
4. Students must be capable of collecting and interpreting relevant data (usually within their area of study) in order to make statements that reflect social, scientific or ethical relevant issues.
5. Understand the concepts of bias and variance, and be able to use data preparation methods and regularisation techniques to obtain generalisable solutions from the available data.
6. Understand the gradient descent algorithm and its variants and how to apply it to solve regression and classification problems.
7. Work cooperatively to achieve common objectives, assuming own responsibility and respecting the role of the different members of the team.

Content

- Machine learning fundamentals
  - Supervised and unsupervised learning
  - Regression and classification
  - Gradient descent
  - Linear and polynomial regression
  - Logistic regression
  - Over/under-fitting and decomposition of bias and variance
  - Regularization
- Basic methods of pattern recognition
  - Nearest Neighbour models and case-based reasoning
  - Data clustering
  - Kernel Methods and Support Vector Machines
  - Decision trees and random forests
  - Experimentation and performance evaluation
- Probabilistic methods
  - Bayesian parameter estimation
  - Gaussian Mixture Models
  - Probability density estimation
  - Expectation Maximization

Methodology

There will be three types of teaching activities: theory classes, solving practical exercises individually (problems) and developing projects in small groups of 2-3 students.

1. Theory classes: Presentation of the theoretical content of the subject. For each of the topics studied, the main theoretical concepts and mathematical formulation are exposed, as well as the corresponding algorithmic solutions.

2. Laboratory sessions: The laboratory sessions aim to facilitate interaction and reinforce the understanding of the topics covered in the theory classes. During the laboratory sessions we will address two types of activities: the resolution of practical exercises (problems) and the monitoring and presentation of projects.

2.1 Problems: A weekly set of problems to work through will be used, starting from the bsis of Python code provided in Jupyter notebooks that exemplifies the coding details of the concepts exposed during theory classes. Work on the problems will begin in class and must be completed individually by each student at home. Students will be required to make a weekly submission of their work, that will comprise the problems portfolio.

2.2 Projects: Two projects will be carried out during the semester, where students will have to solve specific problems of certain complexity. The projects will be solved in small groups of 2-3 students, where each member of the group must contribute a part and put it together with the rest to obtain the final solution. These

working groups must be maintained until the end of the semester and must be self-managed in terms of distribution of roles, work planning, assignment of tasks, management of available resources, conflicts, etc. To develop the project, the groups will work autonomously, while the practical sessions will be used (1) for the teacher to present the projects' theme and discuss possible approaches, (2) for monitoring the status of the project and (3) for the teams to present their final results.

The above activities will be complemented by a system of tutoring and consultations outside class hours.

All the information of the subject and the related documents that the students need will be available at the virtual campus (cv.uab.cat).

Annotation: Within the schedule set by the centre or degree programme, 15 minutes of one class will be reserved for students to evaluate their lecturers and their courses or modules through questionnaires.

Activities

Title	Hours	ECTS	Learning Outcomes
Type: Directed			
Problem Sessions	12	0.48	6, 1, 5, 2
Project Sessions	9	0.36	6, 1, 5, 2, 3, 4, 7
Theory Classes	24	0.96	6, 5, 2
Type: Supervised			
Problem Solving (individual)	10	0.4	6, 1, 5, 2
Tutoring	2	0.08	6, 1, 5, 2, 3, 4
Work on projects (groups)	30	1.2	6, 1, 5, 2, 3, 4, 7
Type: Autonomous			
Individual studying	20	0.8	6, 1, 5, 2
Problem Solving (individual)	10	0.4	6, 1, 5, 2
Work on projects (groups)	20	0.8	6, 1, 5, 2, 3, 4, 7

Assessment

To assess the level of student learning, a formula is established that combines knowledge acquisition, the ability to solve problems and the ability to work as a team, as well as the presentation of the results obtained.

Final grade

The final grade is calculated weighted in the following way and according to the different activities that are carried out:

Final grade = 0.4 * Theory Grade + 0.1 * Problems Grade + 0.5 * Projects Grade

This formula will be applied as long as the theory and the laboratory grades, are higher than 5. There is no restriction on the exercises grade. If doing the calculation of the formula yields >= 5 but does not reach the minimum required in any of the evaluation activities, then a final grade of 4.5 will be given.

Theory Grade

The theory grade aims to assess the individual abilities of the student in terms of the theoretical content of the subject, this is done continuously during the course through two partial exams:

Theory Grade = 0.5 * Grade Exam 1 + 0.5 * Grade Exam 2

The mid-term exam (Exam 1) is done in the middle of the semester and serves to eliminate part of the subject if it is passed. The final exam (Exam 2) is done at the end of the semester and serves to eliminate the rest of subject if it is passed.

In order to obtain a final pass theory grade, it will be required for the partial exam grades 1 and 2 to be both higher than 4.

Recovery exam: In case the theory grade does not reach the adequate level to pass, the students can take a recovery exam, destined to recover the failed part (1, 2 or both) of the continuous evaluation process.

Problems Grade

The aim of the problems is for the student to become familiar with the practical implementation of the theoretical concepts. The weekly submission of problem solutions will be used as evidence of this work.

In order to obtain a grade for exercises, it is necessary that more than 50%of the exercises are submitted during the semester. In the contrary, the exercises grade will be 0.

Problems Grade = Portfolio evaluation

Projects Grade

The part of projects carries an essential weight in the overall mark of the subject. Developing the projects requires that the students work in groups and design an integral solution to the defined challenge. In addition, the students must demonstrate their teamwork skills and present the results to the class.

Each of the two projects is evaluated through its deliverable, an oral presentation that students will make in class, and a self-evaluation process. The participation of students in all three activities (preparing the deliverable, presentation and auto evaluation) is necessary in order to obtain a projects grade. The grade is calculated as follows:

Project Grade X = 0.6 * Grade Deliverables + 0.3 * Grade Presentation + 0.1 * Grade Self-evaluation

If performing the above calculation yields >= 5 but the student did not participate in any of the activities (deliverable, presentation, auto evaluation), then a final grade of 4.5 will be given to the corresponding project.

Projects Grade = 0.5 * Project Grade 1 + 0.5 * Project Grade 2

To obtain a project grade, it will be necessary that the grades of both projects are above 4.

There is no recovery of the projects: in case of not presenting a deliverable or considering it copied, if the final Project Grade does not exceed 5.0, the subject is considered failed.

Important notes

Notwithstanding other disciplinary measures deemed appropriate, and in accordance with the academic regulations in force, evaluation activities will be suspended with zero (0) whenever a student commits any academic irregularities that may alter such evaluation (for example, plagiarizing, copying, letting copy, ...). The

evaluation activities qualified in this way and by this procedure will not berecoverable. If you need to pass any of these assessment activities to pass the subject, this subject will be failed directly, without opportunity to recover it in the same course.

In case there the student does not deliver any exercise solutions, does not attend any project presentation session during the laboratory sessions and does not take any exam, the corresponding grade will be a "non-evaluable". In another case, the "no shows" count as a 0 for the calculation of the weighted average.

In order to pass the course with honours, the final grade obtained must be equal or higher than 9 points. Because the number of students with this distinction cannot exceed 5% of the total number of students enrolled in the course, it is given to whoever has the highest final marks. In case of a tie, the results of the partial exams will be taken into account.

Assessment Activities

Title	Weighting	Hours	ECTS	Learning Outcomes
Individual written exams	40%	4	0.16	6, 5, 2
Presentation of projects	15%	4	0.16	1, 3, 4, 7
Problem deliverables	10%	2	0.08	6, 1, 5, 2
Realization of projects (deliverables and self-assessment)	35%	3	0.12	3, 4, 7

Bibliography

- Data Science from Scratch: First Principles with Python, Joel Grus, O'Reilly Media, 2015, 1st Ed.
- Pattern Recognition and Machine Learning, Christopher Bishop, Springer, 2011
- Model-Based Machine Learning, J. Winn, C. Bishop, early access: http://mbmlbook.com/
- Introduction to Data Science. L. Igual, S. Seguí. Springer, 2017
- Pattern Classification, Duda, Hart, and Stork, 2nd Ed., 2002

Software

For the problems and projects of the course we will use Python, and the Python libraries NumPy, MatPlotLib, SciKit Learn, Pandas